2016-04-15

Multivariate Epi-Splines and Evolving Function Identification Problems

Royset, Johannes O.

J.O. Royset and R. J-B Wets, 2016, "Multivariate Epi-Splines and Evolving Function Identification Problems," Set-Valued and Variational Analysis, to appear. Erratum.
http://hdl.handle.net/10945/48687

MULTIVARIATE EPI-SPLINES AND EVOLVING FUNCTION IDENTIFICATION PROBLEMS*

Johannes O. Royset

Operations Research Department
Naval Postgraduate School
joroyset@nps.edu

Roger J-B Wets

Department of Mathematics
University of California, Davis
rjbwets@ucdavis.edu

Abstract. The broad class of extended real-valued lower semicontinuous (lsc) functions on $I\!R^n$ captures nearly all functions of practical importance in equation solving, variational problems, fitting, and estimation. The paper develops piecewise polynomial functions, called epi-splines, that approximate any lsc function to an arbitrary level of accuracy. Epi-splines provide the foundation for the solution of a rich class of function identification problems that incorporate general constraints on the function to be identified including those derived from information about smoothness, shape, proximity to other functions, and so on. As such extrinsic information as well as observed function and subgradient values often evolve in applications, we establish conditions under which the computed epi-splines converge to the function we seek to identify. Numerical examples in response surface building and probability density estimation illustrate the framework.

Keywords: function approximation, epi-splines, epi-convergence, infinite-dimensional optimization, shape restriction, constrained optimization

AMS Classification: 41A15 41A29 41A63 49J53 65D07 65D15 65K10

Date: April 15, 2015

*This material is based upon work supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant numbers 00101-80683, W911NF-10-1-0246 and W911NF-12-1-0273, and, in part, was carried out while the first author was on a sabbatical leave at the University of California, Department of Mathematics.

1 Introduction

The class of extended real-valued lower semicontinuous (lsc) functions on $I\!R^n$ includes nearly all functions of practical importance in the solution of systems of equations and inequalities, in optimization and variational problems, in statistical estimation, and in numerous applications related to data fitting and smoothing. They are neither required to be integrable nor defined on a compact set. The possibility of function values of $\pm\infty$ facilitates the representation of constrained minimization problems, where infeasible points are assigned the value ∞, and the utilization of nonlinear transformations, for example of the form $g(x) = \exp(f(x))$ where $f(x) = -\infty$ is required to allow for $g(x) = 0$. In this paper, we develop piecewise polynomial functions, defined on an arbitrary partition of $I\!R^n$, that approximate to any level of accuracy a lsc function. The approximations are named *epi-splines* due to their reliance on the epi-topology [4, 26] on the space of lsc functions and their obvious structural similarity with polynomial splines. The choice of lsc functions rather than upper semicontinuous (usc) functions is insignificant because a function f is usc whenever $-f$ is lsc. Consequently, all our results are trivially extended to usc functions.

We deviate from the usual constructions of splines (see, e.g., [24, 10, 27, 16, 13, 20]) by *not* insisting on smoothness, a property not universally satisfied by the lsc functions. Of course, the consideration of splines with less than the standard smoothness is not new, it originates with Curry and Schoenberg [8, 9]; see [11] for a recent study. Our construction on an arbitrary partition of the whole of $I\!R^n$ is novel, however. In fact, even for the one-dimensional special case, the approximation of lsc function on $I\!R$ is new; our previous study [30] dealt with compact intervals of $I\!R$.

Splines are intimately tied to optimization problems through their variational theory pioneered in [18, 12, 1] and further developed up to the present; see for example [21, 25, 6, 7, 5]. We also start with a problem of identifying a function that minimizes a criterion and satisfies given constraints, but depart from the traditional focus on smoothness, interpolation, and (least-squares) approximation. Motivated by applications in curve fitting, regression, probability density estimation, variogram computation, financial curve construction, and building of stochastic processes [31, 30], we instead consider nearly arbitrary criterion and constraints, which leads to a broad class of *function identification problems*. Although recent efforts focus on additional constraints, for example related to shape [32, 19, 15, 14, 22, 23, 17], our treatment is more general. We consider constraints derived from *data* as well as *extrinsic information* about the actual function to be identified, and give specific examples related to continuity, smoothness, convexity, monotonicity, function values, unimodality, integral values, subgradients, and proximity to another function. In contrast to the common setting where optimal solutions of specific function identification problems are polynomial splines, the more general problems, defined over the space of lsc functions as demanded by applications, require direct *construction* of approximating epi-splines.

We establish the notion of *information growth*, which extends the traditional study of limiting properties of splines as additional function and derivative information is acquired. We also consider incrementally arriving information about the criterion and constraints as well as approximations necessitated by algorithmic and computational issues. This leads to a rich class of *evolving* function identification problems and fundamental questions of approximation of optimization problems. We establish that epi-splines obtained as solutions of such evolving function identification problems tend to the actual function to be identified under general assumptions as more information becomes available and approximations are refined.

The contribution of the paper is therefore three-fold: (i) we construct piecewise polynomial approximations of lsc functions, (ii) we consider general function identification problems involving nearly arbitrary constraints, and (iii) we show that epi-splines obtained by solving evolving problems of this general kind approximate the actual function to be identified.

The paper proceeds in §2 by defining the broad class of function identification problems and providing background for the analysis. §3 constructs epi-splines and approximating function identification problems, and establishes their properties. §4 extends the results to approximating functions obtained by the composition of a monotonic function with an epi-spline. §5 discusses formulation of constraints in function identification problems, for example using extrinsic information about the actual function. The paper ends with numerical examples in §6. A summary of some results and an exposition of the framework of analysis are given in the forthcoming tutorial [31].

2 Function Identification Problems

A function identification problem of the general kind considered here aims to determine a function f on $I\!R^n$ that minimizes some criterion given by a functional ψ, while also satisfying constraints given by a set F. This "actual problem" takes the form

$$(FIP): \qquad \min \psi(f) \text{ such that } f \in F \subseteq \mathcal{F},$$

where $\mathcal{F}$ is a space of functions. We would like to handle applications with discontinuous functions that might incorporate implicit constraints represented by the function value ∞. This situation arises in building of response surfaces defined only on a subset set of $I\!R^n$, not known a priori, and that subsequently need to be minimized under constraints. Nonlinear transformations also demand extended real-valuedness of functions in (FIP). For example, a function f might be subject to the transformation $\exp(f)$ to ensure nonnegativity and/or better conditioning but then $f = -\infty$ is required to represent 0. Moreover, we would like to consider functions that are not integrable and that are defined on the whole of $I\!R^n$ or unbounded subsets. These requirements lead to the set of extended real-valued lower semi-continuous (lsc) functions from $I\!R^n$ to $\overline{I\!R} := I\!R \cup \{-\infty, \infty\}$, excluding $f \equiv \infty$, which we denote by lsc-fcns($I\!R^n$). We let $\mathcal{F} =$ lsc-fcns($I\!R^n$), or possibly a subset thereof. The choice of upper semi-continuous (usc) functions would yield similar results, but that development is omitted for simplicity of exposition[2].

We next review some pertinent facts regarding the space of lsc functions and the notion of epi-convergence that is central in the subsequent analysis; see [26, Section 7.I] for details.

We embed the space lsc-fcns($\mathbb{R}^n$) with the *epi-distance* $d\!l$, which for any $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$, not identically equal to ∞, is defined as

$$d\!l(f,g) := \int_0^\infty d\!l_\rho(f,g)e^{-\rho}d\rho,$$

where the *ρ-epi-distance*, $\rho \geq 0$, is given by

$$d\!l_\rho(f,g) := \max_{\|\bar{x}\|\leq\rho} |d(\bar{x}, \operatorname{epi} f) - d(\bar{x}, \operatorname{epi} g)|,$$

the standard distance between a point $\bar{x} = (x, x_0) \in \mathbb{R}^{n+1}$, with $x \in \mathbb{R}^n$ and $x_0 \in \mathbb{R}$ and a set $S \subset \mathbb{R}^{n+1}$ is given by

$$d(\bar{x}, S) := \inf_{\bar{y}\in S} \|\bar{x} - \bar{y}\|,$$

and the *epi-graph* of f is given by

$$\operatorname{epi} f := \{(x, x_0) \in \mathbb{R}^{n+1} \mid f(x) \leq x_0\},$$

with an analogous expression for $\operatorname{epi} g$; see Figure 1. We observe that $d\!l_\infty(f,g)$ coincides with the classical Pompeiu-Hausdorff distance between the sets $\operatorname{epi} f$ and $\operatorname{epi} g$. However, that notion is not appropriate in the present context as both $\operatorname{epi} f$ and $\operatorname{epi} g$ are unbounded and the "truncation" implied by a finite ρ in $d\!l_\rho(f,g)$ becomes essential.

The epi-distance is a metric on lsc-fcns($\mathbb{R}^n$) and induces the *epi-topology* (sometimes called the Attouch-Wets topology), which we use throughout the paper. In fact, by Theorem 7.58 of [26], (lsc-fcns($\mathbb{R}^n$), $d\!l$) is a complete metric space. As we see below, it is also separable and therefore a Polish space.

The space lsc-fcns($\mathbb{R}^n$) is not a vector space[3] as $-f \notin$ lsc-fcns($\mathbb{R}^n$) for $f \in$ lsc-fcns($\mathbb{R}^n$) given by $f(x) = 0$ if $x \in [0,\infty)^n$ and $f(x) = 1$ otherwise. However, it is a cone, i.e., $\lambda f \in$ lsc-fcns($\mathbb{R}^n$) whenever $\lambda \geq 0$ and $f \in$ lsc-fcns($\mathbb{R}^n$), and lsc-fcns($\mathbb{R}^n$) $\cup \{\infty\}$ (i.e., the inclusion of the function $f \equiv \infty$) is convex because $\lambda f + (1-\lambda)g \in$ lsc-fcns($\mathbb{R}^n$) $\cup \{\infty\}$ whenever $\lambda \in [0,1]$ and $f, g \in$ lsc-fcns($\mathbb{R}^n$) $\cup \{\infty\}$.

We say that functions $f^\nu : \mathbb{R}^n \to \overline{\mathbb{R}}$ *epi-converge* to a function $f^0 : \mathbb{R}^n \to \overline{\mathbb{R}}$ if $d\!l(f^\nu, f^0) \to 0$. By Theorem 7.58 in [26], f^ν epi-converges to f^0 if and only if their epi-graphs converge in the sense

[2] Passing to a "mirror" development for usc functions requires only adjusting some signs and inequalities, and replacing epi by hypo.

[3] We define addition of functions and multiplication with a scalar in the usual "pointwise" manner. To handle extended real values, we adopt the conventions that $\infty + a = \infty$ and $-\infty + a = -\infty$ for $a \in \mathbb{R}$, $\infty + \infty = \infty + (-\infty) = -\infty + \infty = \infty$, $\lambda\infty = -\infty$ for $\lambda < 0$, $\lambda\infty = \infty$ for $\lambda > 0$, $0\infty = 0$, and similarly for $\lambda(-\infty)$.

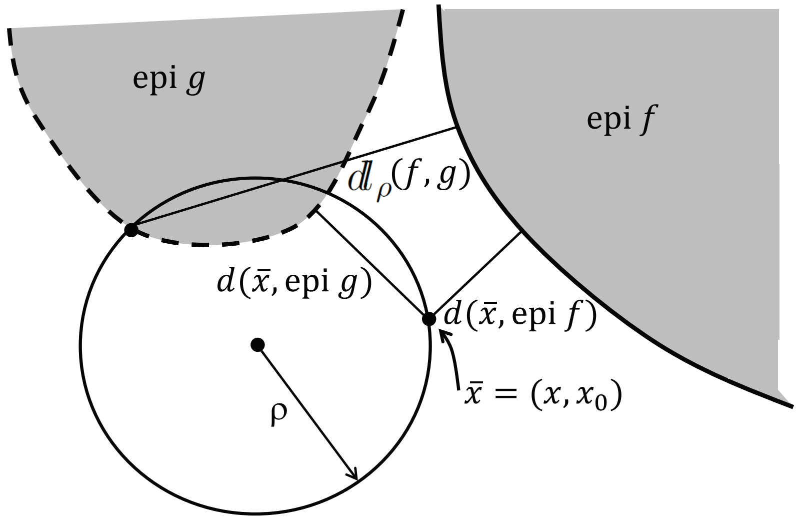


Figure 1: Epigraphs and ρ-epi-distance.

of Painlevé-Kuratowski[4]. An operationally convenient characterization of epi-convergence is resulting from this equivalence:

2.1 Proposition *[26, Theorem 7.2] Functions $f^\nu : I\!R^n \to \overline{I\!R}$ epiconverge to $f : I\!R^n \to \overline{I\!R}$ if and only if for every $x \in I\!R^n$,*

a) $\liminf_\nu f^\nu(x^\nu) \geq f(x)$, *for all* $x^\nu \to x$

b) $\limsup_\nu f^\nu(x^\nu) \leq f(x)$ *for some sequence* $x^\nu \to x$.

Occasionally, we consider the subset of lsc-fcns($I\!R^n$) consisting of functions that take the value ∞ beyond a closed subset B of $I\!R^n$, which we denote by lsc-fcns(B).

An actual problem (FIP) can only exceptionally be solved analytically and even computational approaches need to rely on approximations. Moreover, incremental arrival of data and extrinsic information, representing information growth about the underlying function, might necessitate the consideration of a family of evolving problems. The formulation and analysis of such evolving and approximating problems are the topic of the next section.

3 Evolving Problems and Approximations

In contrast to the classical variational theory of splines, where optimal solutions of certain problems automatically result in piecewise polynomial functions (see for example [7]), which cannot be expected in our general setting, we *construct* piecewise polynomial functions on an arbitrary partition of $I\!R^n$.

[4]We recall that the outer limit of a sequence of sets $\{A^\nu\}_{\nu\in N}$ is the collection of points y to which a subsequence of $\{y^\nu\}_{\nu\in N}$, with $y^\nu \in A^\nu$, converges. The inner limit is the points to which a sequence of $\{y^\nu\}_{\nu\in N}$, with $y^\nu \in A^\nu$, converges. If both limits exist and are identical, we say that the set is the Painlevé-Kuratowski limit of $\{A^\nu\}_{\nu\in N}$; see [4, 26].

The partition is only exceptionally related to the data available unlike the construction of classical univariate splines whose knot placement is often data driven. Our construction leads to epi-splines and fundamental approximation results given in the next subsection. The subsequent subsection turns to the broader question of evolving criteria and constraints driven by changes in approximations and information growth. We conclude the section with a series of estimates of the epi-distance between two epi-splines as well as a discussion of connections with other modes of convergence.

3.1 Epi-splines

Approximation of lsc functions is achieved by *lsc epi-splines* that are piecewise polynomial functions defined on a partition of $I\!R^n$, or a subset thereof. Specific details are given next. We use the notation $\text{cl}\, S$ to denote the closure of a subset S.

3.1 Definition (partition) *A finite collection $R_1, R_2, ..., R_N$ of open subsets of $I\!R^n$ is a* partition *of a closed set $B \subseteq I\!R^n$ if $\cup_{k=1}^N \text{cl}\, R_k = B$ and $R_k \cap R_l = \emptyset$ for all $k \neq l$.*

We observe that the definition abuses slightly standard terminology by having the subsets R_1, ..., R_N open and $\cup_{k=1}^N R_k \neq B$. However, the focus on open sets simplifies the following exposition.

We adopt a "total degree" convention and say that a polynomial in n dimensions is of *total degree* p if it is expressed as a finite sum of polynomial terms each having the sum of the powers of the variables being no larger than p. Another convention would have had only minor consequences for the following results. The set of all such polynomials is denoted by $\text{poly}^p(I\!R^n)$. We note that the total number of terms in such a polynomial is at most

$$n_p := (n+p)!/(n!p!).$$

Consequently, every polynomial in $\text{poly}^p(I\!R^n)$ is given by n_p real parameters. For any $f : I\!R^n \to \overline{I\!R}$, we adopt the notation

$$\liminf_{x' \to x} f(x') := \lim_{\delta \downarrow 0} \left[\inf_{x' \in I\!B(x,\delta)} f(x') \right],$$

where

$$I\!B(x,\delta) := \{x' \in I\!R^n \mid \|x' - x\| \leq \delta\}.$$

Informally, $\liminf_{x' \to x} f(x')$ is the smallest value of f near x. We let $\rho I\!B := I\!B(0,\rho)$ and also use the same notation for balls in other dimensions too as the meaning will be clear from the context. Moreover, $I\!N := \{1, 2, ...\}$ and $I\!N_0 := \{0\} \cup I\!N$.

3.2 Definition (lsc epi-splines) *A (lsc) epi-spline $s : I\!R^n \to \overline{I\!R}$ of order $p \in I\!N_0$, with partition $\mathcal{R} = \{R_k\}_{k=1}^N$ of a closed set $B \subseteq I\!R^n$, is a function that*

$$\begin{aligned} &\textit{on each } R_k,\ k = 1, ..., N, \textit{ is polynomial of total degree } p, \\ &\textit{has } s(x) = \infty \textit{ for } x \notin B, \textit{ and} \\ &\textit{for every } x \in I\!R^n, \textit{ has } s(x) = \liminf_{x' \to x} s(x'). \end{aligned}$$

The family of all such epi-splines is denoted by e-spl$^p_n(\mathcal{R})$.

We stress that B in Definition 3.2 might very well be the whole of $I\!R^n$ or some unbounded subset. As the name indicates, lsc epi-splines are indeed lsc functions, which trivially follows from the definition of such functions.

3.3 Proposition *For any partition $\mathcal{R}$ of a closed set $B \subseteq I\!R^n$, $p \in I\!N_0$, and $n \in I\!N$,* e-spl$^p_n(\mathcal{R}) \subset$ *lsc-fcns*$(B) \subseteq$ *lsc-fcns*$(I\!R^n)$.

We deal exclusively with lsc epi-splines and systematically drop "lsc" in the following. Since an $s \in$ e-spl$^p_n(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$, involves N polynomials of total degree p, it is fully characterized by

$$n_e := Nn_p = N(n+p)!/(n!p!) \text{ parameters.}$$

We next show that epi-splines of any order can approximate lsc functions to an arbitrary level of accuracy. However, this requires a refinement of the partition as follows:

3.4 Definition (infinite refinement) *A sequence $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ of partitions of a closed set $B \subseteq I\!R^n$, with $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$, is an infinite refinement if*

$$\text{for every } x \in B \text{ and } \varepsilon > 0, \text{ there exists } \bar{\nu} \in I\!N \text{ such that}$$
$$R_k^\nu \subset I\!B(x,\varepsilon) \text{ for every } \nu \geq \bar{\nu} \text{ and } k \text{ satisfying } x \in \operatorname{cl} R_k^\nu.$$

In the case of a compact B, there are obvious choices of infinite refinements. A simple example of an infinite refinement on (unbounded) $I\!R$ is to take $N^\nu = 2\nu + 2$, $R_1^\nu = (-\infty, -\sqrt{\nu})$, $R_k^\nu = ((k-\nu-2)/\sqrt{\nu}, (k-\nu-1)/\sqrt{\nu})$ for $k = 2, 3, ..., 2\nu+1$, and $R_{2\nu+2}^\nu = (\sqrt{\nu}, \infty)$. Then $\bar{\nu} > \max\{x^2, \varepsilon^{-2}\}$ satisfies the above condition. Obviously, much flexibility exists in constructing such infinite refinements. We now state a density result, which as elsewhere in the paper is with respect to the epi-topology.

3.5 Theorem (dense approximation) *For any $p \in I\!N_0$ and $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$, an infinite refinement of a closed set $B \subseteq I\!R^n$,*

$$\bigcup_{\nu=1}^{\infty} \text{e-spl}^p_n(\mathcal{R}^\nu) \text{ is dense in } \textit{lsc-fcns}(B).$$

Proof: Let $s^0 \in$ lsc-fcns(B) and $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$. It suffices to construct a sequence of epi-splines of order $p = 0$. For every $\nu \in I\!N$ and R_k^ν, $k = 1, 2, ..., N^\nu$, we define

$$\sigma^\nu(R_k^\nu) := \begin{cases} \inf_{x \in \operatorname{cl} R_k^\nu} s^0(x) & \text{if } \inf_{x \in \operatorname{cl} R_k^\nu} s^0(x) \in [-\nu, \nu] \\ \nu & \text{if } \inf_{x \in \operatorname{cl} R_k^\nu} s^0(x) > \nu, \\ -\nu & \text{otherwise} \end{cases}$$

and construct $s^\nu : I\!R^n \to \overline{I\!R}$ as follows:

$$s^\nu(x) := \min_{k=1,2,...,N^\nu} \left\{ \sigma^\nu(R_k^\nu) \;\middle|\; x \in \operatorname{cl} R_k^\nu \right\}, \quad x \in B,$$

and $s^\nu(x) = \infty$ for $x \notin B$. Clearly, s^ν is constant on each R_k^ν, $k = 1, 2, ..., N^\nu$ and satisfies $\liminf_{x' \to x} s^\nu(x') = s^\nu(x)$ for all $x \in I\!R^n$. Hence, $s^\nu \in \text{e-spl}_n^0(\mathcal{R}^\nu)$ and consequently also in $\text{e-spl}_n^p(\mathcal{R}^\nu)$ for $p \in I\!N$. We next show that the two conditions of Proposition 2.1 holds. Let $x \in I\!R^n$ be arbitrary. By lower semicontinuity of s^0, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $s^0(x') \geq s^0(x) - \varepsilon$ whenever $x' \in I\!B(x, \delta)$. Since $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is an infinite refinement, there also exists a $\bar{\nu}$ such that $R_k^\nu \subset I\!B(x, \delta)$ for every $\nu \geq \bar{\nu}$ and k satisfying $x \in \text{cl}\, R_k^\nu$. Hence, there exists a neighborhood of x on which $s^\nu \geq \max\{s^0(x) - \varepsilon, -\nu\}$ for all $\nu \geq \bar{\nu}$. Thus, for every sequence $x^\nu \to x$,

$$\text{liminf}_\nu\, s^\nu(x^\nu) \geq \text{liminf}_\nu \max\{s^0(x) - \varepsilon, -\nu\} = s^0(x) - \varepsilon.$$

Since ε is arbitrary, $\liminf s^\nu(x^\nu) \geq s^0(x)$ and condition a) of Proposition 2.1 holds. For b), simply set $x^\nu = x$ for all ν. If $x \notin B$, then $s^\nu(x^\nu) = s^0(x) = \infty$. If $x \in B$ and $s^0(x) > -\infty$, then $s^\nu(x) \leq \min\{s^0(x), \nu\}$ for all ν sufficiently large. If $x \in B$ and $s^0(x) = -\infty$, then $s^\nu(x^\nu) = s^\nu(x) = -\nu$. From this follows,

$$\text{limsup}_\nu\, s^\nu(x^\nu) = \text{limsup}_\nu\, s^\nu(x) \leq s^0(x),$$

which concludes the proof. □

A close examination shows that one can consider only *rational epi-splines* of $\text{e-spl}_n^0(\mathcal{R}^\nu)$ in the proof of Theorem 3.5, i.e., functions $s : I\!R^n \to \overline{I\!R}$ with $s(x) = q_k$ for $x \in R_k^\nu$, q_k a rational constant, $k = 1, 2, ..., N^\nu$. Specifically, in that proof one can replace $\sigma^\nu(R_k^\nu) = \inf_{x \in \text{cl}\, R_k^\nu} s^0(x)$ by $\sigma^\nu(R_k^\nu)$ equals any rational number in $[\max\{-\nu, \inf_{x \in \text{cl}\, R_k^\nu} s^0(x) - 1/\nu\}, \inf_{x \in \text{cl}\, R_k^\nu} s^0(x)]$. This implies only minor changes in the proof and we obtain the next result.

3.6 Corollary (separability of the lsc functions) *For $p \in I\!N_0$ and $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$, an infinite refinement of a closed set $B \subseteq I\!R^n$, $(\text{lsc-fcns}(B), d\!l)$ is separable, with the rational epi-splines of $\cup_{\nu=1}^\infty \text{e-spl}_n^0(\mathcal{R}^\nu)$ furnishing a countable dense subset.*

We observe that the restrictions to subsets of lsc functions does not automatically lead to similar density results as the following simple example shows. The 0th-order epi-splines consist of piecewise constant functions and the continuous 0th-order epi-splines are therefore simply the constant functions. Consequently, the *continuous* 0th-order epi-splines cannot be dense in the space of continuous functions. Often, the choice of partition can also lead to the situation that every continuous epi-spline, regardless of order, is simply a constant function. An example is the partition of $I\!R^2$ into $R_1 = \{x \mid e^{x_1} > x_2\}$ and $R_2 = \{x \mid e^{x_1} < x_2\}$. Two polynomials defined on R_1 and R_2, respectively, cannot coincide on $\{x \mid e^{x_1} = x_2\}$ without being identical everywhere. The situation further complicates with restrictions to continuity of derivatives. We here give a result for a partition consisting of simplexes. We recall that a simplex S in $I\!R^n$ is the convex hull of $n+1$ points $x^0, x^1, ..., x^n \in I\!R^n$, with $x^1 - x^0$, $x^2 - x^0$, ..., $x^n - x^0$ linearly independent. We start with the case when only a compact subset of $I\!R^n$ needs to be partitioned.

3.7 Definition (simplicial complex partition of compact set.) *A partition $R_1, R_2, ..., R_N$ of $I\!R^n$ is a* simplicial complex partition *of a compact set $B \subset I\!R^n$ if $\text{cl}\, R_1$, ..., $\text{cl}\, R_N$ are simplexes.*

As usual, we denote by $\mathcal{C}^0(B)$ the continuous functions on $B \subseteq I\!R^n$. This leads us to the following density result for continuous functions, which holds for orders $p \geq 1$.

3.8 Theorem (dense approximation of continuous functions on compact set) *For any $p \in I\!N$ and $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$, an infinite refinement of a compact set $B \subset I\!R^n$ consisting of simplicial complex partitions of B,*

$$\bigcup_{\nu=1}^{\infty} \text{e-spl}_n^p(\mathcal{R}^\nu) \cap \mathcal{C}^0(B) \textit{ is dense in } \mathcal{C}^0(B).$$

Proof: Let $s^0 \in \mathcal{C}^0(B)$ and $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$. It suffices to construct a sequence of epi-splines of order $p = 1$. For $k = 1, ..., N^\nu$, let q_k^ν be the unique affine function on $I\!R^n$ that coincides with the values of s^0 at the $n+1$ extreme points of R_k^ν. We define s^ν to be the epi-spline with partition $\mathcal{R}^\nu$ given by the affine functions q_k^ν on R_k^ν, $k = 1, ..., N$, which is then of first order. We next consider the continuity of s^ν. Consider a facet[5] $\{x \in I\!R^n \mid x \in \text{cl}\, R_k^\nu \cap \text{cl}\, R_l^\nu\}$, $k \neq l$, which is necessarily bounded. By construction, q_k^ν and q_l^ν both coincide with the value of s^0 at the n extreme points of the facet. Since q_k^ν and q_l^ν are affine, they must coincide on the entire facet. Consequently, s^ν is continuous and in e-spl$_n^p(\mathcal{R}^\nu)$ for $p \in I\!N$. We next show that the two conditions of Proposition 2.1 holds. Let $x \in B$ be arbitrary. By continuity of s^0, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|s^0(x') - s^0(x)| < \varepsilon$ whenever $x' \in I\!B(x, \delta) \cap B$. Since $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is an infinite refinement, there also exists a $\bar{\nu}$ such that $R_k^\nu \subset I\!B(x, \delta)$ for every $\nu \geq \bar{\nu}$ and k satisfying $x \in \text{cl}\, R_k^\nu$. We find that $\min_{x' \in I\!B(x,\delta) \cap B} s^0(x') \leq q_k^\nu \leq \max_{x' \in I\!B(x,\delta) \cap B} s^0(x')$ for all every $\nu \geq \bar{\nu}$ and k satisfying $x \in \text{cl}\, R_k^\nu$. Hence, there exists a neighborhood S of x where $|s^\nu(x') - s^0(x)| \leq \varepsilon$ for all $\nu \geq \bar{\nu}$ and $x' \in S \cap B$. This and the fact that both s^ν and s^0 are infinity beyond B, we find that for every sequence $x^\nu \to x$,

$$\text{liminf}_\nu\, s^\nu(x^\nu) \geq s^0(x) - \varepsilon.$$

Since ε is arbitrary, $\liminf s^\nu(x^\nu) \geq s^0(x)$ and condition a) of Proposition 2.1 holds. For b), simply set $x^\nu = x$ for all ν and observe that $s^\nu(x) \leq s^0(x) + \varepsilon$ for all $\nu \geq \bar{\nu}$. From this follows,

$$\text{limsup}_\nu\, s^\nu(x^\nu) = \text{limsup}_\nu\, s^\nu(x) \leq s^0(x) + \varepsilon.$$

Again, since ε is arbitrary condition b) follows, which concludes the proof. □

The case of partition of the whole of $I\!R^n$ requires additional notation. For $x = (x_1, ..., x_n) \in I\!R^n$, we use the notation x_{-i} to denote the $(n-1)$-dimensional vector that excludes the ith component of x.

3.9 Definition (simplex cylinder.) *A simplex cylinder S in $I\!R^n$ generated by a simplex S' in $I\!R^{n-1}$ and the interval $[\eta_1, \eta_2] \subset \overline{I\!R}$ is the set $\{x \in I\!R^n \mid x_{-i} \in S', x_i \in [\eta_1, \eta_2]\}$ for some i. We say that S' is the base of S.*

3.10 Definition (simplicial complex partition of $I\!R^n$.) *A partition $R_1, R_2, ..., R_N$ of $I\!R^n$ is a simplicial complex partition of $I\!R^n$ if for a box $B = [l_1, u_1] \times [l_2, u_2] \times ... \times [l_n, u_n]$, with $-\infty < l_i < u_i < \infty$, $i = 1, 2, ..., n$, and $N_s, N_c, N_r \in I\!N$, with $N_s + N_c + N_r = N$, we have that*

[5]We recall that a *facet* of an n-dimensional simplex is an $(n-1)$-dimensional set defined by one of the faces of the simplex.

a) $\mathrm{cl}\, R_1$, ..., $\mathrm{cl}\, R_{N_s}$ *are simplexes and* $\cup_{k=1}^{N_s} \mathrm{cl}\, R_k = B$

b) R_{N_s+1}, ..., $R_{N_s+N_c}$ *are simplex cylinders generated by the* $(n-1)$*-dimensional simplexes formed by the intersection of* R_k*, for some* $k = 1, ...N_s$*, and a facet of* B $\{x \in I\!R^n \mid l_j \le x_j \le u_j, j \ne i, x_i = l_i\}$*, for some* $i = 1, ..., n$*, and the interval* $[-\infty, l_i]$ *as well as by the simplexes formed by the intersection of* R_k*, for some* $k = 1, ...N_s$*, and* $\{x \in I\!R^n \mid l_j \le x_j \le u_j, j \ne i, x_i = u_i\}$ *for some* $i = 1, ..., n$*, and the interval* $[u_i, \infty]$*.*

c) $R_{N_s+N_c+1}$, ..., R_N *are unbounded* n*-dimensional hyperrectangles given by the* n*-fold cartesian product of intervals of the forms* $[l_i, u_i]$*,* $[u_i, \infty)$*, and* $(-\infty, l_i]$*,* $i = 1, ..., n$*, each using at most* $n-2$ *intervals of the form* $[l_i, u_i]$*.*

A density result for continuous functions on $I\!R^n$ follows next.

3.11 Theorem (dense approximation of continuous functions on $I\!R^n$) *For any* $p \in I\!N$ *and* $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$*, an infinite refinement on* $I\!R^n$ *consisting of simplicial complex partitions of* $I\!R^n$*,*

$$\bigcup_{\nu=1}^{\infty} \text{e-spl}_n^p(\mathcal{R}^\nu) \cap \mathcal{C}^0(I\!R^n) \textit{ is dense in } \mathcal{C}^0(I\!R^n).$$

Proof: Let $s^0 \in \mathcal{C}^0(I\!R^n)$ and $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$, with N_s^ν and N_c^ν the number of simplexes and simplex cylinders, respectively. Again, it suffices to construct a sequence of epi-splines of order $p = 1$. For $k = 1, ..., N_s^\nu$, let q_k^ν be the unique affine function on $I\!R^n$ that coincides with the values of s^0 at the $n+1$ extreme points of R_k^ν. For $k = N_s^\nu + 1, ..., N_s^\nu + N_c^\nu$, let q_k^ν be the unique affine function on $I\!R^n$ that coincides with the values of s^0 at the n extreme points of R_k^ν and that is constant in the directions perpendicular to the base of the simplex cylinder R_k^ν. For $k = N_s^\nu + N_c^\nu + 1, ..., N$, R_k^ν are hyperrectangles with n, $n-1$, ...3, or 2 unbounded directions. If the number of unbounded directions of such R_k^ν is m, then R_k^ν has $n - m + 1$ extreme points. Let q_k^ν be the unique affine function on $I\!R^n$ with value of s^0 at the extreme points of R_k^ν and that is constant in the directions of unboundedness. We define s^ν to be the epi-spline given by the affine functions q_k^ν, $k = 1, ..., N$, which is then of first order. We next consider the continuity of s^ν. First consider a facet $\{x \in I\!R^n \mid x \in \mathrm{cl}\, R_k^\nu \cap \mathrm{cl}\, R_l^\nu\}$, $k \ne l$, that is bounded. By construction, q_k^ν and q_l^ν both coincide with the value of s^0 on the n extreme points of the facet. Since q_k^ν and q_l^ν are affine, the must coincide on the whole facet. Second consider an unbounded facet $\{x \in I\!R^n \mid x \in \mathrm{cl}\, R_k^\nu \cap \mathrm{cl}\, R_l^\nu\}$, $k \ne l$. By construction, q_k^ν and q_l^ν both coincide with the value of s^0 on the extreme points of the facet. This fact and the constancy of q_k^ν and q_l^ν in unbounded directions imply that q_k^ν and q_l^ν coincide on the whole facet. Consequently, s^ν is continuous and in $\text{e-spl}_n^p(\mathcal{R}^\nu)$ for $p \in I\!N$. We next show that the two conditions of Proposition 2.1 holds. Let $x \in I\!R^n$ be arbitrary. By continuity of s^0, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|s^0(x') - s^0(x)| < \varepsilon$ whenever $x' \in I\!B(x, \delta)$. Since $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is an infinite refinement, there also exists a $\bar\nu$ such that $R_k^\nu \subset I\!B(x, \delta)$ for every $\nu \ge \bar\nu$ and k satisfying $x \in \mathrm{cl}\, R_k^\nu$. Thus $\min_{x' \in I\!B(x,\delta)} s^0(x') \le q_k^\nu \le \max_{x' \in I\!B(x,\delta)} s^0(x')$ for all every $\nu \ge \bar\nu$ and k satisfying $x \in \mathrm{cl}\, R_k^\nu$ and there exists a neighborhood S of x where $|s^\nu(x') - s^0(x)| \le \varepsilon$ for all $\nu \ge \bar\nu$ and $x' \in S$. Consequently, for every sequence $x^\nu \to x$,

$$\mathrm{liminf}_\nu\, s^\nu(x^\nu) \ge s^0(x) - \varepsilon.$$

Since ε is arbitrary, $\liminf s^\nu(x^\nu) \geq s^0(x)$ and condition a) of Proposition 2.1 holds. For b), simply set $x^\nu = x$ for all ν and observe that $s^\nu(x) \leq s^0(x) + \varepsilon$ for all $\nu \geq \bar{\nu}$. From this follows,

$$\text{limsup}_\nu\, s^\nu(x^\nu) = \text{limsup}_\nu\, s^\nu(x) \leq s^0(x) + \varepsilon.$$

Again, since ε is arbitrary condition b) follows, which concludes the proof. □

We end the subsection with a discussion of "decomposition" of epi-splines on $I\!R^n$ into sums of those defined on a lower-dimensional space. For given $p, n \in I\!N$ and $x \in I\!R^n$, there are $\binom{n}{p}$ subvectors of x with p components. We denote by $x_{[j]}$, $j = 1, ..., \binom{n}{p}$, these subvectors. The next theorem provides a decomposition of an epi-spline in terms of lower-dimensional polynomials.

3.12 Theorem (decomposition.) *For every $s \in \text{e-spl}_n^p(\mathcal{R})$, with $n > p$ and $\mathcal{R} = \{R_k\}_{k=1}^N$, and $k = 1, 2, ..., N$, there exist $q_k \in \text{poly}^p(I\!R^n)$ and $q_{k,j} \in \text{poly}^p(I\!R^p)$, $j = 1, 2, ..., \binom{n}{p}$, such that*

$$s(x) = q_k(x) = \sum_{j=1}^{\binom{n}{p}} q_{k,j}(x_{[j]}), \text{ for all } x \in R_k.$$

Proof: The first equality follows trivially from the definition of epi-splines. A polynomial $q_k(x)$ is the sum of n_p terms each involving at most p components of x. A term involving $x_{j_1}, ..., x_{j_p}$, say, is also part of the description of a polynomial of total degree p in the p-dimensional subspace of $I\!R^n$ corresponding to dimensions j_1, ..., j_p. Since there are $\binom{n}{p}$ ways of selecting such subspaces, the second equality follows. □

Theorem 3.12 decomposes n-dimensional polynomials into sums of p-dimensional polynomials. Using similar arguments as in this theorem's proof, we obtain a "one-dimensional reduction" as stated next.

3.13 Corollary *For every $s \in \text{e-spl}_n^p(\mathcal{R})$, with $n > p$ and $\mathcal{R} = \{R_k\}_{k=1}^N$, and $k = 1, 2, ..., N$, there exist $q_{k,i} \in \text{poly}^p(I\!R^{n-1})$, $i = 1, 2, ..., n$, such that*

$$s(x) = \sum_{i=1}^n q_{k,i}(x_{-i}), \text{ for all } x \in R_k.$$

3.2 Evolving and Approximating Problems

The previous subsection deals with the approximation of lsc-fcns($I\!R^n$) by $\text{e-spl}_n^p(\mathcal{R})$. This lays the foundation for approximating the infinite-dimensional (FIP) by one involving a finite number of parameters. In practice, however, data and extrinsic information might accumulate gradually and the criterion functional and constraints might also need to be approximated, which lead to a family of evolving function identification problems:

$$(FIP^\nu): \qquad \min \psi^\nu(s) \text{ such that } s \in F^\nu \cap \mathcal{S}^\nu,$$

with $\psi^\nu : \mathcal{F} \to \overline{I\!R}$ being an *approximate criterion functional*, possibly representing an incomplete data set of size ν, $F^\nu \subseteq \mathcal{F}$ being an *approximate constraint set* that substitutes F using currently available information as well as computationally required approximations, and $\mathcal{S}^\nu \subseteq$ e-spl$_n^p(\mathcal{R}) \cap \mathcal{F}$. Since epi-splines are characterized by a finite number of parameters, (FIP^ν) is in one-to-one correspondence with an optimization problem in a (finite-dimensional) Euclidean space. If the approximations are constructed properly, one might hope that solutions of (FIP^ν) approximate solutions of (FIP). This property is ensured by the next results under standard conditions. First, however, we need to recall an extended notion of epi-convergence.

Since we not only need to deal with epi-convergence of functions in lsc-fcns($I\!R^n$), but also of the extended real-valued functionals defined on lsc-fcns($I\!R^n$), giving rise to (FIP) and (FIP^ν), we provide a definition of epi-convergence of the evolving problems (FIP^ν) to the limiting problem (FIP); see [2] for further details about epi-convergence of functionals defined on metric spaces.

3.14 Definition *A sequence of evolving problems* $\{(FIP^\nu)\}_{\nu \in N}$ *epi-converges to an actual problem* (FIP) *if*

a) for every sequence $\{s^\nu\}_{\nu\in K}$*, with* K *an infinite subset of* $I\!N$*,* $s^\nu \in F^\nu \cap \mathcal{S}^\nu$*, and* $d\!l(s^\nu, f) \to 0$*, we have that* $f \in F$ *and* $\liminf_\nu \psi^\nu(s^\nu) \geq \psi(f)$*;*

b) for every $f \in F$*, there exists a sequence* $\{s^\nu\}_{\nu\in N}$*, with* $s^\nu \in F^\nu \cap \mathcal{S}^\nu$*, such that* $d\!l(s^\nu, f) \to 0$ *and* $\limsup_\nu \psi^\nu(s^\nu) \leq \psi(f)$.

With the perspective that (FIP^ν) is an approximation of (FIP), we give two results that justify the use of an epi-spline obtained from (FIP^ν) as an approximation of an actual function given by (FIP). Epi-convergence of (FIP^ν) to the actual problem (FIP) is the central property. We denote the optimal values of (FIP) and (FIP^ν) by V and V^ν, respectively.

3.15 Theorem (convergence of minimizers [2, Theorem 2.5].) *Suppose that* $\{(FIP^\nu)\}_{\nu\in N}$ *epi-converges to* (FIP)*,* s^k *minimizes* (FIP^{ν_k})*,* $k \in I\!N$*, and* $d\!l(s^k, f) \to 0$*. Then,* f *is a minimizer of* (FIP) *and* $\lim V^{\nu_k} = V$.

We next give a sufficient condition for epi-convergence of (FIP^ν) to the actual problem (FIP). We recall that ψ^ν converges continuously to ψ relative to $\mathcal{F}$ if for every $f \in \mathcal{F}$ and sequence $s^\nu \to f$, with $s^\nu \in \mathcal{F}$, $\psi^\nu(s^\nu) \to \psi(f)$. Let $\operatorname{int} S$ denote the interior of a set S.

3.16 Theorem (sufficient condition for epi-convergence.) *If* ψ^ν *converges continuously to* ψ *relative to* $\mathcal{F}$*,* $\cup_{\nu\in N}\mathcal{S}^\nu$ *is dense in* $\mathcal{F}$*, and* F^ν *converge*[6] *to* $F = \operatorname{cl}(\operatorname{int} F)$*, then* $\{(FIP^\nu)\}_{\nu\in N}$ *epi-converges to* (FIP).

Proof. We here follow the argument of Theorem 2.5 of [30], but include it for completeness. We first consider a) in Definition 3.14. For $\{s^\nu\}_{\nu\in K}$, with K an infinite subset of $I\!N$, $s^\nu \in F^\nu \cap \mathcal{S}^\nu$, and

[6] In the sense of Painlevé-Kuratowski; see footnote of §2 and [26, Chapter 4].

$d\!l(s^\nu, f) \to 0$, we immediately find that $f \in F$ from the assumption that $F^\nu \to F$. In view of the continuous convergence of ψ^ν to ψ, we establish part a). Second, we consider part b) and let $f \in F$ be arbitrary. Since $F = \mathrm{cl}(\mathrm{int}\, F)$, there exists a sequence $\{f^\mu\}_{\mu\in N} \subset \mathrm{int}\, F$ such that $d\!l(f^\mu, f) \to 0$. For each μ, the facts that $\cup_{\nu\in N}\mathcal{S}^\nu$ is dense in $\mathcal{F}$, F^ν converges to F, and f^μ belongs to an open subset of F imply that there exists a sequence $\{s^{\mu,\nu}\}_{\nu=\nu_\mu}^\infty \subset F^\nu \cap \mathcal{S}^\nu$ such that $d\!l(s^{\mu,\nu}, f^\mu) \to 0$ as $\nu \to \infty$. Consequently, we can construct a sequence $\{\nu_\mu^*\}_{\mu\in N}$ such that $\nu_{\mu+1}^* > \nu_\mu^*$, $\nu_\mu^* \geq \max\{\nu_\mu, \nu_{\mu+1}\}$, and $d\!l(s^{\mu,\nu}, f^\mu) \leq 1/\mu$ for all $\nu \geq \nu_\mu^*$ and μ. For arbitrary $\varepsilon > 0$, there exists therefore an integer $\mu_0 \geq 2/\varepsilon$ such that for all $\nu \geq \nu_\mu^*$ and $\mu > \mu_0$,

$$d\!l(f^\mu, f) \leq \varepsilon/2 \text{ and } d\!l(s^{\mu,\nu}, f^\mu) \leq \varepsilon/2.$$

We construct the sequence $\{s^\nu\}_{\nu=\nu_1}^\infty \subset F^\nu \cap \mathcal{S}^\nu$ by setting

$$s^\nu = s^{\mu,\nu} \text{ with } \mu \text{ satisfying } \nu \in \{\nu_{\mu-1}^* + 1, \nu_{\mu-1}^* + 2, ..., \nu_\mu^*\}.$$

Then for every $\nu > \nu_{\mu_0}^*$ and some $\mu_\nu > \mu_0$,

$$d\!l(s^\nu, f) = d\!l(s^{\mu_\nu,\nu}, f) \leq d\!l(s^{\mu_\nu,\nu}, f^{\mu_\nu}) + d\!l(f^{\mu_\nu}, f) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Consequently, $d\!l(s^\nu, f) \to 0$ and $\psi^\nu(s^\nu) \to \psi(f)$ by continuous convergence, which establish part b). □

We note that the constraint qualification $F = \mathrm{cl}(\mathrm{int}\, F)$, i.e., F is solid, avoids "isolated" feasible points that cannot easily be approximated.

3.3 Estimating Epi-Distances

We next provide estimates of epi-distances, especially between epi-splines, and make connections between epi-convergence and other modes of convergence. It is immediately clear from [26, Proposition 7.15] that uniform convergence implies epi-convergence, which we utilize repeatedly below. However, the converse fails as illustrated below. In fact, epi-convergence can be viewed as a one-sided convergence.

3.17 Theorem (epi-distance estimates) *For $s, s' \in$ e-spl$_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$, one has for any $\rho \geq 0$,*

(i) $d\!l(s, s') \leq \|s - s'\|_\infty$ and $d\!l_\rho(s, s') \leq \|s - s'\|_\infty$
(ii) $d\!l(s, s') \leq \max_{k=1,\dots,N} \sup_{x\in R_k} |s(x) - s'(x)|$ and $d\!l_\rho(s, s') \leq \max_{k=1,\dots,N} \sup_{x\in R_k} |s(x) - s'(x)|$.

Proof: We first consider $d\!l_\rho(s, s')$. For any $\bar{x} = (x, x_0) \in \rho I\!\!B \subset I\!\!R^{n+1}$, $d(\bar{x}, \mathrm{epi}\, s) \leq d(\bar{x}, \mathrm{epi}\, s') + \|s - s'\|_\infty$. Reversing the roles of s and s' yields that $|d(\bar{x}, \mathrm{epi}\, s) - d(\bar{x}, \mathrm{epi}\, s')| \leq \|s - s'\|_\infty$. Consequently, $d\!l_\rho(s, s') \leq \|s - s'\|_\infty$. Since $\int_0^\infty e^{-\rho} d\rho = 1$, we also conclude that the first part of item (i) holds. From (i) it is clear that item (ii) would hold with R_k replaced by $\mathrm{cl}\, R_k$. The "closure" is superfluous for the following reason. For every $x^0 \in I\!\!R^n$ there exists a sequence $x^\nu \to x^0$ and $k^* \in \{1, 2, ..., N\}$ such that $x^\nu \in R_{k^*}$ for all ν and $\lim_\nu s'(x^\nu) = s'(x^0)$ by the definition of epi-splines and openness of R_k, $k = 1, ..., N$. Since $s(x^0) = \liminf_{x' \to x^0} s(x')$ and s is continuous on R_{k^*}, it then follows that

$$s(x^0) - s'(x^0) \leq \lim_\nu s(x^\nu) - \lim_\nu s'(x^\nu) \leq \sup_{x \in R_{k^*}} |s(x) - s'(x)|.$$

Reversing the roles of s and s', yields that

$$s(x^0) - s'(x^0) \geq - \sup_{x \in R_{k^*}} |s(x) - s'(x)|.$$

Consequently,

$$|s(x^0) - s'(x^0)| \leq \max_{k=1,...,N} \sup_{x \in R_k} |s(x) - s'(x)|,$$

which, because $x^0 \in I\!R^n$ is arbitrary, completes the proof of item (ii) after again using the fact that $\int_0^\infty e^{-\rho} d\rho = 1$. □

We observe that since $s - s'$ is also a polynomial on R_k, say $q \in \text{poly}^p(I\!R^n)$,

$$\sup_{x \in R_k} |s(x) - s'(x)| = \max \left\{ \sup_{x \in \text{cl}\, R_k} q(x), - \inf_{x \in \text{cl}\, R_k} q(x) \right\}.$$

In general, optimization of polynomial over an arbitrary closed set is difficult, but many special, more tractable cases exist, especially in practically important situations with low degree and partitions consisting of boxes.

By Theorem 3.17, a sequence $\{s^\nu\}_{\nu=1}^\infty \subset \text{e-spl}_n^p(\mathcal{R})$ converging to $s^0 \in \text{e-spl}_n^p(\mathcal{R})$ in the L^∞-norm, i.e., $\|s^\nu - s^0\|_\infty \to 0$, also converges in the epi-distance, i.e., $d\!l(s^\nu, s^0) \to 0$. The converse fails, however, as the following counterexample on $I\!R$ demonstrates. For $\mathcal{R} = \{(-\infty, 0), (0, \infty)\}$ and $\nu \in I\!N$, let $s^\nu \in \text{e-spl}_1^1(\mathcal{R})$ be $s^\nu(x) = 0$ if $x \leq 0$ and $s^\nu(x) = x/\nu$ otherwise. We also define $s^0(x) = 0$ for all $x \in I\!R$. Clearly, $\|s^\nu - s^0\|_\infty = \infty$ for all $\nu \in I\!N$. However, for any $\rho \geq 0$, $d\!l_\rho(s^\nu, s^0) \leq \rho/\nu$ and therefore $d\!l(s^\nu, s^0) \leq (1/\nu) \int_0^\infty \rho e^{-\rho} d\rho = 1/\nu$. Consequently, $d\!l(s^\nu, s^0) \to 0$, but $\|s^\nu - s^0\|_\infty \not\to 0$.

Further estimates of the epi-distance are available through a supporting quantity defined next, which as $d\!l$ and $d\!l_\rho$, fully characterize epi-convergence; see [26, Theorem 7.58]. We let for any $f, g : I\!R^n \to \overline{I\!R}$,

$$\hat{d\!l}_\rho(f, g) = \inf\{\eta \geq 0 \mid \text{epi}\, f \cap \rho I\!B \subset \text{epi}\, g + \eta I\!B, \text{epi}\, g \cap \rho I\!B \subset \text{epi}\, f + \eta I\!B\}.$$

As $d\!l_\rho$, $\hat{d\!l}_\rho$ is closely related to the Pompeiu-Hausdorff distance; see [26, Chapters 4 and 7]. Moreover, for f, g convex with $f(0) \leq 0$ and $g(0) \leq 0$, $d\!l_\rho(f, g) = \hat{d\!l}_\rho(f, g)$ for any $\rho \geq 0$ by [26, Exercise 7.60]. The same exercise also provides the following more general estimates.

3.18 Proposition *For $f, g : I\!R^n \to \overline{I\!R}$ not identical to ∞, the following hold with $d_f = d(0, \text{epi}\, f)$ and $d_g = d(0, \text{epi}\, g)$:*

(i) $\hat{d\!l}_\rho(f,g) \leq d\!l_\rho(f,g) \leq \hat{d\!l}_{\rho'}(f,g)$, *when* $\rho' \geq 2\rho + \max\{d_f, d_g\}$

(ii) $d\!l(f,g) \geq (1-e^{-\rho})|d_f - d_g| + e^{-\rho} d\!l_\rho(f,g)$

(iii) $d\!l(f,g) \leq (1-e^{-\rho}) d\!l_\rho(f,g) + e^{-\rho}(\max\{d_f, d_g\} + \rho + 1)$.

Computation of the epi-distance by combining results for each set in a partition is supported by the next result, which especially is useful if the diameter

$$\operatorname{diam}(R_k) := \sup_{x,x' \in \operatorname{cl} R_k} \|x - x'\|$$

of R_k is small. For $s \in \text{e-spl}_n^p(\mathcal{R})$, we adopt the notation

$$s_k := s + \delta_{\operatorname{cl} R_k}, \text{ where } \delta_{\operatorname{cl} R_k}(x) = 0 \text{ if } x \in \operatorname{cl} R_k \text{ and } \infty \text{ otherwise.}$$

3.19 Theorem (additional epi-distance estimates) *For* $s, s' \in \text{e-spl}_n^p(\mathcal{R})$, *with* $\mathcal{R} = \{R_k\}_{k=1}^N$, *one has*

(i) $\hat{d\!l}_\rho(s, s') \leq \max_{k=1,...,N} \hat{d\!l}_\rho(s_k, s'_k)$

(ii) For any $k = 1, ..., N$, $\hat{d\!l}_\rho(s_k, s'_k) = 0$ *if* $\operatorname{epi} s_k \cap \rho I\!B = \operatorname{epi} s'_k \cap \rho I\!B = \emptyset$, *and otherwise*
$\hat{d\!l}_\rho(s_k, s'_k)/\sqrt{2} \leq$
$$\max\left[\operatorname{diam}(R_k), \inf_{x \in \operatorname{cl} R_k} s'(x) - \max\left\{\min_{x \in \operatorname{cl} R_k \cap \rho \boldsymbol{B}} s(x), -\rho\right\}, \inf_{x \in \operatorname{cl} R_k} s(x) - \max\left\{\min_{x \in \operatorname{cl} R_k \cap \rho \boldsymbol{B}} s'(x), -\rho\right\}\right].$$

Proof: First consider part (i). Since for all $k = 1, ..., N$,

$$\operatorname{epi} s_k \cap \rho I\!B \subset \operatorname{epi} s'_k + \hat{d\!l}_\rho(s_k, s'_k) I\!B,$$

we also have that
$$\left(\bigcup_{k=1}^N \operatorname{epi} s_k\right) \cap \rho I\!B \subset \left(\bigcup_{k=1}^N \operatorname{epi} s'_k\right) + \max_{k=1,...,N} \hat{d\!l}_\rho(s_k, s'_k) I\!B.$$

A similar results with the roles of s_k and s'_k reversed and the fact that $\operatorname{epi} s = \cup_{k=1}^N \operatorname{epi} s_k$ and likewise for s' yield the conclusion. Second consider part (ii). The first case follows trivially from the definition of $\hat{d\!l}_\rho$. The more involved case relies on [26, Proposition 7.61] from which we deduce that $\hat{d\!l}_\rho(s_k, s'_k)/\sqrt{2}$ is no larger than the infimum of all $\eta \geq \operatorname{diam}(R_k)$ such that for all $x \in \rho I\!B \cap \operatorname{cl} R_k$,

$$\inf_{x' \in \operatorname{cl} R_k} s(x') \leq \max\{s'(x), -\rho\} + \eta \text{ and also with the roles of } s \text{ and } s' \text{ reversed.}$$

Consequently, η must satisfy

$$\inf_{x' \in \operatorname{cl} R_k} s(x') - \max\left\{\min_{x \in \operatorname{cl} R_k \cap \rho \boldsymbol{B}} s'(x), -\rho\right\} \leq \eta \text{ and also with the roles of } s \text{ and } s' \text{ reversed,}$$

and the conclusion follows. ◻

We again observe that optimizing polynomials in general is challenging, but the formula in Theorem 3.19(ii) requires only minimization, in contrast to the one in Theorem 3.17(ii) where both minimization and maximization are needed, and therefore in the case of convex s_k and s'_k as well as convex R_k is easily implemented.

We end this subsection with a result that highlights the connections between epi-splines and their representation by a finite number of parameters. We start with an intermediate result.

3.20 Proposition (functional convergence of polynomials). *Suppose that $q^\nu \in \mathrm{poly}^p(\mathbb{R}^n)$, $\nu \in \mathbb{N}_0$, and $c = (c_1, c_2, ..., c_{n_p})$ is a basis for such polynomials, i.e., there exists a unique $a^\nu \in \mathbb{R}^{n_p}$ such that $q^\nu = \langle c(\cdot), a^\nu\rangle$, $\nu \in \mathbb{N}_0$. Then,*

$$d\!l(q^\nu, q^0) \to 0 \Longleftrightarrow a^\nu \to a^0.$$

Proof. Since

$$|q^\nu - q^0| = |\langle c(\cdot), a^\nu - a^0\rangle| \le \|c(\cdot)\|\|a^\nu - a^0\|$$

and c is continuous (in fact polynomial), convergence $a^\nu \to a^0$ implies uniform convergence of q^ν to q^0 on any compact subset of $\mathbb{R}^n$. Using Proposition 2.1, a standard argument shows that q^ν epi-converges to q^0 and therefore also $d\!l(q^\nu, q^0) \to 0$. Next, we consider the converse. Suppose for the sake of a contradiction that $d\!l(q^\nu, q^0) \to 0$, but there exists a $\rho > 0$ and an infinite subset N_∞ of $\mathbb{N}_0$ such that $\|a^\nu - a^0\| \ge \rho$ for all $\nu \in N_\infty$. Let $x^1, ..., x^{n_p} \in \mathbb{R}^n$ be a collection of distinct points such that every polynomial $q \in \mathrm{poly}^p(\mathbb{R}^n)$ is uniquely defined by the solution of the system $C^0 a = b^0(q)$, where C^0 is the n_p-by-n_p matrix with rows $c(x^i)$, $i = 1, ..., n_p$, and $b^0(q)$ is the transpose of $(q(x^1), ..., q(x^{n_p}))$. That is, $q = \langle c(\cdot), a\rangle$, where a is the solution of $C^0 a = b^0(q)$. Let $\varepsilon^\mu \searrow 0$, as $\mu \to \infty$, and the open balls

$$B_i^\mu = \{(x, x_0) \in \mathbb{R}^{n+1} \mid \|(x, x_0) - (x^i, q^0(x^i))\| < \varepsilon^\mu\}.$$

For each μ and i, it follows from the "hit-and-miss criterion" of [26, Proposition 4.5(a)] that there exists a ν_i^μ such that for all $\nu \ge \nu_i^\mu$ $\mathrm{epi}\, q^\nu \cap B_i^\mu \ne \emptyset$. Moreover, by [26, Proposition 4.5(b)], there exists a ν_0^μ and a compact set $S \subset \mathbb{R}^n$ containing $x^1, ..., x^{n_p}$ such that for all $\nu \ge \nu_0^\mu$, $q^\nu > q^0 - \varepsilon^\mu/2$ on S. Consequently, with $\nu^\mu = \max[\nu_0^\mu, \nu_1^\mu, ..., \nu_{n_p}^\mu]$, the graph of q^ν intersects all the balls B_i^μ, $i = 1, \ldots, n_p$, for all μ and $\nu \ge \nu^\mu$. Let $(x^{i,\nu,\mu}, q^\nu(x^{i,\nu,\mu}))$, $i = 1, \ldots, n_p$, be such intersection points, i.e., $(x^{i,\nu,\mu}, q^\nu(x^{i,\nu,\mu})) \in B_i^\mu$ for $i = 1, \ldots, n_p, \mu \in \mathbb{N}$, and $\nu \ge \nu^\mu$. Let $C^{\nu,\mu}$ be the n_p-by-n_p matrix with rows $c(x^{i,\nu,\mu})$, $i = 1, ..., n_p$, and $b^{\nu,\mu}(q^\nu)$ be the transpose of $(q^\nu(x^{1,\nu,\mu}), ..., q^\nu(x^{n_p,\nu,\mu}))$. For sufficiently large μ, the unique solution $a^{\nu,\mu}$ of $C^{\nu,\mu} a = b^{\nu,\mu}(q^\nu)$ coincides with the coefficients a^ν of q^ν for $\nu \ge \nu^\mu$, $\mu \in \mathbb{N}$. Moreover, $a^{\nu,\mu}$ and a^0 are the unique optimal solutions of

$$\min_a \|C^{\nu,\mu} a - b^{\nu,\mu}(q^\nu)\|^2 \text{ and } \min_a \|C^0 a - b^0(q^0)\|^2,$$

respectively. Since $\varepsilon^\mu \downarrow 0$, as $\mu \to \infty$, $C^{\nu,\mu} \to C^0$ and $b^{\nu,\mu}(q^\nu) \to b^0(q^0)$, as $\mu \to \infty$ and $\nu \ge \nu^\mu$, the objective function of the first problem epi-converges to that of the second problem as $\mu \to \infty$ and $\nu \ge \nu^\mu$. Theorem 7.31 of [26] then implies that $a^{\nu,\mu} \to a^0$ as $\mu \to \infty$, $\nu \ge \nu^\mu$. Consequently, there exists a $\bar{\mu}$ such that $\|a^{\nu,\mu} - a^0\| < \rho$ for all $\mu \geq \bar{\mu}$, $\nu \geq \nu^\mu$. Since $a^{\nu,\mu} = a^\nu$ for $\nu \geq \nu^\mu$, we have reached a contradiction. $\square$

3.21 Theorem (convergence of the parameters). *Suppose that $s^\nu \in \text{e-spl}_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$, $\nu \in I\!N_0$, and $c = (c_1, c_2, ..., c_{n_p})$ is a basis for polynomials in $\text{poly}^p(I\!R^n)$, i.e., for unique $a^{k,\nu} \in I\!R^{n_p}$,*

$$s^\nu(x) = \langle c(x), a^{k,\nu} \rangle, \text{ when } x \in R_k.$$

Then,

$$a^{1,\nu} \to a^{1,0}, ..., a^{N,\nu} \to a^{N,0} \iff s^\nu \to s^0 \text{ uniformly on compact sets } S \subset I\!R^n \iff d\!l(s^\nu, s^0) \to 0.$$

Proof. Since polynomials of total order p convergence uniformly on compact sets if and only if their coefficient converges, we obtain the first implications. In view of Proposition 3.20 and the fact that uniform convergence implies epi-convergence [26, Proposition 7.15], the second implication holds. $\square$

4 Composite Epi-Spline

Probability density estimation [29, 34] is one of our major incentives for considering problems of the form (FIP) and constructing evolving approximations (FIP^ν). A density is a nonnegative function that sums up to 1 and an estimate is chosen so as to minimize some appropriate criterion; for further details see §6.2, [29], and references therein. In addition, many "standard" densities belong to an exponential family [3], which lead us to building density estimates in terms of an exponential function composed with an epi-spline, which are nonnegative automatically, rather than epi-splines. In other cases, a nonlinear transformation of this kind may improve conditioning of (FIP) and therefore facilitate its numerical implementation. On a theoretical level the results for such *composite epi-splines*, including the formulation of (FIP) in terms of composite epi-splines, follow rather straightforwardly from those in §2 and §3. Still, for easy reference and a better understanding of the specific properties, it is useful to record the central results.

We start by defining the set of composite epi-splines corresponding to a function $\theta : I\!R \to I\!R$.

4.1 Definition (composite epi-splines). *A composite epi-spline $h : I\!R^n \to I\!R$ of order $p \in I\!N_0$, with partition $\mathcal{R} = \{R_k\}_{k=1}^N$ of $I\!R^n$ and function $\theta : I\!R \to I\!R$, is a function*

$$h = \theta \circ s, \text{ where } s \in \text{e-spl}_n^p(\mathcal{R}).$$

The family of all such composite epi-splines is denoted by $\text{c-spl}_n^p(\mathcal{R}, \theta)$.

Composite epi-splines are generally not lsc and may be upper semi-continuous (usc) as we see next. We denote by usc-fcns($I\!R^n$) the space of all usc $f : I\!R^n \to \overline{I\!R}$, excluding $f \equiv -\infty$. After observing that the hypograph of a function f, $\text{hypo} f = \{(x, x_0) \,|\, f(x) \geq x_0\}$ is just a mirror image of the epigraph of $-f$, one can mimic the definitions and constructions described for lsc functions to set up the *hypo-distance* $d\!l_{\rm hypo}(f,g) = d\!l(-f,-g)$, between any two functions f and g and generate the *hypo-topology* which again makes (usc-fcns, $d\!l$) a Polish space. A sequence of functions f^ν *hypo-converge* to f if $-f^\nu$ epi-converge to $-f$; see [26, Chapters 4 & 7] for a broader treatment.

4.2 Proposition *For a continuous $\theta : I\!R \to I\!R$, the following hold:*

(i) If θ is increasing, then c-spl$^p_n(\mathcal{R}, \theta) \subset$ *lsc-fcns($I\!R^n$).*

(ii) If θ is decreasing, then c-spl$^p_n(\mathcal{R}, \theta) \subset$ *usc-fcns($I\!R^n$).*

Proof: These results follow as direct consequences of definitions of lsc and usc functions. □

4.3 Theorem (dense approximations by composite epi-splines). *Under the assumption of Theorem 3.5 and a continuous $\theta : I\!R \to I\!R$, the following hold:*

(i) If θ is increasing and extended to $\overline{I\!R}$ by setting $\theta(\infty) = \sup\theta$ and $\theta(-\infty) = \inf\theta$, then

$$\bigcup_{\nu=1}^{\infty} \text{c-spl}^p_n(\mathcal{R}^\nu, \theta) \textit{ is dense in } \{\theta \circ f \mid f \in \textit{lsc-fcns}(I\!R^n)\} \textit{ under the epi-topology.}$$

(ii) If θ is decreasing and extended to $\overline{I\!R}$ by setting $\theta(\infty) = \inf\theta$ and $\theta(-\infty) = \sup\theta$, then

$$\bigcup_{\nu=1}^{\infty} \text{c-spl}^p_n(\mathcal{R}^\nu, \theta) \textit{ is dense in } \{\theta \circ f \mid f \in \textit{usc-fcns}(I\!R^n)\} \textit{ under the hypo-topology.}$$

Proof: By Theorem 3.5, for every $s^0 \in$ lsc-fcns($I\!R^n$) there exists a sequence $\{s^\nu\}_{\nu\in N}$, with $s^\nu \in$ e-spl$^p_n(\mathcal{R}^\nu)$ for all ν, such that $d\!l(s^\nu, s^0) \to 0$. By Exercise 7.8 of [26], it then follows that $\theta \circ s^\nu$ epi-converges to $\theta \circ s^0$ and the first conclusion follows. The second result is a consequence of an identical argument under a sign change. □

5 Extrinsic information

Information about the nature of solutions of (FIP) may justify the consideration of subsets of lsc-fcns($I\!R^n$) as defined through F and $\mathcal{F}$ and likewise subsets of e-spl$^p_n(\mathcal{R})$. We here provide a few examples of such information and their implementation as constraints in (FIP^ν).

Domain. External information may indicate that the (effective) domain of solution functions of (FIP) is a closed, strict subset B of $I\!R^n$. Then, a partition of B instead of $I\!R^n$ by selecting $\mathcal{R} = \{R_k\}_{k=1}^N$ such that $\cup_{k=1}^N \operatorname{cl} R_k = B$ avoids wasting computational effort on uninteresting parts of $I\!R^n$. In general, the specific choice of $\mathcal{R}$ is guided by the flexibility required, with a small number of subsets needed when solutions of (FIP) are (nearly) polynomials of order p, and by implementation issues, which can become substantial if intricate subsets are combined with constraints of the type given in the remainder of this section as well as other.

Continuity. There may be a need to limit the consideration to solution functions that are continuous on the whole or parts of $I\!R^n$. For a partition $\mathcal{R} = \{R_k\}_{k=1}^N$, an epi-spline defined in terms of the polynomials $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, 2, ..., N$, is continuous on the boundary between R_k and R_l, $k \neq l$, if

$$q_k(x) = q_l(x) \;\text{ for all } x \in \text{cl}\, R_k \cap \text{cl}\, R_l. \tag{1}$$

Certainly, continuity in the case of $p = 0$ implies that $q_k = q_l$, but continuity for intricate R_k and R_l can also force $q_k = q_l$, which may not always be desirable. Less trivial examples are provided by partitions consisting of subsets R_k, $k = 1, ..., N$, that are defined by a finite number of unions and intersections of halfspaces. We refer to such a subset as a *finite polytope*. The next proposition supports the implementation of constraints with this structure.

5.1 Proposition *Suppose that R_k and R_l are finite polytopes with a common facet defined by $\langle c, x\rangle = b$, and $q_k, q_l \in \text{poly}^p(I\!R^n)$. Then, the constraints*

$$q_k(x) = q_l(x) \;\textit{ for all } x \in \text{cl}\, R_k \cap \text{cl}\, R_l \textit{ with } \langle c, x\rangle = b \tag{2}$$

are equivalent to $(n-1+p)!/((n-1)!p!)$ equality constraints on the coefficients of q_k and q_l.

Proof: The constraint $q_k(x) = q_l(x)$ is equivalent to $q(x) = 0$, with $q = q_k - q_l \in \text{poly}^p(I\!R^n)$. Since a polynomial in $\text{poly}^p(I\!R^n)$ vanishes on an infinite number of points if and only if all its coefficients are zeros, the conclusion follows after observing that $(n-1+p)!/((n-1)!p!)$ is the number of coefficients for polynomials in $\text{poly}^p(I\!R^{n-1})$, where the dimensional reduction is caused by the restriction to the facet. □

Naturally, the expressions for the finite number of equality constraints in Proposition 5.1 can get involved and we here provide specifics only for the simple, but useful case $n = p = 2$ with a partition consisting of rectangles aligned with the coordinate axes. We let $q_k(x) = a_0^k + a_1^k x_1 + a_2^k x_2 + a_{12}^k x_1 x_2 + a_{11}^k x_1^2 + a_{22}^k x_2^2$ and similarly for q_l. For a facet defined by $x_i = b$, $i = 1$ or 2, we obtain the three constraints

$$\begin{aligned} a_0^k - a_0^l + (a_i^k - a_i^l)b + (a_{ii}^k - a_{ii}^l)b^2 &= 0 \\ a_j^k - a_j^l + (a_{12}^k - a_{12}^l)b &= 0 \\ a_{jj}^k - a_{jj}^l &= 0 \end{aligned}$$

with $j = 2$ when $i = 1$ and $j = 1$ when $i = 2$.

Continuous Differentiability. Partial derivatives of a polynomial in $\text{poly}^p(I\!R^n)$, $p \in I\!N$, is a polynomial in $\text{poly}^{p-1}(I\!R^n)$ and continuous differentiability simply requires continuity of those derivatives. Consequently, we ensure that an epi-spline is continuously differentiable by imposing the conditions of the previous paragraph for each partial derivative. For the example with $n = p = 2$, each facet requires two constraints per partial derivative for a total of four constraints. Obviously, higher order differentiability follows the same pattern.

Convexity. The convexity of an epi-spline, in general, requires the convexity of the polynomials on R_1, ..., R_N, and also the proper behavior on the boundary between such subsets. The following provides a supporting result, which allows for simplifications if an epi-spline is continuously differentiable.

5.2 Proposition *Suppose that $s \in \text{e-spl}_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$, is defined in terms of the polynomials $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, ..., N$. Then, s is convex if*

(i) q_k are convex on convex sets R_k, $k = 1, ..., N$,

(ii) for all $x \in \text{cl}\, R_k \cap \text{cl}\, R_l$, $y \in R_k$, and $k, l \in \{1, 2, ..., N\}$,

$$q_k(x) = q_l(x) \text{ and } \langle \nabla q_k(x) - \nabla q_l(x), y - x \rangle \geq 0.$$

If s is continuously differentiable, then item (ii) is satisfied automatically.

Proof: Since q_k is continuous on $I\!R^n$, the convexity on R_k implies convexity on $\text{cl}\, R_k$. Consequently, for $x \in \text{cl}\, R_k \cap \text{cl}\, R_l$, $y \in R_k$,

$$\begin{aligned} q_k(y) &\geq q_k(x) + \langle \nabla q_k(x), y - x \rangle \\ &= q_l(x) + \langle \nabla q_k(x) - \nabla q_l(x), y - x \rangle + \langle \nabla q_l(x), y - x \rangle \\ &\geq q_l(x) + \langle \nabla q_l(x), y - x \rangle. \end{aligned}$$

We next establish the convexity of s by showing that for $x \in \text{cl}\, R_l$ and $y \in \text{cl}\, R_k$, $x \neq y$ and $k \neq l$,

$$q_k(y) \geq q_l(x) + \langle \nabla q_l(x), y - x \rangle.$$

We start with the case when the line segment $\{z \in I\!R^n \mid z = \alpha(y - x) + x, \alpha \in (0, 1)\}$ intersects $\text{cl}\, R_l$ and $\text{cl}\, R_k$, and not $\text{cl}\, R_{l'}$ for $l' \neq l, k$. Then, there exists an $\bar{\alpha} \in (0, 1)$ such that $z = \bar{\alpha}(y - x) + x$ is in both $\text{cl}\, R_l$ and $\text{cl}\, R_k$. From above and the convexity of q_l, we obtain that

$$\begin{aligned} q_k(y) &\geq q_l(z) + \langle \nabla q_l(z), y - z \rangle \\ &\geq q_l(x) + \langle \nabla q_l(x), z - x \rangle + \langle \nabla q_l(z), y - z \rangle. \end{aligned}$$

The convexity of q_l also ensures that

$$\begin{aligned} &\langle \nabla q_l(z), y - z \rangle - \langle \nabla q_l(x), y - z \rangle \\ =&\langle \nabla q_l(z) - \nabla q_l(x), y - z \rangle \\ =&(1/\bar{\alpha} - 1)\langle \nabla q_l(z) - \nabla q_l(x), z - x \rangle \geq 0, \end{aligned}$$

because $(1/\bar{\alpha} - 1) > 0$. Consequently,

$$\begin{aligned} q_k(y) &\geq q_l(x) + \langle \nabla q_l(x), z - x \rangle + \langle \nabla q_l(x), y - z \rangle \\ &= q_l(x) + \langle \nabla q_l(x), y - x \rangle, \end{aligned}$$

which establishes the condition for the stated case. The general case is established by applying the above arguments repeatedly at each of the finite number of points on the line segment $\{z \in I\!R^n \mid z = \alpha(y - x) + x, \alpha \in (0,1)\}$ where s is not continuously differentiable.

We next turn to the second claim. If s is continuously differentiable, then $q_k(x) = q_l(x)$ and $\nabla q_k(x) = \nabla q_l(x)$ for all $x \in \operatorname{cl} R_k \cap \operatorname{cl} R_l$ and $k, l \in \{1, 2, ..., N\}$, which confirms that the second conclusion holds. $\square$

We observe that item (ii) in Proposition 5.2 relates to the monotonicity of subgradients of s.

Monotonicity. We say that a function $f : I\!R^n \to I\!R$ is nondecreasing if

$$f(x) \leq f(y) \text{ whenever } x \leq y,$$

where the last inequality is interpreted componentwise. For an epi-spline $s \in \text{e-spl}_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$ and polynomials $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, ..., N$, to be nondecreasing, it is obviously needed that $q_k(x) \leq q_k(y)$ for $x, y \in R_k$ with $x \leq y$. Since q_k is differentiable, this is ensured by $\nabla q_k(x) \geq 0$ for $x \in R_k$. If $p = 2$ and R_k is a finite polytope, then ∇q_k is affine and it suffices to impose $\nabla q_k(x) \geq 0$ for vertices x of $\operatorname{cl} R_k$. It is clear that under continuity, boundary points are immaterial:

5.3 Proposition *If $s \in \text{e-spl}_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$, is defined in terms of the polynomials $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, ..., N$, and is continuous, then s is nondecreasing whenever q_k, $k = 1, ..., N$, are nondecreasing.*

In general, however, an epi-spline may not be nondecreasing even if q_k, $k = 1, ..., N$, are nondecreasing. If R_k, $k = 1, ..., N$, are finite polytopes consisting of *boxes*, i.e., sets of the form $R_k = \{x \in I\!R^n \mid l_k \leq x \leq u_k\}$, then the following proposition provides guidance, where we say that R_k *precedes* R_l if there exist $x \in R_k$ and $y \in R_l$ with $x \leq y$.

5.4 Proposition *If $s \in \text{e-spl}_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$ consisting of boxes, is defined in terms of the polynomials $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, ..., N$, then s is nondecreasing if*

(i) q_k, $k = 1, ..., N$, are nondecreasing

(ii) $q_k(x) \leq q_l(x)$ for $x \in \operatorname{cl} R_k \cap \operatorname{cl} R_l$ and R_k preceding R_l.

A nonincreasing epi-spline is treated similarly with the appropriate reversal of inequalities.

Bounds. The nonnegative of an epi-spline requires conditions that ensure the nonnegativity of polynomials, which except for orders 0 and 1 are nontrivial. A composition, as discussed in §4, constructed from the exponential function guarantees nonnegativity automatically. Specifically, a problem optimizing over $h \in \text{lsc-fcns}(\mathbb{R}^n)$, $h \geq 0$, can be reformulated as one over $f \in \text{lsc-fcns}(\mathbb{R}^n)$ by setting $h = e^f$. Since $f = -\infty$ is included in $f \in \text{lsc-fcns}(\mathbb{R}^n)$ and corresponds to $h = 0$, the reformulation is equivalent. Consequently, instead of approximating the function h of the original problem by $s \in \text{e-spl}_n^p(\mathcal{R})$, $s \geq 0$, we can approximate the function f of the reformulated problem by $s \in \text{e-spl}_n^p(\mathcal{R})$. Nonpositivity and other bounds are treated similarly.

Log-concavity. A function $h : \mathbb{R}^n \to [0, \infty)$, with $h > 0$ on a convex set X, is log-concave on X if $\log h$ is concave on X. Again a composition is convenient. A problem optimizing over $h \in \text{lsc-fcns}(\mathbb{R}^n)$, $h \geq 0$ and log-concave on X, can be reformulated as one over $f \in \text{lsc-fcns}(\mathbb{R}^n)$, f concave on X, by setting $h = e^f$. The concavity constraint is then handled by ensuring convexity, as described above, of the negative of the corresponding approximating epi-spline.

Integral conditions. Requirements that functions $f \in \text{lsc-fcns}(\mathbb{R})$ in (FIP) should satisfy[7]

$$l \leq \int_X f(x)dx \leq u$$

for an open set $X \subset \mathbb{R}^n$ is handled by ensuring that an epi-spline given by polynomials q_k, $k = 1, ..., N$, and partition $\mathcal{R} = \{R_k\}_{k=1}^N$ satisfies

$$l \leq \sum_{k=1}^{N} \int_{R_k \cap X} q_k(x)dx \leq u.$$

Due to the polynomial forms, the integrals are here easily computed analytically whenever the descriptions of X and R_k are "simple."

Proximity. Given an epi-spline $s^0 \in \text{e-spl}_n^p(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$ and corresponding polynomials $q_k^0 \in \text{poly}^p(\mathbb{R}^n)$, $k = 1, ..., N$, applications may require epi-splines in $\text{e-spl}_n^p(\mathcal{R})$ that are "close" to s^0. Constraints on norms between s^0 and $s \in \text{e-spl}_n^p(\mathcal{R})$, given by $q_k \in \text{poly}^p(\mathbb{R}^n)$, $k = 1, ..., N$, is facilitated by the expression

$$\|s - s^0\|_m^m := \int |s(x) - s^0(x)|^m dx = \sum_{k=1}^{N} \int_{R_k} |q_k - q_k^0|^m dx.$$

If m is even, then the right-most integrals are easy to compute analytically as the integrand is polynomial as long as R_k is "simple."

Subgradient bounds. We recall the notion of *subgradients* of a function $f : \mathbb{R}^\nu \to \overline{\mathbb{R}}$, where we need the notation $x^\nu \underset{f}{\to} x$ to denote a sequence $x^\nu \to x$ that also satisfies $f(x^\nu) \to f(x)$. A comprehensive

[7]Since $f \in \text{lsc-fcns}(\mathbb{R}^n)$, the level sets $\{x \in \mathbb{R} \mid f(x) \leq \gamma\}$ are closed. Consequently, f is measurable and for open sets $X \subset \mathbb{R}^n$, the integrals $\int_X f_+(x)dx$ and $\int_X f_-(x)dx$, with $f_+ = \max\{0, f\}$ and $f_- = \max\{0, -f\}$, are well-defined, but possibly infinite. With the usual conventions $\infty - a = \infty$, $a - \infty = -\infty$, and $\infty - \infty = \infty$, $\int_X f(x)dx = \int_X f_+(x)dx - \int_X f_-(x)dx$ is therefore well-defined.

treatment of the topic is found in [26, Chapter 8]. For a function $f : I\!R^n \to \overline{I\!R}$ and $v, x \in I\!R^n$ with $f(x)$ finite, we say that

(i) v is a regular subgradient of f at x, if

$$\liminf_{y \to x, y \neq x} \frac{f(y) - f(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0,$$

with the set of all regular subgradients denoted by $\hat{\partial} f(x)$;

(ii) v is a subgradient of f at x if there are sequences $x^\nu \underset{f}{\to} x$ and $v^\nu \to v$, $v^\nu \in \hat{\partial} f(x^\nu)$, with the set of all subgradients denoted by $\partial f(x)$.

Connections between the subgradients are summarized next.

5.5 Proposition *Suppose that $s \in$ e-spl$^p_n(\mathcal{R})$ and $x \in I\!R^n$. Then, $\hat{\partial} s(x)$ and $\partial s(x)$ are closed subsets of $I\!R^n$, $\hat{\partial} s(x)$ is convex, and*

$$\hat{\partial} s(x) \subset \partial s(x) \neq \emptyset,$$

with equality holding if s is regular[8].

Proof: The only part that requires a proof is the claim about nonemptiness. Let $\mathcal{R} = \{R_k\}_{k=1}^N$ and s be defined in terms of $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, ..., N$. For given $x \in I\!R^n$, let $k \in \{1, ..., N\}$ be such that $s(x) = q_k(x)$ and $x \in \text{cl}\, R_k$. There exists a sequence $x^\nu \to x$ with $x^\nu \in R_k$. Since $\nabla q_k(x^\nu) \to \nabla q_k(x)$ and $\{\nabla q_k(x^\nu)\} = \hat{\partial} s(x^\nu)$, $\nabla q_k(x) \in \partial s(x)$ and $\partial s(x)$ is therefore nonempty. The remaining claims are direct consequences of [26, Theorem 8.6 and Corollary 8.11] due to the finiteness and lower semicontinuity of s. $\square$

Constraints on subgradients are supported by the following results.

5.6 Theorem (subgradients) *Suppose that $s \in$ e-spl$^p_n(\mathcal{R})$, with $\mathcal{R} = \{R_k\}_{k=1}^N$ consisting of Clarke regular*[9] *sets $\text{cl}\, R_k$, $k = 1, 2, ..., N$, is continuous and defined in terms of $q_k \in \text{poly}^p(I\!R^n)$, $k = 1, ..., N$. Then, for a nonempty and closed $S \subset I\!R^n$,*

$$\nabla q_k(x) \in S \text{ for } x \in R_k, k = 1, ..., N \iff \partial s(x) \subset S \text{ for } x \in I\!R^n.$$

Proof: Since $\partial s(x) = \{\nabla q_k(x)\}$ for $x \in R_k$, the implication from right to left follows trivially. Next, we consider the converse. Let $x \in I\!R^n$ and $\hat{\partial} s(x) \neq \emptyset$. If $\hat{\partial} s(x) = \{v\}$, i.e., is a singleton, then there exists a k and $x^\nu \to x$ with $x^\nu \in R_k$ and $\nabla q_k(x^\nu) \to v$. Since $\nabla q_k(x^\nu) \in S$ and S is closed, it follows that $v \in S$. Now, suppose that $\hat{\partial} s(x)$ is not a singleton and let $v \in \hat{\partial} s(x)$ be arbitrary. For the sake

[8] We recall that a function $f : I\!R^n \to I\!R$ is regular if for every $x \in I\!R^n$, epi f is Clarke regular at $(x, f(x))$; see the next footnote and [26, Definitions 6.4,7.25]. In particular, if f is convex, then it is regular.

[9] A set $A \subset I\!R^n$ is Clarke regular if at all $x \in A$, (i) $A \cap \mathcal{B}$ is closed for some closed neighborhood $\mathcal{N}$ of x and (ii) every normal vector v of A at x is regular, i.e., $\langle v, y - x \rangle \leq o(\|y - x\|)$ for $y \in A$. For example, if A is locally convex at x for all $x \in A$, then A is regular; see [26, Definitions 6.3 and 6.4].

of a contradiction, suppose that $v \notin S$. Then, since $\operatorname{cl} R_k$, $k = 1, ..., 2$, are Clarke regular there exist $k^* \in \{1, ..., N\}$, $h \in I\!R^n$, $\delta > 0$, and $t^* > 0$ such

$$x + th \in R_{k^*} \text{ and } s(x + th) = q_{k^*}(x + th) \text{ for all } t \in [0, t^*]$$

and

$$\langle v - \nabla q_{k^*}(x), h \rangle \geq \delta \|h\|,$$

where the last inequality is a consequence from the fact that k^* can be selected such that $v \neq \nabla q_{k^*}(x)$. Let $\varepsilon \in (0, \delta)$ and $t_\varepsilon \in (0, t^*)$ be such that

$$|q_{k^*}(x + th) - q_{k^*}(x) - \langle \nabla q_{k^*}(x), th \rangle| \leq \varepsilon t \|h\| \text{ for all } t \in (0, t_\varepsilon),$$

which follows from the smoothness of q_{k^*}. Using these facts, we obtain that

$$\begin{aligned} s(x) + \langle v, th \rangle &= s(x) + \langle v - \nabla q_{k^*}(x), th \rangle + \langle \nabla q_{k^*}(x), th \rangle \\ &\geq q_{k^*}(x) + \langle \nabla q_{k^*}(x), th \rangle + \delta t \|h\| \\ &\geq s(x + th) + (\delta - \varepsilon) t \|h\| \end{aligned}$$

for all $t \in (0, t_\varepsilon)$. Consequently,

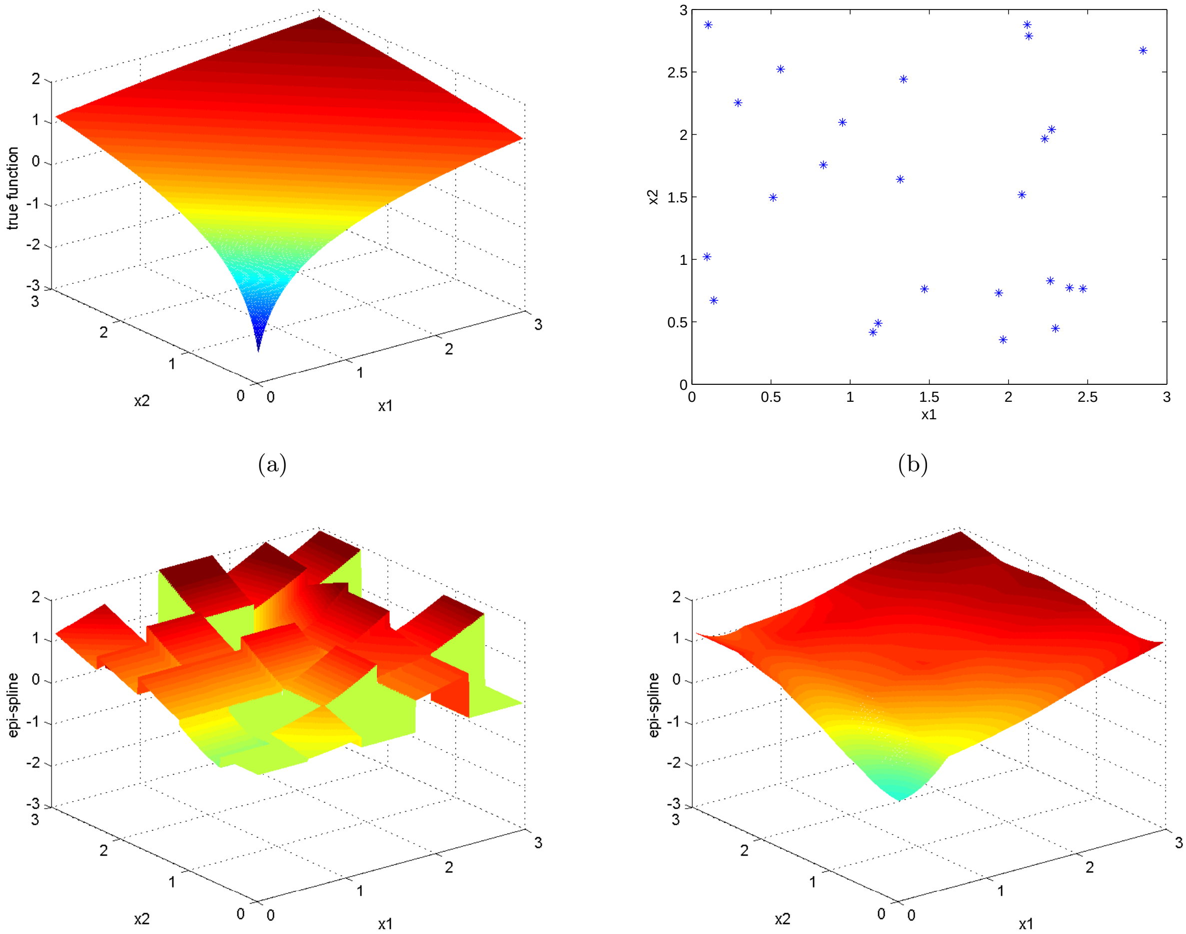
$$\frac{s(x + th) - s(x) - \langle v, th \rangle}{t\|h\|} \leq -(\delta - \varepsilon) < 0$$

for all $t \in (0, t_\varepsilon)$, which contradicts the assumption that v is a regular subgradient. Since the situation with $\hat{\partial} s(x) = \emptyset$ is trivial, this establishes that $\hat{\partial} s(x) \subset S$ for all $x \in I\!R^n$. Finally, we consider $v \in \partial s(x)$ for an arbitrary $x \in I\!R^n$. Then, by definition, there exists $x^\nu \to x$ and $v^\nu \to v$, with $v^\nu \in \hat{\partial} s(x^\nu)$. Since $\hat{\partial} s(x^\nu) \subset S$ for all ν, $v \in S$ due to the closeness of S, which completes the proof. □

In view of the preceding results, it is clear that applications demanding constraints on the size of subgradients of an epi-spline $s \in \text{e-spl}_n^p(\mathcal{R})$ are satisfied by imposing constraints on gradients of the corresponding polynomials $q_k \in \text{poly}^p(I\!R^n)$ on R_k. The partial derivatives of q_k would then be in $\text{poly}^{p-1}(I\!R^n)$, which generally would require an infinite number of constraints to ensure the inclusion in a set for all $x \in R_k$. However, if $p \leq 2$ and R_k is a finite polytope, then it suffices to enforce the constraints at the vertices of $\operatorname{cl} R_k$.

6 Applications

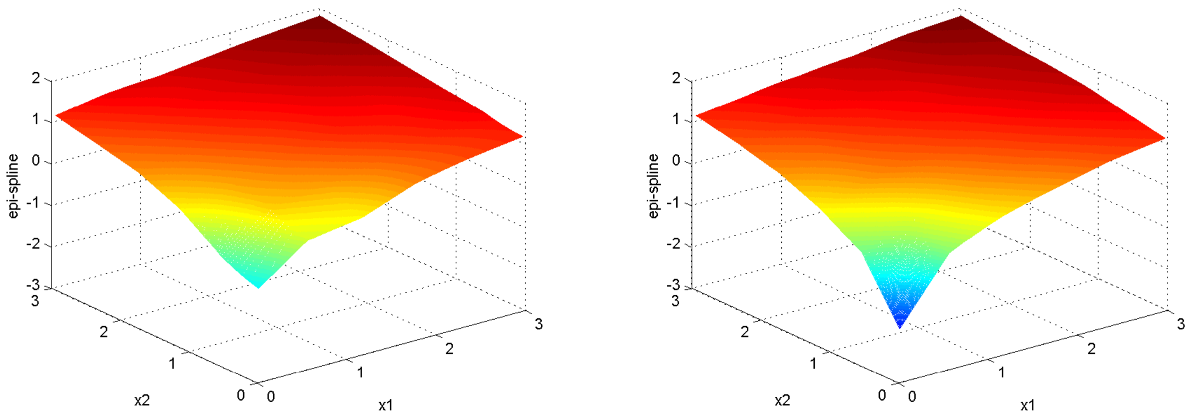
We illustrate epi-splines through a series of examples arising in response surface construction and probability density estimation. Focusing on second-order epi-splines in two, occasionally three, dimensions, it suffices to represent polynomials using the bases $(1, x_1, x_2, x_1^2, x_1 x_2 x_2^2)$ and $(1, x_1, x_2, x_3, x_1^2, x_1 x_2, x_2^2, x_1 x_3, x_2 x_3, x_3^2)$ for two and three dimensions, respectively. Numerical examples in one dimension are found in [30, 29, 28, 33]. We only consider partitions of domains of interest consisting of rectangles of equal size. Though, the number of rectangles varies. CPLEX (12.5.1.0) and CONOPT (3.15L) solve resulting linear and nonlinear programs, respectively, on a 64-bit Windows 7 laptop running at 2.60GHz, with 4GB RAM, after they are formulated in GAMS (24.1.3).



(a) (b)

(c) (d)

Figure 2: Logarithm example: actual function (a), randomly generated data (b), lsc epi-spline (c), and continuous epi-spline (d).



(a)

(b)

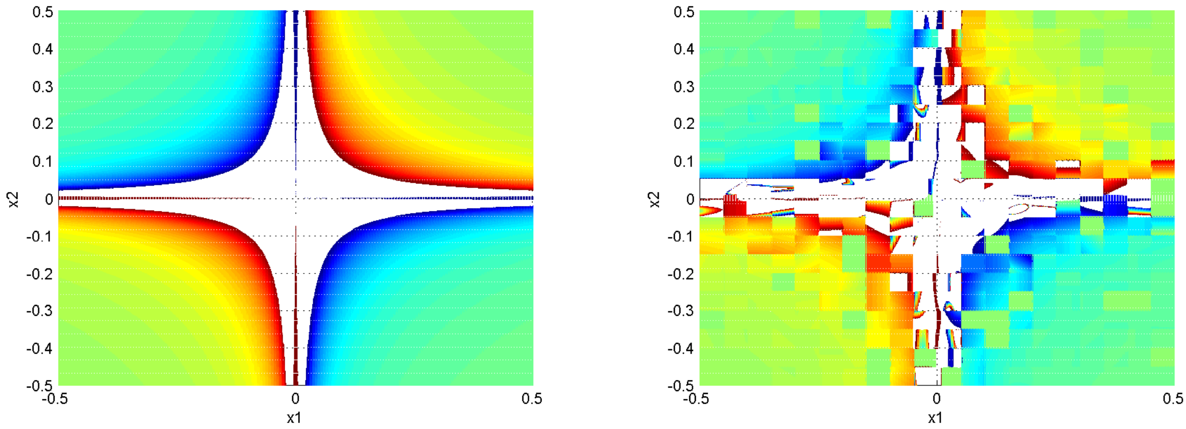
Figure 3: Logarithm example: nondecreasing continuous epi-splines (a) and also with $s(0.1, 0.1) \leq -1.5$ (b).

6.1 Response Surface

We reconstruct four functions based on observations of the functions at a finite number of points and extrinsic information about the smoothness of the functions and other factors.

Logarithmic function. Suppose that $f(x) = \log(x_1 + x_2)$ defined on the domain $[0, 3]^2$; see Figure 2(a). Relying on observed function values at 25 randomly generated points according to a uniform distribution on $[0, 3]^2$ (see Figure 2(b)) and a partition with $N = 25$ open rectangles, we obtain the least-squares minimizing epi-splines of Figures 2(c), 2(d), and 3. All the fits achieve essentially a zero error at the 25 data points, but the values of the epi-splines at other points depend on the extrinsic information included. Figure 2(c) shows the fit for "unconstrained" lsc epi-splines. The fit improves significantly when constrained to continuous epi-splines as described in §5; see Figure 2(d). We achieve further improvement after additionally restricting to the nondecreasing epi-splines (Figure 3(a)) and to a function value of no more than -1.5 at $(0.1, 0.1)$ (Figure 3(b)). The actual value $f(0.1, 0.1) = -1.6$.

Inverse function. Suppose that $f(x) = 1/(x_1 x_2)$ defined on the domain $[-0.5, 0.5]^2$; see Figure 4(a) for a color contour plot, with red and blue indicating areas with high and low function values, respectively, and white corresponding to values above 100 or below -100. We adopt a minimum absolute deviation criterion and unconstrained lsc epi-splines. For $N = 400$ and 900 randomly generated data points from a uniform distribution on the domain, we obtain the epi-spline of Figure 4(b). The total absolute deviation across the data points is only 6.0, but errors appear elsewhere as the "granular" picture indicates. Figure 5 shows epi-splines based on 2500 data points. In some sense, the fit improves as indicated by Figure 5(a). However, total absolute deviation increases to $6 \cdot 10^5$ as the partition is not fine enough to capture the large variation of f near the origin. The situation improves with a finer parti-



(a) (b)

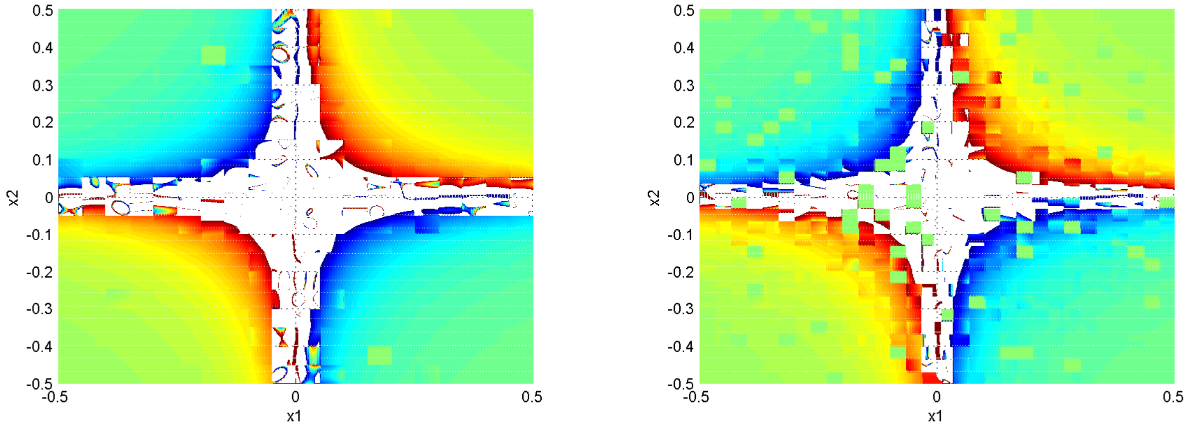
Figure 4: Inverse example: actual function (a) and epi-spline from 900 points (b).

tion of $N = 900$, which results in an essentially negligible total absolute deviation of 1.4; see Figure 5(b).

Trigonometric function. Suppose that $f(x) = (\cos(\pi x_1) + \cos(\pi x_2))^3$ defined on $[-3, 3]^2$; see Figure 6(a). We rely on continuous epi-splines and a partition with $N = 400$. Based on 900 uniformly distributed data points and a max-deviation criterion, we obtain the epi-spline fit of Figure 6(b). Maximum and average errors across the data points are 0.415 and 0.313, respectively. Mean-square and absolute deviation criteria give similar results. A relaxation of the continuity requirement results in essentially perfect interpolation at the expense of a more "rugged" fit.

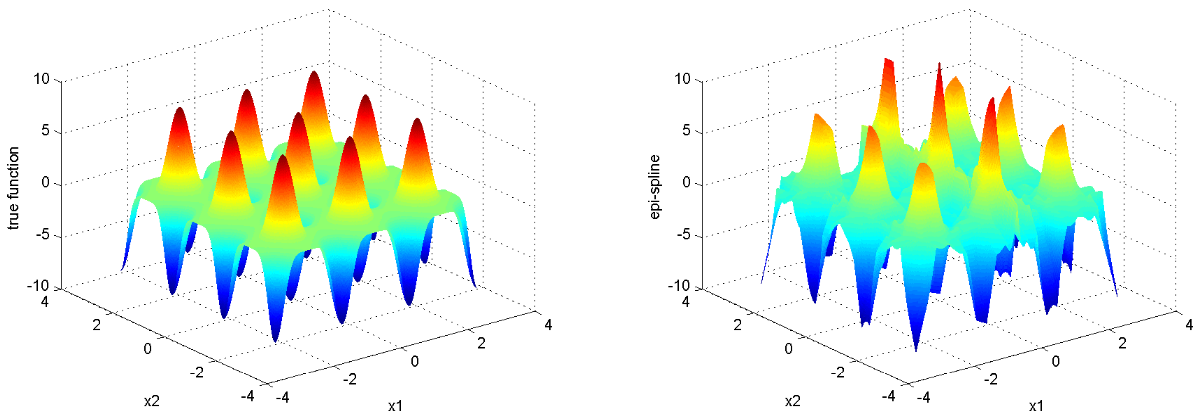
We also examine the 3-dimensional function $f(x) = (\cos(\pi x_1)+\cos(\pi x_2)+\cos(\pi x_3))^3$ defined on $[-3, 3]^3$ and 27,000 data points uniformly generated. Again relying on continuous epi-splines, but now with a partition using $N = 8000$, we obtain maximum and average errors over the data points of 5.16 and 1.68, respectively. The resulting linear program consists of 80,001 variables, 136,800 equality constraints (including redundancies), and 54,000 inequality constraints (reduced dual LP actually solved has 27,086 rows, 111,402 columns) and solves in 27 seconds using the "barrier" option in CPLEX. A switch to an absolute deviation criterion yields maximum and average errors of 7.74 and 1.49, respectively, and solves in 69 seconds. When relaxing the continuity constraints, errors are driven to near zero.

Sinc function. Suppose that $f(x) = \sin(\pi\|x\|)/(\pi\|x\|)$ for $x \neq 0$ and $f(x) = 1$ otherwise. We use the least-squares criterion and continuously differentiable epi-splines. Figure 7 depicts the actual function in part (a) as well as epi-splines estimates based on $N = 400$ and 900 randomly generated data points from a uniform distribution on $[-5, 5] \times [-5, 5]$ in part (b) and $N = 225$ and 600 points in part (c). The mean-squared error in both cases is 0.25.



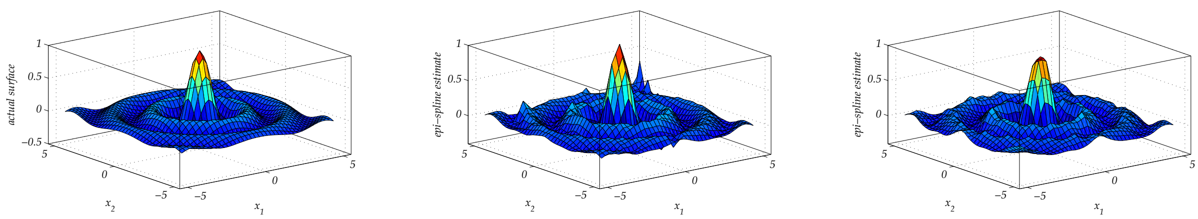
(a) (b)

Figure 5: Inverse example: using 2500 points, epi-splines with $N = 400$ (a) and $N = 900$ (b).



(a) (b)

Figure 6: Trigonometric example: actual function (a) and epi-spline (b).

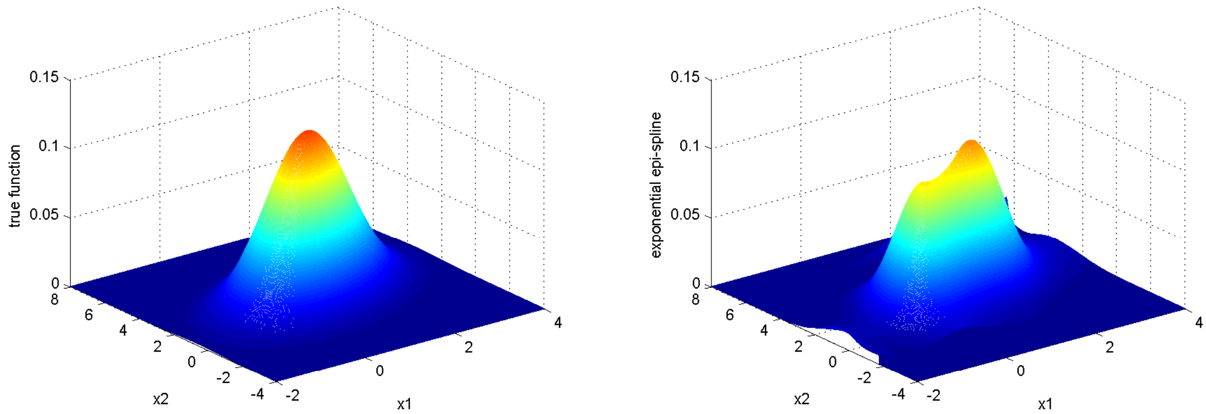


(a) (b) (c)

Figure 7: Sinc example: actual function (a) and epi-spline from 900 (b) and 600 (c) points.

6.2 Probability Density Estimation

Probability density estimation is another arena where epi-splines show promise. We refer to [29] for an in-depth study of one-dimensional density estimation. We here concentrate on two dimensions. Since densities are nonnegative functions, we rely on composite epi-splines $\text{c-spl}_2^2(\mathcal{R}, \exp)$, i.e., *exponential epi-splines* of the form $h = e^{-s}$, with $s \in \text{e-spl}_2^2(\mathcal{R})$.



(a)

(b)

Figure 8: Normal Example: actual density (a) and continuously differentiable exponential epi-spline for sample size 100 (b).

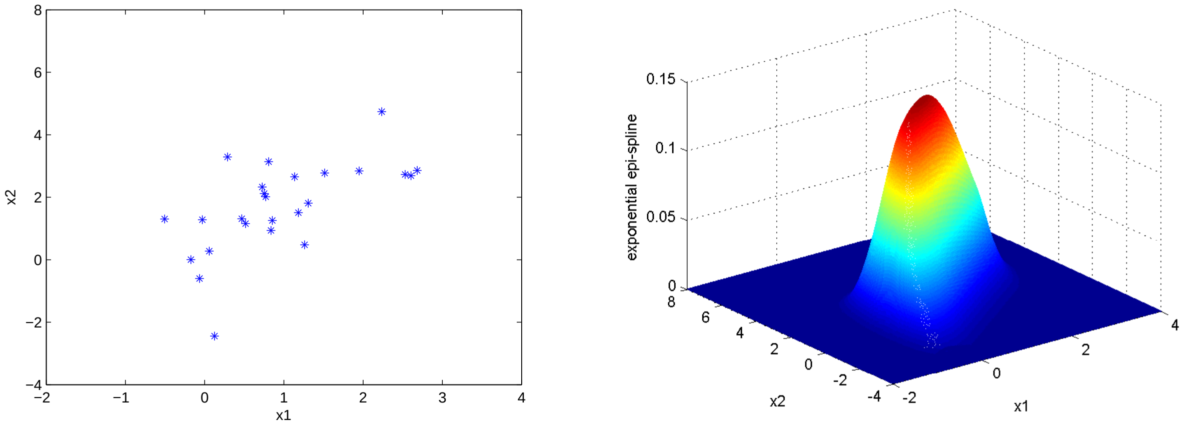
For a sample $X^1, X^2, ..., X^\nu$ that is independently and identically distributed as a "actual distribution," it is well-known (see for example [29]) that a constrained maximum likelihood estimator is an optimal solution of

$$\max \frac{1}{\nu} \prod_{i=1}^{\nu} f(X^i)^{1/\nu} \text{ such that } f \geq 0, \int f(x)dx = 1, f \in \mathcal{F},$$

where $\mathcal{F}$ is an appropriately selected space of functions on $I\!\!R^n$. Passing to exponential epi-splines, we arrive after equivalently maximizing the logarithm of the objective function at an approximate problem

$$\min \frac{1}{\nu} \sum_{i=1}^{\nu} s(X^i) \text{ such that } \int e^{-s(x)}dx = 1, s \in F^\nu \cap \mathcal{S}^\nu$$

where $\mathcal{S}^\nu = \text{e-spl}_n^p(\mathcal{R}) \cap \text{lsc-fcns}(I\!\!R^n)$ and F^ν is a subset of the lsc functions s that satisfies $\int e^{-s(x)}dx = 1$, but could include numerous other restrictions exemplified below and in §5. A solution s of this problem provides a density estimator through the composition e^{-s}, where we observe that the nonnegativity is automatically satisfied. We refer to [29] for further details including simplifications in the approximate problem that ensure its convexity. Two numerical examples illustrate the approach.

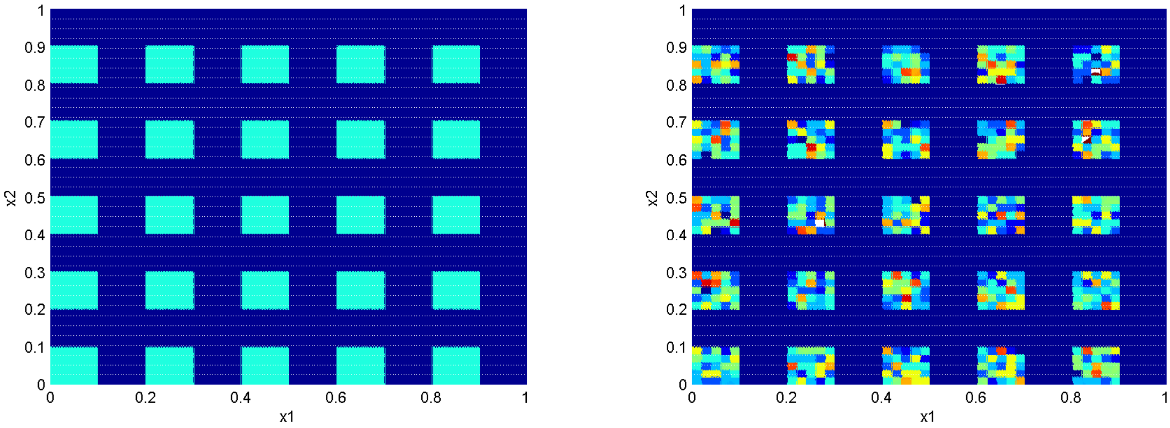


(a) (b)

Figure 9: Normal Example: Sample of size 25 (a) and corresponding exponential epi-spline (b).

Normal density example. Suppose that the actual density h^0 is a bivariate normal with mean vector $(1, 2)$ and variance-covariance matrix $(1, 0.5; 0.5, 2)$; see Figure 8(a). We would like to reproduce this density using a sample of size ν. The partition $\mathcal{R}$ consists of $N = 100$ rectangles of equal size covering a domain defined in each dimension to range from two empirical standard deviations below the lowest observed value to two empirical standard deviations above the highest value. In addition to constraints specified below, we let second-order partial derivatives of the epi-splines to be in the range $[-1000, 1000]$. For a sample size of $\nu = 10,000$, exponential epi-spline estimates are visually good (not displayed) regardless of the combination of additional constraints on continuity, continuous differentiability, and log-concavity with mean-square errors MSE $= \int (e^{-s(x)} - h^0(x))^2 h^0(x) dx$ of approximately 40. A reduced sample size of $\nu = 100$, provides a MSE of $6.1 \cdot 10^5$ and a poor visual fit under a continuity constraint. Under continuous differentiability, the MSE improves to $2.1 \cdot 10^3$ and a good visual fit; see Figure 8(b). However, the estimate is not log-concave. An additional constraint would enforce such a condition easily, but instead of displaying that case we also reduce the sample size. Under the extremely small sample with $\nu = 25$ illustrated in Figure 9(a), a satisfactory fit with MSE of $5.3 \cdot 10^3$ is obtained under continuous differentiability and log-concavity constraints as seen in Figure 9(b). Of course, the normal density is especially well suited for estimation by second-order exponential epi-splines. In fact, it suffices to consider a partition consisting of a single open set, $I\!R^n$. We next consider a more challenging situation.

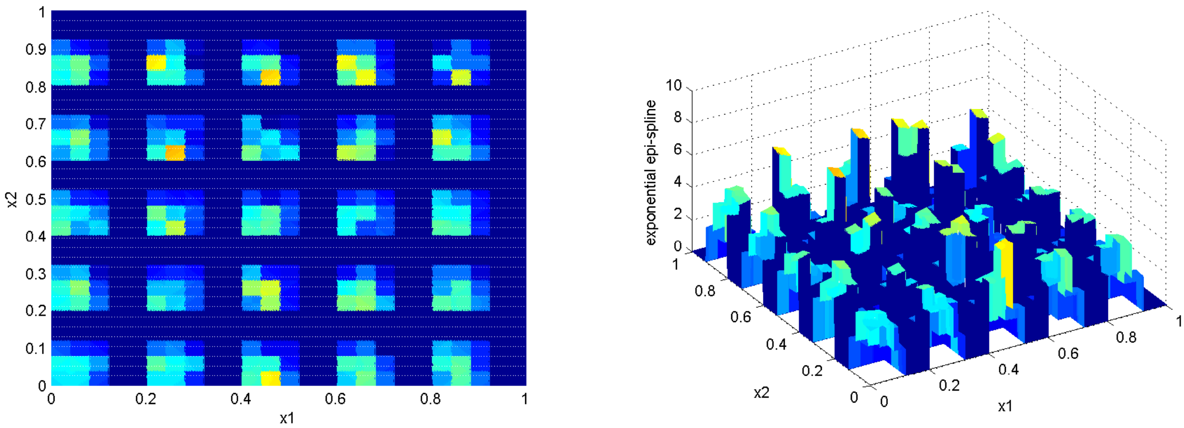
Uniform Mixture Example. Suppose that the actual density $h^0(x) = 4$ if $x = (x_1, x_2)$ satisfies $(k-1)0.2 \leq x_1 \leq (k-1)0.2 + 0.1$, $k = 1, 2, ..., 5$, and $(l-1)0.2 \leq x_2 \leq (l-1)0.2 + 0.1$, $l = 1, 2, ..., 5$, and $h^0(x) = 0$ otherwise; see Figure 10(a) for a color contour plot. This "uniform mixture" density is estimated by an exponential epi-spline defined on $[0, 1]^2$. Additional information about the support of the actual density is ignored. We assume that the partial derivatives of the epi-splines on the open sets



(a) (b)

Figure 10: Uniform Example: True density (a) and exponential epi-spline with $N = 2500$ (b).

of the partition are in the range $[-1, 1]$ and rely on a sample of size $\nu = 2500$. Figure 10(b) shows an exponential epi-spline estimate for a partition with $N = 2500$. Although the MSE is large, the essential nature of the density is captured. The fit is obtained in 693 seconds after solving a problem with 15,000 variables, a convex objective function, and 40,000 linear inequality constraints[10]. Figure 11 provides similar results, obtained in 30 seconds, for $N = 625$ in a contour plot (a) and a titled view (b).



(a) (b)

Figure 11: Uniform Example: Exponential epi-spline with $N = 625$ in contour (a) and tilted view (b).

[10]We refer to [29] for a convex formulation that removes the constraint that ensures that the estimate integrates to one and replaces it by a penalty term in the objective function.

References

[1] M. Atteia. Généralisation de la définition et des propriétés de "splines-fonctions". *Comptes Rendus de l'Académie des Sciences de Paris*, 260:3350–3553, 1965.

[2] H. Attouch and R. Wets. A convergence theory for saddle functions. *Transactions of the American Mathematical Society*, 280(1):1–41, 1983.

[3] O. Bandorff-Nielsen. *Information and Exponential Families in Statistical Theory*. Wiley, 1978.

[4] G. Beer. *Topologies on Closed and Closed Convex Sets*, volume 268 of *Mathematics and its Applications*. Kluwer, 1992.

[5] A. Bezhaev and V. Vasilenko. *Variational Theory of Splines*. Kluwer, 2001.

[6] B. Bojanov, H. Hakopian, and S. Sahakian. *Spline Functions and Multivariate Interpolation*. Kluwer, 1993.

[7] R. Champion, C.T. Lenard, and T.M. Mills. A variational approach to splines. *ANZIAM Journal*, 42:119–135, 2000.

[8] H.B. Curry and I.J. Schoenberg. On spline distributions and their limits: The Polya distribution functions. *Bulletin of the American Mathematical Society*, 53:1114, 1947.

[9] H.B. Curry and I.J. Schoenberg. On Polya frequency functions IV: The fundamental spline functions and their limits. *Journal d'Analyse Mathématique*, 17:71–107, 1966.

[10] W. Dahmen. On multivariate B-splines. *SIAM Journal on Numerical Analysis*, 17(2):179191, 1980.

[11] O. Davydov. Approximations by piecewise constants on convex partitions. *Journal of Approximation Theory*, 164(2):346–352, 2012.

[12] C. de Boor. Best approximation properties of spline functions of odd degree. *Journal of Mathematical Mechanics*, 12:747–749, 1963.

[13] C. de Boor, K. Höllig, and S. Riemenschneider. *Box splines*. Springer, 1993.

[14] I. C. Demetriou and M. J. D. Powell. Least squares smoothing of univariate data to achieve piecewise monotonicity. *IMA Journal of Numerical Analysis*, 11(3):411–432, 1991.

[15] I. C. Demetriou and M. J. D. Powell. The minimum sum of squares change to univariate data that gives convexity. *IMA Journal of Numerical Analysis*, 11(3):433–448, 1991.

[16] N. Dyn and A. Ron. Local approximation by certain spaces of exponential polynomials, approximation order of exponential box splines and related interpolation problems. *Transactions of the American Mathematical Society*, 319:381404, 1990.

[17] R. T. Farouki, C. Manni, M. L. Sampoli, and A. Sestini. Shape-preserving interpolation of spatial data by pythagorean-hodograph quintic spline curves. *IMA Journal of Numerical Analysis*, to appear, 2014.

[18] J.C. Holladay. A smoothest curve approximation. *Math. Tables Aids Comput.*, 11:233–243, 1957.

[19] P. D. Kaklis and D. G. Pandelis. Convexity-preserving polynomial splines of non-uniform degree. *IMA Journal of Numerical Analysis*, 10(2):223–234, 1990.

[20] M.-J. Lai and L. L. Schumaker. *Spline Functions on Triangulations*. Cambridge University Press, 2007.

[21] P.-J. Laurent. *Approximation et Optimization*, volume 13 of *Enseignement des Sciences*. Hermann, 1972.

[22] E. Mammen and C. Thomas-Agnan. Smoothing splines and shape restrictions. *Scandinavian Journal of Statistics*, 26(2):239–252, 1999.

[23] M. Meyer. Constrained penalized splines. *Canadian Journal of Satistics*, 40:190–206, 2012.

[24] C. A. Micchelli. On a numerically efficient method for computing multivariate B-splines. In W. Schempp et al., editor, *Multivariate Approximation Theory*, pages 211–248. Springer, Basel, 1979.

[25] G. Nürnberger. *Approximation by Spline Functions*. Springer, 1989.

[26] R.T. Rockafellar and R. Wets. *Variational Analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaft*. Springer, 3rd printing-2009 edition, 1998.

[27] A. Ron. Exponential box splines. *Constructive Approximation*, 150:357378, 1988.

[28] J.O. Royset, N. Sukumar, and R. J-B Wets. Uncertainty quantification using exponential epi-splines. In *Proceedings of the International Conference on Structural Safety and Reliability*, 2013.

[29] J.O. Royset and R. Wets. Fusion of soft and hard information in nonparametric density estimation. pre-print, University of California, Davis, 2013.

[30] J.O. Royset and R. Wets. On function identification problems. pre-print, University of California, Davis, 2014.

[31] J.O. Royset and R. J-B Wets. From data to assessments and decisions: Epi-spline technology. In A. Newman, editor, *INFORMS Tutorials*. INFORMS, Catonsville, 2014.

[32] L. I. Schumaker. On shape preserving quadratic spline interpolation. *SIAM Journal on Numerical Analysis*, 20(4):854–864, 1983.

[33] D.I. Singham, J.O. Royset, and R. J-B Wets. Density estimation of simulation output using exponential epi-splines. In *Winter Simulation Conference*, 2013.

[34] R. Sood and R. Wets. Information fusion. http://www.math.ucdavis.edu/ prop01, 2011.