Springer Tracts in Electrical and Electronics Engineering

Yash Agrawal
Kavicharan Mummaneni
P. Uma Sathyakam *Editors*

Interconnect Technologies for Integrated Circuits and Flexible Electronics

Springer

Springer Tracts in Electrical and Electronics Engineering

Springer Tracts in Electrical and Electronics Engineering (STEEE) publishes the latest developments in Electrical and Electronics Engineering - quickly, informally and with high quality. The intent is to cover all the main branches of electrical and electronics engineering, both theoretical and applied, including:

- Signal, Speech and Image Processing
- Speech and Audio Processing
- Image Processing
- Human-Machine Interfaces
- Digital and Analog Signal Processing
- Microwaves, RF Engineering and Optical Communications
- Electronics and Microelectronics, Instrumentation
- Electronic Circuits and Systems
- Embedded Systems
- Electronics Design and Verification
- Cyber-Physical Systems
- Electrical Power Engineering
- Power Electronics
- Photovoltaics
- Energy Grids and Networks
- Electrical Machines
- Control, Robotics, Automation
- Robotic Engineering
- Mechatronics
- Control and Systems Theory
- Automation
- Communications Engineering, Networks
- Wireless and Mobile Communication
- Internet of Things
- Computer Networks

Within the scope of the series are monographs, professional books or graduate textbooks, edited volumes as well as outstanding PhD theses and books purposely devoted to support education in electrical and electronics engineering at graduate and post-graduate levels.

Review Process

The proposal for each volume is reviewed by the main editor and/or the advisory board. The books of this series are reviewed in a single blind peer review process.

Ethics Statement for this series can be found in the Springer standard guidelines here https://www.springer.com/us/authors-editors/journal-author/journal-author-helpdesk/before-you-start/before-you-start/1330#c14214.

Yash Agrawal · Kavicharan Mummaneni ·
P. Uma Sathyakam
Editors

Interconnect Technologies for Integrated Circuits and Flexible Electronics

Springer

Editors
Yash Agrawal
Department of Information
and Communication Technology
DA-IICT
Gandhinagar, Gujarat, India

Kavicharan Mummaneni
Department of Electronics
and Communication Engineering
NIT Silchar
Silchar, Assam, India

P. Uma Sathyakam
School of Electrical Engineering
Vellore Institute of Technology
Vellore, Tamil Nadu, India

ISSN 2731-4200 ISSN 2731-4219 (electronic)
Springer Tracts in Electrical and Electronics Engineering
ISBN 978-981-99-4475-0 ISBN 978-981-99-4476-7 (eBook)
https://doi.org/10.1007/978-981-99-4476-7

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Interconnects have become essentially important in integrated circuit (IC) and printed-circuit board (PCB) designs. In the present state-of-the-art technologies such as network-on-chip, system-on-chip and 3D staggered integrated circuit designs, interconnects constitute large portion of the entire system. Interconnects dominantly control the signal integrity and are the major source for causing latency, crosstalk and power dissipation in the system. These non-idealities exaggerate deeper in nanotechnology dimension nodes as technology scales down.

In an advent to improve the performance of these systems, exploration of emerging on-chip/off-chip interconnects has become highly important. This book is dedicated towards investigation and modelling of advanced interconnects that can suitably match the current as well as future technology requirements. Different aspects of interconnects such as materials, its physical characteristics, parasitic extraction, design, structure, modelling, machine learning and neural network-based models, signalling schemes, crosstalk, signal integrity performance analysis, variability, reliability and associated electronic-design automation tools are explored and systematically detailed.

The book also covers interconnect technologies for flexible electronic system. The flexible electronics have gained much importance in today's electronic industry and are expected to cover a major portion of all consumable electronic products in near future. The interconnect technology for flexible electronics system design is highly vital and different from the traditional rigid-based interconnects those are laid over the hard bound PCBs. The flexible electronics-based systems require assimilation of advanced designs such as stretchable interconnects. These require different sets of materials, fabrication process as well as consideration of new added design parameters such as elasticity, twisting, folding and bendability. In order to realize a wide range of applications where stretching movements are prevalent, such as wearable healthcare monitoring, flexible electronic systems comprising of stretchable interconnects are essential. The research in flexible electronics and stretchable interconnects have progressed gradually over time. However, hunt for sustainable flexibility and improved electrical performance still offers a vast field of research and opportunities to potential researchers.

The various concepts for the aforesaid topics and technologies have been systematically covered in the following chapters. The basic analytical modelling of interconnects is handled in Chaps. 1–4. Chapter 1 discusses the importance of reducing the complexity of interconnect models that are used to calculate the delay and run time, by decreasing the model order through pole clustering and Pade's approximation. Delay and overshoot modelling of T-tree interconnects is outlined in Chap. 2. The finite-difference time-domain (FDTD) model is employed to find the performance of copper and SWCNT bundle interconnects in Chap. 3. Performance metrics like delay, power dissipation and power delay product are modelled for the two cases of voltage and current mode signalling schemes. In Chap. 4, factors like delay, crosstalk delay, overshoot, peak noise and timing instance are analysed for copper and CNT interconnects.

Long interconnects suffer from propagation delay and signal degradation. So, repeaters must be inserted to reduce signal delay and decay. In Chap. 5, the optimal repeater number for an interconnect segment is found by using particle swarm optimization and artificial neural networks-based approach. Chapter 6 is devoted to the analysis of Through Silicon Vias (TSVs). They are used for connecting the metal layers in a 3D IC. Here the fabrication challenges, geometry, shape and filler materials associated with TSVs are discussed. Chapters 7 and 8 discuss about application of neural networks and swarm intelligence-based algorithms for on-chip interconnect modelling.

Encoding schemes and test sequences using three level pulse amplitude modulation are discussed in Chap. 9. The PAM3 is compared to non-return to zero (NRZ) and four level PAM. Analysis is carried out to find the reflection coefficient, crosstalk and noise interconnects and circuits.

Chapters 10–16 provide an exhaustive account of flexible and stretchable interconnects for emerging applications such as flexible electronics, wearable electronics, portable medical devices, defence and aerospace applications. This also includes review chapters, which explore the recent, intensive and trendy development of flexible electronics and stretchable interconnects. The study specifically discusses the substrate, device and encapsulation approaches. In addition, an overview of stretchable interconnect geometries, fabrication techniques and materials used to fabricate substrate materials are presented. Further, development of sensors on flexible substrate is presented to demonstrate the concept of low-cost, lightweight system design for application such as plant disease monitoring. Finally, the key developments in the field of flexible electronics and stretchable interconnects have been emphasized by including a wide variety of applications.

Gandhinagar, India — Yash Agrawal
Silchar, India — Kavicharan Mummaneni
Vellore, India — P. Uma Sathyakam

Contents

About the Editors

Dr. Yash Agrawal is faculty at Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT) Gandhinagar, Gujarat, India. He has accomplished his Postdoc in network-on-chip design and graphene materials from University of Rennes, France. He graduated with his Ph.D. and M.Tech. degrees from the National Institute of Technology (NIT) Hamirpur, HP, India. He attained university rank from Nagpur University during his B.E. degree in Electronics and Communication in 2009. He achieved the third place in all India Mentor Graphics design contest held at Bengaluru, India, in 2011.

His research interest includes VLSI interconnects, EMC and signal integrity of high speed interconnects, optimization schemes, numerical method analysis, nanotechnology, network-on-chip designs, flexible electronics, FPGA-based designs, image processing using VLSI architectures, VLSI design for bio-medical devices, etc. He has conducted various Government funded workshops and conferences. He has several publications in contributed books, journals, and various national and international conferences of repute. He has received IETE K. S. Krishnan Memorial Award for best system oriented paper in 2017. He is reviewer of various reputed journals. He is a member of IEEE and chair of IEEE SSC&ED societies, Gujarat section 2023.

Dr. Kavicharan Mummaneni received B.Tech. degree in Electronics and Communication Engineering Department from JNTU Hyderabad, M.Tech. degree in Electronics Design and Technology specialization from NIT Calicut and Ph.D. in VLSI domain from NIT Warangal. Currently working as an Assistant Professor-I at National Institute of Technology Silchar. His research interests include: Microelectronic devices, VLSI interconnects, Signal Integrity, Power Integrity, Device Modelling, Carbon Nanotube Interconnects, Flexible electronic design. He has published several research papers and book chapters in the domain of Micro/ Nanoelectronic devices and circuits. He is a Senior IEEE member and Fellow of IETE India.

Dr. P. Uma Sathyakam is working as Associate Professor in the School of Electrical Engineering, Vellore Institute of Technology, Vellore, where he earned his M.S. (by Research) in Nanoelectronics and VLSI in 2011 and his Ph.D. in Nanoelectronics in 2018. Prior to this, he earned his B.Sc. Electronics from University of Kerala, Thiruvananthapuram in 2005 and M.Sc. Applied Electronics from Bharathiar University, Coimbatore with distinction and Gold medal in 2007.

He served as a Research Fellow in a MHRD sponsored project on 'Development of online Lab in Microelectronics and VLSI' at VIT. He was also a Research Associate during his M.S. at VIT. He has more than 13 years of teaching and research experience.

His current research interests are nano-electrodes for supercapacitors, Plasmonics, VLSI interconnects, carbon nanotube electronics, graphene, electronic materials and renewable energy materials. He has authored three books and 25 papers in peer reviewed journals and conferences. He is an Associate Editor of *Nanotechnology for Environmental Engineering*, a Springer Nature journal. He is also a peer reviewer for many journals. He has reviewed more than 75 papers for journals from IEEE, Elsevier, Springer, IET, Nature, Taylor & Francis, Trans Tech Publishers, and Wiley. He is a Senior Member of IEEE EDS/EPS.

Chapter 1
An Efficient Model Order Reduction of Interconnects Using Machine Learning for Timing Analysis

Kavicharan Mummaneni, Malvika, and Vivek Kumar

Abstract An interconnect can be modeled as a linear RLC equivalent, and this accurate interconnect structure can be modeled by using complex and higher-order transfer functions. These complex and higher-order transfer functions generally require large resources in terms of memory and time, especially for a timing analysis tool to calculate the interconnect delay (net delay) and run time, which is very high because of higher number of poles. This chapter describes the methodology to reduce the higher-order transfer function to lower-order transfer function keeping in view that stability and response of the reduced system should be matched with the original higher-order system. The model order reduction achieves an advantage in terms of memory and time complexity. This work utilizes pole clustering and Pade's approximation techniques to reduce the order of the system.

Keywords Machine learning · Pole clustering · Interconnects · Pade's approximation · Markov parameters

Introduction

The research of algorithms which get better over practice is known as machine learning. Utilizing sample data, machine learning aims to establish a model known as training dataset. It is mostly chosen for predicting and decision-making which is not externally programmed. In machine learning (ML), mainly two types of clustering techniques are useful in these applications. Those are partitional clustering and hierarchical clustering. Along with these clustering techniques, numerator reduction has been performed using Pade's approximation. The objective of this work is to condense the order of higher-order original transfer function so that the reduced ordered system should mimic the original system. The correlation between two

K. Mummaneni (✉) · Malvika · V. Kumar
Department of Electronics and Communication, NIT Silchar, Silchar, Assam, India
e-mail: kavicharan@ece.nits.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_1

systems can be observed by the metrics such as the integral square error of step responses and transient delay parameters like rise time, settling time, and delay time differences. In addition, timing metrics correlation in frequency response has been calculated in terms of bandwidth.

The proposed machine learning technique is based on pole clustering approach (Beyene 2017) and Pade's approximation (Jay Singha and Vishwakarma 2016) for reducing the order of transfer function. The denominator of the higher-order transfer function has been simplified by pole clustering approach and for numerator reduction, the moments matching technique is used. However, order reduction should be done in such a way that stability and response of the system must be retained in the reduced model. To compare both the reduced system with original system, integral square error (ISE) and timing metrics are introduced in the work.

Pole clustering is a technique that clusters a given set of poles by searching the hidden patterns that may be present in the set. It is a process of grouping "similar" poles into disjointed pole clusters. Clustering is generally of two types: partitional clustering and hierarchical clustering. Partitional clustering divides random poles into a set of clusters called as partitions. In this kind of clustering, initially the user must define the required number of clusters such as K-Means clustering. In hierarchical clustering, the algorithms repeat the cycle of either merging smaller clusters into larger ones or dividing larger clusters into smaller ones. Finally, it produces a hierarchy of clusters called a dendrogram (Beyene 2017), which is a graphical representation of the clusters to display the hierarchy very clearly.

Both K-means and agglomerative hierarchical clustering algorithms require some distance measure to be defined in the data space. Euclidean distance, Manhattan distance, Minkowski distance, and cosine similarity are some of the most used metrics of similarity for numeric data. For pole clustering, the distance measure that preserves the dominant poles has been used to generate stable reduced order models (Han and Kamber 2012; Vishwakarma and Prasad 2008a). The methods calculate the cluster centers using the poles in each cluster by applying the Inverse Distance Measure (IDM) criteria.

Proposed Model Order Reduction

A SISO causal LTI dynamic system is represented below.

Let the transfer function of higher-order original system be of the order "n" (Jay Singha and Vishwakarma 2016).

$$G_n(s) = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_{n-1} s^{n-1}}{b_0 + b_1 s + b_2 s^2 + \cdots + b_n s^n} \tag{1.1}$$

The corresponding desired reduced order model for kth order system (Vishwakarma and Prasad 2008b) is

$$G_k(s) = \frac{c_0 + c_1 s + c_2 s^2 + \ldots a_{k-1} s^{k-1}}{d_0 + d_1 s + d_2 s^2 + \cdots + d_n s^k} \tag{1.2}$$

The resultant denominator from pole clustering results as

$$D_k(s) = (s - p_{e1})(s - p_{e2}) \ldots\ldots (s - p_{ek}) \tag{1.3}$$

The reduced numerator is obtained by applying Pade's approximations.

Denominator Reduction

The denominator of transfer function is reduced to kth order using pole clustering approach by using inverse distance metric. The initial step of denominator reduction is to divide the total poles in the form of clusters. The outcome of clustering will be groups of poles (clusters) with disjoint poles, i.e., clusters that are formed must be different; any two clusters should not have even a single common pole. In the formation of clusters, one needs to arrange the poles in an order so that they can be easily divided. The poles are arranged in the decreasing order of their dominance and clusters are formed in such a way that the first cluster has the most dominant poles and subsequent clusters can have lesser dominant poles than their previous cluster (Vishwakarma and Prasad 2009). The cluster center is derived by using the following algorithm and the output of algorithm is a reduced order pole.

Algorithm

Let us assume each cluster has "r" number of poles p_1, p_2, p_3 ….. p_r respectively.

1. Let "r" denotes number of poles in the cluster such that $|p_1| < |p_2| < |p_3| \ldots\ldots |p_r|$
2. Set j = 1
3. Find the pole cluster using

$$c_j = \left[\left(\sum_{j=2}^{r} \left(-\frac{1}{|p_i - p_1|} \right) + \left(-\frac{1}{|p_1|} \right) \right) \div r \right]^{-1} \tag{1.4}$$

4. Set j = j + 1.
5. Now find a modified or new cluster center from Vishwakarma and Prasad (2008b).

$$c_j = \left[\left(\left(-\frac{1}{|p_1|} \right) \right) + \left(-\frac{1}{|c_{j-1}|} \right) \right) \div 2 \right]^{-1} \tag{1.5}$$

6. Check whether j = r + 1, if not go to step 4
7. Else consider the result in 5th step as cluster center.

Let us consider an example transfer function

$$G_n(s) = \frac{40320 + 185760s + 222088s^2 + 122664s^3 + 36380s^4 + 5982s^5 + 514s^6 + 18s^7}{40320 + 109584s + 118124s^2 + 67284s^3 + 2244s^4 + 4536s^5 + 514s^6 + 36s^7 + s^8} \tag{1.6}$$

The above transfer function has 8 poles, and it can be reduced to 2nd order by reducing the effective number of poles to 2.

Total number of poles (n) = 8.
Order of required reduced model (k) = 2.
As can be seen now, number of clusters = order of reduced model.
Number of clusters required = 2.
Maximum number of Poles per cluster = n/k = 4.
1st cluster: (−1, −2, −3, −4), 2nd cluster: (−5, −6, −7, −8).
Cluster centers obtained from algorithm are
1st cluster center: −1.0186, 2nd cluster center: −4.5606.
2nd order Denominator is (S + 1.0186) (S + 4.5606)

$$= S^2 + 5.5792S + 4.6454 \tag{1.7}$$

Numerator Reduction

Numerator of reduced order model is obtained from Pade's approximations. The original nth order system can be expanded in power series about s = 0 as

$$G_n(s) = \frac{a_0 + a_1 s + a_2 s^2 + \ldots + a_{n-1} s^{n-1}}{b_0 + b_1 s + b_2 s^2 + \ldots + b_n s^n} = t_0 + t_1 s + t_2 s^2 + \ldots \tag{1.8}$$

where t_0, t_1, t_2 are called timing moments.

The timing moments capture the low-frequency dynamics of the system response. In general, the low-frequency variations are observed during the steady state behavior of the system. These coefficients are also termed as low-frequency moments, and purposefully introduced to match steady-state moments of reduced model to that of original model. These parameters can be obtained from simulations in a simple method by dividing the numerator with its denominator of the transfer function by arranging the terms from lower powers of "s" to higher powers of "s".

The transfer function considered in (1.6) has timing moments as 1.0000, 1.8893, and 2.5563 respectively. In contrast to these timing moments, original transfer function can be expanded in power series about S = ∞ as

$$G_n(s) = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_{n-1} s^{n-1}}{b_0 + b_1 s + b_2 s^2 + \cdots + b_n s^n} = M_0 s^{-1} + M_1 s^{-2} + M_2 s^{-3} + \ldots \tag{1.9}$$

where M_0, M_1, and M_2 are called Markov parameters (Vishwakarma and Prasad 2008b).

The Markov parameters capture the high-frequency dynamics of system at expansion point $S = \infty$. In general, the high-frequency variations are observed during the initial transient behavior of the system.

$$M_i = \frac{d^i h(t)}{d^i t} \tag{1.10}$$

The first Markov moment is nothing but the impulse response at $t = 0$. The transfer function considered in (1.6) has Markov parameter values as 18, −134, and 978 respectively.

To find out the reduced order model numerator coefficients, equate the assumed reduced model transfer function in (1.2) with power series about $S = 0$ and $S = \infty$ (Jay Singha and Vishwakarma 2016).

$$G_k(s) = \frac{c_0 + c_1 s + c_2 s^2 + \ldots a_{k-1} s^{k-1}}{d_0 + d_1 s + d_2 s^2 + \ldots + d_n s^k} = t_0 + t_1 s + t_2 s^2 + \ldots \tag{1.11}$$

Reduced order denominator is obtained with pole clustering as in (1.3). Cross-multiply (1.11) is followed by comparison of coefficients. The reduced numerator coefficients are

$$c_0 = d_0 t_0$$

$$c_1 = d_0 t_1 + d_1 t_0$$

$$c_2 = d_0 t_2 + d_1 t_1 + d_1 t_0 \tag{1.12}$$

These equations are valid only if timing moments are used in the numerator reduction. The method to find timing moments is observed as dividing numerator by denominator when terms are arranged from lower powers of "s" to higher powers of "s". In the other way, Markov parameters are observed when numerator is divided by denominator when terms are arranged from higher powers of "s" to lower powers of "s". As a result, in reduced numerator coefficients, lower powers of "s" coefficients are associated with timing moments and higher powers of "s" coefficients are associated with Markov parameters as follows:

Lower power "s" coefficients: $c_0 = d_0 t_0$

$$c_1 = d_0 t_1 + d_1 t_0$$

Higher power "s" coefficients: $c_{k-1} = d_k M_0$

$$c_{k-2} = d_k M_1 + d_{k-1} M_0 \tag{1.13}$$

These equations can be generalized to any higher-order transfer function reduction.

Results and Discussions

Consider the timing moments in numerator reduction from the values in (1.12) and (1.7).

$C_0 = 4.6454, C_1 = 14.3557$, which results in 2nd order transfer function (Fig. 1.1):

$$G_2(s) = \frac{14.3557s + 4.6454}{s^2 + 5.5792s + 4.6454} \tag{1.14}$$

Consider both timing and Markov moments, from the Eqs. (1.13) and (1.7)

$c_0 = 4.6454$, $c_1 = 18$, which results in 2nd order transfer function (Fig. 1.2):

$$G_2(s) = \frac{18s + 4.6454}{s^2 + 5.5792s + 4.6454} \tag{1.15}$$

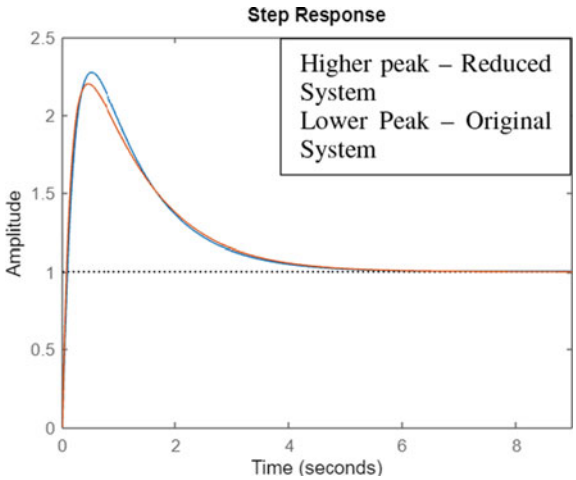


Fig. 1.1 Step response comparison of original system and reduced system (with timing moments)

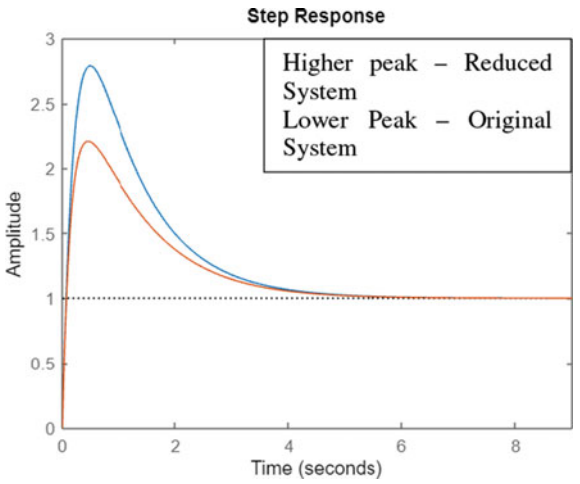


Fig. 1.2 Step response comparison of original system and reduced system (with timing moments and Markov parameters)

The Integral Square Error (ISE) is considered to compare the step response of reduced system and original system, because square of the error value gives an actual idea about accuracy and the difference may nullify the actual error by canceling the positive errors and negative errors. It can be observed from Table 1.1 that the error value is less as compared to others when only timing moments are used. The rise time, peak time, and settling time of Markov moments-based method provides close match with the original system (Vishwakarma and Prasad 2009, 2011; Komarasamy 2011).

Table 1.2 presents the comparison of step response of original system and reduced systems. As discussed already, Markov moments capture the high frequency dynamics of the system in terms of rise time and bandwidth parameters which are closer to the original system when Markov moments are introduced as compared to other methods. The purpose of order reduction is mainly used in timing analysis, as paramount importance is given to match the response parameters.

Table 1.1 Comparison of step response of proposed method with existing method

S.no	Method	Rise time	Peak time
1	Timing moments	0.0692	0.5049
2	Timing and Markov parameters	0.0525	0.4847
3	Existing method (Ramawat et al. 2012)	0.082	0.531

Table 1.2 Comparison of step response of original system and reduced systems

S. no	Response parameter	Original system	Reduced system with timing moments	Reduced system with timing and Markov moments
1	Rise time (sec)	0.0569	0.0692	0.0525
2	Peak time (sec)	0.4490	0.5049	0.4847
3	Settling time (sec)	4.8201	4.5992	4.5781
4	Bandwidth (Hz)	24.3644	19.7336	24.9933
5	ISE		0.299	0.0094

Conclusion

This chapter presented significant improvement in estimation of rise time, peak time, and bandwidth with the introduction of timing moments and Markov moments as compared to the existing methods. The proposed method mainly depends upon developing an expansion point and cluster center formula. Expansion points are considered here as $S = 0$ and $S = \infty$. Having $S = 0$, same steady state accuracy and gain is attained by accounting timing moments, while, at $S = \infty$, a relatively similar behavior by taking Markov moments into consideration is obtained, thereby, preserving the transient response. By choosing further optimized proper expansion point, i.e., region of frequency of interest, better accuracy can be attained.

References

Beyene WT (2017) Reduced-order modeling of high-speed channels using machine learning techniques: partitional and hierarchical clustering. IEEE EPEPS.

Bhasker J, Chadha R (2009) Static timing analysis for nanometer designs. A Practical Approach, Springer

Han J, Kamber M (2012) Data mining concepts and techniques, Waltham. Morgan Kaufmann Publishers, MA

Komarasamy R, Albhonso N, Gurusamy G (2011) Order reduction of linear systems with an improved pole clustering. J Vib Control, 1876–1885

Ramawat K, Bhandari M, Kumar A (2012) Model order reduction by pade approximation and improved pole clustering technique. In: Energy Technology & Ecological Concerns: A Contemporary Approach, pp 103–107.

Singha J, Vishwakarma CB, Chattterjeea K (2016) Biased reduction method by combining improved modified pole clustering and improved pade approximations. Appl Math Model 1418–1426

Vishwakarma CB, Prasad R (2008a) System reduction using modified pole clustering and pade approximation. In: xxii national system conference.

Vishwakarma B, Prasad R (2008b) Clustering method for reducing order of linear system using padé approximation. IETE J Res 54:326–330

Vishwakarma CB, Prasad R (2009) MIMO system reduction using pole clustering and genetic algorithm. Model Simul Eng 1–5.

Waste NHE, Harris DM (2011) CMOS VLSI design a circuits and systems perspective. Pearson Education

Chapter 2
Delay and Overshoot Modelling of Asymmetric T-Tree Interconnects

Malvika, Vivek Kumar, and Kavicharan Mummaneni

Abstract With lower technology nodes and higher data rates, signal integrity has become a big challenge nowadays. The chip performance and signal integrity depend on interconnect delay in deep submicron (DSM) technology nodes. The performance and packing density of high functionality chips are limited by the interconnect delay than the overall circuit delay. This paper discusses the delay and overshoot modeling of asymmetric T-tree interconnect. These delay and overshoot modeling are performed based on the numerical Laplace inversion method. The proposed model is based on the Gaver-Stehfest method, which provides a relationship between the time-varying function and the linear combination of values of transform function. In this work, 50% delay and overshoot are estimated for asymmetric T-tree interconnect for different line lengths, source resistance, and load capacitance. In order to validate the accuracy of the proposed model, the estimated 50% delay and overshoot values are compared with the standard LT Spice model and are found to be in good agreement.

Keywords Delay · Signal integrity · Asymmetric T-tree interconnect · ABCD-matrix · Gaver-stehfest based model

Introduction

Nowadays due to scaling down the transistor sizes, i.e., lower technology node, owing to higher data rate and larger operating frequency of devices, signal integrity has become a big challenge. Since the area of chip is decreasing and the number of transistors in a chip is increasing, interconnect modeling has become more challenging. Due to this, the spacing between interconnect is decreasing which leads to higher coupling between interconnects and signal integrity issues. Delay is an important factor for determining the performance of silicon on chip (SoCs) where

Malvika (✉) · V. Kumar · K. Mummaneni
Department of Electronics and Communication, NIT Silchar, Silchar, Assam, India
e-mail: malvikathakur16@gmail.com

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_2

long interconnects are used for the connection of cores of Intellectual property blocks. This is because as technology scaled down, interconnect delay does not scale along with gate delay. The performance and packing density of high-functionality chips are limited by the interconnect delay than the overall circuit delay. One in five chips fails due to signal integrity. Due to the lower technology node, millions of transistors are used in a single chip and correspondingly millions of interconnects are required to connect those transistors that make the distance between interconnect very less and causing crosstalk induced delay which makes interconnect analysis very challenging. Hence, calculation of delay and overshoot is highly important for the performance of the chip and reliability.

Delay and overshoot analyses can be implemented through analytical models or simulation techniques (tools) like SPICE to get accurate results. For layout optimization, these simulation techniques are computationally expensive. Hence analytical models can be used for finding delay and overshoot (Sakurai 1993; Kobayashi and Sakurai 1991; Friedman et al. 2000; Mehrotra and Bannerjee 2002; Muddu and Kahng 1997; Coulibaly and Kadim 2005; Asai and Tanji 2004). Earlier, the signal delay is calculated by using Elmore's model which uses only resistance and capacitance where inductance is neglected. Elmore's models are used for two-pole transfer functions. To find the signal delay and to capture the high-frequency effect of the dominant inductive RLC line, two poles are not enough, hence these models are limited. Nowadays due to higher operating speed and longer length of interconnect, wire inductance cannot be ignored. To obtain higher accuracy in a high-speed VLSI circuit, effect of inductance needs to be included along with resistance and capacitance. Several research works are going on for analytical modeling of RLC interconnect. Model order reduction techniques such as asymptotic waveform evaluation (AWE) and passive reduced-order interconnect macro-modeling algorithm (PRIMA) have been used for accurate modeling of interconnects in Pillage and Rohrer (1990); Odabasioglu et al. 1998). These two methods yield the higher-order transfer function, which makes the computation complex. These model order reduction techniques are restricted to two and multi-conductor lines which are identical and having the same driver and load. In (Dounavis et al. 1999, 2001), interconnect analysis is developed which utilizes exponential matrix rational approximation models such as Pade approximation (Roy and Dounavis 2009) and matrix Pade approximation method (Kavicharan et al. 2015) that can be obtained from predetermined co-efficient. Many delay models have been proposed to calculate the delay of symmetric single, tree, and multi-conductor lines (Pillage and Rohrer 1990; Odabasioglu et al. 1998; Dounavis et al. 1999, 2001; Roy and Dounavis 2009; Kavicharan et al. 2015). Hence, there is a great demand for developing accurate models for the analysis of asymmetric T-tree interconnect.

In this work, an analytical method for on-chip RLC distributed interconnect is proposed for signal delay evaluation. The proposed model is more accurate than above mentioned analytical models (Roy and Dounavis 2009; Kavicharan et al. 2015). In this paper, delay and overshoot of asymmetric T-tree interconnect are estimated. These delay and overshoot modeling are performed based on the numerical Laplace inversion method. It is based on the Gaver-Stehfest method, which provides

a relationship between the time-varying function and the linear combinations of the transform functions values. The paper is organized as follows. Section "Proposed Model" contains the analysis of 1:2 asymmetric T-tree interconnect network. The overall voltage transfer function is calculated with the help of ABCD-matrix analyses and the expression for final output voltage is derived. Then the Gaver-Stehfest algorithm is employed for the calculation of delay and overshoot. Section "Simulation and Results" contains the simulation results. Finally, Sect. "Conclusion" concludes the work.

Proposed Model

Analysis of 1:2 Asymmetric T-Tree InterConnect

In this section, 1:2 asymmetrical T-tree interconnect is considered. The distributed asymmetric T-tree interconnect is considered as single input multiple output (SIMO) topology. The asymmetrical T-tree interconnect transfer function is extracted with the help of ABCD-matrix analysis. The elements of the T-tree interconnect topology are considered as transmission lines which are defined with the physical length and characteristic impedance. The asymmetrical T-tree interconnect network comprises a single input node and two output nodes. The equivalent circuit of input and output path is represented with simplified single input and single output (SISO) ABCD-matrix. The analytical expression of single input single output transfer function is the behavioral model of T-tree interconnect. This transfer function is used for the prediction of time domain and frequency domain responses of the 1:2 asymmetrical T-tree interconnect for any electrical path (Modelling and of asymmetrical interconnect T-tree laminated on flexible substrate (2015); Ravelo et al (2014)).

Figure 2.1 shows the representation of 1:2 asymmetrical distributed T-tree interconnect network with single input and two output branches. The input transmission line is denoted by characteristic impedance Z_{in} and physical length d_{in}. The output transmission line branches are denoted by characteristic impedance Z_{Ck} and physical length d_k where k = 1, 2. The source voltage V_{in} is connected to input transmission line TL_{in} through source impedance Z_s. TL_1 and TL_2 denote first and second output branch transmission line with load impedances Z_{L1} and Z_{L2}.

The characteristic impedance of output transmission lines "TL_k" is given as

$$Z_{Ck}(S) = \sqrt{R(s) + sL(s)/G(s) + sC(s)} \quad where\, k = 1,\ 2 \tag{2.1}$$

The propagation constant of output transmission lines "TL_k" is given as

$$\gamma_k(S) = \sqrt{(R(s) + sL(s)) * (G(s) + sC(s))} \tag{2.2}$$

Electrical length of input transmission line "TL_{in}" is obtained as

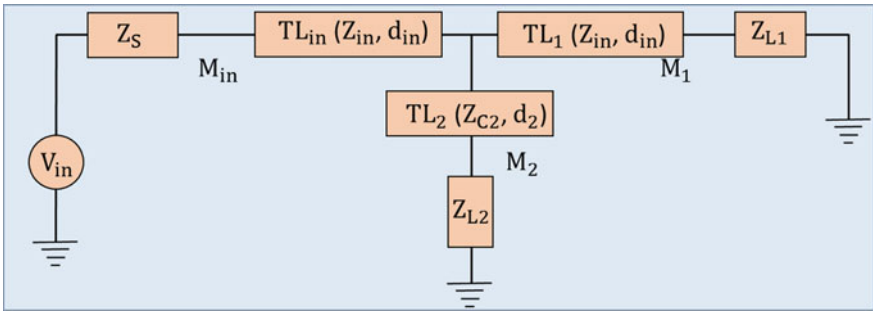


Fig. 2.1 1:2 asymmetrical distributed T-tree interconnect network

$$\theta_{in}(s) = \gamma_{in}(s) * \gamma_{in} \tag{2.3}$$

Electrical length of output transmission line "TL_K" is evaluated as

$$\theta_k(s) = \gamma_k(s) * \mathrm{d}_k \tag{2.4}$$

The ABCD matrices of input transmission line "TL_{in}" is given as

$$[\mathrm{ABCD}]_{in} = \begin{bmatrix} \cosh(\theta \mathrm{in}) & \mathrm{Zinsinh}(\theta \mathrm{in}) \\ \frac{\sinh(\theta \mathrm{in})}{Zin} & \cosh(\theta \mathrm{in}) \end{bmatrix} \tag{2.5}$$

The ABCD matrices of output transmission lines "TL_k" is defined as

$$[\mathrm{ABCD}]_{TL_k} = \begin{bmatrix} \cosh(\theta_k) & \mathrm{Zcksinh}(\theta_k) \\ \sinh(\theta_k)/Z_{Ck} & \cosh(\theta_k) \end{bmatrix} \tag{2.6}$$

The asymmetric T-tree network output load impedance ABCD-matrix is given as

$$[\mathrm{ABCD}]_{Z_k} = \begin{bmatrix} 1 & 0 \\ 1/Z_k & 1 \end{bmatrix} \tag{2.7}$$

The ABCD matrices of equivalent output branches "$TL_k - Z_k$" is obtained from (1.6)*(1.7) as

$$[\mathrm{ABCD}]_{TL_k - Z_k} = \begin{bmatrix} \cosh(\theta_k) & \mathrm{Zcksinh}(\theta_k) \\ \sinh(\theta_k)/Z_{Ck} & \cosh(\theta_k) \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 1/Z_k & 1 \end{bmatrix}$$

$$[\mathrm{ABCD}]_{TL_k - Z_k} = \begin{bmatrix} \cosh(\theta_k) + (Z_{Ck}/Z_k)\sinh(\theta_k) & Z_{Ck}\sinh(\theta_k) \\ (\sinh(\theta_k)/Z_{Ck}) + (\cosh(\theta_k)/Z_k) & \cosh(\theta_k) \end{bmatrix} \tag{2.8}$$

In order to reduce SIMO (single input multiple output) into SISO (single input single output), the remaining branches of the network can be considered as an

equivalent input admittance with respect to M_{in} to M_k,

$$Yk,\ node = \frac{\left(\frac{\sinh(\theta k)}{Zck}\right) + \left(\frac{\cosh(\theta k)}{Zk}\right)}{\cosh(\theta k) + (Z_{Ck}/Z_k)\sinh(\theta k)} \tag{2.9}$$

The total ABCD-matrix of any $M_{in}\_M_k$ electrical path of the network shown in Fig. 2.1 can be determined with the overall matrix relation as

$$[\text{ABCD}]_{M_{in}\_M_k} = \begin{bmatrix} 1 & R \\ 0 & 1 \end{bmatrix} * [\text{ABCD}]_{in} * \begin{bmatrix} 1 & 0 \\ Yk, node & 1 \end{bmatrix} * [\text{ABCD}]_{TL_k} * [\text{ABCD}]_{Z_k} \tag{2.10}$$

Here A_{11} is the first matrix element of above 2*2 matrix multiplication

$$\begin{aligned} A_{11} = {} & [\cosh(\theta_{in}) * \cosh(\theta_k)] + \left[\frac{R}{zin} \sin\ h(\theta_{in}) *) \cosh(\theta_k)\right] \\ & + [Yk,\ node\ Zin\ \sinh(\theta in)\ \cosh(\theta_k)] + [Yk,\ node\ R\ \cosh(\theta_{in})\ \cosh(\theta_k)] \\ & + \left[\frac{Zin}{Zck}\ \sinh(\theta_k)\sinh(\theta_{in}) *) +\right] \\ & \left[\frac{R}{Zck}\sinh(\theta_k)\cosh(\theta_{in}) *)\right] + \left[\frac{Zck}{Zk}\cosh(\theta_{in}) \cdot \sinh(\theta_{in})\right] \\ & + \left[\frac{RZck}{ZkZin}\sinh(\theta_{in}) *)\sinh(\theta_{in})\right] + \left[\frac{Yk,\ node\ Zin\ Zck}{Zk}\sinh(\theta_{in})\sinh(\theta_k)\right] \\ & \left[\frac{R\ Zck}{Zk}\text{Cosh}(\theta in)\sinh(\theta_k)\right] + \left[\frac{Zin}{Zk}\sinh(\theta_{in})\cosh(\theta_k)\right] + \left[\frac{R}{Zk}\cosh(\theta_{in})\cosh(\theta_k)\right] \end{aligned} \tag{2.11}$$

$$\text{The overall transfer function is given by } T_K = \frac{Vk(S)}{Vin(S)} = \frac{1}{A_{11}},\ where\ k = 1,\ 2 \tag{2.12}$$

The Gaver-Stehfest Algorithm

This method is the one-dimensional inversion algorithm in the unified framework for numerically inverting Laplace transforms (Sakurai 1993). The unified framework for Laplace inversion is based on the given expressions as

$$f(t) \approx f_n(t) = \frac{1}{t} \sum_{k=0}^{n} \omega_k f^{\wedge}\left(\frac{\alpha k}{t}\right),\ 0 < t < \infty \tag{2.13}$$

where ω_k represents weights and α_k denotes nodes

These weights and nodes are complex numbers and derived from

$$e^z \approx \sum_{k=0}^{n} \frac{\alpha_k}{\omega_k - Z} \tag{2.14}$$

The inversion formula for a Laplace transform F(s) is given as

$$f(t) \approx \frac{\ln(2)}{t} \sum_{k=1}^{2M} \zeta_k F\left(\frac{k \ln(2)}{t}\right) \tag{2.15}$$

where $\zeta_k = (-1)^{M+k} \sum_{j=[(k+1)/2]}^{\min\{k,M\}} \frac{j^{M+1}}{M!} \binom{M}{j} \binom{2j}{j} \binom{j}{k-j}$.

The Pade approximation and matrix Pade-type approximation (MPTA) algorithms require expressions to calculate poles and residues (Noda et al. 1991; Friedman et al. 2000), where the Gaver-Stehfest inversion formula uses only numerical calculations. The precision and significant digits produced by the above algorithm are directly proportional to M. For all the simulations, we considered M = 10. For discrete time calculation, we considered t, ζ_k and k as vectors of dimensions 1 × T, 1 × 2 M and 1 × 2 M, respectively.

$$\frac{In(2)}{t} = \left[\frac{In(2)}{t_1}, \frac{In(2)}{t_2}, \ldots \frac{In(2)}{t_T}\right]$$

$$\zeta_k = [\zeta_1, \zeta_2, \ldots.., \zeta_{2M}], k = [1,\ 2,\ 3,\ \ldots 2M]$$

For single line interconnect, the output function in frequency domain is given as

$$V_o(s) = H(S) * V_i(S) \tag{2.16}$$

The final output is obtained as

$$V_o \approx \frac{\ln(2)}{t} \circ \zeta_k V_o\left(k^T \frac{\ln(2)}{t}\right) \tag{2.17}$$

where ○ represents element-wise multiplication of two matrices of same dimensions.

Simulation and Results

1:2 Asymmetrical T-Tree Interconnect

The interconnect parameters are provided in (Table 2.1Kavicharan et al. 2013), which are used to estimate delay and overshoot for single input two outputs 1:2 asymmetric T-tree interconnect network as shown in Fig. 2.1. In this section, delay and overshoot values of asymmetric T-tree interconnect are compared with SPICE values and average error percentages are calculated. The 50% delay and overshoot of 1:2 asymmetric T-tree are calculated using proposed Gaver-Stehfest-based model for various lengths, source resistance (R_S), and load capacitance (C_l) of first and second output lines. Tables 2.2 and 2.3 illustrate the 50% delay and overshoot of the proposed model and LT spice results for various transmission (Tx) line lengths of first output line, respectively. It can be observed that the proposed model results match very well with LT SPICE except for longer lengths. Similarly, Tables 2.4 and 2.5 demonstrate the 50% delay and overshoot of the proposed model and LT SPICE results for various transmission (Tx) line lengths of second output line, respectively. In this case, the estimated 50% delay and overshoot values of the proposed model are in good agreement with LT SPICE.

The time domain analyses are carried out by injecting a unit ramp input with amplitude of 1 V, rise time of 0.1 ns. Figures 2.2 and 2.4 show the time domain analysis for input Tx line length = 0.2 cm, $R_S = 50\Omega$, $C_l = 50$fF of first output line and second output line, respectively. Figures 2.3 and 2.5 present the time domain analysis for input Tx line length = 0.5 cm, $R_S = 50\Omega$, $C_l = 50$fF of first output line

Table 2.1 Typical parameters of asymmetric T-tree interconnect

Parameters	Input Tx line values	First output Tx line values	Second Output Tx line values
V_{DD}−V	1	–	–
Length (l)−cm	0.2 to 1	0.2	0.3
Resistance (R)−Ohm	88.29	100	100
Inductance (L)–nH	15.38	16.15	16.15
Capacitance (C)–pF	1.8	1.8	1.8
Load capacitance (C_l)−fF	–	50	50
Source resistance (R_S)–Ohm	50	–	–
Input rise time (T_r)−ns	0.1	–	–

Table 2.2 Comparison of 50% delay of proposed Gaver-Stehfest-based model and LT spice for various input transmission (Tx) line lengths of first output line

Input Tx line length (cm)	R_S (Ω)	C_l(fF)	LT Spice	Proposed model	Proposed model
			50% Delay(ps)	50% Delay(ps)	Error %
0.2	50	50	139.74	137.56	1.56
0.3	50	50	159	156.96	1.28
0.4	50	50	181.38	176.22	2.84
0.5	50	50	201.72	195.92	2.87
0.6	50	50	223.48	216.14	3.28
1	50	50	320.38	301.85	5.78
Average % error					2.93%

Table 2.3 Comparison of overshoot of proposed Gaver-Stehfest-based model and LT spice for various input transmission (Tx) line lengths of first output line

Input Tx line length (cm)	R_S (Ω)	C_l (fF)	LT spice	Proposed model	Proposed model
			Overshoot	Overshoot	Error %
0.2	50	50	1.077	1.066	1.02
0.3	50	50	1.066	1.069	0.28
0.4	50	50	1.059	1.064	0.47
0.5	50	50	1.057	1.055	0.18
0.6	50	50	1.050	1.045	0.47
1	50	50	1.00	1.01	1.00
Average % error					0.57%

Table 2.4 Comparison of 50% delay of proposed Gaver-Stehfest-based model and LT spice for various input transmission (Tx) line lengths of second output line

Input Tx line length (cm)	R_S (Ω)	C_l (fF)	LT spice	Proposed model	Proposed model
			50% delay(ps)	50% delay(ps)	Error %
0.2	50	50	159.86	157.69	1.35
0.3	50	50	177.99	178.32	0.18
0.4	50	50	197.77	198.87	0.55
0.5	50	50	217.54	219.86	1.06
0.6	50	50	237.32	241.45	1.7
1	50	50	332.25	334.12	0.56
Average % error					0.9%

Table 2.5 Comparison of overshoot of proposed Gaver-Stehfest-based model and LT spice for various input transmission (Tx) line lengths of second output line

Input Tx line length (cm)	R_S (Ω)	C_l (fF)	LT spice	Proposed model	Proposed model
			Overshoot	Overshoot	Error %
0.2	50	50	1.061	1.035	2.5
0.3	50	50	1.068	1.037	2.90
0.4	50	50	1.076	1.034	3.90
0.5	50	50	1.070	1.027	4.01
0.6	50	50	1.055	1.020	3.31
1	50	50	1.00	0.997	0.3
Average % error					2.82%

and second output line, respectively. It is apparent that the proposed model results well match with LT SPICE results except during transition of signals, which is due to the approximation of transfer function for simplicity of analysis. Furthermore, the proposed models deliver benefits in terms of simplicity and accuracy.

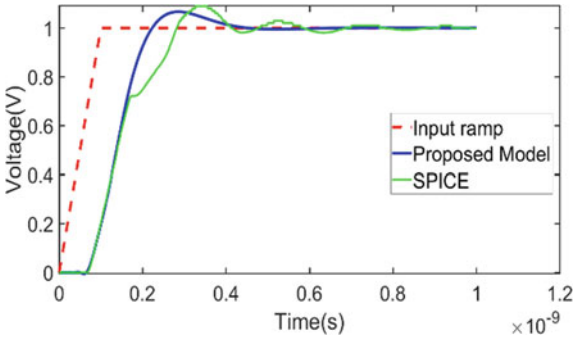


Fig. 2.2 Time domain analysis of first output line, for input Tx line length = 0.2 cm, $R_s = 50\Omega$, $C_l = 50$fF

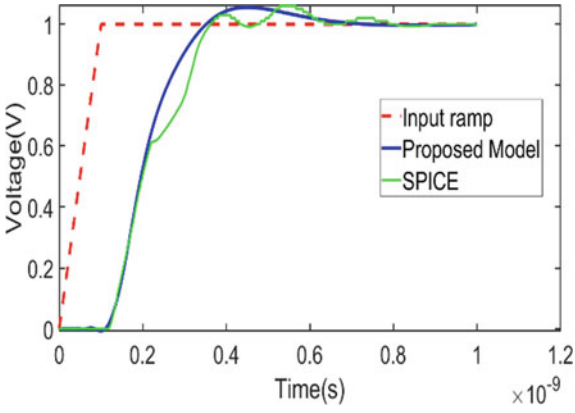


Fig. 2.3 Time domain analysis of first output line, for input Tx line length = 0.5 cm, $R_s = 50\Omega$, $C_l = 50$fF

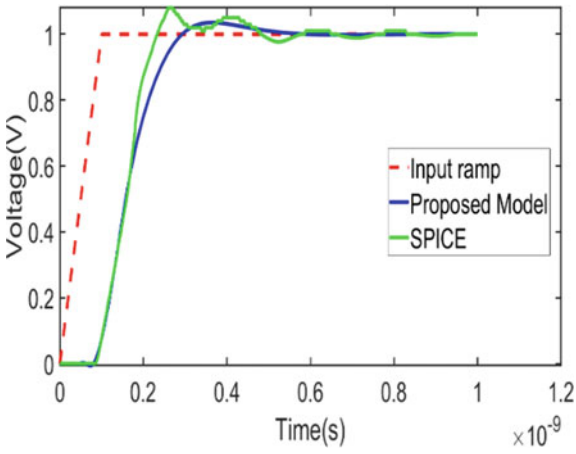


Fig. 2.4 Time domain analysis of second output line, for input Tx line length = 0.2 cm, $R_s = 50\ \Omega$, $C_l =$ 50fF

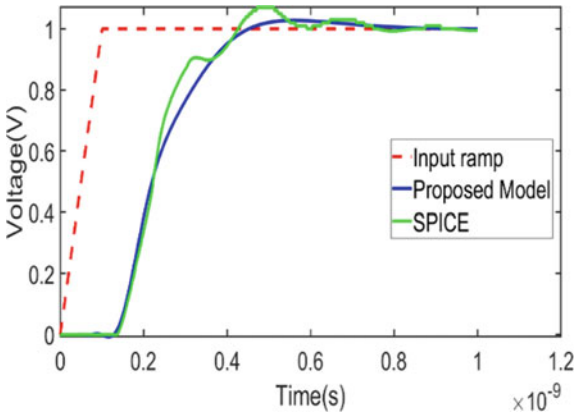


Fig. 2.5 Time domain analysis of second output line, for input Tx line length = 0.5 cm, $R_s = 50\Omega$, $C_l =$ 50fF

Conclusion

The behavioral modeling of asymmetrical T-tree interconnect has been investigated. The delay and overshoot of single input-two outputs 1:2 asymmetrical T-tree interconnect are estimated. Overshoot and 50% delay of 1:2 asymmetrical T-tree interconnect by using proposed Gaver-Stehfest-based model for various lengths, source resistance, and load capacitance of first and second output lines are estimated and respective graphs are plotted and compared with SPICE values. It is investigated that, the proposed Gaver-Stehfest-based model is more accurate and seems to be in good agreement with SPICE values. This research may serve as a base for future studies of crosstalk analysis of various asymmetrical tree branches.

References

Asai H, Tanji Y (2004) Closed-form expressions of distributed RLC interconnects for analysis of on-chip inductance effects. In: Proceedings of the 41st annual Design Automation Conference

Coulibaly LM, Kadim HJ (2005) Analytical crosstalk noise and its induced-delay estimation for distributed RLC interconnects under ramp excitation. In: IEEE International Symposium on Integrated Circuits and Systems

Dounavis A, Achar R, Nakhla M (2001) A general class of passive macromodels for lossy multiconductor transmission lines. IEEE Trans Microw Theory Tech 49:1686–1696

Dounavis A, Achar R, Xin L (1999) Passive closed-form transmissionline model for general-purpose circuit simulators. IEEE Trans Microw Theory Tech 47:2450–2459

Friedman EG, Ismail YI, Neves JL (2000) Equivalent elmore delay for RLC trees. IEEE Transaction Comput-Aided Des Integr Circuits Syst.

Kavicharan M, Murthy NS, Rao NB (2013) An efficient delay estimation model for high speed VLSI interconnects. In: 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)

Kavicharan M, Murthy NS, Rao NB, Prathima A (2015) Modeling and analysis of on-chip single and H-tree distributed RLC interconnects. Circuits Syst Signal Process 35(9):3049–3065

Mehrotra M, Bannerjee K (2002) Analysis of on-chip inductance effects for distributed RLC interconnects. IEEE Trans Comput-Aided Des Integr Circuits Syst

Muddu S, Kahng AB (1997) An analytic delay model for RLC interconnects. IEEE Transacion Comput-Aided Des Integrated Circuits Syst

Noda M, Kobayashi S, Sakurai T (1991) Simple expressions for interconnection delay, coupling and crosstalk in VLSI's. In: Proceedings of IEEE International Symposium on Circuits and Systems

Odabasioglu A, Celik M, Pileggi LT (1998) PRIMA: passive reduced order interconnect macro-modeling algorithm. IEEE Trans Comput-Aided Design Integr Circuits Syst 17: 645–654

Pillage LT, Rohrer RA (1990) Asymptotic waveform evaluation for timing analysis. IEEE Trans Comput-Aided Design Integr Circuits Syst 9: 352–366.

Raveloa B (2015) Modelling of asymmetrical interconnect T-tree laminated on flexible substrate. Eur Phys J Appl Phys 72: 20103

Ravelo B, Vurpillot F, Jastrzebski AK (2014) Asymmetrical PCB interconnect tree modelling with coupling effect. In: Proceedings of 2014 International Symposium on Electromagnetic Compatibility (EMC Europe 2014). Gothenburg, Sweden

Roy S, Dounavis A (2009) Closed-form delay and crosstalk models for RLC on-chip interconnects using a matrix rational approximation. IEEE Trans Comput Aided Des Integr Circuits Syst 28(10):1481–1492

Sakurai T (1993) Closed-form expressions for interconnection delay, coupling, and crosstalk in VLSIs. IEEE Transactions Electron Devices.

Chapter 3
Explicit Power-Delay Models for On-Chip Copper and SWCNT Bundle Interconnects

Yash Agrawal, Vinay Palaparthy, Mekala Girish Kumar, Kavicharan Mummaneni, and Rajeevan Chandel

Abstract In the present chapter, the transient behavior and delay investigation of high-performance interconnects is modeled using finite-difference time-domain (FDTD) technique. Further, an explicit power estimation model for both on-chip conventional copper and futuristic single-wall carbon nanotube (SWCNT) bundle interconnects using FDTD technique is proposed. The model deals with the analysis of two signaling schemes, namely, voltage-mode signaling (VMS) and current-mode signaling (CMS). Power-dissipation, delay, and power_delay_product are the interconnect performance metrics considered. The interconnect is characterized by equivalent single conductor model and CMOS inverter gate is used to drive it. It is investigated that CMS scheme is good for delay-centric designs while for power-centric designs, VMS scheme can be adopted. However, owing to lower power_delay_product, CMS outperforms VMS scheme. Further, it is analyzed that SWCNT bundle

Y. Agrawal (✉) · V. Palaparthy
VLSI and Embedded Systems Research Group, Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, Gujarat 382 007, India
e-mail: yash_agrawal@daiict.ac.in

V. Palaparthy
e-mail: vinay_shrinivas@daiict.ac.in

M. G. Kumar
Department of Electronics and Communication Engineering, TKR College of Engineering and Technology, Hyderabad, Telangana 500 097, India
e-mail: giri.frds@gmail.com

K. Mummaneni
Department of Electronics and Communication Engineering, National Institute of Technology, Silchar, Assam 788 010, India
e-mail: kavicharan@ece.nits.ac.in

R. Chandel
Department of Electronics and Communication Engineering, National Institute of Technology, Hamirpur, Himachal Pradesh 177 005, India
e-mail: rchandel@nith.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_3

interconnects are better in terms of energy efficiency as compared to copper interconnects. The results of the proposed analytical model are validated using SPICE, with a maximum error of less than 3%.

Keywords Current-mode signaling · Energy · Finite-difference time-domain (FDTD) · Interconnects · SWCNT bundle · Voltage-mode signaling

Introduction

The incessant and stupendous advancements in semiconductor technology have led to the giga-scale integration of nano-scale devices and transistors in integrated circuits (ICs) (Kumar et al. 2016). At scaled technologies, local interconnects that connect devices in ICs shrink. However, global interconnects which transmit power and clock signals in an IC, exaggerate in length due to larger chip size and added functionalities on the same chip area for highly miniaturized technologies (Naeemi and Meindl 2007). The operation of interconnects at multi-gigahertz frequency range in nanometer regime causes huge power dissipation in ICs which makes it a leveraging factor in determining the overall system performance.

Traditionally aluminum and consecutively copper have been the choice of the designers as interconnect materials in ICs. However, at scaled dimensions the global interconnects possess high resistance and are more prone to electro-migration, hillock, and void formations (Kumar et al. 2016). Unlike, copper and aluminum, single-wall carbon nanotube (SWCNT) bundle possesses better current capability, longer mean free path of charge carriers, greater thermal conductivity, remarkable mechanical and thermal stability (Naeemi and Meindl 2007). SWCNT bundle has been investigated by researchers and is found to be a suitable replacement to metallic copper and aluminum for on-chip interconnect applications (Naeemi and Meindl 2007; Rai et al. 2014; Zhang et al. 2012; Sahoo et al. 2014).

Apart from incorporating appropriate interconnect materials, employing proper signaling schemes plays a vital part in evaluating the overall system performance in ICs. Current-mode signaling (CMS) scheme is characterized by small voltage swing, has higher speed and data-rate over traditional full voltage rail voltage-mode signaling (VMS) scheme (Katoch et al. 2005; Agrawal et al. 2016; Dave et al. 2013; Venkatraman and Burleson 2007; Agrawal and Chandel 2016; Bashirullah et al. 2003; Tuuna et al. 2012; Lee et al. 2010). Hence, CMS scheme is well matched for high-speed applications in ICs.

The advanced CMS scheme has been relatively lesser explored compared to traditional VMS scheme. To improve the performance of CMS scheme, equalization circuits are used for 5, 10, and 15 mm interconnect lengths (Kim 2010). Equalization is used to minimize inter-symbol interference and throughput. However, equalization circuits lead to large area overhead and also require additional power supply. Moreover, in repeated equalization interconnects, optimum length needs

to be determined to achieve optimum results. The essential benefits of equalization are more explicit at very long wire lengths where added area overhead and power of equalization circuit is marginal in comparison to the overall system performance improvements. Katoch et al. have presented various circuits for current-mode driver and receiver and observed that a factor of three can be gained in delay using CMS over VMS scheme (Katoch et al. 2005). In (Agrawal et al. 2016; Dave et al. 2013; Venkatraman and Burleson 2007), multi-level current-mode signaling scheme has been used which is effective for attaining smaller latency along with higher throughput in the system. In (Agrawal and Chandel 2016; Bashirullah et al. 2003; Tuuna et al. 2012; Lee et al. 2010), CMOS inverter is incorporated as driver for interconnects in CMS scheme. CMOS inverter has simple design, high signal integrity, and consumes less silicon chip area (Lee et al. 2010; Kang and Leblebici 2003). Literature reports about the various research works carried out in copper interconnects using CMS scheme (Katoch et al. 2005; Agrawal et al. 2016; Dave et al. 2013; Venkatraman and Burleson 2007; Agrawal and Chandel 2016; Bashirullah et al. 2003; Tuuna et al. 2012; Lee et al. 2010). However, no substantial work has been informed in SWCNT bundle interconnects and hence needs more research investigations.

Performance of SWCNT bundle interconnects has been investigated by researchers in Naeemi and Meindl (2007); Rai et al. (2014); Zhang et al. (2012); Sahoo et al. (2014). Naeemi et al. have presented physics-related circuit models for SWCNTs (Naeemi and Meindl 2007). In (Rai et al. 2014), it is investigated that crosstalk noise in CNT bundles is significantly lesser than the copper interconnects. In (Zhang et al. 2012; Sahoo et al. 2014), modeling and performance of SWCNT bundle interconnect has been performed using equivalent single conductor model. The work reported in literature on SWCNT bundle interconnects has been performed for VMS scheme. However, modeling and investigation of SWCNT bundle interconnects have not been explored using CMS scheme. Thus, research in this area is still needed.

Modeling the transient behavior, crosstalk noise, signal integrity, and propagation delay for an interconnect system has been performed by various techniques, namely, moment-based (Bashirullah et al. 2003), ABCD (Sahoo et al. 2015), FDTD models (Agrawal and Chandel 2016; Paul 1994; Li et al. 2011; Liang et al. 2012). Moment-based models have limited accuracy as the system poles are approximated by a few dominant poles (Bashirullah et al. 2003). ABCD is based on two-port network model which requires approximation of non-linear CMOS driver gate as lumped capacitive and resistive elements. This approximation causes lesser accurate results (Sahoo et al. 2015). FDTD is a numerical-based technique and beneficial for attaining accurate results (Paul 1994). FDTD-based models have been presented for copper interconnects in Agrawal and Chandel (2016); Paul (1994); Li et al. (2011) and for CNT interconnects in Liang et al. (2012). The models in Li et al. (2011) have incorporated practical CMOS driver for copper interconnects using VMS scheme. This work has been extended to CMS scheme in Agrawal and Chandel (2016). Till date, the models for CNT structures using FDTD technique have used simplified resistive driver model (Liang et al. 2012). However, for accurate analysis, a practical

CMOS driver gate is needed. Hence, further comprehensive model formulation is essential.

The power modeling in interconnects is typically a non-linear function and has been meagerly explored by researchers. The power model for metallic copper interconnects has been presented in Uchino and Cong (2002); Aswatha et al. (2008). These models have used linear resistive driver model and neglected inductive parasitics while modeling the interconnect structures. These approximations lead to less accurate results. Consequently, more accurate power modeling is required.

The above discussed literature reveals various gaps in the current research. The aforementioned issues have been appropriately alleviated in this chapter. The major objectives of the current work are stated as follows. (i) A novel unified model for VMS and CMS schemes along with copper and SWCNT bundle interconnects has been formulated. (ii) Performance evaluation of CNT interconnects with CMS scheme has been innovatively taken up in this chapter. (iii) Most of the research and model formulations are focused toward transient and delay analysis. However, analysis of power dissipation is also significant in an interconnect system. Consequently, in the present work, analytical models for both power and delay have been formulated, using FDTD technique.

This chapter comprises four sections. The present section briefly introduces on-chip interconnects. Section "Model Formulation" presents the novel unified model for copper and SWCNT bundle interconnects. Furthermore, analytical modeling for transient response, delay, and power dissipation have been formulated. The results are presented and discussed in Sect. "Results and Discussion". Finally, concluding remarks are made in Sect. "Conclusion".

Model Formulation

The unified model and interconnect equivalent electrical circuit for copper and SWCNT bundle using VMS and CMS schemes are illustrated in Fig. 3.1. CMOS inverter driver is used in both VMS and CMS schemes. The analysis is performed for global interconnect length of 1 mm. This length is sufficient to provide signal connections in 14 nm technology node (Rai et al. 2014; Sahoo et al. 2014). Hence, buffer insertion for VMS and equalization circuits for CMS are not required in the present analysis. The CMOS inverter characteristics is defined using nth power-law model (Sakurai and Newton 1991). The interconnect is characterized by equivalent single conductor (ESC) model (Amore et al. 2010) and quasi-transverse electromagnetic (TEM) model is used for analytical formulation (Paul 1994). The receiver of the interconnect system is equivalently modeled by a capacitive load (C_L) and a load resistance (R_L). Constant α denotes material of the interconnect. Its value is "1" and "0" for SWCNT bundle and copper interconnects respectively. Parameter β delineates the signaling scheme. It is zero and unity for VMS and CMS schemes respectively. The current (I) and voltage (V) in the distributed sections of interconnect are formulated using TEM model as

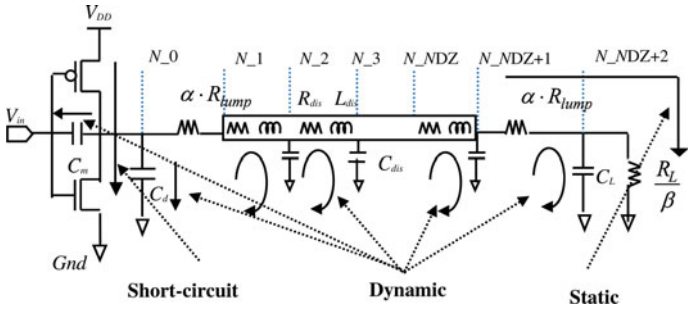


Fig. 3.1 Unified model and equivalent circuit for copper and SWCNT bundle interconnects

$$\frac{\partial V(z,t)}{\partial z} + L_{dis}\frac{\partial I(z,t)}{\partial t} + R_{dis} I(z,t) = 0 \tag{3.1}$$

$$\frac{\partial I(z,t)}{\partial z} + C_{dis}\frac{\partial V(z,t)}{\partial t} = 0 \tag{3.2}$$

where I and V are function of time (t) and position (z) along the interconnect.

Equations (3.1) and (3.2) are discretized in time and position along the interconnect which are solved as

$$\frac{V_{k+1}^{n+1} - V_k^{n+1}}{\Delta z} + L_{dis}\left(\frac{I_k^{n+3/2} - I_k^{n+1/2}}{\Delta t}\right) + R_{dis}\left(\frac{I_k^{n+3/2} + I_k^{n+1/2}}{2}\right) = 0 \tag{3.3}$$

$$\frac{I_k^{n+1/2} - I_{k-1}^{n+1/2}}{\Delta z} + C_{dis}\frac{V_k^{n+1} - V_k^n}{\Delta t} = 0 \tag{3.4}$$

On solving (3.3) and (3.4), recursive expressions for current and voltage are derived and given as below

$$I_k^{n+3/2} = BDI_k^{n+1/2} + B(V_k^{n+1} - V_{k+1}^{n+1}) \tag{3.5}$$

where

$$B = \left(\frac{\Delta z}{\Delta t}L_{dis} + \frac{\Delta z}{2}R_{dis}\right)^{-1} \tag{3.6}$$

and

$$D = \left(\frac{\Delta z}{\Delta t}L_{dis} - \frac{\Delta z}{2}R_{dis}\right) \tag{3.7}$$

where k in (3.5), (3.6) and (3.7) ranges from 1, 2, 3,….., *NDZ* while n denotes a positive integer.

$$V_k^{n+1} = V_k^n + A(I_{k-1}^{n+1/2} - I_k^{n+1/2}) \tag{3.8}$$

where

$$A = \left(\frac{\Delta t}{\Delta z}.\frac{1}{C_{dis}}\right) \tag{3.9}$$

for $k = 2, 3, 4,\ldots,$ *NDZ* and n is a positive integer.

The discretized current and voltage varying with position and time can be represented as

$$V_i^n = V[(i)\Delta z, (n)\Delta t] \tag{3.10}$$

$$I_i^{n+1//2} = I[(i + 1/2)\Delta z, (n + 1/2)\Delta t] \tag{3.11}$$

where i ranges as 1, 2, 3,….., *NDZ* while n denotes a positive integer.

The near-end, recursive and far-end current and voltage along interconnect are evaluated according to Fig. 3.2. The boundary condition associated with near-end case is defined by CMOS inverter. It is evaluated at nodes *N*_0 and *N*_1. The boundary condition associated with far-end is defined by load capacitance (C_L) and load resistance (R_L). These are computed at nodes *N_NDZ* + 1 and *N_NDZ* + 2. The current and voltage at the intermediate nodes are evaluated alternatively and recursively at $\Delta t/2$ period apart in temporal domain and $\Delta z/2$ distance apart in spatial domain. The Courant condition ascertains the stability of FDTD model which is defined as $\Delta t \leq \Delta z \big/ v$ (Paul 1994), where v denotes the signal velocity in interconnects.

The boundary condition associated with near-end case is defined at node *N*_0 by variables V_0 and I_0. The current I_0 for different time instant is evaluated using Kirchhoff's current law (KCL) at node *N*_0 as

$$I_0^{n+1} = I_p^{n+1} - I_n^{n+1} + C_m\left(\frac{V_{in}^{n+1} - I_{in}^n}{\Delta t}\right) + (C_m + C_d)\left(\frac{V_0^{n+1} - V_0^n}{2}\right) \tag{3.12}$$

where I_p^{n+1} and I_n^{n+1} denote PMOS and NMOS currents, respectively. These are obtained using nth power law model (Sakurai and Newton 1991).

Using Ohm's law, V_0^{n+1} is investigated and given as

$$V_0^{n+1} = V_1^{n+1} + \alpha R_{lump}{}' \cdot I_0^{n+1} \tag{3.13}$$

where V_1^{n+1} in (3.13) is derived by substituting k equal *N*_1 in (3.5). Since I_0 and I_1 are separated by distance $\Delta z/2$, therefore, Δz is replaced as $\Delta z/2$ in parameter A. V_1^{n+1} is obtained as

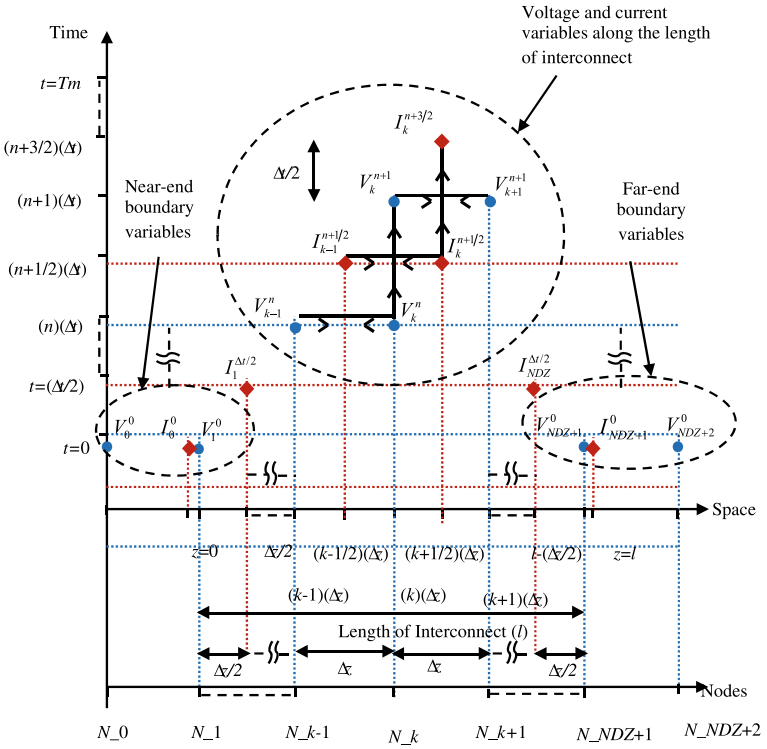


Fig. 3.2 Representation of current and voltage in spatial, temporal, and nodal domains

$$V_1^{n+1} = V_1^n + 2A(I_0^{n+1/2} - I_1^{n+1/2}) \tag{3.14}$$

where

$$I_0^{n+1/2} = \frac{I_0^n + I_0^{n+1}}{2} \tag{3.15}$$

Using (3.12) and (3.13), the equivalent discretized current at node N_0 is formulated as

$$I_0^{n+1} = \begin{pmatrix} I_p^{n+1} - I_n^{n+1} + C_m\left(\dfrac{V_{in}^{n+1} - V_{in}^n}{\Delta t}\right) \\ -(C_m + C_d)\left(\dfrac{V_1^{n+1} - V_1^n}{\Delta t}\right) \\ +\alpha R_{lump}{}'(C_m + C_d)\left(\dfrac{I_0^n}{\Delta t}\right) \end{pmatrix} \tag{3.16}$$

where

$$H = \left(1 + \frac{\alpha R_{lump}'(C_m + C_d)}{\Delta t}\right)^{-1} \tag{3.17}$$

Using (3.14), (3.15), (3.16) and (3.17), V_1 is further modified as

$$V_1^{n+1} = \begin{pmatrix} V_1^n + EAH\left[C_m\left(\frac{V_{in}^{n+1} - V_{in}^n}{\Delta t}\right) + \left(\frac{1}{H} + \frac{\alpha R_{lump}'(C_m + C_d)}{\Delta t}\right) I_0^n\right] \\ -2EAI_1^{n+1} + EAH(I_p^{n+1} - I_n^{n+1}) \end{pmatrix} \tag{3.18}$$

where

$$E = \left(1 + \frac{AH(C_m + C_d)}{\Delta t}\right)^{-1} \tag{3.19}$$

Similarly, boundary condition associated with far-end case is defined at node $N\_NDZ + 2$ by variables V_{NDZ+2} and I_{NDZ+1}. I_{NDZ+1} for varying time instant is obtained using KCL at the node $N\_NDZ + 2$:

$$I_{NDZ+1}^{n+1} = C_L\left(\frac{V_{NDZ+2}^{n+1} - V_{NDZ+2}^n}{\Delta t}\right) + \beta\left(\frac{V_{NDZ+2}^{n+1}}{R_L}\right) \tag{3.20}$$

At far-end node $N\_Nz + 2$, using Ohm's law the obtained voltage is

$$V_{NDZ+2}^{n+1} = V_{NDZ+1}^{n+1} - \alpha R_{lump}' \cdot I_{NDZ+1}^{n+1} \tag{3.21}$$

V_{NDZ+1}^{n+1} in (3.21) is derived by substituting $k = NDZ + 1$ in (3.5) and Δz by $\Delta z/2$ in A. This gives

$$V_{NDZ+1}^{n+1} = V_{NDZ+1}^n + 2A\left(I_{NDZ}^{n+1/2} - I_{NDZ+1}^{n+1/2}\right) \tag{3.22}$$

where

$$I_{NDZ+1}^{n+1/2} = \frac{I_{NDZ+1}^n + I_{NDZ+1}^{n+1}}{2} \tag{3.23}$$

Using (3.20), (3.21) is further formulated as

$$I_{NDZ+1}^{n+1} = J\left(\begin{array}{l}\left(\dfrac{\alpha R_{lump}{}'C_L}{\Delta t}\right)I_{NDZ+1}^{n} + \left(\dfrac{C_L}{\Delta t} + \dfrac{\beta}{R_L}\right)\left(V_{NDZ+1}^{n+1}\right) \\ -\left(\dfrac{C_L}{\Delta t}\right)\left(V_{NDZ+1}^{n}\right)\end{array}\right) \tag{3.24}$$

where

$$J = \left(1 + \frac{\alpha R_{lump}{}'C_L}{\Delta t} + \frac{\alpha\beta R_{lump}{}'}{R_L}\right)^{-1} \tag{3.25}$$

V_{NDZ+1}^{n+1} in (3.22) is further solved as

$$V_{NDZ+1}^{n+1} = FGV_{NDZ+1}^{n} + 2FAI_{NDZ}^{n+1/2} - FA\left(1 + \frac{\alpha J R_{lump}{}'C_L}{\Delta t}\right)I_{NDZ+1}^{n} \tag{3.26}$$

where

$$F = \left(1 + \frac{AJC_L}{\Delta t} + \frac{\beta AJ}{R_L}\right)^{-1} \tag{3.27}$$

and

$$G = \left(1 + \frac{AJC_L}{\Delta t}\right) \tag{3.28}$$

The major components of power-dissipation in interconnects are dynamic, static, and short-circuit. The charging and discharging of interconnect load capacitance leads to dynamic-power dissipation. At time value n, dynamic power is given as

$$P_{dyn}^{n} = \alpha \cdot f\left[(C_m + C_d)\left(V_0^n\right)^2 + (C_L)\left(V_{NDZ}^n\right)^2 + \sum_{k=2}^{NDZ+1}(C_k)\left(V_k^n\right)^2\right] \tag{3.29}$$

where f denotes frequency of signal and α represents switching activity factor.

The static power dissipation at time instant n occurs due to static current flow through the interconnect and load resistance. It is formulated as

$$P_{stat}^{n} = \frac{\beta(V_{DD})^2}{2\left[R_L + \beta \cdot NDZ \cdot R_{dis} + 2\beta \cdot \alpha \cdot R_{lump}{}' + \beta\left(\frac{V_{DD}-V_0^n}{I_p^n}\right)\right]} \tag{3.30}$$

As β approaches zero in (3.30), the static component becomes zero.

The short-circuit power dissipation $\left(P_{sc}^{n}\right)$ at varying time instant n occurs owing to direct flow of current between supply voltage (V_{DD}) and ground. It is given as fraction (x) of total dynamic-power dissipation (Gowan et al. 1998),

$$P_{sc}^{n} = x\, P_{dyn}^{n} \tag{3.31}$$

where x varies from 0.1 to 0.2.

Results and Discussion

Copper and SWCNT bundle interconnects are analyzed using VMS and CMS schemes for 14 nm technology node (International Technology Roadmap for Semiconductors 2012; Predictive Technology Models 2015). The analyses are performed using the proposed FDTD model and are validated using SPICE (2015). The interconnect parasitics for SWCNT bundle are determined using the formulations presented in Naeemi and Meindl (2007); Rai et al. (2014); Zhang et al. (2012); Sahoo et al. (2014). The interconnect dimensions are taken according to ITRS (International Technology Roadmap for Semiconductors 2012). The SWCNT bundle interconnect parasitics, viz., R_{dis}, $C_{dis,}$ and L_{dis} are computed as 16.67 MΩ/m, 5.45 pF/m, and 10.4 μH/m, respectively. The imperfect contact resistance per SWCNT is 6.5KΩ. R_{lump} comprises imperfect contact resistance and quantum resistance. R_{lump} for the bundled structure is 27.87Ω. R_{dis}, $C_{dis,}$ and L_{dis} for copper interconnect are 23.50 MΩ/m, 25.66 pF/m, and 2.16 μH/m, respectively (Predictive Technology Models 2015). R_{lump} is zero for copper interconnect as the contact resistance between the substrate and copper is negligible (Kumar et al. 2016). The load capacitance (C_L) and load resistance (R_L) for CMS scheme are 1KΩ and 0.5fF, respectively. C_L for VMS scheme is 1fF (Agrawal and Chandel 2016). The terminal load in VMS is very large and hence R_L is considered as infinity (Bashirullah et al. 2003).

The nodal voltages and branch currents computed by (3.5)–(3.28) are used for transient analysis and propagation delay computation. Equations (3.29), (3.30) and (3.31) are used to evaluate power dissipation in the interconnect system. Variations in power dissipation, delay, and power_delay_product (PDP) with interconnect length are illustrated in Figs. 3.3, 3.4, and 3.5, respectively. The interconnect length is changed from 200 to 1000 μm for experimentation.

Figure 3.3 depicts delay in both the interconnect materials using VMS and CMS schemes. It is observed from the figure that, irrespective of the interconnect material and signaling schemes, propagation delay increases with interconnect length, in all the cases. This is so, owing to higher parasitics of interconnect for longer wire lengths. Further, it is seen that propagation delay is smaller in CMS scheme over VMS scheme. For example, at interconnect length of 1000 μm, percentage reduction in delay using CMS over VMS scheme with SWCNT bundle interconnect is 139.8%. This is nearly 170.5% for copper interconnects. The smaller propagation delay in CMS scheme is due to reduced voltage swing which allows faster charging and

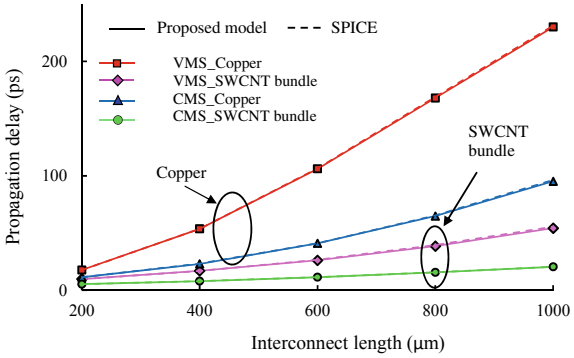


Fig. 3.3 Variation of propagation delay with length of interconnect

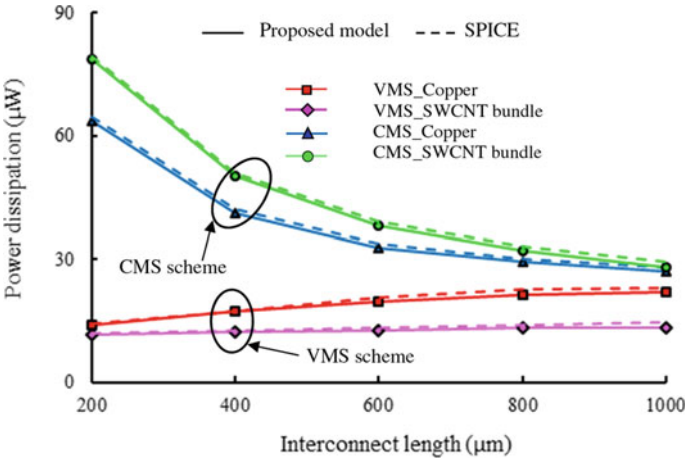


Fig. 3.4 Variation of power dissipation with interconnect length

discharging of capacitances associated at interconnect node. It is also observed that SWCNT bundle interconnects have benefit of smaller delay compared to copper interconnects in both the signaling schemes.

The variations of power dissipation for all the considered signaling schemes and interconnect materials have been presented in Fig. 3.4. The input is a pulse signal of width 1 ns ($f = 1$ GHz) and signal rise/fall time of 50 ps. The activity factor (α) is unity. It is seen from the figure that, with escalation in length of the interconnect, power-dissipation in VMS scheme surges nominally, while it reduces for CMS scheme. However, it is also analyzed that power dissipation is higher in CMS scheme. This occurs due to high static power dissipation in CMS scheme. Further, it is seen that the results of the proposed FDTD-based model are in good agreement with SPICE, with an average percentage error of 2.3%.

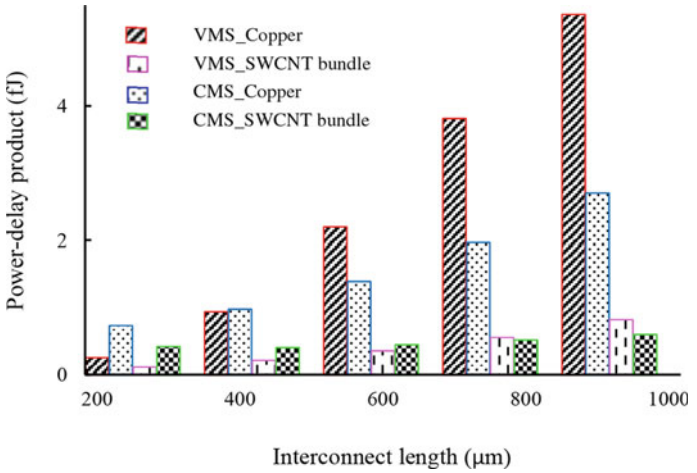


Fig. 3.5 Power_delay_product variation with interconnect length

From Figs. 3.3 and 3.4, a trade-off is observed between power dissipation and propagation delay for VMS and CMS schemes. The efficacy between the VMS and CMS schemes can be investigated by analyzing PDP. PDP gives the total energy dissipated in the system and hence is a vital figure of merit.

A lower value of PDP implies low energy dissipation and is therefore desirable for high-performance applications. From Fig. 3.5, it is seen that CMS has advantage of having lower PDP over VMS scheme. Further, it is inferred that SWCNT bundle interconnect outperforms its counterpart conventional copper interconnect. For instance, at wire length of 1000 μm, PDP in copper interconnect is 6.5 times higher than SWCNT bundle interconnect for VMS scheme. This is about 4.5 times in case of CMS scheme.

Figure 3.6 shows the percentage error in analytical model with respect to SPICE for PDP at different interconnect length. It is observed that maximum percentage error between the two is within 3% for varying interconnect structures and different signaling schemes.

The comparative analysis of computational efficiency of the FDTD-based model with respect to SPICE is illustrated in Fig. 3.7. The analysis is performed on Intel Pentium Dual-Core processor running at 1.73 GHz with 1.5 GB RAM PC and Windows 7 operating system. For the analysis, the interconnect line segmentation is varied from 1 to 1000. The number of time segment is kept as 1000. It is analyzed that the computational runtime is significantly lesser in FDTD technique. The average percentage reduction in CPU runtime using FDTD over SPICE is nearly 64%.

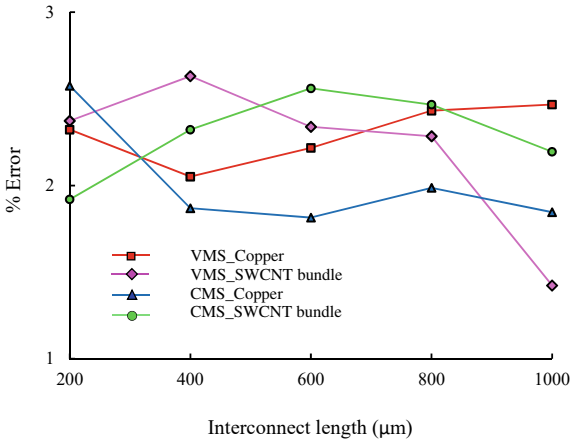


Fig. 3.6 Percentage error in analytical model with respect to SPICE for PDP at different interconnect length

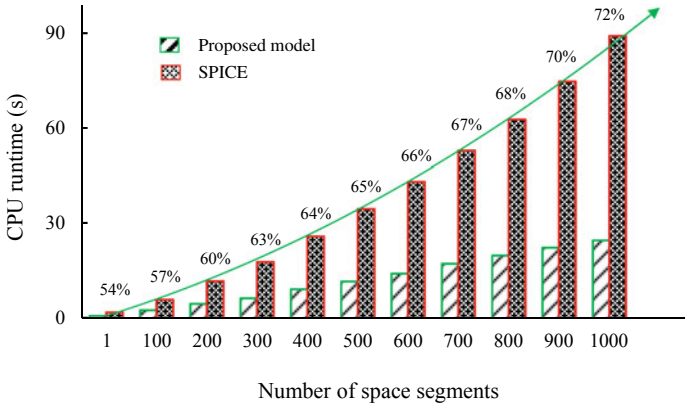


Fig. 3.7 Comparative analysis of CPU runtime of the proposed analytical model with respect to SPICE

Conclusion

The chapter presents a unified power-dissipation model for SWCNT bundle and copper interconnects using VMS and CMS schemes. Concurrently, the model is applicable for evaluation of propagation delay and hence power_delay_product in interconnects. It is analyzed that VMS is efficient for power-centric designs. The percentage reduction in dissipated power in VMS over CMS scheme at 1000 μm for SWCNT bundle interconnect is nearly 98% while this is 20.8% for copper interconnect. Further it is seen that CMS is good for delay-centric designs. Delay in CMS scheme is nearly 170.6% lesser over VMS for SWCNT bundle interconnect. This is

nearly 139.9% for copper interconnect. The analysis also reveals that PDP is lower in CMS scheme. The reduction in PDP in CMS over VMS for SWCNT bundle and copper interconnects is 6.5 and 4.5 times, respectively. Thus, CMS has higher efficacy over VMS scheme. It is also investigated that, irrespective of signaling schemes, SWCNT bundle outperforms copper interconnect. Thus, it may be summarized that SWCNT interconnects along with CMS scheme are delay and energy efficient and hence a better solution for high-performance on-chip interconnects. The present study shall be beneficial for VLSI designers and further research work in this area.

References

Agrawal Y, Chandel R (2016) Crosstalk analysis of current-mode signaling-coupled RLC interconnects using FDTD technique. IETE Tech Rev 33(2):148–159. https://doi.org/10.1080/02564602.2015.1056258

Agrawal Y, Kumar MG, Chandel R (2016) Comprehensive model for high-speed current-mode signaling in next generation MWCNT bundle interconnect using FDTD technique. IEEE Trans Nanotechnol 15(4):590–598. https://doi.org/10.1109/TNANO.2016.2558475

Amore MD, Sarto MS, Tamburrano A (2010) Fast transient analysis of next-generation interconnects based on carbon nanotubes. IEEE Trans Electromagn Compat 52(2):496–503. https://doi.org/10.1109/TEMC.2010.2045383

Aswatha R, Basavaraju T, Kalpana AB (2008) Efficient power modeling for on-chip global interconnects. In: Proceedings of IEEE Midwest Symposium on Circuits and Systems, pp. 458–46. https://doi.org/10.1109/MWSCAS.2008.4616835

Bashirullah R, Liu W, Cavin RK (2003) Current-mode signaling in deep submicrometer global interconnects. IEEE Trans Very Large Scale Integr Syst 11(3):406–417. https://doi.org/10.1109/TVLSI.2003.812366

Dave M, Jain M, Baghini MS, Sharma D (2013) A variation tolerant current mode signaling scheme for on-chip interconnects. IEEE Trans Very Large Scale Integr Syst 21(2):342–353. https://doi.org/10.1109/TVLSI.2012.2185835

Gowan MK, Biro LL, Jackson DB (1998) Power considerations in the design of the alpha 21264 microprocessor. In: Proceedings of IEEE Design Automation Conference, pp. 726–731. https://doi.org/10.1145/277044.277226

International Technology Roadmap for Semiconductors (2012) http://public.itrs.net

Kang SM, Leblebici Y (2003) CMOS digital integrated circuits, TMH, New Delhi

Katoch A, Veendrick H, Seevinck E (2005) High speed current-mode signaling circuits for on-chip interconnects. In: Proceedings of IEEE International Symposium on Circuits and Systems, vol 4, pp 4138−4141. https://doi.org/10.1109/ISCAS.2005.1465542

Kim (2010) Equalized on-chip interconnect: modeling, analysis, and design. Ph.D. Dissertation, Massachusetts Institute of Technology. https://www.rle.mit.edu/isg/documents/BSKimPhDThesis.pdf

Kumar MG, Agrawal Y, Chandel R (2016) Carbon nanotube interconnects–a promising solution for VLSI circuits. IETE J Educ 57(2):46–64. https://doi.org/10.1080/09747338.2016.1158129

Lee K, Qazi M, Kong J, Chandrakasan AP (2010) Low-swing signaling on monolithically integrated global graphene interconnects. IEEE Trans Electron Devices 57(12):3418–3425. https://doi.org/10.1109/TED.2010.2083667

Li XC, Mao JF, Swaminathan M (2011) Transient analysis of CMOS-gate-driven *RLGC* interconnects based on FDTD. IEEE Trans Comput-Aided Des 30(4):574–583. https://doi.org/10.1109/TCAD.2010.2095650

Liang F, Wang G, Lin H (2012) Modeling of crosstalk effects in multiwall carbon nanotube interconnects. IEEE Trans Electromagn Compat 54(1):133–139. https://doi.org/10.1109/TEMC.2011.2172982

Naeemi A, Meindl JD (2007) Design and performance modeling for single-walled carbon nanotubes as local, semiglobal, and global interconnects in gigascale integrated systems. IEEE Trans Electron Devices 54(1):26–37. https://doi.org/10.1109/TED.2006.887210

Paul R (1994) Incorporation of terminal constraints in the FDTD analysis of transmission lines. IEEE Trans Electromagn Compat 36(2):85–91. https://doi.org/10.1109/15.293284

Predictive Technology Models (2015). http://ptm.asu.edu

Rai MK, Khanna R, Sarkar S (2014) Crosstalk analysis in CNT bundle interconnects for VLSI applications. IEEJ Trans Electr Electron Eng 9(4):391–397. https://doi.org/10.1002/tee.21984

Sahoo M, Ghosal P, Rahaman H (2014) Performance modeling and analysis of carbon nanotube bundles for future VLSI circuit applications. J Comput Electron 13(3):673–688. https://doi.org/10.1007/s10825-014-0587-7

Sahoo S, Ghosal P, Rahaman H (2015) Modeling and analysis of crosstalk induced effects in multiwalled carbon nanotube interconnects: An ABCD parameter based approach. IEEE Trans Nanotechnol 14(2):259–274. https://doi.org/10.1109/TNANO.2014.2388252

Sakurai T, Newton AR (1991) A simple MOSFET model for circuit analysis. IEEE Trans Electron Devices 38(4):887–894. https://doi.org/10.1109/16.75219

Tanner EDA tools for SPICE simulation (2015) http://www.tannereda.com

Tuuna S, Nigussie E, Isoaho J, Tenhunen H (2012) Modeling of energy dissipation in RLC current-mode signalling. IEEE Trans Very Large Scale Integr Syst 20(6): 1146−1151. https://doi.org/10.1109/TVLSI.2011.2140345

Uchino T, Cong J (2002) An interconnect energy model considering coupling effects. IEEE Trans Comput-Aided Des 21(7):763–776. https://doi.org/10.1145/378239.379022

Venkatraman V, Burleson W (2007) An energy efficient multi-bit quaternary current-mode signaling for on-chip interconnects. In: Proceedings of IEEE Custom Integrated Circuits Conference, pp. 301–304. https://doi.org/10.1109/CICC.2007.4405738

Zhang K, Tian B, Wang F, Wei J (2012) Crosstalk analysis of carbon nanotube bundle interconnects. Nanoscale Res Lett 7(1):1–5. https://doi.org/10.1186/1556-276X-7-138

Chapter 4
Modelling and Analysis of Copper and Carbon Nanotube VLSI Interconnects

Kavicharan Mummaneni, Vivek Kumar, Malvika, and P. Uma Sathyakam

Abstract In this chapter, a numerical inverse Laplace Transform-based model is presented to estimate the delay and crosstalk noise of Copper (Cu) and Carbon Nanotube (CNT) interconnects. The proposed model is formulated on the Gaver-Stehfest method, which provides a relation between the time domain function and linear combination of the transfer function values. The values of delay and overshoot in single Cu and CNT lines are estimated for different line lengths, load capacitance and source resistances. The proposed model is utilized to calculate crosstalk delay, peak noise voltage and timing instance of peak voltage. All the estimated parameters are compared with the existing Padé model and SPICE and are found to be in good agreement. The 50% delay, overshoot and crosstalk results of Cu interconnects match to SPICE results with 0.37%, 0.65% and 0.58%, respectively. In CNT interconnects, the average error of 50% delay, overshoot and crosstalk of the proposed model are 1.2%, 0.06% and 0.22%, respectively.

Keywords Cu & CNT interconnects · Laplace transform · Dynamic crosstalk · Pade model · FDTD method · LTSPICE

Introduction

In VLSI technology, the overall performance of the chip significantly relies on interconnect delay in deep submicron technology (DSM). Aggressive scaling of interconnects has led to significant coupling capacitance among adjacent interconnects (Srivastava and Banerjee 2005). Capacitive and inductive coupling between

K. Mummaneni (✉) · V. Kumar · Malvika
Electronics and Communication Engineering Department, National Institute of Technology, Silchar, Assam, India
e-mail: kavicharan@ece.nits.ac.in

P. Uma Sathyakam
School of Electrical Engineering, Vellore Institute of Technology, Vellore, India
e-mail: umasathyakam.p@vit.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_4

contiguous lines results in crosstalk noise which impacts the reliability and signal integrity of circuits. Thus, an accurate computational model is required for determining the interconnect delay and crosstalk (Li, 2008). Continuous decrease in size of electronic devices accompanied with increase in lengths of interconnects in VLSI circuits has led to the research of new materials for interconnects such as Carbon Nanotubes (CNTs). CNTs are graphene sheets folded into cylinders having diameter of the order of a nanometer. There are mainly two types of CNTs which will have higher structural perfection. Single-walled CNTs (SWCNTs) contain single graphite sheet continuously encased into a cylindrical tube. Multi-walled CNTs (MWCNTs) contain an array of such concentrically nested nanotubes. Because of the unique physical properties of Carbon Nanotubes such as high current carrying capability, higher mechanical strength, CNTs have caught the eye of researchers in modern times (Elmore 1948). In order to use CNTs as interconnects, comprehensive investigation for various CNT structures need to be carried out.

Only resistance and capacitance are used in Elmore delay model to calculate the interconnect delay (Elmore 1948). The predictions of Elmore delay model are accurate only when the inductance value is very small since for small inductance, the step response of the circuit will be independent of the effect of inductance. To obtain better accuracy in high-speed VLSI circuits, the effect of inductance needs to be included along with them. Also, the signal integrity losses are due to the distributed nature of parasitic resistance, capacitance and inductance.

In practical VLSI circuits, the load is capacitive in nature (Ravindra and Srinivas 2007a). The driver's series impedance may not necessarily match with the characteristic impedance of the transmission line. This results in overshoots or undershoots of the voltage waveforms. The circuit becomes less reliable because of the overshoots produced in the output voltage level and may cause glitches or false transitions. Therefore, accurate estimation of overshoots is necessary in DSM technology.

Earlier, the delay and crosstalk phenomenon were estimated by several analytical models Sahoo and rahaman (2013), Abate and Whitt (2006), Agrawal et al. (2017), Das and Rahaman (2011), Fathi and Forouzandeh (2012), Dounavis et al. (2001), (1999), Elmore (1948), Fathi and Forouzandeh (2009), Fathi et al. (2009), Haji-Nasiri et al. (2012), Hermite (1908). This chapter introduces an analytical model for single and coupled Cu and CNT lines using a numerical Laplace inversion method. The proposed model is based on the Gaver-Stehfest Algorithm (Abate and Whitt 2006). The Gaver-Stehfest algorithm can be put into the unified framework for numerical inversion of Laplace transform. The result of the inverse transform depends only on the number of weights used; it does not depend on the Laplace Transform or the independent variable, time t. The unified framework is derived by approximating the exponential function by a rational function, which is further represented as a partial fraction (Abate and Whitt 2006). The proposed model is easy to implement. The time domain analysis for single, coupled Cu and CNT interconnects are performed by applying the proposed model. The results of the proposed model are compared with existing Padé model (Roy and Dounavis 2009) and SPICE.

Analysis of Cu and CNT Interconnects

A. Analysis of Single Cu Interconnect

In this type of interconnect, single line Cu based RLC is considered. The single Cu-RLC line is described by Telegrapher's equations which are given as

$$\frac{\partial}{\partial z}V(z,t) + L\frac{\partial}{\partial t}I(z,t) + RI(z,t) = 0$$

$$\frac{\partial}{\partial z}I(z,t) + C\frac{\partial}{\partial t}V(z,t) = 0 \tag{4.1}$$

where V and I represent transmission line voltage and current, respectively and R, L, C are per unit length resistance (Ω/m), inductance (H/m) and capacitance (F/m), respectively. Equation (4.1) can be expressed in the frequency domain as

$$\frac{\partial}{\partial z}V(z,s) + (R + sL)I(z,s) = 0$$

$$\frac{\partial}{\partial z}I(z,s) + sCV(z,s) = 0 \tag{4.2}$$

The solution of (4.2) is written in exponential matrix form (Roy and Dounavis 2009) as

$$e^{\varnothing l} = \begin{bmatrix} \cosh(\sqrt{ZYl}) & Z_o\sinh(\sqrt{YZl}) \\ Y_o\sinh(\sqrt{YZl}) & \cosh(\sqrt{ZYl}) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{4.3}$$

where

$$Z_o\sqrt{(YZ)/Y} \; and \; Y_o = Y/\sqrt{YZ} \tag{4.4}$$

$$\varnothing = \begin{bmatrix} 0 & Z \\ Y & 0 \end{bmatrix}$$

B. Delay and Crosstalk Model

The circuit model for a single Cu line is shown in Fig. 4.1. It depicts a point-to-point interconnection model where the input voltage is connected to a source resistance R_S and output has a load capacitance C_l. The final transfer function using ABCD matrix is

$$\begin{bmatrix} V_i(s) \\ V_i(s) \end{bmatrix} = \begin{bmatrix} 1 & R_s \\ 0 & 1 \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ sC_t & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ sC_t & 1 \end{bmatrix}\begin{bmatrix} V_o(s) \\ V_o(s) \end{bmatrix} \tag{4.5}$$

Fig. 4.1 Circuit model for single line interconnect

Simplifying Eq. (4.5) results in the frequency-time domain solution as

$$V_o(s) = H(s)V_i(S) \tag{4.6}$$

where

$$H(S) = (A_{11} + R_s A_{21} + sA_{12}C_1 + sR_s A_{22}C_1)^{-1} \tag{4.7}$$

C. Proposed Approximation Technique

The Gaver-Stehfest algorithm (Abate and Whitt 2006) is one among the precise inversion algorithms within the unified framework used for inversion of Laplace transforms numerically. The unified framework is derived by approximating the exponential function by a rational function, which is further expressed as a partial fraction. The unified framework for Laplace inversion is based on the formula,

$$f(t) \approx f_n(t) = \frac{1}{t}\sum_{k=0}^{n} \omega_k f'\left(\frac{\alpha_k}{t}\right), 0 < t < \infty \tag{4.8}$$

where ω_k and α_k are weights and nodes respectively and are complex numbers. They are independent of transform or argument t. They rely upon n. The nodes and weights are obtained from

$$e^z \approx \sum_{k=0}^{n} \frac{\omega_k}{\alpha_k - z}$$

The Gaver-Stehfest inversion formula for a Laplace transform *F(s)* is given as

$$f(t) = \frac{\ln(2)}{t}\sum_{k=0}^{2M} \zeta_K F\left(\frac{k\ln(2)}{t}\right) \tag{4.9}$$

where

$$\zeta_K = (-1)^{M+k} \sum_{j=\left[\frac{(k+1)}{2}\right]}^{min\{k,M\}} \frac{j^{M+1}}{M!}\binom{M}{j}\binom{2j}{j}\binom{j}{k-j} \tag{4.10}$$

Unlike the exponential matrix approximation techniques such as Pade Approximation which requires symbolic expressions to find poles and residues, the Gaver-Stehfest inversion formula works using only numerical calculations. The precision and significant digits produced by the above algorithm is directly proportional to M. For all the simulations, M = 10 has been considered.

For discrete time calculation, we consider $\boldsymbol{t}$, $\boldsymbol{\zeta_k}$ and $\boldsymbol{k}$ as vectors of dimensions $1 \times T$, $1 \times 2M$ and $1 \times 2M$, respectively.

$$\frac{\ln(2)}{t} = \left[\frac{\ln(2)}{t_1}, \frac{\ln(2)}{t_2}, \ldots . \frac{\ln(2)}{t_T}\right] \tag{4.11}$$

$$\zeta_K = [\zeta_1, \zeta_2, \ldots . \zeta_{2M}] \; where \; k = 1, 2, 3, \ldots 2M \tag{4.12}$$

For single interconnect line with output function in frequency domain is

$$\mathrm{V_o(s)} = \mathrm{H(s)V_i(s)} \tag{4.13}$$

The output obtained from (4.9) expressed in matrix form is

$$V_0 = \frac{\ln(2)}{t} o \zeta_K V_0 \left(K^T \frac{\ln(2)}{t}\right) \tag{4.14}$$

where ○ represents element-wise multiplication of two matrices of the same dimension.

D. **Analysis for N-Coupled Cu Interconnect**

Here, N distributed coupled RLC lines are considered and the circuit model is shown in Fig. 4.2. The voltages and currents are in $N\,X\,1$ matrix form. The R, L, C are also in $N\,X\,N$ matrix form. For N-coupled lines, in (4.3), the exponential matrix is further divided into four submatrices A_{11}, A_{12}, A_{21} and A_{22} each of dimensions $N\,X\,N$. The source resistance and load capacitance matrix is $R_S = diag(R_{S1}, R_{S2}, \ldots R_{SN})$ and $C_l = diag(C_{l1}, C_{l2}, \ldots C_{lN,})$, respectively. For N-coupled interconnects, the output voltage can be represented in matrix form as

$$\begin{bmatrix} V_{o1}(s) \\ V_{o2}(s) \\ . \\ . \\ V_{oN}(s) \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \ldots & H_{1N} \\ H_{21} & H_{22} & \ldots & H_{21} \\ & & . & \\ & & . & \\ & & . & \\ & & . & \\ H_{N1} & H_{N2} \ldots & H_{NN} \end{bmatrix} \begin{bmatrix} V_{i1}(s) \\ V_{i2}(s) \\ . \\ . \\ V_{iN}(s) \end{bmatrix} \tag{4.15}$$

$$V_{0i} = \frac{\ln(2)}{t} o \zeta_K V_{0i} \left(K^T \frac{\ln(2)}{t}\right) \tag{4.16}$$

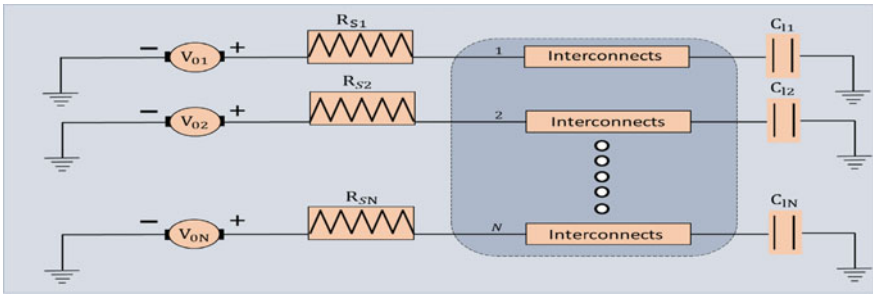


Fig. 4.2 Circuit model for n-coupled line interconnects

The output from (4.15) is found by modifying (4.14) for the *i*th interconnect line as an approximate inverse Laplace Transform which has been derived using the Gaver-Stehfest algorithm. This approximated output is used for the calculation of interconnect delay, overshoot and crosstalk induced noise of Cu and CNT RLC lines.

E. Analysis of Single Line CNT Interconnects

The transfer functions of the SWCNT (Single-Walled) and MWCNT (Multi-Walled) interconnect distributed lines are extracted to find the 50% delay and overshoot at different interconnect lengths by employing the equivalent single conductor CNT interconnect in the Driver-Interconnect-Load system (Mekala Girish Kumar 2017). Figure 4.3a assumes the single line distributed CNT interconnect as an equivalent single conductor.

The system takes into account the driver parasitics, i.e. resistance (R_d) and capacitance (C_d) as shown in Fig. 4.3a. In order to find the transfer function, the DIL system is considered as a cascaded connection of two-port networks as shown in Fig. 4.3b.

The two-port system is represented by ABCD parameters. The transfer function has been realized using the transmission matrix

$$\begin{bmatrix} V_i(s) \\ I_i(s) \end{bmatrix} = \begin{bmatrix} 1 & R_d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ sC_d & 1 \end{bmatrix} \begin{bmatrix} 1 & R_{lump} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} 1 & R_{lump} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ sC_1 & 1 \end{bmatrix} \begin{bmatrix} V_o(s) \\ I_o(s) \end{bmatrix} \tag{4.17}$$

Simplifying (4.16) gives

$$\mathrm{V_o(s) = H(s)V_i(s)} \tag{4.18}$$

where

$$\mathrm{H(s) = \left(\left(1 + sC_1R_{lump}\right) DA_{11} + sC_{1D}A_{12} + \left(1 + sC_1R_{lump}\right) Y\, A_{21} + sC_1 YA_{22}\right)^{-1}} \tag{4.19}$$

Here,

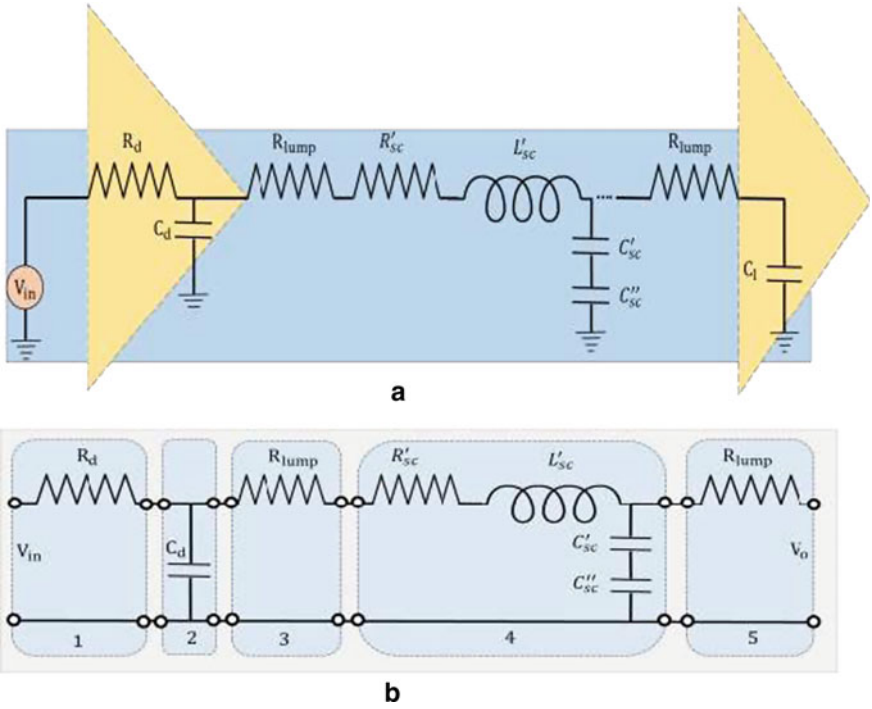


Fig. 4.3 **a** Driver-Interconnect-load system. **b** Cascaded connections of distributed interconnect load (DIL) of 3a

$$D = 1 + sC_d R_d$$

$$\mathrm{Y} = (1 + \mathrm{sC_dR_d})\mathrm{R_{lump}} + \mathrm{R_d} \tag{4.20}$$

F. **Crosstalk Analysis of CNT Interconnect**

The two-coupled line CNT interconnect structure is shown in Fig. 4.4. Here, R_1, R_2 are scattering resistances; L_{K1} and L_{K2} are kinetic inductances; L_{E1} and L_{E2} are electrostatic inductances; C_{E1}, C_{E2} are electrostatic capacitances; C_{Q1}, C_{Q2} are electrostatic capacitances and C_{L1}, C_{L2} represent the load capacitances of each line. The parameters L_{12} and C_{12} are mutual inductances and coupling capacitances, respectively. All the line parasitics are per unit length. The voltages and currents in the transmission lines are in *2 X 1* matrix form. The line parasitics are in *2 X 2* per unit length matrix form.

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}, L = \begin{bmatrix} L_{K1} + L_{E1} & L_{12} \\ L_{12} & L_{K2} + L_{E2} \end{bmatrix}$$

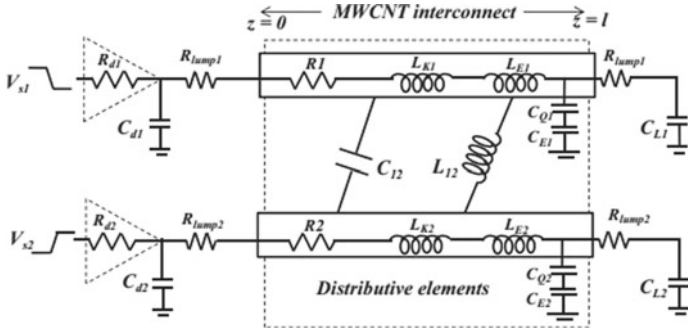


Fig. 4.4 Coupled-two CNT interconnect (Vobulapuram Ramesh Kumar 2014)

$$C = \begin{bmatrix} \left(\frac{1}{C_{Q1}} + \frac{1}{C_{E1}}\right)^{-1} + C_{12} & -C_{12} \\ -C_{12} & \left(\frac{1}{C_{Q2}} + \frac{1}{C_{Q1}}\right)^{-1} + C_{E2} \end{bmatrix} \tag{4.21}$$

The driver parasitics, i.e. resistances (R_{d1} and R_{d2}) and capacitances (C_{d1} and C_{d2}) are diagonal matrices of dimension *2 X 2*. The output is obtained by modifying (4.21) for interconnect line as

$$V_{01} = \frac{\ln 2}{t} o\zeta_K V_{01}\left(K^T \frac{\ln 2}{t}\right)$$
$$V_{02} = \frac{\ln 2}{t} o\zeta_K V_{02}\left(K^T \frac{\ln 2}{t}\right) \tag{4.22}$$

The approximated output is used to examine both functional and dynamic crosstalk. It is used to calculate the interconnect delay and crosstalk noise of the CNT- interconnect lines.

Results and Calculations

Single line and 2-coupled lines for both Cu interconnects and CNT Interconnects are presented to validate the proposed model based on Gaver-Stehfest algorithm. The results obtained are compared with existing models such as Padé model and SPICE results.

A. **Single Line Copper (Cu) Interconnect**

A single line interconnect is employed to check the accuracy and efficiency of the proposed model. The interconnect parameters are given in Table 4.1 which is collected from Roy and Dounavis (2009).

Table 4.1 Typical values for single line copper interconnect parameters

Parameters	Values
V_{DD}	1 V
Length (l)	0.2 to 0.5 cm
Resistance (R)	88.29 Ω/cm
Capacitance (C)	1.8 pF/cm
Inductance (L)	15.38 nH/cm
Load capacitance	50 to 100 fF
Source resistance	50 to 100Ω
Input ramp rise time (T_r)	0.1 ns

The obtained values of 50% time delay and overshoot are compared with the results of the Padé model and SPICE results in Tables 4.2 and 4.3, respectively. It is clearly seen that the results obtained by the proposed model are more accurate than the previously proposed Padé model. The response of a rising ramp is shown in Figs. 4.5 and 4.6 with different source resistance and load capacitance. The average error percentage in the 50% delay and overshoot calculations using the proposed Gaver-Stehfest model are 0.37% and 0.63% respectively.

B. **Two-Coupled Cu Interconnect Lines**

Here, a line of length 2 mm is considered. The interconnect wire parameters are obtained from Sharma et al. (2012) as $R_S = diag(75, 50)$ kΩ, $C_l = diag(30, 30)$ fF, R_1 varies from 0 to 90 kΩ.

$$R = \begin{bmatrix} 653.67 & 0 \\ 0 & 653.67 \end{bmatrix} \frac{k\Omega}{m}, L = \begin{bmatrix} 14.83 & 0.61 \\ 0.61 & 14.83 \end{bmatrix} \frac{\mu H}{m}$$

Table 4.2 Comparison between 50% delay of LTSPICE, Pade model and proposed Gaver-Stehfest-based model for various lengths, source resistance and load capacitance of a single line interconnect

L (cm)	R_s *(Ω)*	C_l *(fF)*	LT SPICE	Pade model (3/3)	Proposed model (M = 10)
			50% delay (ps)	50% delay (ps) (% Error)	50% delay (ps) (% Error)
0.2	50	50	79.8	80.2 (0.5)	80.1 (0.37)
	100	100	98.7	98.6 (0.1)	98.6 (0.1)
0.3	50	50	98.8	99.3 (0.5)	99.0 (0.1)
	100	100	116.9	118.6 (1.45)	117.1 (0.17)
0.5	50	50	135.2	137.8 (1.92)	135.4 (0.14)
	100	100	156.6	151.9 (3.00)	154.5 (1.34)
Average error percentage				1.245	0.37

Table 4.3 Comparison between overshoot of proposed Gaver-Stehfest-based model, Pade and LTSPICE for various lengths, source resistance and load capacitance of a single line interconnect

l (cm)	R_s *(Ω)*	C_l *(fF)*	LT SPICE	Pade model	Proposed model (M = 10)
			Over shoot (V)	Overshoot (V) (% Error)	Overshoot (V) (% Error)
0.2	50	50	1.125	1.14 (1.33)	1.11 (0.89)
	100	100	1	1 (0)	1 (0)
0.3	50	50	1.15	1.15 (0)	1.15 (0)
	100	100	1.00	1.01(0.99)	1.00 (0)
0.5	50	50	1.01	1.057 (4.653)	1.05(2.94)
	100	100	1.00	1.01(0.99)	1.00(0)
Average error percentage				1.327	0.63

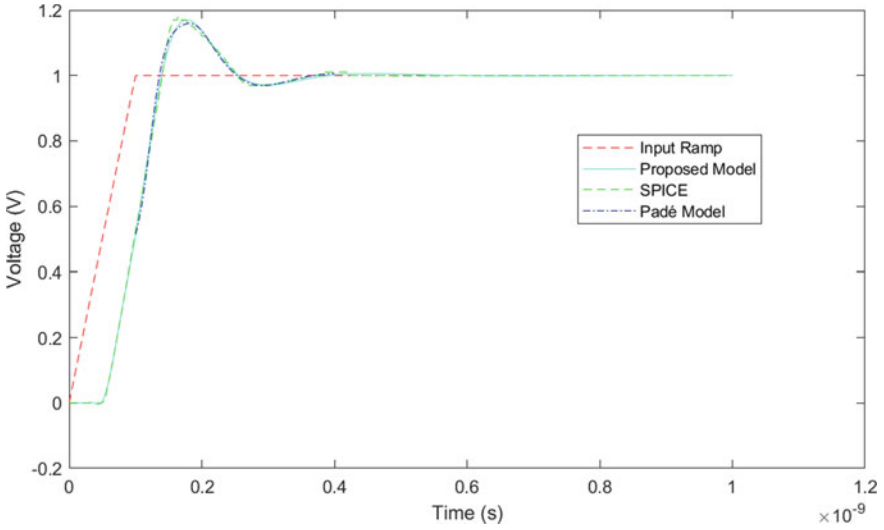


Fig. 4.5 Time domain analysis of single copper line, for length = 0.3 cm, $R_s = 50\Omega$, $C_l = 50$fF

$$C = \begin{bmatrix} 93.33 & -71.50 \\ -71.50 & 93.33 \end{bmatrix} \frac{p\text{F}}{m} \tag{4.23}$$

In coupled-2-line systems, line 1 and line 2 are named victim and aggressor lines, respectively. There are two types of crosstalk, namely, functional crosstalk and dynamic crosstalk. For the case of functional crosstalk, the aggressor line is switched from logic low to logic high while the victim line is kept quiet in logic low level. For the case of dynamic crosstalk, both the lines are switched simultaneously. Firstly, functional crosstalk is analysed. The aggressor line is excited by a ramp signal with

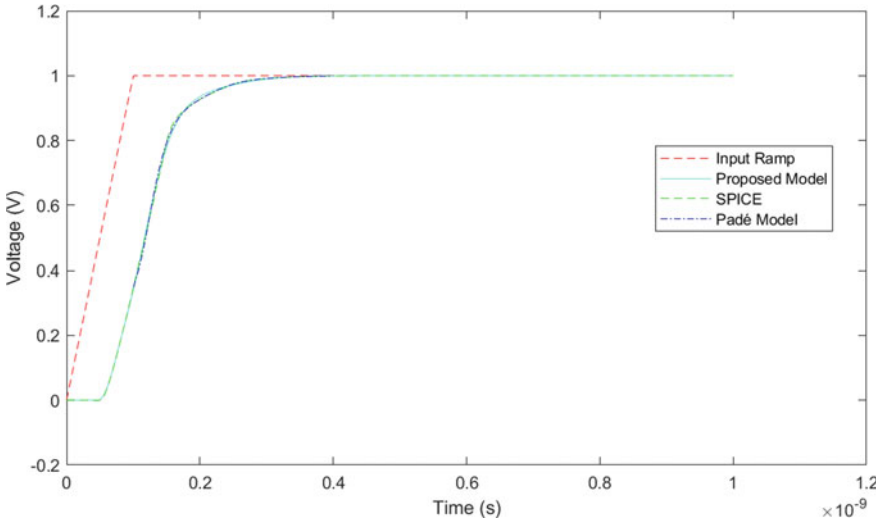


Fig. 4.6 Time domain analysis of single Cu line, for length = 0.3 cm, $R_s = 100\Omega$, $C_l = 100$fF

a rise time of 50 ps and $V_{DD} = 1.2$ V, while the victim line is grounded. The time domain response for the victim line is shown in Fig. 4.7.

In the second case, dynamic crosstalk is considered. Both victim and aggressor lines are driven with the same phase inputs at the same time. The time domain response is plotted in Fig. 4.8.

Table 4.4 shows the comparison for crosstalk induced delay in the aggressor line between the proposed model and the existing models with LTSPICE results. It is observed that the average error in estimation of 50% delay in the aggressor line is 0.61%. Table 4.5 presents the 50%-time delay for dynamic crosstalk. The results obtained using the proposed model are compared with the results of Pade model and the results of LTSPICE simulation tool. The average error percentage in the calculation using the proposed model is 0.55%.

C. Single Line CNT Interconnect

The CNT parameters in Table 4.6 are used to calculate the 50% delay and overshoot of SWCNT and MWCNT (Mekala Girish Kumar 2017). The values of overshoot and delay for SWCNT and MWCNT are compared with the results of Pade model and LTSPICE. The response of a rising ramp is shown in Figs. 4.9 and 4.10 for SWCNT and MWCNT, respectively. Table 4.7 presents the 50% delay for both Single-walled CNTs and Multi-walled CNTS of the proposed model, existing Pade model and LTSPICE.

Figure 4.11 represents a plot between delay and interconnect length for Single-Walled CNTs and Multi-Walled CNTs which is further utilized for comparative study of SWCNT and MWCNT.

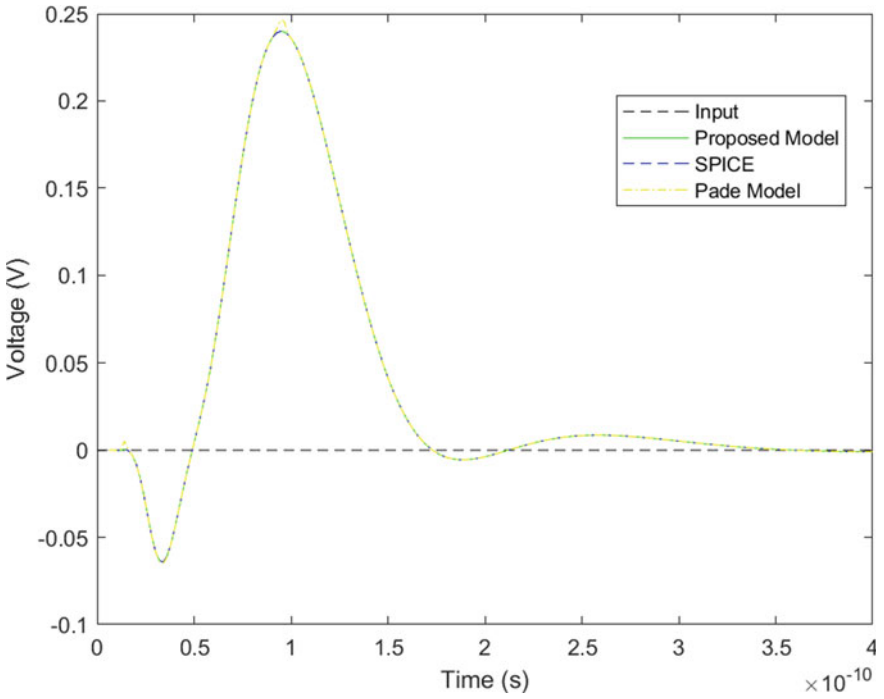


Fig. 4.7 Functional crosstalk in victim line when aggressor is rising ramp while victim is grounded for the case of 2-coupled Cu

It can be clearly seen from Table 4.7 that, by using the proposed model, the average error associated with the estimation of 50%-time delay for SWCNTs is 1.118% and for MWCNTs is 1.344%. Therefore, the delay is predicted better by the proposed model and is easier to implement than the existing Pade Model. It is observed from Fig. 4.11 that Multi-walled CNTs exhibit lesser time delay as compared to Single-Walled CNTs. Hence, MWCNTs are considered as potential candidates for global interconnects (Table 4.8).

D. **Two-Coupled Line CNT Interconnect**

The length of the interconnect lines have been taken as 1 mm. The corresponding line parasitics, obtained from Vobulapuram Ramesh Kumar (2014) are

$$R = \begin{bmatrix} 653.67 & 0 \\ 0 & 653.67 \end{bmatrix} \frac{k\Omega}{m}, L = \begin{bmatrix} 14.83 & 0.61 \\ 0.61 & 14.83 \end{bmatrix} \frac{\mu H}{m}$$

$$C = \begin{bmatrix} 93.33 & -71.50 \\ -71.50 & 93.33 \end{bmatrix} \frac{p\mathrm{F}}{m} \tag{4.24}$$

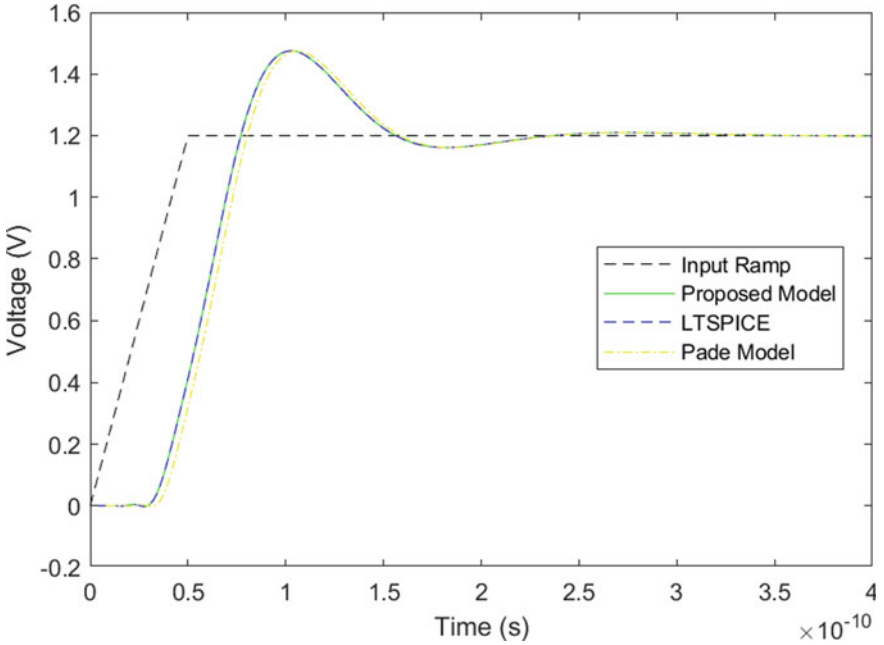


Fig. 4.8 Dynamic crosstalk in 2-coupled Cu lines when both inputs are rising ramp

Table 4.4 Comparison between 50% delay of LTSPICE, Pade model and proposed Gaver-Stehfest-based model for functional crosstalk in the aggressor line for Cu interconnect

Line resistance (Ω)	50% delay		
	LTSPICE (ps)	Pade model) (ps) (% Error)	Proposed model (M = 10) (ps) (% Error)
0	33.45	33.64 (0.5)	32.69 (2.3)
15	34.66	35.62 (2.15)	34. 64(0.09)
30	36.57	37.59 (1.57)	36.60 (0.05)
45	38.76	38.88 (0.3)	38.54 (0.56)
60	40.44	40.58 (0.32)	40.47 (0.05)
75	42.18	42.29 (0.23)	42.41(0.52)
90	44.03	43.99 (0.11)	44.37 (0.75)
Average error percentage		0.87	0.61

The driver parasitics are $R_d = diag(13.85,13.85)$ kΩ and $C_d = diag$ (0.07, 0.07) fF. The load capacitance is varied to verify the robustness of the model. Here, the two lines, viz., line 1 and line 2 are named as aggressor line and victim line, respectively. The crosstalk effects are examined for two cases. In the first case, i.e. functional crosstalk, the aggressor line is switched from logic low level to logic high level

Table 4.5 Comparison between 50% delay of LTSPICE, Pade model and proposed Gaver-Stehfest-based model for dynamic crosstalk

	50% delay		
Line resistance (Ω)	LT SPICE (ps)	Pade model(2/2) (ps) (% Error)	Proposed model (ps) (% Error)
0	30.49	31.67 (3.87)	30.26
15	31.53	32.85 (4.15)	31.33
30	32.71	33.95 (3.76)	32.44
45	33.58	35.00 (4.2)	33.55
60	34.77	36.02 (3.57)	34.67
75	35.98	37.00 (2.86)	35.81
90	37.18	37.96 (2.19)	36.95
Average error percentage		3.52	0.55

Table 4.6 Typical values for CNT interconnect parameters

Parameters		Values
V_{DD}		1 V
Length (l)		500 to 2500 μm
Resistance (R)	SWB	4.4 Ω/μm
	MWB	1.6 Ω/μm
Inductance (L)	SWB	2.75 pH/μm
	MWB	2.17 pH/μm
Capacitance (C)	SWB	14.97 aF/μm
	MWB	13.37 aF/μm
Driver resistance (R_d)		13.80 kΩ
Driver capacitance (C_d)		0.07 fF
Load capacitance (C_l)		1 fF

while the victim line is kept quiet in logic low level. For the next case, i.e. dynamic crosstalk, both lines are switched at the same time, either in-phase or out-phase.

The functional crosstalk is shown in Figs. 4.12 and 4.13 for victim and aggressor lines, respectively. The aggressor line is excited with a ramp input signal having rising time of 50 ps and $V_{DD} = 1$ V. Table 4.9 presents the error associated with the prediction of crosstalk induced peak voltage (crosstalk noise) and timing instant on the victim line using the proposed Gaver-Stehfest based model with respect to LTSPICE simulations and existing Pade model (Vobulapuram Ramesh Kumar 2014). The average error percentage for peak voltage between the obtained analytical results and simulation (LTSPICE) results is 0.17%. The average error in estimating the peak voltage timing is 0.27%. This is more accurate than the existing FDTD method (Vobulapuram Ramesh Kumar 2014). Hence, both peak voltage and peak voltage timing instance are predicted accurately using the proposed model.

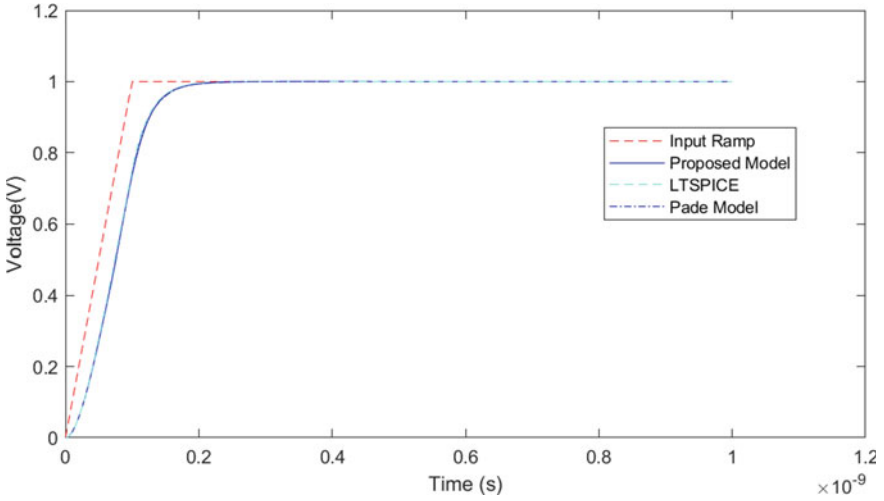


Fig. 4.9 Time domain analysis of SWCNT (single-walled CNT) line, for length = 500 μm

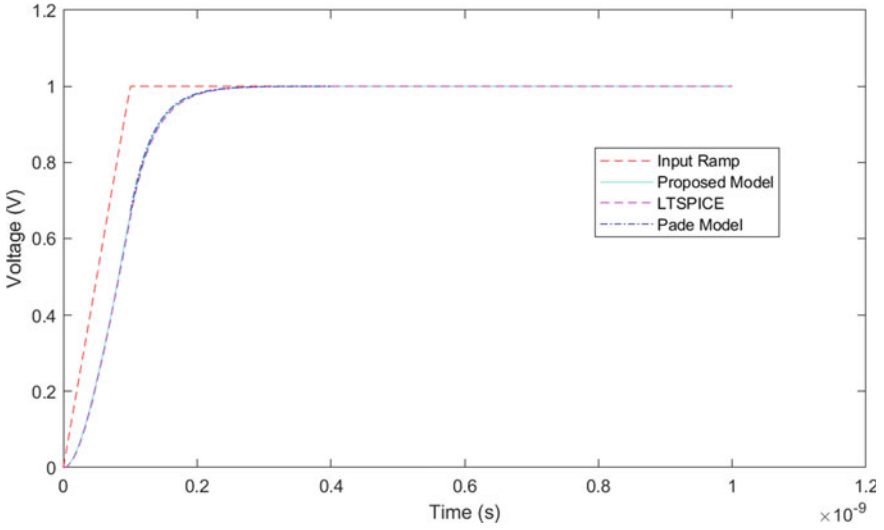


Fig. 4.10 Time domain analysis of MWCNT (multi-walled CNT) line, for length 1000 μm

For dynamic crosstalk, both the lines give the same phase inputs at the same time. The output for dynamic crosstalk is shown in Fig. 4.14.

As per Table 4.10, it is found that the average error percentage in the calculation of propagation delay in aggressor line for dynamic crosstalk using the proposed model is 0.39% which is less than average error percentage for existing Pade method. Hence, the proposed model predicts the crosstalk in CNTs accurately with less than 1% error.

Table 4.7 Comparison between 50% delay of LTSPICE, Pade model and proposed model for different lengths of single-walled CNT and multi-walled CNT

Inter connect length (μm)	50% delay (single-walled CNT)			50% delay (multi-walled CNT)		
	LT SPICE (ps)	Pade model (% error)	Proposed model (M = 10) (% error)	LTSPICE (ps)	Pade model (% error)	Proposed model (M = 10) (% error)
500	74.442	75.449 (1.35)	75.584 (1.53)	73.17	74.264 (1.49)	74.421 (1.7)
1000	83.249	84.194 (1.13)	84.006 (0.90)	80.28	81.97 (2.1)	81.859 (1.96)
1500	91.085	92.177 (1.19)	91.845 (0.83)	88.52	88.996 (0.53)	88.689 (0.19)
2000	98.792	99.589 (0.80)	99.432 (0.64)	94.177	95.489 (1.39)	95.187 (1.07)
2500	105.215	106.82 (1.52)	107.00 (1.69)	99.756	101.57 (1.81)	101.56 (1.80)
Average error percentage		1.198	1.118	Average error percentage	1.464	1.344

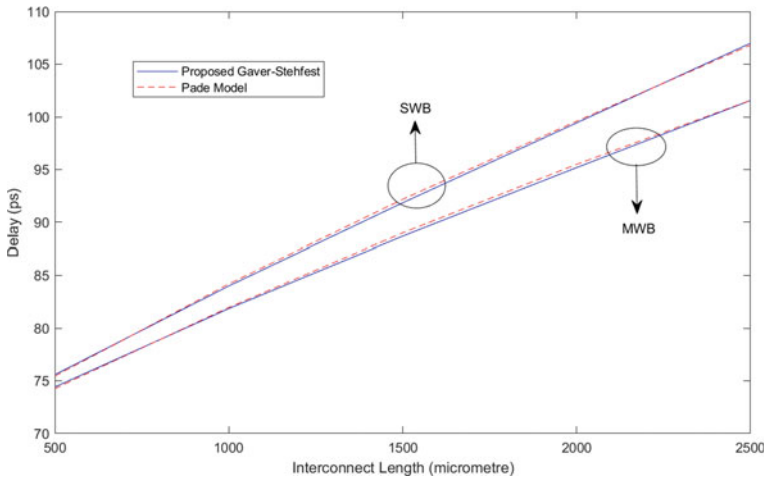


Fig. 4.11 50% Time delay comparison between SWB and MWB for various lengths

Table 4.8 Comparison between overshoot of proposed Gaver-Stehfest-based model and Pade model for several lengths of Single-Walled CNT and Multi-Walled CNT

Interconnect length (μm)	Overshoot (SWCNT)				Overshoot (MWCNT)	
	LTSPICE (V)	Pade model (V) (% error)	Proposed model (M = 10) (% Error)	LT SPICE (V)	Pade model (% error)	Proposed model (M = 10) (% error)
500	1.003	1.021 (1.79)	1.0002 (0.09)	1.001	1.002 (0.09)	1.002 (0.09)
1000	1.0013	1.022 (−0.88)	1.0001 (0.12)	1.003	1.007 (0.39)	1.001 (0.19)
1500	1.0001	1.0001 (0)	1.0002 (0.01)	1.0001	1.001 (0.18)	1.0001 (0)
2000	1	1.001 (0.1)	1 (0)	1.001	1.006 (−0.49)	1 (0.09)
2500	1.001	1.004 (0.21)	1 (0.09)	1	1(0)	1 (0)
Average error percentage		0.5	0.6	Average error percentage	0.23	0.07

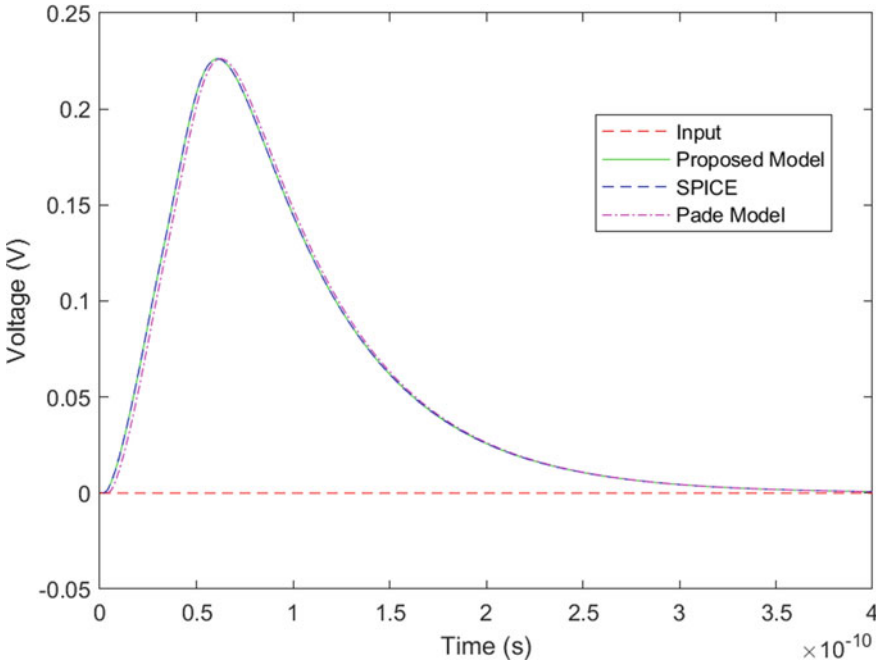


Fig. 4.12 Functional crosstalk in victim line of MWCNT when aggressor input is rising ramp while victim is grounded ($C_l = 40$fF)

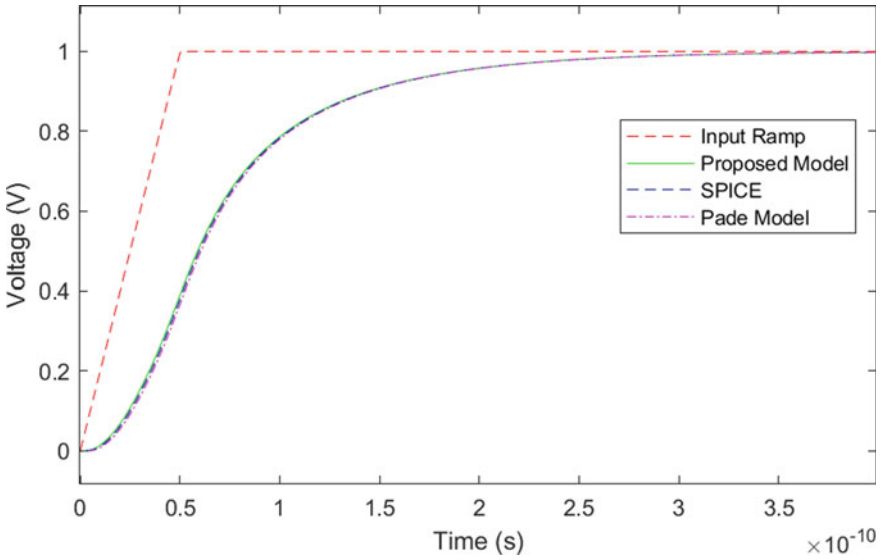


Fig. 4.13 Output in aggressor line of MWCNT in functional crosstalk

Table 4.9 Error associated with the calculation of peak voltage and timing instant of peak voltage on victim line for the case of functional crosstalk

Load capacitance (C_L) (fF)	Peak voltage on the victim line			Timing of peak voltage on victim line		
	LTSPICE (V)	Pade model (V) (% error)	Proposed model (V) (% error)	LTSPICE (ps)	Pade model (ps) (% error)	Proposed model (ps) (% error)
10	0.39	0.390 (0)	0.390 (0)	72.5	73.3 (1.1)	72.9 (0.50)
20	0.318	0.318 (0)	0.318 (0)	77.6	77.9 (0.3)	78.0 (0.51)
30	0.263	0.262 (0.38)	0.263 (0)	80.9	80.8 (0.12)	80.9 (0.00)
40	0.223	0.221 (0.89)	0.222 (0.44)	83.1	82.3 (0.96)	82.9 (0.24)
50	0.193	0.191 (1.03)	0.192 (0.51)	84.4	84.9 (0.6)	84.5 (0.11)
Average error percentage		0.46	0.19	Average error percentage	0.61	0.27

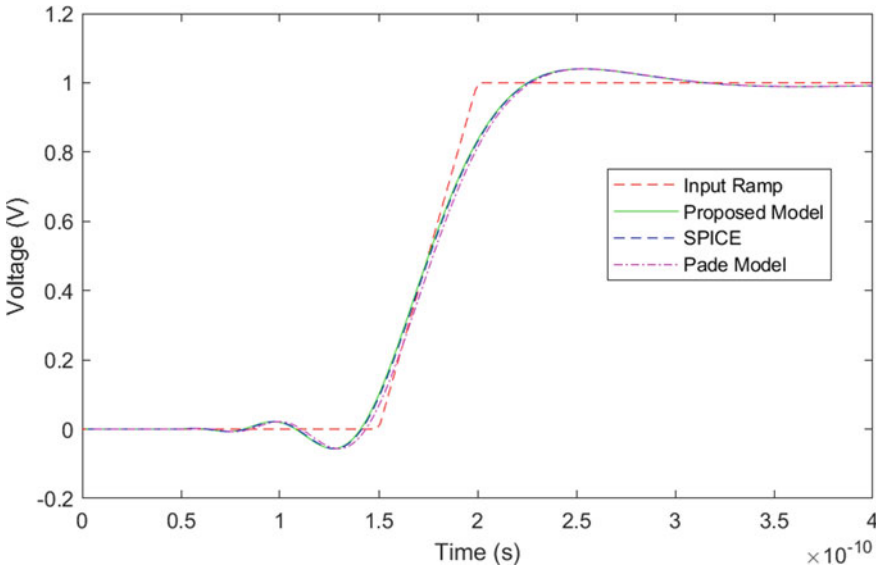


Fig. 4.14 Dynamic crosstalk in MWCNT when both inputs are rising ramp

Table 4.10 Comparison of 50% delay for dynamic crosstalk when inputs are in-phase

Load capacitance (C_L) (fF)	Aggressor line delay		
	SPICE (ps)	Pade model (ps) (% error)	Proposed model (ps) (% error)
10	20.85	21.25 (1.91)	21.12 (1.29)
20	29.49	30.15 (2.23)	29.52 (0.10)
30	36.44	37.19 (2.05)	36.52 (1.21)
40	43.58	43.96 (0.87)	43.61 (0.06)
50	47.41	48.36 (2.00)	48.50 (2.29)
Average error percentage		1.82	0.39

Conclusion

This chapter investigated the delay and crosstalk effects in single line and coupled line interconnects for both copper and carbon nanotube interconnects using the proposed Gaver-Stehfest based method. The delay, overshoot and crosstalk are estimated analytically and compared with existing models and SPICE simulation results. The study of functional crosstalk and dynamic crosstalk in two coupled lines has been presented. Furthermore, the proposed model based on Gaver-Stehfest algorithm can be broadened to analyse n-coupled lines. The proposed model is found to be more accurate than the prevailing models and is in good agreement with SPICE values. The analysis suggests that the proposed model can play a big role in performance analysis of Cu and CNT interconnects and possible to integrate in TCAD simulators.

References

Abate J, Whitt W (2006) A unified framework for numerically inverting laplace transform. INFORMS J Comput 18(4):408–421

Agrawal Y, Kumar MG, Chandel R (2017) A unified delay, power and crosstalk model for current mode signaling multiwall carbon nanotube interconnects. Springer Science+Business Media

Das D, Rahaman H (2011) Analysis of crosstalk in single-and multiwall carbon nanotube interconnects and its impact on gate oxide reliability. IEEE Trans Nanotechnol 10(6).

Dounavis A, Achar R, Nakhla M (2001) A general class of passive micromodels for lossy multiconductor transmission lines. IEEE Microw Theory Tech 49(10):1686–1696

Dounavis A, Achar R, Xin L (1999) Passive closed-form transmission line model for general-purpose circuit simulators. IEEE Trans Microw Theory Tech 47:2450–2459

Elmore WC (1948) The transient response of demand linear networks with particular regard to wideband amplifiers. J Appl Phys 19(1):55–63

Fathi D, Forouzandeh B (2009) A novel approach for stability analysis in carbon nanotube interconnects. IEEE Electron Device Lett 30(5).

Fathi D, Forouzandeh B, Mohajerzadeh S, Sarvari R (2009) Accurate analysis of carbon nanotube interconnects using transmission line model. Micro & Nano Lett.

Fathi D, Forouzandeh B (2012) Time domain analysis of carbon nanotube interconnects based on distributed RLC model. World Scientific Publishing Company

Haji-Nasiri S, Faez R, Moravvej-Farshi MK (2012) Stability analysis in multiwall carbon nanotube bundle interconnects.

Hermite C (1908) Sur la résolution de l'équation du cinquème degré. Œuvres De Charles Hermite 2:5–21

Kumar MG, Agrawal Y, Chandel R (2017) Stability analysis of carbon nanotube interconnects. Springer Science+Business Media Singapore.

Kumar VR, Kaushik BK, Patnaik A (2014) Crosstalk noise modeling of multiwall carbon nanotube (MWCNT) interconnects using finite-difference time-domain (FDTD) technique. Microelectron Reliab

Li H, Yin WY, Banerjee K, Mao JF (2008) Circuit modeling and performance analysis of multi-walled carbon nanotube interconnects. IEEE Trans Electron Devices 55(6).

Ravindra JV, Srinivas MB (2007a) Analytical crosstalk model with inductive coupling in VLSI interconnects. In: IEEE Workshop on Signal Propagation on Interconnects.

Ravindra JV, Srinivas MB (2007b) Analytical crosstalk model with inductive coupling in VLSI interconnects. In: IEEE Workshop on Signal Propagation on Interconnect.

Roy S, Dounavis A (2009) Closed-form delay and crosstalk models for RLC on-chip interconnects using a matrix rational approximation. IEEE Trans Comput-Aided Des Integr Circuits Syst 28(10)

Sahoo M, Rahaman H (2013) Modeling of crosstalk delay and noise in single-walled carbon nanotube bundle interconnects. IEEE

Sharma DK, Mittal S, Kaushik B, Sharma R, Yadav K, Majumder M (2012) Dynamic crosstalk analysis in RLC modeled interconnects using FDTD method. In: Third International Conference in Computer and Communication Technology.

Srivastava N, Banerjee K (2005) Performance analysis of carbon nanotube interconnects for VLSI applications. In: IEEE/ACM International Conference on Computer-Aided Design

Sathyakam PU, Mallick PS, Singh P (2020) Geometry-based crosstalk reduction in CNT interconnects. J Circuits, Syst, Comput 29(6)

Chapter 5
Repeater Insertion for Carbon Nanotube Interconnects

Wen-Sheng Zhao

Abstract Optimal repeater insertion techniques are usually employed in long interconnects to improve the performance. As a promising alternative interconnect material, carbon nanotube (CNT) has attracted much attention in past several years. However, it is worth noting that the implementation of CNT interconnects is hindered by large contact resistance, which would affect the number and size of repeaters. In this chapter, the repeater insertion technique is explored for CNT interconnects, with the contact effect treated appropriately. It is demonstrated that multi-walled CNT (MWCNT) is susceptible to the contact resistance, and special attention should be paid to the repeater insertion in MWCNT interconnects. Further, both the delay and power consumption are considered in the design of repeaters. The particle swarm optimization (PSO) algorithm is employed to capture the optimal number and size of repeaters, and the results are used to train an artificial neural network (ANN). The developed procedure can be applied to the optimal design of interconnect system in nano-CMOS and future nano-carbon-based ICs.

Keywords Carbon nanotube · Interconnect · Circuit model · Repeater insertion

Introduction

With the rapid development of semiconductor manufacturing technologies, the feature size of integrated circuits (ICs) is continually shrinking and the device dimensions are continually scaling down. Unlike transistors, interconnects suffer from limitations of both diminishing current-carrying capacity and a nonlinear increase in the resistivity, mainly attributed to the decreased conductor area. The interconnect delay has become much larger than the transistor delay and hence the interconnect era has

W.-S. Zhao (✉)
Zhejiang Provincial Key Lab of Large-Scale Integrated Circuit Design, School of Electronics and Information, Hangzhou Dianzi University, Hangzhou 310018, China
e-mail: wsh.zhao@gmail.com

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_5

been launched (Meindl 2003; Venkatesan et al. 2001; Wong et al. 2020). More importantly, the current density of on-chip interconnects increases dramatically, thereby threatening the reliability of the integrated circuits (Zhao et al. 2022a; Sathyakam and Mallick 2017).

Carbon nanotube (CNT) are considered as a promising candidate for interconnect applications due to their extraordinary electrical, thermal, and mechanical properties (Srivastava et al. 2005; Close and Wong 2007; Zhao et al. 2019a; Todri-Sanial et al. 2017; Chiariello et al. 2013). For instance, the ampacity of CNTs could be about an order of magnitude higher than Cu counterpart (Radosavljevic et al. 2001; Wei et al. 2001), thereby significantly improving the IC reliability. According to the number of layers of carbon atoms, CNTs can be divided into two categories: single-walled carbon nanotubes (SWCNTs) and multi-walled carbon nanotubes (MWCNTs) (Avourise et al. 2003). An SWCNT is a rolled-up graphene sheet with diameter of a few nanometers, and an MWCNT can be viewed as a parallel assembly of SWCNTs (Li et al. 2008). In general, the intrinsic resistance of an SWCNT is too large for interconnect applications, the monolayer or bundled SWCNTs are usually employed to reduce resistive losses (Zhao et al. 2015; Naeemi and Meindl 2007; Maffucci et al. 2009). It was demonstrated that the monolayer SWCNTs could reduce the power dissipation and crosstalk, which are mainly attributed to their low parasitic capacitance, while bundled SWCNTs exhibit better transmission performance than their Cu counterpart due to low effective resistivity. It is worth noting that SWCNTs could be metallic or semiconducting depending on their chirality. In contrast, all the MWCNTs are metallic and therefore, they are more suitable for interconnect applications (Close and Wong 2008; Majumder et al. 2012; Liang et al. 2011a). After developing a multi-conductor circuit model of MWCNT interconnects (Li et al. 2008), it is demonstrated that MWCNT interconnects could provide better performance than conventional Cu and bundled SWCNTs at the intermediate and global interconnect levels. To simplify the modeling procedure, an equivalent single-conductor (ESC) transmission line model was established later for MWCNTs (Sarto and Tamburrano 2009), where the equivalent kinetic inductance and quantum capacitance were extracted from inductance and capacitance networks. Further, the application range of ESC model for characterizing the MWCNT interconnects was also examined (Tang and Mao 2014).

Figure 5.1 plots the diagram of interconnect delay versus length, and the schematic of driver-interconnect-load (DIL) system is shown in the figure inset. In the DIL model, R_{d} and C_{d} denote the resistance and capacitance of the driver, C_{L} is the load capacitance, and r_{int} and c_{int} represent the per-unit-length resistance and capacitance of the interconnect, respectively. Here, the interconnect length is denoted by l. It is evident that both the interconnect resistance and capacitance are proportional to the length. For interconnect of very short length, the delay is mainly determined by the driver's elements, and the corresponding region is defined as sub-linear region. As the interconnect length exceeds a certain value, the interconnect delay reaches linear region, i.e., the delay increases linearly with the length. For interconnect length longer than 50 gate pitches, the influences of both interconnect resistance and capacitance becomes nonnegligible, i.e., the interconnect delay is determined by $r_{\mathrm{int}} c_{\mathrm{int}} l^2$

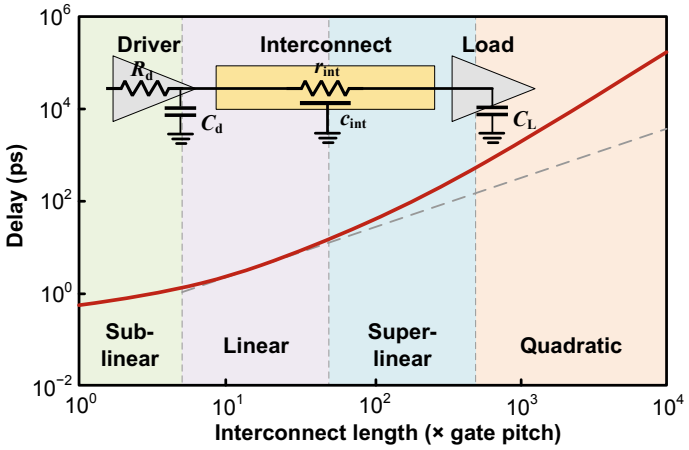


Fig. 5.1 Interconnect delay versus length

(Rakheja and Naeemi 2010). It is evident that there is significant quadratic pattern between the interconnect delay and length in quadratic region. It is evident that by inserting repeaters in a long interconnect, the interconnect delay of each segment can be suppressed from quadratic region to linear region or super-linear region, thereby reducing the total delay significantly. Several studies have been carried out to investigate the repeater insertion methodology for on-chip interconnects (Li et al. 2005; Ismail and Friedman 2000; Banerjee and Mehrotra 2002, 2006). For instance, considering the influence of the inductance, the analytical expressions were derived in Ismail and Friedman (2000) for predicting the optimal number of repeaters and repeater size. Aiming at achieving low power dissipation, the power-optimal repeater insertion in nanometer interconnects was studied in Banerjee and Mehrotra (2002). It was demonstrated that the total interconnect delay could be effectively suppressed by inserting repeaters in a long interconnect.

The repeater insertion techniques were also investigated for CNT interconnects. It was claimed (Guistininai et al. 2010) that the performance improvement by inserting repeaters in bundled SWCNTs is not significant. Further, the MWCNT interconnects were investigated using finite-difference time-domain (FDTD) method (Liang et al. 2011b), and the optimal repeater insertion in intermediate and global level MWCNT interconnects were calculated using an analytical expressions (Ismail and Friedman 2000). However, it is worth noting that CNT interconnects, which are different to their Cu counterpart, have significant contact resistance although considerable efforts have been devoted to improving contact performance (Chai et al. 2012). Therefore, the implementation of each repeater inserted in the CNT interconnect would introduce two additional contact resistances, which must be treated in an appropriate way.

Repeater Insertion in CNT Interconnects

Figure 5.2 shows the schematic diagram of a CNT interconnect inserted with a series of equispaced repeaters. Here, the number of segments divided by the repeaters is denoted as k, the size of repeaters with respect to the minimum-sized gate is h. That is, the repeaters are h times the minimum size, with the output resistance R_{d0}/h, output capacitance hC_{d0}, and input capacitance hC_{L0}, where R_{d0}, C_{d0}, and C_{L0} are the output resistance, output capacitance, and input capacitance of the minimum-sized repeater, respectively. In the figure, R_c denotes the contact resistance of a CNT interconnect, and as described earlier, when a repeater is inserted, two additional contact resistances would be introduced at two sides of the repeater.

Figure 5.3a shows the cross-sectional schematic of a Cu wire placed above the ground plane. Here, ε_r represents the relative permittivity of surrounding dielectric, h is the distance between Cu wire and ground plane, and w and t are the width and thickness of the Cu wire, respectively. The effective resistance of nanoscale Cu wire can be calculated by $R_{Cu} = \rho_{Cu} l/(wt)$, and the effective resistivity is given as (Im et al. 2005)

$$\rho_{Cu} = \rho_0 \left[\frac{1}{3} \frac{1}{1/3 - \alpha/2 + \alpha^2 - \alpha^3 \ln\left(1 + \alpha^{-1}\right)} + \frac{3}{8} C(1-p) \frac{1 + \mathrm{AR}}{\mathrm{AR}} \frac{\lambda}{w} \right] \quad (5.1)$$

where ρ_0 is the bulk resistivity, $\alpha = \lambda R / \left[d_g (1 - R) \right]$, λ is the mean free path (MFP), $\mathrm{AR} = t/w$, p is the specularity parameter, R is the reflectivity coefficient at grain boundaries, d_g is the average distance between grain boundaries, and $C = 1.2$. Here, both p and R are set as 0.5, and the line edge roughness is about 40% (Ceyhan and Naeemi 2013). The parasitic capacitance of Cu wire could be extracted using full-wave electromagnetic solver (e.g., ANSYS Q3D Extractor).

Similarly, the parasitic capacitances of monolayer SWCNTs and bundled SWCNTs could be obtained using full-wave electromagnetic solver as the capacitance of SWCNTs is close to that of Cu wires with equal cross-sectional area (Naeemi and Meindl 2007). The quantum capacitance of SWCNT is usually neglected as it is much larger than electrostatic one (Zhao et al. 2014). Different with Cu wire, a CNT interconnect has a quantum contact resistance R_Q and imperfect contact resistance R_{mc}, i.e., $R_C = R_Q + R_{mc}$. The imperfect contact resistance highly depends on the fabrication process, and is usually neglected in the modeling and performance analysis. The quantum contact resistance of an SWCNT can be calculated by

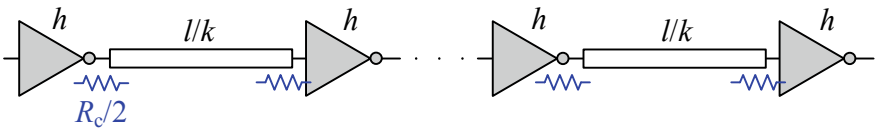


Fig. 5.2 Schematic of repeaters inserted into a CNT interconnect

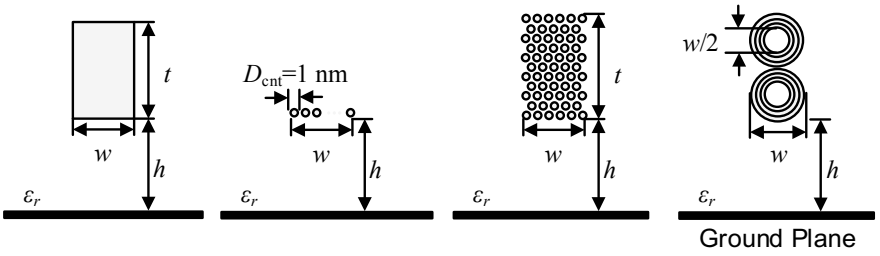


Fig. 5.3 Cross-sectional views of (**a**) Cu, (**b**) monolayer SWCNTs, (**c**) bundled SWCNTs, and (**d**) MWCNT interconnects

$$R_Q = \frac{h}{2q^2}\frac{1}{N_{\text{ch}}} \tag{5.2}$$

where h is the Planck's constant, q is the electron charge, and N_{ch} denotes the number of conducting channels. The per-unit-length (p.u.l.) scattering resistance of an SWCNT is given as (Naeemi and Meindl 2007)

$$R_S = \frac{h}{2q^2}\frac{1}{N_{\text{ch}}}\frac{1}{\lambda}. \tag{5.3}$$

In general, N_{ch} is 2 for the metallic SWCNT, and λ is about 1 μm. In general, the fraction of metallic SWCNTs in a bundle is about 1/3, and it could be improved through specific fabrication process (Harutyunyan et al. 2009). Therefore, the monolayer SWCNTs shown in Fig. 5.3b are assumed as metallic, while the fraction of metallic SWCNTs in the bundle shown in Fig. 5.3c is set as 1/3. The number of SWCNTs in monolayer and bundled SWCNTs could be calculated by $\text{Int}[w/(D_{\text{cnt}}+\delta)]$ and

$$n_w n_t - \text{Int}[N_t/2]$$

where $n_w = \text{Int}[(w - D_{\text{cnt}})/(D_{\text{cnt}}+\delta)] + 1$, $n_t = \text{Int}\left[2(t - D_{\text{cnt}})/\left(\sqrt{3}(D_{\text{cnt}}+\delta)\right)\right] + 1$, and $\delta = 0.34$ nm is Van der Waal's gap (Srivastava et al. 2005). $\text{Int}[\cdot]$ denotes that only the integer part is considered. The contact resistance and p.u.l. scattering resistances of monolayer SWCNTs and bundled SWCNTs could be obtained by dividing the corresponding resistance of an isolated SWCNT by the number of SWCNTs.

As shown in Fig. 5.3d, two MWCNTs are usually employed as on-chip interconnects, and the aspect ratio is therefore set as 2. As shown in Fig. 5.4, the diameter of the MWCNT is equal to the interconnect width w, and the innermost diameter is set as half of the MWCNT diameter. The number of shells in an MWCNT can be calculated by Liu et al. (2018)

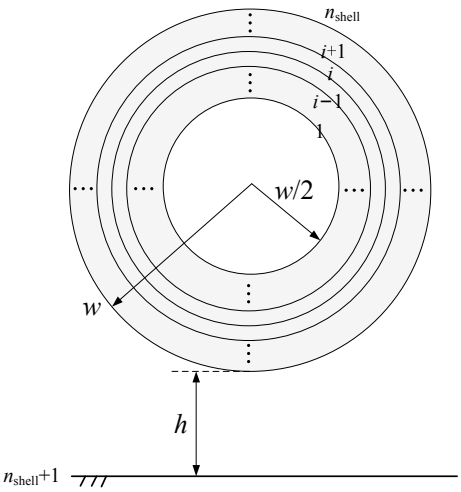


Fig. 5.4 Cross-sectional diagram of an MWCNT

$$n_{\text{shell}} = 1 + \text{Int}\left[\frac{1}{2}\frac{w}{\delta}\right]. \tag{5.4}$$

The radius of i th shell is $r_i = r_1 + (i-1)\delta$. The terminal contact resistance of an MWCNT is given as

$$R_C = \left[\sum\nolimits_{i=1}^{n_{\text{shell}}}\left(\frac{h}{4q^2 N_{\text{ch},i}} + R_{\text{mc},i}\right)^{-1}\right]^{-1} \tag{5.5}$$

where $N_{\text{ch},i} = a \cdot r_i + b$, $a = 0.122\,\text{nm}^{-1}$, and $b = 0.425$. As an MWCNT can be viewed as a parallel assembly of SWCNTs, its p.u.l. scattering resistance could be calculated by

$$R_S = \frac{h}{2q^2}\left(\sum\nolimits_{i=1}^{n_{\text{shell}}} N_{\text{ch},i}\lambda_i\right)^{-1} \tag{5.6}$$

where $\lambda_i = 2000 r_i$ is the MFP of the i th shell.

After the values of resistance and capacitance are determined, the interconnect can be modeled as a simple RC model, and the 50% time delay of one segment in CNT interconnect can be calculated as

$$\begin{aligned}\tau_{\text{seg}}(h,k) &= 0.69(R_C h C_{\text{L0}} + R_{\text{d0}} C_0) \\ &\quad + 0.69\left(\frac{R_C C}{2} + R_S h C_{\text{L0}} + \frac{R_{\text{d0}} C}{h}\right)\frac{l}{k} + 0.38 R_S C\left(\frac{l}{k}\right)^2\end{aligned} \tag{5.7}$$

where C denotes the parasitic capacitance of the on-chip interconnect, and $C_0 = C_{\text{d0}} + C_{\text{L0}}$. The total time delay is $T(h,k) = k\tau_{\text{seg}}(h,k)$, and by satisfying

$\partial T(h,k)/\partial h = 0$ and $\partial T(h,k)/\partial k = 0$, an equation can be obtained as

$$69(R_{\mathrm{d0}}C_0 + R_C C_{\mathrm{L0}} h) = 38 R_S C \left(\frac{R_C C_{\mathrm{L0}} h}{R_{\mathrm{d0}} C - R_S C_{\mathrm{L0}} h^2} \right)^2. \tag{5.8}$$

The optimal value of h can be calculated as h_{opt} by solving the above expression, and the optimal number of segments k_{opt} can be obtained by

$$k_{\mathrm{opt}} = \mathrm{Int}\left[\left(\frac{R_{\mathrm{d0}} C}{C_{\mathrm{L0}} h_{\mathrm{opt}}^2} - R_S \right) \frac{l}{R_C} \right]. \tag{5.9}$$

By substituting h_{opt} and k_{opt} into $T(h,k) = k\tau_{\mathrm{seg}}(h,k)$, the minimum time delay of the CNT interconnect can be obtained accordingly. It is evident that h_{opt} is independent with the interconnect length, while k_{opt} increases linearly with the length.

A transition length is defined as a value where the repeaters begin to work, and it is given as

$$l_{\mathrm{transition}} = \frac{69}{38} \left(\frac{R_{\mathrm{d0}}}{R_S h_{\mathrm{opt}}} + \frac{R_C}{2R_S} + \frac{C_{\mathrm{L0}} h_{\mathrm{opt}}}{C} \right). \tag{5.10}$$

Taking the interconnect parameters at the 14 nm technology node as an example (see Table 5.1), it is calculated that the transition length is about 100 μm, and increases with the contact resistance, in particular for MWCNT interconnects. Figure 5.5 shows the total time delay of the bundled SWCNTs and MWCNT interconnects with different lengths. In the figure, h_{opt} is calculated by solving (5.8), and the cases with $h = 50$ are also plotted. When the length is shorter than $l_{\mathrm{transition}}$, as shown in Fig. 5.5a, the time delay increases almost linearly with the number of repeaters. That is, the inserted repeaters have little influence on the time delay. However, for a long CNT interconnect, as shown in Fig. 5.5b, there exists a minimum value of time delay, implying that the repeaters could reduce the total time delay effectively. Moreover, it is found in Fig. 5.5b that the optimal number of repeaters k_{opt} remains unchanged for different values of h for ideally contacted CNT interconnects. This is because that the contact resistance can be neglected for ideally contacted CNT interconnects, and therefore, k_{opt} and h_{opt} become irrelevant with each other. Then, the optimal value of repeater size is plotted in Fig. 5.6 for different imperfect contact resistance per channel. It is evident that h_{opt} decreases with the increasing imperfect contact resistance, in particular for MWCNT interconnects.

In previous studies, the contact resistance of CNT interconnect was usually combined into the total interconnect resistance in the repeater design. The optimal values of repeater size and number of repeaters could be calculated by Liang et al. (2011b)

Table 5.1 Geometrical and physical parameters adopted from ITRS 2013

Technology node		14 nm	7 nm
Intermediate	Width (nm)	14	7
	Aspect ratio	2.1	2.3
	Dielectric thickness (nm)	26.6	13.3
	Cu resistivity ($\mu\Omega$·cm)	7.43	11.41
	Capacitance for Cu wire (pF/μm)	130.8	110.3
	Capacitance for MWCNT (pF/μm)	115.8	92.4
Global	Width (nm)	21	11
	Aspect ratio	2.34	2.34
	Dielectric thickness (nm)	31.5	16.5
	Cu resistivity ($\mu\Omega$·cm)	6.11	8.97
	Capacitance for Cu wire (pF/μm)	143.6	115
	Capacitance for MWCNT (pF/μm)	124.2	99.7
Minimum-sized gate	Output resistance (kΩ)	30.3	69.7
	Input capacitance (fF)	0.22	0.13
Relative permittivity of the ILD		2.0	1.6

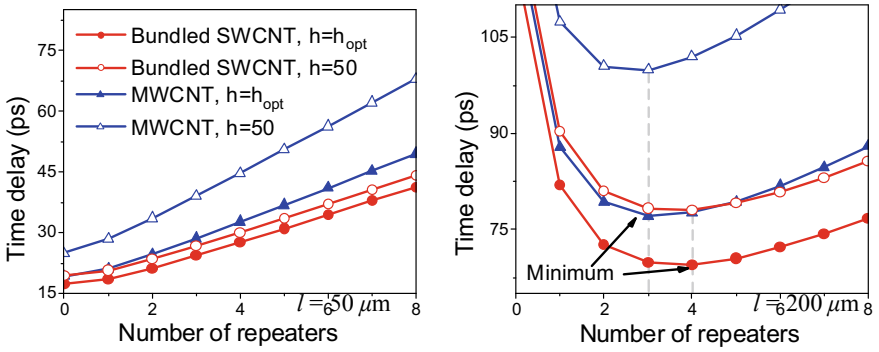


Fig. 5.5 Time delay against number of repeaters for ideally contacted SWCNT bundle and MWCNT interconnects with length of **a** 50 μm and **b** 200 μm

$$h_{\text{opt}}{}' = \sqrt{\frac{R_{\text{d0}}Cl}{R_{\text{t}}C_{\text{L0}}}} \tag{5.11}$$

$$k_{\text{opt}}{}' = \text{Int}\left[\sqrt{\frac{38R_{\text{t}}Cl}{69R_{\text{d0}}C_0}}\right] \tag{5.12}$$

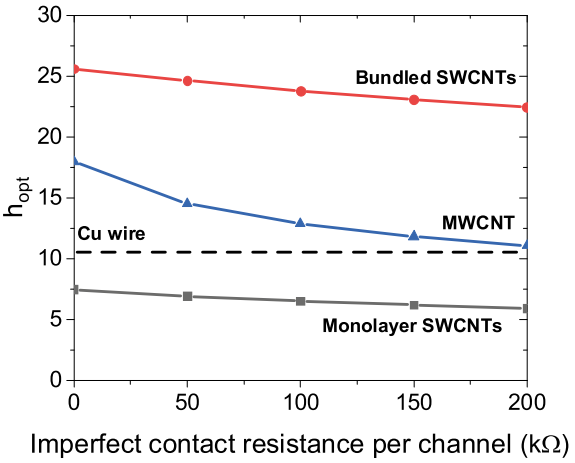


Fig. 5.6 Optimal value of repeater size h_{opt} versus imperfect contact resistance per channel

where $R_t = R_S l + R_C$. As R_C do exist in each segment, the percentage error induced by the above simplified treatment is defined as

$$\text{Error} = \left[\frac{T\left(h_{\text{opt}}\prime, k_{\text{opt}}\prime\right)}{T\left(h_{\text{opt}}, k_{\text{opt}}\right)} - 1 \right] \times 100\%. \tag{5.13}$$

It is evident that the percentage error is independent with the interconnect length. Figure 5.7 shows the percentage error versus imperfect contact resistance per channel for CNT interconnect with length of 1000 μm. It is evident that the percentage error increases slightly with the imperfect contact resistance for monolayer SWCNTs and bundled SWCNTs, but it is still acceptable. However, as the imperfect contact resistance is larger than ~70 kΩ, the percentage error in the case of MWCNT interconnect would exceed 10%, implying that the optimal repeater insertion for MWCNT interconnects must consider the influence of contact resistance.

Influence of Inductance on the Repeater Insertion in MWCNT Interconnects

As indicated in Ismail and Friedman (2000), the inductance has a certain influence on the time delay and repeater design of on-chip interconnect. Here, the influence of kinetic inductance on the repeater design in MWCNT interconnect would be discussed in the following. As shown in Fig. 5.8, when a repeater is inserted in a long MWCNT interconnect, two contact resistance would be introduced, and each segment of MWCNT could be equivalent to a RLC model. The p.u.l. inductance of an MWCNT is calculated as $L = L_{\text{K}} + L_{\text{M}}$, where L_{M} and L_{K} are the magnetic inductance and kinetic inductance, respectively. The magnetic inductance could be

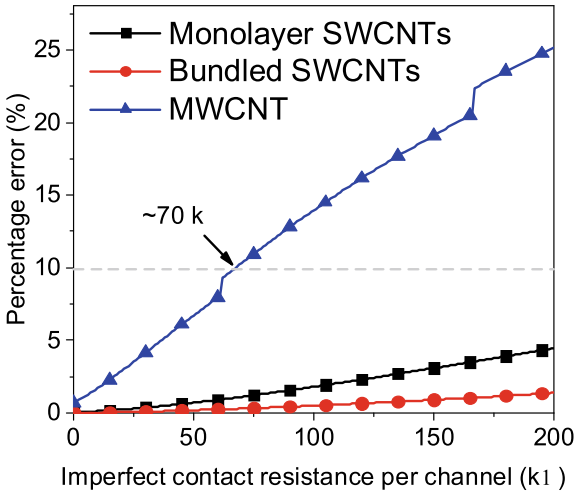


Fig. 5.7 Percentage error versus imperfect contact resistance per channel

obtained by $L_M = \mu_0 \varepsilon_0 \varepsilon_r / C$, and the kinetic inductance is given by $L_K = L_{K/channel} \cdot \left(\sum_{i=1}^{n_{shell}} N_{ch,i}\right)^{-1}$, where $L_{K/channel} \approx 8\text{nH}/\mu\text{m}$ is the kinetic inductance per channel.

The 50% time delay of one segment in MWCNT interconnect can be given by

$$\tau_{seg}(h,k) = \tau_{RC} + (\tau_{RLC} - \tau_{RC}) \cdot M \tag{5.14}$$

where $M = \left(1 + 0.005\left(\cos\left(9\cos\left((a_1 + a_2)^{0.294}/(1.2 - 0.0498a_3)\right)\right)\right)^2\right)^{-1/2}$, $\tau_{RC} = \tau_f \cdot (0.4a_1 + 0.7(a_2 + a_1a_3 + a_2a_3))$, and $\tau_{RLC} = \tau_f \cdot \left(1 + a_3\left(0.342 + 0.397a_2 + 0.229a_1^{2.34} + 0.642a_1^{1.3}a_2^{1.35}\right)\right)$. . The three basic coefficients are the normalized ratio defined as (Lu et al. 2016)

$$a_1 = \frac{\text{line resistance}}{\text{characteristic impedance}} = \frac{1}{Z_0}\left(R_C + \frac{R_S l}{k}\right) \tag{5.15}$$

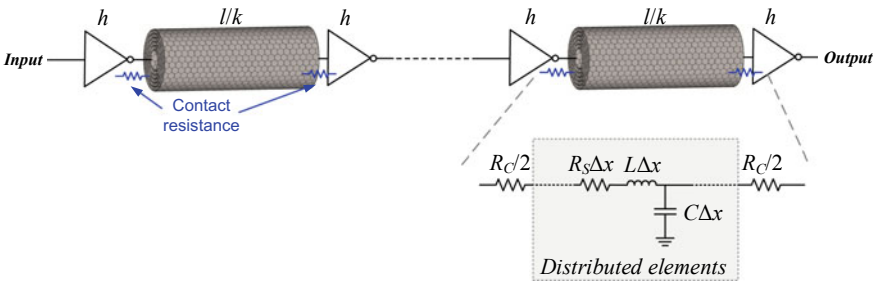


Fig. 5.8 Schematic of MWCNT interconnect inserted with equispaced repeaters

$$a_2 = \frac{\text{driver output resistance}}{\text{characteristic impedance}} = \frac{R_{\text{d0}}}{hZ_0} \tag{5.16}$$

$$a_3 = \frac{\text{load resistance}}{\text{characteristic impedance}} = \frac{hkC_{\text{L0}}}{Cl} \tag{5.17}$$

where $Z_0 = \sqrt{L/C}$ is the characteristic impedance and $\tau_{\text{f}} = \sqrt{LC}l/k$ is the time-of-flight (Venkatesan et al. 2003). It is evident that τ_{seg} is a function of a_1, a_2, and a_3, and it is written as $\tau_{\text{seg}} = \tau_{\text{f}} \cdot F(a_1, a_2, a_3)$. The total time delay of the MWCNT interconnect is given as

$$T_{\text{total}} = k\tau_{\text{f}} \cdot F(a_1, a_2, a_3) = \sqrt{LC}l \cdot F(a_1, a_2, a_3). \tag{5.18}$$

According to Ismail and Friedman (2000), the optimal repeater size and optimal number of segments can be expressed as

$$h_{\text{opt}} = \sqrt{\frac{R_{\text{d0}}Cl}{C_{\text{L0}}(R_S l + R_C)}}h\prime \tag{5.19}$$

$$k_{\text{opt}} = \sqrt{\frac{4}{7}\frac{(R_S l + R_C)Cl}{R_{\text{d0}}C_{\text{L0}}}}k\prime \tag{5.20}$$

where $h\prime$ and $k\prime$ are the correction factors. By substituting h_{opt} and k_{opt} into T_{total}, the total time delay can be written as

$$T_{\text{total}} = \sqrt{LC}l \cdot F\left(m + \sqrt{\frac{7}{4}} \cdot \frac{1}{k\prime pq}, \frac{q}{h\prime p}, \sqrt{\frac{4}{7}} \cdot \frac{1}{h\prime k\prime}\right) \tag{5.21}$$

where $p = \sqrt{L/(R_S R_{\text{d0}} C_{\text{L0}})}$, $q = \sqrt{1 + R_C/(R_S l)}$, and $m = R_C/Z_0$. As R_C is usually smaller than $R_S l$, q, and p could be limited to [1, 1.4] and [0, 10], respectively. Moreover, after calculating Z_0 according to the parameters adopted from ITRS 2013, m is limited to the range of [0, 10]. To minimize the time delay, the expressions of $\partial F/\partial h\prime = 0$ and $\partial F/\partial k\prime = 0$ should be solved simultaneously. By utilizing the multivariable curve fitting technique, the closed-form expressions of $h\prime$ and $k\prime$ could be obtained as

$$h' = \cos(q-1)\frac{0.839 + 0.269m - 0.021m^2 - 0.917\ln p + 0.345(\ln p)^2 - 0.037(\ln p)^2}{1 + 0.209m + 0.045m^2 - 0.005m^3 - 0.854\ln p + 0.361(\ln p)^2} \tag{5.22}$$

$$k' = \cos(q-1)\frac{0.835 + 0.307m - 0.047m^2 + 0.002m^3 - 0.606\ln p + 0.116(\ln p)^2}{1 + 0.739m - 0.073m^2 + 0.002m^3 - 0.474\ln p + 0.161(\ln p)^2}. \tag{5.23}$$

The final values of h_{opt} and k_{opt} could be captured by substituting (5.22) and (5.23) into (5.19) and (5.20), respectively, and then, the minimum time delay T_{total} could be obtained.

To explore the influence of contact resistance and inductance, three types of MWCNT interconnects are investigated, as listed in Table 5.2. MWCNT 1 denotes that only the quantum contact resistance is considered, i.e., the MWCNT interconnects with minimum contact resistance. However, the imperfect contact resistance is inevitable in real-world applications. Therefore, a contact resistance of 10 kΩ is considered in the case of MWCNT 2. It is worth noting that the kinetic inductance of an isolated CNT may be 15 times larger than the theoretical value. Therefore, MWCNT 3 denotes an MWCNT interconnect with contact resistance of 10 kΩ and kinetic inductance per channel of 60 nH/μm.

Figure 5.9 shows the time delay for 1000 μm-long intermediate level Cu and MWCNT interconnects with different number of inserted repeaters. The cases of h_{opt} and $h = 50$ are considered, and it is evident that minimum total time delay could be achieved by employing h_{opt}. The estimated values from the derived analytical expressions are also compared with the simulated results at 14 and 7 nm technology nodes, as listed in Tables 5.3 and 5.4. It is found that the minimum time delay increases significantly as the IC feature size scales down, and at the same time, the required number of repeaters increases. Moreover, it is found that MWCNT interconnects require less repeaters than their Cu counterpart, and therefore, the implementation of MWCNT interconnects could save power consumption and chip area. However, the contact resistance has a significant influence on the time delay of MWCNT interconnects at the 14 nm node, and its influence could be suppressed at 7 nm node.

Then, the optimal repeater insertion in global Cu and MWCNT interconnects are explored, as shown in Fig. 5.10. Similarly, the simulated and estimated k_{opt} for global level interconnect at the 14 and 7 nm technology nodes are compared in Tables 5.5 and 5.6. It is evident that the estimated values agree well with the simulated results. The contact resistance has a certain influence on the repeater design in global level MWCNT interconnects, implying that reducing the contact resistance is always desired.

Table 5.2 Circuit parameters of three types of MWCNT interconnects

Type	MWCNT 1	MWCNT 2	MWCNT 3
R_C (kΩ)	0.55 (intermediate level) 0.29 (global level)	10	10
$L_{K/channel}$ (nH/μm)	8	8	60

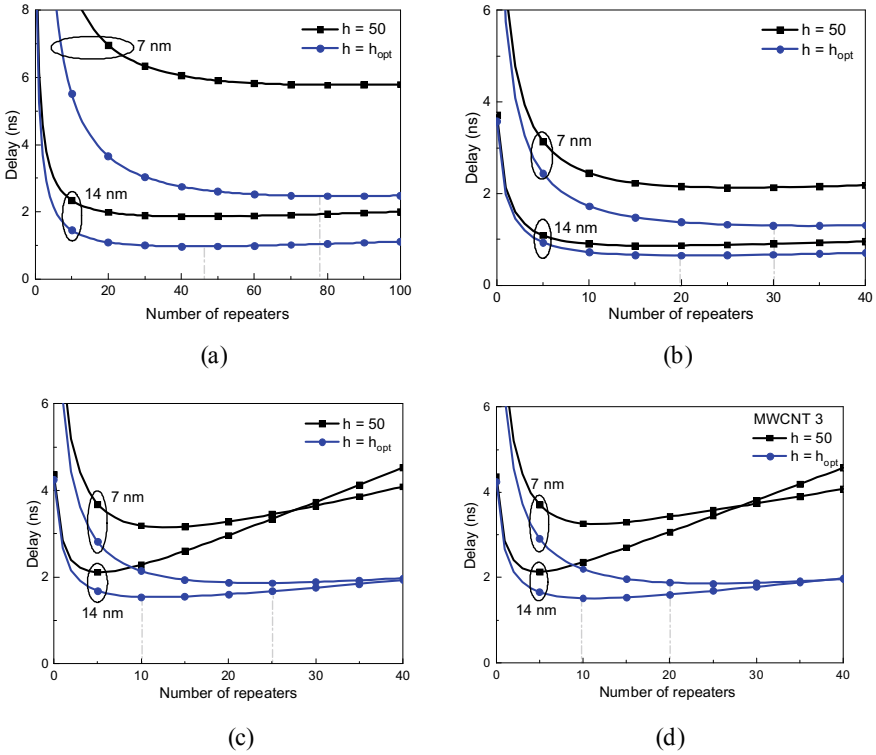


Fig. 5.9 Time delay versus number of repeaters for 1000 μm-long intermediate level Cu and MWCNT interconnects. **a** Cu; **b** MWCNT 1; **c** MWCNT 2; **d** MWCNT 3

Table 5.3 k_{opt} and h_{opt} for intermediate level interconnect at the 14 nm technology node

Type	Cu	MWCNT 1	MWCNT 2	MWCNT 3
Simulated k_{opt}	45	22	13	12
Estimated k_{opt}	44	20	12	10
Estimated h_{opt}	10	17	10	10
T_{total} (ns)	0.974	0.651	1.538	1.522

Table 5.4 k_{opt} and h_{opt} for intermediate level interconnect at the 7 nm technology node

Type	Cu	MWCNT 1	MWCNT 2	MWCNT 3
Simulated k_{opt}	84	37	25	26
Estimated k_{opt}	80	32	25	20
Estimated h_{opt}	8	12	11	7
T_{total} (ns)	2.473	1.303	1.865	1.890

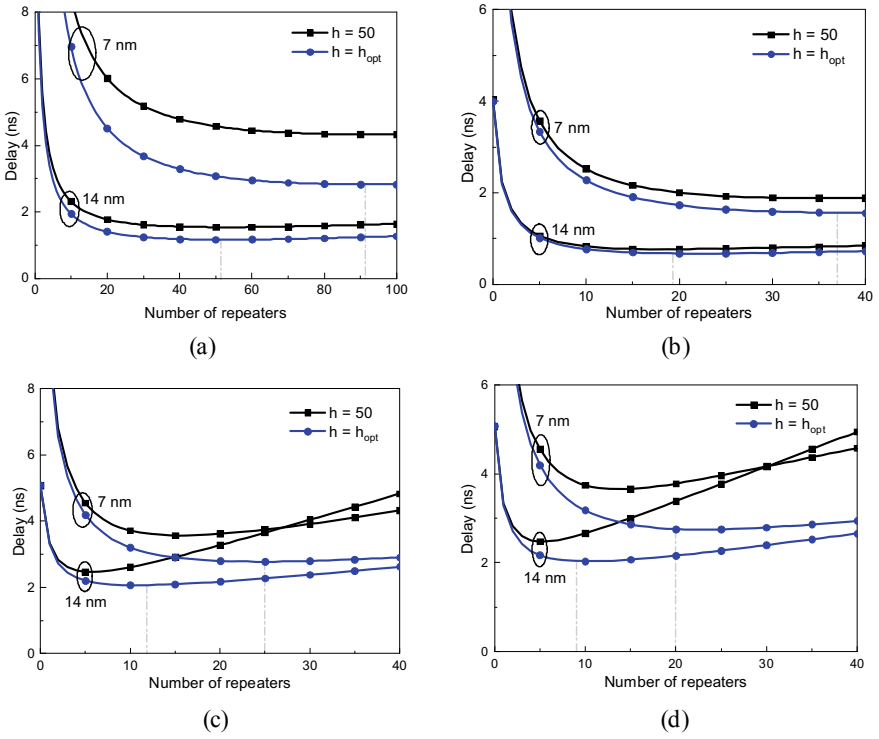


Fig. 5.10 Time delay versus number of repeaters for 2000 μm-long global level Cu and MWCNT interconnects. **a** Cu; **b** MWCNT 1; **c** MWCNT 2; **d** MWCNT 3

Table 5.5 k_{opt} and h_{opt} for global level interconnect at the 14 nm technology node

Type	Cu	MWCNT 1	MWCNT 2	MWCNT 3
Simulated k_{opt}	54	23	12	11
Estimated k_{opt}	52	20	13	10
Estimated h_{opt}	18	25	14	14
T_{total} (ns)	1.172	0.678	2.067	2.038

Table 5.6 k_{opt} and h_{opt} for global level interconnect at the 7 nm technology node

Type	Cu	MWCNT 1	MWCNT 2	MWCNT 3
Simulated k_{opt}	96	42	26	24
Estimated k_{opt}	92	38	26	21
Estimated h_{opt}	14	22	17	15
T_{total} (ns)	2.828	1.562	2.776	2.758

Repeater Design in CNT Interconnects to Reduce Delay and Power

The above analysis aims at reducing the time delay of CNT interconnects with the consideration of contact resistance. However, such delay-minimal repeater insertion methodology might result in overestimating the number of repeaters, which would lead to excessive power dissipation. Therefore, it is highly desirable to develop repeater design in on-chip interconnects with the consideration of both time delay and power dissipation. Similar with above analysis, equispaced repeaters are inserted in a long interconnect, as shown in Fig. 5.2, and the time delay of one segment is given as (Ismail and Friedman 2000)

$$\tau_{\text{seg}} = \left(1.48\xi + \exp\left(-2.9\xi^{1.35}\right)\right)\sqrt{\frac{Ll}{k}\left(\frac{Cl}{k} + hC_{\text{L0}}\right)} \tag{5.24}$$

where

$$\xi = \frac{R_t}{2}\sqrt{\frac{C}{L}}\frac{R_t + C_t + R_tC_t(1 + C_{\text{d0}}/C_{\text{L0}}) + 0.5}{\sqrt{1 + C_t}} \tag{5.25}$$

$R_t = R_{\text{d0}}/(h(R_Sl/k + 2R_C))$ and $C_t = hkC_{\text{L0}}/(Cl)$. The total time delay $T_{\text{total}} = k\tau_{\text{seg}}$. The power dissipation consumed by one segment can be approximated as

$$P_{\text{seg}} = \left(\frac{Cl}{k} + h(C_{\text{d0}} + C_{\text{L0}})\right)V_{\text{dd}}^2. \tag{5.26}$$

The total power dissipation is $P_{\text{total}} = kP_{\text{seg}}$. . Here, a figure-of-merit (FOM) is defined as (Li et al. 2020; Zhao et al. 2019b)

$$\text{FOM} = (P_{\text{total}})^p \cdot (T_{\text{total}})^q \tag{5.27}$$

where p and q are weighting factors, and they are (0, 1) for delay-optimal and (1, 0) for power-optimal repeater designs, respectively. For simplicity, they are set as 1 in this study.

Here, the particle-colony optimization (PSO) algorithm is employed for delay-power-product-optimal repeater designs in Cu and CNT interconnects. The PSO algorithm is a stochastic procedure and could realize search process in the space by mutual cooperation and competition among particles (Zhao et al. 2022b). The total number of particles in the swarm is $popsize$, the particle dimension is m, and the termination condition of the algorithm is $maxiter$. The flight speed and position of i th particle at time t in the search space are defined as

$$V_i(t) = [V_{i1}(t), V_{i2}(t), \ldots, V_{im}(t)]^T \tag{5.28}$$

$$X_i(t) = [X_{i1}(t), X_{i2}(t), \ldots, X_{im}(t)]^T. \tag{5.29}$$

The individual extremum and the population extremum of i th particle at time t in the search space are

$$pbest_i(t) = [p_{i1}(t), p_{i2}(t), \ldots, p_{im}(t)]^T \tag{5.30}$$

$$gbest_i(t) = [g_{i1}(t), g_{i2}(t), \ldots, g_{im}(t)]^T. \tag{5.31}$$

The update formulas for all particles can be expressed as

$$V_{i+1}(t+1) = W_{ei}V_i(t) + c_1r_1(pbest_i(t) - X_i(t)) + c_2r_2(gbest(t) - X_i(t)) \tag{5.32}$$

$$X_{i+1}(t+1) = X_i(t) + V_{i+1}(t+1) \tag{5.33}$$

where $W_{ei} = W_{\max} - (W_{\max} - W_{\min}) \cdot iter/maxiter$ is the inertia weighting coefficient, and $iter$ is the current number of iteration. Here, $W_{\max}$ and $W_{\min}$ are set as 0.9 and 0.4, respectively. c_1 and c_2 are the learning factors of the PSO algorithm, and they affect the self-learning ability and social learning ability of the particles. r_1 and r_2 are the random numbers in the range of [0, 1].

Figure 5.11 shows the flowchart of the PSO algorithm. The parameters of the particle such as m, W_{ei}, c_1, and c_2 as well as a group of particles including the flight speed and information of position are initialized. Here, m, c_1, and c_2 are set as 2, and popsize and maxiter are 50 and 1000, respectively. The fitness value of i th particle and best position of the swarm are set as current individual extremum $p\text{best}_i$, and the total extremum $g$best, respectively. The flight speed and position of the particle are updated by comparing the current value with best value found previously. The searching is finally stopped as the termination condition is satisfied, and the desired result would be output.

To verify the feasibility of the PSO algorithm, the global interconnects at 14 and 7 nm technology nodes are considered. The delay-optimal repeater designs ($p = 0$ and $q = 1$) are conducted using the PSO algorithm, and the results are compared with the following analytical expressions (Ismail and Friedman 2000)

$$h_{\text{opt}} = \sqrt{\frac{R_{\text{d0}}Cl}{(R_Sl + R_C)C_{\text{L0}}}} \frac{1}{\left[1 + 0.18\left(T_{\text{L/R}}\right)^3\right]^{0.26}} \tag{5.34}$$

$$k_{\text{opt}} = \text{Int}\left[\sqrt{\frac{(R_sl + R_C)Cl}{2R_{\text{d0}}(R_{\text{d0}} + C_{\text{L0}})}} \frac{1}{\left[1 + 0.21\left(T_{\text{L/R}}\right)^3\right]^{0.28}}\right] \tag{5.35}$$

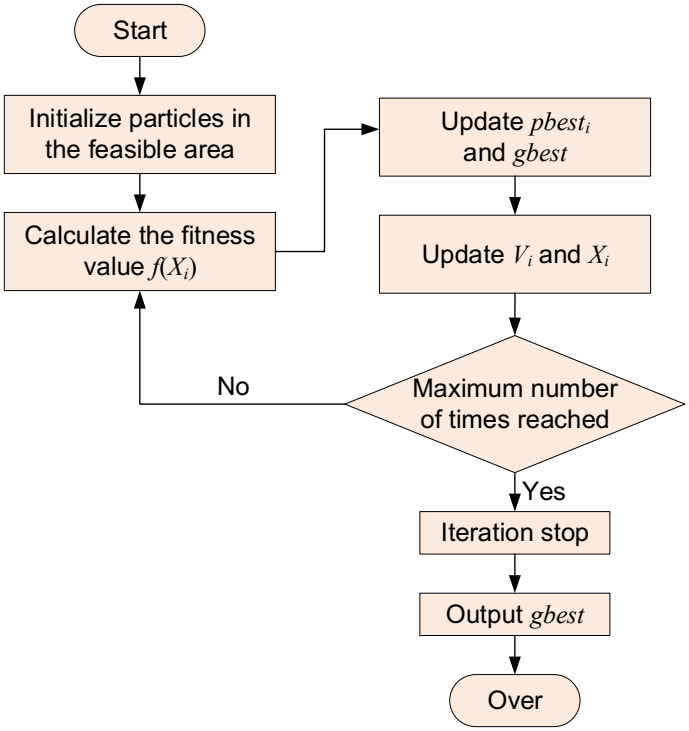


Fig. 5.11 Flowchart of the PSO algorithm

where

$$T_{\mathrm{L/R}} = \sqrt{\frac{Ll}{R_{\mathrm{d0}}(R_S l + R_C)(C_{\mathrm{d0}} + C_{\mathrm{L0}})}}. \tag{5.36}$$

As shown in Fig. 5.12, the optimal repeater numbers obtained using the PSO algorithm agree well with the values estimated by analytical expressions. It is evident that CNT interconnects require less repeaters than their Cu counterpart, thereby consuming less power. Moreover, as the technology node advances, the required number of repeaters would increase, implying that the power dissipation becomes more important in advanced technology node. Then, the optimal repeater designs aiming at minimum power-delay-product (PDP) are performed using the PSO algorithm, as shown in Fig. 5.13. The optimal repeater numbers predicted using the PSO algorithm are compared with those from the genetic algorithm (GA).

Furthermore, the influence of kinetic inductance on the optimal number of repeaters in CNT interconnects is explored. As aforementioned, the measured kinetic inductance of CNT might be 15 times larger than the theoretical value. As shown in Fig. 5.14, the variation in the kinetic inductance has negligible influence on

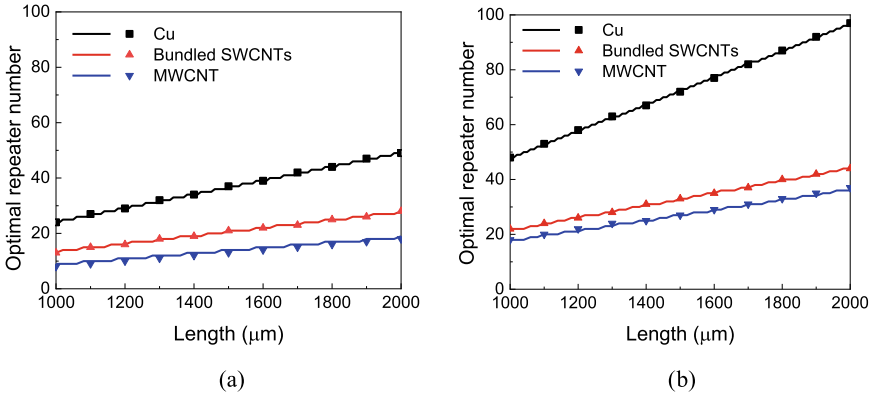


Fig. 5.12 Optimal repeater number in delay-optimal repeater design obtained by the PSO algorithm and analytical expression at **a** 14 nm node and **b** 7 nm node (line: PSO; symbol: expression)

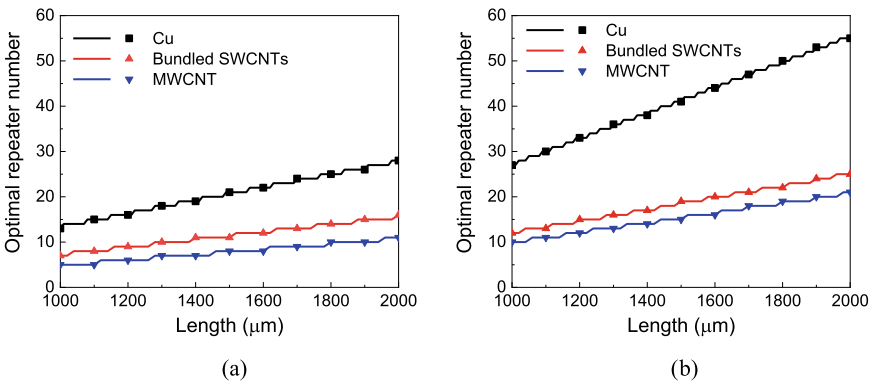


Fig. 5.13 Optimal repeater number in PDP-optimal repeater design obtained by the PSO and GA algorithms at **a** 14 nm node and **b** 7 nm node (line: PSO; symbol: GA)

the optimal number of repeaters in bundled SWCNTs due to its large number of conducting channels. However, the optimal repeater number decreases dramatically with increasing kinetic inductance for MWCNT interconnects, which is because that MWCNT has high MFP but low number of conducting channels. Therefore, MWCNTs are susceptible to the variation in the kinetic inductance. It is well known that the influence of inductance on the interconnect response tends to be trivial as the damping factor increases. As shown in Fig. 5.15, the influence of kinetic inductance would tend to decline with the increasing of imperfect contact resistance.

It is worth noting that although the PSO algorithm could be used in the repeater design, it is time-consuming for real-world applications. Artificial neural network (ANN) is a typical class of machine learning algorithm, and it gained attention as a fast and flexible vehicle to electronic designs. ANN mimics the structure of human

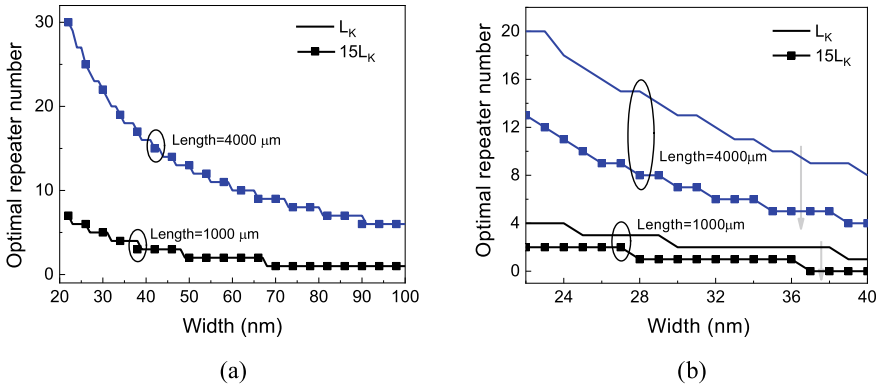


Fig. 5.14 Optimal number of repeaters in **a** bundled SWCNTs and **b** MWCNT with $R_{mc} = 0$ at the 14 nm technology node

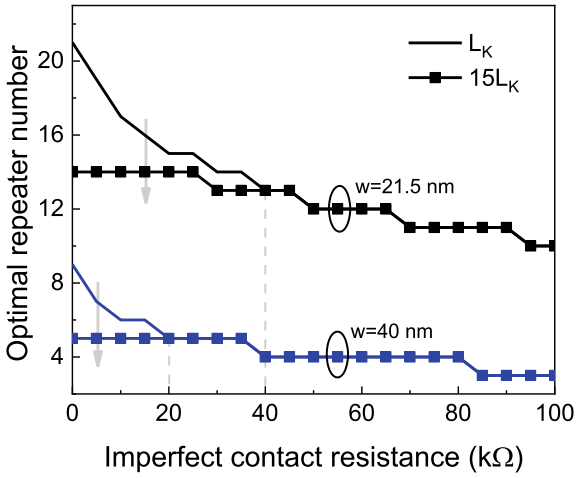


Fig. 5.15 Optimal number of repeaters in MWCNT interconnect with length of 4000 μm at the 14 nm technology node

synapse connections to imitate the brain's function. As shown in Fig. 5.16, a two-layer feedforward ANN is trained using the backpropagation algorithm, and the number of nodes in these layers are denoted as m and n, respectively. The process of information forward propagation and error backpropagation goes round and round. The input variables pass from the input layer to the hidden layers during the forward propagation phase, and they are processed by activation function and then enter to the output layer. The error backpropagation is utilized for updating the network parameters and minimizing the error function. The Levenberg–Marquardt (LM) algorithm is used for training the network as it could overcome the drawback that training the ANN using gradient method converges slowly and may fall into local minimum point.

The repeater designs in global level Cu, bundled SWCNTs, and MWCNT interconnects at the 14 nm technology node are conducted for establishing training

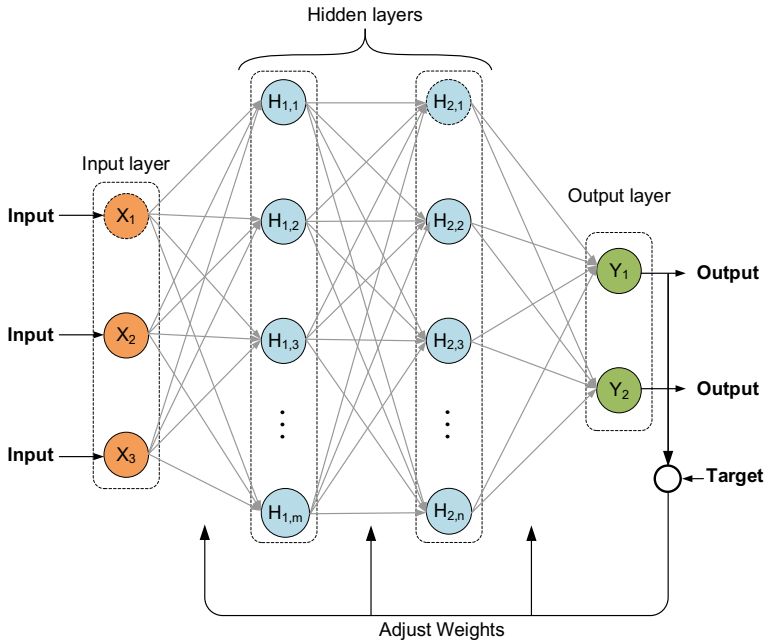
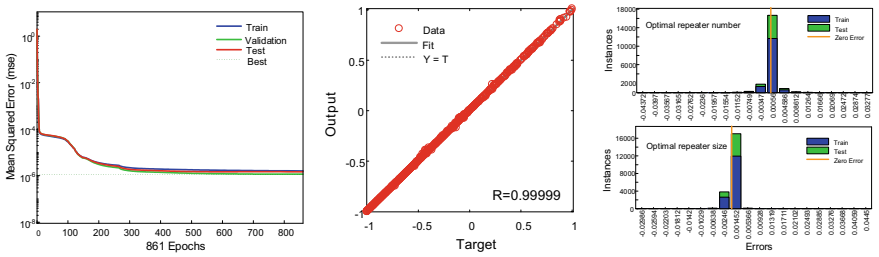


Fig. 5.16 Schematic of artificial neural network

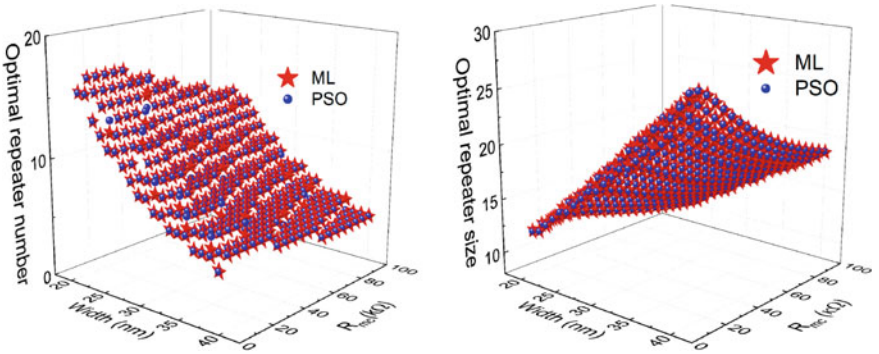
dataset. The input variables consist of interconnect width, interconnect length, and metal-CNT contact resistance, while the outputs include the optimal number of repeaters and optimal repeater size. The outputs are obtained using the PSO algorithm. Although the MWCNT interconnect could provide superior performance over the Cu wire, its diameter is limited, and therefore, its maximum value is set as 40 nm in the preparation of dataset. As normalization is a way to simplify the calculations and to speed up the convergence of the training network, the dataset is normalized to the range of [0, 1] by linearization. The number of nodes in two hidden layers are $m = 40$ and $n = 15$, respectively. Here, 70% of the input data is used as the training dataset, and the rest is used as the test dataset.

Figure 5.17a shows the number of iterations of the ANN to achieve convergence for MWCNT interconnects at the 14 technology, and the goal of mean square error (MSE) is set as 10^{-8}. The recursive process of the learning algorithm stops when the error cannot be reduced anymore. A comparison between the ANN outputs and desired targets is shown in Fig. 5.17b. The solid line represents the best-fit linear regression line between the outputs and targets, and the dashed line denotes the perfect results. The solid and dashed linear completely overlap with each other, implying a good fit for the repeater designs in MWCNT interconnects. The relative error between ANN outputs and targets is shown in Fig. 5.17c, and it is evident that the error of most samples is close to 0 (Fig. 5.18).



(a) (b) (c)

Fig. 5.17 **a** Squared error versus curves, **b** ANN output versus target regression curves, and **c** relative errors between outputs and targets for PDP-optimal repeater designs in MWCNT interconnects at the 14 nm technology node



(a) (b)

Fig. 5.18 **a** Optimal number of repeaters and **b** optimal repeater size of MWCNT interconnects obtained using PSO algorithm and trained ANN

Using the trained ANN, the optimal number of repeaters and optimal repeater size can be obtained readily, and the ANN outputs overlaid on the results from the PSO algorithm. It is found that ANN requires less simulation time than conventional methods. For instance, by using the PSO algorithm for handling tens of thousands of data, the CPU times would exceed 25, 25, and 55 h for repeater designs in Cu, bundled SWCNTs, and MWCNT interconnects, respectively. However, by using the trained ANN, the CPU times could be reduced to 4.34, 4.07, and 4.43, implying that the implementation of the ANN could effectively save computational time.

Conclusion

In this chapter, the repeater designs in CNT interconnects were investigated based on the distributed RC model. It was demonstrated that CNT interconnects require less repeaters than their Cu counterpart, thereby saving the chip area and power consumption. However, MWCNTs are susceptible to the contact resistance, which must be taken into account in the repeater designs in MWCNT interconnects. Then, the inductance was considered in the circuit model of MWCNT interconnect, and closed-form expressions were derived for predicting the optimal number of repeaters and repeater size for MWCNT interconnects. It was found that the influence of contact resistance on the repeater designs in MWCNT interconnect becomes more significant at advanced technology node, and therefore more fabrication efforts need to be pursued. Finally, the product of time delay and power dissipation of on-chip interconnects were used as the optimization objective. The PSO algorithm was developed to extract the optimal number of repeaters and repeater size, and it was verified against analytical expressions and GA. To avoid time-consuming numerical optimization, the dataset was developed by using the PSO algorithm for a specific technology node. An ANN was employed and trained to establish the relationship between the physical parameters and the optimal repeater design, and it was found that this procedure could dramatically reduce the computational time, and it could be applied to the optimal designs in future nanoscale CMOS and carbon-based ICs.

Acknowledgements This study was supported in part by the National Natural Science Foundation of China under Grant No. 62222401, and in part by the Zhejiang Provincial Natural Science Foundation under Grant No. LXR22F040001.

References

Avourise P, Appenzeller J, Martel R, Wind SJ (2003) Carbon nanotube electronics. Proc IEEE 91(11):1772–1784

Banerjee K, Mehrotra A (2002) A power-optimal repeater insertion methodology for global interconnects in nanometer designs. IEEE Trans Electron Devices 49(11):2001–2007

Ceyhan A, Naeemi A (2013) Cu interconnect limitations and opportunities for SWNT interconnects at the end of the roadmap. IEEE Trans Electron Devices 60(1):374–382

Close GF, Wong HSP (2008) Assembly and electrical characterization of multiwall carbon nanotube interconnects. IEEE Trans Nanotechnol 7(5):596–600

Close GF, Wong HSP (2007) Fabrication and characterization of carbon nanotube interconnects. IEDM Tech Dig 203–206

Chai C, Hazeghi A, Takei K, Chen HY, Chan PCH, Javey A, Wong HSP (2012) Low-resistance electrical contact to carbon nanotubes with graphitic interfacial layer. IEEE Trans Electron Devices 59(1):12–19

Chiariello AG, Maffucci A, Miano G (2013) Circuit models of carbon-based interconnects for nanopackaging. IEEE Trans Compon Packag Manuf Technol 3(11):1926–1937

Chen G, Friedman FG (2006) Low-power repeaters driving RC and RLC interconnects with delay and bandwidth constraints. IEEE Trans Very Large Scale Integr (VLSI) Syst 14(2): 161–172

Guistininai A, Tucci V, Zamboni W (2010) Modeling issues and performance analysis of high-speed interconnects based on a bundle of SWCNT. IEEE Trans Electron Devices 57(8):1978–1986
Harutyunyan AR, Chen G, Paronyan TM, Pigos EM, Kuznetsov OA, Hewaparakrama K, Kim SM, Zakharov D, Stach EA, Sumanasekera GU (2009) Preferential growth of single-walled carbon nanotubes with metallic conductivity. Science 326(5949):116–120
Ismail YI, Friedman EG (2000) Effects of inductance on the propagation delay and repeater insertion in VLSI circuits. IEEE Trans Very Large Scale Integr (VLSI) Syst 8(2): 195–206
Im S, Srivastava N, Banerjee K, Goodson KE (2005) Scaling analysis of multilevel interconnect temperatures for high-performance ICs. IEEE Trans Electron Devices 52(12):2710–2719
Liang F, Wang G, Lin H (2011a) Modeling of crosstalk effects in multiwall carbon nanotube interconnects. IEEE Trans Electromagnetic Compatibility 54(1):133–139
Liang F, Wang G, Ding W (2011b) Estimation of time delay and repeater insertion in multiwall carbon nanotube interconnects. IEEE Tarns Electron Devices 58(8):2712–2720
Li XC, Mao JF, Huang HF, Liu Y (2005) Global interconnect width and spacing optimization for latency, bandwidth and power dissipation. IEEE Trans Electron Devices 52(10):2272–2279
Li H, Yin WY, Banerjee K, Mao JF (2008) Circuit modeling and performance analysis of multi-walled carbon nanotube interconnects. IEEE Trans Electron Devices 55(6):1328–1337
Li W, Zhao WS, Liu PW, Wang J, Wang G (2020) Optimal repeater insertion for horizontal and vertical graphene nanoribbon interconnects. Int J Numer Model 33:e2696
Lu Q, Zhu Z, Yang Y, Ding R (2016) Analysis of propagation delay and repeater insertion in single-walled carbon nanotube bundle interconnects. Microelectron J 54:85–92
Liu PW, Cheng ZH, Zhao WS, Lu Q, Zhu Z, Wang G (2018) Repeater insertion for multi-walled carbon nanotube interconnects. Appl Sci 8:236
Meindl JD (2003) Beyond Moore's law: the interconnect era. Comput Sci Eng 5:20–24
Majumder MK, Pandya ND, Kaushik BK, Manhas SK (2012) Analysis of MWCNT and bundled SWCNT interconnects: Impact on crosstalk and area. IEEE Electron Device Lett 33(8):1180–1182
Maffucci A, Miano G, Villone F (2009) A new circuit model for carbon nanotube interconnects with diameter-dependent parameters. IEEE Trans Nanotechnol 8(3):345–354
Naeemi A, Meindl JD (2007) Design and performance modeling for single-walled carbon nanotubes as local, semiglobal, and global interconnects in gigascale integrated systems. IEEE Trans Electron Devices 54(1):26–37
Radosavljevic M, Lefebvre J, Johnson AT (2001) High-field electrical transport and breakdown in bundles of single-wall carbon nanotubes. Phys Rev B 64(24):241307
Rakheja S, Naeemi A (2010) Interconnects for novel state variables: performance modeling and device and circuit implications. IEEE Trans Electron Devices 57(10):2711–2718
Sarto MS, Tamburrano A (2009) Single-conductor transmission-line model of multiwall carbon nanotubes. IEEE Trans Nanotechnol 9(1):82–92
Sathyakam PU, Mallick PS (2017) Carbon nanotube interconnects with air-gaps: effect on thermal stability, delay and area. J Nano Res 48:29–37
Srivastava N, Joshi RV, Banerjee K (2005) Carbon nanotube interconnects: implications for performance, power dissipation and thermal management. IEDM Tech Dig 249–252
Tang M, Mao J (2014) Modeling and fast simulation of multiwalled carbon nanotube interconnects. IEEE Trans Electromagnetic Compatibility 57(2):232–240
Todri-Sanial A, Dijon J, Maffucci A (2017) Carbon nanotubes for interconnects. Springer, Zurich, Switzerland
Venkatesan R, Davis JD, Meindl JD (2003) Compact distributed RLC interconnect models—Part IV: Unified models for time delay, crosstalk and repeater insertion. IEEE Trans Electron Devices 50(4):1094–1102
Venkatesan JAR, Kaloyeros A, Beylansky M, Souri SJ, Banerjee K, Saraswat KC, Rahman A, Reif R, Meindl JD (2001) Interconnect limits on gigascale integration (GSI) in the 21st century. Proc IEEE 89(3):305–324

Wei BQ, Vajtai R, Ajayan PM (2001) Reliability and current carrying capacity of carbon nanotubes. Appl Phys Lett 79(8):1172–1174

Wong HSP, Akarvardar K, Antoniadis D, Bokor J, Hu C, King-Liu TJ, Mitra S, Plummer JD, Salahuddin S (2020) A density metric for semiconductor technology. Proc IEEE 108(4):478–482

Zhao WS, Fu K, Wang DW, Li M, Wang G, Yin WY (2019a) Mini-review: modeling and performance analysis of nanocarbon interconnects. Appl Sci 9(11):2174

Zhao WS, Liu PW, Yu H, Hu Y, Wang G, Swaminathan M (2019b) Repeater insertion to reduce delay and power in copper and carbon nanotubes-based nanointerconnects. IEEE Access 7:13622–13633

Zhao WS, Zhang R, Wang DW (2022a) Recent progress in physics-based modeling of electromigration in integrated circuit interconnects. Micromachines 13(883):1–31

Zhao WS, Wang BX, Wang DW, You B, Liu Q, Wang G (2022b) Swarm intelligence algorithm-based optimal design of microwave microfluidic sensors. IEEE Trans Industrial Electron 69(2):2077–2087

Zhao WS, Wang G, Hu J, Sun L, Hong H (2015) Performance and stability of monolayer single-walled carbon nanotube interconnects. Int J Numer Model 28:456–464

Zhao WS, Wang G, Sun L, Yin WY, Guo YX (2014) Repeater insertion for carbon nanotube interconnects. Micro Nano Lett 9(5):337–339

Chapter 6
Through Silicon Vias for 3D Integration—A Mini Review

Yachana Arora, Vandana Boora, Rohit Dhiman, and Rajeevan Chandel

Abstract Three-dimensional integration of integrated circuits (ICs) has shown a better utilization of space on the substrate by stacking the chips one onto another than its planar counterparts. This effective use of space and thus miniaturization has led to a holistic performance enhancement of the very large-scale integration (VLSI) systems. Vertical interconnects or through silicon vias (TSVs) are the foundation of 3D integration. The fabrication technologies required to incorporate the through silicon vias in the substrate and the challenges associated with their fabrication are significantly important, as these may hugely affect the final product performance and cost. To increase the reliability and maintain the signal integrity of an overall IC, various novel structures of TSVs have been devised by the researchers, wherein it is attempted to find the better geometry, shape, and filler materials for the vias. Subsequently, a brief overview of the different aspects of TSVs for 3D integration is presented in the current chapter. The electrical characteristics of the vias are highlighted, considering the lumped equivalent model which may be analyzed using various numerical techniques.

Keywords Through Silicon Vias (TSVs) · Signal Integrity · 3D Integration · Reliability · Via Fabrication · Equivalent Circuit · Stacking · Modeling

Y. Arora (✉) · V. Boora · R. Dhiman · R. Chandel
Department of Electronics and Communication Engineering, National Institute of Technology Hamirpur, Hamirpur, Himachal Pradesh 177 005, India
e-mail: 184533@nith.ac.in

V. Boora
e-mail: 184535@nith.ac.in

R. Dhiman
e-mail: rohitdhiman@nith.ac.in

R. Chandel
e-mail: rchandel@nith.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_6

Introduction to 3D IC

In order to meet the goals of power reduction and high integration in very large-scale integration (VLSI) systems, a novel chip-stacking packaging technique known as the three-dimensional integrated circuit (3D IC) has been developed recently. When many layers of active devices are stacked and connected vertically to create 3D ICs, the process is known as 3D integration packaging technology which is many times also referred to as vertical integration.

This approach has potential to greatly enhance chip functionality, packing density, and performance (Guarini et al. 2002). 3D ICs may facilitate the integration of diverse materials, devices, and signals and also offer the microchip architecture (Topol et al. 2006). By utilizing quick, dense inter die vias, 3D ICs aim to break through interconnect scaling obstacles, opening the door for further improvements in complementary metal-oxide semiconductors (CMOS) performance.

Furthermore, 3D ICs allow for the fusion of several fabrication techniques to create a genuine system on chip (SoC) by integrating them on the same chip and shrinking the form factor (Lee and Chakrabarty 2009).

3D integration presents a graceful approach to the off-chip connection difficulty by physically bringing the communication ICs closer to one another and concurrently lowering the area footprint, which is critical for handheld devices. Given this, 3D integration is thought to be an effective way to deal with the off-chip connectivity problems (2018).

Challenges in 3D Integration

Though 3D ICs offer better space utilization, decreased resistance-capacitance delay, miniaturization, etc., but the fact that these are complex structures to fabricate, and test cannot be ruled out. There are several challenges in the way to design a highly reliable 3D structure. The following are some of the problems that a designer needs to address before jumping to a final design for its system on chip.

Testing of IC and Its Yield

When dies are fabricated on the wafer, it is tested for each of the dies using a probe. The die that passes the test is called a known good die (KGD) and only KGD are involved to form the 3D stack. Another method of assembly is done at wafer level where the die is packaged while it is still on wafer and then diced. This method is far cheaper than individual testing of die, but it can lead to loss in yield if any of the die in the wafer breaks down.

Heat Removal

Power consumed by the IC and how effectively it is cooled down are the important factors that determine efficiency and longevity of the electronic gadgets. With the 3D integrated topology, the distance between the die at the top level and heat sink is significantly increased. Improper thermal management may lead to junction leakage and electro-migration may deteriorate device performance. For every 10 ºC rise in junction temperature, there is a decrease in performance of clock buffer by 1.2% (Im and Banerjee 2000). Hence, thermal management in IC is a critical issue that needs to be addressed to decrease the thermal hotspots in 3D IC. Ref. (Tavakkoli et al. 2016) suggested the use of liquid cooling technique using microfluidic channels to address this issue.

Reliability

3D stacking of chips requires vertical interconnections for signal and power supply. This vertical interconnection goes directly through the substrate of the chip, hence the name through silicon via (TSV). There are already fabricated components present on the substrate, hence the TSV fabrication must be carefully done without disturbing the surrounding environment of the substrate.

TSVs can be fabricated using various processes. The selection of the fabrication process must be carefully done as it could affect the stress profile of the substrate, thereby posing severe challenges to the already fabricated transistor. For these 3D ICs to function properly and possibly set a performance ceiling, the on-chip wires connecting the inputs/outputs (I/Os) to the TSVs must be reliable.

Power Distribution Networks

In 3D ICs, power needs to be vertically transported from package pins through the stacks. Vertical vias called TSVs are used to distribute the power/ground to the die present higher up in the stack. Improving the internal power grid/network will greatly influence the voltage drop across the chip and the chip area requirements (Kim et al. 2010).

Many dedicated power/grounds TSVs may be required for the chip stack, but this will add to increased crosstalk problems. Hence, an innovative 3D power delivery solution is essential to provide high-quality voltage levels.

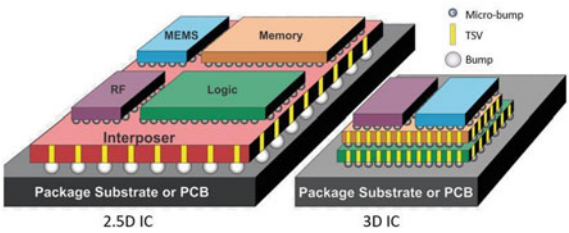


Fig. 6.1 Different placement of chips in 2.5D and 3D IC (Medhat 2022)

Placement of Chips and Vias

The goal of the designer is to get a significantly improved utilization of substrate area, and improved wire length to have good signal integrity, lesser stress, and a good thermal profile. Therefore, the designer needs to pre-decide the placement of chips and vias onto the substrate to get better performance benefits. Figure 6.1 shows the placement of chips in 2.5D and 3D integration depicting better space utilization and lesser footprint due to the different placement of chips.

Clock Distribution Network

The performance of the entire SoC is highly dependent on how the clock network is distributed in the system. Data is transmitted synchronously between the data lines using a clock distribution network (CDN). To satisfy performance parameters like clock skew/slew and power, conventional clock network designs must be expanded to include the resistance-capacitance properties of the via.

Signal Integrity

There are several TSVs used in the 3D stack. If placed closer to each other, it will lead to TSV-to-TSV noise coupling and crosstalk. To get a good signal integrity (SI) measure, both insertion loss and noise coupling need to be taken care of. Ref. (Cho and Kim 2013) analyzed that TSV noise coupling is a more significant factor than insertion loss. It was observed that for TSV diameter of 30μm and height of 100μm, noise coupling is about −35dB at 1GHz reflecting the fact that interconnect geometry, its placement inside the substrate, and signaling model play a significant role in maintaining the integrity of the signal.

Via Reliability

Vertical vias form the fundamental technology of 3D integration and its reliability needs to be addressed. Thermal profile, stress induced, via placement, via fabrication process and techniques add up to the reliability of the via. Numerical methods that are conventionally used to capture these via failures are not always computationally feasible for chip-level analysis and optimization. Therefore, researchers are coming up with many possible structures which could improve the signal integrity issues of vertical vias (Rahman et al. 2013; Qian et al. 2016, 2017; Rao 2020; Hu et al. 2020; Kim et al. 2015; Kumar et al. 2022; Kumar and Dhiman 2022; Nilsson et al. 2009, 2021; Liang et al. 2013; Lu et al. 2018; Ruehli 1996).

Via Fabrication Process and Technologies

Via Fabrication Steps

Various processes are carried out to fabricate a TSV onto the substrate. Etching for via production is one of the numerous phases in the fabrication of vias and the deposition of oxide layer is carried out to form insulation between the two conducting materials. A barrier or seed layer is laid down followed by the addition of filler material, such as copper (Cu). The unwanted materials from the Cu plating are removed using the chemical mechanical polishing (CMP) technique. Each of the steps is explained pictorially in Fig. 6.2 and detailed in the following sub-sections.

Etching

The silicon is cut into deep pits with a large aspect ratio, and then filler material is inserted into them. A trench typically measures 100–150 μm in height and 1–5 μm in breadth. Etching is the method used to create these trenches. A thin film of a substrate is selectively removed, allowing the material to be reconfigured into the desired pattern. Depending on the desired profile, the etching method may be wet, dry, laser drilled, reactive ion, or deep reactive ion. Anisotropic wet etching is typically utilized for TSV production with a very big pitch to create a trench. However, the desired profile can be altered by combining isotropic wet etching and dry etching (Garrou et al. 2014). Laser drilling and reactive-ion etching (RIE) are the main processes used in dry etching. The RIE technique typically uses a high-density plasma source, where the plasma's ions and radial species chemically and physically etch the surface.

Following etching, the residual byproducts are eliminated, resulting in continuous plasma reactivity. Scientist Bosch patented the Bosch technique, also known as DRIE

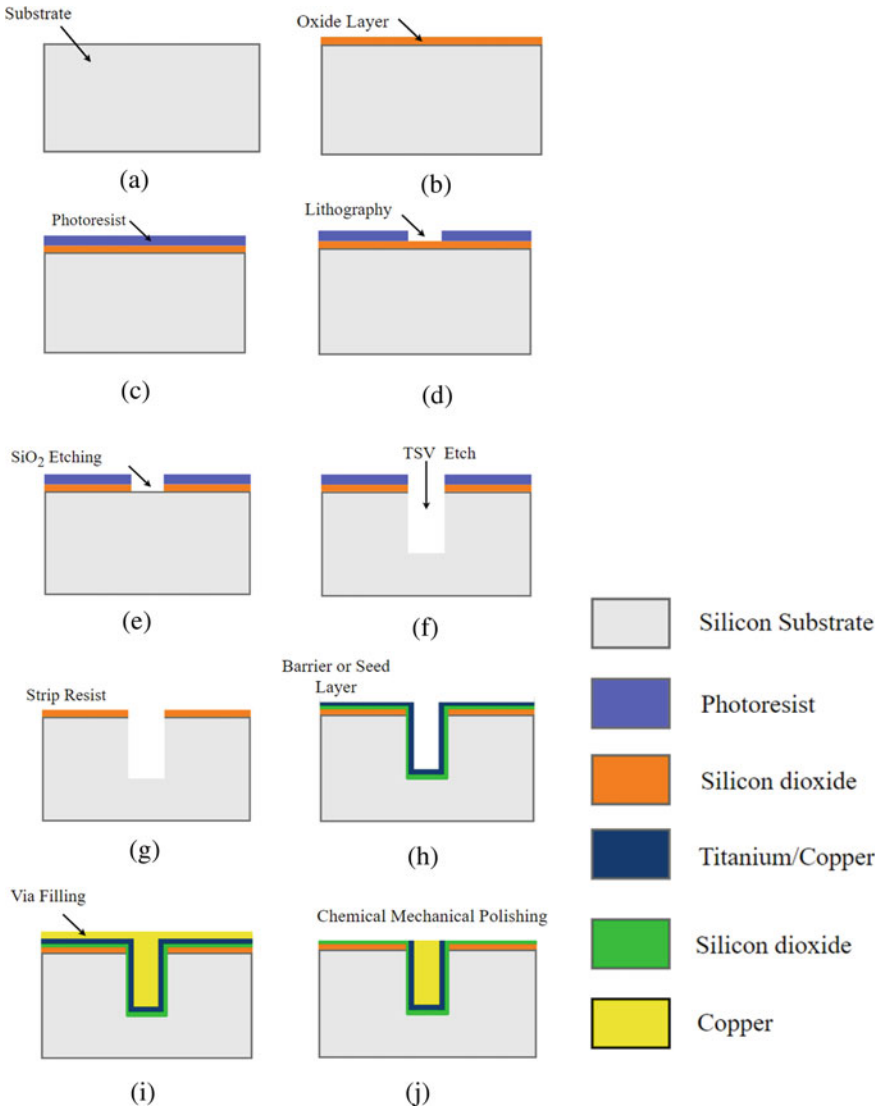


Fig. 6.2 Via fabrication steps: **a** Silicon substrate **b** Deposition of oxide layer on the substrate by thermal oxidation **c** Photoresist **d** Lithography to mark pattern of via **e** Oxide etch **f** Etching of TSV using technique such as dry etching **g** Strip resist **h** Seed/barrier layer of titanium (Ti), oxide layer deposition by plasma etch technique **i** Copper plating/Via filling **j** CMP to remove unwanted material from top

or the Bosch etching alternative, in response to the slowness of RIE method (Laermer et al. 2003).

Deposition of the Oxide Layer

The signal will be lost if the filler material, which is metallic, is placed directly into a silicon trench. A further insulating layer is therefore needed between the trench and filler material which requires some deposition techniques. A few of the deposition techniques include atomic layer deposition (ALD) and plasma-enhanced chemical vapor deposition (PECVD). These methods offer good adhesion with silicon at temperature range between 100 and 200 °C.

ALD produces oxide layers with low breakdown voltage. Therefore, for the deposition of oxide layers in TSVs, high-quality PECVD is preferred. The chemical vapor deposition (CVD) method is perfect for dielectric applications in TSV-based interconnects. This method demonstrates an inherently conformal deposition that is essential for covering the titanium metal barrier and copper seed in the next steps. Organic dielectrics can be processed using deposition methods such as spin coating and spraying. Because of their low modulus and ability to be coated at ambient temperature, these materials function as a stress buffer layer.

Seed Layer

A metal barrier or seed layer is added after the oxide isolation layer has been deposited. These layers are there to stop metal from diffusing into the silicon or the oxide layer, hence the layer should be uniform. To achieve equal thicknesses on the top and sides during the application of the barrier layer, the oxide must be deposited. The most difficult and expensive activity in the TSV fabrication process flow is the application of the barrier layer and subsequent through-filling. Titanium (Ti) and tantalum (Ta) are the best barrier materials. These materials, which can provide extremely uniform step coverage and sheet resistance, are deposited via a PVD method. As a result, the barrier for depositing a film is thinner. A physical vapor deposition (PVD) of tantalum nitride and seed copper can be used, followed by a full metal deposition (Riedel et al. 2000). Ionized metal plasma (IMP) sputtering offers superior step coverage due to the directionality of the deposited atoms and utilization of ion bombardment to sputter materials from the bottom of the via to the sidewall. IMP-based PVD enables copper seed layer conformality either on the via sidewall or at the via bottom.

Via Filling/Plating

The traditional wafer manufacturing procedures use tungsten as via filler to fill the connection of 1μm diameter and a depth of 2 μm. Sputtered titanium is typically

used as the contact and barrier layer for the filling. For the via to be filled with tungsten, more than half of its diameter must be deposited since tungsten filling is conformal. It can be referred to as the common, high-throughput wafer production process. However, film stress and wafer bow are the potential concerns with tungsten TSV filling, which limits the thickness that can be deposited. Hence, only tungsten is used for small-diameter holes and annular vias, while polysilicon is no longer commonly considered for use in TSVs due to concerns over via resistance (2016).

A layer of oxide, such as silicon dioxide (SiO_2), is necessary for the commonly used electrodeposition of copper to create a conducting channel in the substrate. However, it is complicated in several ways, including dependability, throughput, and process controllability. Implementing high-aspect-ratio TSVs with conductive TSV metal cores that are void-free is very challenging. Therefore, research is necessary to discover some alternative plating processes. The use of solder balls, filling with conductive metal pastes, wire-bonded gold cores, and other potential methods are a few of these.

Chemical Mechanical Polishing

After filler material is added to the via, an unwanted layer of metal or oxide is deposited on the ends of TSV, which is removed using the process called CMP. The name is so because CMP uses both mechanical abrasions of the surface metal with a mechanical polishing pad and an abrasive colloidal slurry. This combines mechanical material removal and a chemical reaction to remove excessive metal. Since alumina is more similar in hardness to tungsten than most other abrasives, it is the most often utilized abrasive material for tungsten CMP. The removal of tungsten involves mechanical abrasion followed by continual, self-limiting oxidation of the tungsten surface (Liu et al. 1996). An aqueous solution with nanoparticles with a diameter of several hundred nanometers is used to polish copper. Nitric acid, hydrogen peroxide, and ammonium hydroxide are common ingredients in slurry solutions. Because copper is a soft material as opposed to tungsten, mechanical effects play a big part in the polishing process.

There are two steps to the polishing process: An oxide with a higher removal rate, excellent polarization, and little nonuniformity is utilized in the first step. The second phase, which often selectively dishes to the barrier and has a lower removal rate, continues the process (Xu and Lu 2013). One of the challenges this polishing approach has is the quick removal of thick materials without compromising wafer topography. By reducing the size of the TSV and developing bottom-up copper plating, it is possible to reduce the copper overburden and the polishing time.

Wafer Thinning

Wafer thinning is done to maintain the total thickness variation (TTV) of the wafer. Minimization of this thickness and adhesive are done at the initial bonding step,

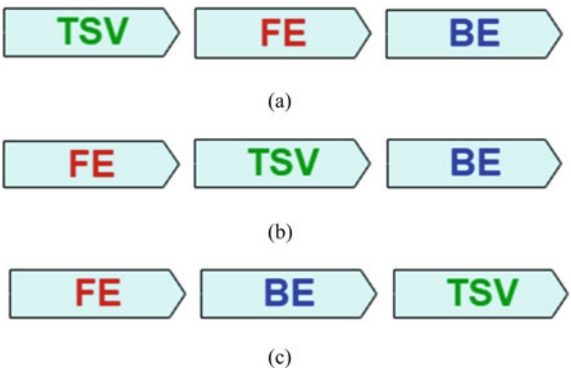


Fig. 6.3 Via fabricating techniques **a** Via-first **b** Via-middle **c** Via-last

hence a careful measurement of TTV is required. Wafers are typically thinned in two steps: (i) adding metal films to the back to facilitate backside contact with the device, and (ii) lowering the TSV depth. Silicon substrates can be thinned after being bonded together to either make room for the formation of an interconnect on the back of the bonded wafer or to make an existing connection visible for bonding to another substrate. Wafer thinning generally involves a procedure called back grinding. To protect the front side of the device wafer during this operation, a "back grind tape" with a thickness of roughly 100–300 μm is placed (Hosali et al. 2008).

TSV Process Technologies

Vias are fabricated mainly by three techniques, namely, via-first, via-middle, and via-last. These are depicted in Fig. 6.3 and are detailed in further sub-sections.

Via-First

This approach produces TSVs by printing the transistors in silicon prior to the front-end (FE) process as shown in Fig. 6.3a. For TSVs created utilizing the via-first method, high temperatures (about 1000°C) are reached during transistor manufacture. Consequently, the coefficient of thermal expansion (CTE) of the TSV materials must be decreased.

When employing the via-first technique, traditional materials with high CTEs such as copper (177 ppm/K) and tungsten (4.6 ppm/K) are used. Polysilicon is widely used as a filler because of its tolerance to higher temperatures and fewer impurities. However, the TSV resistance of polysilicon material is very high which limits its use, leading TSV fabrication designers to ignore the via-first method.

Via-Middle

Figure 6.3b depicts the fabrication of TSVs after the front-end (FE) process but before the back end (BE) process, or after the transistors are created but prior to the design of the metallization layers. The ideal materials for the via-middle approach are copper or tungsten. Tungsten is widely used as the filler material to produce on-chip metallization layers, which require temperature of about 400 °C, due to its lower CTE than Cu.

Via-Last

Using via-last technology, TSVs are produced by following FE and BE methods as seen in Fig. 6.3c. Because high-temperature FE and BE processes are completed before TSV production, Cu can be used as the filler material due to its low resistivity. However, further limitations on the usage of Cu include material diffusion in the silicon substrate and electro-migration effects at high frequencies. All these issues make the system unreliable. The materials employed in the current research initiatives must therefore be carbon-based and possess low CTE, resistivity, electro-migration effect, and contamination.

The via-last technique has been determined to be the most efficient way to make TSVs with regular Cu filler. This approach can also be used with image sensors and layered dynamic random-access memory (DRAM). The low electrical resistance of the vias serves as the primary defense for this method. For instance, a CMOS image sensor displays vias with a 2:1 aspect ratio and a diameter greater than 40 μm. Other devices have a 5:1 aspect ratio and through diameters between 10 and 25 μm.

Progress in 3D ICs

3D ICs offer enhanced performance owing to utilization of the z-direction of the chips. As the boundaries for scaling CMOS continue to rise, system designers are increasingly looking for multichip packages to improve performance without packing more devices on a single chip. To connect the various chips in stacked-die packages, wire bonds are frequently formed from the chip edges to the substrate. This technique still has latency, bandwidth, and power issues due to increasing circuit complexity and density, as well as the resulting increase in signal traffic on the interconnect. New 3D chip-stacking techniques that create multilayer connections directly through active circuitry now promise to significantly improve system performance. The best way to progress toward “More-than-Moore” technologies is through 3D IC packaging, which stacks dies vertically. Overall, three-dimensional integration can be divided into two approaches as follows:

One-dimensional approach: It entails sequential device processing. Prior to the back-end processing, that connects the devices via interconnects, the front-end processing, which creates layers of devices, is repeated multiple times on a single wafer to produce multiple active device layers.

Stacking approach: Wafer-to-wafer, wafer-to-die, and die-to-die are the three stacking techniques. In stacking approach, each layer is handled independently using traditional manufacturing methods. Bonding technology is used to join these processed layers together to form a 3D IC.

Traditional packaging technology uses wire bonding technique while TSVs and micro bumps are used for 3D stacking of chips. TSVs have been able to provide the highest integration density as it has led to shorter signal paths with improved electrical performance due to reduced capacitive, inductive, and resistive components. The TSVs in 3D ICs also have the benefit of having fine pitches. Due to its smallest space, connectivity density, and smallest form factor, this technology offers the best performance.

As a result, the primary goal of the current scenario is to develop TSVs that are both trustworthy and cost-effective. Wafer thinning and die bonding are the two main procedures used in TSV fabrication for 3D stacking. The wafer thinning procedure helps to reduce the impact of TSVs on the surrounding environment. It is not expected that the scaling rate of TSV dimensions in TSV bonding will match feature size. This is due to the alignment tolerance during bonding, which limits TSV scaling. However, the monolithic approach to 3D integration allows TSV scaling to be equivalent to feature size. This is because the connections make use of local wires.

Literature Survey on Via Modeling

To keep up with the demand of modern technology requirements, the advanced ICs require major changes both in terms of their design and the way these are packaged. By reducing resistance-capacitance delays in comparison to a conventional planar connection design, the 3D integration has been viewed for enhanced chip performance. Use of TSV for 3D integration is viewed as a solution to reach "More than Moore" demands. These interconnects travel through the substrate directly, thereby decreasing the interconnect length required. Since TSV conducts signal through it, it is filled with some conducting materials. TSV with copper fillers are currently being used in industry as they have been able to provide important benefits for increasing chip functionality while maintaining the same die size. However, it has electrical and mechanical issues as well. Carbon nanotubes (CNTs) are seen as an alternative to copper as they offer higher current carrying capacity, high resistance toward electro-migration, and very high mechanical and thermal strength compared to copper (Awano et al. 2010). Besides material, geometry of TSV and signaling scheme in TSV are also considered to have a significant effect on the performance.

In Ref. (Rahman et al. 2013), TSV is analyzed with optimization in its structural geometry such as cylindrical, conical, and square. It is noted that the performance of TSV is greatly influenced by design parameters and structural geometry and concluded that insertion loss varied least in cylindrical TSV. The performance of the cylindrical TSV is superior to that of other types of structures in terms of TSV characteristics (height, diameter, and insulation gap). An equivalent lumped model is also suggested which can be used to simulate 3D circuits.

The behavior of TSV with differential signaling scheme was studied in Qian et al. (2016) and compared it with single-ended TSV. The transmission characteristics of TSV were found to be highly affected by the thickness of dielectric at high frequencies and conductor loss at lower frequencies due to finite conductivity of via material. An increase in dielectric thickness from 100 to 400 nm led to an improvement in magnitude of insertion loss (S_{21}) at 5 GHz. This so happened because increase in dielectric thickness decreases the TSV conductance G and thereby dielectric loss. However, the effect of dielectric thickness is relatively smaller in the frequency region of under 1 GHz and over 50 GHz, since the overall loss is dominated by resistance-dependent conductor loss and substrate dependent dielectric loss, respectively, at these regions.

The doping concentration of the silicon substrate is another important factor. When the doping concentration varies from 1×10^{22} to 4×10^{22} m^{-3}, the magnitude of S_{21} at a frequency of 50 GHz decreases by 205.76% and 203.57% in odd- and even-mode excitations, respectively. In contrast, the effect of doping concentration is negligible below 5 GHz due to dominance of conductor and dielectric losses.

Rao has modeled a cylindrical profile that uses both copper and CNT (Rao 2020). After CNT (both single-walled and multi-walled) is grown inside TSV, it is filled with copper. Now the performance of TSV also depends on the ratio in which Cu and CNT are filled and the proportion of multi-walled and single-walled CNT.

It is recommended that to get an improved signal, a higher proportion of MWCNTs, and copper must be filled. To achieve higher current density along the surface model, a higher proportion of SWCNTs and lesser copper must be present. This model showed similar values for crosstalk issues, but improved results were seen for delay model.

Qian et al. (2017) have investigated the electrical–thermal performance of through glass vias filled with CNTs and a tapered geometry in S-G-S type signaling scheme. A property of glass via was observed that inductive element of conductor loss is dominant when used a glass substrate unlike silicon TSV. The signal loss in differential scheme is found to be less than in common mode, while it increases with increase in via pitch.

Fabricating several TSVs in the substrate adds to crosstalk issues but providing signal, power, ground to each layer of the stack is also essential. So, aiming at the problem of input/output (I/O) limitation, a differential multi-bit carbon nanotube TSV is reported in Hu et al. (2020) where each TSV carries two types of signals, hence the name multibit. Single ended TSV carries only one signal (G-S-S-G) through it, but a two-bit TSV will have G-SS-G or GS-SG type of signaling scheme. Here S stands for signal and G stands for ground. This resulted in improved integration density,

lesser crosstalk and noise propagation problems and increased number of I/O in 3D IC.

The RLGC circuit model of single-ended signal through glass via (TGV) assuming it to be cylindrical is given in Kim et al. (2015). In this, instead of using silicon substrate, a separate layer made up of glass is added which is called glass interposer. The via passing through this interposer is accordingly referred to as through glass via. Glass is used because it offers higher resistivity (10^{12}–10^{16} Ω-cm) than silicon (0.16–0.60 Ω-cm). There is no need to provide insulation between substrate and via etch, due to very less conductivity of glass, hence this structure is less complex than TSV. The proposed model is verified up to 40 GHz with full-3D simulation. Using the proposed model with parasitic components of through glass via, it is reported that the inductance of via is the most dominant parameter to the result of s-parameter. It confirmed that the insertion loss of TGV is minimized with minimum height and pitch between the GS pair and a maximum diameter of through glass via. It is observed that inductance of via greatly influences the S-parameters. Variation in geometrical parameters such as pitch between adjacent via, height and diameter of via was studied in detail and they concluded that insertion loss will be minimum when height and pitch of adjacent vias, i.e., G-S pair is kept to its minimum and diameter of via is kept high.

Kumar et al. (2022) proposed a cylindrical differential multibit TGV(DM-TGV) design with CNT filler. An improvement of around 44% is reported in DM-TGCVs under differential mode with respect to differential TSVs for varying via radius and nearly 22% for differential multibit CNT TSVs. An improvement in crosstalk up to 89% is reported when compared to the differential TSVs at 20 GHz.

Various other novel structures such as tapered multibit differential multibit TGV, three-bit TSV, coaxial TSV with variation in filler materials and signal schemes are explored in Kumar and Dhiman (2022), Nilsson et al. (2009, 2021), Liang et al. (2013), Lu et al. (2018).

Electrical Modeling of Vias

At high frequencies, via is modeled as a transmission line cable due to increase in rise time of the signal through it. A transmission line is modeled as a distributed passive network of *R*, *L*, *C,* and *G* parameters, where *R*, *L*, *C,* and *G* are resistance, inductance, capacitance, and conductance respectively. Hence, via is an example of distributed element model. Figure 6.4 shows the front view of single-ended TSVs inside a substrate. In Fig. 6.4, the thickness of the oxide is represented by t_{ox}, while d_{TSV} and h_{TSV} represent the diameter and height of the TSV, respectively. Further, the top view of coupled signal-ground TSVs is shown in Fig. 6.5. In this figure, C_{ox} is the oxide capacitance, whereas C_{Si} and G_{si} represent the capacitance and conductance offered by the silicon substrate, respectively.

An electrical equivalent model of single-ended TSV is presented in Fig. 6.6. It is modeled as a series of resistances and inductances. The copper TSV is covered by

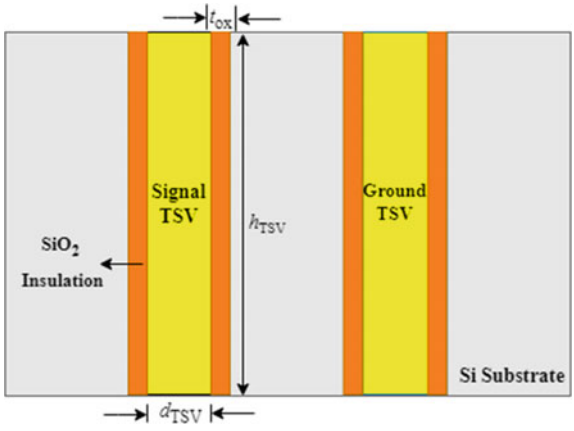


Fig. 6.4 Front view of TSVs inside the substrate

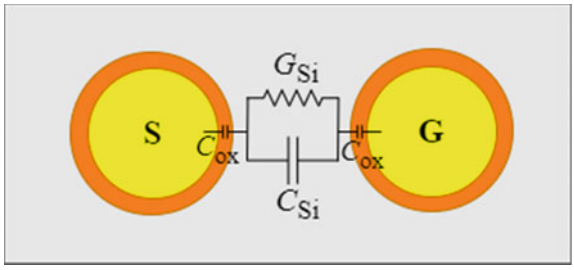


Fig. 6.5 Top view of coupled signal-ground TSVs

an oxide layer which acts as an insulation between the conducting channel and Si substrate. Due to the presence of oxide layer, a MOS capacitance is formed (metal—Cu of TSV, oxide—dielectric layer, semiconductor—silicon substrate) modeled as C_{ox}. This comes into play when dielectric is present. The finite resistivity of substrate material is modeled by C_{Si} and G_{si} in parallel. In Fig. 6.6, Δx is the infinitesimal segment of h_{TSV}.

Numerical Techniques

The operational characteristics of a via are explored using equivalent circuit model. Frequency and time domain analyses of the model can be performed to deduce the behavior of the via in terms of its RLGC parameters. Some of the numerical techniques useful for these analyses include finite-difference time-domain (FDTD), partial element equivalent circuit (PEEC), and ABCD (Ruehli 1996). FDTD is a

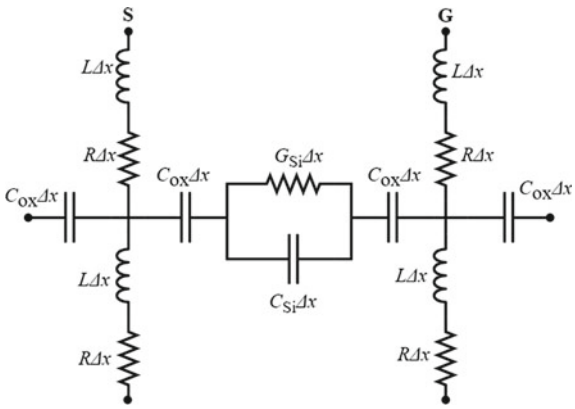


Fig. 6.6 Electrical equivalent model of single-ended TSV

popular computational electrodynamics modeling technique. In this technique, time-dependent partial differentials are discretized using central-difference approximations to space and time. To know the effect on signal due to variation in frequency, researchers have been popularly using the techniques, namely, PEEC and ABCD.

The PEEC method is a type of electromagnetic simulation that relies on the integral formulation of Maxwell's equation. As vias are modeled as transmission line cables, ABCD technique is a popular choice to solve two-port network under high-frequency operations. There are several TSVs etched in the substrate which will eventually form an equivalent circuit which is cascade of several of two TSVs. ABCD technique can easily solve such series of cascade of two-port network.

The design and modeling of high frequency electromagnetic structure can be carried out using commercially available electronic design automation (EDA) tool, namely, high frequency structure simulator (HFSS) provided by Ansys (Ansys EDA tool 2022).

Ansys is a software company that provides engineering simulation platforms and delivers 3D design software. HFSS is a solution for electromagnetic structures using the finite element method. Since the impact of different flaws on the TSV performance parameters can be thoroughly described, it is perfect for TSV fault analysis. It may be utilized to determine how electric and magnetic fields are distributed and how intense they are not just within the entire TSV, but also in the surrounding medium, which includes the substrate and the dielectric.

Summary

Traditionally, the packaging of integrated circuits was meant to provide mechanical and environmental support from any kind of abrasions. But today the world of packaging has got a whole new meaning. The way the ICs are packaged defines the performance of SoC. Packaging techniques such as 2D, 2.5D, and 3D integration,

wafer-level packaging, chip-scale integration are some industry-level solutions to get increased performance from SoC.

3D integration promises better performance than 2D or 2.5D integration. In the past few years, 3D integrated circuits have emerged as a solution to miniaturization, offering benefits of improved performance, higher memory bandwidth, smaller form factor, and heterogeneous integration. It has provided a holistic approach to multi-die integration, for an optimal system to SoC solution. However, choice of adapting the packaging techniques solely depends on the requirement of the designer and the cost considerations. 3D integration is a more complex technique than 2D or 2.5D integration. This is because it involves the fabrication of vertical vias or TSVs in addition to interposer technology like in 2.5D.

The different methods to integrate a via in the IC be it TSV first, middle, or last largely affect the performance of signal through it. The different approaches such as monolithic or stacking to integrate different die onto a single substrate are important choices that need to be made prior that could best enable the multilayer connections to considerably boost system performance. Therefore, these integration methods are important choices that must be pre-decided from all angles as these could affect the final product performance and cost dramatically. Thus, being a very essential requirement of present-day integrated chips, different aspects of 3D integration and via modeling are presented in the current chapter which will significantly benefit the VLSI designers and academia.

Acknowledgements The authors sincerely acknowledge with gratitude the technical and financial support received from the Science and Engineering Research Board, Department of Science and Technology (SERB-DST), GoI, New Delhi, through the Core Research Grant [Ref. No.: CRG/2021/000780].

References

Ansys EDA tool (2022) https://www.ansys.com

Awano Y, Sato S, Nihei M, Sakai T, Ohno Y, Mizutani T (2010) Carbon nanotubes for VLSI: interconnect and transistor applications. In: Proceedings of the IEEE, vol 98, no 12, pp 2015–2031

Cho J, Kim J (2013) Signal integrity design of TSV and interposer in 3D-IC. In: IEEE 4th Latin American symposium on circuits and systems (LASCAS). Cusco, Peru, pp 1–4. https://doi.org/10.1109/LASCAS.2013.6519092

Garrou P, Koyanagi M, Ramm P (2014) Handbook of 3D integration volume 3: 3D process technology. Wiley-VCH, Weinheim, Germany

Guan W, Lu H, Zhang Y, Zhang Y (2021) A novel differential multi-bit carbon nanotube through silicon vias. In: 4th international conference on advanced electronic materials, computers and software engineering (AEMCSE). Changsha, China, pp 1–5. https://doi.org/10.1109/AEMCSE51986.2021.00042

Guarini KW, Topol KW, Ieong M, Yu R, Shi L, Newport MR, Frank DJ, Singh DV, Cohen GM, Nitta SV, Boyd DC, O'Neil PA, Tempest SL, Pogge HB, Purushothaman S, Haensch WE (2002) Electrical integrity of state-of-the-art 0.13μm SOI CMOS devices and circuits transferred for three-dimensional (3D) integrated circuit (IC) fabrication. In: Int electron devices meeting (IEDM). San Francisco, CA, USA, pp 943–945. https://doi.org/10.1109/IEDM.2002.1175992

Hosali S, Smith G, Smith L, Vitkavage S, Arkalgud S (2008) Through-silicon via fabrication, backgrind and handle wafer technologies. In: Tan C, Gutmann R, Reif L (eds) Wafer level 3-D ICs process technology. Integrated circuits and systems. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-76534-1_5

Hu Q-H, Zhao W-S, Fu K, Wang G (2020) Modeling and characterization of differential multibit carbon-nanotube through-silicon vias. In: IEEE transactions on components, packaging and manufacturing technology, vol 10, no 3, pp 534–537. https://doi.org/10.1109/TCPMT.2020.2972340

Im S, Banerjee K (2000) Full chip thermal analysis of planar (2-D) and vertically integrated (3-D) high performance ICs. In: International electron devices meeting. Technical Digest. IEDM (Cat. No.00CH37138). San Francisco, CA, USA, pp 727–730. https://doi.org/10.1109/IEDM.2000.904421

Kaushik BK, Kumar VR, Majumder MK, Alam A (2016) Through silicon vias materials, models, design, and performance, 1st edn. CRC Press. https://doi.org/10.1201/9781315368825

Kim J, Hwang I, Kim Y, Cho J, Sundaram V, Tummala R, Kim J (2015) Precise RLGC modeling and analysis of through glass via (TGV) for 2.5D/3D IC. In: IEEE 65th electronic components and technology conference (ECTC). San Diego, CA, USA, pp 254–259. https://doi.org/10.1109/ECTC.2015.7159601

Kim K et al. (2010) Analysis of power distribution network in TSV-based 3D-IC. In: 19th Topical meeting on electrical performance of electronic packaging and systems. Austin, TX, USA, pp 177–180. https://doi.org/10.1109/EPEPS.2010.5642575

Kumar A, Chandel R, Dhiman R (2022) Proposal and analysis of carbon nanotube based differential multibit through glass vias. Microelectronics J 126: 105500. https://doi.org/10.1016/j.mejo.2022.105500

Kumar A, Dhiman R (2022) Tapered differential multibit through glass vias for three dimensional integrated circuits. In: IEEE electrical design of advanced packaging and systems (EDAPS). Urbana, IL, USA, pp 1–3. https://doi.org/10.1109/EDAPS56906.2022.9995142

Laermer F, Schilp A (2003) Method of anisotropic etching of silicon. US Patent 6531068B2

Lee H-HS, Chakrabarty K (2009) Test challenges for 3D integrated circuits. In: IEEE design and test of computers, vol 26, no 5, pp 26–35. https://doi.org/10.1109/MDT.2009.125

Liang F, Wang G, Zhao D, Wang B-Z (2013) Wideband impedance model for coaxial through-silicon vias in 3-D integration. IEEE Trans Electron Dev 60(8): 2498–2504. https://doi.org/10.1109/TED.2013.2268869

Liu CW, Tseng WT, Dai BT, Lee CY, Yeh CF (1996) A perspective on the wear mechanism during CMP of tungsten thin films. In: Proceedings of CMP VLSI/ULDI multilevel interconnection conference, pp 138–144

Lu Q, Zhu Z, Liu Y, Liu X, Yin X (2018) Wideband electromagnetic modeling of coaxial-annular through-silicon vias. IEEE Trans Electromagnetic Compatibility 60(6): 1915–1922. https://doi.org/10.1109/TEMC.2017.2771293

Medhat D (2022) 2.5/3D IC reliability verification has come A long way, semiconductor engineering. https://semiengineering.com/2-5-3d-ic-reliability-verification-has-come-a-long-way

Nilsson P, Ljunggren A, Thorslund R, Hagstrom M, Lindskog V (2009) Novel through-silicon via technique for 2d/3d SiP and interposer in low-resistance applications. In: 59th electronic components and technology conference. San Diego, CA, USA, pp 1796–1801. https://doi.org/10.1109/ECTC.2009.5074261

Qian L, Xia Y, Shi G (2016) Electrical modeling and analysis of a mixed carbon nanotube-based differential through silicon via in 3-D integration. IEEE Trans Nanotechnol 15(2): 155–163. https://doi.org/10.1109/TNANO.2015.2509019

Qian L, Xia Y, Shi G, Wang J, Ye Y, Du S (2017) electrical–thermal characterization of through packaging vias in glass interposer. IEEE Trans Nanotechnol 16(6): 901–908. https://doi.org/10.1109/TNANO.2017.2722686

Rahman T, Miao J, Youcef H (2013) Structure optimization of through silicon via (TSV) interconnect as transmission channel for 3D integration. In: 5th IEEE international symposium on microwave, antenna, propagation and EMC technologies for wireless communications. pp 668–671. https://doi.org/10.1109/MAPE.2013.6689929

Rao M (2020) Electrical modeling of copper and mixed carbon bundles as a composite for 3D interconnect applications. In: IEEE 15th international conference on nano/micro engineered and molecular system (NEMS). San Diego, CA, USA, pp 495–499. doi: https://doi.org/10.1109/NEMS50311.2020.9265633

Riedel S et al (2000) Investigation of plasma treatment in a multistep TiN MOCVD process. Microelectron Eng 50:533–540. https://doi.org/10.1016/S0167-9317(99)00324-X

Ruehli A (1996) Partial element equivalent circuit (PEEC) method and its application in the frequency and time domain. In: Proceedings of symposium on electromagnetic compatibility. Santa Clara, CA, USA, pp 128–133. https://doi.org/10.1109/ISEMC.1996.561214

Satomi Y, Hachiya K, Kanamoto T, Kurokawa A (2018) Optimization of full-chip power distribution networks in 3D ICs. In: IEEE 3rd international conference on integrated circuits and microsystems (ICICM). Shanghai, China, pp 134–138. https://doi.org/10.1109/ICAM.2018.8596650

Tavakkoli F, Ebrahimi S, Wang S, Vafai K (2016) Analysis of critical thermal issues in 3D integrated circuits. Int J Heat Mass Transf 97:337–352. https://doi.org/10.1016/j.ijheatmasstransfer.2016.02.010

Topol AW et al (2006) Three-dimensional integrated circuits. IBM J Res Dev 50(4.5): 491–506. https://doi.org/10.1147/rd.504.0491

Xu Z, Lu J-Q (2013) Through silicon-via fabrication technologies, passives extraction, and electrical modeling for 3-D integration/packaging. IEEE Trans Semicond Manuf 26(1): 23–34. https://doi.org/10.1109/TSM.2012.2236369

Chapter 7
Neural Networks for Fast Design Space Exploration of On-Chip Interconnect Networks

Somesh Kumar and Rahul Kumar

Abstract In this chapter, we will look at how artificial neural network (ANN) models may be used to design on-chip interconnects for integrated circuits and systems. Conduct of on-chip copper (Cu) interconnect networks is restricted by dispersive processes such as grain boundary scattering, surface roughness scattering, top/bottom surface, and sidewall scattering when minimum interconnect width scales below the 22 nm mark. The resistance of the interconnects, measured over a unit of length, is greatly inflated over the bulk value due to these scattering phenomena. Enhanced signal attenuation, delay, and power loss are all consequences of such extremely resistive interconnects. Augmenting to causing discontinuities in the interconnect line, the migration of copper ions into the dielectric layer increases dielectric conductivity and leakage losses. To prevent copper ions from diffusing away from the copper conductor, a barrier layer is often positioned around it. Recently, many studies have investigated the potential of employing graphene nanoribbons as a replacement barrier material for Cu interlinks. At ambient temperature, the mean free route for electrons in graphene nanoribbons is much greater than in copper. Therefore, in addition to the copper conductor, the barrier layer provides extra low-resistance routes for conducting electrons. It is becoming more important to do SPICE-based simulations of hybrid Cu-Graphene on-chip interconnect networks due to their rising popularity. Simulations of interconnects using SPICE are notoriously time-consuming and computationally intensive. Using surrogate models is one way to deal with the computational overhead of exploring design space. Due to their capacity to mimic the extremely nonlinear input–output correlations of electronic packaging systems, surrogate templates based on machine learning (ML) regression are now quite sought-after. We will also collate the results of our ML-based models to those of our comprehensive EM and SPICE simulations.

S. Kumar (✉)
ABV-Indian Institute of Information Technology and Management Gwalior, Gwalior, M.P, India
e-mail: somesh@iiitm.ac.in

R. Kumar
Georgia Institute of Technology, Atlanta, USA
e-mail: fmr5281@psu.edu

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_7

Keywords Copper · EM solver · Graphene · Interconnect · Neural networks · On-chip

Introduction

Performance characteristics including attenuation, latency, and power loss of on-chip copper (Cu) interconnect networks increase when any dimension (width, length, thickness) of interconnects scales below the 22 nm threshold. This is because of scattering at the grain boundaries, top and bottom surfaces, and sidewalls (Meindl 2001; Steinhögl et al. 2005; Todri-Sanial et al. 2017). Because copper ions diffuse into the dielectric layer, conductivity and leakage losses increase. The barrier layer surrounding the copper layer is a common strategy for preventing diffusion. Materials like tantalum *(Ta)* and tantalum nitride *(TaN)* are often employed for the barrier layers (Yang et al. 2004). These capping layers are not conducive to electron conduction because they have greater resistance than copper. The maximum current carrying capacity of the connecting wires also drops because of the decrease in the effective cross-section area. Graphene nanoribbons have been the subject of much research as a possible replacement barrier material for Cu interconnects in recent years (Zhao et al. 2015; Achar and Nakhla 2001). We observed that the ampacity of copper conductors increases when covered with ultra-thin graphene sheets. In addition, graphene nanoribbons are remarkably resistant to thermal oxidation even when subjected to extreme heat. At ambient temperature, graphene nanoribbons have far longer electron mean free pathways than copper. Because of this, the barrier layer complements the copper conductor by providing a second, low-impedance channel for the flow of electrons (Zhao et al. 2015). These low-resistance routes prevent copper ions from diffusing into the dielectric and greatly reduce the equivalent resistance per unit length (p.u.l.) of the whole conductor and junction. Due to these factors, researchers have shown a strong desire to create hybrid Cu-graphene-on-chip interconnects for use at the sub-22 nm technology node.

Simulating hybrid Cu-graphene on-chip interconnect networks using SPICE is becoming more important as their use rises. Typically, there are two stages involved in running a network simulation using SPICE. Parameters p.u.l. of networks are initially calculated using quasi-TEM approximations of linked structures, often with electromagnetic (EM) solvers (Achar and Nakhla 2001). Transient analysis of networks is then performed by plugging the collected characteristics into preexisting multiconductor transmission line (MTL-SPICE) models (Achar and Nakhla 2001). It is common knowledge that the first of the aforementioned stages has a higher temporal complexity than the second. When doing a design space exploration, the time required for this first stage increases dramatically. Simply expressed, design space exploration is the process of running SPICE simulations using the aforementioned two phases over thousands of iterations, each time varying the input geometric, physical, and material characteristics of the network. Using surrogate models is one approach to coping with the computing work required to explore the design space.

A surrogate model is a closed-form mathematical formula that mimics or emulates the input–output relationships of the particular system under test. Machine learning (ML) regression-based surrogate models are widely used nowadays (Torun et al. 2020) because they can simulate the extremely nonlinear input–output connection of electronic packaging architectures. Artifical neural networks (ANNs) are created to simulate the functional dependency of parameters for the values of geometric, physical, and material-technical characteristics of networks. It is hypothesized that ANNs, if taught to faithfully imitate this input–output function dependence, would be far more effective than full-wave EM solvers.

Despite ANNs' numerical efficiency, training data must be retrieved from rigorous full-wave EM solvers, a process that entails several iterations and a significant amount of computer overhead. In order to cut down on the high price tag associated with this lengthy ANN training period, several methods have been documented in the literature, such as source difference method, prior knowledge integration, and space-mapping techniques.

Each of these methods relies on the premise that the pulse has strong functional dependencies. Empirical models may provide close approximations of the linked network parameters of the input geometric, physical, and material characteristics. So, a very little quantity of training data derived from more stringent full-wave EM solvers may be used to efficiently train ANNs, whereas a relatively big volumes of training data collected from lighter numerical models can be used instead. Training time for conventional ANNs is prohibitively time-expensive, but we can drastically cut that cost here without compromising accuracy. Knowledge-based neural networks (KBNN) (Zhang et al. 2003) are ANNs that may be trained on such data combinations. In comparison to classic ANNs, KBNNs have been found to have superior convergence qualities; nevertheless, they do have some significant drawbacks. The empirical model required to train the KBNN must be known in advance. This is a substantial difficulty for interconnect networks based on hybrid Cu-graphene materials.

ANNs for Interconnect Modeling

In order to do transient analysis on high-speed connections, several different SPICE models of MTLs have been created over the years (Achar and Nakhla 2001; Paul 2007). The electrical behavior of networks is modeled utilizing the Telegraph and partial differential equations. For the performance benchmarking of any network, it is of utmost importance to extract RLC p.u.l. parameters. Because of this, we must first extricate network p.u.l. RLC attributes based on our existing knowledge of the network's shape and architecture in order to run a SPICE simulation of the linked network. The use of a full-wave rigorous EM solver is common place for this goal. Extraction of RLC parameters is a time-consuming procedure due to the high computational intricacy of such solvers. Parametric sweeps and investigations of design space increase this problem. The electronic packaging industry has taken an interest in ML regression models (Park et al. 2020) because of their capacity to

simulate the complicated input–output nonlinear behavior of devices, circuits, and systems. Particularly, ANNs have risen in popularity as a widely used ML regression model. The fragmental interdependence of p.u.l. RLC attributes of interconnections in terms of values of geometric, material, and physical design characteristics of network are the subject of most current efforts on building ANNs. The fundamental benefit of these ANN models is that, unlike conventional full-wave EM solvers, they are intrinsically closed and can be queried for attributed coverage and design space searches in snatch of time required by the CPU.

a. Artificial Neural Networks for Interlink Network P.U.L. Modeling

Think of a linked network in general that has N geometric, physical, and material design characteristics. Each design parameter may have a value anywhere within a certain range of values called a "support interval" $[p_{i,\min}, p_{i,\max}]$ to optimize the operation of a network. That being said, we may use a mathematical representation of the input parameters, which looks like this:

$$p_i = \frac{\left(p_{i,\min} + p_{i,\max}\right)}{2} + \frac{\left(p_{i,\max} - p_{i,\max}\right)}{2}\lambda_i; \quad i = 1, ..., N \tag{7.1}$$

here the variables λ_i represent uncertainties in design parameters, and typically value lies between −1 and 1. RLC attributes of the interlinked network are denoted through matrices R, L, and C with constant values $\lambda = [\lambda_1, \ldots, \lambda_N]$, at very high time-costs. Similar functional relationships may be modeled using a multilayer perceptron (MLP) implementation of ANN (2000), as seen in Fig. 7.1a. Neurons in this MLP design are stacked in levels, with information from every level being passed on to the next. The measures of λ at a given design spot are fed into the network at the input layer, a group of neurons located on the network's left side. The output layer is the last stage of the MLP, and it gives the values of the R, L, or C matrices that go with them.

Next, to assimilate the process through which data passes through the MLP, Fig. 7.1 shows the structure and connectivity of individual neurons within the MLP. As per Fig. 7.1b, the neuron accepts as inputs the weighted sum S as

$$\mathbf{W} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}; \ \mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}; \ \mathbf{S} = b + \mathbf{W}^t\mathbf{X} = b + \sum_{j=1}^{n} w_j x_j \tag{7.2}$$

The weight vector, input data vector, bias term, and the total number of input terms are all denoted by the letters W, X, b, and n, respectively. To get the scalar output, we take the scalar sum S and run it through a nonlinear activation function, which is given as

$$y = \sigma(\mathbf{S}) = \sigma\left(b + \sum_{j=1}^{n} w_j x_j\right) \tag{7.3}$$

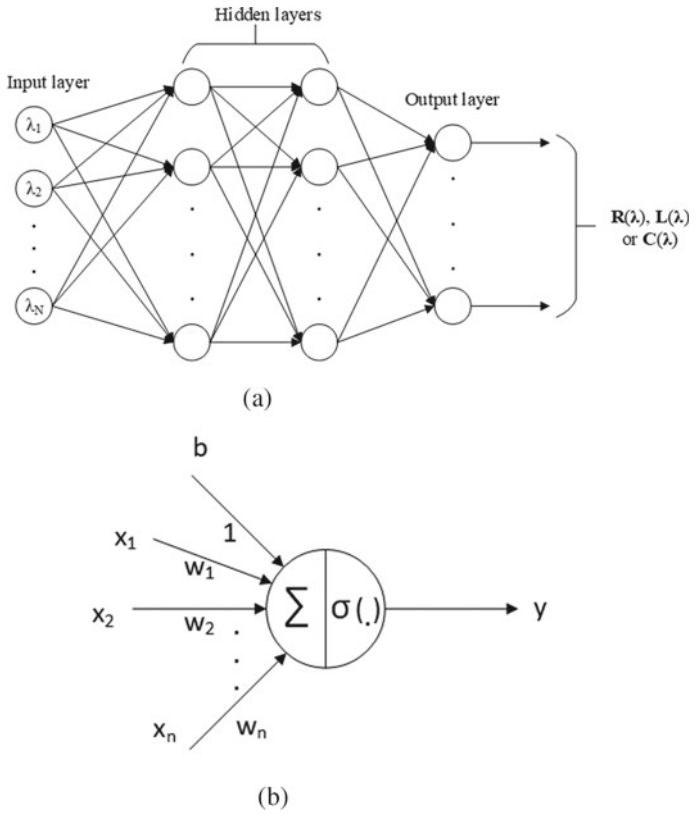


Fig. 7.1 Predicting the p.u.l. characteristics of interlink networks using artificial neural networks (ANNs). **a** The interconnections between neurons in a multilayer perceptron ANN. **b** Transmission of information inside a single neuron

Thus, when the input data X passed through a neuron, the nonlinear output y(X) is produced. The functional dependencies {R(λ), L(λ), C(λ)} are thus recorded by the ANN as shown in Fig. 7.1.

Importantly, each neuron's weight and bias values are adjusted so that errors within the predicted ANN output and parameters computed using the full-wave EM solver are minimized when λ is covered in the attributed space of the multidimensional hypercube. "Training" an ANN refers to the process of optimizing its performance. Various gradient or non-gradient-based schemes are currently known for training ANNs, of which the class of gradient-based back-propagation techniques is very popular. Usually, for the interconnect problem, we need to train three different ANNs to reproduce the functional dependencies of R(λ), L(λ), and C(λ) separately. If the ANN is trained such that the error between the predicted ANN output and the results of the full-wave EM solver is small enough in the multidimensional parametric space, then ANN behaves as a closed-form surrogate for the full-wave EM solver.

b. Knowledge-Based Neural Networks (KBNNs)

The key difficulty in training the above ANN is running the full-wave EM solver many times (let's say K) to extract the training data and additional problem-specific knowledge needed to speed up the training of the ANN (Bakr et al. 2000) as shown in Fig. 7.2.

The key difficulty in training the above ANN is the need to run the full-wave EM solver many times (say, K times) to extricate the problem-specific auxiliary knowledge needed to speed up the training of the ANN (Bakr et al. 2000). Knowledge-based neural networks are networks that can be taught to utilize more knowledge to converge more quickly. In order to train KBNNs more quickly than conventional ANNs, a number of methods have been devised, including source difference, previous knowledge, and space-mapping. The source differencing approach, for instance, considers the inter-dependencies between p.u.l. functional components. Empirical models like as (Watson et al. 1996) are used to fit R, L, and C matrices on the input vector λ.

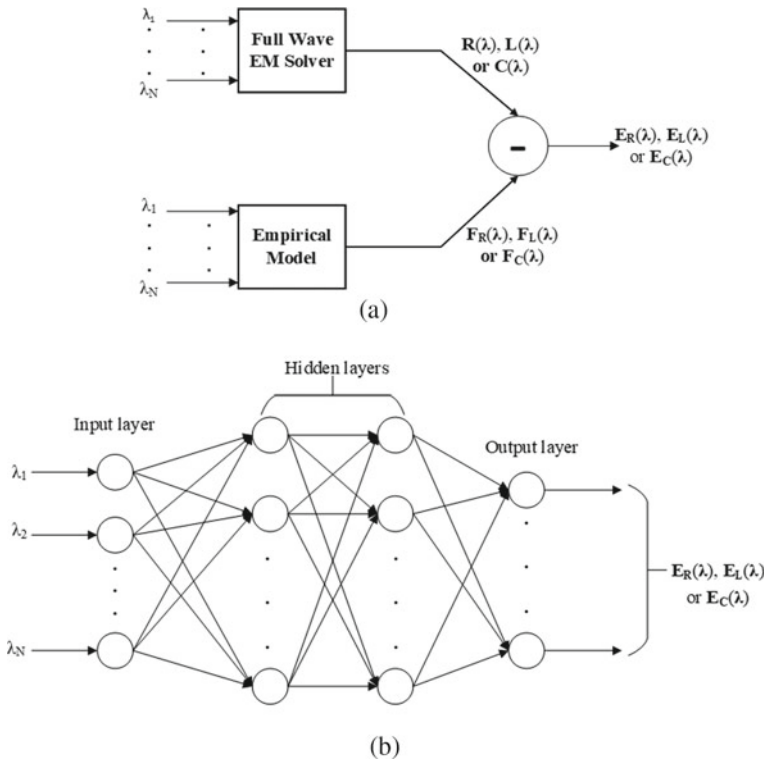


Fig. 7.2 Informed neural networks. **a** The functional dependence of the error matrices in (7.5) on the input variables λ, as shown by the source difference formalism of KBNNs. **b** Artificial neural network (ANN) design based on multilayer perceptrons to model error matrices (7.5)

$$\tilde{R} = F_R(\lambda),\ \tilde{L} = F_L(\lambda),\ \tilde{C} = F_C(\lambda) \tag{7.4}$$

where $\left\{\tilde{R}, \tilde{L}, \tilde{C}\right\}$ indicate the estimated values for the R, L, and C matrices that may be inferred through the empirical models $F_R(\lambda)$, $F_L(\lambda)$, and $F_C(\lambda)$ respectively. The empirical model in (7.4) represents auxiliary knowledge about p.u.l. matrices unavailable from network-structured quasi-TEM solutions (that is, training data obtained from EM solvers).

$$\begin{aligned} E_{\mathrm{R}}(\boldsymbol{\lambda}) &= R(\boldsymbol{\lambda}) - \mathbf{F}_{\mathrm{R}}(\boldsymbol{\lambda}); \\ E_{\mathrm{L}}(\boldsymbol{\lambda}) &= L(\boldsymbol{\lambda}) - \mathbf{F}_{\mathrm{L}}(\boldsymbol{\lambda}); \\ E_{\mathrm{C}}(\boldsymbol{\lambda}) &= C(\boldsymbol{\lambda}) - \mathbf{F}_{\mathrm{C}}(\boldsymbol{\lambda}) \end{aligned} \tag{7.5}$$

The norms of the error matrices in (7.5) are the equivalent R, L, and C matrices if the findings of the empirical model in (7.4) agree well with those of the full-wave EM solver. Therefore, the ANN needs far less training samples to correctly imitate an error matrix of (7.5) than p.u.l. R, L, and C matrices. Because of this, the source differencing method suggests ANN instead of p.u.l. R, L, and C matrices to imitate the error matrix in (7.5).

$$\begin{aligned} \mathbf{R}(\boldsymbol{\lambda}) &= \mathbf{E}_R(\boldsymbol{\lambda}) + \mathbf{F}_R(\boldsymbol{\lambda}); \\ \mathbf{L}(\boldsymbol{\lambda}) &= \mathbf{E}_L(\boldsymbol{\lambda}) + \mathbf{F}_L(\boldsymbol{\lambda}); \\ \mathbf{C}(\boldsymbol{\lambda}) &= \mathbf{E}_C(\boldsymbol{\lambda}) + \mathbf{F}_C(\boldsymbol{\lambda}) \end{aligned} \tag{7.6}$$

where the second component in RHS (7.6) is the well-known empirical model of (7.4), and the first term is the ANN depiction of the error matrix in (7.5). Figure 7.3 depicts KBNN, which is the combination of this ANN with the empirical model of (7.6).

Despite KBNNs' capacity to train quicker than traditional ANNs, their key difficulty is the need to know in advance a sufficient empirical model of (7.4). This is especially difficult for hybrid Cu-graphene interconnects since the well-established empirical model of (7.4), which applies to pure copper interconnects, is grossly wrong

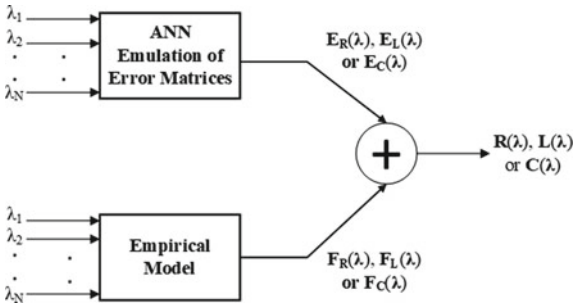


Fig. 7.3 True measures of the p. u. l. **R, L,** and **C** matrices using the source difference formulation in Fig. 7.2

when dealing with heterogeneous topologies. At now, there is no established empirical template of hybrid Cu-graphene linked networks for training KBNNs. Therefore, there is a deficiency in the existing research about the viability of KBNN for swift design space analysis of hybrid Cu-graphene interlink networks.

KBNNs for Hybrid *Cu*-Graphene Interlink Networks

a. Recognizing Potential Empirical Models for KBNNs

The paper (Cheng et al. 2018) mentions a closed-form model which was developed to estimate the dependence of p.u.l. R, L, and C matrices of the hybrid Cu-graphene interlink network for the input variable λ. The resistance of Cu-graphene interconnect networks is given as (Cheng et al. 2018):

$$\frac{1}{R_{hybrid}} = \frac{1}{R_{Cu}} + \sum_{k=t,b,l,r} \frac{1}{R_g^k} \tag{7.7}$$

where the quantities $\left\{R_g^t, R_g^b, R_g^l, R_g^r\right\}$ in (7.7) mean the upper, lower, left, and right graphene resistive layer p.u.l equivalent resistive effects. Following this, the effective parameters of capacitance and inductance with respect to every conductor in a hybrid connection network are determined as (Zhao et al. 2015):

$$\begin{aligned} C_{hybrid} &= \left[\sum_{k=t,b,l,r} \left(C_{rec}^k\right)^{-1} + \frac{1}{C_e}\right]^{-1} \\ L_{hybrid} &= \left(\sum_{k=t,b,l,r} \frac{1}{L_{rec}^k}\right)^{-1} + L_e \end{aligned} \tag{7.8}$$

here C_e and L_e represent the *p.u.l.* electrostatic capacitance and magnetic inductance components and $L_e C_e = \mu_0 \varepsilon_0 \varepsilon_r$. The quantities $\left\{C_{rec}^t, C_{rec}^b, C_{rec}^r, C_{rec}^l\right\}$ and $\left\{L_{rec}^t, L_{rec}^b, L_{rec}^r, L_{rec}^l\right\}$ in (7.8) for each conductor are the recursive approach used to compute the p.u.l. equivalent capacitance and inductance from the top, bottom, left, and right sides of the graphene barrier layers as described in Zhao et al. (2015).

Then, for every potential structure, a set of N_{test} testing nodes are selected where $N_{test} = \{100, 200, 500, 1000\}$. An error matrix of (7.5) is computed for each test node and a fixed measure of N_{test}. The error matrix in (7.5) is then converted to a scalar quantity normalized with respect to the nominal p.u.l. parameter as

$$
\begin{aligned}
e_R &= \frac{1}{\sqrt{N_{test}}\|R(0)\|}\sqrt{\sum_{k=1}^{N_{test}}\|E_R(\lambda_K)\|^2};\\
e_L &= \frac{1}{\sqrt{N_{test}}\|L(0)\|}\sqrt{\sum_{k=1}^{N_{test}}\|E_L(\lambda_K)\|^2};\\
e_C &= \frac{1}{\sqrt{N_{test}}\|C(0)\|}\sqrt{\sum_{k=1}^{N_{test}}\|E_C(\lambda_K)\|^2};
\end{aligned}
\tag{7.9}
$$

where the Frobenius norm is utilized as the matrix norm in (7.9). Applying (7.9)'s calculations to a variety of interconnect test configurations, we may draw the following three conclusions as

(i) Error quantities of (7.9) normally saturate for $N_{test} = 200$. This is justified through the fact that for $N_{test} \geq 200$, the variation in magnitude of error in (7.9) is negligible.
(ii) For all investigated interconnect test structures, *p.u.l.* inductance (e_L) error is lower than 1%. This suggests that the *p.u.l.* inductance template described in Zhao et al. (2015) is correct in its closed-form approximation, (i.e., $F_L(\lambda)$) is adequately precise in serving as the emulator of a full-wave EM solver (i.e., $\mathbf{F}_L(\lambda) \approx \mathbf{L}(\lambda)$ or as per the converse $\|E_L(\lambda)\| \approx 0$).
(iii) The error regarding *p.u.l.* resistance (e_R) is constant within 4–8% while regarding *p.u.l.* capacitance (e_C) precisely within 10–25% for all the distinct interlink test structures observed.

Since the e_R and e_C errors are not insignificant, it is anticipated that the *p.u.l.* R, L, and C matrices of hybrid Cu-graphene interlink networks cannot be predicted directly using the closed-form models of Cheng et al. (2018). However, the closed-form templates in Cheng et al. (2018) aren't ruled out as viable empirical templates in (7.4) since the quantities of errors e_R and e_C aren't too huge in (7.5). This crucial feature of the closed-form models of Cheng et al. (2018) makes them attractive for training KBNNs—something that is not known about the work of Cheng et al. (2018).

Numerical Example

Here, we demonstrate how the proposed KBNN can be used to conduct design space survey for hybrid Cu-graphene interconnects with improved accuracy and numerical efficiency compared to traditional ANNs and the closed-form models in Cheng et al. (2018). The commercial full-wave EM solver ANSYS Q3D Extractor is used to collect the training data for this purpose. All of the three-layer ANNs and KBNNs implemented in some illustrated instances were constructed, trained, and evaluated with the help of MATLAB R2019a machine-learning toolbox. As

illustrated in Fig. 7.4, we take into account an MTL network with three lines of Cu-graphene hybrid conductors. Lines 1 and 3 are activated by a saturated ramp voltage source with a rise time $T_r = 0.1$ ps and an amplitude of 1 V, respectively. Table 7.1 displays the design parameters used in this instance. Analyzing the data shown in Table 7.1, one can infer that $N = 5$ dimensions are needed for this example's design space.

The first section of this example employs the full-wave ANSYS Q3D Extractor tool, a traditional ANN, and suggested KBNN to extract p.u.l. RLC attributes for the network are shown in Fig. 7.4. The closed model of Cheng et al. (2018) has errors of $e_L = 0.66\%$, $e_R = 3.63\%$, and $e_C = 11.85\%$. Just because such errors are relatively small, we expect the closed model of Cheng et al. (2018) to serve as a valid empirical model. Importantly, this example assumes that the closed model

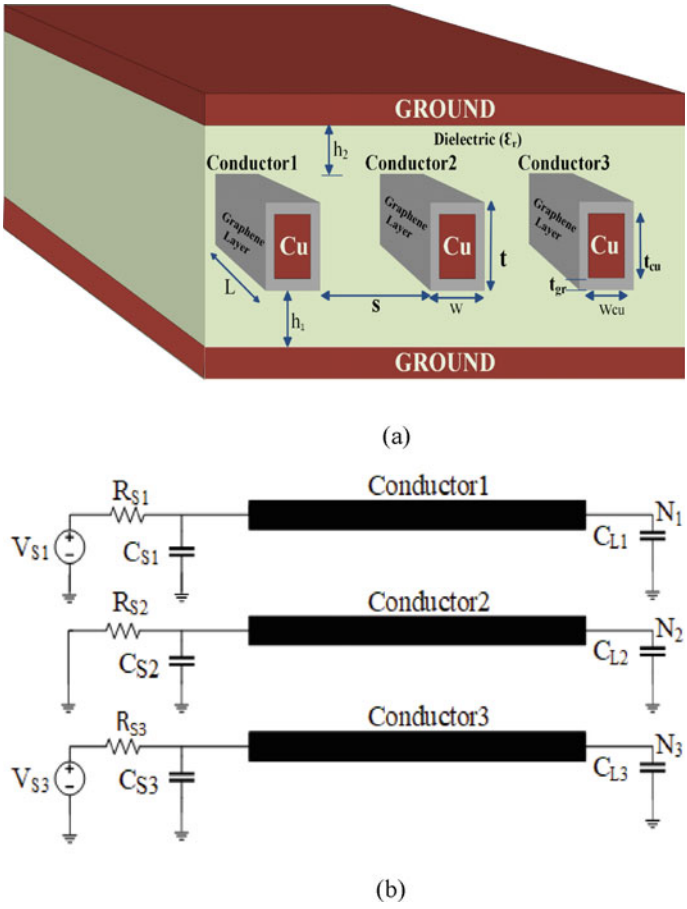


Fig. 7.4 **a** Schematic of Hybrid copper and graphene interconnects **b** Interconnect with driver and load circuits

Table 7.1 Design parameters for the network

	Design attributes	Nominal measures (nm)	% Variable perimeter
1	Width (w)	13.7	±10%
2	Thickness (t)	27.31	
3	Spacing (S)	13	
4	Height of interconnect from GDN layer (h_1)	26	
5	Barrier layer thickness (t_{gr})	0.83	

of Cheng et al. (2018) for matrix L exactly emulates the ANSYS Q3D extraction tool. However, e_R and e_C are not trivial mistakes, necessitating the use of ANNs to simulate their error matrices $E_R(\lambda)$ and $E_C(\lambda)$.

By thoroughly exploration of all the above ANNs, it is conceived that when it comes to ANNs, a hyperbolic tangent function makes for the ideal activation function. Regular ANN training calls for 3 hidden neurons and 4000 epochs. However, an ANN mimicking error matrices $E_R(\lambda)$ and $E_C(\lambda)$ only needed 500 epochs to train and 3 hidden neurons. Each ANN's training phase involves an incremental increase in the number of training nodes: Nt = [15, 25, 35, 70, 140, 350, 700, 1000]. 500 test nodes are randomly chosen in the design space without crossing into the area occupied by the training nodes.

Figure 7.5 illustrates the relationship between the test error and the number of training nodes (N_t) for the suggested KBNN and the traditional ANN. The selected settings increase the error between the closed model of Cheng et al. (2018) and the results anticipated by ANSYS Q3D Extractor. This indicates that mimicking the matrices $E_R(\lambda)$ and $E_C(\lambda)$ for these parameters is the most challenging (i.e., it takes the most training nodes) and thus, the suggested KBNNs exhibit the slowest convergence. Now, Fig. 7.5 demonstrates that for the same number of training nodes (N_t), the suggested KBNNs are consistently capable of attaining a lower testing error norm than the traditional ANNs of Fig. 7.1a. This data conspicuously demonstrates that the suggested KBNNs have a quicker convergence rate than standard ANNs. In fact, Fig. 7.5 reveals that if a minute error perimeter of 2% is required while computing the p.u.l. attributes of the hybrid Cu-graphene interlink network of Fig. 7.4, the suggested KBNNs entail roughly P = 25 training nodes as opposed to the traditional ANNs in Fig. 7.1a, which entail roughly Q = 125 training nodes. It indicates that the proposed KBNNs will give a training speedup of Q/P = 5 times that of standard ANNs. It is a decrease of 80% with respect to count of training samples required to obtain equivalent inaccuracy. SPICE-based design space explorations are then done to evaluate the precision of the suggested KBNNs and collate it to that provided by traditional ANNs.

The short-term crosstalk behavior at node N2 in Fig. 7.4 is first predicted using the HSPICE W element template at two random design spots, having λ =[0.54 −0.91 0.49 −0.90 −0.39] or λ = [−0.97 −0.67 0.45 0.25 0.95]. The W element template's parameters are retrieved using the KBNN, standard ANN, and closed-form models

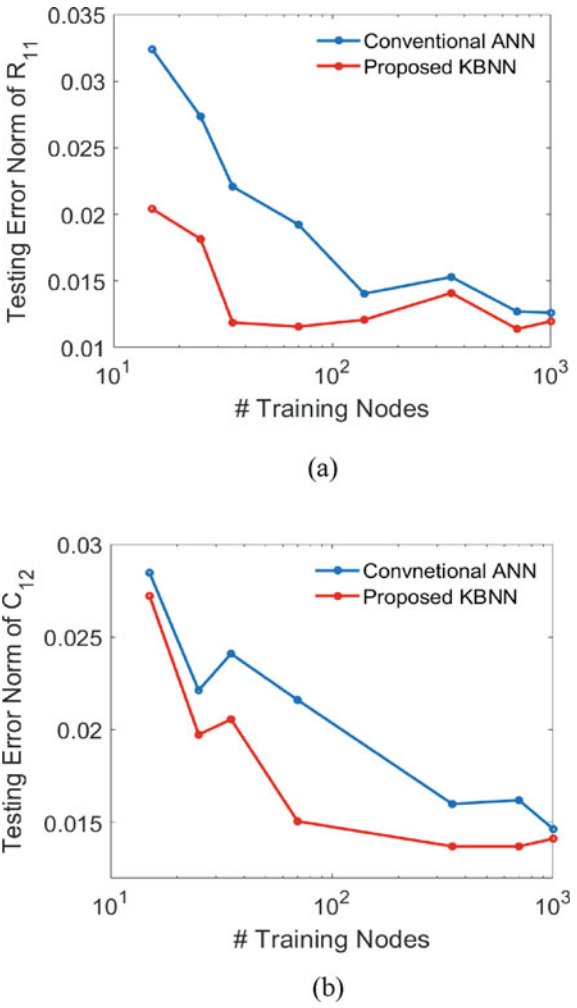


Fig. 7.5 Comparing the convergence of proposed KBNNs with conventional ANNs **a** Scaling of testing error norm with respect to number of training nodes for R_{11}. **b** Scaling of testing error norm with respect to number of training nodes for C_{12}

suggested in Cheng et al. (2018). Figure 7.6 illustrates the imprecision of the closed-form model (Cheng et al. 2018) through transient findings acquired using all three approaches. Because of this, the model presented in Cheng et al. (2018) cannot be used directly to examine the design spectrum of hybrid Cu-graphene networks. On the other hand, Fig. 7.6 demonstrates that the findings of the suggested KBNN are in excellent agreement with those of the standard ANN, even though they were trained from atleast Q/P = 5 times fewer nodes. As a result, the suggested KBNN becomes very attractive to use for design space exploration.

Finally, we utilize the HSPICE W-Element template for doing a full search of the 5-dimensional design spectrum with ten-thousand symmetrically sampled points. Each sample point's W-element model parameters are retrieved using the aforementioned

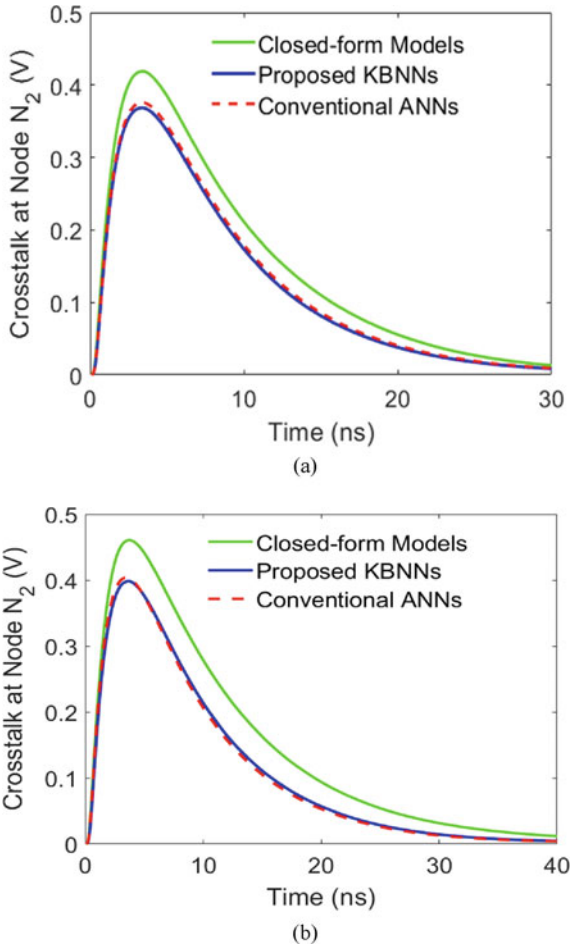


Fig. 7.6 Accuracy comparison of proposed KBNNs, conventional ANNs with models in Cheng et al. (2018); Transient Response at N_2 for **a** $\lambda = [0.54\ -0.91\ 0.49\ -0.90\ -0.39]$ **b** $\lambda = [-0.97\ -0.67\ 0.45\ 0.25\ 0.95]$

three techniques. Three crucial signal integrity (SI) qualities are tracked during the design space sweep: the maximal crosstalk at node N2, the rising time of the transient response at node N1, and the settling time of the transient response at node N1. When comparing the findings produced from the standard ANN to those acquired from the proposed KBNN, Fig. 7.6 shows the RMS errors reported for the SI values above using the closed model of Cheng et al. (2018).

Finally, 10,000 evenly spaced sample points are utilized to do a full sweep of the 5-dimensional design spectrum using the HSPICE W-element template. The same three procedures are used to extract the p.u.l. attributes concerning the W-element template at each sample spot. Time consumption for the response at node N1 to rise from 0.1V to 0.9V (the rise time), time taken for response in node N1 to reach 0.5V (the time delay), and the peak crosstalk with respect to the response at node N2 are 3 crucial SI entities measured during the design space sweep. Table 7.2 displays the

root-mean-square error (RMS) reported for the aforementioned SI values produced using the suggested KBNNs along with the closed-form models from Cheng et al. (2018) in comparison to the results gathered through the traditional ANNs in Fig. 7.1a. Additionally, Table 7.2 also displays the total CPU time spent exploring design space for each approach. Time spent on processing the training data and carrying out the optimization are included in this overall CPU cost. To compute p.u.l. set attributes for each design point and conduct a SPICE simulation at that design point, the suggested KBNN requires Ne = 500 epochs, whereas the traditional ANN requires Ne = 4000 epochs.

The suggested KBNN produces comparable errors (less than 3% error) to the standard ANN, but with substantially reduced CPU cost, as shown in Table 7.2, which also shows that the speedup gained by KBNN compared to traditional ANN is somewhat lower than the predicted Q/P = 5 times. Reasonable, here we taking into account the expense of training in terms of time. Here, we calculate the total amount of CPU time spent on a ten-thousand-point design space coverage using a SPICE simulation, and we see that KBNN has no effect on this time. Meanwhile, the suggested KBNN is substantially slower than the closed-form model in Cheng et al. (2018), but the latter results in much greater errors for all three SI values. This exemplifies how the proposed KBNN outperforms both traditional ANN and the work of Cheng et al. (2018).

Table 7.2 Comparing the precision and numerical efficacy of proposed KBNNs to traditional ANN along with closed-form models of Cheng et al. (2018)

SI quantitites for example 1	Suggested KBNNs (P = 25)			Closed-form model of Cheng et al. (2018)		
	Mean squared error (%)	CPU time (min)	Speedup compared to traditional ANNs (Q = 125)	Mean squared error (%)	CPU time (min)	Speedup compared to traditional ANNs (Q = 125)
Node response time rise N1	2.904	103.352	4.713	11.884	20.8	24.341
Response delay time at node N1	2.526			11.713		
Peak crosstalk for response at node N2	1.052			13.091		

Summary

In this chapter, we develop a neural network to rapidly explore the layout options regarding hybrid Cu-graphene interlink networks on integrated circuits. This chapter's cardinal handout is the spotting of the closed model of Cheng et al. (2018) as a thoroughly efficacious empirical model, which are utilized in the efficient extraction of information about the effects of geometric, material, and physical dimensions of interconnect structures on p.u.l. R, L, C attributes. When combined with data gathered through full-wave EM solvers with the received information, this helps traditional ANNs learn faster. Such ANNs are called knowledge-based neural networks (KBNN).

References

Achar R, Nakhla MS (2001) Simulation of high-speed interconnects. Proc IEEE 89(5):693–728

Bakr MH, Bandler JW, Ismail MA, Rayas-Sánchez JE, Zhang QJ (2000) Neural space-mapping optimization for EM-based design. IEEE Trans Microw Theory Tech 48(12): 2307–2315

Cheng Z et al (2018) Analysis of Cu-graphene interconnects. IEEE Access 6:53499–53508

Meindl JD (2001) Interconnect limits on gigascale integration (GSI). In: 2001 6th international symposium on plasma- and process-induced damage (IEEE Cat. No.01TH8538)

Park H et al (2020) Deep reinforcement learning-based optimal decoupling capacitor design method for silicon interposer-based 2.5-D/3-D ICs. IEEE Trans Components, Packaging Manuf Technol 10(3):467–478

Paul CR (2007) Analysis of multiconductor transmission lines. Wiley

Steinhögl W, Schindler G, Steinlesberger G, Traving M, Engelhardt M (2004) Comprehensive study of the resistivity of copper wires with lateral dimensions of 100 nm and smaller. J Appl Phys 97(2): 023706

Todri-Sanial A, Dijon J, Maffucci A (2017) Carbon nanotubes for Interconnects. Springer

Torun HM, Durgun AC, Aygün K, Swaminathan M (2020) Causal and passive parameterization of s-parameters using neural networks. IEEE Trans Microw Theory Tech 68(10):4290–4304

Watson PM, Gupta KC (1996) EM-ANN models for microstrip vias and interconnects in dataset circuits. IEEE Trans Microw Theory Tech 44(12): 2495–2503

Yang LY, Zhang DH, Li CY, Foo PD (2004) Comparative study of Ta, TaN and Ta/TaN bi-layer barriers for Cu-ultra low-k porous polymer integration. Thin Solid Films 462–463: 176–181

Zhao W, Wang D, Wang G, Yin W (2015) Electrical modeling of on-chip cu-graphene heterogeneous interconnects. IEEE Electron Device Lett 36(1):74–76

Zhang QJ et al (2003) Artificial neural networks for RF and microwave design—from theory to practice 51(4): 1339–1350

Zhang QJ, Gupta KC (2000) Neural networks for RF and microwave design (Book+ Neuromodeler Disk). Artech House, Inc.

Chapter 8
A Comprehensive Analysis of Emerging Variants of Swarm Intelligence for Circuits and Systems

Surendra Hemaram, Aksh Chordia, and Jai Narayan Tripathi

Abstract With the growing complexity in semiconductor technology, the design and optimization of VLSI circuits are automated with a high degree of reliability and precision. There are many applications of VLSI circuits and systems that require solving optimization problems of large-scale degrees. While analytical methods may suffer from slow convergence and the curse of dimensionality, metaheuristics-based swarm intelligence algorithms are proven to be an efficient alternative. This chapter presents the applications of emerging swarm intelligence-based techniques for the optimization of VLSI circuits and systems. In this chapter, the recent advances in particle swarm optimization algorithms are used to solve the optimization problem encountered in the VLSI industry. The efficacy of these algorithms is illustrated using two practical case studies. The first case study is related to the design of a 2.4 GHz CMOS LC tank oscillator, and the second one is optimizing a practical power delivery network. A comparative analysis of the performances of these algorithms is presented. This study helps the reader to choose an appropriate PSO optimization algorithm for their required application.

Keywords VLSI circuits and systems · Swarm intelligence · Variants of particle swarm optimization

S. Hemaram · A. Chordia (✉) · J. N. Tripathi
Indian Institute of Technology, Jodhpur, India
e-mail: chordia.2@iitj.ac.in

S. Hemaram
e-mail: surendra.1@iitj.ac.in

J. N. Tripathi
e-mail: jai@iitj.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_8

Introduction

The continuous scaling down of switching devices increases the complexity of modern-day Integrated Circuits (ICs). The design and optimization of Very Large-Scale Integration (VLSI) circuits have become complex with increasing complexity. Due to the downscaling of the devices, the sensitivity of circuit response to technological variation increases. For getting a high parametric yield, it is very important to reduce the performance variations so that VLSI circuits with close tolerance can be manufactured. Thus, computational intelligence techniques emerge to fulfill this requirement for optimizing the technological parameter values (Hocevar et al. 1988; Hoenigschmid et al. 1997; Chen et al. 2000; Rutenbar et al. 2007; Lyu et al. 2018). These techniques formulate the design problems as constrained optimization problems to optimize the technological parameters in circuits.

The primary issue lying in optimization is searching for the global optimum results, for which swarm intelligence-based algorithms are proven to be effective. One of these algorithms, namely the Particle Swarm Optimization (PSO), is a widely popular swarm intelligence-based metaheuristic optimization algorithm (Kennedy and Eberhart 1995). The PSO is proven to be one of the prominent optimization algorithms in the field of VLSI circuits and systems (Park and Allstot 2004; Park et al. 2004; Kamisetty et al. 2011; Khan and Khalifa 2011; Sabat et al. 2009; Tripathi et al. 2013; Kim et al. 2018; Hussein and Jarndal 2018; Garbaya et al. 2018). PSO is becoming more popular, because of its relatively fast convergence, ease of implementation, and has more effective memory capacity (del Valle et al. 2008). In this chapter, the applications of different PSO variants are demonstrated to solve the practical problems of circuits and systems. The recent advances in PSO techniques have resulted in various variants of the standard PSO algorithm having a better performance. The recently introduced PSO variants used in this chapter are Linear Decreasing Inertia Weight PSO (LDIW-PSO), Chaos Enhanced PSO (CE-PSO), Adaptive Inertia Weight PSO (AIW-PSO), and Gaussian Quantum Behaved PSO (GQPSO). The efficacy of these algorithms is performed using two practical case studies; one is at the system level, and another one is at the circuit level. This study also employs two population-based state-of-the-art metaheuristic optimization algorithms, namely real coded genetic algorithm (GA) and differential evolution (DE), and their comparison with different PSO variants for a holistic analysis (Man et al. 1996; Deb and Kumar 1995; Deb and Deb 2014; Chuang et al. 2016; Storn and Price 1997; Das and Suganthan 2011).

There are two case studies presented for the purpose of comparison and benchmarking. The first case study is a circuit-level problem that is related to the design of a 2.4 GHz CMOS LC tank oscillator with minimized phase noise. This circuit is considered for study as the oscillators are integral parts of almost all VLSI circuits having essential applications in phase-locked loops, clock recovery circuits, frequency synthesizers, etc. The most important quality metric of these oscillators is phase noise which is defined as the noise power relative to the carrier power occurring at a particular offset frequency away from the carrier frequency (Lee and Hajimiri

2000). Here, the automated framework of optimization using PSO variants is used to minimize oscillator phase noise.

The second case study is a system-level problem related to optimizing the power delivery network (PDN) to minimize supply noise. In general, as a solution to this problem, decoupling capacitors are used to minimize the supply noise by minimizing the cumulative impedance of PDN (Swaminathan and Engin 2007). When the number of capacitors and the number of the port for placing them is more, then intuitive selections and placements of decoupling capacitors become very difficult in designing a PDN. In such cases, the importance of optimization algorithms becomes crucial to solving this problem. In the literature, there are several methodologies available for the optimal selection and placement of decoupling capacitors using different optimization techniques (Kahng 2006; Piersanti et al. 2018; Choi and Swaminathan 2011; Tripathi et al. 2014; Erdin and Achar 2019; Xu et al. 2020). Here, the cumulative impedance of PDN is minimized by optimal selections and placements of decoupling capacitors by the different strategies of PSO.

The main contributions of this chapter can be summarized as

1. The chapter provides an insight to recent and emerging variants of the PSO algorithms and their applications for the optimization of Circuits and Systems. Two diverse case studies of optimization are presented to compare and benchmark these algorithms. Passive design.
2. It introduces an efficient strategy of optimization for the circuits and system-level problems. For the circuits level, a framework is developed that interfaces the simulation engine (Cadence-Spectre) and equation engine (MATLAB) in real time. The circuit is implemented in the simulation engine, and the optimization algorithm is implemented in the equation engine. Such an optimization strategy is generic in nature and can easily be extended for the optimization of the other case studies as well.
3. A comparative analysis of the performance of these algorithms is presented. Also, other state-of-the-art optimization strategies like Genetic Algorithm and Differential Evolution methods have been considered for solving the problem to prove the efficacy of PSO-based optimization.
4. Such a ubiquitous demonstration interests the reader as this optimization problem and its implementation are generic in nature and can be extended to any other optimization problem easily.

Particle Swarm Optimization

Particle swarm optimization (PSO) is a class of swarm intelligence-based metaheuristic techniques invented by Kennedy and Eberhart in 1995 (Kennedy and Eberhart 1995; Shi and Eberhart 1998). The PSO is a population-based optimization algorithm simulating swarm behavior such as fish and bird schooling. In PSO, swarms are also

called particles, and these particles move toward the optimum position by following a specific trajectory based upon their memory.

The algorithms start with random initialization of the population within the defined search space. As time progresses, a particle moves its position in the search space based on two factors: its best position so far ('local best' or 'lbest') and the best position obtained by any particle so far ('global best' or 'gbest'). The new positions and velocities of all the particles are defined based on these two equations and are given as follows:

$$v_{id}^{k+1} = v_{id}^{k} + r_1 c_1 (x_{id}^{p} - x_{id}^{k}) + r_2 c_2 (x_{d}^{g} - x_{id}^{k}) \tag{8.1}$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \tag{8.2}$$

where n represents population size, the maximum number of iterations is denoted by N, D represents the number of dimensions, i.e. $d = 1, 2, \ldots, D$; k represents the current iteration, x_{id}^{p} represents local best of each particle and x_{d}^{g} is the current global best, c_1 and c_2 are acceleration coefficients, and r_1 and r_2 represent random numbers in the interval (0, 1). The steps for PSO are summarized as a pseudocode in Algorithm 8.1 (Kennedy and Eberhart 1995; Shi and Eberhart 1998).

Algorithm 8.1 PSO

```
1: begin
2: Objective function f, initialize v and x for n particles having D number of dimensions, define
   maximum number of iterations N.
3: Calculate the initial minimum fitness f_min^{k=0}.
4: while (k < N)
5:     for i = 1, 2, ..., n do
6:         for d = 1, 2, ..., D do
7:             Update the velocity (v) and position (x).
8:         end for
9:     Calculate f at new position x_id^{k+1}.
10:    Compute minimum fitness f_min^{k+1}.
11:    end for
12:    Compute the current personal best x_id^p.
13:    Compute the current global best x_d^g and f(x_d^g).
14: end while
15: Output: x_d^g and f(x_d^g).
16: end
```

Recent Advances in Particle Swarm Optimization

In this section, different variants of the PSO are briefly discussed. The different variants of PSO used in this chapter are Linear Decreasing Inertia Weight PSO (LDIW-PSO), Chaos Enhanced PSO (CE-PSO), Adaptive Inertia Weight PSO (AIW-PSO), and Gaussian Quantum Behaved PSO (GQPSO).

Linear Decreasing Inertia Weight PSO

The inertia weight in PSO is able to control the global exploration and local exploitation abilities. The inertia weight strategy in PSO was introduced in Shi and Eberhart (1998). PSO with linearly decreasing inertia weight (LDIW-PSO) is endorsed in Shi and Eberhart (1998), Xin et al. (2009), Shi and Eberhart (1999) due to its good performance over a large number of optimization problems. The velocity and position vector for inertia weight PSO is defined as follows:

$$v_{id}^{k+1} = w\ v_{id}^{k} + r_1 c_1 (x_{id}^{p} - x_{id}^{k}) + r_2 c_2 (x_{d}^{g} - x_{id}^{k}) \tag{8.3}$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \tag{8.4}$$

where w represents the inertia weight and the remaining parameters are similar to the classical PSO algorithm.

In linear decreasing inertia weight (LDIW) strategy, weight is a function of maximum number of iterations as well as current time step (i.e. iteration), and it is defined as follows (Xin et al. 2009):

$$w = w_f + (w_i - w_f) \times \frac{N - k}{N} \tag{8.5}$$

where N represents the maximum number of iterations, w_i and w_f represent the initial and final inertial weights, respectively, and k denotes the current iteration. The steps to implement LDIW-PSO are similar to Algorithm 8.1 with updated velocity Eq. (8.3) which incorporates the inertia weight strategy define by Eq. (8.5).

Chaos Enhanced PSO

Chaos is a bounded unstable dynamic behavior that shows sensitive dependency on initial conditions and includes infinite unstable periodic movements in nonlinear systems. Chaos mechanism became a very helpful mechanism to avoid the PSO from getting trapped into the local optimum in a searching process (Alatas et al. 2009; Feng et al. 2007). The basic idea of chaos-based PSO is the incorporation of chaotic mapping to the inertia weight coefficient. In general, logistic mapping is used as chaotic mapping, and it is defined as follows (Feng et al. 2007):

$$y_{k+1} = y_k \times (1 - y_k)\mu \tag{8.6}$$

where $3.57 < \mu \leq 4$.

Generally, $\mu = 4$ is used in the logistic mapping. But it does not assure chaos on initial values of $y_0 \in (0, 0.25, 0.5, 0.75, 1)$ that may arise during the initial generation process (Hong et al. 2016). In chaos enhanced PSO (CE-PSO), an enhanced version

of chaotic inertia weight is used which employs sine chaotic mapping (Hong et al. 2016; Arasomwan and Adewumi 2014).

The inertia weight strategies for CE-PSO are given as follows (Hong et al. 2016):

- Choose the random number r in the interval (0, 1) and make $y_0 = r$.
- Formulate the chaotic mapping as

$$y_{k+1} = |\sin(\frac{\pi y_k}{r})| \tag{8.7}$$

  where y_k, $y_{k+1} \in (0, 1)$, k is current iteration, and r represents the random number in the interval (0, 1). The absolute sign in the expression ensures that the next iteration process in chaos space has $y_{k+1} \in (0, 1)$.
- Formulate the inertia weight w as follows:

$$w = 0.5 \times r + 0.5 \times y_{k+1} \tag{8.8}$$

The steps to implement the CE-PSO are similar to Algorithm 8.1 with updated velocity Eq. (8.3) which incorporates the inertia weight strategy define by Eq. (8.8).

Adaptive Inertia Weight PSO

The aim of adaptive inertia weight PSO (AIW-PSO) is also to enhance the exploration and exploitation ability of PSO algorithms. The exploration means the ability of the algorithm to abandon the current peak and search for better solutions. The exploitation characteristic of PSO states that all the particles converge to the same peak and remain there. To have an adaptive inertia weight strategy, firstly, it is required to determine the swarm situation at each iteration. The adaptive inertia weight (AIW) strategy involves the concept of particle success rate, which acts as a feedback quantity to discover the swarm's position in the search space.

The success of particle i at iteration k in a minimization problem is defined as follows (Nickabadi et al. 2011):

$$S(i,k) = \begin{cases} 1, & \text{if } f(pbest_{id}^{k}) \leq f(pbest_{id}^{k-1}) \\ 0, & \text{if } f(pbest_{id}^{k}) = f(pbest_{id}^{k-1}) \end{cases} \tag{8.9}$$

where $pbest_{id}^{k}$ is the best position found by particle i until iteration k, i.e. (x_{id}^{p}), and $d = 1, 2, \ldots, D$ is variable for dimension space. The success percentage of a swarm is calculated using this success value and is defined as

$$P_s(k) = \frac{\sum_{i=1}^{n} S(i,k)}{n} \tag{8.10}$$

where n represents the number of particles and $P_s \in (0, 1)$ is success percentage of the swarm. The inertia weight w can be formulated as a function of P_s and is defined as follows (Nickabadi et al. 2011):

$$w = w_f + (w_i - w_f) \times P_s \tag{8.11}$$

where w_i and w_f represent the initial and final inertia weight, respectively.

The steps to implement the AIW-PSO are summarized as a pseudocode in Algorithm 8.2 (Nickabadi et al. 2011).

Algorithm 8.2 AIW-PSO

begin
Initialize the velocity v and location x of n particles having D dimensions ($d = 1, 2, \ldots, D$), define the maximum number of iterations N.
Initialize personal best to current position of particle, $x_{id}^p = x_{id}$.
Calculate the initial global best x_d^g.
Initialize inertia weight function $w = 1$.
while ($k < N$)
    $S = 0$.
    for $i = 1, 2, \ldots, n$
        for $d = 1, 2, \ldots, D$
            Update Velocity (v) and Position (x).
        end for
        if ($f(x_{id}) < f(x_{id}^p)$)
            $S = S + 1$
            Update personal best $x_{id}^p = x_{id}$.
            if ($f(x_{id}) < f(x_d^g)$).
                Update global best $x_d^g = x_{id}$.
            end if
        end if
    end for
    Compute: $P_S = \frac{S}{n}$ and Update inertia weight (8.11).
end while
end

Gaussian Quantum Behaved PSO

The quantum-behaved PSO was developed by Sun et al. (2011). In the quantum model of a PSO, the situation of a particle is obtained by a wave function $\psi(x, k)$ (Schrödinger equation) instead of computation of the position (x) and velocity (v) vectors. In this context, the probability of the particle appearing in position x_i is given by the probability density function $|\psi(x, k)|^2$, the form of which depends on the potential field the particle lies in. The quantum-PSO algorithm has been shown to solve a wide range of continuous optimization problems successfully, and many effi-

cient strategies have been proposed to improve the algorithm (Mikki and Kishk 2006; Omkar et al. 2009; dos Santos Coelho 2008). In Gaussian quantum PSO (GQPSO), random numbers are generated using the absolute value of Gaussian probability distribution with zero mean and unit variance, i.e. $G = |(N(0, 1))|$ instead of the uniform probability distribution functions in the range (0, 1). Generating random numbers using the Gaussian probability distribution function for the stochastic coefficients of PSO may provide a good compromise between the probability of having a large number of small amplitudes around the current points and a small probability of having higher amplitudes, which may allow particles to move away from the current point and escape from local minima. The update of the position of particles in GQPSO is given as follows (Sun et al. 2011; Coelho 2010):

$$x_i(k+1) = \begin{cases} P_l + \beta.|Mbest_i - x_i(k)|.\ln(\frac{1}{G}), & \text{if } r \geq 0.5 \\ P_l - \beta.|Mbest_i - x_i(k)|.\ln(\frac{1}{G}), & \text{if } r < 0.5 \end{cases} \tag{8.12}$$

where $G = |(N(0, 1))|$, β is contraction–expansion coefficient, r is random number in the interval (0, 1), and the $Mbest$ is mean of personal best (x_i^p) position of all the particles and it can be defined as follows:

$$Mbest_i = \frac{1}{N}\sum_{i=1}^{n} x_i^p(k) \tag{8.13}$$

P_l is local attractor and is defined as follows:

$$P_l = \frac{G\, x_{id}^p + g\, x_d^g}{G + g} \tag{8.14}$$

where $g = |(N(0, 1))|$, x_{id}^p is personal best, x_d^g is global best, $d = 1, 2, \ldots, D$ and $i = 1, 2, \ldots, n$.

The steps for GQPSO are summarized as a pseudocode in Algorithm 8.3 (Sun et al. 2011; Coelho 2010).

Case Studies

In this section, two case studies which are used as examples for the purpose of benchmarking are explained. Both case studies are having practical relevance in the VLSI industry—one at the circuit level and another one at the system level.

Algorithm 8.3 GQPSO

begin
Objective function f, initialize x for n particles having D dimensions ($d = 1, 2, \ldots, D$), define N.
Calculate the initial minimum fitness $f_{min}^{k=0}$.
while ($k < N$)
  for $i = 1, 2, \ldots, n$ **do**
    Update the position as follows:
    if $r \geq 0.5$
      $x_{id}(k+1) = P_l + \beta.|Mbest_{id} - x_{id}(k)|.ln(\frac{1}{G})$
    else
      $x_{id}(k+1) = P_l - \beta.|Mbest_{id} - x_{id}(k)|.ln(\frac{1}{G})$
    end if
    Calculate f at new position $x_i(k+1)$.
    Compute minimum fitness f_{min}^{k+1}.
  end for
  Compute the current personal best x_{id}^{p}.
  Compute the current global best x_d^g and $f(x_d^g)$.
end while
Output: x_d^g and $f(x_d^g)$.
end

Case Study I: Phase Noise Minimization in an Oscillator

Circuit Description

In most of the VLSI systems, oscillators fulfill the clock signal requirements. An important class of oscillators is CMOS LC tank oscillators that are widely used for their relative ease of implementation and superior output characteristics such as jitter and phase noise (Lee and Hajimiri 2000; Hajimiri and Lee 1998). This case study aims to improve the phase noise quality of a 2.4 GHz cross-coupled CMOS LC tank oscillator as shown in Fig. 8.1. Phase noise is defined as the irregularities in the phase of the periodic output signal arising due to short-term fluctuations (Koukab 2011; Lee and Hajimiri 2000). It is represented as the noise power relative to the carrier power, measured at a certain offset frequency away from the carrier frequency. Phase noise in an oscillator directly disturbs the timing accuracy and affects the signal-to-noise ratio as well (Hajimiri and Lee 1998; Tripathi et al. 2011).

It is worth noting that phase noise of oscillators is directly affected by the lengths and widths of the constituting transistors along with inductor and capacitor values (Tripathi et al. 2011). Hence, this study aims to optimize the phase noise when these parameters are varied within a predefined design space. Here, the objective function, i.e. the phase noise is considered at a specific offset frequency (100 kHz) which needs to be minimized. In this chapter, all the phase noise results are presented for this specific frequency point. Phase noise at only a specific offset is reported for ease of presentation; however, it is implied that minimizing the phase noise at a single frequency point automatically minimizes the phase noise for the complete

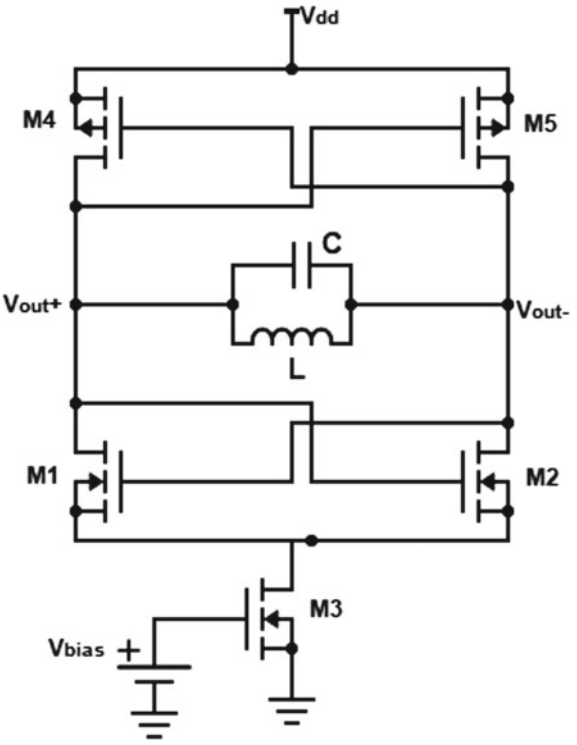


Fig. 8.1 CMOS LC tank oscillator

frequency range. Such an offset frequency is chosen to be reported due to its applications in wireless communication domain. The technology node for this study is 180 nm. The optimization is performed in MATLAB acting as a computational tool, and the Cadence-Spectre is chosen to be the simulation engine.

Automated Framework for Phase Noise Minimization

The objective is to minimize the phase noise with a constraint that the fundamental frequency of the oscillator does not deviate significantly. The parameters that directly affect the phase noise are taken into consideration as shown in Table 8.1. These parameters become the decision variables for the optimization problem.

An automated framework is designed in the C shell scripting language that enables the simulation engine to talk to the computational tool in real time. It enables the computational tool to provide the inputs for simulations, invoke the simulations for these parameters, and parse back the simulation results to the computational tool. Figure 8.2 shows the steps involved in one data cycle from optimization engine to simulation engine to giving back the results to the optimization engine. The same framework can be used for all the variants of PSO. Such a framework is generic in nature and carries the potential to be used for other optimization problems as well.

Table 8.1 List of decision variables and their specifications for test circuit I

Parameter	Unit	Lower bound	Upper bound
Width of $M1$ (W1)	μm	4.08	5.04
Width of $M2$ (W2)	μm	4.08	5.04
Width of $M3$ (W3)	μm	4.08	5.04
Width of $M4$ (W4)	μm	4.08	5.04
Width of $M5$ (W5)	μm	4.08	5.04
Inductor value (L)	nH	4.00	4.80
Capacitor value (C)	pF	0.90	1.10

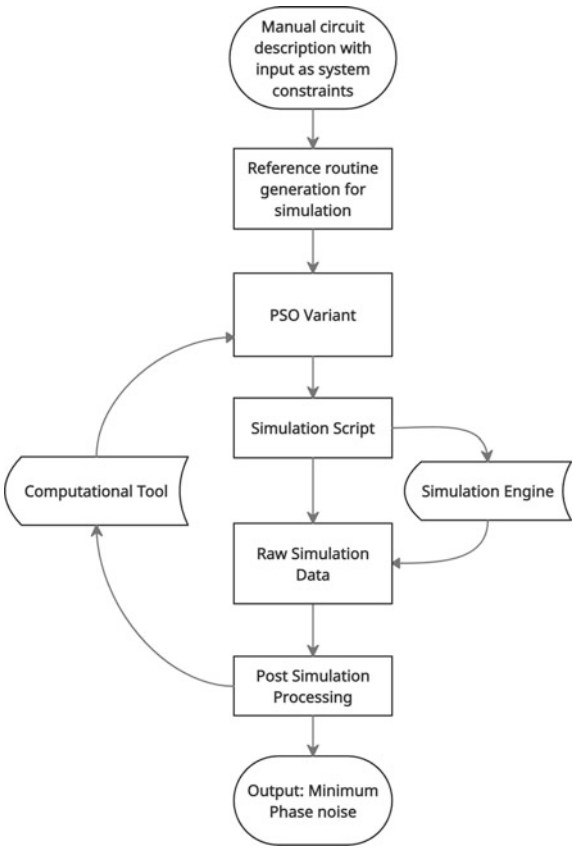


Fig. 8.2 Framework of design automation between the computational tool and simulation engine

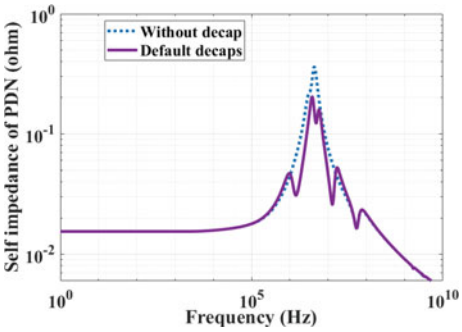


Fig. 8.3 Self-impedance of PDN with default decoupling capacitors

Case Study II: Optimization of Decoupling Capacitors in a PDN

System Description

In an IC, a PDN is responsible for supplying the power to the entire system by providing the V_{dd} and the ground voltages. Typically, a PDN consists of several components which includes DC–DC converters (also called voltage regulator modules (VRM)), decoupling capacitors, interconnects onboard and packages, etc. (Tripathi et al. 2019). In a PDN, there exist inductive and capacitive effects associated with all these components, which affects the system performances in different frequency ranges (Swaminathan and Engin 2007). It is crucial to control the variations in the supply voltage to system assertively because these variations can lead to false switching in dynamical digital circuits along with timing and delay variations (Haihua Su et al. 2003). In analog circuits, these imperfections can lead to improper biasing of the transistors leading to signal corruptions (Tripathi et al. 2019).

The most common and powerful technique to minimize the power supply noise is by intentionally placing the on-chip capacitors ('*decaps*') onto the PDN. These decap helps reduce the cumulative impedance of PDN below a certain limit called the target impedance (Z_T). However, selecting and placing the capacitors intuitively is not feasible when there are innumerable capacitors available commercially. In such cases, optimization techniques are required to solve such problems.

Figure 8.3 shows the self-impedance of a practical PDN at the pad of the IC (i.e. at silicon). This impedance is the equivalent impedance of PDN component modeling such as VRM, package, PCB, and the on-chip load. For the PCB and the package, *S*-parameters are extracted from a commercial 3D solver, and for the on-chip PDN, a Chip Power Model (CPM) is used (Kulali et al. 2007).

The reduction in cumulative impedance of the PDN by appropriate placement of the decaps can be evidenced by Fig. 8.3. Here, two cases are shown, one shows the self-impedance of the PDN without decaps (indicated by a dotted line) and the another one shows the self-impedance of the PDN with intuitive placement of decaps (indicated by a solid line). The maximum value of the self-impedance of the PDN without decaps is 361.2 mΩ. With decaps, the value reduces to 203.9 mΩ.

Decap Optimization Framework

This optimization problem is solved by considering the S-parameter data of both PDN and the decaps. This data (equivalent S-parameters data) can be extracted for the available ports (without placing decoupling capacitors) using the commercially available 3D solvers. The S-parameter data is then converted to the corresponding Z-parameter (Z_{pdn}) data by simple transform relationship.

Now, in order to reduce the impedance of PDN, the decaps are placed at the corresponding ports and the resultant equivalent impedance Z_{eq} (at the port where impedance is observed) of the network is given as follows (Kahng 2006):

$$Z_{eq} = (Z_{pdn}^{-1} + Z_{decap}^{-1})^{-1} \tag{8.15}$$

where Z_{decap} represent the Z-parameter matrix of the decaps which only carry the diagonal elements. Alternately, (8.15) can be be written as

$$Z_{eq} = (Y_{pdn} + Y_{decap})^{-1} \tag{8.16}$$

where

$$Y_{decap} = Z_{decap}^{-1} = \begin{bmatrix} \frac{1}{z_1} & 0 & 0 \ldots & 0 \\ 0 & \frac{1}{z_2} & 0 \ldots & 0 \\ \vdots & \vdots & \vdots \ddots & \vdots \\ 0 & 0 & 0 \ldots & \frac{1}{z_n} \end{bmatrix}$$

where z_k are z_{11} values of decoupling capacitors at k_{th} port.

It is worth noting that the Z-parameter data considered here has a dimension with respect to frequency also. The shape of Z-matrix is $p \times p \times f$ where p represents the number of ports and f is the number of frequency points. Generally, the impedance is measured at the port which is nearer to the chip. Throughout this case study, measurement of the impedance is done at port 1 of the considered PDN. Correspondingly, the objective function becomes to maximize the self-impedance at this port. This can be formulated as

$$Z_{obj} = \max(Z_{eq}(1, 1)) \tag{8.17}$$

Here, the number of decision variables D is equal to 2 for each capacitor: port number and the capacitor index (the index of decap to be considered).

Algorithm 8.4 summarizes the steps involved for the computation of the optimum number of decaps to stabilize the PDN using the different variants of PSO. Here, $Z_{p\times p\times f}$ is the impedance of PDN without decaps, $Z_{1\times f}$ is the impedance of decoupling capacitors, p is the number of ports available for the placement of decaps, f is the number of frequency points taken for analysis, C represents the index of decaps, P represents the port numbers corresponding to capacitor indices, and N_D represents the number of decaps.

Algorithm 8.4 Decap Optimization Framework

begin
Input: $Z_{p\times p\times f}$, $Z_{1\times f}$, Z_T.
Output: C, P, $max(Z_{eq}(1,1))$.
Objective function: $Z_{obj} = max(Z_{eq}(1,1))$,
$(Z_{eq})_f = (Z_{pdn}^{-1} + Z_{Decap}^{-1})_f^{-1}, \forall\ f \in [0, f_{max}]$.
Initialize $N_D = 0$.
while $(Z_{obj} > Z_T)$ **do**
  $N_D = N_D + 1$.
  Initialize $k = 0$, define N.
  for $i = 1, 2, \ldots, n$ **do**
    for $i_c = 1, 2, \ldots, N_D$ **do**
      Generate initial population (v_i and x_i).
    end for
  end for
  Calculate initial minimum $Z_{obj}^{k=0}$.
  while $(k < N)$ **do**
    $k = k + 1$.
    Update inertia weight w.
    for $i = 1, 2, \ldots, n$ **do**
      for $i_c = 1, 2, \ldots, N_D$ **do**
        Update velocity v_i and position x_i
      end for
      Compute current minimum of Z_{obj}^{k+1}.
    end for
    Compute personal best (x_{id}^p) and global best (x_d^g) .
    Compute Z_{obj} at x_d^g.
  end while
end while
Out: Z_{obj}, x_d^g:(P, C).
end

Results and Comparative Analysis

In order to demonstrate that these algorithms can adequately evaluate and reproduce the results, each PSO variant is executed ten times. The population size n is taken to be 30, and the maximum number of iterations N is set to 25. The c_1 and c_2 are taken as 1.5 each, and the values of w_i and w_f are 0.9 and 0.4, respectively. The algorithms were implemented using MATLAB R2019b and were run on a personal portable computer with 8 GB RAM and Intel $i5$ 9th Gen 2.4 GHz cores.

Case Study I

Table 8.2 compares the performance of all the algorithms for ten independent runs. It gives the average, the best case result of phase noise (P_n) for the considered oscillator out of these ten runs along with the average CPU time for each of the algorithms. Also, the fundamental frequency (f_0) of the considered oscillator is reported at the best case result for each algorithm. Table 8.3 reports the optimized values of

Table 8.2 Performance comparison of the algorithms for Case Study I out of 10 runs

Criterion	PSO	LDIW-PSO	CE-PSO	AIW-PSO	GQPSO	DE	GA
Average phase noise (dBc/Hz)	−115.16	−115.24	−115.22	−115.30	−115.50	−115.28	−114.76
Minimum phase noise (dBc/Hz)	−115.54	−115.52	−115.61	−115.45	−115.62	−115.81	−115.30
Frequency (f_0) at min. Phase noise (GHz)	2.37	2.37	2.38	2.28	2.34	2.33	2.35
Average CPU time (s)	1180.6	1182.7	1176.5	1163.8	1175.3	1168.1	1171.4

Table 8.3 Best case (out of 10 runs) phase noise and optimized value of decision variables for different optimization algorithms

Algorithm	P_n (dBc/Hz)	f_0 (GHz)	W1 (μm)	W2 (μm)	W3 (μm)	W4 (μm)	W5 (μm)	L (nH)	C (pF)
PSO	−115.54	2.37	4.08	4.08	4.74	4.08	4.22	4.00	1.10
LDIW-PSO	−115.52	2.37	4.08	4.08	4.66	4.08	4.08	4.00	1.10
CE-PSO	−115.61	2.38	4.08	4.11	4.84	4.08	4.10	4.00	1.10
AIW-PSO	−115.45	2.28	4.08	4.08	4.08	4.08	4.59	4.35	1.10
GQPSO	−115.62	2.34	4.12	4.16	5.04	4.87	4.11	4.14	1.10
DE	−115.81	2.33	4.08	4.08	4.08	4.67	5.04	4.80	1.10
GA	−115.30	2.35	4.08	4.08	4.08	4.21	4.08	4.11	1.10

the considered seven decision variables as obtained from the best case result from each study. Figure 8.4 exhibits the best case phase noise convergence plots of all the considered algorithms. In order to understand the phase noise optimization for the complete spectrum (10 Hz–1 GHz offset), Fig. 8.5 is plotted. It shows the minimum noise obtained at each offset frequency point for each of the considered algorithm.

For this optimization study at circuit level, the following observations can be made:

- The best case results obtained by all the PSO algorithms are comparable to each other. The average CPU time for all the PSO algorithms is also similar. In case of CPU time, the time required to interface the data between EDA tool and computation tool is neglected.

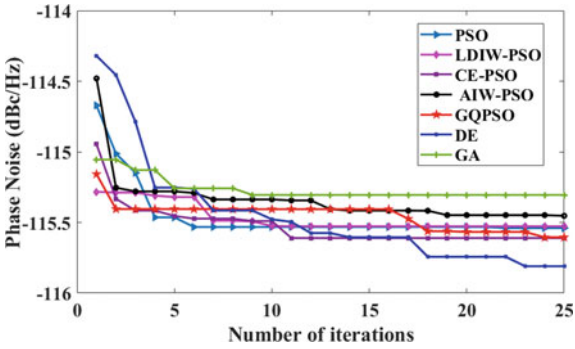


Fig. 8.4 Phase noise convergence plot of all the algorithms

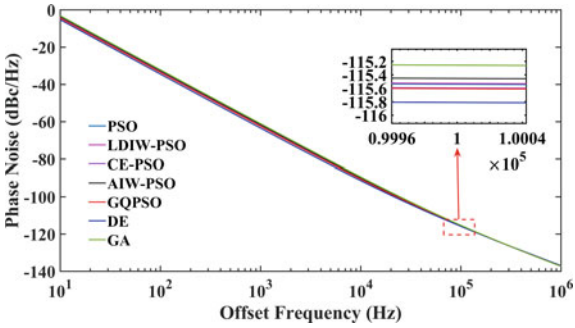


Fig. 8.5 Phase noise plot

- CE-PSO and GQPSO perform better than the rest of PSO variants in an average sense and can be considered the most reliable for such studies.
- CE-PSO and GQPSO were able to minimize the phase noise to −115.62 and −115.61 dBc/Hz at 100 kHz, with the fundamental frequency being 2.38 and 2.34 GHz, respectively.
- The results obtained by PSO variants are comparable to DE and are better than GA.

Case Study II

In this work, the Z-parameter data used for the analysis is having dimensions of 22 × 22 × 1571, where 22 represents the number of ports and 1571 is the number of frequency points. For minimizing the equivalent impedance of the PDN, the observing port is marked as port-1 and the impedance Z_{11} is measured at port-1. Thus, there are 21 ports available for placing decaps. The target impedance of the PDN is defined as 60 mΩ as per the system requirement. Table 8.4 shows the performance comparison of all the algorithms for 10 independent runs. Table 8.5 shows the optimized impedance obtained with the minimum number of decaps by different algorithms. The best case convergence plots of the algorithms are shown in Fig. 8.6.

Table 8.4 Performance comparison of the algorithms for Case study II out of 10 runs

Criterion	PSO	LDIW-PSO	CE-PSO	AIW-PSO	GQPSO	DE	GA
Minimum no. of decaps	6	4	4	4	4	4	4
Average no. of decaps	8	5	5	5	5	6	6
Average value of Imp (mΩ)	57.1	57.5	56.9	56.8	56.1	57.6	57.5
Imp (mΩ) with Min no. of decaps	59.4	57.5	56.3	55.9	55.5	56.9	56.3
Average computation (CPU) time (s)	791.6	438.7	485.2	590.3	412.5	553.2	549.7

Table 8.5 Default and optimized PDN impedance

Criterion	Self-impedance
Without decaps	361.2
Default decaps	203.9
Optimized (with 6 decaps by PSO)	59.4
Optimized (with 4 decaps by LDIW-PSO)	57.5
Optimized (with 4 decaps by CE-PSO)	56.3
Optimized (with 4 decaps by AIW-PSO)	55.9
Optimized (with 4 decaps by GQPSO)	55.5
Optimized (with 4 decaps by DE)	56.9
Optimized (with 4 decaps by GA)	56.3

For this optimization problem:

- The minimum number of decaps obtained by PSO variants are equal to four only, but GQPSO gives the best optimized impedance value compared to others algorithm for the same number of decaps which is shown in Fig. 8.7.
- The GQPSO algorithm reduced the PDN impedance from 203.9 to 55.5 mΩ with only four decaps and is shown in Table 8.5. The GQPSO also achieves the minimum average impedance of the PDN and the minimum average computation time which is shown in Table 8.4.
- The results obtained by PSO variants are better than the classical PSO (without inertia weight) and also comparable to other state-of-the-art algorithms, namely GA and DE which are shown in Table 8.4.
- The optimized impedance value obtained by GQPSO is better than the GA and DE for the same number of decaps and is shown in Table 8.5.

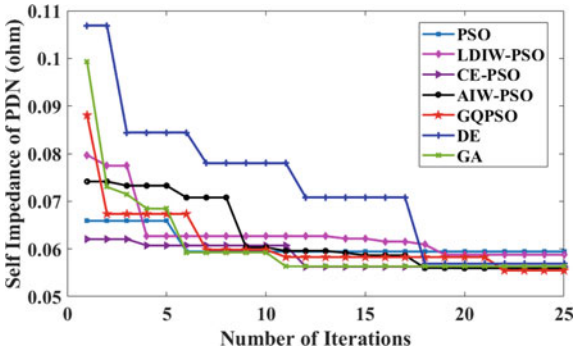


Fig. 8.6 Convergence plot of self-impedance for all the algorithms

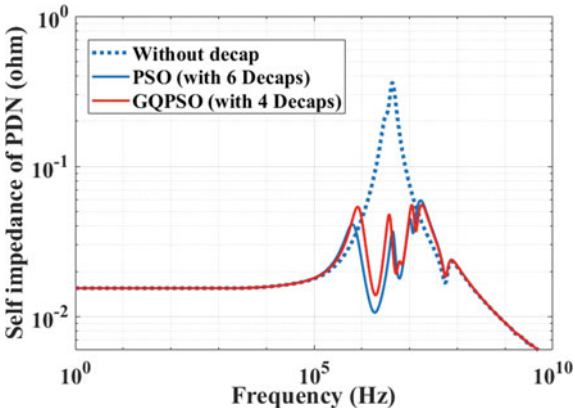


Fig. 8.7 Self-impedance of PDN with optimum number of decaps

Table 8.6 Optimized decoupling capacitors with their corresponding port numbers by GQPSO

Capacitor index	Capacitor model	Manufacturer	Port number
7	C1005X5R0J475M050BC	TDK	6
2	GRT155R61E105KE01	Murata	10
4	GRT155R61A684KE01	Murata	11
3	CL05B104KO5NNN	Samsung	16

Table 8.6 shows the list of the optimum number of decaps with their index numbers (C) and corresponding port numbers (P). Figure 8.7 shows the self-impedance profile of the PDN without decaps, default decaps, and optimized decaps by GQPSO.

Conclusion

In this work, the application of different PSO variants is illustrated to optimize the VLSI circuits and systems. The efficacy of each algorithm is illustrated using two

practical case studies; the first one is a circuit level optimization problem, i.e. phase noise minimization in a oscillator circuit. Also, a system-level optimization, i.e. optimal selection and placement of decoupling capacitors in a power delivery network, is illustrated. The circuit level optimization strategy discussed here introduces the utilization of a real-time interfacing strategy between the simulation and equation engine. Such an optimization strategy is generic in nature and can be adapted for similar optimization problems easily. The results obtained by PSO variants are also comparable to two other population-based state-of-the-art metaheuristic algorithms, namely differential evolution (DE) and real coded genetic algorithm (GA). The GQPSO outperforms the other PSO variants.

References

Alatas B, Akin E, Ozer AB (2009) Chaos embedded particle swarm optimization algorithms. Chaos Solit Fractals 40(4):1715–1734. https://doi.org/10.1016/j.chaos.2007.09.063

Arasomwan AM, Adewumi AO (2014) An investigation into the performance of particle swarm optimization with various chaotic maps. Math Probl Eng 2014:178959

Chen HM, Samudra GS, Chan DSH, Ibrahim Y (2000) Global optimization for digital MOS circuits performance. IEEE Trans Comput Aided Des Integr Circuits Syst 19(1):161–164. https://doi.org/10.1109/43.822628

Choi JY, Swaminathan M (2011) Decoupling capacitor placement in power delivery networks using MFEM. IEEE Trans Compon Packag Manuf Technol 1(10):1651–1661. https://doi.org/10.1109/TCPMT.2011.2165954

Chuang YC, Chen CT, Hwang C (2016) A simple and efficient real-coded genetic algorithm for constrained optimization. Appl Soft Comput 38:87–105. https://doi.org/10.1016/j.asoc.2015.09.036

Coelho LdS (2010) Gaussian quantum-behaved particle swarm optimization approaches for constrained engineering design problems. Expert Syst Appl 37(2):1676–1683. https://doi.org/10.1016/j.eswa.2009.06.044

Das S, Suganthan PN (2011) Differential evolution: a survey of the state-of-the-art. IEEE Trans Evol Comput 15(1):4–31

Deb K, Deb D (2014) Analysing mutation schemes for real-parameter genetic algorithms. Int J Artif Intell Soft Comput 4:1–28. https://doi.org/10.1504/IJAISC.2014.059280

Deb K, Kumar A (1995) Real-coded genetic algorithms with simulated binary crossover: studies on multimodal and multiobjective problems. Complex Syst 9

del Valle Y, Venayagamoorthy GK, Mohagheghi S, Hernandez J, Harley RG (2008) Particle swarm optimization: basic concepts, variants and applications in power systems. IEEE Trans Evol Comput 12(2):171–195. https://doi.org/10.1109/TEVC.2007.896686

dos Santos Coelho L (2008) A quantum particle swarm optimizer with chaotic mutation operator. Chaos Solit Fractals 37(5):1409–1418. https://doi.org/10.1016/j.chaos.2006.10.028

Erdin I, Achar R (2019) Multi-objective optimization of decoupling capacitors for placement and component value. IEEE Trans Compon Packag Manuf Technol 9(10):1976–1983. https://doi.org/10.1109/TCPMT.2019.2930565

Feng Y, Teng G, Wang A, Yao Y (2007) Chaotic inertia weight in particle swarm optimization. In: Second international conference on innovative computing, information and control (ICICIC 2007), pp 475–475. https://doi.org/10.1109/ICICIC.2007.209

Garbaya A, Kotti M, Drira N, Fakhfakh M, Tlelo-Cuautle E, Siarry P (2018) An RBF-PSO technique for the rapid optimization of (CMOS) analog circuits. In: 2018 7th international confer-

ence on modern circuits and systems technologies (MOCAST), pp 1–4. https://doi.org/10.1109/MOCAST.2018.8376657
Haihua S, Sapatnekar SS, Nassif SR (2003) Optimal decoupling capacitor sizing and placement for standard-cell layout designs. IEEE Trans Comput Aided Des Integr Circuits Syst 22(4):428–436. https://doi.org/10.1109/TCAD.2003.809658
Hajimiri A, Lee TH (1998) Phase noise in CMOS differential LC oscillators. In: 1998 symposium on VLSI circuits. Digest of technical papers (Cat. No. 98CH36215), pp 48–51. https://doi.org/10.1109/VLSIC.1998.687999
Hocevar DE, Cox PF, Yang P (1988) Parametric yield optimization for MOS circuit blocks. IEEE Trans Comput Aided Des Integr Circuits Syst 7(6):645–658. https://doi.org/10.1109/43.3204
Hoenigschmid H, Miura-Manausch M, Prigge O, Rahm A, Savignac D (1997) Optimization of advanced MOS technologies for narrow distribution of circuit performance. IEEE Trans Comput Aided Des Integr Circuits Syst 16(2):199–204. https://doi.org/10.1109/43.573834
Hong YY, Beltran AA, Paglinawan AC (2016) A Chaos-enhanced particle swarm optimization with adaptive parameters and its application in maximum power point tracking. Math Probl Eng 2016:6519678
Hussein AS, Jarndal AH (2018) Reliable hybrid small-signal modeling of GaN HEMTs based on particle-swarm-optimization. IEEE Trans Comput Aided Des Integr Circuits Syst 37(9):1816–1824. https://doi.org/10.1109/TCAD.2017.2782779
Park J, Choi K, Allstot DJ (2004) Parasitic-aware RF circuit design and optimization. IEEE Trans Circuits Syst I: Regular Papers 51(10):1953–1966. https://doi.org/10.1109/TCSI.2004.835691
Kahng S (2006) GA-optimized decoupling capacitors damping the rectangular power-bus' cavity-mode resonances. IEEE Microw Wirel Compon Lett 16(6):375–377
Kamisetty S, Garg J, Tripathi JN, Mukherjee J (2011) Optimization of analog RF circuit parameters using randomness in particle swarm optimization. In: 2011 world congress on information and communication technologies, pp 274–278. https://doi.org/10.1109/WICT.2011.6141257
Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of ICNN'95: International conference on neural networks, vol 4, pp 1942–1948
Khan BK, Khalifa YA (2011) An evolutionary method for analog circuits optimization utilizing Mosfet-C filters. In: 2011 18th IEEE international conference on electronics, circuits, and systems, pp 362–365. https://doi.org/10.1109/ICECS.2011.6122288
Kim K, Kang S, Sim J, Park H, Kim B (2018) A search algorithm for the worst operation scenario of a cross-point phase-change memory utilizing particle swarm optimization. IEEE Trans Very Large Scale Integr (VLSI) Syst 26(12):2591–2598. https://doi.org/10.1109/TVLSI.2018.2855959
Koukab A (2011) Reactive power imbalances in LC VCOS and their influence on phase-noise mechanisms. IEEE Trans Microw Theory Tech 59(12):3118–3128. https://doi.org/10.1109/TMTT.2011.2169496
Kulali E, Wasserman E, Zheng J (2007) Chip power model—a new methodology for system power integrity analysis and design. In: 2007 IEEE electrical performance of electronic packaging, pp 259–262. https://doi.org/10.1109/EPEP.2007.4387176
Lee TH, Hajimiri A (2000) Oscillator phase noise: a tutorial. IEEE J Solid-State Circuits 35(3):326–336. https://doi.org/10.1109/4.826814
Lyu W, Xue P, Yang F, Yan C, Hong Z, Zeng X, Zhou D (2018) An efficient Bayesian optimization approach for automated optimization of analog circuits. IEEE Trans Circuits Syst I: Regular Papers 65(6):1954–1967. https://doi.org/10.1109/TCSI.2017.2768826
Man KF, Tang KS, Kwong S (1996) Genetic algorithms: concepts and applications [in engineering design]. IEEE Trans Ind Electron 43(5):519–534
Mikki SM, Kishk AA (2006) Quantum particle swarm optimization for electromagnetics. IEEE Trans Antennas Propag 54(10):2764–2775. https://doi.org/10.1109/TAP.2006.882165
Nickabadi A, Ebadzadeh MM, Safabakhsh R (2011) A novel particle swarm optimization algorithm with adaptive inertia weight. Appl Soft Comput 11(4):3658–3670. https://doi.org/10.1016/j.asoc.2011.01.037

Omkar S, Khandelwal R, Ananth T, Narayana Naik G, Gopalakrishnan S (2009) Quantum behaved particle swarm optimization (QPSO) for multi-objective design optimization of composite structures. Expert Syst Appl 36(8):11312–11322. https://doi.org/10.1016/j.eswa.2009.03.006

Jinho P, Allstot DJ (2004) RF circuit synthesis using particle swarm optimization. In: 2004 IEEE international symposium on circuits and systems (IEEE Cat. No. 04CH37512), vol 5, pp V–V. https://doi.org/10.1109/ISCAS.2004.1329466

Piersanti S, de Paulis F, Olivieri C, Orlandi A (2018) Decoupling capacitors placement for a multichip PDN by a nature-inspired algorithm. IEEE Trans Electromag Compat 60(6):1678–1685. https://doi.org/10.1109/TEMC.2017.2770089

Rutenbar RA, Gielen GGE, Roychowdhury J (2007) Hierarchical modeling, optimization, and synthesis for system-level analog and RF designs. Proc IEEE 95(3):640–669. https://doi.org/10.1109/JPROC.2006.889371

Sabat SL, Kumar KS, Udgata SK (2009) Differential evolution and swarm intelligence techniques for analog circuit synthesis. In: 2009 world congress on nature biologically inspired computing (NaBIC), pp 469–474. https://doi.org/10.1109/NABIC.2009.5393356

Shi Y, Eberhart R (1998) A modified particle swarm optimizer. In: 1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360), pp 69–73

Shi Y, Eberhart RC (1998) Parameter selection in particle swarm optimization. In: Porto VW, Saravanan N, Waagen D, Eiben AE (eds) Evolutionary programming VII. Springer, Berlin, Heidelberg, pp 591–600

Shi Y, Eberhart RC (1999) Empirical study of particle swarm optimization. In: Proceedings of the 1999 congress on evolutionary computation-CEC99 (Cat. No. 99TH8406), vol 3, pp 1945–1950. https://doi.org/10.1109/CEC.1999.785511

Storn R, Price K (1997) Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11:341–359. https://doi.org/10.1023/A:1008202821328

Sun J, Fang W, Palade V, Wu X, Xu W (2011) Quantum-behaved particle swarm optimization with gaussian distributed local attractor point. Appl Math Comput 218(7):3763–3775. https://doi.org/10.1016/j.amc.2011.09.021

Swaminathan M, Engin E (2007) Power integrity modeling and design for semiconductors and systems. Prentice Hall Modern semiconductor design series, Pearson Education

Tripathi JN, Mukherjee J, Apte PR (2011) Optimizing phase noise of 2 GHz RF oscillator with minimum frequency deviation. In: 2011 IEEE 54th international midwest symposium on circuits and systems (MWSCAS), pp 1–4. https://doi.org/10.1109/MWSCAS.2011.6026275

Tripathi JN, Mukherjee J, Apte PR, Chhabra NK, Nagpal RK, Malik R (2013) Selection and placement of decoupling capacitors in high speed systems. IEEE Electromag Compat Mag 2(4):72–78

Tripathi JN, Mahajan A, Mukherjee J, Nagpal RK, Malik R, Gupta N (2014) Decoupling network optimization in high speed systems by mixed-integer programming. In: 2014 IEEE international symposium on circuits and systems (ISCAS), pp 1010–1013. https://doi.org/10.1109/ISCAS.2014.6865309

Tripathi JN, Sharma VK, Shrimali H (2019) A review on power supply induced jitter. IEEE Trans Compon Packag Manuf Technol 9(3):511–524. https://doi.org/10.1109/TCPMT.2018.2872608

Xin J, Chen G, Hai Y (2009) A particle swarm optimizer with multi-stage linearly-decreasing inertia weight. In: 2009 international joint conference on computational sciences and optimization, vol 1, pp 505–508

Xu Z, Wang J, Fan J (2020) Decoupling capacitor placement optimization with Lagrange multiplier method. In: 2020 IEEE international symposium on electromagnetic compatibility signal/power integrity (EMCSI), pp 22–25. https://doi.org/10.1109/EMCSI38923.2020.9191625

Chapter 9
PAM3: History, Algorithm, and Performance Comparison to NRZ and PAM4

Sherman Shan Chen, Francesco de Paulis, David R. Stauffer, and Brian Holden

Abstract The origin and evolution of Three-level Pulse Amplitude Modulation (PAM3) are introduced. Two PAM3 variant encoding schemes (3-bit-to-2-trit (3B2T) and 11-bit-to-7-trit (11B7T)) are described, along with the trit test sequence for PAM3. The high-speed transmission performance of PAM3 is investigated under the influences of five major channel impairments, with NRZ and PAM4 as comparisons. The work reveals that PAM3 has a certain performance advantage over NRZ and PAM4 in terms of reflection coefficient, insertion loss, and mode conversion, while being relatively more sensitive to crosstalk and noise.

Keywords PAM3 · NRZ · PAM4 · High-speed transmission · Channel impairment · Noise

Introduction

Three-level Pulse Amplitude Modulation (PAM3) is a high-speed modulation technique that converts the binary bit stream into three-level symbols so as to reduce the bandwidth requirement on the frequency spectrum, at the cost of a lowered eye height (EH) and associated complexity. The use of a three-level signal to transmit

S. S. Chen (✉)
Northampton, UK
e-mail: shan@kandou.com

F. de Paulis
University of L'Aquila, L'Aquila, Italy

D. R. Stauffer
Essex Junction, VT, USA

B. Holden
Los Gatos, CA, USA

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_9

information emerged as early as 1961 when a three-level code, Alternate Mark Inversion (AMI), was used in the first T-Carrier system (Wikipedia 2022). This code sent a zero input as a 0 line code and a one input as either + or − line code, alternating between the two values when a one was the input. One's density was ensured by the constraint that the 00 code was not used for voice samples. AMI was a Return to Zero (RZ) code to help along the low-performance clock recovery circuits of the day.

In 1963, a three-level encoding scheme called Duobinary was proposed with the purpose of reducing the bandwidth requirement (Lender 1963). The encoding algorithm of Duobinary is essentially determining the output level −1, 0, or 1, based on whether there is an odd/even number of 0s between the two preceding non-zero bits. Therefore, it does not follow a fixed n-binary-to-m-ternary mapping relation. Interested readers can refer to Lender (1963) for more details. While reducing the Nyquist frequency to only a half of NRZ, Duobinary has some drawbacks including low eye height, high power, and others (David 2022). In 1995, Fast Ethernet 802.3u (100Base-TX) was released which brought another three-level modulation technology called Multi-Level Transition (MLT-3) to the public (IEEE Computer Society 2018). The encoding process of MLT-3 is to cycle sequentially through the voltage levels [−1, 0, +1] in accordance with the following rules:

- If the next bit is 1, move to the next state;
- If the next bit is 0, stay in the same state.

Such an encoding algorithm reduces the number of transitions and consequentially lowers the power consumption and electromagnetic interference (EMI). MLT-3 however does not reduce the Nyquist frequency of the signal which still remains at 31.25 MHz (1/4 of the aggregate bit rate) for each pair in a four-pair channel. Also, it is not an n-binary-to-m-ternary encoding scheme either.

The earliest adoption of PAM3 dates back to also 1995 when 100BASE-T4 (100 Mbit/s over four-pair Cat3 or better cable) was released. 100BATE-T4 uses the 8B6T code which converts 8 binary bits into 6 base-3 digits, i.e., 3-level symbols. After this first time standard-adoption, PAM3 was not used by any prevalent standards until 2011. Due to the increasing popularity of electric vehicles, the demand for upgrading the traditional Controller Area Network (CAN) bus or Local Interconnect Network (LIN) to Automotive Ethernet led to the revival of the technique of PAM3. Originated from the BroadR-Reach (BR) standard (published by the One-Pair Ethernet (OPEN) Alliance Special Interest Group (SIG) in 2011 (Broadcom Corporation 2014)), both 100BASE-T1 (100 Mbit/s Ethernet over a single twisted pair for automotive applications, released in March 2016 (IEEE Computer Society 2016a)), and 1000BASE-T1 (1000 Mbit/s Ethernet over a single twisted pair for automotive applications, released in September 2016 (IEEE Computer Society 2016b)) adopted PAM3 as the line code. The reason for such a decision could be that PAM3 is relatively better balanced between the various requirements including the full-duplex operation over a single unshielded pair (UTP), the bandwidth requirement, and the EMI which is extremely stringent in automotive applications (Thain et al. 2016).

The Universal Serial Bus 4 (USB4™) Specification Version 2.0 (shorthanded as "USB4v2.0") was released in October 2022 (USB-IF 2022), doubling the speed of USB4v1 (20 Gbps/lane) to 40 Gbps/lane. In addition, the line code was changed from non-return-to-zero (NRZ) to PAM3, resulting in a baud rate of 25.6 Gbd/lane.

In December 2022, Samsung announced its plan for the 36 Gbps-rated next-generation graphics double data rate 7 (GDDR7) memory (VideoCardz.com 2022). Despite the GDDR6X (18 Gbps to 23 Gbps), which was co-developed by NVIDIA and Micron Technology, adopting four-level pulse amplitude modulation (PAM4), Samsung chose PAM3 as the line code for GDDR7 (TechPowerUp 2022). It should be noted that neither GDDR6X nor GDDR7 has made it into the JEDEC specification yet. This may indicate that, at least for the DDR applications, which modulation technique, PAM3 or PAM4, will triumph remains yet to be determined. Even in the applications where differential pair dominates, although there seems to be a trend that most of the latest high-speed standards turned to PAM4 as the line code including 200/400 Gbps Ethernet (IEEE Computer Society 2018), and Optical Internetworking Forum Common Electrical I/O (OIF CEI) middle-reach (MR) and long-reach (LR) 56/112 Gbps (OIF 2022), the latest USB specification, i.e., USB4v2.0, however, selected PAM3 as its encoding scheme for PAM3's better performance in terms of handling the USB Type-C connector, as well as other comparative advantages over PAM4. The variations in selecting PAM3 or PAM4 existing among the latest high-speed standards reflect that PAM3 possesses some distinct features that bolster its position in the competition with PAM4 for the high-speed modulation technique, at least in particular application scenarios.

Unlike NRZ and PAM4, which have been extensively studied, there are significantly fewer research works carried out on PAM3, let alone making an explicit comparison between PAM3 and other modulation techniques. The work in Holden (2013) compared the bandwidth requirements for NRZ, PAM3, PAM4, and other line codes in order to achieve a 400 Gbps backplane high-speed transmission system. The performances of ensemble NRZ (ENRZ), NRZ, PAM3, and PAM4 in terms of the sensitivity to crosstalk are evaluated by Chen et al. (2021). The tolerance of ENRZ, NRZ, PAM3, and PAM4 to the channel skew is examined in Chen and Xu (2021). In this chapter, we investigate the high-speed transmission performance of PAM3, with NRZ and PAM4 as the comparison. The remainder of the chapter is structured as follows: Section "The Encoding Schemes and Test Sequences of PAM3" describes the two encoding schemes of PAM3 (3-bit-to-2-trit (3B2T) and 11-bit-to-7-trit (11B7T)), along with the two test sequences for PAM3 (pseudo-random binary sequence of n-th order (PRBSn), pseudo-random ternary sequence of n-th order (PRTSn)). Section "Performance Analysis: In Comparison to NRZ and PAM4" examines the performances of the four line codes (including two PAM3 encoding schemes) under the influences of various channel impairments. Section "Summary" summarizes the chapter.

Table 9.1 USB4v2.0 3B2T (USB-IF 2022)

Binary	Ternary
000	−1, −1
001	−1, +1
010	−1, +1
011	+1, −1
100	0, +1
101	+1, −1
110	+1, 0
111	+1, +1

Table 9.2 100BASE-T1 3B2T (IEEE Computer Society 2016a)

Binary	Ternary
000	−1, −1
001	−1, 0
001	−1, 1
011	0, −1
100	0, +1
101	+1, −1
110	+1, 0
111	+1, +1

The Encoding Schemes and Test Sequences of PAM3

3B2T Versus 11B7T

When migrating from a binary system to a ternary system, corresponding to the "bit" used to represent one symbol in binary systems, the term "trit" is used to refer to one ternary symbol. Each trit can be at one of three voltage levels in [−1 0 1], or [0 1 2] when dealing with a single-ended system. The basic encoding format for PAM3 is 3B2T which maps 3 bits to 2 trits and achieves a transmission efficiency of 1.5 bits/symbol. In terms of the specific mapping codebook, variants exist. For example, the 3B2T mapping scheme used in USB4v2.0, 100BASE-T1, and 1000BASE-T1 are all different from each other, as presented by Tables 9.1, 9.2, and 9.3.

It can be seen that in both 3B2T mapping schemes, the ternary combination "00" is not employed. There are two reasons for doing so:

1. DC balance: Without the (0, 0) code, the DC balance is maintained. Adding the (0, 0) code will however inevitably breach the DC balance.

Table 9.3 1000BASE-T1 3B2T (IEEE Computer Society 2016b)

Binary	Ternary
000	−1, −1
001	0, −1
010	−1, 0
011	−1, +1
100	0, +1
101	+1, −1
110	+1, +1
111	0, +1

2. According to 802.3 section "96.3.3.3.5 Generation of ternary pair (TAn, TBn)" (IEEE Computer Society 2018), the ternary symbol pair (0, 0) is used in certain special, non-data codes.

Based on our experiment, replacing either (−1, −1) and (1, 1) by (0, 0) indeed improves the resulting eye diagrams—due to the less inter-symbol interference (ISI) as a result of the replacement. But because of the above two reasons, the 802.3 committees (and USB-IF as well) decided not to use the (0, 0) code for data transferring and instead use it for the control signals.

In USB4v2.0, based on the fundamental 3B2T codebook, additional processing is conducted to formulate a new encoding scheme called 11B7T, i.e., mapping 11 bits into 7 trits. By doing so, the transmission efficiency is increased to 1.57 bits/symbol (USB-IF 2022). In addition, the 11B7T code was reported to have a better noise immunity performance than 3B2T (and other schemes such as 10B7T) (Liu and Marvel Corp. 2014). Interested readers can consult the USB4v2.0 specification for the implementation details of the 11B7T encoding (USB-IF 2022).

PRBS Versus PRTS

For binary systems, PRBSn is widely used as the stimulus source, along with other test patterns each carrying specific characterization purpose (OIF 2022). When the line code is changed to PAM3, the binary sequence accordingly needs to be replaced by a ternary sequence. This can be done in two ways. The first is utilizing a PRTSn sequence generator which directly outputs the PRTS trit stream (USB-IF 2022). Equation (9.1) gives the generation equation of PRTS19

$$Trit_{PRTS19} = 1 + X^2 + X^{19} \tag{9.1}$$

As can be seen, the structure of a PRTS generator is very similar to that of a PRBS generator, with the following differences:

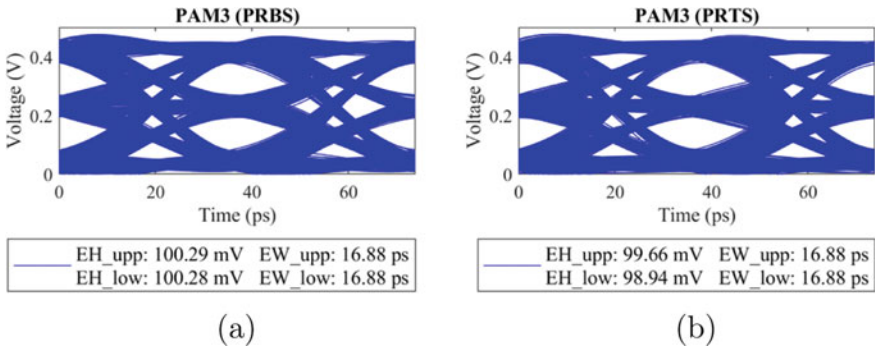


Fig. 9.1 Eye diagram: PRBS11 versus PRTS11 (Channel loss: 8 dB@10 GHz. Only linear equalizations are applied)

1. Both addition and multiplication are modulo-3 operations.
2. All signals are of three levels, i.e., [−1 0 1].

The second method is to use the conventional PRBSn to generate the binary pseudo-random sequence and then convert it into PAM3 ternary stream using a BIN-to-PAM3 encoder.

Both methods are adopted by USB4v2.0. In USB4v2.0, PRBS11 and PRTS7 are used as the transmitter data source for the victim channel, while PRTS19 is used as the source for the aggressor channel. The reason for such an arrangement might be that PRBS11 has a sequence length of 2047 (2^{11}) and PRTS7 has a similar sequence length of 2187 (3^7). In the mean time, PRTS19 has a much larger length of 1,162,261,467 trits. Figure 9.1 compares the eye diagrams obtained with PRBS11 and PRTS7. We can see that both sequences produce quite similar eye diagrams with PRTS7 negligibly worse than PRBS11. Throughout this study, PRBS11 is used as the data source for all the simulations.

Scrambling and Pre-coding

In order to limit the running length, i.e., the length of consecutive same bits/trits, it is necessary to add scrambling to the bit/trit stream. In this study, we use the scrambler used by 10/100/200GBASE-T for the NRZ bit stream scrambling. Equation (9.2) gives the construction formula of the 58-tap scrambler

$$G(x) = 1 + x^{39} + x^{58} \tag{9.2}$$

It has been proven that by applying a partial response adjustment to the transmit signal, a certain level of known ISI can be added to the signal. At the receiver side, apply the reverse operation and the known ISI can be removed. The benefit of doing

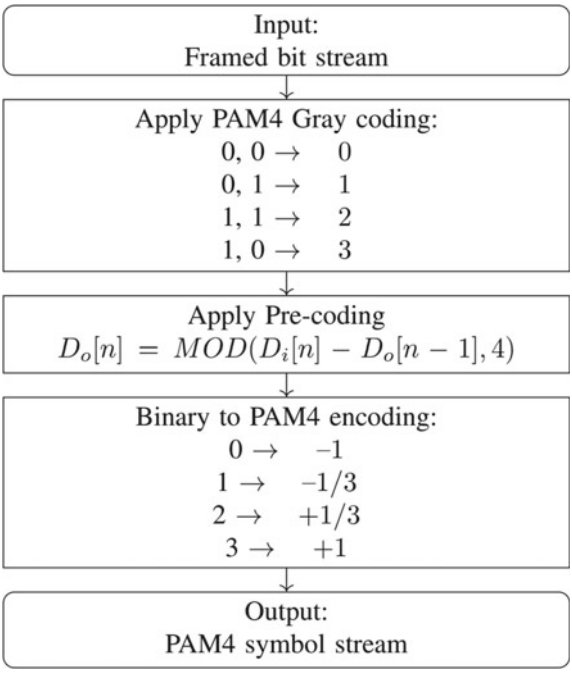


Fig. 9.2 100GBASE-KP4 PMA flow (IEEE Computer Society 2018)

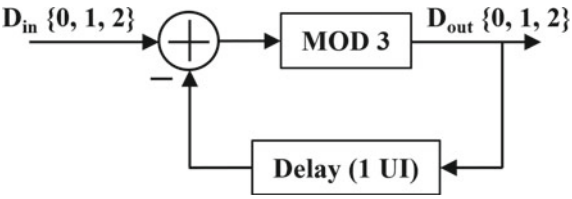


Fig. 9.3 PAM3 pre-coder (USB-IF 2022)

so is that the channel transmission SNR can be improved (Law and Masouros 2018). This technique is called Pre-coding which normally utilizes a Tomlinson-Harashima Pre-coder at the transmitter (TX), and a corresponding decoder at the receiver (RX). The structures of the binary/ternary/quarternary pre-coders remain the same except for the modulo operation being different. Figures 9.3 and 9.4 give the block diagrams of PAM3 and PAM4 pre-coders. The NRZ pre-coder can be obtained by simply changing the modulus to 2.

For PAM4, we follow the PMA transmit process specified in 100GBASE-KP4 as illustrated in Fig. 9.2. Note that in the PAM4 PMA encoding process, scrambling is applied before the Gray coding.

Instead of using the conventional scrambler as adopted by the majority of standards, USB4v2.0 employs a unique structure for its trit scrambler. The USB4v2.0 scrambler first performs a special operation (called F-function operation) on the pre-built PRBS11 and PRTS19 bit/trit streams, the result then is performed modulo 3

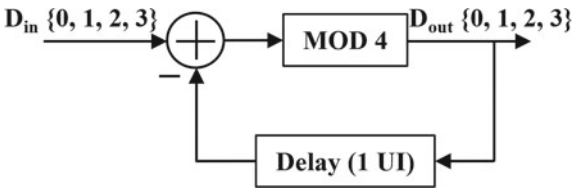


Fig. 9.4 PAM4 pre-coder (IEEE Computer Society 2018)

with the input trit stream to generate the scrambled output trit stream. Interested readers can refer to "Figure 4-40. Scrambler Operation" in USB-IF (2022) for more details.

As a default configuration, throughout this study, the scrambler and pre-coder corresponding to each modulation technique are applied in all simulations.

Performance Analysis: In Comparison to NRZ and PAM4

In this section, we investigate the performances of PAM3, in comparison to NRZ and PAM4. For PAM3, the 3B2T scheme will be examined throughout the study and the 11B7T scheme will be included in the analyses of reflection coefficient (RC, or return loss (RL)), insertion loss (IL), and crosstalk (XT). The performance comparison is conducted by comparing the best eye diagrams obtained using each modulation technique, in the presence of various channel impairments. For a high-speed channel, the major performance-impacting factors are RL, IL, XT, mode conversion (MC), and noise. In the following contents, these impairments will be examined individually.

Reflection Coefficient

The reflection coefficient refers to the ratio of the reflected energy versus the incident energy. It may be necessary to distinguish the reflection coefficient from a widely misused term "return loss". According to the standard definitions (Everything RF 2023; Electronics Notes 2023)

$$\textit{Return Loss (dB)} = -\textit{Reflection Coefficient (dB)} \tag{9.3}$$

Based on this, for a reflective system, its dB[1]-valued reflection coefficient is always less than zero and its dB-valued return loss should always be greater than zero. However, in much literature return loss is often used where the reflection coefficient actually should be used. For example, many articles present negative return loss results which should actually be the reflection coefficient. A possible reason for such

[1] Decibel.

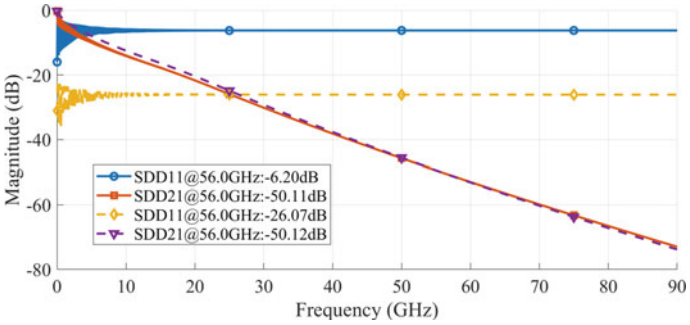


Fig. 9.5 S_{dd11}, S_{dd21} of Ch1, Ch2. Yellow: Ch1 S_{dd11}; Purple: Ch1 S_{dd21}; Blue: Ch2 S_{dd11}; Red: Ch2 S_{dd21}

a widespread misuse could be that return loss is shorter and more distinctive as a term than reflective coefficient which can be confused with the resistor-capacitor network (also shorthanded as RC), and therefore is preferable to use. In this chapter, we use the reflection coefficient as the parameter of the reflection level.

Two channels aiming to run at 112 Gbps, Ch1 and Ch2, are built representing the low/high reflection levels, respectively. Figure 9.5 plots the S_{dd11} and S_{dd21} of the two test channels, with the values at 56 GHz annotated. We can see that Ch1 has a reflective coefficient of −26.07 dB while Ch2 has a level of −6.2 dB, indicating that Ch2 has a significantly higher reflection level than Ch1. Meanwhile, both channels have a nearly identical insertion loss about 50.1 dB@56 GHz. This way, we eliminate the influence of the insertion loss on the simulation results obtained with these two channels since the deviations in the results can only come from the different levels of reflection.

A PRBS11 pseudo-random sequence running at 112 Gbps is used as the transmitter data for the simulation. Both rise/fall times are set to 0.15 UI, as a representation of the typical situation seen in most applications (OIF 2022; IEEE Computer Society 2018). A 3-tap FIR, an IEEE 802.3 CTLE, and a 2-tap DFE are implemented as the equalization devices. As mentioned, the scramblers and the pre-coders suiting each modulation technique are implemented. In addition, a fourth-order Butterworth filter is also implemented in accordance with the requirements from IEEE 802.3/USB4v2.0. Except in the study of the influences of noise, noises are not considered in the simulation. Ideal clocking is assumed which means there will be no impact of jitter. The settings of all three equalizers i.e., FIR, CTLE, and DFE, are swept so as to obtain the optimal eye diagrams. All eye diagrams are plotted in a single-ended manner so as to align with the waveforms seen by the slicers in the RX. In all sections except for the "Noise", the eye height and eye width (EW) are measured over the 40–60% range centered at the peak point. Throughout this chapter, the above settings are used as the baseline configuration for all simulations unless otherwise indicated (such as in the study of crosstalk).

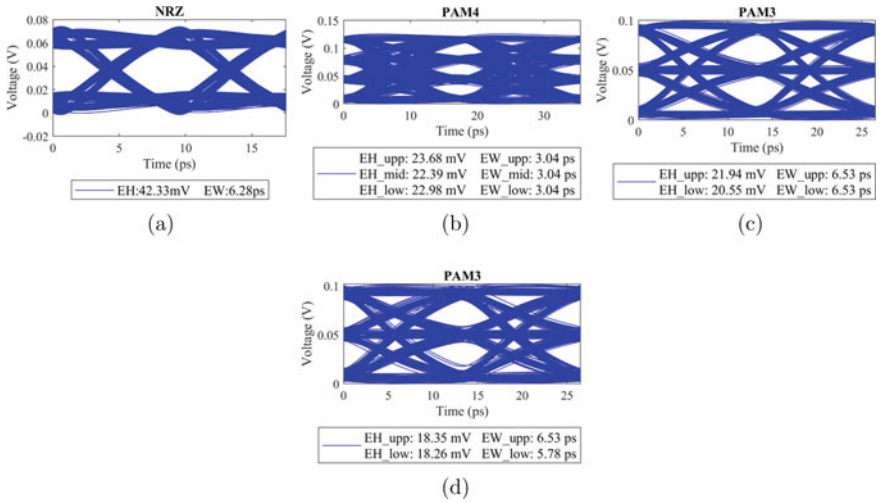


Fig. 9.6 Eye diagrams obtained with the high reflection channel (Ch2). **a** NRZ; **b** PAM4; **c** PAM3 (3B2T); **d** PAM3 (11B7T)

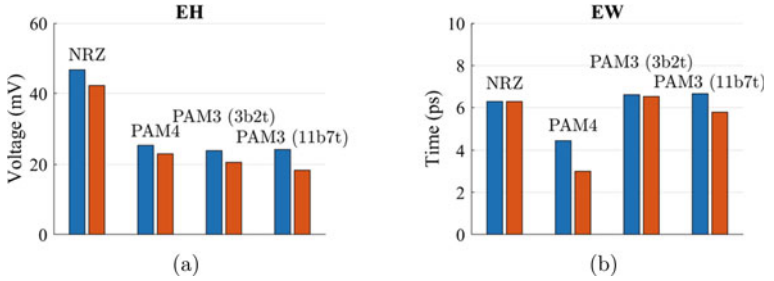


Fig. 9.7 Sensitivity to reflection coefficient **a** EH; **b** EW. Blue: Low RL; Red: High RL

Figure 9.6 presents the eye diagrams obtained with Ch2, i.e., the high reflection channel,[2] and Fig. 9.7 shows the comparison of the EH and the EW obtained with the two channels using the four codes. The following observations can be made:

1. PAM3 (3B2T) and NRZ appear to be relatively more robust against reflection compared to PAM3 (11B7T) and PAM4, evidenced by similar degradation on EH, and no or nearly no degradation on EW as the level of reflection increases.
2. PAM3 (11B7T), despite it producing eye diagrams slightly better than PAM3 (3B2T) with the low reflection channel, degraded more significantly when the reflection level increased.

Overall, as the reflection level increases, PAM3 (3B2T) and NRZ demonstrate a certain level of performance edge against PAM3 (11B7T) and PAM4.

[2] The eye diagrams obtained with Ch1 are omitted due to space limitation.

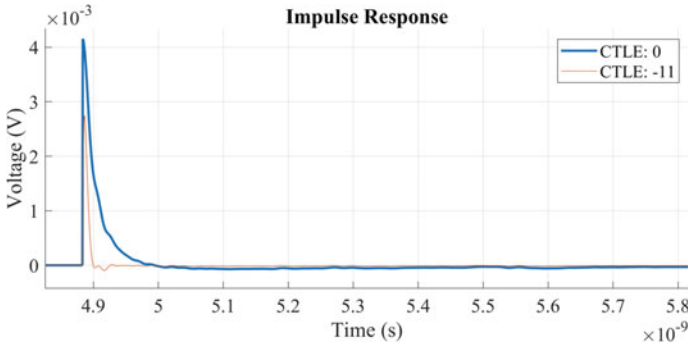


Fig. 9.8 IR versus CTLE. Obtained with the 49 dB@56 GHz channel. Blue: Impulse response with CTLE = 0 dB. Red: Impulse response with CTLE = −11 dB

Loss

Next, we look at the influences of channel loss on the four codes. For an adequately designed high-speed transmission channel, channel loss is often the most critical property of the channel. This is because the insertion loss (IL) usually imposes the largest impairment on the channel performance, demanding the application of sophisticated equalization in order to counter the attenuation and signal distortion caused by the IL. Figure 9.8 shows the impulse response (IR) applied with the continuous time linear equalization (CTLE) of 0 dB and −11 dB, respectively, with the two values of the DC gain providing no equalization and a quite high level of attenuation of the low-frequency transfer function to balance the higher frequency channel loss. Figure 9.9 shows the channel output waveforms of an 112 Gbps PAM3 data stream passing through a 49 dB@56 GHz channel, under the same values of CTLE. We have the following observations:

1. With no CLTE applied (i.e., 0 dB), the IR presents a fairly slow-falling edge, denoting a large ISI. This is evidenced by the channel output waveform which shows significant ISI and is unlikely to produce an open eye diagram.
2. With adequate CLTE applied (−11 dB in this case), the ISI is significantly reduced. Despite this improvement coming at the cost of a lowered signal amplitude, the equalized channel output shows a smaller but less ISI-disturbed waveform which is able to produce an open eye diagram.[3]

To evaluate the performance of the line codes under the influences of varying channel losses, two channels, Ch3 and Ch4, are built as the test vehicles for this study. Ch3 is a 26 dB@56 GHz channel sourced from 802.3df task force (IEEE P802.3df 2022) and Ch4 is created by cascading two Ch3 channels so as to obtain a higher insertion loss. Figure 9.10 plots the S_{dd11} and S_{dd21} of the two test channels, with the values at 56 GHz annotated. Except that the channels are different, other simulation settings

[3] Due to space limitation, the eye diagrams are omitted.

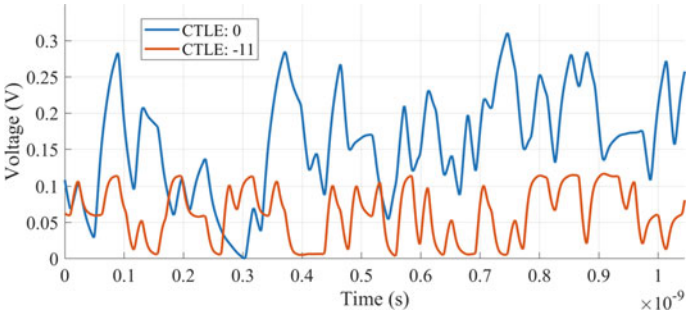


Fig. 9.9 PAM3 ISI versus CTLE. Obtained with the 49 dB@56 GHz channel. Blue: Channel output with CTLE = 0 dB. Red: Channel output with CTLE = −11 dB

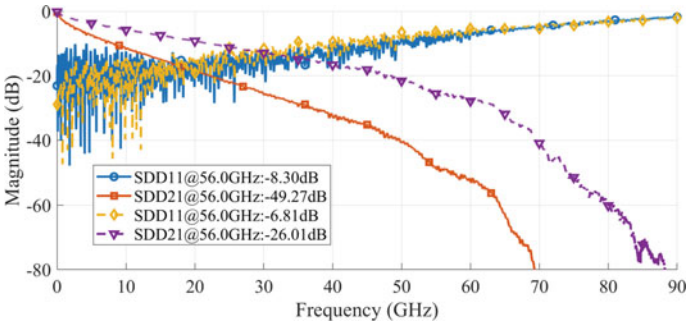


Fig. 9.10 S_{dd11}, S_{dd21} of Ch3, Ch4. Yellow: Ch3 S_{dd11} (−6.81 dB@56 GHz); Purple: Ch3 S_{dd21} (−26.01 dB@56 GHz). Blue: Ch4 S_{dd11} (−8.3 dB@56 GHz); Red: Ch4 S_{dd21} (−49.27 dB@56 GHz)

follow the baseline simulation configurations (with equalizers swept to achieve the optimal settings). Figures 9.11 and 9.12 present the eye diagrams obtained with Ch3 and Ch4 using the four codes. Figure 9.13 depicts the comparison of the EH and the EW obtained with the two channels. We have the following observations:

1. As insertion loss goes up, NRZ and PAM4 have worse EH degradation than PAM3 (3B2T).
2. Less EW degradation is seen with NRZ and PAM3 (3B2T) than PAM4.
3. PAM3 (3B2T) demonstrates a better performance than PAM3 (11B7T), especially with the high loss channel.

In summary, both PAM3 (3B2T) and PAM3 (11B2T)) are less sensitive to channel loss than NRZ and PAM4, whereas PAM3 (3B2T) and PAM3 (11B7T) have similar performances.

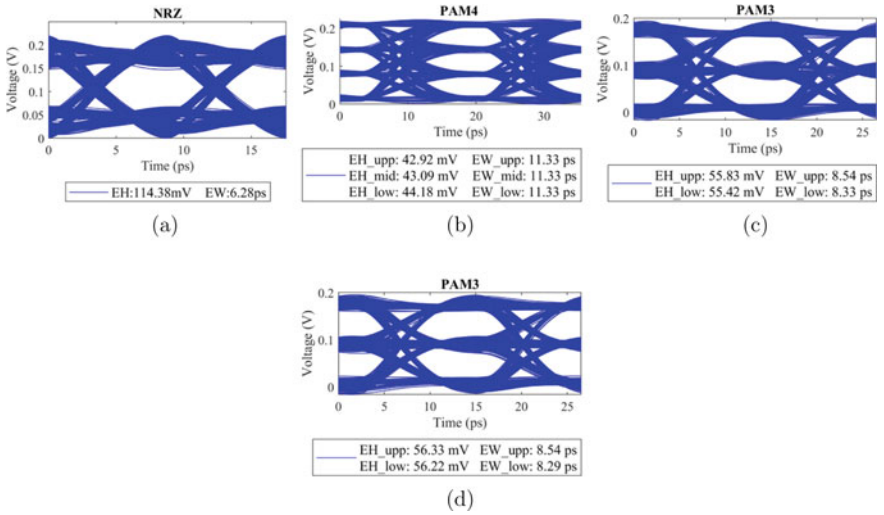


Fig. 9.11 Eye diagrams obtained with Ch3 (26 dB@56 GHz). **a** NRZ; **b** PAM4; **c** PAM3 (3B2T); **d** PAM3 (11B7T)

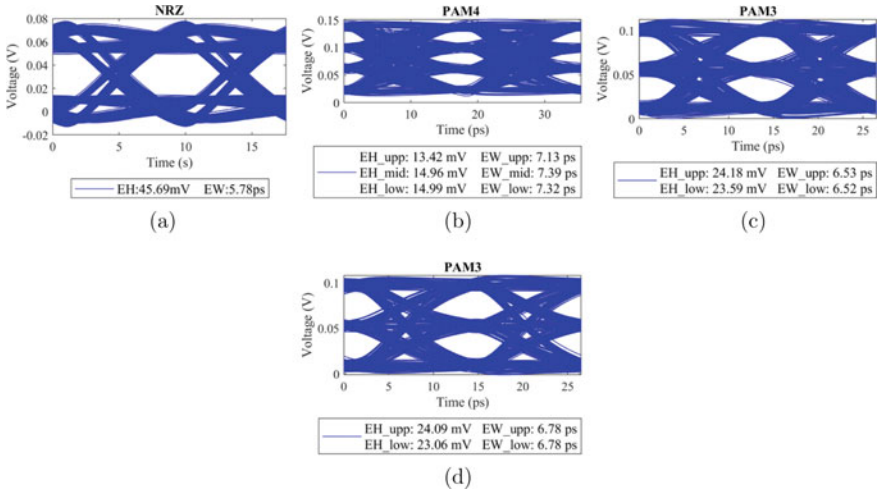


Fig. 9.12 Eye diagrams obtained with Ch4 (49 dB@56 GHz). **a** NRZ; **b** PAM4; **c** PAM3 (3B2T); **d** PAM3 (11B7T)

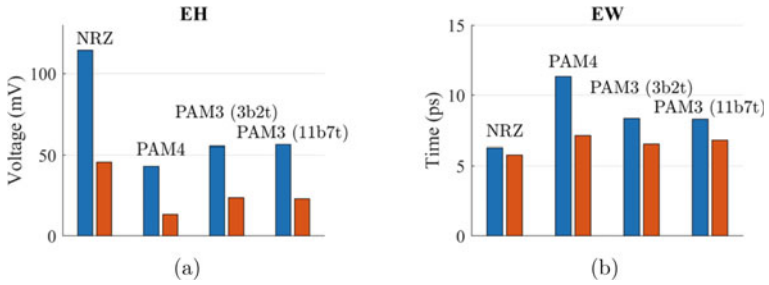


Fig. 9.13 Sensitivity to channel loss **a** EH; **b** EW. Blue: Low loss (Ch3); Red: High loss (Ch4)

Crosstalk

The subsequent analysis involves the investigation of the performances of the four codes under the influence of crosstalk. Two 12-port channels, Ch5 and Ch6, are created as the test vehicles with the center differential pair being the victim and the pairs on each side as the aggressors. Figure 9.14 shows the differential reflection coefficient, IL, near-end crosstalk (NEXT[4]), and far-end crosstalk (FEXT[5]). As can be seen, both channels have an insertion loss of 28 dB at 14 GHz, with Channel 1 showing a normal level of NEXT/FEXT (−39 dB/−35 dB), in contrast to the significantly high level (−26 dB/−27 dB) of Channel 2. The simulation is performed with the following configurations:

1. PRBS11 with different seeds are used as the source data for the victim, aggressor 1, and aggressor 2. Equations (9.4) to (9.6) give the formulas used to produce the three PRBS11 source streams:

$$bits = mxlfsr([11101110000], [1, 9, 11], 2^{11} - 1) \tag{9.4}$$

$$bits_{agg1} = mxlfsr([110011], [1, 9, 11], 2^{11} - 1) \tag{9.5}$$

$$bits_{agg2} = mxlfsr([110001], [1, 9, 11], 2^{11} - 1) \tag{9.6}$$

2. The data rate is 28 Gbps.
3. Other settings are the same as the baseline simulation configuration.

The eye diagrams obtained with the four codes are shown in Figs. 9.15 and 9.16. The EH and EW results are summarized in Fig. 9.17. The following observations are made:

1. PAM3 (3B2T) and PAM3 (11B7T) demonstrate very similar performances under the influences of crosstalk.
2. The order of sensitivity (EH and EW jointly) to crosstalk from high to low is PAM3 (3B2T)/PAM3 (11B7T) → PAM4 → NRZ.

[4] S_{dd31}.
[5] S_{dd41}.

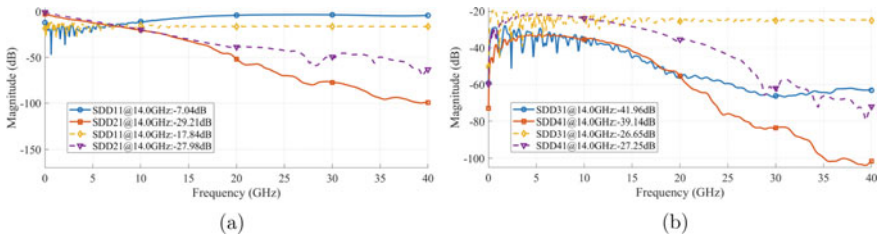


Fig. 9.14 S-parameters of Ch5 (Normal crosstalk level), Ch6 (High crosstalk level). **a** Blue: Ch5 S_{dd11}; Red: Ch5 S_{dd21}; Yellow: Ch6 S_{dd11}; Purple: Ch6 S_{dd21}; **b** Blue: Ch5 S_{dd31}; Red: Ch5 S_{dd41}; Yellow: Ch6 S_{dd31}; Purple: Ch6 S_{dd41}

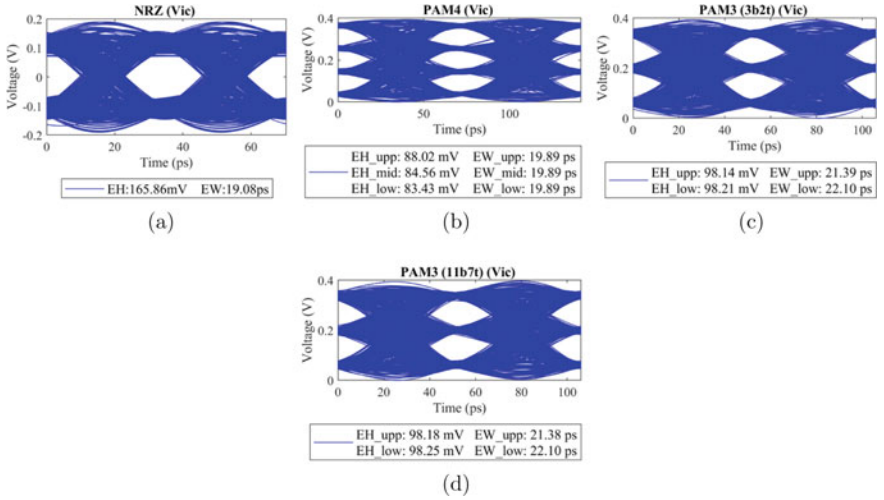


Fig. 9.15 Eye diagrams obtained with Ch 5 (Normal crosstalk level)

In summary, NRZ demonstrates the strongest robustness in the presence of crosstalk among the four codes, and PAM4 is slightly superior to both variants of PAM3.

Mode Conversion

Mode conversion refers to the conversion between the common mode energy and the differential mode energy in a signal. When mode conversion exists, the differential part of the signal, which is utilized to transmit information in differential signaling techniques, will be compromised. This will not only reduce the effective amplitude of the differential signal but also can increase the risk of potential EMI problems due

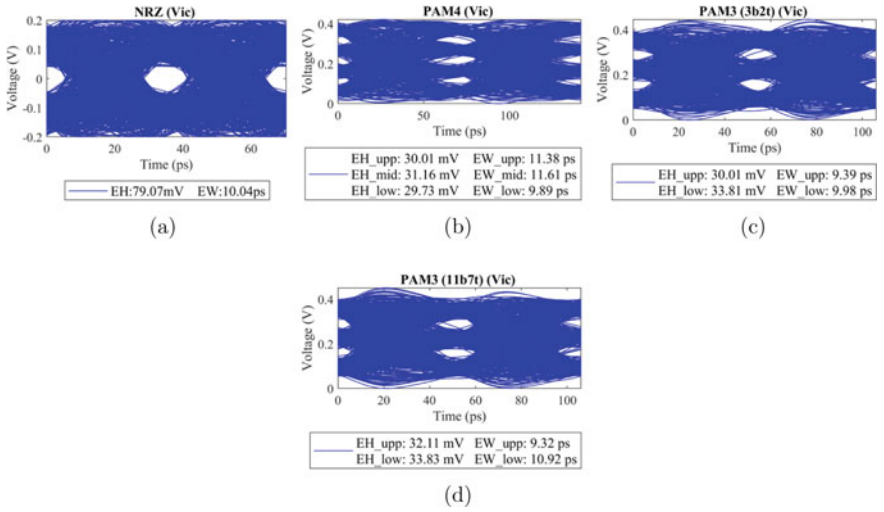


Fig. 9.16 Eye diagrams obtained with Ch 6 (High crosstalk level)

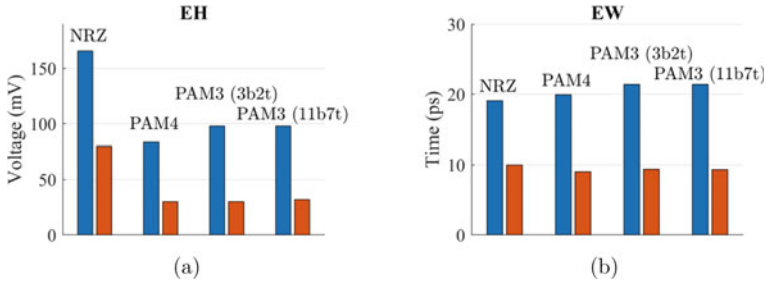


Fig. 9.17 Sensitivity to crosstalk **a** EH; **b** EW. Blue: Normal crosstalk (Ch5); Red: High crosstalk (Ch6)

to the energy leakage from the differential to the common mode. Popular reasons causing mode conversion are

1. Intra-pair skew: This is the unequal propagation time between the P and N leads within one differential pair.
2. Asymmetric design: Any asymmetries on the structure of a component in the channel, such as connector, differential via, and routing, as well as the property of the materials used by the channel, can lead to a certain level of mode conversion.

Since both the above mechanisms have similar impacts on the resulting level of mode conversion, we use the first mechanism to create the mode conversion for this study. Based on Ch4, by adding 10 and 20 ps intra-pair skews, we create two channels, Ch7 and Ch8, with two levels of mode conversion. Figure 9.18 shows the

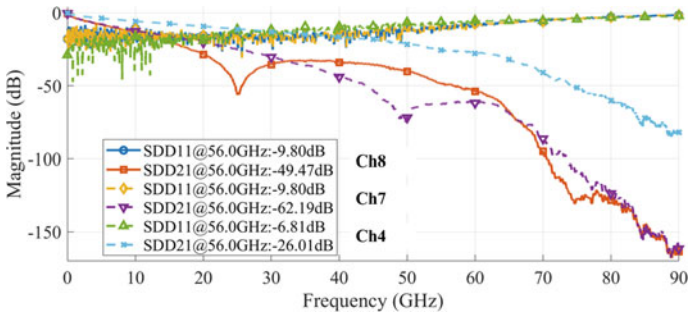


Fig. 9.18 S_{dd11}, S_{dd21} of Ch4 (no skew), Ch7 (10 ps skew), and Ch8 (20 ps skew)

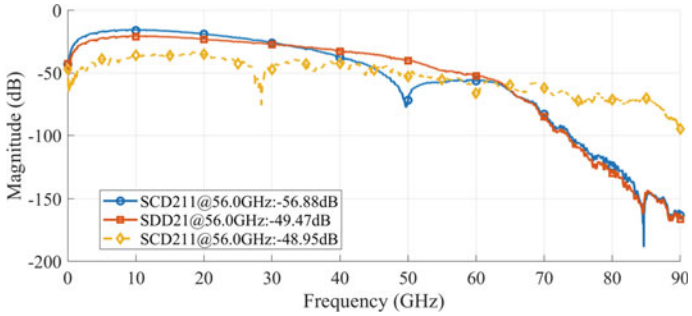


Fig. 9.19 S_{cd21} of Ch4 (no skew), Ch7 (10 ps skew), and Ch8 (20 ps skew). Blue: 20 ps skew; Red: 10 ps skew; Yellow: No skew

S_{dd11} and S_{dd21} of the two channels, along with Ch4 as the baseline. Figure 9.19 plots out the S_{cd21} of the three channels.

In previous studies conducted on reflection, loss, and crosstalk, we have sufficiently examined and demonstrated the performance of the 11B7T PAM3. Confined by the space limitation of the chapter, in the subsequent studies, we will only use 3B2T as the representative of PAM3. By running the baseline simulation on Ch7 and Ch8, we obtain the eye diagrams as plotted in Fig. 9.20 and the bar graph comparison is plotted in Fig. 9.21. The following observations are made:

1. PAM3 (3B2T) demonstrates less degradation than NRZ and PAM4, on both EH and EW.
2. With 20 ps skew, which is 2.24 times the unit interval of NRZ 112 Gbps (8.929 ps), PAM3 produces the highest EH and EW among the three line codes.

Overall, PAM3 is less sensitive to the intra-pair skew (and the consequently produced mode conversion) compared to NRZ and PAM4.

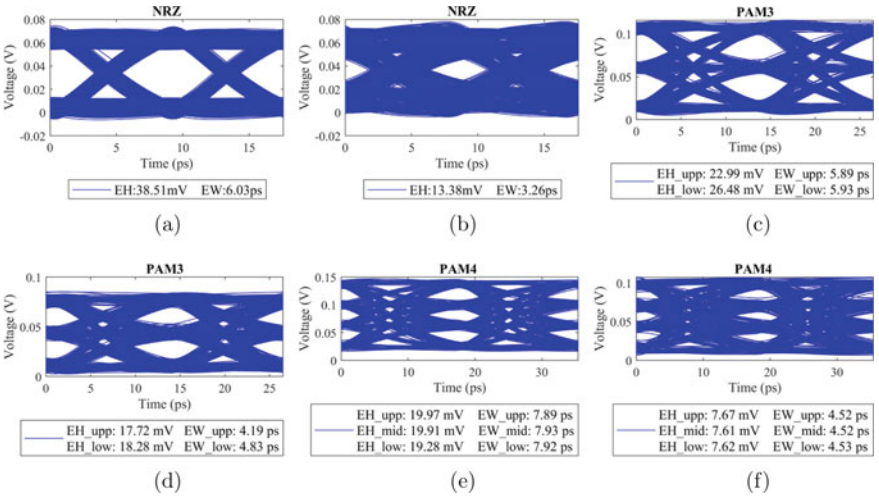


Fig. 9.20 Eye diagrams obtained with Ch7 (10 ps skew) and Ch8 (20 ps skew). **a**, **c**, **e**: Ch7; **b**, **d**, **f**: Ch8

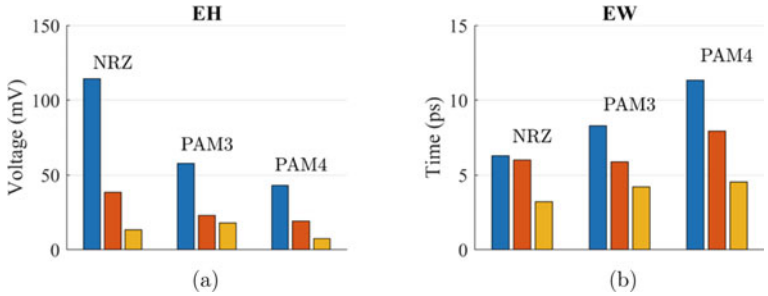


Fig. 9.21 Sensitivity to skew **a** EH; **b** EW. Blue: No skew; Red: 10 ps skew; Yellow: 20 ps skew

Noise

For an electronic communication system, the dominant noise types are typically flicker noise in the low-frequency range and thermal noise in the high-frequency range. Since the corner frequency, which is the intersection of the two types of noises, is normally below 1 MHz, thermal noise should have a greater impact on a high-speed transmission system than flicker noise (Analog Devices 2013). In this study, we will focus on the impact of thermal noise. Equation (9.7) gives the equation for calculating thermal noise

$$P_n = \langle i^2 \rangle \cdot R = 4 \cdot K \cdot T \cdot N_F \cdot B \cdot R \tag{9.7}$$

where P_n is the noise power, $\langle i^2 \rangle$ is the variance of the noise current, K is Boltzmann's constant, T is the Kelvin temperature, N_F is the noise figure, B is the -3 dB bandwidth, and R is the real part of the termination impedance.

The noise figure N_F denotes the SNR degradation of a system, which can be calculated by Eq. (9.8)

$$N_F = \frac{SNR_i}{SNR_o} \tag{9.8}$$

where SNR_i and SNR_o are the input and output signal-to-noise ratios, respectively. Equation (9.7) tells us that the thermal noise power is linearly proportional to absolute temperature, noise figure, bandwidth, and resistance. Here, the noise figure N_F can be considered as representing the encoding/decoding system of each line code, as we will show in the passages to follow.

For a standard-type high-speed transmission system, the noise at the RX analog input, which is usually the input of automatic gain control (AGC), has the predominant impact on the overall signal-to-noise ratio (SNR) of the high-speed transmission system. As a high-level evaluation of the noise performance, as is the case for this study, the test noise only needs to be applied at the input of AGC since such an application of noise is able to characterize the majority of the noise performance of a system.

For an encoding system, its SNR is defined as

$$SNR_{code} = \frac{P_s}{P_n} \tag{9.9}$$

where P_s is the signal power, and P_n is the noise power.

We are not able to find literature that clearly defines the SNR of a channel consisting of a number of cascaded components. Based on the method presented in Analog Devices (2013); Yang et al. (2012), we derive the recursive definition of SNR at the kth component's output in a channel

$$SNR_k = \frac{P_{s\_k}}{P_{n\_k}} = \frac{P_{s\_0} \prod_{i=0}^{k} \alpha_i}{P_{n\_k-1}\alpha_k + P_{n\_k}} \tag{9.10}$$

where P_{s_k} and P_{n_k} are the signal power and the noise power at the output of the kth segment, respectively. P_{s_0} is the signal power generated by the TX silicon, i.e., the 0th component. α_k is the attenuation/gain of the kth component. P_{n_k-1} is the noise power of the $(k-1)th$ component.

Before we start to compute the SNR of each line code, the following assumptions are made as the foundation of the noise performance comparison:

1. All line codes use the same termination impedance and temperature.
2. The thermal noise is a Gaussian stochastic process with a steady distribution.
3. The noises applied to each wire are of equal levels and are fully uncorrelated.

Due to the Gaussian nature of the noise, the combined effect of multiple processing circuits, with each of them corresponding to a wire, should be calculated using the variance (or standard deviation) rather than the peak voltage of the noises. Also based on Eq. (9.7), the noise level relates to the bandwidth of the processing circuit which is in proportion to the Nyquist frequency of the line code; thus, the bandwidth of each line code must be reflected in the SNR calculation. The three codes' SNRs are calculated as follows:

$$\begin{aligned} SNR_{NRZ} &= \frac{(V_{swing}^{NRZ})^2}{(V_{noise}^{NRZ})^2} \\ &= \frac{1}{\sum_{i=1}^{2} \sigma_i^2} = \frac{1}{2\sigma_n^2} \end{aligned} \tag{9.11}$$

$$\begin{aligned} SNR_{PAM3} &= \frac{(V_{swing\_min}^{PAM3})^2}{(V_{noise}^{PAM3})^2} \\ &= \frac{\frac{1}{4}}{\sum_{i=1}^{2} \sigma_i^2 \cdot B_{PAM3}} = \frac{1}{\frac{16}{3}\sigma_n^2} \end{aligned} \tag{9.12}$$

$$\begin{aligned} SNR_{PAM4} &= \frac{(V_{swing\_min}^{PAM4})^2}{(V_{noise}^{PAM4})^2} \\ &= \frac{\frac{1}{9}}{\sum_{i=1}^{2} \sigma_i^2 \cdot B_{PAM4}} = \frac{1}{9\sigma_n^2} \end{aligned} \tag{9.13}$$

where σ_i is the standard deviation of the noise on wire i, and σ_n is the uniform value chosen for σ_i—as aforementioned, all wires have an equal noise level. B_{PAM3}, B_{PAM4}, B_{ENRZ} are the bandwidths required by each code normalized to NRZ, i.e., 0.667, 0.5, and 0.667.

The SNR here is calculated in a voltage-based manner. The same results should be obtained if changed to the power-based manner, with the equations modified accordingly.

The SNR ratio of the four signalings (using NRZ as the baseline) can be further calculated as follows:

$$\begin{aligned} NRZ : PAM3 : PAM4 \\ = 1 : 0.612 : 0.471 \\ = 0\,\text{dB} : -4.26\,\text{dB} : -6.53\,\text{dB} \end{aligned} \tag{9.14}$$

It should be noted that the SNRs calculated above have taken into account the amplitude penalty of PAM3 and PAM4 with respect to NRZ, i.e., 6.02 dB for PAM3, and 9.54 dB for PAM4. Therefore, the reduced noise penalty of PAM3 (−4.26 dB) and PAM4 (−6.53 dB) shows that if the amplitude penalty is taken out of the equation,

then PAM3 and PAM4 have a better noise performance than NRZ—thanks to their smaller Nyquist frequencies.

The simulation is performed on Ch3 with the baseline simulation configuration. An AGC thermal noise of varying levels over 3 dB BW, i.e., 1.5 Nyquist frequency, or 0.75 bit rate for NRZ, is applied to each of the wires during the simulation of each line code. Since the AGC thermal noise is specified in RMS, the amplitude needs to be multiplied by 1.414 when applied to the wires. Equation (9.15) gives the computation of the AGC noise

$$V_{n\_AGC} = 1.414 \cdot V_{n\_white\_rms} \cdot B \tag{9.15}$$

where V_{n_AGC} is peak-to-peak noise voltage produced by the AGC circuit at RX, $V_{n\_white\_rms}$ is the root mean square (RMS) thermal noise of the AGC circuit, and B is the −3 dB bandwidth.

With the AGC noise voltage obtained, we can calculate the channel output under the influences of noise. As an example, the output of wire 0 is calculated as follows:

$$V_{w0} = \sqrt{V_{w0\_in}^2 + V_{n\_AGC}^2} \tag{9.16}$$

where V_{w0} is the output of wire 0 under the influence of the AGC noise, V_{w0_in} is the wire 0 signal arriving at the input of AGC where it has not been impacted by the RX noise yet. All three signals are of RMS values.

For the three differential signalings NRZ, PAM3, and PAM4, independent Gaussian noises are applied to the P and N wires.

In both the OIF CEI and the Ethernet 802.3 standards, there is a channel operating margin (COM) parameter called "one-sided noise spectral density" used to represent the noise level at RX (IEEE 802.3 ck Work Group 2020).

For 200GAUI-4 and 400GAUI-8 chip-to-chip Ethernet channel, this noise parameter (η_0) is defined as $2.6e^{-8}V^2/GHz$ which can be translated into the amplitude spectrum density as shown by Eq. (9.17)

$$\begin{aligned} \eta_0 = 2.6 \cdot 10^{-8} \cdot V^2/GHz = 2.6 \cdot 10^{-8} \cdot 10^{-9} V^2/Hz \\ \Leftrightarrow 5.10nV/\sqrt{Hz} \end{aligned} \tag{9.17}$$

In this study, only thermal noise is considered, and all other types of noises, as well as all types of jitters, are ignored. For this reason, our experiments show that an AGC thermal noise of 5.18 nV/$\sqrt{\text{Hz}}$ does not produce much notable degradation. Such a result, we deem, should be reasonable since many other impairments are not included in this study. After increasing the noise level to 10 nV/$\sqrt{\text{Hz}}$, significant degradation starts to show. Therefore, we use 10 nV/$\sqrt{\text{Hz}}$ and 20 nV/$\sqrt{\text{Hz}}$ as two noise levels to evaluate the performances of the four codes.

The simulation results obtained with 20 nV/$\sqrt{\text{Hz}}$ noise applied are presented in Fig. 9.22. The EHs and EWs are measured at 1e–12 bit error rate. Figure 9.23 presents

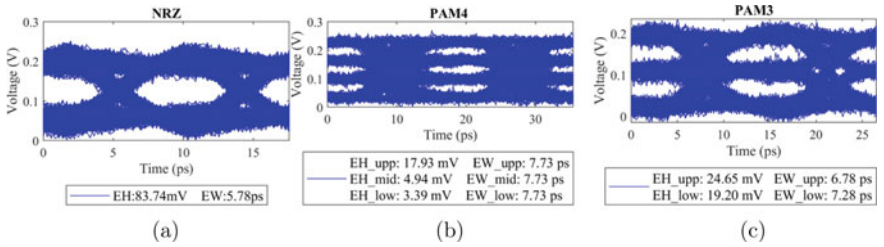


Fig. 9.22 Eye diagrams obtained with the three encoding schemes, with 20 nV/$\sqrt{\text{Hz}}$ noise applied

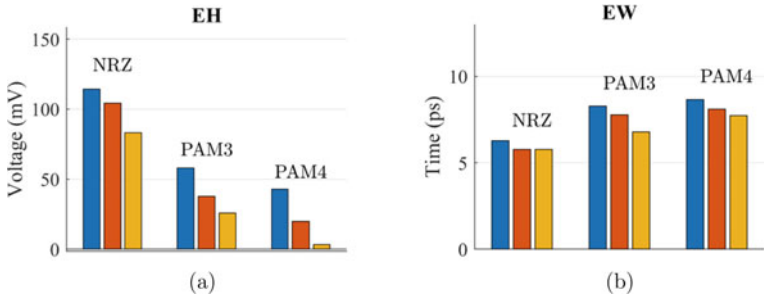


Fig. 9.23 Sensitivity to noise **a** EH; **b** EW. Blue: No skew; Red: 10 ps skew; Yellow: 20 ps skew

the comparison of the EH and EW obtained with the three line codes. We have the following observations:

1. EH-wise, the order of sensitivity with respect to noise is PAM4, PAM3, and NRZ, with PAM4 being the most sensitive code scheme.
2. EW-wise, the order of sensitivity with respect to noise is PAM3, PAM4, and NRZ, with all three codes being relatively insensitive to noise, at least for the noise levels below 20 nV/$\sqrt{\text{Hz}}$.

In summary, NRZ demonstrates the best noise performance among the three, and PAM3 is ranked in the middle, with PAM4 having the worse performance under the influence of noise.

Summary

In this chapter, we start with a brief history of the 3-level modulation technique, then move on to the discussion of PAM3. Two encoding scheme variants 3B2T and 11B7T, and the two ways of producing the pseudo-random test sequence for PAM3 are explained. The high-speed transmission performance of PAM3 is analyzed, in the presence of various channel impairments including reflection coefficient, insertion

loss, crosstalk, skew, and noise. For the purpose of performance comparison, NRZ and PAM4 are included in the simulation and analysis. The results show that while PAM3 has a certain level of robustness in terms of reflection coefficient, insertion loss, and mode conversion, it also has a relatively higher sensitivity with respect to crosstalk and noise. This work and its conclusion can be used as a reference when deciding on the optimal high-speed modulation method for a certain implementation.

References

Analog Devices (2013) Managing noise in the signal chain, Part 1: annoying semiconductor noise, preventable or inescapable? https://www.analog.com/en/technical-articles/managing-noise-in-the-signal-chain.html/. Accessed 17 September 2022

Broadcom Corporation (2014) BroadR-Reach? Physical layer transceiver specification for automotive applications, V3.0, May 2014

Chen SS, Xu Z (2021) The sensitivity of ENRZ to Skew—in comparison to NRZ, PAM3, and PAM4. In: 2021 IEEE 25th workshop on signal and power integrity (SPI), pp 1–4

Chen SS, Xu Z, Holden B, Tajalli A (2021) The sensitivity of ENRZ to crosstalk—in comparison to NRZ, PAM3, and PAM4. In: IEEE international joint EMC/SI/PI and EMC Europe symposium, pp 1116–1121

David B (2015) The mysterious disappearance of duo-binary signaling. https://www.edn.com/the-mysterious-disappearance-of-duo-binary-signaling/. Accessed 17 September 2022

Holden B (2013) An exploration of the technical feasibility of the major technology options for 400GE backplanes. http://www.ieee802.org/3/400GSG/public/13_07/holden_400_01_0713.pdf

https://www.electronics-notes.com (2023) Understanding reflection coefficient. https://www.electronics-notes.com/articles/antennas-propagation/vswr-return-loss/reflection-coefficient.php. Accessed 7 January 2023

https://www.everythingrf.com (2023) Return loss to reflection coefficient calculator—everything RF. https://www.everythingrf.com/rf-calculators/return-loss-to-reflection-coefficient-calculator. Accessed 7 January 2023

http://www.techpowerup.com (2022) Samsung reveals GDDR7 memory uses PAM3 signalling to achieve 36 Gbps data-rate. https://www.techpowerup.com/301767/samsung-reveals-gddr7-memory-uses-pam3-signalling-to-achieve-36-gbps-data-rate. Accessed 11 December 2022

https://www.videocardz.com (2022) SAMSUNG confirms GDDR7 uses PAM3 signaling. https://videocardz.com/newz/samsung-gddr7-memory-to-offer-36-gbps-bandwidth-uses-pam3-signaling. Accessed 11 December 2022

IEEE P802.3df 200 Gb/s, 400 Gb/s, 800 Gb/s, and 1.6 Tb/s Ethernet Task Force (2022) Tools and channel data area. https://www.ieee802.org/3/df/public/tools/index.html. Accessed 7 October 2022

IEEE 802.3 ck Work Group (2020) IEEE 802.3 ck COM script. http://www.ieee802.org/3/ck/public/tools/. Accessed 11 October 2022

IEEE Computer Society (2016a) IEEE standard for ethernet amendment 1: physical layer specifications and management parameters for 100 Mb/s operation over a single balanced twisted pair cable (100BASE-T1). IEEE Standard for Ethernet, March 2016

IEEE Computer Society (2016b) IEEE standard for ethernet amendment 4: physical layer specifications and management parameters for 100 Mb/s operation over a single balanced twisted pair cable (1000BASE-T1). IEEE Standard for Ethernet, September 2016

IEEE Computer Society (2018) IEEE Standard for Ethernet 802.3-2018, December 2018

Law KL, Masouros C (2018) Symbol error rate minimization precoding for interference exploitation. IEEE Trans Commun 66(11):5718–5731

Lender A (1963) The duobinary technique for high-speed data transmission. IEEE Trans Commun Electron 82(2):214–218
Liu Z, Marvel Corp. (2014) Line code and FEC performance. In: IEEE 802.3bp RTPGE, July 2014 Plenary Meeting
OIF (Optical Internetworking Forum) (2022) Common electrical I/O (CEI)-electrical and jitter interoperability agreements for 6G+ bps , 11G+ bps, 25G+, 56G+ bps and 112G+ bps I/O (OIF-CEI-05.0), pp 1–300
Thain A, Le Sergent F, Marot C, Pasquier B (2016) Suitability of 2-wire ethernet solutions for aerospace applications. In: Proceedings of 2016 ESA workshop on aerospace EMC, Aerospace EMC 2016, no 1, pp 4–7
USB-IF (2022) Universal serial bus 4 (USB4) Retimer Specification, Version 2.0, Revision 1.0, October 2022
Wikipedia (2022) T-carrier. https://www.en.wikipedia.org/wiki/T-carrier. Accessed 7 January 2022
Yang N, Yeoh PL, Elkashlan M, Yuan J, Collings IB (2012) Cascaded TAS/MRC in MIMO multiuser relay networks. IEEE Trans Wirel Commun 11(10):3829–3839

Chapter 10
Emerging Interconnect Technologies for Integrated Circuits and Flexible Electronics

Hansel Desmond D'Silva and Amit Kumar

Abstract The chapter examines the latest advancements in electronic interconnect technologies and their effects on system design. It discusses the increasing need for advanced interconnect technologies in catering to faster data transfer, advanced image processing, and stronger computing power, emphasizing the significance of integrating various components into one device. The chapter also analyzes current trends in chip-packages, circuit boards, cables, and connectors, and their contribution to the electronic industry's growth. A brief overview on the trends in the technologies pertaining to high-speed serializer/deserializer, memory and printed/flexible electronics is presented. The chapter offers a comprehensive overview of the most significant trends and advancements in high-speed electronics, making it an indispensable resource for designers and engineers in the field.

Keywords High-speed · Data Rate · Chip · Package · Connector · Cable · Printed Circuit Board

Introduction

As high-speed data rates continue to increase, reaching 25 Gbps and beyond, the application of microwave theory to signal propagation becomes increasingly important (Gao et al. 2020a, 2010; Mingmin et al. 2006; Ng and Yeo 1998). The need for high bandwidth drives data rates closer to 100 Gbps, creating new challenges in effectively managing heat and routing signals through various components such as the package, cable, connector, and printed circuit board.

H. D. D'Silva (✉) · A. Kumar
Achronix Semiconductor Corporation, Bangalore, India
e-mail: hanseldsilva@achronix.com

A. Kumar
e-mail: amitkumar@achronix.com

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_10

To meet these challenges, innovation in signaling and channel topology is required. This requires a thorough understanding of technology trends for different components in a high-speed channel. It has become crucial for designers to integrate different modules into a single device while choosing to create a robust design that can support a variety of products through multiple suppliers. The result is a growing number of innovations aimed at delivering higher throughput and improved performance in electronic products (Kim and Noquil 2005).

The trend toward higher data rates and the need for innovation in high-speed channel technology highlights the importance of staying current with the latest advancements and techniques in the field of electronics. By doing so, designers and engineers can create products that meet the ever-growing demands of technology and provide users with a seamless experience. This chapter delves into the forefront of advancements within package technology, printed circuit boards (PCBs), and flexible electronic technologies. Its purpose is to comprehensively comprehend the transformative innovations occurring in these domains, innovations that play a pivotal role in equipping us to effectively meet the escalating demands for elevated data rates and enhanced bandwidth and throughput capabilities.

Package Technology

The increasing demand for compact and portable devices is driving the need to miniaturize individual components. Moore's law, named after Gordon Moore, co-founder of Intel, stated that the number of transistors on a microprocessor would double approximately every two years, leading to exponential growth in computing power and miniaturization of electronic devices. However, as transistors have approached the physical limits of miniaturization, the rate of growth has slowed down, and experts predict that the law has reached its limits. This has resulted in a shift in focus from shrinking transistors to miniaturizing and integrating components at the package level (Acito 2019).

Packaging technology has become increasingly important in recent years as it plays a critical role in accommodating more and more components at the chip-package level. By integrating components that were previously located on the printed circuit board as on-package components, the design of the device can be simplified while making it more power-efficient. The advancements in semiconductor packaging are helping to meet the demands of emerging technologies, both in terms of application and design, while still considering cost considerations.

As technology advances, the trend in semiconductor packaging is shifting toward a more three-dimensional (3D) approach from the earlier two-dimensional (2D) and two-and-a-half dimensional (2.5D) realized approaches, providing greater design density and performance in a compact form factor (Wu 2015). This shift toward 3D packaging has been driven by the need for ever-increasing data rates and miniaturization of components at the chip-package level. The 1970–2050 semiconductor packaging roadmap highlights the evolution of packaging technology and the need for

greater integration of components at the chip-package level (Status of The Advanced Packaging Industry 2021).

One approach to achieve greater integration is through the use of the chiplet-based design at the chip-package level by the heterogeneous integration of components. This design approach has led to innovations such as the 3D NAND, High Bandwidth Memory (HBM), 3D System-on-Chip (SoC), and large body sized packages for supporting AI, HPC, and networking applications. The trend toward high-density fan-out packaging accelerates in these areas, as well as in mobile devices.

Additionally, advancements in packaging technologies are also supporting the growth of 5G mm Wave technology in mobile devices, such as through the use of double-side molded Ball Grid Array (BGA), low dielectric loss materials, and Antenna in Package (AiP). The trend toward larger package designs using all-side molded Wafer Level Chip Scale Packages (WLCSP) is also expected to increase.

The current trend in the roadmap of packaging technology presents the need for greater integration of components from the perspective of functionality, miniaturization, interconnect density, reliability, and yield, while maintaining cost-effectiveness as the main driving factor. With the increasing demand for high-speed data rates, the development and advancement of packaging technology plays a critical role in the future of electronics.

In summary, the integration of components at the chip-package level has become increasingly important as electronic devices become smaller. Moreover, power consumption is also of prime importance. The trend toward miniaturization and increased data rates has driven the development of a range of package technologies that can meet the demands of emerging technologies while still maintaining performance, reliability, and cost-effectiveness.

Combining Multiple Technologies—Chiplets

The use of chiplets in semiconductor packaging at the chip-package provides a new and innovative solution to the integration of multiple different technologies in miniaturizing components. Instead of trying to integrate multiple technologies on a single substrate, which can result in design and performance limitations, chiplets offer a more efficient approach. This approach involves combining chiplets of different technologies from different suppliers, such as different materials and functions, from different sources, including fabless design houses, foundries, wafer sizes, feature sizes, and companies, using different packaging technologies (Lau 2019a, b). This provides a way to bring together the best of each technology in one integrated system, resulting in improved performance and reduced costs. The ability to integrate chiplets of different technologies allows for greater flexibility and customization, enabling the development of innovative solutions to meet the specific needs of different applications across different technologies.

The various companies in the chip industry are pushing for the standardization of the necessary infrastructure for integrating different IP blocks using chiplets.

The chiplet-based design makes it easier to design more efficient systems for specific applications by allowing designers to choose IP blocks from various vendors across different technologies. Chiplets have emerged as a way to continue or even surpass Moore's Law. The introduction of the Universal Chiplet Interconnect Express (UCIe) Specification 1.0 provides designers with a standardized method of integrating diverse IP blocks of different technologies using chiplets (https://www.uciexpress.org/why-choose-us; https://semiengineering.com/standardizing-chiplet-interconnects). UCIe follows a similar path as the Peripheral Component Interconnect Express (PCIe), which standardizes the interface for add-in card functionality, enabling vendors to mix and match devices for different functions such as graphics, memory, and storage.

The traditional System-on-Chip (SoC) package typically incorporates multiple functional blocks such as processors, co-processors, accelerators, memories, and other input/output functions. However, a chiplet-based approach breaks down these functional blocks into smaller IP blocks which may be manufactured using different technologies, called chiplets. The advantage of this approach is that during manufacturing, any faulty chiplet can be easily replaced, leading to improved yield, cost-effectiveness, and efficiency. The increased level of integration through the use of chiplets allows for a more flexible and scalable design that goes beyond a vendor-specific implementation. By discarding unneeded IP blocks, the chiplet-based approach enables designers to focus on specific applications, resulting in a more robust design.

The AMD 4th generation of EPYC server CPUs, features the use of chiplets in the design. The AMD 4th Gen EPYC 9004 series processors, referred to as Genoa, are based on the Zen 4 DNA technology also used in the company's Ryzen 7000 series of desktop CPUs (https://www.amd.com/en/partner/4th-generation-amd-epyc). The Zen 4 CPU dies are produced with 5 nm technology, while the I/O die is created using 6 nm processes (https://www.amd.com/system/files/documents/4th-gen-epyc-processor-architecture-white-paper.pdf). The use of chiplets in the design of the EPYC Genoa processors enhances the scalability and robustness of the design, allowing for the integration of different IP blocks and the possibility of discarding and replacing individual chiplets during the manufacturing process, resulting in increased efficiency, reduced costs, and improved yield.

Die Stacking

Die stacking is an assembly technique where two or more dies are stacked and bonded inside a single package. It was initially designed to increase memory density by stacking two memory chips together. The process of die stacking provides benefits such as improved performance, enhanced reliability, reduced time-to-market, and the ability to integrate multiple technologies into one package (Die stacking and the system 2012, 2018). There are two types of die stacking: 3D stacking and 2.5D stacking based on interposer, both of which are depicted in Fig. 10.1.

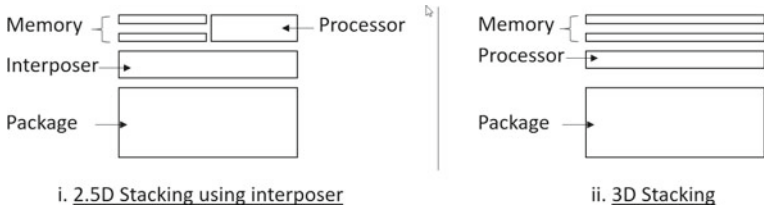


Fig. 10.1 Two flavors of die stacking

The electronics industry is moving toward advanced techniques at the chip-package level instead of placing discrete chips on a board. These techniques include 2.5D, bridges, fan-out, and multi-chip modules (Marte et al. 2018). In 2.5D, the dies are stacked adjacent to each other and connected through through-silicon vias (TSVs) on top of an interposer (Gao et al. 2020b). However, the cost of the interposer can be relatively high, limiting the adoption of 2.5D packaging (Chen et al. 2015; Duan et al. 2021). As a cost-effective alternative, silicon bridges have been introduced as a solution for integrating multiple dies for both logic and memory on a single package (Wu et al. 2018).

Intel's Embedded Multi-die Interconnect Bridge (EMIB) is a solution to the problems posed by the high cost and limited die size of 2.5D packaging with silicon interposers (Duan et al. 2021). EMIB employs a small silicon bridge embedded within a substrate to provide connections between multiple dies in a single package.

The advantage of using a silicon bridge with interposer is that it allows for higher bandwidth communication between the two dies than traditional packaging methods. This is because the interposer can be designed to have a large number of connections and can be made using advanced manufacturing techniques to achieve high density and high performance.

Silicon bridges offer customers a cost-effective solution for interconnecting IP blocks manufactured through different technologies for specific applications. Silicon bridge with interposer technology is commonly used in high-end applications such as graphics processing units (GPUs), high-performance computing (HPC), and data center applications where high-speed communication between chips is critical.

Silicon bridges, along with interposers, help to address the bottleneck in memory bandwidth, particularly in DRAM (https://www.chipestimate.com/Addressing-Memory-Performance-for-100G-Ethernet-Networking/Memoir-Systems/Technical-Article/2012/09/18). The current double data rate four (DDR4) only offers a limited improvement in data rates compared to its predecessor, while Ethernet port speeds have increased significantly over the past decade. To address this issue, the industry has developed high bandwidth memory (HBM), a 3D DRAM technology aimed at high-end systems. HBM stacks multiple DRAM dies on top of each other, providing more IOs. The speeds of DDR4, DDR5, LPDD4, LPDD5, and HBM2 are 3.2 Gbps, 6.4 Gbps, 4.267 Gbps, 6.4 Gbps, and 2.0 to 2.4 Gbps, respectively. The corresponding bandwidths are 25.6 GBps, 51 GBps, 34 GBps, 51 GBps, and 307 GBps. HBM presents significantly greater bandwidth, even though it operates

at lower data rates in comparison to DDRx interfaces. This disparity arises from HBM's utilization of dense routing across an interposer or silicon bridge, enabling the interface to accommodate a substantial 2048 I/Os per HBM interface. However, HBM2 memory, which offers high bandwidth and low capacity, is expensive (price per GB). The industry is exploring other options to address this issue. The use of DDR4 and HBM2 memory helps to balance high performance and low latency, with DDR4 used as high performance memory and HBM2 used as L4 cache memory (https://www.techdesignforums.com/practice/technique/choosing-between-ddr4-and-hbm-in-memory-intensive-applications).

Printed Circuit Board Technology

The advancements in manufacturing technology have allowed for significant growth in the printed circuit board industry. These advancements have resulted in a more flexible design process, enabling designers to incorporate a wider range of structures onto the circuit board (Pfahl et al. 2013). The ease of manufacturing has led to the incorporation of a number of structures on a printed circuit board, which earlier would have been difficult to fabricate. This evolution of printed circuit board technology has created new opportunities for innovation and has opened up new avenues for the development of advanced electronics systems.

The research and development of printed circuit board materials is critical, as each field of application has unique environmental conditions that affect performance. Even small imperfections in printed circuit boards operating at high frequencies can result in various mechanisms that can degrade system margins (Zhang et al. 2019; Hou et al. 2019; Li et al. 2009). To mitigate these effects, a variety of mathematical models have been developed to capture the different mechanisms involved in material properties, which are then incorporated into high-frequency electromagnetic simulations (Engin et al. 2019, 2011).

In the design and fabrication of a printed circuit board, the layer count and the thickness of each layer play a critical role in the overall performance and functionality of the board. The number of layers in a stackup is not a fixed limit and there can be more or fewer layers as per the requirements and capabilities of the manufacturer. As the layer count increases, so does the board thickness, which can impact the board's rigidity. The core layer thickness contributes to the board's overall rigidity, while the thickness of the prepreg layers affects the insulation properties and the thickness of the copper layers impacts the conductivity, resistance, and overall functionality of the board. It is crucial to work with the printed circuit board manufacturer to determine the optimal layer count and thickness for a given application, as each application has unique requirements and environmental conditions.

Impedance discontinuities can have a significant impact on the overall performance of the circuit board and the devices connected to it. As the frequency of the signal increases, the impact of impedance discontinuities become more pronounced, leading to further degradation of the signal quality (https://www.nwengineeringllc.com/article/how-to-design-your-hdi-pcb-stackup.php; Zhang et al. 2014). It is therefore important to carefully design and fabricate the trace and VIA components to minimize impedance discontinuities and maintain signal integrity.

This requires a deep understanding of the electrical properties of the materials used in the circuit board, as well as a knowledge of the various manufacturing processes and techniques involved when creating the trace and VIA components. In addition, careful planning of the routing layers and the placement of the trace and VIA components is essential to minimize the impact of impedance discontinuities on the signal integrity.

To ensure optimal performance, it is advisable to use simulation tools and techniques to model and predict the impact of trace and VIA design on signal integrity, and to validate the design with measurements taken during the manufacturing process. In this way, the impact of impedance discontinuities can be reduced and the signal integrity of the printed circuit board can be maintained, even at high frequencies.

Figure 10.2 illustrates the routing of 16 Gbps GDDR6 signals on a printed circuit board, which was a part of a layout viewed in the Ansys EM Desktop (https://www.ansys.com/en-in/products/electronics). The increasing complexity of signal routing on PCBs is driving the need for innovative routing structures. The design of a PCB requires careful planning to accommodate an increasing number of high-speed signals pertaining to memory due to the growing demand for bandwidth. Figures 10.3 and 10.4 demonstrate the routing of data and command signals for 16 Gbps GDDR6 on a PCB which is was part of a layout viewed in the Ansys EM Desktop (https://www.ansys.com/en-in/products/electronics). Each signal route on a PCB may vary, and ongoing research focuses on finding innovative ways to route signals in dense areas.

The rise of design complexity and density requires exploration of new structures for routing on different layers. One such structure is the Vertical Conductive Structure (VeCS), a slot technology that presents a potential alternative to the VIA technology

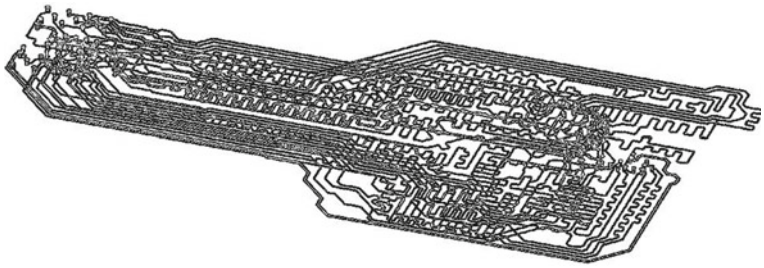


Fig. 10.2 Routing of signals for 16 Gbps GDDR6 across different layers

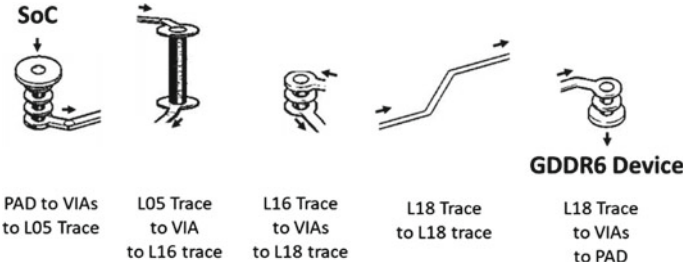


Fig. 10.3 Routing of a data signal for 16 Gbps GDDR6

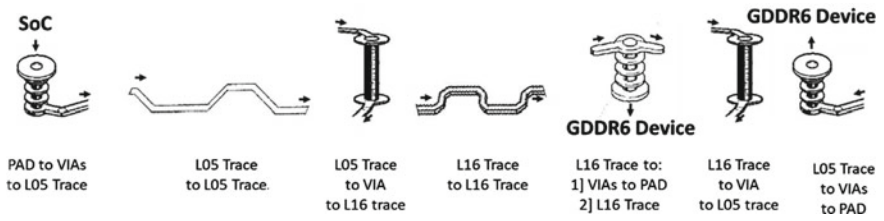


Fig. 10.4 Routing of a command signal for 16 Gbps GDDR6

for accommodating denser trace routing with lower plane inductance (https://www.nextgin-tech.com/innovation-areas/vecs). VeCS technology allows traces to travel vertically through the layer stack-up. Figure 10.5 displays a three-dimensional view in comparing a VIA and VeCS for differential signaling, which was drawn in Ansys EM Desktop (https://www.ansys.com/en-in/products/electronics).

VeCS fabrication involves standard fabrication procedures, including slot drilling in the layer stack-up, followed by metallization and plating of the slot, and finally, hole drilling to remove excess metal, resulting in the creation of vertical trace structures (https://resources.pcb.cadence.com/blog/2020-vertical-conductive-structures-vecs-for-pcb-hd-trace-routing). Figure 10.6 illustrates the fabrication steps of VeCS for differential signaling.

The slot structure of VeCS provides a well-defined ground return path, resulting in improved impedance control at layer transitions and optimized crosstalk at the pin field. The close proximity of the signal and ground in VeCS results in lower plane impedance and improved power delivery. The adoption of VeCS allows for denser trace routing at transition layers, reducing the number of required layers (https://resources.pcb.cadence.com/blog/2020-vertical-conductive-structures-vecs-for-pcb-hd-trace-routing).

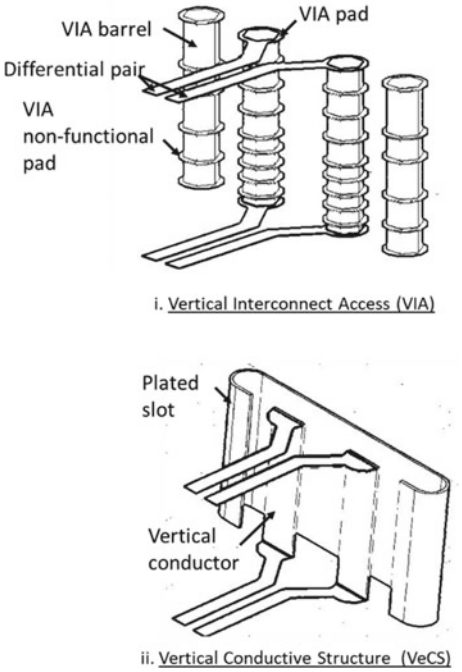


Fig. 10.5 Three-dimensional view of a VIA and VeCS for differential signaling

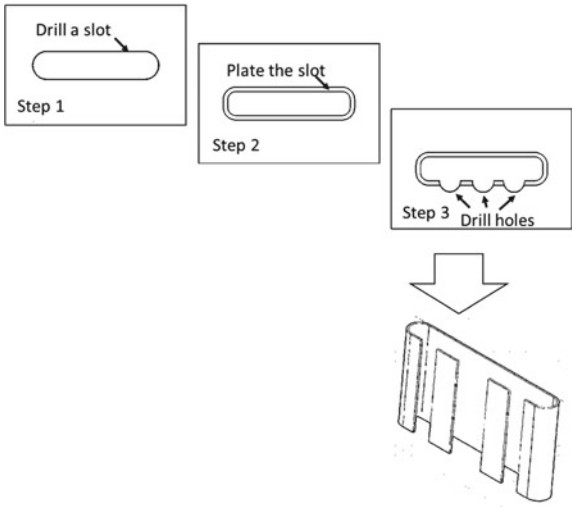


Fig. 10.6 Fabrication of a VeCS for differential signaling

Trend in Data Rate of High-Speed Interfaces

The exponential growth in the demand for data, fueled by the proliferation of applications, is putting pressure on high-speed interfaces to keep pace with increasing data rates. To meet this demand, engineers and researchers are exploring various innovations at both the chip-package and printed circuit board level.

At the same time, there is work in exploring the different signaling and modulation schemes that are tailored to the unique requirements of each application, given the wide range of form-factors that exist at the system level. Table 10.1 provides a visual representation of the different SerDes high-speed interfaces, their corresponding data rates, signaling methods, and unit intervals (Dsilva et al. 2021).

As data rates continue to increase, the unit interval, or the time between consecutive bits, also decreases. This makes it increasingly challenging to detect the signal accurately within the clock cycle. To overcome this challenge, various equalization schemes aimed at detecting the received signal with the maximum accuracy and ensuring error-free transmission are incorporated into the system.

The exponential growth in bandwidth-intensive applications such as artificial intelligence, virtual and augmented reality, and the internet of things are fueling the demand for faster and more efficient memory interfaces. To keep up with the high-speed SerDes interfaces, high-performance memory technologies are becoming increasingly important. Table 10.2 shows the trend in the development of high-speed memory interfaces across DDR (Double Data Rate), GDDR (Graphics Double and LPDDR (Low Power Double Data Rate).

Table 10.1 Different high-speed SerDes interfaces

IO standard	Released year	Signaling	Data rate [Gbps]	Unit interval [psec]
PCIe Gen3	2010	NRZ	8	125.00
10G Ethernet	2002	NRZ	10.3125	96.97
PCIe Gen4	2017	NRZ	16	62.50
25G Ethernet	2014	NRZ	25.78125	38.79
50G Ethernet	2018	PAM4	53.125	37.65
PCIe Gen5	2019	NRZ	32	31.25
PCIe Gen6	2021	PAM4	64	31.25
100G Ethernet	2010	PAM4	106.25	18.82

Table 10.2 Trend in the speed of high-speed memory interfaces

Interface	Released year	Maximum Gbps per pin
DDR1	1998	0.278 (for desktop/laptop modules) 0.213 (for server modules)
DDR2	2003	1.066
DDR3	2007	2.133
DDR4	2014	3.2
DDR5	2020	4.8
GDDR1	1998	3.2
GDDR2	2002	6.4
GDDR3	2004	8
GDDR4	2006	10
GDDR5	2008	8
GDDR5X	2016	12.8
GDDR6	2018	20
GDDR6X	2020	24
LPDDR1	2003	2.5
LPDDR2	2011	4.26
LPDDR3	2012	8.5
LPDDR4	2014	4.26
LPDDR4X	2016	6.4
LPDDR5	2018	6.4
LPDDR5X	2021	8.5

One of the most popular high-speed memory interfaces is GDDR, which offers a high data rate compared to other memory technologies such as DDR and LPDDR. This high-speed memory interface provides reduced latency and high bandwidth as an on-board memory solution, making it an attractive solution for applications that require high performance and reliability, such as self-driving vehicles (https://www.chipestimate.com/Going-Beyond-GPUs-with-DDR6/Rambus/Technical-Article/2019/04/09).

GDDR6 is a newer and faster version of GDDR, with a bandwidth of 20 Gbps per pin (https://www.gamesradar.com/gddr6-memory-explained), which is more than $\times$ 6.25 that of DDR4 and $\times$3.125 that of DDR5 and LPDDR5. This technology provides a significant increase in performance and is well-suited for the ever-growing data-intensive applications that require high bandwidth and low latency.

High-Speed SerDes Channel Topologies

A SerDes Physical Layer (PHY) is a crucial component in a high-speed data transmission system, as it is responsible for transmitting digital signals over a channel. The channel topology refers to the physical arrangement of the elements in the channel, including the cable or connector in defining the environment in which it operates. The topology has a significant impact on the overall performance of the channel, including the amount of attenuation and reflections experienced by the signal as it travels from the transmitter to the receiver.

The SerDes PHY is a critical component in high-speed data transmission systems, as it is responsible for transmitting digital signals over a channel with varying channel topologies. The PHY must be robust enough to support a wide range of equalization schemes in accommodating the different channel topologies for loss and reflections, including chip-to-chip, chip-to-module, and chip-to-chip over connector or cable topologies (IEEE: IEEE Standard for Ethernet 2016; Teng et al. 2022; IEEE 2022). As SerDes technology pushes for higher data rates to meet the demands of modern communication systems, channel topology must adapt in tandem to accommodate the increased frequency and complexity of signal transmission. The synergistic evolution of SerDes data rates and channel topologies underscores the critical interdependence between these two domains in the pursuit of seamless high-speed communication.

The chip-to-chip topology is a common interconnect solution used in high-speed communication between two integrated circuits on a printed circuit board (PCB). The communication between the two chips is established over a medium-reach channel that consists of interconnect elements such as traces and vias. Figure 10.7 shows an example of this topology where two chips communicate over a PCB.

The chip-to-module topology is designed for high-bandwidth switch applications where communication is established over a short-reach channel. The interconnect consists of a combination of elements such as traces, vias, and connectors. Figure 10.8 represents the chip-to-module topology in which a chip communicates with a pluggable optical or copper cable module.

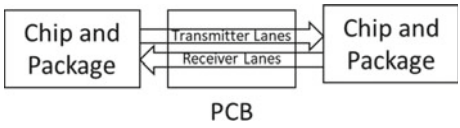


Fig. 10.7 Chip-to-chip topology

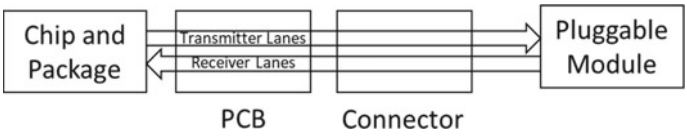


Fig. 10.8 Chip-to-module topology

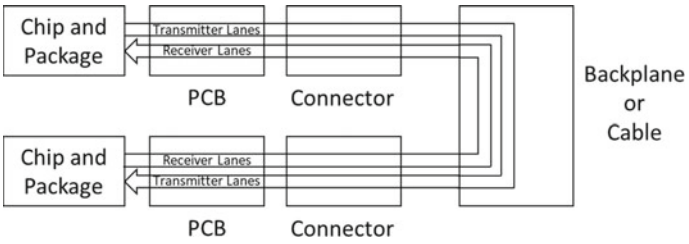


Fig. 10.9 Chip-to-chip over backplane or cable topology

In chip-to-chip over connector or cable topology, the communication is established over a long-reach channel that could be either a cabled solution or advanced PCB technology. The interconnect between the two chips is established over a backplane connector or cable and consists of elements such as traces, vias, and connectors. Figure 10.9 represents the chip-to-chip over connector or cable topology where two chips communicate over a printed circuit board and backplane connector or cable.

Overall, the choice of channel topology depends on the communication requirements, reach, and performance. A robust SerDes PHY needs to be able to support a range of equalization to accommodate the different channel topologies and ensure error-free communication.

The next generation of Ethernet is aiming to reach a data rate of 224 Gbps, which is more than double of the current 100G standard. However, this increased speed also presents new technical challenges, particularly for the chip-to-chip over cable topology. This topology involves connecting two chips using a one-meter cable and two connectors, with printed circuit board trace routing between the chips measuring seven to twelve inches.

To successfully implement this topology at this high data rate, it is important to carefully consider the components and requirements in meeting the specifications. This includes analyzing the link budget, which is a comprehensive evaluation of all the losses and gains in the signal transmission path. This analysis is crucial in determining what is needed to meet the tentative target end-to-end loss of 40 dB up to 53 GHz for 224 Gbps PAM4 signaling.

The goal of this development is to allow faster communication between chips and to ensure the stability of the transmission.

At 224 Gbps PAM4 signaling, for the PCB trace routing, it is recommended to use a ball pitch of 0.8 mm or less to facilitate trace breakout in a dense pin field area (Jiang et al. 2021, 2022). To reduce conductor loss and ensure low loss, it is important to work with the manufacturer to ensure a smooth copper surface roughness.

In the package, the core thickness should be less than thousand micrometers and the total number of routing layers should be no more than twenty-two to meet the one-decibel vertical loss target. In the PCB, accommodation for up to eight signal routing layers is required, with a controlled maximum VIA length of less than sixty five mils and a VIA stub length smaller than six middles to meet the one-decibel of vertical loss target.

The loss budget of the cable assembly is fifteen decibels for a one-meter cable with two connectors. System designers are working with connector and cable suppliers to optimize the cable assembly for both attenuation and reflections along with reliability. The described link budgeting can be used to guide the routing on the PCB to ensure compliance with the end-to-end specification.

High-Speed Memory Channel Topologies

The integration of equalization into memory PHYs is a response to the growing demand for low latency and high bandwidth memory channels in today's modern computing applications (Kim et al. 2019; Na and To 2019; Nitin et al. 2018; Kabat et al. 2022). This integration is necessary to overcome the limitations posed by the increase in attenuation, reflections, and crosstalk at higher frequencies during the routing of traces to memory modules. This integration helps to optimize the system's performance and ensure that the data transmission remains reliable, even at high speeds.

In the past, equalization was only implemented in SerDes PHYs, which handle data transmission between two integrated circuits. However, with the growing demand for memory channels with higher bandwidth requirements, more complex equalizations are now being incorporated into memory PHYs. This helps to mitigate the impact of losses due to attenuation and crosstalk, which can significantly degrade the overall performance of the system.

The GDDR6 interface is a cutting-edge solution for memory systems in response to the need for high bandwidth and low latency. The data rate of 16 Gbps offered by the GDDR6 interface is able to support a range of demanding applications such as high-performance computing and machine learning operations, which require large amounts of data to be processed quickly (https://www.chipestimate.com/Going-Beyond-GPUs-with-DDR6/Rambus/Technical-Article/2019/04/09; https://www.gamesradar.com/gddr6-memory-explained). The high frequency of operation associated with this interface means that the design of both the package and PCB must be optimized to minimize losses, reflections, and crosstalk. The GDDR6 interface also employs equalization techniques to counteract the degraded signal quality at higher frequencies. By meeting the requirements of high-bandwidth and low-latency applications, the GDDR6 interface is well suited for systems that require efficient memory communication.

Memory systems based on GDDR6 SDRAM technology are structured into channels to increase the overall performance and efficiency of the memory system. Each channel is designed around a 16-bit wide data path, which provides a high-speed data transfer rate. The channels can be configured in different ways to meet the specific requirements of the system (https://www.gamesradar.com/gddr6-memory-explained).

One of the configurations is the $\times 16$ bit data channel configuration, which consists of two independent $\times 16$ bit data channels. Figure 10.10 shows GDDR6 signals in

Fig. 10.10 GDDR6 signals in × 16 mode

×16 mode. In this configuration, each channel is connected directly to the memory controller, which allows for a maximum data transfer rate of sixteen bits per clock cycle. The two channels are independent, which means that they can operate simultaneously and independently, providing increased performance and flexibility for the system.

Another configuration is the ×8 channel configuration, which consists of two devices, each with ×8 channels, in a back-to-back clamshell configuration. In this configuration, each device operates as a separate channel, and the two channels are placed back-to-back, which reduces the signal trace length and the associated signal degradation. The ×8 channel configuration provides a more compact and efficient memory solution for systems that require a lower memory capacity or for systems that operate at a lower frequency.

For the ×8 configuration, the devices are typically positioned on opposite sides of the printed circuit board, with one device on the top layer and the other on the bottom layer, forming a clamshell layout. In the ×16 configuration, a 16-bit channel uses two devices in a clamshell design, where byte 0 comes from channel A of the top device and byte 1 comes from channel B of the bottom device, appearing as a single ×16 channel to the controller.

In the ×8 configuration for data connections, only one of the two data bytes per channel is enabled (byte 0 of channel 0 and byte 1 of channel 1), while the other two data bytes are disabled during data transfer. This configuration follows a T-topology, as shown in Fig. 10.11, where the command/address (CA) bytes for both channels are routed together to both DRAMs located on the bottom and top layers.

Overall, the different channel configurations available with GDDR6 SDRAM-based memory systems provide system designers with the flexibility to design memory solutions that meet their specific requirements. The optimized channel design and the high-speed data transfer rate offered by GDDR6 technology make it a popular choice for high-performance applications, such as high-performance computing and machine learning operations.

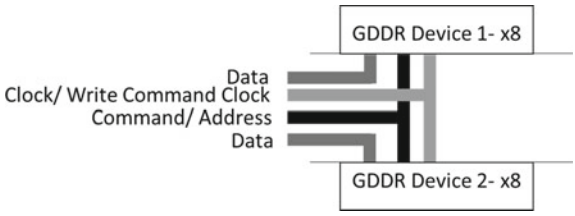


Fig. 10.11 GDDR6 signals in ×8 mode

Printed/Flexible Electronics

Printed and flexible electronics are a growing field in the electronics industry that have the potential to revolutionize the way we access and use electronics. The trend of printed and flexible electronics has been rapidly increasing, moving from being an in-development concept to becoming a commercially viable product (Electronics and in 2021: New Applications Emerging. 2021).

One of the key benefits of printed and flexible electronics is their ability to be easily incorporated into various forms, allowing for unique and innovative applications. For instance, in the healthcare sector, flexible thin-film devices such as skin patches have gained popularity as they enable continuous monitoring of vital signs while providing comfort to patients (Electronics and in 2021: New Applications Emerging. 2021). The recent COVID-19 pandemic has increased the demand for remote consultation and monitoring, which has in turn resulted in a growing need for more printed electronics in the healthcare industry. The ability of printed and flexible electronics to be used in a variety of applications, combined with the recent advancements in the technology, makes it an exciting and promising field for the future.

The use of printed and flexible electronics in the automotive industry is growing as the trend toward electrification in vehicles increases (https://www.idtechex.com/en/research-report/electronic-skin-patches-2021-2031/821#:~:text=It%20reveals%20significant%20opportunity%2C%20with,products%20attached%20to%20the%20skin). The application of these technologies offers several benefits to vehicle manufacturers, such as increased customization options and improved functionality. For example, the integration of printed and flexible electronics can enhance the in-car experience by adding high-performance displays and capacitive control surfaces. These technology advancements have already resulted in the commercialization of backlit capacitive touch sensors in Volkswagen models, which utilize transparent printed metal mesh films and thermoformed parts (https://www.idtechex.com/en/research-report/printed-and-flexible-electronics-for-automotive-applications-2021-2031-technologies-and-markets/806). This highlights the potential of printed and flexible electronics to help vehicle manufacturers provide innovative and high-quality experiences for their customers. Furthermore, the use of printed and flexible electronics in automotive radar systems is also increasing, demonstrating the versatility of these technologies in different aspects of the automotive industry.

As the demand for printed/flexible electronics continues to grow, the industry is expected to experience further commercialization and expansion in 2022. The unique feature of flexibility provided by printed/flexible electronics is attracting its adoption across various industries. In response to this growing trend, simulation methodologies are also being updated to accommodate the adoption of printed/flexible electronics. This involves incorporating new mathematical models into existing simulation infrastructure to better predict and understand the behavior of these devices. This helps companies to design and develop printed/flexible electronics with improved performance, reliability, and cost-effectiveness. As a result, the widespread adoption of

printed/flexible electronics likely continue to drive innovation and growth across a variety of industries in the coming years.

Summary

The trends in the technologies at the chip-package level, printed circuit board, cable, and connector levels indicate the growth of electronic industry. The availability of different technologies enables designers to develop chips that can be customized for performance and area. High data rates require optimization at the system level, not just at the component level. Current chip-packages need to support different channel topologies while supporting different data rates. As the data rates for SerDes approach 100G+, designers need to work with connector and cable vendors to optimize the product. Innovation in memory design is necessary to keep up with the ever-increasing data rate of high-speed SerDes interfaces.

Printed/flexible electronics are being adopted in various applications and are moving from a prototype phase to commercialization. New mathematical models are necessary for simulating printed/flexible electronics. The current trend in the adoption of printed/flexible electronics is driving an update in simulation methodologies using existing infrastructure with the acceptance of new mathematical models.

Acknowledgements The author, Hansel Desmond D'Silva, would like to express his gratitude to his son, Howard, his wife, Chrisenn, for their encouragement and support throughout the writing of this chapter. He would also like to acknowledge his parents and godparents for instilling in him the values and wisdom that have contributed to his success and strong faith. The author Amit Kumar would like to express his gratitude to his son, Avyan, his wife, Anupriya, his parents, brother, and the entire family for always standing with him and supporting him in every phase of life.

References

Acito B (2019) Leveraging the best of package and IC design for system enablement. In: 2019 international wafer level packaging conference (IWLPC). IEEE, San Jose, CA, USA, pp 1–4. https://doi.org/10.23919/IWLPC.2019.8914109

AMD EPYC™ 9004 series processors. https://www.amd.com/en/partner/4th-generation-amd-epyc

Addressing memory performance for 100G ethernet networking. https://www.chipestimate.com/Addressing-Memory-Performance-for-100G-Ethernet-Networking/Memoir-Systems/Technical-Article/2012/09/18

Ansys electronics-electromagnetic, signal integrity, thermal and electro-mechanical simulation solutions. https://www.ansys.com/en-in/products/electronics

Chen S, Li C, Chen Y, Zhang W, Zhong J, Cai J (2015) Study on electrical performance of stacking die package with silicon interposer. In: 2015 16th international conference on electronic packaging technology (ICEPT). IEEE, Changsha, China, pp 1158–1161. https://doi.org/10.1109/ICEPT.2015.7236785

Chunquan L, Xiaole K (2010) Effects of ground vias on high-speed signal transmission in high-speed PCB design. In: 2010 11th international conference on electronic packaging technology &

high density packaging. IEEE, Xi'an, China, pp 88–91. https://doi.org/10.1109/ICEPT.2010.5582357

Choosing between DDR4 and HBM in memory-intensive applications. https://www.techdesignforums.com/practice/technique/choosing-between-ddr4-and-hbm-in-memory-intensive-applications

Duan G, Kanaoka Y, McRee R, Nie B, Manepalli R (2021) Die embedding challenges for EMIB advanced packaging technology. In: 2021 IEEE 71st electronic components and technology conference (ECTC). IEEE, San Diego, CA, USA, pp 1–7. https://doi.org/10.1109/ECTC32696.2021.00012

Die stacking and the system(2012) In: 2012 IEEE hot chips 24 symposium (HCS). IEEE, Cupertino, CA, USA, pp 1–18. https://doi.org/10.1109/HOTCHIPS.2012.7476475

Dsilva H, McMorrow S, Gregory A, Krooswyk S, Mellitz R, Lee B (2021) De-mystifying the impact of Intra-pair Skew on high-speed SerDes Interconnect. In: 2021 IEEE 25th workshop on signal and power integrity (SPI). IEEE, Siegen, Germany, pp 1–4. https://doi.org/10.1109/SPI52361.2021.9505171

Electronic skin patches 2021–2031. https://www.idtechex.com/en/research-report/electronic-skin-patches-2021-2031/821#:~:text=It%20reveals%20significant%20opportunity%2C%20with,products%20attached%20to%20the%20skin

Engin AE, Ndip I, Lang K-D, Aguirre J (2019) Closed-form multipole debye model for time-domain modeling of lossy dielectrics. IEEE Trans Electromagn Compat 61:966–968. https://doi.org/10.1109/TEMC.2018.2838522

Gao H, De S, Payne S (2020a) Impact of surface roughness on return planes in high speed signal propagation. In: 2020a IEEE international symposium on electromagnetic compatibility & signal/power integrity (EMCSI). IEEE, Reno, NV, USA, pp 150–154. https://doi.org/10.1109/EMCSI38923.2020.9191511

Gao G, Theil J, Fountain G, Workman T, Guevara G, Uzoh C, Suwito D, Lee B, Bang KM, Katkar R, Mirkarimi L (2020b) Die to wafer hybrid bonding: multi-die stacking with Tsv integration. In: 2020b international wafer level packaging conference (IWLPC). IEEE, San Jose, CA, USA, pp 1–8. https://doi.org/10.23919/IWLPC52010.2020.9375884

GDDR6 memory explained: what it is, and why it matters for next-gen. https://www.gamesradar.com/gddr6-memory-explained

Going beyond GPUs with GDDR6. https://www.chipestimate.com/Going-Beyond-GPUs-with-DDR6/Rambus/Technical-Article/2019/04/09

Hou F, Wang W, Lin T, Cao L, Zhang GQ, Ferreira JA (2019) Characterization of PCB embedded package materials for SiC MOSFETs. IEEE Trans Compon, Packag Manuf Technol 9: 1054–1061. https://doi.org/10.1109/TCPMT.2019.2904533

How to design your HDI PCB stackup. https://www.nwengineeringllc.com/article/how-to-design-your-hdi-pcb-stackup.php

Institute of electrical and electronics engineers (IEEE) (2022) IEEE standard for ethernet amendment 4: physical layer specifications and management parameters for 100 Gb/s, 200 Gb/s, and 400 Gb/s electrical interfaces based on 100 Gb/s signaling, In: IEEE Std 802.3ck-2022 (Amendment to IEEE Std 802.3–2022 as amended by IEEE Std 802.3dd-2022, IEEE Std 802.3cs-2022, and IEEE Std 802.3db-2022), pp 1–316. https://doi.org/10.1109/IEEESTD.2022.9999414

IEEE (2016) IEEE Standard for Ethernet. In: IEEE Std 802.3–2015 (Revision of IEEE Std 802.3–2012). USA, pp 1–4017. https://doi.org/10.1109/IEEESTD.2016.7428776

Jiang X et al (2022) 224 Gbps-PAM4 end-to-end channel solutions for high-density networking system. In: DesignCon, USA

Jiang X et al (2021) Designing 224G High performance FPGA package and board with confidence. In: DesignCon, USA

Kabat AK, Pandey S, Gopalakrishnan VT (2022) Performance evaluation of high bandwidth memory for HPC workloads. In: 2022 IEEE 35th international system-on-chip conference (SOCC). IEEE, Belfast, United Kingdom, pp 1–6. https://doi.org/10.1109/SOCC56010.2022.9908071

Kim S, Kim S, Kim H, Cho K, Kim J (2019) Design methodology of passive equalizer for GDDR6 memory test. In: 2019 joint international symposium on electromagnetic compatibility, sapporo and Asia-Pacific international symposium on electromagnetic compatibility (EMC Sapporo/APEMC). IEEE, Sapporo, Japan, pp 270–273. https://doi.org/10.23919/EMCTokyo.2019.8893936

Kim J, Noquil JA, keng Tan T, Wu C-L, Choi S-Y (2005) Multi-flip chip on lead frame overmolded IC package: a novel packaging design to achieve high performance and cost effective module package. In: Proceedings electronic components and technology, 2005. ECTC '05. IEEE, Lake Buena Vista, FL, USA, pp 1819–1821. https://doi.org/10.1109/ECTC.2005.1442043

Lau JH (2019a) Heterogeneous integrations. Springer Singapore, Singapore

Lau JH (2019b) Recent advances and trends in heterogeneous integrations. J Microelectron Electron Packag 16:45–77. https://doi.org/10.4071/imaps.780287

Li L, Lee DW, Hwang K-P, Min Y, Wang SX (2009) On-package magnetic materials for embedded inductor applications. In: 2009 international conference on electronic packaging technology & high density packaging. IEEE, Beijing, China, pp 471–474. https://doi.org/10.1109/ICEPT.2009.5270707

Marte A, Ernst U, Wu J, Yauw O, Clauberg H, Buergi D, Chylak B, Friederichs U, Barthold U (2018) Advances in memory die stacking. In: 2018 IEEE 68th electronic components and technology conference (ECTC). IEEE, San Diego, CA, pp 407–418. https://doi.org/10.1109/ECTC.2018.00068

Mingmin Z, Donglin S, Fei D, Shenqing K, Xiaoxiao W (2006) The effect of impedance matching to high-speed connector's signal integrity. In: 2006 7th international symposium on antennas, propagation & EM theory. IEEE, Guilin, China, pp 1–4. https://doi.org/10.1109/ISAPE.2006.353324

Muraoka S, Shinkai G, Yagyu M, Uematsu Y, Ogihara M, Sezaki N, Osaka H (2011) PCB trace modeling and equalizer design method for 10 Gbps backplane. In: 2011 IEEE electrical design of advanced packaging and systems symposium (EDAPS). IEEE, Hanzhou, pp 1–4. https://doi.org/10.1109/EDAPS.2011.6213769

Na N, To H (2019) Effectiveness of equalization and performance potential in DDR5 channels with RDIMM(s). In: 2019 IEEE 69th electronic components and technology conference (ECTC). IEEE, Las Vegas, NV, USA, pp 1208–1214. https://doi.org/10.1109/ECTC.2019.00187

Nitin B, Randy W, Shinichiro I, Eiji F, Shibata R, Yumiko S, Megumi O (2018) DDR5 design challenges. In: 2018 IEEE 22nd workshop on signal and power integrity (SPI). IEEE, Brest, pp 1–4. https://doi.org/10.1109/SaPIW.2018.8401666

Ng CS, Yeo TS (1998) Effects of interline coupling on bit error rate for high-speed digital signal propagation on PCB. IEEE Trans Ind Electron 45:362–364. https://doi.org/10.1109/41.681238

Pfahl B, Fu H, Richardson C (2012) Highlights of iNEMI 2013 technology roadmaps. In: 2012 35th IEEE/CPMT international electronics manufacturing technology conference (IEMT). IEEE, Ipoh, Perak, Malaysia, pp 1–5. https://doi.org/10.1109/IEMT.2012.6521822

Printed electronics in 2021: new applications emerging. https://screenprintingmag.com/printed-electronics-in-2021-new-applications-emerging

Printed and flexible electronics for automotive applications 2021–2031: technologies and markets. https://www.idtechex.com/en/research-report/printed-and-flexible-electronics-for-automotive-applications-2021-2031-technologies-and-markets/806

Status of The Advanced Packaging Industry (2021) https://s3.i-micronews.com/uploads/2021/09/YINTR21223-Status-of-the-Advanced-Packaging-Industry-2021-flyer.pdf

Standardizing chiplet interconnects. https://semiengineering.com/standardizing-chiplet-interconnects

Teng T, Wang M, Hsu J, Lee W, Su T, Kan J, Chen B (2022) PCIe-Express channel design optimization for out-of-guideline three connector topology. In: 2022 17th international microsystems, packaging, assembly and circuits technology conference (IMPACT). IEEE, Taipei, Taiwan, pp. 1–4. https://doi.org/10.1109/IMPACT56280.2022.9966687

Universal Chiplet Interconnect Express. https://www.uciexpress.org/why-choose-us

Vertical conductive structure a new interconnect technology. https://www.nextgin-tech.com/innovation-areas/vecs

Vertical conductive structures (VeCS) for PCB HD trace routing. https://resources.pcb.cadence.com/blog/2020-vertical-conductive-structures-vecs-for-pcb-hd-trace-routing

Wu X (2015) 3D-IC technologies and 3D FPGA. In: 2015 international 3D systems integration conference (3DIC). p. KN1.1-KN1.4. IEEE, Sendai, Japan. https://doi.org/10.1109/3DIC.2015.7334564

Wu J, Yauw O, Tan A, Clauberg H, Buergi D (2018) Thin die stacking technologies for 3D memory packages. In: 2018 China semiconductor technology international conference (CSTIC). IEEE, Shanghai, pp 1–3. https://doi.org/10.1109/CSTIC.2018.8369299

Zhang W, Li N, Kasper E, Zheng Z, Shi L (2019) Impact analysis of high-frequency material and pcb fabrication technology on antenna design for 77/79 GHz automotive radar. In: 2019 IEEE Asia-Pacific microwave conference (APMC). IEEE, Singapore, Singapore, pp 655–657. https://doi.org/10.1109/APMC46564.2019.9038424

Zhang J, Lim J, Yao W, Qiu K, Brooks R (2014) PCB via to trace return loss optimization for >25 Gbps serial links. In: 2014 IEEE international symposium on electromagnetic compatibility (EMC). IEEE, Raleigh, NC, pp 619–624. https://doi.org/10.1109/ISEMC.2014.6899045

4TH Gen AMD EPYC™ processor architecture. https://www.amd.com/system/files/documents/4th-gen-epyc-processor-architecture-white-paper.pdf

Chapter 11
Contact and Interconnect Considerations for Organic and Flexible Electronics

Gargi Konwar and Shree Prakash Tiwari

Abstract Flexible electronics has emerged as a promising technology for demonstration of smart and wearable products due to inherent capabilities for low temperature, low cost, and large-area processing along with multi-functional capabilities in devices. Moreover, this technology has shown other important advantages over conventional electronics such as possibility of fabrication with natural materials and biodegradable substrates leading a path toward eco-sustainable, edible, and green electronics. Organic field-effect/thin-film transistors (FETs/TFTs) are widely explored active devices for flexible electronics due to their suitability for flexible circuits and sensing applications. These devices can act as a crucial block for smart wearable applications such as real-time health monitoring, electronic textiles, and electronic skin due to flexibility and conformability. However, in real systems, these devices face various mechanical, thermal, electrical, and other environmental stimulations during operation. Hence, these devices should be highly stable and reliable under internal or external influences. For integration of the flexible devices in real systems, various contact resistances have to be minimized. Moreover, parasitic effects arising from overlaps and interconnects have to be carefully looked into. In this chapter, firstly, a historical perspective along with potential application areas of flexible electronics are summarized. Further, crucial aspects for designing high-performance organic TFTs with low contact resistance are discussed. New ways adapted by researchers for designing interconnects for flexible and skin-like systems are discussed. Finally, options for alternative natural and biodegradable materials and processes are discussed aiming for eco-sustainable, ingestible, and green electronics.

Keywords Biodegradability · Contact resistance · E-skin · Edible electronics · Flexible electronics · Organic field-effect transistors (OFETs) · Thin-film transistors (TFTs)

G. Konwar · S. P. Tiwari (✉)
Department of Electrical Engineering, Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan 342037, India
e-mail: sptiwari@iitj.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_11

Introduction

In the era of modern internet of things (IoT) enabled smart systems, flexible and stretchable electronics have exhibited massive potential in making human lives easier through various innovative devices (Shi et al. 2020). This field has gained significant exposure in last three decades as one of the emerging research areas with promising applications toward real-time sensors, radio frequency identification (RFID) tags, rollable display screens, biomedical, e-textile, and wearable as well as implantable electronics (Kim et al. 2020; Liu et al. 2022; Rahmanudin et al. 2020; Shi et al. 2020; Singh et al. 2017). As an alternative to traditional rigid electronics, both industry and academia have explored this area for developing smaller and smarter devices with their inherent advantages of lightweight, bendability, large-area processability, and cost-effectiveness as well as compatibility with various flexible and unconventional substrates including thin plastics, textile, paper, etc. Moreover, the use of organic materials adds inherent advantages including solution processability, low-temperature processing, ease of integration, etc. Thin-film transistors (TFTs) are considered to be the most crucial component for flexible electronic systems, especially in circuit components such as backplane drivers and analog interfacing circuits. Intense efforts have been made toward the development of high-performance flexible organic TFTs to reach a level of performance equivalent to amorphous silicon transistors (Sirringhaus 2014). In general, heterogeneous integration is needed to achieve multi-functionality. However, flexible organic TFTs can demonstrate multi-functional properties from a single device, simplifying the preparation process and lowering the overall cost of systems (Kumaresan et al. 2021; Liu et al. 2020; Maddirala et al. 2022; Park et al. 2021; Shang et al. 2021). Moreover, rapid growth of disposable electronics has exacerbated the electronic-waste issue over time, causing serious health and environmental concerns. For this, a variety of environmentally friendly, biodegradable, eco-sustainable smart devices have emerged as a result of the integration of nature-induced materials such as silk, gelatin, egg albumen, chitosan, cellulose, starch, etc., which all offer good degradability, biocompatibility, no requirement of chemical synthesis, that makes them crucial for future sustainable and green electronics (Han et al. 2020; Irimia-Vladu 2014b; Liu et al. 2020; Petritz et al. 2013). They can also be employed as implantable devices in real-time health care and medical diagnosis without harming people's health.

Performance of flexible organic TFTs is highly dependent on various factors starting from material properties, fabrications procedures, quality of films, and crucial interfaces of devices such as semiconductor: dielectric and metal: semiconductor interfaces (Lamport et al. 2018). To achieve high-performance parameters, a suitable selection of these is highly recommended. Low voltage operation with high field-effect mobility (μ), high on–off current ratio (I_{on}/I_{off}), low subthreshold swing (SS), and high environmental stability are the important requirements for high-quality transistors (Sirringhaus 2014). For consistent and stable characteristics in flexible organic TFTs, thermally evaporated organic semiconductors (OSCs) are preferable;

however, solution-processable OSCs are required for low-cost and large-area production capabilities such as inkjet or roll-to-roll printing. In addition, dielectric materials with smooth surface morphology and low gate leakage concerns are highly desirable for these devices as charge conductions happen near the few monolayers of dielectric: semiconductor interfaces (Ortiz et al. 2010). The improvement in interfacial properties of semiconductor/dielectric by introducing hydrophobic self-assembled monolayers (SAMs) or ultrathin polymer films is widely explored. Moreover, high-quality OSCs film can be grown over smooth dielectric surfaces with lower charge trapping sites there. Various high-k inorganics, polymers, and biopolymer dielectrics were introduced as gate insulators over the years to accomplish high electrical performances in flexible devices. High-k dielectrics were widely studied since they provide high gate capacitance for low-voltage operation. However, these have some limitations of charge trapping and lower flexibility as compared to polymer dielectrics. The low dielectric constant (2–5) of a majority of these polymer hinders obtaining low operating voltage TFTs. Bilayer of inorganic/organic hybrid gate dielectric combinations was introduced, where polymer nanolayers were coated over high-k metal oxide dielectrics to reduce surface scattering and eventually help to grow highly oriented π-conjugated semiconductor molecules on top of the dielectrics (Ortiz et al. 2010; Petritz et al. 2013; Rahmanudin et al. 2020). Another important interface that affects the device performance is semiconductor: metal interface. In general, a Schottky contact is seen at the semiconductor metal interface affecting the injection of carriers from source to semiconductor which is needed for charge transport in devices. Larger contact resistances impact more for short-channel devices. Moreover, device geometries also affect contact resistance (Lamport et al. 2018). Devices with staggered configurations have lower contact resistance as the charges are injected over a large area of semiconductors. Several interface engineering has been introduced to suppress the overall contact resistance values (Hou et al. 2016; Lamport et al. 2018). Moreover, in flexible devices, printable nanocomposite and ink based stretchable interconnects were investigated to make it feasible for flexible substrates while providing an extra degree of freedom (Dang et al. 2017; Li et al. 2019; Zhao et al. 2022). Finally, these devices are connected through interconnect metal lines for implementation of circuits which have to be suitably designed for flexible and stretchable applications.

This chapter firstly discusses the operation and development of TFT devices with various organic/inorganic material components. Further, the two crucial aspects, i.e., contact resistance in the TFT device and interconnects suitable for flexible circuits and systems for e-textile and large-area applications will be discussed in detail. The new designs for reducing contact resistances and interconnect reliability for flexible and stretchable devices, along with new easy and cost-effective processes for roll-to-roll and large-area processing, will be finally discussed.

Organic and Flexible Electronics: Historical Developments

Though some initial explorations had earlier happened, in 1977, a demonstration of electrical conductivity in polyacetylene, conjugated polymer, played a huge role in revolutionizing organic electronics (Chiang et al. 1977). In the year 2000, the Chemistry Nobel Committee awarded A. J. Heeger, A. G. MacDiarmid and H. Shirakawa Nobel prize in chemistry (Chiang et al. 1977; Shirakawa et al. 1977) for the discovery of conductive polymers. Organic transistors with polythiophene (Koezuka et al. 1987; Tsumura et al. 1988) and small conjugate oligomers were demonstrated in the late 1980s (Garnier et al. 1990). On the other hand, development for the fabrication of devices on flexible substrates had started much earlier in 1960s, when the first flexible solar cell arrays were demonstrated by thinning a single crystal silicon wafer up to 100 μm and then integrating on a plastic substrate to offer flexibility (Crabb and Treble 1967; Ray 1967). Brody and colleagues developed the first flexible TFT in 1968, when they fabricated a tellurium TFT on a strip of paper for displays (Brody 1969). Further, they had developed TFTs on a wide range of flexible substrates, including polyethylene, mylar, and anodized aluminum foil. Moreover, in the early 1980s, solar cells on organic polymer substrates were also demonstrated with inorganic active materials (Okaniwa et al. 1982a, b). For organic electronics, Burroughes and coworkers demonstrated the first conjugated polymer-based light-emitting diode (PLED) in year 1990 (Burroughes et al. 1990). Apart from organic light-emitting diodes (OLEDs), organic FETs (Mannsfeld et al. 2010; Meager et al. 2014) and circuits based on organic FETs (Di et al. 2013; Khim et al. 2013; Li et al. 2018), solar cells (Chen et al. 2017; Liu et al. 2016), photodiodes (Lamprecht et al. 2005; Ng et al. 2008; Rauch et al. 2009), and lasers (Gaal et al. 2003) with organic materials have been demonstrated, leading the organic and flexible electronics to the commercial level. Developments toward flexible inorganic electronics excelled mid-1980s for the display industry in Japan. Backplane driving circuitry for flexible displays with Si:H TFTs sparked huge interest in flexible electronics. Since then, various research groups and industries have started to explore technologies for flexible electronics extensively. As compared to rigid technology, factors including ruggedness, portability, lightweight, and low production costs contribute to growth of this area. Later researchers demonstrated high-performance flexible transistors with various vacuum evaporated semiconductors making these devices a real competitor to Si:H TFTs which were earlier considered to be suitable for flexible transistors (Bisoyi et al. 2014; Geiger et al. 2022; Leise et al. 2021). Moreover, high processing temperatures required for inorganic semiconductors made it more challenging for flexible electronics. Printed electronics, polymer electronics, plastics electronics, flexible large-area electronics, etc. are some of the common names used today for flexible organic electronics. Foldable displays, flat panel display drivers, radio frequency identification (RFID) tags, solar cells, printable batteries, smart packaging, flexible sensors, printable circuits are examples of the advancement of flexible electronics (Singh et al. 2017; Sirringhaus 2014).

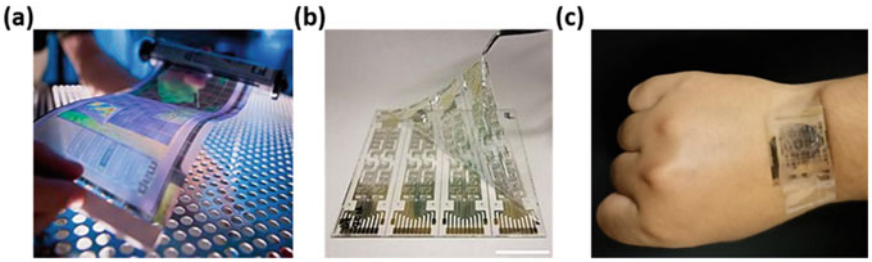


Fig. 11.1 Examples of some flexible electronic devices developed by academia. **a** A flexible display developed by Arizona State University (CC BY 2.0), **b** An organic CMOS logic circuit with total thickness is less than 3 μm by Yasunori Takeda et al. at Yamagata University (CC BY 4.0), and **c** Typical organic transistor devices on transparent flexible substrates by FLAME Research Group, IIT Jodhpur

Moreover, today's flexible electronics has reached to a stage where heterogeneous integration of inorganic and organic devices can be utilized to develop a flexible system for various smart applications (Kumaresan et al. 2021; Ma et al. 2022). Examples of some flexible electronic devices developed by academic institutes are given in Fig. 11.1. These demonstrations range from developing large-area displays to high-performance CMOS-like circuits and multi-functional devices suitable for real-time health monitoring applications (Konwar et al. 2022a, b; Maddirala et al. 2022). Though high performance, stability, and flexibility are demonstrated from many of these devices, multiple implementation challenges are faced at system level when multiple devices are connected together. Contact resistance and interconnect issues become very important for development of reliable systems.

Flexible Electronics: Advantages and Potential Applications

The primary advantages of flexible electronics include bendability, low-temperature processing, large-area processing, multi-functionality, and cost-effectiveness. The property of flexibility allows roll-to-roll fabrications in which batch fabrication is possible at a high speed. In addition, mechanical flexibility broadens the applications of devices in various fields like health care, wearable, implantable, and skin-like electronics, etc. Also, nature-originated materials were recently used for the development of flexible devices to suppress the negative impact of technology on the environment and human health. The major applications are flexible OLED displays, bendable sensors, and arrays, smart cards, flexible circuitry, organic memory devices, artificial skin, etc. Developing OLED-based displays with TFT pixel driver circuits allows them to be fabricated on suitable lightweight and flexible substrates providing desirable features like portability, foldability, rollability, wide viewing angle, etc. Moreover, flexible printed components such as batteries and RFIDs offer various benefits of energy autonomy and IoT capabilities. In 2007, PolyIC had presented the first

printed RFID tag for frequency of 13.56 MHz. These tags have useful for various emerging applications for IoT and wearable devices. Low-cost flexible sensors based on TFT, resistor, or capacitor devices have shown potential applications for human activity monitoring, electronic skin, implantable systems, etc. Organic TFT-based sensors exhibit higher selectivity and sensitivity than the two-terminal-based sensors, because of specific properties of wide range of component materials which are used, and the unique signal amplification capacity of device configuration. The inherent advantages of organic materials have led to the development of lightweight, cost-effective, and easily processable sensors applicable for various smart and stretchable applications.

Flexible Organic TFTs

Solution processability of organic semiconductors helps the TFT devices to be eventually fabricated using printing, making it a suitable candidate for future printed electronics. Organic semiconductors have been deposited by various solution casting techniques, including spin coating, spray coating, printing, etc. at temperatures <100 °C which is much lower than that for polycrystalline Si (~600 °C) and the hydrogenated amorphous Si (~300 °C). This property makes it suitable and compatible to be integrated on flexible substrates like plastics, cloth, paper, etc., which glass transition temperature is around 150 °C. These organic materials can be processed on bendable and foldable substrates due to their soft nature and help to overcome the limitations of existing silicon technology, where skin-like systems are required for health care monitoring devices. The overall advantages and simple processing techniques reduce the budget requirement and help to produce cost-effectively. In 1983, the first organic TFT was demonstrated by Ebisawa where the device was made using polyacetylene as an active semiconductor (Ebisawa et al. 1983). Even though organic TFTs have shown a lot of advantages, achieving high carrier mobility while environmental stability is still a critical challenge. Significant efforts have been made in last decade for achieving mobility higher than 1 $cm^2\ V^{-1}\ s^{-1}$ (Konwar et al. 2022a, b), which is typical for flexible Si:H TFTs. Organic TFTs comprise organic semiconductors as an active layer in place of inorganic active layer. Similar to conventional MOSFET devices they have three terminals naming gate, source, and drain, however, the source and drain are metal in these TFTs contrary to MOSFETs where these electrodes are created by doping or implantation.

Device Structure and Operation

Depending on the position of the electrodes, i.e., gate and source-drain, organic TFTs can have four possible structures as schematically shown in Fig. 11.2. These are called

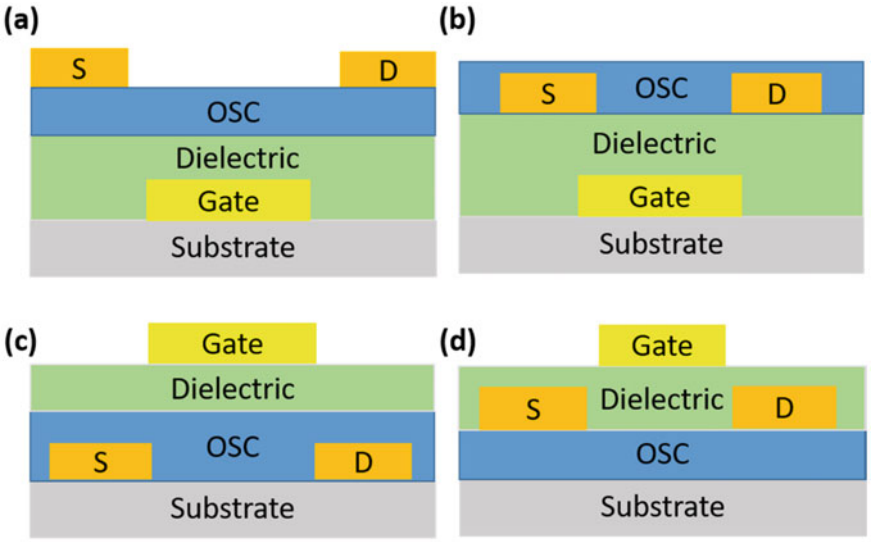


Fig. 11.2 Various configurations of organic TFTs depend on the relative positions of gate, source, and drain contacts. **a** bottom gate top contact, **b** bottom gate bottom contact, **c** top gate bottom contact, and **d** top gate top contact

(a) bottom gate top contact (BGTC), (b) bottom gate bottom contact (BGBC), (c) top gate bottom contact (TGBC), and (d) top gate top contact (TGTC).

Each device structures have their own advantages and disadvantages. For example, top contact structures are preferred for better contact formation and low contact resistance by introducing a large uniform area for charge injection into semiconductor channel. In this configuration, OSCs were deposited over the smooth dielectric surface and do not face many dissimilarities in interfaces hence able to grow better semiconductor film. One of the most significant drawbacks is that photolithography contact patterning is not possible since the contacts must be deposited on the top of OSCs, which may be dissolved by the used solvents in the lithography.

On the other hand, pre-patterned bottom contact electrodes are mostly used for material screening offering simplicity but having issue of large contact resistance due to bad contacts at the electrode edges. Here, photolithographic contact patterning is possible. Top gate structures offer better environmental stability due to the protection layers formed over the semiconductor but it has a lot of process integration limitations as both the gate dielectric and gate contact must be deposited on top of semiconductor layer. Similarly, TGTC configuration is rarely used where physical deposition method like sputtering is not compatible with the gate dielectric layer. Suitable selection of device geometry is recommended to achieve high-performance parameters. Moreover, it eventually depends on the processing methods used for the fabrication of devices.

Operation and Parameter Extraction

It is important to note that organic TFTs work in accumulation mode. However, the characteristics achieved from these devices are identical to those of conventional enhancement mode MOSFETs (Bharti et al. 2016). Since the source-drain electrodes are mostly metals, the charge injection from source to semiconductor is very crucial for operation of these devices. Unlike crystalline silicon which has conduction and valance bands, the energy levels of organic semiconductors are termed as lowest unoccupied molecular orbital (LUMO) and highest occupied molecular orbital (HOMO) corresponding to these levels. For the injection of electrons, the work function of the metal has to be aligned with LUMO level of semiconductor. Similarly, for injection of holes, the work function of the metal has to be aligned with HOMO level of semiconductor. Moreover, these energy levels are shifted with help of applied gate and drain voltages to achieve an n-channel or p-channel device operation. The charge conduction for both types of charge transport is summarized through schematic diagrams in Fig. 11.3. On varying the input gate voltage $V_{GS,}$ the level of HOMO and LUMO were changed to match with the Fermi level of source metal contact as explained in Fig. 11.3.

Initially, when $V_{DS} = V_{GS} = 0$ V, p-type TFT is in an equilibrium state where no current flows between source and drain, and device is in OFF condition as shown in Fig. 11.3a. When $V_{DS} < 0$; $V_{GS} = 0$ V, holes present in p-type OSC migrate

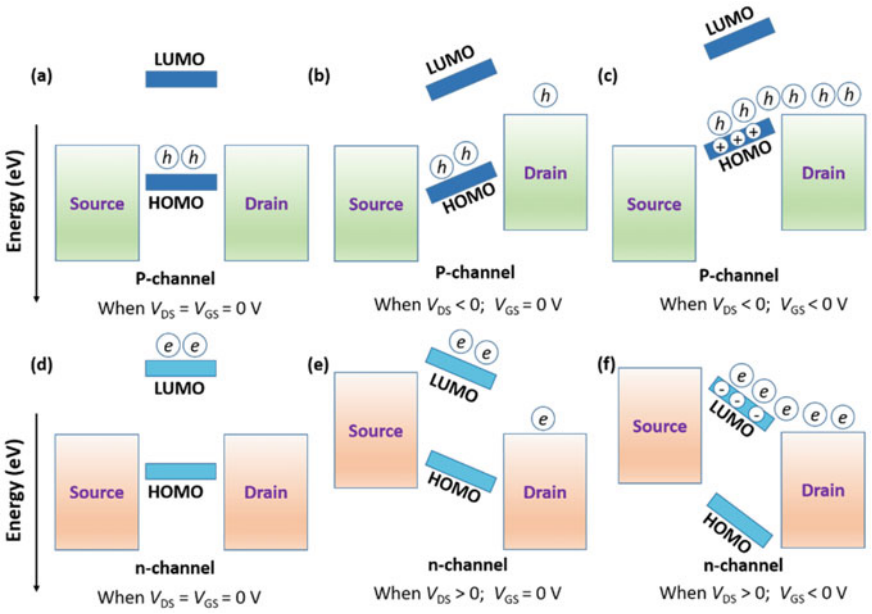


Fig. 11.3 Demonstration of the operating principle of TFTs with p-type OSC **a**–**c** and n-type OSC **d**–**f** depending on applied V_{GS}

toward negative biased drain contact. However, the current at this condition is very negligible since no mobile charges are present due to zero $V_{GS.}$ On applying negative gate voltage (Fig. 11.3c), mobile charges accumulate at the dielectric: semiconductor interface through the dielectric layer, and the transistor starts to conduct. These mobile charges resulted in a high electric field at the interface and shifted HOMO and LUMO levels upward to match the Fermi energy level of drain contact with HOMO level. Similarly, for n-type OSCs (Fig. 11.3d–f) positive biased is applied to turn ON the transistor. Due to similarity of achieved characteristics of organic TFTs with conventional MOSFET, the standard drain current equations are used to model the operation of TFTs in linear and saturation regime.

$$\text{Linear regime, } I_{DS,Lin} = \frac{\mu C_i W}{L}(V_{GS} - V_{TH})V_{DS};\ V_{DS} < (V_{GS} - V_{TH}) \quad (11.1)$$

$$\text{Saturation regime, } I_{DS,Sat} = \frac{\mu C_i W}{2L}(V_{GS} - V_{TH})^2;\ V_{DS} > (V_{GS} - V_{TH}) \quad (11.2)$$

where C_i is gate capacitance density; L and W are length and width of conducting channel, V_{GS} is applied gate input, V_{DS} is drain to source voltage and V_{TH} is threshold voltage.

Device Performance Parameters

Crucial performance parameters of organic TFTs are μ, V_{TH}, I_{on}/I_{off}, subthreshold swing (SS), interface trap density (D_{it}), operational stability, and speed of the device. μ is one of the crucial parameters to define the performance efficiency of any organic transistors, which measures the average charge carrier drift velocity per unit of electrical field. This can be extracted in both linear and saturation regime from their respective (I_{DS}) drain current equations, as given below.

$$\mu_{Lin} = \frac{L}{WC_iV_{DS}}\left(\frac{\partial I_{DS,Lin}}{\partial V_{DS}}\right) \quad (11.3)$$

$$\mu_{Sat} = \frac{2L}{WC_i}\left(\frac{\partial\sqrt{I_{DS,Sat}}}{\partial V_{DS}}\right)^2 \quad (11.4)$$

Low V_{TH} values are preferred for low voltage operation of the device. It defines the minimum V_{GS} required to accumulate sufficient charge at the semiconductor: dielectric interface to turn on the device. This value is generally from the maximum slope of $\sqrt{I_{DS}}$ versus V_{GS} plot of the transfer characteristics in saturation regime along with the extraction of μ. Moreover, a higher C_i value is required to achieve lower V_{TH} and lower voltage operation. Various dielectrics like high-k inorganic or inorganic/organic bilayer hybrid gate dielectric combinations were used to achieve this high C_i

value (Rahi et al. 2022). I_{on}/I_{off} is another important factor which defines as the drain to source current ratio between on and off current. For high-performance devices, it should be high so that both on and off states should be clearly distinguishable. I_{off} can be suppressed through the proper choice of gate material, interface engineering, and patterning of gate electrodes, etc. Similarly, I_{on} is dependent on numerous factors like μ, C_i, operating voltage, trap density, etc. *SS* or inverse subthreshold slope is a measure of how sharply the devices are turned on from the off state which can be defined by following formula

$$\mathrm{SS} = \frac{\partial V_{GS}}{\partial(\log_{10}(I_{DS}))} \tag{11.5}$$

The minimum theoretical *SS* value for any transistors is 60 mV/dec. In general, lower value of *SS* is preferable to turning on the device at a smaller V_{GS} value. Moreover, D_{it} is another crucial parameter to define the number of trapping densities at the dielectric semiconductor interfaces. In general, lower trapping sites are advantageous to operate the device at high charge carrier conductions. D_{it} is calculated by following formula (Held et al. 2015)

$$D_{it} = \frac{c_i}{q}\left[SS\frac{\log(e)}{kT/q} - 1\right] \tag{11.6}$$

where q is electronic charge, k is Boltzmann's constant, and T is temperature at Kelvin. A bad interface leads to slower devices, whereas a high-quality interface helps toward performance and speed of the devices. An appropriate selection of semiconductor: dielectric interface is needed to achieve high performance. Moreover, high operational and electrical stability are other crucial aspects required from the devices in real-world applications. For that, various stability studies are performed including multiple scanning of transfer characteristics, bias-stress, cyclic, electromechanical, and shelf-life stability. For most reported bias-stress studies, a constant V_{GS} is applied for long duration and normalized decay in I_{DS} is observed as given in Eq. (11.7).

$$\frac{I_{DS}(t)}{I_{DS}(0)} = \exp\left[-2\left(\frac{t}{\tau}\right)^{\beta}\right] \tag{11.7}$$

where τ is relaxation time, which defines the charge carrier trapping time and β is stretched exponential factor indicating the involved trap distribution width (Bharti et al. 2016).

Moreover, the operating speed of organic TFTs is much lower compared to inorganic transistors due to various factors including low μ values and higher overlap areas causing parasitic capacitances. Speed can be estimated through the transit frequency (f_T), the maximum frequency at which device can operate, is given by the following equation (Klauk 2018):

$$f_T = \frac{\mu_{eff} V_{DS}}{2\pi L(L + L_{ov,GS} + L_{ov,GD})} \tag{11.8}$$

where μ_{eff} is effective carrier mobility, L is length of the channel, and $L_{OV,\,GS}$, $L_{OV,\,GD}$ are the gate to source and gate to drain overlaps respectively. High mobility is a prerequisite to operating the device at a high switching speed. Other parameters including supply voltage, channel length, and gate overlapping area should be taken care of for overall speed enhancement. Channel length and the overlapping area must be minimized to obtain high-speed transistor. This reduced dimension can only be achieved by patterning the contacts with the help of lithography which is not possible for top contact structures where solvent used in lithography can damage the property of organic semiconductors. Generally, for the flexible TFT devices, demonstrated in laboratories, a common unpatterned bottom gate is used, which is entirely overlapped the drain and source region lowering the speed significantly. As these devices suffer from injection issues, the extracted mobility is often lowered by the actual mobility values in the channel. Moreover, upon scaling the devices to lower L, the contact effects become more dominant affecting the effective mobility values to much lower levels (Tiwari et al. 2009). Hence contact resistance extraction and elimination methods are very important. These aspects are discussed in the next section of the chapter.

Contact Resistance

Contact resistance (R_c) arises primarily from the Schottky barrier of source-drain metal to semiconductor contact. The total voltage dropped across the channel (V_{DS}) is a combination of dropped across source (ΔV_S), drain (ΔV_D), and channel (ΔV_{Ch}). The various resistances associated with any device are illustrated in Fig. 11.4a, showing the resistance between source to semiconductor (R_S), drain to semiconductor (R_D), and channel resistance (R_{Ch}) (Lamport et al. 2018). The contact resistance is contributed by two crucial factors: injection barriers between metal contacts and semiconductor ($R_{C,int}$), and the charge transported from electrode interface to channel ($R_{C,bulk}$). For a reduced value of $R_{C,int}$, proper alignment of metal work function and semiconductor HOMO/LUMO levels is desired. $R_{C,bulk}$ caused by the thickness and surface morphology of OSCs (Matsumoto et al. 2013) can be reduced through good device structure. Controlling charge injection in TFTs while limiting contact effects is a challenging task, and various strategies have been applied to address this issue (Tiwari et al. 2010). In analysis, this barrier is treated as additional resistance to the channel resistance as shown in Fig. 11.4, which schematically represents the variation of channel length and the impact on the resistances of the device between source and drain electrodes. As it can be seen, the ratio of the contact resistance to channel resistance is higher for lower channel devices.

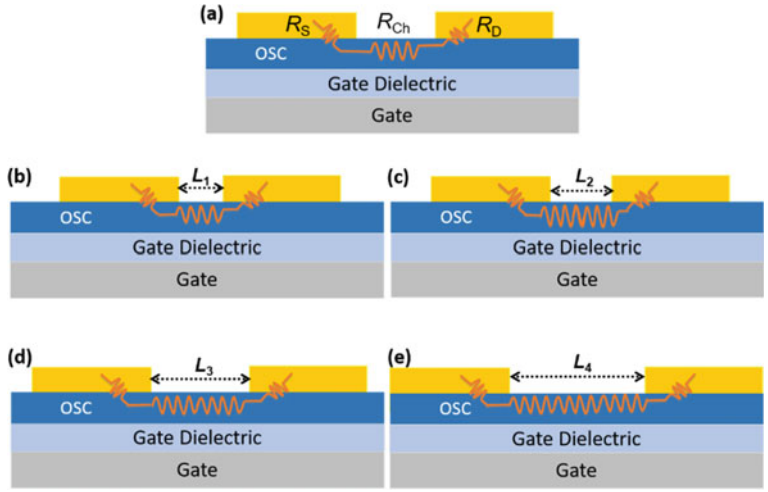


Fig. 11.4 **a** Illustration of various resistances present in between source and drain of a TFT; demonstration of the impact of contact resistance for increased value of channel length **b–e**. As it can be seen, the ratio of the contact resistance to channel resistance is higher for lower channel devices

Contact Resistance Extraction

Measuring contact resistance (R_C) is crucial to overcome the underestimation of extracted μ value. Several techniques including direct and indirect methods were introduced to evaluate R_C (Lamport et al. 2018). One of the most commonly used indirect methods is gated transmission line method (TLM), which has been used as a simple method to extract from electrical characteristics of devices with varying L. It assumes uniform and isotropic semiconductor where R_{Ch} varies linearly with L (Gundlach et al. 2006; Kim et al. 2015; Natali et al. 2016). Figure 11.4 b-e represents the devices in ascending order of L ($L_1 < L_2 < L_3 < L_4$). In TLM, total device resistance (R_{on}) is extracted in a linear regime transfer curve for a particular V_{GS}, by V_{DS} divided I_{DS} when $V_{DS} \to 0$. Further, it is repeated for devices with different L. R_{on} is sum of R_C, and R_{Ch}, where R_{Ch} is directly proportional to L as given below.

$$R_{on} = R_{Ch}(L) + R_C$$

$$= R_{Sh}\frac{L}{W} + R_C$$

$$R_{on}W = R_{sh}L + R_C W \tag{11.9}$$

where R_{Sh} is sheet resistance. The width normalized contact resistance $R_C W$ be estimated by extrapolating $R_{on}W$ to $L = 0$, using the y-intercept of plots of $R_{on}W$ versus L, similar to a line equation y = m.x + C, where y corresponds to $R_{on}W$, x

corresponds to L, and C corresponds to $R_C W$. $R_C W$ can be extracted for multiple V_{GS} values. These plots will show the decreasing $R_C W$ upon increasing V_{GS}. As shown in Fig. 11.4 b–e, it can be observed that as the channel length increases from L_1 to L_4, the value of R_{Ch} is also increased; however, value of R_C is independent and consistent for all structures. For a TFT to function properly, R_C should be negligible compared to R_{Ch}. In a short-channel device, for L_1, both R_C and R_{Ch} values are almost similar, and the impact of R_C is very dominant. The R_C extraction is graphically represented in Fig. 11.5.

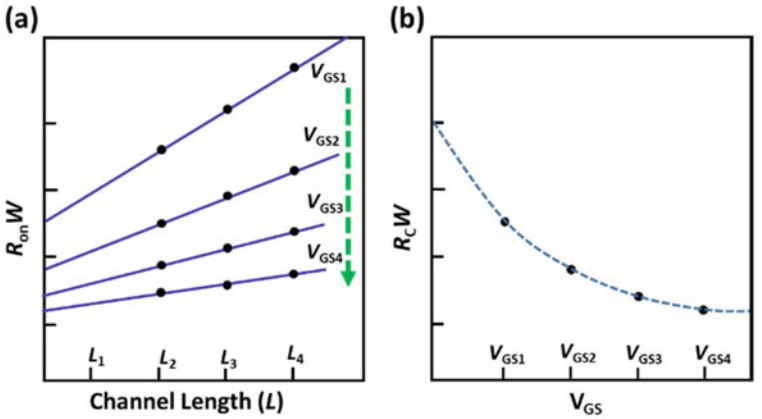


Fig. 11.5 **a** Schematic of $R_{on}W$ versus L plots for various V_{GS} values. The y-intercepts of this plot denote $R_C W$ values. **b** Plot of $R_C W$ values with increasing V_{GS}. As it can be observed, $R_C W$ decreases with increasing V_{GS} values

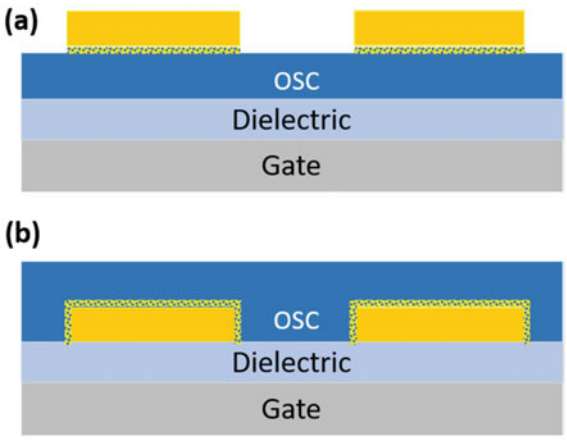


Fig. 11.6 Contact resistance reduction approach **a** selective doping under source/drain contacts for top contacts, and **b** self-assembled monolayer for bottom contact-based TFTs

Approaches to Reduce Contact Resistance

Significant efforts have been made by the researchers to mitigate the $R_{C,int}$ with the help of various approaches such as contact doping (Hou et al. 2016; Tiwari et al. 2010), buffer or self-assembled monolayers (Stoliar et al. 2007), metal oxide interlayer, etc.

Among all these, for top contact devices, selective doping under source/drain is one of the promising techniques to change the Schottky barrier to a graded or ohmic contact. Co-evaporation of host and dopant materials has been done beneath the source/drain electrodes for contact doping in organic devices (Tiwari et al. 2010). These techniques have resulted in complete elimination of contact barriers from devices. Figure 11.6a shows a structure of a contact doped device. Further, Ante et al. have reported organic TFTs with channel length of 100 nm with this approach (Ante et al. 2011). The selective doping forms a hole rich regime near the contacts, which is placed only at the selective area to prevent channel leakage. Though these techniques require thermal evaporation, which is not suitable for solution processing methods, the doping strategies can be explored for solution processing also with suitable mixing of materials. For bottom contact devices, the treatment of metal electrodes with suitable self-assembled monolayers (SAM) enhances the charge injection properties. Au electrodes are invariably treated with SAM layer of octane-1-thiol or other thiols to reduce the contact resistance and enhance the charge injection (Casalini et al. 2012). This technique modifies the morphology of semiconductors and assists to grow semiconductor films with higher crystallinity. In addition, SAMs significantly decrease the charge injection barrier of the semiconductor–metal electrode interface in favor of TFT operation (Youn et al. 2012).

Flexible and Stretchable Interconnects

High flexibility and stretchability of the electrical interconnects are very important for interconnects for flexible and stretchable electronic systems. These properties can be achieved through engineered shapes or using stretchable rubber-like materials. Figure 11.7 represents two types of interconnects, one which has been conventionally used across the technologies, i.e., use for straight metal lines with a suitable cross-sectional area depending on the length of the interconnect. The cross-sectional areas of the interconnect lines control the conductivity and overall resistance of the interconnect. Moreover, aspect ratios of these lines can be tuned to have lower parasitic capacitances caused by overlaps. However, we will focus on the flexibility and stretchability aspects here. For any straight conventional metal line, as shown in Fig. 11.7a, any stretching will cause the breaking of the interconnect which will lead to failure of the circuit or systems where this interconnect is used. Figure 11.7b represents a potential method for designing a reliable interconnect with maximum possible elongation for flexible, large area, and stretchable applications.

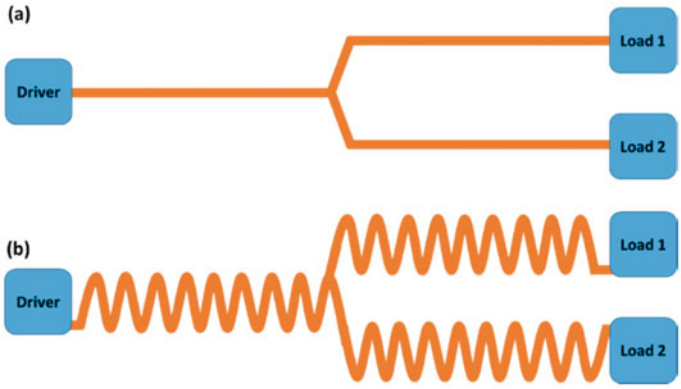


Fig. 11.7 **a** A typical interconnect design for conventional rigid technology or flexible electronics. **b** An engineered interconnect design with 2D or 3D helical structure to induce stretchability. The shapes of this design can vary as per the requirement

Extended versions of this engineered geometry, i.e., various types of helical wires in different shapes have been used for stretchable interconnects for flexible electronics with organic and inorganic semiconductors (Dang et al. 2017; Yu et al. 2017). Helical conductive interconnect design is simple as it resembles the design of a two or three-dimensional spring. This shape induces the stretchability or expandability in the interconnect. The stretchable interconnects have been very useful in the design of skin-like electronic systems. Conductive thread-based clothing has been used in stretchable wearable electronics such as an electronic patch for monitoring strain, pressure, sweat, and other parameters (Ge et al. 2016; Terse-Thakoor et al. 2020). Even with engineered designs, conventional metals still face issues of breaking and failure. To address this, various solution-processed metal inks and nanocomposites are demonstrated in recent years to offer better suitability for flexible, printable, and stretchable electronics (Kumar et al. 2022; Nair et al. 2021a; Widdascheck et al. 2021). Owing to their flexibility, stretchability, conductivity, sensitivity, polymeric nanocomposites with conductive fillers have been recently demonstrated for applications in bioelectronics and wearable devices (Kumar et al. 2022). Silver nanowire-based printable electrothermochromic inks are demonstrated for applications like flexible touch-displays (Nair et al. 2021b).

Toward Green and Biodegradable Electronics

Nowadays, the growing demand for advanced technology has continuously shortened the lifespan of devices and the use of non-sustainable components resulting in a significant increase in electronics-waste (e-waste) issues. The negative impact of e-waste on the environment has led the research community to think of eco-sustainable

alternatives for materials used for producing these electronic systems. To address this, various biodegradable and nature-inspired materials are proposed to be integrated into the manufacturing processes for these technologies. This integration will lead to lower technology footprints on environment after the lifespan of devices. In recent years, various biocompatible, biodegradable materials like natural proteins have been used for fabrication of high-performance flexible TFT devices. Paper is considered to be one of the most abundantly available biodegradable substrates which offers inherent advantages of low cost, lightweight and renewable natures with flexibility, foldability, and recyclability functionalities. In this perspective, developing paper-based devices is very exciting as it could be further extrapolated into high-efficiency printing technologies for large-area applications. Till date, TFTs were successfully developed on a variety of paper substrates including commercial packaging paper, photo paper, starch paper, and bank notes (Casula et al. 2020; Lee et al. 2019; Li et al. 2012; Raghuwanshi et al. 2019; Wang et al. 2017). Similar application have been studied for other biodegradable and biocompatible synthetic polymers, including poly(3-hydroxybutyrate) (PHB), polycaprolactone (PCL), poly 3-hydroxyoctanoate, poly(4-hydroxybutyrate), poly(l-lactide), poly(glycolic acid), poly(glycolide-co-caprolactone), poly(lactide-co-e-caprolactone), and poly(glycerol sebacate) (PGS), etc. (Irimia-Vladu 2014a). Some of the natural proteins used as gate dielectric for high-performance TFTs are silk fibroin, chicken albumen, cross-linked DNA, collagen, etc. (Chang et al. 2011; Hsieh et al. 2013; Kim et al. 2010; Ko et al. 2017; Lee et al. 2018; Tsai et al. 2013; Wang et al. 2011; Zhu et al. 2016). Among these materials, gelatin has been the most widely researched low-cost natural protein used for multi-functional and high-performance TFT devices (Konwar et al. 2022a, b). Devices with gelatin were used to sense stimuli like human breath suitable for real-time breath rate monitoring along with sensing environmental parameters like humidity, UV, and visible light, along with suitability to circuit applications. There are other potential biodegradable materials for these applications. Some of these matrials are derived from edible sources, for example, polylactide is obtained from natural plant sugars and its derivatives are used as flexible substrates and/or dielectrics (Bettinger and Bao 2010).

The edible property of biopolymers have helped to increase the applicability of these eco-friendly devices in emerging biomedical ingestible and edible electronics, where the majority of device material components remained inactive or degraded into non-toxic and environmentally friendly byproducts via enzymes, environmental factors, or water after a certain amount of time (Konwar et al. 2022a, b; Sharova et al. 2021). Silk fibroin is one of the strongest fibers and has a very high dielectric constant, which is needed for high-performance low voltage TFTs. Other nature-inspired diverse organic materials also have the potential to be used in eco-friendly devices such as ecoflex, caramelized glucose, and hard gelatin capsule as substrates, adenine, guanine, glucose, lactose, sucrose, and caffeine for dielectrics, while beta (β)-carotene, indigo, vat orange 3, and perylene diimide as organic semiconductors (Irimia-Vladu et al. 2010). A typical TFT device fabricated on a paper substrate is

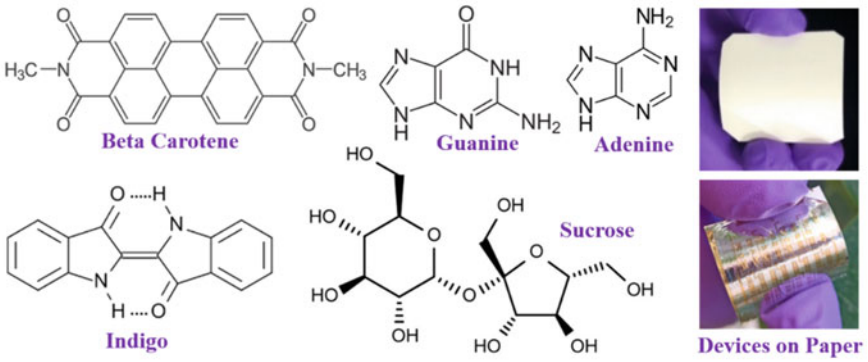


Fig. 11.8 Chemical structures of some nature-inspired edible and biodegradable organic materials for fabrication of electronic devices, and thin-film transistors on paper substrate

shown in Fig. 11.8. In addition, some of the chemical structures for edible organic materials are also shown here. The use of nature-inspired materials in fabrication of flexible devices will help in developing eco-sustainable flexible electronic products.

Summary

In this chapter, firstly, the topic of flexible electronics is comprehensively introduced in the introduction section. A historical review of the development of flexible electronics in last four decades is summarized starting from the demonstration of first flexible TFT and conducting polymer to the current state of the art devices where multi-functionality in these devices with low voltage operation. Further, the advantages and applications of flexible devices are discussed. In the next section, structure and operation of flexible organic TFTs are discussed in detail. Performance parameters along with extraction methods are explained with fundamental aspects. Issues like speed and contact resistance are highlighted and mathematically analyzed. Further, the origin of the contact resistance and extraction methods along with reduction methods are discussed. Moreover, the challenges at system level in designing interconnects for flexible and stretchable electronics are highlighted and possible solutions are presented. Finally, materials and devices for the biodegradable, edible, and green electronics are briefly discussed.

In summary, this chapter discusses performance and challenges for flexible devices, specifically for circuit and system applications including contact and interconnect considerations for organic and flexible electronics, and suitability for biodegradable and green electronics.

References

Ante F, Kälblein D, Zschieschang U, Canzler TW, Werner A, Takimiya K, Ikeda M, Sekitani T, Someya T, Klauk H (2011) Contact doping and ultrathin gate dielectrics for nanoscale organic thin-film transistors. Small 7(9): 1186–1191. https://doi.org/10.1002/smll.201002254

Bettinger CJ, Bao Z (2010) Organic thin-film transistors fabricated on resorbable biomaterial substrates. Adv Mater 22(5):651–655. https://doi.org/10.1002/adma.200902322

Bharti D, Raghuwanshi V, Varun I, Mahato AK, Tiwari SP (2016) High performance and electro-mechanical stability in small molecule: polymer blend flexible organic field-effect transistors. IEEE Electron Device Lett 37(9):1215–1218. https://doi.org/10.1109/led.2016.2592943

Bisoyi S, Zschieschang U, Kang MJ, Takimiya K, Klauk H, Tiwari SP (2014) Bias-stress stability of low-voltage p-channel and n-channel organic thin-film transistors on flexible plastic substrates. Org Electron 15(11):3173–3182. https://doi.org/10.1016/j.orgel.2014.08.057

Brody TP (1969) Letters: paper transistors [letter]. Radio and Electronic Engineer 38(6):326–327. https://doi.org/10.1049/ree.1969.0120

Burroughes JH, Bradley DDC, Brown AR, Marks RN, Mackay K, Friend RH, Burns PL, Holmes AB (1990) Light-emitting diodes based on conjugated polymers. Nature 347(6293): 539–541. https://doi.org/10.1038/347539a0

Casalini S, Shehu A, Destri S, Porzio W, Pasini MC, Vignali F, Borgatti F, Albonetti C, Leonardi F, Biscarini F (2012) Organic field-effect transistors as new paradigm for large-area molecular junctions. Organ Electron 13(5): 789–795. https://doi.org/10.1016/j.orgel.2012.01.020

Casula G, Lai S, Matino L, Santoro F, Bonfiglio A, Cosseddu P (2020) Printed, low-voltage, all-organic transistors and complementary circuits on paper substrate. Adv Electron Mater 6(5), Article 1901027. https://doi.org/10.1002/aelm.201901027

Chang J-W, Wang C-G, Huang C-Y, Tsai T-D, Guo T-F, Wen T-C (2011) Chicken albumen dielectrics in organic field-effect transistors. Adv Mater 23(35):4077–4081. https://doi.org/10.1002/adma.201102124

Chen Y, Ye P, Zhu ZG, Wang X, Yang L, Xu X, Wu X, Dong T, Zhang H, Hou J, Liu F (2017) Achieving high-performance ternary organic solar cells through tuning acceptor alloy. Adv Mater 29(6):1603154. https://doi.org/10.1002/adma.201603154

Chiang CK, Fincher Jr, CR, Park YW, Heeger AJ, Shirakawa H, Louis EJ, Gau SC, MacDiarmid AG (1977) Electrical conductivity in doped polyacetylene. Phys Rev Lett 39(17):1098. https://doi.org/10.1103/PhysRevLett.39.1098

Crabb RL, Treble FC (1967) Thin silicon solar cells for large flexible arrays. Nature 213(5082):1223–1224. https://doi.org/10.1038/2131223a0

Dang W, Vinciguerra V, Lorenzelli L, Dahiya R (2017) Printable stretchable interconnects. Flexible Printed Electron 2(1), Article 013003. https://doi.org/10.1088/2058-8585/aa5ab2

Di CA, Zhang F, Zhu D (2013) Multi-functional integration of organic field-effect transistors (OFETs): advances and perspectives. Adv Mater 25(3):313–330. https://doi.org/10.1002/adma.201201502

Ebisawa F, Kurokawa T, Nara S (1983) Electrical properties of polyacetylene/polysiloxane interface. J Appl Phys 54(6):3255–3259. https://doi.org/10.1063/1.332488

Gaal M, Gadermaier C, Plank H, Moderegger E, Pogantsch A, Leising G, List EJW (2003) Imprinted conjugated polymer laser. Adv Mater 15(14):1165–1167. https://doi.org/10.1002/adma.200305047

Garnier F, Horowitz G, Peng X, Fichou D (1990) An all-organic "soft" thin film transistor with very high carrier mobility. Adv Mater 2(12):592–594. https://doi.org/10.1002/adma.19900021207

Ge J, Sun L, Zhang FR, Zhang Y, Shi LA, Zhao HY, Zhu HW, Jiang HL, Yu SH (2016) A stretchable electronic fabric artificial skin with pressure-, lateral strain-, and flexion-sensitive properties. Adv Mater 28(4): 722–728.https://doi.org/10.1002/adma.201504239

Geiger M, Lingstädt R, Wollandt T, Deuschle J, Zschieschang U, Letzkus F, Burghartz JN, van Aken PA, Weitz RT, Klauk H (2022) Subthreshold swing of 59 mV decade−1 in nanoscale

flexible ultralow-voltage organic transistors. Adv Electron Mater https://doi.org/10.1002/aelm.202101215

Gundlach DJ, Zhou L, Nichols JA, Jackson TN, Necliudov PV, Shur MS (2006) An experimental study of contact effects in organic thin film transistors. J Appl Phys 100(2), Article 024509. https://doi.org/10.1063/1.2215132

Han WB, Lee JH, Shin JW, Hwang SW (2020) Advanced materials and systems for biodegradable transient electronics. Adv Mater 32(51):e2002211. https://doi.org/10.1002/adma.202002211

Held M, Schießl SP, Miehler D, Gannott F, Zaumseil J (2015) Polymer/metal oxide hybrid dielectrics for low voltage field-effect transistors with solution-processed, high-mobility semiconductors. Appl Phys Lett 107(8). https://doi.org/10.1063/1.4929461

Hou J-L, Kasemann D, Widmer J, Günther AA, Lüssem B, Leo K (2016) Reduced contact resistance in top-contact organic field-effect transistors by interface contact doping. Appl Phys Lett 108(10). https://doi.org/10.1063/1.4943646

Hsieh C-Y, Hwang J-C, Chang T-H, Li J-Y, Chen S-H, Mao L-K, Tsai LS, Chueh YL, Lyu PC, Hsu SSH (2013) Enhanced mobility of organic thin film transistors by water absorption of collagen hydrolysate gate dielectric. Appl Phys Lett 103(2): 023303. https://doi.org/10.1063/1.4813075

Irimia-Vladu M (2014a) Green" electronics: biodegradable and biocompatible materials and devices for sustainable future. Chem Soc Rev 43(2):588–610. https://doi.org/10.1039/c3cs60235d

Irimia-Vladu M (2014b) Green" electronics: biodegradable and biocompatible materials and devices for sustainable future. Chem Soc Rev 43(2):588–610. https://doi.org/10.1039/c3cs60235d

Irimia-Vladu M, Troshin PA, Reisinger M, Shmygleva L, Kanbur Y, Schwabegger G, Bodea M, Schwödiauer R, Mumyatov A, Fergus JW, Razumov VF, Bauer S (2010) Biocompatible and biodegradable materials for organic field-effect transistors. Adv Funct Mater 20(23): 4069–4076.https://doi.org/10.1002/adfm.201001031

Khim D, Han H, Baeg KJ, Kim J, Kwak SW, Kim DY, Noh YY (2013) Simple bar-coating process for large-area, high-performance organic field-effect transistors and ambipolar complementary integrated circuits. Adv Mater 25(31):4302–4308. https://doi.org/10.1002/adma.201205330

Kim CH, Hlaing H, Hong JA, Kim JH, Park Y, Payne MM, Anthony JE, Bonnassieux Y, Horowitz G, Kymissis I (2015) Decoupling the effects of self-assembled monolayers on gold, silver, and copper organic transistor contacts. Adv Mater Interfaces 2(2), Article 1400384. https://doi.org/10.1002/admi.201400384

Kim H, Kang T-H, Ahn J, Han H, Park S, Kim SJ, Park MC, Paik SH, Hwang DK, Yi H, Lim JA (2020) Spirally wrapped carbon nanotube microelectrodes for fiber optoelectronic devices beyond geometrical limitations toward smart wearable E-textile applications. ACS Nano 14(12):17213–17223. https://doi.org/10.1021/acsnano.0c07143

Kim YS, Jung KH, Lee UR, Kim KH, Hoang MH, Jin J-I, Choi DH (2010) High-mobility bio-organic field effect transistors with photoreactive DNAs as gate insulators. Appl Phys Lett 96(10):103307. https://doi.org/10.1063/1.3299022

Klauk H (2018) Will we see gigahertz organic transistors? Adv Electron Mater 4(10). https://doi.org/10.1002/aelm.201700474

Ko J, Nguyen LTH, Surendran A, Tan BY, Ng KW, Leong WL (2017) Human hair keratin for biocompatible flexible and transient electronic devices. ACS Appl Mater Interfaces 9(49):43004–43012. https://doi.org/10.1021/acsami.7b16330

Koezuka H, Tsumura A, Ando T (1987) Field-effect transistor with polythiophene thin film. Synth Met 18(1–3):699–704. https://doi.org/10.1016/0379-6779(87)90964-7

Konwar G, Saxena P, Raghuwanshi V, Rahi S, Tiwari SP (2022a) Multifunctional flexible organic transistors with high-k/natural protein bilayer gate dielectric for circuit and sensing applications. ACS Appl Electron Mater, 4(5):2525-2533. https://doi.org/10.1021/acsaelm.2c00303

Konwar G, Saxena P, Rahi S, Tiwari SP (2022b) Edible dielectric composite for the enhancement of performance and electromechanical stability of eco-friendly flexible organic transistors. ACS Appl Electron Mater 4(10):5055–5064. https://doi.org/10.1021/acsaelm.2c01082

Kumar A, Shaikh MO, Kumar RKR, Dutt K, Pan CT, Chuang CH (2022) Highly sensitive, flexible and biocompatible temperature sensor utilizing ultra-long Au@AgNW-based polymeric nanocomposites. Nanoscale 14(5):1742–1754. https://doi.org/10.1039/d1nr05068k

Kumaresan Y, Ozioko O, Dahiya R (2021) Multifunctional electronic skin with a stack of temperature and pressure sensor arrays. IEEE Sens J 21(23):26243–26251. https://doi.org/10.1109/JSEN.2021.3055458

Lamport ZA, Barth KJ, Lee H, Gann E, Engmann S, Chen H, Guthold M, McCulloch I, Anthony JE, Richter LJ, DeLongchamp DM, Jurchescu OD (2018) A simple and robust approach to reducing contact resistance in organic transistors. Nat Commun 9(1): 5130.https://doi.org/10.1038/s41467-018-07388-3

Lamport ZA, Haneef HF, Anand S, Waldrip M, Jurchescu OD (2018) Tutorial: organic field-effect transistors: materials, structure and operation. J Appl Phys 124(7). https://doi.org/10.1063/1.5042255

Lamprecht B, Thünauer R, Ostermann M, Jakopic G, Leising G (2005) Organic photodiodes on newspaper. Physica Status Solidi (A), 202(5):R50–R52

Lee CJ, Chang YC, Wang LW, Wang YH (2019) Biodegradable materials for organic field-effect transistors on a paper substrate. IEEE Electron Dev Lett 40(2): 236–239, Article 8598734. https://doi.org/10.1109/LED.2018.2890618

Lee JH, Kwak HW, Park MH, Hwang J, Kim JW, Jang HW, Jin HJ, Lee WH (2018) Understanding hydroscopic properties of silk fibroin and its use as a gate-dielectric in organic field-effect transistors. Organ Electron 59: 213–219. https://doi.org/10.1016/j.orgel.2018.05.012

Leise J, Pruefer J, Darbandy G, Nikolaou A, Giorgio M, Caironi M, Zschieschang U, Klauk H, Kloes A, Iñiguez B, Borchert JW (2021) Flexible megahertz organic transistors and the critical role of the device geometry on their dynamic performance. J Appl Phys 130(12), Article 125501. https://doi.org/10.1063/5.0062146

Li M, Mangalore DK, Zhao J, Carpenter JH, Yan H, Ade H, Yan H, Müllen K, Blom PW, Pisula W, De Leeuw DM (2018) Integrated circuits based on conjugated polymer monolayer. Nat Commun 9(1): 1–8. https://doi.org/10.1038/s41467-017-02805-5

Li M, Wu Y, Zhang L, Wo H, Huang S, Li W, Zeng X, Ye Q, Xu T, Luo J, Dong S, Wang X (2019) Liquid metal-based electrical interconnects and interfaces with excellent stability and reliability for flexible electronics. Nanoscale 11(12): 5441–5449.https://doi.org/10.1039/c8nr09503e

Li Y, Liu C, Xu Y, Minari T, Darmawan P, Tsukagoshi K (2012) Solution-processed organic crystals for field-effect transistor arrays with smooth semiconductor/dielectric interface on paper substrates. Org Electron 13(5):815–819. https://doi.org/10.1016/j.orgel.2012.01.021

Liu J, Chen S, Qian D, Gautam B, Yang G, Zhao J, Bergqvist J, Zhang F, Ma W, Ade H, Inganäs O, Ade H (2016) Fast charge separation in a non-fullerene organic solar cell with a small driving force. Nat Energy 1(7):1–7. https://doi.org/10.1038/nenergy.2016.89

Liu K, Ouyang B, Guo X, Guo Y, Liu Y (2022) Advances in flexible organic field-effect transistors and their applications for flexible electronics. NPJ Flexible Electron 6(1). https://doi.org/10.1038/s41528-022-00133-3

Liu Q, Wang Y, Arunagiri L, Khatib M, Manzhos S, Feron K, Bottle SE, Haick H, Yan H, Michinobu T, Sonar P (2020) Versatile nature of anthanthrone based polymers as active multifunctional semiconductors for various organic electronic devices. Adv Mater 1(9): 3428–3438. https://doi.org/10.1039/d0ma00728e

Ma S, Kumaresan Y, Dahiya AS, Lorenzelli L, Dahiya R (2022) Flexible tactile sensors using AlN and MOSFETs based ultra-thin chips. IEEE Sens J. https://doi.org/10.1109/JSEN.2022.3140651

Maddirala G, Searle T, Wang X, Alici G, Sencadas V (2022) Multifunctional skin-compliant wearable sensors for monitoring human condition applications. Appl Mater Today 26, Article 101361. https://doi.org/10.1016/j.apmt.2021.101361

Mannsfeld SCB, Tee BCK, Stoltenberg RM, Chen CVHH, Barman S, Muir BVO, Sokolov AN, Reese C, Bao Z (2010) Highly sensitive flexible pressure sensors with microstructured rubber dielectric layers. Nat Mater 9(10): 859–864.https://doi.org/10.1038/nmat2834

Matsumoto T, Ou-Yang W, Miyake K, Uemura T, Takeya J (2013) Study of contact resistance of high-mobility organic transistors through comparisons. Org Electron 14(10):2590–2595. https://doi.org/10.1016/j.orgel.2013.06.032

Meager I, Nikolka M, Schroeder BC, Nielsen CB, Planells M, Bronstein H, Rumer JW, James DI, Ashraf RS, Sadhanala A, Hayoz P Sadhanala A (2014) Thieno [3, 2-b] thiophene flanked isoindigo polymers for high performance ambipolar OFET applications. Adv Funct Mater 24(45):7109–7115. https://doi.org/10.1002/adfm.201402307

Nair NM, Jahanara MM, Ray D, Swaminathan P (2021a) Photoresponse of a printed transparent silver nanowire-zinc oxide nanocomposite. Flexible Printed Electron 6(4), Article 045004. https://doi.org/10.1088/2058-8585/ac2e48

Nair NM, Khanra I, Ray D, Swaminathan P (2021b) Silver nanowire-based printable electrothermochromic ink for flexible touch-display applications. ACS Appl Mater Interfaces 13(29):34550–34560. https://doi.org/10.1021/acsami.1c09115

Natali D, Chen J, Maddalena F, García Ferré F, Di Fonzo F, Caironi M (2016) Injection length in staggered organic thin film transistors: assessment and implications for device downscaling. Adv Electron Mater 2(8), Article 1600097. https://doi.org/10.1002/aelm.201600097

Ng TN, Wong WS, Chabinyc ML, Sambandan S, Street RA (2008) Flexible image sensor array with bulk heterojunction organic photodiode. Appl Phys Lett 92(21):191. https://doi.org/10.1063/1.2937018

Okaniwa H, Nakatani K, Asano M, Yano M, Hamakawa Y (1982a) Production and properties of a-Si: H solar cell on organic polymer film substrate. In: 16th Photovoltaic specialists conference

Okaniwa H, Nakatani K, Yano M, Asano M, Suzuki KJJJOAP (1982b) Preparation and properties of a-Si: H solar cells on organic polymer film substrate 21(S2): 239

Ortiz RP, Facchetti A, Marks TJ (2010) High-k organic, inorganic, and hybrid dielectrics for low-voltage organic field-effect transistors. Chem Rev 110(1):205–239. https://doi.org/10.1021/cr9001275

Park SB, Han JW, Kim JH, Wibowo AF, Prameswati A, Park J, Lee J, Moon MW, Kim MS, Kim YH (2021) Multifunctional stretchable organic–inorganic hybrid electronics with transparent conductive silver nanowire/biopolymer hybrid films. Adv Opt Mater 9(14). https://doi.org/10.1002/adom.202002041

Petritz A, Wolfberger A, Fian A, Irimia-Vladu M, Haase A, Gold H, Rothländer T, Grießer T, Stadlober B (2013) Cellulose as biodegradable high-k dielectric layer in organic complementary inverters. Appl Phys Lett 103(15). https://doi.org/10.1063/1.4824701

Raghuwanshi V, Bharti D, Mahato AK, Varun I, Tiwari SP (2019) Solution-processed organic field-effect transistors with high performance and stability on paper substrates. ACS Appl Mater Interfaces 11(8):8357–8364. https://doi.org/10.1021/acsami.8b21404

Rahi S, Raghuwanshi V, Saxena P, Konwar G, Tiwari SP (2022) High performance flexible transistors with polyelectrolyte/polymer bilayer dielectric. IEEE Trans Electron Dev 1–7. https://doi.org/10.1109/ted.2022.3148700

Rahmanudin A, Tate DJ, Marcial-Hernandez R, Bull N, Garlapati SK, Zamhuri A, Khan RU, Faraji S, Gollu SR, Persaud KC, Turner ML (2020) Robust high-capacitance polymer gate dielectrics for stable low-voltage organic field-effect transistor sensors. Adv Electron Mater 6(3). https://doi.org/10.1002/aelm.201901127

Rauch T, Böberl M, Tedde SF, Fürst J, Kovalenko MV, Hesser G, Lemmer U, Heiss W, Hayden O (2009) Near-infrared imaging with quantum-dot-sensitized organic photodiodes. Nat Photon 3(6):332. https://doi.org/10.1038/nphoton.2009.72

Ray KA (1967) Flexible solar cell arrays for increased space power. IEEE Trans Aerospace Electron Syst, AES 3(1): 107–115. https://doi.org/10.1109/TAES.1967.5408720

Shang X, Song I, Jung GY, Choi W, Ohtsu H, Lee JH, Ahn J, Koo JY, Kawano M, Kwak SK, Oh JH (2021) Micro-/nano-sized multifunctional heterochiral metal–organic frameworks for high-performance visible–blind UV photodetectors. J Mater Chem C 9(23):7310–7318. https://doi.org/10.1039/D1TC01333E

Sharova AS, Melloni F, Lanzani G, Bettinger CJ, Caironi M (2021) Edible electronics: the vision and the challenge. Adv Mater Technol 6(2):2000757. https://doi.org/10.1002/admt.202000757

Shi W, Guo Y, Liu Y (2020) When flexible organic field-effect transistors meet biomimetics: a prospective view of the Internet of Things. Adv Mater 32(15):1901493. https://doi.org/10.1002/adma.201901493

Shirakawa H, Louis EJ, MacDiarmid AG, Chiang CK, Heeger AJ (1977) Synthesis of electrically conducting organic polymers: halogen derivatives of polyacetylene,(CH) x. J Chem Soc, Chem Commun (16):578–580. https://doi.org/10.1039/C39770000578

Singh R, Singh E, Nalwa HS (2017) Inkjet printed nanomaterial based flexible radio frequency identification (RFID) tag sensors for the internet of nano things. RSC Adv 7(77):48597–48630. https://doi.org/10.1039/c7ra07191d

Sirringhaus H (2014) 25th anniversary article: organic field-effect transistors: the path beyond amorphous silicon. Adv Mater 26(9):1319–1335. https://doi.org/10.1002/adma.201304346

Stoliar P, Kshirsagar R, Massi M, Annibale P, Albonetti C, de Leeuw DM, Biscarini F (2007) Charge injection across self-assembly monolayers in organic field-effect transistors: odd−even effects. J Am Chem Soc 129(20):6477–6484. https://doi.org/10.1021/ja069235m

Terse-Thakoor T, Punjiya M, Matharu Z, Lyu B, Ahmad M, Giles GE, Owyeung R, Alaimo F, Shojaei Baghini M, Brunyé TT, Sonkusale S (2020) Thread-based multiplexed sensor patch for real-time sweat monitoring. NPJ Flexible Electron 4(1), Article 18. https://doi.org/10.1038/s41528-020-00081-w

Tiwari SP, Potscavage WJ, Sajoto T, Barlow S, Marder SR, Kippelen B (2010) Pentacene organic field-effect transistors with doped electrode-semiconductor contacts. Organ Electron 11(5): 860–863. https://doi.org/10.1016/j.orgel.2010.01.029

Tiwari SP, Zhang XH, Potscavage WJ, Kippelen B (2009) Study of electrical performance and stability of solution-processed n -channel organic field-effect transistors. J Appl Phys 106(5), Article 054504. https://doi.org/10.1063/1.3204655

Tsai L-S, Hwang J-C, Lee C-Y, Lin Y-T, Tsai C-L, Chang T-H, Chueh YL, Meng H-F (2013) Solution-based silk fibroin dielectric in n-type C60 organic field-effect transistors: mobility enhancement by the pentacene interlayer. Appl Phys Lett 103(23): 233304. https://doi.org/10.1063/1.4841595

Tsumura A, Koezuka H, Ando T (1988) Polythiophene field-effect transistor: its characteristics and operation mechanism. Synth Met 25(1):11–23. https://doi.org/10.1016/0379-6779(88)90318-9

Wang C-H, Hsieh C-Y, Hwang J-C (2011) Flexible organic thin-film transistors with silk fibroin as the gate dielectric. Adv Mater 23(14):1630–1634. https://doi.org/10.1002/adma.201004071

Wang CY, Fuentes-Hernandez C, Chou WF, Kippelen B (2017) Top-gate organic field-effect transistors fabricated on paper with high operational stability. Org Electron 41:340–344. https://doi.org/10.1016/j.orgel.2016.11.026

Widdascheck F, Bischof D, Witte G (2021) Engineering of printable and air-stable silver electrodes with high work function using contact primer layer: from organometallic interphases to sharp interfaces. Adv Funct Mater 31(49), Article 2106687. https://doi.org/10.1002/adfm.202106687

Youn J, Dholakia GR, Huang H, Hennek JW, Facchetti A, Marks TJ (2012) Influence of thiol self-assembled monolayer processing on bottom-contact thin-film transistors based on n-type organic semiconductors. Adv Func Mater 22(9):1856–1869. https://doi.org/10.1002/adfm.201102312

Yu KJ, Yan Z, Han M, Rogers JA (2017) Inorganic semiconducting materials for flexible and stretchable electronics. NPJ Flexible Electron 1(1), Article 4. https://doi.org/10.1038/s41528-017-0003-z

Zhao Y, Zhou W, Shi Y, Yang X, Bai Y, Li L, Feng S, Zhang T (2022) Superelastic alloy based electrical interconnects for highly stretchable electronics. NPJ Flexible Electron 6(1). https://doi.org/10.1038/s41528-022-00142-2

Zhu B, Wang H, Leow WR, Cai Y, Loh XJ, Han M-Y, Chen X (2016) Silk fibroin for flexible electronic devices. Adv Mater 28(22):4250–4265. https://doi.org/10.1002/adma.201504276

Chapter 12
Stretchable Interconnects: Materials, Geometry, Fabrication, and Applications

Vivek Kumar, Malvika, Yash Agrawal, and Kavicharan Mummaneni

Abstract In recent days, stretchable devices gained utmost importance due to their stretchable, bendable, flexible characteristics and also their comfort toward human skin. Stretchable materials are undeniable compared to conventional materials due to its ability to adjust to the curvilinear surface of human skin. When stress or strain is applied it maintains its electrical functionality which is useful for stretchable electronics devices. Stretchable interconnect plays a vital role in these modern stretchable devices by providing an extra degree of stretchability. This survey chapter presents an overview of stretchable interconnect geometries, fabrication techniques, and materials used to fabricate substrate materials. In addition, a wide range of applications are discussed such as strain sensor, temperature sensor, and energy harvesters. The key developments of stretchable interconnects have been highlighted in the domains of healthcare monitoring and energy storage devices.

Keywords Stretchable interconnects · Strain sensor · PDMS · Microfabrication · Substrate materials · e-skin

V. Kumar (✉) · Malvika · K. Mummaneni
Electronics and Communication Enggineering Department, National Institute of Technology, Silchar, Assam, India
e-mail: mail.kumar.vivek@gmail.com

K. Mummaneni
e-mail: kavicharan@ece.nits.ac.in

Y. Agrawal
VLSI and Embedded Systems Research Group, Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, Gujarat 382 007, India
e-mail: yash_agrawal@daiict.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_12

Introduction

In recent days, stretchable electronics gained utmost importance due to their advantages like flexibility and comfortableness. Stretchable electronics has various applications in the fields like medical and consumer applications and provides multiple features like bending, twisting, folding, and stretching. In general, electronic devices cannot support stretching applications due to the components used in them are rigid. There are two ways to make the devices stretchable: (i) by utilizing the multilayer stretchable components (stretchable conductor, stretchable semiconductor, and stretchable dielectrics), also known as the fully stretchable system, and (ii) by partitioning the device into simple blocks termed as rigid islands and joining them with the stretchable interconnects known as the semi-stretchable system. With the help of stretchable interconnects, if any strain is applied to the system, the systems expand without any stress on rigid blocks and any breaks in the system (Dahiya and Valle 2013) (Fig. 12.1).

While designing the semi-stretchable devices, the whole system can be embedded on a stretchable and elastic substrate to give mechanical stability to the system from the outside environment when it is subjected to stretching. In general, traditional interconnects like copper, silver, and gold had limited stretchability (Lacour et al. 2005; Hess-Dunning et al. 2013) since their Young's modulus is in the range of a few hundred Giga pascal. However, stretchable materials like polyaniline (PANi) (Stoyanov et al. 2013), poly(3,4-ethylene dioxythiophene): poly(4-styrene sulfonate) (PEDOT: PSS) (Lipomi et al. 2012), and PDMS (Jeong et al. 2016) (Polydimethylsiloxane) exhibit Young's modulus below one mega pascal's. To design stretchable interconnects, various factors like selection of suitable substrate material, the geometry of the interconnect structure and materials, etc., have to be taken into consideration (Plovie et al. n.d.; Embedded and in Thermoplastic Polymers 2017; Jinno et al. n.d.; Takimiya and Someya 2017; Wang et al. 2018; Johnston et al. 2014; Blau et al. 2011; Salvatore et al. 2017; Tybrandt et al. 2018). Such design can be realized by computer-aided tools with minimum design cost and time. Therefore, a precise flexible-electronic computer-simulation environment is presented in Fig. 12.2. In contrast with rigid electronic-system design, the flexible-electronic design platform must co-simulate both mechanical as well as electronic aspects to consider the design of stretchable interconnects, stretchable electronic circuits, and stretchable circuit-layout wiring.

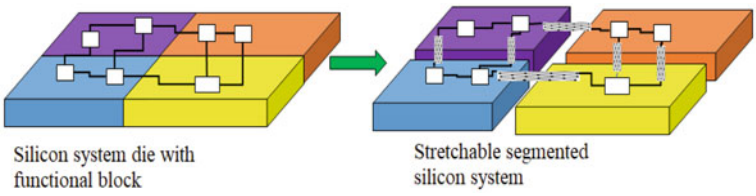


Fig. 12.1 Silicon die with functional blocks (left), Stretchable segmented silicon system (right)

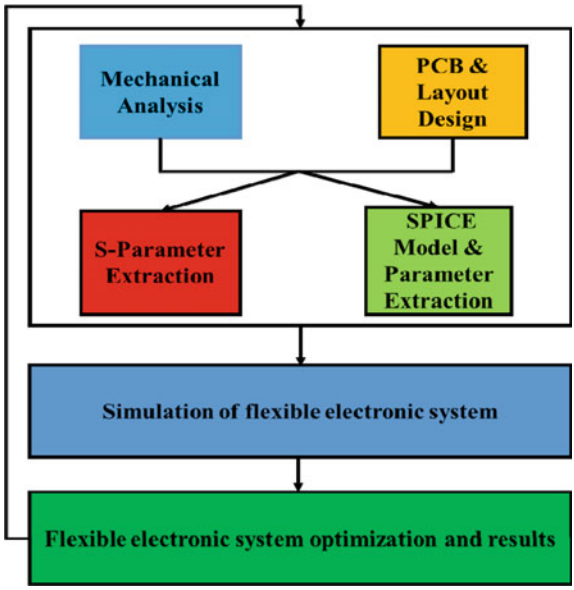


Fig. 12.2 Stretchable-interconnect research design flowchart

Substrate Materials

The electronic systems that are stretched using elastic polymers commonly use various polymers such as polyimide, polyester, and Parylene as substrate materials. The most widely used substrate material is PDMS, which is also known as polydimethylsiloxane. In addition, PBAT (Polybutylene adipate terephthalate) also known as Ecoflex is commonly used stretchable substrate. PBAT (Carlson et al. 2012) had compostable features and is used in biodegradable electronics. Stretchability of PBAT is more than PDMS material, PBAT can be stretched up to 650%, whereas PDMS has 150% stretchability (Dong et al. 2018). Although polyurethane has good mechanical properties, it is not compatible with many organic solvents such as ethanol. In addition, its properties are not as strong as those of other polyurethanes.

Parylene can be used in the form of a liquid or solid base material for wearable electronic applications. It is a highly versatile material that can be used in various industries such as textiles. This material is mainly used for the design of flexible printed circuit board. Its high-quality and robust properties make it a good choice for the various applications when compared to other polymers the stretchability of parylene is less. Polyimide is the popularly used material for design of flexible PCB substrate (Vroman and Tighzert 2009). It is a high strength material and its Young's modulus lies in the range of all the polymeric substrates only. A comparison between the various substrate materials is shown in Table 12.1. Among the various substrate materials, the polydimethylsiloxane (PDMS) has more advantages than the other polymers. This is mainly due to its chemical stability and biocompatible properties.

Table 12.1 Comparison of typical parameters of various substrate materials

Substrate materials	PDMS	PBAT (Ecoflex)	Polyurethane	SEBS	Parylene	Polyimide	Polyethylene naphthalate	Polyethylene terephthalate
Young's modulus	1–7 MPa Dang et al. (2017a, b)	22 kPa Case et al. (2015)	1.1 MPa Case et al.(2015)	10–100 MPa Zhai and McKenna (2014), Koerner et al. (2005)	1.3–3.5 GPa Weiss et al. n.d.)	2.84 GPa Metzen and Stieglitz (2013)	5 GPa (Chang et al. (2015)	2–2.7 GPa Kahouli et al. (2009)
Stretchability	150% Huang et al. (2014)	650% Huang et al. (2014)	200% Du et al. (2011)	700% Zhai and McKenna (2014), Koerner et al. (2005)	18% [67]	19% Metzen and Stieglitz (2013)	145% Chang et al. (2015)	150% Kahouli et al. (2009)
Dielectric constant (at 1 kHz)	2.32–2.4 Jeong et al. (2016)	2.1 Case et al. (2015)	5.68 Case et al. (2015)	2.45 Koerner et al. 2005)	3.1 Grigorescu et al. (2016)	3.1–3.4 Ranjan and Mertig (2008)	3.0 Chang et al. (2015)	3.4 Kahouli et al. (2009)

Fig. 12.3 Mesh-shaped interconnect geometry

Geometry of Stretchable Interconnect Structures

Interconnects can be made stretchable by making the substrate and wiring stretchable. In stretchable interconnects nanomaterials and composite materials are being used as conductive materials compared to conventional bulk metals, such as silver and copper, whose elastic deformation is bounded to very small range. To use these bulk metals, patterning strategies are being utilized (2011). There are several types of stretchable interconnect geometries such as mesh geometry, spiral geometry, meander geometry, and horseshoe geometry.

Mesh Geometry

The mesh geometry consists of a series of metal beams that are in a repeated pattern. The geometries may be different such as rectangular, diamond, or honeycomb, but the effect of applied tensile strains across it remains the same. The shape of the element changes as strains are applied. Strain resulting in plane is accommodated by rotation in plane. These types of interconnects offer advantages such as it maintains its electrical conductivity, if some parts are damaged. Nevertheless, the number of interconnects is limited due to its size in a given area. This type of interconnects best suits for unsupported interconnects compared to supporting silicon since the maximum elongation that can be achieved using metals is reduced. Therefore, it offers advantage of its strong nature, but it also has disadvantage with its size (Abdelhalim et al. 2011) (Fig. 12.3).

Spiral Geometry

A spiral geometry consists of two parallel spiral planes that are connected at the center as shown in Fig. 12.4. When the system gets stretched, the two spiral planes move in opposite directions and support the system's stretching (Kaiju et al. 2010; Kolahchi et al. 2015). Island is connected with spiral springs or at the center of the two spirals. It can also be fabricated as in the case of the mesh geometry. The island acts as a connector and holds the functionality between two spirals and a significant amount of stretchability is provided by the spiral interconnect during the deformation and

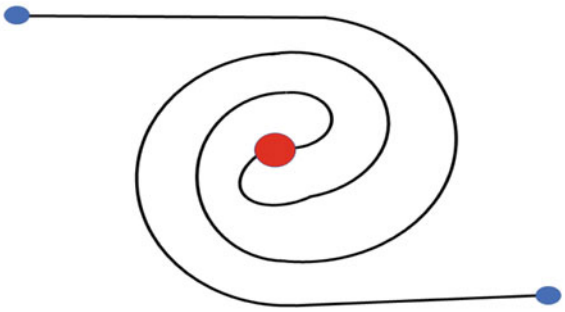


Fig. 12.4 Spiral-shaped interconnect geometry

stretching. Therefore, these islands work as buffer and stretchability is maintained. To unfurl the structure the spiral geometry rotates. The unfurling capacity is restricted if patterning strategies are used. As a supporting silicon structure, the arms width must be small to uncoil it to a straight line (Abdelhalim et al. 2011).

Meander Geometry

Meander-shaped geometry consists of series of half-circles and vertical segments, as shown in Fig. 12.5. There is a curvature between the curved and vertical segments. These vertical segments allow more significant deformations and also reduce the stress concentration on the geometry.

Meander-shaped interconnect geometry can be stretched into a straight line if its material is ductile; thus, it gives higher stretchability. The disadvantage of this structure is that it occupies more area. Meander geometry is also commonly referred to as self-serpentine structure. Figure 12.6 shows the second-order serpentine design (Abdelhalim et al. 2011; Ri Yim and Park et al. 2021), and it provides large extensions. These interconnects gained utmost importance because of their scalability.

Fig. 12.5 Meander-shaped interconnect geometry

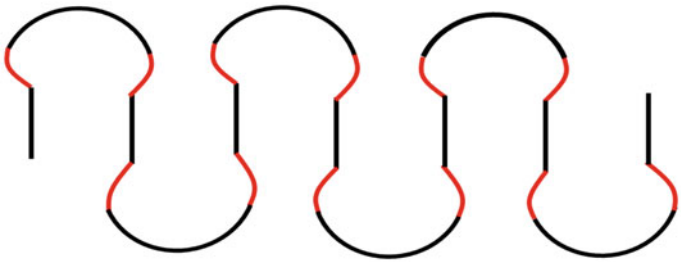


Fig. 12.6 Second-order meander-shaped interconnect geometry

Horseshoe Geometry

Horseshoe-shaped geometry shown in Fig. 12.7 is the most widely used stretchable interconnect geometry, which supports sizeable longitudinal deformation due to the angle of the pivot segment (Abdelhalim et al. 2011). It comprises a series of half-circles, which reduces the stress concentration and allow more significant deformations. Geometry-wise it is similar to meander-shaped geometry except there are no vertical segments. Compared to meander geometry, horseshoe interconnect geometry occupies less area.

Fabrication Techniques

There are several techniques for fabrication of stretchable interconnects, stretchable substrates, and each process has its own advantages and disadvantages (Vroman and Tighzert 2009). The selection of fabricating material plays an important role to achieve multiple features like bending, twisting, folding, and stretching, and the discussion on choosing various materials is already presented in the previous section. Based on the resolution requirement, processed material, and the material used for the fabrication of the substrate, a unique fabrication technique can be chosen. Sometimes two or more techniques are also used in combination to get optimized structure and

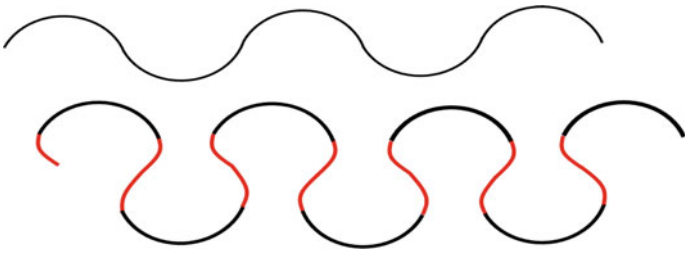


Fig. 12.7 Horseshoe-shaped interconnect geometry

performance (Gonzalez et al. 2008, 2018; Qu et al. 2018; Qin et al. 2006; García Núñez et al. 2018; Núñez et al. 2018; Fan et al. 2008; Abdelhalim et al. 2013; Béduer et al. 2012).

Microfabrication

Figure 12.8a depicts the Microfabrication process. It is an industrial-level process, and it has advantages like scalability and compatibility. It consists of several steps, viz. cleaning, spin coating, baking, ion-implantation, deposition, photolithography, and etching, etc. For the fabrication of stretchable interconnect, the significant steps involved are spin coating, photolithography, and etching. Microfabrication technology has several limitations due to unavoidable solvents, acid, and base solutions. These materials react with substrate materials and affect the quality and resolution of the interconnect structures. This technology is advantageous in terms of its scalability and compatibility with contemporary silicon device fabrication technique. Nevertheless, the scope of this technology is the limited material availability, compatible only to planar substrates and unavoidable to solvent, acid, and base solutions. Owing to these restrictions, many researchers merged this method with other printing technologies for the fabrication of the stretchable electronic systems (Vroman and Tighzert 2009).

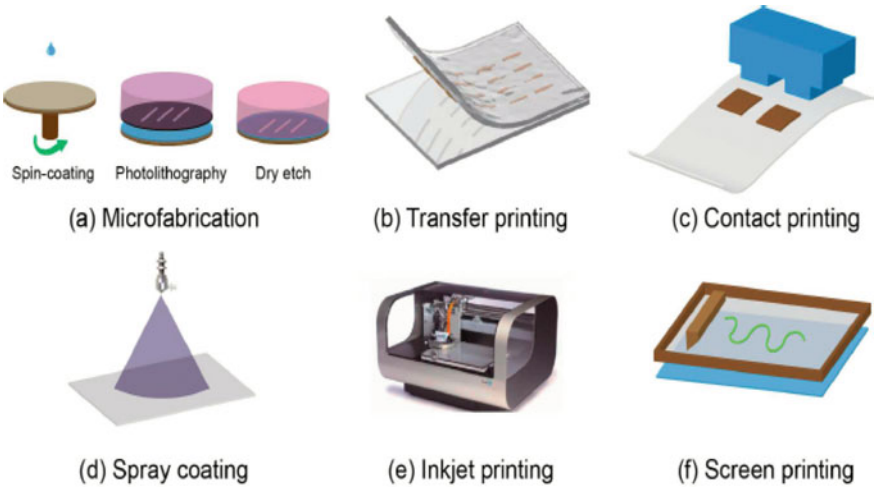


Fig. 12.8 Interconnect fabrication techniques **a** Microfabrication **b** Transfer printing **c** Contact/Stamp printing **d** Spray Printing **e** Inkjet printing **f** Screen Printing (Vroman and Tighzert 2009)

Transfer Printing Technology

Figure 12.8b depicts the Transfer printing process. It consists of two substrates (I) Donor substrate (II) Receiver substrate. This technique describes the process in which the donor substrate and the receiver substrate comes in contact with each other. It eliminates the need for the use of a separate substrate for fabrication. High temperatures and organic solvents are often used to produce nanowires. These are typically fabricated on a silicon substrate and then printed on a desired flexible substrate. Applications of this printing include transfer and alignment of the nano-structures (e.g., nanowires, nanoribbons, etc.) from the donor substrate to the receiver (e.g., polymeric) substrate. The transfer process, which is shown in Fig. 12.8b, can avoid incompatibility issues arising in polymeric substrate to some fabrication processes. This printing technology also assists to align the nanowires in a particular direction to the receiver substrate.

Contact Printing

Figure 12.8c depicts the process of contact printing. It is a type of printing technology, which consists of a pre-fabricated stamp that contact with a specific substrate. Numerous fabrication technologies such as photolithography, micromachining, and nano-imprinting are used to realize the stamp (Vroman and Tighzert 2009). This process is usually done by printing a film using conductive ink. The film's resolution depends on several factors, such as the ink's viscosity and the substrate's surface energy. The range of resolution varies from nano-meter to micro-meter. This technology helps the ink bond more strongly to the substrate. More the surface energy more strongly is the bond. It is also used to develop higher resolution printing.

Spray Coating Technology

Figure 12.8d depicts the process of spray coating. This technology is very famous because of high efficiency. It is a widely used technique to deposit various materials on a substrate. It works by using air pressure and an electrical field to spray the materials through a nozzle. Normally it is used for whole substrate coating along with other techniques of printing, viz. transfer printing or contact printing. Despite its high efficiency, it has a poor resolution (Vroman and Tighzert 2009).

Inkjet Printing Technology

By using this printing technology, more versatile designs can be made. The process is depicted in Fig. 12.8e. The required geometry format is imported into the software in the form of graphical designs. This is an advanced version of conventional ink printer.

Moreover, it is similar to the technology based on dot-matrix printer, which comprises the patterns of tiny dots. An electronic circuit activates the piezoelectrical crystals, which drive these dots. The resolution of the inkjet printing depends on the ink and the diameter of the nozzle. In fabricating LEDs and solar cells (Vroman and Tighzert 2009; Khan et al. 2016), inkjet printing technology is the most widely used.

Screen Printing Technology

Figure 12.8f depicts Screen printing technology. It is a simple printing technology that uses a stencil over which predefined patterns can be available. These patterns allow the manufacturing material to pass through it. The major disadvantage of this printing technology is its resolution, and also the quality of the printing, which depends on various factors like solution viscosity, type of the substrate, angle, and the distance between stencil and substrate, etc. (Khan et al. 2015a, b; Singh et al. 2010; Dang et al. 2015). The highest resolution of this printing technology is about 50 μm (Table 12.2).

Table 12.2 Comparison of various printing technologies

	Spray coating	Screen printing	Inkjet printing
Thickness of printed film (μm)	0.012–50 (Verilhac et al. 2010; Gans and Schubert 2003; Pham et al. 2005)	>0.5 (industry) (Zhao et al. 2019) 0.109 ± 0.004	5–10 (Rogers 2014)
Viscosity of inks (mPa.s)	5–440 (Verilhac et al. 2010)	500–50,000 (Girotto et al. 2009)	1–40 (Girotto et al. 2009)
Resolution (μm)	–	50–500 (Singh et al. 2010)	20–50 (Rogers 2014; Zhao et al. 2019)
Accuracy (μm)	–	±25 (Zhao et al. 2019)	±0.5–25 (Moonen et al. 2012)

Applications of Stretchable Interconnects

The various applications of stretchable interconnects are shown in Fig. 12.9 (Vroman and Tighzert 2009). Due to their flexibility, these devices can be used in diverse applications such as sensors and electronic systems. A stretchable interconnect-based smart prosthetic hand which is equipped with artificial skin can detect various signals such as temperature, humidity, and pressure. This makes it an ideal tool for medical procedures.

For instance, a balloon catheter, which is used in surgery to remove blood blockage, requires around 130% stretching when inflated. In Fig. 12.9, a complete stretchable device is shown that allows users to analyze biofluids like sweat. It can also be stretched using the device's flexible interface. For consumer electronic applications, the flexible displays used in these devices are connected to ILEDs and non-coplanar stretchable connectors. The data transmission and amplifier circuits are used to create a flexible network which collects the data. This network is robust and can be used against various types of movement. In addition to the collection of data, this network can also be used to detect and transmit weak electrophysiological signals.

The stretchable interconnects can also be utilized for the electronic skin (e-skin) for robotics since it has been vastly adopted in many fields, like precise manufacturing and agriculture. Although the robotics are programmed and accurately controlled through computers with precise and infallible programming instructions yet they lack the real-time sensing (e.g., proximity or chemical sensing) from the working environment, restricting them to provide exact dynamic response. Hence,

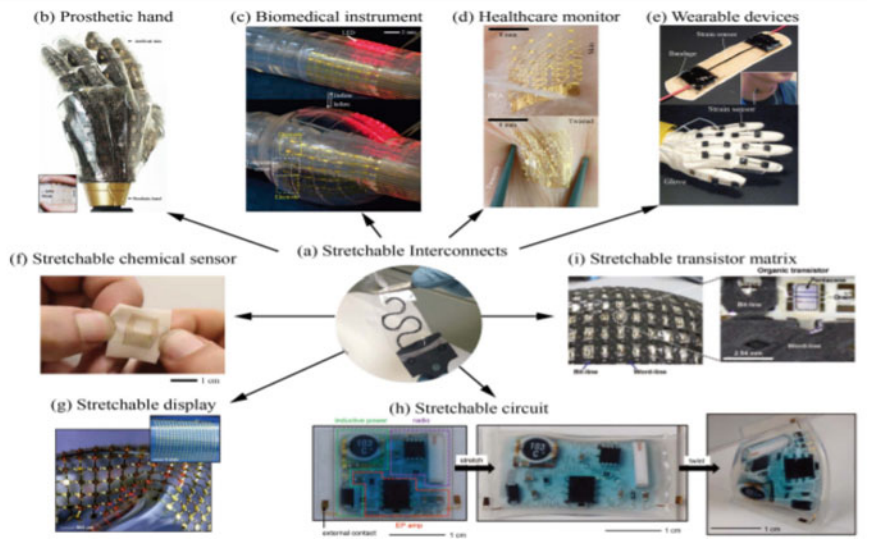


Fig. 12.9 Various applications of stretchable interconnects (b)–(h) (Vroman and Tighzert 2009)

to get exact sensing information, the idea of "e-skin" was proposed, as it can withstand up to certain deformation and perform some complex movements in contrast to conventional rigid electronics, to integrate sensors with robotics. Healthcare monitoring system is one of the key areas which is gaining more attention from stretchable electronics point of view owing to propinquity between human body and electronic devices. The accuracy and faithfulness of the non-invasive physiological parameters, viz. blood pressure, respiration rate, electrolyte concentration, heart rate, and skin temperature of the body, are more with stretchable devices. As it is easy to use, these non-invasive health monitoring systems gaining more popularity, especially in chronic illness (Vroman and Tighzert 2009). The electrical feedback provided by stretchable interconnects helps to improve the performance of a medical device by giving surgeons feedback on the quality of the operation. For instance, if a temperature sensor is placed on the human skin, it can collect data about the temperature within 24 h. Electrodes for precise monitoring and surgery are developed for use in stretchable neural electrodes. These are also being used in wearable electronics, such as motion detection devices. Because of their flexible nature, these devices can be integrated into various types of clothing, such as gloves and bandages.

The rapid development of stretchable interconnects has led to the creation of wearable and skin-mountable electronic devices. One of the most important features of these devices is their ability to monitor the environment and physical status. Strain sensors made with these devices and interconnects can be used to monitor various environmental and chemical conditions. A strain caused by the AgNWs-elastomer can lead to the disconnection of the connecting wires between the different types of conductors. This resulted in an increase in the resistivity of the sensor. Due to the high stretchability of the 1D conductive nanofillers, this type of sensor can be used for various applications. One of the most important factors that can be considered when it comes to the performance of a strain sensor is its linearity. Unlike other types of strain sensors, which have poor linearity, the type of strain sensor known as the capacitance type exhibits good linearity. This type of sensor is made up of a layer of flexible dielectric which is sandwiched between two stretchable electrodes. The device was able to detect various human motions, such as blinking and clenching, which are commonly used in medical procedures. Its sensitivity is also enough to identify target motion. This technology can be useful in monitoring the blood flow rate in order to check the condition of patients with high blood pressure. Another performance parameter that can be measured is the increased sensitivity. One of the strategies that can be used to improve the sensitivity of the device Is by using low-density networks of conductive materials. This can help to minimize the strain on the junction points of the devices. Another strategy is to treat nanostructured materials by exposing them to intense light, which increases the gauge factor. Aside from the physical properties of the device, other factors such as its biocompatibility and low power consumption are also taken into account to achieve high performance.

Energy harvesting devices can be developed using stretchable interconnects. The creation of energy harvesting technology is a progress in the direction of eco-friendly society. Researchers have spent the last ten years working on energy harvesting devices, which are developing with stretchable electronics. Many flexible energy

harvesters function with a low strain ratio (0.1%) before adopting stretchable conductors. Therefore, by utilizing the various techniques, materials, structural arrangement, integration strategy, different stretchable and flexible energy harvesting devices can be investigated. Energy harvesting devices have been designed according to their unique promising applications, primarily in wearable and stretchable electronics. These devices primarily use triboelectricity and piezoelectricity, commonly known as nanogenerators, to transform mechanical energy to electrical energy. Moreover, shape-adaptive triboelectric nanogenerator (saTENG) has numerous applications in day-to-day life.

Figure 12.10a–c depicts the various applications of saTENG as energy harvester from human walking where foot motion being monitored. As shown in Fig. 12.10a, a saTENG has been mounted beneath shoe pad and for electrolyte, tap water (conductivity: 0.248 mS/cm) is used. Human body is used as ground and tap water being used as electrolyte for wearable saTENGs. The foot motion, shown in Fig. 12.10b, is visualized by two sets of LEDs having reversed polarity. The up and down motions of foot are reflected in the LEDs sets lighting alternatively. For the up motion of foot, the electric potential of water electrode decreases and the electrons flow from water to ground causing upper set of LEDs to glow. In downward pressing of foot reverse happens, i.e., electrons flow from ground to water electrode and hence lower set of LEDs glow which is shown in Fig. 12.10c. In the field of robotics this type of saTENG plays a crucial role. Figure 12.10d shows a saTENG looped over a subject's arm to act as a self-powered arm motion sensor and collect energy from pressing action. For a cycle of tapping, more than 80 LEDs can be driven demonstrating its effectiveness. The generated electricity through this saTENG can be rectified and stored in batteries or capacitors as shown in Fig. 12.10e, i.e., capacitor voltage rises quickly while tapping and becomes zero when tapping ceases. Figure 12.10f shows the application of saTENG in the shape of bracelet as wearable power source. The bracelet is worn around arm and it detects the biceps bending angle. With the bending of elbow, the biceps volume increases, and vice versa, thereby changing arm circumference. Due to this, the contact area between arm and saTENG changes. This alteration of contact area causes electrons to flow between water electrode of saTENG and ground and again between ground and water electrode. The rate of transfer of electrons and hence current is proportional to bending angle which is depicted in Fig. 12.10f (Susanna et al. 2011).

It was reported that, a hyper-stretchable elastic-composite generator (SEG) made of extremely long Ag nanowires (VAgNWs), stretchable electrodes had a great stretchability and the output was seven times more than that of piezo-generators (Tobjörk and Österbacka 2011). The fabrication of a shape-adaptive triboelectric nanogenerator (saTENG) comprising conductive liquid enclosed in a polymer cover was developed as a scaling technique for extremely flexible and stretchable energy harvesters and self-powered sensors. It is also possible to design a stretchy energy harvester by combining nanomaterials. IS-TENG exhibits high transparency, extremely long uniaxial strain, and good self-healing capability that can regain its performance after 300 full bifurcation cycles. In addition, Piezoresistive sensors are

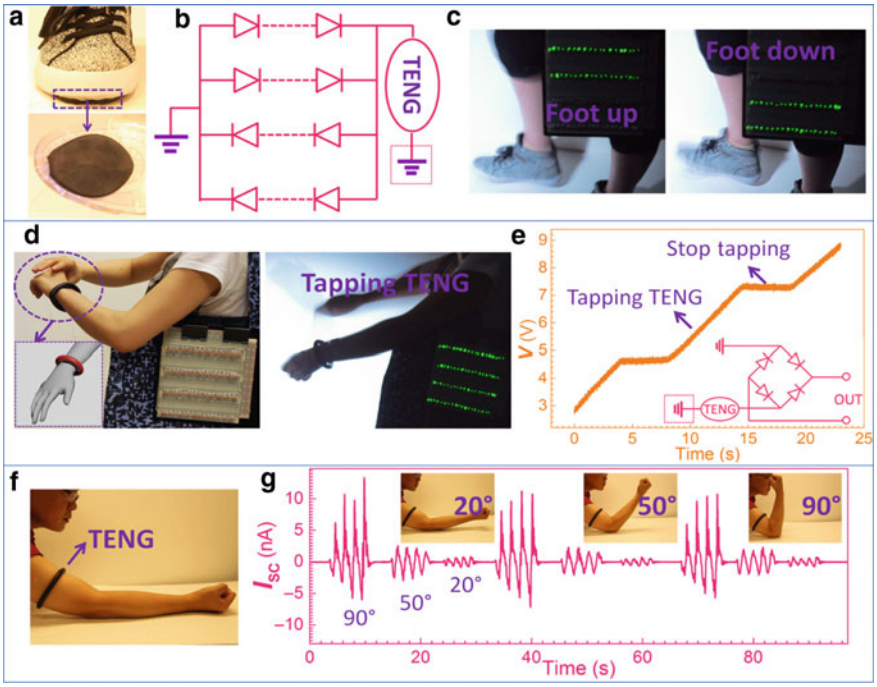


Fig. 12.10 Energy harvesting-based stretchable devices. **a** saTENG attached to shoe **b** Equivalent LED circuit diagram **c** LED motion illustration **d** Bracelet like saTENG demonstration **e** storing energy with bracelet saTENG **f** saTENG worn on arm **g** output of arm saTENG at various bend angle (Susanna et al. 2011)

frequently employed to measure the movement of specific joints or muscles in wireless wearable devices for human movement analysis (Yi et al. 2016; Corzo et al. 2020; Doo Ri Yim 2021; Zhao et al. 2022). Analyzing the gathered data precisely might be challenging if the connection resistance that transmits the sensor signal is likewise sensitive to human movement.

Conclusion

In this chapter, an overview of different interconnect substrate materials, geometric structures, and the fabrication techniques of stretchable interconnects for elastic electronic surface structures have been discussed. Various fabrication techniques for stretchable interconnects and its comparison also have been discussed. The applications of these stretchable electronics, viz. electronic skin, healthcare monitoring system, and energy storage for consumer have been discussed. The ability to conform to the shape of a human body, healthcare monitoring devices have been widely used

for this purpose to collect various physiological data. The wide range of applications that stretchable interconnects can play in electronics has shifted the paradigm in the health care and consumer electronics industry.

References

Abdelhalim A, Abdellah A, Scarpa G, Lugli P (2013) Fabrication of carbon nanotube thin films on flexible substrates by spray deposition and transfer printing. Carbon n. y. 61:72–79

Béduer A, Seichepine F, Flahaut E, Vieu C (2012) A simple and versatile micro contact printing method for generating carbon nanotubes patterns on various substrates. Microelectron Eng 97:301–305

Blau A, Murr A, Wolff S, Sernagor E, Medini P, Iurilli G, Ziegler C, Benfenati F (2011) Flexible, all-polymer microelectrode arrays for the capture of cardiac and neuronal signals. Biomaterials 32(7):1778–1786

Carlson A, Bowen AM, Huang Y, Nuzzo RG, Rogers JA (2012) Transfer printing techniques for materials assembly and micro/nanodevice fabrication. Adv Mater 24(39):5284–5318

Case JC, White EL, Kramer RK (2015) Soft material characterization for robotic applications. Soft Robot 2(2):80–87

Chang Y, Chiu T, Yang Y, Tseng Y, Chen X, Huang P, Rd C (2015) Fatigue response of polyimide thin film under cyclic loading. In: International microsystem, packaging, assembly and circuits technology conference, pp 233–236

Corzo D, Tostado-Blázquez G, Baran D (2020) Flexible electronics: status, challenges and opportunities. Front Electron 1:594003

Dahiya RS, Valle M (2013) Robotic tactile sensing—technologies and system. Springer, Dordrecht

Dang W, Vinciguerra V, Lorenzelli L, Dahiya R (2017a) Printable stretchable interconnects. Flexible Printed Electron 2(1):013003

Dang W, Khan S, Lorenzelli L, Vinciguerra V, Dahiya R (2015) Stretchable interconnects using screen printed nanocomposites of MWCNTs with PDMS and P(VDF-TrFE). Conference PhD Res. in Microelec. Elec. (PRIME), UK, pp 247–250

Dang W, Vinciguerra V, Lorenzelli L, Dahiya R (2017b) Printable stretchable interconnects. Flex Print Electron 2:013003

De Gans BJ, Schubert US (2003) Inkjet printing of polymer micro-arrays and libraries: instrumentation, requirements, and perspectives. Macromol Rapid Commun 24(11):659–666

Dong Z, Duan B, Li J, Yang Y (2018) A stretchable flexible electronic platform for mechanical and electrical collaborative design. Sci China Inf Sci 61 060418:1–060418:15

Du P, Lin X, Zhang X (2011) Dielectric constants of PDMS nanocomposites using conducting polymer nanowires. IEEE Transducers, 645–648

Fan Z, Ho JC, Jacobson ZA, Yerushalmi R, Alley RL, Razavi H, Javey A (2008) Wafer-scale assembly of highly ordered semiconductor nanowire arrays by contact printing. Nano Lett 8(1):20–25

García Núñez C, Liu F, Navaraj WT, Christou A, Shakthivel D, Dahiya R (2018) Heterogeneous integration of contact-printed semiconductor nanowires for high performance devices on large areas. Microsystems Nanoeng 4(1):22

Girotto C, Rand BP, Genoe J, Heremans P (2009) Exploring spray coating as a deposition technique for the fabrication of solution-processed solar cells. Sol Energy Mater Sol Cells 93(4):454–458

Gonzalez M, Axisa F, Vanden Bulcke M, Brosteaux D, Vandevelde B, Vanfleteren J (2008)Design of metal interconnects for stretchable electronic circuits. Microelectron Reliab 48(6):825–832

Grigorescu RM, Ciuprina F, Ghioca P, Ghiurea M, Iancu L, Spurcaciu B, Panaitescu DM (2016) Mechanical and dielectric properties of SEBS modified by graphite inclusion and composite interface. J Phys Chem Solids 89:97–106

Hess-Dunning AE, Tyler DJ, Zorman CA (2013) Stretchable thin-film metal structures on a stimuli-responsive polymer nanocomposite for mechanically-dynamic microsystems. In Transducers 2013, Barcelona, pp 2229–2232

Huang Y, Wang Y, Xiao L, Liu H, Dong W, Yin Z (2014) Microfluidic serpentine antennas with designed mechanical tunability. Lab Chip 14(21):4205–4212

Jeong SH, Zhang S, Hjort K, Hilborn J, Wu Z (2016) PDMS-based elastomer tuned soft, stretchable, and sticky for epidermal electronics. Adv Mater 28:5830–5836

Jinno H, Fukuda K, Xu X, Park S, Suzuki Y, Koizumi M, Yokota T, Osaka I, Takimiya K, Someya T (2017) Stretchable and waterproof elastomer-coated organic photovoltaics for washable electronic textile applications. Nat Energy 2(10):780–785

Johnston ID, McCluskey DK, Tan CKL, Tracey MC (2014) Mechanical characterization of bulk Sylgard 184 for microfluidics and microengineering. J Micromechan Microeng 24(3):0350117 (1–7)

Kahouli A, Sylvestre A, Ortega L, Jomni F, Yangui B, Maillard M, Berge B, Robert JC, Legrand J (2009) Structural and dielectric study of parylene C thin films. Appl Phys Lett 94(15):152901

Kaiju H, Basheer N, Kondo K, Ishibashi A (2010) Surface roughness and magnetic properties of Ni and Ni78Fe22 thin films on polyethylene naphthalate organic substrates. IEEE Trans Magn 46:1356–1359

Khan S, Yogeswaran N, Taube W, Lorenzelli L, Dahiya R (2015a) Flexible FETs using ultrathin Si microwires embedded in solution processed dielectric and metal layers. J Micromechan Microeng 25(12):125019

Khan S, Tinku S, Lorenzelli L, Dahiya RS (2015b) Flexible tactile sensors using screen-printed P(VDF-TrFE) and MWCNT/PDMS composites. IEEE Sens J 15(6):3146–3155

Khan S, Member LL, Dahiya R, Member S (2016) Flexible MISFET devices from transfer printed Si microwires and spray coating. J Electron Devices Soc. 4(4):189–196

Koerner H, Liu W, Alexander M, Mirau P, Dowty H, Vaia RA (2005) Deformation–morphology correlations in electrically conductive carbon nanotube— thermoplastic polyurethane nanocomposites. Polymer (Guildf) 46(12):4405–4420

Kolahchi AR, Ajji A, Carreau PJ (2015) Improvement of PET surface hydrophilicity and roughness through blending In: AIP Conference proceedings (ed. Sadhan C. Jana) 030001 (AIP Publishing, Cleveland, Ohio, USA, 2015)

Lacour SP, Jones J, Wagner S, Li T, Suo Z (2005) Stretchable interconnects for elastic electronic surfaces. Proc IEEE 93(8):1459–1467

Lipomi DJ, Lee JA, Vosgueritchian M, Tee BC-K, Bolander JA, Bao Z (2012) electronic properties of transparent conductive films of PEDOT: PSS on stretchable substrates. Chem Mater 24(2):373–382

Moonen PF, Yakimets I, Huskens J (2012) Fabrication of transistors on flexible substrates: from mass-printing to high-resolution alternative lithography strategies. Adv Mater 24(41):5526–5541

Núñez CG, Liu F, Xu S, Dahiya R (2018) Integration techniques for micro/nanostructure-based large-area electronics (Elements in Flexible and Large-Area Electronics). Cambridge University Press

Pham NP, Burghartz JN, Sarro PM (2005) Spray coating of photoresist for pattern transfer on high topography surfaces. J. Micromech Microeng 15(4):691–697

Plovie BB, Yang Y, Guillaume J, Dunphy S, Dhaenens K, Van Put S, Vervust T, Bossuyt F, Vanfleteren J, Arbitrarily shaped 2.5D circuits using stretchable interconnects embedded in thermoplastic polymers. Adv Eng Mater 19(8):1700032 (1–8)

Qin JJ, Chung TS, Cao Y (2006) Solvent selection for manufacture of fluorinated polyimide composite membranes. Desalination 193:8–13

Qu Y, Nguyen-Dang T, Page AG, Yan W, Das Gupta T, Rotaru GM, Rossi RM, Favrod VD, Bartolomei N, Sorin F (2018) Superelastic Multimaterial electronic and photonic fibers and devices via thermal drawing. Adv Mater 30(27):1707251 (1–8)

Ranjan N, Mertig M (2008) Dielectrophoretically assembled carbon nanotube-metal hybrid structures with reduced contact resistance. Phys Status Solidi Basic Res 245(10):2311–2314
Ri Yim D, Park CW (2021) Hybrid-type stretchable interconnects with double-layered liquid metal-on-polyimide serpentine structure. ETRI J 44:147–154
Rogers JA (2014) Materials for semiconductor devices that can bend, fold, twist, and stretch. MRS Buletinl 39:549–556
Salvatore GA, Sülzle J, Valle FD, Cantarella G, Robotti F, Jokic P, Knobelspies S, Daus A, Büthe L, Petti L, Kirchgessner N, Hopf R, Magno M, Tröster G (2017) Biodegradable and highly deformable temperature sensors for the internet of things. Adv Funct Mater 27(35):1702390 (1–10)
Singh M, Haverinen HM, Dhagat P, Jabbour GE (2010) Inkjet printing-process and its applications. Adv Mater 22(6):673–685
Sosin S (2011) Interconnect schemes for stretchable array-type microsystems. Mater Sci [Online: repository.tudelft.nl]
Stoyanov H, Kollosche M, Risse S, Waché R, Kofod G (2013) Soft conductive elastomer materials for stretchable electronics and voltage controlled artificial muscles. Adv Mater 25(4):578–583
Susanna G, Salamandra L, Brown TM, Di Carlo A, Brunetti F, Reale A (2011) Airbrush spray-coating of polymer bulk-heterojunction solar cells. Sol Energy Mater Sol Cells 95(7):1775–1778
Tobjörk D, Österbacka R (2011) Paper electronics. Adv Mater 23:1935–1961
Tybrandt K, Khodagholy D, Dielacher B, Stauffer F, Renz AF, Buzsáki G, Vörös J (2018) High-density stretchable electrode grids for chronic neural recording. Adv Mater 30(15):1706520 (1–7)
Verilhac JM, Benwadih M, Seiler AL, Jacob S, Bory C, Bablet J, Heitzman M, Tallal J, Barbut L, Frère P, Sicard G, Gwoziecki R, Chartier I, Coppard R, Serbutoviez C (2010) Step toward robust and reliable amorphous polymer field-effect transistors and logic functions made by the use of roll to roll compatible printing processes. Org Electron Phys Mater Appl 11(3):456–462
Von Metzen RP, Stieglitz T (2013) The effects of annealing on mechanical, chemical, and physical properties and structural stability of Parylene C. Biomed Microdevices 15(5):727–735
Vroman I, Tighzert L (2009) Biodegradable polymers. Materials 2(2):307–344
Wang S, Xu J, Wang W, Wang G-, Rastak R, Molina-Lopez F, Chung JW, Niu S, Feig VR, Lopez J, Lei T, Kwon S-K, Kim Y, Foudeh AM, Ehrlich A, Gasperini A, Yun Y, Murmann B, Tok JB, Bao Z (2018) Skin electronics from scalable fabrication of an intrinsically stretchable transistor array. Nature 555:83–88
Weiss RA, Sen A, Pottick LA, Willis CL, Block copolymer ionomers: 2. Viscoelastic and mechanical properties of sulphonated poly(styreneethylene/butylene-styrene). Polymer (Guildf) 32(15):2785–2792, 199
Yang F, Wang Y, Kuang X, Nie Y, Xia W (2018) Design and analysis of self-similar serpentine interconnects for stretchable electronics. Inter Conf Electro Packaging Tech (ICEPT), 1741–1743
Yi F, Wang X, Niu S, Li S, Yin Y, Dai K, Zhang G, Lin L, Wen Z, Guo H, Wang J (2016) A highly shape-adaptive, stretchable design based on conductive liquid for energy harvesting and self-powered biomechanical monitoring. Sci Adv 2(6):e1501624
Yim DR, Park CW (2022) Hybrid-type stretchable interconnects with double-layered liquid metal-on-polyimide serpentine structure. ETRI J 44:147–154. https://doi.org/10.4218/etrij.2021-0188
Zhai M, McKenna GB (2014) Elastic modulus and surface tension of a polyurethane rubber in nanometer thick films. Polym. (United Kingdom) 55(11):2725–2733
Zhao Y, Kim A, Wan G, Tee BCK (2019) Design and application of stretchable and self-healable conductors for soft electronics. Nano Convergence 6(1):1–22
Zhao Y, Zhou W, Shi Y et al. (2022) Superelastic alloy based electrical interconnects for highly stretchable electronics. npj Flex Electron 6, 8 (2022). https://doi.org/10.1038/s41528-022-00142-2

Chapter 13
Flexible Electronics: A Critical Review

Gulafsha Bhatti, Yash Agrawal, Vinay Palaparthy, Mummaneni Kavicharan, and Meenu Agrawal

Abstract Progressive changes are occurring at an unprecedented rate in the field of electronics, science, and technology. The conviction and evolutionary development of flexible electronics have been incredibly witnessed in recent years. Flexible electronics or flex circuits refer to technology that can be mounted onto the adaptable substrate. Flexible electronics comprises varying splendid properties such as flexibility, bending, rolling, folding, and stretching without losing its functionality. These characteristics are not fulfilled by present conventional electronic-based systems. Flexible electronics is highly beneficial in varying applications and used in several upcoming devices such as wearable gadgets, medical and industrial products, flexible sensors, and displays. Evolvement of flexible electronics has led to extensive research and development in flexible substrates and stretchable interconnects. Flexible substrate holds the flexible electronics. The substrate materials can be conductive polymers, organic semiconductors, graphene derivatives, and amorphous silicon. Further, to withstand large deformation, stretchable interconnects provide the overall flexibility to the electronic system. The interconnects are the connection between various components to transmit and receive signals. The foremost challenge in developing flexible electronics is not to lose electrical functionality of the stretchable interconnect along with providing good mechanical stability. The purpose of

G. Bhatti (✉) · Y. Agrawal · V. Palaparthy
Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, Gujarat, India
e-mail: 202021005@daiict.ac.in

Y. Agrawal
e-mail: mr.yashagrawal@gmail.com

V. Palaparthy
e-mail: vinay_shrinivas@daiict.ac.in

M. Kavicharan
Department of ECE, National Institute of Technology, Silchar, Assam, India
e-mail: kavicharan.mummaneni@gmail.com

M. Agrawal
Mechanical Engineering Department, Indian Institute of Technology, Hyderabad, Telangana, India
e-mail: meenu13796@gmail.com

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_13

this chapter is to review and address various flexible electronics modules such as components required, relevant materials, interconnect geometries, its electrical and mechanical properties, printing techniques along with its promising applications. Finally, the market growth and potential future research directions have also been discussed in the chapter.

Keywords Flexible electronics (FE) · Interconnects · Polymers · Stretchability · Substrate

Introduction

Today's modern era is highly progressed and dependent on electronic devices and systems. The major breakthrough in the electronic industry started with the invention of transistor by John Bardeen and Walter Brattain in April 1950 (Brinkman et al. 1997). Subsequently in 1956, Jack Kilby from Texas Instruments developed the first integrated circuit (IC) using a mesa transistor. It was an elementary device formed using discrete wire connections of one transistor, a capacitor, and a resistor on a single silicon layer. Thereafter, it took almost a decade to get all these different discrete components in place to start driving the planar process toward making IC (Brinkman et al. 1997). In 1960s, Gorden Moore predicted that the number of transistors on a single chip would grow exponentially with time which is now popularly referred to as Moore's law (Rabaey et al. 2017). Following Moore's law, the number of transistors increased in different processors with year. This is shown in Fig. 13.1. In 1960s, the first metal-oxide semiconductor (MOS) logic was introduced. After two years, the first logic family was launched, which had the advantage of offering higher integration density. This trend continued to take off entirely in the early 1970s. The second age revolution of the digital IC was pioneered with the introduction of the first microprocessors by Intel in 1972. Since then, the growth of electronics in communication, infotainment, and computation has amazingly driven. The patronizing areas such as optoelectronics, radio frequency, bioelectronics, organic light emitting diode (OLED)-based displays, manufacturing, and fabrications still use crystalline materials. However, the rigidity of electronic circuit boards restricts the flexibility of the device. Most of today's electronic systems are made on hard substrates. These systems lack the flexibility to bend or elongate. To mitigate this issue, flexible electronics (FE) technology is emerging. In FE technology, the electronic system can be folded, bended, and elongated. FE are built using flexible or elastic substrates and stretchable interconnects.

The concept of FE has existed for several decades and has a long history. The idea took strike when the solar cells of semiconductor-based silicon wafers used in satellites were thinned to permit a certain degree of bending. This led to the launch of the first single-crystalline flexible silicon solar cell of thickness around 100 μm and assembling them on a plastic substrate to provide the property of flexibility in 1960s (Crabb and Treble 1967; Ray 1967). The semiconductors are at the heart from ancient

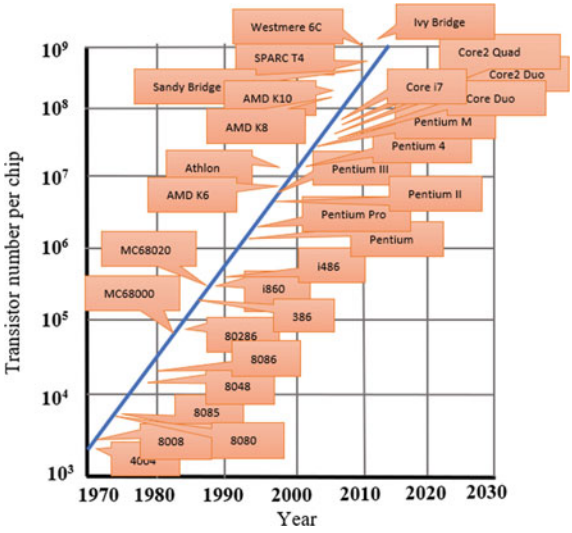


Fig. 13.1 Number of transistors per chip with increase in years following Moore's prediction (Intel® Products 2023)

to modern living that is utilized to build electronic devices such as resistors, diodes, MOS field-effect transistors (MOSFETs), and ICs. Among the most widely used semiconductor materials are silicon, germanium, and gallium arsenide. In which, silicon usage is eminent and high due to its availability in abundance.

Silicon has remained the main steer in developing nanoscale technology to reduce its cost and elongate its reliability. But the rigidity of the material limits its ductility in flexible and stretchable electronics applications. These are the strong motivations for a ubiquitous search for prospective materials with higher potential which can replace and overcome the conventional silicon technology.

As found by the researchers, the foremost emerging materials, such as graphene nanoribbon (GNR) and carbon nanotube (CNT), show outstanding properties in terms of elasticity, thermal capability, reliability, and mechanical strength, when compared to conventional silicon. Afterward, in the following decades, there was a giant stride in the development of different materials such as hydrogenated amorphous silicon (a-Si: H), conductive polymers, organic semiconductors, transition-metal dichalcogenide (TMD), and boron nitride (BN). These gave good base strength to the FE devices that require twisting, bending, rolling which is not satisfied by the conventional electronics. The evolution of substrate materials from rigid to flexible is described in Fig. 13.2. The novel FE holds huge applications in future technologies such as environment monitoring, healthcare, sensor, e-textile, and wearable devices.

This review chapter explores the recent, intensive, and trendy development of FE. The components, characteristics, properties, geometries, applications, and market growth of FE are detailed in this chapter. The present chapter is organized into five sections. The current first section discusses the evolution of electronics era

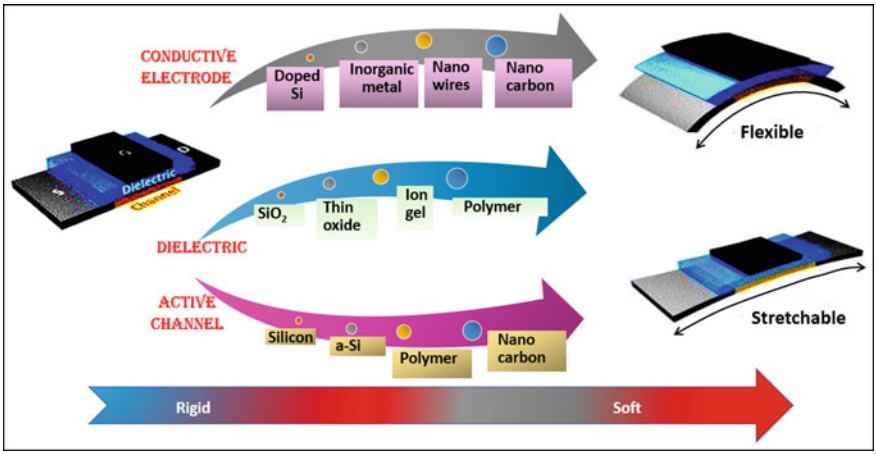


Fig. 13.2 Evolution of electronics from rigid to soft materials (Chae and Lee 2014)

and brief introduction of FE. Section "Components of Flexible Electronics" details about the different components and properties those are used for building FE. In Sect. "Printing Technologies for Fabrication of FE", the latest and most effective printing techniques used for the fabrication of FE are discussed. Next, Sect. "Market Growth and Applications" covers the wide progressive applications of FE and its market growth. Finally, the chapter concludes in Sect. Conclusion with highlighting profound research direction paths toward developing a new era in the world of FE.

Components of Flexible Electronics

The FE comprises four major components as (a) substrate, (b) device, (c) interconnect, and (d) encapsulation. Each of these components must comply to bend and stretch keeping their functionality intact. This section briefly discusses each of the core components. Further, the materials and their properties are also defined.

Substrate

Substrate is a medium that provides foundation for building components over it. A flexible substrate refers to a thin material that should be mechanically robust and have high tolerance to repeated bendability. The factors such as low cost, low mechanical stiffness, high conductivity, simplicity in manufacturing, ease of reparability are the desired features for selecting flexible materials. Silicon has been a foremost essential material used as a substrate in electronics circuits. Thickness of silicon is generally

Table 13.1 Properties and applications of different types of substrate materials (Sankir 2005)

Property	Unit	Metal	Organic polymers/plastic	Thin/ flexible glass
Example		Stainless steel (430)	Plastics (PEN, PI, PET)	Glass (1737)
Thickness	μm	100	100	100
Safe bending radius	cm	4	4	40
Maximum process temperature	°C	1000	180, 300	600
Coefficient of thermal expansion (CTE)	ppm/°C	10	16	4
Elastic modulus	GPa	200	5	70
Electrical conductivity	℧	High	None	None
Thermal conductivity	W/m°C	16	0.1–0.2	1
Deform after device fabrication	–	No	Yes	No
Application	–	OLED	Adhesives, coatings, potting compounds, and sealants	Flat panel display

100 μm and therefore to make this rigid layer flexible, the substrate thickness is required to be lesser than 50 μm so that the chip can be bent without getting cracks. However, reducing the thickness of silicon can make it brittle and non-conductive. Some of the major types of substrate materials preferred for FE applications are metals, organic polymers/plastics, and thin/flexible glass (Sankir 2005). The properties and applications of each type of these substrate materials are listed in Table 13.1. The examples for different types of materials are discussed in the table such as stainless steel (430) is an example of metal, polyethylene naphthalate (PEN), poly-imide (PI), and polyethylene terephthalate (PET) are examples of organic polymers/plastic, and glass (1737) is an example for thin/flexible glass. The properties such as thickness of the material, safe bending radius, maximum process temperature, coefficient of thermal expansion (CTE), elastic modulus, electrical conductivity, thermal conductivity, deform after device fabrication, and applications are discussed in Table 13.1. It can be inferred from the table that the polymeric materials show the highest flexibility and conductivity and hence preferable choice in many FE applications.

Device

Device refers to passive and active components that can be used to build the electronic system. The passive electronic devices comprise resistors, capacitors, and inductors, whereas the active electronic devices are thyristors and transistors which

control the voltage-current characteristics in the circuit. There are different types of transistors such as bipolar junction transistors (BJTs), insulated-gate bipolar transistors (IGBTs), field-effect transistors (FETs) those can be used as switch/power device in the circuit. One of the commonly used FETs is MOSFET. The primary use of MOSFET is to control the flow of electric current between its source and drain terminals based on the voltage supply at its gate terminal. MOSFETs are used to build conventional electronic systems that cannot be stretched or flexed. Therefore, to build the FE system, thin film transistors (TFTs) are majorly used. The TFT is an improved version over MOSFET switch. The fundamental difference between the MOSFETs and TFTs is their supporting substrate. For manufacturing MOSFETs, silicon wafer is commonly used as a substrate while glass/polymers are preferably used for making TFTs.

TFT is a transistor whose fabrication is done by depositing thin films of active semiconductor, dielectric and metallic layers on base of flexible substrate. This can be bent up to 1/16 radius, cut into two halves and still remain functional (Itoh et al. 2016). The fabrication of TFTs is relatively lesser complex and are compatible with wrapping products. In terms of historical point of view, the interest and development of TFTs and FE are traced back to 1960 by Paul Weimer, who showed application of inorganic TFT which turned out into the directed path for the development of novel devices (Weimer 1962). Then gradually, in 1968, Brody and team made a TFT of tellurium on a thin strip of paper. They applied it in forming display matrices further adding up more TFTs on a flexible substrate in enough scale which includes mylar, polyethylene, and anodized aluminum folding foils (Sakuma 2022). TFTs have been extensively used in industries for developing different products such as display, nanotechnology-based bioFET, sensors, IOT devices, RFID, memory, and e-paper. This subsection summarizes the basic structure of assembling TFT, their types, and materials for deformable TFT devices.

The basic structure of TFT comprises following essential elements as (a) substrate, (b) insulator, layer, (c) thin film semiconductor, and (d) three electrodes namely gate, source, and drain. According to the arrangement of these elements, TFT can be divided into staggered (S) and coplanar (C) structure. Further each type of these structures is subdivided into top-gate (TG) and bottom-gate (BG) type. The schematic of each of these structures is illustrated in Fig. 13.3 (Shang et al. 2019). The first layer of TFT is made of flexible substrate. As can be seen from the figure, the source and drain regions are separated by the semiconductor film. Whereas the gate and semiconductor film are separated by the insulating layer to avoid the electrical shorting. The work of gate electrode is to provide signal to the semiconductor layer which develops the contact channel between the source and drain electrodes. The selection of semiconductor material is based on certain criteria. First is that the gate terminal can control semiconductor and still remain separated by an insulating dielectric layer. Second is to provide carrier injection to the conducting contacts. The basic working principle of TFTs is opposite to that of MOSFETs. When a negative gate voltage V_g greater than threshold voltage V_T is applied at gate terminal, the holes get accumulated at the semiconductor-dielectric interface. This interface behaves as a capacitor and thus provides a conducting channel. The holes at interface drift from

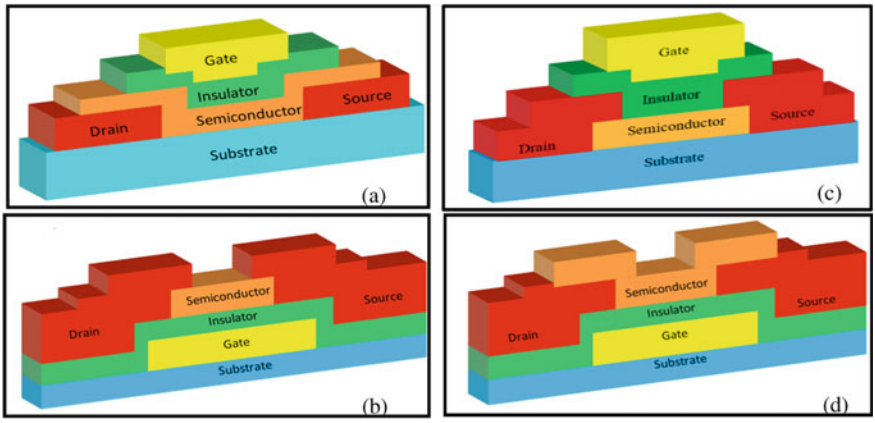


Fig. 13.3 Generalized TFT structures **a** staggered top-gate (S-TG) **b** staggered bottom-gate (S-BG) **c** coplanar top-gate (C-TG) **d** coplanar bottom-gate (C-BG) (Shang et al. 2019)

source toward drain terminal and generate drain current I_D (Sankir 2005). Whereas in TFTs there is no p–n junction in the source and drain regions. In TFT, the modulation is achieved by an accumulation layer, whereas in MOSFETs the inversion region has to be formed to achieve modulation. Thus, TFT operates in accumulation region, unlikely as MOSFET operates in inversion region. Moreover, the manufacturing temperature of MOSFETs and TFTs are also different.

In S-TG and S-BG structures, the contact of source/drain electrodes are on opposite sides of insulating layer. Whereas in case of C-TG and C-BG structures, source/drain electrodes and insulating layer are in direct contact with each other. In S structure configuration, current flows in two planes, first is vertically along the channel and second is horizontally from source to drain electrode. Whereas, in C structure, current flows in a single horizontal plane.

The arrangement of each element has strong influence on advantages/disadvantages, device performance, and manufacturing materials of TFTs. The TG structure is preferable when the semiconductor materials requirement is to have high-quality crystal structure, a flat and continuous film and high processing temperature. The TG structure protects the semiconductor from external damage; hence it may act as a passivation layer. TG structure is not suitable for LCD applications, as it cannot block the backlight unit from reaching the semiconductor layer. On the other hand, the fabrication of BG structure is simpler and has enhanced electrical properties. Therefore, the structure is preferable for light sensitive semiconductor materials. The gate electrode of BG structure blocks the backlight unit effects from reaching the semiconductor materials. Therefore, this is widely used for applications in LCDs. Furthermore, in BG structure the semiconductor layer is directly exposed to air, due to which it becomes simpler to modify its properties. But the exposure can also lead to instability as a result the back-channel surfaces can suffer damage during annealing or plasma treatment processes (Fortunato et al. 2012; Lu et al. 2018).

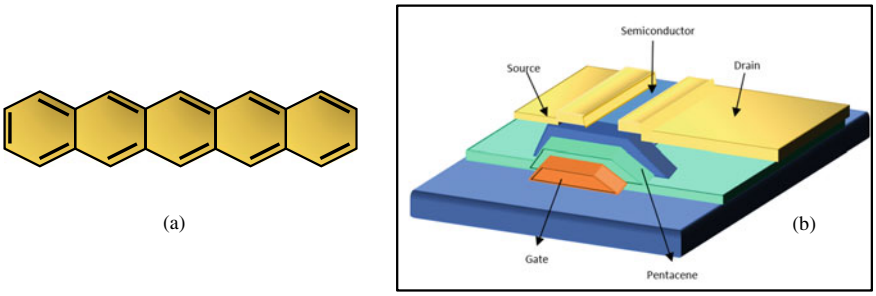


Fig. 13.4 **a** Structure of an individual pentacene molecule comprising of five linear benzene rings **b** pentacene-based TFT (Lin et al. 1997)

The semiconductor active channel layer in TFT can be fabricated using different materials such as zinc oxide, organic semiconductors, and CNTs. For instance, two prominent materials for TFTs are organic semiconductors-based pentacene and CNT which are discussed next in this section.

Pentacene based TFTs: Pentacene is vastly researched as an organic semiconductor due to its large carrier mobility values that is greater than 10 cm^2/Vs on alumina dielectric (Ausanio et al. 2006; Qaiser et al. 2017). Figure 13.4a shows the organic small molecule of pentacene ($C_{22}H_{14}$) which comprises five linear benzene rings. This molecule exhibits the semiconductive properties because of the presence of π-conjugated system in the benzene ring with alternating single and double carbon–carbon bond. The main governing component is molecule–substrate weak Van-der-waal's electrostatic interaction force allows charge transport to occur linearly across the molecule (Hussain and El-Atab 2022; Lin et al. 1997).

The p-type pentacene bottom-gate staggered TFT structure is shown in Fig. 13.4b. In order to generate electric field, a negative or positive gate voltage is applied to the p-type or n-type semiconductor respectively. As a result, p/n type voltage attract either electrons or holes to the dielectric-semiconductor interface surface. Thus, on providing required biasing on source-drain terminals, a conducting channel is produced at this interface which leads to current conduction (Lin et al. 1997). Due to having high mobility, pentacene is widely used for OFET applications.

CNT-based TFTs: CNT can exist in three different structures such as armchair, zigzag, and chiral. These are shown in Fig. 13.5. CNT is the strongest material in terms of tensile strength and stiffest materials in terms of elastic modulus. CNT exhibits extraordinary electrical, thermal, and mechanical performance due to its intrinsic molecular structure which makes it the most promising material for future FE applications (Kumar et al. 2016; Pathade et al. 2021). A schematic architecture of fully printed CNT TFT on elastomeric substrate is shown in Fig. 13.6. It comprises a bilayer high-k dielectric layers of 30 nm MgO and 30 nm of Al_2O_3. These layers are deposited via e-beam evaporation and atomic layer deposition (ALD) which act as dielectric medium. The PET is utilized as a flexible substrate. Above PET substrate, an adhesive layer of silicon dioxide is fictionized. The bilayer structure

enables to achieve higher conduction even at lower temperature of 120 °C. These fully printed CNT TFT devices withstand stretching with tensile strain exceeding 50% while showing no significant electrical degradation (Yu et al. 2018). CNT TFTs are observed to be better than polycrystalline silicon-based TFTs in terms of low cost. Also, it is better than amorphous silicon and organic semiconductor transistors in terms of mobility. Thus, CNT TFTs have acquired wide place in application of flexible displays such as active matrix (AM) for liquid crystal displays (LCD), OLEDs, and radio frequency where it shows highly improvised performance.

The afore discussed structures, their types, and materials when assembled result in a thin sheet array of TFT that can be utilized for numerous novel applications in FE. Figure 13.7 shows the optical image of TFT array on thin flexible film of PET. In addition to the devices presented, a comparison table of a few latest state-of-the-art TFTs is also discussed which is shown in Table 13.2. In the table, the parameters those are considered for comparison of TFTs are channel material, type of substrate material, gate dielectric along with their thickness, channel deposition process, flexibility, effective mobility, ON current, and cost. In the table, M refers to the medium cost and H denotes the high cost of TFTs.

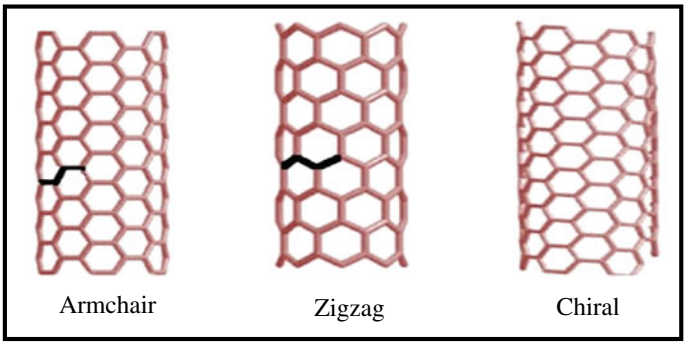


Fig. 13.5 Schematic representing different CNT structures (Singh et al. 2017)

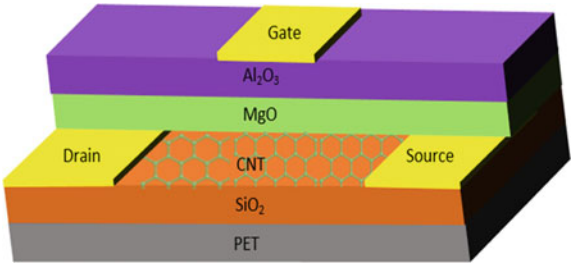


Fig. 13.6 Schematic of the flexible n-type CNT TFT (Yu et al. 2018)

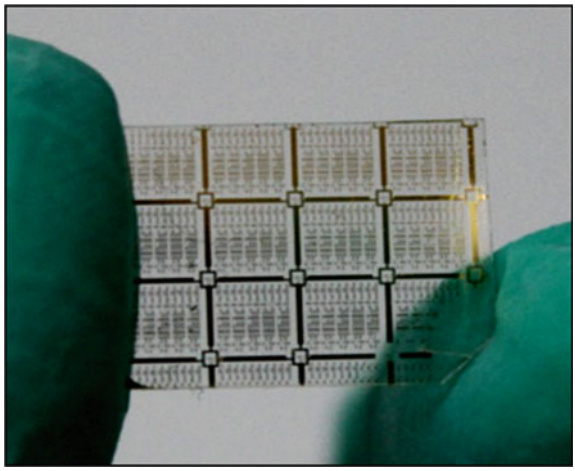


Fig. 13.7 Optical image of n-type TFT arrays on a thin sheet of PET (Li et al. 2015)

Table 13.2 Characteristics of varying TFTs (Hussain and El-Atab 2022)

Channel material	Substrate	Gate dielectric (Thickness)	Channel deposition process	Flexible	Effective mobility cm^2-Vs	I_{ON} (μA/mm)	Cost
s-CNT	Si	SiO_2 (90 nm)	Aerosol Jet Print	No	6.7	102	M
s-CNT	PEN	Al_2O_3 (40 nm)	CVD Transfer	Yes	20	17	M
s-CNT	Si	SiO_2 (500 nm)	Inkjet	No	1.6–4.2	0.11	M
ZnO	PI	ZrO_2 (20 nm)	CAD	Yes	4.5	30	M-H
IGZO	PI	HfLaO (40 nm)	Sputter	Yes	13	128	H
In_2O_3	PI	Al_2O_3 (100 nm)	Spray pyrolysis	Yes	1.25	320	M
IGZO	PI	$ZrAlO_x$ (35 nm)	Spin coat	Yes	8.58	2560	M
Pentacene	Paper	Paralene-C (340 nm)	Thermal evaporation	Yes	0.1	13	H

Interconnect

Interconnect refers to the connection between various devices and components on system. With the miniaturization in technology, interconnects itself have become one of the major limiting factors in attainment of high performance of the system. The advanced technology nodes enable embedding more transistors on the same silicon chip. These lead to corresponding increase in number of interconnects and higher parasitics (Agrawal and Chandel 2015; Agrawal et al. 2016). High speed interconnect channels face issues such as signal degradation, electromagnetic interference (EMI), skin effect, and coupling between interconnects if not modeled efficiently. Therefore,

in order to attain advancing progress in future electronic technologies, the modeling of interconnect is crucial in both rigid and FE system designs.

FE systems which can undergo large deformation need stretchable interconnects. Complete FE system is achieved when the module can sustain large and stretchable deformation. Therefore, the characterization needs to be done by critically focusing on impedance and conductivity of the interconnect. The primary formation of interconnect can be done by combination of two approaches. First approach is by employing deformable materials and second is by tailoring the interconnect geometry designs. Additionally, several approaches have been manifested in (Ausanio et al. 2006; El-Tantawy 2005), which include realization of stretchable interconnects making blend of conductive polymers and "filled" elastomers for forming interconnects between rigid components or blending metal particles.

This subsection briefs about the extraction of parasitic parameters of stretchable interconnect and simultaneously its correlation with the mechanical properties. Secondly, the various geometry structures such as development from straight to meandering, arch-shaped, pre-stretched substrate to fractal/self-similar designs, serpentines designs, horseshoe structure, wavy ribbon, and spiral designs are discussed.

It is likely that while applying stress/strain, stretching, folding, or bending, the electrical properties can get affected. Therefore, the characterization between the electrical and mechanical properties of stretchable interconnect is a great challenge. The parasitic resistance (R), inductance (L), and capacitance (C) have numerous effects on the behavior of the circuit and its reliability. For instance, if these rise, it leads to increase in propagation delay, energy dissipation and power dissipation, noise which tends to degrade device performance. The parasitics of the conductor can be extracted by many ways as through single test access port known as two test points, time domain reflectometry, two port measurement such as Kelvin test (Qaiser et al. 2017) and mathematical method by analytical modeling. In the present chapter, R, L, and C parameters are extracted using analytical modeling method as (Dong et al. 2017).

Resistance: At high frequencies, the resistance of wire increases and skin effect issues get aggravated. In skin effect, the current can pass only through surface peripheral of the conductor which leads to signal degradation. Figure 13.8 depicts the structure of stretchable interconnect. The interconnect can be modeled by equivalent circuit model as shown in Fig. 13.9. The cross section of wire is assumed as rectangular. In Fig. 13.8, parameters R_d and r are the radius of bigger and smaller semi-circle respectively, l_e is the effective length, and l_{wire} is the actual length of wire. The material of the conductor that are considered are Cu and PI. Typically, the resistance of conducting wire can be given as

$$R = \frac{\rho l}{A} \; and \; A = wt \tag{13.1}$$

where A is the cross-sectional area of wire, w is the width, t is the thickness of conducting wire. The modified resistance (R') expression under the skin-depth effect

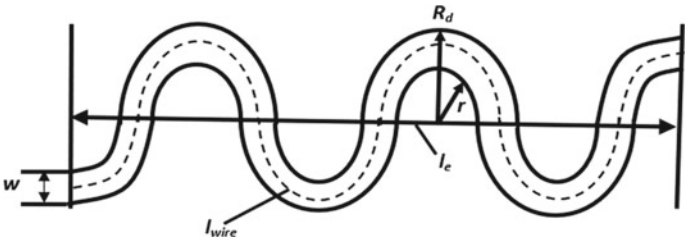


Fig. 13.8 Structure of stretchable interconnects (Agrawal and Chandel 2015)

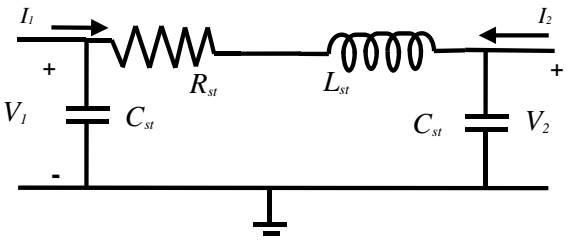


Fig. 13.9 π equivalent circuit model of stretchable interconnect (Agrawal and Chandel 2015)

consideration can be given as

$$R' = \frac{\rho l_{wire}}{wt - (w - 2\delta)(t - 2\delta)} \tag{13.2}$$

Thereafter, resistance of the serpentine stretchable interconnect (i.e., R_{st}) for the structure as shown in Fig. 13.8 can be derived as (Dong et al. 2017):

$$R_{st} = \frac{\rho \pi N \frac{R_d + r}{2}}{wt - (w - 2\delta)(t - 2\delta)} \tag{13.3}$$

where, $\delta = \frac{1}{\sqrt{f \pi \mu \sigma}}$ is the skin depth, ρ is the resistivity of the material, μ is the magnetic permeability of the conductor, and N is the number of considered semicircles.

Inductance: At high frequencies, the inductive effects crop up significantly that affect the performance of system. The generalized expression to obtain inductance (L) can be derived using Neumann's method as (Agrawal and Chandel 2015):

$$L = \frac{\mu}{4\pi} \oint \oint \frac{\overrightarrow{dl\prime} \cdot \overrightarrow{dl}}{|r - r\prime|} \tag{13.4}$$

For considered stretchable interconnect structure as shown in Fig. 13.8, (13.5) is solved to get the final expression as

$$L_{st} = \frac{\mu(l_e + x)}{2\pi} \times \left[\frac{2(l_e + x)}{w + t} + 0.5 + 0.22 \frac{w + t}{(l_e + x)} \right] \tag{13.5}$$

where x is the change in length after stretching the interconnect. From (13.5), it is observed that as interconnect is stretched, the length of wire increases and hence inductance.

Capacitance: The interconnect capacitance leads to coupling effect and crosstalk which results in slowing down of the device and noise generation in the system. The basic expression for capacitance per unit length (p.u.l.) of the wire can be obtained as

$$C = \frac{\mu\varepsilon}{L_d} \tag{13.6}$$

where ε is dielectric constant, L_d is the inductance p.u.l. For the considered stretchable interconnect structure as present in Fig. 13.8, using (13.6) the total capacitance is derived as

$$C_{st} = C\pi N \frac{R_d + r}{2} \tag{13.7}$$

From (13.7), it is observed that with stretching of wire, capacitance increases as the variation in capacitance depends upon cross section of wire.

Along with electrical properties consideration, mechanical properties in FE are also equally important. Parameters such as strength and stiffness are the important measure of mechanical property in FE. To measure the strength and stiffness of material, the fundamental parameters are stress and strain. The stress refers to the force per unit area within the material. The stress within the material is the sum of the measure of internal/residual and external/applied stress. The stress (σ_s) can be given as

$$\sigma_s = \frac{F}{A} \tag{13.8}$$

The tensile or compressive strain is defined as the ratio of final change in length (ΔL) to the original length (L_0). The mathematical expression representing the tensile or compressive strain (ϵ), can be given as

$$\in = \frac{\Delta L}{L_0} \tag{13.9}$$

The critical parameter that is based on the measurement of stress and strain of the material is Young's modulus. Young's modulus represents mechanical property that measures the elasticity of a material. It is calculated by obtaining the relationship between stress and strain of a material. The modulus of elasticity or Young's modulus (E) can be given as

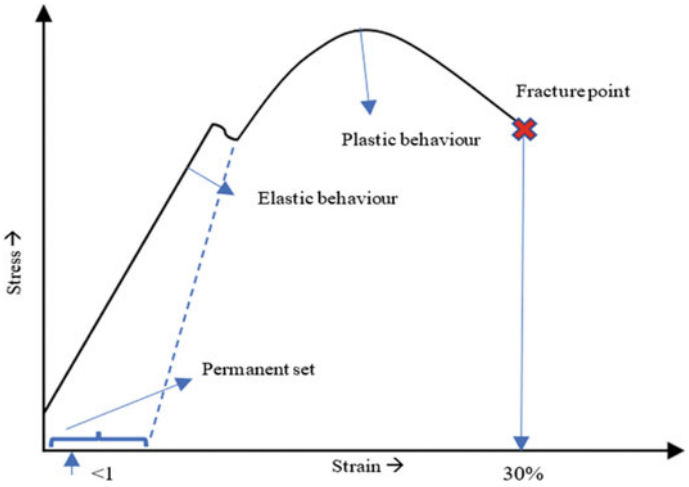


Fig. 13.10 Schematic representing Young's modulus: Stress–strain relationship (Budynas et al. 2021)

$$E = \frac{\sigma_s}{\in} \tag{13.10}$$

Figure 13.10 shows Young's modulus graph which represents the material behavior for a particular slope of stress versus strain. Material that can sustain higher sustainability of strain shows high elasticity. Therefore, the lowest value of Young's modulus depicts the highest elasticity of the material.

The geometry and type of material significantly impacts the elasticity of the stretchable interconnects. Generally, geometry of stretchable systems can be obtained in two ways as (a) engineered shapes and (b) rubber-like shape that are intrinsically stretchable. Figures 13.11 and 13.12 show the different shapes and geometries of stretchable interconnects. Each interconnect geometry and applications are discussed briefly here in this subsection.

Figure 13.11a: The geometry of stretchable interconnect is net or beehive shape. This is realized by implementing the mini-valleys on the surface of pre-strained elastic substrate along with conductive material allocation over it. The advantage of this technique assists toward development of other techniques such as metal-coated net films, 1D metal ribbons, or 2D membrane.

Figure 13.11b–d: Twisted cable, e-textile, and helical coil are another set of geometric designs of stretchable interconnects. In these structures, the winding coils are of large diameter, which limits extension of scalability. Helical coil wires are used in various applications such as in connecting telephones in wired technology. With advent of wireless technology, the approach of stretchable wires has found new applications such as in tactile sensing and e-textile. E-textile, such as conductive thread and stretchable fabrics, has been used in wearable electronics and stretchable

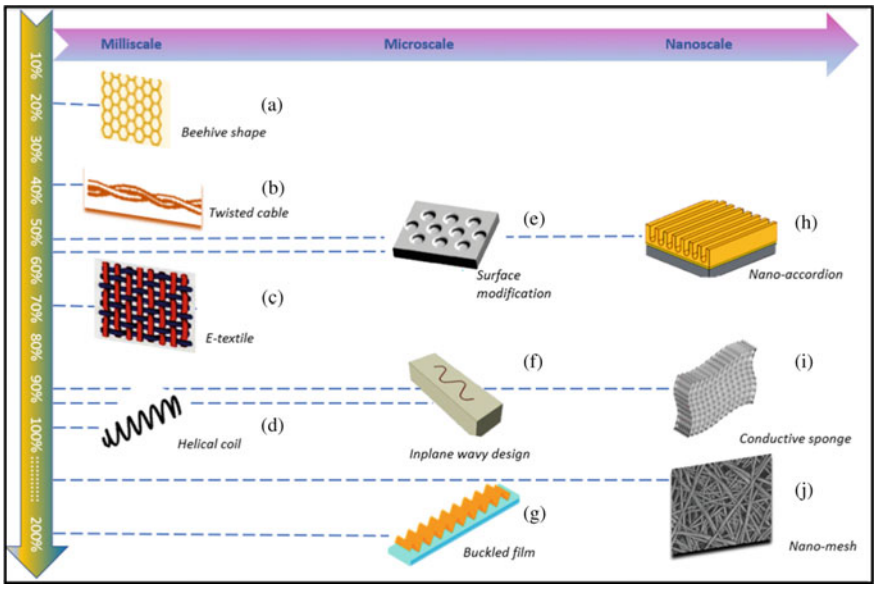


Fig. 13.11 Geometries designs of various stretchable interconnects with respect to their scalability (Dang et al. 2017)

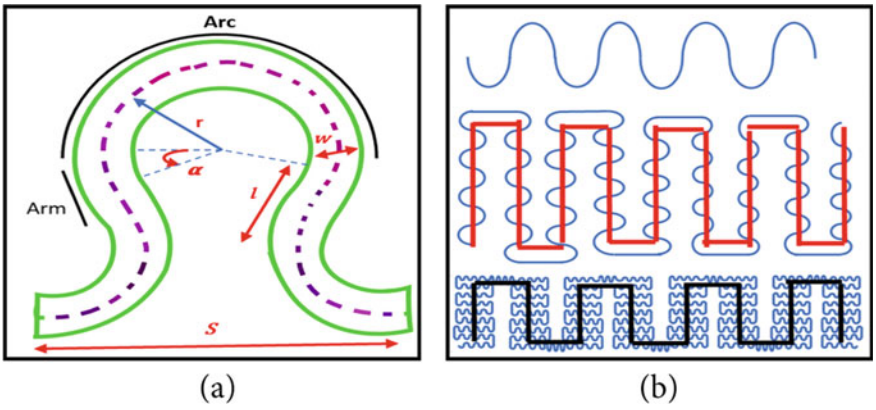


Fig. 13.12 **a** Horseshoe-based interconnect design (Yang et al. 2015) **b** examples of fractal/self-similar designs (Zhang et al. 2014)

tactile sensing-based applications. The failure strain of this geometry structure can be increased up to 100%, however it will lead to increase in resistance by 70%.

Figure 13.11e, h: Next conductive interconnect geometries are surface modification and nano-accordion which are stretchable and bendable. In surface modification geometry, the deformable interconnects that are made on flexible substrate can be modified into the surface that need expansion out of the plane. The nano-accordion

structure is made up of Al-doped ZnO which is highly conductive and optically transparent. This structure can hold the strain of up to 53% of elongation. However, this also leads to change in its resistance significantly.

Figure 13.11f, g: Next geometry is the conductive wavy structure or serpentine that can be reversely stretched without cracks or fractures. The geometry of inplane and outplane wavy designs are also known as buckled film that can be easily scaled down to the microscale as compared to the restricted scalability of helical coil. The wavy structure can be applied on the pre-stretched flexible substrates to improve its stretchability. The wavy geometry can sustain high stretchability of up to 100% and radius bending of up to 1.20 mm with least influence of electrical properties (Dang et al. 2017). The geometric designs and interconnects acquire place in application of systems such as implantable medical technology and smart clothes.

Figure 13.11i: Another geometry for stretchable interconnect is honeycomb lattice structure or conductive sponge-like architecture. These are commercially available as polyurethane (PU) sponge or carbon nanotube sponge. Though there is a great research interest on multilayer graphene sponge, the concept of its application as a stretchable interconnect has limitations due to its restricted scalability. The consistent integration and soldering are difficult due to pores in sponge-based interconnect.

Figure 13.11j: The nano-mesh structure can create very fine structure of interconnect. This nano-mesh structure can withstand strain up to 160%. However, the abrupt change in resistance is observed after 1000 test cycles. Further, some interconnect structures such as serpentine and fractal/self-similar geometries which are more often used in stretchable interconnect applications are shown in Fig. 13.12.

Figure 13.12a, b: The horseshoe shape geometry can accommodate large deformation under mechanical stress and preserve the electrical properties. These interconnect-based geometries when coupled with conductive yarns can improve the robustness in e-textiles circuits. The next exceptional level structure is the fractal geometry-based fractal/self-similar-based interconnects which enables stretchability up to 300%. When the order of fractal is increased, the elasticity of interconnect increases more than double. This is relatively thick interconnect structure which demands applications of low resistance. This structure finds application in stretchable batteries, photovoltaic modules, and super capacitors.

There are numerous geometries, polymeric, and elastic materials used in FE as a stretchable interconnect which does not lose functionality even on repeat bendability. The elastic modulus and electrical conductivity of different materials that are used for stretchable interconnect are listed in Table 13.3. The PU, PDMS, and PET are some of the most commonly used polymers in FE as these possess high elasticity and show moderate conductivity. The conductivity of these materials can be increased by using carbon or metal-based fillers. In the recent research work, a novel composite material such as poly-3,4-ethyl-enedioxythiophene: poly-styrene sulfonic acid (PEDOT: PSS), poly-vinyl alcohol (PVA), phosphoric acid (H_3PO_4), polymer blend, and silver micron sized flakes have been explored which show both good electrical and mechanical properties.

Table 13.3 Elastic modulus and electrical conductivity of materials for stretchable interconnects (Dang et al. 2017)

Material	Elastic modulus (GPa)	Electrical conductivity (S/m)
Polyethylene Terephthalate (PET)	0.0239	1.0×10^{-20}
Polypropylene (PP)	0.2553	1.0×10^{-13}
Polystyrene (PS)	0.9523	1.0×10^{-12}
Gold (Au)	83.000	$4.9 \times 10^{+07}$
Polydimethylsiloxane (PDMS)	0.0025	1.0×10^{-12}
PEDOT: PSS	0.0004	$1.0 \times 10^{+03}$
Polymethyl methacrylate (PMMA)	0.3855	1.0×10^{-19}
Silicon rubber	0.0026	$6.6 \times 10^{+7}$
Polyurethene (PU)	0.0095	1.0×10^{-10}
Graphite	130.00	$1.5 \times 10^{+05}$
Silver (Ag)	81.000	$6.6 \times 10^{+07}$
Copper (Cu)	110.00	$6.4 \times 10^{+07}$
Carbon nanotube (CNT)	250.00	$1.0 \times 10^{+08}$

Encapsulation

To enhance the mechanical and electrical stability of system, each module or component such as substrate, dielectric, interconnect should be able to sustain process variation effects. In order to achieve protection from extrinsic effects, encapsulation layer techniques are applied. The encapsulation is the process where the modules/components are enclosed by protecting materials. The encapsulation layer comprises thermally stable and high mobility organic semiconductors which provides the elevation to the substrate temperature up to 150 °C or higher. Conventional encapsulation techniques are no more feasible for FE due to its inherent rigidity. The absence of encapsulation layer will create obstacle to the stability and materials high cost while using printing technologies on FE. Therefore, to ensure protection of the flexible devices, organic/inorganic hybrid thin-film encapsulation (TFE) is usually considered.

The encapsulation techniques such as single-layer metallic barrier coatings followed by single-layer transparent barrier coatings, multilayer composite (oxide or nitride/polymer) barrier coatings have been developed. These techniques improve bending, twisting, stretching, and resist from direct chemical exposer to the atmosphere. Furthermore, when the interconnects are encapsulated with layer of PDMS, it is observed that the mechanical bendability is raised (Verplancke et al. 2011). The encapsulation material is extremely important as it can dominate the properties of FE systems. A perfect thin layer of inorganic materials such as SiO_2, SiN_x, and Al_2O_3 are highly impermeable to atmospheric gases that can drastically reduce the granular film growth, dust particle, and gas permeation hence good for encapsulation process in FE. These are beneficial in long-term implantable devices, in medical apparatus

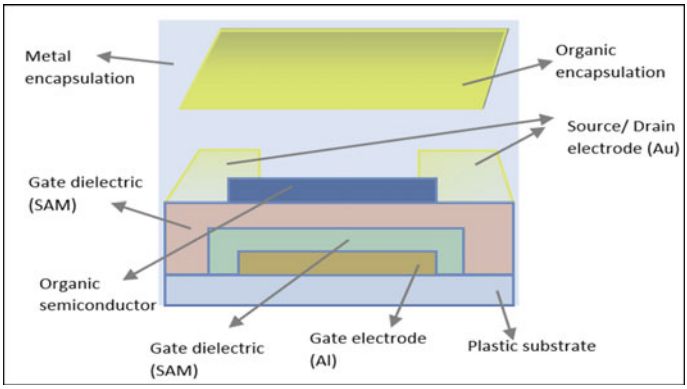


Fig. 13.13 Schematic representing an organic TFT covered with a flexible encapsulation layer (Walsh and Genzer 2012)

as a sensor to detect tumors, swelling, and cancer. Figure 13.13 shows schematic of an organic self-assembled monolayer (SAM) TFT covered with the encapsulation layer. The organic semiconductor dinaphtho-thieno-thiophene (DNTT) and metal composite films are chosen as the heat resistant materials. This film is utilized to encapsulate after the complete manufacturing of SAM transistor. The encapsulation layer improves the thermal stability, substantial deterioration and therefore electronic characteristic does not change even after dipping the transistor in boiling water (Walsh and Genzer 2012).

Printing Technologies for Fabrication of FE

The development of printed electronics to produce devices, components, electronic circuits (e.g., resistors, capacitors, transistors, antennas, and alike) in form of wide array on stretchable thin like paper substrate, plastic, or textiles is very interesting. This can be obtained and fabricated by existing highly efficient printing technologies. The printing FE is based on additive process method, whereas the silicon semiconductor-based conventional electronic manufacturing is based on subtractive-based photolithography method. The overall complexity of the manufacturing process of printed electronics is comparatively lesser, therefore, the printed electronics have the potential for low-cost production whereas the conventional electronics incorporates high production cost. This subsection elaborates on various printing technologies available to fabricate FE.

The classification of various printing technologies is shown in Fig. 13.14. The printing technologies can be classified into two types as contact and non-contact printing. In contact printing technology, there is a main medium that incorporates the circuit information during printing process. In this, the circuit information is obtained

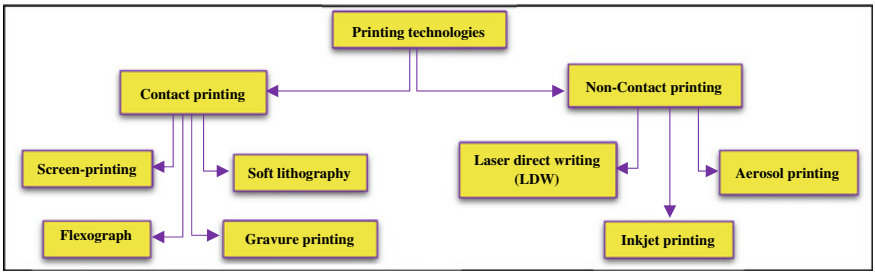


Fig. 13.14 Classification of printing technologies (Cruz et al. 2018)

on the substrate via direct contact or by partial transfer of printing ink. Whereas, in case of non-contact printing there is no such main medium carrying circuit information. Additionally, as the name depicts there is no direct physical contact between the substrate and the printing apparatus. The desired master layout can be laid with different information to print on the flexible substrate. Technologies such as screen printing, flexography, soft lithography, and gravure printing fall under the contact printing methods. The technologies under non-contact printing are laser direct writing (LDW), aerosol printing, and inkjet printing. The schematic of the various printing technologies is provided in Fig. 13.15a–g. Each technology has different potential, specifications, and limitations to print flexible substrate. Tables 13.4 and 13.5 present a detail comparison of various contact and non-contact printing technologies respectively. The comparison is made in terms of mechanism and feature, image resolution, viscosity of ink, film thickness, printing speed, and challenges (Cruz et al. 2018; Walsh and Genzer 2012).

Along with printing technologies, materials for conductive ink to print stretchable interconnects are also very important. These play a vital role in wearable devices, mechanics of human skin consideration for flexible near field communication (NFC) tags, etc. Moreover, the stable contact between ink and substrate is extremely desirable. Table 13.6 gives the comparative conductivity description in descending order of various conductive ink materials. These ink materials are available in market and can be employed by multiple printing fabrication technologies to print FE.

Market Growth and Applications

Market Growth

The benefits of FE such as low cost, portability, energy efficiency, and durability make the previous impossible technical applications possible. These are the concrete reasons why the FE is finding a huge place in development of the latest electronic gadgets. The development of FE products has drastically increased the market growth

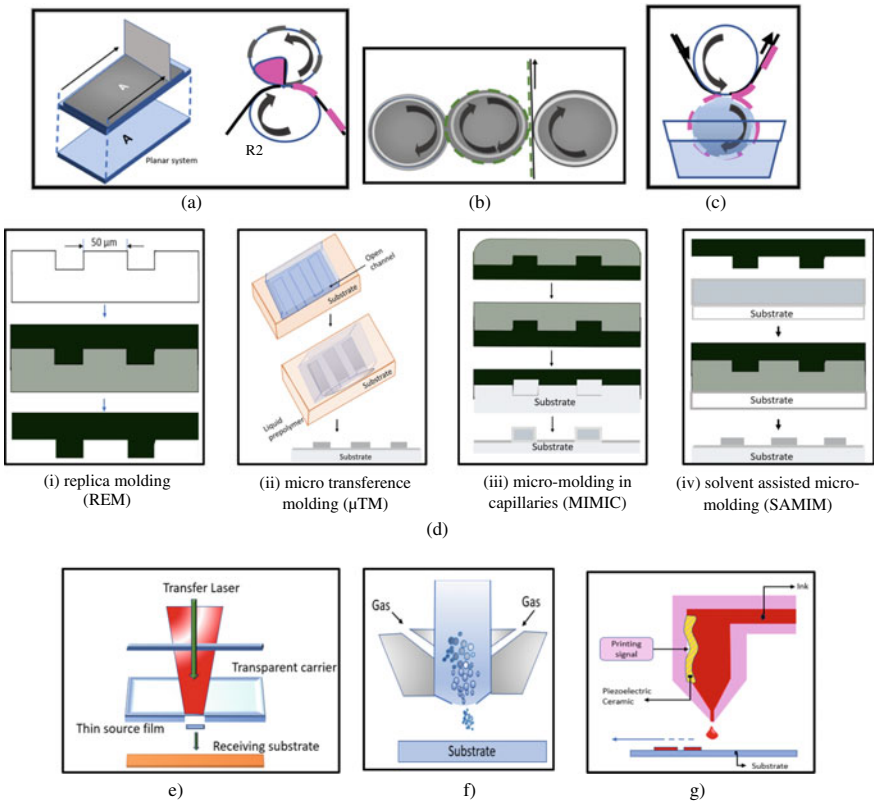


Fig. 13.15 Schematic of contact and non-contact printing technologies. **a** screen printing **b** flexography printing **c** gravure printing **d** soft lithography **e** laser direct writing **f** aerosol printing **g** inkjet printing (Cruz et al. 2018)

rate and is seen to be increasing exponentially. Figure 13.16a, b shows the graph of the FE market growth from 2019 increase up to year 2030. Figure 13.16a shows the growth rate for global flexible display. The presented statistics is developed by Next Move strategy consultancy (Flexible display market size, share n.d.). It is a premier market research and management consulting firm which provide strategically analytics report of the data. From the figure, it can be inferred that there is an increase of market in the flexible display applications integrated with smart wearable devices to be around $250 billion by 2030.

The second graph shown in Fig. 13.16b is the market report forecasted by IDTechEx analysts for FE in healthcare devices (IDTechEx 2020). It shows the market growth for healthcare products valuation to be over $8.3 billion by the year 2030. There is a significant market trend seen toward utilizing technologies for healthcare as was perceived during the COVID-19 pandemic. Thereafter, there was prompt switch from actual services to remote telehealth services in wearable and e-textile form. It is expected that FE will create a huge scope of research and opportunities in numerous fields in the near future.

Table 13.4 Characteristics of varying contact printing technologies (Cruz et al. 2018; Hussain and El-Atab 2022)

Printing technologies	Mechanism and features	Resolution (lines/cm)	Ink viscosity (Pa. s)	Film thickness (μm)	Printing speed (m/min)	Challenges
Screen printing	• Planar or R2R system • Versatile	50	>1–50	up to 12	10 to 15	• Hard to clean • Solvents deteriorate mask patterns • High resolution of uniform line patterns is not possible under 30 μm • Unfeasible use of low viscosity inks to prevent spreading and bleed out • Material wastage
Flexography	• Low-cost patterns plate • High flexibility and low-pressure printing • Better vertical and horizontal pattern quality compared to gravure	60	0.05–0.5	1–2.5	100–500	• Halo effect (patterns with excess of ink) due to printing plate compression to the substrate, despite the applied low pressure • Marbling effect • Complex multi-layers alignment

(continued)

Table 13.4 (continued)

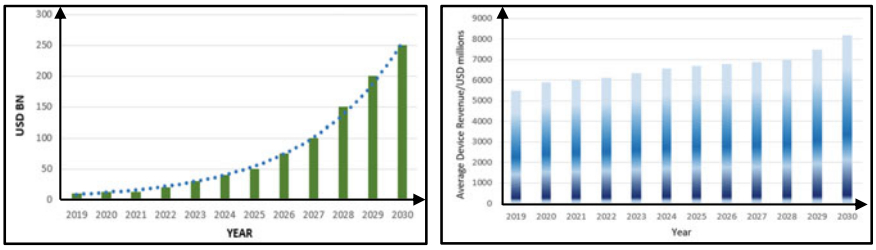
Printing technologies	Mechanism and features	Resolution (lines/cm)	Ink viscosity (Pa. s)	Film thickness (μm)	Printing speed (m/min)	Challenges
Soft lithography	• Encompasses several printing techniques (μCP, REM, μTM, MIMIC, AMIM) • Fabrication of micro-and nanostructures of high quality • Convenient, effective method • Mostly used by the biological science area	30–80	0.01–0.1	0.17–8	Nano-structure device printing	• Require proper adjustment of the surface energies for efficient transfer to the substrate • Common swelling of transferring materials, results in increased features size • Pattern reproduction and resolution is a challenge due to the used forces on stamp • Costly solution
Gravure printing	• High-quality patterns in a cost-effective manner • High speed • Low viscosity inks	100	0.05–0.2	<0.1–5	100–1000	• Cylinder life and high cost • Demanding and careful optimization of the process • A major obstacle is where high resolution is required (e.g., PE)

Table 13.5 Characteristics of varying non-contact printing technologies (Cruz et al. 2018; Hussain and El-Atab 2022)

Printing technologies	Mechanism and features	Resolution (lines/cm)	Ink viscosity (Pa.s)	Film thickness (μm)	Printing speed (m/min)	Challenges
Laser direct printing (LDW)	• 1D to 3D structures • nm to mm magnitude • No mask • Three writing techniques (LDW+ , LDW−, LDWM)	0.7	–	>10	8–12	• High-cost equipment • Not possible to deposit organic substrates, printing only on flat substrates, parallel to the substrate
Aerosol	• Complex design could be printed • Complex conformal surfaces • Many materials and substrates • Non-planar • Low-temperature processing • Local sintering	0.001–1	>0.1	10–250	1.5	• Droplet carrier creates a cloud of powder in surrounding printed area • Sheath gas creates a localized crystallization/ solidification phase at the trace pattern which reduces the quality of the localized bonding layer
Inkjet printing transfer (IPT)	• Low viscosity • Deposition of many types of droplets • Droplets ejection through different actuation phenomena • All type of substrates • Low material wastage • Environmentally friendly	0.002–0.1	0.01–0.5	15–100	15–500	• Printing speed is slow as compared to other techniques • Nozzle clogging

Table 13.6 Conductivity of different ink materials for FE (Reportlinker 2018)

Ink	Conductivity ($\times 10^7$(S/m))
Gold (Au)	6.28
Copper (Cu)	6.01
Silver (Ag)	4.26
Aluminum (Al)	3.77
Platinum (Pt)	0.944
Carbon (C)	0.0002



(a) (b)

Fig. 13.16 **a** Growth prediction in global flexible display market **b** market growth in flexible electronics healthcare devices (IDTechEx 2020; Flexible display market size, share, xxxx)

Applications

The next-generation FE based on advanced materials and mechanical features have ability to immediately integrate components with dynamic surface. As compared to the conventional electronics, the FE are low cost, thin, lightweight, and stretchable that do not require high computational power to operate. FE shows remarkable progress in applications such as wearable electronics, consumer electronics, electronic skin (E-skin), health care, e-textile, robotics, optoelectronics, solar cells, display units, LCD, and OLED. Most of the exciting FE products such as smartphone, smartwatch, and IOT-based devices have been already launched in the market. The stretchability of tactile skin on robotics applications improve adaptability on its curved parts. The FE finds wide application in optoelectronics as well such as in solar cells, LCD, and OLEDs. The FE-based smart wearable devices are useful to monitor daily human activity, and stretchable sensors are used to record plant health. The wide range of endless applications of FE in versatile domains are summarized and shown in Fig. 13.17. The field of FE has been vastly researched in terms of advanced materials, fabrication of novel devices, TFTs, interconnect geometries, and corresponding applications. Therefore, the research and development for FE are exponentially increasing and it shows a breathtaking advancement in world of electronic, science, and research.

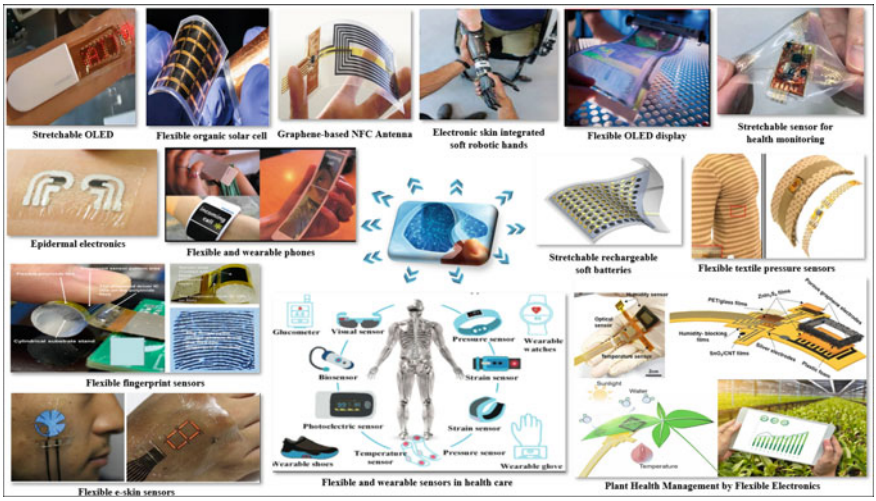


Fig. 13.17 Applications of flexible electronics in versatile domains (Osaka Prefecture University 2020; Lee et al. 2021; Butt et al. 2022; Chang and Lin 2020; Flagship 2016; India 2015; Kim et al. 2019; Nokia phones US. 2023; Grainger Engineering Office 2023; Oliveira 2021; Raphelson 2014; Temporary tattoo offers needle-free way to monitor glucose levels. 2023; Wicaksono et al. 2020; Yokota et al. 2016)

Conclusion

This chapter presents an overview of the various components of FE, materials, printing technologies, applications, and market growth. The interesting evolution from history of VLSI to advanced FE have been briefly presented. Components are the basic building blocks for the formation of FE. The work precisely discusses about the substrate, device, interconnect, and encapsulation methods. In addition, types and materials utilized for different components are reviewed and discussed.

In a process to obtain flexible system, it is found that the conventional electronics based on silicon substrate gets cracked when the device is twisted, folded, or bended. This has been the strong motivation to have research prospective to discover new materials that can overcome the conventional rigidity of device. Moreover, it has been observed that performance can be improved and the cost can also be reduced using the new materials. Materials utilized for the manufacturing of FE are critically important as these are the major foundation for providing flexibility to the FE system. With involvement of intrinsic and stretchable materials such as filler–polymer matrix, organic semiconductor, and rubber-like nanocomposite, it has been observed that the strain of electronic system can be lifted up to 200%. Materials such as organic polymers, CNT-based TFTs, graphene can give high electrical performance even under mechanical stress. Therefore, these have found a great attention in research prospective.

The devices in FE play a vital role. The details of TFTs and their different structures are studied. Stretchable interconnects are another important component in FE. Different stretchable interconnect structures have been explored. Also, *RLC* parasitics of serpentine-based stretchable interconnect have been formulated. Thereafter importance of encapsulation layer and its various materials have also been discussed in the chapter.

In addition to the electrical and mechanical properties of the substrate, device, interconnect, and encapsulation, assembling of these materials is required. For this printing technologies are used. Printing technologies can vary based on their operation and features. FE has the utilization in wide range of applications such as sensors, wearable technology, e-textile, OLED display, flexible batteries, health care, prosthetics, and implantable medical devices. Advancement of FE will lead to the groundbreaking new production of lightweight flexible technologies. However, a wide scope of research is still awaiting for potential researchers in quest of getting the sustainable flexibility along with higher electrical performance. The market growth in FE is huge and can lead to production of various new gadgets and products. From the present chapter, it can be postulated that FE holds huge applications in today's tech-era and is the "*Technology of Tomorrow*".

References

Agrawal Y, Chandel R (2015) Crosstalk analysis of current-mode signalling-coupled rlc interconnects using FDTD technique. IETE Tech Rev 33(2):148–159. https://doi.org/10.1080/02564602.2015.1056258

Agrawal Y, Kumar MG, Chandel R (2016) Comprehensive model for high-speed current-mode signaling in next generation MWCNT bundle interconnect using FDTD technique. IEEE Trans Nanotechnol 15(4):590–598. https://doi.org/10.1109/tnano.2016.2558475

Ausanio G, Barone AC, Campana C, Iannotti V, Luponio C, Pepe GP, Lanotte L (2006) Giant resistivity change induced by strain in a composite of conducting particles in an elastomer matrix. Sens Actuators, A 127(1):56–62. https://doi.org/10.1016/j.sna.2005.12.002

Brinkman WF, Haggan DE, Troutman WW (1997) A history of the invention of the transistor and where it will lead US. IEEE J Solid-State Circuits 32(12):1858–1865. https://doi.org/10.1109/4.643644

Budynas RG, Nisbett JK, Tangchaichit K, Tangchaichit K, Shigley JE (2021) Shigley's Mechanical Engineering Design. McGraw-Hill

Butt MA, Kazanskiy NL, Khonina SN (2022) Revolution in flexible wearable electronics for temperature and pressure monitoring—a review. Electronics 11(5):716. https://doi.org/10.3390/electronics11050716

Chae SH, Lee YH (2014) Carbon nanotubes and graphene towards soft electronics. Nano Convergence 1(1). https://doi.org/10.1186/s40580-014-0015-5

Chang MC, Lin WC (2020) Flexible fingerprint sensor driver IC and bendable evaluation. Int J Circuit Theory Appl 48(12):2190–2201. https://doi.org/10.1002/cta.2874

Crabb RL, Treble FC (1967) Thin silicon solar cells for large flexible arrays. Nature 213(5082):1223–1224. https://doi.org/10.1038/2131223a0

Cruz SM, Rocha LA, Viana JC (2018) Printing Technologies on flexible substrates for printed electronics. Flexible Electron.https://doi.org/10.5772/intechopen.76161

Dang W, Vinciguerra V, Lorenzelli L, Dahiya R (2017) Printable stretchable interconnects. Flexible Printed Electron 2(1):013003. https://doi.org/10.1088/2058-8585/aa5ab2

Dong Z, Duan B, Cao Z, Yang Y (2017) Electromechanical modeling of stretchable interconnects. J Comput Electron 16(1):202–209. https://doi.org/10.1007/s10825-016-0946-7

El-Tantawy F (2005) Development of novel functional conducting elastomer blends containing butyl rubber and low-density polyethylene for current switching, temperature sensor, and EMI shielding effectiveness applications. J Appl Polym Sci 97(3):1125–1138. https://doi.org/10.1002/app.21778

Graphene Flagship (2016) Graphene technology enables fully flexible NFC antennas. ScienceDaily. Retrieved February 23, 2023, from https://www.sciencedaily.com/releases/2016/11/161129084229.htm

Flexible display market size, share, Trends & Forecast 2022–2027. Market Research Company, market research reports and consulting services. Retrieved February 3, 2023, from https://www.imarcgroup.com/flexible-display-market

Fortunato E, Barquinha P, Martins R (2012) Oxide Semiconductor thin-film transistors: a review of recent advances. Adv Mater 24(22):2945–2986. https://doi.org/10.1002/adma.201103228

Grainger Engineering Office of Marketing and Communications. Stretchable battery for flexible circuits. Materials Research Laboratory | UIUC. Retrieved February 23, 2023, from https://mrl.illinois.edu/news/stretchable-battery-flexible-circuits

Hussain MM, El-Atab N (2022) Handbook of flexible and stretchable electronics. CRC Press, Taylor & Francis Group

IDTechEx (2020) Flexible electronics are the future in wearable health monitoring. Flexible electronics are the future in wearable health monitoring. Retrieved February 3, 2023, from https://www.prnewswire.com/news-releases/flexible-electronics-are-the-future-in-wearable-health-monitoring-301049589.html

India PT (2015) Soon, flexible and paper-thin displays with LE-Ofets technology: Study. Gadgets 360. Retrieved February 23, 2023, from https://www.gadgets360.com/laptops/news/soon-flexible-and-paper-thin-displays-with-le-ofets-technology-study-654352

Intel® Products: Processors, Intel® NUC, memory and storage, chipsets. Intel. Retrieved February 3, 2023, from https://www.intel.com/content/www/us/en/products/overview.html

Itoh T, Kobayashi A, Ueno K, Ohta J, Fujioka H (2016) Fabrication of InGaN thin-film transistors using pulsed sputtering deposition. Scientific Reports 6(1). https://doi.org/10.1038/srep29500

Kim YS, Mahmood M, Lee Y, Kim NK, Kwon S, Herbert R, Kim D, Cho HC, Yeo WH (2019) All-in-one, wireless, stretchable hybrid electronics for smart, connected, and ambulatory physiological monitoring. Advanced Science 6(17):1900939. https://doi.org/10.1002/advs.201900939

Kumar MG, Agrawal Y, Chandel R (2016) Carbon nanotube interconnects—a promising solution for VLSI circuits. IETE J Educ 57(2):46–64. https://doi.org/10.1080/09747338.2016.1158129

Lee Y, Chung JW, Lee GH, Kang H, Kim J-Y, Bae C, Yoo H, Jeong S, Cho H, Kang S-G, Jung JY, Lee D-W, Gam S, Hahm SG, Kuzumoto Y, Kim SJ, Bao Z, Hong Y, Yun Y, Kim S (2021) Standalone real-time health monitoring patch based on a stretchable organic optoelectronic system. Sci Adv 7(23). https://doi.org/10.1126/sciadv.abg9180

Li G, Li Q, Jin Y, Zhao Y, Xiao X, Jiang K, Wang J, Fan S (2015) Fabrication of air-stable N-type carbon nanotube thin-film transistors on flexible substrates using bilayer dielectrics. Nanoscale 7(42):17693–17701. https://doi.org/10.1039/c5nr05036g

Lin Y-Y, Gundlach DI, Nelson SF, Jackson TN (1997) Pentacene-based organic thin-film transistors. IEEE Trans Electron Dev 44(8):1325–1331. https://doi.org/10.1109/16.605476

Lu N, Jiang W, Wu Q, Geng D, Li L, Liu M (2018) A review for compact model of thin-film transistors (TFTs). Micromachines 9(11):599. https://doi.org/10.3390/mi9110599

Nokia phones US. The latest Nokia phones and accessories. Retrieved February 23, 2023, from https://www.nokia.com/phones/en_us/

Oliveira C (2021) Advanced materials and technological evolution. AstroPT. Retrieved February 23, 2023, from https://www.astropt.org/2021/07/04/os-materiais-avancados-e-a-evolucao-tecnologica/

Osaka Prefecture University (2020) Plant Health Management by flexible electronics. Retrieved February 3, 2023, from https://phys.org/news/2020-08-health-flexible-electronics.pdf

Pathade T, Agrawal Y, Parekh R, Kumar MG (2021) Structure fortification of mixed CNT bundle interconnects for nano integrated circuits using constraint-based particle swarm optimization. IEEE Trans Nanotechnol 20:194–204. https://doi.org/10.1109/tnano.2021.3058760

Qaiser N, Khan SM, Nour M, Rehman MU, Rojas JP, Hussain MM (2017) Mechanical response of spiral interconnect arrays for highly stretchable electronics. Appl Phys Lett 111(21):214102. https://doi.org/10.1063/1.5007111

Rabaey JM, Chandrakasan AP, Nikolić Borivoje (2017) Digital integrated circuits: a design perspective. Pearson Education

Raphelson S (2014) The prosthetics industry gets a human touch, NPR. NPR. Available at: https://www.npr.org/2014/11/11/358048818/the-prosthetics-industry-gets-a-human-touch. Accessed November 15, 2022

Ray KA (1967) Flexible solar cell arrays for increased space power. IEEE Trans Aerosp Electron Syst AES-3(1):107–115. https://doi.org/10.1109/taes.1967.5408720

Reportlinker (2018) Conductive ink markets 2012–2018 silver and copper inks and pastes and beyond. Conductive Ink Markets 2012–2018 Silver and Copper Inks and Pastes and Beyond. Retrieved February 3, 2023, from https://www.prnewswire.com/news-releases/conductive-ink-markets-2012-2018-silver-and-copper-inks-and-pastes-and-beyond-166091036.html

Sakuma KATSUYUKI (2022) Flexible, wearable, and Stretchable Electronics. CRC Press

Sankir ND (2005) Flexible electronics: Materials and device fabrication

Shang Z-W, Hsu H-H, Zheng Z-W, Cheng C-H (2019) Progress and challenges in P-type oxide-based thin film transistors. Nanotechnol Rev 8(1):422–443. https://doi.org/10.1515/ntrev-2019-0038

Singh E, Srivastava R, Kumar U, Katheria A (2017) Carbon nanotube: a review on introduction, fabrication techniques and optical applications. J Nanosci Nanotechnol 4:120–126. https://doi.org/10.12691/nnr-4-4-1

Temporary tattoo offers needle-free way to monitor glucose levels. Today. Retrieved February 23, 2023, from https://today.ucsd.edu/story/temporary_tattoo_offers_needle_free_way_to_monitor_glucose_levels

Verplancke R, Bossuyt F, Cuypers D, Vanfleteren J (2011) Thin-film stretchable electronics technology based on meandering interconnections: fabrication and mechanical performance. J Micromech Microeng 22(1):015002. https://doi.org/10.1088/0960-1317/22/1/015002

Walsh KMN, Genzer P (2012) The world's first sterilizable flexible organic transistor. Brookhaven National Laboratory. Retrieved February 3, 2023, from https://www.bnl.gov/newsroom/news.php?a=111396

Weimer P (1962) The TFT a new thin-film transistor. Proc IRE 50(6):1462–1469. https://doi.org/10.1109/jrproc.1962.288190

Wicaksono I, Tucker CI, Sun T, Guerrero CA, Liu C, Woo WM, Pence EJ, Dagdeviren C (2020) A tailored, electronic textile conformable suit for large-scale spatiotemporal physiological sensing in vivo. Npj Flexible Electron 4(1). https://doi.org/10.1038/s41528-020-0068-y

Yang S, Ng E, Lu N (2015) Indium Tin Oxide (ITO) serpentine ribbons on soft substrates stretched beyond 100%. Extreme Mech Lett 2:37–45. https://doi.org/10.1016/j.eml.2015.01.010

Yokota T, Zalar P, Kaltenbrunner M, Jinno H, Matsuhisa N, Kitanosako H, Tachibana Y, Yukita W, Koizumi M, Someya T (2016) Ultraflexible organic photonic skin. Sci Adv 2(4). https://doi.org/10.1126/sciadv.1501856

Yu M, Wan H, Cai L, Miao J, Zhang S, Wang C (2018) Fully printed flexible dual-gate carbon nanotube thin-film transistors with tunable ambipolar characteristics for complementary logic circuits. ACS Nano 12(11):11572–11578. https://doi.org/10.1021/acsnano.8b06748

Zhang Y, Fu H, Xu S, Fan JA, Hwang K, Jiang J, Rogers JA, Huang Y (2014) A hierarchical computational model for stretchable interconnects with fractal-inspired designs. J Mech Phys Solids 72:115–130. https://doi.org/10.1016/j.jmps.2014.07.011

Chapter 14
Delay Analysis of Different Stretchable Interconnect Structures

Kavicharan Mummaneni, Vivek Kumar, Malvika, and Yash Agrawal

Abstract Stretchable electronic systems are needed in realizing a wide range of applications, such as wearable healthcare monitoring, where stretching movements are present. Current electronics and sensors are rigid and non-stretchable. However, after integrating with stretchable interconnects, the overall system is able to withstand a certain degree of bending, stretching, and twisting. In this chapter, extraction of the parasitic parameters of a wire and an analytical model is developed based on the skin effect of stretchable interconnects. Analytical models are employed to estimate the delay of various stretchable interconnects and the results are found to be in good accuracy with the simulation results. Finally, the proposed model is employed for comparing the simulation results of circular, rectangular, triangular, and horseshoe stretchable interconnects over a wide frequency range up to 10 GHz.

Keywords Stretchable Interconnects · Healthcare monitoring · ANSYS · Substrate materials · Skin depth

Introduction

Electronic circuits those are flexible and stretchable are a new evolving concept that aims to improve the comfort of consumers. This technology can be used in various applications, such as medical devices that require the circuit to behave as the tissue itself. Currently, traditional power and signal transmission lines are typically placed on rigid or flexible materials. These lines limit the commercial applications

K. Mummaneni (✉) · V. Kumar · Malvika
Electronics and Communication Engineering Department, National Institute of Technology, Silchar, Assam, India
e-mail: kavicharan@ece.nits.ac.in

Y. Agrawal (✉)
VLSI and Embedded Systems Research Group, Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar 382007, Gujarat, India
e-mail: mr.yashagrawal@gmail.com

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_14

of these components due to their size and stiffness. With the use of stretchable electronic circuits, different components can be integrated onto a single flexible surface (Dang 2018). One of the main challenges in the development of stretchable electronic circuits is the reliability of their elastic interconnections (Dong et al. 2017). This is because these components are typically subjected to high mechanical loads. For instance, if a component is placed on a flexible surface, it can potentially bend or even get deformed. In-plane patterned metal conductors, intrinsic conductive polymers, and pre-stressed metal conductors have been proposed as possible solution levels of performance and cost to realize. In this work, we present a method that uses the finite element analysis (FEA) to optimize the shape of metal interconnections. This process allows to create a two-dimensional spring by connecting two points along a periodic metal track. Because of high stretchability and low elastic modulus, polydimethylsiloxanes (PDMS) (Lacour et al. 2005) are widely used in various electronic and medical devices. The first step in this process is to perform 3D FEM simulations to identify the ideal shapes for the various components of the circuit. The results of these simulations were then used to identify the promising structures that could be used to improve the performance of the circuit. The results of mechanical modeling are then used to calculate the integrity of a circuit when it is flexed or stretched.

Conventional electronic products are typically rigid and non-flexible. Due to their high rigidity, semiconductor circuit boards are not ideal for protecting integrated electronic devices. In the past few years, the use of flexible electronics has been widely used in various fields, such as medical devices, sensors, and energy collection (Hess-Dunning et al. 2013; Zhao et al. 2019). Due to the increasing number of achievements in the field of flexible electronics, the market for this technology is expected to grow significantly. There are two main types of flexible electronic circuits: organic and inorganic. The former uses organic materials such as flexible semiconductors, while the latter utilizes mature inorganic semiconductor processes. Despite the rapid development of organic electronics, the electrical properties of these materials remain far apart from those of their mature counterparts. To bridge this gap, flexible materials (Gao et al. 2017; Vroman and Tighzert 2009; Case et al. 2015; Khan et al. 2015) can be used as the interface materials for silicon device modules (Stoyanov et al. 2013). This design concept allows flexible electronic devices to be more robust and flexible than their non-flexible counterparts. Polymers can be used as a replacement for silicon bodies in flexible electronic devices, which are more cost-effective than traditional ones. Due to their properties, flexible electronic devices are expected to have a wide range of applications in various fields. In the past, research on the design of flexible interconnects has been focused on improving their ductility (Lipomi et al. 2012; Jeong et al. 2016). This article presents an analytical model that takes into account the electrical and skin-depth effects of stretchable interconnects. The model takes into account the various factors that affect the inductance and skin-depth effect of the wire (Plovie et al. 2017; Salvatore et al. 2017; Jinno et al. 2017; Wang et al. 2018; Johnston et al. 2014; Blau et al. 2011; Tybrandt et al. 2018; Carlson et al. 2012). An analytical model of a stretchable interconnect equivalent circuit is then established. A flexible electronic design platform is also proposed to study the various aspects

of the design of flexible electronic devices. Through the use of Ansys simulation software, the model can be used to analyze the electrical properties of a stretched wire.

The accuracy of the proposed stretchable interconnect model is verified through this platform. The parameters of the flexible interconnect are extracted from the model to help the user choose a wire that is ideal for their electronic circuit. In addition, the parasitic parameters of the flexible interconnect can be analyzed to improve the signal integrity of the electronic circuit. Due to the nature of the flexible interconnects, it is necessary to analyze their electrical properties when subjected to mechanical deformation. This is beneficial for the development of wearable electronics and other flexible electronic devices. Besides being used in medical and military applications, stretchable electronics can also be used in flexible displays and sensors. One of the most common structures that are ideal for the development of stretchable electronic devices is the island-bridge structure. This type of structure is made up of metal lines that are connected to flexible islands on an elastomeric surface. The metal lines can be configured in various shapes to accommodate different mechanical stresses. Several types of metal lines have been presented in recent years, such as the horseshoe shape, the circular shape, the rectangular shape, and the zigzag shape (Sosin 2011). These are characterized by their unique properties, which allow them to endure multi-axial loads. A combination of these designs can be used to create self-similar structures, which can be easily integrated into a unit cell. Despite the advantages of these types of structures, their design is still a challenge for stretchable interconnects.

Apart from their physical properties, these also have a couple of issues that can affect their reliability. One of these is the formation of metal lines, which can cause adhesive fracture between the stretchable substrate and the metal lines. Various studies have been conducted on the formation of metal lines in stretchable interconnects. One of these studies focused on the stress distribution of these structures. In order to understand the delamination behavior of the horseshoe shape (Sosin 2011), a finite element model was developed based on the zone method. However, this method can be very challenging to implement due to the various convergent problems. In recent times, the selection of substrate materials, fabrication techniques of stretchable interconnect structures, and different geometric structures along with modeling methods have been investigated (Dong et al. 2017; Lacour et al. 2005). Also the equivalent models of stretchable interconnect in order to find out the delay of interconnect structures are reported in Li et al. (2018).

Modeling Methodology of Stretchable Interconnects

Figure 14.1 shows the stretchable interconnect structures of effective length 450 μm. At first interconnect structure has to be designed in the ANSYS geometry or else it can be imported from other design tools like ANSYS Q3d, FreeCAD, Autodesk, etc. In ANSYS workbench, static structural will be used for the analysis of stretchable interconnect structures, by which the parameters like strain, stress, and displacement

can be calculated. Once stretching is performed with the interconnect structure, mechanical model can be extracted. The extracted model then needs to be imported into the Q3D extractor to extract the resistance, inductance, and capacitance values. Once the RLC values were extracted, delay offered by the interconnect structures can be estimated using delay models for different stretching, different widths, and also for different frequencies.

Exploration of Stretchable Interconnects

The design platform for electronic devices is used to perform delay analysis on the structure and flow of stretchable interconnects. This step involves extracting the associated deformation model from the simulation. The model is then imported into the Q3D Extractor, which allows the user to extract the parameters of the deformation model. The simulation is then performed in the Advanced Design System to get the delay of stretchable interconnects.

Equivalent Circuit

The basic equivalent model for an electronic component is the network model. For a single stretchable interconnect, the proposed circuit model is shown in Fig. 14.2. However, since the model excludes the effects of electromagnetic fields on the wire, it is not applicable to high frequencies.

The influence of the signal and the electromagnetic field on the wire is mainly reflected in the R_f and L_{self} of the proposed circuit model. In addition, the deformation of the wire will affect its parasitic parameters. This is an unavoidable problem in flexible circuit design.

Delay Calculation

From the equivalent circuit, the transfer function can be obtained as

$$\mathrm{H(S)} = \frac{1}{1 + b_1 S + b_2 S^2} \tag{14.1}$$

where

$$b_1 = RC_c + RC_L + R_S C_c + R_S C_L \tag{14.2}$$

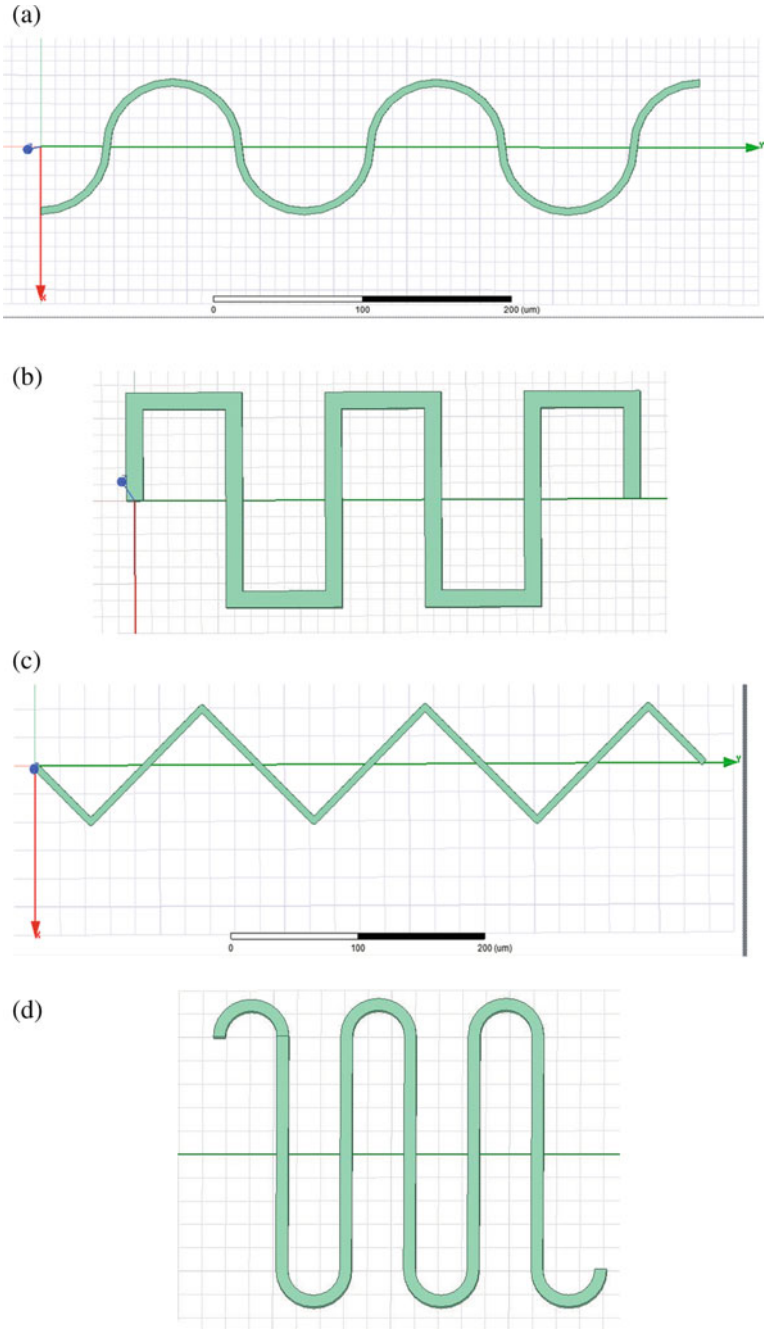


Fig. 14.1 Stretchable interconnect structures in Q3D Extractor **a** Circular, **b** Rectangular, **c** Triangular and **d** Horseshoe

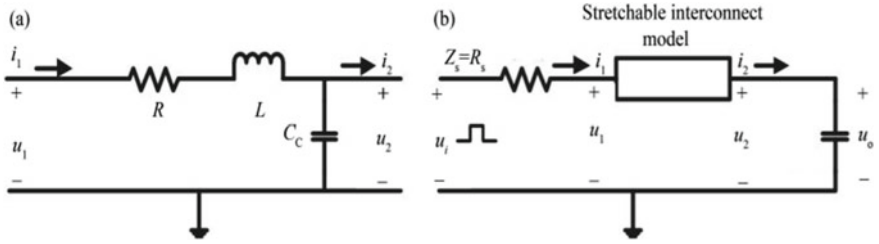


Fig. 14.2 Equivalent circuit **a** RLC network model **b** model including driver and load

$$b_2 = LC_c + LC_L \tag{14.3}$$

The delay of the equivalent circuit is measured by applying step response as input and the Laplace transform of step response is analyzed as

$$U_i(S) = \frac{V_i}{S} \tag{14.4}$$

Then, the output response of the system will be evaluated as

$$U_o(S) = \frac{V_i}{S(1 + b_1 S + b_2 S^2)} \tag{14.5}$$

Finally, the time domain response of the above equation is obtained and given as

$$u_0(t) = V_i\left(1 - e^{\frac{-\alpha t\sqrt{\alpha^2+\beta^2}}{\beta}} \sin \beta t + \rho\right) \tag{14.6}$$

where,

$$\alpha = \frac{b_1}{2b_2}, \quad \beta = \frac{\sqrt{4b_2 - b_1^2}}{2b_2} \quad \rho = \tan^{-1}\left(\frac{\alpha}{\beta}\right) \tag{14.7}$$

Results and Discussion

The interconnect structure is designed in the ANSYS tool, in which static structural are used for the analysis of stretchable interconnect structures. Later, the parameters like strain, stress, and displacement are calculated. Once stretching is performed with the interconnect structure, mechanical model is extracted. The extracted model is then imported into the Q3D extractor to extract the resistance, inductance, and

capacitance values. Once the RLC values are extracted, the delay model is used to estimate the delay offered by the interconnect structure for different stretching, different widths, and also for different frequencies.

Tables 14.1, 14.2, 14.3, and 14.4 show the variation of resistance, inductance, and capacitance with respect to the different stretching percentages of various interconnect structures. At low frequencies, the conductor's exterior and interior contain electromagnetic energy and current conduction. However, at higher frequencies the current flow is primarily focused close to the conductor's surface. The magnetic field is hence limited to the area outside the conductor. The skin-depth effect thus begins to show at high frequency. Very less current flows in the core of the wire as all fields are minimal at a deep enough depth. As a result, the effective cross-section of the conductor gets smaller as the frequency rises. The resistance of the wire becomes frequency dependent as the skin depth falls with increasing frequency. Consequently, the resistance of wire rises with frequency. Hence, the variation in the resistance is more due to the effect of skin depth at higher frequencies.

The magnetic field surrounding the interconnect changes as its shape does. A straight wire of the same length produces stronger magnetic fields than a stretchable interconnect. The inductance, as is widely known, is the quantity of field line rings around a conductor for each amp of current passing through it. As a result, a stretchable interconnect has a higher inductance than a straight wire of the same length. In

Table 14.1 Variation in RLC parameters with w = 5 um wide for circular stretchable interconnect structure

Circular interconnect	Resistance (mΩ)	Inductance (pH)	Capacitance (fF)
0% stretched	545.986	579.996757	6.923595
50% stretched	546.442	679.828540	7.522967
100% stretched	551.845	796.759518	8.302950

Table 14.2 Variation in RLC parameters with w = 5 um wide for rectangular stretchable interconnect structure

Rectangular interconnect	Resistance (mΩ)	Inductance (pH)	Capacitance (fF)
0% stretched	525.016	568.24	7.01
50% stretched	532.562	625.71	7.87
100% stretched	548.265	701.27	8.26

Table 14.3 Variation in RLC parameters with w = 5 um wide for triangular stretchable interconnect structure

Triangular interconnect	Resistance (mΩ)	Inductance (pH)	Capacitance (fF)
0% stretched	544.347	623.965481	7.172989
50% stretched	550.659	757.299361	7.972990
100% stretched	555.258	933.304827	9.291157

Table 14.4 Variation in RLC parameters with w = 5 um wide for horseshoe stretchable interconnect structure

Horseshoe interconnect	Resistance (mΩ)	Inductance (pH)	Capacitance (fF)
0% stretched	540.246	569.17	7.43
50% stretched	542.452	635.40	7.87
100% stretched	550.645	801.28	8.25

other words, as stretchable interconnect is stretched, its inductance rises. Similarly, the capacitance of a stretchable interconnect has a higher capacitance than a straight wire of the same length.

Tables 14.5, 14.6, 14.7, and 14.8 show the delay values of circular, rectangular, triangular, and horseshoe structures for different widths, stretching percentages, and at various frequencies. It has been observed that, as stretching and frequency increases, delay increases due to the effect of skin depth. Stretching the interconnect has a minimal impact on the signal when the signal frequency is small, that is, when the signal period is relatively large. However, when the signal frequency is high, that is, when the signal period is relatively small, the effect of delay on the signal increases. The tables denote the delays of a stretchable interconnect with stretchy interconnects at 1 GHz, 5 GHz, and 10 GHz under various applied stresses. The transmission performance of the stretchy connection declines as applied strain causes the impedance of the conductor to rise. When applied stresses of 0% to 100% are taken into account, it is observed that the delay in interconnect gets worse at 10 GHz frequency. As the phase shift is worse, correspondingly the delay gets longer. This behavior thus demonstrates how, as applied strain grows, a stretchable interconnects impedance rises and its transmission performance degrades.

Table 14.5 Comparison of delay of Circular interconnect structure operating at different widths and operating frequency

Circular interconnect	F = 10 GHz Initial 100% Stretched	F = 5 GHz Initial 100% Stretched	F = 1 GHz Initial 100% Stretched
W = 5 um	51.50 ps 89.10 ps	49.20 ps 84.58 ps	48.10 ps 81.84 ps
W = 10 um	45.52 ps 82.61 ps	43.62 ps 78.70 ps	42.83 ps 76.70 ps
W = 15 um	41.30 ps 75.98 ps	39.74 ps 72.74 ps	31.15 ps 71.20 ps

Table 14.6 Comparison of delay of Rectangular interconnect structure operating at different widths and operating frequency

Rectangular interconnect	F = 10 GHz Initial 100% Stretched	F = 5 GHz Initial 100% Stretched	F = 1 GHz Initial 100% Stretched
W = 5 um	49.12 ps 74.56 ps	48.94 ps 70.01 ps	46.21 ps 69.95 ps
W = 10 um	45.56 ps 65.17 ps	43.27 ps 62.26 ps	41.42 ps 61.23 ps
W = 15 um	40.31 ps 60.39 ps	38.25 ps 59.45 ps	38.03 ps 57.15 ps

Table 14.7 Comparison of delay of Triangular interconnect structure operating at different widths and operating frequency

Triangular interconnect	F = 10 GHz Initial 100% Stretched	F = 5 GHz Initial 100% Stretched	F = 1 GHz Initial 100% Stretched
W = 5 um	46.97 ps 71.01 ps	44.94 ps 67.85 ps	43.90 ps 66.15 ps
W = 10 um	41.35 ps 62.22 ps	39.63 ps 59.56 ps	38.92 ps 58.85 ps
W = 15 um	37.38 ps 57.35 ps	35.95 ps 55.45 ps	35.39 ps 54.65 ps

Table 14.8 Comparison of delay of Horseshoe interconnect structure operating at different widths and operating frequency

Horseshoe interconnect	F = 10 GHz initial 100% stretched	F = 5 GHz initial 100% Stretched	F = 1 GHz initial 100% Stretched
W = 5 um	54.26 ps 92.34 ps	51.81 ps 87.46 ps	49.90 ps 83.25 ps
W = 10 um	48.35 ps 85.42 ps	45.54 ps 81.21 ps	44.12 ps 78.82 ps
W = 15 um	44.19 ps 78.75 ps	41.95 ps 73.24 ps	40.87 ps 71.25 ps

Conclusions

A new analytical model for stretchable interconnects is created taking into account the skin-depth effect. Using a flexible electronics design platform, a model for the delay analysis of stretchable interconnects for various extensions is presented. The analysis of parasitic extraction for stretchable interconnects is established. It is analyzed that if the interconnect structure is stretched, then resistance remains almost constant while inductance and capacitance values increase, which leads to increase in delay. In this chapter, the parasitic comparison of various interconnect structures with different stretching percentages and frequencies is presented. Due to the changes in RLC values, it is investigated that delay decreases as width increases. The parasitic values depend on the frequency at which stretchable device operates. It is observed that as the frequency of operation increases, the delay of the stretchable interconnects increases due to skin effect. As a future scope, this work can be extended to the analysis of crosstalk for different stretchable interconnect structures.

References

Blau A, Murr A, Wolff S, Sernagor E, Medini P, Iurilli G, Ziegler C, Benfenati F (2011) Flexible, all-polymer microelectrode arrays for the capture of cardiac and neuronal signals. Biomaterials 32(7):1778–1786

Carlson A, Bowen AM, Huang Y, Nuzzo RG, Rogers JA (2012) Transfer printing techniques for materials assembly and micro/nanodevice fabrication. Adv Mater 24(39):5284–5318

Case JC, White EL, Kramer RK (2015) Soft material characterization for robotic applications. Soft Robot 2(2):80–87

Dang W (2018) Stretchable interconnects for smart integration of sensors in wearable and robotic applications

Dong Z, Duan B, Boxing C, Zhen Y, Yintang (2017) Electromechanical modeling of stretchable interconnects. J Comput Electron., 16. https://doi.org/10.1007/s10825-016-0946-7

Gao Y, Ota H, Schaler EW, Chen K, Zhao A, Gao W, Fahad HM, Leng Y, Zheng A, Xiong F, Zhang, Tai L-C, Zhao P, Fearing RS, Javey A (2017) Wearable microfluidic diaphragm pressure sensor for health and tactile touch monitoring. Adv Mater 29(39):(1–8) 1701985

Hess-Dunning AE, Tyler DJ, Zorman CA (2013) Stretchable thin-film metal structures on a stimuli-responsive polymer nanocomposite for mechanically-dynamic microsystems. Transducers 2013, Barcelona, SPAIN, pp 2229–2232

Jeong SH, Zhang S, Hjort K, Hilborn J, Wu Z (2016) PDMS-based elastomer tuned soft, stretchable, and sticky for epidermal electronics. Adv Mater 28:5830–5836

Jinno H, Fukuda K, Xu X, Park S, Suzuki Y, Koizumi M, Yokota T, Osaka I, Takimiya K, Someya T (2017) Stretchable and waterproof elastomer-coated organic photovoltaics for washable electronic textile applications. Nat Energy 2(10):780–785

Johnston ID, McCluskey DK, Tan CKL, Tracey MC (2014) Mechanical characterization of bulk Sylgard 184 for microfluidics and microengineering. J Micromechanics Microeng 24(3): 0350117 (1–7)

Khan S, Yogeswaran N, Taube W, Lorenzelli L, Dahiya R (2015) Flexible FETs using ultrathin Si microwires embedded in solution processed dielectric and metal layers. J Micromechan Microeng 25(12):125019

Lacour SP, Jones J, Wagner S, Li T, Suo Z (2005) Stretchable interconnects for elastic electronic surfaces. Proc IEEE 93(8):1459–1467

Li J, Duan B, Dong Z, Yang Y (2018) Analysis of delay from step response based on stretchable flexible interconnects. Sci China Inf Sci 61(10):1–3

Lipomi DJ, Lee JA, Vosgueritchian M, Tee BC-K, Bolander JA, Bao Z (2012) Electronic properties of transparent conductive films of PEDOT: PSS on stretchable substrates. Chem Mater 24(2):373–382

Plovie BB, Yang Y, Guillaume J, Dunphy S, Dhaenens K, Van Put S, Vervust T, Bossuyt F, Vanfleteren J (2017) Arbitrarily shaped 2.5D circuits using stretchable interconnects embedded in thermoplastic polymers. Adv Eng Mater 19(8):1700032 (1–8)

Salvatore GA, Sülzle J, Valle FD, Cantarella G, Robotti F, Jokic P, Knobelspies S, Daus A, Büthe L, Petti L, Kirchgessner N, Hopf R, Magno M, Tröster G (2017) Biodegradable and highly deformable temperature sensors for the internet of things. Adv Funct Mater 27(35):1702390 (1–10)

Sosin S (2011) Interconnect schemes for stretchable array-type microsystems

Stoyanov H, Kollosche M, Risse S, Waché R, Kofod G (2013) Soft conductive elastomer materials for stretchable electronics and voltage controlled artificial muscles. Adv Mater 25(4):578–583

Tybrandt K, Khodagholy D, Dielacher B, Stauffer F, Renz AF, Buzsáki G, Vörös J (2018) High-density stretchable electrode grids for chronic neural recording. Adv Mater 30(15):1706520 (1–7)

Vroman I, Tighzert L (2009) Biodegradable polymers. Materials (basel) 2(2):307–344

Wang S, Xu J, Wang W, Wang G-JN, Rastak R, Molina-Lopez F, Chung JW, Niu S, Feig VR, Lopez J, Lei T, Kwon S-K, Kim Y, Foudeh AM, Ehrlich A, Gasperini A, Yun Y, Murmann B, Tok JB-H, Bao Z (2018) Skin electronics from scalable fabrication of an intrinsically stretchable transistor array. Nature 555:83–88

Zhao Y, Kim A, Wan G, Tee BCK (2019) Design and application of stretchable and self-healable conductors for soft electronics. Nano Convergence 6(1):1–22

Chapter 15
Flexible Sensors for Plant Disease Monitoring

Kamlesh S. Patle, Priyanka Khaparde, Gulafsha Bhatti, Yash Agrawal, and Vinay S. Palaparthy

Abstract Advancement in the flexible substrate led to the sensor fabrication, which offers light weight and low-cost sensors. Such sensors are very important to monitor the plant health on regular intervals and describe the germination of the plant microbial diseases. Numerous climatic and soil characteristics, including rainfall, soil moisture, leaf wetness duration, ambient temperature, and ambient humidity, are used to monitor plant disease. In this paper, we described the reported work for sensing the aforementioned parameters on the flexible substrates. The emphasis has been given on the fabrication of these sensors and their sensor transfer function such as sensitivity, response time, stability, and hysteresis.

Keywords Plant disease · Flexible sensors · Leaf wetness · Soil and ambient temperature · Humidity · Rainfall

Introduction

The world population has been increased by 2 billion people by the year 2019—a huge rise of around 25% (United Nations, Department of Economic and Social Affairs 2019). The Food and Agricultural Organization (FAO) research estimates that an additional 70–90% more food will be needed to feed this population. About 34% of the total crop damaged in the agriculture due to weeds, whereas animal pest contributes around 18%. Further, nearly 16% of the world's total agricultural crop production has been damaged by microbial diseases. An enormous proportion of crop productivity is lost each year due to weed infestations, pest problems, and plant diseases (United Nations, Department of Economic and Social Affairs 2019; Iqbal

K. S. Patle · P. Khaparde · G. Bhatti · Y. Agrawal · V. S. Palaparthy (✉)
System Design Lab, Department of Information and Communication Technology, DAIICT, Gandhinagar, Gujarat 382007, India
e-mail: vinay_shrinivas@daiict.ac.in

Y. Agrawal
e-mail: yash_agrawal@daiict.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_15

et al. 2018; Jain et al. 2019; Chouhan et al. 2018; Fang and Ramasamy 2015). Such losses have a significant negative impact on both the country's economy and the farmer's way of life in nations where farming is the main source of income for a large portion of the population. Crop losses are caused by agriculture's sporadic use of technology methods. For the effective crop growth, it is important to monitor the germination of plant disease, attack of the animal pest, plant transpiration, monitoring drought-like situation in plants, etc. However, this work focuses on the sensors for the plant disease monitoring factors affecting it.

Germination of the plant microbial disease is based on a variety of environmental and meteorological factors that affect a pathogen's ability to survive. When a pathogen gets into contact with a vulnerable host, it has the potential to infect and seriously harm agricultural production. Plant disease lead to a decrease in both the amount and quality of agricultural production (Iqbal et al. 2018). Fungi, which can be found in plant leaves, is one of the most common diseases. Over 70–80% of plant diseases are caused by fungi a major plant pathogen found in the plants (Jain et al. 2019). Disease and disorder are the two main elements that have an impact on crops and their output. The biotic factors that cause disease are either fungi, bacteria, or algae while the abiotic variables that promote disorder are the atmospheric conditions (soil and ambient temperature, rainfall, soil moisture, ambient humidity, and leaf wetness duration (LWD)) (Chouhan et al. 2018). The production could be drastically reduced if these contagious crop illnesses are not promptly treated, endangering the security of the global food supply. There is a necessity for better disease detection to stop crop damage in order to reduce the spread of illnesses, increase productivity, and ensure agricultural sustainability. As a result, experts are keen for early plant predication and provide the instant control measures. Farmers can be assisted in preserving their crops via early disease identification and provision of management methods. These precautions include techniques for diagnosing diseases directly or indirectly. Laboratory-based techniques include the majority of direct detection techniques, while optical sensors are used in indirect detection techniques such hyperspectral imaging, fluorescence imaging, and thermography (Fang and Ramasamy 2015; Golhani et al. 2018; Kumar et al. 2021). The intricacy of the data gathered and the volume of data collected are the limitations of different optical sensing systems. These strategies demand expensive setup and computing costs as well as expertise in data analytics and statistical methodologies, to be used effectively. For this purpose, sensor technology has a potential role to play in plant disease monitoring which is important to monitor soil and ambient temperature, rainfall, soil moisture, ambient humidity, and leaf wetness duration.

The existing sensors-based techniques to monitor the aforementioned environmental and soil parameters have been mounted on the printed circuit board technology (PCB). The PCB-based sensor technology is cost effective and offers high reliability, but the major drawback with this technology is that the fabricated sensor's weight is in grams (typically more than 100 gm). Due to this when such sensors are mounted on the leaf tends to bend the leaf. Additionally, another drawback with the PCB-based sensors is that when these sensors are mounted on the leaf canopy, it does not take the shape of the leaf due to uneven surface of the leaf. As a result, there

is reduced contact resistance between the sensor and leaf, which tends to produce incorrect results. Thus, there is dearth of sensors to be fabricated on the flexible substrates, which offer more contact between the sensor and leaf and also light in weight (Patle et al. 2022, 2021a, b). In further subsequent sections, we will discuss on the fabrication of the flexible humidity, temperature, leaf wetness sensors, rainfall sensor, which is used to monitor the plant microbial disease.

Flexible Humidity Sensors

Sensor Fabrication and Experimental Set-Up

An illustration of the construction of humidity sensors made on a Polyethylene terephthalate (PET) substrate is shown in Fig. 15.1 (Su and Wang 2007). A PET film was used to manufacture the interdigitated gold electrodes using a low-temperature sputtering process. On an alumina substrate, the same interdigitated gold electrodes are created. The following steps were taken in order to prepare the MMA/MAPTAC precursor solution: A highly homogenized solution was created by combining MAPTAC (0.1 ml, 50% solution, Aldrich), MMA (1 ml, 99%, Merck), and azobisisobutyronitrile (AIBN, 0.01 g) and to create a thoroughly homogenized solution, ethanol (0.6 ml) was added after that. Following heating at 90 °C for 1 h in the air to cause co-polymerization to occur, the mixed solution was deposited on the PET substrate using the dip-coating comprising two interdigitated gold electrodes, and the film thickness obtained was around 3.5 mm. This led to the creation of a flexible resistive-type humidity sensor as depicted in Fig. 15.1.

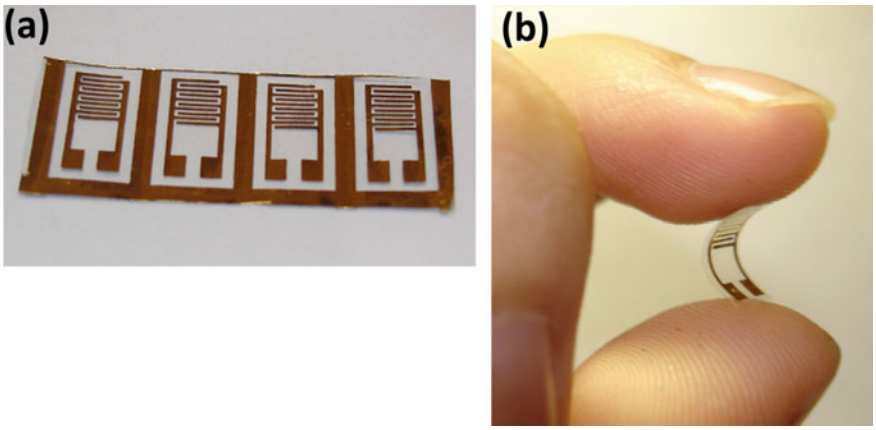


Fig. 15.1 **a** Humidity sensors mounted on a PET substrate, **b** flexible humidity sensor

A test box was used to assess a sensor's complex impedance as a function of relative humidity (RH) with the help of LCR meter. Figure 15.1a A picture of the humidity sensors on a PET substrate. Fabricated humidity sensor on the flexible PET substrates is shown in Fig. 15.1b. By using mass flow controls (Hastings), dry and wet air is combined. A standard hygrometer (that had already been calibrated in the humidity lab of the National Measurement Laboratory (NML)) was used to calibrate the setting humidity and temperature points. At 15, 25, and 35 °C, the humidity ranged from 10 to 90%RH and the frequency for impedance varied from 1 to 100 kHz.

Sensor Transfers Function

Figure 15.2a displays sensor response (sensitivity) where excitation frequency in the LCR is kept at 1 V, and the impedance vs flexible RH sensors has been studied (Su and Wang 2007). For both devices, the frequency is applied from frequency range of 1–100 kHz and effect of impedance with respect to RH is studied. Further from Fig. 15.2a, it can be referred that sensor impedance decreases with an increase in the RH for all the frequencies (Su and Wang 2007).

Further, Fig. 15.2b illustrates that the hysteresis is studied, where it was found that the fabricated sensor offers hysteresis of about 6% RH. Figure 15.2c depicts the flexible humidity sensor's response and recovery characteristics. Response time ranged between 45 and 150 s for desiccation from 73 to 5% RH and 45 to 150 s for humidification from 7 to 78% RH. The sensor's reaction and recovery times took the equilibration period for the water vapor inside the testing chamber. As a result, the flexible humidity sensor's actual response and recovery durations ought to be substantially lower. Figure 15.2d depicts the impact of long-term stability. For at least 120 days, the impedance of the flexible humidity sensor did not differ significantly for the testing points of 10%, 30%, 60%, and 90% RH (Su and Wang 2007).

Flexible Temperature Sensors

Sensor Fabrication and Experimental Set-Up

A fully printable approach is used to create the proposed temperature sensor [12]. Printing has many benefits such as a simple manufacturing method, including low cost, simple patterning, and good precision in device performance. An Ag electrode further PEN substrate, temperature layer and followed by CYTOP encapsulation layer is incorporated to construct the temperature sensor's structure, depicted in Fig. 15.3a. According to Fig. 15.3b, the temperature sensing layer is created using

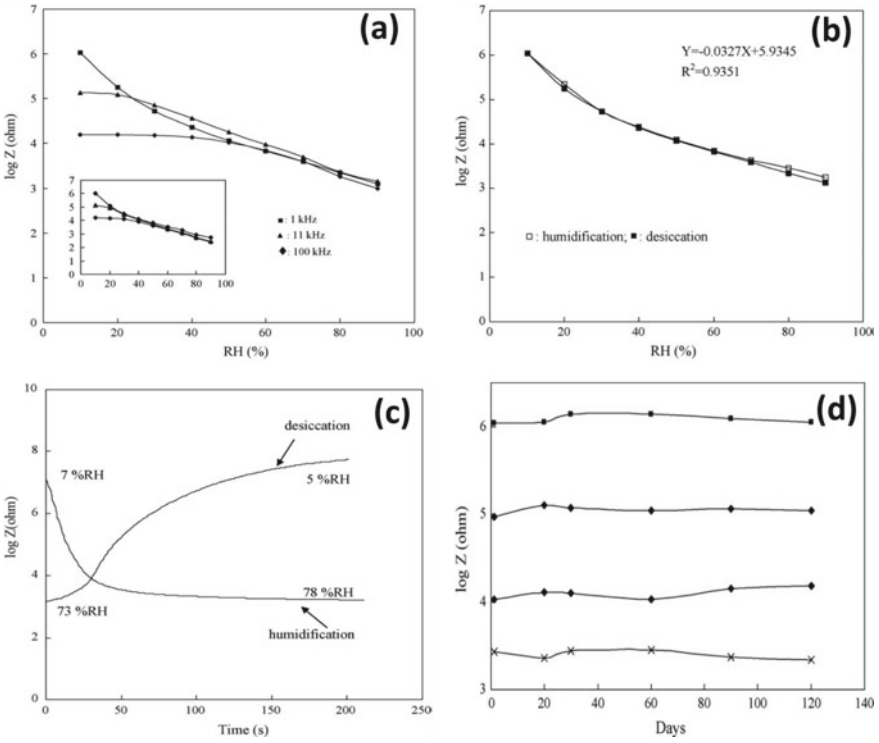


Fig. 15.2 **a** Sensor impedance decreases with increasing RH at all frequencies, **b** observed hysteresis for the fabricated sensor (10–90% RH), **c** response and recovery time of the flexible humidity sensor, **d** effect of long-term stability on the fabricated sensor (min 120 days)

PEDOT:PSS, GOPS as a crosslinker, and Triton X-100 (TX-100). The fabricated temperature sensor fabrication flow is shown in Fig. 15.3c. Initially, Ag electrode with thickness of 100 nm and subsequently with the help of the AgNP ink (50 m PEN film), IDEs patterns are achieved. In this process PEDOT:PSS is used as the sensing layer for the thermal and the dispenser is used to print the CYTOP encapsulation. A photograph of a printed temperature sensor adhered to skin is shown in Fig. 15.3d.

Sensor Transfers Function

Under various humidity circumstances, the resistance variations of fabricated sensors were evaluated in order to analyze the temperature-dependent study. The resistance of the sensor is studied and characteristics remained essentially unchanged when the humidity varied from 30% RH to 80% RH, as illustrated in Fig. 15.4a. The time vs resistance changes have been examined for various humidity and temperature conditions as shown in Fig. 15.4b. The small change in the sensor resistance w.r.t humidity

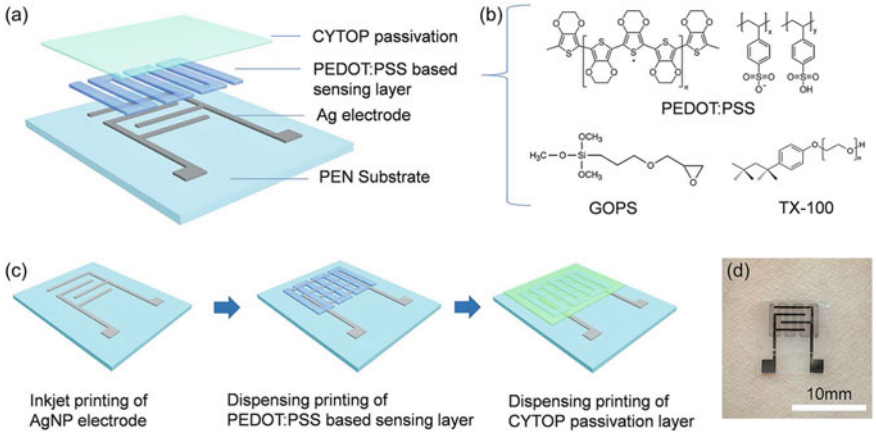


Fig. 15.3 Fabrication flow for the proposed flexible temperature sensors

has proved the potential of the fabricated sensor. Excellent stability was attained in addition to humidity stability. The repeatability was studied for temperatures ranging from 30 to 45 °C as shown in Fig. 15.4c. The little variations indicated that the performance of sensor is consistent in various temperature cycle testing. The temperature-resistance curves of 10 devices were used to characterize the repeatability of the printed sensors. The good repeatability of our printed sensors was demonstrated with small variation in TCR as shown in Fig. 15.4d. The printed electrode's flexibility and the sensing film's PEDOT:PSS foundation also gave the sensor excellent mechanical stability. Figure 15.4e displayed the printed sensor's resistance variation at different bending radii (from 10 to 2 mm). At the aforementioned bending radii, there was just a 1% increase in resistance over flat circumstances. Furthermore, experiments including repeated cyclic bending under 10,000 cycles at a 5 mm bending radius were also carried out. Gradually increasing the bending cycle had no discernible impact on resistance or sensor performance Fig. 15.4f (Wang 2020).

Flexible Leaf Wetness Sensors

Sensor Fabrication and Experimental Set-Up

The fabrication method of micro leaf wetness sensor on the PET flexible substrate is depicted in Fig. 15.5a (Patle et al. 2021b). Initially, IDE patterns were created in the Corel Draw. Next, a flexible polyimide substrate was purchased, as depicted in Fig. 15.5a, and then a layer of aluminum (Al) was put where the electron beam evaporation method is used, as depicted in Fig. 15.5b. Additionally, patterning of the interdigitated electrodes (IDEs) are achieved with the help of laser ablation as

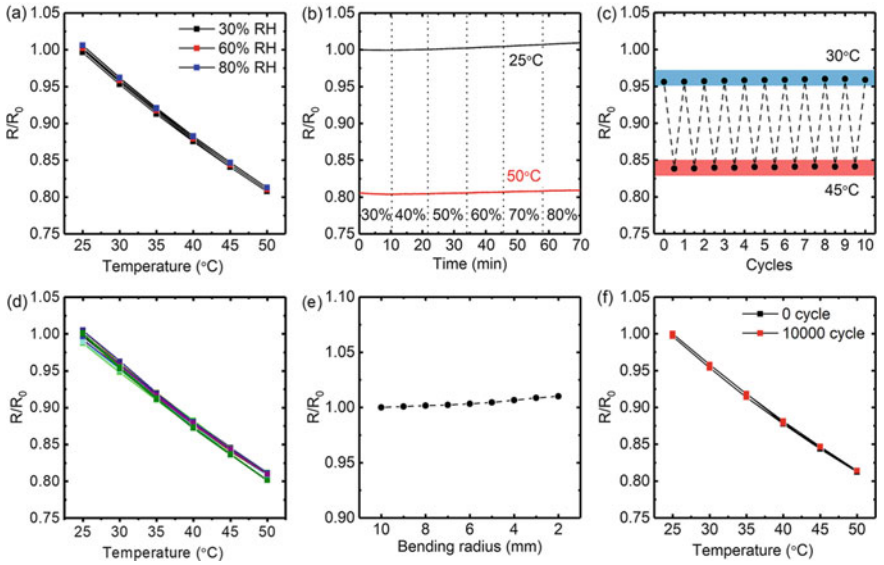


Fig. 15.4 **a** Ambient humidity rising from 30% RH to 80% RH, the temperature-resistance parameters remained essentially same, **b** stability of a relative humidity sensors, **c** temperature repeatability test results for the range of 30 to 45 °C, **d** Excellent mechanical stability of sensors, **e** change in sensor resistance at various bending radii, which is between 2 and 10 mm, **f** effect on resistance or sensor with repetitive cyclic bending under 10,000 cycles at a 5 mm bending radius

indicated in Fig. 15.5c. The constructed IDEs are displayed on the flexible substrates in Fig. 15.5d. The structures were achieved using a 1.06 m Yb fiber laser, and the IDEs were designed to fit on the tulsi leaves. The constructed sensor is mounted using silver epoxy on the PCB to create electrical contact, as illustrated in Fig. 15.5d (Patle et al. 2021b).

The constructed sensor is also coupled to LCR meter, which measures how the capacitance of the sensor changes in response to exposure to moisture as shown in Fig. 15.5c. All of the trials have been completed at a constant room temperature (25 °C) with a relative humidity of about 50% RH. Throughout the tests, the LCR meter's excitation voltage is held constant at 1 Vpp. These leaf wetness sensors can also be used as the rain sensors (Patle et al. 2021b).

Sensor Transfers Function

The sensor is examined in a lab setting to investigate the fabrication's flexible LWS properties, such as response time, sensitivity, accuracy, and hysteresis. For this purpose, the sensing area on the IDEs is separated into various areas to obtain

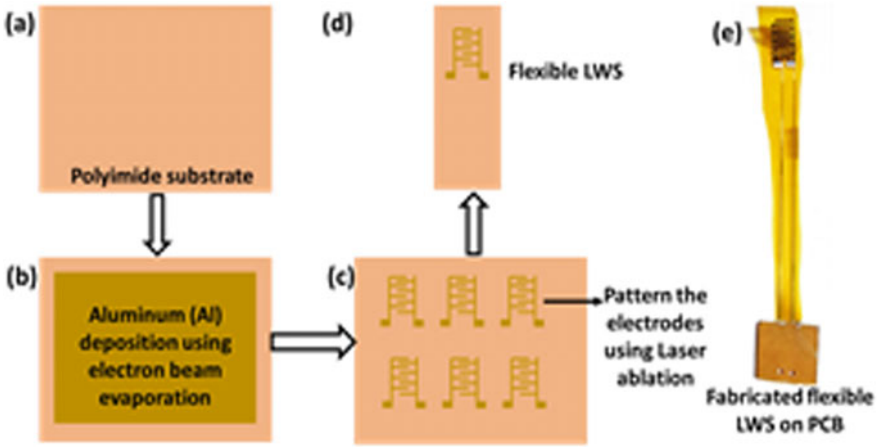


Fig. 15.5 Fabrication of the Al IDEs of the PET substrate

the benchmarking and constructed LWS transfer characteristics. For the measurements, spraying of the water molecule on the constructed LWS is done as depicted in Fig. 15.6. An LCR meter is connected to the sensor and first the sensor's properties w.r.t to frequency response is investigated for various sensing areas as shown in Fig. 15.6. Using an LCR meter, constructed LWS capacitance is measured, as seen in Fig. 15.6a. It is clear from Fig. 15.6a beyond 10 kHz, LWS capacitance remains unaltered. The change in the LWS capacitance with respect to the water molecules covering various sensing areas is depicted in Fig. 15.6b, where it is observed that when number of water molecules covering various sensing areas grows then sensor capacitance grows monotonically. Further, sensor response has been studied ($\Delta C/C$) as represented in Fig. 15.6c. In addition, the hysteresis is studied for the manufactured LWS as shown in Fig. 15.6d. In order to accomplish this, spraying of water is done on sensor, and the LWS capacitance is examined by varying the amount of the detecting area that the water covers as depicted in Fig. 15.4d. The sensor exhibits approximately 5% hysteresis, as seen in Fig. 15.6d (Patle et al. 2021b).

Figure 15.6e represents the response and recovery time, which is studied at 1 kHz. When water was sprayed evenly across the whole surface area of the constructed LWS, the sensor capacitance was then measured and displayed in Fig. 15.6e. According to Fig. 15.6e, the sensor's response time is around 5 s. The impact of temperature on the flexible LWS has also been studied, which is another important issue in the operating exposure of the LWS. In order to achieve this, the sensor is placed in a hot air oven and the temperature is adjusted between 20 and 65 °C, taking into account the in situ agricultural conditions in India. When the temperature changes from 20 to 65 °C, it is clear from Fig. 15.6f that the flexible LWS's sensor capacitance varies by just 6% (Patle et al. 2021b).

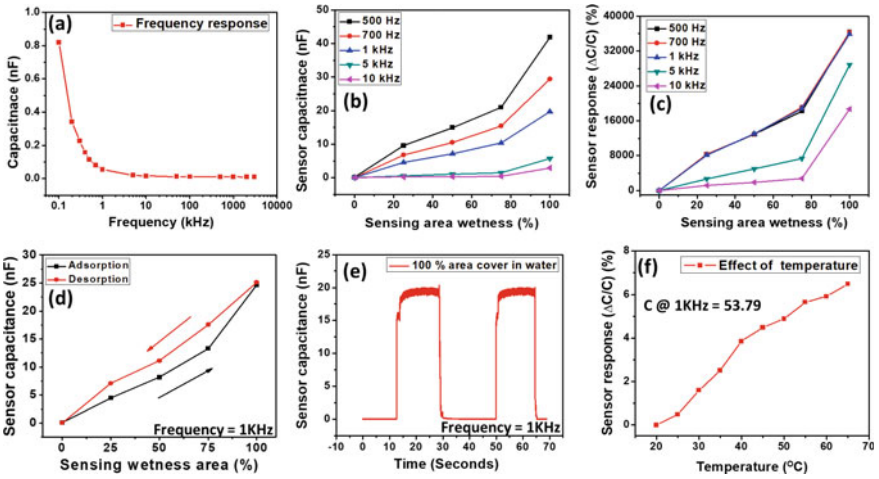


Fig. 15.6 **a** Variation of the sensor with increase in the frequency for dry situation, **b** variation in the sensor capacitance versus increase in the sensing area exposed to wetness, **c** observed response versus increase in the sensing area exposed to wetness, **d** hysteresis observed for the fabricated LWS, **e** time-dependent study of the sensor, **f** temperature study on the developed LWS sensor

Conclusion

Ambient temperature, ambient humidity, leaf wetness, and rain sensors are the primary parameters, which has to be considered for the early plant disease predictions. Considering the affordability for the poor farmers, there is need to develop low-cost sensors and achieve the higher stability and accuracy of detection. This paper provides a review on the fabrication of the sensors to measure ambient temperature, ambient humidity, leaf wetness, and rain, which are widely used in integrated plant disease management which are the basic important parameters for agricultural applications. Additionally, flexible sensors have a potential over the conventional PCB sensors mainly due to their lightweight which further provide the excellent contact between the sensor and the leaf canopy that ensures the more accurate measurements.

References

Chouhan SS, Kaul A, Singh UP, Jain S (2018) Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: an automatic approach towards plant pathology. IEEE Access 6:8852–8863

Fang Y, Ramasamy RP (2015) Current and prospective methods for plant disease detection. Biosensors 4:537–561. www.mdpi.com/journal/biosensors/

Golhani K, Balasundram SK, Vadamalai G, Pradhan B (2018) A review of neural networks in plant disease detection using hyperspectral data. INFORM

Iqbal Z et al (2018) An automated detection and classification of citrus plant diseases using image processing techniques: a review. Comput Electron Agric 153:12–32

Jain A, Sarsaiya S, Wu Q, Lu Y, Shi J (2019) A review of plant leaf fungal diseases and its environment speciation. Bioengineered 409–424

Kumar M, Kumar A, Palaparthy VS (2021) Soil sensors based prediction system for plant diseases using exploratory data analysis and machine learning. IEEE Sens J 21(16):17455–17468

Patle KS, Saini R, Kumar A, Palaparthy VS (2021a) Field evaluation of smart sensor system for plant disease prediction using LSTM network. IEEE Sens J 22(4):3715–3725

Patle KS, Saini R, Kumar A, Surya SG, Palaparthy VS, Salama KN (2021b) IoT enabled, leaf wetness sensor on the flexible substrates for in-situ plant disease management. IEEE Sens J 21(17):19481–19491

Patle KS, Dehingia B, Kalita H, Palaparthy VS (2022) Highly sensitive graphene oxide leaf wetness sensor for disease supervision on medicinal plants. Elsevier Comput Electron Agric 200:107225

Su P-G, Wang C-S (2007) Novel flexible resistive-type humidity sensor. Sens Actuators, B Chem 123(2):1071–1076

United Nations, Department of Economic and Social Affairs, Population Division (2019) World population prospects 2019: highlights (ST/ESA/SER.A/423)

Wang YF, et al. (2020) Fully printed PEDOT: PSS-based temperature sensor with high humidity stability for wireless healthcare monitoring. Sci Rep 10(1):2467

Chapter 16
GaitTracker: A Digital Platform for Measuring, Detecting and Analyzing Gait Changes

Aniruddh Jayant Muley, Kalyan Sasidhar, and Ronak Dhokai

Abstract Gait is an essential bio-marker for long-term health. Traditionally, Gait analysis depends on vision-based or expensive pressure mats where patients are instructed to walk or perform standard limb movements. Early monitoring and detection of gait changes could prevent severe conditions. However, such tests happen late in the onset of the problem. Despite a plethora of wearable devices, such as fitness bands and health trackers, no single device monitors gait and provides an early diagnosis. This work presents a proof-of-concept of our in-house developed wearable inertial measurement unit (IMU) for extracting gait patterns. In addition, the results presented in this work detect changes in gait patterns. The device was tested with ten volunteers (six males and four females, 25 +/− 1.8 years, 163 +/− 8.8 cm) who provided data for both normal and abnormal walking resulting in around 700 gait samples. The results show that sudden changes in gait can be detected with an affordable and portable wearable device.

Keywords Wearable device · Accelerometer · Gait analysis

Introduction

Gait refers to a walking pattern and is unique to every individual (Gafurov et al. 2007). A typical gait pattern consists of multiple events occurring with alternate movement of the limbs. Gait of a person can change due to various reasons including body weight, injuries, age-related cognitive impairments and neurological dysfunctions (Cimolinn et al. 2011). Moreover, gait could change suddenly and if left unnoticed/ignored for a prolonged time, can lead to abnormal gait such as antalgic gait (limping

A. J. Muley · K. Sasidhar (✉) · R. Dhokai
Dhirubhai Ambani Institute of Information and Communication Technology Gandhinagar, Gujarat 382007, India
e-mail: kalyan_sasidhar@daiict.ac.in

Y. Agrawal et al. (eds.), *Interconnect Technologies for Integrated Circuits and Flexible Electronics*, Springer Tracts in Electrical and Electronics Engineering,
https://doi.org/10.1007/978-981-99-4476-7_16

on one leg leading to a bend on one side of the body).[1] Various attempts have been made to detect this form of abnormal gait (Kozlow et al. 2018). However, as revealed by orthopedics, in most cases, individuals realize their abnormalities very late in their lives due to the lack of early assessment or diagnostic facilities and hence resort to kneecaps, medications, and exercises. Still, the gait patterns remain unchanged and abnormal gaits are becoming one of the contributors to knee problems (Gait and Knee Problems 2023) and lower back pains (da Fonseca et al. 2009). The sudden gait changes, if unnoticed, could persist for a longer time, resulting in abnormal gait.

With the evolution of wearable sensing systems, portable fitness bands and smartwatches have already proliferated in the consumer market. However, none of these gadgets is capable of monitoring one's walking and detecting changes or abnormalities in gait patterns. In other words, these devices do not record or extract gait patterns and detect gait changes. This highlights the necessity to encourage regular gait measurement periodically, similar to measuring blood pressure, sugar levels and even oxygen saturation levels at regular intervals. Various survey articles have been published discussing the use of wearable sensing systems for gait analysis (Yang and Hsu 2010; Bhosale et al. 2015). The gap in the literature includes a dedicated device for long-term gait monitoring. To accomplish continuous monitoring and detection, an individual's gait must be stored in a database for which a digital platform is required.

To analyze the importance of the problem, this work conducted a survey across a pool of 30 doctors and physiotherapists. The responses were the following:

- 80% believe that preventive health care is the need of the hour.
- 83.34% believe that a portable device measuring gait parameters with analysis can be useful.
- 86.67% denied when asked if there is any similar device in the market.
- 90% would test and recommend such a device to their patients, if given to them.

Backed by survey responses, we designed and developed a wearable device that can detect gait changes in an individual and store the data in a database. In a natural world setting, analyzing gait abnormalities is limited to a handful of hospitals. The patients have to visit the hospitals for a facility like a Motion Capture system (Estévez-García et al. 2015). This is often an expensive affair and beyond the reach of a majority of the population. With digital health on the rise, there is a pressing need to create a digital repository of individual health data. In a country like India, having a vast population base, there is a need to have arrangements for a huge healthcare database. This is where our affordable device can contribute. Our proposed device can be used to create a gait database that can provide a historical perspective for various stakeholders, including orthopedics, and physiotherapists, who can analyze the data through our algorithms and provide early diagnoses to patients about gait changes.

[1] Auerbach N, Tadi P. Antalgic Gait in Adults. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 January. Available from: https://www.ncbi.nlm.nih.gov/books/NBK559243.

These changes could be subtle, but, if they persist for a long time, could lead to potentially serious medical conditions.

The following are our contributions:

- The wearable device is in itself a novel attempt to provide personalized healthcare diagnostics.
- For the first time in India, the abnormal gait detection system paired with a smartphone application is implemented. Till now, no such work has been observed.
- The differences in gait parameters between normal and abnormal walking within a user are identified.
- We show how changes in walking are identified, detected and presented to the concerned personnel.

Related Work

Analysis of gait is commonly performed by measuring speed, length of rhythm and pitch while moving. The most common approaches are using camera-based systems to extract the gait parameters during locomotion. However, this approach requires a complex set up which is bulky, and incurs infrastructure cost. Such facilities in India are available only in multi-specialty hospitals which are beyond the reach of the general population.

Gait analysis using external sensors has been conducted by placing an accelerometer sensor on the hip, ankles and foot. A wearable body-worn sensing system for analyzing gait and classifying between normal and abnormal has been proposed by Sant'Anna et al. (2012). The gait of subjects with lower back pains is measured and it explores the correlation between improper gait and back pain in Chan et al. (2013). In Cheng et al. (2013), researchers developed GaitTrack, a smartphone app that computes the walking or gait speed of patients and provides a correlation between gait speed and patients with Chronic Obstructive Pulmonary Disease (COPD). Activity recognition, i.e., recognizing walking, running, taking the stairs or sitting still, is an application which has used inertial sensors such as the accelerometer to a great extent. Various research works have proposed unique algorithms for activity recognition (Kwapisz et al. 2011; Sun et al. 2010; Yang 2009; Zhong et al. 2010).

A personalized wearable system that enables the movement of persons with Parkinson's disease is developed by Mazilu et al. (2014). The system detected freezing episodes or abrupt stops in walking using ankle-mounted motion sensors and responded by playing a rhythmic auditory sound that adapted to the patient's regular gait speed, alerting the person to move, and used a smartphone as a Processing Unit. Godfrey et al. validated a low-cost body sensor unit to quantify gait characteristics in a large group of young and older adults (Godfrey 2015). Authors in (Pepa et al. 2015) extract step length using the data collected from a smartphone accelerometer to assess gait of an individual. Perez performed a gait analysis to assess the balance and risk of patients falling with walking difficulties (Perez and Labrador 2016). Miniature-sized

accelerometers and gyroscope sensors (Roy et al. 2016) embedded in smartphones have tremendous application usage.

Works specific to gait analysis through smartphones have been attempted in Chan et al. (2013), Zhong et al. (2010), Pepa et al. (2015), Nishiguchi et al. (2012), Sasidhar and Satyajeet (2017).

We conclude that most of the existing work consists of

- Validating the use of accelerometers in sensing gait parameters.
- Performing kinetic analysis using wearable sensors and understanding gait.
- Identifying the correlation between gait patterns and walking speeds with symptoms of various diseases that have been shown to affect gait.
- Performing classification of different sets of gait samples into multiple classes.

In contrast to the existing work, we propose a system that aims to do the following: a) monitor gait, b) store periodic gait data for identifying and detecting gait changes, c) calculate the extent of the change in gait and d) provide diagnostic information to stakeholders such as physiotherapists and orthopedics.

Methodology

Data Collection

Ten participants, all graduate students residing on the campus of our institute, volunteered to wear the device and provide data. The demographics of the subjects is presented in Table 16.1.

The instruction is given to each user to walk a 10 m path. Users walked with their body posture and pace similar for each walk. However, to induce sudden changes in their walk pattern, a video of antalgic gait has been shown to the users. Under the

Table 16.1 Participating users in experiment

User	Gender	Height (cm)	Age
User 1	Male	171.2	29
User 2	Male	173	24
User 3	Female	154	25
User 4	Male	160	25
User 5	Female	150	24
User 6	Female	155.5	24
User 7	Female	160	26
User 8	Male	170	25
User 9	Male	165	25
User 10	Male	176	23

supervision of the physiotherapist, each user mimicked the action to bring a change in the walking pattern. We obtained a total of 70 gait data sets (50 normal walking and 20 abnormal walking data sets) per user, resulting in a database of 700 gait data sets across ten users.

Measurement Platform

The system was designed and developed in-house using an ARM Cortex M3 core ultra-low power micro-controller with Bluetooth version 2.0, an MPU 6050 3-axis accelerometer and a 3.74 V Lithium polymer (LiPo) battery with a voltage regulator for power supply. The accelerometer consists of microstructures whose capacitance changes due to external forces. This capacitance change is converted to voltage values (Roy et al. 2016). The sensor has a pair of three 16-bit ADCs for digitizing the output. The sensor has a minimum full-scale range of ± 2 g and a maximum full-scale range of ± 16 g.

The sensed data is transmitted via UART from the micro-controller to the Bluetooth module HC-05 V2.0. This module is further paired to a smartphone, where all the computation is performed through a mobile application. Figure 16.1 shows the architecture of the system.

The proof of concept of the device and how it is worn is depicted in Fig. 16.2.

In the next section, the explanation of how the sensor data was preprocessed, how the human walking was recorded and how gait features were extracted is provided.

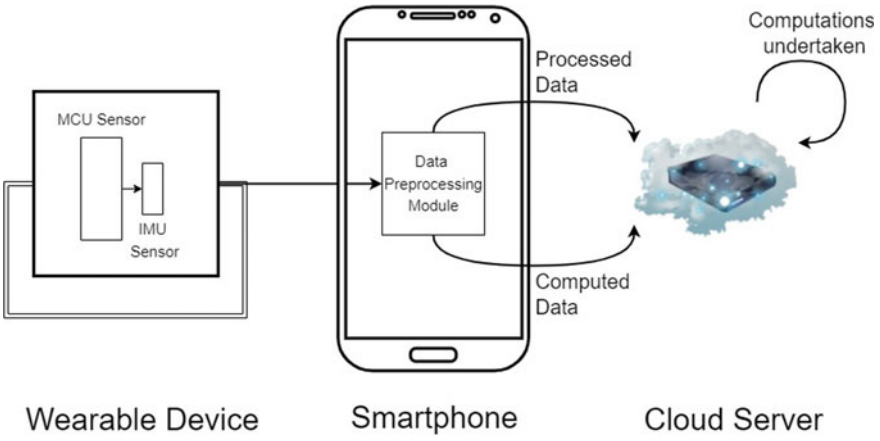


Fig. 16.1 Architecture of flow of the application

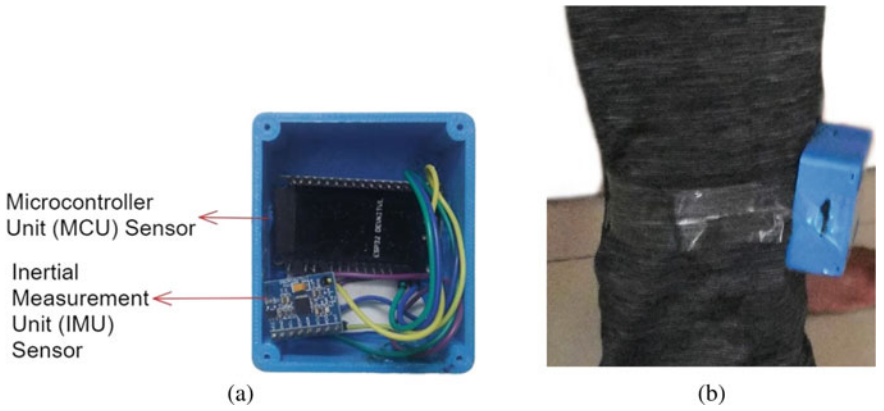


Fig. 16.2 **a** The sensor along with the micro-controller unit is shown on the right and the Bluetooth IC on the left. **b** The device fixed to a band and wrapped around the ankles

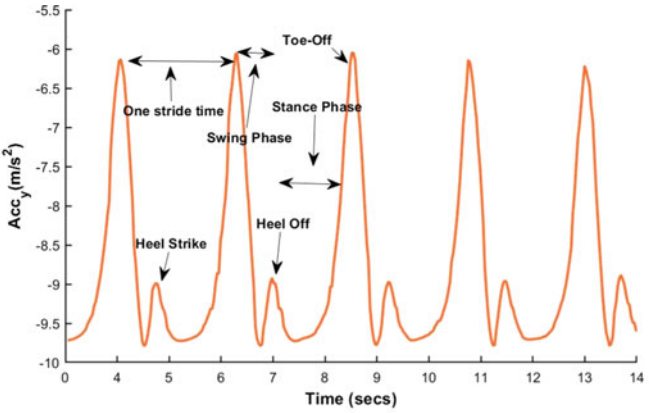
Data Processing

The data from the wearable device was sampled at 50 Hz. The data was sent to a cloud database through a smartphone which was later converted and downloaded in CSV format for analysis in MATLAB (version 2022a). Basic filtering techniques were applied to smoothen the raw data. The data from each user was named *UserID day*1 *normal* and *UserID day*1 *Abnormal* to preserve anonymity. All data was sent to MongoDB, where the data is stored and can be exported in CSV format. Figure 16.3 shows the screenshot of the sample data.

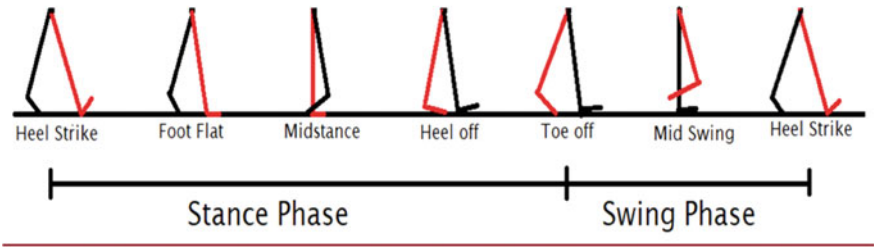
Figure 16.4a illustrates the smoothed accelerometer data captured from one individual who wore the device and walked a distance of 10 m, and the corresponding

acc_x (m/s²)	acc_y (m/s²)	acc_z (m/s²)	Time since start	HH-MI-SS_SSS
-0.1085	9.5662	-0.7045	10	14:56:46:805
0.2361	9.4418	0.042	26	14:56:46:821
0.1667	9.4992	-0.2834	41	14:56:46:836
-0.1516	9.7218	-0.403	57	14:56:46:852
-0.5033	9.4155	-1.1568	74	14:56:46:869
-0.6732	9.6236	-0.8673	90	14:56:46:885
-1.0178	9.65	-0.5036	106	14:56:46:901
-0.9747	9.6524	-0.3887	122	14:56:46:917
-0.8599	9.449	-0.5347	138	14:56:46:933
-0.4555	9.3652	-0.7237	153	14:56:46:948

Fig. 16.3 The accelerometer (*acc x*, *acc y* and *acc z*) data across 3 axes are stored along with the timestamp



(a) Smoothed accelerometer data



(b) Gait phases

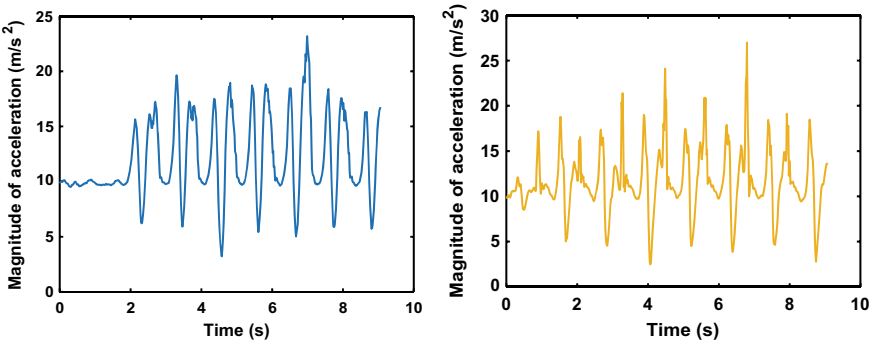
Fig. 16.4 **a** The gait cycle starts at heel strike and ends at heel off. The stance phase is from heel strike to toe-off and the swing phase is from toe-off to heel strike of the other leg. **b** The various gait phases are identified in the smoothed accelerometer data

gait phases are shown in Fig. 16.4b. The device placement was suggested by an orthopedic and sports bio-mechanics expert who tested the device in the lab themselves. The location of the device (Fig. 16.2) assists in capturing the force exerted while walking.

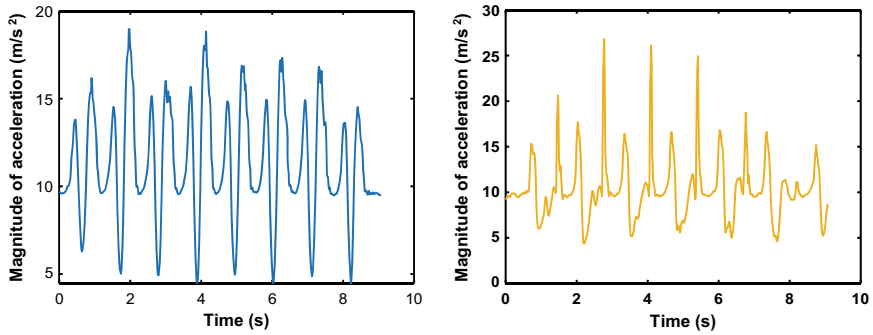
Figure 16.5 compares normal gait and limping gait patterns of a volunteer.

Detection of Gait Changes

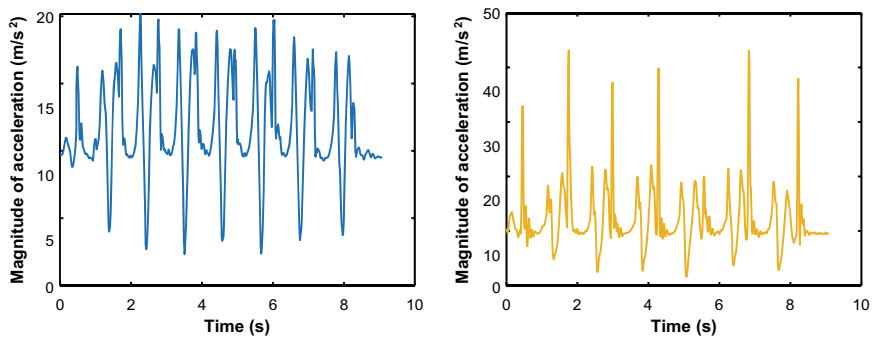
With the proposed device, user's gait can be stored and analyzed for sudden changes so as to diagnose potential chronic ailments early. Let us look at intra-subject differences in stride times between normal and abnormal walking samples. Figure 16.6 provides a statistical comparison of stride duration for both walking patterns of six



(a) User1_day1_normal (b) User1 day1 Abnormal



(c) User2_day1 _normal (d) User2 day1 Abnormal



(e) User3_day1_normal (f) User3_day1_Abnormal

Fig. 16.5 One cycle of gait starts at heel strike and ends at heel off. Stride Length is the distance between successive parts of heel contact of the same foot, and stride time is the corresponding time taken for one stride length. To calculate the stride length, two consecutive heel strikes are considered as the start and end of a stride because when the heel hits the ground, a transient intense force is exerted, resulting in a peak acceleration

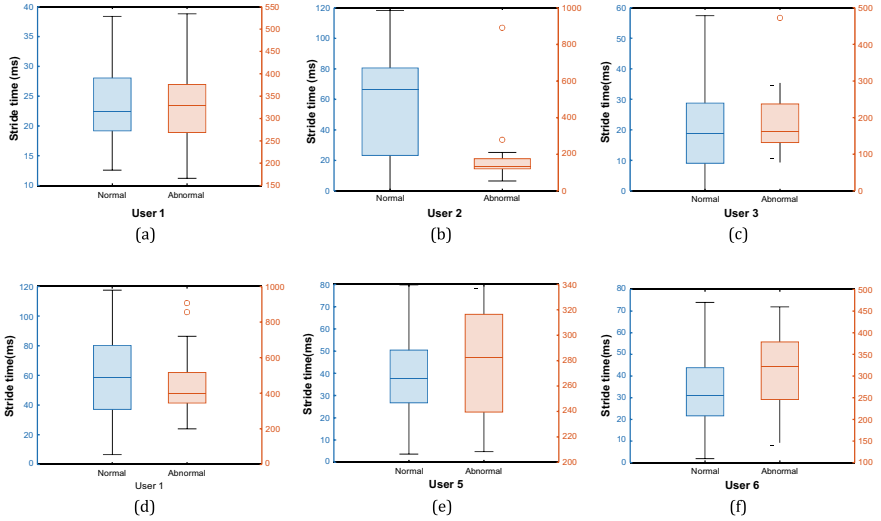


Fig. 16.6 Stride time variations of volunteers: For each box plot, the median and inter-quartile values are in different ranges, confirming that there were significant individual variations in walking between normal and limping

volunteers. This difference clearly provides a quantitative comparison and indication of one's gait changes.

The novelty of this work is to facilitate a personalized classification or identification of abnormality in gait. In other words, classifying a person's gait as normal or abnormal from a pool of gait samples belonging to different persons does not help in analyzing the degree of change in gait. Instead, performing an intra-volunteer gait pattern comparison will help to observe how one's gait patterns change over a period of time.

For instance, the heel strike and heel off features of one user are identified, the stride length is extracted and the stride time is computed. Further, the average change in stride times over each day is computed and changes are observed as shown in Fig. 16.7.

A smartphone app has been designed as explained in section "Measurement Platform". The interface of the app is shown in Fig. 16.8. The interface is self-explanatory. The stakeholder (orthopedic, physiotherapist) has given a list of his patients post log into the app. Once a particular patient is selected, his/her historical gait parameters are available. Further options include visualizing the sensor data along with the gait parameters extracted by the algorithm at the server.

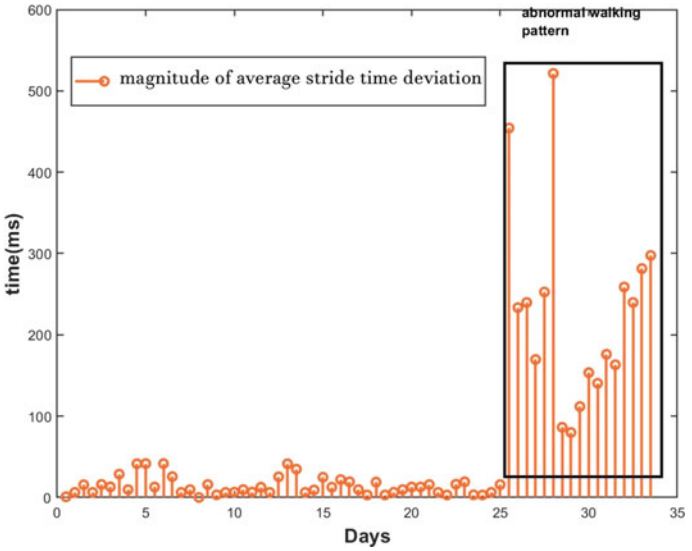
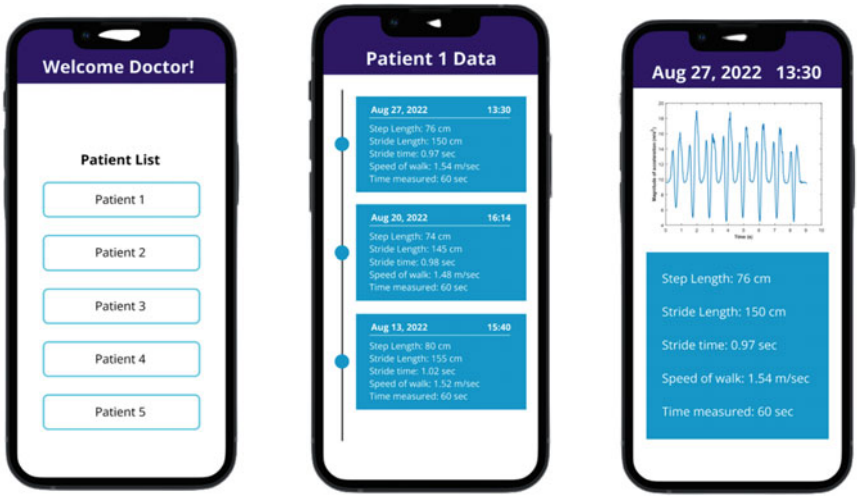


Fig. 16.7 As we can see, most of the deviations are negligible but due to limping the stride times increased. This can be observed between days 25–33. Such sudden but slow changes in stride time could indicate difficulty in walking which if noticed earlier could assist in an early diagnosis by the medical personnel



(a) List of patients whose data is available

(b) History of gait data of the patient

(c) Measured Gait data of a patient

Fig. 16.8 Flow of the screen interface of GaitTracker, the mobile application

Conclusions

This work presented a device that can provide personalized gait monitoring and detect sudden changes. The novelty of the work lies in packaging together a solution that is inexpensive, portable and easily usable. The device can be used at a clinic for regular checkups, where the digital records of each patient's gait patterns are stored for diagnosis. The work is currently looking at predicting when and how much one's gait could change if the current trend of abnormality is observed. With healthcare prediction being a positive trend, this device can play a vital role in the pre-diagnosis and diagnosis of gait irregularities. In future, this work can be extended to the device being worn on both the legs and the gait patterns being measured simultaneously so as to cover a broader scope of gait patterns.

Acknowledgements We thank Abhi Ratnam, Arvind Jayanthi and Prakhar Agarwal for working on the development of the in-house wearable device used for the experiment. We also express our gratitude to Dr. Vivek Trivedi, the physio-therapist under whose supervision the experiment was carried out and the observations were studied. This work is funded by DA-IICT SEED fund under the project titled "WalkSense: Design and Development of a wearable device for abnormal gait detection".

References

Bhosale T, Kudale H, Kumthekar V, Garude S, Dhumal P (2015) Gait analysis using wearable sensors. In: 2015 international conference on energy systems and applications, Oct 2015

Chan H, Zheng H, Wang H, Sterritt R, Newell D (2013) Smart mobile phone based gait assessment of patients with low back pain. In: Ninth international conference on natural computation (ICNC), July 2013

Cheng Q, Juen J, Li Y, Prieto-Centurion V, Krishnan JA, Schatz BR (2013) Gaittrack: health monitoring of body motion from spatio-temporal parameters of simple smart phones. In: Proceedings of the international conference on bioinformatics, computational biology and biomedical informatics

Cimolin V et al (2011) Effects of obesity and chronic low back pain on gait. J Neuro Eng Rehabil

da Fonseca et al (2009) Laboratory gait analysis in patients with low back pain before and after a Pilates intervention. J Sport Rehabil

Estévez-García R, Martín-Gutiérrez J, Mendoza SM, Marante JR, Chinea-Martín P, Soto-Martín O, Lodeiro-Santiago M (2015) Open data motion capture: Mocapull database. Procedia Comput Sci 75:316–326

Gafurov D, Snekkenes E, Bours P (2007) Gait authentication and identification using wearable accelerometer sensor. In: 2007 IEEE workshop on automatic identification advanced technologies, June 2007

Gait and knee problems (2023). http://www.sciencealert.com/gaits-reveal-knee-problems

Godfrey A (2015) Instrumenting gait with an accelerometer: a system and algorithm examination. Med Eng Phys 37(4):400–407

Kozlow P, Abid N, Yanushkevich S (2018) Gait type analysis using dynamic Bayesian networks. Sensors (Basel) 18(10)

Kwapisz JR, Weiss GM, Moore SA (2011) Activity recognition using cell phone accelerometers. ACM SIGKDD Explor Newsl 12(2)

Mazilu S et al (2014) Gaitassist: a wearable assistant for gait training and rehabilitation in Parkinson's disease. In: IEEE international conference on pervasive computing and communication workshops (PERCOM WORKSHOPS), March 2014

Nishiguchi S et al (2012) Reliability and validity of gait analysis by android-based smartphone. Telemed J e-health: Off J Am Telemed Assoc 18

Pepa L, Verdini F, Spalazzi L (2015) Experimental evaluation of a smartphone based step length estimation. In: 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Aug 2015

Perez AA, Labrador MA (2016) A smartphone-based system for clinical gait assessment. In: IEEE international conference on smart computing (SMARTCOMP)

Roy S, Mandal S, Hanumaiah N (2016) Mems accelerometer: from engineering to medicine. IEEE Potentials 35(2):30–33

Sasidhar K, Satyajeet S (2017) iknow how you walk—a smart- phone based personalized gait diagnosing system. In: 2017 international conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 2063–2068

Sant'Anna A, Wickstro¨ m N, Eklund H, Tranberg R (2012) A wearable gait analysis system using inertial sensors part II: Evaluation in a clinical setting. In: BIOSIGNALS

Sun L, Zhang D, Li B, Guo B, Li S (2010) Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In: Proceedings of the 7th international conference on ubiquitous intelligence and computing

Yang J (2009) Toward physical activity diary: Motion recognition using simple acceleration features with mobile phones. In: ACM proceedings of the 1st international workshop on interactive multimedia for consumer electronics

Yang C-C, Hsu Y-L (2010) A review of accelerometry-based wear- able motion detectors for physical activity monitoring. Sensors (Basel, Switzerland)

Zhong S, Wang L, Bernardos AM, Song M (2010) An accurate and adaptive pedometer integrated in mobile health application. In: IET international conference on wireless sensor network (iET-WSN), Nov 2010