Kernel Regression for Vehicle Trajectory Reconstruction under Speed and Inter-vehicular Distance Constraints ⋆

Pierre-Cyril Aubin-Frankowski * **Nicolas Petit** ** **Zoltán Szabó** ***

* *École des Ponts ParisTech and CAS, MINES ParisTech, PSL, France, (email: pierre-cyril.aubin@mines-paristech.fr).*
** *CAS, MINES ParisTech, PSL, France, (email: nicolas.petit@mines-paristech.fr).*
*** *Center of Applied Mathematics, CNRS, École Polytechnique, Institut Polytechnique de Paris, France, (email: zoltan.szabo@polytechnique.edu).*

Abstract
This work tackles the problem of reconstructing vehicle trajectories with the side information of physical constraints, such as inter-vehicular distance and speed limits. It is notoriously difficult to perform a regression while enforcing these hard constraints on time intervals. Using reproducing kernel Hilbert spaces, we propose a convex reformulation which can be directly implemented in classical solvers such as CVXGEN. Numerical experiments on a simple dataset illustrate the efficiency of the method, especially with sparse and noisy data.

Keywords: Non-parametric regression, constraints, Hilbert spaces, convex programming, machine learning, road traffic

1. INTRODUCTION

This article addresses the problem of shape-constrained regression. This problem can be found in many domains of engineering, under various terminologies such as: estimation under constraints, constrained curve fitting or constrained smoothing to name a few. Taking into account the constraints is well known to improve reconstruction performance (Alouani and Blair, 1991; Chang et al., 2009). Prime examples of applications can be found in chemical engineering (Arora and Biegler, 2004), biology (Motulsky and Christopoulos, 2004), among others. Shape-constrained regression is of interest in transportation systems, especially in the context of the vehicle to infrastructure (V2I; Agosta et al. 2016) concept which is central in the research field of Intelligent Transportation Systems (ITS; Work et al. 2009; De Wit et al. 2015; Liang et al. 2016). To illustrate this point, the application problem considered in this article consists in reconstructing the trajectory of several vehicles under constraints, based on a set of noisy position measurements. The vehicles are represented as point-like objects traveling along a one-dimensional path. The constraints are interpreted as some side information, i.e. external or prior knowledge. In our context, the constraints model the fact that the vehicles are non-overtaking and have non-negative curvilinear velocity.

Due to its practical importance, several approaches have long been developed to handle such side information, through Kalman filtering and its advanced forms (EKF, UKF, to name a few), where inequality constraints have been addressed using the notion of Pseudo-Measurements (Tahk and Speyer, 1990). In particle filtering (see Papi et al. (2012) and references therein), affine inequality constraints on dynamics can be dealt with, at the expense of sub-optimality of the solution, using projection of the probability density function with simple saturation functions (Agate and Sullivan, 2004). Other approaches propose to reject estimates that try to escape the region of the state space where the inequality constraints are satisfied (Wang et al., 2002).

Alternatively, in many instances, the problem is recast as a nonlinear programming problem (Arora and Biegler, 2004). A usual approach relies on smoothing splines, which constitutes a principled way to perform regression on measurements (Buisson et al., 2016). However, the discretization of the constraints at the hard-to-select "knot points" is known to be tiresome. For our part, we perform the regression in a reproducing kernel Hilbert space (RKHS). We use fundamental properties of RKHSs to replace the infinite number of constraints (the constraints being defined on a whole time interval) with finite many, without resorting to discretization. The problem is recast with second-order cone constraints. This allows for an efficient implementation of our problem on convex optimization solvers, such as CVXGEN (Mattingley and Boyd, 2012).

⋆ ZSz benefited from the support of the Europlace Institute of Finance and that of the Chair Stress Test, RISK Management and Financial Steering, led by the French École Polytechnique and its Foundation and sponsored by BNP Paribas.

The paper is organized as follows. In Section 2 we formulate the problem of reconstructing trajectories under distance and speed constraints on a time interval as a regularized convex constrained optimization problem in a functional space, first expressed for splines and then in RKHSs. Section 3 is about the optimization of our proposed formulation. Importantly, we explain how this reformulation makes the problem tractable using classical convex programming solvers. In Section 4, we illustrate the approach on a real trajectory dataset, underlining the stability of the reconstruction w.r.t. noise. Conclusions are drawn in Section 5.

2. PROBLEM FORMULATION

In this section we formulate our problem after introducing a few notations.

Notations: Let $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{N} = \{0, 1, \dots\}$ and $\mathbb{N}^* = \{1, 2, \dots\}$ stand for the real, non-negative real, natural numbers and positive integers respectively. We use the $[N] := \{1, \dots, N\}$ shorthand. The j^{th} ($j \in \mathbb{N}$) derivative of a function f is denoted by $f^{(j)}$; we write f' for $j = 1$. The space of continuously differentiable real-valued functions on $\mathcal{T} \times \mathcal{T} \subseteq \mathbb{R}^2$ is denoted by $\mathcal{C}^{(1,1)}(\mathcal{T} \times \mathcal{T})$. The concatenation of vectors $\mathbf{v}_1 \in \mathbb{R}^{d_1}, \dots, \mathbf{v}_M \in \mathbb{R}^{d_M}$ is written as $[\mathbf{v}_1; \dots; \mathbf{v}_M] \in \mathbb{R}^{\sum_{m\in[M]} d_m}$. The zero (resp. all-ones) element of $\mathbb{R}^d$ is denoted by $\mathbf{0}_d$ (resp. $\mathbf{1}_d$). The transpose of a vector $\mathbf{v} \in \mathbb{R}^d$ is denoted by $\mathbf{v}^T$, its Euclidean norm being $\|\mathbf{v}\|_2 = \sqrt{\sum_{i\in[d]} v_i^2}$. Let S be a closed subspace of a Hilbert space $\mathcal{F}$ and $S^\perp = \{f \in \mathcal{F} : \langle f, g\rangle_{\mathcal{F}} = 0 \text{ for all } g \in S\}$ its orthogonal complement. Then every $f \in \mathcal{F}$ has a unique decomposition of the form $f = g + h$ ($g \in S, h \in S^\perp$), i.e $\mathcal{F}$ can be decomposed as $\mathcal{F} = S \oplus S^\perp$.

Our *goal* is to reconstruct the trajectories of a set of vehicles from noisy observations. The vehicles are assumed to form a *convoy* (i.e. they do not overtake and keep a minimum inter-vehicular distance between each other), and we have *speed limits* on the vehicles. These two requirements represent our hard constraints to be fulfilled. We model the trajectories of the vehicles using reproducing kernel Hilbert spaces (RKHS). In the sequel, we introduce this function class and our optimization problem starting from splines (arising from a specific RKHS) and assuming that our convoy has a single vehicle (hence only the speed constraint applies).

Convoy with one vehicle ($Q = 1$): Assume that our convoy is made of a single vehicle. Our datapoints consist of N noisy position measurements $\{x_n\}_{n\in[N]} \subset \mathbb{R}$ recorded at time points $\{t_n\}_{n\in[N]} \subset \mathcal{T} := [0, T]$. In order to capture the (t, x) relation, one can use for example splines. Specifically, let us assume that the modelling class is the Sobolev space[1] $\mathcal{F} := W_2^m(\mathcal{T})$. Then the classical polynomial spline approach can be expressed as the minimization problem

$$\min_{f\in\mathcal{F}} \Bigg[\underbrace{\frac{1}{N} \sum_{n\in[N]} |x_n - f(t_n)|^2}_{\text{approximation error}} + \lambda \underbrace{\int_{\mathcal{T}} |f^{(m)}(t)|^2 \mathrm{d}t}_{\text{smoothing term}} \Bigg], \qquad (1)$$

where $\lambda > 0$ defines the trade-off between the approximation error (first term) and smoothness (second term). It is well-known (Berlinet and Thomas-Agnan, 2004) that the Sobolev space $\mathcal{F}$ can be decomposed as $W_2^m(\mathcal{T}) = \mathcal{F}_1 \oplus \mathcal{F}_2$ with

$$\mathcal{F}_1 = \operatorname{span}\left(1, t, \dots, \frac{t^{m-1}}{(m-1)!}\right),$$

$$\mathcal{F}_2 = \left\{ f \in W_2^m(\mathcal{T}) \mid f^{(j)}(0) = 0 \ (\forall j \in \{0, \dots, m\}\right),$$

$$\|f\|_{\mathcal{F}}^2 = \sum_{j=0}^{m} |f^{(j)}(0)|^2 + \int_{\mathcal{T}} \left| f^{(m)}(t) \right|^2 \mathrm{d}t = \|f_1\|_{\mathcal{F}_1}^2 + \|f_2\|_{\mathcal{F}_2}^2$$

where $f = f_1 + f_2$ ($f_1 \in \mathcal{F}_1$, $f_2 \in \mathcal{F}_2$). In (1), the projection onto $\mathcal{F}_1$ is not penalized, and one can rewrite (1) as

$$\min_{f\in\mathcal{F}} \left[\frac{1}{N} \sum_{n\in[N]} |x_n - f(t_n)|^2 + \lambda \|f_2\|_{\mathcal{F}_2}^2 \right]. \qquad (2)$$

For further details on splines the reader is referred to Wahba (1990); Berlinet and Thomas-Agnan (2004); Wang (2011).

It turns out that not every function class $\mathcal{F}$ with a $\mathcal{F}_1 \oplus \mathcal{F}_2$ decomposition is practically useful. In order to get computationally tractable schemes the relevant assumptions are that (i) $\mathcal{F}_1 = \operatorname{span}\left(\{\varphi_j\}_{j\in[J]}\right)$ is a finite-dimensional space spanned by some basis functions $\{\varphi_j\}_{j\in[J]}$ and (ii) $\mathcal{F}_2 = \mathcal{F}_k$ is a so-called RKHS (Aronszajn, 1950) associated to a kernel $k : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$. This leads to the kernel ridge regression (also called abstract spline) extension of the polynomial spline fitting problem

$$\min_{b\in\mathbb{R},\, f\in\mathcal{F}_k} \left[\frac{1}{N} \sum_{n\in[N]} |x_n - (b + f(t_n))|^2 + \lambda \|f\|_{\mathcal{F}_k}^2 \right], \qquad (3)$$

where $\mathcal{F}_1$ was chosen to be one-dimensional ($J = 1$) containing only the identically constant functions. Using the RKHS norm as a regularizer in (3) ensures the uniqueness of the solution. We will discuss RKHSs ($\mathcal{F}_k$) and kernels (k) at the end of the section.

Finally, let us formulate our proposed optimization task for a single vehicle taking into account the minimum speed constraint[2] ($v_{\min}$) as well:

$$\begin{aligned} &\min_{b\in\mathbb{R},\, f\in\mathcal{F}_k} \left[\frac{1}{N} \sum_{n\in[N]} |x_n - (b + f(t_n))|^2 + \lambda \|f\|_{\mathcal{F}_k}^2 \right] \\ &\quad \text{s.t.} \\ &v_{\min} \le f'(t), \quad \forall t \in \mathcal{T}. \end{aligned} \qquad (4)$$

Convoy with multiple vehicles ($Q > 1$): We now extend our formulation (4) to handle a convoy. Assume that the convoy contains Q vehicles, and that we have N_q noisy position measurements for each of its members $\left\{(t_{q,n}, x_{q,n})_{n\in[N_q]}\right\} \subset \mathcal{T} \times \mathbb{R}$ ($q \in [Q]$) where the time

[1] Classically, the Sobolev space of order m is defined as $W_2^m(\mathcal{T}) := \{f : \mathcal{T} \to \mathbb{R} \mid f^{(j)}$ is absolutely continuous for all $j \in \{0, \dots, m-1\}$, and $\int_{\mathcal{T}} [f^{(m)}]^2(t)\mathrm{d}t < \infty\}$.

[2] We will discuss in Section 3 the similar $v_{\max}$ requirement.

points might differ per vehicle. Assuming (without loss of generality) that the vehicles are ordered by index ($q = Q$ being the last, while $q = 1$ is the first in the lane) the non-overtaking property with given minimum inter-vehicular distance $d_{\min} \geq 0$ can be formulated as $d_{\min} + b_Q + f_Q(t) \leq b_{Q-1} + f_{Q-1}(t), \ldots, d_{\min} + b_2 + f_2(t) \leq b_1 + f_1(t)$ for all $t \in \mathcal{T}$. This gives our final optimization problem describing the trajectory reconstruction of the convoy:

$$\min_{\substack{f_1,\ldots,f_Q \in \mathcal{F}_k,\\ b_1,\ldots,b_Q \in \mathbb{R}}} \frac{1}{Q} \sum_{q=1}^{Q} \left[\left(\frac{1}{N_q} \sum_{n=1}^{N_q} \left| x_{q,n} - (b_q + f_q(t_{q,n})) \right|^2 \right) + \lambda \|f_q\|_{\mathcal{F}_k}^2 \right] \tag{5a}$$

s.t.

$$d_{\min} + b_{q+1} + f_{q+1}(t) \leq b_q + f_q(t), \forall q \in [Q-1],\, t \in \mathcal{T}, \tag{5b}$$

$$v_{\min} \leq f_q'(t), \qquad \forall q \in [Q],\, t \in \mathcal{T}. \tag{5c}$$

We now return to the discussion on the RKHS $\mathcal{F}_k$, the class of functions we use for modelling. A Hilbert space $\mathcal{F}_k$ of $\mathcal{T} \to \mathbb{R}$ functions is called a RKHS (Aronszajn, 1950) with reproducing kernel[3] $k : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ if (i) $k_t(\cdot) := k(\cdot, t) \in \mathcal{F}_k$ for all $t \in \mathcal{T}$ and (ii) $f(t) = \langle f, k_t \rangle_{\mathcal{F}_k}$ for all $t \in \mathcal{T}$ and $f \in \mathcal{F}_k$. This second (reproducing) property expresses that the function evaluation ($f \mapsto f(t)$) in a RKHS is reproduced by taking the inner product with k_t; hence the name. Examples for kernels include the Gaussian (k_G) or the 3/2-Matérn kernel (k_M) defined on $\mathbb{R}$ as

$$k_G(t,s) = e^{-(t-s)^2/(2\sigma^2)}, \tag{6}$$

$$k_M(t,s) = (1 + \sqrt{3}|t-s|/\sigma) e^{-\sqrt{3}|t-s|/\sigma}, \tag{7}$$

where $\sigma > 0$ is the "bandwidth". Constructively, $\mathcal{F}_k = \overline{\mathrm{span}}\,(k_t : t \in \mathcal{T})$; this means that for the Gaussian kernel the elements of $\mathcal{F}_k$ are the limits of sums of Gaussians with real coefficients. For instance $\{k_M(\cdot, t) : t \in \mathbb{R}\}$ spans $W_2^2(\mathbb{R})$.

The properties of the kernel determine that of the elements of the associated RKHS $\mathcal{F}_k$. For example if the kernel is bounded (i.e., $\sup_{t,s \in \mathcal{T}} k(t,s) < \infty$), the same holds for the elements of $\mathcal{F}_k$. A similar conclusion is valid if k is (i) continuous and bounded, (ii) m-times continuously differentiable, or (iii) analytic. A second equivalent definition of kernels is often important from an optimization point of view. A symmetric function $k : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ ($k(t,s) = k(s,t)$ for all $t, s \in \mathcal{T}$) is called kernel if the associated Gram matrix $\mathbf{G} = [k(t_i, t_j)]_{i,j \in [n]}$ is positive semi-definite for any choice of $n \in \mathbb{N}^*$ and $t_1, \ldots, t_n \in \mathcal{T}$. This property of $\mathbf{G}$ will result in a convex objective function. For further details on RKHS and kernels, the reader might consult Berlinet and Thomas-Agnan (2004); Steinwart and Christmann (2008); Saitoh and Sawano (2016).

Having covered our proposed trajectory inference formulation of the convoy (5a)-(5c) and basic properties of RKHSs, in the next section we focus on the optimization of our objective function.

[3] We abbreviate the function $s \in \mathcal{T} \mapsto k(s,t) \in \mathbb{R}$ as $k(\cdot, t)$.

3. OPTIMIZATION

The primary challenge one has to resolve in the optimization problem (5a)-(5c) is the infinite number of constraints (due to $\mathcal{T}$) in (5b) and (5c). In the literature, such requirements are typically tackled by requiring the constraints to hold only at a finite number of time points; unfortunately this discretization approach does not guarantee that the constraints are fulfilled elsewhere. In contrast, we propose a strengthened optimization problem which implies both (5b)-(5c) and is computationally tractable.

We *assume* that k is defined on a set containing $\mathcal{T}$ and that its restriction to $\mathcal{T}$ belongs to $\mathcal{C}^{(1,1)}(\mathcal{T} \times \mathcal{T})$ with bound $\kappa := \sup_{t,s \in \mathcal{T}} \sqrt{k(t,s)}$, which holds for example with $\kappa = 1$ for the Gaussian and Matérn kernels. Let the union of the measured time points be $\{t_m\}_{m \in [M]} := \cup_{q \in [Q], n \in [N_q]} \{t_{q,n}\}$. We look for solutions of the form $f_q = \sum_{m \in [M]} a_{q,m} k_{t_m}$ ($\mathbf{a}_q := (a_{q,m})_{m \in [M]} \in \mathbb{R}^M$).

Given $N \in \mathbb{N}^*$ and a finite family of functions $\{g_n\}_{n \in [N]} \subset \mathcal{F}_k$, let $\mathbf{u} = (u_n)_{n \in [N]} \in \mathbb{R}^N$ such that $u_n \geq \sup_{t \in \mathcal{T}} g_n(t)$. Then, our strengthened convex optimization problem expresses as:

$$\min_{\substack{\boldsymbol{a}_1,\ldots,\boldsymbol{a}_Q \in \mathbb{R}^M,\\ b_1,\ldots,b_Q \in \mathbb{R},\\ \boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_Q \in \mathbb{R}_+^N,\\ \boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_Q \in \mathbb{R}_+^N}} \frac{1}{Q} \sum_{q=1}^{Q} \frac{1}{N_q} \Big[\boldsymbol{a}_q^T (\mathbf{G}_M \boldsymbol{\Pi}_q^T \boldsymbol{\Pi}_q \mathbf{G}_M + \lambda N_q \mathbf{G}_M) \boldsymbol{a}_q + N_q b_q^2 + 2 \left(b_q \mathbf{1}_{N_q} - \boldsymbol{x}_q \right)^T \boldsymbol{\Pi}_q \mathbf{G}_M \boldsymbol{a}_q - 2 b_q \mathbf{1}_{N_q}^T \boldsymbol{x}_q \Big] \tag{8a}$$

s.t.

$$\kappa \left\| \mathbf{G}_0^{1/2} [\boldsymbol{a}_q - \boldsymbol{a}_{q+1}; \boldsymbol{\alpha}_q] \right\|_2 \leq b_q - b_{q+1} - d_{\min} - \boldsymbol{\alpha}_q^T \boldsymbol{u}, \quad \forall q \in [Q-1], \tag{8b}$$

$$\kappa \left\| \mathbf{G}_D^{1/2} [\boldsymbol{a}_q; \boldsymbol{\beta}_q] \right\|_2 \leq -v_{\min} - \boldsymbol{\beta}_q^T \boldsymbol{u}, \, \forall q \in [Q], \tag{8c}$$

where $\boldsymbol{x}_q = (x_{q,n})_{n \in [N_q]} \in \mathbb{R}^{N_q}$, $\boldsymbol{\Pi}_q \in \mathbb{R}^{N_q \times M}$ is the projector from $\mathbb{R}^M$ to $\mathbb{R}^{N_q}$, the $\mathcal{F}_k$-function $\partial_2 k(\cdot, t)$ is the derivative of $k(\cdot, t)$ w.r.t. the second variable, $\mathbf{G}_0 \in \mathbb{R}^{(M+N) \times (M+N)}$ is the Gram matrix of $(k_{t_1}, \ldots, k_{t_M}, g_1, \ldots, g_N)$, $\mathbf{G}_D \in \mathbb{R}^{(M+N) \times (M+N)}$ is the Gram matrix of $(\partial_2 k(\cdot, t_1), \ldots, \partial_2 k(\cdot, t_M), g_1, \ldots, g_N)$, $\mathbf{G}_M \in \mathbb{R}^{M \times M}$ is the Gram matrix of $(k_{t_1}, \ldots, k_{t_M})$.

In practice, one can set for instance $N = M$ and $g_m = -k_{t_m}$ for all $m \in [M]$; this is the choice made in Section 4. Furthermore, constraint (8c) can be overly conservative. An alternative is to cover $\mathcal{T}$ with intervals $I_m := [t_m - \delta_m, t_m + \delta_m]$ with $\delta_m > 0$, set $u_m = \sup_{t \in I_m} g_m(t)$ and replace (8c) by

$$\kappa \left\| \mathbf{G}_m^{1/2} [\boldsymbol{a}_q; \beta_{q,m}] \right\|_2 \leq -v_{\min} - \beta_{q,m} u_m, \, \forall q \in [Q], \, m \in [M] \tag{8d}$$

with $\mathbf{G}_m \in \mathbb{R}^{(M+1) \times (M+1)}$ being the Gram matrix of $(\partial_2 k(\cdot, t_1), \ldots, \partial_2 k(\cdot, t_M), g_m)$.

Remarks:

- While the derivation of the optimization problem (8a)-(8c) is involved (using convex analysis in Hilbert spaces), we can give the intuition with even more strengthened constraints (but which are less useful in practice), the proof of which is straightforward. Consider requirement (5b) for a fixed $q \in [Q-1]$; (5c) can be handled similarly. In case of (5b), one has to satisfy

$$\sup_{t\in\mathcal{T}} \big[\underbrace{f_{q+1}(t) - f_q(t)}_{=\langle f_{q+1}-f_q, k_t\rangle_{\mathcal{F}_k}}\big] \le b_q - b_{q+1} - d_{\min}, \qquad (9)$$

where on the l.h.s. we applied the *reproducing property* of k. Since $\|k_t\|^2_{\mathcal{F}_k} = \langle k_t, k_t\rangle_{\mathcal{F}_k} = k(t,t) \le \kappa^2$ for any $t \in \mathcal{T}$, it is sufficient for (9) to hold that f_q, f_{q+1}, b_q, b_{q+1} satisfy

$$\sup_{g\in\mathcal{F}_k,\, \|g\|_{\mathcal{F}_k}\le\kappa} \langle f_{q+1} - f_q, g\rangle_{\mathcal{F}_k} \le b_q - b_{q+1} - d_{\min}.$$

Applying the Cauchy-Schwarz inequality, one gets

$$\kappa \|f_{q+1} - f_q\|_{\mathcal{F}_k} \le b_q - b_{q+1} - d_{\min}. \qquad (10)$$

Under the parameterization $f_q = \sum_{m\in[M]} a_{q,m} k_{t_m}$ ($a_{q,m} \in \mathbb{R}$) and applying again the *reproducing property* of k, (10) reads, equivalently, as

$$\kappa \left\| \mathbf{G}_M^{1/2}(\mathbf{a}_q - \mathbf{a}_{q+1})\right\|_2 \le b_q - b_{q+1} - d_{\min}. \qquad (11)$$

This inequality is a special case of (8b) with $\boldsymbol{\alpha}_q = \mathbf{0}$.

- Solution of (8a)-(8c) in practice: the convex problem (8a)-(8c) can be readily implemented with classical solvers such as CVXGEN (Mattingley and Boyd, 2012). It has a quadratic objective function and constraints involving a Euclidean norm and affine terms, i.e. second-order cone constraints. Hence, our problem scales at worst as $\mathcal{O}((Q(N+M))^3)$ (Alizadeh and Goldfarb, 2003). Computing the Gram matrices $\mathbf{G}$ boils down to using the reproducing property for function values and for their derivatives. For all $t \in \mathcal{T}$ and $h \in \mathcal{F}_k$, this writes as

$$h(t) = \langle h, k_t\rangle_{\mathcal{F}_k}, \quad h'(t) = \langle h, \partial_2 k_t\rangle_{\mathcal{F}_k},$$

which implies that

$$\langle h_1, h_2\rangle_{\mathcal{F}_k} = \sum_{i\in[N_1]}\sum_{j\in[N_2]} c_i d_j k(t_i, s_j)$$

whenever h_1 and h_2 can be written in the form of $h_1 = \sum_{i\in[N_1]} c_i k_{t_i}$ and $h_2 = \sum_{j\in[N_2]} d_j k_{s_j}$ ($c_i, d_j \in \mathbb{R}$, $t_i, s_j \in \mathcal{T}$). Hence, it is worthwhile to choose $\{g_n\}_{n\in[N]} \subset \operatorname{span}(\{k_t\}_{t\in\mathcal{T}})$; this being a mild assumption as $\operatorname{span}(\{k_t\}_{t\in\mathcal{T}})$ is dense in $\mathcal{F}_k$. Under this requirement on $\{g_n\}_{n\in[N]}$, all the inner products appearing in $\mathbf{G}$ are easy to compute. As these Gram matrices are positive semi-definite (see the end of Section 2), one can take the matrix square roots $\mathbf{G}^{1/2}$ in (8b) and (8c), which can be replaced by the output of their Cholesky decomposition. These computations need to be done only once, prior to the numerical resolution of (8a).

- Maximum speed constraint: in addition to the minimum speed constraint (5c), one might also have an upper bound on the speed of the vehicles:

$$f'_q(t) \le v_{\max}, \ \forall q \in [Q],\, t \in \mathcal{T}.$$

Such a requirement can be encoded similarly to (8c): introducing an additional variable $\boldsymbol{\gamma}_q \in \mathbb{R}^N_+$, it is sufficient to add the

$$\kappa \left\| \mathbf{G}_D^{1/2}[-\boldsymbol{a}_q; \boldsymbol{\gamma}_q]\right\|_2 \le v_{\max} - \boldsymbol{\gamma}_q^T \boldsymbol{u}, \ \forall q \in [Q] \qquad (12)$$

constraint to the optimization problem (8a)-(8c).

- Group of convoys: one could also consider an extension of the presented trajectory inference formulation (8a)-(8c) to groups of cars forming convoys for given time intervals. In this case, one should just replace the time interval $\mathcal{T}$ and the car indices $[Q]$ with some subsets in the constraint (8b), and change the bound $\mathbf{u}$ accordingly. The less restrictive formulation (8d) of (8c) exploits the same idea.

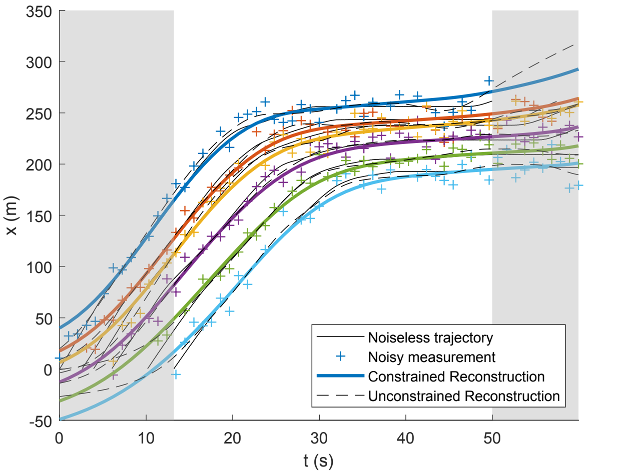
4. NUMERICAL EXPERIMENTS

In this section, we demonstrate the efficiency of the proposed trajectory reconstruction approach with speed and inter-vehicular distance constraints. The goal of the experiments is two-fold:

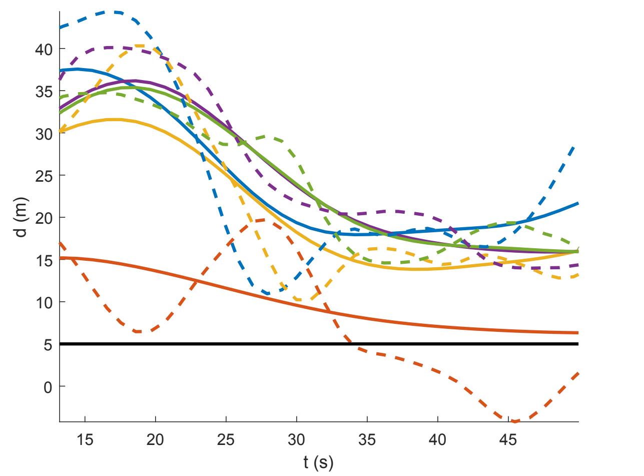
- **Experiment-1**: We show that trajectory reconstruction not taking into account the speed and inter-vehicular distance requirements can easily lead to inconsistencies, whereas our constrained solution remains consistent. We illustrate the idea on GPS-like data from a highway section including a traffic jam.
- **Experiment-2**: We complement the first experiment by investigating how severe the error in the trajectory reconstruction is as a function of the measurement noise. For large levels of noise, enforcing the constraints is quite beneficial.

We start by describing our dataset. We use trajectories from the recent MoCoPo benchmark (Buisson et al., 2016). These trajectories $t \mapsto (x, y)$ correspond to cars driving on a two-lane highway. For illustration purposes, we select six vehicles ($Q = 6$) following each other in the same lane with no overtaking and including a traffic jam. Including the traffic jam section makes the estimation extremely challenging as the measurement noise in the positions can easily lead to the (false) prediction that the vehicles are moving forward and back.

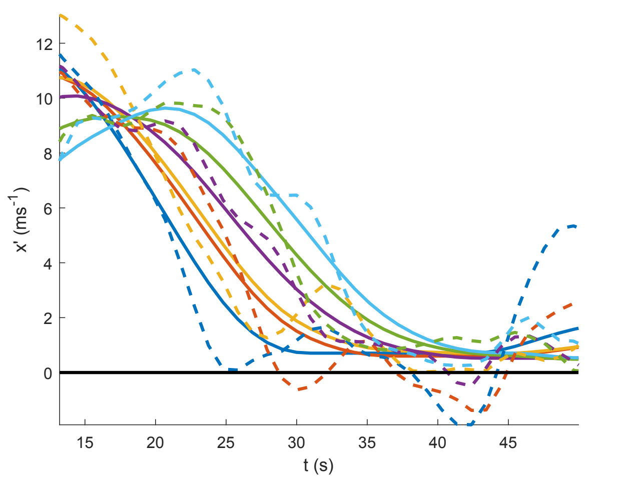
After projecting the two-dimensional position coordinates (x, y) to the distance travelled along the lane (from now on referred to as x), we sub-sample it by decreasing the 25 Hz measurement frequency to 1 Hz which is a usual value for GPS measurements. We then consider 20% of the measurements to be missing resulting in $M = N = 350$ data points, and corrupt the remaining data by adding Gaussian noise to it with standard deviation σ_{noise}. We also apply a pre-processing step, taking out an affine component. This can be interpreted as modelling the discrepancy with respect to a reference trajectory (acting as a virtual vehicle in the middle of the convoy) rather than the trajectories themselves. It corresponds to replacing the position measurement $x_{q,n}$ with $x_{q,n} - f_0(t_{q,n})$ for all $q \in [Q], n \in [N_q]$. Here $f_0(t) = \bar{v}\,t + \bar{x}$ is obtained by performing linear regression on the points $\{(t_{q,n}, x_{q,n})\}_{q\in[Q], n\in[N_q]}$. In the implementation of (8c), we thus change accordingly $v_{\min}$ to $v_{\min} - \bar{v}$.



(a) Measurements and reconstructed trajectories $t \mapsto f_q(t)$. Each colour represents a vehicle. Grey area: when some vehicles are out of the road section.



(b) Reconstructed pairwise distances $t \mapsto f_q(t) - f_{q+1}(t)$ ($q = 1, \ldots, 5$) compared to the $d_{\min} = 10\,\mathrm{m}$ threshold. Solid lines: constrained estimator. Dashed lines: unconstrained estimator. The colour of the lines correspond to that of the leading vehicle.



(c) Reconstructed velocities $t \mapsto f_q'(t)$ compared to the $v_{\min} = 0\,\mathrm{ms}^{-1}$ threshold. The notation is the same as in (b).

Figure 1: Reconstruction of the convoy trajectory.

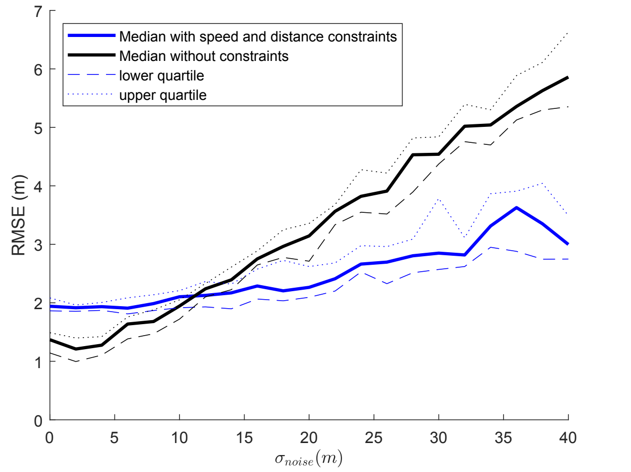


Figure 2: Median, lower and upper quartiles of the RMSE w.r.t. the ground truth trajectories as a function of the noise level σ_{noise}.

Unlike the polar coordinates used by Buisson et al. (2016), and introduced to project 2D to 1D data, we take the projection step for granted and use information from the whole convoy to reconstruct the individual trajectories. We also corrupt the data that came from cameras to resemble GPS data.

In our *first experiment* the measurement noise is at an average level $\sigma_{\text{noise}} = 5\,\mathrm{m}$. We require the vehicles to maintain a distance of at least $d_{\min} = 5\,\mathrm{m}$ and a velocity $v_{\min} = 0\,\mathrm{ms}^{-1}$; the latter bound encodes that cars cannot go backward on a highway. We used the 3/2-Matérn kernel (7) with bandwidth σ equal to the square root of the median of the squared pairwise distance of the time points, and applied leave-one-out cross-validation (see e.g. (Rifkin and Lippert, 2007)) to determine the optimal regularization parameter λ.

We compare our proposed trajectory reconstruction approach taking into account both the inter-vehicular distance and speed constraints (solution of (8a)-(8b)-(8d)) with the unconstrained trajectory estimator (solution of (8a) *without* (8b)-(8d)). The estimated trajectories are depicted in Fig. 1a with the noiseless (used for performance evaluation) and the noisy measurements (used for estimation).

The reconstructed pairwise distances $t \mapsto f_q(t) - f_{q+1}(t)$ ($q \in [Q-1]$) and velocities $t \mapsto f_q'(t)$ ($q \in [Q]$) are illustrated in Fig. 1b and Fig. 1c. As it can be read out from Fig. 1b, one pair of vehicles clearly violates the distance constraint for the unconstrained kernel ridge regression (KRR), while the proposed scheme respects the inter-vehicular distance requirement. The situation is even more severe in case of the estimated speed values: as it can be seen in Fig. 1c many speed trajectories obtained by KRR take negative values. In contrast, the proposed technique correctly handles the speed constraints, even in the challenging traffic jam scenario. Notice that the values of velocity never reach $v_{\min}$ due to the conservatism of our strengthened approach. This experiment demonstrates the efficiency of our trajectory reconstruction method for an average measurement noise level $\sigma_{\text{noise}} = 10\,\text{m}$.

In order to provide further insight into the behaviour of our approach, in the *second experiment* we studied the effect of the noise level σ_{noise} on the accuracy of the trajectory estimation. The accuracy of the estimation was computed as the root-mean-square error (RMSE) w.r.t. the ground truth trajectories at the time points when all the vehicles were on the studied road section (white area in Fig. 1a) at the original 25 Hz frequency. The experiment was repeated 40 times and the resulting median, lower and upper quartiles are reported in Fig. 2 for varying noise level. One can see that for large noise level (say $\sigma_{\text{noise}} \geq 10\,\text{m}$) the added constraints offer a more accurate reconstruction than the unconstrained KRR method even in RMSE sense. For smaller noise level the precise handling of the inter-vehicular distance and speed bounds is the main benefit of the proposed approach while keeping comparable RMSE w.r.t. KRR. These two experiments illustrate the efficiency of our trajectory reconstruction technique which allows taking into account inter-vehicular distance and speed constraints in a principled way.

5. CONCLUSIONS AND PERSPECTIVES

In this article, a method has been presented to recover the curvilinear position of vehicles from noisy measurements while enforcing constraints of minimum inter-vehicular distance and speed limits at all times. The proposed method is guaranteed to provide feasible estimates, benefiting from the RKHS representation of the trajectories. A key feature is that the inequality constraints can be addressed, without any discretization, as a finite dimensional convex problem that can be efficiently solved. The obtained numerical results show that we get a good root-mean-square error in short computation time on a small dataset of real trajectories. We plan to extend this approach to larger transportation datasets, such as NGSIM (US Department of Transportation – FHWA, 2008), and to other application fields where the positivity and monotonicity appear naturally as a valuable side-information when performing a shape-constrained regression.

REFERENCES

Agate, C.S. and Sullivan, K.J. (2004). Road-constrained target tracking and identification using a particle filter. In *Signal and Data Processing of Small Targets*, 532–543.

Agosta, G., Barenghi, A., Brandolese, C., Fornaciari, W., Pelosi, G., Delucchi, S., Massa, M., Mongelli, M., Ferrari, E., Napoletani, L., Bozzi, L., Tieri, C., Cassioli, D., and Pomante, L. (2016). V2I cooperation for traffic management with SafeCop. In *Euromicro Conference on Digital System Design*, 621–627.

Alizadeh, F. and Goldfarb, D. (2003). Second-order cone programming. *Mathematical Programming*, 95(1), 3–51.

Alouani, A.T. and Blair, W.D. (1991). Use of a kinematic constraint in tracking constant speed, maneuvering targets. In *IEEE Conference on Decision and Control*, 2055–2058.

Aronszajn, N. (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68, 337–404.

Arora, N. and Biegler, L.T. (2004). A trust region SQP algorithm for equality constrained parameter estimation with simple parameter bounds. *Computational Optimization and Applications*, 28(1), 51–86.

Berlinet, A. and Thomas-Agnan, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer.

Buisson, C., Villegas, D., and Rivoirard, L. (2016). Using polar coordinates to filter trajectories data without adding extra physical constraints. In *Transportation Research Board 95th Annual Meeting*.

Chang, T., Wang, L., and Chang, F. (2009). A solution to the ill-conditioned GPS positioning problem in an urban environment. *IEEE Transactions on Intelligent Transportation Systems*, 10(1), 135–145.

De Wit, C.C., Morbidi, F., Ojeda, L.L., Kibangou, A.Y., Bellicot, I., and Bellemain, P. (2015). Grenoble traffic lab: An experimental platform for advanced traffic monitoring and forecasting. *IEEE Control Systems Magazine*, 35(3), 23–39.

Liang, K.Y., van de Hoef, S., Terelius, H., Turri, V., Besselink, B., Mårtensson, J., and Johansson, K.H. (2016). Networked control challenges in collaborative road freight transport. In *European Control Conference*, 2 – 14.

Mattingley, J. and Boyd, S. (2012). CVXGEN: a code generator for embedded convex optimization. *Optimization and Engineering*, 13(1), 1–27.

Motulsky, H. and Christopoulos, A. (2004). *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting*. Oxford University Press.

Papi, F., Podt, M., Boers, Y., Battistello, G., and Ulmke, M. (2012). On constraints exploitation for particle filtering based target tracking. In *International Conference on Information Fusion*, 455–462.

Rifkin, R.M. and Lippert, R.A. (2007). Notes on regularized least-squares. Technical report. (`http://cbcl.mit.edu/publications/ps/MIT-CSAIL-TR-2007-025.pdf`).

Saitoh, S. and Sawano, Y. (2016). *Theory of Reproducing Kernels and Applications*. Springer Singapore.

Steinwart, I. and Christmann, A. (2008). *Support Vector Machines*. Springer.

Tahk, M. and Speyer, J.L. (1990). Target tracking problems subject to kinematic constraints. *IEEE Transactions on Automatic Control*, 35(3), 324–326.

US Department of Transportation – FHWA (2008). NGSIM – Next Generation SIMulation. (`https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm`).

Wahba, G. (1990). *Spline Models for Observational Data*. SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics.

Wang, L.S., Chiang, Y.T., and Chang, F.R. (2002). Filtering method for nonlinear systems with constraints. *IEE Proceedings-Control Theory and Applications*, 149(6), 525–531.

Wang, Y. (2011). *Smoothing Splines – Methods and Applications*. CRC Press.

Work, D.B., Tossavainen, O., Jacobson, Q., and Bayen, A.M. (2009). Lagrangian sensing: traffic estimation with mobile devices. In *American Control Conference*, 1536–1543.