CONTROL THEORETIC SMOOTHING SPLINES ARE APPROXIMATE LINEAR FILTERS

C. F. MARTIN, W. DAYAWANSA
and YISHAO ZHOU

REPORT No. 13, 2002/2003, spring

ISSN 1103-467X
ISRN IML-R- -13-02/03- -SE+spring

INSTITUT MITTAG-LEFFLER
THE ROYAL SWEDISH ACADEMY OF SCIENCES

CONTROL THEORETIC SMOOTHING SPLINES ARE APPROXIMATE LINEAR FILTERS

Y. ZHOU, W. DAYAWANSA, AND C. MARTIN

ABSTRACT. The problem of constructing and approximating control theoretic smoothing splines is considered in this paper. It is shown that the optimal approximating function can be given as the solution of a forced Hamiltonian system, that can be explicitly solved using the Riccati transform, and an explicit linear filter can be constructed.

1. INTRODUCTION

In this paper we consider the problem of constructing and approximating control theoretic smoothing splines. Control theoretic splines have proven to be useful in trajectory planning for aircraft, [1], and in various trajectory planning problems in robotics, [3, 12]. The general theory of control theoretic splines and numerous variations on the theme have appeared in [11, 6].Control theoretic splines are related to the theory of L-splines developed originally by R. Varga and M. Shultz, [7], and in the context of control theory by [8]. The theory of smoothing splines is based on the seminal work of G. Wahba. The monograph, [13], is used implicitly throughout this paper.

Key words and phrases. control theoretic smoothing splines, feedback, kernel approximation, forced Hamiltonian system.

The work of the first author is supported in part by the Swedish Research Council (VR), Magnusson's Foundation through Swedish Royal Academy of Sciences and the SEB Bergvall's Foundation. The work of the third author was supported in part by NSF Grants ECS 9720357 and ECS 9705312. The support of KTH in the spring of 2000 and 2003 is gratefully acknowledged. this work was supported in part by a fellowship form the Japanese society for the Promotion of Science while the third author was in residence at Tokyo Denki University and RIKEN. This work was supported in part by a grant from the Texas Coordinating Board for Higher Education through the Advanced Research Projects Program. The third author wishes to thank the Mittag-Leffler Institute for support during the spring of 2003 and the Wenner-Gren Foundation for support during the summer of 2003.

It is well understood that polynomial smoothing splines act as a smoothing filter on noisy data and that they are in some precise sense band limited. That is, small changes in one data point has primary effect on the spline in a small neighborhood of that point. This was studied explicitly by B. Silverman in [9]. In this paper we construct an approximate linear filter for control theoretic splines. The construction is based on linear-quadratic optimal control and related filtering and tracking results. The theory of linear quadratic control is basic to linear control theory and has been studied intensively since the the seminal paper of R. Kalman, [4], appeared in 1960. We use a quadratic cost function that contains the function representing the data to be approximated and the cost function is minimized subject to the constraints of the control system that is being used to generate the approximating curve. In section two we precisely formulate the problem and find the control theoretic spline. In Section 3 we find by performing the optimization the optimal control. The optimal control is in feedback form and feedbacks to the system the difference between the data and the approximating curve. We then construct the operator equation that gives the optimal spline and discuss several of its properties. In section 4 we explicitly construct the approximating curve using the theory of linear quadratic optimal control. We show that the optimal approximating function can be given as the solution of a forced Hamiltonian system. We explicitly solve this system using the Riccati transform, [2], and construct an explicit linear filter that is the solution of the operator equation in Section 3. In Section 5 we consider a large time approximation to the finite horizon optimal control problem and show that it approximates the control theoretic spline quite closely. In Section 6 we present several examples and associated numerical results.

2. Problem Formulation

We consider the control system

$$\dot{x} = Ax + bu \tag{2.1}$$

$$y = c'x \tag{2.2}$$

where $x \in \mathbb{R}^n$, $u, y \in \mathbb{R}$ and A, b and c are constant matrices of compatible dimension and $x(0) = x_0$. We further assume that the system is controllable and observable. Because we are primarily interested in approximation rather than control we make the assumption that

$$c'b = c'Ab = \cdots = c'A^{n-2}b = 0.$$

This assumption simplifies some calculations and is natural for spline approximation.

We assume as given a data set of the form

$$D_N = \{(t_{i,N}, \alpha_{i,N}) : \quad 0 < t_{1,N} \leq t_{2,N} \leq \cdots \leq t_{N,N} < T\}. \tag{2.3}$$

We assume that T is fixed and finite. We assume that there exists a function $g_N \in C[0,T]$ so that $\alpha_{i,N} = g_N(t_{i,N}) + \epsilon_{i,N}$ where $\epsilon_{i,N}$ is a symmetrically distributed random variable. Let $S_N(t)$ be a smooth, piecewise polynomial function such that

$$S_N(t_{i,N}) = \alpha_{i,N},$$

i.e. S_N is an interpolating polynomial spline function. We assume that the data is such that there exists a function $f \in L_2[0,T]$ so that

$$\lim_{N\to\infty} \|f - S_N\|_2 = 0 \tag{2.4}$$

and we further assume that there exists a function $g \in C[0,T]$ such that the sequence $\{g_N\}$ converges uniformly to g.

We assume as given a cost function of the form

$$J_N(u) = \lambda \int_0^T u^2(t)dt + \sum_{i=1}^N w_{i,N}(y(t_{i,N}) - \alpha_{i,N})^2. \tag{2.5}$$

We allow $u \in L_2[0,T]$. We assume that $w_{i,N} > 0$ for all indices i, N and a critical assumption is that for every $g \in C[0,T]$

$$\lim_{N\to\infty} |\int_0^T g(t)dt - \sum_{i=1}^N w_{i,N} g(t_{i,N})| = 0, \tag{2.6}$$

that is, the sampling times and weights form a convergent quadrature algorithm.

We define a second cost function in terms of the function f of equation (2.4).

$$J(u) = \int_0^T \lambda u^2(t) + (y(t) - f(t))^2 dt \tag{2.7}$$

We now formulate two optimal control problems. The first will produce the control that drives the output of the linear system to form the control theoretic smoothing spline and the second will produce an approximation to the spline function and will be the object of interest for this paper.

Problem 1.

$$\min_{u\in L_2[0,T]} J_N(u)$$

subject to the constraints of the system (2.1) *and* (2.2).

Problem 2.

$$\min_{u \in L_2[0,T]} J(u)$$

subject to the constraints of the system (2.1) *and* (2.2).

We define the function $g_t(s)$ as

$$g_t(s) = \begin{cases} c' e^{A(t-s)} b & t > s \\ \\ 0 & t \leq s \end{cases} \tag{2.8}$$

Calculating the derivative of $J_N(u)$ in the direction of w and setting the derivative equal to zero we see that a necessary condition for the optimal control is given by

$$\lambda u(s) + \sum_{i=1}^{N} w_{i,N}(y(t_{i,N}) - \alpha_{i,N}) g_{t_{i,N}}(s) = 0$$

Now $y(t_{i,N})$ is a function of u. However J_N is convex and hence has a unique minimum thus we see that the optimal control must be of the form

$$u(s) = \sum_{i=1}^{N} \tau_i g_{t_{i,N}}(s)$$

and substituting this into $J_N(u)$ and doing some routine manipulation we have the finite dimensional problem of minimizing

$$J_N(\tau) = \lambda \tau' G \tau + \tau' G W_N G \tau - 2\tau' G W_N \alpha_N + \text{constant}$$

where G is the Grammian formed from the linearly independent functions $g_{t_{i,N}}(s)$, τ is the vector of unknown coefficients τ_i, W_N is the diagonal matrix with the weights $w_{i,N}$ on the diagonal and α_N is the vector of data. Calculating the derivative of $J_N(\tau)$ with respect to τ we have the necessary and sufficient condition for the optimal τ to be the solution of

$$\lambda G \tau + G W_N G \tau - G W_N \alpha_N = 0$$

and since G and W_N are non-singular we have that τ is the solution of

$$(\lambda W_N^{-1} + G)\tau = \alpha_N \tag{2.9}$$

Let $u_N(t)$ be the unique solution to Problem (1) and let u be the unique solution to problem (2). It was shown in [5] that the sequence $\{u_N(t)\}$ converges to u in a point wise manner. Thus the solution of Problem (2) is an approximation to the control theoretic spline of Problem (1). Our goal in this paper is to find a solution to Problem (2) so that

$$y(t) = \int_0^T h(t,s) f(s) ds. \tag{2.10}$$

3. Reduction to an Operator Equation

Fro typographical convenience define an operator

$$L_t(u) = \int_0^T g_t(s) u(s) ds. \tag{3.1}$$

It is clear that $y(t) = L_t(u) + c'e^{At}x_0$ and that by replacing the function f by $f(t) - c'e^{At}x_0$ the problem remains unchanged. In this section we will assume without loss of generality that $x_0 = 0$. We begin by calculating the Gatuex derivative of $DJ(u)(w)$. We have after a simple calculation

$$DJ(u)(w) = 2\int_0^T \lambda u(t) w(t) + (L_t(u) - f(t)) L_t(w) dt \tag{3.2}$$

and calculating the second derivative with respect to u and evaluating at w we have

$$D^2(u)(w) = 2\int_0^T \lambda w^2(t) + L_t(w)^2 dt. \tag{3.3}$$

From this we see that the second derivative is nonnegative and is 0 if and only if $w(t) = 0$. Thus the functional is convex and hence has a unique minimum.

We now return to equation (3.2) and set it equal to zero to obtain a necessary and sufficient condition for optimality.

$$\begin{aligned}
0 &= \int_0^T \lambda u(t) w(t) + (L_t(u) - f(t)) L_t(w) dt \\
&= \int_0^T \lambda u(s) w(s) ds + \int_0^T \int_0^t (L_t(u) - f(t)) g_t(s) w(s) ds dt \\
&= \int_0^T \lambda u(s) w(s) ds + \int_0^T \int_s^T (L_t(u) - f(t)) g_t(s) w(s) dt ds \\
&= \int_0^T \left[\lambda u(s) + \int_s^T (L_t(u) - f(t)) g_t(s) dt \right] w(s) ds
\end{aligned}$$

Now this expression is 0 for all w and hence we have that the optimal u satisfies the integral equation

$$\lambda u(s) + \int_s^T (L_t(u) - f(t))g_t(s)dt = 0 \tag{3.4}$$

Multiplying this expression by $g_t(s)$ and integrating we have

$$\lambda y(t) + \int_0^t g_t(s) \int_s^T (y(r) - f(r))g_r(s)drds = 0$$

and after a little reorganization we have

$$\lambda y(t) + \int_0^t \int_s^T g_t(s)g_r(s)y(r)drds = \int_0^t \int_s^T g_t(s)g_r(s)f(r)drds. \tag{3.5}$$

We now define the operator K as

$$K(g) = \int_0^t \int_s^T g_t(s)g_r(s)g(r)drds \tag{3.6}$$

for $g \in L^2[0,T]$. Now for every $g \in L^2[0,T]$ $K(g)$ is smooth and hence in $L^2[0,T]$. Now rewriting equation (3.5) we have

$$(\lambda I + K)(y) = K(f) \tag{3.7}$$

Lemma 3.1. *The operator K is self adjoint.*

Proof: We prove the lemma by direct calculation.

$$\begin{aligned}
\langle w, Ku\rangle &= \int_0^T w(t)K(u)(t)dt \\
&= \int_0^T w(t) \int_0^t \int_s^T g_t(s)g_r(s)u(r)drdsdt \\
&= \int_0^T \int_0^t w(t) \int_s^T g_t(s)g_r(s)u(r)drdsdt \\
&= \int_0^T \int_s^T w(t) \int_s^T g_t(s)g_r(s)u(r)drdtds \\
&= \int_0^T \left[\int_s^T w(t)g_t(s)dt\right]\left[\int_s^T g_r(s)u(r)dr\right] ds \\
&= \langle Kw, u\rangle
\end{aligned}$$

We now decompose K as the sum of two operators by changing the order of integration. An elementary calculation shows that

$$K(u) = \int_0^t \left[\int_0^r g_t(s)g_r(s)ds\right] u(r)dr + \int_t^T \left[\int_0^t g_t(s)g_r(s)ds\right] u(r)dr \tag{3.8}$$

Now define operators F and B (forward and backward) as

$$F(u) = \int_0^t \left[\int_0^r g_t(s) g_r(s) ds\right] u(r) dr \tag{3.9}$$

and

$$B(u) = \int_t^T \left[\int_0^t g_t(s) g_r(s) ds\right] u(r) dr. \tag{3.10}$$

Note that F and B are bounded and hence K is bounded. Also note that from the proof of the lemma the operator K is positive. Thus the spectrum of K is bounded below by 0 and hence the spectrum of $I + K$ is bounded away from 0 and the operator $I + K$ is thus injective.

Lemma 3.2. *The operator $\lambda I + K$ is one to one and onto.*

Proof: It only remains to prove that $I + K$ is onto. Suppose otherwise. Then there exists a function $x \in L^2[0,T]$ such that for all $y \in L^2[0,T]$ $\langle x, (I + K)(y)\rangle = 0$. We use the fact that the operator is self adjoint to conclude that $(I + K)x = 0$. This is equivalent to the fact that x is unique solution to the optimal control problem with cost function

$$J(u) = \int_0^T u^2(t) + y^2(t) dt$$

subject to the constraint of the system defined by equations (2.1) and (2.2). However, it is easy to see that the optimal control is identically 0 and hence that the corresponding $y(t)$ is identically zero. Thus we conclude that $x = 0$ and hence $\lambda I + K$ is onto.

We can thus solve equation (3.7) for to obtain

$$y = (\lambda I + K)^{-1} K(f) \tag{3.11}$$

In the next section we will explicitly construct a representation of the operator $(\lambda I + K)^{-1} K$ in terms of an associated Riccati equation.

4. The optimal control problem

We return to equation (3.4). This representation of the optimal control can be rewritten as

$$u(t) = -\int_s^T \frac{1}{\lambda} c' e^{A(t-s)} b(y(t) - f(t)) dt$$

and in this form we see that it is in dynamic feedback form and is be fed back through the system adjoint to the original system (2.1) and (2.2). Our first goal is to explicitly write out the relationship between the system and its adjoint.

We begin by letting

$$\ell(s) = \int_s^T e^{A'(t-s)} c(y(t) - f(t))dt \tag{4.1}$$

where we have replaced $c'e^{A(t-s)}b$ by its transpose. We calculate the derivative of ℓ to obtain

$$\begin{aligned}\dot{\ell}(s) =& - A'\ell(s) - c(y(s) - f(s))\\ =& - A'\ell(s) - cc'x + cf(s)\end{aligned}$$

where we have used the fact from equation (2.2) that $y = c'x$. We now see that

$$u(s) = -\frac{1}{\lambda} b'\ell(s). \tag{4.2}$$

Now from equations (2.1) and (2.2) we have

$$\begin{aligned}\dot{x}(s) =& Ax(s) + bu(s)\\ =& Ax(s) - \frac{1}{\lambda} bb'\ell(s).\end{aligned}$$

From the definition of ℓ we have

$$\ell(T) = 0$$

and we have that

$$x(0) = x_0$$

from the original definition of the system. Writing this in more conventional form we have

$$\frac{d}{dt}\begin{pmatrix} x \\ \ell \end{pmatrix} = \begin{pmatrix} A & -\frac{1}{\lambda}bb' \\ -cc' & -A' \end{pmatrix}\begin{pmatrix} x \\ \ell \end{pmatrix} + \begin{pmatrix} 0 \\ c \end{pmatrix} f \tag{4.3}$$

with boundary conditions

$$\ell(T) = 0 \text{ and } x(0) = x_0. \tag{4.4}$$

Thus from the solution of this problem we can explicitly construct the approximate spline $y(t)$.

To solve the two point boundary value problem we introduce the Riccati transform.

$$\begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} I & 0 \\ -P(t) & I \end{pmatrix} \begin{pmatrix} x \\ \ell \end{pmatrix} \tag{4.5}$$

Using this change of basis on the two point boundary value problem we have after a considerable amount of matrix multiplication

$$\frac{d}{dt}\begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} A - \frac{1}{\lambda}bb'P(t) & -\frac{1}{\lambda}bb' \\ -\dot{P} - PA - cc' + P\frac{1}{\lambda}bb'P - A'P & -(A - \frac{1}{\lambda}bb'P(t))' \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} + \begin{pmatrix} 0 \\ c \end{pmatrix} f(t)$$

We now set

$$-\dot{P} - PA - cc' + P\frac{1}{\lambda}bb'P - A^T P = 0$$

and assign it the terminal value of $P(T) = 0$. Under the conditions we have imposed of observability and controlability of the original system this nonlinear differential equation (Riccati equation) has a unique solution on the interval $[0, T]$. We thus have the following system of equations to solve

$$\dot{P} = - PA - cc' + P\frac{1}{\lambda}bb'P - A'P, \quad P(T) = 0 \tag{4.6}$$

$$\dot{w} = - (A - \frac{1}{\lambda}bb'P(t))'w + cf, \quad w(T) = 0 \tag{4.7}$$

$$\dot{x} = (A - \frac{1}{\lambda}bb'P(t))x - \frac{1}{\lambda}bb'w, \quad x(0) = x_0. \tag{4.8}$$

We begin by solving and storing the solution of the Riccati equation and substituting this into equation (4.7). We now have a linear time varying terminal value problem to solve. Let $\Phi(t, \tau)$ be the solution of

$$\frac{d}{dt}\Phi(t, \tau) = (A - \frac{1}{\lambda}bb'P(t))\Phi(t, \tau)$$

with initial data given by

$$\Phi(\tau, \tau) = I$$

and let $\Psi(t, \tau)$ be the solution of

$$\frac{d}{dt}\Psi(t, \tau) = -(A - \frac{1}{\lambda}bb'P(t))'\Psi(t, \tau)$$

with initial data given by

$$\Psi(\tau, \tau) = I.$$

So the solution of equation (4.7) is given by

$$w(t) = -\int_t^T \Psi(t,\tau)cf(\tau)d\tau \tag{4.9}$$

and the solution of equation (4.8) is given by

$$x(t) = \Phi(t,0)x_0 - \int_0^t \Phi(t,s)\frac{1}{\lambda}bb'w(s)ds. \tag{4.10}$$

Concatenating the two solutions we have

$$y(t) = c'\Phi(t,0)x_0 + c'\int_0^t \Phi(t,s)\frac{1}{\lambda}bb'\int_s^T \Psi(s,r)cf(r)drds \tag{4.11}$$

Changing the order of integration we have

$$\begin{aligned} y(t) =& c'\Phi(t,0)x_0 + \int_0^t \int_0^r c'\Phi(t,s)\frac{1}{\lambda}bb'\Psi(r,s)cdsf(r)dr \\ &+ \int_t^T \int_0^t c'\Phi(t,s)\frac{1}{\lambda}bb'\Psi(r,s)cdsf(r)dr. \end{aligned} \tag{4.12}$$

Thus we have

$$y(t) = c'\Phi(t,0)x_0 + \int_0^T K(t,\sigma)f(\sigma)d\sigma \tag{4.13}$$

where

$$K(t,\sigma) = \begin{cases} \frac{1}{\lambda}c'\int_0^\sigma \Phi(t,\tau)bb'\Psi(\tau,\sigma)cd\tau & 0 \le \sigma \le t \\ \\ \frac{1}{\lambda}c'\int_0^t \Phi(t,\tau)bb'\Psi(\tau,\sigma)cd\tau & t \le \sigma \le T \end{cases} \tag{4.14}$$

The initial data only has significant effect on the curve for a short time because of the optimality of the solution.

5. Simplification of the formulae

In this section we will give explicit expressions for the state transition matrix in terms of the system parameters and the solution of the Riccati equation. We consider the system (4.3) with $x(0) = x_0$ and $\ell(T) = 0$. By variation of parameters formula, we obtain

$$\begin{pmatrix} x(t) \\ \ell(t) \end{pmatrix} = e^{(t-T)H}\begin{pmatrix} x(T) \\ 0 \end{pmatrix} + \int_T^t e^{(t-s)H}\begin{pmatrix} 0 \\ cf(s) \end{pmatrix} ds, \tag{5.1}$$

where

$$H = \begin{pmatrix} A & -\frac{1}{\lambda}bb' \\ -cc' & -A' \end{pmatrix}.$$

This yields the relation between boundary values $x(T)$ and $\ell(0)$,

$$\begin{pmatrix} x_0 \\ \ell(0) \end{pmatrix} = e^{-TH} \begin{pmatrix} x(T) \\ 0 \end{pmatrix} + \int_T^0 e^{-sH} \begin{pmatrix} 0 \\ cf(s) \end{pmatrix} ds. \tag{5.2}$$

Now we partition the matrix e^{tH} as follows

$$e^{tH} = \begin{pmatrix} X_1(t) & X_2(t) \\ Y_1(t) & Y_2(t) \end{pmatrix}$$

where X_i, Y_i, $i = 1, 2$ are $n \times n$ matrices. Due to the semigroup properties we have the following identities, being used later

$$\begin{aligned} X_1(t-s) &= X_1(t)x_1(-s) + X_2(t)Y_1(-s) \\ 0 &= X_1(t)X_2(-t) + X_2(t)Y_2(-t) \end{aligned} \tag{5.3}$$

The unique positive definite solution to the Riccati equation is given by

$$P(t) = Y_1(t-T)X_1(t-T)^{-1}. \tag{5.4}$$

using the standard Hamitonian argument.

Then, solving the first block of equations in (5.2) yields

$$x(T) = X_1(-T)^{-1} \left(x_0 + \int_0^T X_2(-s)cf(s)ds \right) \tag{5.5}$$

which, together with (5.1), leads to the following expression

(5.6)

$$x(t) = X_1(t-T)X_1(-T)^{-1} \left(x_0 + \int_0^T X_2(-s)cf(s)ds \right) + \int_t^T X_2(t-s)cf(s)ds.$$

By the identities from (5.3) and (5.4) we have

$$
\begin{aligned}
x(t) =&(X_1(t)+X_2(t)P(0))x_0 + \int_0^T (X_1(t+X_2(t)P(0))X_2(-s)cf(s)ds \\
&- \int_t^T (X_1(s)X_2(-s)+X_2(t)Y_2(-s)cf(s)ds \\
=&(X_1(t)+X_2(t)P(0))x_0 + \int_0^t (X_1(t+X_2(t)P(0))X_2(-s)cf(s)ds \\
&+ \int_t^T \left((X_1(t)+X_2(t)P(0))X_2(-s) - (X_1(s)X_2(-s)+X_2(t)Y_2(-s))\right) cf(s)ds \\
=&(X_1(t)+X_2(t)P(0))x_0 + \int_0^t (X_1(t+X_2(t)P(0))X_2(-s)cf(s)ds \\
&+ \int_t^T (X_2(t)(P(0)X_2(-s)-Y_2(-s))cf(s)ds.
\end{aligned}
$$

Finally, we get the kernel

$$
K(t,\sigma) = \begin{cases} c'(X_1(t)+X_2(t)P(0))X_2(-\sigma)c, & 0 \le \sigma \le t, \\ c'X_2(t)(P(0)X_2(-\sigma)-Y_2(-\sigma))c, & t \le \sigma \le T. \end{cases}
$$

It is the same as the following explicit formula in terms of the system parameters and the Riccati solution

$$
K(t,\sigma) = \begin{cases} c' \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I & 0 \\ P(0) & 0 \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix} c, & 0 \le \sigma \le t, \\ c' \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} 0 & 0 \\ P(0) & -I \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix} c, & t \le \sigma \le T. \end{cases} \tag{5.7}
$$

Furthermore, we have, by a simple observation, that

$$
\int_0^\sigma \Phi(t,\tau)\frac{1}{\lambda}bb'\Phi(\sigma,\tau)'d\tau = \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I & 0 \\ P(0) & 0 \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix}, \text{ for } 0 \le \sigma \le t
$$

$$
\int_0^t \Phi(t,\tau)\frac{1}{\lambda}bb'\Phi(\sigma,\tau)'d\tau = \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} 0 & 0 \\ P(0) & -I \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix}, \text{ for } t \le \sigma \le T
$$

and

$$
\Phi(t,0) = \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I \\ P(0) \end{pmatrix}.
$$

Therefore, the transition matrix $\Phi(t,s)$ is

$$\Phi(t,s) = \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I \\ P(0) \end{pmatrix} \left(\begin{pmatrix} I & 0 \end{pmatrix} e^{sH} \begin{pmatrix} I \\ P(0) \end{pmatrix} \right)^{-1}$$

6. Examples and numericla results

In this section we consider themost imprtant of the splines, the cubic spline and construct the explicit linear filter. We begin by deriving the associated Riccati equation.

6.1. **Derivation of Riccati equation of cubic smoothing splines.** Let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad c = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Let

$$J(u) = \int_0^T y(t)^2 + \lambda^{-1} u(t)^2 dt$$

where λ is a positive constant. We now solve the following problem:

$$\min_{u(t)} J(u)$$

subject to the constraints that

$$\dot{x} = Ax + bu, \quad y = cx.$$

The Hamiltonian matrix associated with this problem is

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda} \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

Our immediate task is to calculate

$$e^{Ht}$$

and to do so we need the following

$$H^2 = \begin{pmatrix} 0 & 0 & 0 & -\frac{1}{\lambda} \\ 0 & 0 & \frac{1}{\lambda} & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad H^3 = \begin{pmatrix} 0 & 0 & \frac{1}{\lambda} & 0 \\ -\frac{1}{\lambda} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\lambda} \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad H^4 = \begin{pmatrix} -\frac{1}{\lambda} & 0 & 0 & 0 \\ 0 & -\frac{1}{\lambda} & 0 & 0 \\ 0 & 0 & -\frac{1}{\lambda} & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda} \end{pmatrix}$$

and from this calculation we see that the eigenvalues of H are the forth roots of $-\frac{1}{\lambda}$,

$$(\frac{1}{\lambda})^{\frac{1}{4}}(\pm 2^{\frac{-1}{2}} \pm i2^{\frac{-1}{2}})$$

We now give a different form of $\exp Ht$.

$$\begin{aligned}
e^{Ht} =& \sum_{n=0}^{\infty} \frac{H^n t^n}{n!} \\
=& \sum_{n=0}^{\infty} \frac{H^{4n} t^{4n}}{(4n)!} + \sum_{n=0}^{\infty} \frac{H^{4n+1} t^{4n+1}}{(4n+1)!} + \sum_{n=0}^{\infty} \frac{H^{4n+2} t^{4n+2}}{(4n)!} + \sum_{n=0}^{\infty} \frac{H^{4n+3} t^{4n+3}}{(4n+3)!} \\
=& I \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n}}{(4n)!} + H \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n+1}}{(4n+1)!} + H^2 \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n+2}}{(4n+2)!} + H^3 \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n+3}}{(4n+3)!} \\
=& f_0(t)I + f_1(t)H + f_2(t)H^2 + f_3(t)H^3
\end{aligned}$$

We note the following

$$He^H t = f_0 H + f_1 H^2 + f_2 H^3 - \frac{1}{\lambda} f_3 I = f_0' I + f_1' H + f_2' H^2 + f_3' H^3$$

and hence by linear independence of the powers of H we have the following differential relations.

$$f_0' = -\frac{1}{\lambda} f_3, \quad f_0^{(2)} = -\frac{1}{\lambda} f_2, \quad f_0^{(3)} = -\frac{1}{\lambda} f_1$$

sot that it suffices to find a closed form relation for f_0.

Lemma 6.1. $f_0(t) = \cosh(\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}} t) \cos(\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}} t)$

Proof: Recall that

$$\cosh t = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!}$$

and

$$\cos t = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!}$$

and so

$$\cosh t + \cos t = 2 \sum_{n=0}^{\infty} \frac{t^{4n}}{(4n)!}.$$

Using the fact that $(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})^4 = -1$ we have that

$$\cosh(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})t + \cos(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})t = 2 \sum_{n=0}^{\infty} \frac{(-1)^n (\frac{1}{\lambda})^n t^{4n}}{(4n)!}$$

Now replacing cosh and cos in the above expression by their exponential representations we have after a little manipulation

$$\cosh(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}}+i\frac{1}{\sqrt{2}})t+\cos(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}}+i\frac{1}{\sqrt{2}})t = 2\cosh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\cos\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t$$

thus we have proved the lemma.

Using the differential relations we have after some tedious calculations

$$\begin{aligned}
f_0(t) &= \cosh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\cos\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\\
f_1(t) &= \frac{\lambda^{1/4}}{\sqrt{2}}(\cos\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\sinh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t+\sin\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\cosh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t)\\
f_2(t) &= \lambda^{1/2}\sin\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\sinh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\\
f_3(t) &= \frac{-\lambda^{3/4}}{\sqrt{2}}(\sinh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\cos\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t-\cosh\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\sin\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t).
\end{aligned}$$

Thus we have a closed form representation of $\exp Ht$.

$$e^{Ht}=\begin{pmatrix} f_0 & f_1 & \frac{1}{\lambda}f_3 & -\frac{1}{\lambda}f_2\\ -\frac{1}{\lambda}f_3 & f_0 & \frac{1}{\lambda}f_2 & -\frac{1}{\lambda}f_1\\ -f_1 & -f_2 & f_0 & \frac{1}{\lambda}f_3\\ f_2 & f_3 & -f_1 & f_0\end{pmatrix}$$

Now let

$$F(t)=\begin{pmatrix} -f_1 & -f_2\\ f_2 & f_3\end{pmatrix}\begin{pmatrix} f_o & f_1\\ -\frac{1}{\lambda}f_3 & f_0\end{pmatrix}^{-1}$$

To show that

$$F(t)$$

is a solution to the Riccati equation take the derivative of $F(t)$ using the fact

$$\frac{d}{dt}\exp Ht = H\exp Ht$$

and then note that $F(0)=0$ The particular solution we want is then given by

$$P(t)=F(t-T).$$

We now calculate $F(t)$ explicitly. We make the convention that $\mathrm{S}=\sinh, \mathrm{s}=\sin, \mathrm{C}=\cosh, \mathrm{c}=\cos$ otherwise the following calculations are difficult to typeset.

First

$$F(t) = \frac{1}{f_0^2 + \frac{1}{\lambda} f_1 f_3} \begin{pmatrix} -f_1 f_0 - \frac{1}{\lambda} f_2 f_3 & f_1^2 - f_2 f_0 \\ f_2 f_0 + \frac{1}{\lambda} f_3^2 & -f_1 f_2 + f_3 f_0 \end{pmatrix}$$

after some calculations we have

$$f_1^2 - f_2 f_0 = \frac{\lambda^{1/2}}{\sqrt{2}} (\mathrm{c}^2\mathrm{S}^2 + \mathrm{s}^2\mathrm{C}^2)$$

$$f_1^2 - f_2 f_0 = f_2 f_0 + \frac{1}{\lambda} f_3^3$$

$$-f_1 f_0 - \frac{1}{\lambda} f_2 f_3 = \frac{\lambda^{1/4}}{\sqrt{2}} (-\mathrm{SC} - \mathrm{sc})$$

$$f_3 f_0 - f_1 f_2 = \frac{\lambda^{3/4}}{\sqrt{2}} (-\mathrm{SC} + \mathrm{cs})$$

$$f_0^2 + \frac{1}{\lambda} f_1 f_3 = \frac{1}{2} (\mathrm{C}^2 + \mathrm{c}^2)$$

We have now found an explicit form for $F(t)$

$$F(t) = \begin{pmatrix} -\sqrt{2}\lambda^{1/4} \frac{\mathrm{SC}-\mathrm{cs}}{\mathrm{C}^2+\mathrm{c}^2} & \lambda^{1/2} \frac{\mathrm{S}^2+\mathrm{s}^2}{\mathrm{C}^2+\mathrm{c}^2} \\ \lambda^{1/2} \frac{\mathrm{S}^2+\mathrm{s}^2}{\mathrm{C}^2+\mathrm{c}^2} & \sqrt{2}\lambda^{3/4} \frac{\mathrm{cs}-\mathrm{SC}}{\mathrm{C}^2+\mathrm{c}^2} \end{pmatrix}.$$

Thus the explicit solution of the Riccati equation is

$$P(t) = F(t - T).$$

6.2. **Derivation of Riccati equation of polynomial smoothing splines.** For completeness we derive the solution of the plynomial splines through the matrix exponential of the Hamiltonian matrix. We shell give an explicit form of e^{Ht} for general positive integer n, where

$$H = \begin{pmatrix} A & -\alpha b b' \\ -c c' & -A' \end{pmatrix},$$

with A a forward shift matrix and b, c the unit vector with the last respectively the first entry 1. Unlike the derivation of e^{Ht} in cubic smoothing splines where the entries of this matrix were obtained by solving the differential equations, we use a direct approach by the definition of the matrix exponential.

By a straightforward calculation we obain the following properties.

Proposition 6.2. (1) *The characteristic polynomial of the Hamiltonian matrix H is*

$$\det(zI - H) = z^{2n} + (-1)^n \alpha$$

(2) $H^{2nk+i} = ((-1)^{n-1}\alpha)^k H^i$, $i = 0, 1, 2, ..., 2n-1, k = 0, 1, 2, ...$

Using the second property we obtain the matrix exponential e^{Ht} in the form of series, and by inspection we obtain

Proposition 6.3. *The matrix exponential of H is*

$$e^{Ht} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

where

$$B_{11} = \begin{pmatrix} f_0 & -\frac{1}{\alpha} f_0^{(2n-1)} & -\frac{1}{\alpha} f_0^{(2n-2)} & \cdots & -\frac{1}{\alpha} f_0^{(n+1)} \\ f_0' & f_0 & -\frac{1}{\alpha} f_0^{(2n-1)} & \cdots & -\frac{1}{\alpha} f_0^{(n+2)} \\ f_0'' & f_0' & f_0 & \cdots & -\frac{1}{\alpha} f_0^{(n+3)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_0^{(n-2)} & f_0^{(n-3)} & f_0^{(n-4)} & \cdots & -\frac{1}{\alpha} f_0^{(2n-1)} \\ f_0^{(n-1)} & f_0^{(n-2)} & f_0^{(n-3)} & \cdots & f_0 \end{pmatrix}$$

$$B_{12} = \begin{pmatrix} -f_0' & f_0'' & \cdots & -f_0^{(n-1)} & f_0^{(n)} \\ -f_0'' & f_0''' & \cdots & -f_0^{(n)} & f_0^{(n+1)} \\ -f_0''' & f_0^{(4)} & \cdots & f_0^{(n+1)} & f_0^{(n+2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -f_0^{(n-1)} & f_0^{(n)} & \cdots & -f_0^{(2n-3)} & f_0^{(2n-2)} \\ -f_0^{(n)} & f_0^{(n+1)} & \cdots & -f_0^{(2n-2)} & f_0^{(2n-1)} \end{pmatrix}$$

$$B_{21} = \begin{pmatrix} \frac{1}{\alpha} f_0^{(2n-1)} & \frac{1}{\alpha} f_0^{(2n-2)} & \frac{1}{\alpha} f_0^{(2n-3)} & \cdots & \frac{1}{\alpha} f_0^{(n)} \\ -\frac{1}{\alpha} f_0^{(2n-2)} & -\frac{1}{\alpha} f_0^{(2n-3)} & -\frac{1}{\alpha} f_0^{(2n-4)} & \cdots & -\frac{1}{\alpha} f_0^{(n-1)} \\ \frac{1}{\alpha} f_0^{(2n-3)} & \frac{1}{\alpha} f_0^{(2n-4)} & \frac{1}{\alpha} f_0^{(2n-5)} & \cdots & \frac{1}{\alpha} f_0^{(n-2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{\alpha} f_0^{(n)} & -\frac{1}{\alpha} f_0^{(n-1)} & -\frac{1}{\alpha} f_0^{(n-2)} & \cdots & -\frac{1}{\alpha} f_0' \end{pmatrix}$$

$$B_{22} = \begin{pmatrix} f_0 & -f_0' & \cdots & f_0^{(n-2)} & -f_0^{(n-1)} \\ \frac{1}{\alpha} f_0^{(2n-1)} & f_0 & \cdots & -f_0^{(n-1)} & f_0^{(n-2)} \\ -\frac{1}{\alpha} f_0^{(2n-2)} & \frac{1}{\alpha} f_0^{(2n-2)} & \cdots & f_0^{(n-2)} & -f_0^{(n-3)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\alpha} f_0^{(n+1)} & -\frac{1}{\alpha} f_0^{(n+2)} & \cdots & -\frac{1}{\alpha} f_0^{(2n-1)} & f_0 \end{pmatrix}$$

if n is even, and $f_0(t) = 1 - \frac{\alpha t^{2n}}{(2n)!} + \frac{\alpha^2 t^{4n}}{(4n)!} - \cdots$; *or*

$$B_{11} = \begin{pmatrix} f_0 & \frac{1}{\alpha} f_0^{(2n-1)} & \cdots & \frac{1}{\alpha} f_0^{(n+1)} \\ f_0' & f_0 & \cdots & \frac{1}{\alpha} f_0^{(n+2)} \\ \cdots & \cdots & \cdots & \cdots \\ f_0^{(n-1)} & f_0^{(n-2)} & \cdots & f_0 \end{pmatrix}$$

$$B_{12} = \begin{pmatrix} -f_0' & f_0'' & \cdots & f_0^{(n-1)} & -f_0^{(n)} \\ -f_0'' & f_0''' & \cdots & f_0^{(n)} & -f_0^{(n+1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -f_0^{(n)} & f_0^{(n+1)} & \cdots & f_0^{(2n-2)} & -f_0^{(2n-1)} \end{pmatrix}$$

$$B_{21} = \begin{pmatrix} -\frac{1}{\alpha} f_0^{(2n-1)} & -\frac{1}{\alpha} f_0^{(2n-2)} & \cdots & -\frac{1}{\alpha} f_0^{(n)} \\ \frac{1}{\alpha} f_0^{(2n-2)} & \frac{1}{\alpha} & \cdots & \frac{1}{\alpha} f_0^{(n-1)} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{\alpha} f_0^{(n)} & -\frac{1}{\alpha} f_0^{(n-1)} & \cdots & -\frac{1}{\alpha} f_0' \end{pmatrix}$$

$$B_{22} = \begin{pmatrix} f_0 & -f_0' & \cdots & -f_0^{(n-2)} & f_0^{(n-1)} \\ -\frac{1}{\alpha} f_0^{(2n-1)} & f_0 & \cdots & f_0^{(n-2)} & -f_0^{(n-2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\alpha} f_0^{(n+1)} & -\frac{1}{\alpha} f_0^{(n+2)} & \cdots & -\frac{1}{\alpha} f_0^{(2n-1)} & f_0 \end{pmatrix}$$

if n is odd, and $f_0(t) = 1 + \frac{\alpha t^{2n}}{(2n)!} + \frac{\alpha^2 t^{4n}}{(4n)!} + \cdots$.

Now we shall find the function $f_0(t)$ that has power series shown in the above proposition. Using the property of the unity roots, it is not hard to see that

$$f_0(t) = \begin{cases} \frac{1}{2n} \sum_{k=0}^{n-1} \Big(\exp\,(-\alpha^{\frac{1}{2n}} (-1)^{\frac{2k+1}{2n}} t) + \exp\,(\alpha^{\frac{1}{2n}} (-1)^{\frac{2k+1}{2n}} t) \Big) & \text{if } n = \text{even} \\ \frac{1}{2n} \sum_{k=1}^{n} \Big(\exp\,(-\alpha^{\frac{1}{2n}} (-1)^{\frac{k}{n}} t) + \exp\,(\alpha^{\frac{1}{2n}} (-1)^{\frac{k}{n}} t) \Big) & \text{if } n = \text{odd} \end{cases}$$

Consequently we obtain the real form of the function f_0

Proposition 6.4.

$$f_0(t) = \begin{cases} \frac{2}{n} \sum_{k=1}^{n-1} \cosh(\alpha^{\frac{1}{2n}} (\cos(\frac{k\pi}{2n}) t) \cos(\alpha^{\frac{1}{2n}} \sin(\frac{k\pi}{2n}) t) & \textit{if } n = \textit{even} \\ \frac{1}{n} \cosh(\alpha^{\frac{1}{2n}} t) + \frac{2}{n} \sum_{k=1}^{n-1} \cosh(\alpha^{\frac{1}{2n}} (\cos(\frac{k\pi}{n}) t) \cos(\alpha^{\frac{1}{2n}} \sin(\frac{k\pi}{n}) t) & \textit{if } n = \textit{odd} \end{cases}$$

Now the solution of $F(t)$ can be obtained by taking the first n columns of the last n rows multiplied by the inverse of the first n rows and the first n columns. The solution of the Riccati equation is then $P(t) = F(t-T)$. For the positive definite steady state solution we can either take the limit of $P(t)$ as t tends to minus infinity, or using the eigenvectors corresponding to the eigenvalues with negative real parts. The last mentioned fact can be used to get an approximated kernel.

6.3. **Approximation and bandwidth.** Before showing numerical results we discuss two things. First we show that the kernel in terms of the solution of the Riccati equation can be approximated by its steady state, and second the bandwidth of the kernel. The first issue is based on the well known results from control theory that the solution of Riccati equation tends to its steady state, the positive definite solution of algebraic Riccati equation, is exponential, if the eigenvalues of H are a bit away from the imaginary axis. The steady state Riccati solution in cubic splines is

$$P = \begin{pmatrix} \sqrt{2}\lambda^{1/4} & \sqrt{\lambda} \\ \sqrt{\lambda} & \sqrt{2}\lambda^{3/4} \end{pmatrix}.$$

It is positive definite. Then the transition matrix is

$$\Phi(t,s) = e^{(t-s)(A-\frac{1}{\lambda}bb'P)}$$

since $A-\frac{1}{\lambda}bb'P$ is a constant matrix. It is known that the feedback matrix $A-\frac{1}{\lambda}bb'P$ is Hurwitz, i.e. the eigenvalues of the matrix lie in the left-half complex plane and they are equal to the left-half plane eigenvalues of the matrix H. Consequently the kernel is,

(6.1)

$$\hat{K}(t,\sigma) = \begin{cases} \dfrac{\exp(-\frac{\sigma+t}{\sqrt{2}\lambda^{1/4}})}{2\sqrt{2}\lambda^{1/4}} \Bigg[\left(-2 + \exp(\frac{\sqrt{2}\sigma}{\lambda^{1/4}}) \right) \cos \dfrac{\sigma - t}{\sqrt{2}\lambda^{1/4}} + \cos \dfrac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \\ \quad - \exp(\frac{\sqrt{2}\sigma}{\lambda^{1/4}}) \sin \dfrac{\sigma-t}{\sqrt{2}\lambda^{1/4}} - \sin \dfrac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \Bigg], \quad 0 \le \sigma \le t, \\ \dfrac{\exp(-\frac{\sigma+t}{\sqrt{2}\lambda^{1/4}})}{2\sqrt{2}\lambda^{1/4}} \Bigg[\left(-2 + \exp(\frac{\sqrt{2}t}{\lambda^{1/4}}) \right) \cos \frac{\sigma - t}{\sqrt{2}\lambda^{1/4}} + \cos \frac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \\ \quad + \exp(\frac{\sqrt{2}t}{\lambda^{1/4}}) \sin \dfrac{\sigma-t}{\sqrt{2}\lambda^{1/4}} - \sin \dfrac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \Bigg], \quad t \le \sigma \le T. \end{cases}$$

Now we turn to a short discussion of the bandwidth of the kernel. The term bandwidth is frequently used in statistic literautres. The definition we use here is the interval where one obtains most information. More precisely, Let the number β be determined by the solution of the following equation

$$\frac{\displaystyle\int_{\max(t-\beta,0)}^{\min(t+\beta,T)} K(t,\sigma)d\sigma}{\displaystyle\int_0^T K(t,\sigma)d\sigma} = 0.9. \tag{6.2}$$

Then 2β is the bandwidth of the kernel $K(t,\sigma)$.

This is a nonlinear equation which can be numerically solved by e.g. Newton-Raphson method. After integration of $\hat{K}$, we get the denominator respectively the numerator of the left hand side in (6.2)

$$\frac{1}{2}e^{-\frac{t+T}{\sqrt{2}\lambda^{1/4}}}\left(-2e^T\sqrt{2}\lambda^{1/4}\cos\frac{t}{\sqrt{2}\lambda^{1/4}} - \left(-1+e^{\frac{T}{\sqrt{2}\lambda^{1/4}}}\right)\cos\frac{t-T}{\sqrt{2}\lambda^{1/4}}\right.$$
$$\left.+2\left(e^{\frac{t+T}{\sqrt{2}\lambda^{1/4}}} + \left(-e^{\frac{T}{\sqrt{2}\lambda^{1/4}}} + \cos\frac{T}{\sqrt{2}\lambda^{1/4}}\right)\sin\frac{t}{\sqrt{2}\lambda^{1/4}}\right)\right);$$

$$\frac{1}{2}e^{-\frac{3t}{\sqrt{2}\lambda^{1/4}}}\left(2e^{\frac{3t}{\sqrt{2}\lambda^{1/4}}} - e^{\frac{-\beta+t}{\sqrt{2}\lambda^{1/4}}}\left(-1+e^{\frac{\sqrt{2}\beta}{\lambda^{1/4}}} + 2e^{\frac{\sqrt{2}t}{\lambda^{1/4}}}\right)\cos\frac{\beta}{\sqrt{2}\lambda^{1/4}}\right.$$
$$\left.-(e^{\frac{-\beta+t}{\sqrt{2}\lambda^{1/4}}} + e^{\frac{\beta+t}{\sqrt{2}\lambda^{1/4}}})\sin\frac{\beta}{\sqrt{2}\lambda^{1/4}} + e^{\frac{\beta+t}{\sqrt{2}\lambda^{1/4}}}\sin\frac{\beta-2t}{\sqrt{2}\lambda^{1/4}} + e^{\frac{-\beta+t}{\sqrt{2}\lambda^{1/4}}}\sin\frac{\beta+2t}{\sqrt{2}\lambda^{1/4}}\right).$$

Solving the equation in some special cases, we obtain $\beta_{T,\lambda}$

$\beta_{10,0.01} = 0.547334$, $\beta_{10,0.1} = 0.968005$, $\beta_{10,10} = 3.22425$, $\beta_{10,100} = 4.03161$,
$\beta_{8,0.01} = 0.547444$, $\beta_{8,0.1} = 0.979628$, $\beta_{8,10} = 2.8789$, $\beta_{8,100} = 3.3587$,
$\beta_{4,0.001} = 0.307604$, $\beta_{4,0.01} = 0.564674$, $\beta_{4,0.1} = 1.10534$, $\beta_{4,10} = 1.70457$,
$\beta_{30,10} = 3.71504$.

6.4. **Numerical simulations of the kernel.** Figure 1 to Figure 3 (graphs with dashed line is the kernel computed from the approximation) show clearly that the approximation is remarkably good, especially when λ is small, the approximation has the same curvature change as the kernel. We have observed that there is some small ocsilations at the end of the time interval when λ is very large, and the larger λ is the worse the oscilations are. This is in fact caused by the fact that the eigenvalues of $H$is very close to the imaginary axis. Then the convergence of the Riccati solution is no longer exponential. Note also that the kernel is well-behaved for large T, as shown in Figure 4. However, it may not be realistic to make T too large, because it indicates that we have to use many points in practice.

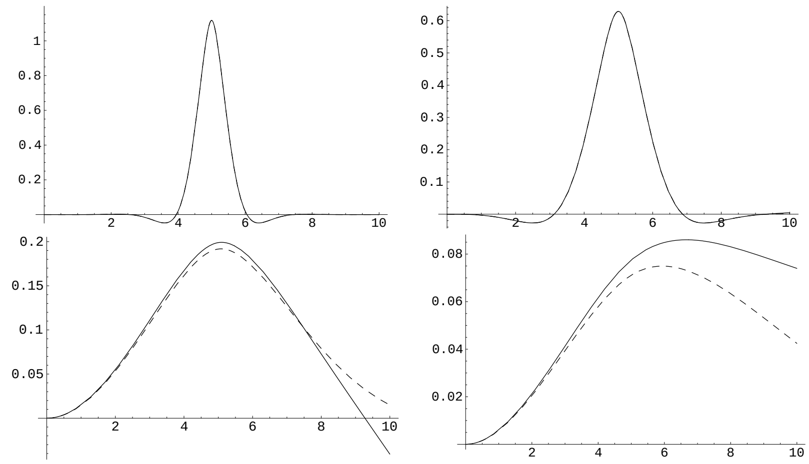

FIGURE 1. $T = 10$, $\lambda = 0.01, 0.1, 10, 100$

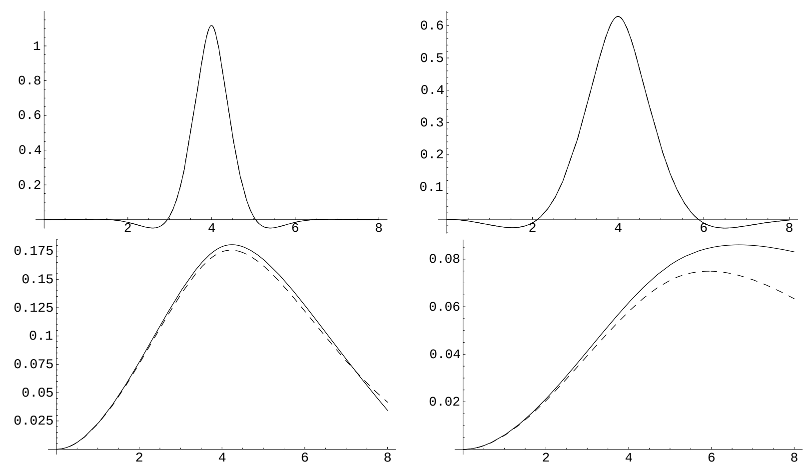

FIGURE 2. $T = 8$, $\lambda = 0.01, 0.1, 10, 100$

Now we compare some of our results to the some of Silverman's [9]. We note that our approximated kernel is close to the kernel (2.10) in [9], but we do not use the density function explicitly, because we do not consider interpolating points as random ones. Perhaps this is one of the main differences between our approach and the one by Silverman. Observe also that the angle $\frac{\pi}{4}$ is implicit in our formula. It is involved in the eigenvalues of H or $A - \frac{1}{\lambda}bb'P$. The behavior of our approximation

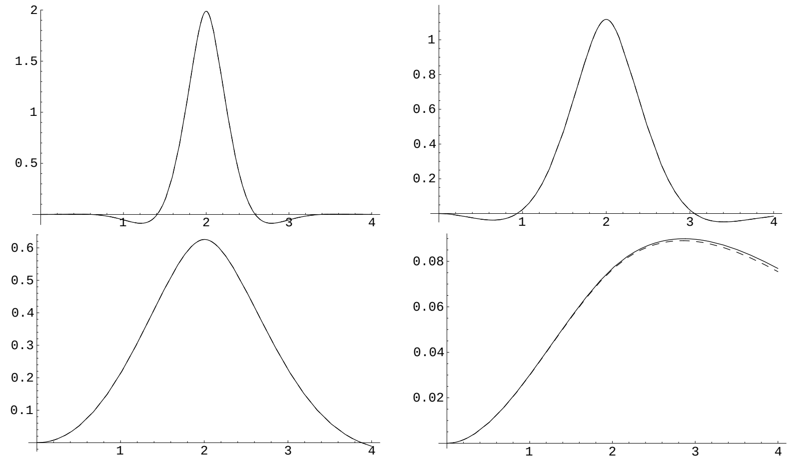

FIGURE 3. $T = 4$, $\lambda = 0.001, 0.01, 0.1, 10$

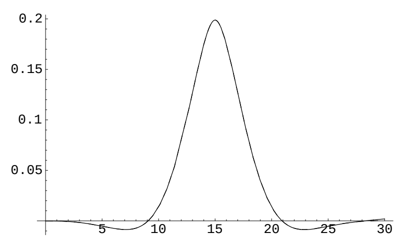

FIGURE 4. $T = 30$, $\lambda = 10$

away from the boundary are also similar to the results of Silverman, and the kernel (6.1) is symmetric. The asymmetric graphs shown in this paper can be in principle overcome by increasing the T. Moreover, the bandwidth according to our definition is not local. Moreover, we want to point out that the choice of $\lambda = 10^{-7}$ seems unncessarily small in many situations, see the graph below. This may indicate that the points one wants to fit are densely accumulated over a very small interval with relatively vertical scale.

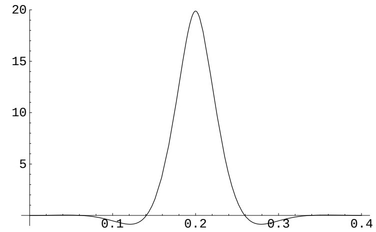

FIGURE 5. $T = 0.4$, $\lambda = 10^{-7}$

6.5. **Torsion smoothing splines.** Finally, we consider another important case where $A = \begin{pmatrix} 0 & 1 \\ 0 & a \end{pmatrix}$, which corresponds to the torsion smoothing splines. So

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & a & 0 & -\frac{1}{\lambda} \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -a \end{pmatrix}$$

A straightforward computation shows that H has

(1) four distinct real eigenvalues, $\pm\sqrt{\frac{a^2}{2} \pm \frac{1}{2}\sqrt{a^4 - \frac{4}{\lambda}}}$, if $a^4 - \frac{4}{\lambda} > 0$,
(2) two double real eigenvalues, each with one Jordan block, $\pm\frac{|a|}{\sqrt{2}}$, if $a^4 - \frac{4}{\lambda} = 0$,
(3) two distinct pairs of complex conjugate eigenvalues, $\pm\lambda^{-1/4}e^{\pm\frac{i}{2}\theta}$, $\theta = \sqrt{\frac{4}{\lambda a^4} - 1}$, if $a^4 - \frac{4}{\lambda} < 0$.

Notice that the root loci of the eigenvalues of $H$are as follows. When a is zero (corresponding to the cubic splines), there are four complex eigenvalues lying on the circle with radius $1/\lambda^{1/4}$ centered at 0, and are placed symmetrically with respect to both real and imaginary axes with an angle to the axes $\theta = \frac{\pi}{4}$. Then these four roots pairwise move towards the real axis along the circle and collapse at the two the real points $\pm 1/\lambda^{1/4}$ as $|a|$ increases to $(4/\lambda)^{1/4}$, i.e. θ tends to 0, then the eigenvalues of H become two double ones with only one Jordan block. Finally as $|a|$continues increasing the double eigenvalues of H bifurcate to four distinct real ones, each of the double eigenvalues bifurcates in opposite directions, one goes toward the imaginary axis and the other to infinity.

Although it is a direct inspection, it is an important fact. It explains why the kernel behaves as shown below.

The approximation is very good as in the case of cubic splines.

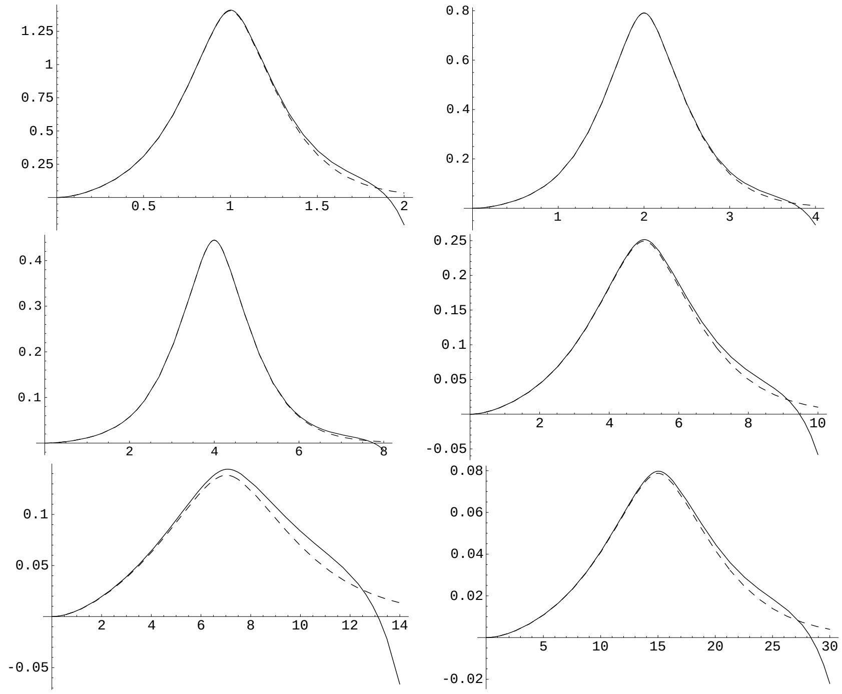

FIGURE 6. The graphs correspond to the cases (i) $T = 2$, $\lambda = 0.001$, (ii) $T = 4$, $\lambda = 0.01$, (iii) $T = 8$, $\lambda = 0.1$, (iv) $T = 10$, $\lambda = 1$, (v) $T = 14$, $\lambda = 10$, (iv) $T = 30$, $\lambda = 100$.

Now we turn to other two non-degenerated cases. Next figure is the case of four real eigenvalues for fixed $a = 1 + 400^{1/4}$, and we vary λ.

Note that the kernel cannot behave well as $|a|$ becomes large, because this forces the two eigenvalues on each half plane to go to zero and the infinity (on each side).

Finally, we show two groups of graphs of the kernel in case there are two pairs of complex conjugate eigenvalues for fixed a. We want to show that the remark we made above applies to the situation as θ tends to zero.

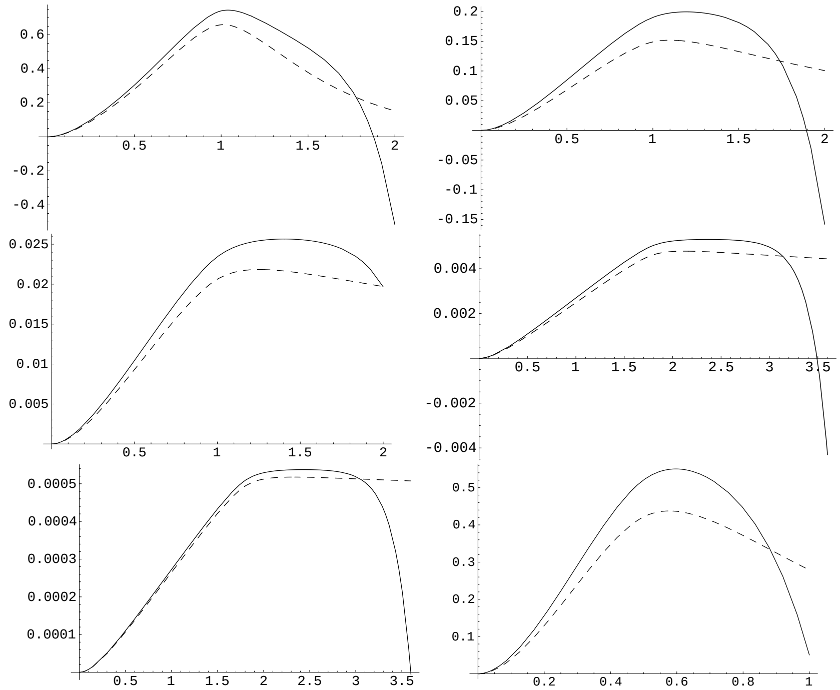

FIGURE 7. $\lambda = 0.01, 0.1, 1, 10, 100$. The first and the last graphs correspond idetical λ and a but different length of T.

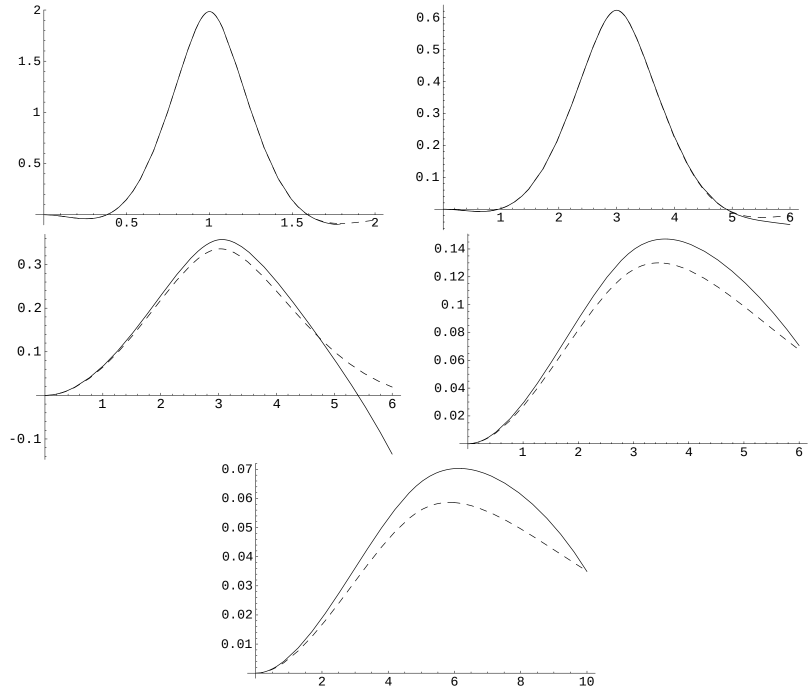

FIGURE 8. $a = 1/3$, $\lambda = 0.01, 0.1, 1, 10, 100$.

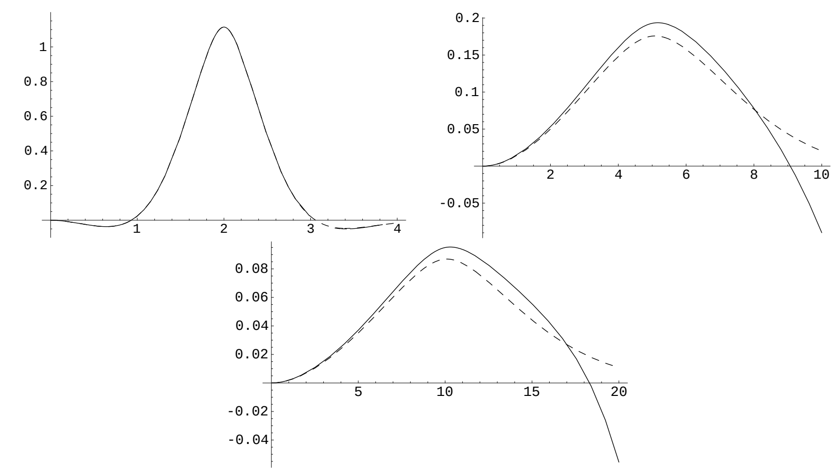

FIGURE 9. $a = 1/3$, $\lambda = 0.01, 10, 100$ with ncreased length T.

References

[1] Crouch, P.; Jackson, J. Dynamic Interpolation and Application to Flight Control, Journal of Guidance, Control and Dynamics, Vol. 14, pp. 814 - 822,1991.

[2] Doolin, Brian F.; Martin, Clyde F. Introduction to differential geometry for engineers. Monographs and Textbooks in Pure and Applied Mathematics, 136. Marcel Dekker, Inc., New York, 1990.

[3] Egerstedt, M; Martin, C. Optimal trajecory planning and smoothing splines. Automatica 37 2001 1057–1064

[4] Kalman, R. E. Contributions to the theory of optimal control. Bol. Soc. Mat. Mexicana (2) 5 1960 102–119.

[5] Martin, C.; Egerstedt, M. Statistical Estimates for Generalized Splines, to appear in COVC.

[6] Martin, C. F.; Sun, S.; Egerstedt, M. Optimal control, statistics and path planning. Computation and control, VI (Bozeman, MT, 1998). Math. Comput. Modelling 33 (2001), no. 1-3, 237–253.

[7] Schultz, M. H.; Varga, R.S. L-splines. Numer. Math. 10 1967 345–369

[8] Silva-Liete, F.; Rodrigues, R. L-splines - a manifestation of optimal control. IMA Journal of Mathematical Control and Information 19 (2002) 313–324.

[9] Silverman, B. W. Spline smoothing: the equivalent variable kernel method. Ann. Statist. 12 (1984), no. 3, 898–916

[10] Silverman, B. W. Some aspects of the spline smoothing approach to non-parametric regression curve fitting J. Roy. Stat. Soc. 47 (1985), no. 1, 1–21.

[11] Sun, S.; Egerstedt, M. B.; Martin, C. F. Control theoretic smoothing splines. IEEE Trans. Automat. Control 45 (2000), no. 12, 2271–2279.

[12] Takahashi, S.; Ghosh, B.; Martin, C. Control Theoretic smoothing splines and its applications, Transactions of the Society of Instrument and Control Engineers, vol. 38, no. 3, pp. 293-298, 2002.

[13] Wahba, G. Spline models for observational data. CBMS-NSF Regional Conference Series in Applied Mathematics, 59. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[14] Zhou Y.; Egerstedt, M; C. Martin, C, Optimal approximation of functions, Com. Inf. Syst. 1 (2001) 101–112.

Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409

Department of Mathematics, Stockholm University, SE-10693, Stockholm

E-mail address: `yishao@math.su.se`

E-mail address: `daya@math.ttu.edu`

E-mail address: `martin@math.ttu.edu`