Optimal Knots Selection for Sparse Reduced Data

Ryszard Kozera[1(✉)] and Lyle Noakes[2]

[1] Faculty of Applied Mathematics and Informatics, Warsaw University of Life Sciences-SGGW, Nowoursynowska Street 159, 02-776 Warsaw, Poland
ryszard.kozera@gmail.com

[2] School of Mathematics and Statistics, The University of Western Australia, 35 stirling Highway, Crawley, Perth, WA 6009, Australia
lyle.noakes@uwa.edu.au

Abstract. We discuss an interpolation scheme (based on optimization) to fit a given ordered sample of reduced data Q_m in arbitrary Euclidean space. Here the corresponding knots are not given and need to be first somehow guessed. This is accomplished by solving an appropriate optimization problem, where the missing knots minimize the cost function measuring the *total squared norm of acceleration* of the interpolant (here a natural spline). The initial infinite dimensional optimization (set to minimize an acceleration within the class of admissible curves) is reduced to the finite dimensional problem, for which the unknown optimal interpolation knots are to be found. The latter introduces a highly non-linear optimization task, both difficult for theoretical analysis and in derivation of computationally feasible optimization scheme (in particular handling medium and large number of data points). The experiments to compare the interpolants based either on optimal knots or on the so-called cumulative chords are performed for 2D and 3D data. The problem of interpolating or approximating reduced data is applicable in computer vision (image segmentation), in computer graphics (curve modeling in computer aided geometrical design) or in engineering and physics (trajectory modeling).

Keywords: Reduced sparse data · Interpolation · Knots selection

1 Introduction

The problem of fitting data points in Euclidean space E^m is a classical problem for which there exist many different interpolation or approximation techniques (see e.g. [1–6]). Most classical interpolation schemes assume a given sequence of ordered data $\mathscr{M} = \{x_0, x_1, \ldots, x_n\}$ (where $x_i \in E^m$) together with the corresponding set of ordered interpolation knots $\{t_i\}_{i=0}^n$ (*parametric interpolation on non-reduced data*). The problem of data fitting and modeling gets more complicated while dealing with the *reduced data* i.e. when only $\mathscr{M}$ is available (termed as *non-parametric interpolation*). Here, for a given fitting scheme, different choices

F. Huang and A. Sugimoto (Eds.): PSIVT 2015 Workshops, LNCS 9555, pp. 3–14, 2016.
DOI: 10.1007/978-3-319-30285-0_1

of ordered interpolation knots $\{\hat{t}_i\}_{i=0}^n$ render different curves. An early work on this topic can be found in [7] which was later extended among all in [8–19], where various quantitative criteria (often for special $m = 2, 3$) are introduced to measure the appropriateness of particular choice of $\{\hat{t}_i\}_{i=0}^n$ (e.g. convergence rate for dense data $\mathscr{M}$ derived from the unknown curve). A more recent work in which different parameterization of the unknown knots are discussed, including the so-called *cumulative chord parameterization*

$$\hat{t}_i = 0, \qquad \hat{t}_{i+1} = \hat{t}_i + \|q_{i+1} - q_i\|, \tag{1}$$

can be found e.g. in [5], [20–26]. The analysis of convergence rates to the unknown curve $\gamma : [0, T] \to E^m$ and its length $d(\gamma)$ (based on different parameterizations and dense samplings) is also recently studied among all in [27–40].

In this paper we introduce a special criterion of choosing the unknown knots (applicable not only to dense but also to sparse data) minimizing the mean squared of norm of the second derivative of the interpolating curve. An initial infinite dimensional optimization problem (see Lemma 1) is reduced to the corresponding finite dimensional one (set to determine the unknown knots). The latter constitutes a constrained highly non-linear optimization task (knots must be ordered) difficult for the theoretical analysis and computationally sensitive to the increase of interpolation points while standard optimization techniques are invoked. An alternative (not analyzed in this paper) is a computationally feasible optimization scheme called *Leap-Frog* (see [41–44]) which is here adapted to compute the suboptimal knots for ordered data in arbitrary Euclidean space E^m. The performance of the *Leap-Frog Algorithm* is illustrated on 2D and 3D reduced data $\mathscr{M}$ (i.e. for $m = 2, 3$) and subsequently compared with the multidimensional analogue of *Secant Method* (see e.g. [2] or [46]). The initial guess is chosen according to cumulative chords (1).

The proposed scheme for knots selection is applicable, in data fitting and curve modeling (e.g. computer graphics and computer vision), in approximation and interpolation (e.g. in trajectory planning, image segmentation, data compression) as well as in many other engineering and physics problems (robotics or particle trajectory estimation). Specific applications for fitting sparse (and dense) reduced data $\mathscr{M}$ in E^m can be found e.g. in [5,6] or [47].

2 Problem Formulation

Assume that *ordered (by indexing) data points* $\mathscr{M} = \{x_0, x_1, x_2, \ldots, x_n\}$ are given (here $x_i \in E^m$ and $x_i \neq x_{i+1}$, for $i = 0, 1, \ldots, n$ with $n \geq 2$). Such $\mathscr{M}$ is called *admissible data*. Define now a class (denoted by $\mathscr{I}_T$) of *admissible curves* γ as piecewise C^2 curves $\gamma : [0, T] \to E^m$ (where $0 < T < \infty$ is fixed) interpolating $\mathscr{M}$ with the ordered *free unknown knots* $\{t_i\}_{i=0}^n$ satisfying $\gamma(t_i) = x_i$ (here $t_i < t_{i+1}$, $t_0 = 0$ and $t_n = T$). More specifically, we assume that any choice of ordered interpolation knots $\{t_i\}_{i=0}^n$ yields a curve $\gamma \in C^1([0, T])$ such that it extends over sub-segment $[t_i, t_{i+1}]$ (for each $i = 0, 1, \ldots, n-1$) to $\gamma \in C^2([t_i, t_{i+1}])$ - i.e. γ is C^2 except of being C^1 only at interpolation knots $\{t_i\}_{i=0}^n$. The reason why we do not confine our analysis within a more natural class of $\gamma \in C^2([t_0, t_n])$ is justified by the subsequent choice of computational scheme (called herein *Leap-Frog* - see [42]) which effectively deals with the optimization problem (3). This scheme is designed to iteratively produce the a sequence of curves $\gamma_{LF}^k \in \mathscr{I}_T$ generically positioned outside of the class $C^2([t_0, t_n])$ (i.e. $\gamma_{LF}^k \notin C^2([t_0, t_n])$). However, the computed optimum by Leap-Frog belongs to the tighter class of functions in $C^2([t_0, t_n])$ - see [41,48].

We look for *an optimal* $\gamma_{opt} \in \mathscr{I}_T$ minimizing the following functional:

$$\mathscr{J}_T(\gamma) = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \|\ddot{\gamma}(t)\|^2 dt \;, \tag{2}$$

i.e. satisfying

$$\mathscr{J}_T(\gamma_{opt}) = \min_{\gamma \in \mathscr{I}_T} \mathscr{J}_T(\gamma) \;. \tag{3}$$

For future needs define also $\mathscr{J}_T^i$ as the i-th *segment energy*:

$$\mathscr{J}_T^i(\gamma) = \int_{t_i}^{t_{i+1}} \|\ddot{\gamma}(t)\|^2 dt \;, \tag{4}$$

obviously satisfying the inequality $\mathscr{J}_T^i(\gamma) \leq \mathscr{J}_T(\gamma)$. Note that for each function $\gamma \in \mathscr{I}_T$ the corresponding sequence of unknown interpolation knots $\{t_i\}_{i=0}^n$ (with t_0 and $t_n = T$ fixed) satisfies with $n-1$ internal components the following:

$$\Omega_{t_0}^T = \{(t_1, t_2, \ldots, t_{n-1}) \in \mathbb{R}^{n-1} : t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = T\}. \tag{5}$$

Evidently (3) defines an infinite dimensional optimization task (considered over $\mathscr{I}_T$) not invariant with respect to an arbitrary C^2 class re-parameterization $\phi : [0, T] \to [0, \tilde{T}]$ (with $\tilde{T} > 0$).

Remark 1. Note that if we confine reparametrizations' class to the affine ones i.e. $\phi(t) = t\tilde{T}/T$ (with $\phi^{-1}(s) = sT/\tilde{T}$) then as $\phi^{-1}(s) = t$, $\phi^{-1'} \equiv T/\tilde{T}$ and $\phi^{-1''} \equiv 0$, formula (2) reads for $\tilde{\gamma}(s) = (\gamma \circ \phi^{-1})(s)$ (upon using integration by substitution) as:

$$\begin{aligned}\mathscr{J}_{\tilde{T}}(\tilde{\gamma}) &= \frac{T^3}{\tilde{T}^3} \sum_{i=0}^{n-1} \int_{\tilde{t}_i}^{\tilde{t}_{i+1}} \phi^{-1'}(s) \|(\ddot{\gamma} \circ \phi^{-1})(s)\|^2 ds = \frac{T^3}{\tilde{T}^3} \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \|\ddot{\gamma}(t)\|^2 dt \\ &= \frac{T^3}{\tilde{T}^3} \mathscr{J}_T(\gamma) \;. \end{aligned} \tag{6}$$

Thus a curve $\gamma_{opt} \in \mathscr{I}_T$ is optimal to $\mathscr{J}_T$ if and only if a corresponding $\tilde{\gamma}_{opt} \in \mathscr{I}_{\tilde{T}}$ is optimal for $\mathscr{J}_{\tilde{T}}$. Therefore we can effectively assume $t_n = T$ to be arbitrary. Similar argument can be applied to $t_0 = 0$ (we set $\phi(t) = t - t_0$ if the latter does not hold). □

In the anticipation of the forthcoming materials we briefly re-introduce different families of *piecewise cubics* interpolating data points $\mathscr{M}$ (see also [10, Chap.IV]) subject to various boundary conditions. In addition, in the remark to follow we also formulate the specific energy formulation (a special case of (2)) for the family of the so-called natural splines.

Remark 2. First recall that *a cubic spline interpolant* $\gamma_{\mathscr{T}}^{C_i} = \gamma_{\mathscr{T}}^{C}|_{[t_i,t_{i+1}]}$, for given temporarily fixed interpolation knots $\mathscr{T} = (t_0, t_1, \ldots, t_{n-1}, t_n)$ (here the knots $\{t_n\}_{i=0}^{n}$ are admissible) is defined as

$$\gamma_{\mathscr{T}}^{C_i}(t) = c_{1,i} + c_{2,i}(t-t_i) + c_{3,i}(t-t_i)^2 + c_{4,i}(t-t_i)^3, \qquad (7)$$

to satisfy (for $i = 0, 1, 2, \ldots, n-1$; $c_{j,i} \in \mathbb{R}^m$, where $j = 1, 2, 3, 4$)

$$\gamma_{\mathscr{T}}^{C_i}(t_{i+k}) = x_{i+k} \ , \quad \dot{\gamma}_{\mathscr{T}}^{C_i}(t_{i+k}) = v_{i+k} \ , \quad k = 0, 1$$

with the velocities $v_0, v_1, v_2, \ldots, v_{n-1}, v_n \in \mathbb{R}^m$ assumed to be temporarily free parameters (*if unknown*). The coefficients $c_{j,i}$ (with $\Delta t_i = t_{i+1} - t_i$) are defined as follows:

$$\begin{aligned} c_{1,i} &= x_i \ , \\ c_{2,i} &= v_i \ , \\ c_{4,i} &= \frac{v_i + v_{i+1} - 2\frac{x_{i+1}-x_i}{\Delta t_i}}{(\Delta t_i)^2} \ , \\ c_{3,i} &= \frac{\frac{(x_{i+1}-x_i)}{\Delta t_i} - v_i}{\Delta t_i} - c_{4,i}\Delta t_i \ . \end{aligned} \qquad (8)$$

The latter comes from (7) and Newton's formula (see e.g. [4, Chap. 1])

$$\begin{aligned} \gamma_{\mathscr{T}}^{C_i}(t) = \gamma_{\mathscr{T}}^{C_i}(t_i) &+ \gamma_{\mathscr{T}}^{C_i}[t_i,t_i](t-t_i) + \gamma_{\mathscr{T}}^{C_i}[t_i,t_i,t_{i+1}](t-t_i)^2 \\ &+ \gamma_{\mathscr{T}}^{C_i}[t_i,t_i,t_{i+1},t_{i+1}](t-t_i)^2(t-t_{i+1}) \end{aligned}$$

combined with $c_{1,i} = \gamma_{\mathscr{T}}^{C_i}(t_i)$, $c_{2,i} = \dot{\gamma}_{\mathscr{T}}^{C_i}(t_i)$, $c_{3,i} = \ddot{\gamma}_{\mathscr{T}}^{C_i}(t_i)/2$, and $c_{4,i} = \gamma_{\mathscr{T}}^{C_i'''}(t_i)/6$. Adding $n-1$ constraints enforcing continuity of second derivatives of $\gamma_{\mathscr{T}}^{C}$ at $x_1, x_2, \ldots, x_{n-1}$ i.e. for $i = 1, 2, \ldots, n-1$ $\ddot{\gamma}_{\mathscr{T}}^{C_{i-1}}(t_i) = \ddot{\gamma}_{\mathscr{T}}^{C_i}(t_i)$ leads (upon using (7) and (8)) to the m tridiagonal linear systems (strictly diagonally dominant) of $n-1$ equations in $n+1$ vector unknowns representing velocities at $\mathscr{M}$ i.e. $v_0, v_1, v_2, \ldots, v_{n-1}, v_n \in \mathbb{R}^m$:

$$\begin{aligned} v_{i-1}\Delta t_i + 2v_i(\Delta t_{i-1} + \Delta t_i) + v_{i+1}\Delta t_{i-1} &= b_i \ , \\ b_i = 3(\Delta t_i \frac{x_i - x_{i-1}}{\Delta t_{i-1}} + \Delta t_{i-1}\frac{x_{i+1}-x_i}{\Delta t_i}) \ . \end{aligned} \qquad (9)$$

The terminal velocities v_0 and v_n (*if unknown*) can be calculated from the conditions $\ddot{\gamma}_{\mathscr{T}}^{C}(0) = \ddot{\gamma}_{\mathscr{T}}^{C}(T_c) = \mathbf{0}$ combined with (8) (this yields *a natural cubic* *spline interpolant* $\gamma_{\mathscr{T}}^{NS}$ - a special $\gamma_{\mathscr{T}}^{C}$) which supplements (9) with two missing linear equations:

$$2v_0 + v_1 = 3\frac{x_1 - x_0}{\Delta t_0}\ , \quad v_{n-1} + 2v_n = 3\frac{x_n - x_{n-1}}{\Delta t_{n-1}}. \tag{10}$$

The resulting m linear systems, each of size $(n+1) \times (n+1)$, (based on (9) and (10)) as strictly row diagonally dominant result in one vector solution $v_0, v_1, v_2, \ldots, v_{n-1}, v_n$ (solved e.g. by Gauss elimination without pivoting - see [4, Chap.4]), which when fed into (8) determines explicitly *a natural cubic spline* $\gamma_{\mathscr{T}}^{NS}$ (with fixed $\mathscr{T}$).

Combining (2) with (7) results in (*in fact it is also true for any spline* $\gamma_{\mathscr{T}}^{C}$ *provided the respective velocities* $\{v_i\}_{i=0}^{n}$ *are somhow prescribed* - e.g. also for a Hermite or a complete spline [4, Chap. 4])):

$$\begin{aligned}\mathscr{J}_T(\gamma_{\mathscr{T}}^{NS}) &= \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \|\ddot{\gamma}_{\mathscr{T}}^{NS_i}(t)\|^2 dt \\ &= \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \langle 2c_{3,i} + 6c_{4,i}(t-t_i)|2c_{3,i} + 6c_{4,i}(t-t_i)\rangle dt \\ &= 4\sum_{i=0}^{n-1}(\|c_{3,i}\|^2 \Delta t_i + 3\|c_{4,i}\|^2(\Delta t_i)^3 + 3\langle c_{3,i}|c_{4,i}\rangle(\Delta t_i)^2)\ .\end{aligned}$$

Upon introducing $c_{5,i} = c_{3,i} + c_{4,i}\Delta t_i$ the latter reduces into:

$$\mathscr{J}_T(\gamma_{\mathscr{T}}^{NS}) = 4\sum_{i=0}^{n-1}(\Delta t_i\|c_{5,i}\|^2 + (\Delta t_i)^3\|c_{4,i}\|^2 + \langle c_{5,i}|c_{4,i}\rangle(\Delta t_i)^2). \tag{11}$$

By (8) we also have $c_{5,i} = ((x_{i+1} - x_i)/(\Delta t_i)^2) - (v_i/\Delta t_i)$ and thus

$$\begin{aligned}\Delta t_i\|c_{5,i}\|^2 &= \frac{\|x_{i+1} - x_i\|^2}{(\Delta t_i)^3} + \frac{\|v_i\|^2}{\Delta t_i} - 2\frac{\langle x_{i+1} - x_i|v_i\rangle}{(\Delta t_i)^2}\ , \\ (\Delta t_i)^3\|c_{4,i}\|^2 &= \frac{\|v_i + v_{i+1}\|^2}{\Delta t_i} + \frac{4\|x_{i+1} - x_i\|^2}{(\Delta t_i)^3} - \frac{4\langle v_{i+1} + v_i|x_{i+1} - x_i\rangle}{(\Delta t_i)^2}\ , \\ \langle c_{5,i}|c_{4,i}\rangle(\Delta t_i)^2 &= \frac{\langle x_{i+1} - x_i|v_i + v_{i+1}\rangle}{(\Delta t_i)^2} - \frac{2\|x_{i+1} - x_i\|^2}{(\Delta t_i)^3} - \frac{\|v_i\|^2}{\Delta t_i} - \frac{\langle v_i|v_{i+1}\rangle}{\Delta t_i} \\ &\quad + \frac{2\langle v_i|x_{i+1} - x_i\rangle}{(\Delta t_i)^2}\end{aligned}$$

which when passed to (11) yields finally

$$\begin{aligned}\mathscr{J}_T(\gamma_{\mathscr{T}}^{NS}) = 4\sum_{i=0}^{n-1}\Big(&\frac{-1}{(\Delta t_i)^3}(-3\|x_{i+1} - x_i\|^2 + 3\langle v_i + v_{i+1}|x_{i+1} - x_i\rangle\Delta t_i \\ &-(\|v_i\|^2 + \|v_{i+1}\|^2 + \langle v_i|v_{i+1}\rangle)(\Delta t_i)^2\Big)\ .\end{aligned} \tag{12}$$

Clearly for a given data points $\mathscr{M}$ and each strictly increasing sequence of fixed knots $\{t_i\}_{i=0}^n$ *the natural spline* $\gamma_{\mathscr{T}}^{NS}$ exists and is unique. Note also that (12) (in contrast with each curve $\gamma_{\mathscr{T}}^{C_i}$ itself) involves only two vector coefficients $c_{3,i}$ and $c_{4,i}$ appearing in (7).

Once we vary the knots $\{t_i\}_{i=0}^n$ (subject to $t_0 < t_1 < t_2 < ... < t_{n-1} < t_n = T$ with $t_n = T$ and t_0 fixed) the corresponding space of such natural splines γ^{NS} (denoted here by $\mathscr{I}^{NS}$) evidently satisfies $\mathscr{I}^{NS} \subset \mathscr{I}_T \cap C^2([0,T_c])$ (also γ^{NS} is $C^\infty([t_i,t_{i+1}])$ for each $i = 0,1,\dots,n-1$ - we omit subscript $\mathscr{T}$ to emphasize that internal knots may vary).

If now we minimize (2) only over a class of natural splines $\mathscr{I}^{NS} \subset \mathscr{I}_T$ (such thinning of (2) and (3) to (12) is justified later in Lemma 1) then

$$\mathscr{J}_T(\gamma_{opt}^{NS}) = \min_{\gamma^{NS}\in\mathscr{I}^{NS}} \mathscr{J}_T(\gamma^{NS}) \tag{13}$$

reduces into finding optimal parameters for (13) $(t_1^{opt}, t_2^{opt}, \dots, t_{n-1}^{opt})$ (here terminal knots $t_n = T$ and t_0 are constant) within the family of natural splines $\mathscr{I}^{NS}$, subject to the constraint $t_0 < t_1^{opt} < t_2^{opt} < ... < t_{n-1}^{opt} < t_n = T$. As we mentioned there is a one-to-one correspondence between natural spline γ^{NS} and the interpolation knots. Consequently (13) can be reformulated (see also (5)) upon introduction $\hat{\mathscr{T}} = (t_1, t_2, \dots, t_{n-1})$ into the following minimization problem in $\hat{\mathscr{T}}$:

$$\begin{aligned}
\mathscr{J}_T(\gamma_{opt}^{NS}) &= \min_{\hat{\mathscr{T}}\in\Omega_{t_0}^T} \mathscr{J}_T^F(t_1,t_2,\dots,t_{n-1}) \\
&= \min_{\hat{\mathscr{T}}\in\Omega_{t_0}^T} \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \|\ddot{\gamma}^{NS}(s)\|^2 ds \\
&= \min_{\hat{\mathscr{T}}\in\Omega_{t_0}^{T_c}} 4\sum_{i=0}^{n-1}\Big(\frac{-1}{(\Delta t_i)^3}(-3\|x_{i+1}-x_i\|^2 + 3\langle v_i + v_{i+1}|x_{i+1}-x_i\rangle \Delta t_i \\
&\qquad -(\|v_i\|^2 + \|v_{i+1}\|^2 + \langle v_i|v_{i+1}\rangle)(\Delta t_i)^2\Big) .
\end{aligned} \tag{14}$$

Thus *an implicit* formula (12) yields, upon feeding v_i from (9) and (10), the *explicit non-linear expression* for $\mathscr{J}_T^F$ (see (14)) ready for minimization over $\mathscr{I}^{NS}$ with respect to free variables $(t_1, t_2, \dots, t_{n-1}) \in \Omega_{t_0}^T$. Note that the value of the energy $\mathscr{J}_T^F(\hat{\mathscr{T}})$ is always finite for each $\hat{\mathscr{T}} \in \Omega_{t_0}^T$ since $\gamma_{\mathscr{T}}^{NS} \in C^2$ and therefore $\|\ddot{\gamma}_{\mathscr{T}}^{NS}\|$ is continuous over a compact set $[t_0, t_n]$ (as $t_n = T < +\infty$).

The class of natural splines $\mathscr{I}^{NS}$ is invoked here since one can reduce (3) to the same optimization confined merely to the subclass of natural splines $\mathscr{I}^{NS} \subset \mathscr{I}_T$ (see [10,41,48]). In fact by [41] the following holds (for arbitrary fixed knots $t_0 < t_n = T$):

Lemma 1. *For a given admissible data points $\mathscr{M}$ in arbitrary Euclidean space E^m the subclass of natural splines $\mathscr{I}^{NS} \subset \mathscr{I}_T$ satisfies*

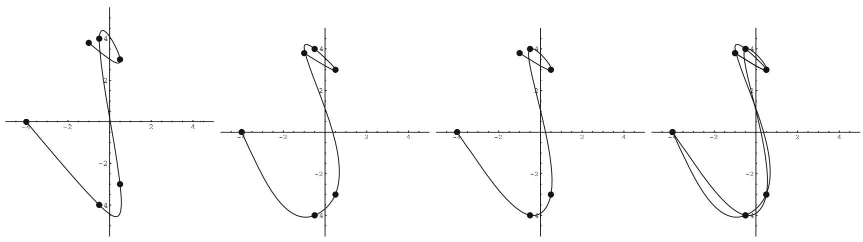
$$\min_{\gamma \in \mathscr{I}_T} \mathscr{J}_T(\gamma) = \min_{\gamma^{NS} \in \mathscr{I}^{NS}} \mathscr{J}_T(\gamma^{NS}) \ . \tag{15}$$

In the next section we test the optimization problem (15) (converted into (12) or (14)) on some 2D and 3D data.

3 Numerical Experiments

Experiments are performed with *Mathematica Package*. We compare now the performance of *Leap-Frog* (see [41]) with the *Secant Method* applied to (12) (as justified by (15)). The experiments are conducted only for sparse reduced data points $\mathscr{M}$ in $E^{2,3}$. They represent special cases of all admissible sparse reduced data in arbitrary E^m. The initial guesses are based on cumulative chords (1).

The first example deals with reduced data $\mathscr{M}$ in E^2 (i.e. for $m = 2$).



(a) (b) (c) (d)

Fig. 1. Natural splines interpolating data points $\mathscr{M}_{2D1}$ (a) $\gamma^{NS}_{\mathscr{T}_{uni}}$ with uniform knots $\mathscr{T}_{uni}$, (b) $\gamma^{NS}_{\mathscr{T}_c}$ with cumulative chords $\mathscr{T}_c$, (c) $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}}$ with optimal knots $\mathscr{T}^{LF}_{opt} = \mathscr{T}^{SM}_{opt}$ (thus $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}} = \gamma^{NS}_{\mathscr{T}^{SM}_{opt}}$) (d) $\gamma^{NS}_{\mathscr{T}_{opt}}$ and $\gamma^{NS}_{\mathscr{T}_c}$ plotted together.

Example 1. Consider for $n = 5$ the following 2D reduced data points (see dotted points in Fig. 1):

$$\mathscr{M}_{2D1} = \{(-4, 0), (-0.5, -4), (0.5, -3), (-0.5, 4), (0.5, 3), (-1, 3.8)\} \ .$$

A blind guess of uniform interpolation knots yields (rescaled to $T_c = \hat{T}_n$ - see (1)):

$$\mathscr{T}_{uni} = \{0, 3.38291, 6.76583, 10.1487, 13.5317, 16.9146\}$$

and the initial guess based on cumulative chord $\mathscr{T}_c = (t_0, \hat{\mathscr{T}}_c, T_c)$ (i.e. based on the geometry of the layout of the data) renders:

$$\mathscr{T}_c = \{0, 5.31507, 6.72929, 13.8004, 15.2146, 16.9146\}.$$

The natural splines $\gamma^{NS}_{\mathscr{T}_{uni}}$ (based on $\mathscr{T}_{uni} = (0, \hat{\mathscr{T}}_{uni}, 16.9146))$ and $\gamma^{NS}_{\mathscr{T}_c}$ (based on $\mathscr{T}_c$) yield the energies $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_{uni}) = 7.18796 < \mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_c) = 7.8536$. Both interpolants $\gamma^{NS}_{\mathscr{T}_{uni}}$ and $\gamma^{NS}_{\mathscr{T}_c}$ are shown in Fig. 1a and b, respectively. The *Secant Method* yields (for (12)) the optimal knots (augmented by terminal times $t_0 = 0$ and $t_5 = T_c$)

$$\mathscr{T}^{SM}_{opt} = \{0, 3.67209, 5.62892, 11.435, 14.5491, 16.9146\}$$

with the optimal energy $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}^{SM}_{opt}) = 4.25388$. The *execution time* amounts to $T^{SM} = 7.037922sec$. The resulting curve $\gamma^{NS}_{\mathscr{T}^{SM}_{opt}}$ is plotted in Fig. 1c. Note that for each free variable *Secant Method* uses here two initial numbers $t^c_i \pm 0.5$. The *Leap-Frog Algorithm* decreases the initial energy $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_c)$ upon 42 iterations to $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}^{SM}_{opt})$ (i.e. as for *Secant Method*) with optimal values satisfying $\mathscr{T}^{LF}_{opt} = \mathscr{T}^{SM}_{opt}$, up to the 6th decimal place - this is the iteration bound). The respective *execution time* $T^{LF} = 3.333620sec < T^{SM}$. The 0th, 1st, 10th, 20th, 30th and 42nd iterations *Leap-Frog* decrease the energy $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_c)$ to:

$$\{7.8536, 4.93366, 4.25839, 4.25389, 4.25388, 4.25388\}$$

with only the first two iterations contributing to major energy decrease (and hence the corrections of the initial guess for knots taken as $\mathscr{T}_c$). The resulting natural spline $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}}$ (clearly the same as $\gamma^{NS}_{\mathscr{T}^{SM}_{opt}}$ yielded by *Secant Method*) based on $\mathscr{T}^{LF}_{opt}$ is shown in Fig. 1c and also visually compared with $\gamma^{NS}_{\mathscr{T}_c}$ in Fig. 1d. Note that if *Leap-Frog* iteration bound condition is changed e.g. to make current *Leap-Frog* energy equal to $\mathscr{J}^F_{\mathscr{T}_c}(\mathscr{T}^{SM}_c)$ (say up to 5th decimal place) then only 22 iterations are needed here with shorter execution time $T^{LF}_E = 2.270440$ *sec* and with optimal knots

$$\mathscr{T}^{LF_E}_{opt} = \{0, 3.67502, 5.63183, 11.436814.5498, 16.9146\}.$$

We miss out now a bit on precise estimation of the optimal knots but we speed up the *Leap-Frog* execution time by obtaining almost the same interpolating curve as the optimal one (as $\mathscr{T}^{LF_E}_{opt} \approx \mathscr{T}^{SM}_{opt}$). The other iteration a posteriori stopping criteria can also be considered which even further accelerate *Leap-Frog* performance at almost no cost in difference between computed curve an optimal curve. □

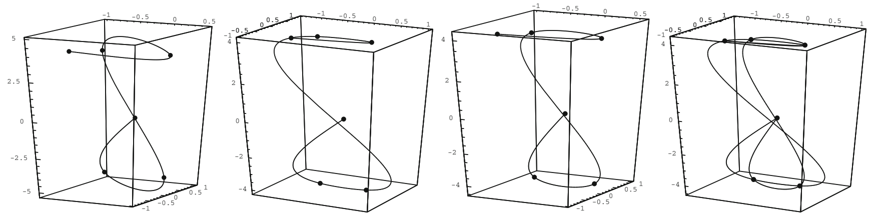
We pass now to an example of reduced data in E^3 (i.e. with $m = 3$).

Example 2. Consider for $n = 5$ the following 3D reduced data points (see dotted points in Fig. 2):

$$\mathscr{M}_{3D1} = \{(0,0,0), (-0.5,0,-4), (0.5,0,-4), (-0.5,0,4), (0.5,0,4), (-1,0,3.8)\}.$$

The uniform interpolation knots read (rescaled to $T_c = \hat{t}_n$ - see (1)) as:

$$\mathscr{T}_{uni} = \{0, 3.12133, 6.24266, 9.364, 12.4853, 15.6067\}$$



(a) (b) (c) (d)

Fig. 2. Natural splines interpolating data points $\mathscr{M}_{3D1}$ (a) $\gamma^{NS}_{\mathscr{T}_{uni}}$ with uniform knots $\mathscr{T}_{uni}$, (b) $\gamma^{NS}_{\mathscr{T}_c}$ with cumulative chords $\mathscr{T}_c$, (c) $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}}$ with optimal knots $\mathscr{T}^{LF}_{opt} = \mathscr{T}^{SM}_{opt}$ (thus $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}} = \gamma^{NS}_{\mathscr{T}^{SM}_{opt}}$) (d) $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}}$ and $\gamma^{NS}_{\mathscr{T}_c}$ plotted together.

and the initial guess based on cumulative chord $\mathscr{T}_c$ is here:

$$\mathscr{T}_c = \{0, 4.03113, 5.03113, 13.0934, 14.0934, 15.6067\}.$$

The natural splines $\gamma^{NS}_{\mathscr{T}_{uni}}$ (based on $\mathscr{T}_{uni}$) and $\gamma^{NS}_{\mathscr{T}_c}$ (based on $\mathscr{T}_c$) yield the following energies $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_{uni}) = 10.145 > \mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_c) = 9.45031$. Again, both interpolants $\gamma^{NS}_{\mathscr{T}_{uni}}$ and $\gamma^{NS}_{\mathscr{T}_c}$ are presented in Fig. 2a,b, respectively.

The *Secant Method* yields (for (12)) the optimal knots (augmented by terminal times $t_0 = 0$ and $t_5 = T_c$)

$$\mathscr{T}^{SM}_{opt} = \{0, 2.91851, 5.12399, 11.1964, 13.507, 15.6067\}$$

with the optimal energy $\mathscr{J}^F_{\mathscr{T}_c}(\mathscr{T}^{SM}_{opt}) = 4.65476$. The *execution time* amounts to $T^{SM} = 6.783365sec$. The resulting curve $\gamma^{NS}_{\mathscr{T}^{SM}_{opt}}$ is plotted in Fig. 2c. Note that for each free variable *Secant Method* uses here two initial numbers $t^c_i \pm 0.1$. *Leap-Frog* decreases the initial energy to $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}^{LF}_{opt}) = \mathscr{J}^F_{\mathscr{T}_c}(\mathscr{T}^{SM}_{opt})$ (as for the *Secant Method*) with the iteration stopping conditions $\mathscr{T}^{LF}_{opt} = \mathscr{T}^{SM}_{opt}$ (up to 6th decimal point) upon 38 iterations. The respective *execution time* amounts to $T^{LF} = 3.757498 < T^{SM}$. The 0th (i.e. $\mathscr{J}^F_{\mathscr{T}_c}(\hat{\mathscr{T}}_c)$), 1st, 2nd, 10th, 13th, and 36th iterations *Leap-Frog* decrease the energy to:

$$\{9.45031, 5.30697, 4.83704, 4.65485, 4.65476, 4.65476\}$$

with again only the first three iterations contributing to major correction of the initial guess knots $\mathscr{T}_c$. The resulting natural spline $\gamma^{NS}_{\mathscr{T}^{LF}_{opt}}$ (clearly the same as $\gamma^{NS}_{\mathscr{T}^{SM}_{opt}}$ yielded by *Secant Method*) based on $\mathscr{T}^{LF}_{opt}$ is shown in Fig. 2c and also visually compared with $\gamma^{NS}_{\mathscr{T}_c}$ in Fig. 2d. Again if *Leap-Frog* iteration bound condition is changed e.g. to make current *Leap-Frog* energy equal to $\mathscr{J}^F_{\mathscr{T}_c}(\mathscr{T}^{SM}_c)$ (say up to 5th decimal place) then only 13 iterations are needed here with shorter execution time $T^{LF}_E = 1.878057 < T^{SM}$ and optimal knots

$$\mathscr{T}^{LF_E}_{opt} = \{0, 2.92093, 5.12632, 11.1981, 13.5079, 15.6067\}.$$

As previously, we miss out here a bit on precise estimation of the optimal knots but we accelerate the *Leap-Frog* execution time by obtaining almost the same interpolating curve as the optimal one (as $\mathscr{T}_{opt}^{LF_E} \approx \mathscr{T}_{opt}^{SM}$). □

4 Conclusions

In this paper we discussed the method of fitting reduced data $\mathscr{M}$ in arbitrary Euclidean space E^m with natural splines $\gamma_{\mathscr{T}}^{NS}$ based on finding the best unknown knots $(t_1^{opt}, t_2^{opt}, \ldots, t_{n-1}^{opt})$ (and thus the best natural spline) to minimize the total mean of squared norm of acceleration of the interpolant. The original optimization problem (2) derived in a wider class of piecewise-C^2 class interpolants is reduced to the class of natural splines $\gamma_{\mathscr{T}}^{NS}$ - see Lemma 1. This in turn reformulates into the finite-dimensional constrained optimization task (14) in $(t_1, t_2, \ldots, t_{n-1})$-variables, subject to the satisfaction of the inequalities $t_0 < t_1 < t_2, < \ldots < t_{n-1} < t_n$. Two computational schemes are deployed to test the quality of the computed interpolants - i.e. Leap-Frog and Secant Method. They both do not rely on large size matrix inversion during the computational procedure. For sparse reduced data $\mathscr{M}$ our optimization set-up together with applied numerical schemes offer a feasible choice (supplemented with computational tools) of approximating the unknown interpolation knots $\{t_i\}_{i=0}^n \approx \{\hat{t}_i^{opt}\}_{i=0}^n$. Future work will include the theoretical analysis of the nature of (14) and convergence of tested iterative schemes to its local (global) minima (minimum).

Some recent related work on fitting reduced data $\mathscr{M}$ in $E^{2,3}$ can also be found in [49,50].

References

1. Bézier, P.E.: Numerical Control: Mathematics and Applications. John Wiley, New York (1972)
2. Davis, P.J.: Interpolation and Approximation. Dover Pub. Inc., New York (1975)
3. Farin, G.: Curves and Surfaces for Computer Aided Geometric Design. Academic Press, San Diego (1993)
4. de Boor, C.: A Practical Guide to Spline. Springer-Verlag, New York (1985)
5. Kvasov, B.I.: Methods of Shape-Preserving Spline Approximation. World Scientific, Singapore (2000)
6. Piegl, L., Tiller, W.: The NURBS Book. Springer-Verlag, Heidelberg (1997)
7. Epstein, M.P.: On the influence of parameterization in parametric interpolation. SIAM J. Numer. Anal. **13**, 261–268 (1976)
8. Barsky, B.A., DeRose, T.D.: Geometric continuity of parametric curves: three equivalent characterizations. IEEE. Comp. Graph. Appl. **9**(6), 60–68 (1989)
9. Boehm, E., Farin, G., Kahmann, J.: A survey of curve and surface methods in CAGD. Comp. Aided Geom. Design **1**(1), 1–60 (1988)
10. de Boor, C., Höllig, K., Sabin, M.: High accuracy geometric Hermite interpolation. Comp. Aided Geom. Design **4**(4), 269–278 (1987)
11. Hoschek, J.: Intrinsic parametrization for approximation. Comp. Aided Geom. Design **5**(1), 27–31 (1988)

12. Lachance, M.A., Schwartz, A.J.: Four point parabolic interpolation. Comp. Aided Geom. Design **8**, 143–149 (1991)
13. Lee, E.T.Y.: Corners, cusps, and parameterization: variations on a theorem of Epstein. SIAM J. Numer. Anal. **29**, 553–565 (1992)
14. Lee, E.T.Y.: Choosing nodes in parametric curve interpolation. Comp. Aided Geom. Design **21**, 363–370 (1989)
15. Marin, S.P.: An approach to data parameterization in parametric cubic spline interpolation problems. J. Approx. Theory **41**, 64–86 (1984)
16. Nielson, G.M., Foley, T.A.: A survey of applications of an affine invariant norm. In: Lyche, T., Schumaker, L.L. (eds.) Math. Methods Comp. Aided Geom. Design. pp. 445–467. Academic Press, New York (1989)
17. Schaback, R.: Interpolation in $\mathbb{R}^2$ by piecewise quadratic visually C^2 Bézier polynomials. Comp. Aided Geom. Design **6**, 219–233 (1989)
18. Sederberg, T.W., Zhao, J., Zundel, A.K.: Approximate parametrization of algebraic curves. In: Strasser, W., Seidel, H.P. (eds.) Theory and Practice in Geometric Modelling, pp. 33–54. Springer-Verlag, Heidelberg (1989)
19. Taubin, T.: Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and range image segmentation. IEEE Trans. Patt. Mach. Intell. **13**(11), 1115–1138 (1991)
20. Kocić, L.M., Simoncinelli, A.C., Della, V.B.: Blending parameterization of polynomial and spline interpolants. Facta Univ. (NIŠ) Ser. Math. Inf. **5**, 95–107 (1990)
21. Mørken, K., Scherer, K.: A general framework for high-accuracy parametric interpolation. Math. Comput. **66**(217), 237–260 (1997)
22. Noakes, L., Kozera, R.: Cumulative chords piecewise-quadratics and piecewise-cubics. In: Klette, R., Kozera, R., Noakes, L., Weickert, J. (eds.) Geometric Properties from Incomplete Data. Computational Imaging and Vision, vol. 31, pp. 59–75. Springer, The Netherlands (2006)
23. Kozera, R., Noakes, L.: Piecewise-quadratics and exponential parameterizations for reduced data. Appl. Maths Comput. **221**, 620–638 (2013)
24. Farouki, R.T.: Optimal parameterizations. Comp. Aided Geom. Design **14**(2), 153–168 (1997)
25. Rababah, A.: High order approximation methods for curves. Comp. Aided Geom. Des. **12**, 89–102 (1995)
26. Schaback, R.: Optimal geometric Hermite interpolation of curves. In: Dæhlen, M., Lyche, T., Schumaker, L. (eds.) Mathematical Methods for Curves and Surfaces II, pp. 1–12. Vanderbilt University Press, Nashville (1998)
27. Floater, M.S.: Point-based methods for estimating the length of a parametric curve. J. Comput. Appl. Maths **196**(2), 512–522 (2006)
28. Floater, M.S.: Chordal cubic spline interpolation is fourth order accurate. IMA J. Numer. Anal. **26**, 25–33 (2006)
29. Kozera, R.: Cumulative chord piecewise-quartics for length and curve estimation. In: Petkov, N., Westenberg, M.A. (eds.) CAIP 2003. LNCS, vol. 2756, pp. 697–705. Springer, Heidelberg (2003)
30. Kozera, R.: Asymptotics for length and trajectory from cumulative chord piecewise-quartics. Fundam. Inf. **61**(3–4), 267–283 (2004)
31. Kozera, R.: Curve modelling via interpolation based on multidimensional reduced data. Stud. Inf. **25**(4B–61), 1–140 (2004)
32. Kozera, R., Noakes, L.: C^1 interpolation with cumulative chord cubics. Fundam. Inf. **61**(3–4), 285–301 (2004)

33. Kozera, R., Noakes, L., Klette, R.: External versus internal parameterizations for lengths of curves with nonuniform samplings. In: Asano, T., Klette, R., Ronse, C. (eds.) Geometry, Morphology, and Computational Imaging. LNCS, vol. 2616, pp. 403–418. Springer, Heidelberg (2003)
34. Kozera, R., Noakes, L.: Smooth interpolation with cumulative chord cubics. In: Wojciechowski, B., Smółka, B., Paulus, H., Kozera, R., Skarbek, W., Noakes, L. (eds.) ICCVG 2004. Computational Imaging and Vision, vol. 32, pp. 87–94. Springer, The Netherlands (2006)
35. Noakes, L., Kozera, R.: More-or-less uniform sampling and lengths of curves. Quar. Appl. Maths **61**(3), 475–484 (2003)
36. Noakes, L., Kozera, R.: Interpolating sporadic data. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part II. LNCS, vol. 2351, pp. 613–625. Springer, Heidelberg (2002)
37. Noakes, L., Kozera, R., Klette, R.: Length estimation for curves with ε-uniform sampling. In: Skarbek, W. (ed.) CAIP 2001. LNCS, vol. 2124, pp. 518–526. Springer, Heidelberg (2001)
38. Noakes, L., Kozera, R., Klette, R.: Length estimation for curves with different samplings. In: Bertrand, G., Imiya, A., Klette, R. (eds.) Digital and Image Geometry. LNCS, vol. 2243, pp. 339–351. Springer, Heidelberg (2002)
39. Noakes, L., Kozera, R.: Cumulative chords and piecewise-quadratics. In: Wojciechowski, K. (ed.) ICCVG 2002. Association for Image Processing Poland, vol. II, pp. 589–595. Silesian University of Technology Gliwice Poland, Institute of Theoretical and Applied Informatics, PAS, Gliwice, Poland (2002)
40. Vincent, S., Forsey, D.: Fast and accurate parametric curve length computation. J. Graphics Tools **6**(4), 29–40 (2002)
41. Kozera, R., Noakes, L.: Fitting data via optimal interpolation knots. To be submitted
42. Noakes, L.: A Global algorithm for geodesics. J. Math. Austral. Soc. Ser. A **64**, 37–50 (1999)
43. Noakes, L., Kozera, R.: Nonlinearities and noise reduction in 3-source photometric stereo. J. Math. Imag. Vision **18**(3), 119–127 (2003)
44. Noakes, L., Kozera, R.: 2D Leap-Frog Algorithm for optimal surface reconstruction. In: Latecki, M.J. (ed.) SPIE 1999. Vision Geometry VIII vol. 3811, pp. 317–328. Society of Industrial and Applied Mathematics, Bellingham, Washington. (1999)
45. Noakes, L., Kozera, R.: Denoising images: non-linear Leap-Frog for shape and light-source recovery. In: Asano, T., Klette, R., Ronse, C. (eds.) Geometry, Morphology, and Computational Imaging. LNCS, vol. 2616, pp. 419–436. Springer, Heidelberg (2003)
46. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
47. Janik, M., Kozera, R., Kozioł, P.: Reduced data for curve modeling - applications in graphics, computer vision and physics. Adv. Sci. Tech. **7**(18), 28–35 (2013)
48. Noakes, L., Kozera, R.: Optimal natural splines with free knots. To be submitted
49. Kuznetsov, E.B., Yakimovich, A.Y.: The best parameterization for parametric interpolation. J. Comp. Appl. Maths **191**, 239–245 (2006)
50. Shalashilin, V.I., Kuznetsov, E.B.: Parametric Continuation and Optimal Parameterization in Applied Mathematics and Mechanics. Kluver Academic Publishers, Boston, Dordrecht, London (2003)