# Adaptive, Anisotropic and Hierarchical cones of Discrete Convex functions

Jean-Marie Mirebeau

▶ **To cite this version:**

Jean-Marie Mirebeau. Adaptive, Anisotropic and Hierarchical cones of Discrete Convex functions. Numerische Mathematik, 2016, 132 (4), pp.807-853. hal-00943096v2

**HAL Id: hal-00943096**
**https://hal.science/hal-00943096v2**

Submitted on 23 Sep 2014

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

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

# Adaptive, Anisotropic and Hierarchical cones of Discrete Convex functions*

Jean-Marie Mirebeau†

September 23, 2014

**Abstract**

We introduce a new class of adaptive methods for optimization problems posed on the cone of convex functions. Among the various mathematical problems which posses such a formulation, the Monopolist problem [24, 10] arising in economics is our main motivation.

Consider a two dimensional domain $\Omega$, sampled on a grid $X$ of $N$ points. We show that the cone $\mathrm{Conv}(X)$ of restrictions to $X$ of convex functions on $\Omega$ is typically characterized by $\approx N^2$ linear inequalities; a direct computational use of this description therefore has a prohibitive complexity. We thus introduce a hierarchy of sub-cones $\mathrm{Conv}(\mathcal{V})$ of $\mathrm{Conv}(X)$, associated to stencils $\mathcal{V}$ which can be adaptively, locally, and anisotropically refined. We show, using the arithmetic structure of the grid, that the trace $U_{|X}$ of any convex function $U$ on $\Omega$ is contained in a cone $\mathrm{Conv}(\mathcal{V})$ defined by only $\mathcal{O}(N \ln^2 N)$ linear constraints, in average over grid orientations.

Numerical experiments for the Monopolist problem, based on adaptive stencil refinement strategies, show that the proposed method offers an unrivaled accuracy/complexity trade-off in comparison with existing methods. We also obtain, as a side product of our theory, a new average complexity result on edge flipping based mesh generation.

A number of mathematical problems can be formulated as the optimization of a convex functional over the *cone of convex functions* on a domain $\Omega$ (here compact and two dimensional):

$$\mathrm{Conv}(\Omega) := \{U : \Omega \to \mathbb{R};\ U \text{ is convex}\}.$$

This includes optimal transport, as well as various geometrical conjectures such as Newton's problem [16, 18]. We choose for concreteness to emphasize an economic application: the Monopolist (or Principal Agent) problem [24], in which the objective is to design an optimal product line, and an optimal pricing catalog, so as to maximize profit in a captive market. The following minimal instance is numerically studied in [1, 10, 21] and on Figure 1. With $\Omega = [1,2]^2$

$$\min\left\{\int_\Omega \left(\frac{1}{2}\|\nabla U(z)\|^2 - \langle \nabla U(z), z\rangle + U(z)\right) dz;\ U \in \mathrm{Conv}(\Omega),\ U \geq 0\right\}. \tag{1}$$

We refer to the numerical section §6, and to [24] for the economic model details; let us only say here that the Monopolist's optimal product line is $\{\nabla U(z);\ z \in \Omega\}$, and that the optimal prices are given by the Legendre-Fenchel dual of $U$. Consider the following three regions, defined for $k \in \{0,1,2\}$ (implicitly excluding points $z \in \Omega$ close to which $U$ is not smooth)

$$\Omega_k := \{z \in \Omega;\ \mathrm{rank}(\mathrm{Hessian}\, U(z)) = k\}. \tag{2}$$

*This work was partly supported by ANR grant NS-LBR ANR-13-JS01-0003-01.

†CNRS, University Paris Dauphine, UMR 7534, Laboratory CEREMADE, Paris, France.

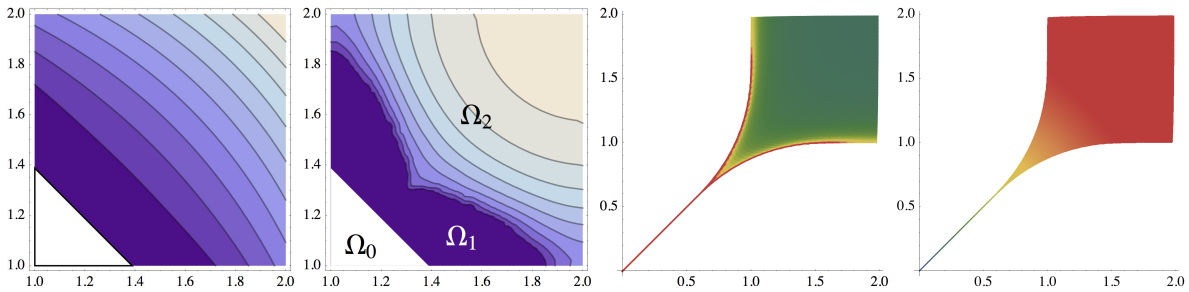



Figure 1: Numerical approximation $U$ of the solution of the classical Monopolist's problem (1), computed on a $50 \times 50$ grid. Left: level sets of $U$, with $U = 0$ in white. Center left: level sets of $\det(\text{Hessian}\, U)$ (with again $U = 0$ in white); note the degenerate region $\Omega_1$ where $\det(\text{Hessian}\, U) = 0$. Center right: distribution of products sold by the monopolist. Right: profit margin of the monopolist for each type of product (margins are low on the one dimensional part of the product line, at the bottom left). Color scales on Figure 11.

Strong empirical evidence suggests that these three regions have a non-empty interior, although no qualitative mathematical theory has yet been developed for these problems. The optimal product line observed numerically, Figure 1, confirms a qualitative (and conjectural) prediction of the economic model [24] called "bunching": low-end products are less diverse than high-end ones, down to the topological sense. (The monopolist willingly limits the variety of cheap products, because they may compete with the more expensive ones, on which he has a higher margin.)

We aim to address numerically optimization problems posed on the cone of convex functions, through numerical schemes which preserve the rich qualitative properties of their solutions, and have a moderate computational cost. In order to put in light the specificity of our approach, we review the existing numerical methods for these problems, which fall in the following categories. We denote by $X$ a grid sampling of the domain $\Omega$, and by $\text{Conv}(X)$ the cone of discrete (restrictions of) convex functions

$$\text{Conv}(X) := \{U_{|X};\, U \in \text{Conv}(\Omega)\}. \tag{3}$$

- (Interior finite element methods) For any triangulation $\mathcal{T}$ of $X$, consider the cone

$$\text{Conv}(\mathcal{T}) := \{u : X \to \mathbb{R};\, \mathrm{I}_{\mathcal{T}}\, u \in \text{Conv}(\Omega)\}.$$

  A natural but *invalid* numerical method for (1) is to fix a-priori a family $(\mathcal{T}_h)_{h>0}$ of regular triangulations of $\Omega$, where $h > 0$ denotes mesh scale, and to optimize the functional of interest over the associated cones. Indeed, the union of the cones $\text{Conv}(\mathcal{T}_h)$ is *not* dense in $\text{Conv}(\Omega)$, see [7]. Let us also mention that for a given generic $u \in \text{Conv}(X)$, there exists *only one* triangulation $\mathcal{T}$ of $X$ such that $u \in \text{Conv}(\mathcal{T})$, see §1.3.

- (Global constraints methods) The functional of interest, suitably discretized, is minimized over the cone $\text{Conv}(X)$ of discrete convex functions [5], or alternatively [10] on the augmented cone

$$\text{GradConv}(X) := \{(U_{|X}, \nabla U_{|X});\, U \in \text{Conv}(\Omega)\}, \tag{4}$$

  in which we refer by $\nabla U$ to arbitrary elements of the subgradient of the convex map $U$.

  Both $\text{Conv}(X)$ and $\text{GradConv}(X)$ are characterized by a family of long range linear inequalities, with domain wide supports, and of cardinality growing *quadratically* with $N := \#(X)$,

see §1.1 and [10]. Despite rather general convergence results, these two methods are impractical due to their expensive numerical cost, in terms of both computation time and memory.

- (Local constraints methods) Another cone $\mathrm{Conv}'(X)$ is introduced, usually satisfying neither $\mathrm{Conv}(X) \subseteq \mathrm{Conv}'(X)$ nor $\mathrm{Conv}'(X) \subseteq \mathrm{Conv}(X)$, but typically characterized by relatively few constraints, with short range supports. Obermann et al. [22, 21] use $\mathcal{O}(N)$ linear constraints, with $N := \#(X)$. Merigot et al. [18] use slightly more linear constraints, but provide an efficient optimization algorithm based on proximal operators. Aguilera et al. [1] consider $\mathcal{O}(N)$ constraints of semi-definite type.

  Some of these methods benefit from convergence guarantees [18, 1] as $N \to \infty$. Our numerical experiments with [1, 22, 21] show however that they suffer from accuracy issues, see §6.3 and Figure 10, which limits the usability of their results.

- (Geometric methods) A polygonal convex set can be described as the convex hull of a finite set of points, or as an intersection of half-spaces. Geometric methods approximate a convex function $U$ by representing its epigraph $\{(z,t);\, z \in \Omega,\, t \geq U(z)\}$ under one of these forms. Energy minimization is done by adjusting the points position, or the coefficients of the affine forms defining the half-spaces, see [26, 16].

  The main drawback of these methods lies in the optimization procedure, which is quite non-standard. Indeed the discretized functional is generally non-convex, and the polygonal structure of the represented convex set changes topology during the optimization.

We propose an implementation of the constraint of convexity via a limited (typically quasi-linear) number of linear inequalities, featuring both short range and domain wide supports, which are selected locally and anisotropically in an adaptation loop using a-posteriori analysis of solutions to intermediate problems. Our approach combines the accuracy of global constraint methods, with the limited cost of local constraint ones, see §6.3. It is based on a family of sub-cones

$$\mathrm{Conv}(\mathcal{V}) \subseteq \mathrm{Conv}(X),$$

each defined by some linear inequalities associated to a family $\mathcal{V}$ of *stencils*, see Definition 1.7. These stencils are the data $\mathcal{V} = (\mathcal{V}(x))_{x \in X}$ of a collection of offsets $e \in \mathcal{V}(x)$ pointing to selected neighbors $x + e$ of any point $x \in X$, and satisfying minor structure requirements, see Definition 1.6. The cones satisfy the hierarchy property $\mathrm{Conv}(\mathcal{V} \cap \mathcal{V}') = \mathrm{Conv}(\mathcal{V}) \cap \mathrm{Conv}(\mathcal{V}')$, see Theorem 1.8. Most elements of $\mathrm{Conv}(X)$ belong to a cone $\mathrm{Conv}(\mathcal{V})$ defined by only $\mathcal{O}(N \ln^2 N)$ linear inequalities, in a sense made precise by Theorem 1.11. Regarding both stencils and triangulations as directed graphs on $X$, we show in Theorem 1.13 (under a minor technical condition) that the cone $\mathrm{Conv}(\mathcal{V})$ is the union of the cones $\mathrm{Conv}(\mathcal{T})$ associated to triangulations $\mathcal{T}$ included in $\mathcal{V}$. Our hierarchy of cones has similarities, but also striking differences as discussed in conclusion, with the other multiscale constructions (wavelets, adaptive finite elements) used in numerical analysis.

The minimizer $u \in \mathrm{Conv}(X)$ of a given convex energy $\mathcal{E}$ can be obtained without ever listing the inequalities defining $\mathrm{Conv}(X)$ (which would often not fit into computer memory for the problem sizes of interest), but only solving a small sequence of optimization problems over sub-cones $\mathrm{Conv}(\mathcal{V}_i)$ associated to stencils $\mathcal{V}_i$, designed through adaptive refinement strategies. Our numerical experiments give, we believe, unprecedented numerical insight on the qualitative behavior of the monopolist problem and its variants. Thanks to the adaptivity of our scheme, this accuracy is not at the expense of computation time or memory usage. See §6.

# 1 Main results

The constructions and results developed in this paper apply to an arbitrary convex and compact domain $\Omega \subseteq \mathbb{R}^2$, discretized on an orthogonal grid of the form:

$$\Omega \cap hR_\theta(\xi + \mathbb{Z}^2), \tag{5}$$

where $h > 0$ is a scale parameter, $R_\theta$ is the rotation of angle $\theta \in \mathbb{R}$, and $\xi \in \mathbb{R}^2$ is an offset. The latter two parameters are used in our main approximation result Theorem 1.11, heuristically to eliminate by averaging the influence of rare unfavorable cases in which the approximated convex function hessian is degenerate in a direction close to the grid axes. For simplicity, and up to a linear change of coordinates, we assume unless otherwise mentionned that these parameters take their canonical values:

$$X := \Omega \cap \mathbb{Z}^2.$$

The choice of a grid discretization provides arithmetic tools that would not be available for an unstructured point set.

**Definition 1.1.** *1. An element $e = (\alpha, \beta) \in \mathbb{Z}^2$ is called irreducible iff $\gcd(\alpha, \beta) = 1$.*

*2. A basis of $\mathbb{Z}^2$ is a pair $(f, g) \in (\mathbb{Z}^2)^2$ such that $|\det(f, g)| = 1$. A basis $(f, g)$ of $\mathbb{Z}^2$ is direct iff $\det(f, g) = 1$, and acute iff $\langle f, g\rangle \geq 0$.*

Considering special (non-canonical) bases of $\mathbb{Z}^d$ is relevant when discretizing anisotropic partial differential equations on grids, such as anisotropic diffusion [11], or anisotropic eikonal equations [19]. In this paper, and in particular in the next proposition, we rely on a specific two dimensional structure called the Stern-Brocot tree [13], also used in numerical analysis for anisotropic diffusion [2], and eikonal equations of Finsler type [20].

**Proposition 1.2.** *The application $(f, g) \mapsto e := f + g$ defines a bijection between direct acute bases $(f, g)$ of $\mathbb{Z}^2$, and irreducible elements $e \in \mathbb{Z}^2$ such that $\|e\| > 1$. The elements $f, g$ are called the parents of $e$. (Unit vectors have no parents.)*

*Proof.* Existence, for a given irreducible $e$ with $\|e\| > 1$, of the direct acute basis $(f, g)$ such that $e = f + g$. We assume without loss of generality that $e = (\alpha, \beta)$ has non-negative coordinates. Since $\gcd(\alpha, \beta) = 1$ and $\|e\| > 1$ we obtain that $\alpha \geq 1$ and $\beta \geq 1$. Classical results on the Stern-Brocot tree [13] state that the irreducible positive fraction $\alpha/\beta$ can be written as the *mediant* $(\alpha' + \alpha'')/(\beta' + \beta'')$ of two irreducible fractions $\alpha'/\beta'$, $\alpha''/\beta''$ (possibly equal to 0 or $+\infty$), with $\alpha', \beta', \alpha'', \beta'' \in \mathbb{Z}_+$ and $\alpha'\beta'' - \beta'\alpha'' = 1$. Setting $f = (\alpha', \beta')$ and $g = (\alpha'', \beta'')$ concludes the proof.

Uniqueness. Assume that $e = f + g = f' + g'$, where $(f, g)$, $(f', g')$ are direct acute bases of $\mathbb{Z}^2$. One has $\det(f, e) = \det(f, f + g) = 1$, and likewise $\det(f', e) = 1$. Hence $\det(f - f', e) = 0$, and therefore $f' = f + ke$ for some scalar $k$, which is an integer since $e$ is irreducible. Subtracting we obtain $g' = e - f' = g - ke$, and therefore

$$\begin{aligned}\langle f', g'\rangle = \langle f + ke, g - ke\rangle &= \langle (k+1)f + kg, -kf + (1-k)g\rangle \\ &= -k(k+1)\|f\|^2 - k(k-1)\|g\|^2 + \langle f, g\rangle(1 - 2k^2).\end{aligned}$$

This expression is negative unless the integer $k$ is zero, hence $f = f'$, and $g = g'$. □

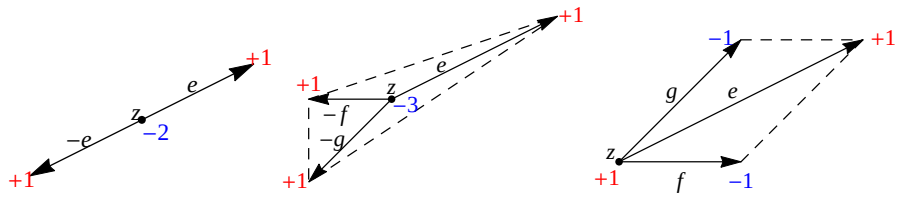



Figure 2: The supports, and weights, of the different linear forms $S^e_x$, $T^e_x$, $P^e_x$.

### 1.1 Characterization of discrete convexity by linear inequalities

We introduce some linear forms on the vector space $\mathcal{F}(X) := \{u : X \to \mathbb{R}\}$, which non-negativity characterizes restrictions of convex maps. The convex hulls of their respective supports have respectively the shape of a segment, a triangle, and a parallelogram, see Figure 2.

**Definition 1.3.** *For each $x \in \mathbb{Z}^2$, consider the following linear forms of $u \in \mathcal{F}(\mathbb{Z}^2)$.*

1. *(Segments) For any irreducible $e \in \mathbb{Z}^2$:*

$$S^e_x(u) := u(x+e) - 2u(x) + u(x-e).$$

2. *(Triangles) For any irreducible $e \in \mathbb{Z}^2$, with $\|e\| > 1$, of parents $f, g$:*

$$T^e_x(u) := u(x+e) + u(x-f) + u(x-g) - 3u(x).$$

3. *(Parallelograms) For any irreducible $e \in \mathbb{Z}^2$, with $\|e\| > 1$, of parents $f, g$:*

$$P^e_x(u) := u(x+e) - u(x+f) - u(x+g) + u(x).$$

*A linear form $L$ among the above can be regarded as a finite weighted sum of Dirac masses. In this sense we define the support* $\mathrm{supp}(L) \subseteq \mathbb{Z}^2$, $\#\,\mathrm{supp}(L) \in \{3, 4\}$. *The linear form $L$ is also defined on $\mathcal{F}(X)$ whenever* $\mathrm{supp}(L) \subseteq X$.

If $u \in \mathrm{Conv}(X)$, then by an immediate convexity argument one obtains $S^e_x(u) \geq 0$ and $T^e_x(u) \geq 0$, whenever these linear forms are supported on $X$. As shown in the next result, this provides a minimal characterization of $\mathrm{Conv}(X)$ by means of linear inequalities. The linear forms $P^e_x$ will on the other hand be used to define strict sub-cones of $\mathrm{Conv}(X)$. The following result corrects[1] Corollary 4 in [5].

**Theorem 1.4.**

- *The cone* $\mathrm{Conv}(X)$ *is characterized by the non-negativity of the linear forms $S^e_x$ and $T^e_x$, introduced in Definition 1.3, which are supported in $X$.*
- *If one keeps only one representative among the identical linear forms $S^e_x$ and $S^{-e}_x$, then the above characterization of* $\mathrm{Conv}(X)$ *by linear inequalities is minimal.*

For any given $x \in X$, the number of linear inequalities $S^e_x$ (resp. $T^e_x$) appearing in the characterization of $\mathrm{Conv}(X)$ is bounded by the number of irreducible elements $e \in \mathbb{Z}^2$ such that $x + e \in X$. Hence the $N$-dimensional cone $\mathrm{Conv}(X)$ is characterized by at most $2N^2$ linear inequalities, where $N := \#(X)$.

If all the elements of $X$ are *aligned*, this turns out to be an over estimate: one easily checks that exactly $N - 1$ inequalities of type $S^e_x$ remain, and no inequalities of type $T^e_x$. This favorable

[1] Precisely, the constraints $T^e_x$ were omitted in [5] for $\|e\| > \sqrt{2}$.

situation does not extend to the two dimensional case however, because irreducible elements arise frequently in $\mathbb{Z}^2$, with positive density [14]:

$$\frac{6}{\pi^2} = \lim_{n\to\infty} n^{-2} \# \left\{(i,j) \in \{1,\cdots,n\}^2;\, \gcd(i,j) = 1\right\}.$$

If the domain $\Omega$ has a non-empty interior, then one easily checks from this point that the minimal description of $\mathrm{Conv}(X)$ given in Theorem 1.4 involves no less than $cN^2$ linear constraints[2], where the constant $c > 0$ depends on the domain shape but not on its scale (or equivalently, not on the grid scale $h$ in (5)). This quadratic number of constraints, announced in the description of global constraint methods in the introduction, is a strong drawback for practical applications, which motivates the construction of adaptive sub-cones of $\mathrm{Conv}(X)$ in the next subsection.

**Remark 1.5** (Directional convexity)**.** *Several works addressing optimization problems posed on the cone of convex functions [5, 22], have in the past omitted all or part of the linear constraints $T^e_x$, $x \in X$, $e \in \mathbb{Z}^2$ irreducible with $\|e\| > 1$. We consider in Appendix A this weaker notion of discrete convexity, introducing the cone* $\mathrm{DConv}(X)$ *of directionally convex functions, defined by the non-negativity of only $S^e_x$, $x \in X$, $e \in \mathbb{Z}^2$ irreducible.*

*We show that elements of* $\mathrm{DConv}(X)$ *cannot in general be extended into globally convex functions, but that one can extend their restriction to a grid coarsened by a factor* 2*. We also introduce a hierarchy of sub-cones of* $\mathrm{DConv}(X)$*, similar to the one presented in the next subsection.*

## 1.2 Hierarchical cones of discrete convex functions

We introduce in this section the notion of stencils $\mathcal{V} = (\mathcal{V}(x))_{x\in X}$ on $X$, and discuss the properties (hierarchy, complexity) of cones $\mathrm{Conv}(\mathcal{V})$ attached to them. The following family $\mathcal{V}_{\max}$ of sets is referred to as the "maximal stencils": for all $x \in X$

$$\mathcal{V}_{\max}(x) := \{e \in \mathbb{Z}^2 \text{ irreducible};\, x + e \in X\}. \tag{6}$$

The convex cone generated by a subset $A$ of a vector space is denoted by $\mathrm{Cone}(A)$, with the convention $\mathrm{Cone}(\emptyset) = \{0\}$.

**Definition 1.6.** *A family $\mathcal{V}$ of stencils on $X$ (or just: "Stencils on $X$") is the data, for each $x \in X$ of a collection $\mathcal{V}(x) \subseteq \mathcal{V}_{\max}(x)$ (the stencil at $x$) of irreducible elements of $\mathbb{Z}^2$, satisfying the following properties:*

- *(Stability) Any parent $f \in \mathcal{V}_{\max}(x)$, of any $e \in \mathcal{V}(x)$, satisfies $f \in \mathcal{V}(x)$.*
- *(Visibility) One has* $\mathrm{Cone}(\mathcal{V}(x)) = \mathrm{Cone}(\mathcal{V}_{\max}(x))$.

*The set of candidates for refinement $\hat{\mathcal{V}}(x)$ consists of all elements $e \in \mathcal{V}_{\max}(x) \setminus \mathcal{V}(x)$ which two parents $f, g$ belong to $\mathcal{V}(x)$.*

In other words, a stencil $\mathcal{V}(x)$ at a point $x \in \Omega$ contains the parents of its members whenever possible (Stability), and covers all possible directions (Visibility). By construction, these properties are still satisfied by the *refined stencil* $\mathcal{V}(x) \cup \{e\}$, for any candidate for refinement $e \in \hat{\mathcal{V}}(x)$. The collection $\hat{\mathcal{V}}(x)$ is easily recovered from $\mathcal{V}(x)$, see Proposition 3.8.

**Definition 1.7.** *We attach to a family $\mathcal{V}$ of stencils on $X$ the cone* $\mathrm{Conv}(\mathcal{V}) \subseteq \mathcal{F}(X)$*, characterized by the non-negativity of the following linear forms: for all $x \in X$*

[2]This number of constraints is empirically (and slightly erroneously) estimated to $\mathcal{O}(N^{1.8})$ in [5].

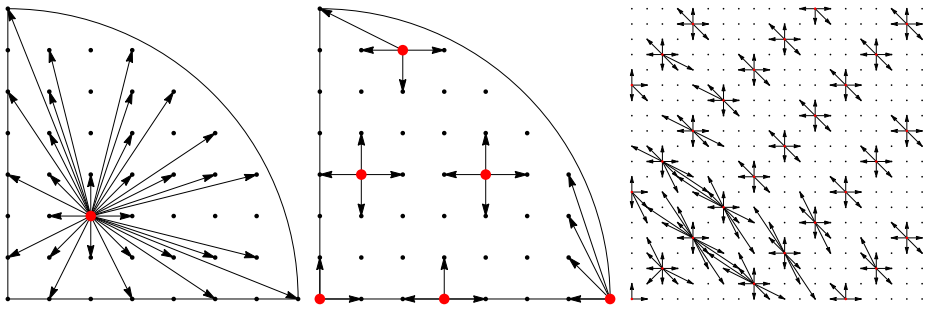


Figure 3: Left: a maximal stencil at a point of a domain. Center: some minimal stencils. Right: some adaptively generated stencils used in the numerical resolution of (1).

1. $S_x^e$, *for all* $e \in \mathcal{V}(x)$ *such that* $\mathrm{supp}(S_x^e) \subseteq X$.
2. $T_x^e$ *for all* $e \in \mathcal{V}(x)$, *with* $\|e\| > 1$, *such that* $\mathrm{supp}(T_x^e) \subseteq X$.
3. $P_x^e$ *for all* $e \in \hat{\mathcal{V}}(x)$ *(by construction* $\mathrm{supp}(P_x^e) \subseteq X$*).*

When discussing unions, intersections, and cardinalities, we (abusively) identify a family $\mathcal{V}$ of stencils on $X$ with a subset of $X \times \mathbb{Z}^2$:

$$\mathcal{V} \approx \{(x, e);\, x \in X,\, e \in \mathcal{V}(x)\}. \tag{7}$$

Note that the cone $\mathrm{Conv}(\mathcal{V})$ is defined by at most $3\#(\mathcal{V})$ linear inequalities. The sets $\mathcal{V}_{\max}$ are clearly stencils on $X$, which are maximal for inclusion, and by Theorem 1.4 we have $\mathrm{Conv}(\mathcal{V}_{\max}) = \mathrm{Conv}(X)$. The cone $\mathrm{Conv}(\mathcal{V})$ always contains the quadratic function $q(x) := \frac{1}{2}\|x\|^2$, for any family $\mathcal{V}$ of stencils. Indeed, the inequalities $S_x^e(q) \geq 0$, $x \in X$, $e \in \mathcal{V}(x)$, and $T_x^e(q) \geq 0$, $\|e\| > 1$, hold by convexity of $q$. In addition for all $e \in \hat{\mathcal{V}}(x)$, of parents $f, g$, one has

$$P_x^e(q) = \frac{1}{2}\left(\|x + f + g\|^2 - \|x + f\|^2 - \|x + g\|^2 + \|x\|^2\right) = \langle f, g\rangle \geq 0,$$

since the basis $(f, g)$ of $\mathbb{Z}^2$ is acute by definition, see Proposition 1.2.

**Theorem 1.8** (Hierarchy)**.** *The union* $\mathcal{V} \cup \mathcal{V}'$*, and the intersection* $\mathcal{V} \cap \mathcal{V}'$ *of two families* $\mathcal{V}, \mathcal{V}'$ *of stencils are also families of stencils on* $X$*. In addition*

$$\mathrm{Conv}(\mathcal{V}) \cap \mathrm{Conv}(\mathcal{V}') = \mathrm{Conv}(\mathcal{V} \cap \mathcal{V}'), \tag{8}$$
$$\mathrm{Conv}(\mathcal{V}) \cup \mathrm{Conv}(\mathcal{V}') \subseteq \mathrm{Conv}(\mathcal{V} \cup \mathcal{V}'). \tag{9}$$

As a result, if two families of stencils $\mathcal{V}, \mathcal{V}'$ satisfy $\mathcal{V} \subseteq \mathcal{V}'$, then

$$\mathrm{Conv}(\mathcal{V}) \subseteq \mathrm{Conv}(\mathcal{V}') \subseteq \mathrm{Conv}(X).$$

The left inclusion follows from (8), and the right inclusion from (9) applied to $\mathcal{V}'$ and $\mathcal{V}_{\max}$. The intersection rule (8) also implies the existence of stencils $\mathcal{V}_{\min}$ minimal for inclusion, which are illustrated on Figure 3 and characterized in Proposition 5.1.

**Remark 1.9** (Optimization strategy)**.** *For any* $u \in \mathrm{Conv}(x)$*, there exists by* (8) *a unique smallest (for inclusion) family of stencils* $\mathcal{V}$ *such that* $u \in \mathrm{Conv}(\mathcal{V})$*. If* $u$ *is the minimizer of an energy* $\mathcal{E}$ *on* $\mathrm{Conv}(X)$*, then it can be recovered by minimizing* $\mathcal{E}$ *on the smaller cone* $\mathrm{Conv}(\mathcal{V})$*, defined by* $\mathcal{O}(\#(\mathcal{V}))$ *linear constraints. Algorithm 1 in §6, attempts to find these smallest stencils* $\mathcal{V}$ *(or slightly larger ones), starting from* $\mathcal{V}_{\min}$ *and performing successive adaptive refinements.*

In the rest of this subsection, we fix a grid scale $h > 0$ and consider for all $\theta \in \mathbb{R}$, and all $\xi \in \mathbb{R}^2$, the grid

$$X_\theta^\xi := \Omega \cap hR_\theta(\xi + \mathbb{Z}^2). \tag{10}$$

The notions of stencils and related cones trivially extend to this setting, see §5 for details. We denote by $|\Omega|$ the domain area, and by $\mathrm{diam}(\Omega) := \max\{\|y - x\|;\, x, y \in \Omega\}$ its diameter. We also introduce rescaled variants, defined for $h > 0$ by

$$|\Omega|_h := h^{-2}|\Omega|, \qquad \mathrm{diam}_h(\Omega) := h^{-1}\,\mathrm{diam}(\Omega).$$

For any parameters $\theta, \xi$, one has denoting $N := \#(X_\theta^\xi)$ (with underlying constants depending only on the shape of $\Omega$)

$$|\Omega|_h \approx N, \qquad \mathrm{diam}_h(\Omega) \approx \sqrt{N}. \tag{11}$$

**Proposition 1.10.** *Let $X := X_\theta^\xi$, for some grid position parameters $\theta \in \mathbb{R}$, $\xi \in \mathbb{R}^2$, and let $N := \#(X)$. Let $u \in \mathrm{Conv}(X)$, and let $\mathcal{V}$ be the minimal stencils on $X$ such that $u \in \mathrm{Conv}(\mathcal{V})$. Then $\#(\mathcal{V}) \leq CN\,\mathrm{diam}_h(\Omega)$, for some universal constant $C$ (i.e. independent of $\Omega, h, \theta, \xi, u$).*

Combining this result with (11) we see that an optimization strategy as described in Remark 1.9 should heuristically not require solving optimization problems subject to more than $N\,\mathrm{diam}_h(\Omega) \approx N^{\frac{3}{2}}$ linear constraints. This is already a significant improvement over the $\approx N^2$ linear constraints defining $\mathrm{Conv}(X)$. The typical situation is however even more favorable: in average over randomized grid orientations $\theta$ and offsets $\xi$, the restriction to $X_\theta^\xi$ of a convex map $U : \Omega \to \mathbb{R}$ (e.g. the global continuous solution of the problem (1) of interest) belongs to a cone $\mathrm{Conv}(\mathcal{V}_\theta^\xi)$ defined by a quasi-linear number $\mathcal{O}(N \ln^2 N)$ of linear inequalities.

**Theorem 1.11.** *Let $U \in \mathrm{Conv}(\Omega)$, and let $\mathcal{V}_\theta^\xi$ be the minimal stencils on $X_\theta^\xi$ such that $U_{|X_\theta^\xi} \in \mathrm{Conv}(\mathcal{V}_\theta^\xi)$, for all $\theta \in \mathbb{R}$, $\xi \in \mathbb{R}^2$. Assuming $\mathrm{diam}_h(\Omega) \geq 2$, one has for some universal constant $C$ (i.e. independent of $h, \Omega, U$):*

$$\int_{[0,1]^2} \int_0^{\pi/2} \#(\mathcal{V}_\theta^\xi)\, d\theta\, d\xi \leq C\,|\Omega|_h\,(\ln \mathrm{diam}_h(\Omega))^2. \tag{12}$$

### 1.3 Stencils and triangulations

We discuss the connections between stencils and triangulations, which provides in Theorem 1.13 a new insight on the hierarchy of cones $\mathrm{Conv}(\mathcal{V})$, and yields in Theorem 1.15 a new result of algorithmic geometry as a side product of our theory. We assume in this subsection and §4 that the discrete domain convex hull, denoted by $\mathrm{Hull}(X)$, has a non-empty interior. All triangulations considered in this paper are implicitly assumed to cover $\mathrm{Hull}(X)$ and to have $X$ as collection of vertices.

**Definition 1.12.** *Let $\mathcal{T}$ be a triangulation, and let $\mathcal{V}$ be a family of stencils on $X$. We write $\mathcal{T} \prec \mathcal{V}$ iff the directed graph associated to $\mathcal{T}$ is included in the one associated to $\mathcal{V}$. In other words iff for any edge $[x, x + e]$ of $\mathcal{T}$, one has $e \in \mathcal{V}(x)$.*

The next result provides a new interpretation to our approach to optimization problems posed on the cone of convex functions, as a relaxation of the naïve (and flawed without this modification) method via interior finite elements.

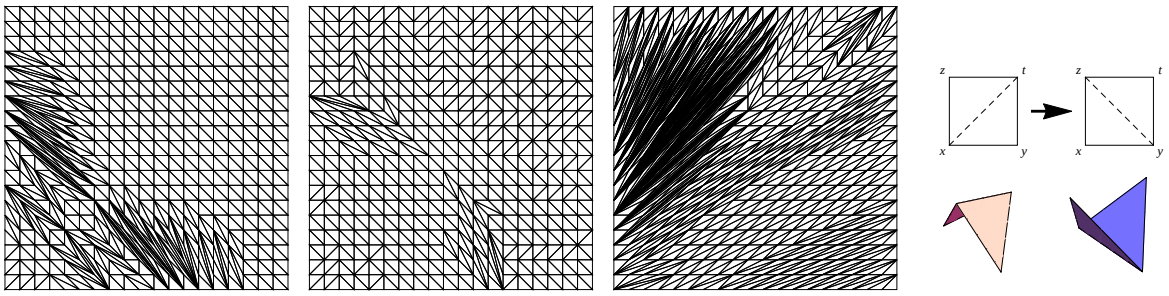



Figure 4: Delaunay triangulations associated to numerical solutions of (variants of) the monopolist problem, see 6.2. Corresponding convex functions shown on Figure 1 (left), Figure 7 (left), and Figure 8 (left) respectively. Right: illustration of edge flipping. The diagonal $[x,t]$ shared by two triangles $T := [x,y,t]$ and $T' := [x,z,t]$ can be flipped into $[y,z]$ if $T \cup T'$ is convex. Right, below: the piecewise linear interpolation of a discrete map $u$ is made convex by this flip.

**Theorem 1.13.** *Let $\mathcal{V}$ be a family of stencils on $X$. If $\mathrm{Conv}(\mathcal{V})$ has a non-empty interior, then*

$$\mathrm{Conv}(\mathcal{V}) = \bigcup_{T \prec \mathcal{V}} \mathrm{Conv}(\mathcal{T}). \tag{13}$$

Delaunay triangulations are a fundamental concept in discrete geometry [9]. We consider in this paper a slight generalization in which the lifting map needs not be the usual paraboloid, but can be an arbitrary convex function, see Definition 1.14. Within this paper Delaunay triangulations are simultaneously (i) a theoretical tool for proving results, notably Proposition 1.10 and Theorem 1.13, (ii) an object of study, since in Theorem 1.15 we derive new results on the cost of their construction, and (iii) a numerical post-processing tool providing global convex extensions of elements of $\mathrm{Conv}(X)$, see Figure 4 and Remark 6.3.

**Definition 1.14.** *We say that $\mathcal{T}$ is an $u$-Delaunay triangulation iff $u \in \mathrm{Conv}(\mathcal{T})$; equivalently the piecewise linear interpolation $\mathrm{I}_{\mathcal{T}}\, u : \mathrm{Hull}(X) \to \mathbb{R}$ is convex. We refer to $u$ as the* lifting map*.*

*A $q$-Delaunay triangulation, with $q(x) := \frac{1}{2}\|x\|^2$, is simply called a Delaunay triangulation.*

Two dimensional Delaunay triangulations, and three dimensional convex hulls, have well known links [9]. Indeed $\mathcal{T}$ is an $u$-Delaunay triangulation iff the map $x \in \mathrm{Hull}(X) \mapsto (x, \mathrm{I}_{\mathcal{T}}\, u(x)) \in \mathbb{R}^3$ spans the bottom part of the convex envelope $K$ of lifted set $\{(x,u(x));\, x \in X\}$. As a result of this interpretation, we find that (i) any element $u \in \mathrm{Conv}(X)$ admits an $u$-Delaunay triangulation, and (ii) generic elements of $\mathrm{Conv}(X)$ admit exactly one $u$-Delaunay triangulation (whenever all the faces of $K$ are triangular). In particular, the union (13) is disjoint up to a set of Hausdorff dimension $N-1$.

Any two triangulations $\mathcal{T}, \mathcal{T}'$ of $X$ can be transformed in one another through a sequence of elementary modifications called *edge flips*, see Figure 4. The minimal number of such operations is called the edge flipping distance between $\mathcal{T}$ and $\mathcal{T}'$. Edge flipping is a simple, robust and flexible procedure, which is used in numerous applications ranging from fluid dynamics simulation [8] to GPU accelerated image vectorization [23]. Sustained research has been devoted to estimating edge flipping distances within families of triangulations of interest [15], although flipping distance bounds are usually quadratic in the number of vertices.

**Theorem 1.15.** *Let $\mathcal{V}$ be a family of stencils on $X$, and let $u \in \mathrm{Conv}(\mathcal{V})$. Then any standard Delaunay triangulation of $X$ can be transformed into an $u$-Delaunay triangulation via a sequence of $C\#(\mathcal{V})$ edge flips. The constant $C$ is universal (in particular it is independent of $u, \mathcal{V}, \Omega$).*

Combining this result with Theorem 1.11, we obtain that only $\mathcal{O}(N \ln^2 N)$ edge flips are required to construct an $U$-Delaunay triangulation of $X$, with $N := \#(X)$, for any convex function $U \in \mathrm{Conv}(\Omega)$, in an average sense over grid orientations. Note that (more complex and specialized) convex hull algorithms [6] could also be used to produce an $U$-Delaunay triangulation, at the slightly lower cost $\mathcal{O}(N \ln N)$. Theorem 1.15 should be understood as a first step in understanding the typical behavior of edge flipping.

### 1.4 Outline

We prove in §2 the characterization of discrete convexity by linear constraints of Theorem 1.4. The hierarchy properties of Theorem 1.8 are established in §3. Triangulation related arguments are used in §4 to show Proposition 1.10 and Theorems 1.13 and 1.15. The average cardinality estimate of Theorem 1.11 is proved in §5. Numerical experiments, and algorithmic details, are presented in §6. Finally, the weaker notion of directional convexity is discussed in Appendix A.

## 2 Characterization of convexity via linear constraints

This section is devoted to the proof of Theorem 1.4, which characterizes discrete convex functions $u \in \mathrm{Conv}(X)$ via linear inequalities. The key ingredient is its generalization, in [5], to arbitrary unstructured finite point sets $X'$ (in contrast with the grid structure of $X$).

**Theorem 2.1** (Carlier, Lachand-Robert, Maury)**.** *Let $\Omega$ be a convex domain, and let $X' \subseteq \Omega$ be an arbitrary finite set. The cone $\mathrm{Conv}(X')$, of all restrictions to $X'$ of convex functions on $\Omega$, is characterized by the following inequalities, none of which can be removed:*

- *For all $x, y, z \in X'$, all $\alpha \in ]0,1[$, $\beta := 1-\alpha$, such that $z = \alpha x + \beta y$ and $[x,y] \cap X' = \{x, y, z\}$:*

$$\alpha u(x) + \beta u(y) \geq u(z). \tag{14}$$

- *For all $p, q, r, z \in X'$, all $\alpha, \beta, \gamma \in \mathbb{R}_+^*$ with $\alpha + \beta + \gamma = 1$, such that $z = \alpha p + \beta q + \gamma r$, the points $p, q, r$ are not aligned and $[p,q,r] \cap X' = \{p, q, r, z\}$:*

$$\alpha u(p) + \beta u(q) + \gamma u(r) \geq u(z). \tag{15}$$

In the following, we establish Theorem 1.4 by specializing Theorem 2.1 to the grid $X := \Omega \cap \mathbb{Z}^2$, and identifying (14) and (15) with the constraints $S_z^e(u) \geq 0$ and $T_z^e(u) \geq 0$ respectively, in Propositions 2.2 and 2.5 respectively. Note that the cases (14) and (15) are unified in [5], although we separated them in the above formulation for clarity. For $x_1, \cdots, x_n \in \mathbb{R}^2$, we denote

$$[x_1, \cdots, x_n] := \mathrm{Hull}(\{x_1, \cdots, x_n\}).$$

**Proposition 2.2.** *Let $x, y, z \in X$, $\alpha \in ]0,1[$, $\beta := 1 - \alpha$, be such that $z = \alpha x + \beta y$ and $[x,y] \cap X = \{x, y, z\}$. Then $\alpha = \beta = 1/2$ and there exists an irreducible $e \in \mathbb{Z}^2$ such that $x = z + e$, $y = z - e$. (Thus $\alpha u(x) + \beta u(y) - u(z) = S_z^e(u)/2$.)*

*Proof.* We define $e := x - z \in \mathbb{Z}^2$, and assume for contradiction that $e$ is not irreducible: $e = ke'$, for some integer $k \geq 2$ and some $e' \in \mathbb{Z}^2$. Then $z + e' \in [z,x] \cap \mathbb{Z}^2 \subseteq [x,y] \cap X$, which is a contradiction. Thus $e$ is irreducible, and likewise $f := y - z \in \mathbb{Z}^2$ is irreducible. Observing that $f$ is negatively proportional to $e$, namely $f = -(\alpha/\beta)e$, we obtain that $f = -e$, which concludes the proof. □

The next lemma, used in Proposition 2.5 to identify the constraints (15), provides an alternative characterization of the parents of an irreducible vector, see Proposition 1.2.

**Lemma 2.3.** *Let $e, f, g \in \mathbb{Z}^2$ be such that $e = f + g$, $\|e\| \geq \max\{\|f\|, \|g\|\}$, and $\det(f, g) = 1$. Then $f, g$ are the parents of $e$.*

*Proof.* Since $\det(f, e) = \det(f, f + g) = 1$, the vector $e$ is irreducible. In addition, $\|e\| > 1$, since otherwise $e, f, g$ would be three pairwise linearly independent unit vectors in $\mathbb{Z}^2$. Let $f_0, g_0$ be the parents of $e$. Observing that $\det(f, e) = \det(f_0, e) = 1$, we obtain that $f = f_0 + ke$ for some scalar $k$, which must be an integer since $e$ is irreducible. If $k > 0$ then $\|f\|^2 = k^2\|e\|^2 + 2k\langle f_0, e\rangle + \|f_0\|^2 > \|e\|^2$ which is a contradiction (recall that $\langle f_0, e\rangle = \langle f_0, f_0 + g_0\rangle \geq 0$). If $k < 0$, then observing that $g = e - f = g_0 - ke$ we obtain likewise a contradiction. Hence $k = 0$ and $f, g$ are the parents of $e$, which concludes the proof. □

**Corollary 2.4.** *Let $(f, g)$ be a basis of $\mathbb{Z}^2$ which is* not *acute. If $\|f\| \geq \|g\|$, then $f + g$ is a parent of $f$, and otherwise it is a parent ot $g$.*

*Proof.* Up to exchanging the roles of $f, g$, we may assume that $\det(f, g) = 1$. Denoting $m := \max\{\|f\|, \|g\|, \|f + g\|\}$, we have by Lemma 2.3 three possibilities: (i) $\|f + g\| = m$, and $f, g$ are the parents of $f + g$, (ii) $\|f\| = m$, and $-g, f + g$ are the parents of $f$, (iii) $\|g\| = m$, and $f + g, -f$ are the parents of $g$. Excluding (i), since $(f, g)$ is not an acute basis, we conclude the proof. □

**Proposition 2.5.** *Let $p, q, r, z \in X$, $\alpha, \beta, \gamma \in \mathbb{R}_+^*$, with $\alpha + \beta + \gamma = 1$, be such that $z = \alpha p + \beta q + \gamma r$, the points $p, q, r$ are not aligned, and $[p, q, r] \cap X = \{p, q, r, z\}$. Then $\alpha = \beta = \gamma = 1/3$ and, up to permuting $p, q, r$, there exists an irreducible $e \in \mathbb{Z}^2$ with $\|e\| > 1$, of parents $f, g$, such that $p = z + e$, $q = z - f$, and $r = z - g$. (Thus $\alpha u(p) + \beta u(q) + \gamma u(r) - u(z) = T_z^e(u)/3$.)*

*Proof.* Let $e := p - z$, $f := z - q$, $g := z - r$. Up to permuting $p, q, r$, we may assume that $\|e\| \geq \max\{\|f\|, \|g\|\}$ and $\det(f, g) \geq 0$. Note that $f$ and $g$ are not collinear since $z$ lies in the interior of $[p, q, r]$. We claim that $(f, g)$ is a basis of $\mathbb{Z}^2$. Indeed, otherwise, the triangle $[0, f, g]$ would contain an element of $e' \in \mathbb{Z}^2$ distinct from its vertices. Since $\Omega$ is convex, this implies $[p, q, r] \cap X \supseteq \{p, q, r, z, z + e'\}$, which is a contradiction.

Thus $(f, g)$, and likewise $(e, f)$ and $(e, g)$, are bases of $\mathbb{Z}^2$, and therefore

$$|\det(e, f)| = |\det(f, g)| = |\det(g, e)| = 1. \tag{16}$$

Injecting in the above equation the identity $e = (\beta/\alpha)f + (\gamma/\alpha)g$, we obtain that $|\beta/\alpha| = 1$ and $|\gamma/\alpha| = 1$. Thus $\alpha = \beta = \gamma = 1/3$ since these coefficients are positive and sum to one. Finally, we have $e = f + g$, $\|e\| \geq \max\{\|f\|, \|g\|\}$, and $f, g$ is a direct basis. Using Lemma 2.3 we conclude as announced that $f, g$ are the parents of $e$. □

# 3 Hierarchy of the cones $\mathrm{Conv}(\mathcal{V})$

This section is devoted to the proof of Theorem 1.8, which is split into two parts: the proof that an intersection (or union) of stencils is still a stencil, and the hierarchy properties (8) and (9).

### 3.1 An intersection of stencils is still a stencil

Let $\mathcal{V}, \mathcal{V}'$ be families of stencils on $X$. Property (Stability) of stencils immediately holds for the intersection $\mathcal{V} \cap \mathcal{V}'$ and union, $\mathcal{V} \cup \mathcal{V}'$, while property (Visibility) is also clear for the union $\mathcal{V} \cup \mathcal{V}'$. In order to establish property (Visibility) for the intersection $\mathcal{V} \cap \mathcal{V}'$, we identify in Proposition 3.7 a family $\mathcal{V}_{\min}$ of stencils which included in any other. From $\mathcal{V}_{\min} \subseteq \mathcal{V}$ and $\mathcal{V}_{\min} \subseteq \mathcal{V}'$ we obtain $\mathcal{V}_{\min} \subseteq \mathcal{V} \cap \mathcal{V}'$, so that (Visibility) for $\mathcal{V}_{\min}$ implies the same property for $\mathcal{V} \cap \mathcal{V}'$.

**Definition 3.1.** *The cyclic strict trigonometric order on $\mathbb{R}^2 \setminus \{0\}$ is denoted by $\prec$.*

In other words $e_1 \prec e_2 \prec e_3$ iff there exists $\theta_1, \theta_2, \theta_3 > 0$, such that $\theta_1 + \theta_2 + \theta_3 = 2\pi$ and $e_{i+1}/\|e_{i+1}\| = R_{\theta_i} e_i / \|e_i\|$ for all $1 \leq i \leq 3$, with $e_4 := e_1$. The following lemma, and Corollary 3.5, discuss the combination of the cyclic ordering with the notions of parents (and children) of an irreducible vector.

**Definition 3.2** (Collection of ancestors of a vector)**.** *For any irreducible $e \in \mathbb{Z}^2$, let $\mathrm{Anc}(e)$ be the smallest set containing $e$ and the parents of any element $e' \in \mathrm{Anc}(e)$ such that $\|e'\| > 1$.*

**Lemma 3.3.** *Let $e \in \mathbb{Z}^2 \setminus \{0\}$, let $(f, g)$ be a direct basis of $\mathbb{Z}^2$ such that $f \prec e \prec g$, and let us consider the triangle $T := [e, f, g]$. Then (i) $f + g \in T$. If in addition $e$ is irreducible, $\|e\| > 1$ and (ii.a) $\langle f, g \rangle \geq 0$ or (ii.b) $e \notin \mathrm{Anc}(f) \cup \mathrm{Anc}(g)$, then the parents of $e$ also belong to $T$.*

*Proof.* Point (i). By construction, we have $e = \alpha x + \beta y$ for some positive integers $\alpha, \beta$. One easily checks that $e + (\beta - 1)f + (\alpha - 1)g = (\alpha + \beta - 1)(f + g)$. This expression of $f + g$ as a weighted barycenter of the points $e, f, g$, establishes (i).

Points (ii.a) and (ii.b). We fix $e$ and show these points by *decreasing* induction on the integer $k = \langle f, g \rangle$. Initialization: Assume that $k = \langle f, g \rangle \geq \frac{1}{2}\|e\|^2$. Then $\|e\|^2 = \|\alpha f + \beta g\|^2 \geq 2\alpha\beta\langle f, g \rangle \geq 2\langle f, g \rangle$, which is a contradiction. No basis $(f, g)$ satisfies simultaneously $f \prec e \prec g$ and $\langle f, g \rangle = k$. The statement is vacuous, hence true.

Case $k = \langle f, g \rangle \geq 0$. If $e = f + g$, then $f, g$ are the parents of $e$, and the result follows. Otherwise, we have either $f \prec e \prec (f + g)$ or $(f + g) \prec e \prec g$. Since $\langle f, f + g \rangle > \langle f, g \rangle$ and $\langle f + g, g \rangle > \langle f, g \rangle$, we may apply our induction hypothesis to the bases $(f, f + g)$ and $(f + g, g)$ which satisfy (ii.a). Thus the parents of $e$ belong to $T_1 := [e, f, f + g]$ or $T_2 := [e, g, f + g]$. Finally, Point (i) implies that $f + g \in T$, thus $T_1 \cup T_2 \subseteq T$ which concludes the proof of this case.

Case $k = \langle f, g \rangle < 0$. Assumption (ii.b) must hold, since (ii.a) contradicts this case. By corollary 2.4, $f + g$ is a parent of $f$ or of $g$. Hence $e \neq f + g$ and $\mathrm{Anc}(f + g) \subseteq \mathrm{Anc}(f) \cup \mathrm{Anc}(g)$. We apply our induction hypothesis to the bases $(f, f + g)$ and $(f + g, g)$ which satisfy (ii.b), and conclude the proof similarly to the case $k \geq 0$. □

**Lemma 3.4.** *Consider an irreducible $e \in \mathbb{Z}^2$, $\|e\| > 1$, and let $f, g$ be its parents. The children of $e$ (i.e. the vectors $e' \in \mathbb{Z}^2$ of which $e$ is a parent) have the form $f + ke$ and $ke + g$, $k \geq 1$.*

*Proof.* Let $e'$ be a children of $e$, and let $f'$, $g'$ be its parents. Without loss of generality, we assume that $g' = e$. Then $\det(f', e) = 1 = \det(f, e)$, thus $f' = f + ke$ for some $k \in \mathbb{R}$. Since $e$ is irreducible, one has $k \in \mathbb{Z}$. Since $0 \leq \langle f', e \rangle = \langle f, e \rangle + k\|e\|^2 \leq (k - 1)\|e\|^2$, one has $k \geq 1$. The result follows. □

**Corollary 3.5.** *Let $e \in \mathbb{Z}^2 \setminus \{0\}$, let $(f, g)$ be a direct acute basis of $\mathbb{Z}^2$ such that $f \prec e \prec g$. Then any child $e'$ of $e$ (i.e. $e$ is a parent of $e'$) satisfies $f \prec e' \prec g$.*

*Proof.* Let $K := \mathrm{Cone}(\{f, g\})$, and let $\mathring{K}$ be its interior. Let also $f', g'$ be the parents of $e$. By assumption $e \in \mathring{K}$, and by Lemma 3.3 (ii.a) one has $f', g' \in [e, f, g] \subseteq K$. Thus $f' + ke, ke + g' \in \mathring{K}$ for any integer $k \geq 1$, which by Lemma 3.4 concludes the proof. □

**Lemma 3.6** (Consecutive elements of a stencil)**.** *Let $\mathcal{V}$ be a family of stencils on $X$, let $x \in X$, and let $f, g$ be two trigonometrically consecutive elements of $\mathcal{V}(x)$. Then either (i) $(f, g)$ form a direct acute basis, or (ii) no element $e \in \mathcal{V}_{\max}(x)$ satisfies $f \prec e \prec g$.*

*Proof.* We distinguish three cases, depending on the value of $\det(f, g)$. In case $\det(f, g) \leq 0$, property Visibility of stencils implies (ii).

Case $\det(f, g) = 1$. Assuming that (i) does not hold, Corollary 2.4 implies that $f + g$ is a parent of $f$ or of $g$. Assuming that (ii) does not hold, we have $f + g \in [e, f, g]$ for some $e \in \mathcal{V}_{\max}(x)$ by Lemma 3.3 (i), thus $f + g \in \mathcal{V}_{\max}(x)$ by convexity of $\Omega$, thus $f + g \in \mathcal{V}(x)$ by (Stability), which contradicts our assumption $f, g$ are trigonometrically consecutive in $\mathcal{V}(x)$.

Case $k := \det(f, g) > 1$. We assume without loss of generality that $\|f\| \geq \|g\|$, hence $\|f\|^2 \geq \det(f, g) > 1$. Let $f', g'$ be the parents of $f$, so that $g = \alpha f' + \beta g'$ for some $\alpha, \beta \in \mathbb{Z}$. We obtain $k = \det(f, \alpha f' + \beta g') = \beta - \alpha$. If $\alpha = 0$ or $\beta = 0$, then $g$ is not irreducible, which is a contradiction. If $\alpha$ and $\beta$ have the same sign, then $\|g\| > \|f\|$, which again is a contradiction. Hence $\alpha, \beta$ have opposite signs, and since $\beta - \alpha = k > 1$ we obtain $\beta > 0 > \alpha$. Finally we have $g = \alpha(f - g') + \beta g'$, thus $g - \alpha f = (\beta - \alpha)g'$, and therefore $g' \in [0, f, g]$. By convexity $g' \in \mathcal{V}_{\max}(x)$, by (Stability) $g' \in \mathcal{V}(x)$, which contradicts our assumption that $f, g$ are trigonometrically consecutive in $\mathcal{V}(x)$. This concludes the proof. □

**Proposition 3.7** (Characterization of the smallest stencils)**.** *For all $x \in X$, define $\mathcal{V}_{\min}(x)$ as the collection of all $e \in \mathcal{V}_{\max}(x)$ which have none or just one parent in $\mathcal{V}_{\max}(x)$ (this includes all unit vectors in $\mathcal{V}_{\max}(x)$). Then $\mathcal{V}_{\min} := (\mathcal{V}_{\min}(x))_{x \in X}$ is a family of stencils, which is contained in any other family of stencils.*

*Proof.* Property (Stability) of stencils. Consider $x \in X$ and $e \in \mathcal{V}_{\min}(x)$. Assume for contradiction that $e$ has one parent $e' \in \mathcal{V}_{\max}(x)$ which is not an element of $\mathcal{V}_{\min}(x)$. Hence $e'$ has two parents $f, g \in \mathcal{V}_{\max}(x)$. By Corollary 3.5 we have $f \prec e \prec g$, thus by Lemma 3.3 (ii.a) the *two* parents of $e$ belong to the triangle $[e, f, g]$, hence also to $\mathcal{V}_{\max}(x)$ by convexity of $\Omega$. This contradicts our assumption that $e \in \mathcal{V}_{\min}(x)$.

Property (Visibility). We consider $e \in \mathcal{V}_{\max}(x)$, and prove by induction on the norm $\|e\|$ that $e \in K := \mathrm{Cone}(\mathcal{V}_{\min}(x))$. If $\|e\| = 1$ or if none of just one parent of $e$ belongs to $\mathcal{V}_{\max}(x)$, then $e \in \mathcal{V}_{\min}(x) \subseteq K$. If both parents $f, g$ of $e$ belong to $\mathcal{V}_{\max}(x)$, then by induction $f, g \in K$, and by additivity $f + g \in K$, which concludes the proof.

Minimality for inclusion of $\mathcal{V}_{\min}$. Let $\mathcal{V}$ be a family of stencils, let $x \in X$, $e \in \mathcal{V}_{\min}(x)$, and let us assume for contradiction that $e \notin \mathcal{V}(x)$. By property (Visibility) of stencils, the vector $e$ belongs to the cone generated by two elements $f, g \in \mathcal{V}(x)$, which can be chosen trigonometrically consecutive in $\mathcal{V}(x)$. By lemma 3.6, $(f, g)$ is a direct acute basis of $\mathbb{Z}^2$. By Lemma 3.3 (ii.a) the parents of $e$ belong to the triangle $[e, f, g]$, hence to $\mathcal{V}_{\max}(x)$ by convexity of $\Omega$, which contradicts the definition of $\mathcal{V}_{\min}(x)$. □

**Proposition 3.8** (Structure of candidates for refinement)**.** *Let $\mathcal{V}$ be a family of stencils on $X$, and let $x \in X$. Then the parents $f, g$, of any candidate for refinement $e \in \hat{\mathcal{V}}(x)$, are consecutive elements of $\mathcal{V}(x)$ in trigonometric order.*

*Proof.* Since $e \notin \mathcal{V}(x)$, there exists by (Visibility) two elements $f, g \in \mathcal{V}(x)$ such that $f \prec e \prec g$, and which we can choose trigonometrically consecutive in $\mathcal{V}(x)$. By Lemma 3.6, $(f, g)$ is a direct acute basis. By Lemma 3.3 (ii.a) the parents $f', g'$ of $e$ between satisfy $f \preceq f' \prec g' \preceq g$. Recalling that $f', g' \in \mathcal{V}(x)$, by definition of $\hat{\mathcal{V}}(x)$, we obtain $f = f'$ and $g = g'$ which concludes the proof. □

### 3.2 Combining and intersecting constraints

The following characterization of the cones $\mathrm{Conv}(\mathcal{V})$ implies the announced hierarchy properties.

**Proposition 3.9.** *For any family $\mathcal{V}$ of stencils on $X$ one has*

$$\mathrm{Conv}(\mathcal{V}) = \{u \in \mathrm{Conv}(X);\ P_x^e(u) \geq 0 \text{ for all } x \in X,\ e \in \mathcal{V}(x)\}. \tag{17}$$

Before turning to the proof of this proposition, we use it to conclude the proof of Theorem 1.8. The sub-cone $\mathrm{Conv}(\mathcal{V})$, of $\mathrm{Conv}(\Omega)$, is characterized by the non-negativity of a family of linear forms indexed by $\mathcal{V}_{\max} \setminus \mathcal{V}$, with the convention (7). Observing that

$$\mathcal{V}_{\max} \setminus (\mathcal{V} \cup \mathcal{V}') = (\mathcal{V}_{\max} \setminus \mathcal{V}) \cap (\mathcal{V}_{\max} \setminus \mathcal{V}'), \qquad \mathcal{V}_{\max} \setminus (\mathcal{V} \cap \mathcal{V}') = (\mathcal{V}_{\max} \setminus \mathcal{V}) \cup (\mathcal{V}_{\max} \setminus \mathcal{V}'),$$

we find that $\mathrm{Conv}(\mathcal{V} \cup \mathcal{V}')$ is characterized, as a subset of $\mathrm{Conv}(X)$, by the intersection of the families of constraints defining $\mathrm{Conv}(\mathcal{V})$ and $\mathrm{Conv}(\mathcal{V}')$, while $\mathrm{Conv}(\mathcal{V} \cap \mathcal{V}')$ is defined by their union. Hence we conclude as announced

$$\mathrm{Conv}(\mathcal{V} \cup \mathcal{V}') \supseteq \mathrm{Conv}(\mathcal{V}) \cup \mathrm{Conv}(\mathcal{V}'), \qquad \mathrm{Conv}(\mathcal{V} \cap \mathcal{V}') = \mathrm{Conv}(\mathcal{V}) \cap \mathrm{Conv}(\mathcal{V}').$$

*Proof of Proposition 3.9.* We proceed by decreasing induction on the cardinality $\#(\mathcal{V})$.

Initialization. If $\#(\mathcal{V}) = \#(\mathcal{V}_{\max})$, then $\mathcal{V} = \mathcal{V}_{\max}$, and therefore $\mathrm{Conv}(\mathcal{V}) = \mathrm{Conv}(\mathcal{V}_{\max}) = \mathrm{Conv}(X)$ and $\mathcal{V}_{\max} \setminus \mathcal{V} = \emptyset$. The result follows.

Induction. Assume that $\#(\mathcal{V}) < \#(\mathcal{V}_{\max})$, thus $\mathcal{V} \subsetneq \mathcal{V}_{\max}$. Let $x \in X$ and $e \in \mathcal{V}_{\max}(x) \setminus \mathcal{V}(x)$ be such that $\|e\|$ is minimal. Since $e \notin \mathcal{V}_{\min}(x) \subseteq \mathcal{V}(x)$, the two parents $f, g$ of $e$ belong to $\mathcal{V}_{\max}(x)$. Since $\|e\| > \max\{\|f\|, \|g\|\}$, and by minimality of the norm of $e$, we have $f, g \in \mathcal{V}(x)$. Hence $e$ is a candidate for refinement: $e \in \hat{\mathcal{V}}(x)$.

Consider the extended stencils $\mathcal{V}'$ defined by $\mathcal{V}'(x) := \mathcal{V}(x) \cup \{e\}$, and $\mathcal{V}'(y) := \mathcal{V}(y)$ for all $y \in X \setminus \{x\}$. Let $\mathcal{L}$ and $\mathcal{L}'$ be the collections of linear forms enumerated in Definition 1.7, which non-negativity respectively defines the cones $\mathrm{Conv}(\mathcal{V})$ and $\mathrm{Conv}(\mathcal{V}')$ as subsets of $\mathcal{F}(X)$. Let also $\mathcal{L}_0 := \mathcal{L} \cap \mathcal{L}'$. Since $e \in \hat{\mathcal{V}}(x)$ we have $\mathcal{L} = \mathcal{L}_0 \cup \{P_x^e\}$. Using Proposition 3.8 we obtain $\hat{\mathcal{V}}'(x) \setminus \hat{\mathcal{V}}(x) \subseteq \{e+f, e+g\}$, hence $\mathcal{L}'$ is the union of $\mathcal{L}_0$ and of those of the following constraints which are supported on $X$:

$$S_x^e,\ T_x^e,\ P_x^{e+f},\ P_x^{e+g}. \tag{18}$$

We next show that $\mathrm{Cone}(\mathcal{L}) = \mathrm{Cone}(\mathcal{L}' \cup \{P_x^e\})$, by expressing the linear forms (18) in terms of the elements of $\mathcal{L}$.

- If $S_x^e$ is supported on $X$, then $-e \in \mathcal{V}_{\max}(x)$. Assuming that $-e \in \mathcal{V}(x)$, we obtain $S_x^e = S_x^{-e} \in \mathcal{L}$. On the other hand, assuming that $-e \notin \mathcal{V}(x)$, we obtain $-f, -g \in \mathcal{V}_{\max}(x)$ by Proposition 3.7, since otherwise $-e \in \mathcal{V}_{\min}(x) \subseteq \mathcal{V}(x)$. Therefore $S_x^f, S_x^g$ are supported on $X$, hence they belong to $\mathcal{L}$. By minimality of the norm of $e$, we have $-f, -g \in \mathcal{V}(x)$, hence $-e \in \hat{\mathcal{V}}(x)$ and therefore $P_x^{-e} \in \mathcal{L}$. As a result $S_x^e = P_x^e + P_x^{-e} + S_x^f + S_x^g \in \mathrm{Cone}(\mathcal{L})$.

- If $T_x^e$ is supported on $X$, then $-f, -g \in \mathcal{V}_{\max}(x)$. Therefore $S_x^f, S_x^g$ are supported on $X$, hence they belong to $\mathcal{L}$. As a result $T_x^e = P_x^e + S_x^f + S_x^g \in \mathrm{Cone}(\mathcal{L})$.

- If $P_x^{e+f}$ is supported on $X$, then $x + e + f \in X$, thus $f \in \mathcal{V}_{\max}(x + e)$ and therefore $f \in \mathcal{V}(x+e)$ by minimality of $\|e\|$. The linear form $S_{x+e}^f$ belongs to $\mathcal{L}$, since it has support $\{x + g, x + e, x + e + f\} \subseteq X$. Observing that the parents of $e + f$ are $e$ and $f$, we find that $P_x^{e+f} = P_x^e + S_{x+e}^f \in \mathrm{Cone}(\mathcal{L})$. The case of $P_x^{e+g}$ is similar.

Denoting by $K^*$ the dual cone of a cone $K$, we obtain

$$\mathrm{Conv}(\mathcal{V}) = \mathrm{Cone}(\mathcal{L})^* = \mathrm{Cone}(\mathcal{L}' \cup \{P_x^e\})^* = \{u \in \mathrm{Conv}(\mathcal{V}');\ P_x^e(u) \geq 0\}.$$

Applying the induction hypothesis to $\mathcal{V}'$, we conclude the proof. □

# 4 Stencils and triangulations

Using the interplay between stencils $\mathcal{V}$ and triangulations $\mathcal{T}$, we prove Proposition 1.10 and Theorems 1.13, 1.15. By convention, all stencils $\mathcal{V}$ are on $X$, and all triangulations $\mathcal{T}$ have $X$ as vertices and cover $\mathrm{Hull}(X)$.

## 4.1 Minimal stencils containing a triangulation

We characterize in Proposition 4.4 the minimal stencils $\mathcal{V}$ containing a triangulation, in the sense of Definition 1.12, and we estimate their cardinality, proving Proposition 1.10. In the way, we establish in Proposition 4.3 "half" (one inclusion) of the decomposition of $\mathrm{Conv}(\mathcal{V})$ announced in Theorem 1.13.

**Lemma 4.1.** *Let $\mathcal{T}$ be a triangulation, and let $u \in \mathrm{Conv}(\mathcal{T})$. Let $p, q, r \in X$. Assume that $[p, q]$ is an edge of $\mathcal{T}$, and that $s := p + q - r \in X$. Then $u(r) + u(s) \geq u(p) + u(q)$.*

*Proof.* The interpolating function $U := \mathrm{I}_{\mathcal{T}}\, u$ is convex on $\mathrm{Hull}(X)$, and linear on the edge $[p, q]$. Introducing the edge midpoint $m := (p + q)/2 = (r + s)/2$ we obtain $u(p) + u(q) = 2U(m) \leq u(r) + u(s)$, as announced. □

The inequalities $u(r) + u(s) \geq u(p) + u(q)$ identified in the previous lemma are closely tied with the linear constraints $P_x^e$, since $[p, q, r, s]$ is a parallelogram, and as shown in the next lemma. The set $\mathrm{Anc}(e)$ of ancestors of an irreducible vector $e \in \mathbb{Z}^2$ was introduced Definition 3.2.

**Lemma 4.2.** *Let $e \in \mathbb{Z}^2$ be irreducible, with $\|e\| > 1$, and let $(f, g)$ be a direct basis such that $f \prec e \prec g$ and $e \notin \mathrm{Anc}(f) \cup \mathrm{Anc}(g)$. Let $x \in X$ be such that $f, g, f + g, e \in \mathcal{V}_{\max}(x)$. If $u \in \mathrm{Conv}(X)$ satisfies $u(x) + u(x + f + g) \geq (x + f) + u(x + g)$, then $P_x^e(u) \geq 0$.*

*Proof.* Without loss of generality, up to adding a global affine map to $u$, we may assume that $u(x + e) = u(x + f) = u(x + g) = 0$. Denoting by $f', g'$ the parents of $e$, we have by Lemma 3.3 (ii.b) $f', g', f+g \in [e, f, g]$, hence by convexity $u(x+f'), u(x+g'), u(x+f+g) \leq 0$. Our hypothesis implies $u(x) \geq -u(x+f+g) \geq 0$, therefore $P_x^e(u) = u(x)-u(x+f')-u(x+g')+u(x+e) \geq 0$. □

**Proposition 4.3.** *If a triangulation $\mathcal{T}$, and stencils $\mathcal{V}$, satisfy $\mathcal{T} \prec \mathcal{V}$, then $\mathrm{Conv}(\mathcal{T}) \subseteq \mathrm{Conv}(\mathcal{V})$.*

*Proof.* The inequalities $S_x^e(u) \geq 0$, and $T_x^e(u) \geq 0$, for $x \in X$, $e \in \mathcal{V}(x)$, hold by convexity of $u$. We thus consider an arbitrary refinement candidate $e \in \hat{\mathcal{V}}(x)$, $x \in X$, and establish below that $P_x^e(u) \geq 0$.

Since the triangulation $\mathcal{T}$ covers $\mathrm{Hull}(X)$, there exists a triangle $T \in \mathcal{T}$, containing $x$, and such that $e \in \mathrm{Cone}(T - x)$. Since $\mathcal{T} \prec \mathcal{V}$ and $e \notin \mathcal{V}(x)$, the segment $[x, x + e]$ is not an edge of $\mathcal{T}$. Denoting the vertices of $T$ by $[x, x + f, x + g]$ we have $f \prec e \prec g$. Since $e \in \hat{\mathcal{V}}(x)$, one has $e \notin (\mathrm{Anc}(f) \cup \mathrm{Anc}(g)) \subseteq \mathcal{V}(x)$. Applying Lemma 4.1 to the edge $[x + f, x + g]$ we obtain $u(x)+u(x+f+g) \geq u(x+f)+u(x+g)$. Finally, Lemma 4.2 implies $P_x^e(u) \geq 0$ as announced. □

**Proposition 4.4.** *Let $\mathcal{T}$ be a triangulation, and for all $x \in X$ let $V_x$ be the collection of all $e \in \mathbb{Z}^2$ such that $[x, x+e]$ is an edge of $T$. The minimal family of stencils satisfying $\mathcal{T} \prec \mathcal{V}$ is given by*

$$\mathcal{V}(x) := \mathcal{V}_{\max}(x) \cap \bigcup_{e \in V_x} \mathrm{Anc}(e).$$

*Proof.* The family of sets $\mathcal{V}$ satisfies the (Stability) property by construction. Since the triangulation $\mathcal{T}$ covers $\mathrm{Hull}(X)$, the sets $(V_x)_{x\in X}$ satisfy the (Visibility) property, hence also the larger sets $\mathcal{V}(x) \supseteq V_x$.

Minimality. Consider *arbitrary* stencils $\mathcal{V}$ such that $\mathcal{T} \prec \mathcal{V}$. Let also $x \in X$, $e \in V_x$, $e' \in \mathcal{V}_{\max}(x) \cap \mathrm{Anc}(e)$, and let us assume for contradiction that $e' \notin \mathcal{V}(x)$. By property (Visibility) there exists $f, g$, trigonometrically consecutive elements of $\mathcal{V}(x)$, such that $f \prec e \prec g$ (where $\prec$ refers to the cyclic trigonometric order, see Definition 3.1). By Lemma 3.6, $(f, g)$ is a direct acute basis of $\mathbb{Z}^2$. By Corollary 3.5, and an immediate induction argument, we have $f \prec e \prec g$, hence $e \notin \mathcal{V}(x)$, which contradicts our assumption that $\mathcal{T} \prec \mathcal{V}$. □

Given a triangulation $\mathcal{T}$, our next objective is to estimate the cardinality of the minimal stencils $\mathcal{V}$ such that $\mathcal{T} \prec \mathcal{V}$. We begin by counting the ancestors of an irreducible vector.

**Lemma 4.5.** *1. Let $(f, g)$ be an acute basis of $\mathbb{Z}^2$. Then either (i) $f$ is a parent of $g$, (ii) $g$ is a parent of $f$, or (iii) $\|f\| = \|g\| = 1$.*

*2. For any irreducible $e \in \mathbb{Z}^2$ one has $\#(\mathrm{Anc}(e)) \leq \|e\|_\infty + 2$, where $\|(\alpha, \beta)\|_\infty := \max\{|\alpha|, |\beta|\}$.*

*Proof.* Point 1. If $\langle f, g\rangle > 0$, then applying Corollary 2.4 to the non-acute basis $(f, -g)$ we find that either (i) $(f, g - f)$ are the parents of $g$, or (ii) $(f - g, g)$ are the parents of $f$. On the other hand if $\langle f, g\rangle = 0$, then $1 = |\det(f, g)| = \|f\|\|g\|$, hence (iii) $\|f\| = \|g\| = 1$.

Before proving Point 2, we introduce the cone $K$ generated by $(1, 0), (1, 1)$, so that $\|(\alpha, \beta)\|_\infty = \alpha$ for any $(\alpha, \beta) \in K$. If $e \in \mathbb{Z}^2$ irreducible belongs to the interior of $K$, then its parents $f, g \in K$, and we have $\|e\|_\infty = \|f\|_\infty + \|g\|_\infty$.

Point 2 is proved by induction on $\|e\|_\infty$. It is immediate if $\|e\|_\infty = 1$, hence we may assume that $\|e\|_\infty \geq 2$, and denote its parents by $f, g$. We have $\|e\|_\infty = \|f\|_\infty + \|g\|_\infty$, since without loss of generality we may assume that $e \in K$. Applying Point 1 we find that either (i) $\mathrm{Anc}(e) = \mathrm{Anc}(g) \cup \{e\}$, (ii) $\mathrm{Anc}(e) = \mathrm{Anc}(f) \cup \{e\}$, or (iii) $\|f\| = \|g\| = 1$, so that $\|e\|_\infty = 1$, a case which we have excluded. Thus $\#\,\mathrm{Anc}(e) \leq \max\{\#\,\mathrm{Anc}(f), \#\,\mathrm{Anc}(g)\} + 1$, which implies the announced result by induction. □

**Proposition 4.6.** *Let $\mathcal{T}$ be a triangulation, and let $\mathcal{V}$ be the minimal family of stencils such that $\mathcal{T} \prec \mathcal{V}$. Then $\#(\mathcal{V}) \leq 6(N-2)(\mathrm{diam}(\Omega) + 2)$, with $N := \#(X)$. A sharper estimate holds for (standard) Delaunay triangulations: $\#(\mathcal{V}) \leq 6(N-2)$,*

*Proof.* Let $E, F$ be respectively the number of edges and faces of $\mathcal{T}$, where faces refer to both triangles and the infinite exterior face. By Euler's theorem, $N - E + F = 2$. Since each edge is shared by two faces, and each face has at least three edges, one gets $2E \geq 3F$, hence $E \leq 3(N-2)$, and therefore, with the notation $V_x$ of Proposition 4.4,

$$\sum_{x\in X} \#(V_x) = 2E \leq 6(N-2). \tag{19}$$

Combining lemma 4.5, Proposition 4.4, and observing that any edge $[x, x+e]$ of $\mathcal{T}$ satisfies $\|e\|_\infty \leq \mathrm{diam}(\Omega)$, we obtain $\#(\mathcal{V}(x)) \leq \#(V_x)(\mathrm{diam}(\Omega) + 2)$, which in combination with (19) implies the first estimate on $\#(\mathcal{V})$.

In the case of a Delaunay triangulation, we claim that $\mathcal{V}(x) = V_x$. Indeed, consider an edge $[x, x+e]$ of $\mathcal{T}$, and a parent $f \in \mathcal{V}_{\max}(x)$ of $e$. Since $\mathcal{T}$ covers $\mathrm{Hull}(X)$, it contains a triangle $[x, x+e, x+f']$ with $f$ and $f'$ on the same side of the edge $[x, x+e]$. Thus the determinants $\det(e, f)$ and $\det(e, f')$ have the same sign, and therefore the same value since their magnitude is 1. As a result $f' = ke + f$, for some integer $k$. Since $\mathcal{T}$ is Delaunay, the point $x+f$ is outside of the circumcircle of $[x, x+e, x+f']$. This property is equivalent to the non-positivity of the following determinant, called the in-circle predicate: assuming without loss of generality that $\det(e, f) = 1$ so that the vertices $(x, x+e, x+f')$ are in trigonometric order

$$\begin{aligned}
\det \begin{pmatrix} e_1 & e_2 & \|e\|^2 \\ f_1 & f_2 & \|f\|^2 \\ ke_1+f_1 & ke_2+f_1 & \|ke+f\|^2 \end{pmatrix} &= \det \begin{pmatrix} e_1 & e_2 & \|e\|^2 \\ f_1 & f_2 & \|f\|^2 \\ 0 & 0 & \|ke+f\|^2 - k\|e\|^2 - \|f\|^2 \end{pmatrix}, \\
&= \|ke+f\|^2 - k\|e\|^2 - \|f\|^2, \\
&= k(k-1)\|e\|^2 + 2k\langle e, f\rangle, \qquad (20)
\end{aligned}$$

where we denoted $e = (e_1, e_2)$, $f = (f_1, f_2)$. Observing that $0 < \langle e, f\rangle \leq \|e\|\|f\| < \|e\|^2$, we find that (20) is non-positive only for $k = 0$. Thus $f = f'$, hence $f \in V_x$, and therefore $\mathcal{V}(x) = V_x$ as announced. Finally, the announced estimate of $\#(\mathcal{V})$ immediately follows from (19). □

Let us conclude the proof of Proposition 1.10. Let $u \in \mathrm{Conv}(X)$, and let $\mathcal{V}_u$ be the minimal stencils such that $u \in \mathrm{Conv}(\mathcal{V}_u)$. Let $\mathcal{T}$ be an $u$-Delaunay triangulation, and let $\mathcal{V}_\mathcal{T}$ be the minimal stencils such that $\mathcal{T} \prec \mathcal{V}$. By Proposition 4.3 we have $u \in \mathrm{Conv}(\mathcal{T}) \subseteq \mathrm{Conv}(\mathcal{V}_\mathcal{T})$, hence $\mathcal{V}_u \subseteq \mathcal{V}_\mathcal{T}$. Estimating $\#(\mathcal{V}_T)$ with Proposition 4.6, we obtain as announced $\#(\mathcal{V}_u) \leq 6(N-2)(\mathrm{diam}(\Omega)+2)$.

### 4.2 Decomposition of the cone $\mathrm{Conv}(\mathcal{V})$, and edge-flipping distances

We conclude in this section the proof of Theorem 1.13, and establish the complexity result Theorem 1.15 on the edge-flipping generation of $u$-Delaunay triangulations.

**Definition 4.7.** *We say (abusively) that a discrete map $u : X \to \mathbb{R}$ is generic iff, for all $x \in X$ and all $e \in \mathbb{Z}^2$ such that the linear form $P_x^e$ is supported on $X$, one has $P_x^e(u) \neq 0$.*

Generic elements are dense in $\mathrm{Conv}(X)$, since this set is convex, has non-empty interior, and since non-generic elements lie on a union of hyperplanes. The quadratic function $q(x) := \frac{1}{2}\|x\|^2$ is not generic however, since choosing $e = (1, 1)$ one gets $P_x^e(q) = 0$.

**Lemma 4.8.** *Consider stencils $\mathcal{V}$, a generic $u \in \mathrm{Conv}(\mathcal{V})$, and an $u$-Delaunay triangulation $\mathcal{T}$. Then $\mathcal{T} \prec \mathcal{V}$.*

*Proof.* Consider an edge $[x, x+e]$ of $\mathcal{T}$. If the linear form $P_x^e$ is not supported on $\mathcal{T}$, then $e \in \mathcal{V}_{\min}(x) \subseteq \mathcal{V}(x)$ by Proposition 3.7. On the other hand if $P_x^e$ is supported on $X$, then $P_x^e(u) \leq 0$ by Lemma 4.1. By genericity of $u$, we have $P_x^e(u) < 0$, hence $e \in \mathcal{V}(x)$ by Proposition 3.9. This concludes the proof. □

We established in Proposition 4.3 that $\mathrm{Conv}(\mathcal{V}) \supseteq \cup_{\mathcal{T} \prec \mathcal{V}} \mathrm{Conv}(\mathcal{T})$. The next corollary, stating the reverse inclusion, concludes the proof of Theorem 1.13.

**Corollary 4.9.** *If $\mathrm{Conv}(\mathcal{V})$ has a non-empty interior, then $\mathrm{Conv}(\mathcal{V}) \subseteq \cup_{\mathcal{T} \prec \mathcal{V}} \mathrm{Conv}(\mathcal{T})$.*

*Proof.* The set $K := \cup_{\mathcal{T} \prec \mathcal{V}} \mathrm{Conv}(\mathcal{T})$ contains all generic elements of $\mathrm{Conv}(\mathcal{V})$, by Lemma 4.8. Observing that $K$ is closed, and recalling that generic elements are dense in $\mathrm{Conv}(\mathcal{V})$, we obtain the announced inclusion. □

The next lemma characterizes the obstructions to the convexity of the piecewise linear interpolant $\mathrm{I}_{\mathcal{T}} u$ of a convex function $u \in \mathrm{Conv}(X)$ on a triangulation $\mathcal{T}$. See also Figure 4 (right).

**Lemma 4.10.** *Consider $u \in \mathrm{Conv}(X)$, and a triangulation $\mathcal{T}$ which is* not *$u$-Delaunay. Then there exists $x \in X$, and a direct basis $(f, g)$ of $\mathbb{Z}^2$, such that the triangles $[x, x+f, x+g]$ and $[x+f, x+g, x+f+g]$ belong to $\mathcal{T}$, and satisfy $u(x) + u(x+f+g) < u(x+f) + u(x+g)$.*

*Proof.* Since convexity is a local property, there exists two triangles $T, T' \in \mathcal{T}$, sharing an edge, such that the interpolant $\mathrm{I}_{\mathcal{T}} u$ is not convex on $T \cup T'$ (i.e. convexity fails on the edge by $T$ and $T'$). Up to a translation of the domain, we may assume that $T = [0, f, g]$ and $T' = [f, g, e]$, for some $e, f, g \in \mathbb{Z}^2$. The pair $(f, g)$ is a basis of $\mathbb{Z}^2$ because the triangle $T$ contains no point of $\mathbb{Z}^2$ except its vertices; up to exchanging $f$ and $g$ we may assume that it is a direct basis. Up to adding an affine function to $u$, we may assume that $u$ vanishes at the vertices $0, f, g$ of $T$.

If $f$ lies in the triangle $[0, e, g]$, then since $u \in \mathrm{Conv}(X)$, and recalling that $u(0) = u(f) = u(g) = 0$, we obtain $u(e) \geq 0$. This implies that $\mathrm{I}_{\mathcal{T}} u$ is convex on $T \cup T'$, which contradicts our assumption. Likewise $g \notin [0, f, e]$, thus $f \prec e \prec g$ and therefore $\det(f, e) > 0$ and $\det(e, g) > 0$. We next observe that

$$\det(f, g) + \det(g - e, f - e) = \det(f, e) + \det(e, g).$$

The four members of this equation are integers, the two left being equal to $2|T| = 2|T'| = 1$, and the two right being positive. Hence $\det(f, e) = \det(e, g) = 1$, and therefore $e = f + g$ as announced. From this point, the inequality $u(x) + u(x+f+g) < u(x+f) + u(x+g)$ is easily checked to be equivalent to the non-convexity of $\mathrm{I}_{\mathcal{T}} u$ on $T \cup T'$. □

**Proposition 4.11.** *Consider stencils $\mathcal{V}$, a triangulation $\mathcal{T} \prec \mathcal{V}$, and $u \in \mathrm{Conv}(\mathcal{V})$. Define a sequence of triangulations $\mathcal{T}_0 := \mathcal{T}$, $\mathcal{T}_1$, $\mathcal{T}_2 \cdots$ as follows: if $\mathcal{T}_i$ is $u$-Delaunay, then the sequence ends, otherwise $\mathcal{T}_{i+1}$ is obtained by flipping an arbitrary edge of $\mathcal{T}_i$ satisfying Lemma 4.10. Then the sequence is finite, contains at most $\#(\mathcal{V})$ elements, and $\mathcal{T}_i \prec \mathcal{V}$ for all $0 \leq i \leq n$.*

*Proof.* Proof that $\mathcal{T}_i \prec \mathcal{V}$, by induction on $i \geq 0$. Initialization: $\mathcal{T}_0 := \mathcal{T} \prec \mathcal{V}$ by assumption. Induction: adopting the notations of Lemma 4.10, the "flipped" edge $[x+f, x+g]$ of $\mathcal{T}_i$ is replaced with $[x, x+e]$ in $\mathcal{T}_{i+1}$, with $e := f+g$. We only need to check that $e \in \mathcal{V}(x)$, and for that purpose we distinguish two cases. If the basis $(f, g)$ is acute, then $f, g$ are the parents of $e$, and we have $P^e_x(u) < 0$ by Lemma 4.10. This implies $e \in \mathcal{V}(x)$ by Proposition 3.9. On the other hand, if the basis $(f, g)$ is not acute, then by Corollary 2.4 the vector $e$ is a parent of either $f$ or $g$, thus $e \in \mathcal{V}(x)$ by property (Stability) of stencils.

Bound on the number $n$ of edge flips. For all $0 \leq i < n$ one has $\mathrm{I}_{\mathcal{T}_{i+1}} u \leq \mathrm{I}_{\mathcal{T}_i} u$ on $\mathrm{Hull}(X)$, and this inequality is strict at the common midpoint of the flipped edges $[x_i + f_i, x_i + g_i]$ and $[x_i, x_i + e_i]$, with the above conventions. Hence the edge $[x_i, x_i + e_i]$ appears in the triangulation $\mathcal{T}_{i+1}$ but not in any of the $\mathcal{T}_j$, for all $0 \leq j \leq i$. It follows that $i \mapsto (x_i, e_i)$ is injective, and since $e_i \in \mathcal{V}(x_i)$ this implies $n \leq \#(\mathcal{V})$. □

We finally prove Theorem 1.15. Consider a Delaunay triangulation $\mathcal{T}$, and the minimal stencils $\mathcal{V}_{\mathcal{T}}$ such that $\mathcal{T} \prec \mathcal{V}$. Let also $u \in \mathrm{Conv}(X)$, and let $\mathcal{V}_u$ be the minimal stencils such that $u \in \mathrm{Conv}(\mathcal{V}_u)$. Then, by Proposition 4.11, $\mathcal{T}$ can be transformed into an $u$-Delaunay triangulation via $\#(\mathcal{V}_{\mathcal{T}} \cup \mathcal{V}_u)$ edge flips. Furthermore $\#(\mathcal{V}_{\mathcal{T}}) = \mathcal{O}(\#(X))$ by Proposition 4.6

and $\#(\mathcal{V}_u) \geq \#(X)$, as follows e.g. from property (Visibility) of stencils. Thus $\#(\mathcal{V}_\mathcal{T} \cup \mathcal{V}_u) = \mathcal{O}(\#(\mathcal{V}_u))$, and the result follows.

# 5 Average case estimate of the cardinality of minimal stencils

The minimal stencils $\mathcal{V}$, such that the cone $\mathrm{Conv}(\mathcal{V})$ contains a given discrete convex map, admit a simple characterization described in the following proposition.

**Proposition 5.1.** *Let $u \in \mathrm{Conv}(X)$, and let $\mathcal{V}$ be the minimal stencils on $X$ such that $u \in \mathrm{Conv}(\mathcal{V})$. For any $x \in X$, and any irreducible $e \in \mathbb{Z}^2$ with $\|e\| > 1$, one has:*

$$e \in \mathcal{V}(x) \Leftrightarrow (P^e_x \text{ is not supported on } X, \text{ or } P^e_x(u) < 0).$$

*Proof.* Proof of implication $\Leftarrow$. If $P^e_x$ is not supported on $X$, then $e \in \mathcal{V}_{\min}(x) \subseteq \mathcal{V}(x)$ by Proposition 3.7. On the other hand if $P^e_x(u) < 0$, then $e \in \mathcal{V}(x)$ by Proposition 3.9.

Proof of implication $\Rightarrow$. Consider $x \in X$, $e \in \mathcal{V}(x)$, with $\|e\| > 1$, and such that $P^e_x$ is supported on X. Assume for contradiction that $P^e_x(u) \geq 0$, and denote by $f, g$ the parents of $e$. Let $E := \{e' \in \mathcal{V}(x);\ f \prec e' \prec g\}$. The parents of any $e' \in E$ belong to $E \cup \{f, g\}$ by Lemma 3.3 (ii), and one has $P^{e'}_z(u) \geq 0$ by Lemma 4.2. Defining new stencils by $\mathcal{V}'(x) := \mathcal{V}(x) \setminus E$, and $\mathcal{V}'(y) := \mathcal{V}(y)$ for $y \neq x$, we contradict the minimality of $\mathcal{V}$. □

The rest of this section is devoted to the proof of Theorem 1.11, and for that purpose we consider the rotated and translated grids $X^\xi_\theta$, defined in (10). For simplicity, but without loss of generality, we assume a unit grid scale $h := 1$. For each rotation angle $\theta \in \mathbb{R}$, and each offset $\xi \in \mathbb{R}^2$, we introduce an affine transform $A^\xi_\theta$: for all $x \in \mathbb{R}^2$

$$A^\xi_\theta(x) := R_\theta(\xi + x).$$

For any set $E \subseteq \mathbb{R}^2$, and any affine transform $A$, we denote $A(E) := \{A(e);\ e \in E\}$. For instance, the displaced grids (10) are given by $X^\xi_\theta := \Omega \cap A^\xi_\theta(\mathbb{Z}^2)$.

The maximal stencils on the grid $X^\xi_\theta$ are defined by: for all $x \in X^\xi_\theta$

$$\mathcal{V}^{\theta,\xi}_{\max}(x) := \{e \in \mathbb{Z}^2 \text{ irreducible};\ x + R_\theta e \in X^\xi_\theta\}.$$

A family $\mathcal{V}^\xi_\theta$ of stencils on $X^\xi_\theta$ is a collection of sets $\mathcal{V}^\xi_\theta(x) \subseteq \mathcal{V}^{\theta,\xi}_{\max}(x)$, $x \in X^\xi_\theta$ which satisfies the usual (Stability) and (Visibility) properties of Definition 1.6 (replacing, obviously, instances of $\mathcal{V}_{\max}$ with $\mathcal{V}^{\theta,\xi}_{\max}$). For $x \in X^\xi_\theta$, and $e \in \mathcal{V}^{\theta,\xi}_{\max}$ we consider the linear forms $S^e_{x,\theta}(u) := u(x + R_\theta e) - 2u(x) + u(x - R_\theta e)$, and likewise $T^e_{x,\theta}$, $P^e_{x,\theta}$, which are used to define cones $\mathrm{Conv}(\mathcal{V}^\xi_\theta) \subseteq \mathrm{Conv}(X^\xi_\theta)$. In a nutshell, when embedding a stencil element $e \in \mathcal{V}^\xi_\theta(x) \subseteq \mathbb{Z}^2$, where $x \in X^\xi_\theta$, into the physical domain $\Omega$ (e.g. considering $x + R_\theta e \in X^\xi_\theta$), one should never forget to apply the rotation $R_\theta$.

Consistently with the notations of Theorem 1.11, we consider a fixed convex map $U \in \mathrm{Conv}(\Omega)$, and study the smallest stencils $\mathcal{V}^\xi_\theta \subseteq \mathbb{Z}^2$ on $X^\xi_\theta$ such that the restriction of $U$ to $X^\xi_\theta$ belongs to $\mathrm{Conv}(\mathcal{V}^\xi_\theta)$. The *midpoints* $m = x + R_\theta e/2$ of "stencil edges" $[x, x + R_\theta e]$, $x \in X^\xi_\theta$, $e \in \mathcal{V}^\xi_\theta(x)$, play a central role in our proof.

**Definition 5.2.** *For any $m \in \Omega$, and any irreducible $e \in \mathbb{Z}^2$, let*

$$\Lambda^e_m := \{(\theta, \xi) \in [0, \pi/2[ \times [0, 1[^2;\ m = x + R_\theta e/2,\ \text{for some } x \in X^\xi_\theta \text{ such that } e \in \mathcal{V}^\xi_\theta(x)\}.$$

We introduce offsetted grids, of points with half-integer coordinates

$$\mathcal{Z} := (\tfrac{1}{2}, 0) + \mathbb{Z}^2, \qquad \mathcal{Z}' := (0, \tfrac{1}{2}) + \mathbb{Z}^2, \qquad \mathcal{Z}'' := (\tfrac{1}{2}, \tfrac{1}{2}) + \mathbb{Z}^2.$$

For any $x, y \in \mathbb{Z}^2$ with $x - y$ irreducible, the midpoint $(x+y)/2$ of the segment $[x, y]$ belongs to the disjoint union $\mathcal{Z} \sqcup \mathcal{Z}' \sqcup \mathcal{Z}''$.

**Lemma 5.3.** *For any $m \in \Omega$ and any $(\theta, \xi) \in \Lambda_m^e$, one has $m \in A_\theta^\xi(\mathcal{Z} \sqcup \mathcal{Z}' \sqcup \mathcal{Z}'')$.*

*Proof.* Let $x \in X_\theta^\xi$, and $e \in \mathcal{V}_\theta^\xi(x)$, be such that $m = x + R_\theta e/2$. Observing that the coordinates of $e$ are not both even, since $e$ is irreducible, we obtain $e/2 \in \mathcal{Z} \sqcup \mathcal{Z}' \sqcup \mathcal{Z}''$. Adding $R_\theta(e/2)$ to $x \in A_\theta^\xi(\mathbb{Z}^2)$ yields as announced a point $m \in A_\theta^\xi(\mathcal{Z} \sqcup \mathcal{Z}' \sqcup \mathcal{Z}'')$. □

For any point $m \in \mathbb{R}^2$, and any angle $\theta$, there exists exactly one offset $\xi \in [0, 1[^2$ such that $m \in A_\theta^\xi(\mathcal{Z})$; and likewise for $\mathcal{Z}'$, $\mathcal{Z}''$. Hence the set $\Lambda_m^e$ contains redundant information, which motivates the following definition: for any $m \in \Omega$, and any irreducible $e \in \mathbb{Z}^2$

$$\Theta_m^e := \{\theta \in [0, \pi/2[;\ \exists \xi \in [0, 1[^2,\ (\theta, \xi) \in \Lambda_m^e \text{ and } m \in A_\theta^\xi(\mathcal{Z})\}. \tag{21}$$

and similarly we define $\Theta_m'^e$, $\Theta_m''^e$, by replacing $\mathcal{Z}$ with $\mathcal{Z}'$, $\mathcal{Z}''$ respectively in (21). By convention, $\Theta_m^e = \Theta_m'^e = \Theta_m''^e = \emptyset$ for *non* irreducible vectors $e \in \mathbb{Z}^2$. The following lemma accounts in analytical terms for a simple combinatorial identity: one can count stencil edges by looking at their endpoints or their midpoints.

**Lemma 5.4.** *The following integrals are equal:*

$$\int_{[0,1]^2} \int_0^{\frac{\pi}{2}} \#(\mathcal{V}_\theta^\xi)\, d\theta d\xi = \sum_{e \in \mathbb{Z}^2} \int_{m \in \Omega} (|\Theta_m^e| + |\Theta_m'^e| + |\Theta_m''^e|)\, dm, \tag{22}$$

*where $|\Theta|$ denotes the Lebesgue measure of a Borel set $\Theta \subseteq \mathbb{R}$.*

*Proof.* Consider $m \in \Omega$, $e \in \mathbb{Z}^2$, and $\theta \in \Theta_m^e$. Then there exists a unique $\xi \in [0, 1[^2$ such that $m \in A_\theta^\xi(\mathcal{Z})$. This uniquely determines the point $x := m - \frac{1}{2} R_\theta e \in X_\theta^\xi$ such that $e \in \mathcal{V}_\theta^\xi(x)$. Likewise for $\Theta_m'^e$, $\Theta_m''^e$. Conversely, the data of $\theta$, $\xi$, $x \in X_\theta^\xi$ and $e \in \mathcal{V}_\theta^\xi(x)$ uniquely determines $m := x + R_\theta e/2$, and also by Lemma 5.3 a unique set among $\Theta_m^e$, $\Theta_m'^e$, $\Theta_m''^e$ containing $\theta$. As a result the left and right hand side of (22) are just two different expressions of the measure of

$$\{(m, e, i, \theta);\ \theta \in \Theta_m^{(i)e}\} \subseteq \Omega \times \mathbb{Z}^2 \times \{0, 1, 2\} \times [0, \pi/2[,$$

where $\Theta_m^{(0)e} := \Theta_m^e$, $\Theta_m^{(1)e} := \Theta_m'^e$, and $\Theta_m^{(2)e} := \Theta_m''^e$. Implicitly, we equipped $\mathbb{Z}^2$ and $\{0, 1, 2\}$ with the counting measure, and $[0, \pi/2[$ and $\Omega$ with the Lebesgue measure (which in the latter case is preserved by the rotations $R_\theta$). □

In order to estimate (22), we bound in the next lemma the size of the sets $\Theta_m^e$, $\Theta_m'^e$, $\Theta_m''^e$.

**Lemma 5.5.** *Let $e \in \mathbb{Z}^2$ be irreducible, with $\|e\| > 1$, of parents $f, g$. Let also $m \in \Omega$. Then for any $\theta, \varphi \in \Theta_m^e$, one has $\sin|\theta - \varphi| \leq 2/\min\{\langle e, f\rangle, \langle e, g\rangle\}$. Likewise for $\Theta_m'^e$, $\Theta_m''^e$.*

*Proof.* Without loss of generality, we may assume that $m$ is the origin of $\mathbb{R}^2$. Let $Q$ be the parallelogram of vertices $\{\pm R_\theta e, \pm R_\varphi e\}$; note that $\frac{1}{2} Q \subseteq \Omega$. A point $x \in \mathbb{R}^2$ belongs to $Q$ iff

$$|\det(x, R_\theta e \pm R_\varphi e)| \leq |\det(R_\theta e, R_\varphi e)| = \|e\|^2 \sin|\theta - \varphi|. \tag{23}$$

Indeed $\sin|\theta-\varphi| = |\sin(\theta-\varphi)|$, since $\theta,\varphi \in [0,\pi/2[$ by construction. We assume without loss of generality that $\langle e,f\rangle \leq \langle e,g\rangle$. Introducing $h := e-2f = g-f$ we observe that $\langle e,h\rangle \geq 0$, and compute

$$|\det(R_\theta h, R_\theta e)| = |\det(h,e)| = |\det(e-2f,e)| = 2.$$
$$|\det(R_\theta h, R_\varphi e)| \leq |\det(h,e)|\cos(\varphi-\theta) + |\langle h,e\rangle|\sin|\varphi-\theta| \leq 2 + (\|e\|^2 - 2\langle e,f\rangle)\sin|\varphi-\theta|.$$

In the second line, we used the identity $\sin(a+b) = \sin(a)\cos(b)+\cos(a)\sin(b)$, where $a$ denotes the angle between $e$ and $h$, and $b := \varphi-\theta$. Combining these two estimates with (23), and assuming for contradiction that $\sin|\theta-\varphi| \geq 2/\langle e,f\rangle$, we obtain that $R_\theta h \in Q$. By symmetry, $-R_\theta h \in Q$, and likewise $\pm R_\varphi h \in Q$.

In the following, we denote $x := -R_\theta e/2$, $y := -R_\varphi e/2$, $p := \pm R_\theta h/2$, $q := \pm R_\varphi h/2$, where the signs for $p$ and $q$ are chosen so that $p,q \in [x,-x,y]$. Denoting by $\alpha,\beta,\gamma$ (resp. $\alpha',\beta',\gamma'$) the barycentric coordinates of $p$ (resp. $q$) in this triangle, convexity implies

$$U(p) \leq \alpha U(x) + \beta U(-x) + \gamma U(y), \tag{24}$$
$$U(q) \leq \alpha' U(x) + \beta' U(-x) + \gamma' U(y). \tag{25}$$

Let $\xi \in [0,1[^2$ be such that $m \in A^\xi_\theta(\mathcal{Z})$. Then $x \in X^\xi_\theta$, $e \in \mathcal{V}^\xi_\theta(x)$, and $m = x + R_\theta e/2$ (recall that we fixed $m=0$). Using the characterization of minimal stencils of Proposition 5.1, we obtain

$$0 > P^e_{x,\theta}(U) = U(x) - U(x+R_\theta f) - U(x+R_\theta g) + U(x+R_\theta e), \tag{26}$$

provided this linear form is supported on $X^\xi_\theta$. Note that $x+R_\theta e = -x \in X^\xi_\theta$, that $x+R_\theta f = \varepsilon p$, and that $x+R_\theta g = -\varepsilon p$ for some $\varepsilon \in \{-1,1\}$. Since $\pm p \in \frac{1}{2}Q \subseteq \Omega$ this confirms that (26) is supported on $X^\xi_\theta := \Omega \cap A^\xi_\theta(\mathbb{Z}^2)$. Inserting in (26) the values of $-x, p, -p$, and proceeding likewise for $y$ and $q$, we obtain

$$U(x) + U(-x) < U(p) + U(-p) \tag{27}$$
$$U(y) + U(-y) < U(q) + U(-q). \tag{28}$$

Up to adding an affine map to $U$, we may assume that $U(x) = U(-x) = 0$, and $U(p) = U(-p)$. From (27) we obtain $U(p) > 0$. Hence also $U(y) > 0$ using (24), and therefore $U(q) \leq \gamma' U(y) \leq U(y)$ using (25). Likewise $U(-q) \leq U(-y)$, which contradicts (28) and concludes the proof. □

We finally conclude the proof of Theorem 1.11, by combining (22) with the next lemma.

**Lemma 5.6.** *For any $m \in \Omega$, with $r := \max\{1, \mathrm{diam}(\Omega)\}$: (likewise for $\Theta'^e_m$, $\Theta''^e_m$)*

$$\sum_{e\in\mathbb{Z}^2} |\Theta^e_m| \leq 2\pi + 4\pi^2(1+\ln r)^2. \tag{29}$$

*Proof.* Note that $\Theta^e_m \subseteq [0,\pi/2[$ for any $e \in \mathbb{Z}^2$, and that $\Theta^e_m = \emptyset$ if the $e$ is not irreducible, or if $\|e\| > r$. For any two vectors $e,e' \in \mathbb{Z}^2$, we write $e' \lhd e$ iff $e'$ is a parent of $e$ (which implies that $e,e'$ are irreducible, and that $\|e\| > 1$). Isolating the contributions to (29) of the four unit vectors, and applying Lemma 5.5 to other vectors, we thus obtain

$$\sum_{e\in\mathbb{Z}^2} |\Theta^e_m| \leq 4\times\frac{\pi}{2} + \sum_{\|e\|\leq r}\sum_{e'\lhd e} \arcsin\left(\frac{2}{\langle e,e'\rangle}\right) \leq 2\pi + \sum_{e'\in\mathbb{Z}^2}\sum_{\substack{e\,\rhd e'\\ \|e\|\leq r}} \frac{\pi}{\langle e,e'\rangle}, \tag{30}$$

where we used the concavity bound $\arcsin(x) \leq \frac{\pi}{2}x$ for all $x \in [0,1]$ (and slightly abused notations for arguments of arcsin larger than 1).

Consider a fixed irreducible $e' \in \mathbb{Z}^2$, and denote by $f, g$ its parents if $\|e'\| > 1$, or the two orthogonal unit vectors if $\|e'\| = 1$, so that $\det(f, e') = 1 = \det(e', g)$. If $e \in \mathbb{Z}^2$ is such that $e' \lhd e$, then $|\det(e, e')| = 1$; assuming $\det(e, e') = 1$ (resp. $-1$) we obtain that $e = f + ke'$ (resp. $e = g + ke'$) for some scalar $k$ which must be (i) an integer since $e'$ is irreducible, (ii) non-negative since $\langle e, e' \rangle \geq 0$, and (iii) positive since $\|e'\| < \|e\|$. Assuming $\|e\| \leq r$, we obtain in addition $k\|e'\| \leq r$, thus $k \leq r$ and $\|e'\| \leq r$. As a result

$$\sum_{\substack{e \rhd e' \\ \|e\| \leq r}} \frac{1}{\langle e, e' \rangle} \leq \sum_{1 \leq k \leq r} \left( \frac{1}{\langle f + ke', e' \rangle} + \frac{1}{\langle g + ke', e' \rangle} \right) \leq \frac{2}{\|e'\|^2} \sum_{1 \leq k \leq r} \frac{1}{k}. \tag{31}$$

Inserting (31) into (30) yields the product of the two following sums, which are easily bounded via comparisons with integrals: isolating the terms for $k = 1$, and for all $\|e'\| \leq \sqrt{2}$

$$\sum_{1 \leq k \leq r} \frac{1}{k} \leq 1 + \int_1^r \frac{dt}{t} = 1 + \ln r, \qquad \sum_{\substack{0 < \|e'\| \leq r \\ \text{irreducible}}} \frac{1}{\|e'\|^2} \leq 4 + 4 \times \frac{1}{2} + \int_{1 \leq \|x\| \leq r} \frac{dx}{\|x\|^2} \leq 6 + 2\pi \ln r.$$

Noticing that $2\pi \geq 6$ we obtain (29) as announced. □

# 6 Numerical experiments

Our numerical experiments cover the classical formulation [24] of the monopolist problem, as well as several variants, including lotteries [17, 25], or the pricing of risky assets [4]. We choose to emphasize this application in view of its appealing economical interpretation, and the often surprising qualitative behavior. Our algorithm can also be applied in a straightforward manner to the computation of projections onto the cone of convex functions defined on some square domain, with respect to various norms as considered in [5, 18, 22, 21] (this amounts to denoising under a convexity prior). It may however not be perfectly adequate for the investigation of geometric conjectures [16, 26, 18], due to the use of a grid discretization.

The hierarchical cones of discrete convex functions introduced in this paper are combined with a simple yet adaptive and anisotropic stencil refinement strategy, described in §6.1. The monopolist model is introduced in §6.2, and illustrated with numerous experiments. We compare in §6.3 our implementation of the constraint of convexity, with alternative methods proposed in the literature, in terms of computation time, memory usage, and solution quality.

## 6.1 Stencil refinement strategy

We introduce two algorithms which purpose is to minimize a given lower semi-continuous proper convex functional $\mathcal{E} : \mathcal{F}(X) \to \mathbb{R} \cup \{+\infty\}$, on the $N$-dimensional cone $\mathrm{Conv}(X)$, $N := \#(X)$, without ever listing the $\mathcal{O}(N^2)$ linear constraints which characterize this cone. They both generate an increasing sequence of stencils $\mathcal{V}_0 \subsetneq \mathcal{V}_1 \subsetneq \cdots \subsetneq \mathcal{V}_n$ on $X$, and minimizers $(u_i)_{i=0}^n$ of $\mathcal{E}$ on cones defined by $\mathcal{O}(\#(\mathcal{V}_i))$ linear constraints. The subscript $i$ refers to the loop iteration count in Algorithms 1 and 2, and the loop ends when the stencils are detected to stabilize: $\mathcal{V}_n = \mathcal{V}_{n+1}$. The final map $u_n$ is guaranteed to be the global minimum of $\mathcal{E}$ on $\mathrm{Conv}(X)$.

Our first algorithm is based on an increasing sequence $\mathrm{Conv}(\mathcal{V}_0) \subseteq \cdots \subseteq \mathrm{Conv}(\mathcal{V}_n)$ of subcones of $\mathrm{Conv}(X)$. If constraints of type $P_x^e$, $x \in X$, $e \in \mathcal{V}_i(x)$ are *active* for the minimizer $u_i$

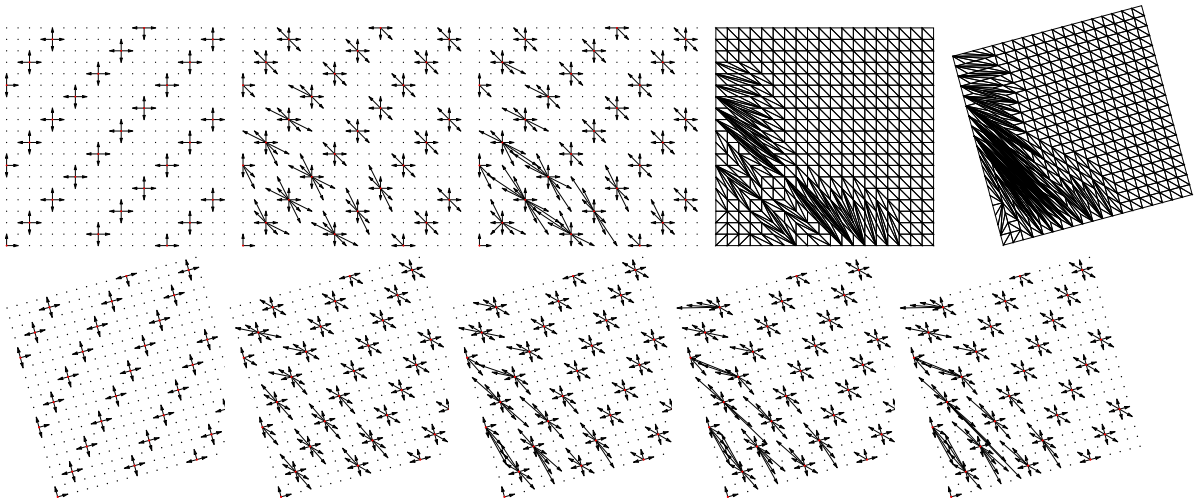


Figure 5: Top: Algorithm 2, for the classical monopolist problem (1) on $[1,2]^2$ with a $20\times20$ grid, converges in 2 stencils refinement steps (using the extended candidates $\hat{\mathcal{V}}_\rho$, $\rho := 1.5$). Bottom: 4 refinement steps are needed with a different density of customers, uniform on the square $[1,2]^2$ rotated by $\pi/12$. Top right: $u$-Delaunay triangulations associated with the respective discrete solutions, for illustration of Theorem 1.13.

of $\mathcal{E}$ on $\mathrm{Conv}(\mathcal{V}_i)$ (i.e. the corresponding Lagrange multipliers are positive), then refined stencils $\mathcal{V}_{i+1}$ are adaptively generated from $\mathcal{V}_i$; otherwise $u_i$ is the global minimizer of $\mathcal{E}$ on $\mathrm{Conv}(X)$, and the method ends. Note that the optimization of $\mathcal{E}$ on $\mathrm{Conv}(\mathcal{V}_{i+1})$ can be hot-started from the previous minimizer $u_i \in \mathrm{Conv}(\mathcal{V}_i) \subseteq \mathrm{Conv}(\mathcal{V}_{i+1})$.

**Algorithm 1** Sub-cones approach to stencil refinement

Start with the minimal stencils: $\mathcal{V} \leftarrow \mathcal{V}_{\min}$. (See Proposition 3.7)
**Until** the stencils $\mathcal{V}$ stabilize
  **Find** a minimizer $u$ of the energy $\mathcal{E}$ on $\mathrm{Conv}(\mathcal{V})$,
    and extract the Lagrange multipliers $\lambda$ associated to the constraints $P^e_x$, $x \in X$, $e \in \hat{\mathcal{V}}(x)$.
  **Set** $\mathcal{V}(x) \leftarrow \mathcal{V}(x) \cup \{e \in \hat{\mathcal{V}}(x);\ \lambda(P^e_x) > 0\}$, for all $x \in X$.

**Definition 6.1.** *For any family $\mathcal{V}$ of stencils on $X$, we denote by* $\mathrm{Conv}'(\mathcal{V}) \subseteq \mathcal{F}(X)$ *the cone defined by the non-negativity of: for all $x \in X$, and all $e \in \mathcal{V}(x)$, the linear forms $S^e_x$ and $T^e_x$ if $\|e\| > 1$, provided they are supported on $X$. Note that* $\mathrm{Conv}(\mathcal{V}) \subseteq \mathrm{Conv}(X) \subseteq \mathrm{Conv}'(\mathcal{V})$.

Algorithm 2 is based on a *decreasing* sequence $\mathrm{Conv}'(\mathcal{V}_0) \supseteq \cdots \supseteq \mathrm{Conv}'(\mathcal{V}_n)$ of super-cones of $\mathrm{Conv}(X)$. The minimizer $u_i$ of $\mathcal{E}$ on $\mathrm{Conv}'(\mathcal{V}_i)$ may not belong to $\mathrm{Conv}(X)$, even less to $\mathrm{Conv}(\mathcal{V}_i)$, and in particular some of the values $P^e_x(u_i)$, $x \in X$, $e \in \hat{\mathcal{V}}_i(x)$, may be negative. In that case, refined stencils $\mathcal{V}_{i+1}$ are adaptively generated from $\mathcal{V}_i$; otherwise, $u_i$ is the global minimizer of $\mathcal{E}$ on $\mathrm{Conv}(X)$, and the method ends.

Algorithms 1 and 2 are provided "as is", without any complexity guarantee. Our numerical experiments are based on algorithm 2, because the numerical test "$P^e_x(u) < 0$" turned out to be more robust than "$\lambda(P^e_x) > 0$". In order to limit the number $n$ of stencil refinement steps, we use a slightly extended set $\hat{\mathcal{V}}_\rho(x)$ of candidates for refinement, with $\rho := 1.5$, see Definition 6.2 below (note that $\hat{\mathcal{V}}_1 = \hat{\mathcal{V}}$). With this modification, $n$ remained below 10 in all our experiments. The constructed stencils were generally sparse, highly anisotropic, and almost minimal for the

**Algorithm 2** Super-cones approach to stencil refinement

Start with the minimal stencils $\mathcal{V}$.
**Until** the stencils $\mathcal{V}$ stabilize
  **Find** a minimizer $u$ of the energy $\mathcal{E}$ on $\mathrm{Conv}'(\mathcal{V})$.
  **Set** $\mathcal{V}(z) \leftarrow \mathcal{V}(z) \cup \{e \in \hat{\mathcal{V}}(z);\ P_z^e(u) < 0\}$, for all $z \in X$.

discrete problem solution $u \in \mathrm{Conv}(X)$ eventually found, see Figure 5. Observation of Figure 9 suggests that $n$ grows logarithmically with the problem dimension $N$, and that the final stencils cardinality $\#(\mathcal{V}_n)$ depends quasi-linearly on $N$, as could be expected in view of Remark 1.9 and Theorem 1.11. However, we could not mathematically establish such complexity estimates.

**Definition 6.2** (Extended candidates)**.** *Let $\mathcal{V}$ be a family of stencils, let $\rho \geq 1$, and let $x \in X$. A vector $e \in \mathcal{V}_{\max}(x) \setminus \mathcal{V}(x)$, of parents $f, g$, belongs to the extended candidates $\hat{\mathcal{V}}_\rho(x)$ iff there exists trigonometrically consecutive $f', g' \in \mathcal{V}(x)$ such that $f' \preceq f \prec g \preceq g'$ and $\|f\|\|g\| \leq \rho \|f'\|\|g'\|$.*

## 6.2 The monopolist problem

A monopolist has the ability to produce goods, which have two characteristics and may thus be represented by a point $q \in \mathbb{R}^2$. The manufacturing cost $\mathrm{Cost}(q) : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ is known and fixed a-priori. Infinite costs account for products which are "meaningless", or impossible to build. The selling price $\pi : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ is fixed unilaterally by the monopolist except for the "null" product $(0, 0)$, which must be available for free ($\pi(0, 0) \leq 0$). The characteristics of the consumers are also represented by a point $z \in \mathbb{R}^2$, and the utility of product $q$ to consumer $z$ is modeled by the scalar product between their characteristics

$$\mathcal{U}(q, z) := \langle q, z \rangle. \tag{32}$$

More general utility pairings $\mathcal{U}$ are considered in [12], yet the numerical implementation of the resulting optimization problems remains out of reach, see [18] for a discussion. All consumers $z$ are rational, "screen" the proposed price catalog $\pi$, and choose the product of maximal *net* utility $\langle q, z \rangle - \pi(q)$ (i.e. raw utility minus price). Introducing the Legendre-Fenchel dual $U$ of the prices $\pi$: for all $z \in \mathbb{R}^2$

$$U(z) := \pi^*(z) := \sup_{q \in \mathbb{R}^2} \langle q, z \rangle - \pi(z),$$

we observe that the optimal product[3] for consumer $z$ is $\nabla U(z)$, which is sold at the price

$$\pi(\nabla U(z)) = \langle \nabla U(z), z \rangle - U(z). \tag{33}$$

The net utility function $U$ is convex and non-negative by construction, and uniquely determines the products bought and their prices. Conversely, any convex non-negative $U$ defines prices $\pi = U^*$ satisfying the admissibility condition $\pi(0) \leq 0$, and such that $\pi^* = U^{**} = U$. The distribution of the characteristics of the potential customers is known to the monopolist, under the form of a bounded measure $\mu$ on $\mathbb{R}^2$. He aims to maximize his total profit: the integrated difference (sales margin) between the selling price (33), and the production cost

$$\sup \left\{ \int_{\mathbb{R}^2} \Big( \langle \nabla U(z), z \rangle - U(z) - \mathrm{Cost}(\nabla U(z)) \Big) d\mu(z);\ U \in \mathrm{Conv}(\mathbb{R}^2),\ U \geq 0 \right\} \tag{34}$$

[3]Strictly speaking, the optimal product $Q(z)$ is an element of the subgradient $\partial_z U$, which (Lebesgue-)almost surely is a singleton $\{\nabla U(z)\}$. Hence we may write (34) in terms of $\nabla U(z)$, provided the density $\mu$ of customers is absolutely continuous with respect to the Lebesgue measure.

If production costs are convex, then this amounts to maximizing a concave functional of $U$ under convex constraints; see [3] for precise existence results. If $U$ maximizes (34), then an optimal catalog of prices is given by $U^*$. Quantities of particular economic interest are the monopolist margin, and the distribution of product sales:

$$\text{Margin} = U^* - \text{Cost}, \qquad \text{SalesDistribution} = (\nabla U)_{\#}\mu, \tag{35}$$

where # denotes the push forward operator on measures. The regions defined by $\{U = 0\}$ and $\{\det(\text{Hessian}\, u) = 0\}$ are also important, as they correspond to different categories of customers, see below. We present numerical results for three instances of the monopolist problem, associated to different product costs. These three models are clearly simplistic idealizations of real economy. Their interest lies in their, striking, qualitative properties, which are stable and are expected to transfer to more complex models.

For implementation purposes, we observe that the maximum profit (34) is unchanged if one considers $U$ only defined on a convex set $K \supseteq \text{supp}(\mu)$, and imposes the additional constraint $\text{Cost}(\nabla U(x)) < \infty$ for all $x \in K$. The chosen discrete domain is a square grid $X$, such that $\text{supp}(\mu) \subseteq \text{Hull}(X)$. This density is represented by non-negative weights $(\mu_x)_{x\in X}$, set to zero outside $\text{supp}(\mu)$. The integral appearing in (34) is discretized using finite differences, see [5] for convergence results. The resulting convex program is solved by combining Mosek software's interior point (for linear problems) or conic (for quadratic[4] problems) optimizer, with the stencil refinement strategy of Algorithm 2, §6.1.

**Classical model.** The produced goods are cars (for concreteness), which characteristics $q = (q_1, q_2)$ are non-negative and account for the engine horsepower $q_1$ and the upholstery quality $q_2$. Production cost is quadratic: $\text{Cost}(q) := \frac{1}{2}\|q\|^2$ for all $q \in \mathbb{R}^2_+$, and $\text{Cost}(q) = +\infty$ otherwise (cars with negative characteristics are unfeasible). Consumer characteristics $x = (x_1, x_2)$ are their appetite $x_1$ for car performance, and $x_2$ for comfort, consistently with (32). The qualitative properties of this model are the following [24]: denoting by $U$ a solution of (34), and ignoring regularity issues in this heuristic discussion

- (Desirability of exclusion) The optimal monopolist strategy often involves neglecting a positive proportion of potential customers - which "buy" the null product 0 at price 0. In other words, the solution of (34) satisfies $U = 0$ on an open subset of $\text{supp}(\mu)$, hence also $\nabla U = 0$. The economical interpretation is that introducing (low end) products attractive to this population would reduce overall profit, because other customers currently buying expensive, high margin products, would change their minds and buy these instead.

- (Bunching) "Wealthy" customers $z$ generally buy products which are specifically designed for them, in the sense that $\nabla U$ is a local diffeomorphism close to $z$. "Poor" potential customers are excluded from the trade: $U = 0$ close to $z$, see the previous point. There also exists an intermediate category of customers characterized by $\det(\text{Hessian}\, U) = 0$ close to $z$, so that the same product $q = \nabla U(z)$ is bought by a one dimensional "bunch" of customers $(\nabla U)^{-1}\{q\}$. The image of this category of customers, by $\nabla U$, is one dimensional product line. From an economic point of view, the optimal strategy limits the variety of intermediate range products in order, again, to avoid competing with high margin sales.

[4]Following the indications of Mosek's user manual, quadratic functionals are implemented under the form of linear functionals involving auxiliary variables subject to conic constraints.

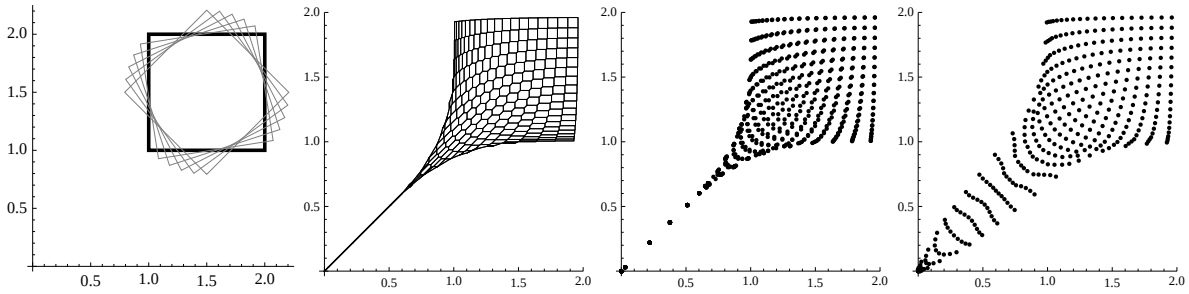


Figure 6: Left: domain $[1,2]^2$ (thick black), and rotated domains used in numerical experiments for the classical principal agent problem, see Figure 11. We computed a minimizer $u \in \mathrm{Conv}(X)$ of a discretization of the classical monopolist problem on $[1,2]^2$, see (1), on a $20 \times 20$ grid $X$, and an $u$-Delaunay triangulation $\mathcal{T}$, see Figure 4. Center left: subgradients cells $\partial_x U$, $x \in X$, with $U := \mathrm{I}_{\mathcal{T}}\, u$. Center right: the gradients $\nabla\, \mathrm{I}_T\, u$, $T \in \mathcal{T}$ (vertices of the previous cells). Right: the less precise numerical method $\mathrm{OF}_3$, see 6.3, thickens the product line and hides the bunching phenomenon.

Considering, as in [24, 18], a uniform density of customers on the square $[1,2]^2$, we illustrate[5] on Figure 1 the estimated solution $U$ (left), $\det(\mathrm{Hessian}\, U)$ (center left), the sales distribution (center right) and the monopolist margin (right), see (35). The phenomena of exclusion $U = 0$ and of bunching $\det(\mathrm{Hessian}\, U) = 0$ are visible (center left subfigure) as a white triangle and as the darkest level set of $\det(\mathrm{Hessian}\, U)$ respectively. The image by $\nabla U$ of customers subject to bunching appears (center right subfigure) as a one dimensional red structure in the product sales distribution.

We also consider variants where the density of customers is uniform on the square $[1,2]^2$ *rotated* by an angle $\theta \in [0, \pi/4]$ around its center, see Figures 6 (left) and 11. Our experiments suggest that exclusion occurs iff $\theta \in [0, \theta_0]$, with $\theta_0 \approx 0.47$ rad. Bunching is always present, yet two regimes can be distinguished: the one dimensional product line, associated to the bunching phenomenon, is included in the boundary of the two dimensional one iff $\theta \in [\theta_1, \pi/4]$, with $\theta_0 < \theta_1 \approx 0.55$ rad. Proving mathematically this qualitative behavior is an open problem.

**Remark 6.3** (Subgradient measure)**.** *Studying the "bunching" phenomenon requires to estimate the hessian determinant* $\det(\mathrm{Hessian}\, U) \geq 0$ *of the solution* $U$ *of* (34)*, and to visualize the degenerate region* $\det(\mathrm{Hessian}\, U) = 0$*. The hessian determinant also appears in the density of product sales* (35)*. These features need to be extracted from a minimizer* $u \in \mathrm{Conv}(X)$ *of a finite differences discretization of* (34)*, which is a delicate problem since (i) the hessian determinant is a "high order" quantity, and (ii) equality to zero is a numerically unstable test. Naïvely computing a discrete hessian* $H_u$ *via second order finite differences, we obtain an oscillating, non-positive and overall imprecise approximation* $\det(H_u)$*, see Figure 10 (right).*

*The following approach gave better results, see Figure 10 (left): compute the largest convex* $U : \mathrm{Hull}(X) \to \mathbb{R}$ *such that* $U \leq u$ *on* $X$ *(if* $u \in \mathrm{Conv}(X)$*, then* $U = \mathrm{I}_{\mathcal{T}}\, u$ *for any* $u$*-Delaunay triangulation). Then for all* $x \in X \setminus \partial\, \mathrm{Hull}(X)$

$$h^2 \det(\mathrm{Hessian}\, U(x)) \approx |\{\nabla U(x+e);\ \|e\|_\infty \leq h/2\}| \approx |\{\nabla \hat{U}(x+e);\ \|e\|_\infty \leq h/2\}| = |\partial_x \hat{U}|,$$

*where* $\partial$ *denotes the sub-gradient, set-valued operator on convex functions, and* $|\cdot|$ *the two dimensional Lebesgue measure. The sub-gradient sets* $\partial_x \hat{U}$ *are illustrated on Figure 6.*

[5] With this customer density, [24] expected the bunching region to be triangular, and the image $\nabla U$ to be the union of the segment $[(0,0),(1,1)]$ and of the square $[1,2]^2$. After discussion with the author, and in view of the numerical experiments, we believe that these predictions are erroneous.

**Product bundles and lottery tickets.** Two types of products $P_1$, $P_2$ are considered, which the consumer of characteristics $x = (x_1, x_2) \in \mathbb{R}^2$ respectively values $x_1$ and $x_2$. The two products are indivisible, and consumers are not interested in buying more than one of each. The monopolist sells them in bundles $q = (q_1, q_2) \in \{0,1\}^2$ which characteristics are the presence ($q_i = 1$) of product $P_i$, or its absence ($q_i = 0$), for $i \in \{1,2\}$. In order to maximize profit, the monopolist also considers probabilistic bundles, or lottery tickets, $q \in [0,1]^2$ for which the product $P_i$ has the probability $q_i$ of being present. This is consistent with (32), provided consumers are risk neutral. Production costs are neglected, so that $\mathrm{Cost}(q) = 0$ if $q \in [0,1]^2$ and $\mathrm{Cost}(q) = +\infty$ otherwise. Three different customer densities were considered, see below. The qualitative property of interest is the presence, or not, of probabilistic bundles in the monopolist's optimal strategy.

- Uniform customer density on $[0,1]^2$. We recover the known exact minimizer [17]:

  $$U(x,y) := \max\{0,\, x-a,\, y-a,\, x+y-b\}, \text{ with } a := 2/3, \text{ and } b := (4-\sqrt{3})/2,$$

  up to numerical accuracy, see Figure 7 (left). This optimal strategy does *not* involve lottery tickets: $\nabla U(x,y) \in \{0,1\}^2$, wherever this gradient is defined. The uselessness of lottery tickets is known for similar 1D problems and was thought to extend to higher dimension, until the following two counter-examples were independently found [17, 25].

- Uniform customer density on the triangle $T := \{(x,y) \in [0,1]^2;\, x + y/2 \geq 1\}$. The monopolist strategy associated to

  $$U(x,y) := \max\{0,\, x+y/2-1,\, x+y-b\}, \text{ with } b = 1 + 1/(2\sqrt{3}), \tag{36}$$

  which involves the lottery ticket $(1, 1/2)$, yields better profits that any strategy restricted to deterministic bundles [17]. The triangle $T$, and the numerical best $U$, are illustrated on Figure 7 (center). These experiments suggest that (36) is a[6] globally optimal solution.

- Uniform customer density on the kite shaped domain $\{(x,y) \in [0,1]^2;\, x+y/2 \geq 1$ or $x/2+y \geq 1\}$, see Figure 7 (right). The optimal monopolist strategy is proved in [25] to involve probabilistic bundles, but it is not identified. Our numerical experiments suggest that it has the form

  $$U(x,y) := \max\{0,\, x+y/2-1,\, x/2+y-1,\, x+y-b\},$$

  which involves the lottery tickets $(1, 1/2)$ and $(1/2, 1)$. Under this assumption, the optimal value $b = 1 + 1/(3\sqrt{2})$ is easily computed.

Numerous qualitative questions remain open. Is there a distribution of customers for which the optimal strategy involves a continuum of distinct lottery tickets $\{(1,\alpha);\, \alpha_0 \leq \alpha \leq \alpha_1\}$ ?

**Pricing of risky assets.** A more complex economic model is considered in [4], where financial products, characterized by their expectancy of gain and their variability, are sold to agents characterized by their risk aversion and their initial risk exposure. We do not give the details of this model here, but simply point out that it fits in the general framework of (34) with the cost function $\mathrm{Cost}(a,b) := -\alpha(\xi a + \sqrt{-(a^2+b)})$, if $a^2 + b \leq 0$, and $+\infty$ otherwise, where $\xi \in \mathbb{R}$ and $\alpha \geq 0$ are parameters, see Example 3.2 in [4]. Observing that, for $a^2 + b < 0$

$$\det(\mathrm{Hessian}\,\mathrm{Cost}(a,b)) = \frac{1}{4(a^2+b)^2},$$

[6] Optimal solutions of (34) are not uniquely determined outside the customer density support $T = \mathrm{supp}(\mu)$.

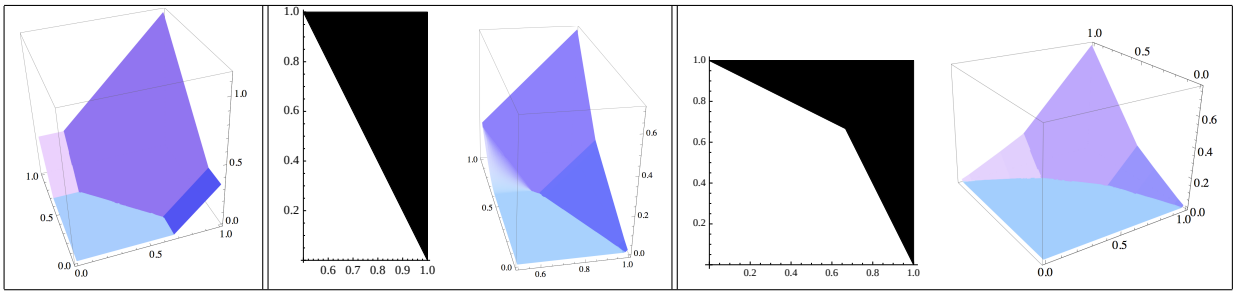


Figure 7: Three dimensional plot of the optimal $U$ for the product-bundles variant of the monopolist problem, with respect to various customer distributions. Left : uniform distribution on $[0,1]^2$. Center, and right: distribution uniform on the illustrated black polygon.

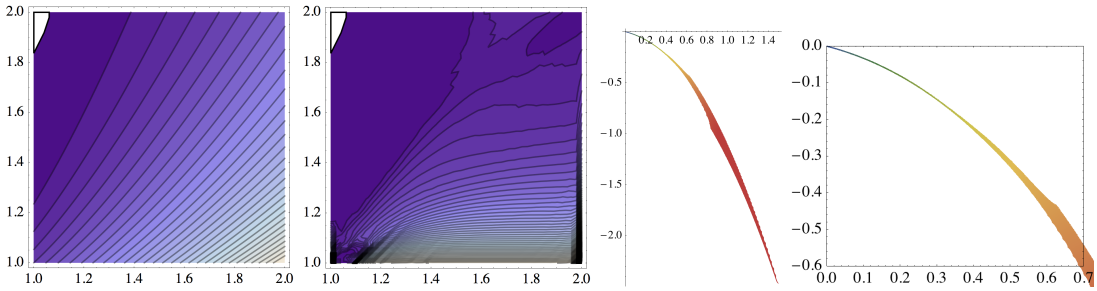


Figure 8: Optimal pricing of risky assets, with the parameters $\alpha = 1$, $\xi = 1$, and a uniform customer density on $[1,2]^2$. Left: level sets of the estimated solution $U$ of (37), with exclusion region $U = 0$ in white. Center left: level sets of $\det(\text{Hessian}\, U)$, the darkest one is the estimated bunching region $\det(\text{Hessian}\, U) = 0$. Center right, and right (detail): optimal product line, colored with the monopolist margin.

we easily obtain that this cost is convex[7]. The lack of smoothness of the square-root appearing in the cost function is a potential issue for numerical implementation, hence the problem (34) is reformulated using an additional variable $V$ subject to a (optimizer friendly) conic constraint

$$\max \left\{ \int \left( \langle \nabla U, z \rangle - U + \alpha\xi\, \partial_x U + \alpha V \right) d\mu;\; U \in \mathrm{Conv}_0(\mathbb{R}^2),\; V^2 + (\partial_x U)^2 + \partial_y U \leq 0 \right\}. \quad (37)$$

A numerical solution, presented Figure 8, displays the same qualitative properties (Desirability of exclusion, Bunching) as the classical monopolist problem with quadratic cost.

### 6.3 Comparison with alternative methods

We compare our implementation of the constraint of convexity with alternative methods that have been proposed in the literature. The compared algorithms are the following:

- (Adaptive constraints) The optimization strategy (Conv) described in Algorithm 2 section §6.1, based on the hierarchy of cones $\mathrm{Conv}(\mathcal{V})$, and used in our numerical experiments §6.2. The adaptation (DConv) of this strategy to the hierarchy $\mathrm{DConv}(\mathcal{V})$ of cones of "directionally convex" functions, see Appendix A.

[7] This property was not noticed in the original work [4].

- (Local constraints) The approach of Aguilera and Morin (AM, [1]) based on semi-definite programming. A method of Oberman and Friedlander ($\mathrm{OF}_2$, $\mathrm{OF}_3$, [22]), where $\mathrm{OF}_k$ refers to minimization over the cone $\mathrm{DConv}(\mathcal{V}_k)$ associated to the fixed stencil $\mathcal{V}_k(x) := \{e \in \mathcal{V}_{\max}(x);\ \|e\| \leq k\}$. A modification of $\mathrm{OF}_3$ by Oberman ($\mathrm{Ob}_3$, [21]), with additional constraints ensuring that the output is truly convex.

- (Global constraints) Direct minimization over the full cone $\mathrm{Conv}(X)$, as proposed by Carlier, Lachand-Robert and Maury (CLRM, [5]). Minimization over $\mathrm{GradConv}(X)$, see (4), following[8] Ekeland and Moreno (EM, [10]).

The numerical test chosen is the classical model of the monopolist problem, with quadratic cost, on the domain $[1,2]^2$, see Figure 1 and §6.2. This numerical test case is classical and also considered in [10, 18, 22]. It is discretized on a $n \times n$ grid, for different values of $n$ ranging from 10 to 100 (10 to 50 for global constraint methods due to memory limitations).

The number of linear constraints of the optimization problems assembled by the methods is shown on Figure 9 (center left). For adaptive strategies, this number corresponds to the final iteration. The semi-definite approach AM is obviously excluded from this comparison. Two groups are clearly separated: Adaptive and Local methods on one side, with quasi-linear growth, and Global methods on the other side, with quadratic growth. Let us emphasize that, despite the similar cardinalities, many constraints of the adaptive methods are not local, see Figure 5. The method DConv generally uses the least number of constraints, followed by $\mathrm{OF}_2$ and then Conv.

**Definition 6.4.** *The convexity defect of a discrete map $u : X \to \mathbb{R}$, is the smallest $\varepsilon \geq 0$ such that $u + \varepsilon q \in \mathrm{Conv}(X)$. The directional convexity defect of $u$ is the smallest $\varepsilon \geq 0$ such that $u + \varepsilon q \in \mathrm{DConv}(X)$, see Appendix A.*

Figure 9 displays the convexity defect of the discrete solutions produced by the different algorithms, at several resolutions. This quantity stabilizes at a positive value for the methods $\mathrm{OF}_2$ and $\mathrm{OF}_3$, which betrays their non-convergence as $n \to \infty$. We expect the convexity defect of the method AM to tend to zero, as the resolution increases, since this method benefits from a convergence guarantee [1]; for practical resolutions, it remains rather high however. Other methods, except DConv, have a convexity defect several orders of magnitude smaller, and which only reflects the numerical precision of the optimizer (some of the prescribed linear constraints are slightly violated by the optimizer's output). Finally, the method DConv has a special status since it often exhibits a large convexity defect, but its *directional* convexity defect vanishes (up to numerical precision).

We attempt on figure 10 to extract, with the different numerical methods, the regions of economical interest: potential customers excluded from the trade $\{U = 0\}$, and of customers subject to bunching $\{\det(\mathrm{Hessian}\, U) = 0\}$. While the features extracted from the method Conv are (hopefully) convincing, the coordinate bias of the method $\mathrm{OF}_2$ is apparent, whereas the method AM does not recover the predicted triangular shape of the set of excluded customers [24]. The other methods DConv, CLRM, EM, (not shown) perform similarly to Conv; the method $\mathrm{OF}_3$ (not shown) works slightly better than $\mathrm{OF}_2$, but still suffers from coordinate bias. The method $\mathrm{Ob}_3$ (not shown) seems severely inaccurate[9]: indeed the hessian matrix condition

[8] We use the description of $\mathrm{GradConv}(X)$ by $\mathcal{O}(N^2)$ linear constraints given in [10], but (for simplicity) not their energy discretization, nor their method for globally extending elements $u \in \mathrm{GradConv}(X)$.

[9] The methods $(\mathrm{Ob}_k)_{k\geq 1}$ are closely related to our approach since they produce outputs with zero convexity defect (up to numerical precision), and the number of linear constraints only grows linearly with the domain cardinality: $\mathcal{O}(k^2 N)$, with $N := \#(X)$. We suspect that better results could be obtained with these methods by selecting adaptively and locally the integer $k$.

number with $\mathrm{Ob}_k$, $k \geq 1$, cannot drop below $\approx 1/k^2$, see [21], which is incompatible with the bunching phenomenon, see the solution gradients on Figure 6.

For each method we compute exactly the monopolist profit (34), associated with the largest global map $U \in \mathrm{Conv}(\Omega)$ satisfying $U \leq u$ on $X$, where $u \in \mathcal{F}(X)$ is the method's discrete output. It is compared on Figure 9 with the best possible profit (which is not known, but was extrapolated from the numerical results). Convergence rate is numerically estimated to $n^{-1.1}$ for all methods[10] except (i) the semi-definite approach AM for which we find $n^{-0.75}$, and (ii) the method $\mathrm{Ob}_3$, for which energy does not seem to decrease.

In terms of computation time[11], three groups of methods can be distinguished. Global methods suffer from a huge memory cost in addition to their long run times. Methods using a (quasi)-linear number of constraints have comparable run times, thanks to the limited number of stencil refinement steps of the adaptive ones (their computation time might be further reduced by the use of appropriate hot starts for the consecutive subproblems). Finally, the semi-definite programming based method AM is surprisingly fast[12], although this is at the expense of accuracy, see above. For $n = 100$, the method CLRM would use $27 \times 10^6$ linear constraints, which with our equipment simply do not fit in memory. The proposed method Conv selects in 5 refinement steps a subset containing $\approx 0.4\%$ of these constraints ($100 \times 10^3$), and which is by construction guaranteed to include all the active ones; it completes in 6 minutes on a standard laptop.

In summary, adaptive methods combine the accuracy and convergence guarantees of methods based on global constraints, with the speed and low memory usage of those based on local constraints.

## Conclusion and perspectives

We in this paper introduced a new hierarchy of discrete spaces, used to adaptively solve optimization problems posed on the cone of convex functions. The comparison with existing hierarchies of spaces, such as wavelets or finite element spaces on adaptively refined triangulations, is striking by its similarities as much as by its differences. The cones $\mathrm{Conv}(\mathcal{V})$ (resp. adaptive wavelet or finite element spaces) are defined through linear inequalities (resp. bases), which become increasingly global (resp. local) as the adaptation loop proceeds. Future directions of research include improving the algorithmic guarantees, developing more applications of the method such as optimal transport, and generalizing the constructed cones of discrete convex functions to unstructured or three dimensional point sets.

**Acknowledgement.** The author thanks Pr Ekeland and Pr Rochet for introducing him to the monopolist problem, and the Mosek team for their free release policy for public research.

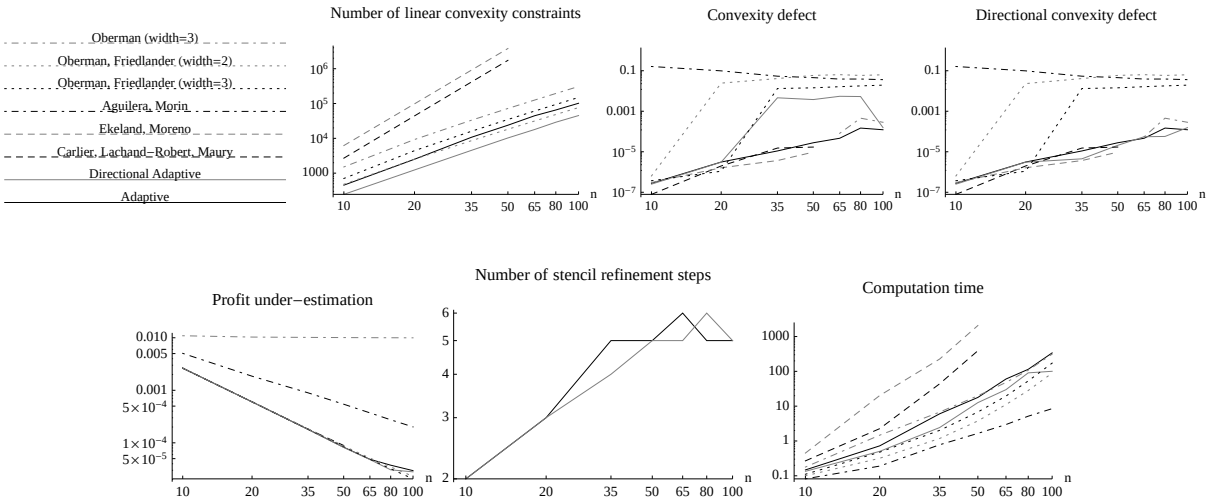
## References

[1] N. E. Aguilera and P. Morin. Approximating optimization problems over convex functions. *Numerische Mathematik*, 2008.

[10] The (presumed) non-convergence of the methods $\mathrm{OF}_2$ and $\mathrm{OF}_3$ is not visible in this graph.

[11] Experiments conducted on a 2.7 GHz Core i7 (quad-core) laptop, equipped with 16 GB of RAM.

[12] The method AM, implemented with Mosek's conic optimizer, takes only 2.5s to solve the product bundles variant of the monopolist problem on a $64 \times 64$ grid, with a uniform density of consumers on $[0, 1]^2$. This contrasts with the figure, 751s, reported in [1] in the same setting but with a different optimizer.




| $n = 50 \setminus$ Method | Conv | DConv | AM | $OF_2$ | $OF_3$ | $Ob_3$ | CLRM | EM |
|---|---|---|---|---|---|---|---|---|
| Constraints $\times 10^{-3}$ | 24 | 10 | NA | 18 | 35 | 72 | 1738 | 3803 |
| Defect $\times 10^{3}$ | 0.03 | 3.8 | 46 | 59 | 14 | 0 | 0.02 | 0.01 |
| Profit under estimation $\times 10^{3}$ | 0.11 | 0.11 | 0.57 | 0.11 | 0.11 | 10 | 0.12 | 0.11 |
| Computation time | 18s | 13s | 1.7s | 3.8s | 6.8s | 20s | 391s | 2070s |

Figure 9: Comparison of different numerical methods for the classical Monopolist problem.

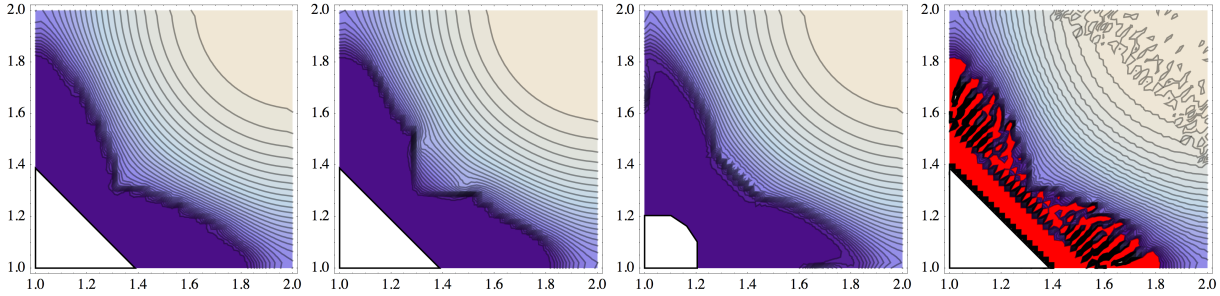


Figure 10: Level set $U < 10^{-4}$, in white, approximating the region $U = 0$ of excluded customers. Other level sets: $\{k\eta \leq \det(\text{Hessian}\, U) \leq (k+1)\eta\}$, with $\eta = 0.07$. The dark blue one, for $k = 0$, approximates the region $\det(\text{Hessian}\, U) = 0$ of customers subject to bunching. Hessian determinant extracted via subgradient measures, see Remark 6.3, and the numerical methods Conv (left), $OF_2$ (center left) and AM (center right). Right: extraction by taking the determinant of a finite differences discrete Hessian (and here the method Conv); this naïve procedure is unstable and produces negative values (in red).

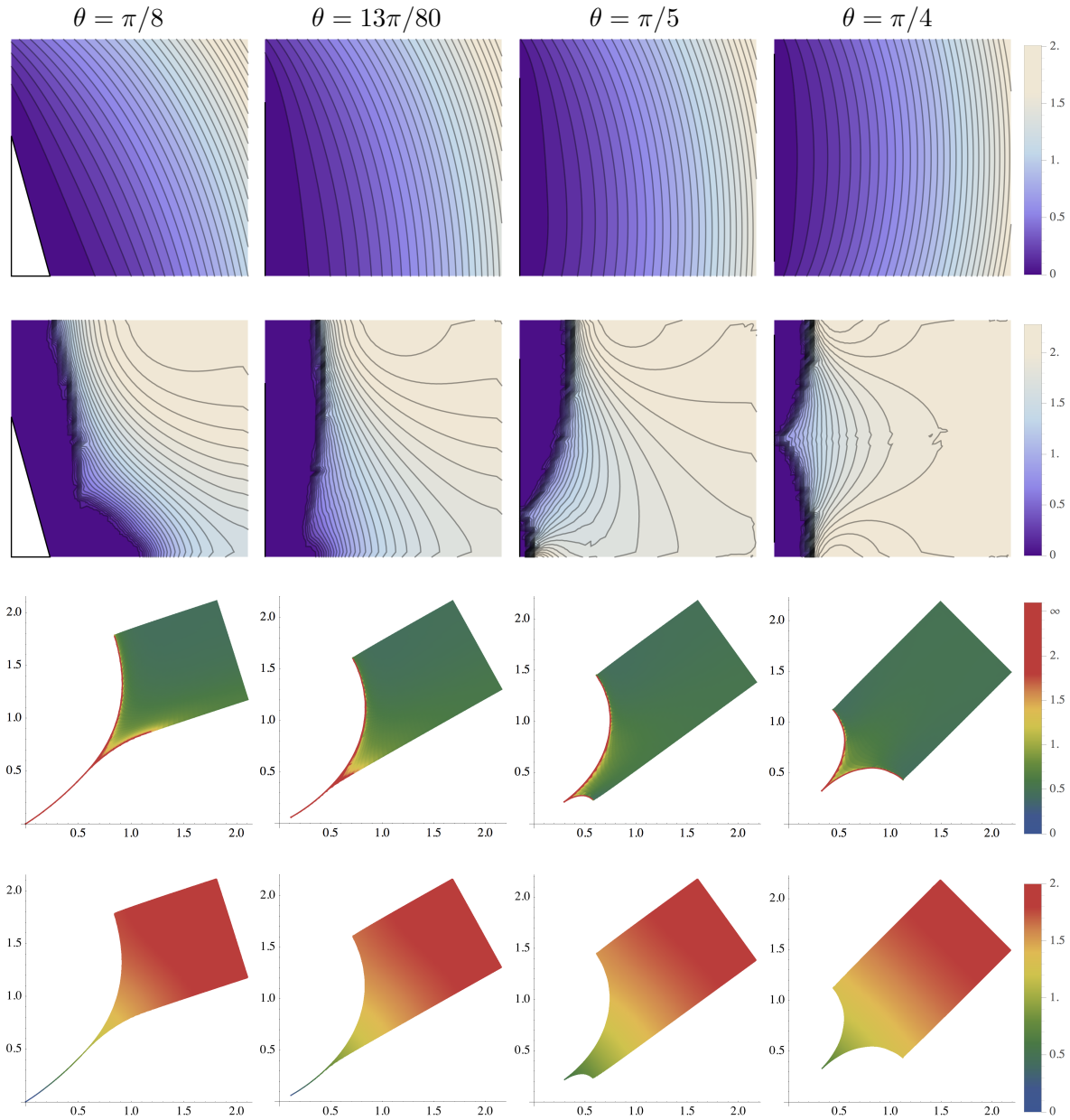



Figure 11: Results of the classical principal agent model, with a uniform density of customers on the domain obtained by rotating the square $[1,2]^2$ around its center $(3/2, 3/2)$ by the indicated angle $\theta$. From top to bottom: level sets of $U$ (vanishing set indicated in white), level sets of $\det(\text{Hessian}\, U)$ (allows to discriminate between the three categories of customers), density of products bought (which explodes on some curves), margin of the monopolist.

[2] J. F. Bonnans, E. Ottenwaelter, and H. Zidani. A fast algorithm for the two dimensional HJB equation of stochastic control. Technical report, 2004.

[3] G. Carlier. A general existence result for the principal-agent problem with adverse selection. *Journal of Mathematical Economics*, 2001.

[4] G. Carlier, I. Ekeland, and N. Touzi. Optimal derivatives design for mean–variance agents under adverse selection. *Mathematics and Financial Economics*, 2007.

[5] G. Carlier, T. Lachand-Robert, and B. Maury. A numerical approach to variational problems subject to convexity constraint. *Numerische Mathematik*, 2001.

[6] B. Chazelle. An optimal convex hull algorithm in any fixed dimension. *Discrete and Computational Geometry*, 1993.

[7] P. Choné and H. V. J. Le Meur. Non-Convergence result for conformal approximation of variational problems subject to a convexity constraint. *Numerical Functional Analysis and Optimization*, 2001.

[8] C. Dobrzynski and P. Frey. Anisotropic Delaunay Mesh Adaptation for Unsteady Simulations. In *Proceedings of the 17th international Meshing Roundtable*. 2008.

[9] H. Edelsbrunner and R. Seidel. Voronoi diagrams and arrangements. *Discrete and Computational Geometry*, 1986.

[10] I. Ekeland and S. Moreno-Bromberg. An algorithm for computing solutions of variational problems with global convexity constraints. *Numerische Mathematik*, 2010.

[11] J. Fehrenbach and J.-M. Mirebeau. Sparse Non-negative Stencils for Anisotropic Diffusion. *Journal of Mathematical Imaging and Vision*, 2013.

[12] A. Figalli, Y. H. Kim, and R. J. McCann. When is multidimensional screening a convex program? *Journal of Economic Theory*, 2011.

[13] R. L. Graham, D. E. Knuth, and O. Patashnik. Concrete mathematics : a foundation for computer science, 2nd ed 1994 GRAHAM Ronald L., KNUTH Donald E., PATASHNIK Oren: Librairie Lavoisier. *Addison & Wesley*, 1994.

[14] G. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford Science Publications, 1979.

[15] F. Hurtado, M. Noy, and J. Urrutia. Flipping Edges in Triangulations. *Discrete and Computational Geometry*, 1999.

[16] T. Lachand-Robert and E. Oudet. Minimizing within Convex Bodies Using a Convex Hull Method. *SIAM Journal on Optimization*, 2005.

[17] A. M. Manelli and D. R. Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 2006.

[18] Q. Merigot and E. Oudet. Handling convexity-like constraints in variational problems (Preprint). *www-ljk.imag.fr*.

[19] J.-M. Mirebeau. Anisotropic Fast-Marching on cartesian grids using Lattice Basis Reduction (Preprint). *arXiv.org*, 2012.

[20] J.-M. Mirebeau. Efficient fast marching with Finsler metrics. *Numerische Mathematik*, 2013.

[21] A. M. Oberman. A Numerical Method for Variational Problems with Convexity Constraints. *SIAM Journal on Scientific Computing*, 2013.

[22] A. M. Oberman and M. P. Friedlander. A numerical method for variational problems over the cone of convex functions (Preprint). *arXiv.org*, 2011.

[23] M. Qi, T.-T. Cao, and T.-S. Tan. Computing 2D constrained Delaunay triangulation using the GPU. In *the ACM SIGGRAPH Symposium*, 2012.

[24] J. C. Rochet and P. Choné. Ironing, Seeping, and Multidimensional Screening. *Econometrica*, 1998.

[25] J. Thanassoulis. Haggling over substitutes. *Journal of Economic Theory*, 2004.

[26] G. Wachsmuth. The Numerical Solution of Newton's Problem of Least Resistance (Preprint). 2013.

# A Directional convexity

We introduce and discuss a weak notion of discrete convexity, which involves slightly fewer linear constraints than (3) and seems sufficient to obtain convincing numerical results, see §6.3.

**Definition A.1.** *We denote by* $\mathrm{DConv}(X)$ *the collection of elements in* $\mathcal{F}(X)$ *on which all the linear forms* $S^e_x$, $x \in X$, $e \in \mathbb{Z}^2$ *irreducible, supported on* $X$, *take non-negative values.*

Some elements of $\mathrm{DConv}(X)$ cannot be extended into global convex maps on $\mathrm{Hull}(X)$. Their existence follows from the second point of Theorem 1.4 (minimality of the collection of constraints $S^e_x$, $T^e_x$), but for completeness we give (without proof) a concrete example.

**Proposition A.2.** *Let* $u \in \mathcal{F}(\mathbb{Z}^2)$ *be defined by* $u(1,1) = 1$, $u(-1,0) = u(0,-1) = -1$, *and* $u(x) := 2\|x\|^2$ *for other* $x \in \mathbb{Z}^2$. *Then for all* $x \in \mathbb{Z}^2$, *and all irreducible* $e \in \mathbb{Z}^2$, *one has* $S^e_x(u) \geq 1$, *and* $T^e_x(u) \geq 2$ *if* $\|e\| > 1$, *with the exception* $T_0^{(1,1)}(u) = -1$.

Elements of $\mathrm{DConv}(\mathbb{Z}^2)$ are nevertheless "almost" convex, in the sense that their restriction to a coarsened grid is convex.

**Proposition A.3.** *If* $u \in \mathrm{DConv}(X)$, *then* $u_{|X'} \in \mathrm{Conv}(X')$, *with* $X' := X \cap 2\mathbb{Z}$.

*Proof.* Let $x \in X$, and let $e \in \mathbb{Z}^2$, $\|e\| > 1$, be irreducible and of parents $f, g$. Assuming that $x + 2e, x - 2f, x - 2g \in X$, and observing that $x \pm e \in X$ by convexity, we obtain

$$u(x+2e) + u(x-2f) + u(x-2g) - 3u(x) = 2S^e_x(u) + S^e_{x+e}(u) + S^{f-g}_{x-e}(u) \geq 0.$$

Likewise $u(x+2e) - 2u(x) + u(x-2e) = S^e_{x-e}(u) + 2S^e_x(u) + S^e_{x+e}(u) \geq 0$. □

The cone $\mathrm{DConv}(X)$ of directionally convex functions admits, just like $\mathrm{Conv}(X)$, a hierarchy of sub-cones $\mathrm{DConv}(\mathcal{V})$ associated to stencils.

**Definition A.4.** *Let* $\mathcal{V}$ *be a family of stencils on* $X$, *and let* $u \in \mathcal{F}(X)$. *The cone* $\mathrm{DConv}(\mathcal{V})$ *is defined by the non-negativity of the following linear forms: for all* $x \in X$

- *For all* $e \in \mathcal{V}(x)$, *the linear form* $S^e_x$, *if supported on* $X$.

- *For all $e \in \hat{\mathcal{V}}(x)$, the linear form $H_x^e := P_x^e + P_x^{-e}$, if supported on $X$.*

**Proposition A.5.**
- *For any stencils $\mathcal{V}, \mathcal{V}'$, one has* $\mathrm{DConv}(\mathcal{V}) \subseteq \mathrm{DConv}(X)$, $\mathrm{DConv}(\mathcal{V}) \cap \mathrm{DConv}(\mathcal{V}') = \mathrm{DConv}(\mathcal{V} \cap \mathcal{V}')$, *and* $\mathrm{DConv}(\mathcal{V}) \cup \mathrm{DConv}(\mathcal{V}') \subseteq \mathrm{DConv}(\mathcal{V} \cup \mathcal{V}')$.

- *For any stencils $\mathcal{V}$, one has*

$$\mathrm{DConv}(\mathcal{V}) = \{u \in \mathrm{DConv}(X);\ H_x^e(u) \geq 0 \text{ for all } x \in X,\ e \in \mathcal{V}_{\max}(x) \setminus \mathcal{V}(x)\}. \tag{38}$$

*Proof.* As observed in §3.2, the second point of this proposition implies the first one. We denote by $\mathbb{P}(\mathcal{V})$ the identity (38), and prove it by decreasing induction over $\#(\mathcal{V})$. Since $\mathbb{P}(\mathcal{V}_{\max})$ clearly holds, we consider stencils $\mathcal{V} \subsetneq \mathcal{V}_{\max}$.

Let $x \in X$, $e \in \mathcal{V}_{\max}(x) \setminus \mathcal{V}(x)$, be such that $\|e\|$ is minimal. Similarly to Proposition 3.9 we find that $e$ belongs to the set $\hat{\mathcal{V}}(x)$ of candidates for refinement at $x$, and define stencils $\mathcal{V}'$ by $\mathcal{V}(x) := \mathcal{V}(x) \cup \{e\}$, and $\mathcal{V}'(y) := \mathcal{V}(y)$ for $y \neq x$. The cones $\mathrm{DConv}(\mathcal{V})$ and $\mathrm{DConv}(\mathcal{V}')$ are defined by a common collection of constraints, with the addition respectively of $H_x^e$ for $\mathrm{DConv}(\mathcal{V})$, and $S_x^e$, $H_x^{e+f}$, $H_x^{e+g}$ for $\mathrm{DConv}(\mathcal{V}')$. Expressing the latter linear forms as combinations of those defining $\mathrm{DConv}(\mathcal{V})$

$$S_x^e = H_x^e + S_x^f + S_x^g, \quad H_x^{e+f} = H_x^e + S_{x+e}^f + S_{x-e}^f, \quad H_x^{e+g} = H_x^e + S_{x+e}^g + S_{x-e}^g,$$

and observing that $\mathbb{P}(\mathcal{V}')$ holds by induction, we conclude the proof of $\mathbb{P}(\mathcal{V})$. □