



Chapter 11

Rail Roughness Profile Identification from Vibration Data via Mixing of Reduced-Order Train Models and Bayesian Filtering

Charikleia D. Stoura, Konstantinos E. Tatsis, and Eleni N. Chatzi

Abstract The increasing demand for mobility worldwide has led to an ever-increasing expansion of railway networks. Such a rapid expansion poses a challenge to guaranteeing quality, reliability, efficiency, and, most importantly, safety of railway infrastructure. Under this perspective, continuous monitoring emerges as a promising alternative to the traditionally adopted visual inspections, or inspections via portable measuring devices, which aim at collecting geometric data for the diagnosis and prognosis of defects in tracks. Recently, railway operators worldwide have adopted the use of dedicated Track Recording Vehicles equipped with optical and inertial sensors to collect data from tracks and assess their condition. Such an approach revolutionized rail condition assessment, introducing a mobile data acquisition platform for track inspection. On the other hand, the deployment of such specialized vehicles requires disruption of regular rail service, which hinders their frequent operation and thus the continuous collection of rail data. This work aims to tackle this limitation by examining an onboard monitoring (OBM) method that hinges on collecting vibration data from in-service trains. The proposed methodology relies on the collection of acceleration data from axle boxes of trains running at normal speeds. Its novelty lies in the usage of realistic train models and the consideration of the dynamic interaction between rails and trains, which is usually simplistically ignored. The adopted train models are reduced so as to decrease the required computational effort. The identification task is based on sequential Bayesian inference for joint input and state estimation, thereby also accounting for uncertainties related to the train model. Estimating the input leads to the identification of the pertinent rail roughness profile, which can subsequently provide information on the existence of isolated defects, for example, welded joints and squats, along the track system. This study is limited to reliable prediction of the dynamics of the train–track system in the vertical direction, but proposes methods and tools of general value.

Keywords System identification · Onboard monitoring · Railway infrastructure · Bayesian inference · Hybrid modeling

11.1 Introduction

Efficient and effective monitoring of railway infrastructures is decisive for the regular maintenance of railways, which, accordingly, guarantees the quality and safety of rail transport. Traditional approaches based on visual inspections and measurements on-site, although reliable, can no longer accommodate the increased need for monitoring along large portions of railways. Therefore, roving implementations have been widely explored during the last few years [1].

Such a roving method constitutes the use of Track Recording Vehicles. Those are vehicles equipped with optical (e.g., laser scanners, high-speed cameras) and inertial sensors (e.g., gyroscopes, inclinometers) that collect geometric data from the rails [2]. These geometric data can then deliver insight into specific irregularities of the rails. The usage of such vehicles can provide very accurate information regarding the characteristics and location of isolated defects on tracks. However, these are limited to operating while normal rail operation is suspended, thus cannot provide continuous information regarding the condition of the track [3].

A viable alternative comprises the usage of in-service trains equipped with simple monitoring systems, for example, accelerometers. Accelerometers can be mounted on different train vehicle components, such as axle boxes, bogies, and car bodies, allowing for a continuous supply of vibration data. On the other hand, railway operators increasingly improve the monitoring systems that are mounted on in-service trains, aiming primarily at vehicle condition monitoring, which can

C. D. Stoura (✉) · K. E. Tatsis · E. N. Chatzi

Institute of Structural Engineering, Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Zürich, Switzerland
e-mail: charikleia.stoura@ibk.baug.ethz.ch; tatsis@ibk.baug.ethz.ch; chatzi@ibk.baug.ethz.ch

though further facilitate the monitoring of the sustaining rail infrastructure (tracks, rail bridges). For instance, the ICN train of Swiss Federal Railways (SBB) is currently equipped with accelerometers in various locations (axle boxes, bogies, car body), while further including two pairs of tensiometric wheels, able to measure contact forces [4]. This offers an abundance of vibration-based monitoring data that need to be appropriately processed.

A first approach to handling such data pertains to the usage of signal decomposition techniques, including wavelet transform [5], or mixed filtering approaches, including Kalman, band-pass, and compensation filters [6]. These approaches rely on solely treating the obtained data without considering the dynamics of the underlying physical problem, that is, the dynamic interaction between trains and tracks. A recent study by Dertimanis et al. [3] incorporated the dynamic train–track interaction into the identification of defects on rails via acceleration data collected by the axle box of a simple train model. This study revealed promising results toward the identification of rail roughness profiles from in-service trains.

In this work, we propose an indirect approach for the identification of rail roughness profiles from data collected by traversing trains. The proposed scheme considers the physics behind the dynamic train–track interaction phenomenon [7] and couples the substructure-based dynamics with Bayesian inference methods [8] to perform identification of rail roughness [9]. The enabling tool is a realistic, three-dimensional (3D) train model running on tracks with rail roughness defined according to the German spectra for high-speed trains [10], respecting a realistic scenario setting.

11.2 Train–Track Interaction Model

The adopted 3D train vehicle is modeled with six rigid bodies corresponding to the wheelsets, bogies, and car body (Fig. 11.1). All bodies comprise five degrees of freedom (DOFs) each; two translations and three rotations. The longitudinal DOF is omitted assuming constant running speed. The equation of motion (EOM) of the train running on the track system is

$$\mathbf{m}^v \ddot{\mathbf{u}}^v(t) + \mathbf{c}^v \dot{\mathbf{u}}^v(t) + \mathbf{k}^v \mathbf{u}^v(t) = \mathbf{W}^v \mathbf{p}(t) \quad (11.1)$$

where $\mathbf{u}^v(t)$ corresponds to the response vector of the train and $\dot{\square}$ indicates differentiation with respect to time t . The mass matrix of the train is denoted by \mathbf{m}^v , while \mathbf{c}^v and \mathbf{k}^v represent the damping and stiffness matrices, respectively, that correspond to the train's suspension system. The right-hand side is factorized with respect to $\mathbf{p}(t)$, the contact force vector between the train and the underlying track system, and \mathbf{W}^v , which is the contact direction matrix connecting the DOFs of the train to the contact force elements. The contact force $\mathbf{p}(t)$ follows a Hertzian contact model and is expressed as

$$\mathbf{p}(t) = k^H \mathbf{r}_c(x) \quad (11.2)$$

where k^H is the contact stiffness between the train wheels and the rail, and $\mathbf{r}_c(x)$ is the vector containing the roughness of the rails. The EOM of the train can be written in state-space form by initially transforming Eq. (11.1) to the state equation as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_c \mathbf{p}(t) \quad (11.3)$$

where $\mathbf{x}(t)$ is the state vector:

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{u}^v(t) \\ \dot{\mathbf{u}}^v(t) \end{bmatrix} \quad (11.4)$$

and $\mathbf{p}(t)$ is the input vector. Lastly, \mathbf{A}_c and \mathbf{B}_c are, respectively, the system and input matrices of the system, defined as

$$\mathbf{A}_c = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -(\mathbf{m}^v)^{-1} \mathbf{k}^v & -(\mathbf{m}^v)^{-1} \mathbf{c}^v \end{bmatrix}, \quad \mathbf{B}_c = \begin{bmatrix} \mathbf{0} \\ -(\mathbf{m}^v)^{-1} \mathbf{W}^v \end{bmatrix} \quad (11.5)$$

Assuming that acceleration measurement data are available, the output vector can be written as

$$\mathbf{y}(t) = \mathbf{C}_c \mathbf{x}(t) + \mathbf{D}_c \mathbf{p}(t) \quad (11.6)$$

where the output matrix \mathbf{C}_c and the feedforward matrix \mathbf{D}_c are, respectively:

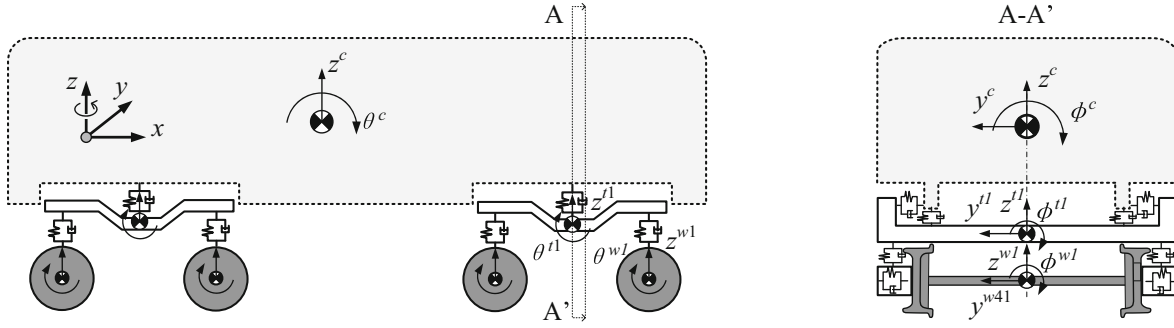


Fig. 11.1 A three-dimensional train vehicle model

$$\mathbf{C}_c = [-\mathbf{W}^a(\mathbf{m}^v)^{-1}\mathbf{k}^v - \mathbf{W}^a(\mathbf{m}^v)^{-1}\mathbf{c}^v], \quad \mathbf{D}_c = \mathbf{W}^a(\mathbf{m}^v)^{-1}\mathbf{W}^v \quad (11.7)$$

with \mathbf{W}^a being the selection matrix of accelerations connecting the output data with the DOFs of the system.

11.3 Train Model Reduction

Train models can be rather complicated, consisting of a large number of DOFs. The number of DOFs can significantly increase the computational cost of the analysis without adding value to the identification task. To this end, a model reduction can be performed, without affecting the quality of the identified input. This study performs model order reduction based on the modal (eigenvalue) analysis of the train vehicle. Such an analysis is performed by the following factorization of the system matrix:

$$\mathbf{A}_c = \mathbf{V}\mathbf{L}\mathbf{V}^{-1} \quad (11.8)$$

where $\mathbf{L} = \text{diag}(\boldsymbol{\Lambda}, \boldsymbol{\Lambda}^*)$ is a diagonal matrix consisting of the complex conjugate eigenvalues of the system and \mathbf{V} is the eigenvector matrix written as

$$\mathbf{V} = \begin{bmatrix} \boldsymbol{\Psi} & \boldsymbol{\Psi}^* \\ \boldsymbol{\Psi}\boldsymbol{\Lambda} & \boldsymbol{\Psi}^*\boldsymbol{\Lambda}^* \end{bmatrix} \quad (11.9)$$

where $\boldsymbol{\Psi}$ is a matrix containing the modal vectors for the physical space. Note that the modal vectors of $\boldsymbol{\Psi}$ are complex as the system is not proportionally damped. The state-space system assumes the form

$$\dot{\boldsymbol{\zeta}}(t) = \mathbf{A}_m\boldsymbol{\zeta}(t) + \mathbf{B}_m\mathbf{p}(t) \quad (11.10)$$

$$\mathbf{y}(t) = \mathbf{C}_m\boldsymbol{\zeta}(t) + \mathbf{D}_m\mathbf{p}(t) \quad (11.11)$$

where $\boldsymbol{\zeta}(t)$ is the modal state vector. The state and measurement matrices, initially defined by Eqs. (11.5) and (11.7), become

$$\mathbf{A}_m = \mathbf{V}^{-1}\mathbf{A}_c\mathbf{V}, \quad \mathbf{B}_m = \mathbf{V}^{-1}\mathbf{B}_c, \quad \mathbf{C}_m = \mathbf{C}_c\mathbf{V}, \quad \mathbf{D}_m = \mathbf{D}_c \quad (11.12)$$

The entries of the state vector in modal basis $\boldsymbol{\zeta}(t)$ of Eq. (11.10) are decoupled. To reduce the size of the model, only a subset $\tilde{\boldsymbol{\zeta}}(t)$ of $\boldsymbol{\zeta}(t)$ is retained, while the rest of the states are truncated. Specifically, only the first few vertical and rolling, in case of different roughness in the left and right rails, modes of the train are considered as the goal is to identify rail roughness profiles in the vertical direction. Thus, the reduced-order system consists of a reduced number of modes encapsulated in $\tilde{\mathbf{A}}$ and $\tilde{\boldsymbol{\Psi}}$, representing the matrices of the selected eigenvalues and eigenvectors, respectively. Accordingly, the state-space matrices of Eq. (11.12) are transformed into $\tilde{\mathbf{A}}_m$, $\tilde{\mathbf{B}}_m$, and $\tilde{\mathbf{C}}_m$. The feedforward matrix \mathbf{D}_m remains unchanged as it is not associated with the truncated state vector.

To discretize the reduced state-space matrices of the system, a sampling rate $F_s = 1/T_s$ is adopted, thus the discretization time is $t = kT_s$. The stochastic discrete-time state-space representation of Eqs. (11.10) and (11.11) is written as

$$\tilde{\boldsymbol{\zeta}}_{k+1} = \tilde{\mathbf{A}}_d \tilde{\boldsymbol{\zeta}}_k + \tilde{\mathbf{B}}_d \mathbf{p}_k + \mathbf{w}_k \quad (11.13)$$

$$\mathbf{y}_k = \tilde{\mathbf{C}}_d \tilde{\boldsymbol{\zeta}}_k + \mathbf{D}_d \mathbf{p}_k + \mathbf{r}_k \quad (11.14)$$

where \mathbf{w}_k and \mathbf{r}_k denote the discrete-time process and measurement noise terms, with known covariance matrices $\mathbf{Q}^w = \mathbb{E}[\mathbf{w}_k \mathbf{w}_k^T]$ and $\mathbf{Q}^r = \mathbb{E}[\mathbf{r}_k \mathbf{r}_k^T]$, respectively. Note that, in this case, for the discretization of the pertinent matrices $\tilde{\mathbf{A}}_d$, $\tilde{\mathbf{B}}_d$, $\tilde{\mathbf{C}}_d$, and \mathbf{D}_d a first-order hold assumption is made. A first-order hold discretization assumes the input as piece-wise linear, different from the commonly adopted zero-order hold assumption that assumes piece-wise constant input. In other words, the relationship between the continuous and digital signal counterparts is based on a linear interpolation between samples as

$$\mathbf{p}(t) = \mathbf{p}_k + \frac{t - kT_s}{T} (\mathbf{p}_{k+1} - \mathbf{p}_k), \quad kT_s \leq t \leq (k+1)T_s \quad (11.15)$$

This assumption emerges from our effort to accurately reconstruct the input vector, which herein comprises the roughness profile (i.e., a random time-series), as indicated by Eq. (11.2), in a reliable and efficient manner. Eventually, the discretized system matrices are given as

$$\tilde{\mathbf{A}}_d = e^{T_s \tilde{\mathbf{A}}_m}, \quad \tilde{\mathbf{B}}_d = \frac{1}{T_s} \tilde{\mathbf{A}}_m^{-2} (\tilde{\mathbf{A}}_d - \mathbf{I})^2 \tilde{\mathbf{B}}_m, \quad \tilde{\mathbf{C}}_d = \tilde{\mathbf{C}}_m, \quad \tilde{\mathbf{D}}_d = \tilde{\mathbf{D}}_m + \tilde{\mathbf{C}}_m \left[\frac{1}{T_s} \tilde{\mathbf{A}}_m^{-2} (\tilde{\mathbf{A}}_d - \mathbf{I}) - \tilde{\mathbf{A}}_m^{-1} \right] \tilde{\mathbf{B}}_m \quad (11.16)$$

11.4 Rail Roughness Profile Identification

For the identification of the roughness profile, the study follows a dual Kalman filter (DKF) approach for joint input-state estimation, as introduced in the study of Eftekhar Azam et al. [9]. This approach performs a sequential identification that first updates the input and then the state based on the updated input value, with the Kalman filter applied to both stages. To this end, for the temporal evolution of the unknown input a random walk model is adopted as follows:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{v}_k \quad (11.17)$$

where \mathbf{v}_k is the discrete-time input noise term, with covariance matrix $\mathbf{Q}^v = \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^T]$. The DKF is described in detail in the study of Eftekhar Azam et al. [9], which shows that the accuracy of the identification task relies on three parameters: the process noise covariance \mathbf{Q}^w , the input noise covariance \mathbf{Q}^v , and the measurement noise covariance \mathbf{Q}^r . The process noise covariance matrix \mathbf{Q}^w represents the accuracy of the adopted physical model, \mathbf{Q}^v depends on the noise of the input, and \mathbf{Q}^r depends on the accuracy of the measurement instruments used. \mathbf{Q}^w and \mathbf{Q}^r can be specified for a given model and measurement device. On the other hand, \mathbf{Q}^v is typically adopted as the tuning parameter. Tuning is performed via an L-curve, which assumes input noise as a regularization parameter, which has to be tuned for optimizing the performance of the adopted filter [9].

11.5 Numerical Application

To validate the efficacy of the proposed approach to recover rail roughness profiles, this chapter adopts a 3D train model (Fig. 11.1) traversing a track system of known roughness. The adopted model represents a practical train with 35 DOFs in total whose properties are derived from the work of Zeng et al. [11]. The track is modeled via rigid beams with constant stiffness and the rail roughness is simulated as a stationary stochastic process with the spectral representation method [12]. The quality of roughness is defined according to the German spectra for high-speed railways [10]. Note that different roughness profiles are assumed for the left and right rails.

Simulated acceleration data are generated by employing the entire train model running on straight portions of track at a constant speed of 150 km/h. The sampling frequency of measured acceleration data is set to $f_s = 1000$ Hz. For the

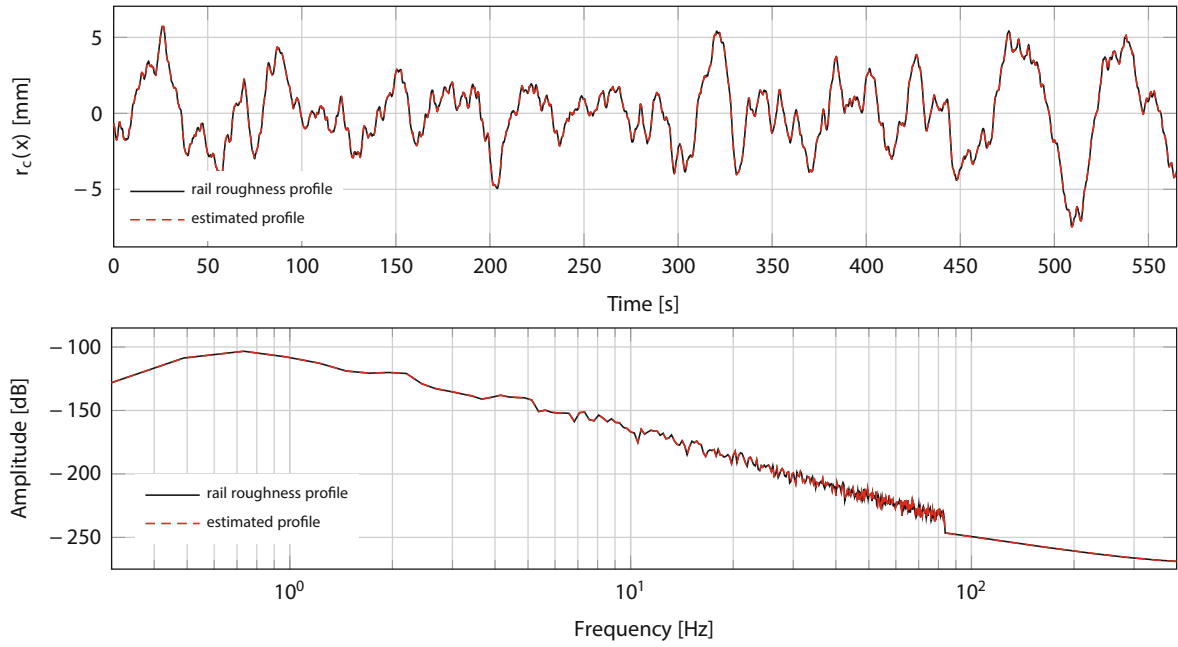


Fig. 11.2 Time-histories (top) and one-sided power spectral densities (bottom) of the true (black) and estimated via DKF (red) roughness profiles of the left rail using the complete vehicle model

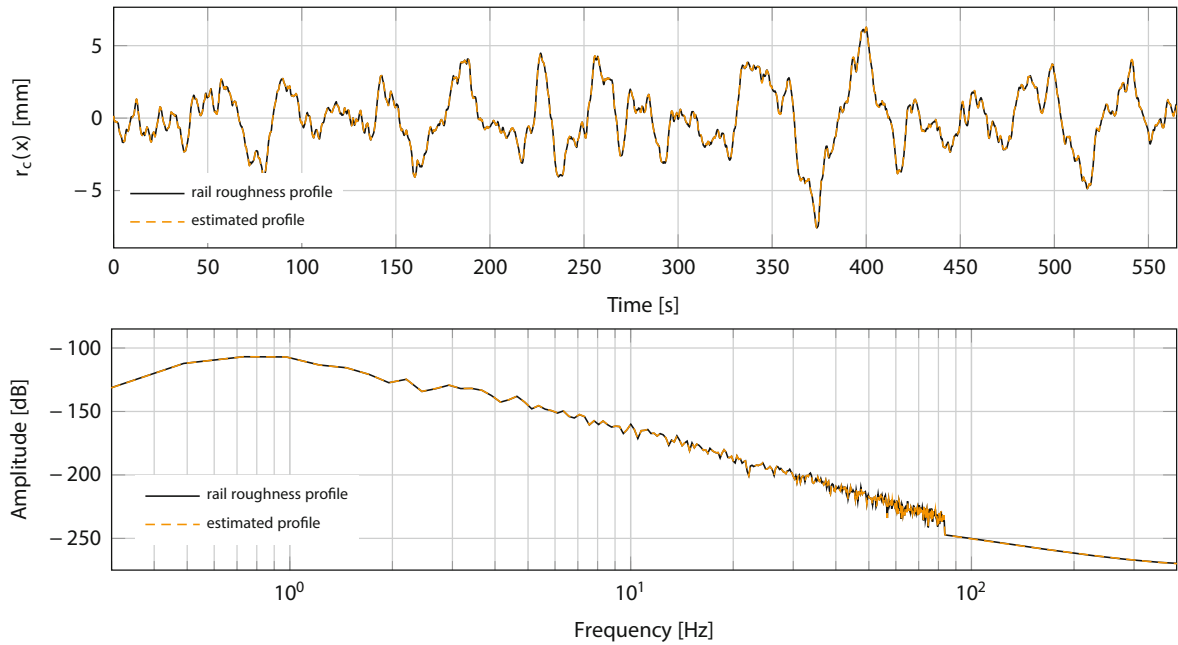


Fig. 11.3 Time-histories (top) and one-sided power spectral densities (bottom) of the true (black) and estimated via DKF (orange) roughness profiles of the right rail using the complete vehicle model

identification task, measurements of the axle boxes (above the wheels of each wheelset) are considered to be available as displacement data are usually more difficult to measure.

For the identification of rail roughness, the numerical application first considers the entire vehicle model (35 DOFs). The discretization of state-space matrices follows a first-order hold, as in Eq. (11.16), with sampling period $T_s = 10^{-3}$ s. The various noise parameters are defined next. Assuming an accurate vehicle model, the covariance matrix of the process noise is set to $\mathbf{Q}^w = 10^{-10} \cdot \mathbf{I}_1$, where \mathbf{I}_1 is an identity matrix with dimension equal to the number of the system states. The covariance matrix of the measurement noise of the obtained acceleration data is considered to be $\mathbf{Q}^f = 10^{-2} \cdot \mathbf{I}_2$, where \mathbf{I}_2

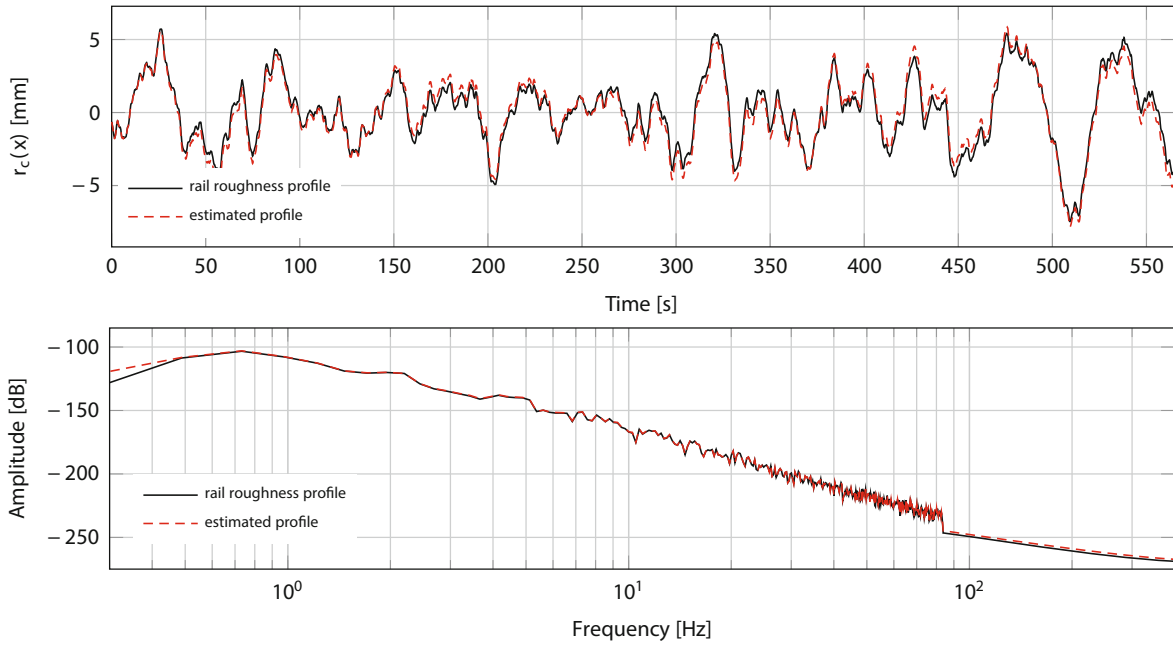


Fig. 11.4 Time-histories (top) and one-sided power spectral densities (bottom) of the true (black) and estimated via DKF (red) roughness profiles of the left rail using the reduced vehicle model

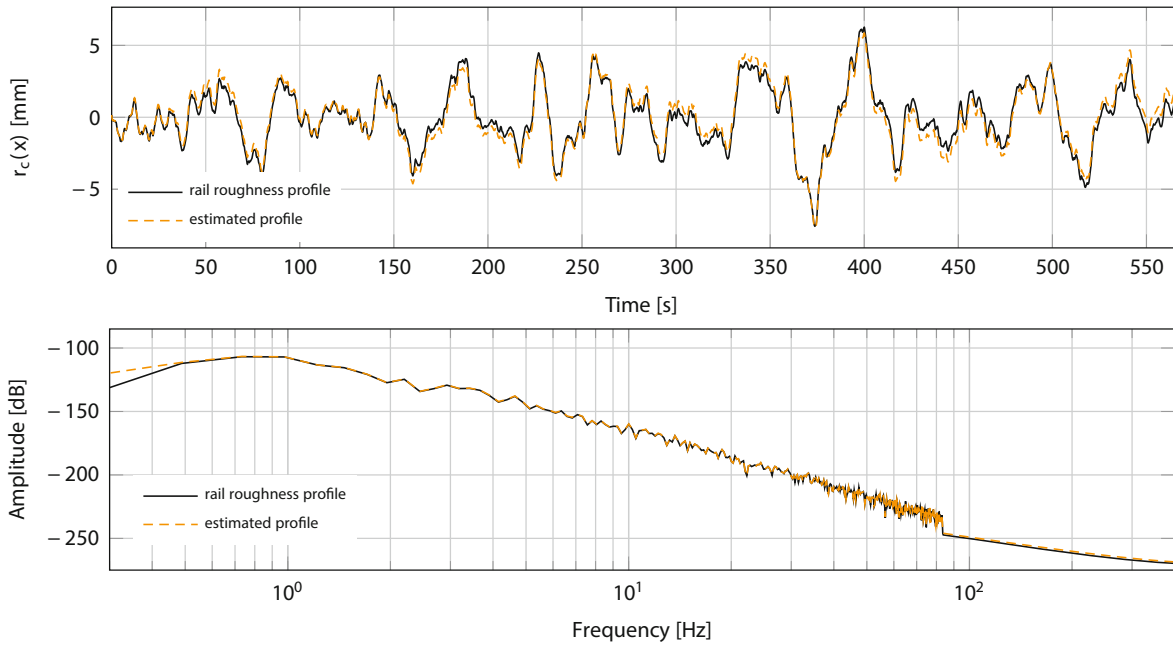


Fig. 11.5 Time-histories (top) and one-sided Power Spectral Densities (bottom) of the true (black) and estimated via DKF (orange) roughness profiles of the right rail using the reduced vehicle model

is an identity matrix with dimension equal to the number of measurements. Lastly, the input noise \mathbf{Q}^y is estimated according to an L-curve. The estimated value is $\mathbf{Q}^y = 10^7 \cdot \mathbf{I}_3$, where \mathbf{I}_3 is an identity matrix with dimension equal to the number of inputs.

Figures 11.2 and 11.3 illustrate the true and estimated time histories and one-sided power spectral densities (PSDs) of the left and right rail roughness profiles considering the entire vehicle model for the application of DKF. There is perfect agreement between the true and estimated rail roughness profiles as no error is added to the measured data and no reduction of the model is performed. Also, according to the PSD graphs of Figs. 11.2 and 11.3 (bottom) all frequencies of interest are identified.

Following, the vehicle model is reduced via modal analysis, and the final number of modes is 16. Since the roughness of the left and right rails are different, both vertical and rolling modes need to be retained to properly capture the dynamics of the train. The number of modes can be reduced even more to further enhance computational efficiency, but the accuracy of the identification task is expected to drop. In addition, in the case that the same roughness profile for both rails is considered, the number of modes can be further decreased (to, e.g., 8 modes) as only vertical modes are necessary. Figures 11.4 and 11.5 demonstrate the true and identified rail roughness profiles of the left and right rails, respectively, in the case of the reduced vehicle model. As the selected modes capture the dynamics of the system in both vertical and rolling directions, the identified roughness profiles agree very well with the true roughness profiles of the rails. Some small discrepancies appear as the dynamics in other directions (lateral, yawing, and pitching) is ignored.

11.6 Conclusion

This study proposes a rail roughness identification scheme that adopts reduced, yet realistic, train–track models and Bayesian filtering. The adoption of realistic train models is motivated by the need for continuous monitoring of railway infrastructure that could rely on the use of in-service trains. To this end, detailed train models can enhance the accuracy and reliability of such schemes. The coupling between train and track is considered via a Hertzian contact model, while continuous contact at all times is assumed.

As practical train models may consist of hundreds of DOFs, this study additionally suggests a model reduction based on modal analysis. This decreases the number of DOFs increasing the computational efficiency of the identification task. The identification of rail roughness relies on the use of a dual Kalman filter for joint input and state estimation, which considers the roughness profile as the input of the estimation problem.

A numerical application on a 3D train traversing practical rail profiles reveals that the proposed method can accurately reconstruct the applied input, that is, the traversed rail profiles. When the entire vehicle model is adopted, there is perfect agreement between the real and identified rail profiles. The proposed approach still returns very good results in the case of a reduced vehicle model. The accuracy of the identification task relies on the number of modes finally retained.

Acknowledgments This study was supported by the Stavros Niarchos Foundation through the ETH Zürich Foundation and the ETH Zürich Postdoctoral Fellowship scheme.

References

1. Weston, P., Roberts, C., Yeo, G., Stewart, E.: Perspectives on railway track geometry condition monitoring from in-service railway vehicles. *Veh. Syst. Dyn.* **53**, 1063–1091 (2015)
2. Urda, P., Aceituno, J.F., Mu noz, S., Escalona, J.L.: Measurement of railroad track irregularities using an automated recording vehicle. *Measurement* **183**, 109765 (2021)
3. Dertimanis, V.K., Zimmermann, M., Corman, F., Chatzi, E.N.: On-board monitoring of rail roughness via axle box accelerations of revenue trains with uncertain dynamics. In: *Model Validation and Uncertainty Quantification*, vol. 3. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 167–171 (2019)
4. Hoelzl, C., Dertimanis, V.K., Landgraf, M., Ancu, L., Zurkirchen, M., Chatzi, E.N.: On-board monitoring for smart assessment of railway infrastructure: a systematic review. In: *The Rise of Smart Cities. Advanced Structural Sensing and Monitoring Systems (2022)*, pp. 223–259
5. Bocciolone, M., Caprioli, A., Cigada, A., Collina, A.: A measurement system for quick rail inspection and effective track maintenance strategy. *Mech. Syst. Signal Process.* **3**, 1242–1254 (2007)
6. Lee, J.S., Choi, S., Kim, S.S., Park, C., Kim, Y.G.: A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie. *IEEE Trans. Instrum. Meas.* **61**, 749–758 (2012)
7. Stoura, C.D., Dimitrakopoulos, E.G.: Additional damping effect on bridges because of vehicle-bridge interaction. *J. Sound Vib.* **476**, 115294 (2020)
8. Tatsis, K., Dertimanis, V., Papadimitriou, C., Lourens, E., Chatzi, E.: A general substructure-based framework for input-state estimation using limited output-only measurements. *Mech. Syst. Signal Process.* **150**, 107223 (2021)
9. Azam, S.E., Chatzi, E.N., Papadimitriou, C.: A dual Kalman filter approach for state estimation via output-only acceleration measurements. *Mech. Syst. Signal Process.* **60–61**, 866–886 (2015)
10. Gao, W.W., Xia, H., De Roeck, G., Liu, K.: Integral model for train-track-bridge interaction on the Sesia viaduct: dynamic simulation and critical assessment. *Comput. Struct.* **112–113**, 205–216 (2012)
11. Zeng, Q., Stoura, C.D., Dimitrakopoulos, E.G.: A localized lagrange multipliers approach for the problem of vehicle-bridge interaction. *Eng. Struct.* **168**, 82–92 (2018)
12. Dimitrakopoulos, E.G., Zeng, Q.: A three-dimensional dynamic analysis scheme for the interaction between trains and curved railway bridges. *Comput. Struct.* **149**, 43–60 (2015)