An iterated projection approach to variational problems under generalized convexity constraints

Guillaume Carlier* Xavier Dupuis†

April 26, 2016

Abstract

The principal-agent problem in economics leads to variational problems subject to global constraints of b-convexity on the admissible functions, capturing the so-called incentive-compatibility constraints. Typical examples are minimization problems subject to a convexity constraint. In a recent pathbreaking article, Figalli, Kim and McCann [19] identified conditions which ensure convexity of the principal-agent problem and thus raised hope on the development of numerical methods. We consider special instances of projections problems over b-convex functions and show how they can be solved numerically using Dykstra's iterated projection algorithm to handle the b-convexity constraint in the framework of [19]. Our method also turns out to be simple for convex envelope computations.

Keywords: Principal-agent problem, b-convexity constraint, convexity constraint, convex envelopes, iterated projections, Dykstra's algorithm.

Mathematics subject classification: 49M25, 65K15, 90C25.

1 Introduction

Variational problems subject to a convexity constraint arise in several different contexts such as mathematical economics [30], Newton's least resistance problem [7], [8], [20], optimal transport for the quadratic cost [6] or shape optimization [9], [21]. Existence of minimizers is generally not an issue since the set of convex functions have good local compactness properties in most reasonable functional spaces. However, on the one hand, the convexity constraint makes it difficult to write optimality conditions in the form of a tractable Euler-Lagrange PDE (see [22]) which could be used for instance to derive regularity results (see [13]). On the other hand, handling the convexity constraint numerically in a consistent and efficient way is also a challenging problem which has received a lot of attention in the last fifteen years. Choné and Le Meur [15] have first identified specific difficulties, one of which being that one cannot use conformal convex

*CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, France carlier@ceremade.dauphine.fr

†Department of Economics and Finance, LUISS Guido Carli, Viale Romania 32, 00197 Roma, Italy xdupuis@luiss.it

finite-elements since they form, as the meshsize vanishes, a set of functions which is not (by far!) dense in the set of convex functions. From this negative result, a lot of proposals have been made: use of interpolates of convex functions [14], supporting hyperplanes [18], finite-differences and semi-definite programming for nonnegativity of the Hessian [1], polyhedral approximation and directional convexity in the spirit of wide stencils for nonlinear PDEs [29], splitting of the directional convexity constraints and proximal algorithms [25]... Whatever method is used, some subtle tradeoff has to be made between provable convergence, accuracy and the computational cost resulting from the number of convexity constraints enforced at the discretized level (typically $O(N^2)$ with an N points grid). A major breakhtrough (for two dimensions) has been made recently by Mirebeau [26] who introduced a hierarchy of subcones of the cone of interpolates of convex functions and an adaptative refinement strategy leading typically on a grid with N points to essentially only $O(N \ln^2(N))$ convexity constraints.

It turns out that in optimal transport (for non quadratic transportation costs) and in principal-agent problems in economics (see section 2 for a brief presentation), the case of the convexity constraint is somehow special and corresponds to a very particular choice of the transport cost (quadratic) or of the valuation function (bilinear). The general form of the principal-agent problem typically involves a (valuation) function b of two variables x and y (in $\mathbb{R}^d$, say) and the relevant constraint on the admissible functions is that they are b-convex (a b-convex function is a function u of the variable x which can be written as a supremum of the form $u(x) = \sup_y \{b(x, y) - v(y)\}$ so that when b is the scalar product, we recover the convexity constraint). One then has to minimize a certain integral functional among b-convex functions; general existence results can be proved but not much more in general and this has been a serious limitation for understanding principal-agent models in several dimensions. It is only recently with the pathbreaking work of Figalli, Kim and McCann [19] that some conditions (intimately related to the conditions of Ma, Trudinger and Wang [24] for the regularity of optimal transport maps) were identified to make the principal-agent problem a convex program and, in the first place, the set of b-convex functions a convex set (see section 4 below for an elementary presentation in a special case). This raised the hope to use convex optimization algorithms to solve numerically some convex minimization problems posed over the set of b-convex functions, provided it is convex. To the best of our knowledge, the present paper is the first one which addresses b-convexity numerically (even for a quite resticted class of b). What makes the analysis of Figalli, Kim and McCann amenable to a computational approach is that, under some conditions detailed in section 3, the fact that u is b-convex and that $q = \nabla u$ can be rewritten as a single condition:

$$u(x') - u(x) \geq \Gamma_b(x', x, q(x)),\ \forall x', x,$$

for a certain function Γ_b depending (in a somehow indirect way) on the primitive datum b and which, provided it is convex in its last argument, immediately implies convexity of the set of b-convex functions. Now, discretizing this constraint on an N-points grid gives N^2 convex constraints. We will restrict ourselves to projection problems, i.e. quadratic minimization problems, so that, once discretized on a grid with N points, we face a projection on the intersection of convex sets, each of whom being given by one of the constraints above corresponding to two grid points x' and x. This can be solved with Dykstra's iterative projection algorithm (see the original papers by Dyskstra [17] for the case of convex cones, Boyle and Dykstra [5] for general closed convex sets; see also the extension to general Bregman projections by Bauschke and Lewis [3] or to monotone operators by Bauschke and Combettes [2], Combettes [16]). Each step of the algorithm consists in a projection onto only one of the convex sets; if these elementary projections are easy (it depends on the complexity and geometry of Γ_b, note that they are explicit in the case of the convexity constraint), Dykstra's algorithm can effectively be used even if it is computationally costly.

In the case of the convexity constraint, the use of Dykstra's iterative projection algorithm is not really competitive with most of the approaches recalled above. Nevertheless, as far as we know, it is the only one which can be quite directly adapted to the case of the b-convexity constraint. It is also worth noting that our numerical approach presents some similarities with the one of Ekeland and Moreno-Bromberg [18] in the sense that it replaces finite families of supporting hyperplanes by finite families of graphs of the nonlinear functions $\Gamma_b(., x_j, q_j)$. It also shares at least in spirit some similarities with the method of Mérigot and Oudet [25]: both methods somehow split the huge number of constraints into smaller subsets on which projections can be performed quickly.

The paper is organized as follows. In section 2, we motivate the study of variational problems with a b-convexity constraint by the principal-agent problem in economics. Section 3 recalls the structural conditions of Figalli, Kim and McCann under which b-convex functions form a convex set and thanks to which existence and uniqueness of a minimizer can be established. In section 4, we specify a particular class of b satisfying the conditions of Figalli, Kim and McCann, which can be seen as perturbations of the scalar product and for which we give an elementary and self-contained presentation. In section 5, we restrict ourselves even further to the case of quadratic objectives, i.e. to projection problems, and give some examples, in particular related to envelope computations. In section 6 we prove a convergence result for discretization of these projection problems. Section 7 explains how to use Dykstra's algorithm in this context. Finally, numerical results are presented in section 8.

2 Motivations

2.1 The principal-agent problem

To motivate variational problems with a b-convexity constraint, let us start by an informal description of the so-called principal-agent problem in economics. Let us consider a population of heterogeneous agents, each of whom has some vector of characteristics x in the sense that an agent with type x derives a utility $b(x, y) - p$ from consuming a good of type y for the price p. The set of possible agents types is denoted X, the set of possible goods types is denoted Y and b is a given function from $X \times Y \to \mathbb{R}$. We now consider a monopolist (the *principal*), proposing to the population of agents a contract menu that is a pair of functions $x \in X \mapsto (y(x), p(x)) \in Y \times \mathbb{R}$ specifying for each type of agent x the good and the price intended for her. The monopolist cannot observe directly the characteristics of agents, so the contract menu should be consistent with the fact that $(y(x), p(x))$ is really prefered by type x; this is the so-called incentive-compatibility condition:

$$b(x, y(x)) - p(x) \geq b(x, y(x')) - p(x'),\ \forall x, x'. \tag{2.1}$$

Moreover the agents have access to an outside option given by a fixed pair (y_0, p_0) and thus accept a contract from the monopolist only if

$$b(x, y(x)) - p(x) \geq u_0(x) := b(x, y_0) - p_0, \ \forall x. \tag{2.2}$$

The aim of the monopolist is then to minimize his total cost, which is given by an expression of the form

$$\int_X \phi(x, p(x), y(x)) \mathrm{d}x,$$

over the $(y(.), p(.))$ satisfying conditions (2.1) and (2.2). For instance, ϕ can be given by $\phi(x, p, y) = (c(y) - p)\rho(x)$ where $c(y)$ is the production cost of y and ρ represents the density of agents types, but other models are possible. Defining $u(x) := b(x, y(x)) - p(x)$ for every x, we see that the incentive-compatibility condition can be rewritten as:

$$u(x') - u(x) \geq b(x', y(x)) - b(x, y(x)), \ \forall x, x'. \tag{2.3}$$

Condition (2.3) of course imposes sharp restictions on the function u. Firstly, it imposes a global shape restriction, namely that $u(.)$ is a supremum of functions of the form $b(., y) - p$: this is the b-convexity constraint on u. Condition (2.3) can also be rephrased by saying that $u(.) - b(., y(x))$ is minimized at x so that if b is differentiable and u is differentiable at x (and in fact, there are well-known conditions on b which guarantee a priori that b-convex functions are differentiable at least almost everywhere, see section 3) then one obtains that $\nabla u(x) = \partial_x b(x, y(x))$ which is a local necessary condition for (2.3) to hold. If we go one step further, as was done in the seminal work of Figalli, Kim and McCann [19], by assuming that the relation $q = \partial_x b(x, y)$ can be inverted in the sense that $q = \partial_x b(x, y) \iff y = y_b(x, q)$ for some map y_b, then one can actually deduce $y(x)$ from the knowledge of $q(x) := \nabla u(x)$ by the relation $y(x) = y_b(x, q(x))$. Replacing $y(x)$ by $y_b(x, q(x))$ in (2.3) and defining $\Gamma_b(x', x, q) := b(x', y_b(x, q)) - b(x, y_b(x, q))$, we obtain the following reformulation of the incentive-compatibility constraint:

$$u(x') - u(x) \geq \Gamma_b(x', x, q(x)), \ \forall x, x'. \tag{2.4}$$

Note that (2.4) implies that, if u is differentiable at x, then $\nabla u(x) = \partial_x b(x, y_b(x, q(x))$, i.e. $q(x) = \nabla u(x)$ so that constraint (2.4) encodes the relation $q = \nabla u$. Using the variables (q, u) instead of (y, p) (so that $p(x) = b(x, y(x)) - u(x)$ and $q(x) = \partial_x b(x, y(x))$, $y(x) = y_b(x, q(x))$), the monopolist's program rewrites

$$\inf_{(u,q)} \int_X L(x, u(x), q(x)) \mathrm{d}x$$

subject to $u \geq u_0$ and (2.4), with $L(x, u, q) := \phi(x, b(x, y_b(x, q)) - u, y_b(x, q))$. Of course, a crucial role is played here by the map y_b but also by the function $\Gamma_b(x', x, q) := b(x', y_b(x, q)) - b(x, y_b(x, q))$ which appears in the right-hand side of the incentive-compatibility condition written in the form (2.4).

2.2 The case of the convexity constraint

Variational problems subject to a convexity constraint arise in different applied settings: economics with Rochet-Choné model [30] (which corresponds to the principal-agent problem with a bilinear b), but also Newton's least resistance problem (see [7], [8], [20]). We shall also see in section 5 how to relate such variational problems with the computations of convex envelopes, a problem with its own interest. Such problems consist in minimizing, given a convex domain X of $\mathbb{R}^d$, an integral functional of the form $\int_X L(x, u(x), \nabla u(x))\mathrm{d}x$ among convex functions. This is in fact equivalent to minimize $\int_X L(x, u(x), q(x))\mathrm{d}x$ subject to the constraint

$$u(x') - u(x) \geq (x' - x) \cdot q(x),\ \forall x', x$$

which automatically implies convexity of u as well as $q \in \partial u$ hence $q = \nabla u$ a.e.. In other words, the constraints that u is convex and $q = \nabla u$ are a particular case of the non-local condition (2.4) with $b(x, y) = x \cdot y$ so that $y_b(x, q) = q$ and $\Gamma_b(x', x, q) = (x' - x) \cdot q$. The idea of writing the convexity constraints in this way for numerical purposes was first used by Ekeland and Moreno-Bromberg [18].

3 Existence and uniqueness

Let us now consider the following assumptions:

(A1) $b \in C^1(\overline{X} \times \overline{Y})$ where X and Y are open subsets of $\mathbb{R}^d$ with X convex and bounded, and b is uniformly semiconvex with respect to x, i.e. there exists $\lambda \geq 0$ such that for every $y \in \overline{Y}$, $x \mapsto b(x, y) + \frac{\lambda}{2}|x|^2$ is convex on $\overline{X}$,

(A2) b is twisted (or satisfies the generalized Spence-Mirrlees condition), i.e. for every $x \in \overline{X}$, $y \mapsto \partial_x b(x, y)$ is a diffeomorphism from $\overline{Y}$ to $Q_x := \partial_x b(x, \overline{Y})$, the set Q_x is closed and convex, and its graph $\mathrm{Graph}(Q) := \{(x, q) \in \mathbb{R}^d \times \mathbb{R}^d,\ x \in \overline{X},\ q \in Q_x\}$ is closed; this enables one to define the map y_b on $\mathrm{Graph}(Q)$ by

$$y = y_b(x, q) \iff \partial_x b(x, y) = q.$$

(A3) For all $(x', x) \in \overline{X}^2$ the map

$$q \in Q_x \mapsto \Gamma_b(x', x, q) := b(x', y_b(x, q)) - b(x, y_b(x, q)), \tag{3.1}$$

is convex and continuous on Q_x.

Assumption **(A1)** ensures that every (finite) b-convex function u is semi-convex and therefore:

- u is subdifferentiable at every point, i.e. for every $x \in X$,

$$\partial u(x) := \{q \in \mathbb{R}^d\ :\ u(x + h) \geq u(x) + q \cdot h + o(h)\} \neq \emptyset.$$

- u is differentiable at every point except possibly a set of Hausdorff dimension at most $d - 1$ (hence of zero measure).

The general form of the variational problems we are interested in is:

$$\inf_{(u,q)} J(u,q) := \int_X L(x, u(x), q(x))\mathrm{d}x \tag{3.2}$$

subject to

$$q(x) \in Q_x, \text{ for a.e. } x \in X, \tag{3.3}$$

and

$$u(x') - u(x) \geq \Gamma_b(x', x, q(x)), \text{ for a.e. } x \in X \text{ and all } x' \in X, \tag{3.4}$$

where Γ_b is defined by (3.1). Note that inequality (3.4) is an equality when $x' = x$ so, under assumption **(A2)**, whenever u is differentiable at x, we have $q(x) = \nabla u(x)$. Thanks to assumption **(A3)**, first outlined by Figalli, Kim and McCann [19], the constraint (3.4) is convex in (u, q) (equivalently in $(u, \nabla u)$); if, in addition $L(x, \cdot, \cdot)$ is convex, then the minimization problem (3.2)-(3.4) is a convex program. Therefore we have the following well-posedness result (existence essentially follows from [10] and uniqueness from [19]; for the sake of completeness we give a short proof).

Theorem 3.1. *Assume that*

- b *satisfies* **(A1)**, **(A2)** *and* **(A3)**,
- L *is a lsc integrand such that, for a.e.* x, $(u, q) \in \mathbb{R} \times Q_x \mapsto L(x, u, q)$ *is strictly convex, and it satisfies for some* $C > 0$ *the coercivity condition*

$$L(x, u, q) \geq C(|u| + |q| - 1), \text{ for a.e. } x \in X \text{ and all } (u, q) \in \mathbb{R} \times Q_x, \tag{3.5}$$

- *there exists* (u_0, q_0) *satisfying (3.4)-(3.3) such that* $J(u_0, q_0) < +\infty$.

Then the problem (3.2)-(3.4) admits a unique solution.

Proof. Let (u_n, q_n) be a minimizing sequence for (3.2)-(3.4). By (3.5), $(u_n)_n$ is bounded in $W^{1,1}(X)$. Since the functions (u_n) are all λ-convex (with the same λ), taking a subsequence if necessary, we may assume that u_n converges locally uniformly to some λ-convex u and $q_n = \nabla u_n$ converges to $q := \nabla u$ a.e. (see for instance [11]). Then by Fatou's Lemma, we deduce that $J(u, q)$ is the infimum of the problem (3.2)-(3.4). Note that for a.e. x, $q(x) \in Q_x$ since this set is closed. It remains to show that (u, q) satisfies the constraint (3.4), but this is obvious by passing to the limit in the inequality

$$u_n(x') - u_n(x) \geq \Gamma_b(x', x, q_n(x))$$

which holds for all x', all n and all x in a set of full measure on which we may further assume that $q_n(x)$ converges to $q(x)$. This shows existence; uniqueness directly follows from the convexity of the constraints and the strict convexity of J.

□

To show consistency of discrete approximations in section 6, we shall need the following elementary result:

Lemma 3.2. *Assume that b satisfies* **(A1)** *and* **(A2)***. Then for all $(x, x', \overline{x}) \in \overline{X}^3$ and $\overline{q} \in Q_{\overline{x}}$, one has*

$$y_b(\overline{x}, \overline{q}) = y_b(x, \partial_x b(x, y_b(\overline{x}, \overline{q}))), \tag{3.6}$$

and

$$\Gamma_b(x', \overline{x}, \overline{q}) - \Gamma_b(x, \overline{x}, \overline{q}) = \Gamma_b(x', x, \partial_x b(x, y_b(\overline{x}, \overline{q}))). \tag{3.7}$$

Proof. Let $q := \partial_x b(x, y_b(\overline{x}, \overline{q}))$. By definition, $\partial_x b(x, y_b(x, q)) = q$, which implies $y_b(x, q) = y_b(\overline{x}, \overline{q})$ (i.e. proves (3.6)) since b is twisted. Then (3.7) immediately follows since $\Gamma_b(x', \overline{x}, \overline{q}) - \Gamma_b(x, \overline{x}, \overline{q}) = b(x', y_b(\overline{x}, \overline{q})) - b(x, y_b(\overline{x}, \overline{q})) = b(x', y_b(x, q)) - b(x, y_b(x, q)) = \Gamma_b(x', x, q)$.

□

4 A tractable specification for b

Checking assumption **(A3)** directly on b is not easy in practice since it involves y_b via Γ_b. There are however (a few) known examples, which can be found in [19], in connection with the seminal work of Ma, Trudinger and Wang [24] on the regularity of optimal transport maps. For the sake of tractability, we shall restrict ourselves to a class of particular b for which computations can be performed explicitly up to a certain point. Namely, from now on, we will only consider perturbations of the scalar product. More precisely, we consider the following specification:

(B1) b is of the form

$$b(x, y) := x \cdot y + f(x)g(y) \tag{4.1}$$

with $(f, g) \in C^1(\overline{X}) \times C^1(\mathbb{R}^d)$ where X is an open bounded convex subset of $\mathbb{R}^d$ (here $Y = \mathbb{R}^d$), f and g are convex, $g \geq 0$, and

$$\inf_{(x,y) \in X \times \mathbb{R}^d} \nabla f(x) \cdot \nabla g(y) =: \kappa > -1. \tag{4.2}$$

Note that (4.2) is trivially satisfied in the following two cases:

- if g is Lipschitz on $\mathbb{R}^d$ and $\|\nabla f\|_{L^\infty(X)} \|\nabla g\|_{L^\infty(\mathbb{R}^d)} < 1$;
- if $\nabla f(x) \cdot \nabla g(y) = 0$ for all $(x, y) \in X \times \mathbb{R}^d$ (which is in particular the case if for instance $f(x) = f(x_1)$ and $g(y) = g(y_2, \ldots, y_d)$).

It is known that suitable perturbations of the scalar product of the form (4.1) satisfy the convexity assumption **(A3)** (see [19], [23], [24]). We shall give in the next paragraph an elementary and self-contained proof as well as some properties of y_b and Γ_b under assumption **(B1)** (which is slightly weaker than the one considered in the previous references).

4.1 Properties of y_b and Γ_b

Note first that **(A1)** is satisfied with $\lambda = 0$ since f is convex and $g \geq 0$, therefore b-convex functions are convex as suprema of convex functions. Next, given $(x, q) \in \overline{X} \times \mathbb{R}^d$, $y = y_b(x, q)$ is obtained by solving

$$y = q - g(y)\nabla f(x),$$

or equivalently by solving the scalar equation

$$\lambda - g(q - \lambda \nabla f(x)) = 0 \tag{4.3}$$

and setting $y = q - \lambda \nabla f(x)$; (4.3) has a unique nonnegative root since its left-hand side, as a function of λ, has derivative not less than $1 + \kappa > 0$ by (4.2) and value $-g(q) \leq 0$ for $\lambda = 0$. This shows that **(A2)** holds with $Q_x = \mathbb{R}^d$. Moreover,

$$y_b(x, q) = q - g_b(x, q)\nabla f(x) \tag{4.4}$$

where

$$g_b(x, q) := g(y_b(x, q)) \tag{4.5}$$

is the unique root of (4.3). Thanks to (4.2), one readily checks that

$$0 \leq g_b(x, q) \leq \frac{g(q)}{1 + \kappa}. \tag{4.6}$$

By (4.2), (4.6) and the fact that g is locally Lipschitz, one also deduces that for any compact subset K of $\mathbb{R}^d$ there exists a positive constant C_K such that, for all $x, x' \in \overline{X}$ and $q, q' \in K$,

$$|g_b(x, q) - g_b(x', q')| \leq C_K(|q - q'| + |\nabla f(x) - \nabla f(x')|). \tag{4.7}$$

which implies the continuity of g_b, as well as that of y_b and Γ_b by (4.4). Let us now check that the convexity assumption **(A3)** is satisfied. By direct computation,

$$\Gamma_b(x', x, q) = (x' - x) \cdot q + D_f(x', x) g_b(x, q)$$

where D_f is the Bregman divergence associated to f:

$$D_f(x', x) := f(x') - f(x) - \nabla f(x) \cdot (x' - x),$$

which is nonnegative since f is convex. The convexity of Γ_b with respect to q thus amounts to showing that $g_b(x, q)$ is convex with respect to q, which can be done as follows. Let $(x, q_0, q_1, t) \in \overline{X} \times \mathbb{R}^d \times \mathbb{R}^d \times [0, 1]$, and set $q_t := (1 - t)q_0 + tq_1$, $\lambda_0 := g_b(x, q_0)$, $\lambda_1 := g_b(x', q')$, and $\lambda_t = (1 - t)\lambda_0 + t\lambda_1$. Then $g_b(x, q_t)$ is the root of $\lambda \mapsto \lambda - g(q_t - \lambda \nabla f(x))$, which is increasing with respect to λ by (4.2) and nonnegative at $\lambda = \lambda_t$ by convexity of g. We deduce that $g_b(x, q_t) \leq \lambda_t$, which shows that **(A3)** is satisfied.

Recall that if u is b-convex, then it is convex hence differentiable a.e. on X. Moreover if we denote by $A \subset X$ the set of points at which u is differentiable, then ∇u is continuous on A (see [31]). The constraints that u is b-convex and $q = \nabla u$ can be expressed as

$$u(x') - u(x) \geq \Gamma_b(x', x, q(x)), \ \forall (x', x) \in X \times A.$$

Now if $x \in X \setminus A$, one can choose a sequence x_n in A, remaining away from ∂X and converging to x. Since u is locally Lipschitz, we may further assume that the bounded sequence $q_n := \nabla u(x_n)$ converges to some $\overline{q}$. Up to redefining q at x by setting $q(x) = \overline{q}$, and using the fact that Γ_b is continuous, we see that there is no loss of generality in imposing that the inequality $u(x') - u(x) \geq \Gamma_b(x', x, q(x))$ actually holds for every $(x', x) \in X^2$.

4.2 Further specification for the numerical examples

We will consider in our simulations (section 8) a special case where g_b (and thus y_b and Γ_b) can be explictly computed. This special case is given by

$$g(y) := \sqrt{1 + |y|^2}$$

and f any C^1 convex function with Lipschitz constant on $\overline{X}$ strictly less than 1 (for instance $f(x) = \sqrt{1 + |x|^2}$ in our simulations). In this case, the solution $g_b(x, q)$ to (4.3) is the unique positive root of the quadratic equation

$$\lambda^2 = 1 + |q - \lambda \nabla f(x)|^2,$$

and thus has the following explicit expression

$$g_b(x, q) = \frac{\left[(q, \nabla f(x))^2 + (1 - |\nabla f(x)|^2)(1 + |q|^2)\right]^{1/2} - q \cdot \nabla f(x)}{1 - |\nabla f(x)|^2}.$$

Note that explicit expressions are also obtained for any f and g satisfying (4.2) with $\nabla f(x) \cdot \nabla g(y) = 0$ for all $(x, y) \in X \times \mathbb{R}^d$. Indeed, in this case, $\lambda \mapsto g(q - \lambda \nabla f(x))$ is constant hence $g_b(x, q) = g(q)$. This is in particular the case when $f(x) = f(x_1)$ and $g(y) = g(y_2, \ldots, y_d)$.

5 A tractable specification for L

We have addressed in section 4 the choice of a tractable class of b so that conditions **(A1)**, **(A2)** and **(A3)** are satisfied. To make the problem (3.2) tractable numerically, we shall restrict ourselves to projection problems, i.e. to quadratic Lagrangians L. From now on, we assume that b is a perturbation of the scalar product of the form (4.1) which satisfies assumption **(B1)**, and we consider the quadratic problem

$$\inf_{(u,q)} J(u, q) := \int_X \Big(\frac{\alpha(x)}{2} |q(x) - q_0(x)|^2 + \frac{\beta(x)}{2} |u(x) - u_0(x)|^2 \Big) \mathrm{d}x \tag{5.1}$$

subject to

$$u(x') - u(x) \geq \Gamma_b(x', x, q(x)), \ \forall (x', x) \in X^2, \tag{5.2}$$

where $(u, q) \in C(X) \cap H^1(X) \times L^2(X, \mathbb{R}^d)$ and $(u_0, q_0, \alpha, \beta) \in C(\overline{X}, \mathbb{R}) \times C(\overline{X}, \mathbb{R}^d) \times C(\overline{X}, (0, +\infty))^2$. We denote

$$K_b := \{(u, q) \in C(X) \cap H^1(X) \times L^2(X, \mathbb{R}^d) \, : (u, q) \text{ satisfies } (5.2)\}. \tag{5.3}$$

We have already seen that if $(u, q) \in K_b$, then necessarily u is b-convex and $q = \nabla u$, and explained at the end of section 4.1 why the b-convexity constraint can be imposed on the whole of $X \times X$. We know from theorem 3.1 that (5.1)-(5.2) admits a unique solution which we denote $(\overline{u}, \overline{q}) = (\overline{u}, \nabla \overline{u})$. Of course, there is no particular extra difficulty in considering additional pointwise convex constraints on u or q such as $u \geq v$, $u \leq v$, $u = v$ on ∂X or $q \in Q$ with Q a closed convex subset of $\mathbb{R}^d$.

The reason why we consider only quadratic Lagrangians L, i.e. projection problems, is that we wish to keep numerical computations as simple as possible. It is in fact possible (but computationally costly) to use Dykstra's algorithm with general Bregman distances to address convex but non quadratic Lagrangians, but we leave this for future research. We may also view the projection problem as a first step to address more general Lagrangians for instance by a projected gradient method, each step of which amounts to solving a problem of the form (5.1)-(5.2). We present in the next paragraphs some examples which, we believe, motivate directly the choice of a quadratic L. Let us also remark that we require α and β to be everywhere strictly positive, that is we ask L to be strongly convex in both variables u and q; this is due to the fact that in Dykstra's iterative projection algorithm, the condition $q = \nabla u$ is not enforced during the iterations, so to guarantee convergence, this nondegeneracy condition is necessary.

5.1 Rochet and Choné principal-agent model

In [30], Rochet and Choné considered the principal-agent problem in the case of a bilinear utility, a quadratic production cost $c(y) := \frac{1}{2}|y|^2$ and 0 reservation utility. The Rochet and Choné problem then consists in the quadratic minimization problem with a convexity and a nonnegativity constraints:

$$\inf \Big\{ \int_X [\frac{1}{2}|\nabla u(x)|^2 - x \cdot \nabla u(x) + u(x)]\rho(x)\mathrm{d}x \ : \ u \text{ convex}, \ u \geq 0 \Big\}$$

which can equivalently be rewritten as

$$\inf \Big\{ \int_X [\frac{1}{2}|q(x)|^2 - x \cdot q(x) + u(x)]\rho(x)\mathrm{d}x \ : \ (u, q) \in K_b, u \geq 0 \Big\} \tag{5.4}$$

with $b(x, y) = x \cdot y$ and K_b defined by (5.3). This is not exactly a projection problem of the form (5.1)-(5.2) because there is no explicit term in u^2 in (5.4), so in our simulations we will regularize it adding an extra quadratic term with a (small) parameter $\varepsilon > 0$. Namely, the regularized version of (5.4) then reads

$$\inf \Big\{ \int_X [\frac{1}{2}|q(x)|^2 - x \cdot q(x) + \frac{\varepsilon}{2} u^2(x) + u(x)]\rho(x)\mathrm{d}x \ : \ (u, q) \in K_b, u \geq 0 \Big\}, \tag{5.5}$$

which corresponds to a projection problem with $q_0 = x$, $u_0 = -\varepsilon^{-1}$, $\alpha = \rho$, $\beta = \varepsilon\rho$, and an additional nonnegativity constraint on u.

5.2 Convex envelopes

Let X be a bounded open convex subset of $\mathbb{R}^d$ and $u_0 \in C(\overline{X}) \cap H^1(X)$, the convex envelope, u_0^{**} of u_0 is by definition the largest convex function which is pointwise below

u_0 in $\overline{X}$. Obviously, u_0^{**} minimizes $\int_X (u-u_0)^2$ among all convex functions u such that $u \le u_0$ in X, but this is not a projection problem of the form (5.1)-(5.2) because here we have no term in $|\nabla u|^2$. One could of course regularize this minimization problem by adding $\frac{\varepsilon}{2}\int_X |\nabla u|^2$ (we will actually do it to approximate more general envelopes in the next paragraph). It turns out however that the convex envelope also minimizes a Dirichlet integral. Indeed, let u be a H^1 convex function such that $u \le u_0$ and $u = u_0^{**}$ on ∂X; then by convexity of the square and an integration by parts,

$$\frac{1}{2}\int_X (|\nabla u|^2 - |\nabla u_0^{**}|^2) \ge \int_X \Delta u_0^{**}(u_0^{**} - u),$$

and the latter is nonnegative since $u \le u_0^{**}$, and u_0^{**} being convex, it has a nonnegative Laplacian. Hence u_0^{**} is also the minimizer of the Dirichlet integral $\frac{1}{2}\int_X |\nabla u|^2$ among all convex functions which agree with u_0^{**} on ∂X and are below u_0 on $\overline{X}$. Therefore provided that $u_0^{**} = u_0$ on ∂X, the convex envelope u_0^{**} can be obtained by solving

$$\inf\Big\{\int_X \frac{\alpha}{2}|\nabla u|^2 + \frac{\beta}{2}|u-u_0|^2 \ : \ u \text{ convex}, \ u \le u_0, \ (u-u_0)|_{\partial X} = 0\Big\}$$

for every positive constants α and β. Of course, to make this observation relevant in practice to compute u_0^{**}, we should already know u_0^{**} on ∂X. It is however well-known that $u_0 = u_0^{**}$ on ∂X (equivalently, u_0 agrees with a convex function on ∂X) in the following cases:

- if ∂X contains no segment, by the following formula for the convex envelope:

  $$u_0^{**}(x) = \inf\Big\{\sum_{i=1}^{d+1}\lambda_i u_0(x_i) \ : \ \lambda_i \ge 0, \ x_i \in \overline{X}, \ \sum_{i=1}^{d+1}\lambda_i = 1, \ \sum_{i=1}^{d+1}\lambda_i x_i = x\Big\},$$

  which implies that $u_0^{**} = u_0$ on ∂X if every boundary point of X is extreme;

- if the dimension $d = 1$, as a particular case of the previous one;

- if u_0 is the sum of a convex function and a function with compact support in X.

It is worth mentioning here some purely PDE approaches to the computation of convex envelopes. Vese [32] showed that the convex envelope can be obtained as the limit for large time of a solution of a nonlinear parabolic equation (also see [12] for exponential convergence results) and Oberman [27], [28] developed a direct approach based on the observation that convex envelopes solve a nonlinear elliptic equation.

5.3 More general envelopes

Since by definition, b-convex functions are stable by suprema, one can also define the b-convex envelope of a function u_0 as the largest b-convex function below u_0. If b is as in section 4 (or more generally if $b(.,y)$ is convex for every y), then b-convex functions are convex hence subharmonic; the same argument as for convex envelopes thus enables to conclude that if u_0 coincides with its b-convex envelope on ∂X, then the b-convex envelope of u_0 can also be obtained by the quadratic minimization problem

$$\inf\Big\{\frac{1}{2}\int_X \frac{\alpha}{2}|q|^2 + \frac{\beta}{2}|u-u_0|^2 \ :\ (u,q)\in K_b,\ u\le u_0,\ (u-u_0)|_{\partial X}=0\Big\}. \tag{5.6}$$

Let us mention two cases where the convex envelope of u_0 coincides with u_0 on ∂X:

- if u_0 is the sum of a b-convex function and a nonnegative function which vanishes on ∂X;
- if the dimension $d = 1$ and u_0 has compact support in X (see below).

Lemma 5.1. *Assume that b satisfies **(B1)** with $X = (0,1)$ and let $u_0 \in C_c((0,1))$. Then u_0 coincides with its b-convex envelope at 0 and 1.*

Proof. Let $\varepsilon > 0$ be such that $u_0 = 0$ on $[0,\varepsilon]\cup[1-\varepsilon,1]$. Let us denote by v the b-convex envelope of u_0; by definition $v \le u_0$ up to the boundary of X. To prove equality at 0 and 1, it is enough to construct a b-convex function w such that $w \le u_0$ and $w(0) = w(1) = 0$. Define w by $w(x) := \max\{w_1(x), w_2(x)\}$ where

$$w_1(x) := b(x,y_1) - b(0,y_1),\ w_2(x) := b(x,y_2) - b(1,y_2),$$

with $y_1 < 0 < y_2$ and $|y_i| \ge M$, $i = 1,2$, with M to be chosen properly. By (4.2), we have $w_1'(x) \le y_1(1+\kappa) + \|f'\|_\infty|g(0)| \le -M(1+\kappa) + \|f'\|_\infty|g(0)|$ where $\kappa > -1$. Then choosing M large enough, $w_1(x) \le -\frac{M}{2}(1+\kappa)x$ for every $x\in[0,1]$. If we further restrict M so that $-\frac{M}{2}(1+\kappa) < \varepsilon^{-1}\min u_0$, we get $w_1 \le u_0$. In a similar way for such an M, we also have $w_2 \le u_0$. Hence $v \ge w$, so that $v(0) \ge w(0) = 0$ and $v(1) \ge w(1) = 0$. □

In the general case where one cannot take for granted that u_0 coincides with its b-convex envelope on ∂X, one can still approximate the b-convex envelope by solving for a small $\varepsilon > 0$ the problem

$$\inf\Big\{\frac{1}{2}\int_X \frac{\varepsilon}{2}|q|^2 + \frac{1}{2}|u-u_0|^2 \ :\ (u,q)\in K_b,\ u\le u_0\Big\}. \tag{5.7}$$

6 Discretization and convergence

We now address the discretization of (5.1)-(5.2) and our aim in this section is to prove a Γ-convergence result. To prove Γ-convergence, it will be useful to have that Lipschitz functions are dense in energy for the problem (5.1)-(5.2), which requires the extra assumption that g is Lipschitz on $\mathbb{R}^d$ (this would not be needed for variants of (5.1)-(5.2) with additional constraints that bound uniformly q). Note first that since here b-convex functions are convex, they are locally Lipschitz hence $K_b \subset W^{1,\infty}_{\mathrm{loc}}(X)\times L^\infty_{\mathrm{loc}}(X)$. Let $(u,q)\in K_b$, $M > 0$ and (u_M, q_M) be defined by

$$u_M(x') := \sup_{x\in X,\ |q(x)|\le M}\{u(x) + \Gamma_b(x',x,q(x))\},\ q_M := \nabla u_M. \tag{6.1}$$

By construction u_M is b-convex, $u_M \leq u$, $u_M = u$ on $A_M := \{x \in X \ : \ |q(x)| \leq M\}$; since q is locally bounded, u_M converges locally uniformly to u and q_M converges to q almost everywhere. If $x \notin A_M$, let $x_n \in A_M$ be such that

$$u_M(x) = \lim_n u(x_n) + \Gamma_b(x, x_n, q(x_n));$$

since $|q(x_n)| \leq M$, passing to a subsequence, we may assume that $(x_n, q_n, y_n) = (x_n, q(x_n), y_b(x_n, q(x_n)))$ converges to some $(\overline{x}, \overline{q}, \overline{y})$ with $\overline{x} \in \overline{X}$, $|\overline{q}| \leq M$ and $\overline{y} = y_b(\overline{x}, \overline{q})$ since y_b is continuous by (4.7). Let now $x' \in X$; since $u_M(x') \geq u(x_n) + \Gamma_b(x', x_n, q_n)$, using the continuity of Γ_b, we get $u_M(x') - u_M(x) \geq \Gamma_b(x', \overline{x}, \overline{q}) - \Gamma_b(x, \overline{x}, \overline{q})$, so that if u_M is differentiable at x (which is the case almost everywhere), we necessarily have $\nabla u_M(x) = q_M(x) = \partial_x b(x, y_b(\overline{x}, \overline{q}))$. Lemma 3.2 thus gives

$$u_M(x') - u_M(x) \geq \Gamma_b(x', \overline{x}, \overline{q}) - \Gamma_b(x, \overline{x}, \overline{q}) = \Gamma_b(x', x, q_M(x)),$$

so that $(u_M, q_M) \in K_b$. Moreover, recalling (4.4) and (4.5), we have

$$\begin{aligned} q_M(x) = \partial_x b(x, y_b(\overline{x}, \overline{q})) &= y_b(\overline{x}, \overline{q}) + g_b(\overline{x}, \overline{q}))\nabla f(x) \\ &= \overline{q} + g_b(\overline{x}, \overline{q})(\nabla f(x) - \nabla f(\overline{x})), \end{aligned}$$

so that by (4.6),

$$|q_M(x)| \leq |\overline{q}| + \frac{2\|\nabla f\|_{L^\infty(X)}}{1+\kappa} g(\overline{q}). \tag{6.2}$$

Lemma 6.1. *Assume that b satisfies **(B1)** with the extra assumption that g is Lipschitz on $\mathbb{R}^d$. Let $(u, q) \in K_b$ and (u_M, q_M) be defined by (6.1). Then there exists a constant C (independent of M and (u, q)) such that $|q_M| \leq C(1 + |q|)$, which implies that u_M converges to u in H^1 as $M \to +\infty$ and that the minimum of J over K_b coincides with its infimum over $K_b \cap (W^{1,\infty}(X) \times L^\infty(X))$.*

Proof. Only the inequality $|q_M| \leq C(1 + |q|)$ has to be shown: the H^1 convergence will directly follow by Lebesgue's dominated convergence Theorem. If $|q(x)| \leq M$, there is nothing to prove since $q_M = \nabla u_M = \nabla u = q$ a.e. in A_M. If $|q(x)| > M$ and u_M is differentiable at x, then by (6.2), the fact that $|\overline{q}| \leq M$ and the assumption that g is Lipschitz, we have $|q_M(x)| \leq C(1 + M) \leq C(1 + |q(x)|)$ for some constant C.

□

6.1 Discretization

We now discretize (5.1)-(5.2) as follows. Denoting by λ the Lebesgue measure on X, we consider for each discretization parameter $h > 0$, a finite subset $X_h = \{x_i^h, \ i = 1, \dots, N_h\}$ of X in such a way that the sequence of discrete measures

$$\lambda^h := \frac{\lambda(X)}{N_h} \sum_{i=1}^{N_h} \delta_{x_i^h}$$

approaches λ, i.e.

$$\lambda^h \overset{*}{\rightharpoonup} \lambda \text{ as } h \to 0. \tag{6.3}$$

We now consider the following discretized version of (5.1)-(5.2):

$$\inf_{(u^h,q^h)\in K_b^h} J^h(u^h,q^h) := \frac{\lambda(X)}{N_h}\sum_{i=1}^{N_h}\left[\frac{\alpha(x_i^h)}{2}|q_i^h - q_0(x_i^h)|^2 + \frac{\beta(x_i^h)}{2}|u_i^h - u_0(x_i^h)|^2\right] \quad (6.4)$$

where

$$K_b^h := \{(u_i^h, q_i^h) \in (\mathbb{R}\times\mathbb{R}^d)^{N_h} \ : \ u_i^h - u_j^h \geq \Gamma_b(x_i^h, x_j^h, q_j^h), \ \forall (i,j) \in \{1,\dots,N_h\}^2\}. \quad (6.5)$$

Note that (6.4)-(6.5) is a finite-dimensional projection problem, and under our assumptions K_b^h is a closed and convex set so that existence and uniqueness of a minimizer which we denote by $(\overline{u}^h, \overline{q}^h)$ is straightforward.

Elements of K_b^h can easily be extended to elements of K_b admissible for the continuous problem as follows:

Lemma 6.2. *Given $(u^h, q^h) \in K_b^h$, let us extend u^h to X by setting*

$$u^h(x) := \max_{i=1,\cdots,N_h}\{u_i^h + \Gamma_b(x, x_i^h, q_i^h)\}, \ \forall x \in \overline{X}.$$

For each $x \in \overline{X}$, select an index i among those for which the maximum above is achieved which minimizes the distance between x and x_i^h and set

$$q^h(x) := \partial_x b(x, y_b(x_i^h, q_i^h)),$$

so that $q^h(x_i^h) = \partial_x b(x_i^h, y_b(x_i^h, q_i^h)) = q_i^h$. Then the obtained extension (which by abusing notations we continue to denote by (u^h, q^h)) belongs to K_b.

Proof. First observe that $(u^h, q^h) \in W^{1,\infty}(X) \times L^\infty(X)$. Let $x \in \overline{X}$ and let i be an index for which $u^h(x) = u_i^h + \Gamma_b(x, x_i^h, q_i^h)$ and $q^h(x) = \partial_x b(x, y_b(x_i^h, q_i^h))$. For $x' \in X$, we have $u^h(x') \geq u_i^h + \Gamma_b(x', x_i^h, q_i^h)$ hence

$$u^h(x') - u^h(x) \geq \Gamma_b(x', x_i^h, q_i^h) - \Gamma_b(x, x_i^h, q_i^h) = \Gamma_b(x', x, q^h(x))$$

where the last equality follows from Lemma 3.2. □

Extending $(u^h, q^h) \in K_b^h$ as in Lemma 6.2, one can write

$$J^h(u^h, q^h) = \int_X L(x, u^h(x), q^h(x))\mathrm{d}\lambda^h(x)$$

whereas the functional J for the continuous problem is, for all $(u, q) \in K_b$,

$$J(u,q) = \int_X L(x, u(x), q(x))\mathrm{d}x$$

where

$$L(x,u,q) := \frac{\alpha(x)}{2}|q - q_0(x)|^2 + \frac{\beta(x)}{2}|u - u_0(x)|^2, \ \forall (x,u,q) \in X \times \mathbb{R} \times \mathbb{R}^d.$$

6.2 Convergence

This section is devoted to a detailed Γ-convergence proof. For the Γ-liminf inequality, the following result will be useful:

Lemma 6.3. *Let $(u^h, q^h) \in K_b^h$ extended as in Lemma 6.2 be such that*

$$\sup_h \|q^h\|_{L^2(\lambda^h)} < +\infty. \tag{6.6}$$

Assume that u^h converges uniformly on compact subsets of X to some function u and that q^h converges a.e. to $q := \nabla u$. Then

1. *$(u, q) \in K_b$,*
2. *for every $F \in C_b(X \times \mathbb{R}^d)$, one has*

$$\lim_{h\to 0} \frac{\lambda(X)}{N_h} \sum_{i=1}^{N_h} F(x_i^h, q_i^h) = \lim_{h\to 0} \int_X F(x, q^h(x))\mathrm{d}\lambda^h(x) = \int_X F(x, \nabla u(x))\mathrm{d}x, \tag{6.7}$$

3. *the following Γ-liminf inequality holds:*

$$\liminf_h J^h(u^h, q^h) \geq J(u, q). \tag{6.8}$$

Proof. Let us define the discrete measure γ^h on $X \times X \times \mathbb{R}^d$ by $\gamma^h := \lambda^h \otimes ((\mathrm{id}, q^h)_\# \lambda^h)$, i.e. for every $F \in C_b(X \times X \times \mathbb{R}^d)$,

$$\int_{X\times X\times\mathbb{R}^d} F(x', x, q)\mathrm{d}\gamma^h(x', x, q) := \int_{X\times X} F(x', x, q^h(x))\mathrm{d}\lambda^h(x')\mathrm{d}\lambda^h(x).$$

The family of measures γ^h is tight because its first and second marginals are λ^h which satisfies (6.3) and its third marginals have bounded second moments by (6.6). Up to a subsequence, we may therefore assume that γ^h narrowly converges to some measure γ. Note that γ is necessarily of the form $\lambda \otimes \theta$ where θ has first marginal λ hence can be disintegrated as $\theta(\mathrm{d}x, \mathrm{d}q) = \lambda(\mathrm{d}x)\theta^x(\mathrm{d}q)$, so that

$$\lim_{h\to 0} \int_{X\times X} F(x', x, q^h(x))\mathrm{d}\lambda^h(x')\mathrm{d}\lambda^h(x) = \int_{X\times X\times} \Big(\int_{\mathbb{R}^d} F(x', x, q)\mathrm{d}\theta^x(q)\Big)\mathrm{d}x'\mathrm{d}x$$

for every $F \in C_b(X \times X \times \mathbb{R}^d)$. Since $(u^h, q^h) \in K_b^h$, we have $u^h(x') - u^h(x) \geq \Gamma_b(x', x, q)$ γ^h-a.e.. Let then $\psi \in C_c(X \times X \times \mathbb{R}^d)$ with $\psi \geq 0$; since $(x', x, q) \mapsto \psi(x', x, q)(u^h(x') - u^h(x) - \Gamma_b(x', x, q))$ converges uniformly to $\psi(x', x, q)(u(x') - u(x) - \Gamma_b(x', x, q))$ and γ^h converges narrowly to γ, we have

$$\int_{X\times X\times\mathbb{R}^d} \psi(x', x, q)(u(x') - u(x) - \Gamma_b(x', x, q))\mathrm{d}\gamma(x', x, q) \geq 0$$

so that the continuous function $u(x')-u(x)-\Gamma_b(x',x,q)$ is nonnegative for every $x' \in X$, a.e. $x \in X$ and θ^x-a.e. q. If x is a point of differentiability of u, this implies in particular that $q = \nabla u(x)$, so θ^x necessarily coincides with the Dirac mass at $\nabla u(x)$. We then have $\gamma = \lambda \otimes ((\mathrm{id}, \nabla u)_{\#}\lambda)$ and the whole family γ^h converges narrowly to γ as $h \to 0$, hence

$$\lim_{h\to 0} \frac{\lambda(X)^2}{N_h^2} \sum_{j=1}^{N_h} \sum_{i=1}^{N_h} F(x_j^h, x_i^h, q_i^h) = \int_{X\times X} F(x', x, \nabla u(x)) \mathrm{d}x' \mathrm{d}x$$

for every $F \in C_b(X \times X \times \mathbb{R}^d)$, which in particular proves that (6.7) holds. Fixing $M > 0$, we have by (6.7) and (6.6)

$$\int_X \frac{\alpha(x)}{2} \min(M, |\nabla u(x)-q_0(x)|^2) \mathrm{d}x \le \liminf_h \int_X \frac{\alpha(x)}{2} |q^h(x)-q_0(x)|^2 \mathrm{d}\lambda^h(x) \le C \quad (6.9)$$

for some C independent of h. Letting M go to ∞, this shows that $q = \nabla u \in L^2$ and thus $(u,q) \in K_b$ (the b-convexity inequality has already been established). In a similar way, taking a family of cutoff functions $\psi_k \in C_c(X)$, $0 \le \psi_k \le 1$ with $\psi_k = 1$ on $\{x \in X \ : \ \mathrm{dist}(x, \partial X) \ge 1/k\}$ and $\psi_k \le \psi_{k+1}$, using the uniform convergence of u^h to u on compact subsets of X, we get

$$\int_X \psi_k(x) \frac{\beta(x)}{2} (u(x) - u_0(x))^2 \mathrm{d}x \le \liminf_h \int_X \frac{\beta(x)}{2} (u^h(x) - u_0(x))^2 \mathrm{d}\lambda^h(x). \quad (6.10)$$

Passing to the supremum in M in (6.9) and in k in (6.10) finally establishes the Γ-liminf inequality (6.8):

$$\begin{aligned}\liminf_h J^h(u^h, q^h) \ge \liminf_h \int_X \frac{\alpha(x)}{2} |q^h(x) - q_0(x)|^2 \mathrm{d}\lambda^h(x) \\ + \liminf_h \int_X \frac{\beta(x)}{2} (u^h(x) - u_0(x))^2 \mathrm{d}\lambda^h(x) \ge J(u,q).\end{aligned}$$

□

Remark 6.4. If in Lemma 6.3 one further assume that u^h is uniformly Lipschitz, i.e.

$$\sup_h \|q^h\|_{L^\infty} < +\infty,$$

then the measures γ^h in the proof are supported by a fixed compact set and thus one can use a quadratic-test function in (6.7) which actually gives

$$\int_X \frac{\alpha(x)}{2} |\nabla u(x) - q_0(x)|^2 \mathrm{d}x = \lim_h \int_X \frac{\alpha(x)}{2} |q^h(x) - q_0(x)|^2 \mathrm{d}\lambda^h(x).$$

In a similar way, in this case u^h converges uniformly to u in X and we also have

$$\int_X \frac{\beta(x)}{2} (u(x) - u_0(x))^2 \mathrm{d}x = \lim_h \int_X \frac{\beta(x)}{2} (u^h(x) - u_0(x))^2 \mathrm{d}\lambda^h(x).$$

Hence with a uniform bound on q^h, the last statement in Lemma 6.3 can be streghtened to:

$$\lim_h J^h(u^h, q^h) = J(u,q).$$

Theorem 6.5. *Assume that b satisfies **(B1)** with the extra assumption that g is Lipschitz on $\mathbb{R}^d$. Let $(\overline{u}^h, \overline{q}^h)$ be the solution of (6.4)-(6.5), extended as in Lemma 6.2, and let $(\overline{u}, \overline{q})$ be the solution of (5.1)-(5.2). Then $\overline{u}^h$ converges locally uniformly to $\overline{u}$ and $\overline{q}^h$ converges in L^2_{loc} and a.e. to $\overline{q}$ as $h \to 0$.*

Proof. Since $(\overline{u}^h, \overline{q}^h)$ solves (6.4)-(6.5) it is easy to see that there is a constant C such that, for all h,

$$\|\overline{q}^h\|_{L^2(\lambda^h)} + \|\overline{u}^h\|_{L^2(\lambda^h)} \leq C. \tag{6.11}$$

We first claim that for every ω convex such that $\omega \subset\subset X$,

$$\sup_h \|\nabla \overline{u}^h\|_{L^\infty(\omega)} = C_\omega < +\infty.$$

Indeed, if it was not the case, using the fact that the all functions $\overline{u}^h$ are convex and arguing as in the proof of Theorem 1 in [14], this would imply that, up to a subsequence, the norm of all elements of $\partial \overline{u}^h$ converge to $+\infty$ uniformly on some open subset of X, which would also imply a uniform explosion for $\overline{q}^h$ since $\overline{q}^h(x) \in \partial \overline{u}^h(x)$, and thus this would contradict (6.11) and (6.3).

Passing to a subsequence if necessary we may therefore assume that for some convex function u, one has $\overline{u}^h \to u$ uniformly on compact subsets of X, $\overline{q}^h \to q = \nabla u$ a.e. and in L^2_{loc}. By Lemma 6.3, $(u, q) \in K_b$ and $J(u, q) \leq \liminf_h J^h(\overline{u}^h, \overline{q}^h)$. To prove that $(u, q) = (\overline{u}, \overline{q})$ (and then that the whole family converges, not only a subsequence), it is enough by Lemma 6.1 to show that $J(u, q) \leq J(v, p)$ for every $(v, p) \in K_b \cap (W^{1,\infty}(X) \times L^\infty(X))$. Let $(v, p) \in K_b \cap W^{1,\infty}(X) \times L^\infty(X)$, then extend as in Lemma 6.2 the discrete data $(v(x_i^h), p(x_i^h))_{i=1,\dots,N_h}$ and denote by (v^h, p^h) this extension. It is easy to check that p^h is uniformly bounded and v^h converges uniformly to v, then using Remark 6.4 and the optimality of $(\overline{u}^h, \overline{q}^h)$ we get

$$J(u, q) \leq \liminf_h J^h(\overline{u}^h, \overline{q}^h) \leq \liminf_h J^h(v^h, p^h) = J(v, p),$$

which gives the desired result.

□

7 Numerical method

7.1 Dykstra's iterative projection algorithm

We focus on the discretized problem (6.4)-(6.5), where for notational simplicity we drop the discretization index h and set $\Gamma_{ij}(q) := \Gamma_b(x_i, x_j, q)$:

$$\inf_{(u,q)} \sum_{k=1}^{N} \left[\frac{\alpha_k}{2} \left| q_k - q_k^0 \right|^2 + \frac{\beta_k}{2} \left| u_k - u_k^0 \right|^2 \right]$$
$$\text{subject to} \quad u_i - u_j \geq \Gamma_{ij}(q_j), \ \forall (i,j)\{1, \dots, N\}^2.$$

Defining the convex subsets of $\mathbb{R}^N \times \mathbb{R}^{dN}$

$$C_{i,j} := \left\{ (u, q) : \ u_i - u_j \geq \Gamma_{ij}(q_j) \right\}, \ \forall (i,j)\{1, \dots, N\}^2,$$

and the weighted squared Euclidean distance in $\mathbb{R}^N \times \mathbb{R}^{dN}$

$$D_{\alpha,\beta}\left((u,q),(u',q')\right) := \sum_{k=1}^{N} \left[\frac{\alpha_k}{2}\left|q_k - q_k'\right|^2 + \frac{\beta_k}{2}\left|u_k - u_k'\right|^2\right],$$

the discrete problem is to find the projection $P_C^{\alpha,\beta}(u^0,q^0)$, solution in $\mathbb{R}^N \times \mathbb{R}^{dN}$ to

$$\inf_{(u,q)} D_{\alpha,\beta}\left((u,q),(u^0,q^0)\right) \quad \text{subject to} \quad (u,q) \in C := \bigcap_{(i,j)} C_{i,j}.$$

This projection problem can be solved iteratively by Dykstra's algorithm [3, 5]. Let us now recall it for a projection problem onto a closed convex subset K of $\mathbb{R}^m$ which can be written as an intersection of *elementary* closed convex subsets $K_1, \dots, K_L$ of $\mathbb{R}^m$:

$$K := \bigcap_{l=1}^{L} K_l.$$

Given z^0, we wish to compute $P_K^{\alpha,\beta}(z^0) = \operatorname{argmin}_{z\in K} D_{\alpha,\beta}(z,z^0)$. First extend by L-periodicity the familly of convex sets $(K_l)_{l=1,\cdots,L}$ by setting

$$K_{l+kL} := K_l,\ \forall (k,l) \in \mathbb{N} \times \{1,\dots,L\}.$$

Then initialize the algorithm by setting

$$\theta^{-L+1} = \cdots = \theta^{-1} = \theta^0 = 0$$

and starting from the point z^0 we want to project, update z^n and θ^n for $n \geq 1$ by:

$$z^n = P_{K_n}^{\alpha,\beta}\Big(z^{n-1} + \theta^{n-L}\Big),\ \theta^n = z^{n-1} - z^n + \theta^{n-L}.$$

The fact that the sequence z^n converges to $P_K^{\alpha,\beta}(z^0)$ has been established by Boyle and Dyskstra [5]; Bauschke and Lewis [3] have extended this result to projection problems with general Bregman distances.

Remark 7.1. For the sake of simplicity, we have omitted above the case of pointwise convex constraints on u or q such as $u_i \geq 0$, $u_i \leq v_i$, $|q_i| \leq r_i$, or $q_i \in \mathbb{R}^d_+$. Such constraints can be handled easily with Dykstra's algorithm by adding elementary convex sets on which the projections are totally explicit.

Remark 7.2. The main drawback of Dystra's algorithm in our context is the a priori very large number N^2 of elementary convex sets on which we have to project. One can however significantly reduce the computational cost if $C = \cap_{(i,j)\in I} C_{i,j}$ where the cardinality of $I \subset \{1,\dots,N\}^2$ is smaller than N^2. This reduction of the number of constraints can be done when the b-convexity constraint *propagates*. This is in particular the case in dimension 1 under the so-called Spence-Mirrlees condition $\partial^2_{xy} b > 0$ and for ordered grid points ($x_i < x_j$ for $i < j$), where it is enough to impose the b-convexity on neighbour points ($j = i \pm 1$), resulting in a number of elementary convex sets $O(N)$ instead of $O(N^2)$. It is indeed easy to see that the constraint propagates in this case: set $y_i := y_b(x_i, q_i)$ and first remark that the two "local" constraints $u_{i+1} - u_i \geq$

$b(x_{i+1}, y_i) - b(x_i, y_i)$ and $u_i - u_{i+1} \geq b(x_i, y_{i+1}) - b(x_{i+1}, y_{i+1})$ together with Spence-Mirrlees condition imply that $y_{i+1} \geq y_i$; then summing the two consecutive constraints $u_{i+1} - u_i \geq b(x_{i+1}, y_i) - b(x_i, y_i)$ and $u_{i+2} - u_{i+1} \geq b(x_{i+2}, y_{i+1}) - b(x_{i+1}, y_{i+1})$ and using the fact that $b(x_{i+2}, y_{i+1}) - b(x_{i+1}, y_{i+1}) \geq b(x_{i+2}, y_i) - b(x_{i+1}, y_i)$ again by Spence-Mirrlees condition and the fact that $y_{i+1} \geq y_i$, we obtain that $u_{i+2} - u_i \geq b(x_{i+2}, y_i) - b(x_i, y_i)$, which is exactly what we mean by propagation of the constraint.

Remark 7.3. In a similar way, for the convexity constraint $b(x, y) = x \cdot y$, $\Gamma_b(x', x, q) = (x' - x) \cdot q$ and for regular grid points, the number of convexity constraints $u_i - u_j \geq (x_i - x_j) \cdot q_j$ can be significantly reduced (for instance for aligned points, it is enough to consider consecutive points x_i and x_j). We refer to [14] and the recent efficient approach of Mirebeau [26] for details.

7.2 Elementary projections

7.2.1 Lagrangian relaxation

In Dyskstra's algorihm, we have to compute at each step an elementary projection onto a single convex $C_{i,j}$, $P^{\alpha,\beta}_{C_{i,j}}(\hat{q}, \hat{u}) := P_{i,j}(\hat{q}, \hat{u})$ for various $(\hat{q}, \hat{u})$. The elementary projection $(\overline{q}, \overline{u}) = P_{i,j}(\hat{q}, \hat{u})$ is obviously given by $\overline{q}_k = \hat{q}_k$ for $k \neq j$ and $\overline{u}_k = \hat{u}_k$ for $k \neq i, j$ with $(\overline{q}_j, \overline{u}_i, \overline{u}_j)$ obtained by solving the low-dimensional projection problem

$$\inf_{(u_i, u_j, q_j)} \left[\frac{\alpha_j}{2} |q_j - \hat{q}_j|^2 + \frac{\beta_i}{2} |u_i - \hat{u}_i|^2 + \frac{\beta_j}{2} |u_j - \hat{u}_j|^2 \right]$$
$$\text{subject to} \quad u_i - u_j \geq \Gamma_{ij}(q_j).$$

If $\hat{u}_i - \hat{u}_j \geq \Gamma_{ij}(\hat{q}_j)$, then this projection problem is trivial and $(\overline{u}_i, \overline{u}_j, \overline{q}_j) = (\hat{u}_i, \hat{u}_j, \hat{q}_j)$, so only the case $\hat{u}_i - \hat{u}_j < \Gamma_{ij}(\hat{q}_j)$ where the constraint is binding requires some specific attention and can conveniently be addressed by Lagrangian relaxation as follows. Since the problem above is convex and qualified (the constraint is linear in u_i and u_j), there is no duality gap with its dual problem, which is

$$\sup_{\lambda \geq 0} l_{ij}(\lambda)$$

where $l_{ij}(\lambda)$ is defined, for $\lambda \geq 0$, by

$$\inf_{(u_i, u_j, q_j)} \left[\frac{\alpha_j}{2} |q_j - \hat{q}_j|^2 + \frac{\beta_i}{2} |u_i - \hat{u}_i|^2 + \frac{\beta_j}{2} |u_j - \hat{u}_j|^2 + \lambda \left(-u_i + u_j + \Gamma_{ij}(q_j) \right) \right]. \quad (7.1)$$

The dual function l_{ij} is concave – as an infimum of affine functions – and differentiable – since the Lagrange problem (7.1) has a unique solution for any $\lambda \geq 0$, that we denote by $(q_j(\lambda), u_i(\lambda), u_j(\lambda))$ – with $l'_{ij}(\lambda) = -u_i(\lambda) + u_j(\lambda) + \Gamma_{ij}(q_j(\lambda))$. Note that $u_i(\lambda), u_j(\lambda)$ are given explicitly by

$$u_i(\lambda) = \hat{u}_i + \frac{\lambda}{\beta_i}, \quad u_j(\lambda) = \hat{u}_j - \frac{\lambda}{\beta_j}, \quad (7.2)$$

whereas $q_j(\lambda)$ is the unique solution (not explicit in general) to

$$\alpha_j(q_j - \hat{q}_j) + \lambda \nabla \Gamma_{ij}(q_j) = 0, \text{ i.e. } q_j(\lambda) = \left(\text{id} + \frac{\lambda}{\alpha_j} \nabla \Gamma_{ij} \right)^{-1} (\hat{q}_j). \quad (7.3)$$

The solution of the projection problem $(\overline{u}_i, \overline{u}_j, \overline{q}_j)$ in the case where $\hat{u}_i - \hat{u}_j < \Gamma_{ij}(\hat{q}_j)$ is then $(u_i(\lambda), u_j(\lambda), q_j(\lambda))$ where $\lambda > 0$ solves

$$l'_{ij}(\lambda) = -u_i(\lambda) + u_j(\lambda) + \Gamma_{ij}(q_j(\lambda) = 0. \tag{7.4}$$

Note that (7.4) is monotone with respect to λ since l_{ij} is concave.

In practice, we have experimented two numerical methods to solve the elementary projection problem in the case $\lambda > 0$. The first one is a duality method: it consists in solving the dual problem via (7.4):

1. Given $\lambda > 0$, compute $q_j(\lambda)$ from (7.3) (by Newton's method if necessary) and $u_i(\lambda), u_j(\lambda)$ from (7.2) (explicitly).
2. Evaluate $l'_{ij}(\lambda) = -u_i(\lambda) + u_j(\lambda) + \Gamma_{ij}(q_j(\lambda))$ and stop if λ is satisfying (7.4) up to a fixed tolerance.
3. Update $\lambda > 0$ by one step of a dichotomy method to solve $l'_{ij}(\lambda) = 0$.

The second method that we have experimented (and actually used for the numerical results in section 8) is a primal-dual method: it consists in finding simultaneously the primal-dual solution $(q_j(\lambda), u_i(\lambda), u_j(\lambda), \lambda)$ to the elementary projection problem via the optimality conditions for $\lambda > 0$. This amounts to computing simultaneously $(q_j(\lambda), \lambda)$ as the unique solution to

$$\begin{cases} \alpha_j\,(q_j - \hat{q}_j) + \lambda \nabla \Gamma_{ij}(q_j) = 0 \\ -\,(\hat{u}_i + \dfrac{\lambda}{\beta_i}) + (\hat{u}_j - \dfrac{\lambda}{\beta_j}) + \Gamma_{ij}(q_j) = 0, \end{cases} \tag{7.5}$$

by Newton's method. See e.g. [4] for variants.

7.2.2 Special cases

In the convexity constraint case, $\Gamma_{ij}(q) = (x_i - x_j) \cdot q$, the solution $q_j(\lambda)$ to equation (7.3) is explicitly given by

$$q_j(\lambda) = \hat{q}_j - \frac{\lambda}{\alpha_j}(x_i - x_j).$$

Plugging it, together with $u_i(\lambda), u_j(\lambda)$ given by (7.2), into the optimality equation (7.4), we get that the solution to the dual problem when $\lambda > 0$ (i.e. when $(\hat{q}_j, \hat{u}_i, \hat{u}_j)$ is not feasible for the primal problem) is

$$\lambda = \frac{-\hat{u}_i + \hat{u}_j + \Gamma_{ij}(\hat{q}_j)}{\frac{1}{\beta_i} + \frac{1}{\beta_j} + \frac{|x_i - x_j|^2}{\alpha_j}}.$$

Therefore the elementary projections are explicit in this case.

In the b-convexity constraint case with $b(x, y) := x \cdot y + f(x)g(y)$, recall that

$$\Gamma_{ij}(q) = (x_i - x_j) \cdot q + D_f(x_i, x_j) g_b(x_j, q).$$

When g_b is known explicitly (as it is in the special cases described in 4.2) and twice differentiable, the elementary projection problems can be solved as explained above. In practice, we solve (7.5) by Newton's method initialized with the explicit solution (q_j, λ) to the convexity constraint case. In the more general case where g_b is not known explicitly, one could imagine to solve simultaneously (4.3) to evaluate it numerically, but we have not explored this way.

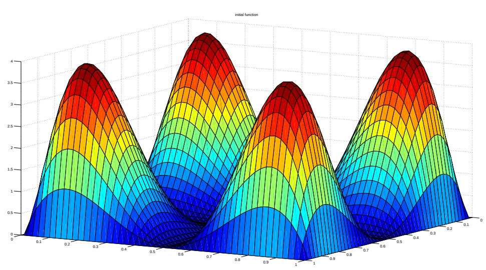
8 Numerical results

We present here various examples of projection problems of the form (5.1)-(5.2) (sometimes on convex subsets of C corresponding to additional pointwise convex constraints) and their numerical resolution by the method described in sections 6 and 7. Since the main limitation to this method (up to the numerical resolution of (4.3) to get g_b) is the fact that Γ_b has to be explicit (very likely y_b too), we have restricted ourselves to two cases for b:

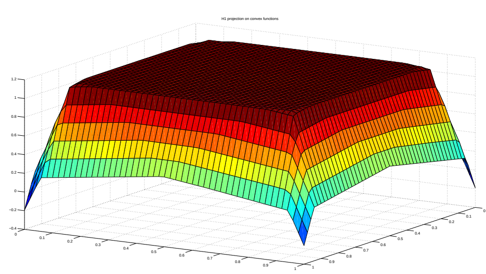
- the bilinear case $b(x, y) := x \cdot y$, which corresponds to the convexity constraint;
- a perturbation of it given by $b(x, y) := x \cdot y + f(x)g(y)$ with $f(x) := \sqrt{1 + |x|^2}$, $g(y) := \sqrt{1 + |y|^2}$.

In 3d figures, the z-axis is directed downward.

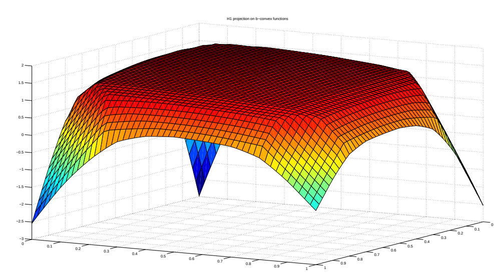
8.1 H^1 projection on b-convex functions



Initial function



Projection on convex functions



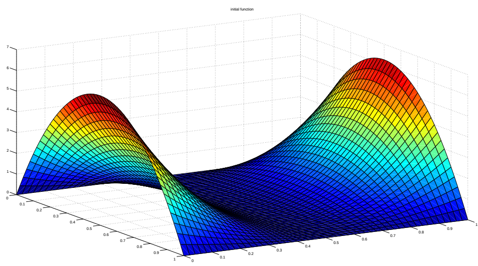
Projection on b-convex functions

Figure 1: $H^1((0, 1)^2)$ projections of $-x(1 - x)(2x - 1)^2 y(1 - y)(2y - 1)^2$

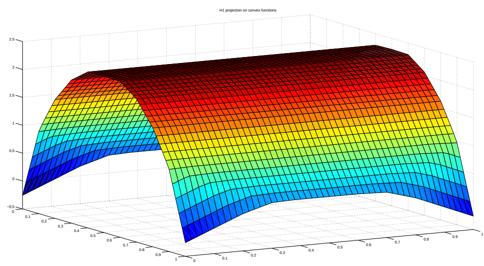
The first example is precisely the projection problem (5.1)-(5.2) with $\alpha = \beta = 1$ and $q_0 = \nabla u_0$. We have solved numerically such problems in dimension $d = 2$, see Figure 1 and 2 for the $H^1((0, 1)^2)$ projection of two different u_0.

8.2 b-convex envelopes in dimension one

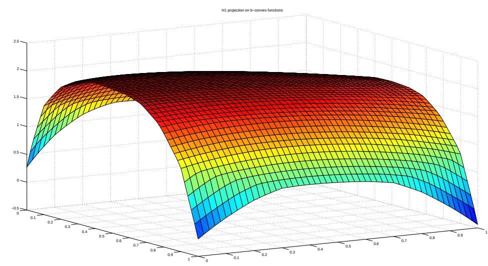
We next have computed convex and b-convex envelopes in dimension $d = 1$ using the formulation (5.6) (both for convex and b-convex envelopes) with $\alpha = \beta = 1$, $q_0 = 0$ and a nonconvex u_0 depicted in Figure 3.



Initial function

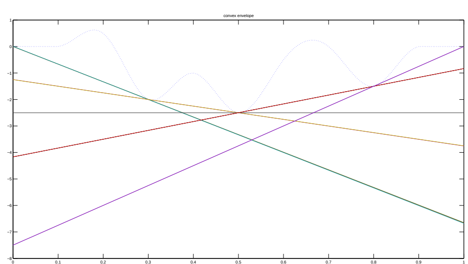


Projection on convex functions

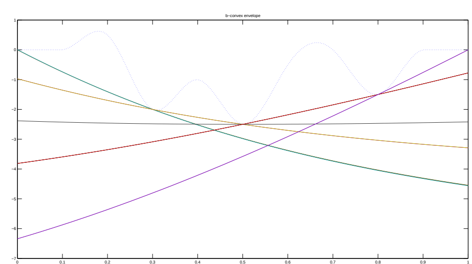


Projection on b-convex functions

Figure 2: $H^1((0,1)^2)$ projections of $-x(1-x)(y-\frac{1}{2})$



Convex envelope



b-convex envelope

Figure 3: Envelopes of a function in dimension 1

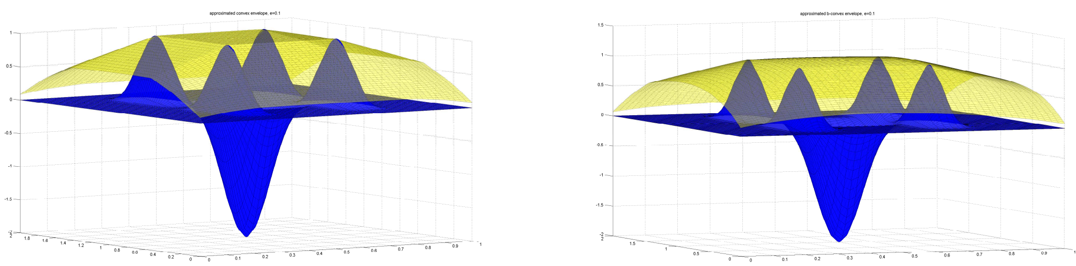
8.3 Approximated b-convex envelopes in dimension two

We now present two-dimensional approximated convex/b-convex envelopes, using the regularized formulation (5.7) with different values of the regularization parameter $\varepsilon > 0$, respresented in Figure 4.

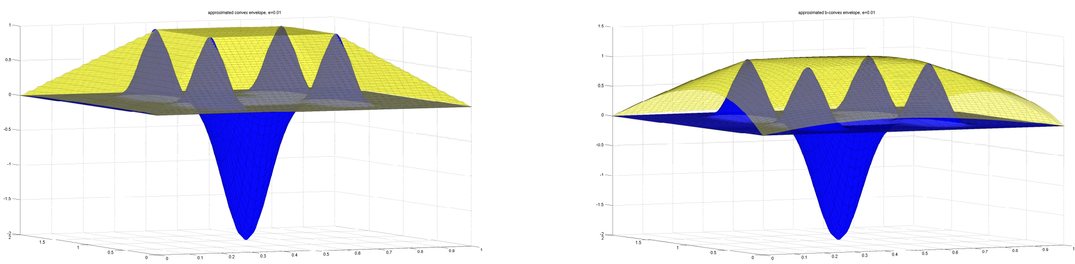
8.4 Regularized Rochet and Choné problem

We finally consider the Rochet and Choné prinicipal-agent problem on the square $[1,2]^2$, regularized as in (5.5) with $\varepsilon = 0.01$, see Figure 5.

Acknowledgements: The authors benefited from the hospitality of the Fields Institute (Toronto, Canada), where part of the present research was conducted during the Thematic Semester on Variational Problems in Physics, Economics and Geometry. They gratefully acknowledge support from the ANR, through the projects ISOTACE



Approximated convex envelope ($\varepsilon = 0.1$) Approximated b-convex envelope ($\varepsilon = 0.1$)



Approximated convex envelope ($\varepsilon = 0.01$) Approximated b-convex envelope ($\varepsilon = 0.01$)

Figure 4: Approximated envelopes of a function in dimension 2

(ANR-12-MONU-0013), OPTIFORM (ANR-12-BS01-0007) and from INRIA through the “action exploratoire" MOKAPLAN and wish to thank J.-D. Benamou for stimulating discussions. G.C. gratefully acknowledges the hospitality of the Mathematics and Statistics Department at UVIC (Victoria, Canada) and support from the CNRS.

References

[1] N. E. Aguilera and P. Morin. Approximating optimization problems over convex functions. *Numer. Math.*, 111(1):1–34, 2008.

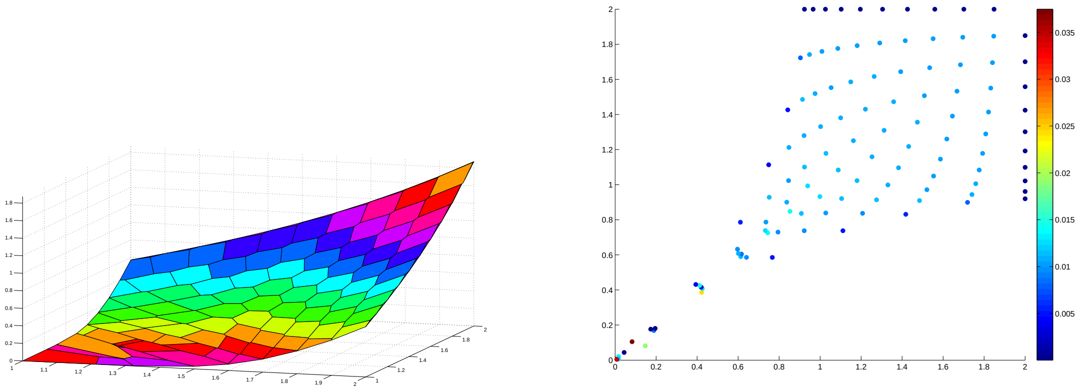
[2] H. H. Bauschke and P. L. Combettes. A Dykstra-like algorithm for two monotone operators. *Pac. J. Optim.*, 4(3):383–391, 2008.

[3] H. H. Bauschke and A. S. Lewis. Dykstra's algorithm with Bregman projections: a convergence proof. *Optimization*, 48(4):409–427, 2000.

[4] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. *Numerical optimization*. Universitext. Springer-Verlag, Berlin, second edition, 2006. Theoretical and practical aspects.

[5] J. P. Boyle and R. L. Dykstra. A method for finding projections onto the intersection of convex sets in Hilbert spaces. In *Advances in order restricted statistical inference (Iowa City, Iowa, 1985)*, volume 37 of *Lecture Notes in Statist.*, pages 28–47. Springer, Berlin, 1986.

[6] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.



Indirect utility of the agents u — Distribution of the types of products $q = \nabla u$

Figure 5: Rochet and Choné problem regularized with $\varepsilon = 0.01$

[7] F. Brock, V. Ferone, and B. Kawohl. A symmetry problem in the calculus of variations. *Calc. Var. Partial Differential Equations*, 4(6):593–599, 1996.

[8] G. Buttazzo, V. Ferone, and B. Kawohl. Minimum problems over sets of concave functions and related questions. *Math. Nachr.*, 173:71–89, 1995.

[9] G. Buttazzo and P. Guasoni. Shape optimization problems over classes of convex domains. *J. Convex Anal.*, 4(2):343–351, 1997.

[10] G. Carlier. A general existence result for the principal-agent problem with adverse selection. *J. Math. Econom.*, 35(1):129–150, 2001.

[11] G. Carlier. Calculus of variations with convexity constraint. *J. Nonlinear Convex Anal.*, 3(2):125–143, 2002.

[12] G. Carlier and A. Galichon. Exponential convergence for a convexifying equation. *ESAIM Control Optim. Calc. Var.*, 18(3):611–620, 2012.

[13] G. Carlier and T. Lachand-Robert. Regularity of solutions for some variational problems subject to a convexity constraint. *Comm. Pure Appl. Math.*, 54(5):583–594, 2001.

[14] G. Carlier, T. Lachand-Robert, and B. Maury. A numerical approach to variational problems subject to convexity constraint. *Numer. Math.*, 88(2):299–318, 2001.

[15] P. Choné and H. V. J. Le Meur. Non-convergence result for conformal approximation of variational problems subject to a convexity constraint. *Numer. Funct. Anal. Optim.*, 22(5-6):529–547, 2001.

[16] P. L. Combettes. Iterative construction of the resolvent of a sum of maximal monotone operators. *J. Convex Anal.*, 16(3-4):727–748, 2009.

[17] R. L. Dykstra. An algorithm for restricted least squares regression. *J. Amer. Statist. Assoc.*, 78(384):837–842, 1983.

[18] I. Ekeland and S. Moreno-Bromberg. An algorithm for computing solutions of variational problems with global convexity constraints. *Numer. Math.*, 115(1):45–69, 2010.

[19] A. Figalli, Y.-H. Kim, and R. J. McCann. When is multidimensional screening a convex program? *J. Econom. Theory*, 146(2):454–478, 2011.

[20] T. Lachand-Robert and M. A. Peletier. An example of non-convex minimization and an application to Newton's problem of the body of least resistance. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 18(2):179–198, 2001.

[21] J. Lamboley, A. Novruzi, and M. Pierre. Regularity and singularities of optimal convex shapes in the plane. *Arch. Ration. Mech. Anal.*, 205(1):311–343, 2012.

[22] P.-L. Lions. Identification du cône dual des fonctions convexes et applications. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 326(12):1385–1390, 1998.

[23] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.

[24] X.-N. Ma, N. S. Trudinger, and X.-J. Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[25] Q. Mérigot and É. Oudet. Handling convexity-like constraints in variational problems. *SIAM J. Numer. Anal.*, 52(5):2466–2487, 2014.

[26] J.-M. Mirebeau. Adaptive, anisotropic and hierarchical cones of discrete convex functions. *arXiv preprint arXiv:1402.1561, to appear in Num. Math.*, 2014.

[27] A. M. Oberman. The convex envelope is the solution of a nonlinear obstacle problem. *Proc. Amer. Math. Soc.*, 135(6):1689–1694 (electronic), 2007.

[28] A. M. Oberman. Computing the convex envelope using a nonlinear partial differential equation. *Math. Models Methods Appl. Sci.*, 18(5):759–780, 2008.

[29] A. M. Oberman. A numerical method for variational problems with convexity constraints. *SIAM J. Sci. Comput.*, 35(1):A378–A396, 2013.

[30] J.-C. Rochet and P. Choné. Ironing, sweeping, and multidimensional screening. *Econometrica*, pages 783–826, 1998.

[31] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[32] L. Vese. A method to convexify functions via curve evolution. *Comm. Partial Differential Equations*, 24(9-10):1573–1591, 1999.