

Estimation and control under constraints through Kernel methods

Pierre-Cyril Aubin-Frankowski

▶ **To cite this version:**

Pierre-Cyril Aubin-Frankowski. Estimation and control under constraints through Kernel methods. Automatic Control Engineering. Université Paris sciences et lettres, 2021. English. NNT : 2021UP-SLM018 . tel-03296985

HAL Id: tel-03296985
https://pastel.hal.science/tel-03296985

Submitted on 23 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à MINES ParisTech

Estimation and Control under Constraints through Kernel Methods

Estimation et Contrôle sous Contraintes par Méthodes à Noyaux

Soutenue par

Pierre-Cyril AUBIN-FRANKOWSKI

Le 5 Juillet 2021

École doctorale n°ED 621

Ingénierie des Systèmes, Matériaux, Mécanique, Energétique

Spécialité

Mathématique et automatique

Composition du jury :

Alban QUADRAT Directeur, INRIA	*Président du jury*
Michel DE LARA Professeur, ENPC	*Rapporteur*
Johan SUYKENS Professeur, KUL	*Rapporteur*
Pauline BERNARD Professeur assistant, MINES Paris	*Examinatrice*
Didier HENRION Directeur, CNRS	*Examinateur*
Laurent PFEIFFER Chargé de recherche, INRIA	*Examinateur*
Nicolas PETIT Professeur, MINES Paris	*Directeur de thèse*

MINES ParisTech | PSL

Estimation and Control under Constraints through Kernel Methods

© 2021

Pierre-Cyril Aubin-Frankowski



This thesis is licensed under a Creative Commons Attribution-Non Commercial-ShareAlike 4.0 International License.

À mes parents, Halina et Jean-Pierre,

ABSTRACT

Pointwise state and shape constraints in control theory and nonparametric estimation are difficult to handle as they often involve convex optimization problem with an infinite number of inequality constraints. Satisfaction of these constraints is critical in many applications, such as path-planning or joint quantile regression. Reproducing kernels are propitious for pointwise evaluations. However representer theorems, which ensure the numerical applicability of kernels, cannot be applied for an infinite number of evaluations. Through constructive algebraic and geometric arguments, we prove that an infinite number of affine real-valued constraints over derivatives of the model can be tightened into a finite number of second-order cone constraints when looking for functions in vector-valued reproducing kernel Hilbert spaces. We show that state-constrained Linear-Quadratic (LQ) optimal control is a shape-constrained regression over the Hilbert space of linearly-controlled trajectories defined by an explicit LQ kernel related to the Riccati matrix. The efficiency of the developed approach is illustrated on various examples from both linear control theory and nonparametric estimation. Finally, we provide some results for general differential inclusions in minimal time problems and identification of the graph of the set-valued map. Most of all we bring to light a novel connection between reproducing kernels and optimal control theory, identifying the Hilbertian kernel of linearly controlled trajectories.

RÉSUMÉ

Les contraintes ponctuelles d'état et de forme en théorie du contrôle et en estimation non-paramétrique sont difficiles à traiter car elles impliquent souvent un problème d'optimisation convexe en dimension infinie avec un nombre infini de contraintes d'inégalité. La satisfaction de ces contraintes est essentielle dans de nombreuses applications, telles que la planification de trajectoires ou la régression quantile jointe. Or, les noyaux reproduisants sont un choix propice aux évaluations ponctuelles. Cependant les théorèmes de représentation qui en sous-tendent les applications numériques ne peuvent pas être appliqués à un nombre infini d'évaluations. Par des arguments algébriques et géométriques constructifs, nous prouvons qu'un nombre infini de contraintes affines à valeur réelle sur les dérivées des fonctions peut être surcontraint par un nombre fini de contraintes coniques du second ordre si l'on considère des espaces de Hilbert à noyau reproduisant de fonctions à valeurs vectorielles. Nous montrons que le contrôle optimal linéaire-quadratique (LQ) sous contraintes d'état est une régression sous contraintes de forme sur l'espace de Hilbert de trajectoires contrôlées linéairement. Cet espace est défini par un noyau LQ explicite lié à la matrice de Riccati. L'efficacité de notre approche est illustrée par divers exemples issus de la théorie du contrôle linéaire et de l'estimation non-paramétrique. Enfin, nous énonçons des résultats pour des inclusions différentielles générales dans des problèmes de temps minimal et d'identification du graphe de la correspondance. Surtout nous faisons ressortir un lien nouveau entre méthodes à noyaux et contrôle optimal en identifiant le noyau hilbertien des espaces de trajectoires contrôlées linéairement.

REMERCIEMENTS

Je remercie mon directeur de thèse, Nicolas Petit, pour sa confiance et l'inestimable liberté de recherche qu'il m'a accordée durant ces trois belles années. Je suis très reconnaissant à Johan Suykens et Michel De Lara de m'avoir fait l'insigne honneur de rapporter sur ma thèse, et à Pauline Bernard, Didier Henrion, Laurent Pfeiffer et Alban Quadrat, d'avoir accepté de faire partie de mon jury.

J'ai été amené aux noyaux reproduisants par le talent pédagogique de Jean-Philippe Vert et y suis resté grâce à l'affabilité de Zoltán Szabó. Je ne saurais assez leur exprimer ma gratitude pour m'avoir formé dans le domaine et d'avoir fait de mes ébauches des articles. Anna Korba et Pierre Ablin, soyez remerciés pour ce bel article entre deux confinements ! Ces collaborations, ainsi que celles avec Alessandro Rudi, Stéphane Gaubert et Thomas Kerdreux, ont illuminé mes années.

J'ai eu la chance d'avoir des anges gardiens, qui ont su par leur humanité extraordinaire, m'accompagner jusqu'à bon port. Je suis immensément reconnaissant à Françoise Prêteux d'avoir rempli ce rôle salutaire. À Stéphane Canu et Michel De Lara pour leurs conseils attentionnés quand j'hésitais. À mes frères, Henri-Jean et Marc, de m'avoir toujours témoigné amitié et soutien. À Georges Haddad et Vladimir Lozève de m'avoir communiqué la passion qui les anime. À mes maîtres, Jean-Louis Basdevant, Alain Pommellet, Jean-François Logeais, Jean-Louis Frot et Agnès Besse, de m'avoir transmis leur goût pour l'enseignement et de m'avoir armé pour descendre dans l'arène.

J'y ai été si bien entouré. Par mon vieux frère d'armes de Bozar, Robin, dont la gentillesse est contagieuse. Par Dhruv et nos discussions si souvent renouvelées sur le monde, la physique ou l'économie. Je vous rends grâce –Denis, Laetitia, José, Son, Thomas, Alexandre- pour notre amitié. Merci à mes camarades de thèse, Dilshad, Hubert, Nils et Mona, dont la présence m'a été trop longtemps privée, comme m'a un temps été ravie la joie d'arpenter les opéras avec mes amis de Juvenilia, absence adoucie par les chocolats de Fitzgerald Addoh.

C'est enfin à mes parents qui m'ont amené à la lumière, dans toutes ses acceptions, que je veux rendre hommage. J'espère mériter toujours leur estime.

CONTENTS

1 INTRODUCTION 1
1.1 Background on estimation and control with constraints . . . 2
1.1.1 Considered framework for optimizing models with constraints . 2
1.1.2 Shape constraints in nonparametric regression . . . 6
1.1.3 State constraints in LQ optimal control . . . 8
1.2 Reproducing kernels in a nutshell . . . 9
1.2.1 History of reproducing kernels: A realm (still) divided . . . 10
1.2.2 Positive definite kernels: A nonlinear embedding with inner products . . . 11
1.2.3 RKHSs: A Hilbertian topology stronger than pointwise convergence 13
1.2.4 Vector-valued RKHSs and reproducing property for derivatives . 15
1.2.5 Two computational tools: the representer theorems and the kernel trick . . . 16
1.3 Contributions of the thesis . . . 19
1.3.1 Structure of the thesis . . . 19
1.3.2 Tightening infinitely many shape constraints into finitely many . 20
1.3.3 A new kernel for (state-constrained) LQ optimal control . . . 30
1.3.4 Differential inclusions: Lipschitz minimal time and kernel-based graph identification . . . 37
Bibliography for Chapter 1 . . . 41
2 KERNEL REGRESSION WITH SHAPE CONSTRAINTS FOR VEHICLE TRAJECTORY RECONSTRUCTION 49
2.1 Introduction . . . 50
2.2 Problem Formulation . . . 51
2.3 Optimization . . . 54
2.4 Numerical Experiments . . . 57
2.5 Conclusions and perspectives . . . 61
Bibliography for Chapter 2 . . . 61
3 REAL-VALUED AFFINE SHAPE CONSTRAINTS IN RKHSS 63
3.1 Introduction . . . 64
3.2 Problem formulation . . . 66
3.3 Results . . . 69
3.4 Numerical experiments . . . 73
3.5 Appendix of the results of Chapter 3 . . . 76

3.5.1 Proof of Theorem 3.1 . . . 76
3.5.2 Shape-constrained kernel ridge regression . . . 81
3.5.3 Examples of handled shape constraints . . . 83
Bibliography for Chapter 3 . . . 83
4 STATE CONSTRAINTS IN LQ OPTIMAL CONTROL THROUGH THE LQ KERNEL 89
4.1 Introduction . . . 90
4.2 Theoretical preliminaries and problem formulation . . . 93
4.3 Revisiting LQ control through kernel methods . . . 96
4.3.1 Case of vanishing Q . . . 98
4.3.2 Case of nonvanishing Q . . . 102
4.4 Theoretical approximation guarantees . . . 103
4.5 Finite-dimensional implementation and numerical example . . . 110
Bibliography for Chapter 4 . . . 118
5 THE LQ REPRODUCING KERNEL AND THE RICCATI EQUATION 123
5.1 Introduction . . . 124
5.2 Vector spaces of linearly controlled trajectories as vRKHSs . . . 125
5.3 Proof of Theorem 5.2 . . . 128
Bibliography for Chapter 5 . . . 131
6 LIPSCHITZ MINIMUM TIME FOR DIFFERENTIAL INCLUSIONS WITH STATE CONSTRAINTS 133
6.1 Introduction . . . 134
6.2 Main results . . . 136
6.3 Discussion on the main results . . . 138
6.3.1 Nonautonomous systems . . . 138
6.3.2 Weakening the hypotheses . . . 139
6.3.3 Considering point targets . . . 140
6.4 Proofs . . . 142
Bibliography for Chapter 6 . . . 150
7 DATA-DRIVEN SET APPROXIMATION FOR DETECTION OF TRANSPORTATION MODES 151
7.1 Introduction . . . 152
7.2 Approximating discrete sets with SVDD . . . 154
7.2.1 Notations . . . 154
7.2.2 Theoretical framework of the SVDD algorithm . . . 154
7.2.3 The SVDD algorithm . . . 155
7.2.4 Extension of SVDD to noisy data . . . 156
7.2.5 Discussion on application of SVDD and illustration on real data 156
7.3 Theoretical guarantees on set estimation . . . 157

7.3.1 Main results 159
7.3.2 Further results 162
7.4 Application to detection of transportation mode on simulated data . . . 163
7.5 Conclusion and perspectives 164
Bibliography for Chapter 7 164

1

INTRODUCTION

> The sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work — that is, correctly to describe phenomena from a reasonably wide area. Furthermore, it must satisfy certain aesthetic criteria — that is, in relation to how much it describes, it must be rather simple.
>
> — John von Neumann, *Method in the Physical Sciences* in *The Unity of Knowledge*, 1955

Abstract This chapter outlines the content of the manuscript. We first present the general idea and necessary background on the topics of state and shape constraints and on reproducing kernels. We particularly stress how Linear-Quadratic optimal control falls within a machine learning formalism. We then discuss the contributions of the thesis. In all that follows, we focus on non-discrete domains and convex optimization problems over vector-valued reproducing kernel Hilbert spaces of functions with an infinite number of affine real-valued constraints on their derivatives.

Résumé Ce chapitre décrit le contenu du manuscrit. Nous présentons d'abord l'idée générale et divers pré-requis et éléments de contexte sur les contraintes de forme et d'état, ainsi que sur les noyaux reproduisants. Nous soulignons en particulier comment le contrôle optimal linéaire-quadratique peut s'écrire dans un formalisme d'apprentissage automatique. Nous discutons ensuite des contributions de la thèse. Dans tout ce qui suit, nous nous concentrons sur des problèmes d'optimisation continue, convexes, portant sur des espaces de Hilbert à noyau reproduisant de fonctions à valeurs vectorielles avec un nombre infini de contraintes affines à valeur réelle sur leurs dérivées.

1.1 BACKGROUND ON ESTIMATION AND CONTROL WITH CONSTRAINTS

1.1.1 *Considered framework for optimizing models with constraints*

What do estimation and control have in common? For linear-quadratic systems, a control theorist may say there is some intrinsic duality between observability and controllability. For an agnostic optimizer, estimation and control share the burden of dealing with choices. For instance, one wants to choose a nonparametric model based on given samples, or choose a path between some given points. In both estimation and control, optimization theory allows to cut the Gordian node, selecting a model that minimizes a functional (the *objective*) over a collection of models (the *hypothesis class*). Like in social selection, it is dubious whether there exists an optimal choice of objective and hypothesis class, there are just many of them, each with its pros and cons. However the class of considered problems should have some minimal properties. Unlike the Pithy of Delphi, the answer from the solicited optimization framework should not be sibylline. It should be unique and, if possible, sensible, that is to say, feasible. The larger the class, the more choices available, but all should abide by the reality principle, and the fabric of reality is its constraints.

Among constraints, there are those that should hold pointwise. For non-discrete domains of definition, they would in principle necessitate Dirac masses to define the evaluation of the model at some points. However one does not not need to involve Laurent Schwartz's distributions for a very special class of Hilbert spaces of functions, namely *reproducing kernel Hilbert spaces* (RKHSs). In short, these spaces $\mathcal{H}_K$ can be defined as the spaces for which the Hilbertian topology is stronger than pointwise convergence. Born in Green functions and rejuvenated under the name of kernel methods in machine learning, these rich spaces nevertheless provide tractable solutions. Most of all they satisfy John von Neumann's "aesthetic criteria". This is the main reason of the interest of the author of this thesis for the elegant mathematical field of kernels. If this introduction achieves to convey this touch of refinement to the reader, then it will have achieved its purpose.

In this thesis, we focus on convex optimization problems with affine real-valued constraints on derivatives of the model, which is sought for in a RKHS of $\mathbb{R}^Q$-valued functions defined over a set $\mathcal{X}$, for $Q \in \mathbb{N}^*$. To properly define derivatives, the set $\mathcal{X}$ is assumed to be a subset of $\mathbb{R}^d$ and to be contained in the closure of its interior. The function family $\mathcal{H}_K$ is used to capture the relation between the input variable x and

the output $\mathbf{y}$, with the optional use of a bias term $\mathbf{b} \in \mathbb{R}^B$, with $B \in \mathbb{N}^*$. Hence our **goal** is to solve constrained optimization problems of the form

$$\begin{aligned} (\bar{\mathbf{f}}, \bar{\mathbf{b}}) \in \quad & \underset{(\mathbf{f},\mathbf{b}) \in \mathcal{H}_K \times \mathbb{R}^B}{\arg\min} \quad \mathcal{L}(\mathbf{f}, \mathbf{b}) && (\mathcal{P}_{\mathrm{ML}}) \\ & \text{s.t.} \\ & \mathbf{b} \in \mathcal{B}, && (1.1a) \\ & (\mathbf{f}, \mathbf{b}) \in C, && (1.1b) \end{aligned}$$

where the *objective* is an extended-valued function $\mathcal{L} : \mathcal{H}_K \times \mathbb{R}^B \to \mathbb{R} \cup \{\infty\}$, the *bias variable* $\mathbf{b}$ has to belong to a nonempty closed convex set $\mathcal{B} \subseteq \mathbb{R}^B$, and C is a nonempty closed convex set of affine constraints over the pairs $(\mathbf{f}, \mathbf{b})$. In a machine learning perspective, the geometric intuition of the pairs $(\mathbf{f}, \mathbf{b})$ follows that of the classical support vector machines where $\mathbf{f}$ controls the direction, whereas $\mathbf{b}$ determines the bias of the hyperplane. This constant bias $\mathbf{b}$ to be estimated can be interpreted as an auxiliary degree of freedom, which can be removed by setting for instance $\mathcal{B} = \{0\}$. It is not considered in the chapters of this thesis pertaining to control theory.

Notations: Since this thesis lies at the confluence between two different fields - machine learning and control theory - it had to draw its notations from two traditions difficult to harmonize. While we removed notation conflicts in this introduction, they may persist between chapters, since the state is usually written as $\mathbf{x}(\cdot)$ in control theory, whereas the sample points in $\mathcal{X}$ are denoted by $\mathbf{x}$ in machine learning. Below we write $[\![n_1, n_2]\!]$ for the set of integers between $n_1, n_2 \in \mathbb{N}$ and use the shorthand $[N] := [\![1, N]\!] = \{1, \dots, N\}$ with $N \in \mathbb{N}$ and the convention that $[0]$ is the empty set. We denote by $\mathbb{R}_+$ the set of nonnegative reals. For a matrix $\mathbf{A} \in \mathbb{R}^{Q \times Q}$, we denote by $\|\mathbf{A}\|$ its operator norm, while Id_Q is the identity matrix of $\mathbb{R}^{Q \times Q}$. The set $\mathbb{B}$ generically designates the closed unit ball, a subscript indicating which set and topology are considered. We chose to keep the output space for the function spaces implicit, in order to avoid cumbersome notations, as it can be always deduced from the context. The space of functions with continuous derivatives up to order $s \in \mathbb{N}$ is denoted by $\mathcal{C}^s$.

In all that follows, given a set $\mathcal{X}$, a *kernel* is a function from $\mathcal{X} \times \mathcal{X}$ to $\mathbb{R}$ (denoted k if real-valued) or $\mathbb{R}^{Q \times Q}$ (denoted K if matrix-valued). The term *trajectory* designates the function $t \mapsto \mathbf{x}(t)$ defining the state. The term *solution* of an optimization problem always refers to an element of the search space achieving the optimal value. We say that we *discretize* a constraint involving a non-finite set $\mathcal{K}$ if we enforce the constraint only on a finite subset of $\mathcal{K}$.

In particular, we formulate our problem in the empirical risk minimization framework. Assume that we have access to some fixed family of samples $S := \{(\mathbf{x}_n, \mathbf{y}_n)\}_{n\in[N]} \subset \mathcal{X} \times \mathbb{R}^Q$. The quality of the estimated pair $(\mathbf{f}, \mathbf{b}) \in \mathcal{H}_K \times \mathbb{R}^B$ w.r.t. S is assessed through a *loss function* $L : \mathbb{R}^B \times (\mathbb{R}^Q)^N \to \mathbb{R} \cup \{\infty\}$. We also consider a *regularizer function* $R : \mathbb{R}_+ \to \mathbb{R} \cup \{\infty\}$ penalizing the model. This prevents, in machine learning terms, "overfitting" and, in control terms, an excessive "cost-to-go". With these notations, our considered *objective function* to be minimized is

$$\mathcal{L}(\mathbf{f}, \mathbf{b}) = L\left(\mathbf{b}, (\mathbf{f}(\mathbf{x}_n))_{n\in[N]}\right) + R\left(\|\mathbf{f}\|_K\right), \tag{1.2}$$

and, for a given $I \in \mathbb{N}$, the affine real-valued state/shape constraints on some given compacts sets $\mathcal{K}_i \subseteq \mathcal{X}$ take the form

$$C = \left\{(\mathbf{f}, \mathbf{b}) \,|\, 0 \leqslant D_i(\mathbf{f} - \mathbf{f}_{0,i})(\mathbf{x}) + \boldsymbol{\beta}_i^\top \mathbf{b} - b_{0,i}, \forall \mathbf{x} \in \mathcal{K}_i, \forall i \in [I]\right\}, \tag{$\mathcal{C}$}$$

where the differential operators $(D_i)_{i\in[I]}$ are defined over $\mathcal{H}_K$ and are such that $D_i = \sum_j \gamma_{i,j} \partial^{\mathbf{r}_{i,j}}$ where $\partial^{\mathbf{r}} f(\mathbf{x}) = \frac{\partial^{\sum_{j=1}^d r_j} f(\mathbf{x})}{\partial x_1^{r_1} \cdots \partial x_d^{r_d}}$ for the multi-index $\mathbf{r}$, e.g. $D_i = \boldsymbol{e}_1^\top(\partial_{\boldsymbol{e}_2} - \partial_{\boldsymbol{e}_1})$ or $D_i = \boldsymbol{e}_1^\top \mathrm{Id}_{\mathcal{H}_K}$ for $\mathbf{r} = \mathbf{0}$. Possible shifts in ($\mathcal{C}$) are expressed by the terms $b_{0,i} \in \mathbb{R}$ and $\mathbf{f}_{0,i} \in \mathcal{H}_K$. The vectors $\boldsymbol{\beta}_i \in \mathbb{R}^B$ describe linear interactions between the bias coordinates. The bias $\mathbf{b} \in \mathbb{R}^B$ can be both variable (e.g. $f_q + b_q$) or constraint-related (for instance, $b_1 \leqslant f(x)$, $b_2 \leqslant \partial_{\boldsymbol{e}_1} f(x)$); hence the dimension B of the bias can differ in general from the dimension Q of the output.

In nonparametric estimation, the constraints ($\mathcal{C}$) are known in the literature as a form of *shape constraints* since they can force the shape of $\mathbf{f}$ to be nonnegative or monotonous component-wise. These constraints typically occur in statistics for density estimation or isotonic regression.[1] In control theory, within the formalism considered in this thesis, ($\mathcal{C}$) correspond to the so-called *pure state constraints*.[2] While the empirical risk minimization framework ($\mathcal{P}_{\mathrm{ML}}$) with (1.2) is ubiquitous in machine learning, it may seem atypical in control theory. We thus underline right away the connection between

1. The word *shape* does not refer to the field of *shape optimization* where the optimized variable would be a set. Note that the requirement of convex components f_q for $d > 1$ would result in a positive semidefinite (SDP) constraint over the Hessian of f_q (see Aubin-Frankowski & Szabó, 2020a, for the extension of our approach to the SDP setting).

2. Control constraints are not considered because we require the continuity of the function $D_i(\mathbf{f} - \mathbf{f}_{0,i})$ to perform our analysis. In other words, constraints involving the control are not regular enough for the kernel formalism that we will investigate, since they bear on a variable $\mathbf{u}(\cdot) \in L^2(0, T)$ which is not continuous. This is also seen from the fact that, unlike "pure state constraints", constraints involving the control do not hold at all times, but only almost everywhere.

($\mathcal{P}_{ML}$) and state-constrained Linear-Quadratic (LQ) optimal control problems. In its simplest form, the problem of time-varying LQ optimal control with finite horizon and affine inequality state constraints writes as

$$\begin{aligned}
\min_{z(\cdot),\mathbf{u}(\cdot)} \quad & g(z(T)) + z(t_{ref})^\top \mathbf{J}_{ref} z(t_{ref}) + \int_{t_0}^{T} \left[z(t)^\top \tilde{\mathbf{Q}}(t) z(t) + \mathbf{u}(t)^\top \tilde{\mathbf{R}}(t) \mathbf{u}(t) \right] dt \\
\text{s.t.} \quad & \\
& z(t_0) = z_0, \\
& z'(t) = \mathbf{A}(t) z(t) + \mathbf{B}(t) \mathbf{u}(t), \text{ a.e. in } [t_0, T], \\
& \boldsymbol{\gamma}_i(t)^\top z(t) \leqslant b_{i,0}(t), \ \forall\, t \in [t_0, T], \forall\, i \in [I],
\end{aligned} \tag{$\mathcal{P}_{LQ}$}$$

where the state $z(t) \in \mathbb{R}^Q$, the control $\mathbf{u}(t) \in \mathbb{R}^P$, $\mathbf{A}(t) \in \mathbb{R}^{Q \times Q}$, $\mathbf{B}(t) \in \mathbb{R}^{Q \times P}$, $\boldsymbol{\gamma}_i(t) \in \mathbb{R}^Q$, $b_{i,0}(t) \in \mathbb{R}$, while $\mathbf{J}_{ref} \in \mathbb{R}^{Q \times Q}$, $\tilde{\mathbf{Q}}(t) \in \mathbb{R}^{Q \times Q}$ and $\tilde{\mathbf{R}}(t) \in \mathbb{R}^{P,P}$ are positive semidefinite matrices. We shall henceforth assume that, for some $r > 0$, for all $t \in [t_0, T]$, $\tilde{\mathbf{R}}(t) \succcurlyeq r \mathrm{Id}_M$, in the sense of positive semidefinite matrices. We assume as well that $\mathbf{A}(\cdot) \in L^1([t_0, T])$, $\mathbf{B}(\cdot) \in L^2([t_0, T])$, $\mathbf{Q}(\cdot) \in L^1([t_0, T])$, and $\mathbf{R}(\cdot) \in L^2([t_0, T])$. To have a finite objective, we restrict our attention to measurable controls satisfying $\mathbf{R}(\cdot)^{1/2} \mathbf{u}(\cdot) \in L^2([t_0, T])$. The terminal cost $g : \mathbb{R}^Q \to R \cup \{\infty\}$ is an extended-valued lower-semicontinuous function. We introduced a reference time t_{ref} and a positive semidefinite matrix $\mathbf{J}_{ref} \in \mathbb{R}^{Q \times Q}$ in order to define the LQ kernel later on.

We consider the following dictionary to highlight that the formalism of ($\mathcal{P}_{ML}$) is rich enough to encompass ($\mathcal{P}_{LQ}$):

— a time point $t \in [t_0, T]$ corresponds to a point $x \in \mathcal{X}$;

— a linearly-controlled trajectory $z(\cdot)$ corresponds to a model $f : x \mapsto f(x) \in \mathbb{R}^Q$;

— the linearly-controlled trajectories define a Hilbert space of functions $\mathcal{H}_K$, equipped with an inner product derived from the quadratic "running cost", i.e. the integral term of ($\mathcal{P}_{LQ}$);

— the initial ($z(t_0) = z_0$), terminal ($g(z(T))$), and any intermediate pointwise costs, even when expressed as indicator functions, can be included in the *loss function* L;

— the sum of the quadratic "running cost" and of the added term in $\mathbf{J}_{ref}$ correspond to the regularizer R that is equal to the squared Hilbert norm $\|z(\cdot)\|^2_{\mathcal{H}_K}$;

— the affine state constraints correspond to the set C of affine constraints, when setting $\mathcal{K}_i = [t_0, T]$. Since the "full state" $z(t)$ is a vector already containing all the derivatives of interest, the differential operators are simply $D_i = \boldsymbol{\gamma}_i^\top \mathrm{Id}_{\mathcal{H}_K}$. The extra constant biases $f_{0,i}$ and $b_{0,i}$ allow to define the boundaries $b_{i,0}(t)$ of the constraints;

— the extra variable $\mathbf{b} \in \mathcal{B}$ in ($\mathcal{P}_{\mathrm{ML}}$) would allow to describe some uncertainty over the definition of the state constraints; we here dismiss it, taking $\mathcal{B} = \{\mathbf{0}\}$, this removes any notation conflict between $\mathbf{B}(t) \in \mathbb{R}^{Q \times P}$ and $B \in \mathbb{N}$.

As straightforward as it may seem, to the author's best knowledge, this connection between ($\mathcal{P}_{\mathrm{ML}}$) and ($\mathcal{P}_{\mathrm{LQ}}$) had not been noticed before Aubin-Frankowski (2021b). This link between kernel methods and LQ optimal control allows to deal with pure state constraints as a form of shape constraints over vector-valued RKHSs, leveraging generic kernel methods principles rather than control-based results for state-constrained systems.

The fundamental challenge one faces when trying to solve ($\mathcal{P}_{\mathrm{ML}}$) is that in most relevant cases, such as ($\mathcal{P}_{\mathrm{LQ}}$), the sets $\mathcal{K}_i$ have non-finite cardinality, and hence there is an infinite number of constraints to satisfy. For instance, in ($\mathcal{P}_{\mathrm{LQ}}$), the sets $\mathcal{K}_i$ correspond to time intervals and, in applications to traffic control, the goal could be to avoid collisions or to respect a speed limit at all times. Since optimization with an infinite number of constraints requests to reduce the problem to a finite number of constraints to make it computationally tractable, one has to either relax or tighten the problem. *Relaxing* corresponds to approaches for which the constraint ($\mathcal{C}$) is not guaranteed to be satisfied at all points of $\mathcal{K}_i$. *Tightening* approaches on the contrary guarantee satisfaction of the constraint over $\mathcal{K}_i$ or a superset. This defines a dichotomy, rooted in optimization principles, to summarize the prior literature, either in nonparametric estimation, for ($\mathcal{P}_{\mathrm{ML}}$) in Section 1.1.2 below, or in LQ control theory, for ($\mathcal{P}_{\mathrm{LQ}}$) in Section 1.1.3 below. We will draw upon this dichotomy of relaxation as opposed to tightening in the next sections. Since we are interested in guaranteeing that the constraints are satisfied, all the contributions of this thesis revolve around a proposed tightening of the constraints, based on a finite compact covering in $\mathcal{H}_K$.

1.1.2 *Shape constraints in nonparametric regression*

We first review some aspects of shape constraints in nonparametric regression. They arise because in many applications one has to deal with a limited number of samples due to the difficulty or the cost of data acquisition. A well-established way to tackle this serious bottleneck and to improve sample-efficiency corresponds to incorporating qualitative priors on the shape of the model, such as non-negativity, monotonicity, convexity or supermodularity, collectively known as shape constraints (Guntuboyina & Sen, 2018). This side information can originate from both physical and theoretical constraints on the model such as "children usually grow" in growth charts or "be nonnegative and integrate to one" in density estimation. In the statistics community, the main emphasis has been on designing consistent estimators and on studying their

rates (Chen & Samworth, 2016; Deng & Zhang, 2020; Freyberger & Reeves, 2018; Han & Wellner, 2016; Kur et al., 2020; Lim, 2020). While these asymptotic results are of significant theoretical interest, imposing shape priors is generally beneficial in the small-sample regime. [3]

Various scientific fields, including econometrics, statistics, biology, game theory or finance, impose shape constraints on their hypothesis classes. For instance, economic theory dictates increasing and concave utility functions, decreasing demand functions, or monotone link functions (Chetverikov et al., 2018; Johnson & Jiang, 2018). In statistics, applying a monotonicity assumption on the regression function (for instance in isotonic regression; Han et al. 2019) dates back at least to Brunk (1955). Density estimation entails non-negativity which can be paired with other constraints (Royset & Wets, 2015), whereas, in quantile regression, conditional quantile functions grow w.r.t. the quantile level (Koenker, 2005). In biology, monotone regression is particularly well-suited to dose-response studies (Hu et al., 2005) and to identification of genome interactions (Luss et al., 2012). Inventory problems, game theory and pricing models commonly rely on the assumption of supermodularity (Simchi-Levi et al., 2014; Topkis, 1998). In financial applications, call option prices should be increasing in volatility, monotone and convex in the underlying stock price (Aït-Sahalia & Duarte, 2003).

We use the "relax vs tighten" dichotomy to stress the diversity of the shape constraints literature. In nonparametric estimation, *relaxing* can correspond to enforcing the constraint only at a finite number of points (Agrell, 2019; Blundell et al., 2012; Takeuchi et al., 2006) by replacing $\mathcal{K}$ with a discretization $\{\mathbf{x}_m\}_{m=1\dots M} \subsetneq \mathcal{K}$ in ($\mathcal{C}$). An alternative approach for relaxing is to add soft penalties to the objective $\mathcal{L}(\mathbf{f})$ (Brault et al., 2019; Koppel et al., 2019; Sangnier et al., 2016). *Tightening* on the contrary restricts the search space of functions $\mathcal{F}$ ($\mathcal{F} = \mathcal{H}_K$ in this thesis) to a smaller and more amenable subset $\mathcal{F}_0 \subsetneq \mathcal{F}$. This principle can be implemented by encoding the requirement ($\mathcal{C}$) into $\mathcal{F}_0$ through algebraic techniques. This approach is feasible for restrictive finite-dimensional $\mathcal{F}_0$ such as subsets of polynomials (Hall, 2018) or polynomial splines (Meyer, 2018; Papp & Alizadeh, 2014; Pya & Wood, 2015; Turlach, 2005; Wu & Sickles,

3. Note that, in statistics, shape constraints are not always stated as in ($\mathcal{P}_{ML}$). Rather than considering a large class of functions such as $\mathcal{H}_K$, sometimes the optimization is taken over the values $f(\mathbf{x}_i)$ under a constraint (such as monotonicity or convexity). The model is then obtained by interpolating between the values in some very specific classes of functions, often taken to be piecewise affine or splines. These classes are chosen so that the constraints are straightforwardly satisfied outside of the samples used for the interpolation. This approach often entails a highly constraint-specific analysis limited to dimension 1 in both input and output. We should also state another *caveat*: in this statistical literature, the word *kernel* can also designate a choice of weights when performing weighted averaging, based on a non-negative kernel ($k(x, y) \geqslant 0$), as for the Nadaraya-Watson estimator. We refer to https://francisbach.com/cursed-kernels/ for a disambiguation between non-negative and positive definite kernels.

2018). For infinite-dimensional sets $\mathcal{F}_0$, the idea has recently been extended elegantly by Marteau-Ferey et al. (2020) for $\mathcal{F}_0 = \mathcal{H}_k$ and a single non-negativity constraint over the whole space ($\mathcal{K} = \mathbb{R}^d$), forcing f to be a kernel analogue of a sum-of-squares. These limitations motivate the design of novel shape-constrained optimization techniques which avoid (i) restricted function classes, (ii) limited out-of-sample guarantees and (iii) lack of modularity in terms of the shape constraints imposed. Designing such a technique is the main objective of this thesis concerning its contributions to shape-constrained nonparametric regression.

1.1.3 *State constraints in LQ optimal control*

A comprehensive review of state constraints in general nonlinear optimal control can be found in Hartl et al., 1995. Since we focus on Hilbert spaces, we restrict our attention to the linear-quadratic context which is the subfield that was arguably most explored in optimal control. Indeed, without state constraints, and under mild assumptions, the unconstrained Linear Quadratic Regulator (LQR) enjoys an explicit solution defined through the Hamiltonian system and the related Riccati equation (see e.g. Speyer & Jacobson, 2010, for a thorough introduction and the bibliography therein). However, with state constraints, little can be said as Pontryagin's Maximum Principle involves not only an adjoint vector but also measures supported on the constraint boundary (Hartl et al., 1995). One has thus to guess beforehand when the state-constraint is active (at the so-called junction times) in order to write the first-order necessary condition (Hermant, 2009). Secondly, to derive mathematical statements such as the PMP, one needs some very restrictive assumptions. This has proven to be arduous and made state-constrained continuous-time optimization a difficult problem. Let us recall the well-known intuition for the appearance of discontinuities. If one follows an optimal trajectory of the LQR starting in the interior of state constraints, one may reach the boundary while the unconstrained Hamiltonian system of the Maximum Principle may incite to use a control leading to violation of the constraint. One has then to apply a different control to remain in the constraint set, possibly generating a discontinuity in the adjoint vector.

Despite its first appearance in the mid-50s, research is still active in the field of LQ optimal control, not only because of its numerous applications (see e.g. the examples of Burachik et al. (2014) and references within), but also for its theoretical aspects, even without constraints for approximation schemes (Bourdin & Trélat, 2017), or just control constraints (Burachik et al., 2014). Many of these improvements are motivated by model predictive control (e.g. Mayne et al., 2000, for a review), considered for instance in a time-invariant discrete-time state-constrained setting in Grüne and Guglielmi

(2018) or continuous-time (van Keulen, 2020). In particular, Kojima and Morari (2004) proved that the solutions of a time-invariant LQR with discretized constraints converge to the solution of ($\mathcal{P}_{\mathrm{LQ}}$), putting emphasis on function spaces of controls. As a matter of fact, the aforementioned approaches focus on the control, used to obtain the trajectories, while, in our approach, trajectories are instead at the core of the analysis.

As discussed above, when seeking a continuous-time numerical solution, one has to face an infinite number of pointwise constraints, and has to either relax the computationally intractable optimization problem or tighten it. In control theory, *relaxing* means either enforcing the constraint only at a finite number of points, without guarantees elsewhere (as with any time discretization method, e.g. Kojima & Morari, 2004), or through soft penalties (Gerdts & Hüpping, 2012), such as approximations of barrier functions (Dower et al., 2019). *Tightening* usually implies either choosing $\mathbf{u}(\cdot)$ in a convenient subspace of $L^2([t_0, T])$, for instance the one of piecewise constant functions [4] or of splines with prescribed knots (Mercy et al., 2016), or through hard penalties, such as the so-called interior penalty methods "where constraints are penalized in a way that guarantees the strict interiority of the approaching solutions" (Chaplais et al., 2011; Malisani et al., 2014). Let us illustrate the difference between relaxing and tightening state constraints. Consider the problem of a traffic regulator whose aim is to enforce a speed limit over a highway. The drivers for their part want to go as far as possible in a given time. Deploying speed cameras ensures at best that the speed constraint is satisfied locally (relaxing). However if a smaller maximum speed, defined by changing the threshold, is imposed at the camera level (tightening), then the cars cannot accelerate enough to break the speed limit before reaching the next camera. In a nutshell, the kernel methods framework we advocate allows to compute both a threshold and the resulting trajectories. Providing feasible interior solutions and proving their convergence to the solution of the original problem is the main objective of this thesis concerning control theory. This is achieved based on the theory of reproducing kernels.

1.2 REPRODUCING KERNELS IN A NUTSHELL

Kernel methods are a wide and mature mathematical field, deeply rooted in functional analysis, with ramifications across many other domains. They can thus be introduced from several perspectives. We shall in this thesis favor the two historical viewpoints, either putting forward the kernel or the function space, as presented by

4. This is known as *sampled-data* or *digital control* (Ackermann, 1985), the *sampled-data* terminology does not refer to machine learning techniques.

Aronszajn (1950). We thus do not cover a third now classical approach describing kernels as covariance functions of Gaussian processes.

For the reader curious to delve deeper into the kernel jungle, we recommend several monographs in connection with the following topics: integral operators theory and partial differential equations (Saitoh, 1997; Saitoh & Sawano, 2016), Gaussian processes in probabilities (Berlinet & Thomas-Agnan, 2004; Kanagawa et al., 2018), feature maps and applications in machine learning (Schölkopf & Smola, 2002), general support vector machines (Steinwart & Christmann, 2008), generating functions for complex variables (Alpay, 2001), and harmonic analysis on semigroups (Berg et al., 1984). Although kernels have known several extensions beyond the Hilbertian case, we shall not use them in this thesis. For some more exotic kernels, we refer to Mary (2003) and Ong et al. (2004) for Krein and Pontryagine's kernels and to Lin et al. (2019) for Banach kernels.

Section 1.2.1 summarizes the history of reproducing kernels as divided between kernel-driven (Section 1.2.2) and operator-driven (Section 1.2.3) approaches. The extension of the reproducing property to derivatives and to vector-valued outputs is covered in Section 1.2.4. The propitious representer theorems that we shall apply in most of our analysis are stated in Section 1.2.5.

1.2.1 *History of reproducing kernels: A realm (still) divided*

From a mathematical standpoint, we can summarize the century-long history of kernels as a succession of theoretical breakthroughs and recessions due to the lack of novel insights. The theory itself was reborn multiple times around 1907 and 1921. As Aronszajn (1950) recounts it, two tendencies prevailed, either putting the function space forward and considering the kernel as a tool to be computed (Zaremba, Szegö, Bergman), or studying the kernel for its properties and then deriving the function space as a by-product (Mercer, Moore). The unifying effort of mathematicians (Aronszajn, 1943, 1950; Schwartz, 1964) more acquainted with modern Bourbaki structures led to the elegant theory as it is still presented today. Meanwhile, Emanuel Parzen in a Promethean gesture [5] brought the torch of kernels to statisticians and passed it at Stanford, notably to Grace Wahba and Thomas Kailath, *in fine* extending the theoretical results to engineering studies.

So far, we have not mentioned machine learning. Are least-squares problems the first example of it? It is obviously hard to draw a line between communities, but machine learning has the peculiarity with respect the aforementioned fields to have at

5. This is charmingly recounted by Persi Diaconis in the preface of Berlinet and Thomas-Agnan (2004).

its core the credo of performance. Since we are sketching the history, let us mention some of the multiple reasons which could be invoked to explain the emergence of machine learning in the early 90's: the economic boom of the private sector after the Cold War, the access to personal computers, the emigration of scientists from Eastern Europe,... To say the least, Vladimir Vapnik, one of the popes of statistical learning, joined the Bell Labs in the early 90s and Support Vector Machines (SVMs) peaked in the early 2000s. But the applicative power reached a glass ceiling as the cubic complexity of SVMs could deal with millions of points but not the billions that the industry now required. Dwelling in the shadow of neural networks, some kernel specialists look either to bypass the complexity issue through kernel approximations or to reassert the importance of kernels as a theoretical limit of neural networks.

Conceptually, the dichotomy individuated by Aronszajn is still valid. *Kernel-driven* approaches where the kernel is taken off-the-shelf have inherited Mercer's viewpoint, whereas *operator-driven* approaches still require the kernel to be computed for a specific task. This illustrates the widening gap in the literature between, on the one hand, Schölkopf and Smola (2002), and, on the other hand, Berlinet and Thomas-Agnan (2004) and Saitoh and Sawano (2016). The two latter books still put much emphasis on the role of Sobolev spaces and of derivatives for their applications (covariance of processes, ODEs, PDEs), with $\mathcal{X}$ often taken as a subset of the real line. Schölkopf and Smola (2002) on the other hand take the kernels for given and use them on various input sets (graphs, probability measures), essentially extending linear algorithms to nonlinear settings. The flexibility of the kernel-driven approach and its simplicity of exposition made it prevalent in courses on the topic, while the operator-driven viewpoint was kept up in the statistics community and in a smaller mathematical one. In this thesis, we shall leverage both the kernel-driven approach, when facing agnostic machine learning problems where all that is required is a function satisfying shape constraints, and the operator-driven viewpoint, for control problems where we have to identify the relevant kernel.

1.2.2 *Positive definite kernels: A nonlinear embedding with inner products*

The kernel-driven approach that is prevalent in machine learning starts with positive definite kernels.

Definition 1.1 (Positive definite kernel)**.** Let $\mathcal{X}$ be a nonempty set. A real-valued kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is positive definite if it is

— symmetrical: $\forall x, x' \in \mathcal{X},\ k(x, x') = k(x', x)$, and

— positive: $\forall\, \mathbf{a} \in \mathbb{R}^N, (x_i)_{i\in[N]} \in \mathcal{X}^N,\ \sum_{i=1}^N \sum_{j=1}^N a_i a_j k(x_i, x_j) \geqslant 0$.

For the Gram matrix $\mathbf{G}$ related to the points $(x_i)_{i\in[N]}$ for the kernel k, positivity writes as

$$\forall\, (x_i)_{i\in[N]} \in \mathcal{X}^N,\ \mathbf{G} := [k(x_i, x_j)]_{i,j\in[N]}, \qquad \forall\, \mathbf{a} \in \mathbb{R}^N,\ \mathbf{a}^\top \mathbf{G}\mathbf{a} \geqslant 0.$$

There is thus a slight abuse w.r.t. the Anglo-Saxon terminology since the positive definite kernels that we will consider are associated with positive semidefinite matrices (i.e. "positive" for non-English speakers). This choice is a predominant trend in the kernel community.

Any inner product on a Hilbert space $\mathcal{X}$ is positive definite. Some less trivial classical examples of positive definite kernels on $\mathbb{R}^Q$ include the Gaussian and Laplacian kernels for bandwidths $\sigma > 0$ and $\lambda \geqslant 0$,

$$k_\sigma(x, y) = \exp\left(-\|x - y\|^2_{\mathbb{R}^Q}/(2\sigma^2)\right), \quad k_\lambda(x, y) = \exp(-\lambda\|x - y\|_{\mathbb{R}^Q}).$$

To prove their positive definiteness, recall that the sum of positive definite kernels, their product and their scaling by a nonnegative scalar are all positive definite kernels as well. As highlighted by Cuturi (2005, p. 5), this kernel-driven matrix viewpoint on kernels "often challenges the functional viewpoint itself, notably in the semi-supervised setting where defining a kernel matrix on a dataset is sufficient to use kernel methods". At its origins, "kernel machines [were] black boxes", a qualification now only used for neural networks. This stemmed from the apparently elaborate relation with feature maps:

Theorem 1.1 (Theorem 4, Berlinet and Thomas-Agnan, 2004)**.** *A kernel k is positive definite if and only if there exists a set $\mathcal{T}$ and an embedding $\Phi : \mathcal{X} \to l^2(\mathcal{T})$, the set of real sequences $\{u_t, t \in \mathcal{T}\}$ such that $\sum_{t\in\mathcal{T}} |u_t|^2 < \infty$, where*

$$\forall x, x' \in \mathcal{X} \times \mathcal{X},\ k(x, x') = \sum_{t\in\mathcal{T}} \Phi(x)_t \Phi(x')_t = \langle \Phi(x), \Phi(x')\rangle_{l^2(\mathcal{T})}.$$

In other words, setting $\mathcal{H} := l^2(\mathcal{T})$, we have $\forall\, x, x' \in \mathcal{X},\ k(x, x') = \langle \Phi(x), \Phi(x')\rangle_{\mathcal{H}}$.

This nonlinear embedding Φ was for long considered as the main vessel for SVMs (Schölkopf & Smola, 2002) and SVMs are still interpreted as reformulating a nonlinear classification problem in the input space $\mathcal{X}$ into a linear separation problem in $\mathcal{H}$.[6]

6. "The feature map formulation, particularly advocated in the early days of SVMs, also misled some observers into thinking that the kernel mapping was but a piece of the SVM machinery. Instead, SVM should be rather seen as an efficient computational approach – among many others – deployed to select a "good" function f in the RKHS." (Cuturi, 2005, p. 7)

However, in Theorem 1.1, $\mathcal{H}$ or Φ have no reason to be unique due to isometry properties of Hilbert spaces w.r.t. l^2. Moreover, the proof of Theorem 1.1 actually relies on the key embedding $x \mapsto k(x, \cdot)$ which favors a specific choice of $\mathcal{H}$ as a Hilbert space of functions over $\mathcal{X}$.

Remark 1.1 (Separability and reproducing kernels)**.** Positive definite kernels as defined in Definition 1.1 do not assume separability of $\mathcal{X}$. A technical necessary and sufficient condition for the existence of a separable $l^2(\mathcal{T})$ in Theorem 1.1, for which $\mathcal{T}$ is countable, and can thus be taken equal to $\mathbb{N}$, can be found in (Fortet, 1995, p. 142) in connection with the Karhunen-Loève representation of stochastic processes. An easier sufficient criterion is that $\mathcal{X}$ is a separable topological space and that k is continuous over $\mathcal{X} \times \mathcal{X}$ (Steinwart & Christmann, 2008, Lemma 4.33), this criterion is always satisfied in this thesis since the considered kernels are continuous and $\mathcal{X} \subseteq \mathbb{R}^d$.

1.2.3 *RKHSs: A Hilbertian topology stronger than pointwise convergence*

The operator-driven approach focuses on a special class of Hilbert space of functions, those that have a topology stronger than pointwise convergence. This makes them well-suited as a *hypothesis class* in optimization problems involving pointwise evaluations of the model.

Definition 1.2 (Real-valued RKHS-1)**.** A real-valued reproducing kernel Hilbert space (RKHS) $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ defined on a nonempty set $\mathcal{X}$ is a Hilbert space of real-valued functions over $\mathcal{X}$ such that the Hilbertian topology is stronger than pointwise convergence, i.e. for all $x \in \mathcal{X}$, the evalutation at x, $\delta_{x,\mathcal{H}} : f \in \mathcal{H} \mapsto f(x)$, is continuous over $\mathcal{H}$ equipped with its Hilbertian topology.

By Riesz's theorem, one immediately derives that, since $\delta_{x,\mathcal{H}}$ is a continuous linear form over $\mathcal{H}$, then one can fix a unique $k_x \in \mathcal{H}$ such that $f(x) = \delta_{x,\mathcal{H}}(f) = \langle f(\cdot), k_x(\cdot) \rangle_{\mathcal{H}}$. This leads to the classical, and equivalent, definition of RKHSs:

Definition 1.3 (Real-valued RKHS-2)**.** A real-valued RKHS $\left(\mathcal{H}_k, \langle \cdot, \cdot \rangle_{\mathcal{H}_k}\right)$ over a nonempty set $\mathcal{X}$ is a Hilbert space of real-valued functions over $\mathcal{X}$ such that there exists a reproducing kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, i.e. a function satisfying:

— $\forall x \in \mathcal{X},\ k_x(\cdot) \in \mathcal{H}_k$ where $k_x : \begin{cases} \mathcal{X} \to \mathbb{R} \\ y \mapsto k(x, y) \end{cases}$,

— $\forall x \in \mathcal{X},\ \forall f \in \mathcal{H}_k,\ f(x) = \langle f(\cdot), k_x(\cdot) \rangle_{\mathcal{H}_k}$.

Let us consider an immediate example of RKHS. For $s \in \mathbb{N}$ and $\mathcal{X}$ an open set of $\mathbb{R}^d$ (or closure of an open set), setting $\partial^{\mathbf{r}} f(\mathbf{x}) = \frac{\partial^{|\mathbf{r}|} f(\mathbf{x})}{\partial x_1^{r_1} \dots \partial x_d^{r_d}}$ for $|\mathbf{r}| := \sum_{i \in [d]} r_i$, recall the following definition of a Sobolev space:

$$H^s(\mathcal{X}, \mathbb{R}) = W^{s,2}(\mathcal{X}, \mathbb{R}) := \left\{ f \in L^2(\mathcal{X}) \,|\, \partial^{\mathbf{r}} f \in L^2(\mathcal{X}),\ \forall \mathbf{r} \in \mathbb{N}^d, \text{ s.t. } |\mathbf{r}| \leqslant s \right\}$$

equipped with the inner product $\langle f, g \rangle_{H^s} = \sum_{|\mathbf{r}| \leqslant s} \langle \partial^{\mathbf{r}} f, \partial^{\mathbf{r}} g \rangle_{L^2(\mathcal{X})}$. Any Sobolev space $H^s(\mathcal{X}, \mathbb{R})$ satisfying the Sobolev inequality $s > d/2$ and equipped with its inner product is a RKHS. On the contrary, for $s \leqslant d/2$, the Hilbertian topology of $H^s(\mathcal{X}, \mathbb{R})$ is too weak to consider pointwise evaluations. Aronszajn's great achievement was to notice that Definitions 1.1, 1.2 and 1.3 coincided and that the embedding $\Phi_k(x) = k_x$ played a specific role:

Theorem 1.2 (Aronszajn, 1950)**.** *If a Hilbert space of functions over a nonempty set $\mathcal{X}$ is a RKHS $\mathcal{H}_k$ with reproducing kernel k, then k is a positive definite kernel. Conversely a positive definite kernel k over $\mathcal{X}$ is reproducing for a unique $\mathcal{H}_k$. Moreover $\mathcal{H}_k$ is the completion of the pre-Hilbert space $\mathcal{H}_{k,0}(\mathcal{X}) := \operatorname{span}(\{k_x(\cdot)\}_{x \in \mathcal{X}})$ for the pre-inner product $\langle k_x, k_y \rangle_{k,0} := k(x, y)$.*

There is thus a one-to-one correspondence between positive definite kernels k and RKHSs $\left(\mathcal{H}_k, \langle \cdot, \cdot \rangle_{\mathcal{H}_k}\right)$. Notice however that it implies that changing $\mathcal{X}$ or $\langle \cdot, \cdot \rangle_{\mathcal{H}_k}$ changes the kernel k. It is in general very hard to identify, given a RKHS $\mathcal{H}_k$ (resp. a kernel k), its kernel k (resp. its RKHS $\mathcal{H}_k$). We illustrate this problem with some examples from Saitoh and Sawano, 2016, p. 11-15, for $\mathcal{X} = \mathbb{R}_+$, where slight differences noticeably change the kernel:

— the Sobolev space H^1 with inner product $\langle f, g \rangle_{H^1} = \frac{2}{\pi} \int_0^\infty f(t)g(t) + f'(t)g'(t)dt$, has for kernel $k(x, y) = \frac{\pi}{4} \left(\exp(-|x - y|) + \exp(-x - y)\right)$;

— the Sobolev space H_0^1 of functions of H^1 such that $f(0) = 0$ with inner product $\langle \cdot, \cdot \rangle_{H^1}$, has for kernel $k(x, y) = \frac{\pi}{4} \left(\exp(-|x - y|) - \exp(-x - y)\right)$;

— the Sobolev space H_0^1 with inner product $\langle f, g \rangle_{H_0^1} = \int_0^\infty f'(t)g'(t)dt$, has for kernel $k(x, y) = \min(x, y)$.

Remark 1.2 (Green functions and reproducing kernels)**.** Since we shall often deal with Sobolev-like spaces, we recall some elements on the conceptual relation with Green functions, following the presentation of Saitoh and Sawano, 2016, p. 60-61. Let D be a differential operator over an open set $\mathcal{X} \subset \mathbb{R}^d$, and set D^* to be its adjoint operator. Define the Green function $G_{D^*D,x} : \mathcal{X} \to \mathbb{R}$ s.t. $D^*D\left(G_{D^*D,x}\right)(y) = \delta_x(y)$ then, if the integrals over the boundaries in Green's formula are null, for any $f \in \mathcal{H}_k$,

$$f(x) = \int_{\mathcal{X}} f(y) D^*D(G_{D^*D,x})(y) dy = \int_{\mathcal{X}} Df(y) D(G_{D^*D,x})(y) =: \langle f, G_{D^*D,x} \rangle_{\mathcal{H}_k},$$

so $k(x,y) = G_{D^*D,x}(y)$. For vector-valued contexts (see Section 1.2.4 below), e.g. $\mathcal{H}_K = W^{s,2}(\mathbb{R}^d, \mathbb{R}^d)$ and $D^*D = (1-\sigma^2\Delta)^s$ component-wise, we refer to Micheli and Glaunés, 2014, p. 9. Alternatively, for $d = 1$, and $G_{D,x}$ defined as $D\,(G_{D,x})(y) = \delta_x(y)$, the kernel associated with the inner product $\int_{\mathcal{X}} Df(y)Dg(y)dy$ for the space of f "null at the border" can be factorized into (see Berlinet and Thomas-Agnan (2004, p. 286) and Heckman (2012))

$$k(x,y) = \int_{\mathcal{X}} G_{D,x}(z)G_{D,y}(z)dz.$$

Aronszajn's formulation of kernels can be seen as a way to bypass Green functions by picking off-the-shelf kernels rather than solving the related PDEs. The latter would provide an explicit orthonormal basis of eigenvectors, whereas the kernel functions k_x form a non-orthonormal generating family better suited for pointwise evaluations. Indeed, applicability of reproducing kernels in machine learning mainly stems from implicit feature maps as discussed in Section 1.2.5.

1.2.4 *Vector-valued RKHSs and reproducing property for derivatives*

When real-valued outputs are not sufficient for the problem at hand, one should resort to vector-valued RKHSs (vRKHSs, see Caponnetto et al., 2008, and references within).[7] They are defined similarly to Definitions 1.1, 1.2 and 1.3. However the kernel is now matrix-valued, and we have to take 'covectors' $\mathbf{p}$ into account for pointwise evaluations:

Definition 1.4 (Vector-valued RKHS)**.** Let $\mathcal{T}$ be a nonempty set. A Hilbert space $(\mathcal{H}_K, \langle\cdot,\cdot\rangle_K)$ of $\mathbb{R}^Q$-vector-valued functions defined on $\mathcal{T}$ is called a vector-valued reproducing kernel Hilbert space (vRKHS) if there exists a matrix-valued kernel $K : \mathcal{T}\times\mathcal{T} \to \mathbb{R}^{Q\times Q}$ such that the *reproducing property* holds: for all $t \in \mathcal{T}$, $\mathbf{p} \in \mathbb{R}^Q$, $K(\cdot,t)\mathbf{p} \in \mathcal{H}_K$ and for all $\mathbf{f} \in \mathcal{H}_K$, $\mathbf{p}^\top \mathbf{f}(t) = \langle \mathbf{f}, K(\cdot,t)\mathbf{p}\rangle_K$. By Riesz's theorem, it is equivalent to saying that, for every $t \in \mathcal{T}$ and $\mathbf{p} \in \mathbb{R}^Q$, the evaluation functional $\mathbf{f} \in \mathcal{H}_K \mapsto \mathbf{p}^\top \mathbf{f}(t) \in \mathbb{R}$ is continuous.

Necessarily, K has a Hermitian symmetry: for all $s,t \in \mathcal{T}$, $K(s,t) = K(t,s)^\top$. This requires more care when manipulating Gram matrices since they are symmetric only when taking into account the necessary transposition. There is again a one-to-one correspondence between K and $(\mathcal{H}_K, \langle\cdot,\cdot\rangle_K)$ (Micheli & Glaunés, 2014, Theorem 2.6), so, again, changing $\mathcal{T}$ or $\langle\cdot,\cdot\rangle_K$ changes K. For instance one can consider a time interval $\mathcal{T} = [0,1]$ and a vector-valued Sobolev space $H^s([0,1],\mathbb{R}^Q)$ for $s \in \mathbb{N}^*$.

7. Vector-valued RKHSs will be used in this thesis to describe trajectories, the components of which cannot be considered independently due to the dynamic constraints.

So far, we have not mentioned the regularity properties of the functions in $\mathcal{H}_K$. In general, they share the regularity of the kernel functions $K(\cdot, x)$ in terms of smoothness or analyticity (Saitoh & Sawano, 2016, p. 73-77). We even have a reproducing property for derivatives which resembles that of Dirac masses, when denoting by ∂_1 (resp. ∂_2) the derivative w.r.t. the first (resp. second) variable of $K(\cdot, \cdot)$.

Theorem 1.3 (Theorem 2.11, Micheli and Glaunés, 2014)**.** *Let $\mathcal{H}_K$ be a vRKHS with kernel $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{Q\times Q}$ for an open subset $\mathcal{X}$ of $\mathbb{R}^d$, and $s \geqslant 0$ be an integer. The following two statements are equivalent:*

- *$\mathcal{H}_K$ is continuously embedded in $\mathcal{C}^s\left(\mathcal{X}, \mathbb{R}^Q\right)$, i.e. the inclusion $\mathcal{H}_K \subseteq \mathcal{C}^s\left(\mathcal{X}, \mathbb{R}^Q\right)$ holds and the identity map from $\mathcal{H}_K$ to $\mathcal{C}^s\left(\mathcal{X}, \mathbb{R}^Q\right)$ is continuous,*
- *the function $\partial_1^r \partial_2^r K$ exists for all multi-indices r with length satisfying $0 \leqslant |r| \leqslant s$, it is continuous in each of the two variables (separately), and it is locally bounded.*

Under the above assumptions, the following also holds:

- *for all $x \in \mathcal{X}$, $p \in \mathbb{R}^Q$, and multi-index r such that $0 \leqslant |r| \leqslant s$, we have $\partial_2^r K(\cdot, x)p \in \mathcal{H}_K$ and*

$$p^\top \partial^r f(x) = \langle \partial_2^r K(\cdot, x)p, f\rangle_{\mathcal{H}_K}, \quad \forall f \in \mathcal{H}_K. \tag{1.3}$$

Remark 1.3 (Reproducing property for derivatives)**.** Specifically, for $s = 0$, one has that $D = p^\top \mathrm{Id}_{\mathcal{H}_K}$, so $D(f)(x) = p^\top f(x)$, and (1.3) reduces to the reproducing property in vRKHSs of Definition 1.4. For real-valued kernels, the reproducing property for kernel derivatives has been studied over open sets $\mathcal{X}$ (e.g. Saitoh & Sawano, 2016), and over compact sets which are the closure of their interior (Zhou, 2008). For practitioners, it cannot be stressed enough that one should apply the differential operator to the variable at which f is evaluated, i.e. if x appears in second position in K in (1.3) then one should use ∂_2. It was shown in Aubin-Frankowski and Szabó (2020a) that (1.3) also holds for, possibly unbounded, sets $\mathcal{X}$ which are contained in the closure of their interior. Note that the derivative $\partial^r f$ has to be continuous for the property (1.3) to hold. Even if functions in $\mathcal{H}_K$ have a derivative in a weak, almost everywhere, distribution sense, (1.3) cannot be applied, as there would be no Riesz's theorem to provide a kernel function $\partial_2^r K(\cdot, x)p \in \mathcal{H}_K$ to consider continuous pointwise evaluations of derivatives. [8]

1.2.5 *Two computational tools: the representer theorems and the kernel trick*

One important perk of kernel methods stems from the fact that, for problems of the form ($\mathcal{P}_{ML}$), with $\mathcal{L}$ defined as in (1.2), the solutions only depend on the observed

8. This explains why for control constraints, whenever the controls are merely L^2, one does not have a reproducing property, as already hinted at in footnote 2 (Page 4).

samples. This behavior is expressed through "representer theorems", which ensure that the solutions of an optimization problem like ($\mathcal{P}_{ML}$) live in a finite-dimensional subspace of $\mathcal{H}_K$ and consequently enjoy a finite representation. These "representer theorems" can be informally summarized as: "a finite number of evaluations implies a finite number of coefficients" or "all the information is contained in the samples". Indeed, their short proofs merely use the fact that any function v added to f, whenever v is orthogonal to the kernel functions $k(x_n, \cdot)$ of the samples x_n, does not change the value of f at these samples. However the norm of the extra v in $f + v$ is penalized by the regularizing function, so v has to be null for an optimal $\bar{f}$. In short, representer theorems are instrumental to obtain a finite-dimensional problem that is exactly equivalent to the infinite-dimensional ($\mathcal{P}_{ML}$).

Representer theorems can be found as early as in Kimeldorf and Wahba (1971, Theorems 3.1 and 5.1) for quadratic norm-penalties and exact or (quadratic) approximate interpolation. This quadratic formulation was prevalent in polynomial splines studies. In machine learning, the most classical formulation for the empirical risk minimization framework is expressed as

Theorem 1.4 (Theorem 2, Schölkopf et al., 2001)**.** *Suppose we are given a nonempty set $\mathcal{X}$, a positive definite real-valued kernel k on $\mathcal{X} \times \mathcal{X}$, samples $S := \{(x_n, y_n)\}_{n\in[\![N]\!]} \subset \mathcal{X} \times \mathbb{R}$, a strictly increasing real-valued function R on $\mathbb{R}_+$, an arbitrary cost function $L : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$, and a family of real-valued functions $\{\psi_m\}_{m\in[\![M]\!]}$ on $\mathcal{X}$ with the property that the $N \times M$ matrix $[\psi_m(x_n)]_{n,m}$ has rank M. Then any $g := f + h$, with $f \in \mathcal{H}_k$ and $h \in \operatorname{span}\{\psi_m\}_{m\in[\![M]\!]}$ minimizing the regularized risk functional $\mathcal{L}(g) = L((g(x_n))_{n\in[\![N]\!]}) + R(\|f\|)$ admits a representation of the form $g(\cdot) = \sum_{n\in[\![N]\!]} \alpha_n k(\cdot, x_n) + \sum_{m\in[\![M]\!]} \beta_m \psi_m(\cdot)$ with coefficients $\alpha_n \in \mathbb{R}$ and uniquely defined $\beta_m \in \mathbb{R}$.*

Remark 1.4 (Parametric families in representer theorems)**.** Why would one need a parametric family $\{\psi_m\}_{m\in[\![M]\!]}$? This comes from the spline background of representer theorems which were originally intended for functions $f \in H^M(\mathbb{R}, \mathbb{R})$, the M-th order Sobolev space, with regularizers of the form $R_M(f) = \int_{\mathbb{R}} |f^{(M)}(x)|^2 dx$. These regularizers R_M do not penalize polynomials of order strictly smaller than M. They are however equal to the squared norm of $H_0^M(\mathbb{R}, \mathbb{R})$, the subspace of $H^M(\mathbb{R}, \mathbb{R})$ of functions with derivatives vanishing at 0, i.e. $f^{(m)}(0) = 0$ for $m \in [\![0, M-1]\!]$. The existence of a "null-space" of the regularizer is a problem we will meet again in optimal control, where the uncontrolled trajectories may be unpenalized by the integral "running cost" of ($\mathcal{P}_{LQ}$).

Theorem 1.4 can be straightforwardly extended to vector-valued contexts and to problems with constraints, as in

Theorem 1.5 (Representer theorem for vRKHSs, Aubin-Frankowski, 2021b)**.** *Let* $(\mathcal{H}_K, \langle\cdot,\cdot\rangle_K)$ *be a vRKHS defined on a nonempty set* $\mathcal{T}$*. Let* $I \in \mathbb{N}$ *and, for* $i \in [\![0, I]\!]$ *and given* $N_i \in \mathbb{N}$, $\{t_{i,n}\}_{n\in[\![1,N_i]\!]} \subset \mathcal{T}$*. Consider the following optimization problem with "loss" function* $L : \mathbb{R}^{N_0} \to \mathbb{R} \cup \{+\infty\}$*, strictly increasing "regularizer" function* $R : \mathbb{R}_+ \to \mathbb{R}$*, and constraint descriptors* $d_i : \mathbb{R}^{N_i} \to \mathbb{R}$, $\lambda_i \geqslant 0$ *and* $\{\mathbf{c}_{i,n}\}_{n\in[\![1,N_i]\!]} \subset \mathbb{R}^Q$*,*

$$\begin{aligned}\bar{\mathbf{f}} \in \underset{\mathbf{f}\in\mathcal{H}_K}{\arg\min} \quad & L\left(\mathbf{c}_{0,1}^\top \mathbf{f}(t_{0,1}), \dots, \mathbf{c}_{0,N_0}^\top \mathbf{f}(t_{0,N_0})\right) + R\left(\|\mathbf{f}\|_K\right) \\ \text{s.t.} \quad & \\ & \lambda_i \|\mathbf{f}\|_K \leqslant d_i(\mathbf{c}_{i,1}^\top \mathbf{f}(t_{i,1}), \dots, \mathbf{c}_{i,N_i}^\top \mathbf{f}(t_{i,N_i})),\ \forall\, i \in [I].\end{aligned}$$

Then, for any minimizer $\bar{\mathbf{f}}$*, there exists* $\{\mathbf{p}_{i,n}\}_{n\in[\![1,N_i]\!]} \subset \mathbb{R}^Q$ *such that* $\bar{\mathbf{f}} = \sum_{i=0}^{I}\sum_{n=1}^{N_i} K(\cdot, t_{i,n})\mathbf{p}_{i,n}$ *with* $\mathbf{p}_{i,n} = \alpha_{i,n}\mathbf{c}_{i,n}$ *for some* $\alpha_{i,n} \in \mathbb{R}$*.*

Remark 1.5 (Representer theorems and the "kernel trick")**.** If one studies problems for which representer theorems exist, then this enables to use the so-called "kernel trick" which underpins most of kernel methods (Schölkopf & Smola, 2002). The latter is merely a remark that the reproducing property allows to write:

$$\langle \sum_{i=0}^{I} K(\cdot, t_i)\mathbf{p}_i, \sum_{j=0}^{J} K(\cdot, t'_j)\mathbf{p}'_j \rangle_K = \sum_{i=0}^{I}\sum_{j=0}^{J} \mathbf{p}_i^\top K(t_i, t'_j)\mathbf{p}'_j. \tag{1.4}$$

Hence one does not need to have an explicit form of $\langle\cdot,\cdot\rangle_K$ on the subspace of $\mathcal{H}_K$ appearing in Theorem 1.5 since, for linear combinations of the kernel functions $K(\cdot, t)\mathbf{p}$, (1.4) allows to compute every pairwise inner product. In this sense the "kernel trick" is at the core of the kernel-driven approach to kernel methods. This is yet one of the discrepancies between the kernel-driven and the operator-driven approach to kernels. As a matter of fact, it considerably restricts the scope of kernel methods to problems of the form ($\mathcal{P}_{ML}$). Indeed, under mild assumptions, only radial regularizers of the form $R(\|\mathbf{f}\|)$ allow for representer theorems (Dinuzzo & Schölkopf, 2012). For more details on this aspect, we refer to Argyriou and Dinuzzo (2014) and references within.

Remark 1.6 (Necessary condition and sparsity for inequalities)**.** Note that a representer theorem, similarly to Pontryagin's Maximum Principle, is only a necessary condition on the form of the solutions. Theorem 1.5 guarantees the existence or uniqueness of an optimal solution only when coupled with other assumptions (e.g. that L and R are lower semi-continuous and R is coercive). This further highlights the analogy with the Maximum Principle in the quadratic case. Besides, we would like to stress that the coefficient $\mathbf{p}_n$ associated with each $\mathbf{x}_n$ can be interpreted as a dual variable

(this is extensively used in Chapter 7). For linear pointwise inequality constraints (e.g. $f(x_n) \geqslant 0$), it is even exactly the Lagrange multiplier of the constraint. This implies that whenever such an inequality is not active, then the corresponding x_n does not appear in the expression of any optimal $\bar{f}$. Hence, for any optimization problem expressed mainly through inequality constraints, the solutions obtained through kernel methods can often be extremely sparsely encoded if most of the constraints are inactive. This is a well-known explanation of the success of SVMs for classification problems.

This sample-focused strength is however recurrently considered as their computational pitfall since kernel methods routinely require handling and inverting matrices, often as large as the number of samples appearing in the problem. In this thesis, we consider the extreme case of an infinite number of samples. This problem is novel since most kernel practitioners (Steinwart & Christmann, 2008) have focused on the objective functions rather than on the constraints. For instance, in least-squares regression, one may have a considerable amount of observed samples, but one seldom has an infinity of them. It is quite the opposite with constraints: a convex constraint set is more often described by an infinite number of inequalities rather than finitely many. The open question is then: how to capitalize on the simplicity of theorems such as Theorems 1.4 and 1.5 when facing an infinite number of constraints?

1.3 CONTRIBUTIONS OF THE THESIS

In this thesis, we shall leverage both the kernel-driven approach, when facing agnostic machine learning problems where all that is requested is to produce a function satisfying shape constraints, and the operator-driven viewpoint, for control problems where we have to identify the relevant kernel. We provide in this section a sketch and summary of the main findings of the thesis. Since the following chapters correspond to published articles and are thus designed to be self-contained, we first state the correspondence between chapters and articles. Then, in order to ease the reading, we give some complementary constructive results of our approach. This should help the reader to see the connections between the various chapters and he/she may return to this overview if necessary.

1.3.1 *Structure of the thesis*

The chapters of this thesis are based on the following articles:

— Chapter 2 was published as a joint work with Nicolas Petit and Zoltán Szabó in the Proceedings of IFAC World Congress 2020, under the title *Kernel Regression for*

Vehicle Trajectory Reconstruction under Speed and Inter-vehicular Distance Constraints (Aubin-Frankowski et al., 2020).

— Chapter 3 was published as a joint work with Zoltán Szabó in Advances in Neural Information Processing Systems (NeurIPS), under the title *Hard Shape-Constrained Kernel Machines* (Aubin-Frankowski & Szabó, 2020b).

— Chapter 4 was accepted for publication with a single author in SIAM Journal on Control and Optimization, under the title *Linearly-constrained Linear Quadratic Regulator from the viewpoint of kernel methods* (Aubin-Frankowski, 2021b).

— Chapter 5 was published with a single author in Comptes Rendus. Mathématique, under the title *Interpreting the dual Riccati equation through the LQ reproducing kernel* (Aubin-Frankowski, 2021a).

— Chapter 6 was published with a single author in Systems & Control Letters, under the title *Lipschitz regularity of the minimum time function of differential inclusions with state constraints* (Aubin-Frankowski, 2020).

— Chapter 7 was published as a joint work with Nicolas Petit in the Proceedings of the European Control Conference (ECC) 2020, under the title *Data-driven approximation of differential inclusions and application to detection of transportation modes* (Aubin-Frankowski & Petit, 2020).

We outline below the structure of Section 1.3:

— Section 1.3.2 deals with the SOC tightening of the shape-constrained empirical risk minimization ($\mathcal{P}_{\mathrm{ML}}$). This corresponds to Chapters 2 and 3.

— Section 1.3.3 presents the novel LQ kernel, designed for (state-constrained) LQ optimal control ($\mathcal{P}_{\mathrm{LQ}}$). This corresponds to Chapters 4 and 5.

— Section 1.3.4 considers differential inclusions, theoretically, to study the regularity of the minimal time function and, numerically, to estimate their graphs. This corresponds to Chapters 6 and 7.

1.3.2 *Tightening infinitely many shape constraints into finitely many*

This section presents two interpretations, algebraic and geometric, of our advocated choice of tightening as either an upper bound of a modulus of continuity in $\mathcal{H}_K$ or as a consequence of a compact covering scheme in $\mathcal{H}_K$. While the algebraic approach allows to derive immediately the tightening, it veils the much richer geometric interpretation. The latter is implicitly used in Chapter 2 when performing a covering based on a ball intersected with half-spaces. In this section, we thus present the advocated

methods in greater generality and in a single formalism, encompassing the two types of coverings used in Chapter 2 and 3. We will follow a constructive approach to show that we can tighten infinitely many shape constraints into finitely many by considering second-order cone (SOC) constraints.

Notations for differential operators: Let the set of linear differential operators of order at most $s \in \mathbb{N}$ on real-valued functions be denoted by

$$O_{1,s} := \left\{ D \,|\, D(f)(x) = \sum_{j \in J} c_j \partial^{r_j} f(x), \operatorname{card}(J) < \infty, |r_j| \leqslant s, c_j \in \mathbb{R} \,(\forall j \in J) \right\}.$$

The set of linear differential operators of order at most $s \in \mathbb{N}$ on $\mathbb{R}^Q$-valued functions is

$$O_{Q,s} := \left\{ D \,|\, D(f)(x) = \sum_{q \in [Q]} \beta_q D_q(f_q)(x), \beta_q \in \mathbb{R}, D_q \in O_{1,s} \right\}.$$

For differential operators $D, \tilde{D} \in O_{Q,s}$ defined as $D(f)(x) = \sum_{q \in [Q]} \gamma_q D_{q,x}(f_q)(x)$ and $\tilde{D}(f)(x') = \sum_{q \in [Q]} \tilde{\gamma}_q \tilde{D}_{q,x'}(f_q)(x')$, and for a kernel $K \in \mathcal{C}^{s,s}\left(\mathcal{X} \times \mathcal{X}, \mathbb{R}^{Q \times Q}\right)$, let

$$DK(x', x) := \sum_{q \in [Q]} \beta_q [D_{q,x} K(x', x)] e_q \in \mathbb{R}^Q, \tag{1.5}$$

$$\tilde{D}^\top DK(x', x) := \sum_{q, q' \in [Q]} \tilde{\gamma}_{q'} \gamma_q e_{q'}^\top \tilde{D}_{q',x'} D_{q,x} K(x', x) e_q \in \mathbb{R}, \tag{1.6}$$

where $D_{q,x} K(x', x) := D_q[x'' \mapsto K(x', x'')](x) \in \mathbb{R}^{Q \times Q}$ and $e_q \in \mathbb{R}^Q$ is the q-th canonical basis vector.

Assumptions: In this section, $K \in \mathcal{C}^{s,s}\left(\mathcal{X} \times \mathcal{X}, \mathbb{R}^{Q \times Q}\right)$ is a smooth matrix-valued kernel and $\mathcal{X} \subseteq \mathbb{R}^d$ is a set which is contained in the closure of its interior. Hence, we can apply the reproducing property for derivatives (1.3) from Theorem 1.3. The compact set $\mathcal{K} \subset X$ is assumed to be non-finite, in order to face the problem of dealing with infinitely many constraints.

We consider at first a single hard affine real-valued shape constraint over $\mathcal{K}$, i.e. ($\mathcal{C}_1$) below, which corresponds to ($\mathcal{C}$) for $I = 1$. Multiple shape constraints ($I > 1$) can be addressed by stacking the results below. Consequently we drop the i index and ($\mathcal{C}$) simplifies to

$$C_1 := \left\{ (f, b) \,|\, 0 \leqslant D(f - f_0)(x) + \beta^\top b - b_0, \forall x \in \mathcal{K} \right\}. \tag{$\mathcal{C}_1$}$$

There are two main challenges to tackle: (i) C_1 cannot be directly implemented since $\mathcal{K}$ is non-finite, (ii) representer theorems, our battle horses showcased in Section 1.2.5, require a finite number of evaluations f. We are thus enticed to discretize the constraint over a collection $\{\tilde{x}_m\}_{m\in[M]} \subset \mathcal{X}$. However this would only give a relaxation. Since Theorem 1.5 holds for some extra norm term on the left-hand side of ($\mathcal{C}_1$), we can tighten ($\mathcal{C}_1$) by considering the following margin buffer in the constraints:

$$C_{1,SOC} := \left\{ (f, b) \,|\, \eta_m \|f - f_0\|_K \leqslant D(f - f_0)(\tilde{x}_m) + \beta^\top b - b_0, \quad \forall m \in [M] \right\}. \quad (\mathcal{C}_{1,SOC})$$

The value of η_m should be chosen depending on D, K and the $\{\tilde{x}_m\}_{m\in[M]}$. For any given finite covering $\mathcal{K} \subseteq \bigcup_{m\in[M]} \mathbb{B}_{\mathcal{X}}(\tilde{x}_m, \delta_m)$, one such value of η_m leading to a tightening is:

$$\eta_m := \sup_{x \in \mathbb{B}_{\mathcal{X}}(\tilde{x}_m, \delta_m)} \|DK(\cdot, \tilde{x}_m) - DK(\cdot, x)\|_K, \; m \in [M]. \quad (1.7)$$

The choice of η_m as in (1.7) can be interpreted in two ways, as:

1. an upper bound of a modulus of continuity of $D(f - f_0)$ defined on a finite covering of the compact set $\mathcal{K}$ (see Section 1.3.2.1);
2. a finite compact covering in $\mathcal{H}_K$. Indeed ($\mathcal{C}_1$) can actually be rewritten as the inclusion in the vRKHS $\mathcal{H}_K$ of a compact set in a half-space. We then tighten this inclusion by taking a finite covering of the compact set in $\mathcal{H}_K$, the value η_m as in (1.7) corresponding to the radius of a specific choice of covering by balls (see Section 1.3.2.2).

Tightening ($\mathcal{C}_1$) into ($\mathcal{C}_{1,SOC}$) leads to second-order cone (SOC) constraints[9] instead of affine inequalities. Hence the set $C_{1,SOC}$ is still convex but, in order to guarantee the shape constraints on a neighborhood of the points $\{\tilde{x}_m\}_{m\in[M]}$, there is a computational price to pay because of the second-order cone programming involved.

1.3.2.1 *Algebraic interpretation: an upper bound of a modulus of continuity in $\mathcal{H}_K$*

We first present an interpretation of ($\mathcal{C}_{1,SOC}$) as an upper bound of the modulus of continuity of $D(f - f_0)$ over a finite covering of a compact $\mathcal{K} \subseteq \bigcup_{m\in[M]} \mathbb{B}_{\mathcal{X}}(\tilde{x}_m, \delta_m)$. Let the modulus of continuity of $D(f - f_0)$ on $\mathbb{B}_{\mathcal{X}}(\tilde{x}_m, \delta_m)$ be defined as

$$\omega_{D(f-f_0)}(\tilde{x}_m, \delta_m) := \sup_{x \in \mathbb{B}_{\mathcal{X}}(\tilde{x}_m, \delta_m)} |D(f - f_0)(x) - D(f - f_0)(\tilde{x}_m)|. \quad (1.8)$$

9. The "second-order cone" terminology is classical in optimization following the similarity between ($\mathcal{C}_{1,SOC}$) and the definition of the Lorentz cone $\{(z, r) \in \mathbb{R}^{N+1} \,|\, \|z\|_2 \leqslant r\}$. In finite dimensions, SOC constraints write as $\{z \in \mathbb{R}^N \,|\, \|Az + b\|_2 \leqslant \beta^\top z - b_0\}$ for $A \in \mathbb{R}^{P\times N}$.

Would it happen that we have an exact finite covering, i.e. $\mathcal{K} = \bigcup_{m\in[M]} \mathbb{B}_{\mathcal{X}}(\tilde{\mathbf{x}}_m, \delta_m)$, then, if $\omega_{D(\mathbf{f}-\mathbf{f}_0)}(\tilde{\mathbf{x}}_m, \delta_m)$ was known for every $m \in [M]$, the constraint $(\mathbf{f}, \mathbf{b}) \in C_1$ would be equivalent to

$$\omega_{D(\mathbf{f}-\mathbf{f}_0)}(\tilde{\mathbf{x}}_m, \delta_m) \leqslant D(\mathbf{f}-\mathbf{f}_0)(\tilde{\mathbf{x}}_m) + \boldsymbol{\beta}^\top \mathbf{b} - b_0, \ \forall\, m \in [M]. \tag{1.9}$$

The equivalence follows from (1.8) since the modulus of continuity is the smallest upper bound of the variations of the values. Applying the reproducing property for derivatives (Theorem 1.3) and the Cauchy-Schwarz inequality, we obtain an upper bound, with η_m defined as in (1.7),

$$\omega_{D(\mathbf{f}-\mathbf{f}_0)}(\tilde{\mathbf{x}}_m, \delta_m) = \sup_{\mathbf{x}\in\mathbb{B}_{\mathcal{X}}(\tilde{\mathbf{x}}_m,\delta_m)} |\langle \mathbf{f}-\mathbf{f}_0, DK(\cdot,\mathbf{x}) - DK(\cdot,\tilde{\mathbf{x}}_m)\rangle_K| \leqslant \eta_m \|\mathbf{f}-\mathbf{f}_0\|_K. \tag{1.10}$$

While the original quantity $\omega_{D(\mathbf{f}-\mathbf{f}_0)}(\tilde{\mathbf{x}}_m, \delta_m)$ can be hard to evaluate, the bound $\eta_m \|\mathbf{f}-\mathbf{f}_0\|_K$ is much more favourable from a computational perspective. Indeed, the term η_m has an explicit finite-dimensional description as, by Theorem 1.3 and the definition (1.6) of $D^\top D$,

$$\|DK(\cdot,\mathbf{x}) - DK(\cdot,\tilde{\mathbf{x}}_m)\|_K^2 = D^\top DK(\mathbf{x},\mathbf{x}) + D^\top DK(\tilde{\mathbf{x}}_m,\tilde{\mathbf{x}}_m) - 2D^\top DK(\mathbf{x},\tilde{\mathbf{x}}_m).$$

By (1.10), the SOC term is larger than the modulus of continuity. From (1.9), the modulus of continuity defines a finite number of inequalities which are equivalent to the original constraint ($\mathcal{C}_1$) with an infinite number of inequalities, i.e. $(\mathbf{f}, \mathbf{b}) \in C_1$. Hence, when replacing $\omega_{D(\mathbf{f}-\mathbf{f}_0)}(\tilde{\mathbf{x}}_m, \delta_m)$ in (1.9) by its upper bound obtained in (1.10), the equivalence becomes an implication and gives rise to the tightened second-order cone (SOC) constraints

$$\eta_m \|\mathbf{f}-\mathbf{f}_0\|_K \leqslant D(\mathbf{f}-\mathbf{f}_0)(\tilde{\mathbf{x}}_m) + \boldsymbol{\beta}^\top \mathbf{b} - b_0, \ \forall\, m \in [M].$$

The value of η_m can be computed analytically in various cases. For instance, for $Q = 1$ and $D = \mathrm{Id}$, with a monotonically decreasing real-valued radial kernel $K(\mathbf{x},\mathbf{y}) = k_0(\|\mathbf{x}-\mathbf{y}\|_{\mathcal{X}})$ (such as the Gaussian kernel), then

$$\eta_m(\delta_m) = \sup_{\mathbf{x}\in\mathbb{B}_{\mathcal{X}}(0,\delta_m)} \sqrt{|2k_0(0) - 2k_0(\|\mathbf{x}\|_{\mathcal{X}})|} = \sqrt{2}\sqrt{|k_0(0) - k_0(\delta_m)|}. \tag{1.11}$$

Depending on the choice of the kernel, similar computations could be carried out for higher-order derivatives. For translation-invariant kernels, η_m can be computed on a single δ_m-ball around the origin as in (1.11). A fast approximation of η_m can also for instance be performed by sampling $\mathbf{x}$ in the ball $\mathbb{B}_{\mathcal{X}}(\tilde{\mathbf{x}}_m, \delta_m)$. Moreover, as η_m is related to the modulus of continuity of DK, the smoother the kernel, the smaller η_m and the tighter the approximation (1.10). As it intuitively follows from (1.10), η_m is *one* of the possible upper bound on the modulus of continuity, enabling guarantees for hard shape constraints. This bound is also tight in the equality case of the Cauchy-Schwarz inequality (1.10).

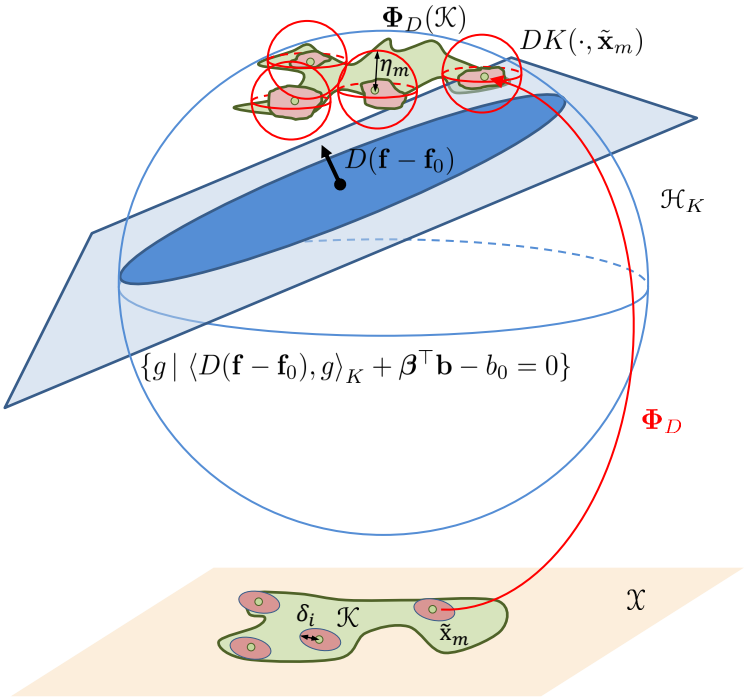


Figure 1.1 – Illustration of the compact covering.

1.3.2.2 *Geometric interpretation: a compact covering scheme in $\mathcal{H}_K$*

Notations for sets and half-spaces: Given a vRKHS $\mathcal{H}_K$, for $\mathbf{f} \in \mathcal{H}_K$ and $\rho \in \mathbb{R}$, let the closed half-spaces and the affine hyperplane associated with the pair $(\mathbf{f}, \rho)$, as well the closed ball in $\mathcal{H}_K$ with center $\mathbf{c} \in \mathcal{H}_K$ and radius $r > 0$ be defined as

$$\mathrm{H}_K(\mathbf{f}, \rho) := \{\mathbf{g} \in \mathcal{H}_K \mid \langle \mathbf{f}, \mathbf{g} \rangle_K = \rho\}, \qquad \mathrm{H}_K^+(\mathbf{f}, \rho) := \{\mathbf{g} \in \mathcal{H}_K \mid \langle \mathbf{f}, \mathbf{g} \rangle_K \geqslant \rho\},$$
$$\mathbb{B}_K(\mathbf{c}, r) := \{\mathbf{f} \in \mathcal{H}_K \mid \|\mathbf{f} - \mathbf{c}\|_K \leqslant r\}, \qquad \mathrm{H}_K^-(\mathbf{f}, \rho) := \{\mathbf{g} \in \mathcal{H}_K \mid \langle \mathbf{f}, \mathbf{g} \rangle_K \leqslant \rho\}.$$

We generically denote open sets in $\mathcal{H}_K$ by Ω, whereas $\bar{\Omega}$ (resp. $\overline{\mathrm{co}}$) refers to their closure (resp. closed convex hull). The open ball is $\mathring{\mathbb{B}}_K(\mathbf{c}, r)$ and $\mathring{\mathrm{H}}_K^+(\mathbf{f}, \rho)$ is the open half-space.

The algebraic approach of Section 1.3.2.1 allowed to derive immediately ($\mathcal{C}_{1,\mathrm{SOC}}$). Nevertheless it veiled a much richer geometric interpretation based on compact cover-

ings in $\mathcal{H}_K$. Applying Theorem 1.3, we rephrase constraint ($\mathcal{C}_1$) as an inclusion of sets using the non-linear embedding $\mathbf{\Phi}_D : \mathbf{x} \in \mathcal{K} \mapsto DK(\cdot, \mathbf{x}) \in \mathcal{H}_K$

$$\begin{aligned}(\mathbf{f}, \mathbf{b}) \in C &\Leftrightarrow b_0 - \boldsymbol{\beta}^\top \mathbf{b} \leqslant D(\mathbf{f} - \mathbf{f}_0)(\mathbf{x}) = \langle \mathbf{f} - \mathbf{f}_0, DK(\cdot, \mathbf{x}) \rangle_K \ \forall \mathbf{x} \in \mathcal{K} \\ &\Leftrightarrow \mathbf{\Phi}_D(\mathcal{K}) := \{DK(\cdot, \mathbf{x}) \,|\, \mathbf{x} \in \mathcal{K}\} \subseteq H_K^+(\mathbf{f} - \mathbf{f}_0, b_0 - \boldsymbol{\beta}^\top \mathbf{b}). \end{aligned} \tag{1.12}$$

The set $\mathbf{\Phi}_D(\mathcal{K})$ is compact in $\mathcal{H}_K$ since $\mathcal{K}$ is compact in $\mathcal{X}$ and $\mathbf{\Phi}_D$ is continuous. However it is intractable to directly ensure the inclusion described in (1.12) whenever $\mathcal{K}$ is not finite. We thus consider an approximation with a "simpler" set $\bar{\Omega}$ containing $\mathbf{\Phi}_D(\mathcal{K})$, and require the inclusion

$$\mathbf{\Phi}_D(\mathcal{K}) \subseteq \bar{\Omega} \subseteq H_K^+(\mathbf{f} - \mathbf{f}_0, b_0 - \boldsymbol{\beta}^\top \mathbf{b}) \tag{1.13}$$

which implies (1.12). Since $\mathbf{\Phi}_D(\mathcal{K})$ is compact, drawing upon compact coverings, we assume that

$$\bar{\Omega} = \cup_{m \in [M]} \bar{\Omega}_m, \tag{1.14}$$

where each $\bar{\Omega}_m$ is the closure of a nonempty finite intersection ($J_{B,m}, J_{H,m} \in \mathbb{N}$) of non-trivial ($r_{m,j} > 0$, $\boldsymbol{\nu}_{m,j} \neq 0$) open balls and open half-spaces

$$\Omega_m = \left(\bigcap_{j \in [J_{B,m}]} \mathring{\mathbb{B}}_K(\mathbf{c}_{m,j}, r_{m,j}) \right) \cap \left(\bigcap_{j \in [J_{H,m}]} \mathring{H}_K^-(\boldsymbol{\nu}_{m,j}, \rho_{m,j}) \right). \tag{1.15}$$

We illustrate in Figure 1.1 one such covering of $\Phi_D(\mathcal{K})$ using only balls in $\mathcal{H}_K$.

Remarks:

— Form of (1.15): The motivation for considering $\bar{\Omega}$ and $\bar{\Omega}_m$ of the form (1.14) and (1.15) is several-fold. Having a finite description enables one to derive a representer theorem. However, only a few sets (mainly points, balls and half-spaces) enjoy explicit convex separation formulas.[10] Focusing on *points* leads to a discretization of $\mathbf{\Phi}_D(\mathcal{K})$ and greedy strategies (such as the Frank-Wolfe algorithm), but without guarantees outside of the points considered. A finite union of *balls* can approximate any compact set in the Hausdorff metric, but balls result in enforcing "buffers" in every direction of $\mathcal{H}_K$. A finite intersection of *half-spaces* can approximate any convex set,[11] but this finite intersection is always unbounded for infinite-dimensional

10. Since Ω_m is convex, the inclusion $\Omega_m \subseteq H_K^+(\mathbf{f} - \mathbf{f}_0, b_0 - \Gamma \mathbf{b})$ in (1.13) is equivalent to $\Omega_m \cap \mathring{H}_K^-(\mathbf{f} - \mathbf{f}_0, b_0 - \Gamma \mathbf{b}) = \emptyset$ which can be interpreted as a convex separation.

11. To motivate the use of half-spaces in (1.15): notice that since the half-space on the r.h.s. of (1.12) is closed and convex, (1.12) is equivalent to the fact that the closed convex hull $\overline{co}(\Phi_D(\mathcal{K}))$ is a subset of $H_K^+(\mathbf{f} - \mathbf{f}_0, b_0 - \boldsymbol{\beta}^\top \mathbf{b})$. Using the support function characterization of closed convex sets, we have that $\overline{co}(\Phi_D(\mathcal{K})) = \bigcap_{g \in \mathcal{H}_K} H_K^-(g, \sigma_{\mathcal{K}}(g))$ where $\sigma_{\mathcal{K}}(g) := \sup_{x \in \mathcal{K}} D(g)(x)$ has to be computed. Considering any finite collection of $g \in \mathcal{H}_K$ provides an over-approximation of $\overline{co}(\Phi_D(\mathcal{K}))$ with finite description.

$\mathcal{H}_K$ resulting in a poor approximation of compact sets. Motivated by obtaining guarantees, we thus consider combinations of balls and half-spaces as in (1.14)-(1.15).

— Non-compact $\mathcal{K}$: Coverings of the form (1.14) and (1.15) exist for any bounded $\mathbf{\Phi}_D(\mathcal{K})$. In particular, if the kernel $D^\top DK(\cdot,\cdot)$ is bounded (recall that $D^\top$ denotes D applied to the first variable of K), then for any set $\mathcal{K} \subseteq \mathcal{X}$, $\mathbf{\Phi}_D(\mathcal{K}) \subset \mathcal{H}_K$ is bounded as well. Consequently the proposed method can be applied to non-compact $\mathcal{K}$ provided that the derivatives of the chosen kernel are bounded.

Owing to the geometric reinterpretation stated in (1.12), one just needs to rewrite $\bar{\Omega} \subseteq H_K^+(\mathbf{f}-\mathbf{f}_0, b_0 - \boldsymbol{\beta}^\top \mathbf{b})$ in (1.13) as a separation formula for convex sets, and choose $\bar{\Omega}$ adequately such that $\mathbf{\Phi}_D(\mathcal{K}) \subseteq \bar{\Omega}$. We will consider the simpler case of Ω_m being either a single ball or a ball intersected with a single half-space. [12] Since $\bar{\Omega}$ is defined as a union in (1.14), when considering its inclusion in $H_K^+(\mathbf{f}-\mathbf{f}_0, b_0 - \boldsymbol{\beta}^\top \mathbf{b})$, one can consider each Ω_m individually and drop the m index. We consider thus the two following cases:

— If $\Omega = \mathring{\mathbb{B}}_K(\mathbf{c}, r)$, then the inclusion is equivalent to saying that the center $\mathbf{c}$ is at a distance r from the hyperplane, i.e.

$$\langle \mathbf{f}-\mathbf{f}_0, \mathbf{c}\rangle_K \geqslant r\,\|\mathbf{f}-\mathbf{f}_0\|_K + b_0 - \boldsymbol{\beta}^\top \mathbf{b}.$$

— If $\Omega = \mathring{\mathbb{B}}_K(\mathbf{c}, r) \cap \mathring{H}_K^-(\boldsymbol{\nu}, \rho)$ then a similar, though more involved, result states that it is necessary and sufficient that there exists $\xi \geqslant 0$ such that

$$\langle \mathbf{f}-\mathbf{f}_0+\xi\boldsymbol{\nu}, \mathbf{c}\rangle_K \geqslant r\,\|\mathbf{f}-\mathbf{f}_0+\xi\boldsymbol{\nu}\|_K + b_0 - \boldsymbol{\beta}^\top \mathbf{b} + \xi\rho.$$

If (1.13) holds, then, because of the inclusion, we have a tightening of ($\mathcal{C}_1$). A natural choice of Ω of the form (1.14) can be obtained by leveraging the compactness of $\mathcal{K}$. Indeed, let us take any finite covering of $\mathcal{K}$ through balls centered at M points $\{\tilde{\mathbf{x}}_m\}_{m\in[M]}$ with radius $\delta_m > 0$. Then one can cover the sets $\Phi_D(\mathbb{B}_\mathcal{X}(\tilde{\mathbf{x}}_m, \delta_m)) \subset \mathcal{H}_K$ by balls $\Omega_m = \mathring{\mathbb{B}}_K(DK(\cdot, \tilde{\mathbf{x}}_m), \eta_m)$ with radii η_m given by (1.7). In other words, $\Phi_D(\mathcal{K}) \subseteq \bigcup_{m\in[M]} \Phi_D(\mathbb{B}_\mathcal{X}(\tilde{\mathbf{x}}_m, \delta_m)) \subseteq \bigcup_{m\in[M]} \bar{\Omega}_m =: \bar{\Omega}$, hence $\bar{\Omega}$ satisfies (1.13) and $(\mathbf{f}, \mathbf{b}) \in C_{1,\mathrm{SOC}}$.

Coming back to Figure 1.1, the constraint ($\mathcal{C}$) was reformulated as requiring that the image $\Phi_D(\mathcal{K})$ of $\mathcal{K}$ under the D-feature map Φ_D is contained in the halfspace 'above' the affine hyperplane defined by its normal vector $D(\mathbf{f}-\mathbf{f}_0)$ and bias $b_0 - \boldsymbol{\beta}^\top \mathbf{b}$. The

12. This more restrictive case is the one considered in Aubin-Frankowski et al. (2020) and Aubin-Frankowski and Szabó (2020b). Extensions of the formulas to the more general (1.15) can be found in Aubin-Frankowski and Szabó (2020a).

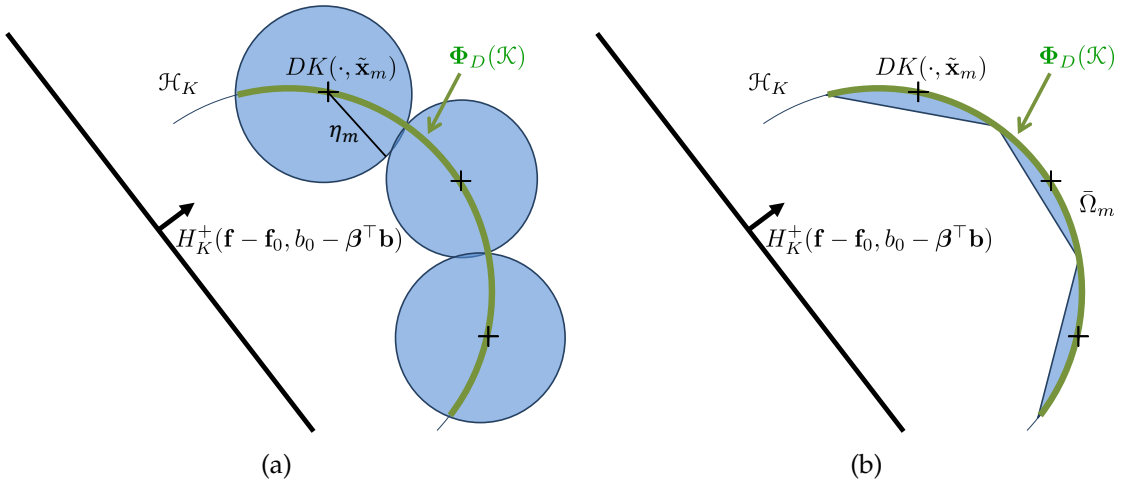


Figure 1.2 – Two examples of coverings in $\mathcal{H}_K$ of $\boldsymbol{\Phi}_D(\mathcal{K})$ by a set $\bar{\Omega} = \cup_{m\in[M]}\bar{\Omega}_m$ contained in the halfspace $H_K^+(\mathbf{f} - \mathbf{f}_0, b_0 - \boldsymbol{\beta}^\top \mathbf{b})$. (a): covering through balls $\Omega_m = \mathring{\mathbb{B}}_K(DK(\cdot, \tilde{x}_m), \eta_m)$. (b): covering through a ball intersected with half-spaces ($J_{B,m} = J_{H,m} = 1$).

discretization, presented below in (1.17), of the constraint ($\mathcal{C}$) at the points $\{\tilde{\mathbf{x}}_m\}_{m\in[M]}$ only requires the images $\Phi_D(\tilde{\mathbf{x}}_m)$ of the points to be above the hyperplane. The constraint ($\mathcal{C}_\eta$) instead inflates each of those points by a radius η_i. With this sketch in mind, we provide a visual illustration of the inclusion relation for Ω_m being a single ball, and for Ω_m being a ball intersected with a single half-space in Figure 1.2(a) and Figure 1.2(b), respectively. Figure 1.2 underlines the fact that more complex structures lead to tighter approximations in Hausdorff distance.

1.3.2.3 *Convergence of the tightening for strongly convex problems*

In this section and the related Chapter 3, we provide guarantees over the recovery of the solution of ($\mathcal{P}_{ML}$) when replacing the constraint ($\mathcal{C}_1$) by ($\mathcal{C}_{1,SOC}$). As the content of Chapter 3 slightly differs from our matrix-valued presentation so far, we sketch here the formalism and the main result of Chapter 3 to make it more readable in light of Section 1.1.1. As a matter of fact, the theorem, published in Aubin-Frankowski and Szabó (2020b), focuses on multitask learning without the vector-valued structure, and thus considers a more restrictive case where K is a "tensor kernel", so the vRKHS is a Cartesian product of a real-valued RKHS $\mathcal{H}_k$, i.e.

$$K(x, y) = k(x, y)\mathrm{Id}_Q, \qquad \mathcal{H}_K = (\mathcal{H}_k)^Q.$$

Aubin-Frankowski and Szabó (2020b) also considered a slightly different objective $\mathcal{L}_\odot$ and regularizer function $R_\odot : \mathbb{R}^Q \mapsto \mathbb{R}_+$, the latter being supposed to be strictly

increasing in each of its arguments. The constraints also involve a matrix $\boldsymbol{W} \in \mathbb{R}^{Q\times Q}$ playing a role similar to the $\boldsymbol{\beta}_i$, and D_i is now defined on the individual components in $\mathcal{H}_k$ only (e.g. $D_i = \partial_{e_1}$):

$$C_\odot = \left\{(\mathbf{f}, \mathbf{b}) \,|\, b_{i,0} - \boldsymbol{\beta}_i^\top \mathbf{b} \leqslant D_i(\boldsymbol{W}\mathbf{f} - \mathbf{f}_0)_i(\mathbf{x}), \forall \mathbf{x} \in \mathcal{K}_i, \forall i \in [I]\right\}. \tag{$\mathcal{C}_\odot$}$$

Assuming there is a unique solution to the problem ($\mathcal{P}_\eta$) below, the SOC tightening at M_i points $\{\tilde{\mathbf{x}}_{i,m}\}_{m\in[M_i]} \subseteq \mathcal{K}_i$ then writes as:

$$(\mathbf{f}_\eta, \mathbf{b}_\eta) = \underset{\substack{\mathbf{f}\in(\mathcal{H}_k)^Q,\, \mathbf{b}\in\mathcal{B}\subset\mathbb{R}^B \\ \text{s.t.}}}{\arg\min} \mathcal{L}_\odot(\mathbf{f}, \mathbf{b}) := L\left(\mathbf{b}, (\mathbf{f}(\mathbf{x}_n))_{n\in[N]}\right) + R_\odot\left((\|f_q\|_k)_{q\in[Q]}\right) \tag{$\mathcal{P}_\eta$}$$

$$b_{i,0} - \boldsymbol{\beta}_i^\top \mathbf{b} + \eta_{i,m}\|(\boldsymbol{W}\mathbf{f} - \mathbf{f}_0)_i\|_k \leqslant \min_{m\in[M_i]} D_i(\boldsymbol{W}\mathbf{f} - \mathbf{f}_0)_i(\tilde{\mathbf{x}}_{i,m})\ \forall i \in [I]. \tag{$\mathcal{C}_\eta$}$$

We also consider the discretization of the I constraints in ($\mathcal{C}_\odot$), which leads to the following relaxation of ($\mathcal{P}_\eta$)

$$v_{\text{disc}} = \min_{\mathbf{f}\in(\mathcal{H}_k)^Q,\, \mathbf{b}\in\mathcal{B}} \mathcal{L}_\odot(\mathbf{f}, \mathbf{b}) \text{ s.t. } b_{i,0} - \boldsymbol{\beta}_i^\top \mathbf{b} \leqslant \min_{m\in[M_i]} D_i(\boldsymbol{W}\mathbf{f} - \mathbf{f}_0)_i(\tilde{\mathbf{x}}_{i,m})\ \forall i \in [I]. \tag{1.17}$$

The theorem concerning the convergence of the tightening, to be found in Chapter 3, shows that ($\mathcal{C}_\eta$) is indeed a tightening of ($\mathcal{C}_\odot$) with a "sandwich property" of the optimal values when paired with a relaxation such as (1.17). A representer theorem is also stated which allows to solve numerically ($\mathcal{P}_\eta$) and give two bounds on the distance in $\mathcal{H}_K \times \mathbb{R}^B$ between the solution of ($\mathcal{P}_\eta$) and that of ($\mathcal{P}_{ML}$) as a function of $(v_\eta - v_{\text{disc}})$ and η respectively. The essential assumptions for the bounds to hold are that $\mathcal{L}_\odot$ is strongly convex in $(\mathbf{f}, \mathbf{b})$ and $\mathbf{b}$ should be constraint-free, i.e. $\mathcal{B} = \mathbb{R}^B$ and $\boldsymbol{\beta}_i^\top \mathbf{b}$ should be able to take any value in $\mathbb{R}$.[13] The intuition behind these assumptions is that, with $\boldsymbol{\beta}_i^\top \mathbf{b}$ left free, one can move up and down the hyperplane of Figure 1.1 until it does not intersect $\boldsymbol{\Phi}_D(\mathcal{K})$ and this move can be performed independently for each constraint i.

In control theory (Section 1.3.3 below), the quadratic regularizer identified in Section 1.3.3.1 gives the strong convexity assumption straightforwardly but one is not allowed to play with the boundaries of the constraints since there is no term playing the role of $\mathbf{b}$. The idea to derive the bound remains similar but the construction is much more sophisticated and relies heavily on results from control theory to construct a feasible trajectory (see Chapter 4).

13. The variable $\mathbf{b}$ is however necessarily subjected to penalties in $\mathcal{L}_\odot$ for the objective to be Lipschitz or strongly convex in $\mathbf{b}$.

1.3.2.4 *Two examples: Joint quantile regression and convoy trajectory reconstruction*

It is instructive to consider a few examples for the problem family ($\mathcal{P}_{\mathrm{ML}}$) for which the results of Section 1.3.2 apply:

Joint quantile regression (JQR; as for instance defined by Sangnier et al. (2016)): Assume that we are given samples $((x_n, y_n))_{n\in[N]}$ from the random variable (X, Y) with values in $\mathcal{X}\times\mathbb{R} \subseteq \mathbb{R}^{d+1}$, as well as Q levels $0 < \tau_1 < \ldots < \tau_Q < 1$. Our goal is to estimate *jointly* the τ_q-quantiles of the conditional distributions $\mathbb{P}(Y|X = x)$ for $q \in [Q]$. In the JQR problem, the estimated τ_q-quantile functions $(f_q + b_q)_{q\in[Q]}$ (modulo the biases $b_q \in \mathbb{R}$) belong to a real-valued RKHS $\mathcal{F}_k$ associated with a kernel $k : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$, and they have to satisfy jointly a monotonically increasing property w.r.t. the quantile level τ. It is natural to require this non-crossing property on the smallest rectangle containing the input points $(x_n)_{n\in[N]}$, in other words on $\mathcal{K} = \prod_{j\in[d]} \left[\min\left\{(x_n)_j\right\}_{n\in[N]}, \max\left\{(x_n)_j\right\}_{n\in[N]}\right]$. Hence, the optimization problem in JQR takes the form

$$\min_{\substack{f\in(\mathcal{F}_k)^Q,\\ b\in\mathbb{R}^Q}} \mathcal{L}(f, b) := \frac{1}{N}\sum_{q\in[Q]}\sum_{n\in[N]} \ell_{\tau_q}\left(y_n - [f_q(x_n) + b_q]\right) + \lambda_b\|b\|_2^2 + \lambda_f \sum_{q\in[Q]} \|f_q\|_{\mathcal{F}_k}^2$$

$$\text{s.t.}$$

$$f_q(x) + b_q \leqslant f_{q+1}(x) + b_{q+1},\ \forall q \in [Q-1],\ \forall x \in \mathcal{K}$$

where $\lambda_b > 0$, $\lambda_f > 0$,[14] and the so-called "pinball loss" is defined as $\ell_\tau(e) = \max(\tau e, (\tau-1)e)$ with $\tau \in (0,1)$. This problem can be obtained as a specific case of ($\mathcal{P}_{\mathrm{ML}}$) by choosing $B = Q$, $s = 0$, $I = Q-1$, $P_i = 1$, $D_i f = f_{i+1} - f_i$, $\beta_i^\top b = b_{i+1} - b_i$ ($\forall i \in [I]$), $K(x, x') = k(x, x')\, I_Q$, $f_{0,i} = 0$, $b_0 = 0$, $\mathcal{B} = \mathbb{R}^B$. Further details and numerical illustration on the JQR problem are provided in Chapter 3.

Convoy trajectory reconstruction: Here, the goal is to estimate vehicle trajectories based on noisy observations. This is a typical situation with GPS measurements, where the imprecision can be compensated through side information, not using only the position of every vehicle but also that of its neighbors. Assume that there are Q vehicles forming a convoy (i.e. they do not overtake and keep a minimum inter-vehicular distance between each other) with speed limit on the vehicles. For each vehicle q we have N_q noisy position measurements $(y_{q,n})_{n\in[N_q]} \subset \mathbb{R}$, each associated with vehicle-specific time points $(x_{q,n})_{n\in[N_q]} \subset \mathcal{X} := [t_0, T]$. Without loss of generality,

14. Sangnier et al. (2016) used the same loss function but a soft non-crossing inducing regularizer inspired by matrix-valued kernels, and also set $\lambda_b = 0$.

let the vehicles be ordered in the lane according to their indices ($q = 1$ is the first, $q = Q$ is the last one). Let $d_{\min} \geqslant 0$ be the minimum inter-vehicular distance, and $v_{\min}$ be the minimal speed to keep.[15] By modelling the location of the q^{th} vehicle at time x as $b_q + f_q(x)$ where $b_q \in \mathbb{R}$, $f_q \in \mathcal{H}_k$ and $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a real-valued kernel, the task at hand can be formulated as

$$\min_{\substack{\mathbf{f}=[f_q]_{q\in[Q]}\in(\mathcal{F}_k)^Q,\\ \mathbf{b}\in\mathbb{R}^Q}} \mathcal{L}(\mathbf{f},\mathbf{b}) := \frac{1}{Q}\sum_{q=1}^{Q}\left[\left(\frac{1}{N_q}\sum_{n=1}^{N_q}|y_{q,n} - (b_q + f_q(x_{q,n}))|^2\right) + \lambda\|f_q\|_{\mathcal{F}_k}^2\right]$$

s.t.

$$d_{\min} + b_{q+1} + f_{q+1}(x) \leqslant b_q + f_q(x), \quad \forall q \in [Q-1],\ \forall x \in \mathcal{X},$$
$$v_{\min} \leqslant f_q'(x) \quad \forall q \in [Q],\ \forall x \in \mathcal{X}.$$

This problem can be obtained as a specific case of ($\mathcal{P}_{\text{ML}}$) by choosing $B = Q$, $s = 1$, $I = 2Q-1$, $P_i = 1$ ($i \in [I]$), $K(x,x') = k(x,x')\,I_Q$, $D_i(\mathbf{f}) = f_i - f_{i+1}$ ($i \in [Q-1]$), $\boldsymbol{\beta}_i^\top \mathbf{b} = b_i - b_{i+1}$ ($i \in [Q-1]$), $b_{0,i} = d_{\min}$ ($i \in [Q-1]$), $D_i(\mathbf{f}) = f'_{i-(Q-1)}$ ($i \in [\![Q, 2Q-1]\!]$), $\boldsymbol{\beta}_i = 0_{1,Q}$ ($i \in [\![Q, 2Q-1]\!]$), $b_{0,i} = v_{\min}$ ($i \in [\![Q, 2Q-1]\!]$), $\mathcal{B} = \mathbb{R}^B$. This application is investigated in Chapter 2.

We next discuss how the approach and insights of Section 1.3.2 can be leveraged for state-constrained LQ optimal control.

1.3.3 *A new kernel for (state-constrained) LQ optimal control*

There have been many attempts at bridging kernel methods and control theory. The two are already related in the works of Kailath (1971) and Parzen (1970). More recently Marco et al. (2017) and Steinke and Schölkopf (2008) have considered kernels for control systems, mainly to encode the input $u(\cdot)$ or for system identification purposes (see e.g. the reviews of Chiuso and Pillonetto (2019) and Pillonetto et al. (2014)). Kernels have also been applied to approximate the Koopman operator over observers of uncontrolled nonlinear systems, in connection with spectral analysis (Fujii & Kawahara, 2019; Williams et al., 2015) or for given controls (Sootla et al., 2018). The kernel Hilbert spaces have also been used to define suitable domains for operators (Giannakis et al., 2019; Rosenfeld et al., 2019). In most cases, the kernel is taken off-the-shelf, as with Gaussian kernels in connection with Bayesian inference (Bertalan et al., 2019; Singh et al., 2018).

15. The requirement $v_{\min} = 0$ means that the vehicles go forward. A maximum speed constraint can be imposed similarly.

On the contrary, departing from the prevalent perspective of using kernel methods as nonlinear embeddings, this thesis rekindles with a long standing tradition of engineering kernels for specific uses. This operator-driven view has been mainly supported by the statistics community, especially in connection with splines and Sobolev spaces (Heckman, 2012; Wahba, 1990). [16] For ($\mathcal{P}_{LQ}$), we show below that the quadratic objective paired with the linear dynamics encodes the relevant kernel, which defines the Hilbert space of controlled trajectories. As kernel methods deal with a special class of Hilbert spaces, they are natural to consider for linear systems or for linearizations of nonlinear systems. Nonetheless the interactions run deeper. For instance we highlight below how the controllability Gramian is directly related to matrix-valued kernels, and we recover the transversality condition merely through a representer theorem. The LQ matrix-valued kernel defined below is also connected to the dual Riccati equation. Since the Riccati equation is often used for the online purpose of finding the optimal control, it enables us to discuss how the kernel formalism effectively allows for optimal synthesis, favoring an offline trajectory-focused viewpoint.

Remark 1.7 (Role of t_{ref})**.** Our extra degree of freedom t_{ref} in ($\mathcal{P}_{LQ}$) allows to define a non-degenerate inner product over the controlled trajectories, covering the case of unpenalized "null control" trajectories. For an optimal control problem such as ($\mathcal{P}_{LQ}$) one works with fixed initial and final times. On the other hand, the solution of the Riccati equation is time-dependent and allows to find optimal trajectories for any initial time. Hence, when working with fixed initial time, we consequently take $t_{ref} = t_0$ and $\mathbf{J}_{ref} = \lambda_0 \mathrm{Id}_Q$ for $\lambda_0 > 0$, in order to cover general terminal costs $g(\cdot)$. For varying initial time, we take instead $t_{ref} = T$, and consider only quadratic terminal costs, setting $\mathbf{J}_{ref} = \mathbf{J}_T \succ \mathbf{0}$ and $g(\cdot) \equiv 0$.

1.3.3.1 *Introducing the LQ kernel*

We first need to bring trajectories to the fore in ($\mathcal{P}_{LQ}$) since the variables are for now both $\boldsymbol{z}(\cdot)$ and $\mathbf{u}(\cdot)$. In his seminal book, Luenberger (Luenberger, 1968, p. 255) already discussed that an optimal control problem such as ($\mathcal{P}_{LQ}$) can be seen as either optimizing over the set of controls $\mathbf{u}(\cdot)$, or jointly over the set of trajectories $\boldsymbol{z}(\cdot)$ and controls $\mathbf{u}(\cdot)$, connected through the dynamic constraint ($\boldsymbol{z}' = \mathbf{A}\boldsymbol{z} + \mathbf{B}\mathbf{u}$). Luenberger also alluded without details to a third possibility, that of optimizing directly over the

16. Drawing inspiration from linear control theory, *control theoretic splines* were devised (Fujioka & Kano, 2013; Magnus Egerstedt, 2009), in particular for path-planning problems (Kano & Fujioka, 2018). Similar approaches for nonlinear systems were proposed earlier, as in Petit et al. (2001). Possibly unbeknownst to non-kernel users, kernel theory, sometimes known as *abstract splines*, is the natural generalization of splines (see e.g. Aubin-Frankowski et al., 2020; Berlinet & Thomas-Agnan, 2004).

controlled trajectories. We follow this last viewpoint and consequently introduce the vector space $\mathcal{S}_{[t_0,T]}$ of controlled trajectories of the linear system:

$$\mathcal{S}_{[t_0,T]} := \Big\{ z : [t_0, T] \to \mathbb{R}^Q \,|\, \exists\, \mathbf{u}(\cdot) \text{ s.t. } z'(t) = \mathbf{A}(t)z(t) + \mathbf{B}(t)\mathbf{u}(t) \text{ a.e.} \\ \text{and } \int_{t_0}^{T} \mathbf{u}(t)^\top \tilde{\mathbf{R}}(t)\mathbf{u}(t)dt < \infty \Big\}. \tag{1.18}$$

There is not necessarily a unique choice of $\mathbf{u}(\cdot)$ for a given $z(\cdot) \in \mathcal{S}_{[t_0,T]}$.[17] Therefore, with each $z(\cdot) \in \mathcal{S}_{[t_0,T]}$, we associate the control $\mathbf{u}(\cdot)$ having minimal norm based on the pseudoinverse $\mathbf{B}(t)^\ominus$ of $\mathbf{B}(t)$ for the $\mathbb{R}^M$-norm $\|\cdot\|_{\tilde{\mathbf{R}}(t)} := \|\tilde{\mathbf{R}}(t)^{1/2} \cdot\|_2$:[18]

$$\mathbf{u}(t) = \mathbf{B}(t)^\ominus [z'(t) - \mathbf{A}(t)z(t)] \text{ a.e. in } [t_0, T]. \tag{1.19}$$

Problem ($\mathcal{P}_{LQ}$) then induces a natural inner product over $\mathcal{S}_{[t_0,T]}$. As a matter of fact, the expression

$$\langle z_1(\cdot), z_2(\cdot)\rangle_K := z_1(t_{ref})^\top \mathbf{J}_{ref} z_2(t_{ref}) + \int_{t_0}^{T} \left[z_1(t)^\top \tilde{\mathbf{Q}}(t) z_2(t) + \mathbf{u}_1(t)^\top \tilde{\mathbf{R}}(t)\mathbf{u}_2(t) \right] dt \tag{1.20}$$

is bilinear and symmetric over $\mathcal{S}_{[t_0,T]} \times \mathcal{S}_{[t_0,T]}$. Since $\|z(\cdot)\|_K^2 = 0$ implies that $\mathbf{u}(\cdot) \overset{a.e}{\equiv} 0$, as $J_{ref} \succ 0$, $z(t_{ref}) = 0$, hence $z(\cdot) \equiv 0$. The description by $\mathcal{S}_{[t_0,T]}$ of the optimization variables effectively pushes controls in the background while bringing forth trajectories as the main object of study. This describes ($\mathcal{P}_{LQ}$) more as a regression problem over $\mathcal{S}_{[t_0,T]}$ than as an optimal control problem over controls. By Cauchy-Lipschitz's theorem, $\|z(\cdot)\|_{K_0} = \|\mathbf{J}_{ref}^{1/2} z(t_{ref})\|_2$ defines a norm over the finite-dimensional subspace $\mathcal{S}_0$ of trajectories with null quadratic cost (hence null control):

$$\mathcal{S}_0 := \left\{ z(\cdot) \,|\, \int_{t_0}^{T} z(t)^\top \tilde{\mathbf{Q}}(t) z(t) dt = 0 \text{ and } z'(t) = \mathbf{A}(t)z(t), \text{ a.e. in } [t_0, T] \right\}. \tag{1.21}$$

As in spline studies and as anticipated in Section 1.2.3, this $\mathcal{S}_0$ has to be treated aside from the controlled trajectories. This further motivates the introduction of the matrix

17. This is the case for instance if $\mathbf{B}(t)$ is not injective for a set of times t with positive measure.

18. For any operator $\mathbf{B}$ from $\mathbb{R}^P$ to $\mathbb{R}^Q$, the pseudoinverse $\mathbf{B}^\ominus$ w.r.t. $\|\cdot\|_{\tilde{\mathbf{R}}}$ is defined as the operator from $\mathbb{R}^Q$ to $\mathbb{R}^P$ that assigns to each $z \in \mathbb{R}^Q$ the element which minimizes the $\|\cdot\|_{\tilde{\mathbf{R}}}$-norm over the affine subspace of minimizers $\bar{\mathbf{u}}$ of $\|z - \mathbf{B}\mathbf{u}\|_2$, i.e. $\mathbf{B}^\ominus z = \arg\min \|\bar{\mathbf{u}}\|_{\tilde{\mathbf{R}}}^2$ for $\bar{\mathbf{u}} \in \arg\min_{\mathbf{u}\in\mathbb{R}^P} \|z - \mathbf{B}\mathbf{u}\|_2$. Consequently, $\mathbf{B}\mathbf{B}^\ominus\mathbf{B} = \mathbf{B}$, and, if $\tilde{\mathbf{R}} = \mathrm{Id}_P$, $\mathbf{B}\mathbf{B}^\ominus$ is the orthogonal projector onto the image $\mathrm{Im}(\mathbf{B})$ of $\mathbf{B}$. As $z(\cdot) \in \mathcal{S}_{[t_0,T]}$, we have that $z'(t) - \mathbf{A}(t)z(t) \in \mathrm{Im}(\mathbf{B})$ in (1.19), so our selection $\mathbf{u}(\cdot)$ is a control which can generate $z(\cdot)$. The quantity $\int_{t_0}^{T} \mathbf{u}(t)^\top \tilde{\mathbf{R}}(t)\mathbf{u}(t)dt$ is necessarily finite since $\mathbf{u}(t)$ has minimal $\|\cdot\|_{\tilde{\mathbf{R}}}$-norm among the controls associated with $z(\cdot) \in \mathcal{S}_{[t_0,T]}$.

$\mathbf{J}_{\text{ref}}$ in the definition of ($\mathcal{P}_{\text{LQ}}$).

As shown in Chapters 4 and 5 (Aubin-Frankowski, 2021a, 2021b), $(\mathcal{S}_{[t_0,T]}, \langle\cdot,\cdot\rangle_K)$ is a vRKHS over $[t_0, T]$ with a reproducing kernel $K_{[t_0,T]}$, dubbed the *LQ kernel*, which can be explicitly computed. Let us denote by $\mathbf{\Phi}_{\mathbf{A}}(t,s) \in \mathbb{R}^{Q\times Q}$ the state-transition matrix of $z'(\tau) = \mathbf{A}(\tau)z(\tau)$, defined from s to t, i.e. $z(t) = \mathbf{\Phi}_{\mathbf{A}}(t,s)z(s)$. In Chapter 4, the LQ kernel is related to the Gramian of controllability,[19] when setting $t_{\text{ref}} = t_0$, $\mathbf{J}_{\text{ref}} = \mathrm{Id}_Q$, $\tilde{\mathbf{Q}}(\cdot) \equiv 0$, since in this case the kernel of the subspace $\{z(\cdot) \in \mathcal{S}_{[t_0,T]} \,|\, z(t_0) = 0\}$ is:

$$K(s,t) = \int_{t_0}^{\min(s,t)} \mathbf{\Phi}_{\mathbf{A}}(s,\tau)\mathbf{B}(\tau)\tilde{\mathbf{R}}(\tau)^{-1}\mathbf{B}(\tau)^\top \mathbf{\Phi}_{\mathbf{A}}(t,\tau)^\top d\tau. \tag{1.22}$$

In Chapter 5, the LQ kernel is related to the dual Riccati equation when setting $t_{\text{ref}} = T$, $\mathbf{J}_{\text{ref}} = \mathbf{J}_T$ and $g(\cdot) \equiv 0$ in ($\mathcal{P}_{\text{LQ}}$). As a matter of fact, in this case, the diagonal initial values $K_{[t_0,T]}(t_0,t_0)$ are equal to the inverse of the Riccati matrices at time t_0. The latter are the solution of the dual Riccati equation, with terminal condition $\mathbf{M}(T) = \mathbf{J}_T^{-1}$,

$$\partial_1\mathbf{M}(t) = \mathbf{A}(t)\mathbf{M}(t) + \mathbf{M}(t)\mathbf{A}(t)^\top - \mathbf{B}(t)\tilde{\mathbf{R}}(t)^{-1}\mathbf{B}(t)^\top + \mathbf{M}(t)\tilde{\mathbf{Q}}(t)\mathbf{M}(t). \tag{1.23}$$

We now discuss how to apply the result of Section 1.3.2 to LQ optimal control.

1.3.3.2 *Dealing with pure state constraints in LQ optimal control*

In this section, we deal with the pure state constraints of ($\mathcal{P}_{\text{LQ}}$). As discussed in Section 1.3.2, we suggest replacing "$\boldsymbol{\gamma}_i(t)^\top z(t) \leqslant b_{0,i}(t)\ (\forall\, t \in [t_0,T])$" by the following strengthened second-order cone (SOC) constraints:

$$\eta_i(\delta_m, t_m)\|z(\cdot)\|_K + \boldsymbol{\gamma}_i(t_m)^\top z(t_m) \leqslant b_{0,i}(\delta_m, t_m),\ \forall\, m \in [M_i], \forall\, i \in [I], \tag{1.24}$$

where the $(t_m)_{m\in[M_i]} \in [t_0,T]^{M_i}$ are time points associated with radii $\delta_m > 0$ satisfying $[t_0,T] \subset \cup_{m\in[M_i]}[t_m - \delta_m, t_m + \delta_m]$. The constants $\eta_i(\delta_m,t_m)$ and $b_{0,i}(\delta_m,t_m)$ are defined as:

$$\begin{aligned}
\eta_i(\delta_m,t_m) &:= \sup_{t\in[t_m-\delta_m,t_m+\delta_m]\cap[t_0,T]} \|K(\cdot,t_m)\boldsymbol{\gamma}_i(t_m) - K(\cdot,t)\boldsymbol{\gamma}_i(t)\|_K,\\
b_{0,i}(\delta_m,t_m) &:= \inf_{t\in[t_m-\delta_m,t_m+\delta_m]\cap[t_0,T]} b_{0,i}(t).
\end{aligned}$$

To make our approach crystal clear in this control context, we repeat the arguments of Section 1.3.2.1. The tightening of the constraints stems from interpreting

19. The Gramian of controllability is defined as $\int_{t_0}^{T} \mathbf{\Phi}_{\mathbf{A}}(T,\tau)\mathbf{B}(\tau)\tilde{\mathbf{R}}(\tau)^{-1}\mathbf{B}(\tau)^\top \mathbf{\Phi}_{\mathbf{A}}(T,\tau)^\top d\tau$, it is thus equal to the specific value $K(T,T)$ of the kernel defined by (1.22).

$\eta_i(\delta_m, t_m)\|z(\cdot)\|_K$ as an upper bound of the modulus of continuity of the unknown $\gamma_i(\cdot)^\top z(\cdot)$ defined as follows

$$\omega_i^{(\gamma_i z)}(\delta_m, t_m) := \sup_{t \in [t_m-\delta_m, t_m+\delta_m]\cap[t_0,T]} \underbrace{|\gamma_i(t)^\top z(t) - \gamma_i(t_m)^\top z(t_m)|}_{|\langle z(\cdot), K(\cdot,t)\gamma_i(t) - K(\cdot,t_m)\gamma_i(t_m)\rangle_K|} \leqslant \eta_i(\delta_m, t_m)\|z(\cdot)\|_K. \tag{1.25}$$

This inequality is obtained applying successively the reproducing property and the Cauchy–Schwarz inequality. Since the intractable modulus of continuity controls the variations of $\gamma_i(t)z(t)$, the SOC upper bound provides a tractable tightening. With the above notations, our strengthened time-varying linear quadratic optimal control problem with finite horizon and finite number of SOC constraints is

$$\begin{aligned} &\min_{\substack{z(\cdot)\in\mathcal{S},\\ z(0)=z_0}} \quad g(z(T)) + \|z(\cdot)\|_K^2 \\ &\text{s.t.} \\ &\eta_i(\delta_m, t_m)\|z(\cdot)\|_K + \gamma_i(t_m)^\top z(t_m) \leqslant b_{0,i}(\delta_m, t_m),\ \forall\, m \in [M_i], \forall\, i \in [I]. \end{aligned} \tag{$\mathcal{P}_{\delta,\text{fin}}$}$$

Under general assumptions, as proven in Chapter 4, ($\mathcal{P}_{\text{LQ}}$) and ($\mathcal{P}_{\delta,\text{fin}}$) have a unique solution for convex and continuous terminal costs $g(\cdot)$. Furthermore one can approximate arbitrarily close, both in $\|\cdot\|_K$ and $\|\cdot\|_\infty$, the solution of ($\mathcal{P}_{\text{LQ}}$) by the solution of ($\mathcal{P}_{\delta,\text{fin}}$) when refining the discretization grid, since each $\eta_i(\delta_m, t_m)$ converges then to zero and each $b_{0,i}(\delta_m, t_m)$ to $b_{0,i}(t_m)$. As seen in Section 1.3.2, the introduction of ($\mathcal{P}_{\delta,\text{fin}}$) as a tightening of ($\mathcal{P}_{\text{LQ}}$) entirely stems from the vRKHS formalism and does not result from optimal control considerations. It leads to a finite number of evaluations of the variable $z(\cdot)$ in ($\mathcal{P}_{\delta,\text{fin}}$) which allows for a representer theorem (Theorem 1.5 in Section 1.2.5).

We now illustrate how to apply the SOC tightening procedure in LQ optimal control in a toy example of a simplified constrained submarine.

1.3.3.3 *Illustration through the example of a simplified constrained submarine*

The trajectory of an underwater vehicle navigating in a two-dimensional cavern is described by a curve $t \in \mathcal{T} := [0,1] \mapsto [x(t); z(t)] \in \mathbb{R}^2$ corresponding to its lateral (x) and depth (z) coordinates at time $t \in \mathcal{T}$. For simplicity, we assume that the lateral component satisfies $x(0) = 0$ and $\dot{x}(t) = 1$ for all $t \in \mathcal{T}$. In this case, $x(t) = t$ for all $t \in \mathcal{T}$ and the control problem reduces to that of ensuring that the depth $z(t)$ stays between the floor and ceiling of the cavern ($z(t) \in \left[z_{\text{low}}(t), z_{\text{up}}(t)\right]$ for all $t \in \mathcal{T}$). We

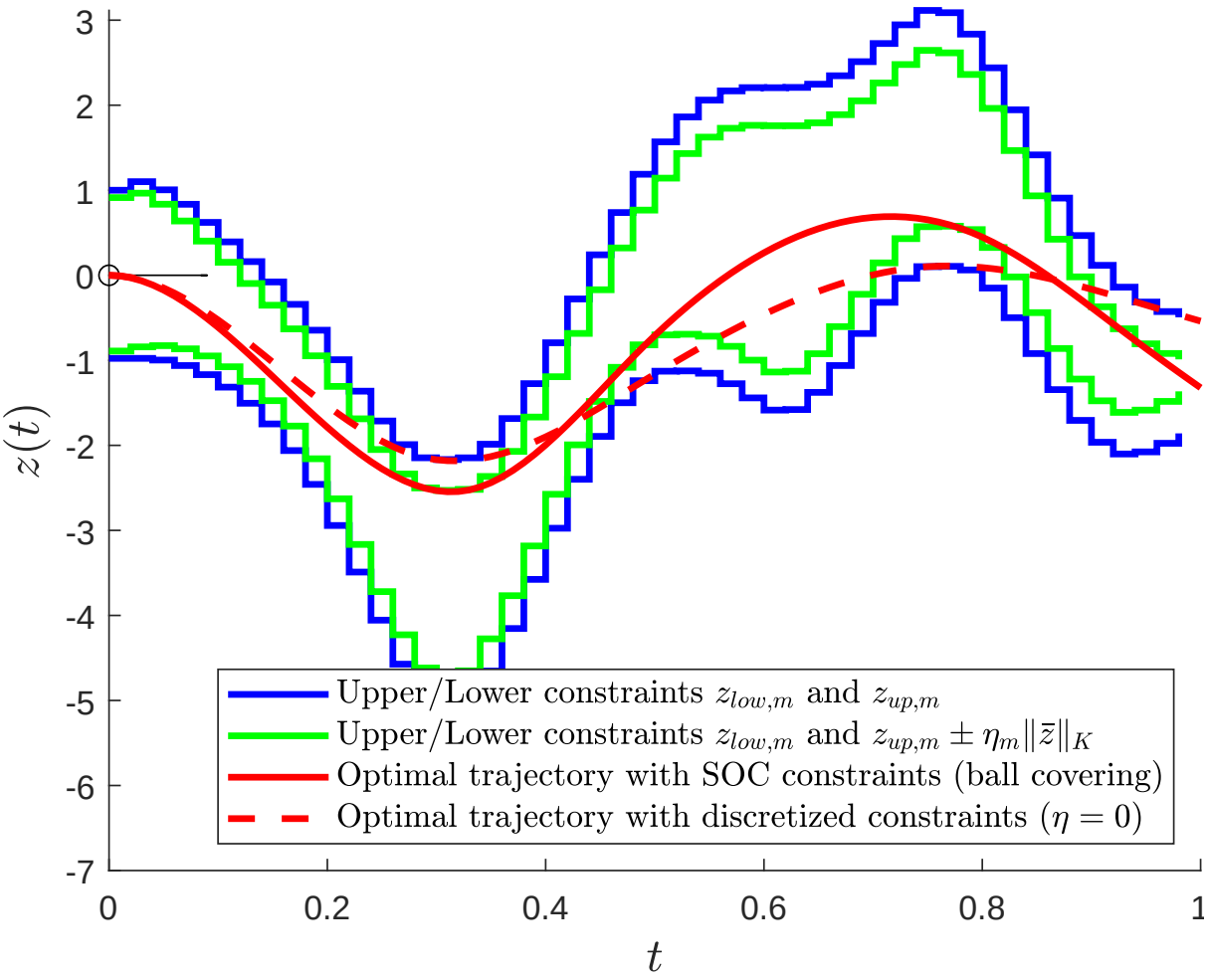


Figure 1.3 – Illustration of the optimal control problem ($\mathcal{P}_{\text{cave}}$) of piloting a vehicle staying between the ceiling (z_{up}) and the floor (z_{low}) of a cavern. Red solid line: SOC-based approach. Red dashed line: solution based on a discretization (formally setting $\eta = 0$). Blue solid lines: constraints ($z_{\text{low},m}$ and $z_{\text{up},m}$). Green solid lines: constraints with buffer $\pm\eta_m \|\bar{z}\|_K$.

take as initial conditions $z(0) = 0$ and $\dot{z}(0) = 0$. Denoting the real-valued control by $u \in L^2(\mathcal{T})$, our control task can be formulated as

$$\begin{aligned} \min_{u(\cdot)\in L^2(\mathcal{T},\mathbb{R})} \quad & \int_{\mathcal{T}} |u(t)|^2 dt \\ \text{s.t.} \quad & \\ & z(0) = 0, \quad \dot{z}(0) = 0, \\ & \ddot{z}(t) = -\dot{z}(t) + u(t), \, \forall\, t \in \mathcal{T}, \\ & z_{\text{low}}(t) \leqslant z(t) \leqslant z_{\text{up}}(t), \, \forall\, t \in \mathcal{T}. \end{aligned} \tag{$\mathcal{P}_{\text{cave}}$}$$

Defining the full state of the vehicle as $\boldsymbol{z}(t) := [z(t); \dot{z}(t)] \in \mathbb{R}^2$, $\boldsymbol{z}(t)$ evolves according to the linear dynamics

$$\dot{\boldsymbol{z}}(t) = \boldsymbol{A}\boldsymbol{z}(t) + \boldsymbol{B}u(t) \in \mathbb{R}^2, \quad \boldsymbol{z}(0) = 0, \quad \boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad \boldsymbol{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$$

According to (1.22), as discussed in Section 1.3.3.1 and Chapter 4, the LQ kernel K of the subspace $\{\boldsymbol{z}(\cdot) \in \mathcal{S}_{\mathcal{T}} \,|\, \boldsymbol{z}(t_0) = 0\}$ is then

$$K(s,t) := \int_0^{\min(s,t)} e^{(s-\tau)\boldsymbol{A}} \boldsymbol{B}\boldsymbol{B}^\top e^{(t-\tau)\boldsymbol{A}^\top} d\tau, \quad s,t \in \mathcal{T}.$$

With our kernel-based formulation, the problem ($\mathcal{P}_{\text{cave}}$) can be rewritten as an optimization problem over full-state trajectories

$$\begin{aligned} \min_{\boldsymbol{z}=[z_1,z_2]\in\mathcal{H}_K} \quad & \|\boldsymbol{z}\|_K^2 \\ \text{s.t.} \quad & \\ & z_{\text{low}}(t) \leqslant z_1(t) \leqslant z_{\text{up}}(t), \, \forall\, t \in \mathcal{T}. \end{aligned}$$

In our experiment we assume that the given bounds z_{low} and z_{up} are piecewise constant: taking a uniform δ-covering $\mathcal{T} = \cup_{m\in[M]}\mathcal{T}_m$ with $\mathcal{T}_m := [t_m - \delta, t_m + \delta]$ and $t_{m+1} = t_m + 2\delta$ for $m \in [M-1]$, this means that $z_{\text{low}}(t) = z_{\text{low},m}$ for all $t \in \mathcal{T}_m$; similarly $z_{\text{up}}(t) = z_{\text{up},m}$ for all $t \in \mathcal{T}_m$.[20] Hence, with the piecewise constant assumption, the control task ($\mathcal{P}_{\text{cave}}$) reduces to

$$\begin{aligned} \min_{\boldsymbol{z}\in\mathcal{H}_K} \quad & \|\boldsymbol{z}\|_K^2 \\ \text{s.t.} \quad & \\ & z_{\text{low},m} \leqslant z_1(t) \leqslant z_{\text{up},m}, \, \forall\, t \in \mathcal{T}_m, \, \forall m \in [M]. \end{aligned}$$

20. For the numerical experiment considered, the piecewise constant bounds are obtained as piecewise approximations of random functions drawn in a Gaussian RKHS.

This optimization problem belongs to the family ($\mathcal{P}_{ML}$) with $Q = 2$, $P = 1$, $D_m(z) = z_1$ and $b_{0,m} = z_{low,m}$ for $m \in [M]$ ($z_{low,m} \leqslant z_1(t)$ for $t \in \mathcal{T}_m$ and $m \in [M]$), $D_{M+m}(z) = -z_1$ and $b_{0,M+m} = -z_{up,m}$ for $m \in [M]$ ($-z_{up,m} \leqslant -z_1(t)$ for $t \in \mathcal{T}_m$ and $m \in [M]$), $z_{0,m} = 0$ and $\Gamma_m = 0$ for $m \in [2M]$, $I = 2M$ and $\mathcal{B} = \{0\}$.

In Fig. 1.3 we compare the optimal trajectory obtained with the proposed SOC tightening (using ball covering) to the one derived when applying discretized constraints (formally corresponding to taking $\eta_m = 0$, as in (1.17)). As illustrated in Fig. 1.3, the vehicle guided with discretized constraints crashes into the blue wall at multiple locations (e.g. $t = 0.5$ or $t = 0.95$), whereas the trajectory resulting from the SOC-based tightening stays within the bounds at all times. The SOC trajectory can be described as solving a problem where $z_{low,m}$ (resp. $z_{up,m}$) was replaced by $z_{low,m} + \eta_m \|\bar{z}\|_K$ (resp. $z_{up,m} - \eta_m \|\bar{z}\|_K$). This acts as a supplementary buffer (green solid line). Even though the SOC trajectory intersects the green boundary, the buffer $\eta_m \|\bar{z}\|_K$ is guaranteed to be large enough for the SOC trajectory to never collide with the blue boundary. On the other hand, the trajectory with discretized constraints is only constrained at the middle points of every piece of the piecewise constant boundaries, it has no information about the constraints in-between. This experiment demonstrates the efficiency of the SOC approach in a safety-critical application where the constraints have to be met at all times.

1.3.4 *Differential inclusions: Lipschitz minimal time and kernel-based graph identification*

This final section is devoted to some aspects of autonomous differential inclusions ($z'(t) \in F(z(t))$) with state constraints. Whereas the linear systems considered in Section 1.3.3.1 form a coherent, albeit restrictive, field, nonlinear control systems are much more diverse. Differential inclusions are to that respect amongst the largest classes of deterministic finite-dimensional systems one can consider. Owing to their nonlinear embedding interpretation, reproducing kernels appear promising to capture some of the nonlinearity of differential inclusions. As shown in Section 1.3.3.1, control-intrinsic Hilbertian kernels have been related only to linear (or linearized) systems. Nevertheless a kernel-based approach could still be used to estimate some relevant quantities, such as the minimal time function in reachability problems with state constraints, or the graph of the set-valued map F based on samples.

Concerning minimal time, Chapter 6 provides sufficient conditions for it to be Lipschitz continuous on its domain. This result is based on inward-pointing conditions, which are also used in the Appendix of Chapter 4. Regularity of the minimal time acts as a conceptual justification of its numerical approximation by smooth functions

such as the ones in a RKHS $\mathcal{H}_k$. It is also a first step to study nonlinear controllability with nonsmooth state constraints. Concerning graph identification, we leverage a kernel method related to one-class SVMs to encode a set $\mathcal{K}$ based on some samples $x_n \in \mathcal{K}$. This corresponds to estimating a pair $(f, \rho) \in \mathcal{H}_k \times \mathbb{R}$ such that $\mathcal{K} \approx \{x \,|\, f(x) \leqslant \rho\}$. We prove that this procedure is "set-consistent" for the Gaussian kernel, i.e. it can approximate any compact set $\mathcal{K}$ in Hausdorff distance provided that the kernel bandwidth is small enough and the samples are dense in $\mathcal{K}$. Graphs of set-valued maps F can thus be approximated through this approach. Since these sections and the corresponding chapters are of a more standalone fashion w.r.t. the rest of the thesis, we only describe here their general objectives and principal contributions.

1.3.4.1 *Lipschitz minimal time of state-constrained differential inclusions*

Studying the regularity of the minimum time function finds its motivation in reachability problems. Let $\mathcal{K}$ and $\mathcal{C}$ be two closed subsets of $\mathbb{R}^Q$ and consider a control system with initial condition $z_0 \in \mathcal{K}$

$$\begin{cases} z'(t) = f(z(t), u(t)), & u(t) \in U, \\ z(0) = z_0, \end{cases} \tag{1.26}$$

where U is a compact subset of $\mathbb{R}^P$, the control $u(\cdot)$ is a measurable function and $f : \mathbb{R}^Q \times U \to \mathbb{R}^Q$ is sufficiently smooth. The state-constrained time optimal control problem consists in finding the minimum time $\tau_{min}(z_0)$ to reach $\mathcal{C}$ along solutions of (1.26) staying in $\mathcal{K}$. Assessing the regularity of τ_{min} allows to answer several questions. For instance, let a time-optimal state-constrained solution $z_{ref}(\cdot)$ of system (1.26) be given. Take as initial condition a point z_1 in a neighborhood of the trajectory set $z_{ref}([0, \tau_{min}(z_0)])$. Under what conditions can z_1 be steered to the target set $\mathcal{C}$ while respecting the state constraints $\mathcal{K}$? How long would it take compared to $\tau_{min}(z_0)$?

More generally, consider the autonomous differential inclusion, with initial condition $z_0 \in \mathcal{K}$:

$$z'(t) \in F(z(t)), \qquad z(0) = z_0, \tag{1.27}$$

where $F : \mathbb{R}^Q \rightsquigarrow \mathbb{R}^Q$ is a set-valued map taking closed, nonempty values. Below, an F-trajectory $z(\cdot)$ on a time interval $[0, T]$ designates an absolutely continuous function satisfying $z'(t) \in F(z(t))$ a.e. on $[0, T]$. A trajectory is called "feasible" if $z([0, T]) \subset \mathcal{K}$.

The capture basin $\mathrm{Capt}_F(\mathcal{K}, \mathcal{C})$ is the set of all points $z_0 \in \mathcal{K}$ such that there exists $T \geqslant 0$ and a feasible F-trajectory starting from z_0 and reaching the target set $\mathcal{C}$ at time T. For a given $z_0 \in \mathcal{K}$, we denote the infimum of such T by $\tau_{min}(z_0)$. By convention,

$\tau_{min}(z_0) = +\infty$ if $z_0 \notin \mathrm{Capt}_{\mathbf{F}}(\mathcal{K}, \mathcal{C})$. By analogy with the control system (1.26), this defines the minimum time function $\tau_{min} : \mathcal{K} \to [0, +\infty]$ associated with the target set $\mathcal{C}$, dynamics $\mathbf{F}$ and state-constraints $\mathcal{K}$. The Lipschitz regularity of the extended function $\tau_{min}(\cdot)$ thus depends on the links existing between both $\mathbf{F}, \mathcal{K}$ and $\mathcal{C}$. In order to exhibit Lipschitz dependence of solutions on initial conditions (through the renowned Filippov's theorem), it is classical to suppose the (local) Lipschitzianity of $\mathbf{F}$. On the other hand, the local Lipschitz continuity of $\tau_{min}(\cdot)$ on its domain $\mathrm{Capt}_{\mathbf{F}}(\mathcal{K}, \mathcal{C})$, in the case without state constraints, has been related to strict inward-pointing conditions on the boundary $\partial\mathcal{C}$ of the target since the 70s (see e.g. Cannarsa & Sinestrari, 2004, Chapter 8, for a modern presentation and the bibliography therein). More recently, it has been shown in Bettiol et al., 2012 that strict inward-pointing conditions on the boundary $\partial\mathcal{K}$ ensure L^∞-distance estimates between arbitrary $\mathbf{F}$-trajectories and the set of feasible ones.

We provide in this thesis sufficient conditions for τ_{min} to be locally Lipschitz continuous on its domain, for state-constrained differential inclusions, under general assumptions on $\mathbf{F}$, $\mathcal{K}$ and $\mathcal{C}$, and a convexified version of the inward-pointing conditions.

1.3.4.2 *SVDD approximation of the graph differential inclusions, applied to detection of transportation modes*

Like in Section 1.3.4.1, we first introduce our problem for control systems, as in (1.26), before extending the framework to differential inclusions, as in (1.27). However, unlike in Section 1.3.4.1, we do not consider the question of the dependence of the minimal time on the initial condition since the unknown is now the function describing the controlled system itself. Given sample trajectories from various fixed systems (e.g. a car and a bike), we would like to determine which of these systems could describe the trajectories underlying some new samples. The possible state constraints are implicit since the sample trajectories are feasible by definition. State constraints are thus incorporated into the definition of the domain of definition of the dynamics. More precisely, consider $J > 1$ forced dynamical systems, each denoted by $(\mathbf{f}_j, U_j)$, for some index $j \in [J]$ that can be interpreted as a discrete-valued parameter, i.e. a *label*. The systems have the governing equations

$$z'(t) = \mathbf{f}_j(z(t), \mathbf{u}(t))$$

where the state vector is $z \in \mathbb{R}^Q$, with the particularity that the input signal $\mathbf{u}(\cdot) \in U_j$ is unknown. Each set U_j is defined as the set of functions containing every possible input signal given the value of j. For example, without further restriction, the sets U_j can be subsets of some functional space (e.g. $\mathcal{C}^0$, L^2) with bounded values in $\mathbb{R}^P$. As

stated above, the forcing signal $\mathbf{u}(\cdot)$ is unknown. Further, despite being unambiguously defined, the sets U_j are unknown as well. To account for this lack of information, the governing equations above are rewritten as differential inclusions:

$$z'(t) \in \mathbf{F}_j(z(t)) := \{\mathbf{f}_j(z(t), \mathbf{u}(t)) \,|\, \mathbf{u}(\cdot) \in U_j\},$$
$$\mathcal{K}_j := \{(z, z') \,|\, z' \in \mathbf{F}_j(z)\} \subset \mathbb{R}^Q \times \mathbb{R}^Q,$$

where the set-valued map $\mathbf{F}_j$ is identified with its graph $\mathcal{K}_j$.[21] We assume that for each value of j, some properly *labeled* values (samples) of the pair $(z(\cdot), z'(\cdot))$ are available, forming J datasets. We first build approximations of the graphs $\mathcal{K}_j$ based on these data points. Once these approximations are available, they can be used to determine which values of j could be compatible with some unlabeled recordings of samples of a new trajectory $(z(\cdot), z'(\cdot))$. In other words, the framework advocated here considers the training data as labeled clouds of points (z, z'), each cloud corresponding to an unknown subset of the unknown U_j. Graphically, constructing an approximation of $\mathcal{K}_j$ amounts to delineating a subset of $\mathbb{R}^{2Q}$ based on the training data (*learning step*). Once the learning step is achieved, when (new) unlabeled data become available, identifying j amounts to testing the membership of the new data to the J learned subsets (*testing step*).

To define approximations of the sets $(\mathcal{K}_j)_{j\in[J]}$, we apply the Support Vector Data Description algorithm (SVDD, Tax & Duin, 2004). SVDD is a kernel method computing a minimal enclosing ball around the data. The output of the algorithm is an indicator function, the evaluation of which allows to readily test membership. The SVDD problem writes as follows

$$(f_N, r_N) := \underset{f\in\mathcal{H}_k,\, r\in\mathbb{R}}{\arg\min}\ r^2 \ \text{ s.t. } \forall n \in [N],\ \|k(x_n, \cdot) - f(\cdot)\|_{\mathcal{H}_k} \leqslant r. \tag{1.28}$$

Based on (f_N, r_N), one can define the closed set

$$\mathcal{K}_N^{\mathrm{SVDD}} := \{x \in \mathbb{R}^{2Q} \mid \|k_x(\cdot) - f_N(\cdot)\|^2_{\mathcal{H}_k} \leqslant r_N^2\}$$

which is the SVDD candidate to approximate $\mathcal{K}$. It is shown in Chapter 7 that $\mathcal{K}_N^{\mathrm{SVDD}}$ for the Gaussian kernel can estimate any compact set $\mathcal{K} \subset \mathbb{R}^{2Q}$ in the Hausdorff metric (a "set-consistency" property of the procedure), provided that the kernel bandwidth is small enough and that the samples are dense in $\mathcal{K}$. Two indexes j_0 and j_1 can then be distinguished if the approximations of their $\mathcal{K}_j$ differ (which depends on both $\mathbf{f}_j$ and U_j). For applications, both learning and testing have to be computationally tractable

21. The graph is bounded whenever the state and its derivatives are bounded for all $\mathbf{u}(\cdot) \in U_j$.

and robust to noise, and SVDD satisfies both conditions. To show the applicability and relevance of the method, this approach is applied to the problem of detecting transportation modes.

BIBLIOGRAPHY FOR CHAPTER 1

Ackermann, J. (1985). *Sampled-data control systems*. Springer Berlin Heidelberg.

Agrell, C. (2019). Gaussian processes with linear operator inequality constraints. *Journal of Machine Learning Research*, *20*, 1–36.

Aït-Sahalia, Y., & Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, *116*(1-2), 9–47.

Alpay, D. (2001). *The Schur algorithm, reproducing kernel spaces, and system theory*. American Mathematical Society.

Argyriou, A., & Dinuzzo, F. (2014). A unifying view of representer theorems. *International Conference on Machine Learning (ICML)*, *32*(2), 748–756.

Aronszajn, N. (1943). La théorie des noyaux reproduisants et ses applications: Première Partie. *Mathematical Proceedings of the Cambridge Philosophical Society*, *39*(3), 133–153.

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

Aubin-Frankowski, P.-C. (2020). Lipschitz regularity of the minimum time function of differential inclusions with state constraints. *Systems & Control Letters*, *139*, 104677.

Aubin-Frankowski, P.-C. (2021a). Interpreting the dual Riccati equation through the LQ reproducing kernel [(https://arxiv.org/abs/2012.12940)]. *Comptes Rendus. Mathématique*, *359*(2), 199–204.

Aubin-Frankowski, P.-C. (2021b). Linearly-constrained Linear Quadratic Regulator from the viewpoint of kernel methods [(to appear, https://arxiv.org/abs/2011.02196)]. *SIAM Journal on Control and Optimization*.

Aubin-Frankowski, P.-C., & Petit, N. (2020). Data-driven approximation of differential inclusions and application to detection of transportation modes. *Proceedings ECC*, 1358–1364.

Aubin-Frankowski, P.-C., Petit, N., & Szabó, Z. (2020). Kernel Regression for Vehicle Trajectory Reconstruction under Speed and Inter-vehicular Distance Constraints [(https://hal-mines-paristech.archives-ouvertes.fr/hal-03021643)]. *IFAC World Congress 2020*.

Aubin-Frankowski, P.-C., & Szabó, Z. (2020a). *Handling hard affine SDP shape constraints in RKHSs* (tech. rep.) [(https://arxiv.org/abs/2101.01519)].

Aubin-Frankowski, P.-C., & Szabó, Z. (2020b). Hard Shape-Constrained Kernel Machines [(http://arxiv.org/abs/2005.12636)]. *Advances in Neural Information Processing Systems (NeurIPS)*.

Berg, C., Christensen, J. P. R., & Ressel, P. (1984). *Harmonic analysis on semigroups*. Springer New York.

Berlinet, A., & Thomas-Agnan, C. (2004). *Reproducing kernel Hilbert spaces in probability and statistics*. Kluwer.

Bertalan, T., Dietrich, F., Mezić, I., & Kevrekidis, I. G. (2019). On learning Hamiltonian systems from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(12), 121107-1 –121107-9.

Bettiol, P., Frankowska, H., & Vinter, R. B. (2012). L^∞ Estimates on trajectories confined to a closed subset. *Journal of Differential Equations*, *252*(2), 1912–1933.

Blundell, R., Horowitz, J. L., & Parey, M. (2012). Measuring the price responsiveness of gasoline demand: economic shape restrictions and nonparametric demand estimation. *Quantitative Economics*, *3*, 29–51.

Bourdin, L., & Trélat, E. (2017). Linear–quadratic optimal sampled-data control problems: Convergence result and Riccati theory. *Automatica*, *79*, 273–281.

Brault, R., Lambert, A., Szabó, Z., Sangnier, M., & d'Alché-Buc, F. (2019). Infinite-task learning with RKHSs. *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1294–1302.

Brunk, H. D. (1955). Maximum likelihood estimates of monotone parameters. *Annals of Mathematical Statistics*, *26*(4), 607–616.

Burachik, R. S., Kaya, C. Y., & Majeed, S. N. (2014). A Duality Approach for Solving Control-Constrained Linear-Quadratic Optimal Control Problems. *SIAM Journal on Control and Optimization*, *52*(3), 1423–1456.

Cannarsa, P., & Sinestrari, C. (2004). *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Birkhäuser Boston.

Caponnetto, A., Micchelli, C. A., Pontil, M., & Ying, Y. (2008). Universal multi-task kernels. *Journal of Machine Learning Research*, *9*(52), 1615–1646.

Chaplais, F., Malisani, P., & Petit, N. (2011). Design of penalty functions for optimal control of linear dynamical systems under state and input constraints. *Conference on Decision and Control (CDC)*, 6697–6704.

Chen, Y., & Samworth, R. J. (2016). Generalized additive and index models with shape constraints. *Journal of the Royal Statistical Society – Statistical Methodology, Series B*, *78*(4), 729–754.

Chetverikov, D., Santos, A., & Shaikh, A. M. (2018). The econometrics of shape restrictions. *Annual Review of Economics*, *10*(1), 31–63.

Chiuso, A., & Pillonetto, G. (2019). System Identification: A Machine Learning Perspective. *Annual Review of Control, Robotics, and Autonomous Systems*, *2*(1), 281–304.

Cuturi, M. (2005). *Etude de noyaux de semigroupe sur objets structurés dans le cadre de l'apprentissage statistique* (Doctoral dissertation) [(https://tel.archives-ouvertes.fr/pastel-00001823/)]. MINES ParisTech.

Deng, H., & Zhang, C.-H. (2020). Isotonic regression in multi-dimensional spaces and graphs. *Annals of Statistics*, *48*(6), 3672–3698.

Dinuzzo, F., & Schölkopf, B. (2012). The representer theorem for Hilbert spaces: a necessary and sufficient condition. *Advances in Neural Information Processing Systems*, *25*, 189–196.

Dower, P. M., McEneaney, W. M., & Cantoni, M. (2019). Game representations for state constrained continuous time linear regulator problems [(http://arxiv.org/abs/1904.05552)]. *arXiv:1904.05552*.

Fortet, R. (1995). *Vecteurs, fonctions et distributions aléatoires dans les espaces de Hilbert : analyse harmonique et prévision*. Hermès.

Freyberger, J., & Reeves, B. (2018). *Inference under shape restrictions* (tech. rep.) [(https://www.ssc.wisc.edu/~jfreyberger/Shape_Inference_Freyberger_Reeves.pdf)]. University of Wisconsin-Madison.

Fujii, K., & Kawahara, Y. (2019). Dynamic mode decomposition in vector-valued reproducing kernel Hilbert spaces for extracting dynamical structure among observables. *Neural Networks*, *117*, 94–103.

Fujioka, H., & Kano, H. (2013). Control theoretic B-spline smoothing with constraints on derivatives, 2115–2120.

Gerdts, M., & Hüpping, B. (2012). Virtual control regularization of state constrained linear quadratic optimal control problems. *Computational Optimization and Applications*, *51*(2), 867–882.

Giannakis, D., Das, S., & Slawinska, J. (2019). Reproducing kernel Hilbert space compactification of unitary evolution groups [http://arxiv.org/abs/1808.01515].

Grüne, L., & Guglielmi, R. (2018). Turnpike properties and strict dissipativity for discrete time linear quadratic optimal control problems. *SIAM Journal on Control and Optimization*, *56*(2), 1282–1302.

Guntuboyina, A., & Sen, B. (2018). Nonparametric shape-restricted regression. *Statistical Science*, *33*(4), 568–594.

Hall, G. (2018). *Optimization over nonnegative and convex polynomials with and without semidefinite programming* (PhD Thesis). Princeton University.

Han, Q., Wang, T., Chatterjee, S., & Samworth, R. J. (2019). Isotonic regression in general dimensions. *Annals of Statistics*, *47*(5), 2440–2471.

Han, Q., & Wellner, J. A. (2016). *Multivariate convex regression: global risk bounds and adaptation* (tech. rep.) [(https://arxiv.org/abs/1601.06844)].

Hartl, R. F., Sethi, S. P., & Vickson, R. G. (1995). A Survey of the Maximum Principles for Optimal Control Problems with State Constraints. *SIAM Review*, *37*(2), 181–218.

Heckman, N. (2012). The theory and application of penalized methods or Reproducing Kernel Hilbert Spaces made easy. *Statistics Surveys*, *6*(0), 113–141.

Hermant, A. (2009). Stability analysis of optimal control problems with a second-order state constraint. *SIAM Journal on Optimization*, *20*(1), 104–129.

Hu, J., Kapoor, M., Zhang, W., Hamilton, S. R., & Coombes, K. R. (2005). Analysis of dose-response effects on gene expression data with comparison of two microarray platforms. *Bioinformatics*, *21*(17), 3524–3529.

Johnson, A. L., & Jiang, D. R. (2018). Shape constraints in economics and operations research. *Statistical Science*, *33*(4), 527–546.

Kailath, T. (1971). RKHS approach to detection and estimation problems–I: Deterministic signals in Gaussian noise. *IEEE Transactions on Information Theory*, *17*(5), 530–549.

Kanagawa, M., Hennig, P., Sejdinovic, D., & Sriperumbudur, B. K. (2018). *Gaussian processes and kernel methods: a review on connections and equivalences* (tech. rep.) [(https://arxiv.org/abs/1807.02582)].

Kano, H., & Fujioka, H. (2018). B-spline trajectory planning with curvature constraint. *2018 Annual American Control Conference (ACC)*, 1963–1968.

Kimeldorf, G., & Wahba, G. (1971). Some results on Tchebycheffian spline functions. *Journal of Mathematical Analysis and Applications*, *33*(1), 82–95.

Koenker, R. (2005). *Quantile regression*. Cambridge University Press.

Kojima, A., & Morari, M. (2004). LQ control for constrained continuous-time systems. *Automatica*, *40*(7), 1143–1155.

Koppel, A., Zhang, K., Zhu, H., & Başar, T. (2019). Projected stochastic primal-dual method for constrained online learning with kernels. *IEEE Transactions on Signal Processing*, *67*(10), 2528–2542.

Kur, G., Dagan, Y., & Rakhlin, A. (2020). *Optimality of maximum likelihood for log-concave density estimation and bounded convex regression* (tech. rep.) [(https://arxiv.org/abs/1903.05315)].

Lim, E. (2020). The limiting behavior of isotonic and convex regression estimators when the model is misspecified. *Electronic Journal of Statistics*, *14*, 2053–2097.

Lin, R., Zhang, H., & Zhang, J. (2019). *On reproducing kernel Banach spaces: generic definitions and unified framework of constructions* (tech. rep.) [(https://arxiv.org/abs/1901.01002)].

Luenberger, D. (1968). *Optimization by vector space methods*. Wiley.

Luss, R., Rossett, S., & Shahar, M. (2012). Efficient regularized isotonic regression with application to gene-gene interaction search. *Annals of Applied Statistics*, *6*(1), 253–283.

Magnus Egerstedt, C. M. (2009). *Control theoretic splines: optimal control, statistics, and path planning*. Princeton University Press.

Malisani, P., Chaplais, F., & Petit, N. (2014). An interior penalty method for optimal control problems with state and input constraints of nonlinear systems. *Optimal Control Applications and Methods*, *37*(1), 3–33.

Marco, A., Hennig, P., Schaal, S., & Trimpe, S. (2017). On the design of LQR kernels for efficient controller learning. *Conference on Decision and Control (CDC)*, 5193–5200.

Marteau-Ferey, U., Bach, F., & Rudi, A. (2020). Non-parametric models for non-negative functions. *Advances in Neural Information Processing Systems (NeurIPS)*, 12816–12826.

Mary, X. (2003). *Hilbertian subspaces, subdualities and applications* (Doctoral dissertation) [(https://tel.archives-ouvertes.fr/tel-00007004)]. INSA de Rouen.

Mayne, D., Rawlings, J., Rao, C., & Scokaert, P. (2000). Constrained model predictive control: stability and optimality. *Automatica*, *36*(6), 789–814.

Mercy, T., Van Loock, W., & Pipeleers, G. (2016). Real-time motion planning in the presence of moving obstacles. *European Control Conference (ECC)*, 1586–1591.

Meyer, M. C. (2018). A framework for estimation and inference in generalized additive models with shape and order restrictions. *Statistical Science*, *33*(4), 595–614.

Micheli, M., & Glaunés, J. A. (2014). Matrix-valued kernels for shape deformation analysis. *Geometry, Imaging and Computing*, *1*(1), 57–139.

Ong, C. S., Mary, X., Canu, S., & Smola, A. J. (2004). Learning with non-positive kernels. *International Conference on Machine Learning (ICML)*, 81–88.

Papp, D., & Alizadeh, F. (2014). Shape-constrained estimation using nonnegative splines. *Journal of Computational and Graphical Statistics*, *23*(1), 211–231.

Parzen, E. (1970). Statistical inference on time series by RKHS methods. *Proceedings 12th Biennial Seminar, Canadian Mathematical Congress*.

Petit, N., Milam, M. B., & Murray, R. M. (2001). Inversion based constrained trajectory optimization. *IFAC Proceedings Volumes*, *34*(6), 1211–1216.

Pillonetto, G., Dinuzzo, F., Chen, T., Nicolao, G. [, & Ljung, L. (2014). Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, *50*(3), 657–682.

Pya, N., & Wood, S. N. (2015). Shape constrained additive models. *Statistics and Computing*, *25*, 543–559.

Rosenfeld, J. A., Kamalapurkar, R., Gruss, L. F., & Johnson, T. T. (2019). Dynamic Mode Decomposition for Continuous Time Systems with the Liouville Operator [http://arxiv.org/abs/1910.03977].

Royset, J. O., & Wets, R. J.-B. (2015). Fusion of hard and soft information in nonparametric density estimation. *European Journal of Operational Research*, *247*(2), 532–547.

Saitoh, S. (1997). *Integral transforms, reproducing kernels and their applications*. Longman.

Saitoh, S., & Sawano, Y. (2016). *Theory of reproducing kernels and applications*. Springer Singapore.

Sangnier, M., Fercoq, O., & d'Alché-Buc, F. (2016). Joint quantile regression in vector-valued RKHSs. *Advances in Neural Information Processing Systems (NIPS)*, 3693–3701.

Schölkopf, B., Herbrich, R., & Smola, A. J. (2001). A generalized representer theorem. *Computational Learning Theory (COLT)*, 416–426.

Schölkopf, B., & Smola, A. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press.

Schwartz, L. (1964). Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). *Journal d'Analyse Mathématique*, *13*, 115–256.

Simchi-Levi, D., Chen, X., & Bramel, J. (2014). *The logic of logistics: theory, algorithms, and applications for logistics management*. Springer.

Singh, S., Sindhwani, V., Slotine, J.-J., & Pavone, M. (2018). Learning stabilizable dynamical systems via control contraction metrics [https://arxiv.org/abs/1808.00113]. *Workshop on Algorithmic Foundations of Robotics (WAFR)*.

Sootla, A., Mauroy, A., & Ernst, D. (2018). Optimal control formulation of pulse-based control using Koopman operator. *Automatica*, *91*, 217–224.

Speyer, J. L., & Jacobson, D. H. (2010). *Primer on optimal control theory*. Society for Industrial; Applied Mathematics.

Steinke, F., & Schölkopf, B. (2008). Kernels, regularization and differential equations. *Pattern Recognition*, *41*(11), 3271–3286.

Steinwart, I., & Christmann, A. (2008). *Support vector machines*. Springer.

Takeuchi, I., Le, Q., Sears, T., & Smola, A. (2006). Nonparametric quantile estimation. *Journal of Machine Learning Research*, *7*, 1231–1264.

Tax, D. M. J., & Duin, R. P. W. (2004). Support vector data description. *Machine Learning*, *54*(1), 45–66.

Topkis, D. M. (1998). *Supermodularity and complementarity*. Princeton University Press.

Turlach, B. A. (2005). Shape constrained smoothing using smoothing splines. *Computational Statistics*, *20*, 81–104.

van Keulen, T. (2020). Solution for the continuous-time infinite-horizon linear quadratic regulator subject to scalar state constraints. *IEEE Control Systems Letters*, *4*(1), 133–138.

Wahba, G. (1990). *Spline models for observational data*. SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics.

Williams, M. O., Rowley, C. W., & Kevrekidis, I. G. (2015). A kernel-based method for data-driven Koopman spectral analysis. *Journal of Computational Dynamics*, *2*(2), 247–265.

Wu, X., & Sickles, R. (2018). Semiparametric estimation under shape constraints. *Econometrics and Statistics*, *6*, 74–89.

Zhou, D.-X. (2008). Derivative reproducing properties for kernel methods in learning theory. *Journal of Computational and Applied Mathematics*, *220*, 456–463.

2

KERNEL REGRESSION WITH SHAPE CONSTRAINTS FOR VEHICLE TRAJECTORY RECONSTRUCTION

This chapter was published as a joint work with Nicolas Petit and Zoltán Szabó in the Proceedings of IFAC World Congress 2020, under the title *Kernel Regression for Vehicle Trajectory Reconstruction under Speed and Inter-vehicular Distance Constraints* (Aubin-Frankowski et al., 2020).

Abstract This chapter 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.

Résumé Dans ce chapitre, on s'intéresse au problème de la reconstruction de trajectoires de véhicules en présence d'informations complémentaires sur des contraintes physiques, telles que la distance inter-véhicules et les limites de vitesse. Il est notoirement difficile d'effectuer une régression tout en appliquant ces contraintes sur des intervalles de temps. En nous appuyant sur des espaces de Hilbert à noyau reproduisant, nous proposons une reformulation convexe qui peut être directement implémentée dans des solveurs classiques tels que CVXGEN. Des expériences numériques sur un jeu de données simplifiées illustrent l'efficacité de la méthode, en particulier pour des données en petit nombre et fortement bruitées.

2.1 INTRODUCTION

This chapter 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 & Blair, 1991; Chang et al., 2009). Prime examples of applications can be found in chemical engineering (Arora & Biegler, 2004), biology (Motulsky & 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; De Wit et al. 2015; Liang et al. 2016; Work et al. 2009). 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 & 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 & Sullivan, 2004). Other approaches propose to reject estimates that try to escape the region of the state space where the inequality constraints are satisfied (L.-S. Wang et al., 2002).

Alternatively, in many instances, the problem is recast as a nonlinear programming problem (Arora & 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 & Boyd, 2012).

The paper is organized as follows. In Section 2.2 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 2.3 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 2.4, we illustrate the approach on a real trajectory dataset, underlining the stability of the reconstruction w.r.t. noise. Conclusions are drawn in Section 2.5.

2.2 PROBLEM FORMULATION

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

Notations: Let $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{N} = \{0, 1, \ldots\}$ and $\mathbb{N}^* = \{1, 2, \ldots\}$ stand for the real, non-negative real, natural numbers and positive integers respectively. We use the $[N] := \{1, \ldots, N\}$ 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 $\boldsymbol{v}_1 \in \mathbb{R}^{d_1}, \ldots, \boldsymbol{v}_M \in \mathbb{R}^{d_M}$ is written as $[\boldsymbol{v}_1; \ldots; \boldsymbol{v}_M] \in \mathbb{R}^{\sum_{m\in[M]} d_m}$. 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 $\boldsymbol{v} \in \mathbb{R}^d$ is denoted by $\boldsymbol{v}^T$, its Euclidean norm being $\|\boldsymbol{v}\|_2 = \sqrt{\sum_{i\in[d]} v_i^2}$. 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

$$\min_{f\in\mathcal{F}}\Bigg[\underbrace{\frac{1}{N}\sum_{n\in[N]}|x_n - f(t_n)|^2}_{\text{approximation error}} + \lambda\underbrace{\int_{\mathcal{T}}|f^{(m)}(t)|^2 dt}_{\text{smoothing term}}\Bigg], \tag{2.1}$$

where $\lambda > 0$ defines the trade-off between the approximation error (first term) and smoothness (second term). It is well-known (Berlinet & 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

$$\begin{aligned}
\mathcal{F}_1 &= \operatorname{span}\left(1, t, \ldots, \frac{t^{m-1}}{(m-1)!}\right),\\
\mathcal{F}_2 &= \left\{f \in W_2^m(\mathcal{T}) \mid f^{(j)}(0) = 0 \ (\forall j \in \{0,\ldots,m\})\right.,\\
\|f\|_{\mathcal{F}}^2 &= \sum_{j=0}^{m}|f^{(j)}(0)|^2 + \int_{\mathcal{T}}\left|f^{(m)}(t)\right|^2 dt = \|f_1\|_{\mathcal{F}_1}^2 + \|f_2\|_{\mathcal{F}_2}^2
\end{aligned}$$

where $f = f_1 + f_2$ ($f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2$). In (2.1), the projection onto $\mathcal{F}_1$ is not penalized, and one can rewrite (2.1) as

$$\min_{f\in\mathcal{F}}\left[\frac{1}{N}\sum_{n\in[N]}|x_n - f(t_n)|^2 + \lambda\|f_2\|_{\mathcal{F}_2}^2\right]. \tag{2.2}$$

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

$$\min_{b\in\mathbb{R},\, f\in\mathcal{F}_k}\left[\frac{1}{N}\sum_{n\in[N]}|x_n - (b + f(t_n))|^2 + \lambda\|f\|_{\mathcal{F}_k}^2\right], \tag{2.3}$$

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 $j \in \{0,\ldots,m-1\}$, and $\int_{\mathcal{T}}[f^{(m)}]^2(t)dt < \infty\}$.

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 (2.3) 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:

$$\begin{aligned} \min_{b \in \mathbb{R},\, f \in \mathcal{F}_k} \quad & \left[\frac{1}{N} \sum_{n \in [N]} |x_n - (b + f(t_n))|^2 + \lambda \|f\|_{\mathcal{F}_k}^2 \right] \\ \text{s.t.} \quad & \\ v_{\min} \leqslant f'(t), \quad & \forall t \in \mathcal{T}. \end{aligned} \tag{2.4}$$

Convoy with multiple vehicles ($Q > 1$): We now extend our formulation (2.4) 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 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 $d_{\min} \geqslant 0$ can be formulated as $d_{\min} + b_Q + f_Q(t) \leqslant b_{Q-1} + f_{Q-1}(t)$, $\ldots$, $d_{\min} + b_2 + f_2(t) \leqslant b_1 + f_1(t)$ for all $t \in \mathcal{T}$. This gives our final optimization problem describing the trajectory reconstruction of the convoy:

$$\min_{\substack{f_1,\ldots,f_Q \in \mathcal{F}_k, \\ b_1,\ldots,b_Q \in \mathbb{R}}} \frac{1}{Q} \sum_{q=1}^{Q} \left[\left(\frac{1}{N_q} \sum_{n=1}^{N_q} |x_{q,n} - (b_q + f_q(t_{q,n}))|^2 \right) + \lambda \|f_q\|_{\mathcal{F}_k}^2 \right] \tag{2.5a}$$

s.t.

$$d_{\min} + b_{q+1} + f_{q+1}(t) \leqslant b_q + f_q(t), \forall q \in [Q-1],\ t \in \mathcal{T}, \tag{2.5b}$$

$$v_{\min} \leqslant f_q'(t), \qquad \forall q \in [Q],\ t \in \mathcal{T}. \tag{2.5c}$$

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

2. We will discuss in Section 2.3 the similar $v_{\max}$ requirement.
3. We abbreviate the function $s \in \mathcal{T} \mapsto k(s, t) \in \mathbb{R}$ as $k(\cdot, t)$.

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

$$k_G(t,s) = e^{-(t-s)^2/(2\sigma^2)}, \tag{2.6}$$

$$k_M(t,s) = (1+\sqrt{3}|t-s|/\sigma)e^{-\sqrt{3}|t-s|/\sigma}, \tag{2.7}$$

where $\sigma > 0$ is the « bandwidth ». Constructively, $\mathcal{F}_k = \overline{\text{span}}(k_t : t \in \mathcal{T})$; 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 $t_1,\dots,t_n \in \mathcal{T}$. 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), Saitoh and Sawano (2016), and Steinwart and Christmann (2008).

Having covered our proposed trajectory inference formulation of the convoy (2.5)-(2.5c) and basic properties of RKHSs, in the next section we focus on the optimization of our objective function.

2.3 OPTIMIZATION

The primary challenge one has to resolve in the optimization problem (2.5)-(2.5c) is the infinite number of constraints (due to $\mathcal{T}$) in (2.5b) and (2.5c). 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 (2.5b)-(2.5c) 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 $u_n \geqslant \sup_{t\in\mathcal{T}} g_n(t)$. Then, our strengthened convex optimization problem expresses as:

$$\min_{\substack{a_1,\dots,a_Q\in\mathbb{R}^M,\\ b_1,\dots,b_Q\in\mathbb{R},\\ \alpha_1,\dots,\alpha_Q\in\mathbb{R}^N_+,\\ \beta_1,\dots,\beta_Q\in\mathbb{R}^N_+}} \frac{1}{Q}\sum_{q=1}^{Q}\frac{1}{N_q}\Big[a_q^T(\mathbf{G}_M\boldsymbol{\Pi}_q^T\boldsymbol{\Pi}_q\mathbf{G}_M+\lambda N_q\mathbf{G}_M)a_q \tag{2.8a}$$

$$+N_q b_q^2+2\left(b_q\mathbf{1}_{N_q}-\mathbf{x}_q\right)^T\boldsymbol{\Pi}_q\mathbf{G}_M a_q-2b_q\mathbf{1}_{N_q}^T\mathbf{x}_q\Big]$$

s.t.

$$\kappa\left\|\mathbf{G}_0^{1/2}[a_q-a_{q+1};\alpha_q]\right\|_2 \leqslant b_q-b_{q+1}-d_{\min}-\alpha_q^T\mathbf{u},\ \forall q\in[Q-1],, \tag{2.8b}$$

$$\kappa\left\|\mathbf{G}_D^{1/2}[a_q;\beta_q]\right\|_2 \leqslant -\nu_{\min}-\beta_q^T\mathbf{u},\ \forall q\in[Q], \tag{2.8c}$$

where $\mathbf{x}_q = (x_{q,n})_{n\in[N_q]} \in \mathbb{R}^{N_q}$, $\boldsymbol{\Pi}_q \in \mathbb{R}^{N_q\times M}$ 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 $(k_{t_1},\dots,k_{t_M},g_1,\dots,g_N)$, $\mathbf{G}_D \in \mathbb{R}^{(M+N)\times(M+N)}$ is the Gram matrix of $(\partial_2 k(\cdot,t_1),\dots,\partial_2 k(\cdot,t_M),g_1,\dots,g_N)$, $\mathbf{G}_M \in \mathbb{R}^{M\times M}$ is the Gram matrix of $(k_{t_1},\dots,k_{t_M})$.

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 2.4. Furthermore, constraint (2.8c) 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 (2.8c) by

$$\kappa\left\|\mathbf{G}_m^{1/2}[a_q;\beta_{q,m}]\right\|_2 \leqslant -\nu_{\min}-\beta_{q,m}u_m,\ \forall q\in[Q],\ m\in[M] \tag{2.8d}$$

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

Remarks:

— While the derivation of the optimization problem (2.8a)-(2.8c) 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 (2.5b) for a fixed $q \in [Q-1]$; (2.5c) can be handled similarly. In case of (2.5b), one has to satisfy

$$\sup_{t\in\mathcal{T}}\Big[\underbrace{f_{q+1}(t)-f_q(t)}_{=\langle f_{q+1}-f_q,k_t\rangle_{\mathcal{F}_k}}\Big] \leqslant b_q-b_{q+1}-d_{\min}, \tag{2.9}$$

where on the l.h.s. we applied the *reproducing property* of k. Since $\|k_t\|^2_{\mathcal{F}_k} = \langle k_t, k_t\rangle_{\mathcal{F}_k} = k(t,t) \leqslant \kappa^2$ for any $t \in \mathcal{T}$, it is sufficient for (2.9) to hold that f_q, f_{q+1}, b_q, b_{q+1} satisfy

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

Applying the Cauchy-Schwarz inequality, one gets

$$\kappa \left\|f_{q+1} - f_q\right\|_{\mathcal{F}_k} \leqslant b_q - b_{q+1} - d_{\min}. \tag{2.10}$$

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, (2.10) reads, equivalently, as

$$\kappa \left\|\mathbf{G}_M^{1/2}(\boldsymbol{\alpha}_q - \boldsymbol{\alpha}_{q+1})\right\|_2 \leqslant b_q - b_{q+1} - d_{\min}. \tag{2.11}$$

This inequality is a special case of (2.8b) with $\boldsymbol{\alpha}_q = \mathbf{0}$.

— Solution of (2.8a)-(2.8c) in practice: the convex problem (2.8a)-(2.8c) can be readily implemented with classical solvers such as CVXGEN (Mattingley & 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 & 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 2.2), one can take the matrix square roots $\mathbf{G}^{1/2}$ in (2.8b) and (2.8c), 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 (2.8a)-(2.8c).

— Maximum speed constraint: in addition to the minimum speed constraint (2.5c), one might also have an upper bound on the speed of the vehicles:

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

Such a requirement can be encoded similarly to (2.8c): introducing an additional variable $\boldsymbol{\gamma}_q \in \mathbb{R}^N_+$, it is sufficient to add the

$$\kappa \left\| \mathbf{G}_D^{1/2}[-\boldsymbol{a}_q; \boldsymbol{\gamma}_q] \right\|_2 \leqslant v_{max} - \boldsymbol{\gamma}_q^T \mathbf{u}, \ \forall q \in [Q] \tag{2.12}$$

constraint to the optimization problem (2.8a)-(2.8c).

— Group of convoys: one could also consider an extension of the presented trajectory inference formulation (2.8a)-(2.8c) 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 (2.8b), and change the bound $\mathbf{u}$ accordingly. The less restrictive formulation (2.8d) of (2.8c) exploits the same idea.

2.4 NUMERICAL EXPERIMENTS

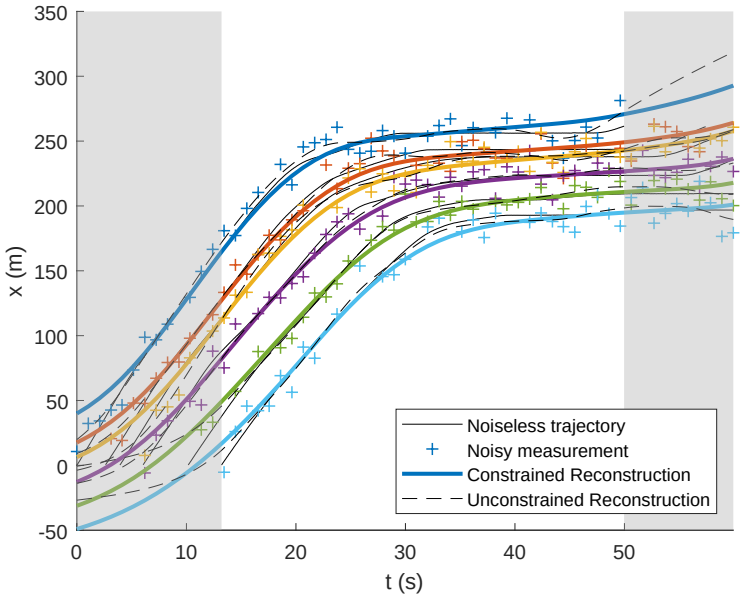
In this section, we demonstrate the efficiency of the proposed trajectory reconstruction approach with speed and inter-vehicular distance constraints. The goal of the experiments is two-fold:

— **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.

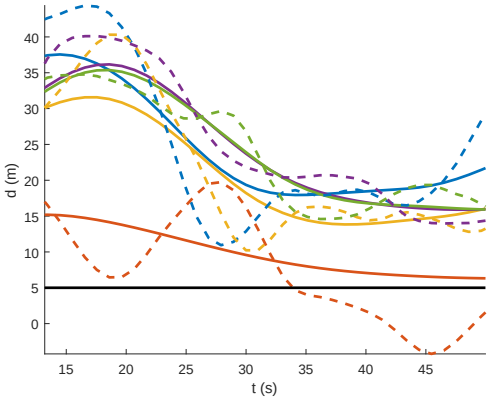
— **Experiment-2**: We complement the first experiment by investigating how severe the error in the trajectory reconstruction is as a function of the measurement noise. For large levels of noise, enforcing the constraints is quite beneficial.

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.

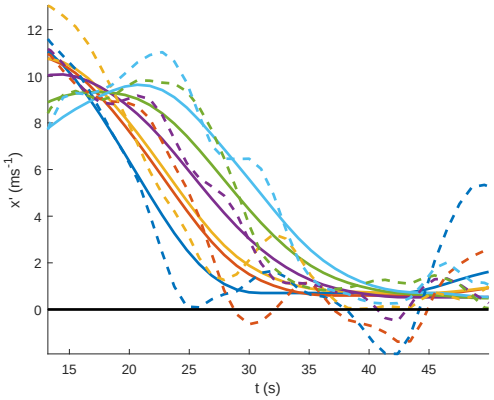
After projecting the two-dimensional position coordinates (x, y) to the distance travelled along the lane (from now on referred to as x), we sub-sample it by decreasing the



(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 $t \mapsto f_q(t) - f_{q+1}(t)$ $(q = 1, \dots, 5)$ 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 2.1 – Reconstruction of the convoy trajectory.

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 (2.8c), we thus change accordingly $v_{\min}$ to $v_{\min} - \bar{v}$. 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 $\sigma_{\text{noise}} = 5\,\text{m}$. We require the vehicles to maintain a distance of at least $d_{\min} = 5\,\text{m}$ and a velocity $v_{\min} = 0\,\text{ms}^{-1}$; the latter bound encodes that cars cannot go backward on a highway. We used the 3/2-Matérn kernel (2.7) 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 & 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 (2.8a)-(2.8b)-(2.8d)) with the unconstrained trajectory estimator (solution of (2.8a) *without* (2.8b)-(2.8d)). The estimated trajectories are depicted in Fig. 2.1a 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. 2.1b and Fig. 2.1c. As it can be read out from Fig. 2.1b, 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. 2.1c 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 $\sigma_{\text{noise}} = 10\,\text{m}$.

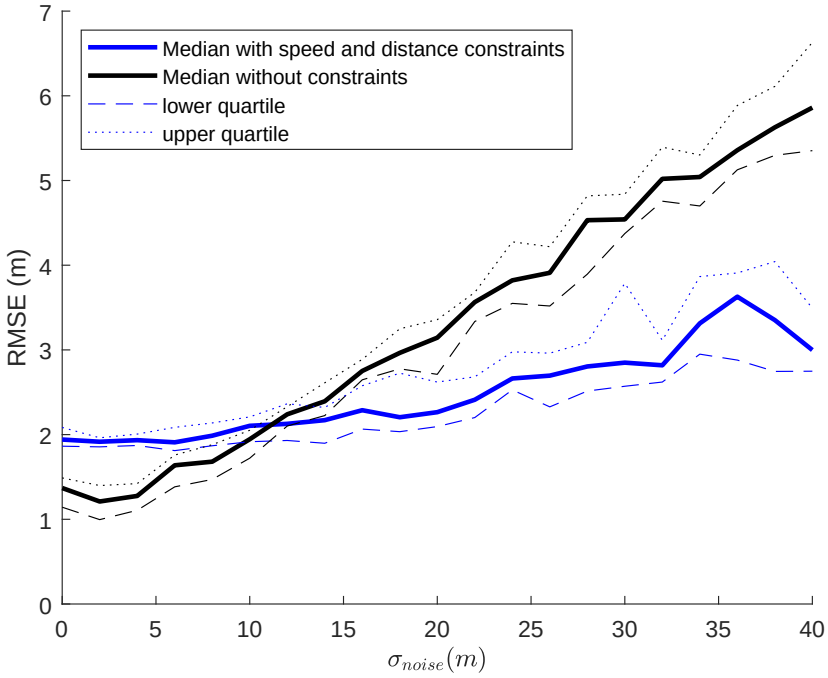


Figure 2.2 – Median, lower and upper quartiles of the RMSE w.r.t. the ground truth trajectories as a function of the noise level σ_{noise}.

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. 2.1a) at the original 25 Hz frequency. The experiment was repeated 40 times and the resulting median, lower and upper quartiles are reported in Fig. 2.2 for varying noise level. One can see that for large noise level (say $\sigma_{\text{noise}} \geqslant 10\,\text{m}$) 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.

2.5 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.

BIBLIOGRAPHY FOR CHAPTER 2

Agate, C. S., & Sullivan, K. J. (2004). Road-constrained target tracking and identification using a particle filter. *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., & Pomante, L. (2016). V2I cooperation for traffic management with SafeCop. *Euromicro Conference on Digital System Design*, 621–627.

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

Alouani, A. T., & Blair, W. D. (1991). Use of a kinematic constraint in tracking constant speed, maneuvering targets. *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., & 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., & Thomas-Agnan, C. (2004). *Reproducing kernel Hilbert spaces in probability and statistics*. Kluwer.

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

Chang, T., Wang, L., & 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., & 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., & Johansson, K. H. (2016). Networked control challenges in collaborative road freight transport. *European Control Conference*, 2–14.

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

Motulsky, H., & 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., & Ulmke, M. (2012). On constraints exploitation for particle filtering based target tracking. *International Conference on Information Fusion*, 455–462.

Rifkin, R. M., & Lippert, R. A. (2007). *Notes on regularized least-squares* (tech. rep.) [(http://cbcl.mit.edu/publications/ps/MIT-CSAIL-TR-2007-025.pdf)].

Saitoh, S., & Sawano, Y. (2016). *Theory of reproducing kernels and applications*. Springer Singapore.

Steinwart, I., & Christmann, A. (2008). *Support vector machines*. Springer.

Tahk, M., & 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., & 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., & Bayen, A. M. (2009). Lagrangian sensing: traffic estimation with mobile devices. *American Control Conference*, 1536–1543.

3

REAL-VALUED AFFINE SHAPE CONSTRAINTS IN RKHSS

This chapter was published as a joint work with Zoltán Szabó in Advances in Neural Information Processing Systems (NeurIPS), under the title *Hard Shape-Constrained Kernel Machines* (Aubin-Frankowski & Szabó, 2020).

Abstract Shape constraints (such as non-negativity, monotonicity, convexity) play a central role in a large number of applications, as they usually improve performance for small sample size and help interpretability. However enforcing these shape requirements in a *hard* fashion is an extremely challenging problem. Classically, this task is tackled (i) in a soft way (without out-of-sample guarantees), (ii) by specialized transformation of the variables on a case-by-case basis, or (iii) by using highly restricted function classes, such as polynomials or polynomial splines. In this chapter, we prove that hard affine shape constraints on function derivatives can be encoded in kernel machines which represent one of the most flexible and powerful tools in machine learning and statistics. Particularly, we present a tightened second-order cone constrained reformulation, that can be readily implemented in convex solvers. We prove performance guarantees on the solution, and demonstrate the efficiency of the approach in joint quantile regression with applications to economics and to the analysis of aircraft trajectories, among others.

Résumé Les contraintes de forme (telles que la positivité, la monotonie, ou la convexité du modèle recherché) jouent un rôle central dans un grand nombre d'applications, car elles améliorent généralement les performances pour les échantillons de petite taille et facilitent l'interprétation des résultats. Cependant, l'application de ces contraintes de forme de manière garantie est un problème extrêmement difficile. Classiquement, cette tâche est abordée (i) par relaxation (sans garanties hors des points de mesure), (ii) par une transformation spécifique des variables au cas par cas, ou (iii) en utilisant des classes de fonctions très restreintes, telles que les polynômes ou les splines polynomiales. Dans ce chapitre, nous prouvons que les contraintes de forme

portant de manière affine sur les dérivées des fonctions peuvent être encodées avec garantie dans des machines à noyau reproduisant. Celles-ci représentent l'un des outils les plus flexibles et les plus puissants en apprentissage automatique et en statistiques. En particulier, nous présentons une reformulation surcontrainte par des contraintes coniques du second ordre, qui peuvent être facilement implémentées dans des solveurs convexes. Nous prouvons des garanties de performance sur la solution, et démontrons l'efficacité de l'approche en régression quantile conjointe, avec des applications à l'économie et à l'analyse des trajectoires d'avion, entre autres domaines.

3.1 INTRODUCTION

Shape constraints (such as non-negativity, monotonicity, convexity) are omnipresent in data science with numerous successful applications in statistics, economics, biology, finance, game theory, reinforcement learning and control problems. For example, in biology, monotone regression techniques have been applied to identify genome interactions (Luss et al., 2012), and in dose-response studies (Hu et al., 2005). Economic theory dictates that utility functions are increasing and concave (Matzkin, 1991), demand functions of normal goods are downward sloping (Blundell et al., 2012; Lewbel, 2010), production functions are concave (Varian, 1984) or S-shaped (Yagi et al., 2020). Moreover cyclic monotonicity has recently turned out to be beneficial in panel multinomial choice problems (Shi et al., 2018), and most link functions used in a single index model are monotone (Balabdaoui et al., 2019; Chen & Samworth, 2016; Li & Racine, 2007). Meanwhile, supermodularity is a common assumption in supply chain models, stochastic multi-period inventory problems, pricing models and game theory (Simchi-Levi et al., 2014; Topkis, 1998). In finance, European and American call option prices are convex and monotone in the underlying stock price and increasing in volatility (Aït-Sahalia & Duarte, 2003). In statistics, the conditional quantile function is increasing w.r.t. the quantile level. In reinforcement learning and in stochastic optimization the value functions are regularly supposed to be convex (Keshavarz et al., 2011; Shapiro et al., 2014). More examples can be found in recent surveys on shape-constrained regression (Chetverikov et al., 2018; Guntuboyina & Sen, 2018; Johnson & Jiang, 2018).

Leveraging prior knowledge expressed in terms of shape structures has several practical benefits: the resulting techniques allow for estimation with smaller sample size and the imposed shape constraints help interpretability. Despite the numerous practical advantages, the construction of shape-constrained estimators can be quite challenging. Existing techniques typically impose the shape constraints (i) in a 'soft'

fashion as a regularizer or at finite many points (Agrell, 2019; Aybat & Wang, 2014; Blundell et al., 2012; Chen & Samworth, 2016; Delecroix et al., 1996; Han et al., 2019; Koppel et al., 2019; Mazumder et al., 2019; Sangnier et al., 2016; Takeuchi et al., 2006; J. Wu et al., 2015; Yagi et al., 2020), (ii) through constraint-specific transformations of the variables such as quadratic reparameterization (Flaxman et al., 2017), positive semi-definite quadratic forms (Bagnell & Farahmand, 2015), or integrated exponential functions (X. Wu & Sickles, 2018), or (iii) they make use of highly restricted functions classes such as classical polynomials (Hall, 2018) or polynomial splines (Koppel et al., 2019; Meyer, 2018; Papp & Alizadeh, 2014; Pya & Wood, 2015; Turlach, 2005; X. Wu & Sickles, 2018). Both the polynomial and spline-based shape-constrained techniques rely heavily on the underlying algebraic structure of these bases, without direct extension to more general function families.

From a statistical viewpoint, the main focus has been on the design of estimators with uniform guarantees (Freyberger & Reeves, 2018; Horowitz & Lee, 2017). Several existing methods have been analyzed from this perspective and were shown to be (uniformly) consistent, on a case-by-case basis and when handling specific shape constraints (Chen & Samworth, 2016; Han et al., 2019; Han & Wellner, 2016; Koppel et al., 2019; Mazumder et al., 2019; J. Wu et al., 2015; Yagi et al., 2020). While these asymptotic results are of significant theoretical interest, applying shape priors is generally beneficial in the small sample regime. In this paper we propose a flexible **optimization framework** allowing multiple shape constraints to be jointly handled in a hard fashion. In addition, to address the bottlenecks of restricted shape priors and function families, we consider general affine constraints on derivatives, and use reproducing kernel Hilbert spaces (RKHS) as hypothesis space.

RKHSs (also called abstract splines; Aronszajn, 1950; Berlinet and Thomas-Agnan, 2004; Wahba, 1990; Wang, 2011) increase significantly the richness and modelling power of classical polynomial splines. Indeed, the resulting function family can be rich enough for instance (i) to encode probability distributions without loss of information (Fukumizu et al., 2008; Sriperumbudur et al., 2010), (ii) to characterize statistical independence of random variables (Bach & Jordan, 2002; Szabó & Sriperumbudur, 2018), or (iii) to approximate various function classes arbitrarily well (Carmeli et al., 2010; Micchelli et al., 2006; Simon-Gabriel & Schölkopf, 2018; Sriperumbudur et al., 2011; Steinwart, 2001), including the space of bounded continuous functions. An excellent overview on kernels and RKHSs is given by Hofmann et al. (2008), Saitoh and Sawano (2016), and Steinwart and Christmann (2008).

In this paper we incorporate into this flexible RKHS function class the possibility to impose *hard* linear shape requirements on derivatives, i.e. constraints of the form

$$0 \leqslant b + Df(\mathbf{x}) \quad \forall \mathbf{x} \in K \tag{3.1}$$

for a bias $b \in \mathbb{R}$, given a differential operator $D = \sum_j \gamma_j \partial^{r_j}$ where $\partial^r f(x) = \frac{\partial^{\sum_{j=1}^d r_j} f(x)}{\partial_{x_1}^{r_1} \cdots \partial_{x_d}^{r_d}}$ and a compact set $K \subset \mathbb{R}^d$. The fundamental technical challenge is to guarantee the pointwise inequality (3.1) at *all* points of the (typically non-finite) set K. We show that one can tighten the infinite number of affine constraints (3.1) into a finite number of second-order cone constraints

$$\eta \|f\| \leqslant b + Df(x_m) \quad \forall m \in \{1, \ldots, M\} \tag{3.2}$$

for a suitable choice of $\eta > 0$ and $\{x_m\}_{m=1\ldots M} \subseteq K$.

Our **contributions** can be summarized as follows.

1. We show that hard shape requirements can be embedded in kernel machines by taking a second-order cone (SOC) tightening of constraint (3.1), which can be readily tackled by convex solvers. Our formulation builds upon the reproducing property for kernel derivatives and on coverings of compact sets.
2. We prove error bounds on the distance between the solutions of the strengthened and original problems.
3. We achieve state-of-the-art performance in joint quantile regression (JQR) in RKHSs. We also combine JQR with other shape constraints in economics and in the analysis of aircraft trajectories.

The paper is structured as follows. Section 3.2 is about problem formulation. Our main result is presented in Section 3.3. Numerical illustrations are given in Section 3.4. Proofs and additional examples are provided in the supplement.

3.2 PROBLEM FORMULATION

In this section we formulate our problem after introducing some notations, which the reader may skip at first, and return to if necessary.

Notations: Let $\mathbb{N} := \{0, 1, \ldots\}$, $\mathbb{N}^* := \{1, 2, \ldots\}$ and $\mathbb{R}_+$ denote the set of natural numbers, positive integers and non-negative real numbers, respectively. We use the shorthand $[n] := \{1, \ldots, n\}$. The p-norm of a vector $v \in \mathbb{R}^p$ is $\|v\|_p = (\sum_{j\in[d]} |v_j|^p)^{\frac{1}{p}}$ $(1 \leqslant p < \infty)$; $\|v\|_\infty = \max_{j\in[d]} |v_j|$. The j-th canonical basis vector is e_j; $0_d \in \mathbb{R}^d$ is the zero vector. Let $\mathbb{B}_{\|\cdot\|}(c, r) = \{x \in \mathbb{R}^d : \|x - c\| \leqslant r\}$ be the closed ball in $\mathbb{R}^d$ with center c and radius r in norm $\|\cdot\|$. Given a norm $\|\cdot\|$ and radius $\delta > 0$, a δ-net of a compact set $K \subset \mathbb{R}^d$ consists of a set of points $\{x_m\}_{m\in[M]}$ such that $K \subseteq \cup_{m\in[M]} \mathbb{B}_{\|\cdot\|}(x_m, \delta)$, in other words $\left\{\mathbb{B}_{\|\cdot\|}(x_m, \delta)\right\}_{m\in[M]}$ forms a covering of K. The identity matrix is I. For a matrix $M \in \mathbb{R}^{d_1 \times d_2}$, $M^\top \in \mathbb{R}^{d_2 \times d_1}$ denotes its transpose, its operator norm is

$\|\boldsymbol{M}\| = \sup_{x\in\mathbb{R}^{d_2}:\|x\|_2=1} \|\boldsymbol{M}x\|_2$. The inverse of a non-singular matrix $\boldsymbol{M} \in \mathbb{R}^{d\times d}$ is $\boldsymbol{M}^{-1} \in \mathbb{R}^{d\times d}$. The concatenation of matrices $\boldsymbol{M}_1 \in \mathbb{R}^{d_1\times d}, \ldots, \boldsymbol{M}_N \in \mathbb{R}^{d_N\times d}$ is denoted by $\boldsymbol{M} = [\boldsymbol{M}_1; \ldots; \boldsymbol{M}_N] \in \mathbb{R}^{\left(\sum_{n\in[N]} d_n\right)\times d}$. Let $\mathcal{X}$ be an open subset of $\mathbb{R}^d$ with a real-valued kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, and associated reproducing kernel Hilbert space (RKHS) $\mathcal{F}_k$. The Hilbert space $\mathcal{F}_k$ is characterized by $f(x) = \langle f, k(x, \cdot)\rangle_k$ $(\forall x \in \mathcal{X}, \forall f \in \mathcal{F}_k)$ and $k(x, \cdot) \in \mathcal{F}_k$ $(\forall x \in \mathcal{X})$ where $\langle\cdot,\cdot\rangle_k$ stands for the inner product on $\mathcal{F}_k$, and $k(x, \cdot)$ denotes the function $y \in \mathcal{X} \mapsto k(x, y) \in \mathbb{R}$. The first property is called the reproducing property, the second one describes a generating family of $\mathcal{F}_k$. The norm on $\mathcal{F}_k$ is written as $\|\cdot\|_k$. For a multi-index $r \in \mathbb{N}^d$ let the r-th order partial derivative of a function f be denoted by $\partial^r f(x) = \frac{\partial^{|r|} f(x)}{\partial x_1^{r_1} \cdots \partial x_d^{r_d}}$ where $|r| = \sum_{j\in[d]} r_j$ is the length of r. When $d = 1$ the $f^{(n)} = \partial^n f$ notation is applied; specifically f'' and f' are used in case of $n = 2$ and $n = 1$. Given $s \in \mathbb{N}$, let $\mathcal{C}^s(\mathcal{X})$ be the set of real-valued functions on $\mathcal{X}$ with continuous derivatives up to order s (i.e., $\partial^r f \in \mathcal{C}(\mathcal{X}) := \mathcal{C}^0(\mathcal{X})$ when $|r| \leqslant s$). Let $I \in \mathbb{N}^*$. Given $(A_i)_{i\in[I]}$ sets let $\prod_{i\in[I]} A_i$ denote their Cartesian product; we write A^I in case of $A = A_1 = \ldots = A_I$.

Our **goal** is to solve hard shape-constrained kernel machines of the form

$$\left(\bar{\boldsymbol{f}}, \bar{\boldsymbol{b}}\right) = \underset{\boldsymbol{f}=(f_q)_{q\in[Q]} \in (\mathcal{F}_k)^Q,\, \boldsymbol{b}=(b_p)_{p\in[P]} \in \mathcal{B},\, (\boldsymbol{f},\boldsymbol{b}) \in C}{\arg\min} \mathcal{L}(\boldsymbol{f}, \boldsymbol{b}), \tag{$\mathcal{P}$}$$

where we are given an objective function $\mathcal{L}$ and a constraint set C (detailed below), a closed convex constraint set $\mathcal{B} \subseteq \mathbb{R}^P$ on the biases, an order $s \in \mathbb{N}$, an open set $\mathcal{X} \subseteq \mathbb{R}^d$ with a kernel $k \in \mathcal{C}^s(\mathcal{X} \times \mathcal{X})$ and associated RKHS $\mathcal{F}_k$, and samples $S = \{(x_n, y_n)\}_{n\in[N]} \subset \mathcal{X} \times \mathbb{R}$. The objective function in ($\mathcal{P}$) is specified by the triplet (S, L, Ω):

$$\mathcal{L}(\boldsymbol{f}, \boldsymbol{b}) = L\left(\boldsymbol{b}, \left(x_n, y_n, (f_q(x_n))_{q\in[Q]}\right)_{n\in[N]}\right) + \Omega\left((\|f_q\|_k)_{q\in[Q]}\right),$$

with loss function $L : \mathbb{R}^P \times \left(\mathcal{X} \times \mathbb{R} \times \mathbb{R}^Q\right)^N \to \mathbb{R}$ and regularizer $\Omega : (\mathbb{R}_+)^Q \to \mathbb{R}$. Notice that the objective $\mathcal{L}$ depends on the samples S which are assumed to be fixed, hence our proposed optimization framework focuses on the empirical risk. The bias $\boldsymbol{b} \in \mathbb{R}^P$ can be both constraint (such as $f(x) \geqslant b_1$, $f'(x) \geqslant b_2$) and variable-related ($f_q + b_q$, see (3.4)-(3.5) later). The restriction of L to $\mathcal{B}$ is assumed to be strictly convex in $\boldsymbol{b}$, and Ω is supposed to be strictly increasing in each of its arguments to ensure the uniqueness of minimizers of ($\mathcal{P}$).

The $I \in \mathbb{N}^*$ hard shape requirements in ($\mathcal{P}$) take the form [1]

$$C = \{(\boldsymbol{f}, \boldsymbol{b}) \,|\, (\boldsymbol{b}_0 - \boldsymbol{U}\boldsymbol{b})_i \leqslant D_i(\boldsymbol{W}\boldsymbol{f} - \boldsymbol{f}_0)_i(x), \forall x \in K_i, \forall i \in [I]\}, \tag{$\mathcal{C}$}$$

1. In constraint ($\mathcal{C}$), $\boldsymbol{W}\boldsymbol{f}$ is meant as a formal matrix-vector product: $(\boldsymbol{W}\boldsymbol{f})_i = \sum_{q\in[Q]} W_{i,q} f_q$.

i.e. (𝒞) encodes affine constraints of at most s-order derivatives ($D_i = \sum_{j\in[n_{i,j}]} \gamma_{i,j}\partial^{\mathbf{r}_{i,j}}$, $|\mathbf{r}_{i,j}| \leqslant s$, $\gamma_{i,j} \in \mathbb{R}$). Possible shifts are expressed by the terms $\mathbf{b}_0 = (b_{0,i})_{i\in[I]} \in \mathbb{R}^I$, $\mathbf{f}_0 = (f_{0,i})_{i\in[I]} \in (\mathcal{F}_k)^I$. The matrices $\mathbf{U} \in \mathbb{R}^{I\times P}$ and $\boldsymbol{W} \in \mathbb{R}^{I\times Q}$ capture the potential interactions within the bias variables $(b_p)_{p\in[P]}$ and functions $(f_q)_{q\in[Q]}$, respectively. The compact sets $K_i \subset \mathcal{X}$ ($i \in [I]$) define the domain where the constraints are imposed.

Remarks:

— Differential operators: As $\mathcal{X} \subseteq \mathbb{R}^d$ is open and $k \in \mathcal{C}^s(\mathcal{X}\times\mathcal{X})$, any differential operator D_i of order at most s is well defined (Saitoh & Sawano, 2016, Theorems 2.5 and 2.6, page 76) as a mapping from $\mathcal{F}_k$ to $\mathcal{C}(\mathcal{X})$. Since the coefficients $\{\gamma_{i,j}\}_{j\in[n_{i,j}]}$ of D_i-s belong to the whole $\mathbb{R}$, (𝒞) can cover inequality constraints in both directions.

— Bias constraint $\mathcal{B}$: Choosing $\mathcal{B} = \{\mathbf{0}_P\}$ leads to constant l.h.s. $\mathbf{b}_0$ in (𝒞). The other extreme is $\mathcal{B} = \mathbb{R}^P$ in which case no hard constraint is imposed on the bias variable $\mathbf{b}$.

— Compactness of K_i-s: The compactness assumption on the sets K_i is exploited in the construction of our strengthened optimization problem (Section 3.3). This requirement also ensures not imposing restrictions "too far" from the observation points, which could be impossible to satisfy. Indeed, let us consider for instance a c_0-kernel k on $\mathbb{R}$, i.e. that $k(x,\cdot) \in \mathcal{C}^0(\mathbb{R})$ for all x and $\lim_{|y|\to\infty} k(x,y) = 0$ for all $x \in \mathbb{R}$ (as for the Gaussian kernel). In this case $\lim_{|y|\to\infty} f(y) = 0$ also holds for all $f \in \mathcal{F}_k$. Hence a constraint of the form "for all $t \in \mathbb{R}_+$, $f(t) \geqslant \epsilon > 0$" can *not* be satisfied for $f \in \mathcal{F}_k$.

— Assumption on $\mathcal{X}$: If $s = 0$ (in other words only function evaluations are present in the shape constraints), then $\mathcal{X}$ can be any topological space.

We give various **examples** for the considered problem family (𝒫). We start with an example where $Q = 1$.

Kernel ridge regression (KRR) with *monotonicity* constraint: In this case the objective function and constraint are

$$\mathcal{L}(f,b) := \frac{1}{N}\sum_{n\in[N]} |y_n - f(x_n)|^2 + \lambda_f \|f\|_k^2, \text{ s.t. } f'(x) \geqslant 0,\ \forall x \in [x_l, x_u] \tag{3.3}$$

with $\lambda_f > 0$. In other words in (𝒫) we have $Q = 1$, $d = 1$, $s = 1$, $P = I = 1$, $K_1 = [x_l, x_u]$, $\Omega(z) = \lambda_f z^2$, $D_1 = \partial^1$, $U = W = 1$, $f_{1,0} = 0$, $b_{1,0} = 0$, and $b \in \mathcal{B} = \{0\}$ is a dummy variable.

Joint quantile regression (JQR; e.g. Sangnier et al., 2016): Given $0 < \tau_1 < \ldots < \tau_Q < 1$ levels the goal is to estimate *jointly* the $(\tau_1, \ldots, \tau_Q)$-quantiles of the conditional distribution $\mathbb{P}(Y|X = \mathbf{x})$ where Y is real-valued. In this case the objective function is

$$\mathcal{L}(\mathbf{f}, \mathbf{b}) = \frac{1}{N} \sum_{q \in [Q]} \sum_{n \in [N]} l_{\tau_q}(y_n - [f_q(\mathbf{x}_n) + b_q]) + \lambda_b \|\mathbf{b}\|_2^2 + \lambda_f \sum_{q \in [Q]} \|f_q\|_k^2, \tag{3.4}$$

where $\lambda_b > 0$, $\lambda_f > 0$,[2] and the pinball loss is defined as $l_\tau(e) = \max(\tau e, (\tau - 1)e)$ with $\tau \in (0, 1)$. In JQR, the estimated τ_q-quantile functions $\{f_q + b_q\}_{q \in [Q]}$ are *not* independent; the joint shape constraint they should satisfy is the monotonically increasing property w.r.t. the quantile level τ. It is natural to impose this *non-crossing* requirement on the smallest rectangle containing the points $\{\mathbf{x}_n\}_{n \in [N]}$, i.e. $K = \prod_{r \in [d]} \left[\min\{(\mathbf{x}_n)_r\}_{n \in [N]}, \max\{(\mathbf{x}_n)_r\}_{n \in [N]}\right]$. The corresponding shape constraint is

$$f_{q+1}(\mathbf{x}) + b_{q+1} \geqslant f_q(\mathbf{x}) + b_q, \ \forall q \in [Q - 1], \ \forall \mathbf{x} \in K. \tag{3.5}$$

One gets (3.4)-(3.5) from ($\mathcal{P}$) by choosing $P = Q$, $I = Q - 1$, $s = 0$, $\mathbf{b}_0 = \mathbf{0}$, $f_0 = 0$, $\mathcal{B} = \mathbb{R}^I$, $K_i = K$ $(\forall i \in [I])$, $\Omega(\mathbf{z}) = \lambda_f \sum_{q \in [Q]} (z_q)^2$, and

$$\mathbf{U} = \mathbf{W} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(Q-1) \times Q}.$$

Further examples: There are various other widely-used shape constraints beyond non-negativity (for which $s = 0$), monotonicity ($s = 1$) or convexity ($s = 2$) which can be taken into account in ($\mathcal{C}$). For instance one can consider n-monotonicity ($s = n$), $(n - 1)$-alternating monotonicity, monotonicity w.r.t. unordered weak majorization ($s = 1$) or w.r.t. product ordering ($s = 1$), or supermodularity ($s = 2$). For details on how these shape constraints can be written as ($\mathcal{C}$), see the supplement (Section 3.5.3).

3.3 RESULTS

In this section, we first present our strengthened SOC-constrained problem, followed by a representer theorem and explicit bounds on the distance to the solution of ($\mathcal{P}$).

2. Sangnier et al. (2016) considered the same loss but a *soft* non-crossing inducing regularizer inspired by matrix-valued kernels, and also set $\lambda_b = 0$.

In order to introduce our proposed tightening, let us first consider the discretization of the I constraints using M_i points $\{\tilde{x}_{i,m}\}_{m\in[M_i]} \subseteq K_i$. This would lead to the following relaxation of ($\mathcal{P}$)

$$v_{\text{disc}} = \min_{f\in(\mathcal{F}_k)^Q,\, b\in\mathcal{B}} \mathcal{L}(f,b) \text{ s.t. } (b_0 - Ub)_i \leqslant \min_{m\in[M_i]} D_i(Wf - f_0)_i(\tilde{x}_{i,m}) \;\forall i\in[I]. \tag{3.6}$$

Our proposed SOC-constrained tightening can be thought of as adding extra, appropriately chosen, η_i-buffers to the l.h.s. of the constraints:

$$(f_\eta, b_\eta) = \operatorname*{arg\,min}_{f\in(\mathcal{F}_k)^Q,\, b\in\mathcal{B}\subset\mathbb{R}^p} \mathcal{L}(f,b) \tag{$\mathcal{P}_\eta$}$$

$$\text{s.t.}$$

$$(b_0 - Ub)_i + \eta_i\|(Wf - f_0)_i\|_k \leqslant \min_{m\in[M_i]} D_i(Wf - f_0)_i(\tilde{x}_{i,m})\;\forall i\in[I]. \tag{$\mathcal{C}_\eta$}$$

The SOC constraint ($\mathcal{C}_\eta$)[3] is determined by a fixed $\eta = (\eta_i)_{i\in[I]} \in \mathbb{R}^I_+$ coefficient vector and by the points $\{\tilde{x}_{i,m}\}$. For each $i\in[I]$, the points $\{\tilde{x}_{i,m}\}_{m\in[M_i]}$ are chosen to form a δ_i-net of the compact set K_i for some $\delta_i > 0$ and a pre-specified norm $\|\cdot\|_X$.[4] Given $\{\tilde{x}_{i,m}\}_{m\in[M_i]}$, the coefficients $\eta_i \in \mathbb{R}_+$ are then defined as

$$\eta_i = \sup_{m\in[M_i],\, u\in\mathbb{B}_{\|\cdot\|_X}(0,1)} \|D_{i,x}k(\tilde{x}_{i,m},\cdot) - D_{i,x}k(\tilde{x}_{i,m} + \delta_i u,\cdot)\|_k, \tag{3.8}$$

where $D_{i,x}k(x_0,\cdot)$ is a shorthand for $y \mapsto D_i(x \mapsto k(x,y))(x_0)$. Problem ($\mathcal{P}_\eta$) has I scalar SOC constraints ($\mathcal{C}_\eta$) over infinite-dimensional variables. Let $\bar{v} = \mathcal{L}(\bar{f},\bar{b})$ be the minimal value of ($\mathcal{P}$) and $v_\eta = \mathcal{L}(f_\eta, b_\eta)$ be that of the discretization ($\mathcal{P}_\eta$). Notice that, when formally setting $\eta = 0$, ($\mathcal{P}_\eta$) corresponds to (3.6).

In our main result below (i) shows that ($\mathcal{C}_\eta$) is indeed a tightening of ($\mathcal{C}$), (ii) provides a representer theorem which allows to solve numerically ($\mathcal{P}_\eta$), and (iii) gives bounds on the difference between the solution of ($\mathcal{P}_\eta$) and that of ($\mathcal{P}$) as a function of $(v_\eta - v_{\text{disc}})$ and η respectively.

Theorem 3.1 (Tightened task, representer theorem, bounds). *Let* $f_\eta = (f_{\eta,q})_{q\in[Q]}$. *Then,*

(i) Tightening: any (f,b) *satisfying* ($\mathcal{C}_\eta$) *also satisfies* ($\mathcal{C}$), *hence* $v_{disc} \leqslant \bar{v} \leqslant v_\eta$.

(ii) Representer theorem: For all $q\in[Q]$, *there exist real coefficients* $\tilde{a}_{i,0,q}, \tilde{a}_{i,m,q}, a_{n,q} \in \mathbb{R}$ *such that* $f_{\eta,q} = \sum_{i\in[I]}\left[\tilde{a}_{i,0,q} f_{0,i} + \sum_{m\in[M_i]} \tilde{a}_{i,m,q} D_{i,x}k(\tilde{x}_{i,m},\cdot)\right] + \sum_{n\in[N]} a_{n,q} k(x_n,\cdot)$.

3. Constraint ($\mathcal{C}_\eta$) is the precise meaning of the preliminary intuition given in (3.2).

4. The existence of finite δ_i-nets ($M_i < \infty$) stems from the compactness of K_i-s. The flexibility in the choice of the norm $\|\cdot\|_X$ allows for instance using cubes by taking the $\|\cdot\|_1$ or the $\|\cdot\|_\infty$-norm on $\mathbb{R}^d$ when covering the K_i-s.

(iii) Performance guarantee: if $\mathcal{L}$ *is* (μ_{f_q}, μ_b)*-strongly convex w.r.t.* $(f_q, \mathbf{b})$ *for any* $q \in [Q]$*, then*

$$\|f_{\eta,q} - \bar{f}_q\|_k \leqslant \sqrt{\frac{2(\nu_\eta - \nu_{disc})}{\mu_{f_q}}}, \qquad \|\mathbf{b}_\eta - \bar{\mathbf{b}}\|_2 \leqslant \sqrt{\frac{2(\nu_\eta - \nu_{disc})}{\mu_b}}. \tag{3.9}$$

If in addition $\mathbf{U}$ *is of full row-rank (i.e. surjective),* $\mathcal{B} = \mathbb{R}^P$*, and* $\mathcal{L}(\bar{\mathbf{f}}, \cdot)$ *is* L_b*—Lipschitz continuous on* $\mathbb{B}_{\|\cdot\|_2}(\bar{\mathbf{b}}, c_f\|\boldsymbol{\eta}\|_\infty)$ *where* $c_f = \sqrt{I}\left\|\left(\mathbf{U}^\top\mathbf{U}\right)^{-1}\mathbf{U}^\top\right\| \max_{i\in[I]}\left\|(\mathbf{W}\bar{\mathbf{f}} - \mathbf{f}_0)_i\right\|_k$*, then*

$$\|f_{\eta,q} - \bar{f}_q\|_k \leqslant \sqrt{\frac{2L_b c_f\|\boldsymbol{\eta}\|_\infty}{\mu_{f_q}}}, \qquad \|\mathbf{b}_\eta - \bar{\mathbf{b}}\|_2 \leqslant \sqrt{\frac{2L_b c_f\|\boldsymbol{\eta}\|_\infty}{\mu_b}}. \tag{3.10}$$

Proof (idea): The SOC-based reformulation relies on rewriting the constraint ($\mathcal{C}$) as the inclusion of the sets $\Phi_{D_i}(K_i)$ in the closed halfspaces $H^+_{\phi_i,\beta_i} := \{g \in \mathcal{F}_k \,|\, \langle \phi_i, g\rangle_k \geqslant \beta_i\}$ for $\forall i \in [I]$ where $\Phi_{D_i}(\mathbf{x}) := D_{i,\mathbf{x}}k(\mathbf{x}, \cdot) \in \mathcal{F}_k$, $\Phi_{D_i}(X) := \{\Phi_{D_i}(\mathbf{x}) \,|\, \mathbf{x} \in X\}$, $\phi_i := (\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i$ and $\beta_i := (\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i$. The tightening is obtained by guaranteeing these inclusions with an η_i-net of $\Phi_{D_i}(K_i)$ containing the δ_i-net of K_i when pushed to $\mathcal{F}_k$. The bounds stem from classical inequalities for strongly convex objective functions. The proof details of (i)-(iii) are available in the supplement (Section 3.5.1).

Remarks:

The **representer theorem** allows one to express ($\mathcal{P}_\eta$) as a finite-dimensional SOC-constrained problem:

$$\min_{\substack{\mathbf{A}\in\mathbb{R}^{N\times Q},\\ \mathbf{b}\in\mathcal{B},\\ \tilde{\mathbf{A}}\in\mathbb{R}^{N\times Q},\\ \tilde{\mathbf{A}}_0\in\mathbb{R}^{I\times Q}}} \mathcal{L}(\mathbf{f}, \mathbf{b}) \text{ s.t. } (\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i + \eta_i \left\|\mathbf{G}^{1/2}\mathbf{g}_i\right\|_2 \leqslant \min_{m\in[M_i]} \left(\mathbf{G}_{D_i}\mathbf{g}_i\right)_{I+N+m} \; \forall i \in [I], \forall q \in [Q],$$

where $\tilde{\boldsymbol{e}}_i \in \mathbb{R}^I$ and $\boldsymbol{e}_i \in \mathbb{R}^{I+N+M}$ are the canonical basis vectors, $\mathbf{g}_i := \left[\tilde{\mathbf{A}}_0; \mathbf{A}; \tilde{\mathbf{A}}\right]\mathbf{W}^\top\tilde{\boldsymbol{e}}_i - \boldsymbol{e}_i$ and the coefficients of the components of $\mathbf{f}$ were collected to $\tilde{\mathbf{A}}_0 = [\tilde{a}_{i,0,q}]_{i\in[I],\, q\in[Q]} \in \mathbb{R}^{I\times Q}$, $\mathbf{A} = [a_{n,q}]_{n\in[N],\, q\in[Q]} \in \mathbb{R}^{N\times Q}$, $\tilde{\mathbf{A}} = \left[\tilde{\mathbf{a}}_{i,q}\right]_{i\in[I],\, q\in[Q]} \in \mathbb{R}^{M\times Q}$ with $M = \sum_{i\in[I]} M_i$ and $\tilde{\mathbf{a}}_{i,q} = \left[\tilde{a}_{i,m,q}\right]_{m\in[M_i]} \in \mathbb{R}^{M_i}$ ($i \in [I]$, $q \in [Q]$). In this finite-dimensional optimization task, $\mathbf{G} \in \mathbb{R}^{(I+N+M)\times(I+N+M)}$ is the Gram matrix of $\{f_{0,i}\}_{i\in I}$, $\{k(\mathbf{x}_n, \cdot)\}_{n\in[N]}$, and $\{D_{i,\mathbf{x}}k(\tilde{\mathbf{x}}_{i,m}, \cdot)\}_{m\in[M_i], i\in I}$), $\mathbf{G}_{D_i} \in \mathbb{R}^{(I+N+M)\times(I+N+M)}$ is the Gram matrix of the differentials D_i of these functions, $\mathbf{G}^{1/2}$ is the matrix square root of the positive semi-definite $\mathbf{G}$.

The **bounds**[5] (3.9)-(3.10) show that smaller $\boldsymbol{\eta}$ gives tighter guarantee on the recovery of $\bar{f}$ and $\bar{b}$. Since $\left|\partial^r f_{\eta,q}(x) - \partial^r \bar{f}_q(x)\right| \leqslant \sqrt{\partial^{r,r} k(x,x)}\left\|f_{\eta,q} - \bar{f}_q\right\|_k$ by the reproducing property and the Cauchy-Schwarz inequality, the bounds on $\|f_\eta - \bar{f}\|_k$ can be propagated to pointwise bounds on the derivatives (for $|r| \leqslant s$). We emphasize again that in our optimization problem ($\mathcal{P}_\eta$) the samples $S = \{(x_n, y_n)\}_{n\in[N]}$ are assumed to be fixed; in other words the bounds (3.9) and (3.10) are meant conditioned on S.

The **parameters** M, δ and $\boldsymbol{\eta}$ are strongly intertwined, their interplay reveals an *accuracy-computation tradeoff*. Consider a shift-invariant kernel ($k(x,y) = k_0(x-y)$, $\forall x, y$), then (3.8) simplifies to $\eta_i := \sup_{u\in\mathbb{B}_{\|\cdot\|_X}(0,1)}\sqrt{\left|2D_{i,x}D_{i,y}k_0(0) - 2D_{i,x}D_{i,y}k_0(\delta_i u)\right|}$, where $D_{i,y}$ is defined similarly to $D_{i,x}$. This expression of η_i implies that whenever $D_{i,x}D_{i,y}k_0$ is L_δ-Lipschitz[6] on $\mathbb{B}_{\|\cdot\|_X}(0,\delta_i)$, then $\eta_i \leqslant \sqrt{2L_\delta}\sqrt{\delta_i}$. By the previous point, a smaller $\boldsymbol{\eta}$ ensures a better recovery which can be guaranteed by smaller δ_i-s, which themselves correspond to a larger number of centers (M_i-s) in the δ_i-nets of the K_i-s. Hence, one can control the computational complexity by the total number M of points in the nets. Indeed, most SOCP solvers rely on primal-dual interior point methods which have (in the worst-case) cubic complexity $\mathcal{O}\left((P+N+M)^3\right)$ per iterations (Alizadeh & Goldfarb, 2003). Controlling M allows one to tackle hard shape-constrained problems in moderate dimensions (d); for details see Section 3.4. In practice, to reduce the number of coefficients in $f_{\eta,q}$, it is beneficial to recycle the points $\{x_n\}_{n\in[N]}$ among the M_i virtual centers, whenever the points belong to a constraint set K_i. This simple trick was possible in all our numerical examples and kept the computational expense quite benign. Supplement (Section 3.5.2) presents an example of the actual computational complexity observed.

While in this work we focused on the optimization problem ($\mathcal{P}$) which contains *solely* infinite-dimensional SOC constraints ($\mathcal{C}_\eta$), the proved ($\mathcal{C}_\eta$) $\Rightarrow$ ($\mathcal{C}$) implication can be of independent interest to tackle problems where other types of constraints are present.[7] For simplicity we formulated our result with uniform coverings (δ_i, η_i, $i \in [I]$). However, we prove it for more general non-uniform coverings ($\delta_{i,m}$, $\eta_{i,m}$, $i \in [I]$, $m \in [M_i]$; see Section 3.5.1). This can beneficial for sets with complex geometry (e.g. star-shaped) or when recyling of the samples was used to obtain coverings (as the samples in S have no reason to be equally spaced); we provide an example (in economics) using a non-uniform covering in Section 3.4.

5. Notice that (3.9) is a computable bound, while (3.10) is not, as the latter depends on properties of the unknown solution of ($\mathcal{P}$).

6. For instance any C^{s+1} kernel satisfies this local Lipschitz requirement.

7. For example having a unit integral is a natural additional requirement beyond non-negativity in density estimation, and writes as a linear equality constraint over the coefficients of $f_{\eta,q}$.

In practice, since the convergence speed of SOCP solvers depends highly on the condition number of $\mathbf{G}^{1/2}$, it is worth replacing $\mathbf{G}^{1/2}$ with $(\mathbf{G} + \epsilon_{\text{tol}}\mathbf{I})^{1/2}$, setting a tolerance $\epsilon_{\text{tol}} \simeq 10^{-4}$. As $\mathbf{G} + \epsilon_{\text{tol}}\mathbf{I} \succcurlyeq \mathbf{G}$ (in the sense of positive semi-definite matrices), this regularization strengthens further the SOC constraint. Moreover, SOCP modeling frameworks (e.g. CVXPY or CVXGEN) suggest to replace quadratic penalties (see (3.4)) with the equivalent $\sqrt{\sum_{q\in[Q]} \|f_q\|_k^2} \leqslant \tilde{\lambda}_f$ and $\|\mathbf{b}\|_2 \leqslant \tilde{\lambda}_\mathbf{b}$ forms. This stems from their reliance on internal primal-dual interior point techniques.

3.4 NUMERICAL EXPERIMENTS

In this section we demonstrate the efficiency of the presented SOC technique to solve hard shape-constrained problems. We focus on the task of joint quantile regression (JQR) where the conditional quantiles are encoded by the pinball loss (3.4) and the shape requirement to fulfill is the non-crossing property (3.5). Supplement (Section 3.5.2) provides an additional illustration in kernel ridge regression (KRR, (3.3)) on the importance of enforcing hard shape constraints in case of increasing noise level.

— **Experiment-1**: We compare the performance of the proposed SOC technique on 9 UCI benchmark datasets with a state-of-the-art JQR solver relying on soft shape constraints.

— **Experiment-2**: We augment the non-crossing constraint of JQR with monotonicity and concavity. Our two examples here belong to economics and to the analysis of aircraft trajectories.

In our experiments we used a Gaussian kernel with bandwidth σ, ridge regularization parameter λ_f and $\lambda_\mathbf{b}$ (or upper bounds $\tilde{\lambda}_f$ on $\sqrt{\sum_{q\in[Q]} \|f_q\|_k^2}$ and $\tilde{\lambda}_\mathbf{b}$ on $\|\mathbf{b}\|_2$). We learned jointly five quantile functions ($\tau_q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$). We used CVXGEN (Mattingley & Boyd, 2012) to solve ($\mathcal{P}_\eta$); the experiments took from seconds to a few minutes to run on an i7-CPU 16GB-RAM laptop.

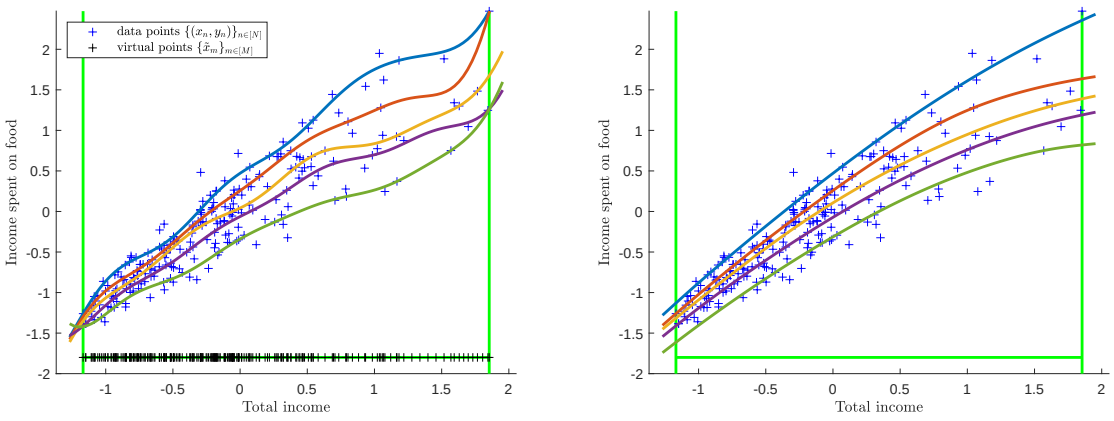
In our **first set of experiments** we compared the efficiency of the proposed SOC approach with the PDCD technique (Sangnier et al., 2016) which minimizes the same loss (3.4) but with a *soft* non-crossing encouraging regularizer. We considered 9 UCI benchmarks. Our datasets were selected with $d \in \{1, 2, 3\}$; to our best knowledge none of the available JQR solvers is able to guarantee in a hard fashion the non-crossing property of the learned quantiles out of samples even in this case. Each dataset was split into training (70%) and test (30%) sets; the split and the experiment were repeated twenty times. For each split, we optimized the hyperparameters $(\sigma, \tilde{\lambda}_f, \tilde{\lambda}_\mathbf{b})$ of SOC, searching over a grid to minimize the pinball loss through a 5-fold cross validation on the training set. Particularly, the kernel bandwith σ was searched over the square

Dataset	d	N	PDCD	SOC
engel	1	235	48 ± 8	53 ± 9
GAGurine	1	314	61 ± 7	65 ± 6
geyser	1	299	105 ± 7	108 ± 3
mcycle	1	133	66 ± 9	62 ± 5
ftcollinssnow	1	93	154 ± 16	148 ± 13
CobarOre	2	38	159 ± 24	151 ± 17
topo	2	52	69 ± 18	62 ± 14
caution	2	100	88 ± 17	98 ± 22
ufc	3	372	81 ± 4	87 ± 6

Table 3.1 – Joint quantile regression on 9 UCI datasets. Compared techniques: Primal-Dual Coordinate Descent (PDCD, Sangnier et al., 2016) and the presented SOC technique. Rows: benchmarks. 2nd column: dimension (d). 3rd column: sample number (N). 4-5th columns: mean $\pm$ std of $100\times$ value of the pinball loss for PDCD and SOC; smaller is better.

root of the deciles of the squared pairwise distance between the points $\{\mathbf{x}_n\}_{n\in[N]}$. The upper bound $\tilde{\lambda}_f$ on $\sqrt{\sum_{q\in[Q]} \|f_q\|_k^2}$ was scanned in the log-scale interval $[-1, 2]$. The upper bound $\tilde{\lambda}_b$ on $\|\mathbf{b}\|_2$ was kept fixed: $\tilde{\lambda}_b = 10 \max_{n\in[N]} |y_n|$. We then learned a model on the whole training set and evaluated it on the test set. The covering of $K = \prod_{r\in[d]} \left[\min\{(\mathbf{x}_n)_r\}_{n\in[N]}, \max\{(\mathbf{x}_n)_r\}_{n\in[N]}\right]$ was carried out with $\|\cdot\|_2$-balls of radius δ chosen such that the number M of added points was less than 100. This allowed for a rough covering while keeping the computation time for each run to be less than one minute. Our results are summarized in Table 3.1. The table shows that while the proposed SOC method guarantees the shape constraint in a *hard* fashion, its performance is on par with the state-of-the-art soft JQR solver.

In our **second set of experiments**, we demonstrate the efficiency of the proposed SOC estimator on tasks with additional hard shape constraints. Our first example is drawn from **economics**; we focused on JQR for the engel dataset, applying the same parameter optimization as in the first experiment. In this benchmark, the $\{(x_n, y_n)\}_{n\in[N]} \subset \mathbb{R}^2$ pairs correspond to annual household income (x_n) and food expenditure (y_n), preprocessed to have zero mean and unit variance. Engel's law postulates a monotone increasing property of y w.r.t. x, as well as concavity. We therefore constrained the quantile functions to be non-crossing, monotonically increasing *and* concave. The interval $K = \left[\min\{x_n\}_{n\in[N]}, \max\{x_n\}_{n\in[N]}\right]$ was covered with a non-uniform partition



(a) (b)

Figure 3.1 – Joint quantile regression on the engel dataset using the SOC technique. Solid lines: estimated conditional quantile functions with $\tau_q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ from bottom (dark green) to top (blue). Left plot: with non-crossing *and* increasing constraints. Right plot: with non-crossing, increasing *and* concavity constraints.

centered at the ordered centers $\{\tilde{x}_{m\in[M]}\}$ which included the original points $\{x_n\}_{n\in[N]}$ augmented with 15 virtual points. The radiuses were set to $\delta_{i,m} := \frac{\tilde{x}_{m+1}-\tilde{x}_m}{2}$ ($m \in [M-1]$, $i \in [I]$). The estimates with or without concavity are available in Fig. 3.1. It is interesting to notice that the estimated curves can intersect outside of the interval K (see Fig. 3.1(a)), and that the additional concavity constraint mitigates this intersection (see Fig. 3.1(b)).

In our second example with extra shape constraints, we focused on the analysis of more than 300 **aircraft trajectories** (Nicol, 2013) which describe the radar-measured altitude (y) of aircrafts flying between two cities (Paris and Toulouse) as a function of time (x). These trajectories were restricted to their takeoff phase (where the monotone increasing property should hold), giving rise to a total number of samples $N = 15657$. We imposed non-crossing and monotonicity property. The resulting SOC-based quantile function estimates describing the aircraft takeoffs are depicted in Fig. 3.2. The plot illustrates how the estimated quantile functions respect the hard shape constraints and shows where the aircraft trajectories concentrate under various level of probability, defining a corridor of normal flight altitude values.

These experiments demonstrate the efficiency of the proposed SOC-based solution to hard shape-constrained kernel machines.

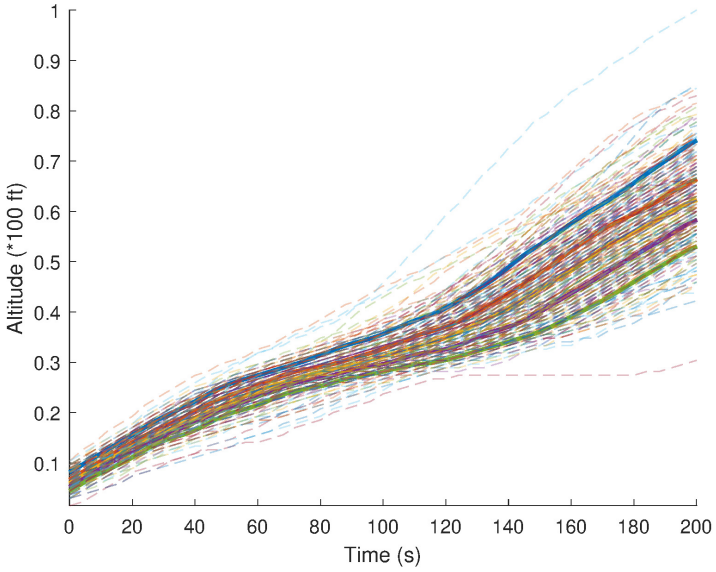


Figure 3.2 – Joint quantile regression on aircraft takeoff trajectories. Number of samples: $N = 15657$. Shape constraints: non-crossing and increasing constraints. Dashed lines: trajectory samples. Solid lines: estimated conditional quantile functions.

3.5 APPENDIX OF THE RESULTS OF CHAPTER 3

We provide the proof (Section 3.5.1) of our main result presented in Section 3.3. Section 3.5.2 is about an additional numerical illustration in the context of kernel ridge regression on the importance of hard shape constraints in case of increasing level of noise. For completeness, reformulations of the additional shape constraint examples for ($\mathcal{C}$) mentioned at the end of Section 3.2 are detailed in Section 3.5.3.

3.5.1 *Proof of Theorem 3.1*

For $i \in [I]$, we shall below denote $\phi_i = (\boldsymbol{W}\mathbf{f} - \mathbf{f}_0)_i$ and $\beta_i = (\mathbf{b}_0 - \boldsymbol{U}\mathbf{b})_i$. The proofs of the different parts are as follows.

(i) **Tightening**: By rewriting constraint ($\mathcal{C}$) using the derivative-reproducing property of kernels (Saitoh & Sawano, 2016; Zhou, 2008) we get

$$\langle \phi_i, D_{i,x} k(\mathbf{x}, \cdot) \rangle_k = D_i \phi_i(\mathbf{x}) \geqslant \beta_i, \ \forall \mathbf{x} \in K_i. \tag{3.11}$$

Let us reformulate this constraint as an inclusion of sets

$$\Phi_{D_i}(K_i) \subseteq H^+_{\phi_i,\beta_i} := \{g \in \mathcal{F}_k \,|\, \langle \phi_i, g \rangle_k \geqslant \beta_i\},$$

where $\Phi_{D_i} : \mathbf{x} \mapsto D_{i,x} k(\mathbf{x}, \cdot) \in \mathcal{F}_k$ and $\Phi_{D_i}(X) := \{\Phi_{D_i}(\mathbf{x}) \,|\, \mathbf{x} \in X\}$.

In order to get a finite geometrical description of $\Phi_{D_i}(K_i)$, we consider a finite covering of the compact set K_i:

$$\{\tilde{\mathbf{x}}_{i,m}\}_{m\in[M_i]} \subseteq K_i \subseteq \bigcup_{m\in[M_i]} \mathbb{B}_{\|\cdot\|_{\mathcal{X}}}(\tilde{\mathbf{x}}_{i,m}, \delta_{i,m}),$$

which implies that

$$\Phi_{D_i}(K_i) \subseteq \bigcup_{m\in[M_i]} \Phi_{D_i}\left(\mathbb{B}_{\|\cdot\|_{\mathcal{X}}}(\tilde{\mathbf{x}}_{i,m}, \delta_{i,m})\right).$$

From the regularity of k, it follows that Φ_{D_i} is continuous from $\mathcal{X}$ to $\mathcal{F}_k$, and we define $\eta_{i,m} > 0$ ($i \in [I]$, $m \in [M_i]$) as

$$\eta_{i,m} := \sup_{\mathbf{u}\in\mathbb{B}_{\|\cdot\|_{\mathcal{X}}}(0,1)} \|\Phi_{D_i}(\tilde{\mathbf{x}}_{i,m}) - \Phi_{D_i}(\tilde{\mathbf{x}}_{i,m} + \delta_{i,m}\mathbf{u})\|_k. \tag{3.12}$$

This means that for all $m \in [M_i]$

$$\Phi_{D_i}\left(\mathbb{B}_{\|\cdot\|_{\mathcal{X}}}(\tilde{\mathbf{x}}_{i,m}, \delta_{i,m})\right) \subseteq \Phi_{D_i}(\tilde{\mathbf{x}}_{i,m}) + \mathbb{B}_k(0, \eta_{i,m}),$$

where $\mathbb{B}_k(0, \eta_{i,m}) := \{g \in \mathcal{F}_k \mid \|g\|_k \leqslant \eta_{i,m}\}$. In other words, for (3.11) to hold, it is sufficient that

$$\Phi_{D_i}(\tilde{\mathbf{x}}_{i,m}) + \mathbb{B}_k(0, \eta_{i,m}) \subseteq H^+_{\phi_i,\beta_i}, \ \forall m \in [M_i]. \tag{3.13}$$

By the definition of $H^+_{\phi_i,\beta_i}$, (3.13) is equivalent to

$$\beta_i \leqslant \inf_{g\in\mathbb{B}_k} \langle \phi_i, D_{i,\mathbf{x}}k(\tilde{\mathbf{x}}_{i,m}, \cdot) + \eta_{i,m}g\rangle_k = D_i(\phi_i)(\tilde{\mathbf{x}}_{i,m}) - \eta_{i,m}\|\phi_i\|_k, \ \forall m \in [M_i].$$

Taking the minimum over $m \in [M_i]$, we get

$$\|\phi_i\|_k \leqslant \min_{m\in[M_i]} \frac{1}{\eta_{i,m}} \left[-\beta_i + D_i(\phi_i)(\tilde{\mathbf{x}}_{i,m})\right]. \tag{3.14}$$

Hence we proved that for any $(\mathbf{f}, \mathbf{b})$ satisfying (3.14), (3.11) also holds. The SOC-based reformulation is illustrated geometrically in Fig. 3.3. Constraint ($\mathcal{C}$) can be reformulated as requiring that the image $\Phi_{D_i}(K_i)$ of K_i under the D_i-feature map $\Phi_{D_i}(\mathbf{x}) := D_{i,\mathbf{x}}k(\mathbf{x}, \cdot) \in \mathcal{F}_k$ is contained in the halfspace 'above' the affine hyperplane defined by normal vector $(\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i$ and bias $(\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i$. The discretization (3.6) of constraint ($\mathcal{C}$) at the points $\{\tilde{\mathbf{x}}_{i,m}\}_{m\in[M_i]}$ only requires the images $\Phi_{D_i}(\tilde{\mathbf{x}}_{i,m})$ of the points to be above the hyperplane. Constraint ($\mathcal{C}_\eta$) instead inflates each of those points by a radius η_i.

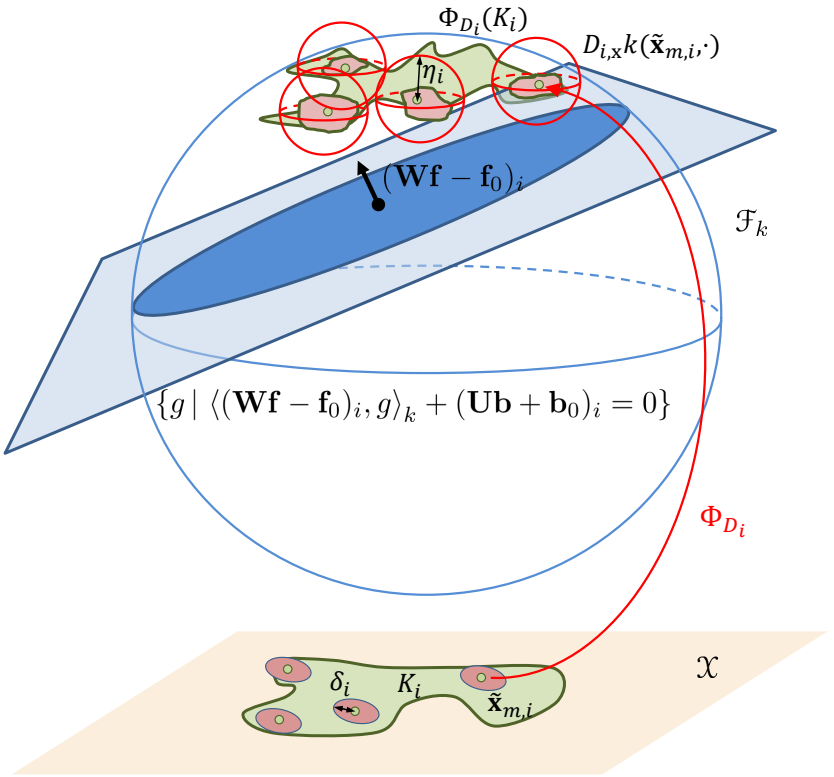


Figure 3.3 – Illustration of the SOC constraint ($\mathcal{C}_\eta$).

(ii) **Representer theorem**: For any $q \in [Q]$, let $f_{\eta,q} = v_q + w_q$ where v_q belongs to[8]

$$V := \operatorname{span}\left(\{f_{0,i}\}_{i\in I}, \{k(x_n, \cdot)\}_{n\in[N]}, \{D_{i,x}k(\tilde{x}_{i,m}, \cdot)\}_{m\in[M_i],i\in[I]}\right) \subset \mathcal{F}_k$$

while w_q is in the orthogonal complement of V in $\mathcal{F}_k$ ($w_q \in V^\perp := \{w \in \mathcal{F}_k : \langle w, v\rangle_k = 0,\ \forall v \in V\}$). Let $\boldsymbol{v} := (v_q)_{q\in[Q]} \in (\mathcal{F}_k)^Q$. As constraint ($\mathcal{C}_\eta$) holds for $(\mathbf{f}_\eta, \mathbf{b}_\eta)$,

$$(\mathbf{b}_0 - \mathbf{U}\mathbf{b}_\eta)_i + \eta_i \|(\boldsymbol{W}\mathbf{f}_\eta - \mathbf{f}_0)_i\|_k \leqslant \min_{m\in[M_i]} D_i(\boldsymbol{W}\mathbf{f}_\eta - \mathbf{f}_0)_i(\tilde{x}_{i,m}), \forall i \in [I].$$

However $(\boldsymbol{v}, \mathbf{b}_\eta)$ also satisfies ($\mathcal{C}_\eta$) since $\|(\boldsymbol{W}\boldsymbol{v} - \mathbf{f}_0)_i\|_k \leqslant \|(\boldsymbol{W}\mathbf{f}_\eta - \mathbf{f}_0)_i\|_k$ and $D_i(\boldsymbol{W}\boldsymbol{v} - \mathbf{f}_0)_i(\tilde{x}_{i,m}) = D_i(\boldsymbol{W}\mathbf{f}_\eta - \mathbf{f}_0)_i(\tilde{x}_{i,m})$:

$$\begin{aligned}
\left\|(\boldsymbol{W}\mathbf{f}_\eta - \mathbf{f}_0)_i\right\|_k^2 &= \left\|\sum_{q\in[Q]} W_{i,q} \underbrace{f_{\eta,q}}_{v_q+w_q} - f_{0,i}\right\|_k^2 = \left\|\underbrace{\sum_{q\in[Q]} W_{i,q}v_q - f_{0,i}}_{\in V} + \underbrace{\sum_{q\in[Q]} W_{i,q}w_q}_{\in V^\perp}\right\|_k^2 \\
&= \left\|\sum_{q\in[Q]} W_{i,q}v_q - f_{0,i}\right\|_k^2 + \left\|\sum_{q\in[Q]} W_{i,q}w_q\right\|_k^2 \geqslant \left\|\sum_{q\in[Q]} W_{i,q}v_q - f_{0,i}\right\|_k^2 \\
&= \|(\boldsymbol{W}\boldsymbol{v} - \mathbf{f}_0)_i\|_k^2, \\
D_i(\boldsymbol{W}\mathbf{f}_\eta - \mathbf{f}_0)_i(\tilde{x}_{i,m}) &= D_i\left(\sum_{q\in[Q]} W_{i,q} \underbrace{f_{\eta,q}}_{v_q+w_q} - f_{0,i}\right)(\tilde{x}_{i,m}) \\
&= D_i(\boldsymbol{W}\boldsymbol{v} - \mathbf{f}_0)_i(\tilde{x}_{i,m}) + D_i\left(\sum_{q\in[Q]} W_{i,q}w_q\right)(\tilde{x}_{i,m}) \\
&= D_i(\boldsymbol{W}\boldsymbol{v} - \mathbf{f}_0)_i(\tilde{x}_{i,m}) + \underbrace{\left\langle \sum_{q\in[Q]} W_{i,q}w_q, D_{i,x}k(\tilde{x}_{i,m}, \cdot)\right\rangle_k}_{=0}
\end{aligned}$$

using the derivative-reproducing property of kernels, and that $\sum_{q\in[Q]} W_{i,q}w_q \in V^\perp$, while $D_{i,x}k(\tilde{x}_{i,m}, \cdot) \in V$. The regularizer Ω is assumed to be strictly increasing in each argument $\|f_{\eta,q}\|_k$. As $\|f_{\eta,q}\|_k^2 = \|v_q\|_k^2 + \|w_q\|_k^2$, and $(\mathbf{f}_\eta, \mathbf{b}_\eta)$ minimizes $\mathcal{L}$, $w_q = 0$ necessarily; in other words $f_{\eta,q} \in V$ for all $q \in [Q]$.

(iii) **Performance guarantee**: From (i), we know that the solution $(\mathbf{f}_\eta, \mathbf{b}_\eta)$ of ($\mathcal{P}_\eta$) is also admissible for ($\mathcal{P}$). Discretizing the shape constraints is a relaxation of ($\mathcal{P}$). Hence $v_{\text{disc}} \leqslant \bar{v} \leqslant v_\eta$.

8. The linear hull of a finite set of points $(v_m)_{m\in[M]}$ in a vector space is denoted by $\operatorname{span}(\{v_m\}_{m\in[M]}) = \{\sum_{m\in[M]} a_m v_m \,|\, a_m \in \mathbb{R},\ \forall m \in [M]\}$.

Let us fix any $(\mathbf{p}_f, \mathbf{p}_b) \in (\mathcal{F}_k)^Q \times \mathbb{R}^P$ belonging to the subdifferential of $\mathcal{L}(\cdot,\cdot) + \chi_{\mathcal{C}}(\cdot,\cdot)$ at point $(\bar{\mathbf{f}}, \bar{\mathbf{b}})$, where $\chi_{\mathcal{C}}$ is the characteristic function of $\mathcal{C}$, i.e. $\chi_{\mathcal{C}}(\mathbf{u},\mathbf{v}) = 0$ if $(\mathbf{u},\mathbf{v}) \in \mathcal{C}$ and $\chi_{\mathcal{C}}(\mathbf{u},\mathbf{v}) = +\infty$ otherwise. Since $(\bar{\mathbf{f}}, \bar{\mathbf{b}})$ is the optimum of ($\mathcal{P}$), for any $(\mathbf{f}, \mathbf{b})$ admissible for ($\mathcal{P}$),

$$\sum_{q\in[Q]} \langle p_{f,q}, f_q - \bar{f}_q\rangle_k + \langle \mathbf{p}_b, \mathbf{b} - \bar{\mathbf{b}}\rangle_2 \geqslant 0, \tag{3.15}$$

where $\mathbf{p}_f = (p_{f,q})_{q\in[Q]}$. Hence using the (μ_{f_q}, μ_b)-strong convexity of $\mathcal{L}$ w.r.t. $(f_q, \mathbf{b})$ we get

$$\mathcal{L}(\mathbf{f}_\eta, \mathbf{b}_\eta) \geqslant \mathcal{L}(\bar{\mathbf{f}}, \bar{\mathbf{b}}) + \sum_{q\in[Q]} \langle p_{f_\eta,q}, f_{\eta,q} - \bar{f}_q\rangle_k + \langle \mathbf{p}_b, \mathbf{b}_\eta - \bar{\mathbf{b}}\rangle_2 + \sum_{q\in[Q]} \frac{\mu_{f_q}}{2}\|f_{\eta,q} - \bar{f}_q\|_k^2 + \frac{\mu_b}{2}\|\mathbf{b}_\eta - \bar{\mathbf{b}}\|_2^2. \tag{3.16}$$

As $\nu_\eta - \nu_{\text{disc}} \geqslant \mathcal{L}(\mathbf{f}_\eta, \mathbf{b}_\eta) - \mathcal{L}(\bar{\mathbf{f}}, \bar{\mathbf{b}})$, using the non-negativity (3.15) for $(\mathbf{f}_\eta, \mathbf{b}_\eta)$, one gets from (3.16) the claimed bound (3.9).

To prove (3.10), recall that $(\bar{\mathbf{f}}, \bar{\mathbf{b}})$ satisfies ($\mathcal{C}$) and that we assume $\mathcal{B} = \mathbb{R}^P$. Let $\eta_i = \max_{m\in[M_i]} \eta_{i,m}$, $i \in [I]$ with $\eta_{i,m}$ defined in (3.12), and $\tilde{\mathbf{b}} = (\tilde{b}_i)_{i\in[I]} \in \mathbb{R}^I$ with

$$\tilde{b}_i := \eta_i \left\|(W\bar{\mathbf{f}} - \mathbf{f}_0)_i\right\|_k. \tag{3.17}$$

As U is full-row rank, one can define its right inverse ($UU^+ = I$) as $U^+ = (U^\top U)^{-1} U^\top$. Then the pair $(\bar{\mathbf{f}}, \bar{\mathbf{b}} + U^+\tilde{\mathbf{b}})$ satisfies ($\mathcal{C}_\eta$) since for any $m \in [M_i]$

$$\begin{aligned}\eta_i \left\|(W\bar{\mathbf{f}} - \mathbf{f}_0)_i\right\|_k = \tilde{b}_i = (UU^+\tilde{\mathbf{b}})_i &\leqslant (UU^+\tilde{\mathbf{b}})_i + \underbrace{(U\bar{\mathbf{b}} - \mathbf{b}_0)_i + D_i(W\bar{\mathbf{f}} - \mathbf{f}_0)_i(\tilde{\mathbf{x}}_{i,m})}_{\geqslant 0} \\ &= (U(\bar{\mathbf{b}} + U^+\tilde{\mathbf{b}}) - \mathbf{b}_0)_i + D_i(W\bar{\mathbf{f}} - \mathbf{f}_0)_i(\tilde{\mathbf{x}}_{i,m}).\end{aligned}$$

Thus, $(\bar{\mathbf{f}}, \bar{\mathbf{b}} + U^+\tilde{\mathbf{b}})$ is admissible for ($\mathcal{P}_\eta$) and as $(\mathbf{f}_\eta, \mathbf{b}_\eta)$ is optimal for ($\mathcal{P}_\eta$), we have

$$\begin{aligned}\mathcal{L}(\mathbf{f}_\eta, \mathbf{b}_\eta) - \mathcal{L}(\bar{\mathbf{f}}, \bar{\mathbf{b}}) &\leqslant \mathcal{L}(\bar{\mathbf{f}}, \bar{\mathbf{b}} + U^+\tilde{\mathbf{b}}) - \mathcal{L}(\bar{\mathbf{f}}, \bar{\mathbf{b}}) \overset{(a)}{\leqslant} L_b\|U^+\tilde{\mathbf{b}}\|_2 \leqslant L_b\|U^+\|\,\|\tilde{\mathbf{b}}\|_2 \\ &\leqslant L_b\|U^+\|\sqrt{I}\|\tilde{\mathbf{b}}\|_\infty \overset{(b)}{\leqslant} L_b\|U^+\|\|\eta\|_\infty\sqrt{I}\max_{i\in[I]}\left\|(W\bar{\mathbf{f}} - \mathbf{f}_0)_i\right\|_k \\ &\overset{(c)}{=} L_b c_f\|\eta\|_\infty,\end{aligned}$$

where (a) stems from the local Lipschitz property of $\mathcal{L}$ ($\|U^+\tilde{\mathbf{b}}\|_2 \leqslant c_f\|\eta\|_\infty$), (b) holds by (3.17), and (c) follows from the definition of c_f. Combined with (3.16), this gives the bound (3.10).

3.5.2 *Shape-constrained kernel ridge regression*

In this section we illustrate in kernel ridge regression (KRR, (3.3)) the importance of enforcing hard shape constraints in case of increasing noise level. We consider a synthetic dataset of $N = 30$ points from the graph of a quadratic function where the values $\{x_n\}_{n\in[N]} \subset \mathbb{R}$ were generated uniformly on $[-2, 2]$. The corresponding y-coordinates of the graph were perturbed by additive Gaussian noise:

$$y_n = x_n^2 + \epsilon_n \ (\forall n \in [N]), \{\epsilon_n\}_{n\in[N]} \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, \xi^2\right).$$

We impose a monotonically increasing shape constraint on the interval $[x_l, x_u] = [0, 2]$, and study the effect of the level of the added noise (ξ) on the desired increasing property of the estimate without (KRR) and with monotonic shape constraint (SOC). Here $\sigma = 0.5$ and $\lambda_f = 10^{-4}$, while ξ varies in the interval $[0, 4]$.

Fig. 3.4(a) provides an illustration of the estimates in case of a fixed noise level $\xi = 1$. There is a good match between the KRR and SOC estimates outside of the interval $[0, 2]$, while the proposed SOC technique is able to correct the KRR estimate to enforce the monotonicity requirement on $[0, 2]$. In order to assess the performance of the unconstrained KRR estimator under varying level of noise, we repeated the experiment 1000 times for each noise level ξ and computed the proportion and amount[9] of violation of the monotonicity requirement. Our results are summarized in Fig. 3.4(b). The figure shows that the error increases rapidly for KRR as a function of the noise level, and even for very low level of noise the monotonicity requirement does not hold. These experiments illustrate that shape constraints can grossly be violated when facing noise if they are not enforced in an explicit and hard fashion. To explicit the tightening property of Theorem 3.1, i.e. that $\nu_{\text{disc}} \leqslant \bar{\nu} \leqslant \nu_\eta$, Fig. 3.4(c) shows the evolution of the optimal values ν_η and ν_{disc} when increasing the number M of discretization points of the constraints on the constraint interval $[0, 2]$. Since by increasing M, we decrease η, the value ν_η decreases when increasing M, whereas ν_{disc} increases as the discretization incorporates more constraints. Increasing M naturally increases the polynomial computation time but not necessarily at the worst cubic expense, as depicted on Fig. 3.4(d), the choice of solver having also importance as it may provide a factor 2 gain.

9. These performance measures are defined as $\frac{1}{2}\int_0^2 \max(0, -f'(x))dx$ and $\int_0^2 \max_{y\in[0,x]}[f(y) - f(x)]dx$. By construction both measures are zero for SOC.

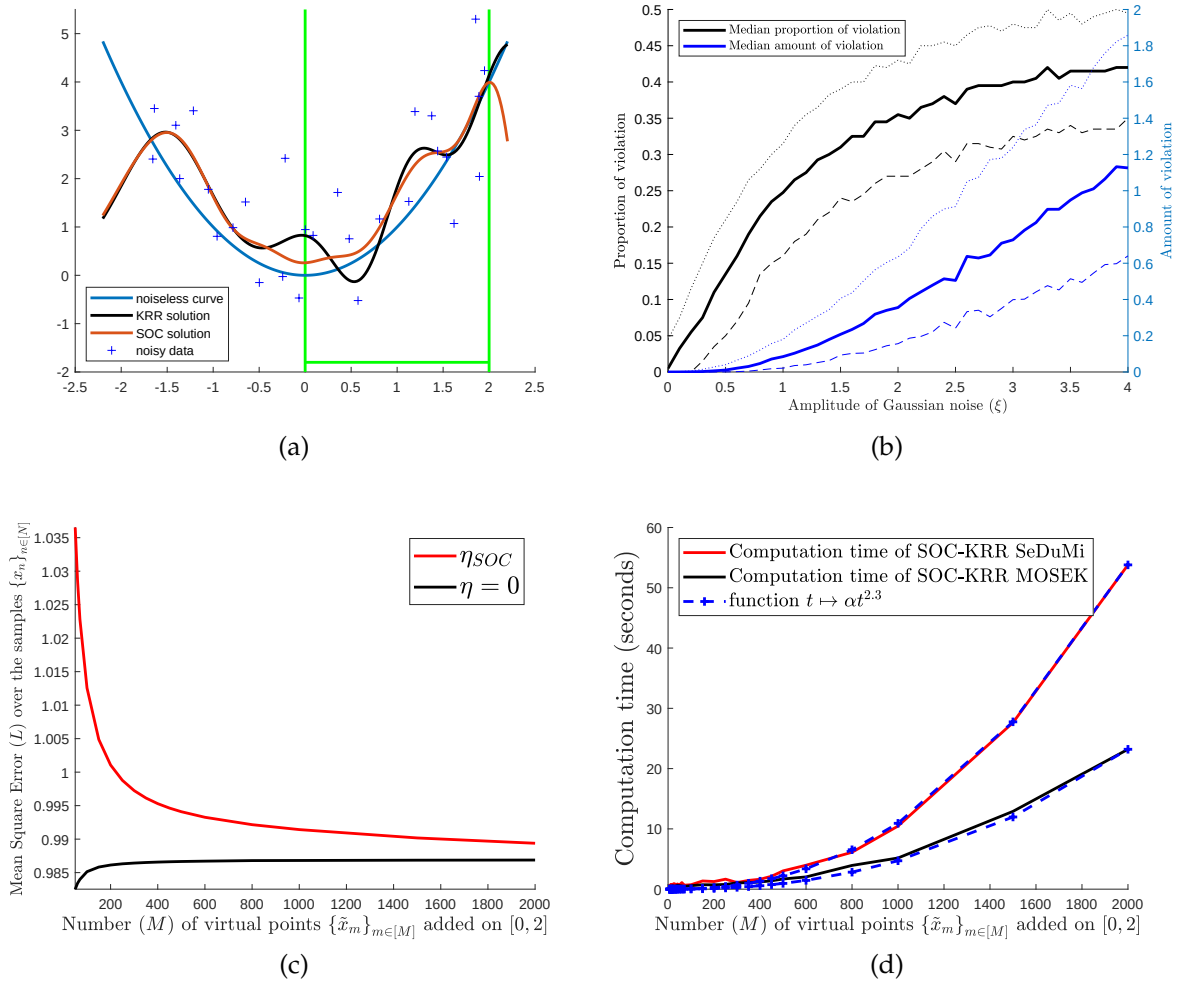


Figure 3.4 – (a): Illustration for kernel ridge regression. Observation: quadratic function perturbed with additive Gaussian noise. Shape constraint: monotone increasing property on $[0,2]$. Compared techniques: regression without (KRR) and with hard shape constraint (SOC). (b): Violation of the shape constraint for the unconstrained KRR estimator as a function of the amplitude of the added Gaussian noise. Error measures: median of the proportion (left) and amount (right) of the violation of the monotone increasing property on $[0,2]$. Dashed lines: lower and upper quartiles. (c): Evolution of the optimal objective values v_η and v_{disc} when increasing the number M of discretization points of the constraints on $[0,2]$. (d): Computation time of ($\mathcal{P}_\eta$) depending on the convex optimization solver selected.

3.5.3 *Examples of handled shape constraints*

In order to make the paper self-contained, in this section we provide the reformulations using derivatives of the additional shape constraints briefly mentioned at the end of Section 3.2: n-monotonicity ($s = n$; Chatterjee et al., 2015), $(n-1)$-alternating monotonicity (Fink, 1982), monotonicity w.r.t. unordered weak majorization ($s = 1$; Marshall et al., 2011, A.7. Theorem) or w.r.t. product ordering ($s = 1$), or supermodularity ($s = 2$; Simchi-Levi et al., 2014, Section 2).

Particularly, n-monotonicity ($n \in \mathbb{N}^*$) writes as $f^{(n)}(x) \geqslant 0$ $(\forall x)$. $(n-1)$-alternating monotonicity [10] ($n \in \mathbb{N}^*$) is similar: for $n = 1$ non-negativity and non-increasing properties are required; for $n \geqslant 2$ $(-1)^j f^{(j)}$ has to be non-negative, non-increasing and convex for $\forall j \in \{0, \ldots, n-2\}$. The other examples are

— **Monotonicity w.r.t. partial ordering**: These generalized notions of monotonicity ($\mathbf{u} \preccurlyeq \boldsymbol{\nu} \Rightarrow f(\mathbf{u}) \leqslant f(\boldsymbol{\nu})$) rely on the partial orderings $\mathbf{u} \preccurlyeq \boldsymbol{\nu}$ iff $\sum_{j\in[i]} u_j \leqslant \sum_{j\in[i]} \nu_j$ for all $i \in [d]$ (unordered weak majorization) and $\mathbf{u} \preccurlyeq \boldsymbol{\nu}$ iff $u_i \leqslant \nu_i$ $(\forall i \in [d])$ (product ordering). For C^1 functions mononicity w.r.t. the unordered weak majorization is equivalent to

$$\partial^{e_1} f(\mathbf{x}) \geqslant \ldots \geqslant \partial^{e_d} f(\mathbf{x}) \geqslant 0 \quad (\forall \mathbf{x}).$$

Monotonicity w.r.t. product ordering for C^1 functions can be rephrased as

$$\partial^{e_j} f(\mathbf{x}) \geqslant 0, \quad (\forall j \in [d], \quad \forall \mathbf{x}).$$

— **Supermodularity**: Supermodularity means that $f(\mathbf{u} \vee \boldsymbol{\nu}) + f(\mathbf{u} \wedge \boldsymbol{\nu}) \geqslant f(\mathbf{u}) + f(\boldsymbol{\nu})$ for all $\mathbf{u}, \boldsymbol{\nu} \in \mathbb{R}^d$, where maximum and minimum are meant coordinate-wise, i.e. $\mathbf{u} \vee \boldsymbol{\nu} := (\max(u_j, \nu_j))_{j\in[d]}$ and $\mathbf{u} \wedge \boldsymbol{\nu} := (\min(u_j, \nu_j))_{j\in[d]}$ for $\mathbf{u}, \boldsymbol{\nu} \in \mathbb{R}^d$. For C^2 functions this property corresponds to

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \geqslant 0 \quad (\forall i \neq j \in [d], \forall \mathbf{x}).$$

BIBLIOGRAPHY FOR CHAPTER 3

Agrell, C. (2019). Gaussian processes with linear operator inequality constraints. *Journal of Machine Learning Research*, *20*, 1–36.

Aït-Sahalia, Y., & Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, *116*(1-2), 9–47.

10. For instance, the generator of a d-variate Archimedean copula can be characterized by $(d-2)$-alternating monotonicity (Malov, 2001; McNeil & Neslehová, 2009).

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

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

Aybat, N. S., & Wang, Z. (2014). A parallel method for large scale convex regression problems. *Conference on Decision and Control (CDC)*, 5710–5717.

Bach, F., & Jordan, M. (2002). Kernel independent component analysis. *Journal of Machine Learning Research*, *3*, 1–48.

Bagnell, J. A., & Farahmand, A. (2015). Learning positive functions in a Hilbert space [(https://www.ri.cmu.edu/pub_files/2015/0/Kernel-SOS.pdf)].

Balabdaoui, F., Durot, C., & Jankowski, H. (2019). Least squares estimation in the monotone single index model. *Bernoulli*, *25*(4B), 3276–3310.

Berlinet, A., & Thomas-Agnan, C. (2004). *Reproducing kernel Hilbert spaces in probability and statistics*. Kluwer.

Blundell, R., Horowitz, J. L., & Parey, M. (2012). Measuring the price responsiveness of gasoline demand: economic shape restrictions and nonparametric demand estimation. *Quantitative Economics*, *3*, 29–51.

Carmeli, C., Vito, E. D., Toigo, A., & Umanitá, V. (2010). Vector valued reproducing kernel Hilbert spaces and universality. *Analysis and Applications*, *8*, 19–61.

Chatterjee, S., Guntuboyina, A., & Sen, B. (2015). On risk bounds in isotonic and other shape restricted regression problems. *Annals of Statistics*, *43*(4), 1774–1800.

Chen, Y., & Samworth, R. J. (2016). Generalized additive and index models with shape constraints. *Journal of the Royal Statistical Society – Statistical Methodology, Series B*, *78*(4), 729–754.

Chetverikov, D., Santos, A., & Shaikh, A. M. (2018). The econometrics of shape restrictions. *Annual Review of Economics*, *10*(1), 31–63.

Delecroix, M., Simioni, M., & Thomas-Agnan, C. (1996). Functional estimation under shape constraints. *Journal of Nonparametric Statistics*, *6*(1), 69–89.

Fink, A. M. (1982). Kolmogorov-Landau inequalities for monotone functions. *Journal of Mathematical Analysis and Applications*, *90*, 251–258.

Flaxman, S., Teh, Y. W., & Sejdinovic, D. (2017). Poisson intensity estimation with reproducing kernels. *Electronic Journal of Statistics*, *11*(2), 5081–5104.

Freyberger, J., & Reeves, B. (2018). *Inference under shape restrictions* (tech. rep.) [(https://www.ssc.wisc.edu/~jfreyberger/Shape_Inference_Freyberger_Reeves.pdf)]. University of Wisconsin-Madison.

Fukumizu, K., Gretton, A., Sun, X., & Schölkopf, B. (2008). Kernel measures of conditional dependence. *Advances in Neural Information Processing Systems (NIPS)*, 498–496.

Guntuboyina, A., & Sen, B. (2018). Nonparametric shape-restricted regression. *Statistical Science*, *33*(4), 568–594.

Hall, G. (2018). *Optimization over nonnegative and convex polynomials with and without semidefinite programming* (PhD Thesis). Princeton University.

Han, Q., Wang, T., Chatterjee, S., & Samworth, R. J. (2019). Isotonic regression in general dimensions. *Annals of Statistics*, *47*(5), 2440–2471.

Han, Q., & Wellner, J. A. (2016). *Multivariate convex regression: global risk bounds and adaptation* (tech. rep.) [(https://arxiv.org/abs/1601.06844)].

Hofmann, T., Schölkopf, B., & Smola, A. J. (2008). Kernel methods in machine learning. *The Annals of Statistics*, *36*(3), 1171–1220.

Horowitz, J. L., & Lee, S. (2017). Nonparametric estimation and inference under shape restrictions. *Journal of Econometrics*, *201*, 108–126.

Hu, J., Kapoor, M., Zhang, W., Hamilton, S. R., & Coombes, K. R. (2005). Analysis of dose-response effects on gene expression data with comparison of two microarray platforms. *Bioinformatics*, *21*(17), 3524–3529.

Johnson, A. L., & Jiang, D. R. (2018). Shape constraints in economics and operations research. *Statistical Science*, *33*(4), 527–546.

Keshavarz, A., Wang, Y., & Boyd, S. (2011). Imputing a convex objective function. *IEEE Multi-Conference on Systems and Control*, 613–619.

Koppel, A., Zhang, K., Zhu, H., & Başar, T. (2019). Projected stochastic primal-dual method for constrained online learning with kernels. *IEEE Transactions on Signal Processing*, *67*(10), 2528–2542.

Lewbel, A. (2010). Shape-invariant demand functions. *The Review of Economics and Statistics*, *92*(3), 549–556.

Li, Q., & Racine, J. S. (2007). *Nonparametric econometrics*. Princeton University Press.

Luss, R., Rossett, S., & Shahar, M. (2012). Efficient regularized isotonic regression with application to gene-gene interaction search. *Annals of Applied Statistics*, *6*(1), 253–283.

Malov, S. V. (2001). On finite-dimensional Archimedean copulas. *Asymptotic Methods in Probability and Statistics with Applications*, 19–35.

Marshall, A. W., Olkin, I., & Arnold, B. C. (2011). *Inequalities: theory of majorization and its applications*. Springer.

Mattingley, J., & Boyd, S. (2012). CVXGEN: a code generator for embedded convex optimization. *Optimization and Engineering*, *12*(1), 1–27.

Matzkin, R. L. (1991). Semiparametric estimation of monotone and concave utility functions for polychotomous choice models. *Econometrica*, *59*(5), 1315–1327.

Mazumder, R., Choudhury, A., Iyengar, G., & Sen, B. (2019). A computational framework for multivariate convex regression and its variants. *Journal of the American Statistical Association*, *114*(525), 318–331.

McNeil, A. J., & Neslehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *Annals of Statistics*, *37*(5B), 3059–3097.

Meyer, M. C. (2018). A framework for estimation and inference in generalized additive models with shape and order restrictions. *Statistical Science*, *33*(4), 595–614.

Micchelli, C., Xu, Y., & Zhang, H. (2006). Universal kernels. *Journal of Machine Learning Research*, *7*, 2651–2667.

Nicol, F. (2013). Functional principal component analysis of aircraft trajectories. *International Conference on Interdisciplinary Science for Innovative Air Traffic Management (ISIATM)*.

Papp, D., & Alizadeh, F. (2014). Shape-constrained estimation using nonnegative splines. *Journal of Computational and Graphical Statistics*, *23*(1), 211–231.

Pya, N., & Wood, S. N. (2015). Shape constrained additive models. *Statistics and Computing*, *25*, 543–559.

Saitoh, S., & Sawano, Y. (2016). *Theory of reproducing kernels and applications*. Springer Singapore.

Sangnier, M., Fercoq, O., & d'Alché-Buc, F. (2016). Joint quantile regression in vector-valued RKHSs. *Advances in Neural Information Processing Systems (NIPS)*, 3693–3701.

Shapiro, A., Dentcheva, D., & Ruszczynski, A. (2014). *Lectures on stochastic programming: modeling and theory*. SIAM - Society for Industrial; Applied Mathematics.

Shi, X., Shum, M., & Song, W. (2018). Estimating semi-parametric panel multinomial choice models using cyclic monotonicity. *Econometrica*, *86*(2), 737–761.

Simchi-Levi, D., Chen, X., & Bramel, J. (2014). *The logic of logistics: theory, algorithms, and applications for logistics management*. Springer.

Simon-Gabriel, C.-J., & Schölkopf, B. (2018). Kernel distribution embeddings: universal kernels, characteristic kernels and kernel metrics on distributions. *Journal of Machine Learning Research*, *19*(44), 1–29.

Sriperumbudur, B., Fukumizu, K., & Lanckriet, G. (2011). Universality, characteristic kernels and RKHS embedding of measures. *Journal of Machine Learning Research*, *12*, 2389–2410.

Sriperumbudur, B., Gretton, A., Fukumizu, K., Schölkopf, B., & Lanckriet, G. (2010). Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, *11*, 1517–1561.

Steinwart, I. (2001). On the influence of the kernel on the consistency of support vector machines. *Journal of Machine Learning Research*, *6*(3), 67–93.

Steinwart, I., & Christmann, A. (2008). *Support vector machines*. Springer.

Szabó, Z., & Sriperumbudur, B. K. (2018). Characteristic and universal tensor product kernels. *Journal of Machine Learning Research*, *18*(233), 1–29.

Takeuchi, I., Le, Q., Sears, T., & Smola, A. (2006). Nonparametric quantile estimation. *Journal of Machine Learning Research*, *7*, 1231–1264.

Topkis, D. M. (1998). *Supermodularity and complementarity*. Princeton University Press.

Turlach, B. A. (2005). Shape constrained smoothing using smoothing splines. *Computational Statistics*, *20*, 81–104.

Varian, H. R. (1984). The nonparametric approach to production analysis. *Econometrica*, *52*(3), 579–597.

Wahba, G. (1990). *Spline models for observational data*. SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics.

Wang, Y. (2011). *Smoothing splines – methods and applications*. CRC Press.

Wu, J., Meyer, M. C., & Opsomer, J. D. (2015). Penalized isotonic regression. *Journal of Statistical Planning and Inference*, *161*, 12–24.

Wu, X., & Sickles, R. (2018). Semiparametric estimation under shape constraints. *Econometrics and Statistics*, *6*, 74–89.

Yagi, D., Chen, Y., Johnson, A. L., & Kuosmanen, T. (2020). Shape-constrained kernel-weighted least squares: estimating production functions for Chilean manufacturing industries. *Journal of Business & Economic Statistics*, *38*(1), 43–54.

Zhou, D.-X. (2008). Derivative reproducing properties for kernel methods in learning theory. *Journal of Computational and Applied Mathematics*, *220*, 456–463.

4

STATE CONSTRAINTS IN LQ OPTIMAL CONTROL THROUGH THE LQ KERNEL

This chapter was accepted for publication with a single author in SIAM Journal on Control and Optimization, under the title *Linearly-constrained Linear Quadratic Regulator from the viewpoint of kernel methods* (Aubin-Frankowski, 2021b).

Abstract The linear quadratic regulator problem is central in optimal control and was investigated since the very beginning of control theory. Nevertheless, when it includes affine state constraints, it remains very challenging from the classical "maximum principle" perspective. In this chapter, we present how matrix-valued reproducing kernels allow for an alternative viewpoint. We show that the quadratic objective paired with the linear dynamics encode the relevant kernel, defining a Hilbert space of controlled trajectories. Drawing upon kernel formalism, we introduce a strengthened continuous-time convex optimization problem which can be tackled exactly with finite dimensional solvers, and which solution is interior to the constraints. When refining a time-discretization grid, this solution can be made arbitrarily close to the solution of the state-constrained Linear Quadratic Regulator. We illustrate the implementation of this method on a path-planning problem.

Résumé Le problème du régulateur linéaire quadratique a un rôle central en contrôle optimal et a été étudié depuis les origines de la théorie du contrôle. Néanmoins, en présence de contraintes d'état affines, il reste très difficile à résoudre dans la perspective classique du "principe du maximum". Dans ce chapitre, nous présentons comment les noyaux reproduisants à valeurs matricielles induisent un autre point de vue sur ce problème. Nous montrons que l'objectif quadratique associé à la dynamique linéaire encodent le noyau pertinent, définissant un espace de Hilbert de trajectoires contrôlées. En s'appuyant sur le formalisme des noyaux, nous introduisons un problème surcontraint d'optimisation convexe en temps continu qui peut être résolu de manière exacte par des solveurs de dimension finie, et dont la solution est à l'intérieur

des contraintes d'origine. Si une grille de discrétisation temporelle des contraintes est raffinée, cette solution peut être rendue arbitrairement proche de la solution du régulateur quadratique linéaire sous contraintes d'état. Nous illustrons la mise en œuvre de cette méthode par un problème de planification de trajectoire.

4.1 INTRODUCTION

In its simplest form, the problem of time-varying linear quadratic optimal control with finite horizon and affine inequality state constraints writes as

$$\begin{aligned} \min_{\mathbf{x}(\cdot),\mathbf{u}(\cdot)} \quad & g(\mathbf{x}(T)) + \int_0^T [\mathbf{x}(t)^\top \mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}(t)\mathbf{u}(t)]dt \\ \text{s.t.} \quad & \\ & \mathbf{x}(0) = \mathbf{x}_0, \\ & \mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t), \text{ a.e. in } [0,T], \\ & \mathbf{c}_i(t)^\top \mathbf{x}(t) \leqslant d_i(t), \forall\, t \in [0,T], \forall\, i \in \{1,\dots,P\}, \end{aligned} \tag{$\mathcal{P}_0$}$$

where the state $\mathbf{x}(t) \in \mathbb{R}^N$, the control $\mathbf{u}(t) \in \mathbb{R}^M$, $\mathbf{A}(t) \in \mathbb{R}^{N,N}$, $\mathbf{B}(t) \in \mathbb{R}^{N,M}$, $\mathbf{C}(t) = [\mathbf{c}_1(t)^\top; \dots; \mathbf{c}_P(t)^\top] \in \mathbb{R}^{P,N}$ ($\mathbf{c}_i(t) \in \mathbb{R}^N$), $\mathbf{d}(t) = (d_i(t))_i \in \mathbb{R}^P$, while $\mathbf{Q}(t) \in \mathbb{R}^{N,N}$ and $\mathbf{R}(t) \in \mathbb{R}^{M,M}$ are positive semidefinite matrices.

Below, for $q \in \{1, 2, \infty\}$, $L^q(0,T)$ denotes the L^q-space of functions over $[0,T]$ with integrable norms of the function values (resp. square-integrable, resp. bounded). We shall henceforth assume that, for all $t \in [0,T]$, $\mathbf{R}(t) \succcurlyeq r\mathrm{Id}_M$ with $r > 0$, as well as $\mathbf{A}(\cdot) \in L^1(0,T)$, $\mathbf{B}(\cdot) \in L^2(0,T)$, $\mathbf{Q}(\cdot) \in L^1(0,T)$, and $\mathbf{R}(\cdot) \in L^2(0,T)$. To have a finite objective, it is natural to restrict our attention to measurable controls satisfying $\mathbf{R}(\cdot)^{1/2}\mathbf{u}(\cdot) \in L^2(0,T)$.

Without state constraints, under mild assumptions, the unconstrained Linear Quadratic Regulator (LQR) enjoys an explicit solution defined through the Hamiltonian matrix and the related Riccati equation (see e.g. Speyer and Jacobson (2010)). With state constraints, little can be said as Pontryagin's Maximum Principle involves not only an adjoint vector but also measures supported on the constraint boundary. A comprehensive review of this approach can be found in Hartl et al., 1995. One has to guess beforehand when the state-constraint is active (at the so-called junction times) in order to write the first-order necessary condition (Hermant, 2009). Secondly one has to impose assumptions to derive the magnitude of the discontinuities of the adjoint

vector. This has proven to be intractable and made state-constrained continuous-time optimization a difficult problem. Let us provide an intuition for the appearance of discontinuities. If one follows an optimal trajectory of the LQR starting in the interior of state constraints, one may reach the boundary while the unconstrained Hamiltonian system of the Maximum Principle may incite to use a control leading to violate the constraint. One has then to apply a different control to remain in the constraint set, possibly generating a discontinuity in the adjoint vector.

Although LQR problems stand at the origin of control theory, research is still active in the field, not only for its numerous applications (see e.g. the examples of Burachik et al. (2014) and references within), but also for its theoretical aspects, even without constraints (Bourdin & Trélat, 2017) or just control constraints (Burachik et al., 2014). Many of these improvements are motivated by model predictive control, considered for instance in a time-invariant discrete-time state-constrained setting in Grüne and Guglielmi (2018) or continuous-time (van Keulen, 2020). In particular, Kojima and Morari (2004) proved that the solutions of a time-invariant LQR with discretized constraints converge to the solution of ($\mathcal{P}_0$), putting emphasis on function spaces of controls. As a matter of fact, the aforementioned approaches focus on the control, used to parametrize the trajectories. In the present study, trajectories are instead at the core of the analysis.

When seeking a continuous-time numerical solution, one has to face an infinite number of pointwise constraints, and has either to relax the computationally intractable optimization problem or tighten it. *Relaxing* means either enforcing the constraint only at a finite number of points, without guarantees elsewhere (as with any time discretization method, e.g. Kojima & Morari, 2004), or through soft penalties (Gerdts & Hüpping, 2012), such as approximations of barrier functions (Dower et al., 2019). *Tightening* usually implies either choosing $\mathbf{u}(\cdot)$ in a convenient subspace of $L^2(0,T)$, such as the one of piecewise constant functions [1] or of splines with prescribed knots (Mercy et al., 2016), or through hard penalties, such as logarithmic barriers (Chaplais et al., 2011). Let us illustrate the difference between relaxing and tightening state constraints. Consider the problem of a traffic regulator whose aim is to enforce a speed limit over a highway. The drivers for their part want to go as far as possible in a given time. Deploying speed cameras ensures at best that the speed constraint is satisfied locally (relaxing). However if a smaller maximum speed is imposed at the camera level (tightening), then the cars cannot accelerate enough to break the speed limit before reaching the next camera. In a nutshell, the kernel methods framework we advocate

1. This is known as *sampled-data* or *digital control* (Ackermann, 1985), the *sampled-data* terminology does not refer to machine learning techniques.

allows to compute both a threshold and the resulting trajectories.

Kernel methods, being related to Green's functions, belong to a branch of functional analysis. Their history was already sketched by Aronszajn (1950), to whom we owe the modern formulation of the theory. The regain of interest thanks to support vector machines (see e.g. Schölkopf & Smola, 2002) has reinstated kernel methods as the most principled technique in machine learning. There have been many attempts since then at bridging kernel methods and control theory. The two are already related in the works of Kailath (1971) and Parzen (1970). More recently Marco et al. (2017) and Steinke and Schölkopf (2008) have considered kernels for control systems, mainly to encode the input $\mathbf{u}(\cdot)$ or for system identification purposes (see e.g. the reviews of Chiuso and Pillonetto (2019) and Pillonetto et al. (2014)). Kernels have also been applied to approximate the Koopman operator over observers of uncontrolled nonlinear systems, in connection with spectral analysis (Fujii & Kawahara, 2019; Williams et al., 2015) or for given controls (Sootla et al., 2018). The kernel Hilbert spaces have also been used to define suitable domains for operators (Giannakis et al., 2019; Rosenfeld et al., 2019). In most cases, the kernel is taken off-the-shelf, as with Gaussian kernels in connection with Bayesian inference (Bertalan et al., 2019; Singh et al., 2018).

On the contrary, departing from the prevalent perspective of using kernel methods as nonlinear embeddings, this article rekindles with a long standing tradition of engineering kernels for specific uses. This view has been mainly supported by the statistics community, especially in connection with splines and Sobolev spaces (Heckman, 2012; Wahba, 1990).[2] For ($\mathcal{P}_0$), we show below that the quadratic objective paired with the linear dynamics encode the relevant kernel, which defines the Hilbert space of controlled trajectories. As kernel methods deal with a special class of Hilbert spaces, they are natural to consider for linear systems or for linearizations of nonlinear systems. Nonetheless the interactions run deeper. For instance we prove below that the controllability Gramian is directly related to matrix-valued kernels, and we recover the transversality condition merely through a representer theorem. This approach was further extended in Aubin-Frankowski (2021) to the connexion between the Linear-Quadratic matrix-valued kernel defined below and the dual Riccati equation. Since the Riccati equation is often used for the online purpose of finding the control, Aubin-Frankowski (2021) also discusses how the kernel formalism effectively allows for optimal synthesis, favoring an offline trajectory-focused viewpoint.

2. Drawing inspiration from linear control theory, *control theoretic splines* were devised (Fujioka & Kano, 2013; Magnus Egerstedt, 2009), in particular for path-planning problems (Kano & Fujioka, 2018). Possibly unbeknownst to non-kernel users, kernel theory, sometimes known as *abstract splines*, is the natural generalization of splines (see e.g. Aubin-Frankowski et al., 2020).

Our main result is to show that through the theory of kernel methods we can solve a novel strengthened version of ($\mathcal{P}_0$), without changing the function spaces involved. This is achieved by introducing a finite number of second-order cone constraints, stronger than the infinite number of affine constraints. We furthermore show that the solution of the strengthened problem can be made arbitrarily close to the solution of ($\mathcal{P}_0$), and that it enjoys a finite representation. One can thus exactly solve the continuous-time problem using only finite dimensional convex optimization solvers. This computable trajectory, which is interior to the affine state constraints, can also foster intuitions on the behavior of the optimal solution of ($\mathcal{P}_0$). The considered tightening of ($\mathcal{P}_0$) relies solely on the kernel formalism. It was first introduced in Aubin-Frankowski and Szabó (2020b) and thoroughly extended in Aubin-Frankowski and Szabó (2020a).

In Section 4.2 we present our strengthened problem and show how the LQR problem can be expressed as a regression problem over a vector-valued reproducing kernel Hilbert space. Section 4.3 details the consequences of this framework and identifies the corresponding Linear-Quadratic kernel. Our main result on approximation guarantees is stated in Section 4.4. Section 4.5 discusses the implementation and the numerical behavior of the strengthened constraints. The setting is also extended to intermediate or terminal equality constraints, as in path-planning problems. The annex pertains to conditions ensuring the existence of interior trajectories for ($\mathcal{P}_0$).

4.2 THEORETICAL PRELIMINARIES AND PROBLEM FORMULATION

In this section, we present the tools from the theory of kernel methods that we shall apply. We then introduce our strengthened problem with second-order cone constraints.

Notations: We use the shorthand $[\![1, P]\!] = \{1, \dots, P\}$. $\mathbb{R}^N_+$ is the subset of $\mathbb{R}^N$ of elements with nonnegative components. $\mathbb{B}_N$ denotes the closed unit ball of $\mathbb{R}^N$ for the Euclidean inner product, $\mathbf{1}_N$ the vector of all-ones. For a matrix $\boldsymbol{A} \in \mathbb{R}^{N,N}$, we write by $\|\boldsymbol{A}\|$ its operator norm. Id_N is the identity matrix of $\mathbb{R}^{N,N}$. We chose not to explicit the output space for the function spaces to avoid cumbersome notations, as it can be always deduced from the context. The space of functions with continuous derivatives up to order s is denoted by $\mathcal{C}^s(0, T)$. For a function $K(\cdot, \cdot)$ defined over a subset of $\mathbb{R} \times \mathbb{R}$, $\partial_1 K(\cdot, \cdot)$ denotes the partial derivative w.r.t. the first variable. For a Hilbert space $(\mathcal{H}_k, \langle \cdot, \cdot \rangle_K)$, $\mathbb{B}_K$ is the closed unit ball of $\mathcal{H}_k$, $\| \cdot \|_K$ denoting the corresponding norm. Given a subspace $V \subset \mathcal{H}_k$, we denote by $V^\perp$ its orthogonal complement w.r.t. $\langle \cdot, \cdot \rangle_K$. A μ-strongly convex function $\mathcal{L} : \mathcal{H}_k \mapsto \mathbb{R}$ is a function satisfying, for all $f_1, f_2 \in \mathcal{H}_k$, $\alpha \in [0, 1]$, $\mathcal{L}(\alpha f_1 + (1-\alpha) f_2) + \alpha(1-\alpha)\frac{\mu}{2}\|f_1 - f_2\|_K^2 \leqslant \alpha \mathcal{L}(f_1) + (1-\alpha)\mathcal{L}(f_2)$.

Definition 4.1. Let $\mathcal{T}$ be a non-empty set. A Hilbert space $(\mathcal{H}_K, \langle\cdot,\cdot\rangle_K)$ of $\mathbb{R}^N$-vector-valued functions defined on $\mathcal{T}$ is called a vector-valued reproducing kernel Hilbert space (vRKHS) if there exists a matrix-valued kernel $K : \mathcal{T}\times\mathcal{T} \to \mathbb{R}^{N,N}$ such that the *reproducing property* holds: for all $t \in \mathcal{T}$, $\mathbf{p} \in \mathbb{R}^N$, $K(\cdot,t)\mathbf{p} \in \mathcal{H}_K$ and for all $\mathbf{f} \in \mathcal{H}_K$, $\mathbf{p}^\top \mathbf{f}(t) = \langle \mathbf{f}, K(\cdot,t)\mathbf{p}\rangle_K$.

Many properties of real-valued RKHSs have been known since Aronszajn (1950), the general theory having been developed by Schwartz (1964). By Riesz's theorem, an equivalent definition of a vRKHS is that, for every $t \in \mathcal{T}$ and $\mathbf{p} \in \mathbb{R}^N$, the evaluation functional $\mathbf{f} \in \mathcal{H}_K \mapsto \mathbf{p}^\top \mathbf{f}(t) \in \mathbb{R}$ is continuous. There is also a one-to-one correspondence between the kernel K and the vRKHS $(\mathcal{H}_K, \langle\cdot,\cdot\rangle_K)$ (see e.g. (Micheli & Glaunès, 2014, Theorem 2.6)), hence modifying $\mathcal{T}$ or changing the inner product changes the kernel. We shall use several classical properties: by symmetry of the scalar product, the matrix-valued kernel has a Hermitian symmetry, i.e. $K(s,t) = K(t,s)^\top$ for any $s,t \in \mathcal{T}$. Moreover, if the vRKHS can be written as $\mathcal{H}_K = \mathcal{H}_0 \oplus \mathcal{H}_1$, then $\mathcal{H}_0$ and $\mathcal{H}_1$ equipped with $\langle\cdot,\cdot\rangle_K$ are also vRKHSs, as closed subspaces of $\mathcal{H}_K$ for $\|\cdot\|_K$, and their kernels K_0 and K_1 satisfy $K = K_0 + K_1$.

Let us define our candidate for a vRKHS, the space $\mathcal{S}$ of trajectories satisfying the dynamical system of ($\mathcal{P}_0$):

$$\mathcal{S} := \{\mathbf{x}(\cdot) \,|\, \exists \mathbf{u}(\cdot) \text{ s.t. } \mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \text{ a.e. and } \int_0^T \mathbf{u}(t)^\top \mathbf{R}(t)\mathbf{u}(t)dt < \infty\}. \tag{4.1}$$

There is not necessarily a unique choice of $\mathbf{u}(\cdot)$ for a given $\mathbf{x}(\cdot) \in \mathcal{S}$ (for instance if $\mathbf{B}(t)$ is not injective for some t). Therefore, with each $\mathbf{x}(\cdot) \in \mathcal{S}$, we associate the control $\mathbf{u}(\cdot)$ having minimal norm based on the pseudoinverse of $\mathbf{B}(t)^\ominus$ of $\mathbf{B}(t)$ for the $\mathbb{R}^M$-norm $\|\cdot\|_{\mathbf{R}(t)} := \|\mathbf{R}(t)^{1/2}\cdot\|$:

$$\mathbf{u}(t) = \mathbf{B}(t)^\ominus[\mathbf{x}'(t) - \mathbf{A}(t)\mathbf{x}(t)] \text{ a.e. in } [0,T]. \tag{4.2}$$

The vector space $\mathcal{S}$ has then a natural scalar-product. As a matter of fact, the expression

$$\langle \mathbf{x}_1(\cdot), \mathbf{x}_2(\cdot)\rangle_K := \mathbf{x}_1(0)^\top \mathbf{x}_2(0) + \int_0^T [\mathbf{x}_1(t)^\top \mathbf{Q}(t)\mathbf{x}_2(t) + \mathbf{u}_1(t)^\top \mathbf{R}(t)\mathbf{u}_2(t)]dt \tag{4.3}$$

is bilinear and symmetric over $\mathcal{S} \times \mathcal{S}$.[3] It is positive definite over $\mathcal{S}$ as $\|\mathbf{x}(\cdot)\|_K^2 = 0$ implies that $\mathbf{u}(\cdot) \overset{a.e}{\equiv} 0$ and $\mathbf{x}(0) = 0$, hence that $\mathbf{x}(\cdot) \equiv 0$. Combining (4.2) and (4.3), we can express $\|\cdot\|_K$ as a Sobolev-like norm split into two semi-norms $\|\cdot\|_{K_0}$ and $\|\cdot\|_{K_1}$

$$\|\mathbf{x}(\cdot)\|_K^2 = \underbrace{\|\mathbf{x}(0)\|^2}_{\|\mathbf{x}(\cdot)\|_{K_0}^2} + \underbrace{\int_0^T [\|\mathbf{x}(t)\|_{\mathbf{Q}(t)}^2 + \|\mathbf{B}(t)^{\ominus}(\mathbf{x}'(t) - \mathbf{A}(t)\mathbf{x}(t))\|_{\mathbf{R}(t)}^2] dt}_{\|\mathbf{x}(\cdot)\|_{K_1}^2}. \tag{4.4}$$

By Cauchy-Lipschitz's theorem, $\|\mathbf{x}(\cdot)\|_{K_0} = \|\mathbf{x}(0)\|$ defines a norm over the finite-dimensional subspace $\mathcal{S}_0$ of trajectories with null quadratic cost (hence null control):

$$\mathcal{S}_0 := \{\mathbf{x}(\cdot) \mid \int_0^T \mathbf{x}(t)^\top \mathbf{Q}(t)\mathbf{x}(t)dt = 0 \text{ and } \mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t), \text{ a.e. in } [0,T]\}. \tag{4.5}$$

Let us define its (infinite-dimensional) orthogonal complement $\mathcal{S}_u := (\mathcal{S}_0)^\perp$ in $\mathcal{S}$ w.r.t. $\|\cdot\|_K$. From now on we equip $\mathcal{S}$ (resp. $\mathcal{S}_0$, $\mathcal{S}_u$) with $\|\cdot\|_K$ (resp. $\|\cdot\|_{K_0}$, $\|\cdot\|_{K_1}$). These will be shown to all be vRKHSs. Suppose we identified the matrix-valued kernels K, K_0 and K_1, spawning them (a procedure to be found in Section 4.3). In a recent approach in kernel methods (Aubin-Frankowski & Szabó, 2020b), developed for regression problems with constraints over derivatives, we suggested replacing "$\mathbf{C}(t)\mathbf{x}(t) \leqslant \mathbf{d}(t)$ $(\forall\, t \in [0,T])$" by the following strengthened second-order cone (SOC)[4] constraints:

$$\eta_i(\delta_m, t_m)\|\mathbf{x}(\cdot)\|_K + \mathbf{c}_i(t_m)^\top \mathbf{x}(t_m) \leqslant d_i(\delta_m, t_m),\ \forall\, m \in [\![1, N_P]\!], \forall\, i \in [\![1, P]\!] \tag{4.6}$$

where the $(t_m)_{m\in[\![1,N_P]\!]} \in [0,T]^{N_P}$ are N_P time points associated to radii $\delta_m > 0$ satisfying $[0,T] \subset \cup_{m\in[\![1,N_P]\!]}[t_m - \delta_m, t_m + \delta_m]$. The constants $\eta_i(\delta_m, t_m)$ and $d_i(\delta_m, t_m)$ are then defined as:

$$\begin{aligned}
\eta_i(\delta_m, t_m) &:= \sup_{t\in[t_m-\delta_m, t_m+\delta_m]\cap[0,T]} \|K(\cdot, t_m)\mathbf{c}_i(t_m) - K(\cdot, t)\mathbf{c}_i(t)\|_K, \\
d_i(\delta_m, t_m) &:= \inf_{t\in[t_m-\delta_m, t_m+\delta_m]\cap[0,T]} d_i(t).
\end{aligned}$$

This tightening of the constraints stems from interpreting $\eta_i(\delta_m, t_m)\|\mathbf{x}(\cdot)\|_K$ as an upper bound of the modulus of continuity of the unknown $\mathbf{C}(\cdot)\mathbf{x}(\cdot)$ defined as follows

$$\omega_i^{(Cx)}(\delta_m, t_m) := \sup_{t\in[t_m-\delta_m, t_m+\delta_m]\cap[0,T]} \underbrace{|\mathbf{c}_i(t)^\top \mathbf{x}(t) - \mathbf{c}_i(t_m)^\top \mathbf{x}(t_m)|}_{|\langle \mathbf{x}(\cdot), K(\cdot,t)\mathbf{c}(t) - K(\cdot,t_m)\mathbf{c}(t_m)\rangle_K|} \leqslant \eta_i(\delta_m, t_m)\|\mathbf{x}(\cdot)\|_K. \tag{4.7}$$

3. The description by $\mathcal{S}$ of the optimization variables effectively pushes controls in the background while bringing forth trajectories as the main object of study. This describes ($\mathcal{P}_0$) more as a regression problem over $\mathcal{S}$ than as an optimal control problem over controls.

4. The "second-order cone" terminology is classical in optimization following the similarity between (4.6) and the definition of the Lorentz cone $\{(z, r) \in \mathbb{R}^{N+1} \mid \|z\| \leqslant r\}$.

This inequality is obtained applying successively the reproducing property and the Cauchy–Schwarz inequality. Since the intractable modulus of continuity controls the variations of $\mathbf{C}(t)\mathbf{x}(t)$, the SOC upper bound provides a tractable tightening. This interpretation through moduli of continuity was extensively developed in Section 3.2 of Aubin-Frankowski and Szabó (2020a).

With the above notations, our strengthened time-varying linear quadratic optimal control problem with finite horizon and finite number of SOC constraints is [5]

$$\begin{aligned} &\min_{\substack{\boldsymbol{x}(\cdot)\in\mathcal{S},\\ \boldsymbol{x}(0)=\boldsymbol{x}_0}} \quad g(\boldsymbol{x}(T)) + \|\boldsymbol{x}(\cdot)\|_K^2 \\ &\text{s.t.} \\ &\eta_i(\delta_m, t_m)\|\boldsymbol{x}(\cdot)\|_K + \mathbf{c}_i(t_m)^\top \boldsymbol{x}(t_m) \leqslant d_i(\delta_m, t_m),\ \forall\, m \in [\![1, N_P]\!], \forall\, i \in [\![1, P]\!]. \end{aligned} \tag{$\mathcal{P}_{\delta,\mathrm{fin}}$}$$

The introduction of ($\mathcal{P}_{\delta,\mathrm{fin}}$) as an approximation of ($\mathcal{P}_0$) entirely stems from the vRKHS formalism and does not result from optimal control considerations. It relies on an inner approximation of a convex set in an infinite-dimensional Hilbert space (see Aubin-Frankowski & Szabó, 2020a, Section 3.1). From a machine learning perspective, the initial condition and the terminal cost act as a "loss" function, whereas the quadratic cost is turned into a norm over $\mathcal{S}$ and can thus be interpreted as a "regularizer". Departing also from optimal control, the tightening is obtained by incorporating the quadratic part of the objective (4.4) in the state constraints to form (4.6). As discussed in Aubin-Frankowski and Szabó (2020b), introducing (4.6) leads to a finite number of evaluations of the variable $\boldsymbol{x}(\cdot)$ in ($\mathcal{P}_{\delta,\mathrm{fin}}$) which allows for a representer theorem (Theorem 4.1 in Section 4.3.1).

Our goal is to show that ($\mathcal{P}_{\delta,\mathrm{fin}}$) is indeed a tightening of ($\mathcal{P}_0$), enjoying a representer theorem providing a finite-dimensional representation of the solution of problem ($\mathcal{P}_{\delta,\mathrm{fin}}$). We also bound the distance between the trajectories solutions of ($\mathcal{P}_0$) and ($\mathcal{P}_{\delta,\mathrm{fin}}$), and prove that it can be made as small as desired by refining the time-discretization grid.

4.3 REVISITING LQ CONTROL THROUGH KERNEL METHODS

This section presents a step-by-step approach to identify the matrix-valued kernel K of the Hilbert space $\mathcal{S}$ of solutions of a linear control system, equipped with the

5. Since $\|\boldsymbol{x}(\cdot)\|_{K_0}^2 = \|\boldsymbol{x}(0)\|^2$ and $\boldsymbol{x}(0)$ is fixed, we replaced the integral $\|\boldsymbol{x}(\cdot)\|_{K_1}^2$ by $\|\boldsymbol{x}(\cdot)\|_K^2$ in the objective. Recently, in Aubin-Frankowski (2021), another choice of inner product was introduced, by incorporating quadratic terminal costs g into $\|\boldsymbol{x}(\cdot)\|_K^2$ for another choice of inner product than (4.3).

scalar product (4.3). This is done independently from the state constraints which effect is only to select a closed convex subset of the space of trajectories. In Section 4.3.1, we consider the case $\mathbf{Q} \equiv 0$ which enjoys explicit formulas. We also express a representer theorem (Theorem 4.1 suited for problems of the form ($\mathcal{P}_{\delta,\text{fin}}$). This allows us to revisit, through the kernel framework, classical notions, such as the solution of the unconstrained LQR problem, or the definition of the Gramian of controllability. In Section 4.3.2, we consider the case $\mathbf{Q} \not\equiv 0$ and relate our solution to an adjoint equation over matrices. Furthermore, the identification of kernels developed in Section 4.3 is by no means restricted to finite T, hence the kernel formalism can also tackle infinite-horizon problems. [6]

Let us denote by $\mathbf{\Phi_A}(t,s) \in \mathbb{R}^{N,N}$ the state-transition matrix of $\mathbf{x}'(\tau) = \mathbf{A}(\tau)\mathbf{x}(\tau)$, defined from s to t. The key property used throughout this section is the variation of constants formula, a.k.a. Duhamel's principle, stating that for any absolutely continuous $\mathbf{x}(\cdot)$ such that $\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$ a.e. we have

$$\mathbf{x}(t) = \mathbf{\Phi_A}(t,0)\mathbf{x}(0) + \int_0^t \mathbf{\Phi_A}(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau. \tag{4.8}$$

Lemma 4.1. $(\mathcal{S}, \langle\cdot,\cdot\rangle_K)$ *is a vRKHS.*

Proof: We have to show that: i) $(\mathcal{S}, \langle\cdot,\cdot\rangle_K)$ is a Hilbert space, ii) for every $t \in [0,T]$ and $\mathbf{p} \in \mathbb{R}^N$, the evaluation functional $\mathbf{x}(\cdot) \in \mathcal{S} \mapsto \mathbf{p}^\top\mathbf{x}(t) \in \mathbb{R}$ is continuous.
i) From (4.3) and the discussion of Section 4.2, it is obvious that $\langle\cdot,\cdot\rangle_K$ is a scalar product. We just have to show that $\mathcal{S}$ is complete. Let $(\mathbf{x}_n(\cdot))_n$ be a Cauchy sequence in $\mathcal{S}$, with associated controls $(\mathbf{u}_n(\cdot))_n$. Then $(\|\mathbf{x}_n(\cdot)\|_K)_n$ is a Cauchy sequence in $\mathbb{R}$ and thus converges, so $(\|\mathbf{x}_n(0)\|)_n$ and $(\|\mathbf{R}(\cdot)^{1/2}\mathbf{u}_n(\cdot)\|_{L^2(0,T)})_n$ are bounded. Since $\mathbf{R}(t) \succcurlyeq r\mathrm{Id}_M$ with $r > 0$, $(\|\mathbf{u}_n(\cdot)\|_{L^2(0,T)})_n$ is thus bounded too, and we can take a subsequence $(\mathbf{u}_{n_i}(\cdot))_i$ weakly converging to some $\mathbf{u}(\cdot)$. Let $s,t \in [0,T]$,

$$\mathbf{x}_n(t) - \mathbf{x}_n(s) \overset{(4.8)}{=} (\mathbf{\Phi_A}(t,s) - \mathrm{Id}_N)\mathbf{x}_n(s) + \int_s^t \mathbf{\Phi_A}(t,\tau)\mathbf{B}(\tau)\mathbf{u}_n(\tau)d\tau. \tag{4.9}$$

Taking $s = 0$, as $(\|\mathbf{x}_n(0)\|)_n$ is bounded, $\mathbf{A}(\cdot) \in L^1(0,T)$ and $\mathbf{B}(\cdot) \in L^2(0,T)$, we have that $\mathbf{\Phi_A}(\cdot,\cdot)$ is continuous and $\{\mathbf{x}_n(\cdot)\}_n$ is uniformly bounded in $\mathcal{C}(0,T)$. Thus (4.9) implies that the sequence $(\mathbf{x}_{n_i}(\cdot))_i$ is equicontinuous. By Ascoli's theorem, we can take a $(\mathbf{x}_{n_{i_j}}(\cdot))_j$ uniformly converging to some $\mathbf{x}(\cdot)$ satisfying (4.8) for $\mathbf{u}(\cdot)$, thus $\mathbf{x}(\cdot) \in \mathcal{S}$.

6. We need to assume T to be finite for the results of Section 4.4 to hold, as we use a procedure based on compact coverings to deal with the state constraints.

ii) Let $t \in [0,T]$, $\mathbf{p} \in \mathbb{R}^N$, and $\mathbf{x}(\cdot) \in \mathcal{S}$. By (4.8) and Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbf{p}^\top \mathbf{x}(t) &= \mathbf{p}^\top \mathbf{\Phi_A}(t,0)\mathbf{x}(0) + \int_0^t \mathbf{p}^\top \mathbf{\Phi_A}(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau \\
|\mathbf{p}^\top \mathbf{x}(t)| &\leqslant \sup_{\tau\in[0,T]} \|\mathbf{p}^\top \mathbf{\Phi_A}(t,\tau)\| \cdot \left(\|\mathbf{x}(0)\| + \| \int_0^t \mathbf{B}(\tau)\mathbf{u}(\tau)d\tau\| \right) \\
&\leqslant \sup_{\tau\in[0,T]} \sqrt{2}\|\mathbf{p}^\top \mathbf{\Phi_A}(t,\tau)\| \cdot \left(\|\mathbf{x}(0)\|^2 + \|\mathbf{B}(\cdot)\|^2_{L^2(0,T)} \|\mathbf{u}(\cdot)\|^2_{L^2(0,T)} \right)^{\frac{1}{2}} \\
&\leqslant \sup_{\tau\in[0,T]} \sqrt{2}\|\mathbf{p}^\top \mathbf{\Phi_A}(t,\tau)\| \cdot \left(1 + \frac{\|\mathbf{B}(\cdot)\|_{L^2(0,T)}}{\sqrt{r}} \right) \left(\|\mathbf{x}(0)\|^2 + \|\mathbf{R}(\cdot)^{\frac{1}{2}}\mathbf{u}(\cdot)\|^2_{L^2(0,T)} \right)^{\frac{1}{2}}.
\end{aligned}$$

Hence the linear map $\mathbf{x}(\cdot) \in \mathcal{S} \mapsto \mathbf{p}^\top \mathbf{x}(t) \in \mathbb{R}$ is continuous.

■

By Definition 4.1, we know that a matrix-valued reproducing kernel $K(\cdot,\cdot)$ exists. We now repeatedly use (4.8) to identify it.

4.3.1 *Case of vanishing Q*

The space $\mathcal{S}$ of controlled trajectories is defined as in (4.1) equipped with the quadratic norm

$$\|\mathbf{x}(\cdot)\|^2_K = \|\mathbf{x}(0)\|^2 + \int_0^T \mathbf{u}(t)^\top \mathbf{R}(t)\mathbf{u}(t)dt = \|\mathbf{x}(\cdot)\|^2_{K_0} + \|\mathbf{x}(\cdot)\|^2_{K_1}, \tag{4.10}$$

where $\mathbf{u}(\cdot)$ is defined as in (4.2). We can further specify its subspaces

$$\mathcal{S}_0 = \{\mathbf{x}(\cdot) \,|\, \mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t), \text{ a.e. in } [0,T]\} \qquad \mathcal{S}_u = \{\mathbf{x}(\cdot) \,|\, \mathbf{x}(\cdot) \in \mathcal{S} \text{ and } \mathbf{x}(0) = 0\}.$$

By uniqueness of the reproducing kernel, we only have to exhibit a candidate $K(s,t)$ satisfying $\mathbf{p}^\top \mathbf{x}(t) = \langle \mathbf{x}(\cdot), K(\cdot,t)\mathbf{p}\rangle_K$ ($\forall t \in [0,T], \mathbf{x} \in \mathcal{S}$, $\mathbf{p} \in \mathbb{R}^N$) and $K(\cdot,t)\mathbf{p} \in \mathcal{S}$ ($\forall t \in [0,T]$, $\mathbf{p} \in \mathbb{R}^N$). The space $(\mathcal{S}_0, \|\cdot\|_{K_0})$ being finite dimensional, we can right away identify its kernel[7]

$$K_0(s,t) = \mathbf{\Phi_A}(s,0)\mathbf{\Phi_A}(t,0)^\top. \tag{4.11}$$

7. This is a classical result for finite dimensional vRKHSs. Fix any family $\{v_j\}_{j\in[\![1,N]\!]}$ spanning $\mathcal{S}_0$, let $\mathbf{V}(s) := [v_j(s)]_{j\in[\![1,N]\!]} \in \mathbb{R}^{N,N}$ and $\mathbf{G}_v := (\langle v_i, v_j\rangle_{K_0})_{i,j\in[\![1,N]\!]}$. The matrix $\mathbf{G}_v$ is invertible as $\|\cdot\|_{K_0}$ is a norm over $\mathcal{S}_0$ and thus $K_0(s,t) = \mathbf{V}(s)^\top \mathbf{G}_v^{-1}\mathbf{V}(t)$. Here we have $\mathbf{V}(s) = \mathbf{\Phi_A}(s,0)^\top$ and $\mathbf{G}_v = \mathrm{Id}_N$.

As $\mathcal{S} = \mathcal{S}_0 \oplus \mathcal{S}_u$, from the properties of sums of kernels, we derive that we should look for K of the form $K_0 + K_1$ for which the reproducing property, with $\langle \cdot, \cdot \rangle_K$ defined in (4.4), writes as follows, for all $t \in [0, T]$, $\mathbf{p} \in \mathbb{R}^N$, $\mathbf{x}(\cdot) \in \mathcal{S}$,

$$\mathbf{p}^\top \mathbf{x}(t) = (K(0,t)\mathbf{p})^\top \mathbf{x}(0) + \int_0^T \left[(\mathbf{B}(t)^\ominus (\partial_1 (K(s,t)\mathbf{p}) - \mathbf{A}(t)K(s,t)\mathbf{p}))^\top \mathbf{R}(s)\mathbf{u}(s) \right] ds. \tag{4.12}$$

Setting $\partial_1 K(s,t) : \mathbf{p} \mapsto \partial_1 (K(s,t)\mathbf{p})$, let us define formally $\mathbf{U}_t(s) := \mathbf{B}(s)^\ominus (\partial_1 K_1(s,t) - \mathbf{A}(s)K_1(s,t))$. By the Hermitian symmetry of K and the fact that $K_0(\cdot, t)\mathbf{p}$ belongs to $\mathcal{S}_0$ and $K_1(0,t)\mathbf{p} = 0$, (4.12) holds if and only if for all $t \in [0,T]$ and $\mathbf{x}(\cdot) \in \mathcal{S}$

$$\mathbf{x}(t) = K_0(t,0)\mathbf{x}(0) + \int_0^T \mathbf{U}_t(s)^\top \mathbf{R}(s)\mathbf{u}(s) ds.$$

This expression can be identified with (4.8) when defining $\mathbf{U}_t(s)$ as follows

$$\mathbf{U}_t(s) := \begin{cases} \mathbf{R}(s)^{-1}\mathbf{B}(s)^\top \mathbf{\Phi}_\mathbf{A}(t,s)^\top & \forall s \leqslant t, \\ 0 & \forall s > t. \end{cases} \tag{4.13}$$

Consequently, to ensure that, for all $t \in [0,T]$, $\mathbf{p} \in \mathbb{R}^N$, $K_1(\cdot, t)\mathbf{p} \in \mathcal{S}_u$, K_1 has to satisfy the following differential equation for any given $t \in [0,T]$:

$$\partial_1 K_1(s,t) = \mathbf{A}(s)K_1(s,t) + \mathbf{B}(s)\mathbf{U}_t(s) \text{ a.e. in } [0,T] \text{ with } K_1(0,t) = 0. \tag{4.14}$$

Since $\mathbf{A}(\cdot) \in L^1(0,T)$, $\mathbf{B}(\cdot) \in L^2(0,T)$, and $\mathbf{R}(t) \succcurlyeq r\mathrm{Id}_M$ with $r > 0$, we have that $\mathbf{U}_t(\cdot) \in L^2(0,T)$. By applying the variation of constants formula to (4.14), with $\mathbf{U}_t(\cdot)$ defined in (4.13), we get an explicit expression for K_1, satisfying a Hermitian symmetry when permuting s and t,

$$K_1(s,t) = \int_0^{\min(s,t)} \mathbf{\Phi}_\mathbf{A}(s,\tau)\mathbf{B}(\tau)\mathbf{R}(\tau)^{-1}\mathbf{B}(\tau)^\top \mathbf{\Phi}_\mathbf{A}(t,\tau)^\top d\tau. \tag{4.15}$$

Remark (Gramian): Formula (4.15) for $K_1(T,T)$ corresponds to the Gramian of controllability. The link is straightforward as the controllability problem of steering a point from $(0,0)$ to $(T, \mathbf{x}_T)$ simply writes as, with $\mathbf{u}(\cdot)$ defined as in (4.2) and $\mathbf{R}(\cdot) \equiv \mathrm{Id}_M$,

$$\begin{aligned} &\min_{\mathbf{x}(\cdot) \in \mathcal{S}} \quad \int_0^T \|\mathbf{u}(t)\|^2 dt \\ &\text{s.t.} \\ &\mathbf{x}(0) = 0, \\ &\mathbf{x}(T) = \mathbf{x}_T, \end{aligned} \tag{4.16}$$

which set of solutions can actually be made explicit.[8] As a matter of fact, in kernel methods, it is classical to look for a "representer theorem", i.e. a necessary condition to ensure that the solutions of an optimization problem live in a finite dimensional subspace of $\mathcal{S}$ and consequently enjoy a finite representation. Such theorems are usually stated without constraints and for real-valued kernels (e.g. Schölkopf et al., 2001). Here we formulate a representer theorem for conic constraints and matrix-valued kernels, as it will prove instrumental to derive a finite formulation for the SOC-strengthening of ($\mathcal{P}_0$).

Theorem 4.1 (Representer theorem). *Let $(\mathcal{H}_k, \langle\cdot,\cdot\rangle_K)$ be a vRKHS defined on a set $\mathcal{T}$. Let $P \in \mathbb{N}$ and, for $i \in [\![0, P]\!]$ and given $N_i \in \mathbb{N}$, $\{t_{i,j}\}_{j\in[\![1,N_i]\!]} \subset \mathcal{T}$. Consider the following optimization problem with "loss" function $L : \mathbb{R}^{N_0} \to \mathbb{R} \cup \{+\infty\}$, strictly increasing "regularizer" function $\Omega : \mathbb{R}_+ \to \mathbb{R}$, and constraints $d_i : \mathbb{R}^{N_i} \to \mathbb{R}$, $\lambda_i \geqslant 0$ and $\{\mathbf{c}_{i,m}\}_{m\in[\![1,N_i]\!]} \subset \mathbb{R}^N$,*

$$\begin{aligned}\bar{\mathbf{f}} \in \underset{\mathbf{f}\in\mathcal{H}_k}{\arg\min} \quad & L\left(\mathbf{c}_{0,1}^\top \mathbf{f}(t_{0,1}), \dots, \mathbf{c}_{0,N_0}^\top \mathbf{f}(t_{0,N_0})\right) + \Omega\left(\|\mathbf{f}\|_K\right) \\ \text{s.t.} \quad & \\ & \lambda_i \|\mathbf{f}\|_K \leqslant d_i(\mathbf{c}_{i,1}^\top \mathbf{f}(t_{i,1}), \dots, \mathbf{c}_{i,N_i}^\top \mathbf{f}(t_{i,N_i})),\ \forall i \in [\![1, P]\!].\end{aligned}$$

Then, for any minimizer $\bar{\mathbf{f}}$, there exists $\{\mathbf{p}_{i,m}\}_{m\in[\![1,N_i]\!]} \subset \mathbb{R}^N$ such that

$$\bar{\mathbf{f}} = \sum_{i=0}^{P} \sum_{m=1}^{N_i} K(\cdot, t_{i,m})\mathbf{p}_{i,m}$$

with $\mathbf{p}_{i,m} = \alpha_{i,m}\mathbf{c}_{i,m}$ for some $\alpha_{i,m} \in \mathbb{R}$.

Proof: Let $\bar{\mathbf{f}}$ be an optimal solution and let $V := \operatorname{span}\left(\{K(\cdot, t_{i,m})\mathbf{c}_{i,m}\}_{m\leqslant N_i,\, i\leqslant P}\right)$. Take $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in V^\perp$ such that $\bar{\mathbf{f}} = \boldsymbol{v} + \boldsymbol{w}$. As $\mathbf{c}_{i,m}^\top \boldsymbol{w}(t_{i,m}) = \langle \boldsymbol{w}(\cdot), K(\cdot, t_{i,m})\mathbf{c}_{i,m}\rangle_K = 0$, the terms appearing in L and d_i are the same for $\bar{\mathbf{f}}$ and $\boldsymbol{v}$. Moreover, $\|\boldsymbol{v}\|_K \leqslant \|\bar{\mathbf{f}}\|_K$, hence $\boldsymbol{v}$ belongs to the constraint set since $\bar{\mathbf{f}}$ does. Furthermore $\Omega\left(\|\boldsymbol{v}\|_K\right) \leqslant \Omega\left(\|\bar{\mathbf{f}}\|_K\right)$, so, by optimality of $\bar{\mathbf{f}}$, $\boldsymbol{w} = 0$ which concludes the proof.

■

In other words, Theorem 4.1 states that to each time t where the variable $\mathbf{f}$ is evaluated corresponds a multiplier $\mathbf{p}_t \in \mathbb{R}^N$ in the expression of the optimal solutions. Hence, if the number of such evaluations is finite, then the representation of $\bar{\mathbf{f}}$ is

8. When choosing the terminal cost $g(\cdot)$ to be equal to the indicator function of $\{x_T\}$, (4.16) does correspond to ($\mathcal{P}_0$) in the absence of state constraints.

finite.[9] Besides, a representer theorem, like Pontryagin's Maximum Principle, is only a necessary condition on the form of the solutions. Theorem 4.1 guarantees the existence or uniqueness of an optimal solution only when coupled with other assumptions (e.g. that L and Ω are convex, and L is lower semi-continuous). This further highlights the analogy with the Maximum Principle in the quadratic case. Theorem 4.1 is instrumental to obtain a finite-dimensional equivalent of ($\mathcal{P}_{\delta,\text{fin}}$).

Theorem 4.1 applied to (4.16) implies that any candidate optimal solution $\bar{x}(\cdot)$ can be written as $\bar{x}(s) = K(s,0)p_0 + K(s,T)p_T$, with $p_0, p_T \in \mathbb{R}^N$. As $\bar{x}(0) = 0$, $\text{proj}_{\mathcal{S}_0}(\bar{x}(\cdot)) = 0$, and as $K_1(\cdot,0) \equiv 0$, $\bar{x}(s) = K_1(s,T)p_T$. So $\bar{x}(T) = x_T$ is satisfied if and only if $x_T \in \text{Im}(K_1(T,T))$ where the operator $K_1(T,T)$ is defined by (4.14), setting $R \equiv \text{Id}$. Hence (4.16) has a solution for any x_T (i.e. the system is controllable) if and only if the Gramian of controllability $K_1(T,T)$ is invertible.

Remark (LQR without state constraints): We derive also from the kernel framework the transversality condition, as well as the classical solution of the LQR problem without state constraints, defined as follows, with $u(\cdot)$ again defined as in (4.2),

$$\begin{aligned} &\min_{x(\cdot)\in\mathcal{S}} \quad g(x(T)) + \frac{1}{2}\int_0^T u(t)^\top R(t)u(t)dt \\ &\text{s.t.} \\ &x(0) = 0. \end{aligned} \tag{$\mathcal{P}_{\text{uncons}}$}$$

Similarly, through the representer theorem, we deduce that $\bar{x}(\cdot) = K_1(\cdot,T)p_T$. Hence, by the reproducing property,

$$\int_0^T \bar{u}^\top(t)R(t)\bar{u}(t)dt = \|K_1(\cdot,T)p_T\|_K^2 = p_T^\top K_1(T,T)p_T.$$

Assume that $g(\cdot) \in \mathcal{C}^1(\mathbb{R}^N,\mathbb{R})$ and that it is convex. Applying the first-order optimality condition, we conclude that

$$0 = \nabla\left(p \mapsto g(K_1(T,T)p) + \frac{1}{2}p^\top K_1(T,T)p\right)(p_T) = K_1(T,T)(\nabla g(K_1(T,T)p_T) + p_T). \tag{4.17}$$

So it is sufficient to take $p_T = -\nabla g(K_1(T,T)p_T) = -\nabla g(\bar{x}(T))$, i.e. to have the transversality condition satisfied. However this formula says more than that, as it covers the

9. The property of having a finite number of evaluations is precisely what distinguishes the unconstrained controllability problem (4.16) or the SOC-constrained problem ($\mathcal{P}_{\delta,\text{fin}}$) from the original state-constrained problem ($\mathcal{P}_0$) which has an infinite number of affine constraints.

problem of degeneracies of the "controllability Gramian" $K_1(T,T)$ and gives an explicit equation (4.17) to be satisfied by $\mathbf{p}_T$. Notice that we do not consider any adjoint equation, only adjoint vectors that are not explicitly propagated. In our framework, the Hamiltonian is implicit.

4.3.2 *Case of nonvanishing Q*

For the case with $\mathbf{Q} \not\equiv 0$, we have a more intricate formula. The reproducing property for K, in which the term $\mathbf{U}_t(\cdot)$ will be explicitly specified below [10], writes as follows, for all $t \in [0,T]$, $\mathbf{p} \in \mathbb{R}^N$, $\mathbf{x}(\cdot) \in \mathcal{S}$,

$$\mathbf{p}^\top \mathbf{x}(t) = (K(0,t)\mathbf{p})^\top \mathbf{x}(0) + \int_0^T (K(s,t)\mathbf{p})^\top \mathbf{Q}(s)\mathbf{x}(s)ds + \int_0^T (\mathbf{U}_t(s)\mathbf{p}))^\top \mathbf{R}(s)\mathbf{u}(s)ds. \tag{4.18}$$

By the Hermitian symmetry of K and the variation of constants formula (4.8), we can rewrite (4.18) as, for all $t \in [0,T]$, $\mathbf{x}(\cdot) \in \mathcal{S}$,

$$\begin{aligned}\mathbf{x}(t) = K(t,0)\mathbf{x}(0) + \int_0^T K(t,s)\mathbf{Q}(s)\left(\mathbf{\Phi}_\mathbf{A}(s,0)\mathbf{x}(0) + \int_0^s \mathbf{\Phi}_\mathbf{A}(s,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau\right)ds \\ + \int_0^T \mathbf{U}_t(s)^\top \mathbf{R}(s)\mathbf{u}(s)ds.\end{aligned}$$

After some regrouping of terms, a change of integration bounds and identification with (4.8), we get the integral equations:

$$\begin{aligned}\mathbf{\Phi}_\mathbf{A}(t,0) &= K(t,0) + \int_0^T K(t,s)\mathbf{Q}(s)\mathbf{\Phi}_\mathbf{A}(s,0)ds, \\ \forall\, s \leqslant t,\ \mathbf{\Phi}_\mathbf{A}(t,s)\mathbf{B}(s) &= \mathbf{U}_t(s)^\top \mathbf{R}(s) + \int_s^T K(t,\tau)\mathbf{Q}(\tau)\mathbf{\Phi}_\mathbf{A}(\tau,s)\mathbf{B}(s)d\tau \\ \forall\, s > t,\ 0 &= \mathbf{U}_t(s)^\top \mathbf{R}(s) + \int_s^T K(t,\tau)\mathbf{Q}(\tau)\mathbf{\Phi}_\mathbf{A}(\tau,s)\mathbf{B}(s)d\tau,\end{aligned}$$

which can be summarized as

$$\begin{aligned}K(t,0) &= \mathbf{\Phi}_\mathbf{A}(t,0) - \tilde{K}(t,0) \text{ with } \tilde{K}(t,s) := \int_s^T K(t,\tau)\mathbf{Q}(\tau)\mathbf{\Phi}_\mathbf{A}(\tau,s)d\tau, \\ \mathbf{U}_t(s)^\top \mathbf{R}(s) &= \begin{cases} (\mathbf{\Phi}_\mathbf{A}(t,s) - \tilde{K}(t,s))\mathbf{B}(s) & \forall s \leqslant t, \\ -\tilde{K}(t,s)\mathbf{B}(s) & \forall s > t. \end{cases}\end{aligned} \tag{4.19}$$

10. Recall that $\mathbf{U}_t(s) := \mathbf{B}(s)^\ominus[\partial_1 K(s,t) - \mathbf{A}(s)K(s,t)]$.

Although not as explicit as formula (4.15) from the case $\mathbf{Q} \equiv 0$, this integral expression for $\mathbf{Q} \not\equiv 0$ will still prove valuable to investigate the regularity of $K(\cdot,\cdot)$ (Lemma 4.2 below). To provide further insight on this expression, again for fixed t, let us introduce formally an adjoint equation for a variable $\mathbf{\Pi}(s,t) \in \mathbb{R}^{N,N}$,

$$\partial_1 \mathbf{\Pi}(s,t) = -\boldsymbol{A}(s)^\top \mathbf{\Pi}(s,t) + \mathbf{Q}(s)K(s,t) \qquad \mathbf{\Pi}(T,t) = \mathrm{Id}_N. \tag{4.20}$$

Again, applying the variation of constants formula to $\mathbf{\Pi}(s,t)$, taking the transpose and owing to the symmetries of $K(\cdot,\cdot)$ and $\mathbf{\Phi}(\cdot,\cdot)$, we derive that

$$\begin{aligned}
\mathbf{\Pi}(s,t) &= \mathbf{\Phi}_{(-\boldsymbol{A}^\top)}(s,T)\mathbf{\Pi}(T,t) + \int_T^s \mathbf{\Phi}_{(-\boldsymbol{A}^\top)}(s,\tau)\mathbf{Q}(\tau)K(\tau,t)\mathrm{d}\tau \\
\mathbf{\Pi}(s,t)^\top &= \mathbf{\Phi}_{\boldsymbol{A}}(T,s) - \int_s^T K(t,\tau)\mathbf{Q}(\tau)\mathbf{\Phi}_{\boldsymbol{A}}(\tau,s)\mathrm{d}\tau = \mathbf{\Phi}_{\boldsymbol{A}}(T,s) - \tilde{K}(t,s).
\end{aligned}$$

Since $\partial_1 K(s,t) = \boldsymbol{A}(s)K(s,t) + \boldsymbol{B}(s)\boldsymbol{U}_t(s)$, by (4.19), for any given time t,

$$\begin{aligned}
\partial_1 K(s,t) &= \boldsymbol{A}(s)K(s,t) + \boldsymbol{B}(s)\boldsymbol{R}(s)^{-1}\boldsymbol{B}(s)^\top \begin{cases} \mathbf{\Pi}(s,t) - \mathbf{\Phi}_{\boldsymbol{A}}(T,s)^\top + \mathbf{\Phi}_{\boldsymbol{A}}(t,s)^\top & \forall s \leqslant t, \\ \mathbf{\Pi}(s,t) - \mathbf{\Phi}_{\boldsymbol{A}}(T,s)^\top & \forall s > t. \end{cases} \\
K(0,t) &= \mathbf{\Pi}(0,t) + \mathbf{\Phi}_{\boldsymbol{A}}(t,0)^\top - \mathbf{\Phi}_{\boldsymbol{A}}(T,0)^\top.
\end{aligned} \tag{4.21}$$

The difference of behavior between the two cases $\mathbf{Q} \equiv 0$ and $\mathbf{Q} \not\equiv 0$ is classical in optimal control. While the control equation runs forward in time, the adjoint equation runs backward. For $\mathbf{Q} \equiv 0$, the adjoint equation can be solved independently from K, which is why $\mathbf{\Pi}_t$ was not introduced. For $\mathbf{Q} \not\equiv 0$, we have two coupled differential equations (4.20)-(4.21) over $K(\cdot,t)$ and $\mathbf{\Pi}_t(\cdot)$. This system does not enjoy an explicit expression, however its solutions can still be computed as a two-point boundary value problem. For quadratic terminal costs, see Aubin-Frankowski (2021), where the connections between computing the kernel and solving the Hamiltonian system or its Riccati equation counterpart are stressed.

4.4 THEORETICAL APPROXIMATION GUARANTEES

In this section, we show that the SOC-constrained problem ($\mathcal{P}_{\delta,\mathrm{fin}}$) is a tightening of the original problem ($\mathcal{P}_0$). We also provide bounds on the $\|\cdot\|_K$-distance between the optimal trajectory of ($\mathcal{P}_0$) and that of ($\mathcal{P}_{\delta,\mathrm{fin}}$). This shows that the SOC-tightening is consistent in a numerical analysis sense, as, for bounded kernels K, convergence in $\|\cdot\|_K$ is stronger than uniform convergence of the states, and also implies convergence of the L^2-norms of the controls. We prove that the kernels $K(\cdot,\cdot)$ identified in Section

4.3 are indeed C^0-continuous.

We shall manipulate various forms of state constraints. We thus write our generic problem $(\mathcal{P}_*)$, with objective $\mathcal{L}$ and constraints defined through the constraint set $\mathcal{V}_*$, as follows

$$\begin{aligned} \bar{\mathbf{x}}^*(\cdot) \in \arg\min_{\mathbf{x}(\cdot)\in\mathcal{V}_*} \quad & \mathcal{L}(\mathbf{x}(\cdot)) := g(\mathbf{x}(T)) + \|\mathbf{x}(\cdot)\|_K^2 \\ \text{s.t.} \quad & \\ & \mathbf{x}(0) = \mathbf{x}_0. \end{aligned} \tag{$\mathcal{P}_*$}$$

Existence and uniqueness of the solution for each $(\mathcal{P}_*)$ will be discussed below. Recall that $\mathcal{S}$ is defined in (4.1). Let $(t_m)_{m\in[\![1,N_0]\!]} \in [0,T]^{N_0}$ be N_0 time points associated to radii $\delta_m > 0$ chosen so that they form a covering $[0,T] \subset \cup_{m\in[\![1,N_0]\!]}[t_m-\delta_m, t_m+\delta_m]$. The vectors $\mathbf{d}_m(\delta_m,t_m) := (d_i(\delta_m,t_m))_{i\in[\![1,P]\!]} \in \mathbb{R}^P$, $\overrightarrow{\boldsymbol{\eta}}(\delta,t) := (\eta_i(\delta,t))_{i\in[\![1,P]\!]} \in \mathbb{R}^P$ and $\overrightarrow{\boldsymbol{\omega}}(\delta,t) := (\omega_i(\delta,t))_{i\in[\![1,P]\!]} \in \mathbb{R}^P$ are defined component-wise: [11]

$$\eta_i(\delta,t) := \sup_{s\in[t-\delta,t+\delta]\cap[0,T]} \|K(\cdot,t)\mathbf{c}_i(t) - K(\cdot,s)\mathbf{c}_i(s)\|_K, \tag{4.22}$$

$$\omega_i(\delta,t) := \sup_{s\in[t-\delta,t+\delta]\cap[0,T]} |d_i(t) - d_i(s)|, \tag{4.23}$$

$$d_i(\delta_m,t_m) := \inf_{s\in[t_m-\delta_m,t_m+\delta_m]\cap[0,T]} d_i(s). \tag{4.24}$$

For $\overrightarrow{\boldsymbol{\epsilon}} \in \mathbb{R}^P_+$, we shall consider the following constraints

$$\begin{aligned} \mathcal{V}_0 &:= \{\mathbf{x}(\cdot)\in\mathcal{S} \,|\, \mathbf{C}(t)\mathbf{x}(t) \leqslant \mathbf{d}(t),\ \forall\, t\in[0,T]\}, \\ \mathcal{V}_{\delta,\text{fin}} &:= \{\mathbf{x}(\cdot)\in\mathcal{S} \,|\, \overrightarrow{\boldsymbol{\eta}}(\delta_m,t_m)\|\mathbf{x}(\cdot)\|_K + \mathbf{C}(t_m)\mathbf{x}(t_m) \leqslant \mathbf{d}(\delta_m,t_m),\ \forall\, m\in[\![1,N_0]\!]\}, \\ \mathcal{V}_{\delta,\text{inf}} &:= \{\mathbf{x}(\cdot)\in\mathcal{S} \,|\, \overrightarrow{\boldsymbol{\eta}}(\delta,t)\|\mathbf{x}(\cdot)\|_K + \overrightarrow{\boldsymbol{\omega}}(\delta,t) + \mathbf{C}(t)\mathbf{x}(t) \leqslant \mathbf{d}(t),\ \forall\, t\in[0,T]\}, \\ \mathcal{V}_\epsilon &:= \{\mathbf{x}(\cdot)\in\mathcal{S} \,|\, \overrightarrow{\boldsymbol{\epsilon}} + \mathbf{C}(t)\mathbf{x}(t) \leqslant \mathbf{d}(t),\ \forall\, t\in[0,T]\}. \end{aligned}$$

To these closed constraint sets correspond the problems $(\mathcal{P}_0)$, $(\mathcal{P}_{\delta,\text{fin}})$ $(\mathcal{P}_{\delta,\text{inf}})$, and $(\mathcal{P}_\epsilon)$. In particular, $\bar{\mathbf{x}}^0(\cdot)$ denotes the optimal solution of $(\mathcal{P}_0)$. When $\mathbf{C}(\cdot)$ and $\mathbf{d}(\cdot)$ are $\mathcal{C}^0$-continuous, we prove right away that $\eta_i(\cdot,t)$ and $\omega_i(\cdot,t)$ converge uniformly in t to 0 as $\delta\to 0^+$, so that the SOC inequalities defining the set $\mathcal{V}_{\delta,\text{inf}}$ converge to the original affine constraints in a pointwise sense.

11. The computation of η_i can be performed using that, by the reproducing property, $\|K(\cdot,t)\mathbf{c}_i(t) - K(\cdot,s)\mathbf{c}_i(s)\|_K^2 = \mathbf{c}_i(t)^\top K(t,t)\mathbf{c}_i(t) + \mathbf{c}_i(s)^\top K(s,s)\mathbf{c}_i(s) - 2\mathbf{c}_i(t)^\top K(t,s)\mathbf{c}_i(s)$. We chose to overload the notation of d_i to define the constants $d_i(\delta_m,t_m)$ in order to draw the parallel with the other perturbations of the constraints, $\eta_i(\delta,t)$ and $\omega_i(\delta,t)$.

Lemma 4.2 (Uniform continuity of K). *If* $\mathbf{A}(\cdot) \in L^1(0,T)$ *and* $\mathbf{B}(\cdot) \in L^2(0,T)$*, then* $K(\cdot,\cdot)$ *is uniformly continuous. Assume furthermore that* $\mathbf{C}(\cdot)$ *and* $\mathbf{d}(\cdot)$ *are* $\mathcal{C}^0$*-continuous. Then for all* $i \in [\![1,P]\!]$*, the increasing functions* $\eta_i(\cdot,t)$ *and* $\omega_i(\cdot,t)$ *converge to* 0 *uniformly w.r.t.* t *as* $\delta \to 0^+$.

Proof: Since $\mathbf{A}(\cdot) \in L^1(0,T)$, $\mathbf{\Phi}_{\mathbf{A}}(\cdot,\cdot)$ is uniformly continuous. Hence, for $\mathbf{Q} \equiv 0$, through the explicit formulas (4.11) and (4.15), we deduce that $K(\cdot,\cdot)$ is uniformly continuous. Consequently, for any $i \in [\![1,P]\!]$, and $s,t \in [0,T]$, recalling that $\mathbb{B}_N$ denotes the closed unit ball of $\mathbb{R}^N$,

$$\begin{aligned}\|K(\cdot,t)\mathbf{c}_i(t) - K(\cdot,s)\mathbf{c}_i(s)\|_K &\leqslant \|(K(\cdot,t)-K(\cdot,s))\mathbf{c}_i(t)\|_K + \|K(\cdot,s)(\mathbf{c}_i(t)-\mathbf{c}_i(s))\|_K \\ &\leqslant \|\mathbf{c}_i(t)\| \sup_{\mathbf{p}\in\mathbb{B}_N} \|(K(\cdot,t)-K(\cdot,s))\mathbf{p}\|_K + \|\mathbf{c}_i(t)-\mathbf{c}_i(s)\| \sup_{\mathbf{p}\in\mathbb{B}_N} \|\mathbf{p}^\top K(s,s)\mathbf{p}\|^{1/2},\end{aligned}$$

which proves the statement for $\eta_i(\cdot,t)$, whereas the result for $\omega_i(\cdot,t)$ stems directly from the uniform continuity of $\mathbf{d}(\cdot)$. Obviously, the components $\eta_i(\cdot,t)$ and $\omega_i(\cdot,t)$ are increasing for any given t.

For $\mathbf{Q} \not\equiv 0$, we do not have explicit formulas such as (4.11) and (4.15). Nonetheless, the $\|\cdot\|_K$-norm (4.4) for $\mathbf{Q} \not\equiv 0$ is stronger than the $\|\cdot\|_K$-norm for $\mathbf{Q} \equiv 0$. Since, for $\mathbf{Q} \equiv 0$, K is uniformly continuous, owing to Schwartz, 1964, Proposition 24, the topology induced by K over $\mathcal{S}$ is stronger than the topology of uniform convergence over $[0,T]$. Hence the topology induced by K for $\mathbf{Q} \not\equiv 0$ is also stronger. Therefore, using again the result of Schwartz (1964), for $\mathbf{Q} \not\equiv 0$, $K(\cdot,\cdot)$ is continuous w.r.t. each variable and locally bounded.[12] Hence $K(\cdot,\cdot)$ is bounded on the compact set $[0,T]\times[0,T]$. Let us prove the continuity of $t \mapsto K(t,t)$. Since K is bounded, by (4.19), $\tilde{K}$ is bounded, so $\mathbf{U}_t(\cdot) \in L^2(0,T)$. Let $\mathbf{p} \in \mathbb{B}_N$ and $t \in [0,T]$, then, by definition of K, $K(\cdot,t)\mathbf{p} \in \mathcal{S}$ is associated to the control $\mathbf{U}_t(\cdot)\mathbf{p}$. Let $\delta > 0$, by the variation of constants formula (4.8), we have

$$\begin{aligned}\|(K(t+\delta,t)-K(t,t))\mathbf{p}\| \leqslant \|\mathbf{\Phi}_{\mathbf{A}}(t+\delta,0)-\Phi_{\mathbf{A}}(t,0)\| \cdot \|K(0,t)\mathbf{p}\| \\ + \int_t^{t+\delta} \|\mathbf{\Phi}_{\mathbf{A}}(t,s)\mathbf{B}(s)\mathbf{U}_t(s)\mathbf{p}\| ds.\end{aligned}$$

Let $\lambda > 0$. With a similar computation when permuting t and $t+\delta$, taking the supremum over $\mathbf{p} \in \mathbb{B}_N$, one can find $\Delta > 0$ such that for any $\delta \in [0,\Delta]$,

$$\max(\|K(t+\delta,t)-K(t,t)\|, \|K(t,t+\delta)-K(t+\delta,t+\delta)\|) \leqslant \lambda/2.$$

12. This allows to derive the continuity of $\eta_i(\cdot,t)$ but does not provide a uniform bound w.r.t to t.

Hence, owing to the Hermitian symmetry of K, for any $\delta \in [0, \Delta]$,

$$\|K(t+\delta, t+\delta) - K(t,t)\| \leqslant \|K(t+\delta, t) - K(t,t)\| + \|K(t, t+\delta) - K(t+\delta, t+\delta)\| \leqslant \lambda.$$

This shows that K is indeed continuous on the diagonal. As underlined by Laurent Schwartz, showing the continuity of $t \mapsto K(t,t)$ is enough to conclude. We reproduce briefly his argument (see Schwartz, 1964, p194): for any $\mathbf{p} \in \mathbb{R}^N$, whenever t converges to t_0, $K(\cdot, t)\mathbf{p}$ weakly converges in $\mathcal{S}$ to $K(\cdot, t_0)\mathbf{p}$, however, by continuity on the diagonal, the norm $\mathbf{p}^\top K(t,t)\mathbf{p}$ converges to $\mathbf{p}^\top K(t_0,t_0)\mathbf{p}$, so $K(\cdot,t)\mathbf{p}$ strongly converges in $\mathcal{S}$ and by extension in $\mathcal{C}^0(0,T)$, exactly showing that $K(\cdot,\cdot)$ is continuous, hence uniformly continuous.

■

Generically, under the minimal assumptions of Lemma 4.2, one can apply the SOC-scheme to obtain $(\mathcal{P}_{\delta,\mathrm{fin}})$, which is equivalent to a finite dimensional problem owing to the representer theorem (Theorem 4.1). By Lemma 4.2, the scheme is coherent since for δ decreasing to zero, the coefficients of the SOC constraints converge uniformly in t to those of the original problem $(\mathcal{P}_0)$. However, ensuring that the solution $\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(\cdot)$ converges to $\bar{\mathbf{x}}^0(\cdot)$ requires a more thorough analysis.

Proposition 4.1 (Nested sequence). *Let $\delta_{max} := \max_{m\in[\![1,N_0]\!]} \delta_m$. For any $\delta \geqslant \delta_{max}$, if, for a given $y_0 \geqslant 0$, $\overrightarrow{\boldsymbol{\epsilon}} \geqslant \sup_{t\in[0,T]}[\overrightarrow{\boldsymbol{\eta}}(\delta,t)y_0 + \overrightarrow{\boldsymbol{\omega}}(\delta,t)]$, then we have a nested sequence*

$$(\mathcal{V}_\epsilon \cap y_0 \mathbb{B}_K) \subset \mathcal{V}_{\delta,inf} \subset \mathcal{V}_{\delta,fin} \subset \mathcal{V}_0. \tag{4.25}$$

Proof: The inclusion $\mathcal{V}_\epsilon \cap y_0\mathbb{B}_K \subset \mathcal{V}_{\delta,\mathrm{inf}}$ stems from the definition of the sets. Since $d_i(\delta_m, t_m) \geqslant d_i(t_m) - \omega_i(\delta_m, t_m)$, $\mathcal{V}_{\delta,\mathrm{inf}} \subset \mathcal{V}_{\delta,\mathrm{fin}}$. Recall that $[0,T] \subset \cup_{m\in[\![1,N_0]\!]}[t_m - \delta_m, t_m + \delta_m]$. Let $t \in [0,T]$ and $\mathbf{x}(\cdot) \in \mathcal{V}_{\delta,\mathrm{fin}}$. Take $m \in [\![1, N_0]\!]$ such that $t \in [t_m - \delta_m, t_m + \delta_m]$. For any $i \in [\![1, P]\!]$, applying the reproducing property and Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbf{c}_i(t)^\top \mathbf{x}(t) &= \mathbf{c}_i(t_m)^\top \mathbf{x}(t_m) + \langle \mathbf{x}(\cdot), K(\cdot,t)\mathbf{c}_i(t) - K(\cdot,t_m)\mathbf{c}_i(t_m)\rangle_K \\
\mathbf{c}_i(t)^\top \mathbf{x}(t) &\leqslant \mathbf{c}_i(t_m)^\top \mathbf{x}(t_m) + \eta_i(\delta_m, t_m)\|\mathbf{x}(\cdot)\|_K \leqslant d_i(\delta_m, t_m) \leqslant d_i(t),
\end{aligned}$$

$d_i(\delta_m, t_m)$ being by definition the infimum of the i-th component $d_i(t)$ of $\mathbf{d}(t)$ on $[t_m - \delta_m, t_m + \delta_m]$. So $\mathbf{x}(\cdot) \in \mathcal{V}_0$, hence $\mathcal{V}_{\delta,\mathrm{fin}} \subset \mathcal{V}_0$.

■

Proposition 4.1 states that, on the one hand, enforcing a finite number of SOC constraints with $\overrightarrow{\boldsymbol{\eta}}$ as in (4.22) is more restrictive than enforcing an infinite number of

affine constraints. On the other hand, SOC constraints are less restrictive than shrinking the affine constraints by some $\overrightarrow{\epsilon} > 0$. The nested property (4.25) is instrumental in our analysis. As a matter of fact, we shall focus on ϵ-perturbations of affine constraints rather than on SOC constraints to construct a trajectory $\mathbf{x}^\epsilon(\cdot) \in \mathcal{V}_\epsilon$ close to $\bar{\mathbf{x}}^0(\cdot)$. We shall then resort to strong convexity arguments to derive bounds on $\|\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(\cdot) - \bar{\mathbf{x}}^0(\cdot)\|_K$.

We now list the hypotheses used to prove our main result.

(H-GEN) $\mathbf{A}(\cdot) \in L^1(0,T)$ and $\mathbf{B}(\cdot) \in L^2(0,T)$, $\mathbf{C}(\cdot)$ and $\mathbf{d}(\cdot)$ are $\mathcal{C}^0$-continuous.

(H-SOL) $\mathbf{C}(0)\mathbf{x}_0 < \mathbf{d}(0)$ and there exists $\overrightarrow{\epsilon} > 0$ such that $\mathcal{V}_\epsilon \cap \{\mathbf{x}(\cdot) \,|\, \mathbf{x}(0) = \mathbf{x}_0\} \neq \emptyset$, i.e. there exists a trajectory $\mathbf{x}^\epsilon(\cdot) \in \mathcal{S}$ satisfying strictly the affine constraints, as well as the initial condition, with $\mathbf{x}_0$ interior to the state constraints.

(H-L) There exists $\mu > 0$ such that the objective function $\mathcal{L} : \mathbf{x}(\cdot) \in \mathcal{S} \mapsto g(\mathbf{x}(T)) + \|\mathbf{x}(\cdot)\|_K^2$ is μ-strongly convex. The terminal cost $g(\cdot)$ is continuous over $\mathbb{R}^N$, $\mathbf{Q}(\cdot) \in L^1(0,T)$, and $\mathbf{R}(\cdot) \in L^2(0,T)$. There exists $r > 0$ such that $\mathbf{R}(t) \succcurlyeq r\mathrm{Id}_M$ for all $t \in [0,T]$.

Discussion of the Assumptions: Assumption (H-gen) ensures the $\mathcal{C}^0$-continuity of the kernel K and of the functions η_i and ω_i (Lemma 4.2). Assumption (H-L) concerns the objective function, whereas (H-sol) ensures that the set of trajectories satisfying the state constraints is non-empty if the latter are shrunk:

— The existence requirement in (H-sol) can be derived from assumptions on the existence of interior viable trajectories. We provide in the Annex an example of such assumptions (Lemma 4.4) based on inward pointing conditions on the boundary and on regularity assumptions on the constraints and the dynamics. The assumption $\mathbf{C}(0)\mathbf{x}_0 < \mathbf{d}_0$ ensures that $\mathbf{x}_0$ is a suitable initial condition for the $\overrightarrow{\epsilon}$-tightening $\mathcal{V}_\epsilon$.[13]

— The strong convexity requirement in (H-L) is obviously satisfied whenever $g(\cdot)$ is convex.[14] It is required in order to bound the distance on solutions since the problems $(\mathcal{P}_*)$ share the same objective but different constraint sets. The terminal cost $g(\cdot)$ is supposed $\mathcal{C}^0$-continuous, but could be taken merely locally continuous in a neighborhood[15] of $\bar{\mathbf{x}}^0(T)$ and lower bounded over any compact subset of $\mathbb{R}^N$. By Lemma 4.3, $(\mathcal{P}_0)$ has a unique optimal solution $\bar{\mathbf{x}}^0(\cdot)$.

13. Since the SOC tightening lies in-between the $\overrightarrow{\epsilon}$-tightening and the original constraints (Proposition 4.1), it cannot be guaranteed that an initial condition on the border of the constraints would be suitable for the SOC tightening.

14. More generally, g could be μ_0-semiconvex (i.e. $g(\cdot) + \frac{\mu_0}{2}\|\cdot\|^2$ is convex) with $2 > \mu_0 \sup_{\mathbf{p}\in\mathbb{B}_N} \|\mathbf{p}^\top K(T,T)\mathbf{p}\|^{1/2}$.

15. This neighborhood is considered with respect to the relative topology of the terminal constraint set $\{\mathbf{x} \in \mathbb{R}^N \,|\, \mathbf{C}(T)\mathbf{x} \leqslant \mathbf{d}(T)\}$.

Lemma 4.3 (Existence and uniqueness of solutions). *Under Assumptions (H-gen), (H-sol), and (H-L), $\bar{\mathbf{x}}^0(\cdot)$ exists and is unique. The same result holds true for $(\mathcal{P}_{\delta,fin})$, $(\mathcal{P}_{\delta,inf})$, and $(\mathcal{P}_\epsilon)$ for any $\delta \in [0, \delta_0]$, where $\delta_0 > 0$ satisfies that $\overrightarrow{\epsilon} \geqslant \sup_{t\in[0,T]}[\overrightarrow{\eta}(\delta_0, t)\|\mathbf{x}^\epsilon(\cdot)\|_K + \overrightarrow{\omega}(\delta_0, t)]$ for $\overrightarrow{\epsilon}$ as in (H-sol).*

Proof: The existence result is a consequence of Tonelli's direct method, usually stated for lower bounded and lower semi-continuous $g(\cdot)$. We detail the proof since our Assumptions are both slightly different and stronger. Since by Assumption (H-sol), $\mathcal{V}_0 \neq \emptyset$, let $(\mathbf{x}_n(\cdot), \mathbf{u}_n(\cdot))$ be a minimizing sequence of $(\mathcal{P}_0)$ converging to the optimal value $\bar{m}$. As $\mathcal{L}(\cdot)$ is μ-strongly convex, $\frac{\mathcal{L}(\mathbf{x}(\cdot))}{\|\mathbf{x}(\cdot)\|_K} \to +\infty$ as $\|\mathbf{x}(\cdot)\|_K \to +\infty$. Hence $\bar{m}$ is finite, and $(\mathbf{x}_n(\cdot))_n$ is a subset of a ball $M_0\mathbb{B}_K \subset \mathcal{S}$, for some $M_0 > 0$. Since $\mathbf{x}(\cdot) \in \mathcal{S} \mapsto \mathbf{x}(T)$ is continuous, $\{\mathbf{x}(T) \,|\, \mathbf{x}(\cdot) \in M_0\mathbb{B}_K\}$ is also bounded. By continuity of $g(\cdot)$, let $m_g := -\inf_{\mathbf{x}(\cdot)\in M_0\mathbb{B}_K} g(\mathbf{x}(T)) < +\infty$. Consequently, for n large enough,

$$r\|\mathbf{u}_n(\cdot)\|^2_{L^2(0,T)} \leqslant \|\mathbf{x}_n(\cdot)\|^2_K \leqslant \bar{m} + 1 - g(\mathbf{x}_n(T)) \leqslant \bar{m} + 1 + m_g,$$

so $(\mathbf{u}_n(\cdot))_n$ is bounded in L^2, and we can take a subsequence $(\mathbf{u}_{n_i}(\cdot))_i$ weakly converging to some $\mathbf{u}(\cdot)$. Let $s, t \in [0, T]$. By the variation of constants formula (4.8), we had derived (4.9). Since $\mathbf{A}(\cdot) \in L^1(0, T)$ and $\mathbf{B}(\cdot) \in L^2(0, T)$, and $\mathbf{x}_n(s)$ is uniformly bounded in n and s by the reproducing property and continuity of K (Lemma 4.2), $(\mathbf{x}_{n_i}(\cdot))_i$ is equicontinuous. By Ascoli's theorem, we thus have a subsequence $(\mathbf{x}_{n_{i_j}}(\cdot))_j$ uniformly converging to some $\mathbf{x}(\cdot)$ satisfying (4.8) for $\mathbf{u}(\cdot)$, thus $\mathbf{x}(\cdot) \in \mathcal{S}$. By continuity of $g(\cdot)$, $\mathcal{L}(\mathbf{x}(\cdot)) = \bar{m}$. Since $\mathcal{L}(\cdot)$ is strongly convex, the optimal trajectory is unique and belongs to the closed set $\mathcal{V}_0$. To conclude, replace $\mathcal{V}_0$ with $\mathcal{V}_{\delta,\text{fin}}$ (resp. $\mathcal{V}_{\delta,\text{inf}}$, and $\mathcal{V}_\epsilon$), the inequality satisfied by δ_0 shows that $\mathbf{x}^\epsilon(\cdot) \in \mathcal{V}_{\delta,\text{fin}}$, consequently the constraint sets are non-empty. The same arguments as above yield the result.

■

Theorem 4.2 (Main result - Approximation by SOC constraints). *Under Assumptions (H-gen), (H-sol), and (H-L), for any $\lambda > 0$, there exists $\bar{\delta} > 0$ such that for all $N_0 > 0$ and $(\delta_m)_{m\in[\![1,N_0]\!]}$, with $[0, T] \subset \cup_{m\in[\![1,N_0]\!]}[t_m - \delta_m, t_m + \delta_m]$ satisfying $\bar{\delta} \geqslant \max_{m\in[\![1,N_0]\!]} \delta_m$, we have*

$$\frac{1}{\gamma_K} \cdot \sup_{t\in[0,T]} \|\bar{\mathbf{x}}^{\delta,fin}(t) - \bar{\mathbf{x}}^0(t)\| \leqslant \|\bar{\mathbf{x}}^{\delta,fin}(\cdot) - \bar{\mathbf{x}}^0(\cdot)\|_K \leqslant \lambda \tag{4.26}$$

with $\gamma_K := \sup_{t\in[0,T],\, \mathbf{p}\in\mathbb{B}_N} \sqrt{\mathbf{p}^\top K(t, t)\mathbf{p}}$.

Proof: Let $\lambda > 0$. Consider any $\tilde{\lambda} > 0$ such that

$$2\tilde{\lambda} + \tilde{\lambda}(\tilde{\lambda} + 2\|\bar{\mathbf{x}}^0(\cdot)\|_{L^\infty(0,T)})\|\mathbf{Q}(\cdot)\|_{L^1(0,T)} \leqslant \lambda. \tag{4.27}$$

By Assumption (H-sol), pick $x^\epsilon(\cdot) \in \mathcal{V}_\epsilon$ such that $x^\epsilon(0) = \bar{x}^0(0)$. Denote by $u^\epsilon(\cdot)$ the associated control. Take $\alpha > 0$ small enough such that $x^{\alpha\epsilon}(\cdot) := \alpha x^\epsilon(\cdot) + (1-\alpha)\bar{x}^0(\cdot) \in \mathcal{S}$ and $u^{\alpha\epsilon}(\cdot) := \alpha u^\epsilon(\cdot) + (1-\alpha)\bar{u}^0(\cdot)$ satisfy

$$\|\bar{x}^0(\cdot) - x^{\alpha\epsilon}(\cdot)\|_{L^\infty(0,T)} = \alpha\|\bar{x}^0(\cdot) - x^\epsilon(\cdot)\|_{L^\infty(0,T)} \leqslant \tilde{\lambda}$$
$$|\|R(\cdot)^{1/2}\bar{u}(\cdot)\|^2_{L^2(0,T)} - \|R(\cdot)^{1/2}u^{\alpha\epsilon}(\cdot)\|^2_{L^2(0,T)}| \leqslant \tilde{\lambda}$$

and, by continuity of $g(\cdot)$, $|g(\bar{x}^0(T)) - g(x^{\alpha\epsilon}(T))| \leqslant \tilde{\lambda}$. Consequently $x^{\alpha\epsilon}(0) = \bar{x}^0(0)$ and for all $t \in [0,T]$, $C(t)x^{\alpha\epsilon}(t) \leqslant \alpha(d(t) - \overrightarrow{\epsilon}) + (1-\alpha)d(t) = d(t) - \alpha\overrightarrow{\epsilon}$, so $x^{\alpha\epsilon}(\cdot) \in \mathcal{V}_{\alpha\epsilon}$. Hence

$$\begin{aligned}
\mathcal{L}(x^{\alpha\epsilon}(\cdot)) - \mathcal{L}(\bar{x}^0(\cdot)) &\leqslant |g(\bar{x}^0(T)) - g(x^{\alpha\epsilon}(T))| + \left|\|\bar{x}^0(\cdot)\|^2_K - \|x^{\alpha\epsilon}(\cdot)\|^2_K\right| \\
&\leqslant \tilde{\lambda} + \int_0^T \left|(\bar{x}^0(t) - x^{\alpha\epsilon}(t))^\top Q(t)(\bar{x}^0(t) + x^{\alpha\epsilon}(t))\right| dt \\
&\quad + \left|\|R(\cdot)^{1/2}\bar{u}(\cdot)\|^2_{L^2(0,T)} - \|R(\cdot)^{1/2}u^{\alpha\epsilon}(\cdot)\|^2_{L^2(0,T)}\right| \\
&\leqslant 2\tilde{\lambda} + \tilde{\lambda}(\tilde{\lambda} + 2\|\bar{x}^0(\cdot)\|_{L^\infty(0,T)})\|Q(\cdot)\|_{L^1(0,T)} \overset{(4.27)}{\leqslant} \lambda.
\end{aligned}$$

Let $\delta_0 > 0$ such that $\alpha\overrightarrow{\epsilon} \geqslant \sup_{t\in[0,T]}[\overrightarrow{\eta}(\delta_0,t)\|x^{\alpha\epsilon}(\cdot)\|_K + \overrightarrow{\omega}(\delta_0,t)]$. Then $x^{\alpha\epsilon}(\cdot) \in \mathcal{V}_{\delta_0,\mathrm{inf}} \subset \mathcal{V}_{\delta_0,\mathrm{fin}}$, the sets thus being non-empty. Notice that, for any $\delta \in [0,\delta_0]$, as $\bar{x}^*(\cdot)$ is optimal for $(\mathcal{P}_*)$, from the nested property (4.25), we derive that

$$\mathcal{L}(\bar{x}^{\delta,\mathrm{fin}}(\cdot)) \leqslant \mathcal{L}(\bar{x}^{\delta,\mathrm{inf}}(\cdot)) \leqslant \mathcal{L}(\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)).$$

As $\mathcal{L}(\cdot)$ is μ-convex, $\mathcal{L}^{-1}(]-\infty, \mathcal{L}(\bar{x}^{\delta_0,\mathrm{inf}}(\cdot))])$ is a bounded set, contained in a ball $M_0\mathbb{B}_K$ for some $M_0 > 0$, and containing all the $\{\bar{x}^{\delta,\mathrm{inf}}(\cdot)\}_{\delta\in[0,\delta_0]}$. Since $\mathcal{S}$ is a vRKHS, $x(\cdot) \in \mathcal{S} \mapsto x(T)$ is continuous. So $\{x(T) \,|\, x(\cdot) \in M_0\mathbb{B}_K\}$ is also bounded. Hence, for any $\delta \in [0,\delta_0]$,

$$g(\bar{x}^{\delta,\mathrm{inf}}(T)) + \|\bar{x}^{\delta,\mathrm{inf}}(\cdot)\|^2_K \leqslant \mathcal{L}(\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)) \leqslant |g(\bar{x}^{\delta_0,\mathrm{inf}}(T))| + \|\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)\|^2_K$$
$$\|\bar{x}^{\delta,\mathrm{inf}}(\cdot)\|_K \leqslant \|\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)\|_K + \sqrt{|g(\bar{x}^{\delta_0,\mathrm{inf}}(T))| + |\inf_{x(\cdot)\in M_0\mathbb{B}_K} g(x(T))|^{\frac{1}{2}}} =: y_0.$$

As $\|\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)\|_K \leqslant y_0$, $x^{\alpha\epsilon}(\cdot)$ and $\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)$ are both admissible for the following problem,

$$\min_{\substack{x(\cdot)\in\mathcal{V}_{\delta_0,\mathrm{inf}} \\ \|x(\cdot)\|_K \leqslant y_0 + \|\bar{x}^{\alpha\epsilon}(\cdot)\|_K}} \mathcal{L}(x(\cdot))$$

with $\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)$ being optimal by definition, hence we have $\mathcal{L}(\bar{x}^{\delta_0,\mathrm{inf}}(\cdot)) \leqslant \mathcal{L}(x^{\alpha\epsilon}(\cdot))$. To conclude, let $\bar{\delta} \in]0,\delta_0]$ such that $\alpha\overrightarrow{\epsilon} \geqslant \sup_{t\in[0,T]}[\overrightarrow{\eta}(\bar{\delta},t)y_0 + \overrightarrow{\omega}(\bar{\delta},t)]$. Then, for any $\delta \in]0,\bar{\delta}]$, by strong convexity of $\mathcal{L}(\cdot)$, $\bar{x}^0(\cdot)$ being optimal for $(\mathcal{P}_0)$,

$$\frac{\mu}{2}\|\bar{x}^{\delta,\mathrm{fin}}(\cdot) - \bar{x}^0(\cdot)\|^2_K \leqslant \mathcal{L}(\bar{x}^{\delta,\mathrm{fin}}(\cdot)) - \mathcal{L}(\bar{x}^0(\cdot)) \leqslant \mathcal{L}(x^{\alpha\epsilon}(\cdot)) - \mathcal{L}(\bar{x}^0) \leqslant \lambda.$$

Replacing λ by $\sqrt{2\lambda/\mu}$, we deduced that $\|\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(\cdot)-\bar{\mathbf{x}}^{0}(\cdot)\|_K \leqslant \lambda$. By Cauchy-Schwarz inequality, for any $t \in [0,T]$, $\|\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(t)-\bar{\mathbf{x}}^{0}(t)\| \leqslant \|\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(\cdot)-\bar{\mathbf{x}}^{0}(\cdot)\|_K \sup_{\mathbf{p}\in\mathbb{B}_N} \sqrt{\mathbf{p}^\top K(t,t)\mathbf{p}}$. By definition of γ_K, taking the supremum over $[0,T]$, we derive the remaining inequality.

■

Theorem 4.2 states that, when the discretization steps $(\delta_m)_m$ go to zero, then the solution $\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(\cdot)$ of the SOC-approximation $(\mathcal{P}_{\delta,\mathrm{fin}})$ can be made arbitrarily close to the solution $\bar{\mathbf{x}}^{0}(\cdot)$ of the original problem $(\mathcal{P}_0)$, uniqueness being ensured by Assumption (H-L). Concerning the stability of solutions under shrinking perturbation of the state constraints, the result of Theorem 4.2 actually also holds when replacing $\bar{\mathbf{x}}^{\delta,\mathrm{fin}}(\cdot)$ by $\bar{\mathbf{x}}^{\epsilon}(\cdot)$, showing that, when $\overrightarrow{\epsilon}$ goes to zero, $\bar{\mathbf{x}}^{\epsilon}(\cdot)$ converges to $\bar{\mathbf{x}}^{0}(\cdot)$.

4.5 FINITE-DIMENSIONAL IMPLEMENTATION AND NUMERICAL EXAMPLE

In this section, we express the finite-dimensional equivalent of problem $(\mathcal{P}_{\delta,\mathrm{fin}})$ owing to the representer theorem (Theorem 4.1) and discuss its implementation on a numerical example. In general, the SOC transformation requires only the minimal hypotheses of Lemma 4.2 to be conceptually grounded. Theorem 4.2 essentially states theoretical guarantees of convergence for small discretization steps. For numerical applications, the problem can thus be extended to incorporate costs or equality constraints at any finite number of intermediate times, as in path planning problems. This situation was already met when discussing the controllability Gramian (4.16). We consequently enrich $(\mathcal{P}_{\delta,\mathrm{fin}})$ to optimization problems considered in Theorem 4.1, of the following form, with $\|\mathbf{x}(\cdot)\|_K^2 = \|\mathbf{x}(0)\|^2 + \|\mathbf{x}(\cdot)\|_{K_1}^2$,

$$\begin{aligned}
&\min_{\mathbf{x}(\cdot)\in\mathcal{S}} \quad L\left(\mathbf{c}_{0,1}^\top \mathbf{x}(t_{0,1}), \dots, \mathbf{c}_{0,N_0}^\top \mathbf{x}(t_{0,N_0})\right) + \|\mathbf{x}(\cdot)\|_K^2 \\
&\text{s.t.} \\
&\eta_i(\delta_{i,m}, t_{i,m})\|\mathbf{x}(\cdot)\|_K + \mathbf{c}_i(t_{i,m})^\top \mathbf{x}(t_{i,m}) \leqslant d_i(\delta_{i,m}, t_{i,m}),\ \forall\, m \in [\![1, N_i]\!], \forall\, i \in [\![1, P]\!],
\end{aligned} \tag{$\mathcal{P}_{\mathrm{SOC}}$}$$

for $\{\mathbf{c}_{0,m}\}_{m\in[\![1,N_0]\!]} \subset \mathbb{R}^N$. The differences between $(\mathcal{P}_{\delta,\mathrm{fin}})$ and $(\mathcal{P}_{\mathrm{SOC}})$ are that $g(\cdot)$, which depended only on the terminal point, is now replaced with a loss $L : \mathbb{R}^{N_0} \to \mathbb{R}\cup\{+\infty\}$, defined on a finite number $(t_{0,m})_{m\in[\![1,N_0]\!]}$ of intermediate points (taken with repetition), and that different discretization grids $(\delta_{i,m}, t_{i,m})_{i,m\in[\![1,N_i]\!]}$ are used for each constraint $i \in [\![1, P]\!]$. As no structural assumptions are imposed on L, it may incorporate indicator functions to account for the initial condition or for rendezvous points. [16] In

16. To write $(\mathcal{P}_{\delta,\mathrm{fin}})$ as $(\mathcal{P}_{\mathrm{SOC}})$, take $N_0 = 2N$, $t_{0,1} = \cdots = t_{0,N} = 0$, $t_{0,N+1} = \cdots = t_{0,N_0} = T$, $c_{0,i} = c_{0,N+i} = e_i$ for $i \in [\![1, N]\!]$. Denoting by $\chi_{x_0}(\cdot)$ the indicator function of x_0, set

Section 4.4, the SOC constraints were introduced to turn an infinite number of affine constraints over $[0, T]$ into a finite number of SOC constraints. The logic is therefore to separate the discrete pointwise requirements (which go to L) from the constraints that should hold on $[0, T]$ (which are approximated by SOC constraints). Since the constraints on $[0, T]$ may apply to different components of the state, we may consider different grid steps for each i.

In general, adding the SOC terms leads to more conservative solutions w.r.t. the one with affine constraints. The SOC constraints formulation only requires estimating $\eta_i(\delta_{i,m}, t_{i,m})$ and $-d_i(\delta_{i,m}, t_{i,m})$ defined in (4.22) and (4.24). Both quantities should in principle be overestimated for the guarantees to hold. However, since the tightening results from a worst-case Cauchy-Schwarz scenario as discussed in (4.7), in practice, underestimation does not affect the numerical results. Moreover the definition (4.22) of η_i is not the only possible formulation.[17] Its choice results from geometrical considerations on coverings[18] of compact sets in infinite-dimensional Hilbert spaces (see Section 3 of Aubin-Frankowski and Szabó (2020a) for more details). Even for other values of η_i than (4.22), considering SOC terms in the constraints proves to be beneficial in terms of local satisfaction of the constraints on a neighborhood of $t_{i,m}$. Besides, the discretization grids considered here are 'static' in the sense that they are fixed before solving ($\mathcal{P}_{SOC}$). Extensions to 'dynamic' grids, refined depending on the optimization steps, can be found in Section 5 of Aubin-Frankowski and Szabó (2020a).

L $\left(\mathbf{c}_{0,1}^\top \mathbf{x}(t_{0,1}), \dots, \mathbf{c}_{0,N_0}^\top \mathbf{x}(t_{0,N_0})\right) := \chi_{x_0}(\mathbf{x}(t_{0,1})) + g(\mathbf{x}(t_{0,N_0}))$ turning a vector notation into its componentwise version.

17. Using that $\mathbf{x}(\cdot) = z_0(\cdot) + z_1(\cdot) \in \mathcal{S}_0 \oplus \mathcal{S}_u$ one could consider two η-terms instead of one to derive a less conservative tightening. In the same spirit $\mathbf{d}(t)$ could be projected onto $\mathcal{S}$, and the projection incorporated in the scalar products of the left hand-side. For $\mathbf{Q}(\cdot) \not\equiv 0$, $\|\mathbf{x}(0)\|^2$ could be replaced in (4.4) by an $\mathcal{S}_0$-norm $\|\mathbf{S}_0\mathbf{x}(0)\|^2$ with a surjective $\mathbf{S}_0 \in \mathbb{R}^{N_0,N}$ where $N_0 = \dim(\mathcal{S}_0)$. This would not change the formulation, but lead to a "tighter" norm $\|\mathbf{x}(\cdot)\|_K$.

18. Definition (4.22) corresponds to a covering made of balls in $\mathcal{S}$.

By Theorem 4.1, for $\mathbf{x}(\cdot) = \sum_{j=0}^{P} \sum_{m=1}^{N_j} K(\cdot, t_{j,m}) \mathbf{p}_{j,m}$ and $z = \|\mathbf{x}(\cdot)\|_K$,[19] ($\mathcal{P}_{\mathrm{SOC}}$) is equivalent to

$$\begin{aligned}
\min_{\substack{z \in \mathbb{R}_+, \\ j \in [\![0,P]\!],\, m \in [\![1,N_j]\!], \\ \mathbf{p}_{j,m} \in \mathbb{R}^N,\, \alpha_{j,m} \in \mathbb{R}}} \quad & L\left(\left(\sum_{j=0}^{P} \sum_{m=1}^{N_j} K(t_{0,n}, t_{j,m}) \mathbf{p}_{j,m} \right)_{n \in [\![1,N_0]\!]} \right) + z^2 \\
\text{s.t.} \quad & \\
z^2 &= \sum_{i=0}^{P} \sum_{n=1}^{N_i} \sum_{j=0}^{P} \sum_{m=1}^{N_j} \mathbf{p}_{i,n}^\top K(t_{i,n}, t_{j,m}) \mathbf{p}_{j,m}, \\
\mathbf{p}_{j,m} &= \alpha_{j,m} \mathbf{c}_j(t_m), \quad \forall m \in [\![1, N_j]\!], \forall j \in [\![1, P]\!], \\
\eta_i(\delta_{i,m}, t_{i,m}) z + \textstyle\sum_{j=0}^{P} \sum_{m=1}^{N_j} \mathbf{c}_i(t_{i,m})^\top K(t_{i,m}, t_{j,m}) \mathbf{p}_{j,m} &\leqslant d_i(\delta_{i,m}, t_{i,m}), \quad \forall m \in [\![1, N_i]\!], \forall i \in [\![1, P]\!].
\end{aligned}$$

For quadratic L with indicator functions, after incorporating the quadratic objective into the constraints through a change of variables, ($\mathcal{P}_{\mathrm{SOC}}$) writes as a second-order cone program (SOCP). Hence this problem can be straightforwardly implemented in convex solvers or modeling frameworks such as CVXGEN (Mattingley & Boyd, 2012). SOCP is slightly more expensive computationally than the quadratic programs (QP) classically derived for LQR (Kojima & Morari, 2004). Both QP and SOCP have polynomial computation times, however the exponent is problem-dependent, so achieving a theoretical comparison is challenging. From a numerical viewpoint, it seems to be also very much framework-dependent. For instance, when using CVXGEN, the baseline cost of calling CVX blurs the difference in computation times between QP and SOCP. As the more time points, the more coefficients, it is beneficial to define the grids as subsets of a 'master grid'. Furthermore when computing η_i or $K(s,t)$, one has to approximate the supremum in (4.22) or the integral in (4.15), for instance through sampling. For time-invariant dynamics and $\mathbf{Q}(\cdot) \equiv 0$, one can use the method proposed in Van Loan (1978) to quickly compute $K_1(s,t)$ in (4.15). If sought for, the analysis of the approximation error would correspond to that of the stability of the solution when perturbing the constants and matrices appearing in the SOCP. This leads to a large body of articles (e.g. Bonnans and Ramirez (2005) and quoting articles) which exceeds the purposes of this study. If one overestimates the value of η and adds a term $\lambda_{\mathrm{cond}} \sum_{i=0}^{P} \sum_{n=1}^{N_i} \|\mathbf{p}_{i,n}\|^2$ to z^2, then the state constraints are further tightened. This is done in practice to improve the conditioning of the matrices, and discussed

19. The reproducing property, $\langle K(\cdot, t_1)\mathbf{p}_1, K(\cdot, t_2)\mathbf{p}_2 \rangle_K = \mathbf{p}_2^\top K(t_2, t_1)\mathbf{p}_1$, applied to $\mathbf{x}(\cdot)$ allows to explicit z.

along other numerical implementation details of a SOC-constrained kernel regression in Aubin-Frankowski and Szabó (2020b). For a sufficiently small error in K, constraints satisfaction could be guaranteed, even when facing numerical errors, through this further tightening involving λ_{cond}.

To highlight the behavior of the SOC-transformation ($\mathcal{P}_{\text{SOC}}$) of problem ($\mathcal{P}_0$), we consider the problem ($\mathcal{P}_{\text{pend}}$) of a 'linear' pendulum with angle $x(t)$ where we control the derivative $u(t)$ of a forcing term $w(t)$, with state constraints both on $w(t)$ and on $\dot{x}(t)$, the full state being $\mathbf{x} := [x, \dot{x}, w] \in \mathbb{R}^3$,

$$
\begin{aligned}
&\min_{x(\cdot),u(\cdot)} \quad -\lambda_T \dot{x}(T) + \lambda_u \|u(\cdot)\|^2_{L^2(0,T)} + \|\mathbf{x}(0)\|^2 \\
&\text{s.t.} \\
&x(0) = 0.5, \quad \dot{x}(0) = 0, \quad w(0) = 0, \\
&x(T/3) = 0.5, \quad x(T) = 0, \\
&\ddot{x}(t) = -10\, x(t) + w(t), \quad \dot{w}(t) = u(t), \text{ a.e. in } [0, T], \\
&\dot{x}(t) \in [-3, +\infty[, \quad w(t) \in [-10, 10], \ \forall\, t \in [0, T].
\end{aligned}
\tag{$\mathcal{P}_{\text{pend}}$}
$$

The objective of ($\mathcal{P}_{\text{pend}}$) can be interpreted as trying to maximize the terminal velocity $\dot{x}(T)$ with an L^2-cost over the control. This example underlines how ($\mathcal{P}_{\text{SOC}}$) effectively allows for more possibilities than ($\mathcal{P}_{\delta,\text{fin}}$). We have both an initial, an intermediate and a final condition on $x(\cdot)$. Since the full state $\mathbf{x} := [x, \dot{x}, w]$ is not fixed at the intermediate time $T/3$, the problem cannot be split into independent problems over two time intervals. We take $\mathbf{R}(\cdot) \equiv \lambda_u \in \mathbb{R}$ in order to have ($\mathcal{P}_{\text{pend}}$) written as ($\mathcal{P}_{\text{SOC}}$). We identify the kernels K_0 and K_1 as in (4.11) and (4.15) and use a uniform grid of N_P points for the two affine constraints over $\dot{x}(t)$ and $w(t)$ and turn them into SOC constraints for all $m \in [\![1, N_P]\!]$

$$
\begin{aligned}
\eta_{\dot{x}} \|\mathbf{x}(\cdot)\|_K - \dot{x}(t_m) &\leqslant -3, \\
\eta_w \|\mathbf{x}(\cdot)\|_K + w(t_m) &\leqslant 10, \\
\eta_w \|\mathbf{x}(\cdot)\|_K - w(t_m) &\leqslant 10,
\end{aligned}
$$

where $\eta_{\dot{x}}$ and η_w are defined as in (4.22) for $\mathbf{C} = [0\text{-}1\,0\,;0\,0\,1\,;0\,0\text{-}1]$. For the experiment of Fig. 4.1, we take $T = 1$, $\lambda_T = 10^6$ and $\lambda_u = 10^4$. All computations took less than 30 seconds.

We first compare SOC constraints with discretized constraints ($\eta_w = 0$) for a moderate value of $N_P = 200$. Whereas the SOC-constrained optimal trajectory is fairly conservative w.r.t. to the bounds over $w(t)$, the optimal trajectory for discretized constraints suffers from chattering and does not satisfy the w-constraints. This was

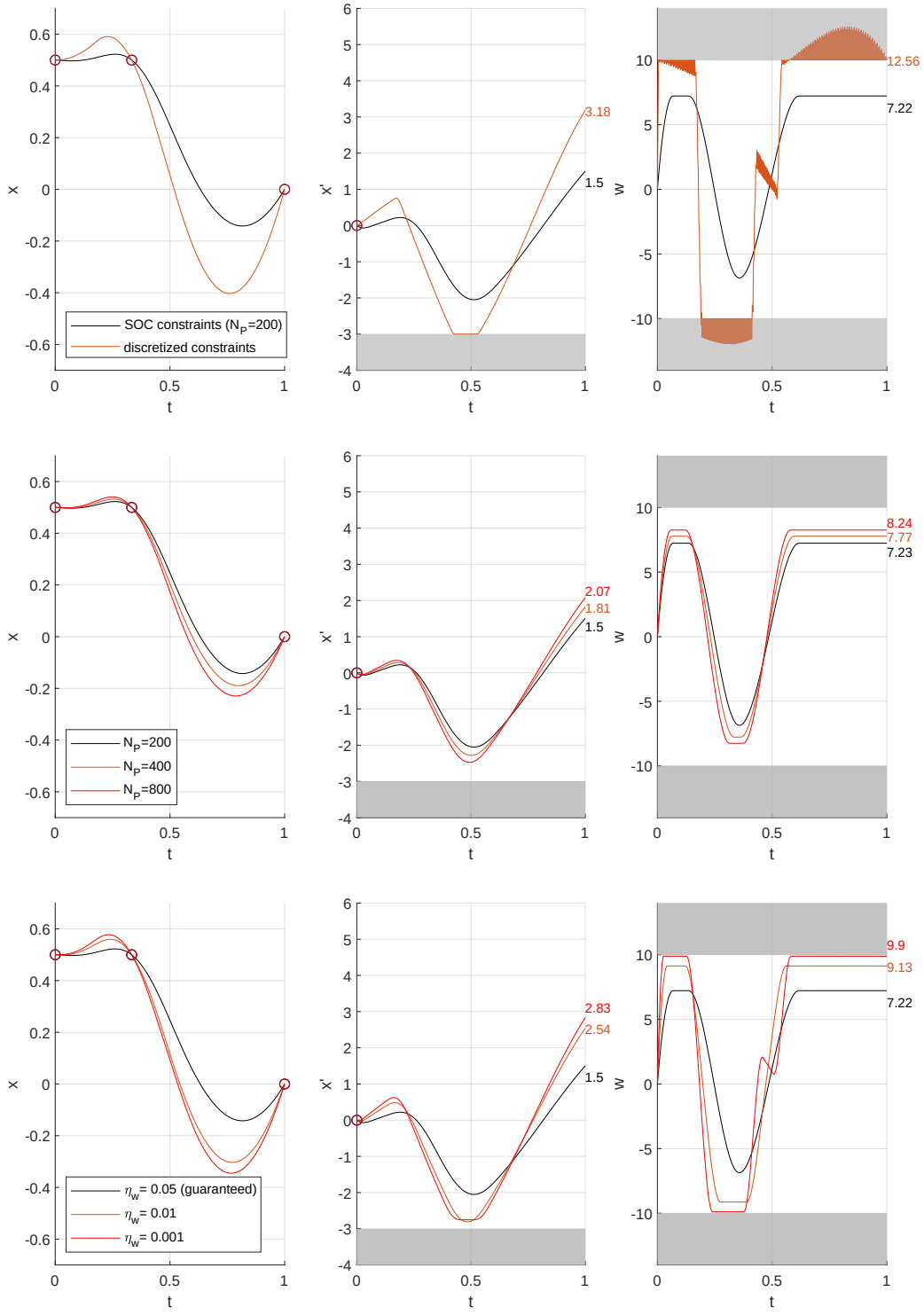


Figure 4.1 – Optimal solutions of ($\mathcal{P}_{\text{pend}}$) for varying N_P and η_w. The red circles indicate the equality-constrained points, the grayed areas the constraints over $[0, T]$. We report the values of $\dot{x}(T)$ and of the maximum of $w(\cdot)$. **Top:** Comparison of SOC constraints (guaranteed η_w) versus discretized constraints ($\eta_w = 0$) for $N_P = 200$. **Center:** Comparison of SOC constraints for varying N_P and guaranteed η_w. **Bottom:** Comparison of SOC constraints for varying η_w and $N_P = 200$.

already hinted at in the speed limit example of the introduction: the control sticks to the w-constraints at the points $(t_m)_m$ but violates them in between. Hence the optimal value of $\dot{x}(T) = 3.18$, corresponding to the optimal solution with discretized constraints (red curve on top row, middle column of Fig. 4.1), is attained only by repeated violation of the w-constraints. We then present the consequences of changing the number of grid points N_P and of lowering η_w w.r.t. to the value of its definition. The parameter $\eta_{\dot{x}}$ is kept fixed to its nominal value as it has little influence on the optimal solutions. We first investigate the effect of changing N_P while keeping the guaranteed value of η_w, defined as in (4.22). We see that the threshold applied to the w-constraints decreases, albeit slowly. Secondly, for $N_P = 200$, we divide η_w first by a factor 5, then by a factor 10. The w-constraints threshold drastically diminishes and, for $\eta_w = 0.001$, the inflexion at $t = 0.5$ appears, as a result of the constrained arc over $\dot{x}(\cdot)$. In the limit case, we would recover the trajectory with discretized constraints ($\eta_w = 0$).

We conclude from this example that incorporating SOC terms in the constraints proves to be beneficial, even for $\overrightarrow{\eta}$ chosen smaller than its nominal value. Nevertheless increasing N_P leads, as stated in Theorem 4.2, to convergence to the optimal trajectory $\bar{x}^0(\cdot)$ with affine state constraints through trajectories $\bar{x}^{\delta,\mathrm{fin}}(\cdot)$ that are always both feasible and interior. Moreover the shapes of the SOC-optimal trajectories provide intuition on the times and properties of constrained arcs.

ANNEX: EXISTENCE OF INTERIOR TRAJECTORIES

We provide here conditions ensuring the existence of interior trajectories for ($\mathcal{P}_0$). For any $\overrightarrow{\epsilon} \in \mathbb{R}^P_+$, let $\mathcal{A}_\epsilon := \{(t,\mathbf{x}) \,|\, t \in [0,T],\ \overrightarrow{\epsilon} + \mathbf{C}(t)\mathbf{x} \leqslant \mathbf{d}(t)\}$ and $\mathcal{A}_{\epsilon,t} := \{\mathbf{x} \,|\, (t,\mathbf{x}) \in \mathcal{A}_\epsilon\}$. Below, for $(t,\mathbf{x}) \in \mathcal{A}_\epsilon$, $T_{\mathcal{A}_\epsilon}(t,\mathbf{x})$ denotes the contingent cone to the set $\mathcal{A}_\epsilon$ at point $(t,\mathbf{x})$ (see e.g. Haddad (1981)).

(H1) $\mathbf{A}(\cdot)$ and $\mathbf{B}(\cdot)$ are $\mathcal{C}^0$-continuous. $\mathbf{C}(\cdot)$ and $\mathbf{d}(\cdot)$ are $\mathcal{C}^1$-continuous and $\mathbf{C}(0)x_0 < \mathbf{d}(0)$.

(H2) There exists $M_u > 0$ such that, for all $t \in [0,T]$ and $\mathbf{x} \in R\mathbb{B}_N$ satisfying $\mathbf{C}(t)\mathbf{x} \leqslant \mathbf{d}(t)$, with $R := (1 + \|x_0\|)e^{T\|\mathbf{A}(\cdot)\|_{L^\infty(0,T)} + TM_u\|\mathbf{B}(\cdot)\|_{L^\infty(0,T)}}$, there exists $\mathbf{u}_{t,x} \in M_u\mathbb{B}_M$ such that

$$\forall i \in I_{t,x} := \{i \,|\, \mathbf{c}_i(t)^\top \mathbf{x} = d_i(t)\},\ \mathbf{C}_i'(t)^\top \mathbf{x} - d_i'(t) + \mathbf{c}_i(t)^\top(\mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}_{t,x}) < 0. \tag{4.28}$$

The inward-pointing condition (H2) is a geometrical assumption on the boundary of the constraints. In particular, (H2) implies that the constraint set is non degenerate, i.e. $\mathcal{A}_{0,t}$ is the closure of its interior at all times t.

Lemma 4.4 (Existence of interior trajectories). *Under Assumptions (H1) and (H2), the following properties are satisfied*

i) there exists $\overrightarrow{\boldsymbol{\epsilon}}_0 > 0$, $M_v > 0$, $\xi > 0$, *and* $\eta > 0$ *such that for all* $\overrightarrow{\boldsymbol{\epsilon}} \leqslant \overrightarrow{\boldsymbol{\epsilon}}_0$ *and all* $(t, x) \in (\partial \mathcal{A}_\epsilon + (0, \eta \mathbb{B}_N)) \cap \mathcal{A}_\epsilon \cap ([0, T] \times (R-1)\mathbb{B}_N)$, *there exists* $u_{t,x} \in M_u \mathbb{B}_M$ *such that* $v = \mathbf{A}(t)x + \mathbf{B}(t)u_{t,x} \in M_v \mathbb{B}_N$ *and*

$$y + \delta(v + \xi \mathbb{B}_N) \subset \mathcal{A}_{\epsilon, t+\delta} \tag{4.29}$$

for all $\delta \in [0, \xi]$ *and all* $y \in x + \xi \mathbb{B}_N$ *such that* $y \in \mathcal{A}_{\epsilon,t}$. *Hence* $v \in T_{\mathcal{A}_\epsilon}(t, x)$, *and*

ii) there exists a trajectory in $\mathcal{S}$ *satisfying strictly the affine constraints, as well as the initial condition.*

The proof below is composed of two parts. The first part translates an inward-pointing condition written in normal form (4.28) to its tangent form (4.29). This allows in the second part to invoke a viability argument to derive the existence of a strictly feasible trajectory. This approach may seem remindful of a result by Soner (1986). However, unlike Soner (1986), we do not need to explicitly construct a feasible trajectory approximating a reference one, nor derive an estimate on their distance. Moreover the proof leverages the linear structure of the dynamics and of the constraints, allowing for time-varying constraints which Soner (1986) did not consider.

Proof: *i)* Define $h(t, x) := \mathbf{C}(t)x - \mathbf{d}(t)$ and $\xi(t, x, u) := \mathbf{C}'(t)x - \mathbf{d}'(t) + \mathbf{C}(t)(\mathbf{A}(t)x + \mathbf{B}(t)u)$.

Step 1: We claim that there exists $\overrightarrow{e} > 0$, $\rho > 0$, and $M_v > 0$ such that for all $t \in [0, T]$ and $x \in R\mathbb{B}_N$ satisfying $h(t, x) \leqslant 0$ with some active constraints (i.e. $I_{t,x} \neq \emptyset$), there exists $u_{t,x} \in \mathbb{R}^M$ such that $v = \mathbf{A}(t)x + \mathbf{B}(t)u \in M_v \mathbb{B}_N$ and

$$\forall\, i \in I^e_{t,x} := \{i \,|\, h_i(t, x) > -e_i\},\ \xi_i(t, x, u_{t,x}) \leqslant -\rho. \tag{4.30}$$

Let $u_{t,x}$ as in (4.28), and set

$$\rho_{t,x} := -\max_{i \in I_{t,x}} \xi_i(t, x, u_{t,x}) > 0 \qquad e_{t,x} := -\max_{i \notin I_{t,x}} h_i(t, x, u_{t,x}) > 0.$$

Since $\xi(\cdot, \cdot, u_{t,x})$ and $h(\cdot, \cdot)$ are continuous, we can find $\Delta_{t,x} > 0$ such that

$$\sup_{\substack{\delta \in [t-\Delta_{t,x}, t+\Delta_{t,x}] \cap [0,T], \\ w \in \mathbb{B}_N}} \|h(t, x) - h(t+\delta, x + \Delta_{t,x} w)\|_\infty \leqslant \frac{e_{t,x}}{2}$$

$$\sup_{\substack{\delta \in [t-\Delta_{t,x}, t+\Delta_{t,x}] \cap [0,T], \\ w \in \mathbb{B}_N,\, i \in I_{t,x}}} |\xi_i(t, x, u_{t,x}) - \xi_i(t+\delta, x + \Delta_{t,x} w, u_{t,x})| \leqslant \frac{\rho_{t,x}}{2}.$$

This implies that the index set of active constraints does not increase in size for points in the open set $\Omega_{t,x} := (t,x) + \Delta_{t,x}]-1,1[\times \mathring{\mathbb{B}}_N$, denoting by $\mathring{\mathbb{B}}_N$ the open unit ball of $\mathbb{R}^N$. Moreover, for any point in $\Omega_{t,x}$, $u_{t,x}$ satisfies our claim for the constants $\rho_{t,x}/2$ and $(e_{t,x}/2)\,1_N$. Since we are considering a compact set, we can select a finite number of $(t_j,x_j)_{j\in[\![1,J]\!]}$ such that

$$([0,T]\times R\mathbb{B}_N)\cap\mathcal{A}_0 \subset \bigcup_{j\in[\![1,J]\!]} \Omega_{t_j,x_j}.$$

To conclude, simply take [20]

$$\rho := \min_{j\in[\![1,J]\!]} \frac{\rho_{t_j,x_j}}{2}, \qquad e := \left(\min_{j\in[\![1,J]\!]} \frac{e_{t_j,x_j}}{2}\right) 1_N, \qquad M_v := R\|A(\cdot)\|_{L^\infty(0,T)} + M_u\|B(\cdot)\|_{L^\infty(0,T)}.$$

Step 2: Thanks to the uniform constants of the normal form (4.30), we derive the constants of the tangent form (4.29). Let us choose $\xi,\eta \in]0,1]$ such that for all $(t,x)\in\partial\mathcal{A}_0\cap([0,T]\times R\mathbb{B}_N)$, setting $v_{t,x} = A(t)x + B(t)u_{t,x} \in M_v\mathbb{B}_N$, for

$$e^{t,x,\delta}_{\alpha,\beta,\gamma} := (C(t+\delta)-C(t))(x+\eta\gamma+\xi\alpha)+\delta C(t+\delta)(v_{t,x}+\eta A(t)\gamma+\xi\beta)+d(t)-d(t+\delta),$$

we have

$$\frac{e}{2} > \sup_{\substack{\delta\in[t-\xi,t+\xi]\cap[0,T]\\ \alpha,\beta,\gamma\in\mathbb{B}_N}} e^{t,x,\delta}_{\alpha,\beta,\gamma} \quad , \quad \frac{-\delta\rho}{2} > \sup_{\substack{\delta\in[t-\xi,t+\xi]\cap[0,T]\\ \alpha,\beta,\gamma\in\mathbb{B}_N,\, i\in I^e_{t,x}}} (e^{t,x,\delta}_{\alpha,\beta,\gamma})_i. \tag{4.31}$$

The first inequality can be derived from the $\mathcal{C}^1$-smoothness of $C(\cdot)$ and $d(\cdot)$, and from the fact that both x, $v_{t,x}$, $A(\cdot)$ and $C(\cdot)$ are uniformly bounded. The second inequality stems from (4.30), as $\lim_{\delta\to0^+}\frac{e^{t,x,\delta}_{0,0,0}}{\delta} = \xi(t,x,u_{t,x})$. The two inequalities of (4.31) state that, for $y := x+\eta\gamma+\xi\alpha$, if $y\in\mathcal{A}_{0,t}$ then

$$y+\delta(v+\xi\mathbb{B}_N)\subset\mathcal{A}_{0,t+\delta} \text{ with } v := v_{t,x}+\eta A(t)\gamma = A(t)(x+\eta\gamma)+B(t)u_{t,x}.$$

Step 3: There just remains to extend the result of Step 2 to the perturbed constraint sets $\mathcal{A}_\epsilon$. For any $t\in[0,T]$, let w_t be an eigenvector of the largest eigenvalue μ_t of $C(t)C(t)^\top$ satisfying $\|w_t\|_\infty = \max_{i\in[\![1,P]\!]} w_{t,i} = 1$. Let $\bar\mu := \min_{t\in[0,T]}\mu_t > 0$. Take any $\overrightarrow{\epsilon}_0 > 0$ such that

$$\sup_{t\in[0,T]} \frac{2\|\overrightarrow{\epsilon}_0\|_\infty\|C(t)^\top w_t\|}{\bar\mu} \leqslant \frac{\eta}{2}.$$

20. If $\mathcal{A}_0$ is bounded, take $M_u := 1+\max_{j\in[\![1,J]\!]}\|u_{t_j,x_j}\|$.

Let $\overrightarrow{\epsilon} \leqslant \overrightarrow{\epsilon}_0$, take $(t,x) \in \partial\mathcal{A}_\epsilon \cap ([0,T] \times (R-1)\mathbb{B}_N)$. Then

$$\exists\, i \in [\![1,P]\!],\ \mathbf{c}_i(t)^\top \left(x + \frac{2\|\overrightarrow{\epsilon}_0\|_\infty \mathbf{C}(t)^\top w_t}{\bar{\mu}} \right) = d_i(t) - \overrightarrow{\epsilon}_i + \frac{2\mu_t}{\bar{\mu}} \|\overrightarrow{\epsilon}_0\|_\infty > 0.$$

Since $\eta \leqslant 1$, one can therefore find $\check{x} \in R\mathbb{B}_N \cap \partial\mathcal{A}_{0,t}$ such that $\|x - \check{x}\| = d_{\mathcal{A}_{0,t}}(x) \leqslant \eta/2$, to which the conclusions of Step 2 apply. So the set considered in (4.29) at x, taking as constants $\xi/2$ and $\eta/2$, is a subset of the one defined by $\check{x}$, ξ and η. Since only differences of $\mathbf{d}(\cdot)$ appeared in Step 2 and that $x \in \partial\mathcal{A}_{\epsilon,t}$, adding $\overrightarrow{\epsilon}$ has no effect on the computations, so it follows that (4.29) is satisfied. For $y = x$, by definition of $T_{\mathcal{A}_\epsilon}(t,x)$, we have that $(1,v) \in T_{\mathcal{A}_\epsilon}(t,x)$.

ii) Let $F(t,x) := \{(1, \mathbf{A}(t)x + \mathbf{B}(t)u) \,|\, u \in M_u\mathbb{B}_M\}$. Since $\mathbf{A}(\cdot)$ and $\mathbf{B}(\cdot)$ are $\mathcal{C}^0$-continuous, the bounded set-valued map $F(\cdot,\cdot)$ is upper semicontinuous. Let $\Omega := \mathcal{A}_\epsilon \cap ([0,T] \times (R-1)\mathring{\mathbb{B}}_N)$. The open ball $\mathring{\mathbb{B}}_N$ being open in $\mathbb{R}^N$, Ω is locally compact in $[0,T] \times \mathbb{R}^N$. In *i)* we have shown that the local viability condition is satisfied for $\mathcal{A}_\epsilon$. The intersection with an open set does not add boundary points, so, for $(t,x) \in \Omega$, $F(t,x) \cap T_\Omega(t,x) \neq \emptyset$. We may therefore apply (Haddad, 1981, Theorem 1) which provides a trajectory $x^\epsilon(\cdot)$ satisfying $x^\epsilon(0) = x_0$, $x^{\epsilon,\prime}(t) \in F(t, x^\epsilon(t))$, and $h(t, x^\epsilon(t)) + \overrightarrow{\epsilon} \leqslant 0$. Let $[0, t_1[$ be the maximal interval of existence of $x^\epsilon(\cdot)$. Since $x^{\epsilon,\prime}(\cdot)$ is measurable, by (Vinter, 1990, Theorem 2.3.13), we can find some measurable $u^\epsilon(\cdot)$ with values bounded by M_u, s.t. $x^{\epsilon,\prime}(\cdot) = \mathbf{A}(t)x + \mathbf{B}(t)u^\epsilon(t)$ a.e. in $[0,T]$. The dynamics being sublinear as $\mathbf{A}(\cdot)$, $\mathbf{B}(\cdot)$, and $u^\epsilon(\cdot)$ are bounded, $x^\epsilon(\cdot)$ can be continuously extended to $(t_1, x^\epsilon(t_1)) \in \Omega$. So $t_1 \geqslant T$, otherwise the viability condition would allow to extend $x^\epsilon(\cdot)$ beyond t_1. Since $u^\epsilon(\cdot)$ is measurable and bounded, $u^\epsilon(\cdot) \in L^2(0,T)$. Hence $x^\epsilon(\cdot) \in \mathcal{S}$ satisfies the required properties.

■

BIBLIOGRAPHY FOR CHAPTER 4

Ackermann, J. (1985). *Sampled-data control systems*. Springer Berlin Heidelberg.

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

Aubin-Frankowski, P.-C. (2021). Interpreting the dual Riccati equation through the LQ reproducing kernel [(https://arxiv.org/abs/2012.12940)]. *Comptes Rendus. Mathématique*, *359*(2), 199–204.

Aubin-Frankowski, P.-C., Petit, N., & Szabó, Z. (2020). Kernel Regression for Vehicle Trajectory Reconstruction under Speed and Inter-vehicular Distance Constraints [(https://hal-mines-paristech.archives-ouvertes.fr/hal-03021643)]. *IFAC World Congress 2020*.

Aubin-Frankowski, P.-C., & Szabó, Z. (2020a). *Handling hard affine SDP shape constraints in RKHSs* (tech. rep.) [(https://arxiv.org/abs/2101.01519)].

Aubin-Frankowski, P.-C., & Szabó, Z. (2020b). Hard Shape-Constrained Kernel Machines [(http://arxiv.org/abs/2005.12636)]. *Advances in Neural Information Processing Systems (NeurIPS)*.

Bertalan, T., Dietrich, F., Mezić, I., & Kevrekidis, I. G. (2019). On learning Hamiltonian systems from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(12), 121107-1 –121107-9.

Bonnans, J. F., & Ramirez, H. (2005). Perturbation analysis of second-order cone programming problems. *Mathematical Programming*, *104*(2-3), 205–227.

Bourdin, L., & Trélat, E. (2017). Linear–quadratic optimal sampled-data control problems: Convergence result and Riccati theory. *Automatica*, *79*, 273–281.

Burachik, R. S., Kaya, C. Y., & Majeed, S. N. (2014). A Duality Approach for Solving Control-Constrained Linear-Quadratic Optimal Control Problems. *SIAM Journal on Control and Optimization*, *52*(3), 1423–1456.

Chaplais, F., Malisani, P., & Petit, N. (2011). Design of penalty functions for optimal control of linear dynamical systems under state and input constraints. *Conference on Decision and Control (CDC)*, 6697–6704.

Chiuso, A., & Pillonetto, G. (2019). System Identification: A Machine Learning Perspective. *Annual Review of Control, Robotics, and Autonomous Systems*, *2*(1), 281–304.

Dower, P. M., McEneaney, W. M., & Cantoni, M. (2019). Game representations for state constrained continuous time linear regulator problems [(http://arxiv.org/abs/1904.05552)]. *arXiv:1904.05552*.

Fujii, K., & Kawahara, Y. (2019). Dynamic mode decomposition in vector-valued reproducing kernel Hilbert spaces for extracting dynamical structure among observables. *Neural Networks*, *117*, 94–103.

Fujioka, H., & Kano, H. (2013). Control theoretic B-spline smoothing with constraints on derivatives, 2115–2120.

Gerdts, M., & Hüpping, B. (2012). Virtual control regularization of state constrained linear quadratic optimal control problems. *Computational Optimization and Applications*, *51*(2), 867–882.

Giannakis, D., Das, S., & Slawinska, J. (2019). Reproducing kernel Hilbert space compactification of unitary evolution groups [http://arxiv.org/abs/1808.01515].

Grüne, L., & Guglielmi, R. (2018). Turnpike properties and strict dissipativity for discrete time linear quadratic optimal control problems. *SIAM Journal on Control and Optimization*, *56*(2), 1282–1302.

Haddad, G. (1981). Monotone trajectories of differential inclusions and functional differential inclusions with memory. *Israel Journal of Mathematics*, *39*(1-2), 83–100.

Hartl, R. F., Sethi, S. P., & Vickson, R. G. (1995). A Survey of the Maximum Principles for Optimal Control Problems with State Constraints. *SIAM Review*, *37*(2), 181–218.

Heckman, N. (2012). The theory and application of penalized methods or Reproducing Kernel Hilbert Spaces made easy. *Statistics Surveys*, *6*(0), 113–141.

Hermant, A. (2009). Stability analysis of optimal control problems with a second-order state constraint. *SIAM Journal on Optimization*, *20*(1), 104–129.

Kailath, T. (1971). RKHS approach to detection and estimation problems–I: Deterministic signals in Gaussian noise. *IEEE Transactions on Information Theory*, *17*(5), 530–549.

Kano, H., & Fujioka, H. (2018). B-spline trajectory planning with curvature constraint. *2018 Annual American Control Conference (ACC)*, 1963–1968.

Kojima, A., & Morari, M. (2004). LQ control for constrained continuous-time systems. *Automatica*, *40*(7), 1143–1155.

Magnus Egerstedt, C. M. (2009). *Control theoretic splines: optimal control, statistics, and path planning*. Princeton University Press.

Marco, A., Hennig, P., Schaal, S., & Trimpe, S. (2017). On the design of LQR kernels for efficient controller learning. *Conference on Decision and Control (CDC)*, 5193–5200.

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

Mercy, T., Van Loock, W., & Pipeleers, G. (2016). Real-time motion planning in the presence of moving obstacles. *European Control Conference (ECC)*, 1586–1591.

Micheli, M., & Glaunès, J. A. (2014). Matrix-valued kernels for shape deformation analysis. *Geometry, Imaging and Computing*, *1*(1), 57–139.

Parzen, E. (1970). Statistical inference on time series by RKHS methods. *Proceedings 12th Biennial Seminar, Canadian Mathematical Congress*.

Pillonetto, G., Dinuzzo, F., Chen, T., Nicolao, G. [, & Ljung, L. (2014). Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, *50*(3), 657–682.

Rosenfeld, J. A., Kamalapurkar, R., Gruss, L. F., & Johnson, T. T. (2019). Dynamic Mode Decomposition for Continuous Time Systems with the Liouville Operator [http://arxiv.org/abs/1910.03977].

Schölkopf, B., Herbrich, R., & Smola, A. J. (2001). A generalized representer theorem. *Computational Learning Theory (CoLT)*, 416–426.

Schölkopf, B., & Smola, A. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press.

Schwartz, L. (1964). Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). *Journal d'Analyse Mathématique*, *13*, 115–256.

Singh, S., Sindhwani, V., Slotine, J.-J., & Pavone, M. (2018). Learning stabilizable dynamical systems via control contraction metrics [https://arxiv.org/abs/1808.00113]. *Workshop on Algorithmic Foundations of Robotics (WAFR)*.

Soner, H. M. (1986). Optimal control with state-space constraint i. *SIAM Journal on Control and Optimization*, *24*(3), 552–561.

Sootla, A., Mauroy, A., & Ernst, D. (2018). Optimal control formulation of pulse-based control using Koopman operator. *Automatica*, *91*, 217–224.

Speyer, J. L., & Jacobson, D. H. (2010). *Primer on optimal control theory*. Society for Industrial; Applied Mathematics.

Steinke, F., & Schölkopf, B. (2008). Kernels, regularization and differential equations. *Pattern Recognition*, *41*(11), 3271–3286.

Van Loan, C. (1978). Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control*, *23*(3), 395–404.

van Keulen, T. (2020). Solution for the continuous-time infinite-horizon linear quadratic regulator subject to scalar state constraints. *IEEE Control Systems Letters*, *4*(1), 133–138.

Vinter, R. (1990). *Optimal control*. Birkhäuser, Basel.

Wahba, G. (1990). *Spline models for observational data*. SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics.

Williams, M. O., Rowley, C. W., & Kevrekidis, I. G. (2015). A kernel-based method for data-driven Koopman spectral analysis. *Journal of Computational Dynamics*, *2*(2), 247–265.

5

THE LQ REPRODUCING KERNEL AND THE RICCATI EQUATION

This chapter was published with a single author in Comptes Rendus. Mathématique, under the title *Interpreting the dual Riccati equation through the LQ reproducing kernel* (Aubin-Frankowski, 2021a).

Abstract In this chapter, we provide an interpretation of the dual differential Riccati equation of Linear-Quadratic (LQ) optimal control problems. Adopting a novel viewpoint, we show that LQ optimal control can be seen as a regression problem over the space of controlled trajectories –an idea already alluded by Luenberger (1968)– and that the latter has a very natural structure as a reproducing kernel Hilbert space (RKHS). The dual Riccati equation then describes the evolution of the values of the LQ reproducing kernel when the initial time changes. This unveils new connections between control theory and kernel methods, a field widely used in machine learning.

Résumé Dans ce chapitre, nous proposons une interprétation de l'équation différentielle de Riccati duale pour des problèmes de contrôle optimal linéaires-quadratiques (LQ). En adoptant un point de vue nouveau, nous montrons que le contrôle optimal LQ peut être considéré comme un problème de régression sur l'espace des trajectoires contrôlées – une formulation déjà évoquée par Luenberger (1968)– et que ce dernier a une structure très naturelle d'espace de Hilbert à noyau reproduisant (RKHS). L'équation de Riccati duale décrit en conséquence l'évolution des valeurs du noyau de reproduction LQ lorsque le temps initial varie. Cette étude dévoile de nouvelles connexions entre la théorie du contrôle et les méthodes à noyau, un domaine largement utilisé en apprentissage automatique.

NB: The invertibility of the Riccati matrices $\mathbf{J}(t,T)$ under our assumptions is a consequence, for instance, of the proof of Caroff and Frankowska (1993, Theorem 3.1). As the Riccati equation preserves positive semidefiniteness, when taking $\mathbf{P}_{1,T} = -\mathbf{J}_T \prec 0$ and $\mathbf{P}_{2,T} = 0$, we obtain that $\mathbf{J}(t,T) = -\mathbf{P}_1(t) \succ -\mathbf{P}_2(t) \succeq 0$. The above theorem was published without its proof in Caroff and Frankowska (1992).

5.1 INTRODUCTION

We consider the problem of finite-dimensional time-varying linear quadratic (LQ) optimal control with finite horizon and quadratic terminal cost as in

$$V(t_0, x_0) := \min_{u(\cdot)} \quad x(T)^\top J_T x(T) + \int_{t_0}^{T} [x(t)^\top Q(t) x(t) + u(t)^\top R(t) u(t)] dt \tag{1}$$

s.t.

$$x'(t) = A(t)x(t) + B(t)u(t), \text{ a.e. in } [t_0, T], \tag{5.1a}$$

$$x(t_0) = x_0, \tag{5.1b}$$

where the state $x(t) \in \mathbb{R}^N$ and the control $u(t) \in \mathbb{R}^M$. We shall henceforth assume that $J_T \succ 0$,[1] and for all $t \in [t_0, T]$, $R(t) \succcurlyeq r\mathrm{Id}_M$ with $r > 0$, as well as $A(\cdot) \in L^1([t_0, T], \mathbb{R}^{N,N})$, $B(\cdot) \in L^2([t_0, T], \mathbb{R}^{N,M})$, $Q(\cdot) \in L^1([t_0, T], \mathbb{R}^{N,N})$, and $R(\cdot) \in L^2([t_0, T], \mathbb{R}^{N,N})$. To have a finite objective, we restrict our attention to measurable controls satisfying $R(\cdot)^{1/2}u(\cdot) \in L^2([t_0, T], \mathbb{R}^N)$. Problem (5.1) is intimately related to the differential Riccati equation,[2] expressed as

$$\begin{aligned} -\partial_1 J(t, T) &= A(t)^\top J(t, T) + J(t, T)A(t) - J(t, T)B(t)R(t)^{-1}B(t)^\top J(t, T) + Q(t), \\ J(T, T) &= J_T, \end{aligned} \tag{5.2}$$

which solution $J(\cdot, T)$ satisfies $V(t_0, x_0) = x_0^\top J(t_0, T)x_0$. It is well-known (e.g. Bensoussan et al., 2007, pp. 31, 408) that under the above positivity assumptions, $J(t, T)$ is a symmetric positive definite matrix, which inverse $M(t, T) := J(t, T)^{-1}$ satisfies a dual Riccati equation

$$\begin{aligned} \partial_1 M(t, T) &= A(t)M(t, T) + M(t, T)A(t)^\top - B(t)R(t)^{-1}B(t)^\top + M(t, T)Q(t)M(t, T), \\ M(T, T) &= J_T^{-1}. \end{aligned} \tag{5.3}$$

This inverse matrix has been used as a tool to obtain a representation formula in infinite-dimensional LQ control (Barbu & Prato, 1992) but it has not received the deserved interest yet. Whereas the solution of (5.2) is equal to the Hessian of the value function $V(t_0, \cdot)$, i.e. $J(t_0, T) = \partial_{x,x}V(t_0, \cdot)$, we show (Theorem 5.2 below) that the solution of (5.3) is equal to the diagonal element of a matrix-valued reproducing kernel $K(\cdot, \cdot)$, naturally associated with (5.1). Owing to this interpretation, the dual Riccati equation (5.3) is thus no less fundamental and effectively allows to reverse the

1. Here $\succcurlyeq$ (resp. $\succ$) denotes the (strict) partial order over positive semi-definite matrices.
2. The index T in $J(\cdot, T)$ is kept as a reminder that (5.2) is defined w.r.t. a given terminal time T. We denote by ∂_1 the derivative w.r.t. the first variable.

perspective between the adjoint vector and the optimal trajectory.

We first need to bring trajectories to the fore in (5.1). In his seminal book, Luenberger (Luenberger, 1968, p255) already discussed that an optimal control problem such as (5.1) can be seen as either optimizing over the set of controls $\mathbf{u}(\cdot)$, or jointly over the set of trajectories $\mathbf{x}(\cdot)$ and controls $\mathbf{u}(\cdot)$, connected through the dynamic constraint (5.1a). Luenberger also alluded without details to a third possibility, that of optimizing directly over the controlled trajectories. We follow this last viewpoint and consequently introduce the vector space $\mathcal{S}_{[t_0,T]}$ of controlled trajectories of the linear system: There is not necessarily a unique choice of $\mathbf{u}(\cdot)$ for a given $\mathbf{x}(\cdot) \in \mathcal{S}_{[t_0,T]}$.[3] Therefore, with each $\mathbf{x}(\cdot) \in \mathcal{S}_{[t_0,T]}$, we associate the control $\mathbf{u}(\cdot)$ having minimal norm based on the pseudoinverse $\mathbf{B}(t)^{\ominus}$ of $\mathbf{B}(t)$ for the $\mathbb{R}^M$-norm $\|\cdot\|_{\mathbf{R}(t)} := \|\mathbf{R}(t)^{1/2}\cdot\|$:

$$\mathbf{u}(t) = \mathbf{B}(t)^{\ominus}[\mathbf{x}'(t) - \mathbf{A}(t)\mathbf{x}(t)] \text{ a.e. in } [t_0, T]. \tag{5.4}$$

Problem (5.1) then induces a natural inner product over $\mathcal{S}_{[t_0,T]}$. As a matter of fact, the expression

$$\langle \mathbf{x}_1(\cdot), \mathbf{x}_2(\cdot)\rangle_K := \mathbf{x}_1(T)^\top \mathbf{J}_T \mathbf{x}_2(T) + \int_{t_0}^{T} [\mathbf{x}_1(t)^\top \mathbf{Q}(t)\mathbf{x}_2(t) + \mathbf{u}_1(t)^\top \mathbf{R}(t)\mathbf{u}_2(t)]dt \tag{5.5}$$

is bilinear and symmetric over $\mathcal{S}_{[t_0,T]} \times \mathcal{S}_{[t_0,T]}$. It is positive definite over $\mathcal{S}_{[t_0,T]}$ as $\|\mathbf{x}(\cdot)\|_K^2 = 0$ implies that $\mathbf{u}(\cdot) \overset{a.e}{\equiv} 0$ and, as $\mathbf{J}_T \succ 0$, $\mathbf{x}(T) = 0$, hence $\mathbf{x}(\cdot) \equiv 0$. Therefore

$$\begin{aligned} V(t_0, \mathbf{x}_0) = \min_{\mathbf{x}(\cdot)\in\mathcal{S}_{[t_0,T]}} \quad & \|\mathbf{x}(\cdot)\|_K^2 \\ \text{s.t.} \quad & \\ \mathbf{x}(t_0) = \mathbf{x}_0. & \end{aligned} \tag{5.6}$$

In other words the value function $V(t_0, \mathbf{x}_0)$ of (5.1) coincides with the optimal value of a constrained norm minimization over $\mathcal{S}_{[t_0,T]}$. The solution of (5.6) can be made explicit as $(\mathcal{S}_{[t_0,T]}, \langle\cdot,\cdot\rangle_K)$ is not an arbitrary Hilbert space, but a vector-valued reproducing kernel Hilbert space (vRKHS).

5.2 VECTOR SPACES OF LINEARLY CONTROLLED TRAJECTORIES AS VRKHSS

Definition 5.1. Let $\mathcal{T}$ be a non-empty set. A Hilbert space $(\mathcal{H}_K(\mathcal{T}), \langle\cdot,\cdot\rangle_K)$ of $\mathbb{R}^N$-vector-valued functions defined on $\mathcal{T}$ is called a vRKHS if there exists a matrix-valued kernel $K_\mathcal{T} : \mathcal{T} \times \mathcal{T} \to \mathbb{R}^{N,N}$ such that the *reproducing property* holds: for all $t \in \mathcal{T}$, $\mathbf{p} \in \mathbb{R}^N$, $K_\mathcal{T}(\cdot, t)\mathbf{p} \in \mathcal{H}_K(\mathcal{T})$ and for all $\mathbf{f} \in \mathcal{H}_K(\mathcal{T})$, $\mathbf{p}^\top \mathbf{f}(t) = \langle \mathbf{f}, K_\mathcal{T}(\cdot, t)\mathbf{p}\rangle_K$.

3. This is the case for instance if $\mathbf{B}(t)$ is not injective for a set of times t with positive measure.

Remark: It is well-known that by Riesz's theorem, an equivalent definition of a vRKHS is that, for every $t \in \mathcal{T}$ and $\mathbf{p} \in \mathbb{R}^N$, the evaluation functional $\mathbf{f} \in \mathcal{H}_K(\mathcal{T}) \mapsto \mathbf{p}^\top \mathbf{f}(t) \in \mathbb{R}$ is continuous. There is also a one-to-one correspondence between the kernel $K_\mathcal{T}$ and the vRKHS $(\mathcal{H}_K(\mathcal{T}), \langle \cdot, \cdot \rangle_K)$ (see e.g. (Micheli & Glaunès, 2014, Theorem 2.6)). Moreover, by symmetry of the scalar product, the matrix-valued kernel has a Hermitian symmetry, i.e. $K_\mathcal{T}(s,t) = K_\mathcal{T}(t,s)^\top$ for any $s, t \in \mathcal{T}$.

Lemma 5.1. *$(\mathcal{S}_{[t_0,T]}, \langle \cdot, \cdot \rangle_K)$ is a vRKHS over $[t_0, T]$ with a reproducing kernel $K_{[t_0,T]}$ which we call the* LQ kernel.

Proof of Lemma 5.1: The proof is identical to the one of Lemma 1 in (Aubin-Frankowski, 2021) where $\mathcal{S}_{[t_0,T]}$ was equipped with the following inner product

$$\langle \mathbf{x}_1(\cdot), \mathbf{x}_2(\cdot) \rangle_{K,\text{init}} := \mathbf{x}_1(t_0)^\top \mathbf{x}_2(t_0) + \int_{t_0}^{T} [\mathbf{x}_1(t)^\top \mathbf{Q}(t) \mathbf{x}_2(t) + \mathbf{u}_1(t)^\top \mathbf{R}(t) \mathbf{u}_2(t)] dt. \quad (5.7)$$

■

Owing to Lemma 5.1, we can look for a "representer theorem", i.e. a necessary condition to ensure that the solutions of an optimization problem like (5.6) live in a finite dimensional subspace of $\mathcal{S}_{[t_0,T]}$ and consequently enjoy a finite representation.

Theorem 5.1. *(Aubin-Frankowski, 2021) Let $(\mathcal{H}_K(\mathcal{T}), \langle \cdot, \cdot \rangle_K)$ be a vRKHS defined on a set $\mathcal{T}$. For a given $I \in \mathbb{N}$, let $\{t_i\}_{i \in [\![1,I]\!]} \subset \mathcal{T}$. Consider the following optimization problem with "loss" function $L : \mathbb{R}^I \to \mathbb{R} \cup \{+\infty\}$, strictly increasing "regularizer" function $\Omega : \mathbb{R}_+ \to \mathbb{R}$, and vectors $\{\mathbf{c}_i\}_{i \in [\![1,I]\!]} \subset \mathbb{R}^N$*

$$\min_{\mathbf{f} \in \mathcal{H}_K(\mathcal{T})} \quad L\left(\mathbf{c}_1^\top \mathbf{f}(t_1), \dots, \mathbf{c}_I^\top \mathbf{f}(t_I)\right) + \Omega\left(\|\mathbf{f}\|_K\right).$$

Then, for any minimizer $\bar{\mathbf{f}}$, there exists $\{\mathbf{p}_i\}_{i \in [\![1,I]\!]} \subset \mathbb{R}^N$ such that $\bar{\mathbf{f}} = \sum_{i=1}^I K_\mathcal{T}(\cdot, t_i)\mathbf{p}_i$ with $\mathbf{p}_i = \alpha_i \mathbf{c}_i$ for some $\alpha_i \in \mathbb{R}$.

Taking $L(\mathbf{e}_1^\top \mathbf{x}(t_0), \dots, \mathbf{e}_N^\top \mathbf{x}(t_0)) := \chi_{x_0}(\mathbf{x}(t_0))$ and $\Omega(y) = y^2$, with $\mathbf{e}_i$ the i-th basis vector of $\mathbb{R}^N$, χ_{x_0} the indicator function of $\mathbf{x}_0$, we apply Theorem 5.1 to (5.6). Since $\|\cdot\|_K^2$ is strongly convex and there exists $\mathbf{x}(\cdot) \in \mathcal{S}_{[t_0,T]}$ satisfying $\mathbf{x}(t_0) = \mathbf{x}_0$, the solution of (5.6) is unique and can be written as $\bar{\mathbf{x}}(t) = K_{[t_0,T]}(t, t_0)\mathbf{p}_0$, with $\mathbf{p}_0 = K_{[t_0,T]}(t_0, t_0)^\ominus \mathbf{x}_0 \in \mathbb{R}^N$, where $K_{[t_0,T]}(t_0, t_0)^\ominus$ is the pseudoinverse of $K_{[t_0,T]}(t_0, t_0)$ for the $\mathbb{R}^N$-seminorm $\|K_{[t_0,T]}(t_0, t_0)^{1/2} \cdot\|$. Thus, owing to the reproducing property,

$$\begin{aligned} V(t_0, \mathbf{x}_0) = \|\bar{\mathbf{x}}(\cdot)\|_K^2 &= \langle K_{[t_0,T]}(\cdot, t_0)\mathbf{p}_0, K_{[t_0,T]}(\cdot, t_0)\mathbf{p}_0 \rangle_K = \mathbf{p}_0^\top K_{[t_0,T]}(t_0, t_0)\mathbf{p}_0 \quad (5.8) \\ &= \mathbf{x}_0^\top K_{[t_0,T]}(t_0, t_0)^\ominus \mathbf{x}_0 = \mathbf{p}_0^\top \mathbf{x}_0. \end{aligned}$$

So we conjecture that $K_{[t_0,T]}(t_0, t_0)^\ominus = \mathbf{J}(t_0, T)$. We actually have a stronger result:

Theorem 5.2. *Let* $K_d : t_0 \in]-\infty, T] \mapsto K_{[t_0,T]}(t_0, t_0)$. *Then* $K_d(t_0) = \mathbf{J}(t_0, T)^{-1}$.

The proof of Theorem 5.2 (in Section 5.3 below) boils down to identifying the reproducing kernel of $(\mathcal{S}_{[t_0,T]}, \langle\cdot,\cdot\rangle_K)$. Informally, the inverse relation comes from inverting the graph of the $(\mathbf{x}, \mathbf{p})$-relation. As a matter of fact, consider the solution $\mathbf{p}(t)$ of the adjoint equation

$$\mathbf{p}'(t) = -\mathbf{A}(t)^\top \mathbf{p}(t) + \mathbf{Q}(t)\bar{\mathbf{x}}(t) \qquad \mathbf{p}(T) = -\mathbf{J}_T\,\bar{\mathbf{x}}(T). \tag{5.9}$$

Then we have $\mathbf{p}(t) = -\mathbf{J}(t, T)\bar{\mathbf{x}}(t)$. In other words, the solution $\mathbf{J}(\cdot, T)$ of the differential Riccati equation maps the optimal trajectory $\bar{\mathbf{x}}(\cdot)$ to its adjoint vector $\mathbf{p}(\cdot)$. On the contrary, since $\bar{\mathbf{x}}(t) = K_{[t_0,T]}(t, t_0)\mathbf{p}_0$, the kernel $K_{[t_0,T]}(\cdot, t_0)$ maps an initial covector $\mathbf{p}_0 \in \mathbb{R}^n$ to the optimal trajectory $\bar{\mathbf{x}}(\cdot)$. This effectively inverts the graph of the relation between $\bar{\mathbf{x}}(\cdot)$ and $\mathbf{p}(\cdot)$. The inversion performed is related to yet another change of perspective, from an online and differential approach to an offline and integral one.

Through Pontryagine's Maximum Principle (PMP), it is well known that the optimal control $\bar{\mathbf{u}}(\cdot)$ satisfies $\bar{\mathbf{u}}(t) = \mathbf{R}(t)^{-1}\mathbf{B}(t)^\top \mathbf{p}(t) = -\mathbf{R}(t)^{-1}\mathbf{B}(t)^\top \mathbf{J}(t, T)\bar{\mathbf{x}}(t) =: \mathbf{G}(t)\bar{\mathbf{x}}(t)$. Hence, based on $\mathbf{J}(t, T)$, one has a closed feedback loop, with gain matrix $\mathbf{G}(t)$, and knows the control to apply based only on the present time and state. However the optimal trajectory $\bar{\mathbf{x}}(\cdot)$ is not encoded as simply as in the kernel formula $\bar{\mathbf{x}}(t) = K_{[t_0,T]}(t, t_0)\mathbf{p}_0$. It has to be derived through numerical approximations of the dynamics (5.1a). Conversely, the kernel $K_{[t_0,T]}$ performs the integration of the Hamiltonian system (5.1a)-(5.9) and sparsely encodes $\bar{\mathbf{x}}(\cdot)$ over $[t_0, T]$ by $\mathbf{p}_0$. This sparsity partly stems from the smaller number of constraints in (5.6) w.r.t. (5.1) since the dynamics (5.1a) were incorporated in the definition of $\mathcal{S}_{[t_0,T]}$. Unlike in the PMP, the adjoint vector $\mathbf{p}(t)$ disappears in the kernel perspective and only the initial condition (or some intermediate rendezvous points) induce a covector $\mathbf{p}_i$.

More generally, for a given interval $[t_0, T]$, Theorem 5.1 states that to encode the optimal trajectories one needs at most as many covectors $\mathbf{p}_i$ as there are points t_i where the trajectory is evaluated in the optimization problem. It is a classical property of "kernel machines", frequently leveraged in classification tasks (e.g. SVMs in Schölkopf and Smola, 2002). This result was exploited in (Aubin-Frankowski, 2021) to tackle affine state constraints. From the PMP perspective, it resulted in focusing only on the measures supported on the constraint boundary. Unlike the adjoint vector $\mathbf{p}(t)$ associated with the equality constraint (5.1a), which never vanishes except for abnormal trajectories, the covectors corresponding to inequality constraints are null whenever the constraint is not active. This led to extremely sparse encoding of the optimal trajectory

by specifying only the active covectors on the $[t_0, T]$ time interval.[4] Offline computation of the kernel is indeed well suited for path-planning problems. The kernel formalism however conflicts with the online perspective since varying t_0 changes the domain of $K_{[t_0,T]}$. As the correspondence between the kernel $K_{\mathcal{T}}$ and the vRKHS $(\mathcal{H}_K(\mathcal{T}), \langle\cdot,\cdot\rangle_K)$ is one-to-one (e.g. (Micheli & Glaunès, 2014, Theorem 2.6)), varying $\mathcal{T} = [t_0, T]$ or modifying the inner product changes the kernel. In general, restricting the domain leads to complicated relations between a vRKHS and its kernel Saitoh and Sawano, 2016, pp.78-80. In our case, the dual Riccati equation (5.3) precisely describes how the values of the LQ kernel change when varying t_0.

5.3 PROOF OF THEOREM 5.2

The proof corresponds to the identification of the reproducing kernel of $(\mathcal{S}_{[t_0,T]}, \langle\cdot,\cdot\rangle_K)$. Since we shall proceed with fixed initial time t_0, we drop the corresponding index and set $K(\cdot,\cdot) = K_{[t_0,T]}(\cdot,\cdot)$. By existence and unicity of the reproducing kernel, we just have to exhibit a function $K(\cdot,\cdot)$ which satisfies the requirements of Definition 5.1.

Let us denote by $\mathbf{\Phi_A}(t,s) \in \mathbb{R}^{N,N}$ the state-transition matrix of $z'(\tau) = \mathbf{A}(\tau)z(\tau)$, defined from s to t, i.e. $z(t) = \mathbf{\Phi_A}(t,s)z(s)$. The key property used throughout this section is the variation of constants, a.k.a. Duhamel's principle, stating that for any absolutely continuous $\mathbf{x}(\cdot)$ such that $\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$ a.e., we have for any $\sigma, t \in [t_0, T]$

$$\mathbf{x}(t) = \mathbf{\Phi_A}(t,\sigma)\mathbf{x}(\sigma) + \int_\sigma^t \mathbf{\Phi_A}(t,\tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau. \tag{5.10}$$

Setting $\partial_1 K(s,t) : \mathbf{p} \mapsto \frac{d}{ds}(K(s,t)\mathbf{p})$, let us define formally $\mathbf{U}(s,t) := \mathbf{B}(s)^{\ominus}[\partial_1 K(s,t) - \mathbf{A}(s)K(s,t)]$. The reproducing property for K then writes as follows, for all $t \in [t_0, T]$, $\mathbf{p} \in \mathbb{R}^N$, $\mathbf{x}(\cdot) \in \mathcal{S}_{[t_0,T]}$,

$$\mathbf{p}^\top \mathbf{x}(t) = (K(T,t)\mathbf{p})^\top \mathbf{J}_T \mathbf{x}(T) + \int_{t_0}^T (K(s,t)\mathbf{p})^\top \mathbf{Q}(s)\mathbf{x}(s)ds + \int_{t_0}^T (\mathbf{U}(s,t)\mathbf{p}))^\top \mathbf{R}(s)\mathbf{u}(s)ds. \tag{5.11}$$

4. In (Aubin-Frankowski, 2021), we considered the inner product (5.7) rather than (5.5) which assumes a terminal quadratic cost. The choice of (5.7) was more appropriate for fixed initial time and general lower semicontinuous convex terminal costs, alongside affine state constraints.

By the Hermitian symmetry of K and the variation of constants (5.10) written for $\sigma = T$, we can rewrite (5.11) as, for all $t \in [t_0, T]$, $x(\cdot) \in \mathcal{S}_{[t_0,T]}$,

$$x(t) = K(t,T)J_T x(T) + \int_{t_0}^{T} K(t,s)Q(s)\left(\Phi_A(s,T)x(T) + \int_T^s \Phi_A(s,\tau)B(\tau)u(\tau)d\tau\right)ds + \int_{t_0}^{T} U(s,t)^\top R(s)u(s)ds.$$

Regrouping terms,

$$x(t) = \left(K(t,T)J_T + \int_{t_0}^{T} K(t,s)Q(s)\Phi_A(s,T)ds\right)x(T) + \int_{t_0}^{T} K(t,s)Q(s)\int_T^s \Phi_A(s,\tau)B(\tau)u(\tau)d\tau ds + \int_{t_0}^{T} U(s,t)^\top R(s)u(s)ds$$

Setting $\tilde{K}(t,s) := \int_{t_0}^{s} K(t,\tau)Q(\tau)\Phi_A(\tau,s)d\tau$, and applying Fubini's theorem, we get

$$x(t) = \left(K(t,T)J_T + \tilde{K}(t,T)\right)x(T) + \int_{t_0}^{T}\left[U(s,t)^\top R(s) - \int_{t_0}^{s} K(t,\tau)Q(\tau)\Phi_A(\tau,s)B(s)d\tau\right]u(s)ds.$$

Identifying with (5.10) for $\sigma = T$, we derive that

$$\begin{aligned} K(t,T)J_T &= \Phi_A(t,T) - \tilde{K}(t,T), \\ U(s,t)^\top R(s) &= \begin{cases} (-\Phi_A(t,s) + \tilde{K}(t,s))B(s) & \forall s \geqslant t, \\ \tilde{K}(t,s)B(s) & \forall s < t. \end{cases} \end{aligned} \tag{5.12}$$

Let us introduce formally an adjoint equation defined for a variable $\Pi(s,t) \in \mathbb{R}^{N,N}$. For any given $t \in [t_0, T]$,

$$\partial_1\Pi(s,t) = -A(s)^\top\Pi(s,t) + Q(s)K(s,t) \qquad \Pi(t_0,t) = -\mathrm{Id}_N.$$

This is the matrix version of (5.9) but with an initial rather than terminal condition. Again, applying the variation of constants (5.10) with $\sigma = t_0$ to $\Pi(s,t)$, taking the transpose and owing to the symmetries of $K(\cdot,\cdot)$ and $\Phi(\cdot,\cdot)$, we derive that

$$\Pi(s,t) = \Phi_{(-A^\top)}(s,t_0)\Pi(t_0,t) + \int_{t_0}^{s} \Phi_{(-A^\top)}(s,\tau)Q(\tau)K(\tau,t)d\tau$$

$$\Pi(s,t)^\top = -\Phi_A(t_0,s) + \int_{t_0}^{s} K(t,\tau)Q(\tau)\Phi_A(\tau,s)d\tau = -\Phi_A(t_0,s) + \tilde{K}(t,s).$$

Since $\partial_1 K(s,t) = \mathbf{A}(s)K(s,t) + \mathbf{B}(s)\mathbf{U}(s,t)$, by (5.12), for any given time $t \in [t_0, T]$, we have two coupled differential equations for $K(\cdot,t)$ and $\mathbf{\Pi}(\cdot,t)$

$$\begin{aligned}
\partial_1 K(s,t) &= \mathbf{A}(s)K(s,t) + \mathbf{B}(s)\mathbf{R}(s)^{-1}\mathbf{B}(s)^\top \begin{cases} \mathbf{\Pi}(s,t) + \mathbf{\Phi}_{\mathbf{A}}(t_0,s)^\top - \mathbf{\Phi}_{\mathbf{A}}(t,s)^\top \ \forall s \geqslant t, \\ \mathbf{\Pi}(s,t) + \mathbf{\Phi}_{\mathbf{A}}(t_0,s)^\top \quad \forall s < t. \end{cases} \\
\partial_1 \mathbf{\Pi}(s,t) &= -\mathbf{A}(s)^\top \mathbf{\Pi}(s,t) + \mathbf{Q}(s)K(s,t), \\
\mathbf{\Pi}(t_0,t) &= -\mathrm{Id}_N \qquad K(t,T)\mathbf{J}_T = -\mathbf{\Pi}(T,t)^\top + \mathbf{\Phi}_{\mathbf{A}}(t,T) - \mathbf{\Phi}_{\mathbf{A}}(t_0,T).
\end{aligned} \tag{5.13}$$

Equations (5.13) seem quite intricate but they become simpler for $t = t_0$, and as seen in (5.8), t_0 is actually the only time that interests us to solve the LQ optimal control problem (5.6). For $t = t_0$, the equations boil down to

$$\begin{aligned}
\partial_1 K(s,t_0) &= \mathbf{A}(s)K(s,t_0) + \mathbf{B}(s)\mathbf{R}(s)^{-1}\mathbf{B}(s)^\top \mathbf{\Pi}(s,t_0) \\
\partial_1 \mathbf{\Pi}(s,t_0) &= -\mathbf{A}(s)^\top \mathbf{\Pi}(s,t_0) + \mathbf{Q}(s)K(s,t_0) \\
\mathbf{\Pi}(t_0,t_0) &= -\mathrm{Id}_N \quad ; \quad \mathbf{\Pi}(T,t_0) = -\mathbf{J}_T K(T,t_0).
\end{aligned} \tag{5.14}$$

Let us solve (5.14) by variation of the constant $\mathbf{J}_T$. We thus look for a function $\mathbf{J}(\cdot,T)$, which we will prove solves (5.2), such that $\mathbf{J}(T,T) = \mathbf{J}_T$ and $\mathbf{\Pi}(s,t_0) = -\mathbf{J}(s,T)K(s,t_0)$. We take the derivative in s of the latter expression to obtain

$$\begin{aligned}
-\mathbf{A}(s)^\top \mathbf{\Pi}(s,t_0) + \mathbf{Q}(s)K(s,t_0) = -\mathbf{J}(s,T)\Big(\mathbf{A}(s)K(s,t_0) + \mathbf{B}(s)\mathbf{R}(s)^{-1}\mathbf{B}(s)^\top \mathbf{\Pi}(s,t_0)\Big) \\
- (\partial_1 \mathbf{J}(s,T))K(s,t_0)
\end{aligned}$$

Therefore, applying to the right of the equations any pseudo-inverse of $K(s,t_0)$,

$$\begin{aligned}
0 = \Big[&\mathbf{A}(s)^\top \mathbf{J}(s,T)^\top + \mathbf{J}(s,T)\mathbf{A}(s) \\
&-\mathbf{J}(s,T)\mathbf{B}(s)\mathbf{R}(s)^{-1}\mathbf{B}(s)^\top \mathbf{J}(s,T) + \mathbf{Q}(s) + \partial_1 \mathbf{J}(s,T)\Big] \operatorname{proj}_{\operatorname{Im} K(s,t_0)}.
\end{aligned}$$

So it suffices that $\mathbf{J}(\cdot,T)$ solves the differential Riccati equation (5.2) and that, by symmetry of K,

$$\mathbf{\Pi}(t_0,t_0) = -\mathrm{Id}_N = -\mathbf{J}(t_0,T)K(t_0,t_0) = -K(t_0,t_0)\mathbf{J}(t_0,T).$$

Consequently $K(t_0,t_0) = \mathbf{J}(t_0,T)^{-1}$ which formalizes our intuition of the inverse relation between $\mathbf{J}$ and K. Now let us vary the initial time and consider the function $K_d : t_0 \mapsto K_{[t_0,T]}(t_0,t_0)$. Taking the derivative w.r.t. t_0 of $\mathbf{\Pi}(t_0,t_0) = -\mathrm{Id}_N = -\mathbf{J}(t_0,T)K_{[t_0,T]}(t_0,t_0)$, we get

$$0 = (\partial_1 \mathbf{J}(t_0,T)K_d(t_0) + \mathbf{J}(t_0,T)(\partial_1 K_d(t_0)),$$

applying $K_d(t_0)$ to the left of the equation and using that $K_d(t_0) = \mathbf{J}(t_0, T)^{-1}$, we obtain

$$0 = -K_d(t_0)\mathbf{A}(t_0)^\top - \mathbf{A}(t_0)K_d(t_0) + \mathbf{B}(t_0)\mathbf{R}(t_0)^{-1}\mathbf{B}(t_0)^\top - K_d(t_0)\mathbf{Q}(t_0)K_d(t_0) + \partial_1 K_d(t_0).$$

This concludes our proof as $K_d(\cdot)$ solves the dual matrix Riccati equation (5.3) which has a unique solution.

BIBLIOGRAPHY FOR CHAPTER 5

Aubin-Frankowski, P.-C. (2021). Linearly-constrained Linear Quadratic Regulator from the viewpoint of kernel methods [(to appear, https://arxiv.org/abs/2011.02196)]. *SIAM Journal on Control and Optimization*.

Barbu, V., & Prato, G. D. (1992). A representation formula for the solutions to operator Riccati equation. *Differential Integral Equations*, *5*(4), 821–829.

Bensoussan, A., Prato, G. D., Delfour, M. C., & Mitter, S. K. (2007). *Representation and control of infinite dimensional systems*. Birkhäuser Boston.

Caroff, N., & Frankowska, H. (1992). Optimality and characteristics of Hamilton-Jacobi-Bellman equations. *Optimization, optimal control and partial differential equations: first franco-romanian conference, Iasi* (pp. 169–180). Birkhäuser Basel.

Caroff, N., & Frankowska, H. (1993). *Optimality and characteristics of Hamilton-Jacobi-Bellman equations* (IIASA Working Paper) [http://pure.iiasa.ac.at/id/eprint/3761/]. WP-93-053.

Luenberger, D. (1968). *Optimization by vector space methods*. Wiley.

Micheli, M., & Glaunès, J. A. (2014). Matrix-valued kernels for shape deformation analysis. *Geometry, Imaging and Computing*, *1*(1), 57–139.

Saitoh, S., & Sawano, Y. (2016). *Theory of reproducing kernels and applications*. Springer Singapore.

Schölkopf, B., & Smola, J., Alexander. (2002). *Learning with kernels : support vector machines, regularization, optimization, and beyond*. MIT Press.

6

LIPSCHITZ MINIMUM TIME FOR DIFFERENTIAL INCLUSIONS WITH STATE CONSTRAINTS

This chapter was published with a single author in Systems & Control Letters, under the title *Lipschitz regularity of the minimum time function of differential inclusions with state constraints* (Aubin-Frankowski, 2020).

Abstract For control systems, the local regularity of the minimum time function τ_{min} in the absence of state constraints has been extensively studied and related both to inward-pointing conditions and to small-time controllability in the neighborhood of a closed target C. In the presence of state constraints, assessing this regularity is crucial to ensure the existence of solutions when perturbing the initial condition. In this chapter, we prove, without imposing the inclusion $C \subset \operatorname{Int} K$, that, for differential inclusions with closed state constraints K and under general assumptions, τ_{min} is locally Lipschitz continuous on its domain which is open in K. We discuss as well extensions to nonautonomous systems and to point targets.

Résumé Pour les systèmes contrôlés, la régularité locale de la fonction temps minimum τ_{min} en l'absence de contraintes d'état a été largement étudiée et liée à la fois aux conditions rentrantes au bord (*inward-pointing*) et à la contrôlabilité en temps court dans le voisinage d'une cible fermée C. En présence de contraintes d'état, cette régularité est cruciale pour garantir l'existence de solutions en cas de perturbation de la condition initiale. Dans ce chapitre, sans imposer l'inclusion $C \subset \operatorname{Int} K$, nous prouvons, pour des inclusions différentielles avec des contraintes d'état fermées K et sous des hypothèses générales, que τ_{min} est localement Lipschitz sur son domaine, qui est ouvert dans K. Nous discutons également des extensions aux systèmes non autonomes et aux cibles ponctuelles.

6.1 INTRODUCTION

Studying the regularity of the minimum time function finds its motivation in reachability problems. Let K and C be two closed subsets of $\mathbb{R}^n$ and consider a control system with initial condition $x_0 \in K$

$$\begin{cases} x'(t) = f(x(t), u(t)), & u(t) \in U, \\ x(0) = x_0, \end{cases} \tag{6.1}$$

where U is a compact subset of $\mathbb{R}^m$, the control $u(\cdot)$ is a measurable function and $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is sufficiently smooth. The state-constrained time optimal control problem consists in finding the minimum time $\tau_{min}(x_0)$ to reach C along solutions of (6.1) staying in K. Assessing the regularity of τ_{min} allows to answer several questions. For instance, let a time-optimal state-constrained solution $x_{ref}(\cdot)$ of system (6.1) be given. Take as initial condition a point x_1 in a neighborhood of the trajectory set $x_{ref}([0, \tau_{min}(x_0)])$. Under what conditions can x_1 be steered to the target set C while respecting the state constraints K? How long would it take compared to $\tau_{min}(x_0)$?

More generally, consider the autonomous differential inclusion [1] with initial condition $x_0 \in K$:

$$x'(t) \in F(x(t)) \qquad x(0) = x_0 \tag{6.2}$$

where $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ is a set-valued map taking closed, nonempty values. Below, an F-trajectory $x(\cdot)$ on a time interval $[0, T]$ designates an absolutely continuous function satisfying $x'(t) \in F(x(t))$ a.e. on $[0, T]$. A trajectory is called "feasible" if $x([0, T]) \subset K$.

The capture basin $\mathrm{Capt}_F(K, C)$ is the set of all points $x_0 \in K$ such that there exists $T \geqslant 0$ and a feasible F-trajectory starting from x_0 and reaching the target set C at time T. For a given $x_0 \in K$, we denote the infimum of such T by $\tau_{min}(x_0)$. By convention, $\tau_{min}(x_0) = +\infty$ if $x_0 \notin \mathrm{Capt}_F(K, C)$. By analogy with the control system (6.1), this defines the minimum time extended function $\tau_{min} : K \to [0, +\infty]$ associated with the target set C, dynamics F and state-constraints K. The Lipschitzianity of the extended function $\tau_{min}(\cdot)$ thus depends on the "Lipschitz regularity" of both F, K and C.

In order to exhibit Lipschitz dependence of solutions on initial conditions (through the renowned Filippov's theorem), it is classical to suppose the (local) Lipschitzianity of F. On the other hand, the local Lipschitz continuity of $\tau_{min}(\cdot)$ on its domain

1. We shall consider nonautonomous systems in Section 6.3.1, when F is also locally Lipschitz in the time variable.

$\mathrm{Capt}_F(K, C)$, in the case without state constraints, has been related to strict inward-pointing conditions on the boundary ∂C of the target since the 70s (see e.g. Cannarsa and Sinestrari, 2004, Chapter 8 for a modern presentation and the bibliography therein). More recently, it has been shown in Bettiol et al., 2012 that strict inward-pointing conditions on the boundary ∂K ensure L^∞-distance estimates between arbitrary F-trajectories and the set of feasible ones. These three ingredients allow us to prove in this paper the openness of $\mathrm{Capt}_F(K, C)$ in K and the local Lipschitz continuity of $\tau_{min}(\cdot)$.

Inquiries on the regularity of the minimum time function using constraint qualifications (albeit without state-constraints) go back at least to the early 70s. As a matter of fact, the latter regularity can be related to a controllability property in the vicinity of the target. For control systems, (local) Lipschitz continuity was already obtained for C^2-regular target sets in a general setting of differential games in Friedman, 1970, Theorem 5 and in a neighborhood of point targets in Petrov, 1970, Theorem 4.1. Through viscosity solutions theory, Lipschitz continuity on $\mathrm{Capt}_F(\mathbb{R}^n, C)$ was then shown for compact piecewise-C^2 targets in Bardi and Falcone, 1990, Theorem 5.4 and in a neighborhood of general closed target sets in Soravia, 1993, Corollary 3.7. Based on nonsmooth analysis and moving to differential inclusions, Veliov, 1997, Theorem 3.1 showed the Lipschitz continuity of τ_{min} on $\mathrm{Capt}_F(\mathbb{R}^n, C)$ for nonautonomous convex-valued F (measurable in time), while Wolenski and Zhuang, 1998, Theorem 6.1 revisited the regularity in a neighborhood of a closed target set C. Finally, for state-constrained nonautonomous control systems (Lipschitz in time), the local Lipschitz continuity of τ_{min} was shown in Cannarsa and Castelpietra, 2008, Theorem 3.8 under assumptions similar to ours, although more stringent as they bore on $f(x, U)$ rather than its convex hull $\mathrm{co}\,(f(x, U))$. We also discard the strong assumption of Cannarsa and Castelpietra, 2008 of having C interior to K. Though we only consider compact-valued F, we have to mention that the minimum time problem has also been studied for control-affine systems where $U = \mathbb{R}^m$ as in Motta and Sartori, 2003 and references therein.

In this article on state-constrained differential inclusions, under general assumptions on F, K and C, and a convexified version of the inward-pointing conditions, we prove the local Lipschitz continuity of τ_{min}, that had been shown for control systems without state constraints. Such a property is a first step in studying nonlinear controllability with nonsmooth state constraints. Furthermore our results encompass those of Cannarsa and Castelpietra, 2008 on nonautonomous control systems, as presented in Section 6.3.1. In Section 6.3.3, we show as well that, for point targets that are interior to the constraints, the classical small-time controllability condition is sufficient to retrieve the local Lipschitzianity of τ_{min} on its domain $\mathrm{Capt}_F(K, C)$.

6.2 MAIN RESULTS

Notation: We denote by $\mathbb{B}$ the closed unit ball in $\mathbb{R}^n$, by $\mathbb{S}^{n-1}$ the unit sphere in $\mathbb{R}^n$ and by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ the Euclidean norm and scalar product. We denote by $\mathbb{R}_+$ the set of nonnegative real numbers. We write $\Pi_K(x)$ for the (possibly set-valued) projection of a point x into K. The function $d_K(\cdot)$ designates the distance to K. The set $\operatorname{Int} K$ stands for the interior of K and the set ∂K for its boundary. We denote by $T_K(x)$ (resp. $N_K(x)$) the Clarke tangent (resp. normal) cone to the subset K at point x. We use the notation $\operatorname{co} F$ for the set-valued map that maps x to the convex hull of $F(x)$.

Assumption 6.1. *General assumptions*

$$\begin{cases} F \text{ takes closed, nonempty values on } \mathbb{R}^n, \\ C \text{ and } K \text{ are two closed nonempty subsets of } \mathbb{R}^n. \end{cases}$$

Assumption 6.2. *Sublinear growth and local Lipschitz continuity of* F

$$\exists\, A \geqslant 0,\ \forall\, x \in \mathbb{R}^n,\ F(x) \subset A(1+\|x\|)\mathbb{B}$$

$$\forall\, R > 0,\ \exists\, k_F \geqslant 0,\ \forall\, x, y \in R\mathbb{B},\ F(y) \subset F(x) + k_F\|x-y\|\mathbb{B}$$

Assumption 6.3. *Strict inward-pointing condition on* ∂K

$$\forall\, x \in \partial K,\ \operatorname{co} F(x) \cap \operatorname{Int} T_K(x) \neq \emptyset$$

Assumption 6.4. *Strict inward-pointing condition on* $\partial C \cap K$

$$\forall\, x \in \partial C \cap K,\ \operatorname{co} F(x) \cap \operatorname{Int} T_K(x) \cap \operatorname{Int} T_C(x) \neq \emptyset$$

Remark 6.1. It stems directly from Rockafellar, 1979, Theorem 2 that for any $x \in C \cap K$ such that $\operatorname{Int} T_K(x) \cap \operatorname{Int} T_C(x) \neq \emptyset$, we have

$$\operatorname{Int} T_{C\cap K}(x) = \operatorname{Int} T_K(x) \cap \operatorname{Int} T_C(x).$$

Remark 6.2. For any closed subset K of $\mathbb{R}^n$, at a given $x \in \partial K$, for $v \in \mathbb{R}^n$ and a fixed $\epsilon > 0$, we have:

$$(v + \epsilon\mathbb{B}) \subset T_K(x) \Leftrightarrow \max_{n\in N_K(x)} \left\langle v, \frac{n}{\|n\|} \right\rangle \leqslant -\epsilon \tag{6.3}$$

The existence at $x \in \partial K$ of such $v \in \operatorname{co} F(x)$ and ϵ is implied by Assumption 6.3. Relation (6.3) allows us to juggle the two translations of strict inward-pointing conditions (based either on the normal cone or on the Clarke tangent cone to the sets). As a matter of fact, while the tangent cone is suitable to build trajectories staying in a set, the normal cone is easier to use when designing trajectories outside a set. In our case, we have both to stay in K and to reach C, leading us to use both perspectives.

Remark 6.3. Assumptions 6.3 and 6.4 actually have deeper implications on the regularity of the sets K and C. If for every $x \in \partial K$, $\operatorname{Int} T_K(x) \neq \emptyset$ (i.e. K is *wedged*), then ∂K is epi-Lipschitzian (it can be represented locally as the epigraph of a Lipschitz function after a nonsingular linear transform, see Rockafellar, 1979, Theorem 3), and the same observation applies to $\partial C \cap K$. Furthermore, the characterization of the interior of the Clarke tangent cone (e.g. Rockafellar, 1979, Theorem 2) implies that both K and $C \cap K$ are the closure of their interiors. Therefore Assumptions 6.3 and 6.4 implicitly require ∂K and $\partial C \cap K$ to be "Lipschitzian surfaces" and K and $C \cap K$ to be the closure of open sets of $\mathbb{R}^n$. In particular C cannot be a point target. Notice that we did not require C to be a subset of K, unlike Cannarsa and Castelpietra, 2008 where the inclusion $C \subset \operatorname{Int} K$ was assumed.

Remark 6.4. Owing to Aubin and Frankowska, 1990, Theorem 10.1.6, the above assumptions also imply that K is viable under $\operatorname{co} F$ (i.e. for any $x_0 \in K$ there exists a feasible $\operatorname{co} F$-trajectory defined on $[0, +\infty[$ starting at x_0). Bettiol et al., 2012, Theorem 2.3 shows that K is even viable under F.

The two following theorems are the main results of this article. Examples 6.1, 6.2 and 6.3 illustrate when they apply.

Theorem 6.1. *Under Assumptions 6.1, 6.2 and 6.3, Assumption 6.4 implies the following property of the minimum time function $\tau_{\min}$ to reach the target C subject to the state constraints K:*

$$\forall R > 0,\ \exists \delta > 0,\ k > 0,\ \forall x \in (C + \delta \mathbb{B}) \cap R\mathbb{B} \cap K,\ \tau_{\min}(x) \leqslant k d(x) := k d_{C \cap K}(x) \quad (6.4)$$

where, by convention, $d_\emptyset(x) = +\infty$.

Theorem 6.2. *Under Assumptions 6.1, 6.2, 6.3 and 6.4, $Capt_F(K, C)$ is open in K and $\tau_{\min}$ is locally Lipschitz continuous on $Capt_F(K, C)$.*

Example 6.1. Consider a two-dimensional simplified lunar landing module with orientable exhaust nozzle, subject to the lunar gravity pulling downwards. Define the constraints as being above the surface of the moon $K = \mathbb{R} \times \mathbb{R}_+$ and the target as a box $C = [-\epsilon, \epsilon] \times [0, \epsilon]$ with $\epsilon > 0$. When close to the ground, the pilot is gradually allowed to activate emergency boosters, strong enough to overcome the gravity, so that its dynamics are

$$F(x) := \underbrace{(0, -1)}_{\text{gravity}} + \underbrace{[-1, 1] \times \{0\}}_{\text{nozzle}} + \underbrace{\max(0, 1/2 - x_2)\{(0, 0), (0, 4)\}}_{\text{boosters}} \subset \mathbb{R}^2.$$

It can be easily checked that the triplet (F, K, C) satisfies all the Assumptions 6.1-6.4, and that $Capt_F(K, C) = K$.

Example 6.2. Define a scalar potential $g_y(x) = \max(0, \min(1 - \|x-y\|, \|x-y\|))$ centered at y. Consider a navigation problem: a child in a two-dimensional stream $\mathbb{R} \times [-2,2]$ wants to reach an aquatic slide $C = \mathbb{B}((0,0), 1/2)$ which creates a local whirl attractor. This defines the following dynamics

$$x' \in F(x) := \underbrace{(1,0)}_{\text{flow}} + \underbrace{\mathbb{B}((0,0), \frac{1}{2})}_{\text{swimmer's controls}} \underbrace{- 2 g_{(0,0)}(x) \frac{x}{\|x\|}}_{\text{whirl attractor}} \subset \mathbb{R}^2,$$

with the whirl continuously extended as $(0,0)$ at $x = (0,0)$. A wave generator centered at $x_K = (-2,0)$ is added. The constraint set is defined as $K = (\mathbb{R} \times [-2,2]) \setminus \mathrm{Int}(\mathbb{B}(x_K, 1/2))$ and the new dynamics, when the wave generator is on, are

$$x' \in \tilde{F}(x) := F(x) + \underbrace{g_{x_K}(x) \frac{x - x_K}{\|x - x_K\|}}_{\text{wave generator}}.$$

One can verify that $(\tilde{F}, K, C)$ satisfies the assumptions of Theorem 6.2. One can also check that $\mathrm{Capt}_{\tilde{F}}(K, C) \subsetneq K$ is a proper open subset of K. On the other hand, when the wave generator is off, (F, K, C) does not satisfy the strict-inward pointing condition at $(-5/2, 0) \in \partial K$.

6.3 DISCUSSION ON THE MAIN RESULTS

6.3.1 *Nonautonomous systems*

Consider the nonautonomous differential inclusion with initial condition $x_0 \in K$:

$$x'(t) \in F(t, x(t)) \qquad x(0) = x_0$$

where $F : \mathbb{R}_+ \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ is a set-valued map taking closed, nonempty values on $\mathbb{R}_+ \times \mathbb{R}^n$. Below, we shall make the same assumption as in Cannarsa and Castelpietra, 2008 that $F(\cdot, x)$ is locally Lipschitz continuous to show that Theorem 6.2 encompasses the results of Cannarsa and Castelpietra, 2008 for nonautonomous systems. This may appear as a restriction [2], however there are comparatively much less attempts at considering nonautonomous systems in their full generality (i.e. when $F(\cdot, x)$ is merely measurable). For differential inclusions and a moving target $C(t)$ but without state

2. Time-Lipschitzianity was acknowledged in Cannarsa and Castelpietra, 2008 as too restrictive. Nonetheless Cannarsa and Castelpietra, 2008 also considered the case of $F(t,x) = c(t,x)\mathbb{B}$ where c is a bounded scalar function, globally Lipschitz in x and merely measurable in t.

constraints, we should mention Veliov, 1997. For differential inclusions with fixed state constraints K but without target, we refer to Bettiol et al., 2012. As a midway, Cannarsa and Castelpietra, 2008 considered a nonautonomous control system with compact control set and fixed K and C.

Assumption 6.5. *Sublinear growth and local Lipschitz continuity of nonautonomous* F

$$\begin{gathered}\exists\, A \geqslant 0,\ \forall\, (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n,\ F(t,x) \subset A(1+\|x\|)\mathbb{B}\\ \forall\, R > 0, \forall\, T \geqslant 0\ \exists\, k_F \geqslant 0,\ \forall\, x,y \in R\mathbb{B},\ \forall\, t,s \in [0,T]\\ F(s,y) \subset F(t,x) + k_F(\|x-y\| + |t-s|)\mathbb{B}\end{gathered}$$

By augmenting the dynamics, our results encompass those of Cannarsa and Castelpietra, 2008. As a matter of fact, let K and C be two closed subsets of $\mathbb{R}^n$. We define the augmented system $\hat{x}'(t) \in \hat{F}(\hat{x})$ under state constraints $\hat{K}$ and with target $\hat{C}$ as

$$\begin{cases}\hat{K} = \mathbb{R}_+ \times K,\ \hat{C} = \mathbb{R}_+ \times C\\ \hat{x}(\cdot) = (\tau(\cdot), x(\cdot))\\ \hat{F}(\hat{x}) = (1, F(\hat{x}))\\ \hat{x}(0) = (0, x_0)\end{cases}$$

As $\tau'(\cdot) = 1$ and $\hat{K}$ and $\hat{C}$ are unbounded on the right in the time variable, Assumptions 6.1, 6.3, 6.4 and 6.5 on F, K and C jointly imply Assumptions 6.1, 6.2, 6.3, 6.4 bearing on $\hat{F}$, $\hat{K}$ and $\hat{C}$. So we may apply Theorem 6.2 to the augmented system. This shows that $Capt_{\hat{F}}(\hat{K},\hat{C})$ is open in $\hat{K}$ for the relative Euclidean topology of $\mathbb{R}_+ \times \mathbb{R}^n$ which is stronger than the Cartesian product topology. Furthermore the minimum time functions coincide for the two systems. We obtain therefore the conclusions of Theorem 6.2 for the original nonautonomous system.

In the proof of Theorems 6.1 and 6.2, a key ingredient, namely Bettiol et al., 2012, Theorem 2.3, was proven for nonautonomous differential inclusions when $F(\cdot,x)$ is absolutely continuous from the left. As a consequence, our results could eventually be extended to this class of systems.

6.3.2 *Weakening the hypotheses*

Among the other hypotheses, few could be relaxed. Indeed Assumptions 6.1 and 6.2 are fairly general and related to the global existence of solutions of (6.2) and to their Lipschitz dependence on initial conditions. Hence, they could hardly be weakened when we seek Lipschitz regularity.

Assumption 6.4 cannot be replaced by Assumption 6.3 expressed for C instead of K, as shown in (counter)-Example 6.3, so we require a jointly inward-pointing condition on K and C.

Example 6.3. Consider a modification of Example 6.1, where the controller over the nozzle of the landing module broke. The dynamics are now $x' \in F(x) = \{(0, \pm 1)\} \subset \mathbb{R}^2$, the constraints are still $K = \mathbb{R} \times \mathbb{R}_+$, and the target is now $C = \mathrm{Hyp}(x_1 \mapsto 1 - x_1^2)$, where Hyp stands for the hypograph. It is clear that K and C satisfy inward-pointing conditions with respect to F, while $C \cap K$ does not. Furthermore $\mathrm{Capt}_F(K, C) = K \cap ([-1, 1] \times \mathbb{R})$ is closed in K and τ_{min} is discontinuous on K. In other words, for some initial conditions, an infinitesimal perturbation may impede reaching the target.

We cannot either drop altogether assuming inward-pointing conditions on C. Indeed, the problem without state constraints is a special case of a state-constrained problem. So the necessary and sufficient conditions from Veliov, 1997, Theorem 2.1 have to remain valid. Let $\Gamma = \mathrm{Graph}\, C(\cdot)$ where $(C(t))_{t \geqslant 0}$ is a moving target. In the simpler case of $F(\cdot, x)$ being continuous and $F(t, \cdot)$ being locally Lipschitz, the necessary condition for the local Lipschitzianity of the minimum time, as shown in Veliov, 1997, Corollary 2.1, Theorem 3.1 can be stated as follows[3]: for every compact set $G \subset \mathbb{R}_+ \times \mathbb{R}^n$, there exists $\epsilon > 0$ such that, for every $(t, x) \in \partial\Gamma \cap G$,

$$\sup_{(p^0, p) \in N_\Gamma(t,x) \cap S^n} \left(\min_{v \in F(t,x)} \left(p^0 + \langle p, v \rangle \right) \right) \leqslant -\epsilon \tag{6.5}$$

Condition (6.5) would then have to be satisifed. On the other hand, if the set K is invariant under F (i.e. for any $x_0 \in K$ all the solutions of the differential inclusion (6.2) are feasible), then no inward-pointing condition on K is required as all the trajectories are already feasible. Nevertheless we still need a strict inward-pointing condition on C, such as (6.5).

6.3.3 *Considering point targets*

Originally, Petrov, 1970 considered point targets (i.e. $C = \{\bar{x}\}$) and devised a necessary and sufficient condition ($0 \in \mathrm{Int}(\mathrm{co}\, F(\bar{x}))$) for the Lipschitz continuity of the minimum time in a neighborhood of $\bar{x}$, in the case without constraints. Since then, research focused mainly on inward-pointing conditions which preclude point targets as discussed in Remark 6.3. We show below that for point targets $\bar{x} \in \mathrm{Int}\, K$, we still

3. The original condition of Veliov, 1997 bore on the proximal normal cone to Γ but, by taking the limit and the closed convex hull, it can be restated for the Clarke normal cone.

have the results of Theorems 6.1 and 6.2. Whenever $\bar{x} \in \partial K$, it is still an open question as to formulating sufficient conditions for local Lipschitzianity of the minimum time.

We begin with an extension of Petrov, 1970, Theorem 4.1 to differential inclusions without state constraints.

Proposition 6.1. *Let $\bar{x} \in \mathbb{R}^n$ and suppose that $0 \in \mathrm{Int}(\mathrm{co}\, F(\bar{x}))$. Then, under Assumptions 6.1 and 6.2, the minimum time function $\tau_{\min}$ to reach the target $\bar{x}$ without state constraints has the following property:*

$$\exists\, \delta > 0,\ k > 0,\ \forall\, x \in (\bar{x} + \delta \mathbb{B}),\ \tau_{\min}(x) \leqslant k\|x - \bar{x}\| \tag{6.6}$$

A proof of Proposition 6.1 appears in the next section. We first state our result for point targets $\bar{x} \in \mathrm{Int}\, K$.

Proposition 6.2. *Let $\bar{x} \in \mathrm{Int}\, K$ and suppose that $0 \in \mathrm{Int}(\mathrm{co}\, F(\bar{x}))$. Then, under Assumptions 6.1, 6.2 and 6.3, $\tau_{\min}$ satisfies (6.6) and is locally Lipschitz continuous on $\mathrm{Capt}_F(K, \{\bar{x}\})$, which is open in K.*

Proof of Proposition 6.2: Fix $\eta > 0$ such that $(\bar{x} + \eta \mathbb{B}) \subset \mathrm{Int}\, K$. Apply Proposition 6.1 to the system (6.2) without constraints, this gives $\delta_0 > 0$ and $k > 0$ such that, for any $\epsilon > 0$ and any $x \in (\bar{x} + \delta_0 \mathbb{B})$, there exists $\tau \in [0, k\|x - \bar{x}\| + \epsilon]$ and an F-trajectory $y(\cdot)$ defined on $[0, \tau]$ satisfying:

$$y(0) = x \qquad y(\tau) = \bar{x} \tag{6.7}$$

Let $A > 0$ be as in Assumption 6.2. Fix any $\delta \in]0, \delta_0]$ and $\epsilon > 0$ satisfying:

$$(\delta + (k\delta + \epsilon) \cdot (A + A\|\bar{x}\|))\, e^{A(k\delta+\epsilon)} \leqslant \eta$$

Take $x \in (\bar{x} + \delta \mathbb{B})$ and consider τ and $y(\cdot)$ as in (6.7). Then for any $t \in [0, \tau]$ we have:

$$\begin{aligned}
\|\bar{x} - y(t)\| &\leqslant \|\bar{x} - x\| + \|x - y(t)\| \leqslant \delta + \int_0^t \|y'(s)\| ds \\
&\leqslant \delta + \int_0^t (A + A\|\bar{x}\| + A\|\bar{x} - y(s)\|) ds \\
&\leqslant \delta + \tau(A + A\|\bar{x}\|) + A \int_0^t \|\bar{x} - y(s)\| ds
\end{aligned}$$

Using Gronwall's lemma, we get that for any $t \in [0, \tau]$, as $\tau \leqslant k\delta + \epsilon$

$$\|\bar{x} - y(t)\| \leqslant (\delta + \tau(A + A\|\bar{x}\|))\, e^{A\tau} \leqslant \eta$$

Consequently $y([0, \tau]) \subset \bar{x} + \eta \mathbb{B} \subset \mathrm{Int}\, K$, so $y(\cdot)$ is feasible. As ϵ is arbitrary, we get that the minimum time with state constraints coincides on $(\bar{x} + \delta \mathbb{B})$ with the minimum time without state constraints, and that both are therefore Lipschitz continuous on $(\bar{x} + \delta \mathbb{B})$. The results of Theorem 6.2 follow immediately, which concludes the proof.

■

6.4 PROOFS

Proof of Theorem 6.1: This proof is partially inspired by Cannarsa and Sinestrari, 2004, pp. 239-243 where the result was proven for a control system without state constraints. The proof differs however due to Assumption 6.4 which is weaker than in Cannarsa and Sinestrari, 2004 as it bears on $\operatorname{co} F$ rather than F, and as we consider a general differential inclusion rather than a control system. Moreover the presence of constraints requires to design feasible trajectories (i.e. respecting the state constraints). This leads to applying both the celebrated relaxation theorem and a "correction" theorem Bettiol et al., 2012, Theorem 2.3 to build F-trajectories staying in K.

Fix any $R > 0$. Let $k_F > 0$ such that:

$$\forall x, y \in 2R\mathbb{B},\ F(y) \subset F(x) + k_F\|x-y\|\mathbb{B}$$

Let $c_F := 5nk_F$ and define

$$M = n + \sup_{x\in 2R\mathbb{B}} \sup_{v\in F(x)} \|v\|.$$

By Assumptions 6.3 and 6.4 and Bettiol et al., 2012, Lemma 5.3, there exists $\epsilon \in]0,1]$, $\eta_0 > 0$ such that:

$$\begin{cases} \forall x \in (\partial(C\cap K)+\eta_0\mathbb{B})\cap 2R\mathbb{B}\cap C\cap K,\ \exists v\in \operatorname{co} F(x), \\ \quad \forall y \in (x+\eta_0\mathbb{B})\cap C\cap K,\ y+[0,\epsilon](v+\epsilon\mathbb{B})\subset C\cap K \end{cases} \tag{6.8}$$

$$\begin{cases} \forall x \in (\partial C\cap \partial K+\eta_0\mathbb{B})\cap 2R\mathbb{B}\cap C\cap K,\ \exists v\in \operatorname{co} F(x), \\ \quad \forall y \in (x+\eta_0\mathbb{B})\cap K,\ y+[0,\epsilon](v+\epsilon\mathbb{B})\subset K \\ \quad \forall z \in (x+\eta_0\mathbb{B})\cap C\cap K,\ z+[0,\epsilon](v+\epsilon\mathbb{B})\subset C\cap K. \end{cases} \tag{6.9}$$

In order to use Gronwall's lemma later on, we require a technical condition on ϵ, which has to be chosen small enough as to satisfy:

$$\epsilon e^{c_F\epsilon/(16M^2)} \leqslant 8MR \tag{6.10}$$

Since $\operatorname{co}(F)$ is locally Lipschitz continuous, there exists $\eta_1 > 0$, such that for every $\tilde{x} \in \partial C\cap K\cap 2R\mathbb{B}$ and $\tilde{v}\in \operatorname{co} F(\tilde{x})$

$$\forall x \in (\tilde{x}+\eta_1\mathbb{B}),\ \exists v\in \operatorname{co} F(x),\ \|v-\tilde{v}\| \leqslant \epsilon/4. \tag{6.11}$$

Let $\eta := \min(\eta_0, \eta_1, R)/2$. For any $\delta \in]0, \eta/3]$, we now define the following sets and quantities:

$$\Delta = \partial C \cap \partial K \cap 2R\mathbb{B} \qquad \tilde{C} = (\partial C \cap K \cap 2R\mathbb{B}) \setminus (\Delta + \frac{\eta}{2}\mathbb{B}) \qquad \tilde{\rho} = \inf_{x \in \tilde{C}} d_{\partial K}(x)$$

$$C_\delta = \partial C \cap K + \delta\mathbb{B} \qquad \tilde{C}_\delta = \left(C_\delta \cap K \cap \frac{3R}{2}\mathbb{B}\right) \setminus (\Delta + \eta\mathbb{B}) \qquad \tilde{\rho}_\delta = \inf_{x \in \tilde{C}_\delta} d_{\partial K}(x)$$

where, by convention, $\emptyset + \mathbb{B} = \emptyset$. The initial conditions x_0 of interest are in $C_\delta \cap K \cap R\mathbb{B}$, hence they either belong to $(\Delta + \eta\mathbb{B})$ or to $\tilde{C}_\delta$.

Claim 6.1. *Whenever $\tilde{C}_\delta \neq \emptyset$, $\tilde{\rho} > 0$ is finite, independently of Δ being empty or not.*

Proof of Claim 6.1: Suppose $\tilde{C}_\delta \neq \emptyset$ and let $x \in \tilde{C}_\delta \subset C_\delta$. Fix any $\bar{x} \in \partial C \cap K$ such that $\|x - \bar{x}\| \leqslant \delta$. As $\delta \leqslant \frac{\eta}{3} \leqslant \frac{R}{6}$ and $x \in \frac{3R}{2}\mathbb{B}$, we have that $\bar{x} \in (x + \delta\mathbb{B}) \subset 2R\mathbb{B}$. Further,

$$d_\Delta(\bar{x}) \geqslant d_\Delta(x) - \|x - \bar{x}\| \geqslant \eta - \delta > \eta/2.$$

Therefore $\bar{x} \in \tilde{C}$, implying it is not empty. If $\Delta \neq \emptyset$, by definition, $\tilde{\rho} > 0$. If $\Delta = \emptyset$, then $\tilde{C} = \partial C \cap K \cap 2R\mathbb{B}$ is compact and $\tilde{C} \cap \partial K = \emptyset$, so $\tilde{\rho} > 0$. In both cases, as $\tilde{C}_\delta \subset \tilde{C} + \delta\mathbb{B}$, $\tilde{\rho}_\delta \geqslant \tilde{\rho} - \delta$.

We now define the key constants that the proof of Theorem 6.1 will require:

$$k_0 := 4/\epsilon \qquad \gamma := \sqrt{1 - \left(\frac{\epsilon}{16M}\right)^2}$$

$$k := \frac{2}{\epsilon}(1 + \gamma) \qquad \tilde{\gamma} := \frac{1 + \gamma}{2}$$

Notice that

$$k(1 - \tilde{\gamma}) = \frac{k(1 - \gamma)}{2} = \frac{\epsilon}{(16M)^2}. \tag{6.12}$$

Define

$$\delta := \min\left(1, \frac{\eta}{3}, \frac{\tilde{\rho}}{4 + 16k_0M}, \frac{\epsilon}{8c_FM}, \frac{R}{32kM}\right). \tag{6.13}$$

As $\delta \leqslant 1$, we may replace ϵ by $\epsilon\delta$ in (6.10).

Let us show that for any $x_0 \in (C_\delta \cap K \cap R\mathbb{B})$, we can define, for $j \in \mathbb{N}$, a feasible F-trajectory $y_j(\cdot)$ on $[0, t_j]$ with $t_j > 0$, satisfying the following properties for $d(x) := d_{C \cap K}(x)$

$$x_{j+1} := y_{j+1}(0) = y_j(t_j) \in \frac{3R}{2}\mathbb{B} \qquad \sum_{j=0}^{\infty} t_j \leqslant 16kd(x_0)$$

and the two inequalities

$$d(x_j) \leqslant \tilde{\gamma}^j d(x_0) \qquad \|x_j - x_0\| \leqslant \frac{\epsilon}{16M} \sum_{k=0}^{j-1} \tilde{\gamma}^k d(x_0). \tag{6.14}$$

Let us first make sure that the $(x_j)_j$ satisfying (6.14) belong to $\frac{3R}{2}\mathbb{B}$, by applying (6.12). Indeed

$$\|x_j\| \leqslant \|x_0\| + \|x_j - x_0\| \leqslant R + \frac{\epsilon}{16M} \frac{1}{1-\tilde{\gamma}} d(x_0) \leqslant R + 16kM\delta \leqslant \frac{3R}{2}.$$

Let $j \in \mathbb{N}$. The inequalities (6.14) are obviously satisfied for $j = 0$, so we will proceed by induction on j. Assume we have already constructed our trajectories up to step j and have not yet reached C (i.e. $x_j \notin C$). Let $\bar{x}_j \in \partial(C \cap K)$ be such that $d(x_j) = \|x_j - \bar{x}_j\|$ (i.e. $\bar{x}_j \in \Pi_{C\cap K}(x_j)$).

As $x_j \in C_\delta \cap K \cap (3R/2)\mathbb{B} = \tilde{C}_\delta \cup (\Delta + \eta\mathbb{B})$, we distinguish two cases. In Case 1, we consider the situation where $\Delta \neq \emptyset$ and $x_j \in (\Delta + \eta\mathbb{B})$. If the x_{j+1} that we design below belongs to $\tilde{C}_\delta$, then we move to Case 2, otherwise the induction proceeds according to Case 1. In Case 2, $x_j \in \tilde{C}_\delta$ and we build by induction $(x_m)_{m\geqslant j+1} \in C_\delta$. Owing to Claim 1, we will show that, for any $m \geqslant j+1$ and $t \in [0, t_m]$, the designed trajectory $y_m(\cdot)$ satisfies $d_{\partial K}(y_m(t)) > 0$. This latter property ensures that once the designed trajectory is far enough from ∂K (i.e. $x_j \in \tilde{C}_\delta$), it stays so, and we can focus on reaching C.

Case 1: Suppose $\Delta \neq \emptyset$ and $x_j \in (\Delta + \eta\mathbb{B})\backslash C$, then take any $\tilde{x}_j \in \Delta \cap \mathbb{B}(x_j, \eta)$. Consider $\tilde{v}_j \in \operatorname{co} F(\tilde{x}_j)$ satisfying (6.9). As $\tilde{x}_j \in \Delta \subset C \cap K \cap 2R\mathbb{B}$,

$$\begin{gathered}\|x_j - \bar{x}_j\| = d(x_j) \leqslant \|x_j - \tilde{x}_j\| \leqslant \eta \\ \|\tilde{x}_j - \bar{x}_j\| \leqslant \|\tilde{x}_j - x_j\| + \|x_j - \bar{x}_j\| \leqslant 2\eta \leqslant \min(\eta_1, \eta_0) \\ \|\bar{x}_j\| \leqslant \|x_j\| + \|x_j - \bar{x}_j\| \leqslant \frac{3R}{2} + \eta \leqslant 2R\end{gathered}$$

we may thus apply (6.11) at $\tilde{x}_j$ and $\bar{x}_j$ (instead of $\tilde{x}$ and x) and then at $\bar{x}_j$ and x_j. In this way, we get $\bar{v}_j \in \operatorname{co} F(\bar{x}_j)$ and $v_j \in \operatorname{co} F(x_j)$ both satisfying (6.11)

$$\|\tilde{v}_j - \bar{v}_j\| \leqslant \epsilon/4 \qquad \|v_j - \bar{v}_j\| \leqslant \epsilon/4$$

and the second (resp. first) line of (6.9) at $\tilde{x}_j$ and $\bar{x}_j$ (resp. at $\tilde{x}_j$ and x_j)

$$\bar{x}_j + [0, \epsilon](\tilde{v}_j + \epsilon\mathbb{B}) \subset C \cap K \qquad x_j + [0, \epsilon](\tilde{v}_j + \epsilon\mathbb{B}) \subset K$$

which implies that

$$\bar{x}_j + [0, \epsilon](v_j + \epsilon/2\mathbb{B}) \subset C \cap K \qquad x_j + [0, \epsilon](v_j + \epsilon/2\mathbb{B}) \subset K.$$

Using that $(x_j - \bar{x}_j) \in N_{C\cap K}(\bar{x}_j)\backslash\{0\}$, relation (6.3) gives $\langle v_j, x_j - \bar{x}_j\rangle \leqslant -\|x_j - \bar{x}_j\|\epsilon/2$. Hence we have shown the two following formulas relating x_j and v_j:

$$\langle v_j, x_j - \bar{x}_j\rangle \leqslant -\|x_j - \bar{x}_j\|\epsilon/2 \qquad x_j + [0,\epsilon](v_j + \epsilon/2\mathbb{B}) \subset K \tag{6.15}$$

Let us now design a trajectory starting at x_j. Define the duration $t_j > 0$ as follows

$$t_j := \frac{\epsilon}{16M^2} d(x_j) \tag{6.16}$$

By Aubin and Frankowska, 1990, Theorem 9.5.3, there exists a c_F-Lipschitz selection f from the set-valued map $\operatorname{co} F$ defined on $2R\mathbb{B}$, satisfying $v_j = f(x_j)$. Let $\check{y}_j(\cdot)$ be the unique solution of the differential equation $x'(t) = f(x(t))$ on $[0, t_j]$ with initial condition $x(0) = x_j$. Let $w_j(\cdot) := f(\check{y}_j(\cdot))$ on $[0, t_j]$. Then

$$\check{y}_j(t) = x_j + \int_0^t w_j(s)ds \tag{6.17}$$

Furthermore we have $\|w_j(s) - v_j\| \leqslant c_F\|\check{y}_j(s) - x_j\|$, and

$$\|\check{y}_j(t) - x_j\| \leqslant \|tv_j\| + \int_0^t \|w_j(s) - v_j\| \leqslant Mt_j + c_F\int_0^t \|\check{y}_j(s) - x_j\|$$

We apply Gronwall's lemma as $\check{y}_j(t)$ is continuous and we take into account (6.10), recalling that $\delta \leqslant 1$

$$\|\check{y}_j(t) - x_j\| \leqslant Mt_j e^{c_F t_j} \leqslant \frac{\epsilon\delta}{16M} e^{\epsilon\delta c_F/(16M^2)} \leqslant R/2 \tag{6.18}$$

Therefore $\check{y}_j(t) \in 2R\mathbb{B}$, which implies that $w_j(t) \in M\mathbb{B}$. Thus (6.17) gives $\|\check{y}_j(s) - x_j\| \leqslant Mt_j$. As $M \geqslant 1$ and $\epsilon \leqslant 1$:

$$\begin{aligned}\|\check{y}_j(t) - x_j - tv_j\| &\leqslant \int_0^t \|w_j(s) - v_j\| \leqslant c_F\int_0^t \|\check{y}_j(s) - x_j\| \leqslant c_F Mt_j \cdot t \\ &\leqslant c_F M\frac{\epsilon\delta}{16M^2}\cdot t \leqslant c_F M\frac{\epsilon}{16M^2}\frac{\epsilon}{8c_F M}\cdot t \leqslant \frac{\epsilon}{2}\cdot t\end{aligned}$$

Thanks to (6.15), this ensures that $\check{y}_j(t) \in \big(x_j + t(v_j + \epsilon/2\mathbb{B})\big) \subset K$. We have so far designed a feasible co F-trajectory. Applying (6.15), we obtain furthermore for any $t \in [0, t_j]$:

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|\check{y}_j(t) - \bar{x}_j\|^2 &= \langle w_j(t), \check{y}_j(t) - \bar{x}_j\rangle \\
&= \langle v_j, x_j - \bar{x}_j\rangle + \langle w_j(t), \check{y}_j(t) - x_j\rangle + \langle w_j(t) - v_j, x_j - \bar{x}_j\rangle \\
&\leqslant -\frac{\epsilon}{2}\|x_j - \bar{x}_j\| + M\|\check{y}_j(t) - x_j\| + c_F\|\check{y}_j(t) - x_j\|\|x_j - \bar{x}_j\| \\
&\leqslant -\frac{\epsilon}{2}d(x_j) + (M + c_F d(x_j))Mt \\
&\leqslant \left(-\frac{\epsilon}{2} + c_F M t_j\right) d(x_j) + M^2 t_j \\
&= \left(-\frac{\epsilon}{2} + \frac{c_F\delta}{M}\frac{\epsilon}{16} + \frac{\epsilon}{16}\right) d(x_j) \leqslant -\frac{3\epsilon}{8}d(x_j) \leqslant -\frac{\epsilon}{32}d(x_j) = -\frac{M^2 t_j}{2}
\end{aligned}$$

Consequently, by integration:

$$d^2(\check{y}_j(t_j)) \leqslant \|\check{y}_j(t_j) - \bar{x}_j\|^2 \leqslant d^2(x_j) - M^2 t_j^2 = \left(1 - M^2\left(\frac{\epsilon}{16M^2}\right)^2\right) d^2(x_j) = \gamma^2 d^2(x_j)$$

However we cannot set $\check{y}_j(t_j)$ as the next x_{j+1}, as it is only a co F-trajectory for the time being. In order to apply the relaxation theorem, we need a globally Lipschitz continuous set-valued map. Fix any $\epsilon_j > 0$. Let $\tilde{F}$ be defined in any $x \in \mathbb{R}^n$ as $\tilde{F}(x) := F(\Pi_{(2R+\epsilon_j)\mathbb{B}}(x))$ where $\Pi_{(2R+\epsilon_j)\mathbb{B}}(x)$ is the unique projection of x into $(2R + \epsilon_j)\mathbb{B}$. Since the projection on a ball is Lipschitz, we deduce that $\tilde{F}$ is globally Lipschitz. As $\check{y}_j([0, t_j]) \subset 2R\mathbb{B}$, $\check{y}_j(\cdot)$ is also an $\tilde{F}$-trajectory. Thanks to the relaxation theorem, we may thus build an $\tilde{F}$-trajectory $\hat{y}_j(\cdot)$ starting from x_j and enjoying the following property:

$$\|\hat{y}_j - \check{y}_j\|_{L_\infty([0,t_j])} \leqslant \epsilon_j$$

As $\check{y}_j([0, t_j]) \subset 2R\mathbb{B}$, $\hat{y}_j([0, t_j]) \subset (2R + \epsilon_j)\mathbb{B}$, on which $\tilde{F}$ and F coincide. Hence $\hat{y}_j(\cdot)$ is an F-trajectory.

If $\hat{y}_j([0, t_j]) \subset K$, then we keep it as our feasible F-trajectory. Otherwise, if it leaves K even during a short time, we correct it into an F-feasible trajectory $y_j(\cdot)$ staying in Int K, through Bettiol et al., 2012, Theorem 2.3 which we apply on $2R\mathbb{B}$ and time interval $[0, t_j] \subset [0, t_{max}]$ with $t_{max} = \epsilon\delta/(16M^2) \geqslant t_j$. This implies the existence of a constant $L \geqslant 1$ (depending only on R and F) with the following property (as $\check{y}_j([0, t_j]) \subset K$):

$$\begin{gathered}
\|y_j - \hat{y}_j\|_{L_\infty([0,t_j])} \leqslant L\|d_K(\hat{y}_j(\cdot))\|_{L_\infty([0,t_j])} \leqslant L\|\hat{y}_j - \check{y}_j\|_{L_\infty([0,t_j])} \leqslant L\epsilon_j \\
d(y_j(t_j)) \leqslant d(\hat{y}_j(t_j)) + \|y_j(t_j) - \hat{y}_j(t_j)\| \leqslant \gamma d(x_j) + L\epsilon_j
\end{gathered}$$

The above relations remain true even if $\hat{y}_j([0,t_j]) \subset K$. We set $\epsilon_j := \frac{1-\gamma}{2L} d(x_j) \leqslant \delta$ and $x_{j+1} := y_j(t_j)$, thus:

$$d(x_{j+1}) \leqslant (\gamma + \frac{1-\gamma}{2}) d(x_j) = \tilde{\gamma} d(x_j)$$

Moreover we derive from (6.14) and (6.16):

$$\|x_{j+1} - x_0\| \leqslant \|x_j - x_0\| + \|x_{j+1} - x_j\| \leqslant \frac{\epsilon}{16M} \sum_{k=0}^{j-1} \tilde{\gamma}^k d(x_0) + Mt_j \leqslant \frac{\epsilon}{16M} \sum_{k=0}^{j} \tilde{\gamma}^k d(x_0)$$

If $x_{j+1} \in (\Delta + \eta\mathbb{B})\backslash C$, then we remain in Case 1. Otherwise, if $x_{j+1} \in \tilde{C}_\delta \backslash C$, we move to Case 2.

Case 2: Suppose that $x_j \in \tilde{C}_\delta \backslash C$. The trajectory construction is similar to Case 1 and even simpler as we do not have to consider the point $\tilde{x}_j$. As $\eta \geqslant \delta$, we can still select $\bar{v}_j \in \operatorname{co} F(\bar{x}_j)$ and $v_j \in \operatorname{co} F(x_j)$ satisfying (6.8) and $\|v_j - \bar{v}_j\| \leqslant \epsilon/4$. We define t_j and $\check{y}_j(\cdot)$ as in (6.16), and apply (6.18). This leads to the same computations for $\frac{d}{dt}\|\check{y}_j(t) - \bar{x}_j\|^2$ and for $d(\check{y}_j(t_j))$. We define again through relaxation an F-trajectory $\hat{y}_j(\cdot)$ with the same ϵ_j and we set $x_{j+1} = \hat{y}_j(t_j) \in C_\delta$, which satisfies (6.14). Then repeating the above steps, we build a sequence $(x_m)_{m \geqslant j+1} \in C_\delta$ connected by trajectories $\hat{y}_m(\cdot)$.

We no longer have to check the feasibility of such $\hat{y}_m(\cdot)$. As a matter of fact, recall that owing to Claim 6.1, as $\tilde{C}_\delta \neq \emptyset$, $\tilde{\rho}$ is finite. Let $m \geqslant j+1$. Using that $\gamma \leqslant \tilde{\gamma} \leqslant 1$, $k \leqslant k_0$ and $d(x_j) \leqslant \delta \leqslant \tilde{\rho}/(4 + 16k_0M)$, we deduce

$$\begin{aligned}
d_{\partial K}(\hat{y}_m(t)) &\geqslant d_{\partial K}(x_j) - \|\hat{y}_m(t) - x_j\| \\
&\geqslant \tilde{\rho}_\delta - \|\hat{y}_m(t) - \check{y}_m(t)\| - \|\check{y}_m(t) - \bar{x}_m\| - \|\bar{x}_m - x_m\| - \|x_m - x_j\| \\
&\geqslant \tilde{\rho} - \delta - L\epsilon_m - 2d(x_m) - \frac{\epsilon}{16M(1-\tilde{\gamma})} d(x_j) \\
&\geqslant \tilde{\rho} - \delta - \frac{1-\gamma}{2} d(x_m) - 2\tilde{\gamma}^{m-j} d(x_j) - 16kMd(x_j) \\
&\geqslant \tilde{\rho} - \delta - \frac{1}{2}\tilde{\gamma}^{m-j} d(x_j) - 2d(x_j) - 16kMd(x_j) \\
&\geqslant \tilde{\rho} - (\frac{7}{2} + 16k_0M)\delta > 0
\end{aligned}$$

The above computation ensures that whenever an x_j is both close enough to C and far enough from ∂K, we can focus only on reaching C.

To conclude, we have built a bounded sequence $(x_j)_{j \geqslant 0}$ connected by feasible F-trajectories $y_j(\cdot)$, satisfying both (6.14) and:

$$\lim_{j \to +\infty} d(x_j) = 0$$

Using (6.12), define τ as follows:

$$\tau := \sum_{j=0}^{\infty} t_j = \frac{\epsilon}{16M^2} \sum_{j=0}^{\infty} d(x_j) \leqslant \frac{\epsilon}{16M^2} \frac{1}{1-\tilde{\gamma}} d(x_0) = 16k d(x_0)$$

Concatenating all the feasible F-trajectories $y_j(\cdot)$, we get a feasible F-trajectory $y(\cdot)$ starting at x_0, defined on $[0,\tau]$ and reaching $C \cap K$ at time τ. Hence assertion (6.4) follows:

$$\tau_{min}(x_0) \leqslant \tau \leqslant 16k d(x_0)$$

which concludes the proof, replacing $16k$ by k to recover (6.4).

■

Proof of Theorem 6.2: Let $x_0 \in \mathrm{Capt}_F(K,C)\backslash C$ and $\xi > 0$ such that $\mathbb{B}(x_0,\xi) \cap C = \emptyset$. Let $y_0(\cdot)$ be a feasible F-trajectory starting at x_0 and reaching C at $\bar{x}_0$ at some time τ_0 where $\tau_{min}(x_0) \leqslant \tau_0 \leqslant 2\tau_{min}(x_0)$. Let $x_1 \in K \cap \mathbb{B}(x_0,\xi)$. Define

$$R := (\|\bar{x}_0\| + 1)\, e^{M(2\tau_{min}(x_0)+1)} \qquad \text{where } M := \sup_{z\in\mathbb{B},\, t\in[0,\tau_0]} \sup_{v\in F(y_0(t)+z)} \|v\|$$

Let k_F be as in Assumption 6.2 for this value of R. Let $\tilde{F}$ be defined in any $x \in \mathbb{R}^n$ as $\tilde{F}(x) := F(\Pi_{R\mathbb{B}}(x))$. Since the projection on a ball is Lipschitz with constant 1, we deduce that $\tilde{F}$ is globally Lipschitz with constant k_F. As $y_0([0,\tau_0]) \subset R\mathbb{B}$, $y_0(\cdot)$ is also an $\tilde{F}$-trajectory. We may then apply the Filippov's existence theorem (see e.g. Aubin and Frankowska, 1990, Theorem 10.4.1), to design an $\tilde{F}$-trajectory $\hat{y}_1(\cdot)$ on $[0,\tau_0]$ starting from x_1 such that:

$$\|y_0 - \hat{y}_1\|_{L_\infty([0,\tau_0])} \leqslant \|x_0 - x_1\| e^{k_F\tau_0} \leqslant c\cdot\xi \text{ where } c := e^{k_F(2\tau_{min}(x_0)+1)}$$

Take from now on $\xi \leqslant 1/c$. Therefore, for any $t \in [0,\tau_0]$, $\hat{y}_1(t) \in (y_0(t) + \mathbb{B})$ and in particular $\hat{y}_1(\tau_0) \in (\bar{x}_0 + \mathbb{B})$), from which we derive that $\hat{y}_1([0,\tau_0]) \subset R\mathbb{B}$, on which $\tilde{F}$ and F coincide. We thus conclude that $\hat{y}_1(\cdot)$ is an F-trajectory. If it stays within K, we keep it. Otherwise we apply Bettiol et al., 2012, Theorem 2.3 to retrieve an F-trajectory $y_1(\cdot)$ on $[0,\tau_0]$ starting from x_1 and staying in K, satisfying in both cases for an $L \geqslant 0$ (depending only on R and F):

$$\begin{aligned}\|y_0 - y_1\|_{L_\infty([0,\tau_0])} &\leqslant \|y_0 - \hat{y}_1\|_{L_\infty([0,\tau_0])} + \|\hat{y}_1 - y_1\|_{L_\infty([0,\tau_0])} \\ &\leqslant (1+L)\|y_0 - \hat{y}_1\|_{L_\infty([0,\tau_0])} \leqslant c(1+L)\xi\end{aligned}$$

Let $\delta > 0$, $k > 0$ and $d(\cdot)$ be as in (6.4). We choose ξ such that $c(1+L)\xi \leqslant \min(1/k, \delta, 1)$. As $y_1(\tau_0) \in (C + \delta\mathbb{B}) \cap K \cap R\mathbb{B}$, we deduce from the dynamic programming principle that:

$$\begin{aligned}\tau_{min}(x_1) &\leqslant \tau_0 + \tau_{min}(y_1(\tau_0)) \leqslant \tau_0 + kd(y_1(\tau_0)) \\ &\leqslant 2\tau_{min}(x_0) + ck(1+L)\xi \leqslant 2\tau_{min}(x_0) + 1\end{aligned}$$

We have thus shown that $x_1 \in \mathrm{Capt}_F(K, C)$. Since x_1 is an arbitrary point in K in a neighborhood of x_0, $\mathrm{Capt}_F(K, C)$ is open in K.

We now repeat the above strategy for $x_1, x_2 \in \mathbb{B}(x_0, \xi/2) \cap K$. Let $\epsilon_1 \geqslant 0$. Let $y_1(\cdot)$ be a feasible F-trajectory starting at x_1 and reaching C at $\bar{x}_1$ at time $\tau_1 \leqslant \tau_{min}(x_1) + \epsilon_1$. Then design a feasible F-trajectory $y_2(\cdot)$ on $[0, \tau_1]$ starting from x_2. With the same arguments as above, $y_2(\tau_1) \in R\mathbb{B}$ and:

$$\begin{aligned}d(y_2(\tau_1)) \leqslant \|y_2(\tau_1) - \bar{x}_1\| \leqslant c(1+L)\|x_2 - x_1\| \leqslant c(1+L)\xi \leqslant \delta \\ \tau_{min}(x_2) \leqslant \tau_1 + \tau_{min}(y_2(\tau_1)) \leqslant \tau_{min}(x_1) + \epsilon_1 + c(1+L)k\|x_2 - x_1\|.\end{aligned}$$

As the roles of x_1 and x_2 can be permuted and ϵ_1 is arbitrary, $\tau_{min}(\cdot)$ is Lipschitz continuous on $\mathbb{B}(x_0, \xi/2) \cap K$.

■

Proof of Proposition 6.1: This constructive proof is largely similar to that of Theorem 6.1, we thus focus only on the few differences as here $K = \mathbb{R}^n$ and $d(x) = \|x - \bar{x}\|$.

Let $R = \|\bar{x}\| + 1$ and define the constants k_F, c_F and M accordingly. Fix $\epsilon > 0$ satisfying both $\epsilon\mathbb{B} \subset \mathrm{co}\, F(\bar{x})$ and (6.10), then define the other constants η_1, k_0, k, γ and $\tilde{\gamma}$ as in the proof of Theorem 6.1. Set δ as follows

$$\delta := \min\left(1, \frac{\eta_1}{6}, \frac{R}{6}, \frac{\epsilon}{8c_F M}, \frac{R}{32kM}\right).$$

Let $x_0 \in (\bar{x} + \delta\mathbb{B})$, $j \in \mathbb{N}$ and suppose that (6.14) is satisfied at x_j. Let $\bar{v}_j = -\epsilon(x_j - \bar{x})/\|x_j - \bar{x}\|$. As $\bar{v}_j \in \mathrm{co}\, F(\bar{x})$, through (6.11), we can fix a $v_j \in \mathrm{co}\, F(x_j)$ such that $\|v_j - \bar{v}_j\| \leqslant \epsilon/4$. Hence

$$\langle v_j, x_j - \bar{x}\rangle \leqslant \langle \bar{v}_j, x_j - \bar{x}\rangle + \|x_j - \bar{x}\|\|v_j - \bar{v}_j\| \leqslant -\frac{\epsilon}{2}\|x_j - \bar{x}\|$$

and we recover (6.15). Let t_j as in (6.16), and construct similarly $\hat{y}_j(\cdot)$ through (6.17) and relaxation. This defines the next point $x_{j+1} = \hat{y}_j(t_j)$ for the induction. In conclusion, the bounded sequence $(x_j)_{j\geqslant 0}$ converges to $\bar{x}$, which is reached in time less than $16k\|x_0 - \bar{x}\|$ by an F-trajectory starting at x_0.

■

BIBLIOGRAPHY FOR CHAPTER 6

Aubin, J.-P., & Frankowska, H. (1990). *Set-valued analysis*. Birkhäuser Boston.

Bardi, M., & Falcone, M. (1990). An approximation scheme for the minimum time function. *SIAM Journal on Control and Optimization*, *28*(4), 950–965.

Bettiol, P., Frankowska, H., & Vinter, R. B. (2012). L^∞ Estimates on trajectories confined to a closed subset. *Journal of Differential Equations*, *252*(2), 1912–1933.

Cannarsa, P., & Castelpietra, M. (2008). Lipschitz continuity and local semiconcavity for exit time problems with state constraints. *Journal of Differential Equations*, *245*(3), 616–636.

Cannarsa, P., & Sinestrari, C. (2004). *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Birkhäuser Boston.

Friedman, A. (1970). Existence of value and of saddle points for differential games of pursuit and evasion. *Journal of Differential Equations*, *7*(1), 92–110.

Motta, M., & Sartori, C. (2003). Minimum time with bounded energy, minimum energy with bounded time. *SIAM Journal on Control and Optimization*, *42*(3), 789–809.

Petrov, N. (1970). On the Bellman function for the time-optimal process problem. *Journal of Applied Mathematics and Mechanics*, *34*(5), 785–791.

Rockafellar, R. (1979). Clarke's tangent cones and the boundaries of closed sets in $\mathbb{R}^n$. *Nonlinear Analysis: Theory, Methods & Applications*, *3*(1), 145–154.

Soravia, P. (1993). Pursuit-evasion problems and viscosity solutions of isaacs equations. *SIAM Journal on Control and Optimization*, *31*(3), 604–623.

Veliov, V. M. (1997). Lipschitz continuity of the value function in optimal control. *Journal of Optimization Theory and Applications*, *94*(2), 335–363.

Wolenski, P. R., & Zhuang, Y. (1998). Proximal analysis and the minimal time function. *SIAM Journal on Control and Optimization*, *36*(3), 1048–1072.

7

DATA-DRIVEN SET APPROXIMATION FOR DETECTION OF TRANSPORTATION MODES

This chapter was published as a joint work with Nicolas Petit in the Proceeding of the European Control Conference 2020, under the title *Data-driven approximation of differential inclusions and application to detection of transportation modes* (Aubin-Frankowski & Petit, 2020).

Abstract This chapter applies the Support Vector Data Description (SVDD) algorithm to approximate the graph of differential inclusions. It is proven that Gaussian SVDD can recover any compact graph if a large enough dataset is available. This data-driven approach can be used to identify discrete-valued parameters of nonlinear dynamical systems with unknown input signal. For illustration, the presented method is applied here both on real and synthetic data for detection of transportation modes based on linear velocity measurements.

Résumé Ce chapitre applique l'algorithme de description des données par vecteurs supports (SVDD) pour approcher le graphe d'inclusions différentielles. Nous prouvons que le SVDD gaussien permet d'approcher tout graphe compact si un ensemble de données suffisamment important est disponible. Cette approche fondée sur des mesures empiriques peut être utilisée pour identifier des paramètres à valeur discrète de systèmes dynamiques non linéaires dont le signal d'entrée est inconnu. À titre d'illustration, la méthode présentée est appliquée ici à la fois sur des données réelles et synthétiques pour la détection des modes de transport à partir de mesures de vitesse linéaire.

7.1 INTRODUCTION

Numerous recent studies have envisioned extending Machine Learning (ML) techniques to the automatic control domain (Khargonekar and Dahleh, 2018; Lamnabhi-Lagarrigue et al., 2017). Profitable pairing has already been achieved for anomaly detection Laouti et al., 2014 and system identification Bagge Carlson, 2018, among others. This article follows this trend, and, for its part, relates one aspect of the vast question of nonlinear dynamical system identification to a ML task in the framework of set-valued analysis.

The studied problem is as follows. Consider $N_0 > 1$ forced dynamical systems, each denoted by (f_j, U_j), for some index $j = 1, ..., N_0$ that can be interpreted as a discrete-valued parameter (or as a *label* in the ML terminology). The systems have as governing equations

$$q'(t) = f_j(q(t), u(t))$$

where the state vector is $q \in \mathbb{R}^n$, with the particularity that the input signal $u(\cdot) \in U_j$ is unknown. Each set U_j is defined as the set of functions containing every input signal that is possible given the value of j. For example, without further restriction, the sets U_j can be subsets of some functional space (e.g. C_0, L_2) with bounded values in $\mathbb{R}^m$. As stated above, the forcing signal $u(\cdot)$ is unknown. Further, despite being unambiguously defined, the sets U_j are unknown as well. To account for this lack of information, the governing equations above are rewritten as differential inclusions:

$$\begin{aligned} q'(t) \in F_j(q(t)) &\triangleq \{f_j(q(t), u(t)) \,|\, u(\cdot) \in U_j\} \\ K_j &\triangleq \{(q, q') \,|\, q' \in F_j(q)\} \subset \mathbb{R}^n \times \mathbb{R}^n \end{aligned} \tag{7.1}$$

where the set-valued map F_j is identified with its graph[1] K_j.

In this article, it is assumed that for each value of j, some properly *labeled* values (samples) of $(q(\cdot), q'(\cdot))$ are available, forming N_0 datasets. We first build approximations of the graphs K_j from these data. Once these approximations are available, they can be used for parameter estimation when presented with unlabeled recordings of samples of $(q(\cdot), q'(\cdot))$. This central question here boils down to determining the possible values of j.

Classically (see e.g. Walter and Pronzato, 1997), most techniques in parameter estimation assume a smooth and known dependence of the measurements on the unknown i, and consider the data as (ordered) time series. In many cases, estimation is treated as an extended-state reconstruction problem, using for instance Kalman filters or state observers with unknown inputs. Numerous references develop techniques of

1. The graph is bounded whenever the state and its derivatives are bounded for $u(\cdot) \in U_j$.

this type (see Besançon, 2007 and references therein), however we do not make these assumptions. Consequently, the proposed approach is not to be mistaken up with extensions of the Kalman perspective to set-valued mappings (Brogliato and Heemels, 2009; Doris et al., 2008).

The framework advocated here considers the data as labeled clouds of points (q, q') [2], each cloud corresponding to an unknown subset of the unknown U_j. Graphically, constructing an approximation of K_j amounts to delineating a subset of $\mathbb{R}^{2n}$ based on the data (*learning step*). Once the learning step is achieved, when (new) unlabeled data become available, identifying j amounts to testing the membership of the new data to the N_0 learned subsets (*testing step*).

To define approximations of the sets $(K_j)_j$, we apply the Support Vector Data Description (SVDD Tax and Duin, 2004) algorithm. SVDD is a kernel method that computes a minimal enclosing ball around the data. The output of the algorithm is an indicator function, the evaluation of which allows to readily test membership. In this article, it is shown that the output function for the Gaussian kernel can estimate any compact set of $\mathbb{R}^{2n}$ (we refer to this property by "set-consistency"). Two indexes i_0 and i_1 can then be distinguished if the approximations of their K_j differ (which depends on both f_j and U_j). For applications, both learning and testing have to be computationally tractable and robust to noise. As highlighted in the article, SVDD satisfies both conditions. To show the applicability and relevance of the method, we consider the problem of detecting transportation modes. This topic has recently attracted much attention (Bolbol et al., 2012,Martin et al., 2017), especially using smartphone inertial data, with reference datasets such as Geolife Zheng et al., 2010.

The paper is organized as follows. In Section 7.2, the problem under consideration is stated, the SVDD algorithm is presented, and illustrated on GPS-measured linear velocity data for transportation mode detection. The main theoretical contribution (Proposition 7.1 on the set-consistency of Gaussian SVDD) is to be found in Section 7.3. Finally, a simulation is performed in Section 7.4 to discuss the numerical behavior of the algorithm.

2. If q' is not directly available from measurements, it can be estimated through filtering Levant, 1998; Luo et al., 2005, Kriging Vazquez and Walter, 2005, numerical differentiation Diop et al., 2000 or high gains observers Dabroom and Khalil, 1999 among other possibilities, possibly with a non-negligible level of noise.

7.2 APPROXIMATING DISCRETE SETS WITH SVDD

7.2.1 *Notations*

Let F be a closed set-valued map defining a differential inclusion (7.1). Its graph K can be represented by an indicator function ϕ, taking non-positive values only in K. We denote K_N by an approximation of K based on a labeled dataset $X_N \triangleq \{(q_i, q_i')\}_{i\leqslant N}$ composed of N couples $x_i = (q_i, q_i') \in \mathbb{R}^d$ with $d = 2n$.

7.2.2 *Theoretical framework of the SVDD algorithm*

The graph estimation problem for K can be tackled through the SVDD algorithm, which is a nonlinear version of the minimal enclosing ball problem (MEB Elzinga and Hearn, 1972). The latter defines the ball of smallest volume of $\mathbb{R}^d$ containing the set X_N. The center of the MEB and its radius are the solutions of the following convex optimization problem:

$$\min_{c\in\mathbb{R}^d, R\in\mathbb{R}} R^2 \ \text{ s.t. } \forall i \leqslant N,\ \|x_i - c\|_{\mathbb{R}^d} \leqslant R \tag{7.2}$$

The SVDD algorithm offers the flexibility and nonlinearity required to fit general sets (as shown later in Proposition 7.1). Rather than looking for a ball in $\mathbb{R}^d$ endowed with its usual topology, $\mathbb{R}^d$ is here embedded in a reproducing kernel Hilbert space (RKHS) $\mathcal{H}_k(\mathbb{R}^d)$ where one seeks a minimal enclosing ball. For clarity, we first sketch RKHS theory before describing further the SVDD problem and its solution.

Definition 7.1. An RKHS $\left(\mathcal{H}_k, (\cdot,\cdot)_{\mathcal{H}_k}\right)$ defined on a set X is a Hilbert space of real-valued functions on X such that there exists a reproducing kernel $k : X \times X \to \mathbb{R}$, i.e. a function satisfying:

- $\forall x \in X,\ k_x(\cdot) \in \mathcal{H}_k(X)$ where $k_x : \begin{cases} X \to \mathbb{R} \\ y \mapsto k(x,y) \end{cases}$
- $\forall x \in X,\ \forall f \in \mathcal{H}_k(X),\ f(x) = (f, k_x)_{\mathcal{H}_k}$

The following summarized fundamental characterization allows for all the computations performed below.

Theorem 7.1 (Aronszajn, 1950)**.** *If a Hilbert space of functions on X is an RKHS $\mathcal{H}_k(X)$, then $k(\cdot,\cdot)$ is a positive definite kernel, i.e. a kernel being both:*
-positive: $\forall m,\ \forall (a_i, x_i) \in (\mathbb{R}\times X)^m,\ \sum a_i a_j k(x_i, x_j) \geqslant 0$
-and symmetrical: $\forall x, y \in X,\ k(x,y) = k(y,x)$

Conversely a positive definite kernel k *on* X *is reproducing for a unique* $\mathcal{H}_k(X)$*. Or, equivalently, there exists a Hilbert space* $\mathcal{H}$ *of real-valued functions and an embedding* $\Phi_k : X \to \mathcal{H}$ *s.t.* $\forall x, x' \in X,\ k(x, x') = (\Phi_k(x), \Phi_k(x'))_{\mathcal{H}}$.

Choosing a positive definite kernel k on X, one can represent any point $x \in X$ as a function $k_x(\cdot) \in \mathcal{H}_k(X)$. Some classical examples of positive definite kernels on $\mathbb{R}^d$ include the Gaussian, Laplacian and linear kernels:

$$k_\sigma(x, y) = \exp\left(-\|x - y\|^2_{\mathbb{R}^d}/(2\sigma^2)\right) \text{ for } \sigma > 0, \tag{7.3}$$

$$k_\lambda(x, y) = \exp(-\lambda\|x - y\|_{\mathbb{R}^d}) \text{ for } \lambda \geqslant 0 \tag{7.4}$$

$$k_{lin}(x, y) = (x, y)_{\mathbb{R}^d} \tag{7.5}$$

7.2.3 *The SVDD algorithm*

The SVDD algorithm recasts the MEB problem (7.2) in an RKHS[3] and considers

$$\min_{f \in \mathcal{H}_k, R \in \mathbb{R}} R^2 \ \text{ s.t. } \forall i \leqslant N,\ \|k(x_i, \cdot) - f(\cdot)\|_{\mathcal{H}_k} \leqslant R \tag{7.6}$$

As shown in Tax and Duin, 2004, the solution (f_k, R_k) of (7.6) is of the form $f_k(\cdot) \triangleq \sum_{i=1}^N \alpha_i k(x_i, \cdot)$, where $(\alpha_i)_i$ solves the dual problem of (7.6), which can be solved through quadratic programming:

$$\min_{\alpha \in \mathbb{R}^N_+} \alpha^T G \alpha - \alpha^T \operatorname{diag}(G) \ \text{ s.t. } \sum_{i=1}^N \alpha_i = 1 \tag{7.7}$$

where the matrix G is the Gram matrix of the $(x_i)_{i \leqslant N}$ (i.e. $G \triangleq (k(x_i, x_j))_{i,j \leqslant N}$) and $\operatorname{diag}(G)$ is the diagonal matrix extracted from G. Based on f_k, one can readily define a membership function ϕ_k on $\mathbb{R}^d$:

$$\begin{aligned} \phi_k(x) &\triangleq \|k_x(\cdot) - f_k(\cdot)\|^2_{\mathcal{H}_k} - R_k^2 \\ &= k(x, x) - R_k^2 + \sum_{i,j \leqslant N} \alpha_i \alpha_j k(x_i, x_j) - 2 \sum_{i \leqslant N} \alpha_i k(x_i, x) \end{aligned} \tag{7.8}$$

In turn, the function ϕ_k defines the sought-after closed set K_N:

$$K_N \triangleq \{x \in \mathbb{R}^d \mid \phi_k(x) \leqslant 0\}$$

Interestingly, the complementarity slackness of the Karush–Kuhn–Tucker (KKT) conditions of (7.7) ensure that for x_i interior to K_N, we have $\alpha_i = 0$. Thus, except for the

3. The MEB in $\mathbb{R}^d$ corresponds to SVDD with the linear kernel k_{lin}.

so-called *support vectors* x_i on the boundary of K_N, most coefficients are null in (7.8), which leads to a sparse representation of the set K_N by means of an indicator function that is quick to evaluate. By construction, $X_N \subset K_N$.

7.2.4 *Extension of SVDD to noisy data*

Because outliers can be present in the data [4], not all the points should be included in the estimate K_N. Following Tax and Duin, 2004, problem (7.6) can be adapted by introducing some slack variables $\xi_i \geqslant 0$:

$$\min_{\substack{f\in\mathcal{H}_k,\, R\in\mathbb{R},\, \xi\in\mathbb{R}^N_+ \\ \forall i\leqslant N, \|k(x_i,\cdot)-f(\cdot)\|^2_{\mathcal{H}_k}\leqslant R^2+M+\xi_i}} \left(R^2 + \frac{1}{\nu N}\sum_{i=1}^{N}\xi_i \right) \tag{7.9}$$

where the parameter $\nu \in]0,1]$ and the margin $M \in \mathbb{R}$ allow to adjust the level of conservatism. The dual problem of (7.9) writes in a form similar to (7.7):

$$\min_{\substack{\alpha\in\mathbb{R}^N \\ \forall i\leqslant N,\, \frac{1}{\nu N}\geqslant\alpha_i\geqslant 0}} \alpha^T G\alpha - \alpha^T\cdot \mathrm{diag}(G) \;\text{ s.t. }\; \sum_{i=1}^{N}\alpha_i = 1 \tag{7.10}$$

The value of the parameter ν can be related to the quantiles and minimum volume sets of the distribution of points. For $1 \geqslant \nu N$, every point is considered as a true point. Moreover the solution α is piecewise linear in ν as shown in Sjöstrand and Larsen, 2006, and ν is an upper bound on the proportion of outliers Schölkopf et al., 2001. Some sparsity is necessarily lost as the set of support vectors is enlarged compared to the noiseless case.

7.2.5 *Discussion on application of SVDD and illustration on real data*

Among the available set approximation procedures, SVDD has many pros making it a suitable choice for the problem of identification of control systems, among which the detection of transportation modes, further exposed in Section 7.4. The sparsity of the solution, mentioned in § 7.2.3, holds for all support vector machines and allows for the quick computation of the indicator functions (7.8), irrespective of the offline training cost. SVDD is non-parametric (with the obvious exception of the choice of kernels) as

4. especially if the filtering employed to generate the q' samples are dealing with missing or incorrectly time-stamped data

the model class is a Hilbert space of functions, so it can be applied to compare systems with the same variables but vastly different governing equations.

However, SVDD has some classical pitfalls. Some of them have been circumscribed as follows: limiting the effect of noise on the training step by taking out sheer outliers through an L0-penalty El Azami et al., 2014; accelerating the computation for online training Laskov et al., 2006; performing interpretation of the output f_k by converting the membership scores to probabilities El Azami et al., 2017. Interestingly, none of these problems occurs in our context. Although, the considered time-series are ridden with noise, sheer anomalies can be smoothed out through continuity assumptions bearing on the underlying dynamics. In contrast, the low-power additive noise, due to sensors or to numerical differentiation, has little impact on the SVDD output. Below, we used a Savistky-Golay filter to smooth the data $q(\cdot)$ and compute its derivative.

Moreover, the typical applications of travel mode detection are in low dimension, and allow for offline computation of the boundaries, taking advantage of the online efficiency. As an illustration, anticipating on further studies conducted in Section 7.4, SVDD is applied to a dataset provided in Martin et al., 2017, where the GPS-measured linear velocity was recorded at 1Hz for a variety of transportation modes. The linear acceleration was computed differentiating the velocity after smoothing the data (to avoid problems related to accelerometer measurements). On Fig.7.1, the five modes of transportation are represented by their SVDD envelopes in the (speed-acceleration) phase space (v, a). The envelopes overlap, so it is only when a trajectory leaves a set that it is made clear it does not belong to the corresponding class. To resolve ambiguities, one can for instance assign a trajectory to the class of smallest volume in $\mathbb{R}^d$ among the compatible ones, as it is the class that is easiest to leave. Developing such ambiguity resolution techniques is left out of the discussion and is postponed to future work.

In this example, SVDD was computed with Gaussian kernels k_σ, which variances σ^2 were adapted to each dynamic to get sharp estimates. However manipulating multiple σ entails considering functions in several RKHSs, with different norms. Furthermore, for every mode, the SVDD estimates are somewhat ill-behaved: when increasing σ, the boundary does not expand monotonically (in the inclusion sense). This calls for studies on comparison of multiple RKHSs and the increasingness property, which are the topic of the next section.

7.3 THEORETICAL GUARANTEES ON SET ESTIMATION

Let $X_N \triangleq \{x_i\}_{i \leqslant N}$ be a finite set of distinct points of X, k be a positive definite kernel on X and $\mathcal{H}_k(X)$ be its associated RKHS. In studies on RKHSs, the kernel is usually

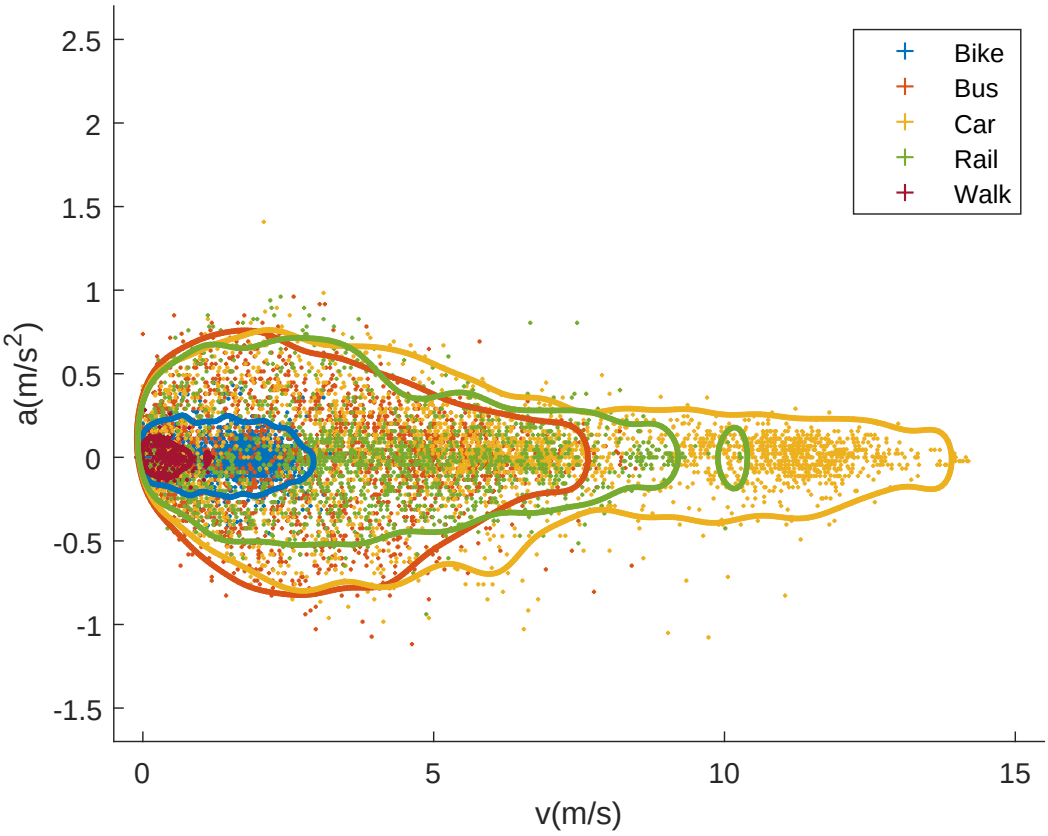


Figure 7.1 – Graph approximations obtained on a reference dataset Martin et al., 2017. The envelopes of various modes of transportation are computed through SVDD in the (speed-acceleration) phase space (v, a) with parameter $\nu = 5\%$ and various Gaussian kernels of width $\sigma \in [0.1, 0.5]$

kept fixed. However, we seek some robustness of the estimate K_N with respect to the kernel. This is the path we explore below considering multiple kernels.

Notations: We denote by $B_{\mathbb{R}^d}(x, R)$ the closed ball in $\mathbb{R}^d$ of center x and radius R and by $B_k(f, R)$ the closed ball in $\mathcal{H}_k(X)$. For any subset K, the set $\mathrm{co}(K)$ designates its convex hull and ∂K stands for its boundary. We denote by $B_k^{SVDD} \triangleq B_k(f_k, R_k)$ the minimal enclosing ball of $\{k_{x_i}\}_i = \Phi_k(X_N)$ in $\mathcal{H}_k(X)$. Let $K_k^{SVDD} \triangleq \Phi_k^{-1}(B_k^{SVDD}) \subset X$.

We make two assumptions on the kernel k:

Assumption 7.1. $0 \notin \mathrm{co}(\{k_{x_i}(\cdot)\}_{i \leqslant N})$

Assumption 7.2. $\exists \kappa_k \in \mathbb{R},\ \forall x \in X,\ k(x, x) = \kappa_k$

Assumption 7.1 is verified for $X = \mathbb{R}^d$ by the Gaussian (7.3) and Laplacian kernels (7.4), due to the positivity of the functions, but not by the linear kernel (7.5).

Assumption 7.2 is verified for $X = \mathbb{R}^d$ by all the translation-invariant continuous positive definite kernels, which, owing to Bochner's theorem Bochner, 1933, are of the form $h(\|x-y\|_{\mathbb{R}^d})$ with h being the Fourier transform of a finite positive measure on $\mathbb{R}^d$.

7.3.1 *Main results*

We first obtain the following result for noiseless SVDD.

Lemma 7.1 (SVDD as an orthogonal projection). *Under Assumptions 7.1 and 7.2, the center f_k of B_k^{SVDD} is the orthogonal projection of 0 for the norm $\|\cdot\|_k$ onto $\text{co}(\{k_{x_i}\}_i)$. The solution (f_k, R_k) of the noiseless SVDD problem (7.6) satisfies:*

$$f_k \triangleq \sum_{i=1}^{N} \overline{\alpha}_i k_{x_i} = \underset{f \in \text{co}(\{k_{x_i}\}_i)}{\arg\min} \|f\|_k^2 \; ; \; R_k = \sqrt{\kappa_k - \|f_k\|_k^2} \tag{7.11}$$

Furthermore:

$$\forall x \in X, \; x \in K_k^{\text{SVDD}} \Leftrightarrow \|f_k\|_k^2 \leqslant f_k(x) \tag{7.12}$$

Proof. The Lagrangian $\mathcal{L}$ corresponding to (7.6) is

$$\mathcal{L}(f, \alpha, R) \triangleq R^2 + \sum_{i=1}^{N} \alpha_i \left(\|k_{x_i} - f\|_k^2 - R^2 \right) \tag{7.13}$$

The KKT conditions require the solution $(\overline{f}, \overline{R}, \overline{\alpha})$ to satisfy:

$$\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial R}(\overline{f}, \overline{R}, \overline{\alpha}) = 2R \left(1 - \sum_{i=1}^{N} \overline{\alpha}_i \right) \\
0 &= \frac{\partial \mathcal{L}}{\partial f}(\overline{f}, \overline{R}, \overline{\alpha}) = 2 \left(\overline{f} - \sum_{i=1}^{N} \overline{\alpha}_i k_{x_i} \right) \\
0 &\leqslant \overline{\alpha}_i \text{ and } 0 \leqslant \overline{R} \\
0 &= \overline{\alpha}_i \left(\|k_{x_i} - \overline{f}\|_k^2 - \overline{R}^2 \right)
\end{aligned}$$

Therefore, the solution $\overline{f}$ belongs to $\text{co}(\{k_{x_i}\}_i) \subset \mathcal{H}_k$. For the $p \in [\![1, N]\!]$ points $(k_{x_{i_j}})_{j \leqslant p} \in \partial B(\overline{f}, \overline{R})$ having a non-null coefficient α, we have:

$$\forall j \leqslant p, \; \|k_{x_{i_j}} - \overline{f}\|_k^2 = \overline{R}^2 = \kappa_k + \|\overline{f}\|_k^2 - 2(k_{x_{i_j}}, \overline{f})_k \tag{7.14}$$

Hence, the scalar product $(k_{x_{i_j}}, \bar{f})_k$ is constant w.r.t. j.

$$\forall j \leqslant p,\ (k_{x_{i_j}}, \bar{f})_k = (\sum_{j=1}^{p} \bar{\alpha}_j k_{x_{i_j}}, \bar{f})_k = \|\bar{f}\|_k^2$$

So $\bar{R}^2 = \kappa_k - \|\bar{f}\|_k^2$, and the constraints of (7.6) become:

$$\|k_{x_i} - \bar{f}\|_k^2 \leqslant \kappa_k - \|\bar{f}\|_k^2 \Leftrightarrow 0 \leqslant (k_{x_i} - \bar{f}, \bar{f})_k \tag{7.15}$$

Owing to Assumption 7.1 and as $\bar{f}$ lies in $\mathrm{co}(\{k_{x_i}\}_i)$, we can apply the Best Approximation Theorem: the constraints require $\bar{f}$ to be the orthogonal projection of 0 for $\|\cdot\|_k$ onto $\mathrm{co}(\{k_{x_i}\}_i)$. Conversely, $\bar{f}$ satisfies the constraints. Finally, (7.12) stems from the same calculus as (7.15). □

For the Gaussian kernel (7.3) we write for conciseness $\mathcal{H}_\sigma \triangleq \mathcal{H}_{k_\sigma}$ and $K_\sigma^{SVDD} \triangleq K_{k_\sigma}^{SVDD}$ (same for B_σ and Φ_σ). We recall the result of Steinwart et al., 2006, Corollary 3.14;

$$\begin{cases} \forall \sigma_2 \geqslant \sigma_1 > 0,\ \mathcal{H}_{\sigma_2}(\mathbb{R}^d) \subset \mathcal{H}_{\sigma_1}(\mathbb{R}^d) \\ \forall f \in \mathcal{H}_{\sigma_2}(\mathbb{R}^d),\ \|f\|_{\sigma_1} \leqslant \left(\frac{\sigma_2}{\sigma_1}\right)^{d/2} \|f\|_{\sigma_2} \end{cases} \tag{7.16}$$

from which we derive the following result.

Lemma 7.2 (Norm inequalities for Gaussian kernels). *Let $\sigma_2 \geqslant \sigma_1 > 0$. The optimal f_{σ_1}, f_{σ_2} defined in Lemma 7.1 satisfy the inequality:*

$$\left(\frac{\sigma_2}{\sigma_1}\right)^d \|f_{\sigma_1}\|_{\sigma_1}^2 \geqslant \|f_{\sigma_2}\|_{\sigma_2}^2 \geqslant \frac{1}{N} \tag{7.17}$$

Proof. Let $\sigma_2 \geqslant \sigma_1 > 0$, and $f_{\sigma_2} = \sum_{j=1}^{N} \alpha_{i,\sigma_2} k_{\sigma_2, x_i}$ be as in (7.11). Define analogously $(\alpha_{i,\sigma_1})_i$. Below, we use the classical result that

$$\frac{1}{N} = \min_{\{\alpha_i \geqslant 0 \mid \sum_{i \leqslant N} \alpha_i = 1\}} \sum_{i \leqslant N} \alpha_i^2$$

and, as the Gaussian kernel has positive values and $\alpha_{i,\sigma_2} \geqslant 0$, we lower bound $\|f_{\sigma_2}\|_{\sigma_2}^2$ by the diagonal terms:

$$\|f_{\sigma_2}\|_{\sigma_2}^2 = \sum_{i,j \leqslant N} \alpha_{i,\sigma_2} \alpha_{j,\sigma_2} k_{\sigma_2}(x_i, x_j) \geqslant \sum_{i \leqslant N} \alpha_{i,\sigma_2}^2 \geqslant \frac{1}{N}$$

Set $\tilde{f}_{\sigma_1} = \sum_{j=1}^{N} \alpha_{i,\sigma_1} k_{\sigma_2,x_i}$. Owing to Lemma 7.1, f_{σ_2} is the element of smallest norm $\|\cdot\|_{\sigma_2}$ in $\mathrm{co}(\{k_{\sigma_2,x_i}\}_i)$. Hence

$$\|\tilde{f}_{\sigma_1}\|_{\sigma_2}^2 \geqslant \|f_{\sigma_2}\|_{\sigma_2}^2$$

Then, after some calculus, and after introducing the dual RKHS space which we omit for brevity, one gets

$$\left(\frac{\sigma_2}{\sigma_1}\right)^d \|f_{\sigma_1}\|_{\sigma_1}^2 \geqslant \|\tilde{f}_{\sigma_1}\|_{\sigma_2}^2$$

This concludes the proof. □

Relation (7.16) states that the Gaussian RKHSs form an increasing sequence of embeddings when σ decreases. This is a first step to compare the indicator functions defined by (7.8). Furthermore the SVDD algorithm with the Gaussian kernel is "set-consistent", in the sense that with enough data it can recover any compact set of $\mathbb{R}^d$:

Proposition 7.1 (Main result: Set-consistency of Gaussian SVDD). *The estimate* K_σ^{SVDD} *of* X_N *by the SVDD algorithm for Gaussian kernels satisfies the following two properties*

$$\begin{aligned} &i): \forall \epsilon > 0,\ \exists \sigma_0 > 0,\ s.t. \\ &\quad \forall 0 < \sigma \leqslant \sigma_0,\ K_\sigma^{SVDD} \subset X_N + B_{\mathbb{R}^d}(0,\epsilon) && (7.18) \\ &ii): \exists M > 0,\ \forall \sigma > 0,\ K_\sigma^{SVDD} \subset X_N + B_{\mathbb{R}^d}(0,M) && (7.19) \end{aligned}$$

Proof. Let $\epsilon > 0$, $\sigma > 0$ and $x \in K_\sigma^{SVDD}$. We combine (7.12) and (7.17):

$$\begin{aligned} \frac{1}{N} \leqslant \|f_\sigma\|_\sigma^2 \leqslant f_\sigma(x) &= \sum_{i \leqslant N} \alpha_{i,\sigma} e^{-\|x-x_i\|_{\mathbb{R}^d}^2/(2\sigma^2)} \\ &\leqslant e^{-\min_i \|x-x_i\|_{\mathbb{R}^d}^2/(2\sigma^2)} \end{aligned}$$

Thus, $\min_i \|x-x_i\|_{\mathbb{R}^d}^2 \leqslant 2\sigma^2 \ln N$. Set $\sigma_0 \triangleq \epsilon/\sqrt{2\ln N}$. Therefore, for any $\sigma \in]0,\sigma_0]$, we have that $K_\sigma^{SVDD} \subset X_N + B_{\mathbb{R}^d}(0,\epsilon)$. Fix $M \triangleq \sup_{i,j \leqslant N} \|x_i - x_j\|_{\mathbb{R}^d}$ and $f_\sigma = \sum_{i \leqslant N} \alpha_i k_{x_i}$ as in (7.11). Let $y \in \mathbb{R}^d \setminus (X_N + B_{\mathbb{R}^d}(0,M))$, then:

$$\begin{aligned} \forall i,j \leqslant N,\ k_\sigma(y,x_i) &= e^{-\|y-x_i\|_{\mathbb{R}^d}^2/(2\sigma^2)} \\ &< e^{-M^2/(2\sigma^2)} \\ &\leqslant e^{-\|x_i-x_j\|_{\mathbb{R}^d}^2/(2\sigma^2)} = k_\sigma(x_i,x_j) \\ \forall i \leqslant N,\ \sum_{i \leqslant N} \alpha_i k_\sigma(y,x_i) &= (\sum_{j \leqslant N} \alpha_j)(\sum_{i \leqslant N} \alpha_i k_\sigma(y,x_i)) \\ &< \sum_{i,j \leqslant N} \alpha_i \alpha_j k_\sigma(x_i,x_j) = \|f_\sigma\|_\sigma^2 \end{aligned}$$

We conclude from (7.12) that $y \notin K_\sigma^{SVDD}$. This yields $K_\sigma^{SVDD} \subset X_N + B_{\mathbb{R}^d}(0, M)$. □

Relations (7.18) and (7.19) show that the sequence $(K_\sigma^{SVDD})_{\sigma>0}$ is bounded and, for σ small enough, lies in a neighborhood of X_N for the norm of $\mathbb{R}^d$. In the limit case, when N tends to infinity, if X_∞ is dense in a given compact $K \subset \mathbb{R}^d$, then K_σ^{SVDD} is dense as well and lies in a neighborhood of K.

7.3.2 *Further results*

Numerical experiments confirm that the sequence $(K_\sigma^{SVDD})_{\sigma>0}$ of sets produced by the SVDD algorithm for Gaussian kernels is not increasing with σ w.r.t the inclusion. The increasingness property of a sequence of sets is akin to a "stability property" in the sense that the predictions drawn for a "large" kernel should contain the predictions obtained for "narrower" kernels.

As an extension, for a given kernel k_σ, we may look for a modified ball $B_\sigma(\tilde{f}_\sigma, \tilde{R}_\sigma)$, that should at least contain the set $\{k_{x_i}\}_i$. We may also wish for the modified ball to include $\Phi_k(K_\sigma^{SVDD})$, i.e. the set of all the k_x that are in the minimal enclosing ball of $\{k_{x_i}\}_i$ in $\mathcal{H}_\sigma$. One possibility is to keep the center fixed and to expand the radius accordingly. We first state a proposition for general kernels.

Proposition 7.2. *Let k_1 and k_2 be two positive definite kernels on X, satisfying Assumptions 7.1 and 7.2, and such that for some $\gamma > 0$ the kernel $\gamma^2 k_1 - k_2$ is positive definite (or equivalently that $\mathcal{H}_{k_2}(X) \subset \mathcal{H}_{k_1}(X)$). Then:* $\mathrm{co}(\{k_2(x_i, \cdot)\}_i) \subset B_{k_2}(f_{k_2}, R_{k_2}) \subset B_{k_1}(f_{k_2}, \gamma R_{k_2}) \subset \mathcal{H}_{k_1}(X)$.

Proof. From Saitoh and Sawano, 2016, Th 2.17, we deduce that the existence of $\gamma > 0$ s.t. the kernel $\gamma^2 k_1 - k_2$ is positive definite is equivalent to the inclusion $\mathcal{H}_{k_2}(X) \subset \mathcal{H}_{k_1}(X)$, the identity being continuous, of norm smaller than γ. Therefore we have $B_{k_2}(f_{k_2}, R_{k_2}) \subset B_{k_1}(f_{k_2}, \gamma R_{k_2}) \subset \mathcal{H}_{k_1}(X)$ and, by definition of the minimal enclosing ball coupled with the triangular inequality, we obtain $\mathrm{co}(\{k_2(x_i, \cdot)\}_i) \subset B_{k_2}(f_{k_2}, R_{k_2})$. This completes the proof. □

Corollary 7.1 (Increasingness of σ-concentric SVDD)**.** *Let $\sigma_0 > 0$, then the sequence $(\Phi_\sigma^{-1}(B_\sigma(f_{\sigma_0}, (\sigma/\sigma_0)^{d/2} R_{\sigma_0})))_{\sigma \geqslant \sigma_0}$ is increasing w.r.t. the inclusion when σ increases.*

Proof. Inequality (7.17) implies that the Gaussian kernel satisfies the assumptions of Proposition 7.2 with $\gamma^2 = (\sigma/\sigma_0)^d$. This proof requires as well introducing dual RKHSs. □

7.4 APPLICATION TO DETECTION OF TRANSPORTATION MODE ON SIMULATED DATA

In this section, we illustrate the practicability of SVDD for approximation of differential inclusions. We consider a general model (7.20) representing two types of vehicles: cars and bikes. In this model, the input exerted onto the vehicles is such that it produces asymptotic tracking of a reference velocity signal $v_{req}(\cdot)$ stemming from an urban-part of the NEDC cycle (New European Driving Cycle) generating a small set ($N \approx 500$) and a large set ($N \approx 15000$) of points. The chosen parameters are listed in Table 7.1 along with the magnitude of the speed.

$$m\dot{v}(t) = -kv^2(t) + u(t) \tag{7.20}$$
$$\text{where } u(t) \triangleq \begin{cases} -F_{max} \text{ if } k_p(v_{req}(t) - v(t)) < -F_{max} \\ F_{max} \text{ if } k_p(v_{req}(t) - v(t)) > F_{max} \\ k_p(v_{req}(t) - v(t)) \text{ otherwise} \end{cases}$$

	m	k	k_p	F_{max}	$\max(v_{req})$
Car	1 T	0.27	20	2 kN	80 km/h
Bike	100 kg	0.5	20	30 N	30 km/h

Table 7.1 – List of parameters of the NEDC simulation

We apply the SVDD algorithm to the pairs $(v, a) \triangleq (v, \dot{v})$ with a Gaussian kernel with a diagonal co-variance matrix $\sigma_v = 1.8$, $\sigma_a = 0.185$. There are between 10 and 30 support vectors in all cases, the learning step computation time ranging from 0.5 sec. to 25 sec. Fig.7.2a and 7.2b show the tightness of our estimate for noiseless data. As soon as enough points have been treated (Fig.7.2b), the SVDD outputs approximate well the theoretical boundaries, i.e. the graph of the differential inclusion (7.1) (which can be derived from (7.20) using a very general v_{req}). Furthermore the estimate is robust to 10 % outliers drawn in a uniform fashion (which results in a relatively strongly corrupted dataset compared to the typical noise present in velocity measurements and the subsequent filtered derivatives), by setting $\nu = 10$ % in (7.9) (Fig.7.2d) according to the guidelines recalled in §7.2.4. As shown in Corollary 7.1, the σ-concentric SVDD balls generate an increasing sequence that rapidly expands (Fig.7.2e). To test the membership of a trajectory to either class (Fig.7.2c), one can check the signs of the functions (7.8). When a function takes a positive value, the trajectory has violated the empirical boundary and thus does not belong to the corresponding

class (Fig.7.2f). Using a representative test trajectory, the car can be easily identified after approximately 1-2 min of measurement.

7.5 CONCLUSION AND PERSPECTIVES

The example of detection of transportation modes highlights the relevance for parameter estimation of a set-valued framework involving SVDD, a kernel method. Our investigation has shown the consistency of the algorithm for Gaussian kernels, as it allows to recover any set-valued map with bounded graph. The dependence of SVDD on its parameters was also discussed. Future work will concern identification of continuous parameters, such as mass estimation as is considered in weigh-in-motion applications, a key technology for improving ground transportation safety Jacob and Feypell-de La Beaumelle, 2010. This problem is more complex that the one treated here. We believe it may require to incorporate the sequentiality of time series into the above framework, which would speed up mode detection as well.

ACKNOWLEDGMENTS

The authors express their gratitude to Stéphane Canu for the enlightning discussions and for the code provided to compute efficiently the SVDD algorithm.

BIBLIOGRAPHY FOR CHAPTER 7

Aronszajn, N. (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society*, *68*(3), 337–337.

Bagge Carlson, F. (2018). *Machine learning and system identification for estimation in physical systems* (Doctoral dissertation). Department of Automatic Control. Department of Automatic Control, Faculty of Engineering LTH, Lund University.

Besançon, G. (2007). *Nonlinear observers and applications, volume 363, chapter an overview on observer tools for nonlinear systems*. Springer.

Bochner, S. (1933). Monotone funktionen, stieltjessche integrale und harmonische analyse. *Mathematische Annalen*, *108*(1), 378–410.

Bolbol, A., Cheng, T., Tsapakis, I., & Haworth, J. (2012). Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification. *Computers, Environment and Urban Systems*, *36*(6), 526–537.

Brogliato, B., & Heemels, M. (2009). Observer design for lure systems with multivalued mappings: a passivity approach. *IEEE Transactions on Automatic Control*, *54*(8), 1996–2001.

Dabroom, A. M., & Khalil, H. K. (1999). Discrete-time implementation of high-gain observers for numerical differentiation. *International Journal of Control*, *72*(17), 1523–1537.

Diop, S., Grizzle, J. W., & Chaplais, F. (2000). On numerical differentiation algorithms for nonlinear estimation. *IEEE Conference on Decision and Control*, *2*, 1133–1138.

Doris, A., Juloski, A. L., Mihajlovic, N., Heemels, M., van de Wouw, N., & Nijmeijer, H. (2008). Observer designs for experimental non-smooth and discontinuous systems. *IEEE Transactions on Control Systems Technology*, *16*(6), 1323–1332.

El Azami, M., Lartizien, C., & Canu, S. (2014). Robust outlier detection with Lo-SVDD. *22th European Symposium on Artificial Neural Networks, ESANN 2014, Bruges, Belgium, April 23-25, 2014*.

El Azami, M., Lartizien, C., & Canu, S. (2017). Converting SVDD scores into probability estimates: application to outlier detection. *Neurocomputing*, *268*, 64–75.

Elzinga, D. J. ., & Hearn, D. W. (1972). The minimum covering sphere problem. *Management Science*, *19*(1), 96–104.

Jacob, B., & Feypell-de La Beaumelle, V. (2010). Improving truck safety: potential of weigh-in-motion technology. *IATSS Research*, *34*(1), 9–15.

Khargonekar, P. P., & Dahleh, M. A. (2018). Advancing systems and control research in the era of ML and AI. *Annual Reviews in Control*, *45*, 1–4.

Lamnabhi-Lagarrigue, F., Annaswamy, A., Engell, S., Isaksson, A., Khargonekar, P., Murray, R. M., Nijmeijer, H., Samad, T., Tilbury, D., & Van den Hof, P. (2017). Systems & control for the future of humanity, research agenda: current and future roles, impact and grand challenges. *Annual Reviews in Control*, *43*, 1–64.

Laouti, N., Othman, S., Alamir, M., & Sheibat-Othman, N. (2014). Combination of model-based observer and support vector machines for fault detection of wind turbines. *International Journal of Automation and Computing*, *11*(3), 274–287.

Laskov, P., Gehl, C., Krüger, S., & Müller, K.-R. (2006). Incremental support vector learning: analysis, implementation and applications. *Journal of Machine Learning Research*, *7*, 1909–1936.

Levant, A. (1998). Robust exact differentiation via sliding mode technique. *Automatica*, *34*(3), 379–384.

Luo, J., Ying, K., He, P., & Bai, J. (2005). Properties of Savitzky-Golay digital differentiators. *Digital Signal Processing*, *15*(2), 122–136.

Martin, B., Addona, V., Wolfson, J., Adomavicius, G., & Fan, Y. (2017). Methods for real-time prediction of the mode of travel using smartphone-based GPS and accelerometer data. *Sensors*, *17*(9), 2058.

Saitoh, S., & Sawano, Y. (2016). *Theory of reproducing kernels and applications*. Springer Singapore.

Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, *13*(7), 1443–1471.

Sjöstrand, K., & Larsen, R. (2006). The entire regularization path for the support vector domain description. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, *9*, 241–248.

Steinwart, I., Hush, D., & Scovel, C. (2006). An explicit description of the reproducing kernel hilbert spaces of gaussian RBF kernels. *IEEE Transactions on Information Theory*, *52*(10), 4635–4643.

Tax, D. M. J., & Duin, R. P. W. (2004). Support vector data description. *Machine Learning*, *54*(1), 45–66.

Vazquez, E., & Walter, E. (2005). Estimating derivatives and integrals with Kriging. *Proceedings of the 44th IEEE Conference on Decision and Control*, 8156–8161.

Walter, E., & Pronzato, L. (1997). *Identification of parametric models: from experimental data)*. Springer.

Zheng, Y., Xie, X., & Ma, W.-Y. (2010). Geolife: a collaborative social networking service among user, location and trajectory. *IEEE Data Engineering Bulletin*, *33*(2), 32–39.

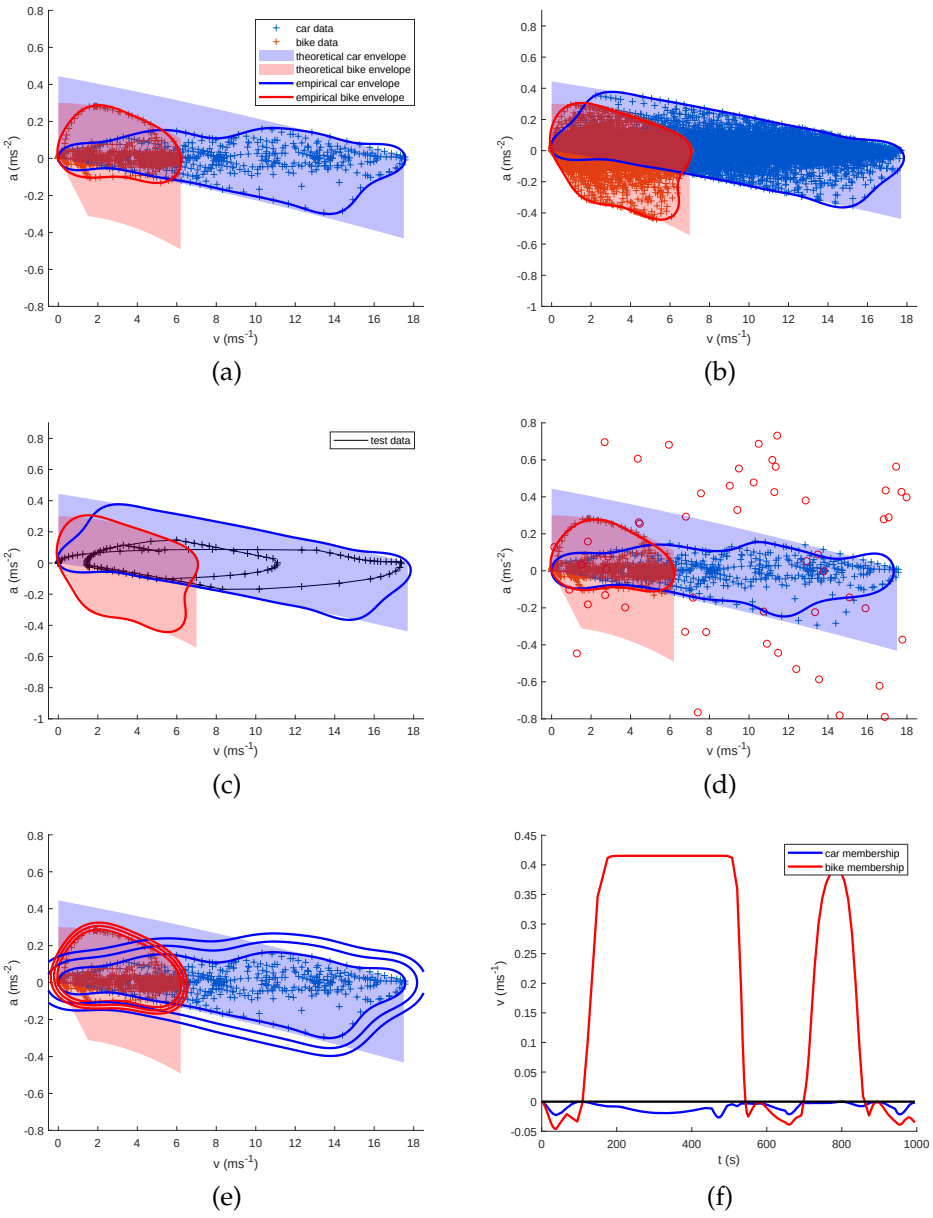


Figure 7.2 – Estimate sets K_σ^{SVDD} of theoretical dynamical limits (filled areas) by the SVDD algorithm on simulation data. (a): noiseless case with a small number of data (~500 pts). (b): noiseless case with a large number of data (~15k pts) (c): a test trajectory is compared to the SVDD boundaries. (d): case with noise on the small data and 10% uniformly distributed outliers. The boundaries are mildly altered. (e): applying σ-concentric SVDD on the noiseless small data with varying σ (100%,102%,104% of the previous σ).(f): the indicator functions show that the test trajectory of (c) is a car as the car values are negative while some bike values are positive.

RÉSUMÉ

Les contraintes ponctuelles d'état et de forme en théorie du contrôle et en estimation non-paramétrique sont difficiles à traiter car elles impliquent souvent un problème d'optimisation convexe en dimension infinie avec un nombre infini de contraintes d'inégalité. La satisfaction de ces contraintes est essentielle dans de nombreuses applications, telles que la planification de trajectoires ou la régression quantile jointe. Or, les noyaux reproduisants sont un choix propice aux évaluations ponctuelles. Cependant les théorèmes de représentation qui en sous-tendent les applications numériques ne peuvent pas être appliqués à un nombre infini d'évaluations. Par des arguments algébriques et géométriques constructifs, nous prouvons qu'un nombre infini de contraintes affines à valeur réelle sur les dérivées des fonctions peut être surcontraint par un nombre fini de contraintes coniques du second ordre si l'on considère des espaces de Hilbert à noyau reproduisant de fonctions à valeurs vectorielles. Nous montrons que le contrôle optimal linéaire-quadratique (LQ) sous contraintes d'état est une régression sous contraintes de forme sur l'espace de Hilbert de trajectoires contrôlées linéairement. Cet espace est défini par un noyau LQ explicite lié à la matrice de Riccati. L'efficacité de notre approche est illustrée par divers exemples issus de la théorie du contrôle linéaire et de l'estimation non-paramétrique. Enfin, nous énonçons des résultats pour des inclusions différentielles générales dans des problèmes de temps minimal et d'identification du graphe de la correspondance. Surtout nous faisons ressortir un lien nouveau entre méthodes à noyaux et contrôle optimal en identifiant le noyau hilbertien des espaces de trajectoires contrôlées linéairement.

MOTS CLÉS

Méthodes à noyaux, Contrôle optimal, Contraintes de formes, Contraintes d'état, Inférence de dynamiques.

ABSTRACT

Pointwise state and shape constraints in control theory and nonparametric estimation are difficult to handle as they often involve convex optimization problem with an infinite number of inequality constraints. Satisfaction of these constraints is critical in many applications, such as path-planning or joint quantile regression. Reproducing kernels are propitious for pointwise evaluations. However representer theorems, which ensure the numerical applicability of kernels, cannot be applied for an infinite number of evaluations. Through constructive algebraic and geometric arguments, we prove that an infinite number of affine real-valued constraints over derivatives of the model can be tightened into a finite number of second-order cone constraints when looking for functions in vector-valued reproducing kernel Hilbert spaces. We show that state-constrained Linear-Quadratic (LQ) optimal control is a shape-constrained regression over the Hilbert space of linearly-controlled trajectories defined by an explicit LQ kernel related to the Riccati matrix. The efficiency of the developed approach is illustrated on various examples from both linear control theory and nonparametric estimation. Finally, we provide some results for general differential inclusions in minimal time problems and identification of the graph of the set-valued map. Most of all we bring to light a novel connection between reproducing kernels and optimal control theory, identifying the Hilbertian kernel of linearly controlled trajectories.

KEYWORDS

Kernel methods, Optimal control theory, Shape constraints, State constraints, Vehicle dynamics inference.