

# Chebyshev Approximation by Linear Combinations of Fixed Knot Polynomial Splines with Weighting Functions

Nadezda Sukhorukova<sup>1,2</sup> · Julien Ugon<sup>3</sup>

Received: 30 June 2015 / Accepted: 27 January 2016 / Published online: 5 February 2016  
© Springer Science+Business Media New York 2016

**Abstract** In this paper, we derive conditions for best uniform approximation by fixed knots polynomial splines with weighting functions. The theory of Chebyshev approximation for fixed knots polynomial functions is very elegant and complete. Necessary and sufficient optimality conditions have been developed leading to efficient algorithms for constructing optimal spline approximations. The optimality conditions are based on the notion of alternance (maximal deviation points with alternating deviation signs). In this paper, we extend these results to the case when the model function is a product of fixed knots polynomial splines (whose parameters are subject to optimization) and other functions (whose parameters are predefined). This problem is nonsmooth, and therefore, we make use of convex and nonsmooth analysis to solve it.

**Keywords** Uniform approximation · Polynomial splines · Alternance · Weighting functions

**Mathematics Subject Classification** 90C51 · 41A15 · 41A50

---

✉ Nadezda Sukhorukova  
nsukhorukova@swin.edu.au

Julien Ugon  
j.ugon@federation.edu.au

<sup>1</sup> Swinburne University of Technology, Melbourne, Australia

<sup>2</sup> Federation University Australia, Ballarat, Australia

<sup>3</sup> Centre for Informatics and Applied Optimization (CIAO), Federation University Australia, Ballarat, Australia

## 1 Introduction

In this paper, we obtain optimality conditions for the problems of approximating a function by linear combinations of fixed knots polynomial splines (whose parameters have to be optimized) and other continuous functions with predefined parameters. This problem is an extension of the theory of polynomial and polynomial spline approximation developed throughout the second half of the twentieth century.

The theory of Chebyshev approximation appeared in the mid-nineteenth century (Chebyshev) and was further developed in the twentieth century [1–8]. Most optimality conditions are based on the notion of alternating extreme deviation point (that is, maximal deviation points with alternating deviation signs). In the first part of this paper, we extend this theory to the approximation by weighted functions to obtain similar alternance conditions.

The objective functions appearing in optimization problems modelling Chebyshev approximation (that is, minimization of the maximal absolute deviation) are typically nonsmooth. In this paper, we propose an approach based on the notion of convex subdifferentials, developed in [9]. Another powerful nonsmooth analysis tool, quasidifferentiability [10, 11], has been successfully applied to improve the existing optimality conditions in the case of free knot polynomial spline approximation [6]. Quasidifferentiability is one of the modern nonsmooth optimization techniques, which can be considered as a generalization of subdifferentials to the case of nonconvex functions (the corresponding objective functions in [6] are nonconvex, and therefore, convex functions subdifferentials are not very efficient). Note also that the problem of best polynomial (and fixed knots polynomial spline) approximation (in its discrete version) can be reduced to a linear programming problem [7] and solved using linear optimization methods.

In general, there are many interesting interconnections between approximation and optimization. It has been pointed out in [12] that the development of powerful optimization techniques is essential for approximation, since it is heavily relying on optimization.

In the second part of this paper, we consider two applications of the theory of approximation by weighted functions, based on two different weighting functions. The first problem arises in the area of signal processing [13–15] and the other in the area of optimal tax modelling.

The paper is organized as follows. In Sect. 2, we introduce necessary definitions and provide existing results from the theory of fixed knot polynomial spline approximation, in particular, alternance-based necessary and sufficient optimality conditions. Then, in Sect. 3, we demonstrate how these conditions can be generalized to the case when polynomial splines combined with other functions. In Sect. 4, we develop alternance-based necessary optimality conditions for the case when more than one spline is involved in approximation. Section 5 provides examples of relevant real-life applications in signal processing and taxation modelling. Finally, in Sect. 6, we give conclusions and further research directions.

## 2 Preliminaries

### 2.1 Polynomial Splines

We start with a number of definitions.

**Definition 2.1** (*Polynomial Spline*) A polynomial spline is a piecewise polynomial. Each polynomial piece lies on an interval

$$[\xi_i, \xi_{i+1}], \quad i = 0, \dots, N - 1.$$

The points  $\xi_0$  and  $\xi_N$  are the *external knots*, and  $\xi_i, i = 1, \dots, N - 1$  are the *internal knots* of the polynomial spline.

Generally, the spline is not infinitely differentiable at its knots; moreover, it can be discontinuous [1]. In this paper, only continuity of the spline is required.

**Definition 2.2** A function that is used to approximate (model) the original function or data is called the modelling function.

Polynomials and polynomial splines as well as trigonometric polynomials are very common modelling functions. The choice of modelling functions is based on the nature of the problem and the goals of approximation.

**Definition 2.3** The difference between the modelling function and the function to approximate is called the *deviation*.

Our aim is to minimize the maximal absolute deviation.

**Definition 2.4** Points in the interval  $[\xi_0, \xi_N]$ , where maximal absolute deviation occurs, are called maximal deviation points or extreme points.

Polynomial splines can be constructed in different ways. The most common one is through the truncated power function [1, Appendix, p. 191]:

$$(t - \tau)_+^j = \begin{cases} 0, & \text{if } t < \tau; \\ (t - \tau)^j, & \text{if } t \geq \tau. \end{cases}$$

Let  $X = (a_{00}, x_0, \xi_1, x_1, \dots, \xi_{N-1}, x_{N-1}) \in \mathbb{R}^{(m+1)N}$ , where

$$x_i = (a_{i1}, \dots, a_{im}) \in \mathbb{R}^m, \quad i = 0, \dots, N - 1$$

and

$$a = \xi_0 \leq \xi_1 \leq \dots \leq \xi_{N-1} \leq \xi_N = b, \quad (1)$$

then

$$s(t) = s[X](t) = a_{00} + \sum_{i=0}^{N-1} \sum_{j=1}^m a_{ij} (t - \xi_i)_+^{m+1-j}. \quad (2)$$

In this paper, we use an alternative formulation, proposed in [16], which is more flexible and enables one to have different degrees for different intervals:

$$s(t) = s[X](t) = a_{00} + \sum_{i=0}^{N-1} \sum_{j=1}^{m_i} a_{ij} (\min\{t, \xi_{i+1}\} - \xi_i)_+^{m_i+1-j}, \tag{3}$$

where  $m_i$  is the polynomial degree at the  $i$ th interval. This formulation also allows one to work with arbitrary knot multiplicities, that is, the difference between the polynomial degree and the order of the highest continuous derivative of the spline at a given knot (see also [8] where a different formulation has been used).

### 2.2 Fixed Knots Polynomial Spline Approximation

In [1], the authors developed necessary and sufficient optimality conditions for fixed knots polynomial splines with the same degree at every subinterval. Then in [16], these conditions have been generalized to the case when each subinterval can have different degree  $m_i$ ,  $i = 1, \dots, N$ . This result can be reformulated as follows:

**Theorem 2.1** ([16]) *A polynomial spline  $s$  with knots  $\xi_0, \dots, \xi_N$  and degree  $m_i$  on the subinterval  $[\xi_{i-1}, \xi_i]$  is a best approximation to a function  $f$  if and only if there exists a subinterval*

$$[\xi_{p-1}, \xi_q], \quad 0 \leq p \leq q \leq N$$

with  $2 + \sum_{i=p}^q m_i$  maximal deviation points, such that the sign of the deviation function  $\text{sign}(s - f)$  is alternating at these points.

The proof is based on reformulating the necessary condition for a minimum for a convex problem as a linear system involving a matrix whose diagonal blocks are Vandermonde-type matrices and on Cramer’s rules of solving linear systems. Note that in case of free knots spline approximation the corresponding optimization problems are much more complicated. In particular, they are nonconvex [1, 6, 12]. This type of problems are out of scope of this paper.

## 3 Optimality Conditions

### 3.1 Objective Function

In this paper, we develop optimality conditions for the following optimization problem:

$$\min f(X) = \max_{t \in [a,b]} |y(t) - (w(t)s[X](t) + h(t))|, \text{ s.t } X \in \mathbb{R}^{\gamma+1}, \tag{4}$$

where

$m_i$ ,  $i = 1, \dots, N$  is the highest possible degree of the corresponding polynomials at the  $i$ th subinterval;

$N$  is the number of subintervals (knots are fixed);  
 $[a, b]$  is the approximation interval;  
 $\gamma = \sum_{i=1}^N m_i$ ;  
 $y(t)$  is the function to approximate;  
 $s[X](t)$  is a polynomial spline of degree  $M = (m_1, \dots, m_N)$ , whose parameters ( $X$ ) are subject to optimization and knots ( $\xi_i, i = 0, \dots, N$ ) are fixed;  
 $w(t)$  and  $h(t)$  are the weighting functions defined on  $[a, b]$ ;  
 $z(t) = w(t)s[X](t) + h(t)$  is the modelling function.

**Definition 3.1** Function  $s[X](t)$  whose parameters  $X$  are subject to optimization is also called core approximation function, while  $w(t)$  and  $h(t)$  are given weighting functions.

Note that the additional term  $h(t)$  can be omitted by replacing  $y(t)$  by  $y(t) - h(t)$ , and therefore, it is enough to study a simpler problem:

$$\text{minimize } f(X) = \max_{t \in [a, b]} |y(t) - w(t)s[X](t)|, \text{ s.t } X \in \mathbb{R}^{\gamma+1}. \quad (5)$$

Since  $y(t) - (w(t)s[X](t))$  is linear with respect to  $X$ , one can conclude that  $q(X) = |y(t) - w(t)s[X](t)|$  is convex in  $X$ . Therefore, the objective function (5) is convex as supremum of convex functions.

When the function  $w \equiv 1$ , Problem (5) is a classical polynomial spline approximation problem and Theorem 2.1 applies.

Therefore, one can apply convex analysis from [9] to solve this problem or reformulate this problem as a linear programming problem [7, 17, 18] and solve it using a linear optimization method. The linear programming approach may be preferable if the available convex optimization method is not adapted to this problem. In Sect. 3.2, we apply these techniques to develop necessary and sufficient optimality conditions for the optimization problem (5).

**Definition 3.2** We denote the deviation function at point  $t$  by

$$\psi_t(s) = \psi(s, t) = y(t) - w(t)s[X](t).$$

**Definition 3.3** The function  $w(t)\psi(X, t)$  for a chosen  $X$  is called the *weighted deviation function*. The function  $w$  is the *weighting function*.

### 3.2 Optimality Conditions

Consider a convex function  $f : \mathbb{R}^{\gamma+1} \rightarrow \mathbb{R}$ . A point  $x^*$  is a minimizer of  $f$  if and only if

$$\mathbf{0}_{\gamma+1} \in \partial f(x^*), \quad (6)$$

where  $\partial f(x^*)$  is the subdifferential of the convex function  $f$  at point  $x^*$  [9].

The subdifferential of the objective function (5) at point  $X$  can be constructed as follows (using subdifferential calculus, see [9]):

$$\text{co}\{w(\tau)\text{sign}(\psi(X, \tau))T(\tau), \tau \in \tau_i, i = 1, \dots, N\}, \quad (7)$$

where  $\tau_i \in I_i$ ,  $i = 1, \dots, N$  are the sets of maximal deviation points that belong to the subinterval  $[\xi_{i-1}, \xi_i]$  and  $T(\tau)$  are as follows:

$$\begin{aligned} (1, \Delta(\tau, m_1, \xi_0), \mathbf{0})^T &\in \mathbb{R}^{\gamma+1}, & \tau &\in [\xi_0, \xi_1]; \\ (1, \Delta(\xi_1, m_1, \xi_0), \Delta(\tau, m_2, \xi_1), \mathbf{0})^T &\in \mathbb{R}^{\gamma+1} & \tau &\in [\xi_1, \xi_2]; \\ \vdots & & \vdots & \\ (1, \Delta(\xi_1, m_1, \xi_0), \Delta(\xi_2, m_2, \xi_1), \dots, \Delta(\tau, m_N, \xi_{N-1}))^T &\in \mathbb{R}^{\gamma+1}, & \tau &\in [\xi_{N-1}, \xi_N], \end{aligned}$$

where

$$\Delta(\alpha, \beta, \lambda) = ((\alpha - \lambda), (\alpha - \lambda)^2, \dots, (\alpha - \lambda)^\beta) \in \mathbb{R}^\beta.$$

In general, each  $I_i$ ,  $i = 1, \dots, n$  can be infinite. However, due to Carathéodory’s theorem [19], if (6) holds, it is possible to construct  $\mathbf{0}_{\gamma+1}$  using at most  $\gamma + 2$  points from the subdifferential  $\partial f(X^*)$ . Therefore, there exists a set of at most  $\gamma + 2$  maximal deviation points  $\tau_1, \dots, \tau_\sigma$  and a set of corresponding positive coefficients  $\lambda_1, \dots, \lambda_\sigma$ , such that

$$[w(\tau_1)\text{sign}(\psi(X, \tau_1))T(\tau_1), \dots, \tau_\sigma \text{sign}(\psi(X, \tau_\sigma))T(\tau_\sigma)]\Lambda = \mathbf{0}_\sigma, \tag{8}$$

where  $\Lambda = (\lambda_1, \dots, \lambda_\sigma) \in \mathbb{R}^\sigma$  and  $\sigma \leq \gamma + 2$ .

A similar system has been considered in [16], where  $w(t) \equiv 1$  for any  $t \in [a, b]$  (approximation by spline functions). One additional difficulty here is that  $w(t)$  may take zero values at maximal deviation points.

**Theorem 3.1** *A spline  $s[X]$  provides a best weighted approximation to  $y$  in (4) if and only if at least one of the following conditions are satisfied.*

1.  $w(t)$  is zero at one or more maximal deviation points, or
2. there exists a subinterval

$$[\xi_{p-1}, \xi_q], \quad 0 \leq p \leq q \leq N$$

with  $2 + \sum_{i=p}^q m_i$  maximal deviation points, such that the sign of the weighted deviation function

$$w(t)\text{sign}(\psi(X, t))$$

is alternating at these points.

*Proof* First assume that

$$w(\tau_i) \neq 0, \quad i = 1, \dots, \sigma.$$

Follow the same procedure as in [16]. The columns of the system in (8) are the same as in system (11) from [16]. Therefore, the study of the existence of a strictly positive solution is very similar, since the magnitude of the corresponding components of

$\Lambda \in \mathbb{R}^\sigma$  can be changed (providing that they remain nonzero). Similar to [16], obtain that

- there exists a sequence of  $2 + \sum_{i=p}^q m_i$  maximal deviation points located at  $[\xi_p, \xi_q]$  and
- the sign of  $w(t)\psi(X, t)$  is alternating at these points.

Now assume that there exists a maximal deviation point  $t^*$  and  $w(t^*) = 0$ . In this case, the optimality condition (8) is automatically satisfied. Another way to see it is based on the fact that the deviation cannot be changed at the point  $t$  where  $w$  vanishes, and therefore, it cannot be improved at this points. Hence, the current maximal deviation cannot be improved, and therefore, the current solution is optimal. □

Note that

- the same result can be obtained in the case of discrete (data) approximation, where only discrete values of  $t$  from  $[a, b]$  are considered;
- the obtained results resemble the classical alternance-based optimization conditions developed for polynomials and fixed knots polynomial splines.

### 4 Generalization to Linear Combinations with More Than One Spline

Since the sets of polynomials and polynomial splines are closed under the sum, in classical approximation theory it is not necessary to consider approximating a function by a sum of polynomials or polynomial splines. In our case, however, it makes sense to consider several weighting functions. In this section, we extend the results of Section 3 to the case of linear combinations with more than one spline. We start with the corresponding optimization problem introduction.

$$\min f(X) = \max_{t \in [a, b]} \left| y(t) - \sum_{j=1}^l w_j(t) s_j[X](t) \right|, \text{ s.t. } X \in \mathbb{R}^{\gamma+1} \tag{9}$$

where, similar to Sect. 3,

$\gamma = \sum_{i=1}^N m_i$ ;

$l$  is the number of components in the sum of the modelling function;

$N$  is the number of subintervals (knots are fixed);

$m_i, i = 1, \dots, N$  is the highest possible degree of the corresponding polynomials at the  $i$ th subinterval;

$y(t)$  is the function to approximate;

$z(t) = \sum_{j=1}^l w_j(t) s_j[X](t)$  is the modelling function;

$s_j[X](t), j = 1, \dots, l$  are polynomial spline of degree  $M = (m_1, \dots, m_N)$ , whose parameters are subject to optimization and knots are fixed;

$[a, b]$  is approximation interval;

$w_j(t), j = 1, \dots, l$  are continuous functions defined on  $[a, b]$ .

The objective function (9) is convex as a supremum of convex functions, and therefore, one can formulate necessary and sufficient conditions for such a function using subdifferential calculus. The coordinates of the points from the corresponding subdifferential consist of  $l$  blocks (one block per spline), and each block is similar to the subdifferential in the case of one spline linear combination.

Again, similar to Sect. 3, one can write a necessary and sufficient optimality condition

$$\mathbf{0}_{l(\gamma+1)} \in \partial f(X), \tag{10}$$

where  $f(X)$  is defined in (9). Essentially, this means that now we need to solve  $l$  linear systems simultaneously, seeking a nonzero nonnegative solution. Since each system is similar to one spline combination studied in Sect. 3, one can formulate the following necessary optimality condition.

**Theorem 4.1** *If the splines  $s_1[X], \dots, s_l[X]$  provide a best weighted approximation to the function  $y$  in (9), then for each  $j = 1, \dots, l$  there exists a subinterval  $[\xi_{p-1}, \xi_q]$ ,  $0 \leq p \leq q \leq N$  with  $2 + \sum_{i=p}^q m_i$  maximal deviation points, such that the sign of the deviation function  $w_j(t)\text{sign}(\psi(X, t))$  is alternating at these points.*

For sufficiency, one needs to ensure that all the sub-systems are solved simultaneously and, clearly, this is not always the case. The investigation of how this topic can be elaborated is one of our future research directions. The main obstacle for progressing in this direction is the fact that the behaviour of the deviation function depends on the properties of the weighting functions. Therefore, it is very hard to develop any sufficient condition that would hold for any type of weighting function  $w_j$   $j = 1, \dots, l$ . In some cases, however, it is easy to formulate optimality conditions that are necessary and sufficient.

1. Assume that all the weighting functions  $w_j$ ,  $j = 1, \dots, l$  vanish at one or more maximal deviation points. In this case, the current spline is optimal, since the current maximal deviation can not be improved.
2. Assume that none of the weighting functions vanishes at the maximal deviation points and

$$w_j(t)/w_1(t), \quad j = 2, \dots, l \tag{11}$$

is constant at any maximal deviation point, then the condition of Theorem 4.1 is a necessary and sufficient optimality condition, since it guarantees that the systems are satisfied simultaneously. Note that the condition (11) can be relaxed; namely, it is enough if it is satisfied in the  $2 + \sum_{i=p}^q m_i$  maximal deviation points, where the sign of the deviation function  $w_j(t)\text{sign}(\psi(X, t))$  is alternating.

3. In general, the number of sequences to consider in Theorem 4.1 is at most the rank of the matrix  $W = (w_j(t_i))$ .

In general, the choice of core and weighting approximation functions are application driven. Therefore, it is essential to develop optimality conditions to various types of functions.



## 5 Applications

In this section, we consider two real-life applications where linear combinations of polynomial splines with other functions are used.

### 5.1 Taxation Modelling

Optimal tax theory [20,21] is concerned with obtaining a function which represents the optimal average taxation rate for achieving certain goals (typically redistribution of wealth). This function represents the ratio of the total tax to the total (taxable) income. It is not practical, however, to use this function for personal tax calculation, and most systems are based on a tax-free threshold and a number of tax income intervals with the corresponding tax rates (positive numbers between zero and one), while  $a_{00} = 0$ . To obtain the taxation system that approximates best the optimal tax function, we can solve the following optimization problem:

$$\min_X \max_{t \in [a,b]} \left| y(t) - \frac{s[X](t)}{t} \right|,$$

where variable  $t$  is the total taxable income,  $a > 0$  is the tax-free threshold and  $s[X]$  is a linear spline, whose parameters are the taxation rates (one per subinterval) and the knots are the taxation subintervals borders. This problem is a weighted approximation problem, where the weighting function is

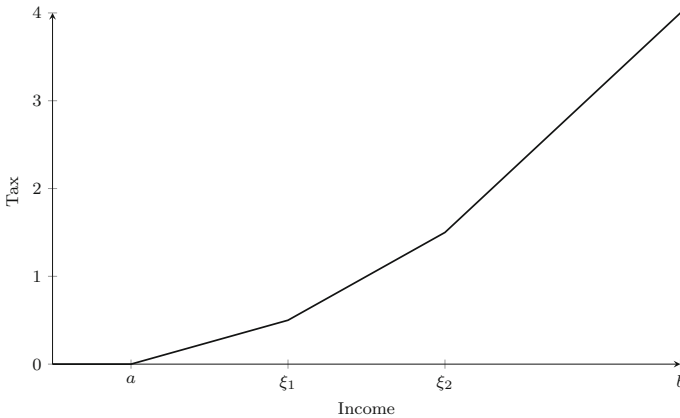
$$w(x) = \frac{1}{t}.$$

**Definition 5.1** A table that connects the taxation intervals and the corresponding taxation rates is called the taxation table.

Consider linear splines in detail (see also Fig. 1)

$$s[X](t) = a_{00} + \sum_{j=1}^N x_j (\min\{t, \xi_{j+1}\} - \xi_j)_+, \quad (12)$$

where  $a_{00} = 0$ ,  $\xi_0 = a$  is the tax-free threshold,  $\xi_N = b$  is the highest possible income considered in this problem,  $\xi_j$ ,  $j = 1, \dots, N$ ,  $\xi_0 = a$  are the borders of the tax intervals and  $x_j$  are the corresponding rates (for simplicity, we remove the index that corresponds to the spline degree, since the degree is always one). Therefore, if  $t < a$ , the income is below the tax-free threshold and there is no tax to pay. For an income in the first interval, that is  $a = \xi_0 \leq t \leq \xi_1$ , then the first  $a$  dollars are untaxed, and the rest of the income ( $(t - a)$  dollars) is taxed at the rate  $x_1$ . For an income in the second interval, that is  $\xi_1 \leq t \leq \xi_2$ , then again the first  $a$  dollars of income is untaxed, the next  $(\xi_1 - \xi_0)$  dollars of income are taxed at the rate  $x_1$ , and finally, the rest of the income ( $(t - \xi_1)$  dollars) is taxed at the rate  $x_2$ . Continuing this process through all



**Fig. 1** Average taxation rates function with thresholds  $a$  (tax-free threshold),  $\xi_1$  and  $\xi_2$

income thresholds, it can be seen that indeed the total tax is calculated using a linear spline whose knots are the borders of the taxation subintervals and the coefficients are the rates. Therefore, if the knots (taxation subintervals) are known, then the total tax amount as a function of income  $t$  can be modelled by optimizing the corresponding spline coefficients (taxation rates).

In most taxation systems, there are a number of requirements on spline coefficients. For example, for progressive taxation systems the rates that corresponds to higher income are higher, namely,

$$0 < x_1 < \dots < x_N.$$

Nonstrict inequalities are also possible, in this case two or more taxation subintervals are joined together as a single subinterval. Therefore, the optimization problems are constrained optimization problems and require some additional care to solve them. In [22], the author develops an algorithm that can take into account the fact that the corresponding spline coefficients are increasing. However, this algorithm has to be adjusted, since our modelling function is not a polynomial spline ( $w(t) = \frac{1}{t} \neq 1$ ).

Note also that the following variable substitution:

$$\bar{\xi}_1 = \xi_1, \bar{\xi}_2 = \xi_2 - \xi_1, \dots, \bar{\xi}_N = \xi_N - \xi_{N-1}$$

leads to an equivalent constraint optimization problem, where the variables are positive (or nonnegative and some of the intervals can be joined together). This variable substitution is equivalent to spline presentation (2). In this case, the knots are still interpreted as the borders of the taxation subintervals, but the spline coefficients are not interpreted as the taxation rates anymore. Then the optimization problems remain convex and we still can apply convex analysis to deal with this problem. In particular, if a minimum occurs at a point from the interior of the feasible set (that is  $X^* \in \mathbb{R}^{N+}$  for taxation problem,  $\mathbb{R}^{N+}$  is the positive quadrant in  $\mathbb{R}$ , where all the coordinates are strictly positive), then a necessary and sufficient optimality conditions is

$$\mathbf{0}_N \in \partial f(\bar{X}), \tag{13}$$

where  $f(\bar{X})$  is the maximal deviation function and the corresponding variables are

$$\bar{\xi}_1, \dots, \bar{\xi}_N.$$

These variables are strictly positive if none of the adjusted subintervals has the same taxation rate. Also note that the dimension of the problem is  $N$ , since the value of the corresponding linear spline is fixed at  $a$  and equal to zero at this point (tax-free threshold). Such kind of splines is also called fixed tails splines. In the case of fixed tail polynomial spline approximation, similar to the results with free tails, the number of alternance points is  $2 + \sum_{j=p}^q m_j$  if  $\xi_p > a$  or  $1 + \sum_{j=p}^q m_j$  if  $\xi_p = a$  (see [16] for details). Therefore, the following theorem holds.

**Theorem 5.1** *Assume that there exists an optimal taxation table for a given number of taxation subintervals  $N$ . Then the corresponding taxation table is optimal if and only if there exists a subinterval*

$$[\xi_p, \xi_q], \quad a = \xi_0 \leq \xi_p < \xi_q \leq \xi_N = b,$$

that contains

- $q - p + 2$  maximal deviation points if  $\xi_p > a$ ;
- $q - p + 1$  maximal deviation points if  $\xi_p = a$

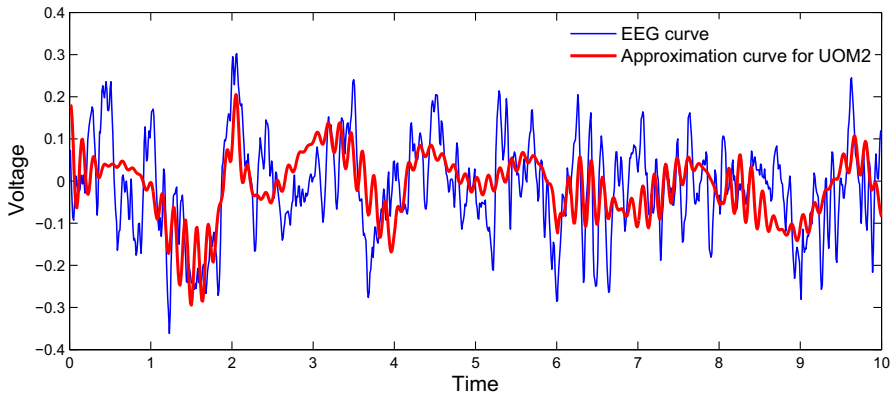
of the modelling function from the approximation function and the signs of deviation are alternating.

*Proof* Note that our assumption on the existence of the optimal taxation table implies that the solution lies in the interior of the feasible set of Problem (12), where the same optimality condition on convex functions applies. The weighting approximations function is  $\frac{1}{t} > 0$ , since  $a > 0$ . Therefore, the sign of the deviation function and the sign of the deviation function multiplied by the weighting function is the same. Then, since it is assumed that global minima achieved in the interior of the feasible region, one can follow the same steps as in Theorem 3.1. To take into account that  $y(a) = 0$ , we use fixed tails results from [16]. □

### 5.2 Signal Processing

In these problems, the original signal is modelled as a sine wave and the amplitude is chosen to be a polynomial spline whose parameters are subject to optimization [13–15]. The optimization problem is

$$\min_X \max_{t \in [a,b]} |y(t) - s[X](t) \sin(\omega t + \tau)|, \tag{14}$$



**Fig. 2** Signal processing: brain wave signal (blue) and approximating weighted spline (red)

where  $\omega$  (frequency) and  $\tau$  (phase) are known parameters. Therefore, the corresponding weighting approximation function is

$$w(x) = \sin(\omega t + \tau).$$

In most practical applications, the parameters frequency and phase are not known. One way to obtain them is to include them as additional variables in (14). However, it is not easy to solve such problems, due to its nonconvexity and a high number of local minima. To overcome this problem, there are a number of ways to estimate the frequency and phase. In some applications, where, for example, the frequency is restricted to integers and the phase does not have to be identified very precisely (sleep scoring [14]). In this case, it is possible (and efficient) to establish a fine grid for these two parameters and consider all the possible combinations, selecting the one that gives the smallest maximal deviation to (14).

In some applications, one needs to approximate the original signal by a sum of two or more waves (similar to Sect. 4). In some cases, one of the waves represents the actual signal, while the second one is the noise and has to be removed (or ignored). In this case, the problem remains convex and therefore can be solved relatively inexpensively (providing that frequencies and phases are known and the convex optimization technique can find local minima to nonsmooth functions), but the optimality conditions are not so easy to verify.

On Fig. 2, the blue curve is the original signal (brain wave, also known as EEG) and the red is the approximation. A K-complex is a special sleep event, characterized by a prominent increase in the amplitude, which then immediately returns back to the original value. The identification of such events is essential for sleep scoring. There is a clear K-complex appearing two seconds after the recording started.

In this application, the computer is attempting to “mimic” doctor’s decision on whether a K-complex occurred or not. This decision is based on a number of rules that the doctor has to apply, in particular, the signal’s frequency and amplitude that describe the “shape” of the wave (phase is not very important, since it does not contribute

to the shape of the wave). While the amplitude can be clearly identified through the scales appearing on the computer screen, the frequency identification is much harder: it is based on visual recognition between wave patterns that correspond to different frequencies. These patterns are normally learnt by the scorer at the time of training. Therefore, the frequency identification (when it is done by a medical doctor) is imprecise and limited to just a few possibilities (normally integers, not exceeding 15 Hz).

## 6 Conclusions

In this paper, we have developed necessary and sufficient optimality conditions for approximation by a model function, that is, constructed as a linear combination of a fixed knot polynomial spline and continuous functions. These conditions are generalizations of alternance type optimality conditions developed for polynomial spline approximation.

In future research, we are planning to extend this study to other types of core approximation functions, in particular, trigonometric functions. Also, it is important to address additional requirements appearing in constrained problems.

Another important research direction is to investigate free knots spline approximation. This additional flexibility will increase the quality of approximation. However, the corresponding optimization problem is much more complicated; in particular, they are nonconvex. This research direction is extremely challenging.

It is also important to develop computational methods for constructing optimal approximations. In particular, it is interesting to see whether the famous Remez algorithm [1, 23], developed for polynomial and polynomial spline approximation, can be extended to the case of linear combinations where the core approximation function is a polynomial spline (or another important class of functions).

Finally, we are planning to elaborate our results for approximating by linear combinations with more than one splines. Currently, we have developed an alternance-based necessary optimality condition and we will work towards developing a necessary and sufficient alternance-based optimality condition.

## References

1. Nürnberger, G.: *Approximation by Spline Functions*. Springer, Berlin (1989)
2. Nürnberger, G., Schumaker, L., Sommer, M., Strauss, H.: Approximation by generalized splines. *J. Math. Anal. Appl.* **108**, 466–494 (1985)
3. Rice, J.: Characterization of Chebyshev approximation by splines. *SIAM J. Numer. Anal.* **4**(4), 557–567 (1967)
4. Schumaker, L.: Uniform approximation by Chebyshev spline functions. II: free knots. *SIAM J. Numer. Anal.* **5**, 647–656 (1968)
5. Sukhorukova, N., Ugon, J.: Characterization theorem for best linear spline approximation with free knots. *Dyn. Contin. Discrete Impuls. Syst.* **17**(5), 687–708 (2010)
6. Sukhorukova, N., Ugon, J.: Characterization theorem for best polynomial spline approximation with free knots. *Trans. Am. Math. Soc.* [arXiv:1412.2323](https://arxiv.org/abs/1412.2323) (2016)
7. Gavurin, M., Malozemov, V.: *Extremum problems with linear constraints*. Leningrad University Press, Leningrad (1984)

8. Vasil'ev, A.A., Malozemov, V.N., Pevnyi, A.B.: Interpolation and approximation by spline functions of arbitrary defect. *Vestn. Leningr. Univ. [Mat.]* **12**, 262–271 (1980)
9. Rockafellar, R.: *Convex Analysis*. Princeton University Press, Princeton (1970)
10. Demyanov, V., Rubinov, A.: *Constructive Nonsmooth Analysis*. Peter Lang, Frankfurt am Main (1995)
11. Demyanov, V., Rubinov, A. (eds.): *Quasidifferentiability and Related Topics, Nonconvex Optimization and Its Applications*, vol. 43. Kluwer Academic, London (2000)
12. Nürnberger, G.: Bivariate segment approximation and free knot splines: research problems 96–4. *Constr. Approx.* **12**(4), 555–558 (1996)
13. Zamir, Z.R., Sukhorukova, N., Amiel, H., Ugon, A., Philippe, C.: Optimization-based features extraction for k-complex detection. *ANZIAM* **55**, C384–C398 (2014)
14. Zamir, Z.R., Sukhorukova, N., Amiel, H., Ugon, A., Philippe, C.: Convex optimisation-based methods for k-complex detection. *Appl. Math. Comput.* **268**, 947–956 (2015)
15. Zamir, Z.R., Sukhorukova, N.: Linear least squares problems involving fixed knots polynomial splines and their singularity study. [arXiv:1503.01176](https://arxiv.org/abs/1503.01176) (2015)
16. Sukhorukova, N.: Uniform approximation by the highest defect continuous polynomial splines: necessary and sufficient optimality conditions and their generalisations. *J. Optim. Theory Appl.* **147**(2), 378–394 (2010)
17. Demyanov, V., Malozemov, V.: *Introduction to Minimax*. Wiley, New York (1974)
18. Stiefel, E.: Note on jordan elimination, linear programming and tchebycheff approximation. *Numer. Math.* **2**(1), 1–17 (1960). doi:[10.1007/BF01386203](https://doi.org/10.1007/BF01386203)
19. Carathéodory, C.: Über den variabilitätsbereich der fourierschen konstanten von positiven harmonischen funktionen. *Rend. Circ. Mat. Palermo (2) Suppl.* **32**(1), 193–217 (1911). doi:[10.1007/BF03014795](https://doi.org/10.1007/BF03014795)
20. Mirrlees, J.A.: Optimal tax theory: a synthesis. *J. Public Econ.* **6**(4), 327–358 (1976)
21. Tuomala, M.: *Optimal Income Tax and Redistribution*. Oxford University Press, Oxford (1990)
22. Sukhorukova, N.: A generalization of the remez algorithm to a class of linear spline approximation problems with constraints on spline parameters. *Optim. Methods Softw.* **23**(5), 793–810 (2008)
23. Remez, E.Y.: *General Computation Methods of Chebyshev Approximation. The Problems with Linear Real Parameters*. Publishing House of the Academy of Science of the Ukrainian SSR, Kiev (1957). English translation, AEC-TR-4491, United States Atomic Energy Commission