Numerical Calibration of Steiner trees

Annalisa Massaccesi, Edouard Oudet, Bozhidar Velichkov

April 18, 2016

Abstract

In this paper we propose a variational approach to the Steiner tree problem, which is based on calibrations in a suitable algebraic environment for polyhedral chains which represent our candidates. This approach turns out to be very efficient from numerical point of view and allows to establish whether a given Steiner tree is optimal. Several examples are provided.

Keywords: Calibration, Minimal surface, Steiner tree problem.
2010 Mathematics Subject Classification: 49J45, 35R35, 49M05, 35J25

1 Introduction

In the present paper we consider variational problems related to one dimensional models arising in several fields among which graph theory, shape optimization, optimal transport and elasto-plasticity. More specifically, we address the question to certify the optimality of a given one dimensionnal structure for some specific costs. Whereas the identification of an optimal structure may be an NP-hard problem like in the case of the optimal Steiner tree problem for instance, we introduce in this article a new algorithm based on the so called notion of calibration which, on cases under study, has been able to establish optimality from both a numerical and a theoretical point of view.

Our approach is rather flexible and applies to models with some common features: the objects under investigation can be represented by means of 1-dimensional rectifiable sets with suitable "labels" and the cost to be optimized, depending both on the support of the object and its labels, has a "subadditive" behavior. Among these problems, we can list k-means clustering, image processing, irrigation and the Steiner tree problem. Here below we discuss the statement of the problem in some of them, the original motivation for the paper being the latter. The reason for such a treatment of the aforementioned problems is the availability of well-established results and tools which are typical of Calculus of Variations and Geometric Measure Theory.

The celebrated Steiner tree problem can be summarized as follows: in the Euclidean space $\mathbb{R}^d$ we are given n points $\{x_1, \dots, x_n\}$ and we have to find the shortest connected set containing these points. More precisely, one would like to solve the following variational problem

$$\inf \Big\{ \mathcal{H}^1(\Sigma) : \Sigma \subset \mathbb{R}^d \text{ is a connected set and } \{x_1, \dots, x_n\} \subset \Sigma \Big\} . \tag{1.1}$$

From an exquisitely theoretical point of view, the Steiner tree problem is well understood : a solution always exists, though it may not be unique, in general. Moreover, the graph which minimizes the problem (1.1) is a union of segments, meeting either at the boundary points $\{x_1, \dots, x_n\}$ or in the so-called "Steiner points". In $\mathbb{R}^2$, the segments meeting at a Steiner point are always three, forming angles of $2\pi/3$ radiants, and Steiner points are at most $n-2$ (see, for instance, [3] and [4]).

The situation may seem complete: indeed, in $\mathbb{R}^2$ there are very efficient algorithms which allow to explicitly construct the solution of (1.1) (see for instance [8]). Nonetheless, the computability of the Steiner tree problem is a very difficult task. In fact, the problem is NP-complete, as one can see in [5].

We propose to attack the Steiner tree problem borrowing some tools from Geometric Measure Theory. More precisely, we show that the Steiner tree problem is equivalent to the classical mass-minimization problem with prescribed boundary (see Section 2.3 or [6], too), provided that the object of our study is 1-dimensional polyhedral chains (that is, finite collections of segments) with "weights" chosen in a suitable normed group G. The key-point is the meaning we assign to the concepts of mass and boundary for such a class of objects. In particular, the mass of a 1-dimensional polyhedral G-chain is the weighted length of its support.

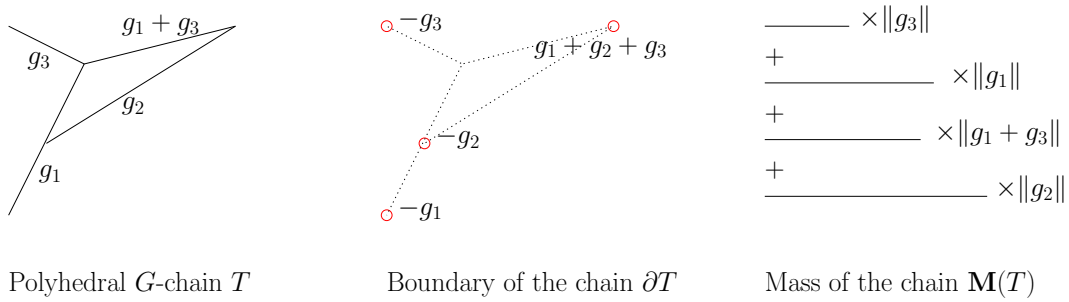


Figure 1: A polyhedral chain with its boundary and mass

Our ideal reference is the classical variational problem of mass-minimization among currents with prescribed boundary, also known as Plateau problem. Nonetheless, these kind of problems are significantly simpler when the dimension of the current is 1: in particular, we can forget about currents and simply deal with polyhedral chains.

Moreover, it turns out that sometimes the mass-minimization problem has good computability properties in the (relatively) easy case of dimension 1. The reason is the calibration technique, that is, a sufficient (but, unfortunately, not necessary) condition for the minimality of a polyhedral chain. Indeed, the existence of a calibration associated with a certain polyhedral chain guarantees that the latter is a mass-minimizer for its boundary. Roughly speaking, having a "robust" candidate for the mass-minimization problem, one may try to explicitly build a calibration associated with this candidate: if we succeed, then the candidate is a genuine solution to the mass-minimization problem.

The constructions of the group and the equivalent objects are completely explicit, and the construction of a possible calibration is explicit, too. Thus we can pass to a numerical implementation of an algorithm which looks for calibrations of a candidate minimizer (in many cases, it is well-known that we are looking for calibration of true minimizers).

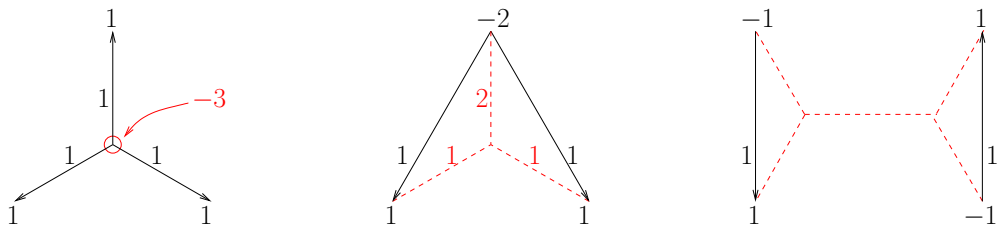
Back to a wider class of variational problem, from the point of view of applications this method is very flexible and its flexibility is based on the choice of the normed group G. The algebraic relations among the elements of the group, which are also the coefficients for the polyhedral chains, and the norm, which plays a crucial role in the computation of the mass of a polyhedral chain, can be chosen in such a way that the candidates in the mass-minimization problem have specific properties.

2 Steiner's candidates as polyhedral chains with coefficients in a group

2.1 Origin of difficulties: why coefficients in a group are needed?

It is easy to imagine how the Steiner tree problem can be rephrased as a mass-minimization problem for polyhedral chains, but it is less trivial to exhibit a group that preserves the two main features of the problem: connectedness of the set and the prescribed boundary datum at

$\{x_1, \ldots, x_n\}$. In fact, the first "natural" attempt $G = \mathbb{Z}$ gives problems even in the simplest cases, as one can see in the pictures below.



"Symmetric" - but not admissible! See (2.1) - attempt to set a boundary. Extra boundary point with multiplicity -3.

The "good" solution is not selected because of multiplicity 2 on the vertical segment.

The mass-minimizer is not even a connected set!

Figure 2: The mass-minimization problem with $G = \mathbb{Z}$

With easy heuristic arguments, one can see that the group G should be chosen with the following properties (see Remark 2.14):

(1) $\exists\, g_1, \ldots, g_n \in G$ (with $g_i \neq g_j$ if $i \neq j$) such that $g_1 + \ldots + g_n = 0$ and $\|g_{i_1} + \ldots + g_{i_k}\| = 1$ for every $k \leq n-1$ and $1 \leq i_1 < \ldots < i_k \leq n$;

(2) $\forall\, g \in G \setminus \{0\} \quad \|g\| \geq 1$.

In Section 2.3 we build a normed group with the properties above and we obtain an equivalence result between the Steiner tree problem and the variational problem given by the mass-minimization among polyhedral G-chains with a suitably chosen boundary.

2.2 Polyhedral chains with coefficients in a group

We will denote by $d \in \mathbb{N}$ the dimension of the ambient space $\mathbb{R}^d$ where the Steiner tree problem is posed. In this section we consider the following general setting: let $\|\cdot\|$ be a norm in $\mathbb{R}^m$, for some $m \in \mathbb{N}$, and $G < \mathbb{R}^m$ an additive subgroup of $\mathbb{R}^m$ that inherits the same norm $\|\cdot\|$. In our construction we will set $m = n$, where n is the number of boundary points, and

Definition 2.1. *A* 1*-dimensional* polyhedral G-chain T *is a collection of oriented and G-weighted segments $\{\sigma_i\}_{i=1}^{N_\sigma}$ in $\mathbb{R}^d$ which intersect only at endpoints*[1]*. More precisely, we will denote by T a triple (Σ, τ, θ) where*

1. *$\Sigma = \bigcup_{i=1}^{N_\sigma} \sigma_i \subset \mathbb{R}^d$ is the* support *of the chain T and $\sigma_i = \overline{x_i y_i}$;*
2. *$\tau : \bigcup_{i=1}^{N_\sigma} \mathring{\sigma}_i \to \mathbb{S}^{d-1}$ is the* orientation *of T (thus $\tau(x) = (y_i - x_i)/|y_i - x_i|$ for every $x \in \mathring{\sigma}_i$);*
3. *$\theta : \Sigma \to G$ is the* multiplicity *of T. For our convenience, we can assume θ to be constantly equal to $\theta_i \in G$ on each segment σ_i.*

Definition 2.2. *Given a* 1*-dimensional polyhedral G-chain $T = (\Sigma, \tau, \theta)$, with $\Sigma = \bigcup_{i=1}^{N_\sigma} \sigma_i$, we define the* mass *of T as*

$$\mathbf{M}(T) := \sum_{i=1}^{N_\sigma} \|\theta_i\| \mathrm{Length}(\sigma_i)\,.$$

[1]The endpoints of each segment σ_i will be denoted by x_i and $y_i \in \mathbb{R}^d$ (with orientation from x_i to y_i).

Definition 2.3. *Given a 1-dimensional polyhedral G-chain $T = (\Sigma, \tau, \theta)$ in $\mathbb{R}^d$, with $\Sigma = \bigcup_{i=1}^{N_\sigma} \sigma_i$ and $\sigma_i = \overline{x_i y_i}$, we define the* boundary *of T as the pair $(B, g) =: \partial T$, where*

1. *$\tilde{B} = \bigcup_{i=1}^{N_\sigma} \{x_i, y_i\} \subset \mathbb{R}^d$ is the* unsuitable support of the boundary *of T;*
2. *$g : \tilde{B} \to G$ is the* multiplicity of the boundary *of T and maps*

$$x \quad \mapsto \quad g(x) := -\sum_{\{i:\, x_i = x\}} \theta_i + \sum_{\{j:\, y_j = x\}} \theta_j \,;$$

3. *$B = \{x \in \tilde{B} :\, g(x) \neq 0\}$ is the* support of the boundary *of T.*

Remark 2.4. Given a finite number of points $\{x_1, \dots, x_n\} \subset \mathbb{R}^d$ with a multiplicity $g : \{x_1, \dots, x_n\} \to G$, the pair $(\{x_i\}_{i=1}^n, g)$ is the boundary of some polyhedral G-chain T if and only if

$$\sum_{i=1}^n g_i = 0\,. \tag{2.1}$$

Indeed, if $(\{x_i\}_{i=1}^n, g)$ is the boundary of a polyhedral G-chain T, then

$$\sum_{i=1}^n g_i = \sum_{x \in B} \left(-\sum_{\{i:\, x_i = x\}} \theta_i + \sum_{\{j:\, y_j = x\}} \theta_j \right) = \sum_{i=1}^{N_\sigma} (-\theta_i + \theta_i) = 0\,.$$

On the other hand, if (2.1) holds (that is, $g_n = -\sum_{i=1}^{n-1} g_i$) then we can build a polyhedral G-chain with $\Sigma = \bigcup_{i=1}^{n-1} \overline{x_i x_n}$ and multiplicity $\theta_i = -g_i$ in every segment $\overline{x_i x_n}$. Condition (2.1) will be called the *admissibility condition.* In this context, a quite natural variational problem arises: namely, given a pair $(\{x_i\}_{i=1}^n, g)$ satisfying (2.1), find a mass-minimizing 1-dimensional G-chain with boundary $(\{x_i\}_{i=1}^n, g)$.

Given a bounded matrix-valued function $M : \mathbb{R}^d \to \mathbb{R}^{m \times d}$, we can define the action of a G-chain T on M as

$$T(M) := \int_\Sigma (M_x \tau(x)) \cdot \theta(x)\, d\mathcal{H}^1(x)\,, \tag{2.2}$$

where $\cdot$ is the standard scalar product in $\mathbb{R}^m$. In this sense, T is a current with coefficients in G, i.e., a continuous, linear functional on the space of smooth $\mathbb{R}^m$-valued 1-forms. Moreover, as a current T has finite mass, rectifiable support and multiplicity in G, and its boundary ∂T have the same properties, too, thus T is a rectifiable current with coefficients in G. This is the (more abstract) approach followed in [6]. Let us briefly recall what is important for us of that approach:

Definition 2.5. *Given a map $M \in C^1(\mathbb{R}^d; \mathbb{R}^{m \times d})$, we denote by $M_{(j)} \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ the jth row of the matrix M, for $j = 1, \dots, m$. M is said to be* curl free *if* $\operatorname{curl} M_{(j)} = 0$[2] *for every $j = 1, \dots, m$.*

Definition 2.6. *Assume $d = 2$ and consider a locally finite partition $\{C_r\}_{r \geq 1}$ of $\mathbb{R}^2$, where the cells C_r are pairwise disjoint, connected open sets with Lipschitz boundary. Given a piecewise constant map $M : \mathbb{R}^2 \to \mathbb{R}^{m \times 2}$ associated with $\{C_r\}_{r \geq 1}$ (i.e., $M_{|C_r} \equiv M_r \in \mathbb{R}^{m \times 2}$), we say that M fulfills the* compatibility condition *if*

$$(M_r - M_l) t(x) = 0 \quad \forall\, x \in \partial C_r \cap \partial C_l\,, \tag{2.3}$$

where $t \in \mathbb{R}^2$ is the tangent vector to ∂C_r.

[2] When $d > 3$, the curl above stands for the exterior derivative of the 1-form associated with $M_{(j)}$.

Proposition 2.7. *If S is a polyhedral G-chain without boundary and $M : \mathbb{R}^d \to \mathbb{R}^{m\times d}$ is either*

- *curl free or*
- *piecewise constant and compatible in the sense of* (2.3) *for some partition $\{C_r\}_{r\geq 1}$ of $\mathbb{R}^2$,*

then $S(M) = 0$.

Proof. Observe that this fact is nothing but a generalized version of Stokes' Theorem. We give here a short proof when $d = 2$ and M is curl free, for the complete proof of this fact (which requires to handle higher dimensional currents) we refer to [6].

Since $\partial S = 0$, we can decompose[3] $S = (\Sigma, \tau, \theta)$ in a finite sum of 1-dimensional polyhedral G-cycles $S_k = (\Sigma_k, \tau_k, \theta_k)$ with constant multiplicity $\theta_k = (\theta_k^{(1)}, \dots, \theta_k^{(m)}) \in G$. Thus, using the definition in (2.2), we obtain

$$\begin{aligned} S(M) &= \int_\Sigma M_x\tau(x)\cdot\theta(x)\, d\mathcal{H}^1(x) = \sum_{k=1}^{N_c}\int_{\Sigma_k} M_x\tau_k(x)\cdot\theta_k\, d\mathcal{H}^1(x) \\ &= \sum_{k=1}^{N_c}\sum_{j=1}^{m}\theta_k^{(j)}\int_{\Sigma_k} M_{(j)}\tau_k(x)\, d\mathcal{H}^1(x) \end{aligned}$$

Nonetheless, each Σ_k is a closed Lipschitz curve, so it is the (topological) boundary of a polygon P_k. The Stokes' Theorem allows to conclude that

$$S(M) = \sum_{k=1}^{N_c}\sum_{j=1}^{m}\theta_k^{(j)}\int_{P_k} \operatorname{curl} M_{(j)}\, d\mathcal{L}^2 = 0\,. \qquad \square$$

Definition 2.8. *A bounded, matrix-valued map $\omega : \mathbb{R}^d \to \mathbb{R}^{m\times d}$ is a* calibration *associated with a polyhedral G-chain $T = (\Sigma, \tau, \theta)$ if the following conditions hold:*

(i) *either ω is curl free*

(i') *or $d = 2$ and ω satisfies the compatibility condition* (2.3) *for some partition $\{C_r\}_{r\geq 1}$;*

(ii) *for $\mathcal{H}^1$-a.e. $x \in \Sigma$ we have that*

$$\omega(x)\tau(x)\cdot\theta(x) = \|\theta(x)\|; \tag{2.4}$$

(iii) *finally*

$$\sup_{x\in\mathbb{R}^d}\left(\sup_{\substack{t\in\mathbb{S}^{d-1}\\ g\in G,\|g\|=1}} \omega(x)t\cdot g\right) \leq 1\,.$$

Definition 2.9. *Given $\varepsilon > 0$, a bounded, matrix-valued map $\omega : \mathbb{R}^d \to \mathbb{R}^{m\times d}$ is an ε-*quasi calibration *associated with a polyhedral G-chain $T = (\Sigma, \tau, \theta)$ if conditions (i) (or (i')) and (ii) in Definition* 2.8 *are fulfilled and*

(iii')

$$\sup_{x\in\mathbb{R}^d}\left(\sup_{\substack{t\in\mathbb{S}^{d-1}\\ g\in G,\|g\|=1}} \omega(x)t\cdot g\right) \leq 1+\varepsilon\,.$$

[3]The existence of such a decomposition is known for 1-dimensional currents (see [2]) and consequently for 1-dimensional polyhedral G-chains (see [6]). Heuristically, for a finite number of segments (as in the case of polyhedral G-chains), one may explicitly perform the decomposition following the path of a fixed coefficient. This path has to close up to prevent the formation of boundary.

Theorem 2.10. *Consider a 1-dimensional polyhedral G-chain $T = (\Sigma, \tau, \theta)$ with boundary ∂T. If there exists a calibration ω associated with T in the sense of Definition 2.8, then T minimizes the mass among 1-dimensional polyhedral G-chain with boundary ∂T.*

Proof. Consider a polyhedral G-chain $T' = (\Sigma', \tau', \theta')$ with $\partial T' = \partial T$ as a competitor. Then the sum $S := T - T'$ has no boundary and we can apply Proposition 2.7, obtaining

$$T(\omega) = T'(\omega)\,. \tag{2.5}$$

Therefore

$$\begin{aligned}\mathbf{M}(T) &= \sum_{i=1}^{N_\sigma} \|\theta_i\| \mathrm{Length}(\sigma_i) \overset{\text{(ii)}}{=} \int_\Sigma \omega(x)\tau(x)\cdot\theta(x)\,d\mathcal{H}^1(x) = T(\omega)\\ &\overset{(2.5)}{=} T'(\omega) = \int_{\Sigma'} \omega(x)\tau'(x)\cdot\theta'(x)\,d\mathcal{H}^1(x) \overset{\text{(iii)}}{\leq} \sum_{i=1}^{N_\sigma} \|\theta_i'\| \mathrm{Length}(\sigma_i') = \mathbf{M}(T')\,.\end{aligned}$$

□

Proposition 2.11. *If ω is an ε-quasi calibration for some polyhedral G-chain T and some $\varepsilon > 0$, then*

$$\mathbf{M}(T) - \mathbf{M}(T_0) \leq \varepsilon \mathbf{M}(T_0)\,,$$

where T_0 is the mass-minimizing polyhedral G-chain with boundary ∂T.

Proof. As in the proof of Theorem 2.10:

$$\mathbf{M}(T) \overset{\text{(ii)}}{=} T(\omega) \overset{\text{(i)}}{=} T_0(\omega) \overset{\text{(iii}')}{\leq} (1+\varepsilon)\mathbf{M}(T_0)\,.$$

□

Remark 2.12. Theorem 2.10 gives a powerful sufficient condition for the mass-minimization: the existence of a calibration. Whether this condition is also necessary is a delicate issue. In fact, with purely functional analytical arguments, one can prove the following: given a mass-minimizing *normal current N with coefficients in $\mathbb{R}^m$*, there exists a *generalized calibration* Ω associated to N. A normal current N with coefficients in $\mathbb{R}^m$ is a continuous, linear functional on the space of smooth $\mathbb{R}^m$-valued forms (see (2.2)) which can be represented as

$$N(M) = \int M_x\tau(x)\cdot m(x)\,d\mu(x)\,,$$

where τ is a vectorfield in $\mathbb{R}^d$, $m \in L^1(\mathbb{R}^d;\mathbb{R}^m)$ and μ is a Radon measure on $\mathbb{R}^d$. A generalized calibration for N is a continuous functional on the space of normal currents with the following properties:

(i) $\Omega(S) = 0$ for every normal current S with coefficients in $\mathbb{R}^m$ with $\partial S = 0$;

(ii) $\Omega(N) = \mathbf{M}(N)$;

(iii) $\Omega(N') \leq \mathbf{M}(N')$ for every normal current N' with coefficients in $\mathbb{R}^m$.

This result has major drawbacks:

1. First of all, the construction of such a Ω is completely abstract (it relies on Hahn-Banch Theorem) and gives very little information on N (notice that $\Omega(N) = \mathbf{M}(N)$ is far from being as significant as (2.4)).

2. The mass-minimization problem among normal currents is different from the mass-minimization problem among rectifiable currents with coefficients in G (or even polyhedral G-chains), the latter being a smaller class of the class of normal currents.

2.3 Calibration in Steiner's context

In Section 2 we have introduced a variational problem for 1-dimensional polyhedral G-chains: namely, the minimization of the mass among G-chains with a prescribed boundary. Moreover, a powerful sufficient condition for minimality has been introduced: the calibration (see Definition 2.8 and Theorem 2.10). These tools can be brought back to the framework of the Steiner tree problem, provided we have a result of equivalence between the Steiner tree problem and the mass-minimization problem for 1-dimensional polyhedral G-chains (see Theorem 2.15): roughly speaking, we describe a way to convert a candidate for the Steiner tree problem in a 1-dimensional polyhedral G-chain and viceversa. This construction depends on a suitable choice of the group G.

Definition 2.13. *Fix $n \in \mathbb{N} \setminus \{0\}$ and consider the following seminorm in $\mathbb{R}^n$:*

$$\|u\|_\star := \max_{i=1,\dots,n} u_i - \min_{i=1,\dots,n} u_i \quad \forall\, u = (u_1, \dots, u_n) \in \mathbb{R}^n\,.$$

Now consider the map $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ associating $(u_1, \dots, u_n) \mapsto (u_1 - u_n, \dots, u_{n-1} - u_n)$. The seminorm $\|\cdot\|_\star$ induces a norm $\|\cdot\|$ in $\mathbb{R}^{n-1}$ via π. If $\{\mathbf{e}_1, \dots, \mathbf{e_n}\}$ is a standard orthonormal basis of $\mathbb{R}^n$, we take $g_i := \pi(\mathbf{e}_i)$ for $i = 1, \dots, n-1$ and we define the suitable group for a Steiner tree problem of n points *as*

$$G = \operatorname{span}(g_1, \dots, g_{n-1})\,.$$

The group G is naturally endowed with the norm $\|\cdot\|$.

Remark 2.14. By construction, the following properties hold:

(P1) for every $k \le n-1$ and $1 \le i_1 < \ldots < i_k \le n-1$ we have that

$$\|g_{i_1} + \ldots + g_{i_k}\| = 1\,.$$

In particular, $\|g_i\| = 1$ for every $i = 1, \dots, n$;

(P2) $\|g\| \ge 1$ for every $g \in G$.

The basic construction for the equivalence theorem 2.15 describes how to pass from a connected set $\Sigma \subset \mathbb{R}^d$ to a 1-dimensional polyhedral G-chain: we put it here for the convenience of the reader. In fact, it is very useful to know this construction for the numerical implementation of the problem. It is not restrictive to assume that Σ is a finite union of segments $\{\sigma_i\}_{i=1}^{N_\sigma}$ and has no loops, its endpoints are exactly $\{x_1, \dots, x_n\}$. Then we define the 1-dimensional polyhedral G-chain $T_\Sigma = (\Sigma, \tau, \theta)$ associated with Σ in the following way: for every $i = 1, \dots, n-1$ take the (injective) path Π_i in Σ connecting x_i with x_n and define the auxiliary G-chain $P_i := (\Pi_i, \tau_i, g_i)$ with orientation τ_i going from x_i to x_n and constant multiplicity g_i. Now set $T_\Sigma := \sum_{i=1}^{n=1} P_i$ and notice that

- the support of T is the original set Σ;
- the multiplicity of T_Σ has always unit norm thanks to property (P1) in Remark 2.14.

Theorem 2.15. *Consider n points $x_1, \dots, x_n \in \mathbb{R}^d$ and the normed group $(G, \|\cdot\|)$ in Definition 2.13, which depends*[4] *only on n. Set $g_n := -(g_1 + \ldots + g_{n-1})$ and consider the admissible boundary $(B, g) = \{(x_i, g_i)\}_{i=1}^n$. A connected set $\Sigma \supset \{x_1, \dots, x_n\}$ solves the Steiner tree problem with datum $\{x_1 \dots, x_n\}$ if and only if the 1-dimensional G-chain T_Σ described above solves the mass-minimization problem with boundary (B, g).*

This theorem is proved in [6], we provide here only the main ideas of the relevant implication, that is, if T_Σ is a mass-minimizer then Σ solves the Steiner tree problem.

[4] Notice that the group G is independent from the position of the points $x_1, \dots, x_n \in \mathbb{R}^d$.

Proof. Consider a connected competitor $\Sigma' \subset \mathbb{R}^d$: as before, it is not restrictive to assume that Σ' is a finite union of segments, has no loops and its endpoints are exactly $\{x_1, \dots, x_n\}$. Therefore we can construct the associated polyhedral G-chain $T_{\Sigma'}$; since the multiplicity has always unit norm, then $\mathcal{H}^1(\Sigma') = \mathbf{M}(T_{\Sigma'})$. On the other hand, property (P2) from Remark 2.14 guarantees that $\mathbf{M}(T_\Sigma) \geq \mathcal{H}^1(\Sigma)$. Finally, by construction $\partial T_\Sigma = \partial T_{\Sigma'}$ and since T_Σ is a mass-minimizer among polyhedral G-chains with the same boundary we conclude that

$$\mathcal{H}^1(\Sigma') = \mathbf{M}(T_{\Sigma'}) \geq \mathbf{M}(T_\Sigma) \geq \mathcal{H}^1(\Sigma)\,. \qquad \square$$

Once we established this equivalence, we can try to solve the Steiner tree problem via calibrations. In fact, given a "candidate minimizer" Σ for a set a points $\{x_1, \dots, x_n\} \subset \mathbb{R}^d$ and the corresponding 1-dimensional polyhedral G-chain T_Σ, it is sufficient to check the existence of a calibration associated with T_Σ. To this aim, it is useful to rephrase condition (iii) in Definition 2.8 as

$$\sup_{x \in \mathbb{R}^d} \sup_{t \in \mathbb{S}^{d-1}} \|\omega(x)t\|^* \leq 1\,, \tag{2.6}$$

where $\|\cdot\|^*$ is the dual norm in $(\mathbb{R}^{n-1}, \|\cdot\|)$.

Proposition 2.16. *Given $n \in \mathbb{N} \setminus \{0\}$ and the norm $\|\cdot\|$ on $\mathbb{R}^{n-1}$ introduced in Definition 2.13, the dual norm $\|\cdot\|^*$ on $\mathbb{R}^{n-1}$ is characterized by the formula*

$$\|w\|^* = w^P \vee w^N\,,$$

where $w = (w_1, \dots, w_{n-1})$ and $w^p := \sum_{i=1}^{n-1} w_i \vee 0$ and $w^N := -\sum_{i=1}^{n-1} w_i \wedge 0$.

The proof is very short and can be found in [6]. Nonetheless, we rewrite the proof here for the reader's ease.

Proof. Given $w \in \mathbb{R}^{n-1}$, we have to compute $\sup_{\|v\|=1} w \cdot v$. If $w^P > w^N$, then the supremum is reached for $v = \sum_{i=1}^{n-1}(w_i \vee 0)/w_i g_i$, which implies $w \cdot v = w^P$. If $w^N > w^P$ instead, the supremum is reached for $v = \sum_{i=1}^{n-1}(w_i \wedge 0)/w_i g_i$, which implies $w \cdot v = w^N$. $\square$

3 Discrete parametrization of calibrations

Consider $T = (\Sigma, \tau, \theta)$ a polyhedral one dimensional G-chain where G is the normed group defined in Definition 2.13. In this section, we are interested in the numerical investigation of the optimality of the Steiner tree defined by the support of T. To this purpose, we consider a polygonal partition $\mathcal{P}$ associated to the support of the polyhedral chain T. More precisely, we require that the weighted segments $\{\sigma_i\}_{i=1}^{N_\sigma}$ are all covered by the edges of $\mathcal{P}$. Notice that this assumption is not mandatory but it could help to describe discontinuous matrix fields. Consider now the matrix fields ω which are piecewise constant on the N_C polygons $\{C_r\}_{1 \leq r \leq N_C}$ of $\mathcal{P}$. If we are able to identify a matrix field ω which satisfies the divergence constraints (i') and the additional conditions (ii) and (iii) of Definition 2.8, Theorem 2.10 ensures that T is globally optimal. Since we can not expect to satisfy exactly this set of constraints from a numerical point of view, we relax this constraint satisfaction problem in a convex non smooth optimization problem which consists in minimizing the real ε with respect to couples $(\varepsilon, (\omega_r)_{1 \leq r \leq N_C})$ which satisfy

(i) $\omega_r \in \mathbb{R}^{m \times 2}$, $\forall\, 1 \leq r \leq N_C$,

(i') $(\omega_r - \omega_l).\tau_{r,l} = 0 \quad \forall\, x \in \partial C_r \cap \partial C_l$, where $\tau_{r,l}$ is a tangent vector of the common edges of C_r and C_l,

(ii) $\omega_{r(k)}\tau_k \cdot \theta_k = \|\theta_k\|$, $\forall\, 1 \leq k \leq N_\sigma$ and all $C_{r(k)}$ cells such that $C_{r(k)} \cap \sigma_k$ is not empty,

(iii) $\sup_{1\leq r\leq N_C}\left(\sup_{t\in\mathbb{S}^1}||\omega_r t||^*\right)\leq 1+\varepsilon.$

In this setting, the closeness of 0 of parameter ε can be interpreted as a qualitative indicator of the optimality of the support of T for our original Steiner tree problem. This interpretation is supported by our stability Proposition 2.11. Notice that we do not claim that every Steiner tree admits a calibration. Nevertheless, Theorems 2.10 and 2.11 assert that, if we are able to exhibit numerically a ε-calibration with ε small enough, then T is close to be optimal in term of its cost value.

We observe that conditions (i') and (ii) represent a finite set of linear equality constraints with respect to the unknown discrete matrix field $(\omega_r)_{1\leq r\leq N_C}$. On the other hand, condition (iii) cannot easily be taken into account numerically due to the mixing of l_1 and l_2 types norms. In order to tackle that difficulty, we define a new set of linear constraints which ensures that condition (iii) is satisfied: let $(t_i)_{1\leq i\leq N_t}$ be the vertices of a convex regular exterior tangent polygon to $\mathbb{S}^1$ which contains the unit disk. We look for a field $(\omega_r)_{1\leq r\leq N_C}$ which satisfies the linear equality constraints

$$\sup_{1\leq r\leq N_C}\left(\sup_{t_1,\ldots,t_{N_t}}||\omega_r t_i||^*\right)\leq 1+\varepsilon. \tag{3.1}$$

for ε as small as possible. By convexity of the norm $||\cdot||^*$, it is clear that condition (3.1) implies condition (iii), which makes $(\varepsilon,(\omega_r)_{1\leq r\leq N_C})$ be an admissible discrete ε-calibration.

4 A non smooth convex programming framework

From the previous section, we deduce that the identification of an exact discrete calibration reduces to solve a large scale finite dimensional convex minimization problem of type

$$F(K((\omega_r)_{1\leq r\leq N_C}))+H((\omega_r)_{1\leq r\leq N_C}) \tag{4.1}$$

where F is the convex characteristic function of the set of equality constraints (i') and (ii), the linear operator $K:(\mathbb{R}^{n\times 2})^{N_C}\to(\mathbb{R}^n)^{N_C\times N_t}$ is induced by the product with vectors $(t_i)_{1\leq i\leq N_t}$ and H is the vectorial convex characteristic function of the product of ball of radius 1 for the norm $||\cdot||^*$, that is, $B(||\cdot||^*,1)^{N_C\times N_t}$. H corresponds to the strongest constraint (3.1). As it has been mentioned, there is no proof of the fact that every optimal Steiner tree has an associated discrete calibration. As a consequence, the previous optimization problem may be ill posed, in the sense that it does not exist a matrix field in the intersection of the convex sets defined by (i'), (ii) and (3.1). In order to consider a well posed constraint satisfaction problem we relax problem (4.1) in considering the minimization problem

$$L((\omega_r^1)_{1\leq r\leq N_C}),(\omega_r^2)_{1\leq r\leq N_C}))=F(K((\omega_r^1)_{1\leq r\leq N_C}))+H((\omega_r^2)_{1\leq r\leq N_C})+||(\omega_r^1-\omega_r^2)_{1\leq r\leq N_C}||^2. \tag{4.2}$$

Defining the vectorial variable $\Omega=((\omega_r^1)_{1\leq r\leq N_C},(\omega_r^2)_{1\leq r\leq N_C})$ and the projection operators π_1 and π_2 on the first and second matrix fields, problem (4.2) can be seen as a minimization problem of type

$$F(\tilde{K}((\Omega)))+\tilde{H}(\Omega) \tag{4.3}$$

where $\tilde{K}=K\circ\pi_1$, $\tilde{H}(\Omega)=H(\pi_2(\Omega))+||\pi_1(\Omega)-\pi_2(\Omega)||^2$. We recall in the next section an efficient algorithm adapted to the minimization of previous non smooth convex functions.

4.1 Proximal algorithms

Several algorithms have been developed in the last decades to solve non smooth convex satisfaction problems of type (4.3). A standard approach is to use the so called ADMM algorithm

Algorithm 1 Alternating-direction method of multipliers (ADMM)

Input $\sigma, \tau, \theta > 0$

Initialization $\Omega_0 \in ((\mathbb{R}^{n\times 2})^{N_C})^2, Y_0 \in ((\mathbb{R}^n)^{N_C\times N_t})^2$, $\Omega_t = \Omega_0$

For $n = 0, 1, \dots$
$Y_{n+1} = \text{prox}_\sigma F^*(Y_n + \sigma\tilde{K}(\Omega_t))$
$\Omega_{n+1} = \text{prox}_\tau \tilde{H}(\Omega_n - \tau\tilde{K}^*(Y_{n+1}))$
$\Omega_t = \Omega_n + \theta(\Omega_{n+1} - \Omega_n)$

(Alternating-direction method of multipliers). The version of the algorithm that we describe below (see [1] for a convergence study) is able to handle functions that are *easily proximable* (see Definition 4.1) and do not require to solve a system involving K^*K, which is a crucial issue in our large scale setting. To describe ADMM algorithm more in details, we recall the definition of proximal operator:

Definition 4.1. *The proximal operator associated to a convex function $M : \mathbb{R}^m \to \mathbb{R}$ is defined by:*

$$\text{prox}_\gamma M(y) = \arg\min_{x\in\mathbb{R}^m} M(x) + \frac{1}{\gamma}||x - y||^2. \tag{4.4}$$

In our case, F is the indicator function of a convex set, and $\text{prox}_\gamma F$ coincides with the projection operator on this set. Analogously, the proximal operator of function $\tilde{H}$ is very closely related to the projection operator on the affine subspace defined by conditions (i') and (ii) (see Section 4.2). We give a complete description of ADMM approach applied to our context in 1, where we used the notation F^* to denote the Legendre transform of F. Notice that Moreau's identity makes it possible to deduce $\text{prox}_\gamma F^*$ from $\text{prox}_\gamma F$ for every τ, x by a simple computation, since:

$$x = \text{prox}_\tau M^*(x) + \tau\,\text{prox}_{\frac{1}{\tau}} M(x/\tau)$$

In all our experiments we used the parameters $\sigma = 10, \tau = \frac{0.9}{\sigma n_K}$ and $\theta = 1$, where n_K stands for a numerical evaluation of the operator norm of $\tilde{K}^*\tilde{K}$.

4.2 The proximal operator of $\tilde{H}$

For completeness, we recall a very standard approach to project on an affine subspace which corresponds to the constraints induced by equalities (i') and (ii). In this context, we need to solve a projection problem of the type

$$\min_{(x,y),\, Ax=b} ||x - x_0||^2 + ||y - y_0||^2 + ||x - y||^2 \tag{4.5}$$

where $x_0 \in \mathbb{R}^p$, $A \in \mathbb{R}^{q\times p}$ and $b \in \mathbb{R}^q$ are the data of the problem. The first order optimality condition with respect to y gives

$$y = \frac{x + y_0}{2} \tag{4.6}$$

while the first optimality condition with respect to x ensures that

$$2(2x - x_0 - y) + A^*\lambda = 0 \tag{4.7}$$

for some unknown Lagrange multiplier $\lambda \in \mathbb{R}^q$. Multiplying this last equality and using the fact that $Ax = b$, we observe that the identification of λ is equivalent to solve the linear system

$$AA^*\lambda = 3b - A(2x_0 + y_0).$$

In our case it is straightforward to establish that the matrix AA^* is non singular, by providing an algorithm to compute λ and then x by equality (4.7). Thus, projecting on an affine subspace is equivalent to solve a linear equation. Since the linear system does not change during iterations, we used a prefactorization approach to speed up the resolution of these linear systems.

4.3 Projection on the unit ball of E^*

In this short section we describe the projection operator with respect to the l_2 norm onto the unite ball of E^*. Let $v \in \mathbb{R}^{n-1}$. We want to find $w \in \mathbb{R}^{n-1}$ which minimizes $||v - w||_2^2$ and satisfies $||w||^* \leq 1$ where $|| \cdot ||^*$ is described in Proposition 2.16. Up to a permutation of the coordinates, we can assume that $v = (v_+, v_-)$ where v_+ and $-v_-$ are vectors with non-negative coefficients. Notice that v_+ and v_- do not have necessarily the same number of elements. In this setting, it is straightforward to check that the components of the optimal vector w have the same sign as the ones of v, because removing a non zero component of w preserves the vector in the unit ball of E^*. Thus, we have to solve the minimization problem

$$\min_{\substack{\sum w_+ \leq 1, -\sum w_- \leq 1, \\ w_+ \geq 0, -w_- \geq 0}} ||v_+ - w_+||^2 + ||v_- - w_-||^2 .$$

Since the latter problem is separable with respect to the positive and negative parts, we conclude that the projection w is simply equal to $(\tilde{w}_+, \tilde{w}_-)$ where $\tilde{w}_+$ and $\tilde{w}_-$ are respectively the projections on the unit ball of vectors w_+ and w_-.

5 Numerical results

We present below some results obtained by our approximation strategy for geometrical configurations drawn in Figure 7. Motivated by the theoretical results obtained in [6], we drew a special attention in identifying calibrations which are piecewise constant on large subsets of our fixed mesh. Considering the well known fact that a calibration for a given tree is also a calibration for every other globally optimal tree, we enforce not only the mesh to cover the edges of the tree but also to cover every edges obtained by "symmetric" trees which are also known to be globally optimal. By imposing these constraints to the mesh, we hope to minimize dramatically the number of triangles needed to approximate precisely piecewise constant calibrations in symmetric configurations. As it can be observed in Figures 3 and 4, we were able to recover the theoretical piecewise constant matrix field which had been introduced in [6] to prove the optimality of Steiner trees associated to the vertices of an equilateral triangle and the vertices of a square. Observe that in the latter case we recover the symmetries imposed by both global optimal solutions. As mentioned previously, by imposing that edges of the mesh cover exactly all optimal trees, we reduce the number of parameters to describe piecewise constant calibrations.

In Figures 5 and 6 we illustrate the behavior with respect to non optimal trees. The optimal values obtained after the same number of iterations were of order 10^{-2} which is significantly different from best values obtained by optimal trees.

When the set of data points has many symmetries, we may suspect that it is more likely that a calibration does not exist since a calibration has to be optimal for every globally optimal tree. It has been suspected by many authors that the vertices of a regular pentagon or of a regular hexagon may be one of the simplest situation where non existence occurs. Surprisingly, our experiments do not confirm this intuition. Actually in all our experiments, the minimal value of our cost function 4.2 was less than 10^{-5} after 10^6 iterations which may suggest that in all these cases a calibration exists. We represent the coefficients of the optimal matrix fields obtained by our algorithm: see the figure 4 and the figures of the appendix.

One unexpected fact is the existence of apparently non piecewise constant calibration for the case of the of the pentagon and of the hexagon (see Figures 3 and 4).

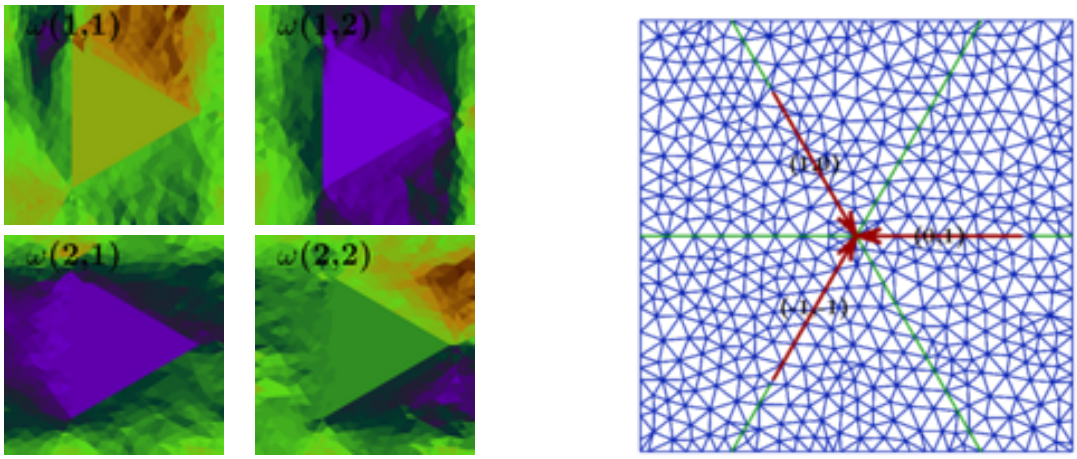


Figure 3: Representation of an optimal calibration associated to three equilateral points

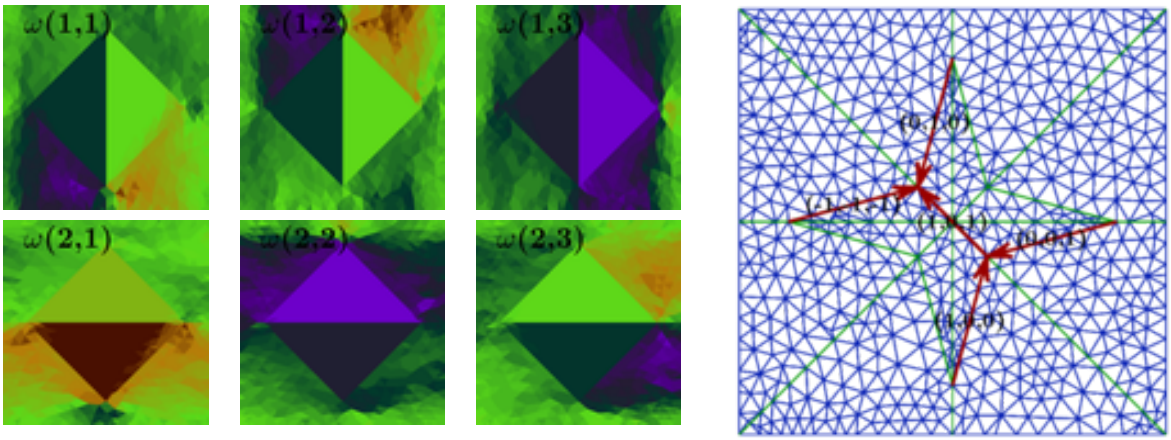


Figure 4: Representation of an optimal calibration associated to the four vertices of a square

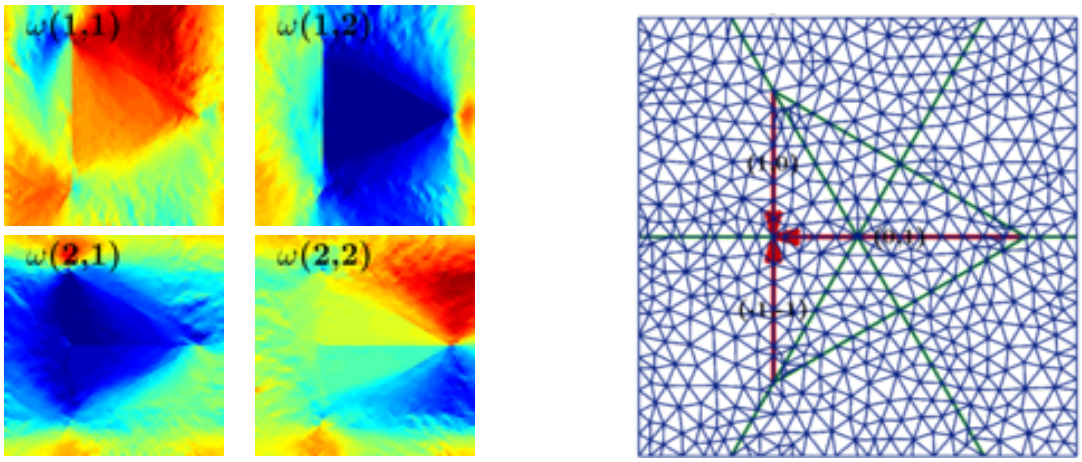


Figure 5: Representation of an optimal matrix field which does not calibrate the non optimal tree on the right (optimal value >> 1)

Acknowledgements

We would like to thank the MmgTools team (see `https://github.com/MmgTools/mmg`): Charles Dapogny, Cécile Dobrzynski, Pascal Frey and Algiane Froehly) for providing the remeshing soft-

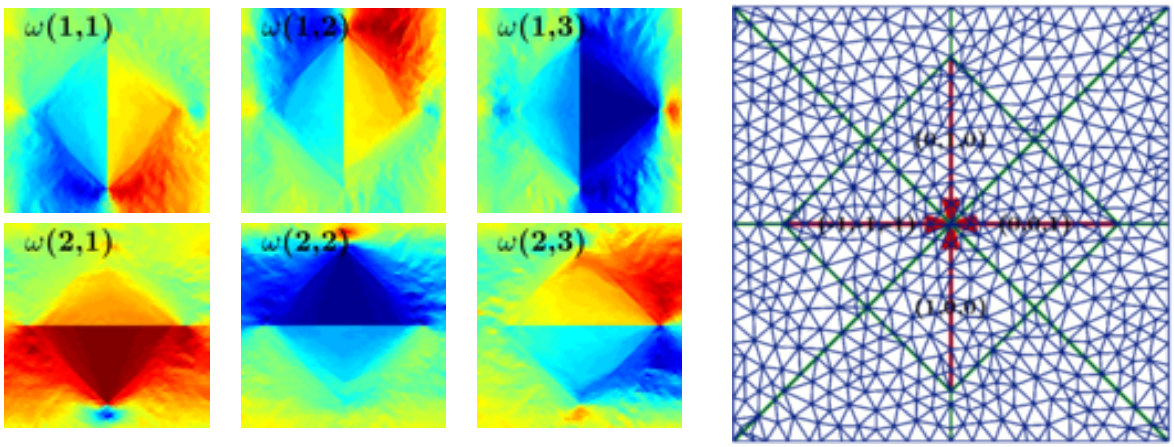


Figure 6: Representation of an optimal matrix field which does not calibrate the non optimal tree on the right (optimal value >> 1)

ware required to generate constraint meshes. Édouard Oudet gratefully acknowledges the support of the ANR, through the projects COMEDIC, GEOMETRYA, PGMO and OPTIFORM.

References

[1] Chambolle, Antonin and Pock, Thomas, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, **40.1**, (2011), 120–145.

[2] Federer, Herbert, *Geometric measure theory*, (Springer-Verlag New York Inc., New York), (1969).

[3] Gilbert, Edgar N. and Pollak, Henry O., *Steiner minimal trees*, SIAM Journal on Applied Mathematics, **16**, (1968), 1–29.

[4] Ivanov, Alexandr O. and Tuzhilin, Alexei A., *Minimal networks. The Steiner problem and its generalizations* (CRC Press, Boca Raton, FL), (1994).

[5] Karp, Richard M., *Reducibility among combinatorial problems*, Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972), (1972), 85–103.

[6] Marchese, Andrea and Massaccesi, Annalisa, *The Steiner tree problem revisited through rectifiable G-currents*, Advances in Calculus of Variations, **9.1**, (2016), 19–39.

[7] Raguet, Hugo and Fadili, Jalal and Peyré, Gabriel, *A generalized forward-backward splitting*, SIAM Journal on Imaging Sciences, **6.3**, (2013), 1199–1226.

[8] Warme, DM and Winter, P and Zachariasen, M, *GeoSteiner 3.1*, Department of Computer Science, University of Copenhagen (DIKU), (2001).

Annalisa Massaccesi: Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190 8057 Zürich - Switzerland
`annalisa.massaccesi@math.uzh.ch`

Edouard Oudet: Laboratoire Jean Kuntzmann (LJK), Université Joseph Fourier
Tour IRMA, BP 53, 51 rue des Mathématiques, 38041 Grenoble Cedex 9 - FRANCE
edouard.oudet@imag.fr
http://www-ljk.imag.fr/membres/Edouard.Oudet/

Bozhidar Velichkov: Laboratoire Jean Kuntzmann (LJK), Université Joseph Fourier
Tour IRMA, BP 53, 51 rue des Mathématiques, 38041 Grenoble Cedex 9 - FRANCE
bozhidar.velichkov@imag.fr
http://www.velichkov.it

Appendices

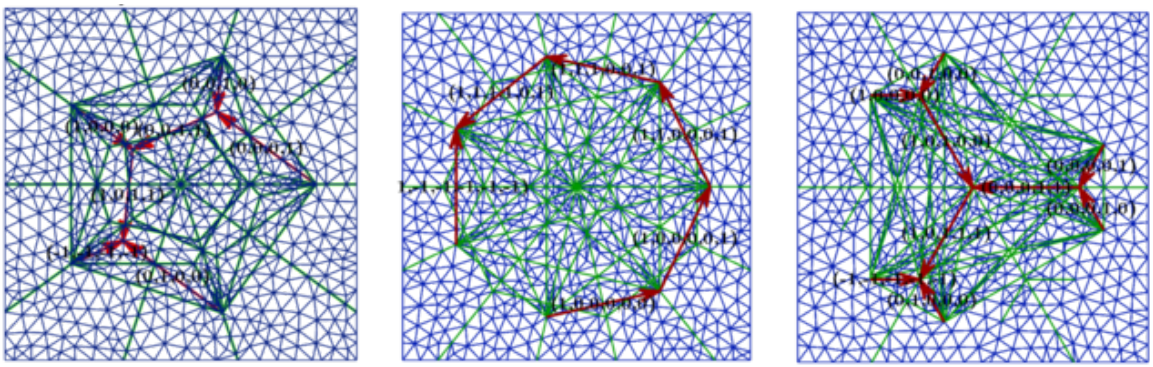


Figure 7: Three optimal Steiner trees and the associated constrained meshes

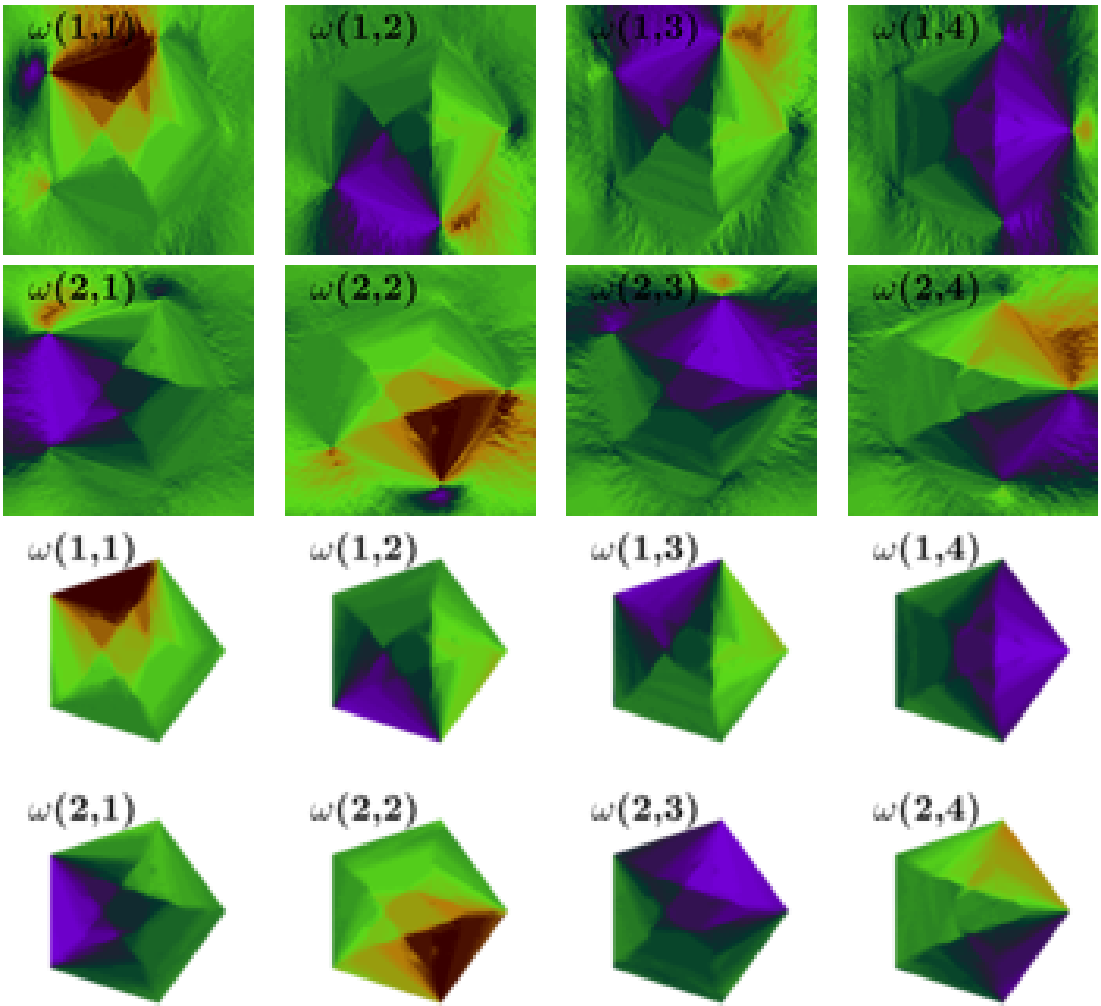


Figure 8: Representation of an optimal calibration associated to the five vertices of a Pentagon (see the first mesh of figure 7)

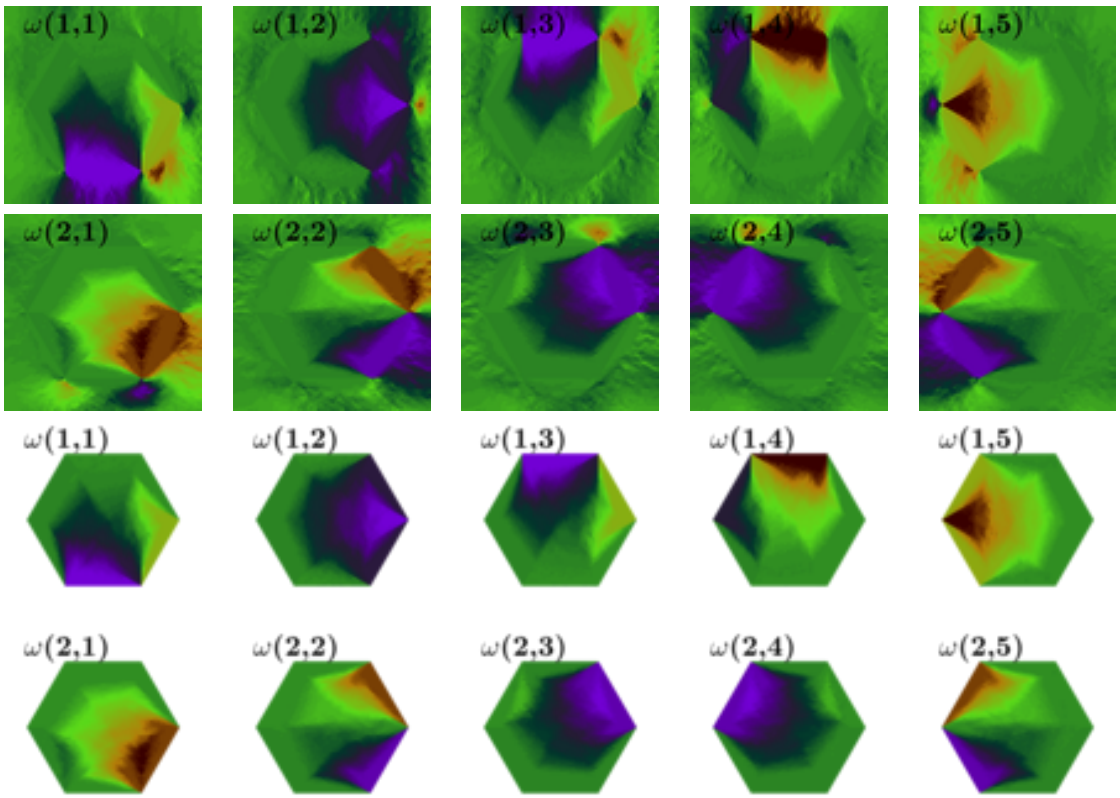


Figure 9: Representation of an optimal calibration associated to the six vertices of an Hexagon (see the second mesh of figure 7)

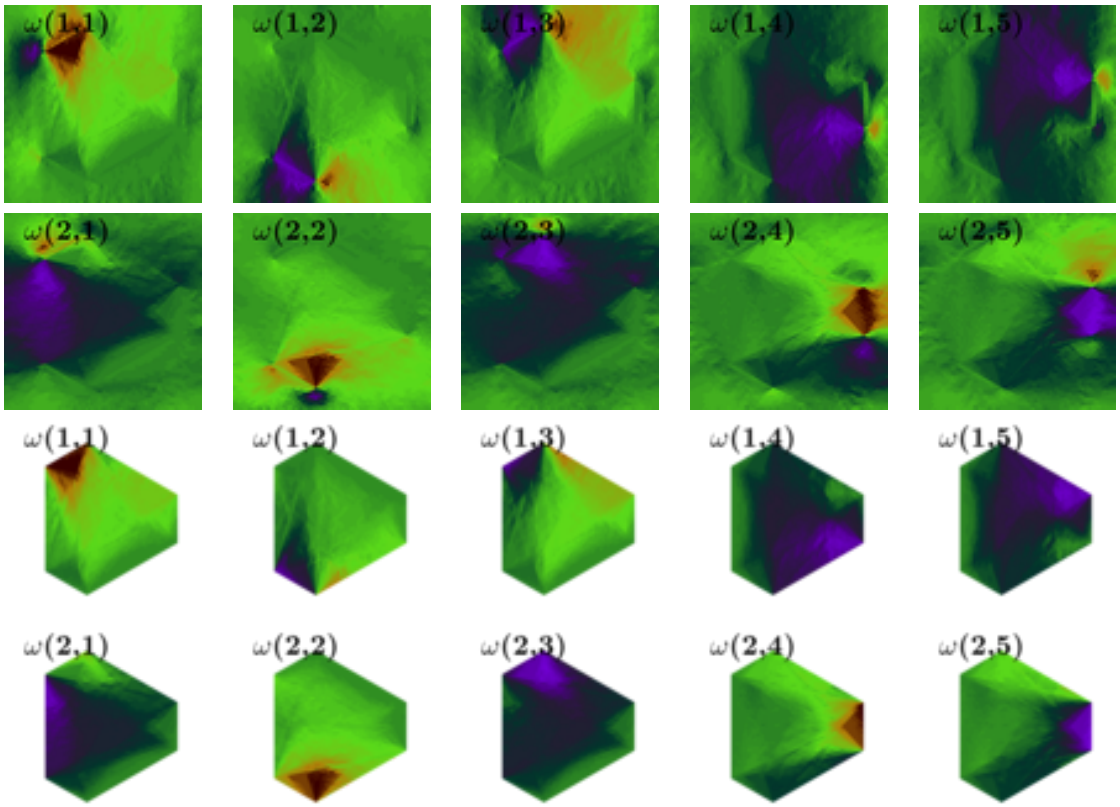


Figure 10: Representation of an optimal calibration associated to the six vertices (see the third mesh of figure 7)