HANDLING CONVEXITY-LIKE CONSTRAINTS IN VARIATIONAL PROBLEMS

QUENTIN MÉRIGOT AND ÉDOUARD OUDET

Abstract. We provide a general framework to construct finite dimensional approximations of the space of convex functions, which also applies to the space of c-convex functions and to the space of support functions of convex bodies. We give precise estimates of the distance between the approximation space and the admissible set. This framework applies to the approximation of convex functions by piecewise linear functions on a mesh of the domain and by other finite-dimensional spaces such as tensor-product splines. We show how these discretizations are well suited for the numerical resolution of problems of calculus of variations under convexity constraints. Our implementation relies on proximal algorithms, and can be easily parallelized, thus making it applicable to large scale problems in dimension two and three. We illustrate the versatility and the efficiency of our approach on the numerical resolution of three problems in calculus of variation : 3D denoising, the principal agent problem, and optimization within the class of convex bodies.

1. Introduction

Several problems in the calculus of variations come with natural convexity constraints. In optimal transport, Brenier theorem asserts that every optimal transport plan can be written as the gradient of a convex function, when the cost is the squared Euclidean distance. Jordan, Kinderlehrer and Otto showed [9] that some evolutionary PDEs such as the Fokker-Planck equation can be reformulated as a gradient flow of a functional in the space of probability densities endowed with the natural distance constructed from optimal transport, namely the Wasserstein space. In the corresponding time-discretized schemes, each timestep involves the resolution of a convex optimization problem over the set of gradient of convex functions. In a different context, the principal agent problem proposed by Rochet and Choné [20] in economy also comes with natural convexity constraints. Despite the possible applications, the numerical implementation of these variational problems has been lagging behind, mainly because of a non-density phenomenon discovered by Choné and Le Meur [5].

Choné and Le Meur discovered that some convex functions cannot be approximated by piecewise-linear convex functions on a regular grid (such as the grid displayed in Figure 1). More precisely, they proved that piecewise-linear convex functions on the regular grid automatically satisfy the inequality $\frac{\partial^2 f}{\partial x \partial y} \geqslant 0$ in a weak sense. Since there exists convex functions that do not satisfy this inequality, this implies that the union of the spaces of piecewise-linear convex functions on the regular grids $(G_\delta)_{\delta>0}$ is not dense in the space of convex functions on the unit square. Moreover, this difficulty is local, and

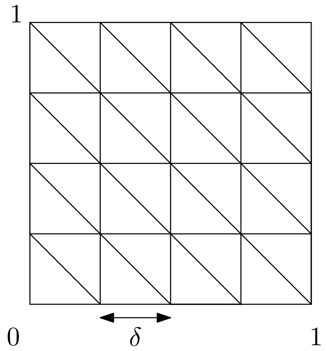

FIGURE 1. The regular grid (G_δ) of $[0,1]^2$. Convex functions that are piecewise linear on the triangles of this grid automatically satisfy the inequality $\frac{\partial^2 f}{\partial x \partial y} \geqslant 0$ in the sense of distributions [5].

it is likely that for any fixed sequence of meshes, one can construct convex functions f that cannot be obtained as limits of piecewise-linear convex functions on these meshes. This phenomenon makes it challenging to use P^1 finite elements to approximate the solution of variational problems with convexity constraints.

1.1. **Related works.** In this section, we briefly discuss approaches that have been proposed in the last decade to tackle the problem discovered by Choné and Le Meur.

Mesh versus grid constraints. Carlier, Lachand-Robert and Maury proposed in [4] to replace the space of P^1 convex functions by the space of the space of *convex interpolates*. For every fixed mesh, a piecewise linear function is a convex interpolate if it is obtained by linearly interpolating the restriction of a convex function to the node of the mesh. Note that these functions are not necessarily convex, and the method is therefore not interior. Density results are straightforward in this context but the number of linear constraints which have to be imposed on nodes values is rather large. The authors observe that in the case of a regular grid, one needs $\simeq m^{1.8}$ constraints in order to describe the space of convex interpolates, where m stands for the number of nodes of the mesh.

Aguilera and Morin [1] proposed a finite-difference approximation of the space of convex functions using discrete convex Hessians. They prove that it is possible to impose convexity by requiring a linear number of nonlinear constraints with respect to the number of nodes. The leading fully nonlinear optimization problems are solved using semidefinite programming codes. Whereas, convergence is proved in a rather general setting, the practical efficiency of this approach is limited by the capability of semidefinite solvers. In a similar spirit, Oberman [16] considers the space of function that satisfy local convexity constraints on a finite set of directions. By changing the size of the stencil, the author proposed different discretizations which lead to exterior or interior approximations. Estimations of the quality of the approximation is given for smooth convex functions.

Higher order approximation by convex tensor-product splines. An important number of publications have been dedicated last years to solve shape preserving least square problems. For instance, different sufficient conditions have been introduced to force the convexity of the approximating functions. Whereas this problem is well understood in dimension one, it is still an active field of research in the context of multivariate polynomials like Bézier or tensor spline functions. We refer the reader to Jüttler [10] and references therein for a detailed description of recent results. In this article, Jüttler describes an interior discretization of convex tensor-product splines. This approach is based on the so called "Blossoming theory" which makes it possible to linearize constraints on the Hessian matrix by introducing additional variables. Based on this framework, the author illustrates the method by computing the L^2 projection of some given function into the space of convex tensor-product splines. Two major difficulties have to be pointed out. First, the density of convex tensor splines in the space of convex functions is absolutely non trivial, and one may expect phenomena similar to those discovered by Choné and Le Meur. Second, the proposed algorithm leads to a very large number of linear constraints.

Dual approaches. Lachand-Robert and Oudet [12] developed a strategy related to the dual representation of a convex body by its support functions. They rely on a simple projection approach that amounts to the computation of a convex hull, thus avoiding the need to describe the constraints defining the set of support functions. To the best of our knowledge, this article is the first one to attack the question of solving problems of calculus of variations within convex bodies. The resulting algorithm can be interpreted as a non-smooth projected gradient descent, and gave interesting results on difficult problems such as Newton's or Alexandrov's problems. In a similar geometric framework, Oudet studied in [18] approximations of convex bodies based on Minkowski sums. It is well known in dimension two that every convex polygon can be decomposed as a finite sum of segments and triangles. While this result cannot be generalized extend to higher dimension, this approach still allows the generation of random convex polytopes. This process was used by the author to study numerically two problems of calculus of variations on the space of convex bodies with additional width constraints.

Ekeland and Moreno-Bromberg [6] proposed a dual approach for parameterizing the space of convex functions on a domain. Given a finite set of points S in the domain, they parameterize convex functions by their value f_s *and their gradient* v_s at those points. In order to ensure that these couples of values and gradients $(f_s, v_s)_{s\in S}$ are induced by a convex function, they add for every pair of points in S the constraints $f_t \geqslant f_s + \langle t-s|v_s\rangle$. This discretization is interior, and it is easy to show that the phenomenon of Choné and Le Meur does not occur for this type of approximation. However, the high number of constraints makes it difficult to solve large-scale problems using this approach. Mirebeau [15] is currently investigating an adaptative version of this method that would allow its application to larger problems.

1.2. **Contributions.** We provide a general framework to construct approximations of the space of convex functions on a bounded domain that satisfies a Lipschitz bound. Our approximating space is a finite-dimensional polyhedron, that is a subset of a finite-dimensional functional space that satisfies a finite number of linear constraints. The main theoretical contribution of this article is a bound on the (Hausdorff) distance between the approximating polyhedron and the admissible set of convex functions, which is summarized in Theorem 2.4. Our discretization is not specific to approximation by piecewise linear functions on a triangulation of the domain, and can easily be extended to approximations of convex functions within other finite-dimensional subspaces, such as the space of tensor product splines. This is illustrated numerically in Section 6.

This type of discretization is well suited to the numerical resolution of problems of calculus of variations under convexity constraints. For instance, we show how to compute the L^2 projection onto the discretized space of convex functions in dimension $d = 2, 3$ by combining a proximal algorithm [2] and an efficient projection operator on the space of 1D discrete convex functions. Because of the structure of the problem, these 1D projection steps can be performed in parallel, thus making our approach applicable to large scale problems in higher dimension. We apply our non-smooth approach to a denoising problem in dimension three in Section 5. Other problems of calculus of variations under convexity constraints, such as the principal-agent problem, can be solved using variants of this algorithm. This aspect is illustrated in Section 6.

Finally, we note in Section 3 that the discretization of the space of convex functions we propose can be generalized to other spaces of functions satisfying similar constraints, such as the space of support functions of convex bodies. The proximal algorithm can also be applied to this modified case, thus providing the first method able to approximate the projection of a function on the sphere onto the space of support functions of convex bodies. Section 7 presents numerical computations of L^p projections (for $p = 1, 2, \infty$) of the support function of a unit regular simplex onto the set of support functions of convex bodies with constant width. We believe that these projection operators could be useful in the numerical study of a famous conjecture due to Bonnesen and Fenchel (1934) concerning the so-called Meissner's convex bodies.

2. A relaxation framework for convexity

In this section, we concentrate on the relaxation of usual convexity for clarity of exposition. However, most of the propositions and theorems presented below can be extended to the the generalizations of convexity presented in Section 3.

2.1. **General setting.** We consider a metric space X, and $\mathcal{C}(X)$ the space of bounded continuous functions on X endowed with the norm of uniform convergence $\|.\|_\infty$. Every subset L of the space of affine forms on $\mathcal{C}(X)$ defines a convex subset of the space of continuous functions by duality:

$$\mathcal{H}_L := \{g \in \mathcal{C}(K);\ \forall \ell \in L,\ \ell(g) \leqslant 0\}. \tag{2.1}$$

Let M, L be two sets of affine forms on $\mathcal{C}(X)$. The set M is called an α-*relaxation* of L, where α is a function from $\mathcal{C}(X)$ to $\mathbb{R}\cup\{+\infty\}$, if the following holds: $\forall \ell \in L,\ \exists \ell_g \in M,\ |\ell(g) - \ell_g(g)| \leqslant \alpha(g)$.

2.2. **Convexity constraints.** In the remaining of this section, we suppose that X is a bounded open convex subset of $\mathbb{R}^d$, and we let L_k be the set of linear forms ℓ on $\mathcal{C}(K)$ that can be written as

$$\ell(g) = g\left(\sum_{i=1}^{k} \lambda_i x_i\right) - \left(\sum_{i=1}^{k} \lambda_i g(x_i)\right) \tag{2.2}$$

for a set of k points $x_1, \ldots, x_k$ in X, and where $(\lambda_i)_{1\leqslant i\leqslant k}$ belongs to the $(k-1)$-dimensional simplex Δ^{k-1}, i.e. $\lambda_1, \ldots, \lambda_k$ are non-negative numbers whose sum is equal to one. Since we are only considering continuous functions, the space $\mathcal{H}_{L_k}$ coincides with the space of convex functions when k is at least two. We denote this space by $\mathcal{H}$.

Theorem 2.1. *Consider an α-relaxation M of L_2. If g lies in $\mathcal{H}_M$, there must exist a convex function $\overline{g}$ in $\mathcal{H}$ such that $\|g - \overline{g}\|_\infty \leqslant d\alpha(g)$.*

Proof. Let us show first that, assuming that g is in $\mathcal{H}_M$, the following inequality holds for any form ℓ in L_k:

$$\ell(g) \leqslant k\alpha(g). \tag{2.3}$$

For $k = 2$, this follows at once from our hypothesis. Indeed, there must exist a linear form ℓ_g in M that satisfies (2.1), so that $\ell(g) \leqslant \ell_g(g) + \alpha(g)$. Since g lies in $\mathcal{H}_M$, $\ell_g(g)$ is non-positive and we obtain (2.3). The case $k > 2$ is proved by induction. Consider λ in the simplex Δ^{k-1} and points $x_1, \ldots, x_k$ in X. We assume $\lambda_1 < 1$ and we let $\mu_i = \lambda_i/(1-\lambda_1)$ for any $i \geqslant 2$. The vector $\mu = (\mu_2, \ldots, \mu_k)$ lies in Δ^{k-2}, and therefore $y = \sum_{i\geqslant 2} \mu_i x_i$ belongs to X. Applying the inductive hypothesis (2.3) twice, we obtain:

$$\begin{aligned}
&g\left(\lambda_1 x_1 + (1-\lambda_1) y\right) - \left(\lambda_1 g(x_1) + (1-\lambda_1) g\left(y\right)\right) \leqslant \alpha(g),\\
&g(y) - \left(\sum_{i=2}^{k} \mu_i g(x_i)\right) \leqslant (k-1)\alpha(g).
\end{aligned}$$

The sum of the first inequality and $(1-\lambda_1)$ times the second one gives (2.3). Now, consider the convex envelope of the function g, that is

$$\overline{g}(x) := \min\left\{\sum_{i=1}^{d+1} \lambda_i g(x_i);\ x_i \in X, \lambda \in \Delta^d \text{ and } \sum_i \lambda_i x_i = x\right\}. \tag{2.4}$$

This function is convex and its graph lies below the graph of g. Given any family of points (x_i) and coefficients (λ_i) such that $\sum \lambda_i x_i = x$, we consider the form $\ell(f) := f(x) - \sum_i \lambda_i f(x_i)$. Applying equation (2.3) to ℓ gives

$$g(x) - d\alpha(g) \leqslant \sum \lambda_i g(x_i)$$

Taking the minimum over the (x_i), (λ_i) such that $\sum \lambda_i x_i = x$, we obtain the desired inequality $|g(x) - \overline{g}(x)| \leqslant d\alpha(g)$. $\square$

2.3. **Discretization of convexity constraints.** We introduce here a family of affine forms M^c_ε which discretize the the convexity constraints L^c_2. We start with a notation: for any triple of points x,y and z such that $z \in [x, y]$, we define the linear form ℓ_{xyz} by the formula

$$\ell_{xyz}(g) := g(z) - \frac{\|zy\|}{\|xy\|} g(x) - \frac{\|xz\|}{\|xy\|} g(x). \tag{2.5}$$

By convention, when we write ℓ_{xyz}, we implicitly assume that z lies on the segment $[x, y]$. Consider an ε-sample of the boundary ∂X, that is a subset $U_\varepsilon \subseteq \partial X$ such that for every point x in ∂X there exists a point x_ε in U_ε with $\|x - x_\varepsilon\| \leqslant \varepsilon$. Given any pair of distinct points (p, q) in U_ε, we let c_{pq} be the discrete segment defined by

$$c_{pq} := \left\{ p + \varepsilon i \frac{(q-p)}{\|q-p\|}; i \in \mathbb{N},\, 0 \leqslant i \leqslant \|q-p\| / \varepsilon \right\}.$$

Finally, we define the following discretized set of constraints

$$M^c_\varepsilon := \{\ell_{xyz};\ x, y, z \in c_{pq} \text{ for some } p, q \in U_\varepsilon \text{ and } z \in [x, y]\}.$$

THEOREM 2.2. *For any function g in the space $\mathcal{H}_{M^c_\varepsilon}$, there exists a convex function $\overline{g}$ on X such that $\|g - \overline{g}\|_\infty \leqslant \mathrm{const}(d)\mathrm{Lip}(g)\varepsilon$.*

We start the proof by a technical lemma that gives an upper bound on the difference between two linear forms corresponding to convexity constraints:

LEMMA 2.3. *Let x, y, z and x', y', z' be six points in X. Assume the following*
(i) $\max(\|x - x'\|, \|y - y'\|, \|z - z'\|) \leqslant \eta$;
(ii) $z \in [x, y]$, $z' \in [x', y']$.
Then, $\left|\ell_{xyz}(g) - \ell_{x'y'z'}(g)\right| \leqslant 6\eta\mathrm{Lip}(g)$.

Proof. We define λ by the relation $z = \lambda x + (1 - \lambda)y$, and λ' is defined similarly. We also define $\ell_i(g) := g(z) - (\lambda' g(x) + (1 - \lambda')g(y))$. Then,

$$\left|\ell_{xyz}(g) - \ell_{x'y'z}(g)\right| \leqslant \left|\ell_{x'y'z}(g) - \ell_i(g)\right| + |\ell_i(g) - \ell_{xyz}(g)|$$

The first term is easily bounded by $2\eta\mathrm{Lip}(g)$, while the second term is bounded by $|\lambda - \lambda'|\,\mathrm{Lip}(g)\,\|xy\|$.

$$\begin{aligned} \left|\lambda - \lambda'\right| &= \left| \frac{\|zy\|}{\|xy\|} - \frac{\|z'y'\|}{\|x'y'\|} \right| \\ &\leqslant \left| \frac{\|zy\| - \|z'y'\|}{\|xy\|} \right| + \left| \frac{\|z'y'\|}{\|x'y'\|} \cdot \frac{\|x'y'\| - \|xy\|}{\|xy\|} \right| \leqslant 4\eta / \|xy\| \end{aligned}$$

Overall, we get the desired upper bound. □

Proof. Our goal is to show that M^c_ε is an α-relaxation of L^c_2. Consider three points x, y, z in X such that z is in the segment $[x, y]$. The straight line (x, y) intersects the boundary of X in two points a and b. By hypothesis, there exists two points p and q in U_ε such that the distances $\|a - p\|$ and $\|b - q\|$ are bounded by ε. The maximum distance between the segments $[a, b]$ and $[p, q]$ is then also bounded by ε and the maximum distance between the segment $[a, b]$ and the finite set c_{pq} by 2ε. This means that there exists three points $x_\varepsilon, y_\varepsilon$ and z_ε in c_{pq} such that $\max(\|x - x_\varepsilon\|, \|y - y_\varepsilon\|, \|y - y_\varepsilon\|) \leqslant 2\varepsilon$. Using Lemma 2.3, we deduce that $\|\ell_{xyz}(g) - \ell_{x_\varepsilon y_\varepsilon z_\varepsilon}(g)\| \leqslant \alpha(g) := 12\varepsilon\mathrm{Lip}(g)$. This implies that M_ε^c is an α-relaxation of L_2^c, and the statement follows from Theorem 2.1. □

2.4. **Finite-dimensional discretization.** Choné and Le Meur proved in [5] that the space of piecewise-linear convex functions on a regular grid of the square is *not dense* in the space of convex functions on this domain. This means that we need to be very careful in order to apply the convexity constraints M_c^ε to finite-dimensional space of functions.

For instance, consider the space E_δ of piecewise-linear functions on a triangulation of the domain with edgelength bounded by δ. It is not difficult to realize that if the ratio ε/δ is small enough, any piecewise linear function on this triangulation that satisfies the relaxed convexity constraints M_c^ε is automatically convex. For such a choice of δ, one can fall in the pitfall discovered by Choné and Le Meur, as illustrated in Figure 2.4. We consider the convex function $f(x,y) = \max(0, x+y-1)$ on the unit square $X = [0,1]^2$ and its projection g on the intersection $\mathcal{H}_{M_\varepsilon^c} \cap E_\delta$, where E_δ is the space of piecewise linear functions on a regular grid with edge length $\delta \gg \varepsilon$. The error $|f-g|$ is displayed for three different choices of grid size. One can observe that the maximum error $\|f-g\|_\infty$ remains almost constant regardless of δ.

As a consequence, we discuss below how to choose ε as a function of δ so as to obtain the density of the union of spaces $(\mathcal{H}_{M_\varepsilon^c} \cap E_\delta)_{\delta>0}$ in $\mathcal{H}$.

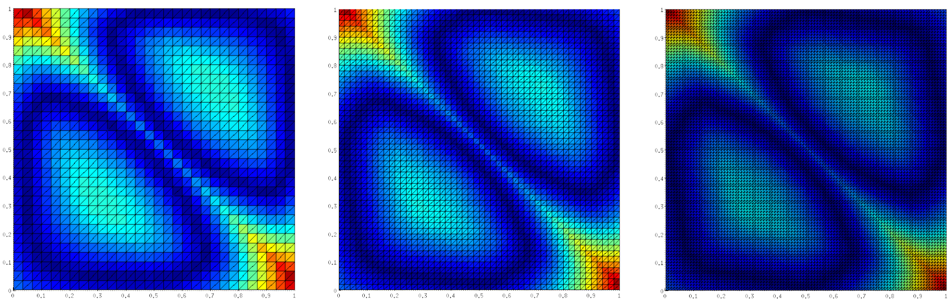

FIGURE 2. Illustration of the non-convergence result phenomena of Choné and Le Meur on a regular grid: stability of the inf norm error on multiple grids. In all three figures, the dark red corresponds to the value 0.2.

Definition 2.1. We call *linear interpolation operator* a linear map $\mathcal{I}_\delta$ from the space $\mathcal{C}(X)$ of continuous functions to a finite-dimensional subspace E_δ of $\mathcal{C}(X)$. Moreover, we will assume that the interpolation operator $\mathcal{I}_\delta$ enjoys the following properties:

$$\mathrm{Lip}(\mathcal{I}_\delta f) \leqslant \mathrm{Lip} f \tag{L1}$$

$$\|f - \mathcal{I}_\delta f\|_\infty \leqslant \delta \mathrm{Lip}(f), \tag{L2}$$

$$\|f - \mathcal{I}_\delta f\|_\infty \leqslant \frac{1}{2}\delta^2 \mathrm{Lip}(\nabla f), \tag{L3}$$

An obvious example of linear interpolation operator is given by linear interpolation on a mesh. Consider a triangulation of a polyhedral domain X, and assume that the diameter of all triangles in is bounded by δ. Then, the operator $\mathcal{I}_\delta$ defined by linear interpolation of the values of the function on the triangles satisfies (L1)–(L3). One could also consider other interpolation spaces, such as higher-order finite elements, or even spaces of functions constructed from tensor-product splines.

2.5. **Hausdorff approximation results.** The results of the previous paragraphs can be combined in order to show that the finite dimensional polyhedron $\mathcal{H}_{M_\varepsilon^c} \cap I_\delta$ converges to the whole set of convex functions for a suitable choice of ε. Given two subsets A, B of a functional space $\mathcal{C}(X)$, we define the half-Hausdorff distance $\mathrm{h_H}(A, B)$ as follows:

$$\mathrm{h_H}(A,B) = \min\left\{r \geqslant 0; \forall f \in A,\ \exists g \in B,\ \|f-g\|_\infty \leqslant r\right\}.$$
$$\mathrm{d_H}(A,B) = \max\left(\mathrm{h_H}(A,B), \mathrm{h_H}(B,A)\right)$$

In the following theorem, the set of γ-Lipschitz functions on X is denoted $\mathrm{B}^\gamma_{\mathrm{Lip}}$, while the set of functions with γ-Lipschitz gradient is denoted $\mathrm{B}^\gamma_{\mathcal{C}^{1,1}}$.

THEOREM 2.4. *Suppose given a bounded convex set X, and an interpolation operator $\mathcal{I}_\delta : \mathcal{C}(X) \to E_\delta$. Then,*

(1) $\mathrm{h_H}(\mathrm{B}^\gamma_{\mathrm{Lip}} \cap E_\delta \cap \mathcal{H}_{M_\varepsilon^c}, \mathrm{B}^\gamma_{\mathrm{Lip}} \cap \mathcal{H}) \leqslant \mathrm{const}(d)\gamma\varepsilon.$

(2) *Assuming $\varepsilon = \mathrm{const}(d)\,\mathrm{diam}(X)^{2/3}\delta^{1/3}$, one has*

$$\mathrm{d_H}(\mathrm{B}^\gamma_{\mathrm{Lip}} \cap E_\delta \cap \mathcal{H}_{M_\varepsilon^c}, \mathrm{B}^\gamma_{\mathrm{Lip}} \cap \mathcal{H}) \leqslant \mathrm{const}(d)\gamma\,\mathrm{diam}(X)^{2/3}\delta^{1/3}, \tag{2.6}$$

(3) *Assuming $\varepsilon = \mathrm{const}(d)\,\mathrm{diam}(X)^{1/3}\delta^{2/3}$, one has*

$$\mathrm{h_H}(\mathrm{B}^\gamma_{\mathcal{C}^{1,1}} \cap \mathcal{H}, E_\delta \cap \mathcal{H}_{M_\varepsilon^c}) \leqslant \mathrm{const}(d)\gamma\,\mathrm{diam}(X)^{1/3}\delta^{2/3}. \tag{2.7}$$

The following easy lemma shows that the space $\mathcal{H}_{M_\varepsilon^c} \cap E_\delta$ has non-empty interior as soon as $\varepsilon < \delta$. While very simple, this fact is the key to the proof of the theorem.

LEMMA 2.5. *Consider the function $s(x) := \|x - x_0\|^2$ on X, where x_0 is a point in X, and the interpolating function $s_\delta := \mathcal{I}_\delta s$. Then,*

$$\max_{\ell \in M_\varepsilon^c} \ell(s_\delta) \leqslant \delta^2 - \varepsilon^2.$$

Proof. Consider three points $x < z < y$ on the real line, such that $|x - z| \geqslant \varepsilon$ and $|y - z| \geqslant \varepsilon$ and $z = \lambda x + (1-\lambda)z$. Then,

$$\begin{aligned} z^2 - \lambda x^2 - (1-\lambda)y^2 &= z^2 - \lambda(z + (x-z))^2 - (1-\lambda)(z + (y-z))^2 \\ &= -[\lambda(x-z)^2 + (1-\lambda)(y-z)^2] \leqslant -\varepsilon^2 \end{aligned}$$

Since the gradient of s is 1-Lipschitz, using (L2) we get $\|s - s_\delta\|_\infty \leqslant \delta^2$. Combining with the previous inequality, we get $\ell s_\delta \leqslant \delta^2 - \varepsilon^2$ for every linear form ℓ in M_ε^c. □

Proof of Theorem 2.4. Theorem 2.1 implies that for every function g in the intersection $\mathcal{H}_{M_\varepsilon^c} \cap \mathrm{B}^\gamma_{\mathrm{Lip}}$, there exists a convex function $\overline{g}$ in $\mathcal{H}$ such that $\|g - \overline{g}\|_\infty \leqslant \mathrm{const}(d)\mathrm{Lip}(g)\varepsilon$. The Lipschitz constant of a function is not increased by taking its convex envelope, and thus $\overline{g}$ belongs to $\mathcal{H} \cap \mathrm{B}^\gamma_{\mathrm{Lip}}$. This gives us the upper bound

$$\mathrm{h_H}(\mathrm{B}^\gamma_{\mathrm{Lip}} \cap E_\delta \cap \mathcal{H}_{M_\varepsilon^c}, \mathrm{B}^\gamma_{\mathrm{Lip}} \cap \mathcal{H}) \leqslant \mathrm{const}(d)\gamma\varepsilon. \tag{2.8}$$

On the other hand, given a convex function f in $\mathrm{B}^\gamma_{\mathrm{Lip}}$, we consider the function $g := \mathcal{I}_\delta f$ defined by the interpolation operator. Using the property (L2), we can show that for any linear form ℓ in L_2,

$$\begin{aligned}\ell(g) &= g(\lambda x + (1-\lambda)y) - (\lambda g(x) + (1-\lambda)g(y)) \\ &\leqslant f(\lambda x + (1-\lambda)y) - (\lambda f(x) + (1-\lambda)f(y)) + 2\delta\gamma \leqslant 2\delta\gamma\end{aligned}$$

Assuming $\delta \leqslant \varepsilon/2$, this implies that for any linear form ℓ in M^c_ε,

$$\ell(g + \eta s_\delta) \leqslant 2\delta\gamma + \eta(\delta^2 - \varepsilon^2) \leqslant 2\delta\gamma - \eta\varepsilon^2/4$$

Consequently, the inequality $\ell(g) \leqslant 0$ holds for any linear form ℓ in M^c_ε, assuming $8\delta\gamma \leqslant \eta\varepsilon^2$. This implies, using property (L1), that the function $h := g + \eta s_\delta$ belongs to the intersection $\mathrm{B}^\gamma_{\mathrm{Lip}} \cap \mathcal{H}_{M^c_\varepsilon}$. We therefore fix $\eta = 8\delta\gamma/\varepsilon^2$. The distance between f and h is bounded by

$$\begin{aligned}\|f - h\|_\infty &\leqslant \|f - g\|_\infty + \eta\,\|s_\delta\|_\infty \\ &\leqslant \gamma\delta + \eta\,\mathrm{diam}(X)^2 = \gamma\delta(1 + 8\,\mathrm{diam}(X)/\varepsilon^2) \\ &\leqslant 9\gamma\delta\,\mathrm{diam}(X)/\varepsilon^2,\end{aligned}$$

thus implying the following upper bound on the half-Hausdorff distance:

$$\mathrm{h_H}(\mathrm{B}^\gamma_{\mathrm{Lip}} \cap \mathcal{H}, \mathrm{B}^\gamma_{\mathrm{Lip}} \cap E_\delta \cap \mathcal{H}_{M^c_\varepsilon}) \leqslant \gamma\delta. \tag{2.9}$$

Combining Equations (2.8)–(2.9), we get

$$\mathrm{d_H}(\mathrm{B}^\gamma_{\mathrm{Lip}} \cap E_\delta \cap \mathcal{H}_M, \mathrm{B}^\gamma_{\mathrm{Lip}} \cap \mathcal{H}) \leqslant \mathrm{const}(d)\gamma \max(\varepsilon, \delta\,\mathrm{diam}(X)^2/\varepsilon^2)$$

We choose ε so as to equate the two terms in the maximum, i.e. $\varepsilon = \delta^{1/3}\,\mathrm{diam}(X)^{2/3}$. This proves assertion (2). Assertion (3) is proved in the same way, using (L3) instead of (L2). □

3. Generalization to convexity-like constraints

3.1. **Support functions.** Recall that a compact convex set K in $\mathbb{R}^d$ is uniquely determined by its *support function*, defined by

$$\underline{\mathrm{h}}_K : x \in \mathbb{R}^d \mapsto \max_{p \in K} \langle x | p \rangle.$$

This function is is positively 1-homogeneous and is therefore completely determined by its restriction h_K on the unit sphere. We consider the space $\mathcal{H}^s \subseteq \mathcal{C}(\mathcal{S}^{d-1})$ of support functions of compact convex sets. This space coincides with the space of bounded functions on the sphere whose homogeneous extensions to the whole space $\mathbb{R}^d$ are convex.

Lemma 3.1. *A bounded function g on the unit sphere is the support function of a bounded convex set if and only if for every $x_1, \dots, x_k$ in the sphere, and every $(\lambda_1, \dots, \lambda_k) \in \Delta^{k-1}$,*

$$\|x\|\, g\left(\frac{x}{\|x\|}\right) \leqslant \sum_i \lambda_i g(x_i), \quad \textit{where } x := \sum_i \lambda_i x_i. \tag{3.10}$$

Moreover, g is the support function of a convex set if it satisfies the inequalities for $k = 2$ only.

Following this lemma, we define L^s_k as the space of all linear forms that can be written as

$$\ell(g) := \left\|\sum_i \lambda_i x_i\right\| g\left(\frac{\sum_i \lambda_i x_i}{\|\sum_i \lambda_i x_i\|}\right) - \sum_i \lambda_i g(x_i), \tag{3.11}$$

where $x_1, \dots, x_k$ are points on the sphere $\mathcal{S}^{d-1}$, and $(\lambda_1, \dots, \lambda_k)$ lies in Δ^{k-1}. With this notation at hand, we have another characterization of the space of support functions: $\mathcal{H}^s$ coincides with the spaces $\mathcal{H}_{L^s_k}$ for any $k \geqslant 2$.

Discretization of the constraints. The discretization of the set L^s_2 of constraints satisfied by support functions follows closely the discretization of the convexity constraints described earlier. Consider three points x, y and z such that x and y are not antipodal and such that z belongs to the minimizing geodesic between x and y. We let z' be the radial projection of z on the extrinsic segment $[xy]$, i.e. such that $z'/\|z'\| = z$. Finally, we let $\lambda = \|zy\| / \|xy\|$ and define:

$$\ell_{xyz}(g) := \left\|z'\right\| g(z) - \lambda g(x) - (1-\lambda) g(y).$$

As before, we discard the constraint ℓ_{xyz} if z does not lie on the minimizing geodesic arc between x and y. Let U_ε be a subset of the sphere that satisfies the sampling condition

$$\forall u \in \mathcal{S}^{d-1},\ \exists(\sigma, v) \in \{\pm 1\} \times U_\varepsilon,\ \text{s.t. } \|u - \sigma v\| \leqslant \varepsilon. \tag{3.12}$$

Then, for every vector u in U_ε we construct an ε-sampling c_u of the great circle orthogonal to u that is also $\frac{\varepsilon}{2}$-sparse, i.e. $\|x - y\| \geqslant \frac{\varepsilon}{2}$ for any pair of distinct points x, y in c_u. The space of constraints we consider is the following:

$$M^s_\varepsilon = \{\ell_{xyz};\ x, y, z \in c_u \text{ for some } u \in U_\varepsilon\}.$$

The proof of the following theorem follows the proof of Theorem 3.2, and even turns out to be slightly simpler as one does not need to take care of the boundary of the domain.

Theorem 3.2. *For any function h in the space $\mathcal{H}_{M^s_\varepsilon}$, there exists a bounded convex set K such that $\|h - h_K\|_\infty \leqslant \mathrm{const}(d)\mathrm{Lip}(g)\varepsilon$.*

It is possible to define a notion of interpolation operator on the sphere as in §2.4, and to obtain Hausdorff approximation results similar to those presented in Theorem 2.4. The statement and proofs of the theorem being very similar, we do not repeat them. However, we show that the indicator function of the unit ball, i.e. the constant function equal to one, belongs to the interior of the set $\mathcal{H}_{M^c_\varepsilon}$. This is the analogous of Lemma 2.5, which was the crucial point of the proof of convergence for the usual convexity.

Lemma 3.3. *With $s(x) := 1$, one has $\max_{\ell \in M^c_\varepsilon} \ell(s) \leqslant -\mathrm{const} \cdot \varepsilon^2$.*

Proof. For every ℓ in M^c_ε, there exists three (distinct) points x, y, z in c_u for some u in U_ε. Let z' denote the radial projection of z on the segment $[x, y]$. Then, $\ell_{xyz}(s) = \|z'\| - 1$. By construction, $\|x - z\|$ and $\|y - z\|$ are at least $\varepsilon/2$, and therefore $\|z'\| \leqslant 1 - \mathrm{const} \cdot \varepsilon^2$ thus proving the lemma. □

Support function as c-convex functions. Oliker[17] and Bertrand[3] introduced another characterization of support functions of convex sets, inspired by optimal transportation theory. They show that logarithm of support functions coincide with c-convex functions on the sphere for the cost function $c(x, y) = -\log(\max(\langle x|y\rangle, 0))$ (see §3.2 for a definition of c-convexity):

LEMMA 3.4. *A bounded positive function $h : \mathcal{S}^{d-1} \to \mathbb{R}$ is convex and positively 1-homogeneous if and only if the function $\varphi := \log(h)$ can be written as*

$$\varphi(x) = \sup_{y\in\mathcal{S}^{d-1}} -\psi(y) - c(x, y)$$

where $c(x, y) = -\log(\max(\langle x|y\rangle, 0))$ and $\psi : \mathcal{S}^{d-1} \to \mathbb{R}$.

Proof. We show only the direct implication, the reverse implication can be found in [3]. By assumption $h = h_K$, where K is a bounded convex set that contains the origin in its interior, and let ρ_K be the radial function of K i.e. $\rho_K(y) := \max\{r;\, ry \in K\}$. Then,

$$h_K(x) = \max_{p\in K}\langle x|p\rangle = \max_{y\in\mathcal{S}^{d-1}} \rho_K(y)\langle x|y\rangle$$

Since $h_K > 0$, the maximum in the right-hand side is attained for a point y such that $\langle x|y\rangle > 0$. Taking the logarithm of this expression, we get:

$$\varphi(x) = \max_{y\in\mathcal{S}^{d-1}} \log(\rho_K(y)) - c(x, y) \qquad \square$$

which concludes the proof of the direct implication.

3.2. **c-Convex functions.** In this paragraph, we show how the discretizations of the spaces of convex and support functions presented above can be extended to c-convex functions. This extension is motivated by a family of generalizations of the principal-agent problem proposed by Figalli, Kim and McCann [7]. In this article, the authors show that the set of c-convex functions is convex if and only if c satisfy the so-called non-negative cross-curvature condition. Under the same assumption, we identify the linear inequalities that define this convex set of functions. The numerical implementation of this section will be presented in future work.

Given a cost function $c : X \times Y \to \mathbb{R}$, where X and Y are two open and bounded subsets of $\mathbb{R}^d$, the c-transform and c^*-conjugate of lower semi-continuous functions $\varphi : X \to \mathbb{R}$ and $\psi : Y \to \mathbb{R}$ are defined by

$$\varphi^{c^*}(y) := \sup_{x\in Y} -c(x, y) - \varphi(x),$$
$$\psi^{c*}(x) := \sup_{y\in Y} -c(x, y) - \psi(y).$$

A function is called c-convex if it is the c^*-transform of a lower semi-continuous function $\psi : Y \to \mathbb{R}$. The space of c-convex functions on X is denoted $\mathcal{H}_c$. We will need the following usual assumptions on the cost function c:

(A0) $c \in \mathcal{C}^4(\overline{X} \times \overline{Y})$, where $X, Y \subseteq \mathbb{R}^n$ are bounded open domains.

(A1) For every point y_0 in Y and x_0 in X, the maps

$$x \in \overline{X} \mapsto -\nabla_y c(x, y_0)$$
$$y \in \overline{Y} \mapsto -\nabla_x c(x_0, y)$$

are diffeomorphisms onto their range (bi-twist).

(A2) For every point y_0 in Y and x_0 in X, the sets $X_{y_0} := -\nabla_y c(X, y_0)$ and $Y_{x_0} := -\nabla_x c(x_0, Y)$ are convex (bi-convexity).

These conditions allow one to define the c-exponential map. Given a point y_0 in the space Y, the c-exponential map $\exp^c_{y_0} : X_{y_0} \to X$ is defined as the inverse of the map $-\nabla_y c(., y_0)$, i.e. it is the unique solution of

$$\exp^c_{y_0}(-\nabla_y c(x, y_0)) = x. \tag{3.13}$$

The following formulation of the non-negative cross-curvature condition is slightly non-standard, but it agrees to the usual formulation for smooth costs under conditions (A0)–(A2), thanks to Lemma 4.3 in [7].

(A3) For every pair of points (y_0, y) in Y the following map is convex:

$$v \in X_{y_0} \mapsto c(\exp^c_{y_0} v, y_0) - c(\exp^c_{y_0} v, y). \tag{3.14}$$

The main theorem of [7] gives a necessary and sufficient condition for the space $\mathcal{H}_c$ of c-convex functions be convex.

THEOREM 3.5. *Assuming (A0)–(A2), the space of c-convex functions $\mathcal{H}_c$ is itself convex if and only if c satisfies (A3).*

The proof that (A0)–(A3) implies the convexity of $\mathcal{H}_c$ given in [7] is direct but non-constructive, as the authors show that the average of two functions φ_0 and φ_1 in $\mathcal{H}_c$ also belongs to $\mathcal{H}_c$. The following proposition provides a set of linear inequality constraints that are both necessary and sufficient for a function to be c-convex.

PROPOSITION 3.6. *Assuming the cost function satisfies (A0)–(A3), a function $\varphi : X \to \mathbb{R}$ is c-convex if and only if it satisfies the following constraints:*

(i) *for every y in Y, the map $\varphi_y : v \in X_y \mapsto \varphi(\exp^c_y v) + c(\exp^c_y v, y)$ is convex.*

(ii) *for every x in X, the subdifferential $\partial\varphi(x)$ is included in Y_x.*

Note that while the first set of constraints (i) can be discretized in an analogous way to the previous sections. On the other hand, the second constraint concerns the subdifferential of φ in the sense of semiconvex functions. It is not obvious how to handle this constraint numerically, except in the trivial case where Y is the whole space $\mathbb{R}^d$.

Proof. Suppose first that φ is c-convex. Then, there exists a function ψ such that $\varphi(x) = \psi^{c*}$ and one has

$$\varphi_y(v) = \sup_z[-\psi(z) - c(\exp^c_y v, z)] + c(\exp^c_y v, y).$$

Equation 3.14 implies that φ_y is convex as a maximum of convex functions.

Conversely, suppose that a map $\varphi : X \to \mathbb{R}$ is such that the maps φ_y are convex for any point y in Y, and let us show that φ is c-convex. Using the definition of φ_y, and the definition of the c-exponential (3.13), one has

$$\varphi(x) = \varphi_y(-\nabla_y c(x, y)) - c(x, y)$$

for any pair of point (x, y) in $X \times Y$. This formula and the convexity of φ_y imply in particular that the map φ is semiconvex. Consequently, for every point x in X, the subdifferential $\partial\varphi(x)$ is non-empty, and there must exist a point y in Y such that $v := -\nabla_x c(x, y)$ belongs to $\partial\varphi(x)$. Hence, x is a critical point of the map $\varphi - c(., y)$, and therefore v is a critical point of φ_y. By convexity, v is also a global minimum of φ_y, i.e. for every w in $\overline{X}_y$,

$$\varphi(\exp_y^c w) + c(\exp_y^c w, y) \geqslant \varphi(x) + c(x, y).$$

Letting $x' = \exp_y^c w$, we get $\varphi(x') \geqslant \varphi(x) + c(x, y) - c(x', y)$. The function $\varphi(x) + c(x, y) - c(., y)$ is thus supporting φ at x. Since φ admits such a supporting function at every point x in X, it is a c-convex function. □

4. Numerical implementation

In this section, we give some details on how to apply the relaxed convexity constraints presented in Section 2 to the numerical resolution of problems of calculus of variation with (usual) convexity constraints. Our goal is to minimize a functional $\mathcal{F}$ over the set of convex functions $\mathcal{H}$. For any convex set K, we will denote i_K the convex indicator function of K, that is the function that vanishes on K and take value $+\infty$ outside of K. The problem above can then be reformulated as

$$\min_{g \in \mathcal{C}(X)} \mathcal{F}(g) + \mathrm{i}_{\mathcal{H}}(g), \tag{4.15}$$

The method that we present in this paragraph can be applied with minor modifications to the other types of convexity-like constraints presented in Section 3.

4.1. **Finite-dimensional setting.** We are given a finite-dimensional subspace E of $\mathcal{C}(X)$, such as the space of piecewise-linear functions on a triangulation of X, and a parameterization $\mathcal{P} : \mathbb{R}^N \to E$ of this space. For every point x in X, we consider the (linear) evaluation map $\mathcal{P}_x : \mathbb{R}^N \to \mathbb{R}$, defined by $\mathcal{P}_x\xi := (\mathcal{P}\xi)(x)$, with the convention that $\mathcal{P}_x\xi = +\infty$ outside of X. We suppose that we are given a certain $\varepsilon > 0$, set $M := M_\varepsilon^c(X)$ and consider the following relaxation of (4.15):

$$\min_{\xi \in \mathbb{R}^N} \mathcal{F}(\mathcal{P}\xi) + \mathrm{i}_{\mathcal{H}_M}(\mathcal{P}\xi). \tag{4.16}$$

In the following, we use the notation introduced in §2.3. Consider a discrete segment c_{pq}, where p and q are distinct points in the ε-dense subset U_ε of the boundary of X. The evaluation of a function $\mathcal{P}\xi$ on such a discrete segment c_{pq} is a vector

$$\mathcal{P}_{pq}\xi = \left(\mathcal{P}\xi(p + \varepsilon i \frac{(q-p)}{\|q-p\|})\right)_{i \in \mathbb{N}}$$

Note that this vector takes takes finite values for indices in $\{0, \dots, |c_{pq}| - 1\}$. This allows us to rewrite the indicator function of the discretized convexity constraints $\mathrm{i}_{\mathcal{H}_M}$ as a sum of indicator functions, one for each of the discrete segment c_{pq}. Define $\mathcal{H}_1$ as the cone of vectors (f_i) indexed by $\mathbb{N}$ that satisfy the discrete convexity conditions $f_i \leqslant \frac{1}{2}(f_{i-1} + f_{i+1})$ for $i \geqslant 1$. The relaxed problem (4.16) is then equivalent to the following minimization problem:

$$\min_{\xi \in \mathbb{R}^N} \mathcal{F}(\mathcal{P}\xi) + \sum_{(p,q) \in U_\varepsilon^2} \mathrm{i}_{\mathcal{H}_1}(\mathcal{P}_{pq}\xi). \tag{4.17}$$

Algorithm 1 Simultaneous-direction method of multipliers (SDMM)

Input: $\gamma > 0$
Initialization: $(y_{1,0},\, z_{1,0}) \in \mathbb{R}^{2N_1}, \dots, (y_{m,0},\, z_{m,0}) \in \mathbb{R}^{2N_m}$
For: $n = 0, 1, \dots$
$\quad x_n = Q^{-1} \sum_{i=1}^m L_i^T (y_{i,n} - z_{i,n})$
$\quad$ **For:** $i = 1, \dots, m$
$\quad\quad s_{i,n} = L_i x_n$
$\quad\quad y_{i,n+1} = \operatorname{prox}_\gamma g_i(s_{i,n} + z_{i,n})$
$\quad\quad z_{i,n+1} = z_{i,n} + s_{i,n} - y_{i,n+1}$

4.2. **Proximal methods.** When the optimized functional $\mathcal{F}$ is convex, Problem (4.17) is a standard convex optimization problem under a large set of inequality constraints. The projection on the set of 1D discrete convex functions $\mathcal{H}_1$ is easy to compute due to its specific structure, as shown in the next paragraph. This particular structure allows us to use a proximal algorithm when the function $\mathcal{F}$ is smooth or when there exists a fast algorithm to compute its proximal operator. Recall that the proximal operator associated to a convex function f is defined by:

$$\operatorname{prox}_\gamma f(y) = \arg\min_{x \in \mathbb{R}^N} f(x) + \frac{1}{\gamma} \|x - y\|^2 . \tag{4.18}$$

For instance, when f is the indicator function of a convex set, $\operatorname{prox}_\gamma f$ coincide with the projection operator on this set regardless of the value of γ. The simultaneous-direction method of multipliers (SDMM) algorithm is designed to solve convex optimization problems of the following type :

$$\min_{x \in \mathbb{R}^N} g_1(L_1 x) + \dots g_m(L_m x)$$

where the $(L_i)_{1 \leqslant i \leqslant m}$ are matrices of dimensions $N_1 \times N, \dots, N_m \times N$ and the function $(g_i)_{1 \leqslant i \leqslant m}$ are convex and easily proximable. Moreover, it assumes that the matrix $Q := \sum_{i=1}^m L_i^T L_i$ is invertible, where L_i^T stands for the transpose of the matrix L_i. A summary of the SDMM algorithm is given in Algorithm 1. More details, and variants of this algorithm can be found in the book [2]. Note that when applied to (4.17), every iteration of the outer loop of the SDMM algorithm involves the computation of several projection on the cone of 1D discrete functions $\mathcal{H}_1$. These projection can be computed independently, thus allowing an easy parallelization of the optimization.

4.3. **Hinge algorithm.** In the algorithm above, we need to compute the l^2 projection of a vector (f_i) on the cone of discrete 1D convex functions $\mathcal{H}_1$. In practice, (f_i) is supported on a finite set $\{0, \dots, n\}$, and one needs to compute the ℓ^2 projection of this vector onto the convex cone

$$\mathcal{H}_1^n = \{g : \{0, \dots, n\} \to \mathbb{R};\ \forall i \in \{1, \dots, n-1\}, 2g_i \leqslant g_{i-1} + g_{i+1}\}.$$

This problem is classical, and several efficient algorithm have been proposed to solve it. Since the number of conic constraints is lower than the dimension of the ambient space $(n+1)$, the number of extreme rays of the cone $\mathcal{H}_1^n$ is bounded by n. Moreover, as noted by Meyer [14], these extreme rays can be computed easily. This remark allows one to parameterize the cone $\mathcal{H}_1^n$ by the space $\mathbb{R} \times \mathbb{R}_+^n$, thus recasting the projection onto $\mathcal{H}_1^n$ into a much simpler

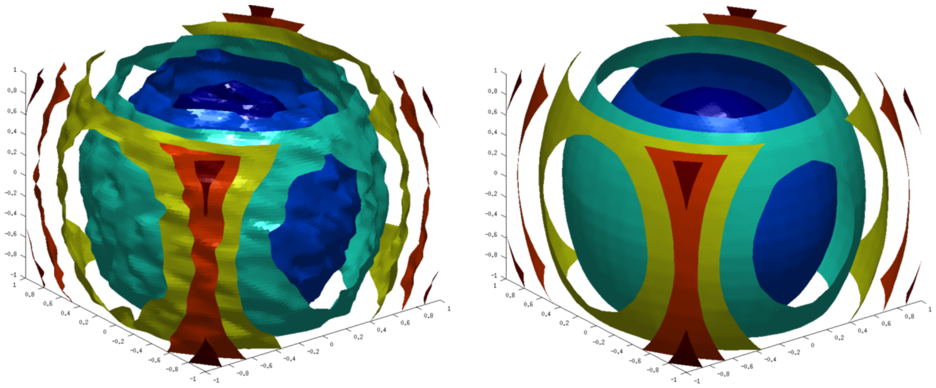

FIGURE 3. Denoising a convex graph by one dimensional projections.

non-negative least squares problem. To solve this problem, we use a simple and efficient exact active set algorithm proposed by Meyer [14].

5. Application I : Denoising

Our first numerical application focuses on the L^2 projection onto the set of convex functions on a convex domain. We illustrate the efficiency of our relaxed approach in the context of denoising. Let u^* be a convex function on a domain X in $\mathbb{R}^d$. We approximate this function by a piecewise linear function on a mesh, and the values of the function at the node of the mesh are additively perturbed by Gaussian noise: $u_0(p) = u^*(p) + c\mathcal{N}(0,1)$, where $\mathcal{N}(0,1)$ stands for the standard normal distribution and c is a small constant. Our goal is then to solve the following projection problem in order to estimate the original function u^*:

$$\min_{u \in \mathcal{H}} \|u - u_0\|_{L^2(X)} .$$

As described in previous sections, our discretization of the space of convex functions is not interior. However, thanks to Theorem 2.4, we obtain a converging discretization process that uses fewer constraints than previously proposed interior approaches. More explicitly, we illustrate below our method on the following three-dimensional denoising setting. Let $u_0(x,y,z) = \frac{x^2}{3} + \frac{y^2}{4} + \frac{z^2}{8}$, $X = [-1,1]^3$ with $c = \frac{1}{40}$. We carried our computation on a regular grid made of 80^3 points and we look for an approximation in the space of piecewise-linear functions. The parameter used to discretize the convexity constraints is set to $\varepsilon = 0.02$. Figure 3 displays the result of the SDMM algorithm after 10^4 iterations. This computation took less than five minutes on a standard computer.

To illustrate the versatility of the method, we performed the same denoising experience in the context of support functions, using the discretization explained in Section 3. As in the previous example, we consider a support function perturbed by additive Gaussian noise $\mathrm{h}_0(p) = \mathrm{h}^*(p) + c\mathcal{N}(0,1)$. In the numerical application, h^* is the support function of the unit isocaedron and $c = 0.05$, as shown on the left of Figure 4. Our goal is to compute the

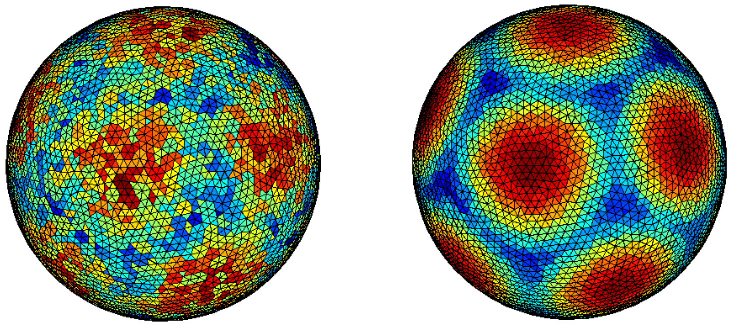

FIGURE 4. Denoising the support function of a convex body. On the left the perturbed support function of the icosaedron. On the right its projection into the set of support functions.

projection of h_0 to the space of support functions:

$$\min_{\mathrm{h}\in\mathcal{H}^s} \|\mathrm{h}-\mathrm{h}_0\|^2_{L^2(\mathcal{S}^{d-1})}.$$

In order to relax the constraint $\mathcal{H}^s$, we imposed one dimensional constraints on a family of 2000 great circles of $\mathcal{S}^{d-1}$ uniformly distributed and a step discretization of every circular arc equal to 0.02. We obtained a very satisfactory reconstruction of h_i after 10^4 iterations of the SDMM algorithm, as displayed on the right of Figure 4.

6. Application II : Principal-agent problem

6.1. **Geographic principal agent problem.** Let X be a bounded convex domain of $\mathbb{R}^d$, a distribution of agent $\rho : X \to \mathbb{R}$ and a finite subset $K \subseteq X$. The *monopolist* or *principal* needs to determine a *price menu* π for pick-up or deliveries, so as to maximize its revenue. The principal has to take into account the two following constraints: (i) the agents will try to maximize their utility and (ii) there is a finite subset $K \subseteq X$ of facilities, that compete with the principal and force him to set its price $\pi(y)$ to zero at any y in K. For a given price menu π, the utility of a location y for an agent located at a position x in X is given by $u_\pi(x,y) = -\frac{1}{2}\|x-y\|^2 - \pi(y)$. The fact that each agent tries to maximize his utility means that he will choose a location that balances closeness and price. The maximum utility for an agent x is given by:

$$u_\pi(x) := \max_{y\in X} u(x,y) = -\frac{1}{2}\|x\|^2 + \max_{y\in X}\left[\langle x|y\rangle - \frac{1}{2}\|y\|^2 - \pi(y)\right]$$

Let us denote $\overline{u}_\pi(x)$ the convex function $u_\pi(x) + \frac{1}{2}\|x\|^2$. This function is differentiable almost every point x in X, and at such a point the gradient $\nabla\overline{u}_\pi(x)$ agrees with the best location for x, i.e. $\nabla\overline{u}_\pi(x) = \arg\max_y u(x,y)$. This implies the following equality:

$$\overline{u}_\pi(x) = \langle x|\nabla\overline{u}_\pi(x)\rangle - \frac{1}{2}\|\nabla\overline{u}_\pi(x)\|^2 - \pi(\nabla\overline{u}_\pi(x))$$

Our final assumption is that the cost of a location for the principal is constant. Our previous discussion implies that the total revenue of the principal, given a price menu π, is computed by the following formula

$$\begin{aligned} R(\pi) &= \int_X \pi(\nabla \overline{u}_\pi(x))\rho(x)\mathrm{d}x \\ &= -\int_X \left[\overline{u}_\pi(x) - \langle x | \nabla \overline{u}_\pi(x) \rangle + \frac{1}{2} \left\| \nabla \overline{u}_\pi(x) \right\|^2 \right] \rho(x)\mathrm{d}x \end{aligned} \tag{6.19}$$

Changing the unknown from π to $v := \overline{u}_\pi$, the assumption that the price vanishes on the set K translates as $u_\pi \geqslant \max_{y\in K} -\frac{1}{2} \left\| \cdot - y \right\|^2$ or equivalently

$$v(x) = \overline{u}_\pi(x) \geqslant \max_{y\in K} \langle x|y\rangle - \frac{1}{2} \left\| y \right\|^2.$$

Thus, we reformulate of the principal's problem in term of v as the maximization of the following functional:

$$L(v) := \int_X \left[v(x) + \frac{1}{2} \left\| \nabla v(x) - x \right\|^2 \right] \rho(x)\mathrm{d}x$$

where the maximum is taken over the set of convex functions $v : X \to \mathbb{R}$ that satisfy the lower bound $v \geqslant \max_{y\in K} \langle . | y\rangle - \frac{1}{2} \left\| y \right\|^2$.

6.2. **Results.** In order to evaluate the accuracy of our algorithm, we first solve the principal-agent problem on the unit disk, with $K = \{(0,0)\}$ and ρ constant. It is known that the optimal profile is radial in this setting, and it is possible to obtain a very accurate description of the optimal radial component by solving a standard convex quadratic programming problem. In parallel, we computed an approximation of the solution without the radial assumption. To perform this computation we looked for an optimal function in the space of P^3 finite elements defined on a regular mesh of the square of size 60×60. The ε parameter for the discretization of convexity is 0.02. As shown on the left picture of Figure 5, our solution perfectly matches the line of the one dimensional profile (displayed on the graph) after 10^3 iteration of the SDMM algorithm. In the standard setting proposed by Rochet and Choné i.e. $X = [1,2]^2$, we recover the bunching phenomena. This can be seen on the right of Figure 5. This result agrees with previous numerical simulations that can be found in the literature [6, 15, 1].

7. Application III : Closest convex set with constant width

A convex compact set K of $\mathbb{R}^d$ has constant width $\alpha > 0$ if all its projection on every straight line are segments of length α. This property is equivalent to the following constraints on the support function of K :

$$\forall \nu \in \mathcal{S}^{d-1},\ \mathrm{h}_K(\nu) + \mathrm{h}_K(-\nu) = \alpha. \tag{7.20}$$

Surprisingly, balls are not the only bodies having this property. In dimension two for instance, Reuleaux's triangles, which are obtained by intersecting three disks of radius α centered at the vertices of an equilateral triangle have constant width α. Moreover, Reuleaux's triangles have been proved by Lebesgue and Blaschke to minimize the area among two dimensional bodies of constant width.

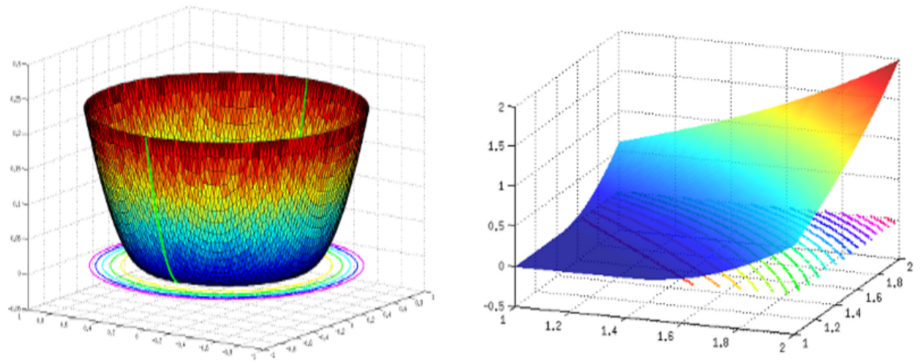

FIGURE 5. Approximation of the radial solution in the disk (left) and approximation of the solution in $[1,2]^2$ by P^3 functions (right).

In dimension three, this problem is more difficult. Indeed the mere existence of non trivial three-dimensional bodies of constant width is not so easy to establish. In particular, no finite intersection of balls has constant width, except balls themselves [19]. As a consequence, contrarily to the two dimensional case, the intersection of four balls centered at the vertices of a regular simplex is not of constant width. In 1912, E. Meissner described in [13] a process to turn this spherical body into an asymmetric bodies with constant width, by smoothing three of its circular edges. This famous body is called "Meissner tetrahedron" in the literature [11]. It is suspected to minimize the volume among three dimensional bodies with the same constant width. Let us point out that Meissner construction is not canonical in the sense that it requires to choose the set of three edges that have to be smoothed. As a consequence, there actually exists two kinds of "Meissner tetrahedron" having the same measure.

In these two constructions, the regular simplex seems to play a crucial role in the optimality (see also [8] for a more rigorous justification of this intuition). It is therefore natural to search for the body with constant width that is the closest from a regular simplex. Note that by uniqueness of the projection in an Hilbert space, Meissner tetrahedra cannot be the projections of a regular simplex with respect to the L^2 norm between support functions. Such an obstruction does not hold for the L^1 and L^∞ norm, which are not strictly convex. We illustrate below that our relaxed approach can be used to numerically investigate these questions. The optimization problem that we have to approximate is of the type

$$\min_{\mathrm{h}\in\mathcal{H}^s\cap\mathcal{W}} \|\mathrm{h}_0 - \mathrm{h}\|_{L^p(\mathcal{S}^2)}, \ 1 \leqslant p \leqslant \infty$$

where $\mathcal{W}$ is the set of function of $\mathcal{S}^2$ which satisfy the width constraints (7.20).

As explained is Section 3, we relax the constraint of being a support function, by imposing convexity-like conditions on a finite family of great circles of the sphere. In the experiments presented below the number of vertices in our mesh of $\mathcal{S}^2$ is 5000. We choose a family of 2000 great circles of $\mathcal{S}^2$ uniformly distributed (with respect to their normal direction) and a

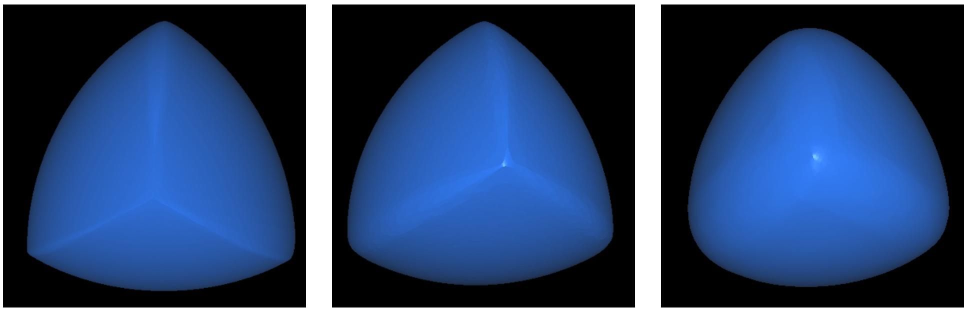

FIGURE 6. Reconstruction of the convex bodies associated to the L^1, L^2 and L^∞ projection of h_S without prescribing the width value.

step discretization of every circular arc equal to 0.02. Finally, the constraint $\mathcal{W}$ is approximated by imposing that antipodal values of the mesh must satisfy a set of linear equality constraints, which can be easily implemented in the proximal operator framework depicted in Section 4.2. Note that in this first experience, the value of the width constraint is not imposed.

We present in Table 1 and Figure 6, our numerical description of the projections of the support function of a regular simplex in the set of support function of constant width bodies for the L^1, L^2 and L^∞ norms. One can observe that the resulting support functions describe a body with constant width within an error of magnitude 0.1%. In other words the gap between the minimal width and the diameter is relatively less than 0.001. In the L^1 case we obtain a convex body whose surface area and volume are close to those of a Meissner body of same width, within a relative error of less than 0.01. We also performed the same experiment starting from the support functions of others platonic solids. In this case, where the value of the width is not imposed, it seems that whatever the norm and the solid, the closest body with constant width seems to always be a ball (whose radius can depend on the type of solid and of the exponent).

REFERENCES

[1] Néstor E Aguilera and Pedro Morin. Approximating optimization problems over convex functions. *Numerische Mathematik*, 111(1):1–34, 2008.

[2] Heinz H Bauschke, Regina S Burachik, Patrick L Combettes, Veit Elser, D Russell Luke, and Henry Wolkowicz. Fixed-point algorithms for inverse problems in science and engineering. 2011.

[3] J. Bertrand. Prescription of Gauss curvature using optimal mass transport. Preprint, 2010.

TABLE 1. Numerical information related to the projections of h_S

	Surface	Volume	Width	Relative width error
L^1 projection of h_S	2.6616	0.36432	0.951	< 0.001
L^2 projection of h_S	2.5191	0.34312	0.920	< 0.003
L^∞ projection of h_S	2.1351	0.28081	0.835	< 0.001

[4] Guillaume Carlier, Thomas Lachand-Robert, and Bertrand Maury. A numerical approach to variational problems subject to convexity constraint. *Numerische Mathematik*, 88(2):299–318, 2001.
[5] Philippe Choné and Hervé VJ Le Meur. Non-convergence result for conformal approximation of variational problems subject to a convexity constraint. 2001.
[6] Ivar Ekeland and Santiago Moreno-Bromberg. An algorithm for computing solutions of variational problems with global convexity constraints. *Numerische Mathematik*, 115(1):45–69, 2010.
[7] Alessio Figalli, Young-Heon Kim, and Robert J McCann. When is multidimensional screening a convex program? *Journal of Economic Theory*, 146(2):454–478, 2011.
[8] HaiLin Jin and Qi Guo. Asymmetry of convex bodies of constant width. *Discrete & Computational Geometry*, 47(2):415–423, 2012.
[9] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.
[10] B. Jüttler. Surface fitting using convex tensor-product splines. *J. Comput. Appl. Math.*, 84(1):23–44, 1997.
[11] Bernd Kawohl and Christof Weber. Meissner's mysterious bodies. *The Mathematical Intelligencer*, 33(3):94–101, 2011.
[12] Thomas Lachand-Robert and Édouard Oudet. Minimizing within convex bodies using a convex hull method. *SIAM Journal on Optimization*, 16(2):368–379, 2005.
[13] E. Meissner. Über Punktmengen konstanter Breite: Drei Gipsmodelle von Flächen konstanter Breite. *Zeitschrift der Mathematik und Physik*, 60:92–94, 1912.
[14] Mary C Meyer. An algorithm for quadratic programming with applications in statistics. Technical report, Technical Report, Colorado State University, 2010.
[15] Jean-Marie Mirebeau. Personal communication (2013).
[16] Adam M Oberman. A numerical method for variational problems with convexity constraints. *SIAM Journal on Scientific Computing*, 35(1):A378–A396, 2013.
[17] V. Oliker. Embedding $\mathcal{S}^n$ into $\mathbb{R}^{n+1}$ with given integral Gauss curvature and optimal mass transport on $\mathcal{S}^n$. *Advances in Mathematics*, 213(2):600–620, 2007.
[18] Édouard Oudet. Shape optimization under width constraint. *Discrete & Computational Geometry*, 49(2):411–428, 2013.
[19] Édouard Oudet et al. Bodies of constant width in arbitrary dimension. *Mathematische Nachrichten*, 280(7):740–750, 2007.
[20] Jean-Charles Rochet and Philippe Choné. Ironing, sweeping, and multidimensional screening. *Econometrica*, pages 783–826, 1998.