A Projection Method on Measures Sets

Nicolas Chauffert, Philippe Ciuciu, Jonas Kahn, Pierre Weiss

A projection method on measures sets

Nicolas Chauffert [*], Philippe Ciuciu [†], Jonas Kahn [‡], Pierre Weiss [§]

January 31, 2016

Abstract

We consider the problem of projecting a probability measure π on a set $\mathcal{M}_N$ of Radon measures. The projection is defined as a solution of the following variational problem:

$$\inf_{\mu \in \mathcal{M}_N} \|h \star (\mu - \pi)\|_2^2,$$

where $h \in L^2(\Omega)$ is a kernel, $\Omega \subset \mathbb{R}^d$ and $\star$ denotes the convolution operator. To motivate and illustrate our study, we show that this problem arises naturally in various practical image rendering problems such as stippling (representing an image with N dots) or continuous line drawing (representing an image with a continuous line). We provide a necessary and sufficient condition on the sequence $(\mathcal{M}_N)_{N \in \mathbb{N}}$ that ensures weak convergence of the projections $(\mu_N^*)_{N \in \mathbb{N}}$ to π. We then provide a numerical algorithm to solve a discretized version of the problem and show several illustrations related to computer-assisted synthesis of artistic paintings/drawings.

Keywords — Constructive quantization, measure theory, nonconvex optimization, halftoning, continuous line drawing.

1 Introduction

Digital Halftoning consists of representing a grayscale image with only black and white tones [30]. For example, a grayscale image can be approximated by a variable distribution of black dots over a white background. This technique, called stippling, is the cornerstone of most printing digital inkjet devices. A stippling result is displayed in Figure 1b. The lion in Figure 1a can be recognized from the dotted image shown in Figure 1b. This is somehow surprising since the differences between the pixel values of the two images are far from zero. One way to explain this phenomenon is to invoke the multiresolution feature of the human visual system [8, 24]. Figures 1c and 1d are blurred versions of Figures 1a and 1b respectively. These blurred images correspond to low-pass versions of the original ones and are nearly impossible to distinguish.

---

[*]e-mail: nicolas.chauffert@gmail.com
[†]e-mail: philippe.ciuciu@gmail.com
[‡]e-mail: jonas.kahn@math.univ-lille1.fr
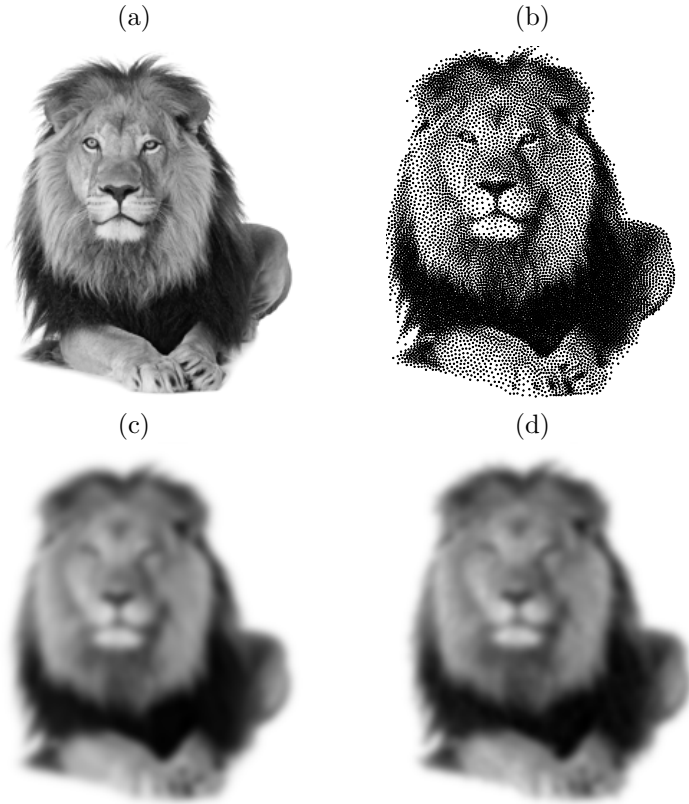[§]e-mail: pierre.armand.weiss@gmail.com

Figure 1: Explanation of the stippling phenomenon. Images (a) and (b) are similar while the norm of their difference is large. Figures (c) and (d) are obtained by convolving (a) and (b) with a Gaussian of variance equal to 3 pixels. After convolution, the images cannot be distinguished.

Assuming that the dots correspond to Dirac masses, this experiment suggests placing the dots at locations $p_1, \ldots, p_N$ corresponding to the minimizer of the following variational problem:

$$\min_{(p_1,\ldots,p_N)\in\Omega^N} \left\| h \star \left(\pi - \frac{1}{N} \sum_{i=1}^{N} \delta_{p_i} \right) \right\|_2^2 \tag{1}$$

where $\Omega \subset \mathbb{R}^2$ denotes the image domain, δ_{p_i} denotes the Dirac measure at point $p_i \in \mathbb{R}^2$, π denotes the target probability measure (the lion) and h is a convolution kernel that should depend on the point spread function of the human visual system. By letting

$$\mathcal{M}(\Omega^N) = \left\{ \mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{p_i}, \ (p_i)_{1 \leq i \leq N} \in \Omega^N \right\} \tag{2}$$

denote the set of N-point measures, problem (1) rereads as a projection problem:

$$\min_{\mu \in \mathcal{M}(\Omega^N)} \| h \star (\pi - \mu) \|_2^2 . \tag{3}$$

This variational problem is a prototypical example that motivates our study. As explained later, it is intimately related to recent works on image halftoning by means of attraction-repulsion potentials proposed in [26, 28, 14]. In references [12, 10, 11] this principle is shown to have far reaching applications ranging from numerical integration, quantum physics, economics (optimal location of service centers) or biology (optimal population distributions).

In this paper, we extend this variational problem by replacing $\mathcal{M}(\Omega^N)$ with an arbitrary set of measures denoted $\mathcal{M}_N$. In other words, we want to approximate a given measure π by another measure in the set $\mathcal{M}_N$. We develop an algorithm that is shown to converge to critical points of this projection problem in a general setting.

To motivate this extension, we consider a practical problem: how to perform continuous line drawing with a computer? Continuous line drawing is a starting course in all art cursus. It consists of drawing a picture without ever lifting the pencil from the page. Figure 2 shows two drawings obtained with this technique. It is also used in marketing, quilting designs, steel wire sculptures, connect the dot puzzles,... A few algorithms were already proposed in [20, 33, 15, 5, 32]. We propose an original solution which consists of setting $\mathcal{M}_N$ as a space of pushforward measures associated with sets of parameterized curves.

Apart from the two rendering applications discussed above, the proposed methodology has potential for diverse applications in fields such as imaging, finance, biology,... As an application example, the interested reader can have a look at our recent preprint on the generation of sampling schemes in magnetic resonance imaging [6].
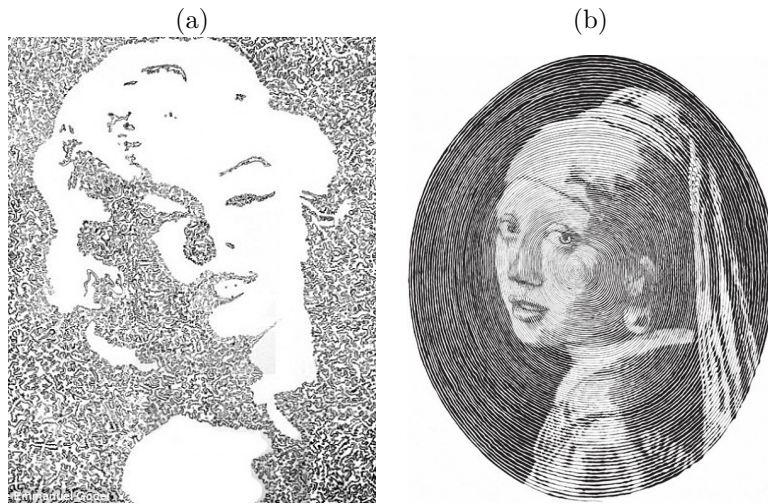
Figure 2: Two examples of continuous line drawing. (a) A sketch of Marylin Monroe by Pierre Emmanuel Godet `http://pagazine.com/` using a continuous line. A close inspection reveals that the line represents objects and characters. (b) Meisje met de Parel, Vermeer 1665, represented using a spiral with variable width. Realized by Chan Hwee Chong `http://www.behance.net/Hweechong`.

The remaining of this paper is structured as follows. We first describe the

notation and some preliminary remarks in Section 2. We propose a mathematical analysis of the problem for generic sequences of measures spaces $(\mathcal{M}_N)_{N\in\mathbb{N}}$ in Section 3. We propose a generic numerical algorithm in Section 4 and derive some of its theoretical guarantees. In Section 5, we study the particular problem of continuous line drawing from a mathematical perspective. Finally, we present some results in image rendering problems in Section 6.

2 Notation and preliminaries

In this paper, we work on the measurable space (Ω, Σ), where $\Omega = \mathbb{T}^d$ denotes the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. An extension to other spaces such as $\mathbb{R}^d$ or $[0,1]^d$ is feasible but requires slight adaptations. Since drawing on a donut is impractical, we will set $\Omega = [0,1]^d$ in the numerical experiments.

The space of continuous functions on Ω is denoted $\mathcal{C}(\Omega)$. The Sobolev space $(W^{m,p}([0,T]))^d$, where $m \in \mathbb{N}$, is the Banach space of d dimensional curves in Ω with derivatives up to the m-th order in $L^p([0,T])$. Let $\mathcal{M}_\Delta$ denote the space of probability measures on Ω, i.e. the space of nonnegative Radon measures p on Ω such that $p(\Omega) = 1$. Throughout the paper $\pi \in \mathcal{M}_\Delta$ will denote a *target measure*. Let $\mathcal{M}$ denote the space of signed measures on Ω with bounded total variation, that is $\mu = \mu_+ - \mu_-$ where μ_+ and μ_- are two finite nonnegative Radon measures and $\|\mu\|_{TV} = \mu_+(\Omega) + \mu_-(\Omega) < \infty$.

Let $h : \Omega \to \mathbb{R}$ denote a continuous function. Let $\mu \in \mathcal{M}$ denote an arbitrary finite signed measure. The convolution product between h and μ is defined for all $x \in \Omega$ by:

$$\mu \star h(x) := \int_\Omega h(x-y)d\mu(y) \tag{4}$$
$$= \mu(h(x - \cdot))$$

In the Fourier space, the convolution (4) translates to, for all $\xi \in \mathbb{Z}^d$ (see e.g., [16]):

$$\widehat{\mu \star h}(\xi) = \hat{\mu}(\xi)\hat{h}(\xi),$$

where $\hat{\mu}$ is the Fourier-Stieltjes series of μ. The Fourier-Stieltjes series coefficients are defined for all $\xi \in \mathbb{Z}^d$ by:

$$\hat{\mu}(\xi) := \int_\Omega e^{-2i\pi\langle \xi, x\rangle} \, d\mu(x).$$

We recall the Parseval formula:

$$\int_\Omega |h(x)|^2 \, dx = \sum_{\xi \in \mathbb{Z}^d} \left|\hat{h}(\xi)\right|^2.$$

Let $J : \mathbb{R}^n \to \mathbb{R}$ denote a function and ∂J its limiting-subdifferential (or simply subdifferential) [22, 1]. Let $C \subseteq \mathbb{R}^n$ denote a closed subset. The indicator function of C is denoted i_C and defined by

$$i_C(x) = \left\{ \begin{array}{ll} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{array} \right.$$

The set of projections of a point $x_0 \in \mathbb{R}^n$ on C is denoted $P_C(x_0)$ and defined by

$$P_C(x_0) = \underset{x \in C}{\text{Arg min}} \|x - x_0\|_2^2.$$

The notation Arg min stands for the whole set of minimizers while $\arg \min$ denotes one of the minimizers. Note that P_C is generally a point-to-set mapping except if C is convex closed, since the projection on a closed convex set is unique. The normal cone at $x \in \mathbb{R}^n$ is denoted $N_C(x)$. It is defined as the limiting-subdifferential of i_C at x. A critical point of the function $J + i_C$ is a point x^* that satisfies $0 \in \partial J(x^*) + N_C(x^*)$. This condition is necessary (but not sufficient) for x^* to be a local minimizer of $J + i_C$.

3 Mathematical analysis

Let

$$\mathcal{N}_h(\mu) := \|h \star \mu\|_2. \tag{5}$$

In this section, we study some basic properties of the following projection problem:

$$\min_{\mu \in \mathcal{M}_N} \mathcal{N}_h(\pi - \mu), \tag{6}$$

where $(\mathcal{M}_N)_{N \in \mathbb{N}}$ denotes an arbitrary sequence of measures sets in $\mathcal{M}_\Delta$.

3.1 Norm properties

We first study the properties of $\mathcal{N}_h$ on the space $\mathcal{M}$ of signed measures with bounded total variation. The following proposition shows that it is well defined provided that $h \in \mathcal{C}(\Omega)$.

Proposition 1. *Let $h \in \mathcal{C}(\Omega)$ and $\mu \in \mathcal{M}$. Then $h \star \mu \in L^2(\Omega)$.*

Proof. It suffices to remark that $\forall x \in \Omega$, $|h \star \mu(x)| \leq \|\mu\|_{TV} \|h\|_\infty < +\infty$. Therefore, $h \star \mu \in L^\infty(\Omega)$. Since Ω is bounded, $h \in L^\infty(\Omega)$ implies that $h \in L^2(\Omega)$. $\square$

Remark 1. *In fact, the result holds true for weaker hypotheses on h. If $h \in \mathcal{L}^\infty(\Omega)$, the set of bounded Borel measurable functions, $h \star \mu \in L^2(\Omega)$ since*

$$\forall x \in \Omega, \ |h \star \mu(x)| \leq \|\mu\|_{TV} \left(\sup_{x \in \Omega} |h(x)| \right) < +\infty.$$

Note that the L^∞-norm is defined with an $\operatorname{ess\,sup}$ while we used a $\sup$ in the above expression. We stick to $h \in \mathcal{C}(\Omega)$ since this hypothesis is more usual when working with Radon measures.

The following proposition gives a necessary and sufficient condition on h ensuring that $\mathcal{N}_h$ defines a norm on $\mathcal{M}$.

Proposition 2. *Let $h \in \mathcal{C}(\Omega)$. The mapping $\mathcal{N}_h$ defines a norm on $\mathcal{M}$ if and only if all Fourier series coefficients $\hat{h}(\xi)$ are nonzero.*

Proof. Let us assume that $\hat{h}(\xi) \neq 0$, $\forall \xi \in \mathbb{Z}^d$. The triangle inequality and absolute homogeneity hold trivially. Let us show that $\mu \neq 0 \Rightarrow \mathcal{N}_h(\mu) \neq 0$. The Fourier series of a nonzero signed measure μ is nonzero, so that there is $\xi \in \mathbb{Z}^d$ such that $\hat{\mu}(\xi) \neq 0$. According to our hypothesis $\hat{h}(\xi) \neq 0$, hence $\widehat{\mu \star h}(\xi) \neq 0$ and $\mathcal{N}_h(\mu) \neq 0$.

On the contrary, if there exists $\xi_0 \in \mathbb{Z}^d$ such that $\hat{h}(\xi_0) = 0$. The non-zero measure defined through its Fourier series by

$$\hat{\mu}(\xi) = \left\{ \begin{array}{ll} 1 & \text{if } \xi = \xi_0 \\ 0 & \text{otherwise} \end{array} \right.$$

satisfies $\mathcal{N}_h(\mu) = 0$ and belongs to $\mathcal{M}$. $\qquad \square$

From now on, owing to Proposition 2, we will systematically assume - sometimes without mentioning - that $h \in \mathcal{C}(\Omega)$ and that $\hat{h}(\xi) \neq 0$, $\forall \xi \in \mathbb{Z}^d$. Finally, we show that $\mathcal{N}_h$ induces the weak topology on $\mathcal{M}$. Let us first recall the definition of weak convergence.

Definition 1. *A sequence of measures $(\mu_N)_{N \in \mathbb{N}}$ is said to weakly converge to $\mu \in \mathcal{M}$, if*

$$\lim_{N \to \infty} \int_\Omega f(x) d\mu_N(x) = \int_\Omega f(x) d\mu(x)$$

for all continuous functions $f : \Omega \to \mathbb{R}$. The shorthand notation for weak convergence is

$$\mu_N \underset{N \to \infty}{\rightharpoonup} \mu.$$

Proposition 3. *Assume that $h \in \mathcal{C}(\Omega)$ and that $\hat{h}(\xi) \neq 0$, $\forall \xi \in \mathbb{Z}^d$. Then for all sequences $(\mu_N)_{N \in \mathbb{N}}$ in $\mathcal{M}$ satisfying $\|\mu_N\|_{TV} \leq M < +\infty$, $\forall N \in \mathbb{N}$,*

$$\lim_{N \to \infty} \mathcal{N}_h(\mu_N) = 0 \quad \Leftrightarrow \quad \mu_N \underset{N \to \infty}{\rightharpoonup} 0.$$

Proof. Let $(\mu_N)_{N \in \mathbb{N}}$ be a sequence of signed measures in $\mathcal{M}$.

If $\mu_N \rightharpoonup 0$, then $\hat{\mu}_N(\xi) = \mu_N(e^{i2\pi\langle \xi, \cdot \rangle}) \to 0$ for all $\xi \in \mathbb{Z}^d$. Since $|\hat{\mu}_N(\xi)\hat{h}(\xi)| \leq 2M|\hat{h}(\xi)|$ for all $\xi \in \mathbb{Z}^d$ and $\displaystyle\sum_{\xi \in \mathbb{Z}^d} |2M\hat{h}(\xi)|^2 < \infty$, dominated convergence yields that $\mathcal{N}_h(\mu_N) \to 0$.

Conversely, assume that $\mathcal{N}_h(\mu_N) \to 0$. Since the μ_N are bounded, there are subsequences μ_{N_s} that converge weakly to a measure ν that depends on the subsequence. We have to prove that $\nu = 0$ for all such subsequences. Since $\mathcal{N}_h(\mu_N) \to 0$, we have $\hat{\mu}_N(\xi) \to 0$ for all $\xi \in \mathbb{Z}^d$. Therefore, $\hat{\nu}(\xi) = 0$, $\forall \xi \in \mathbb{Z}^d$. This is equivalent to $\nu = 0$ (see e.g. [16, p.36]), ending the proof. $\qquad \square$

3.2 Existence of solutions

The first important question one may ask is whether Problem (6) admits a solution or not. Proposition 4 provides sufficient conditions for existence to hold.

Proposition 4. *If $\mathcal{M}_N$ is weakly compact, then Problem (6) admits at least a solution. In particular, if $\mathcal{M}_N$ is weakly closed and bounded in TV-norm, Problem (6) admits at least a solution.*

Proof. Assume $\mathcal{M}_N$ is weakly compact. Consider a minimizing sequence $\mu_n \in \mathcal{M}_N$. By compacity, there is a $\mu \in \mathcal{M}_N$ and a subsequence $(\mu_{n_k})_{k \in \mathbb{N}}$ such that $\mu_{n_k} \underset{k \to +\infty}{\rightharpoonup} \mu$. By Proposition 3, $\mathcal{N}_h$ induces the weak topology on any TV-bounded set of signed measures, so that $\lim_{k \to \infty} \mathcal{N}_h(\pi - \mu_k) = \mathcal{N}_h(\pi - \mu)$.

Since closed balls in TV-norms are weakly compact, any weakly closed TV-bounded set is weakly compact. $\square$

A key concept that will appear in the continuous line drawing problem is that of pushforward or *empirical measure* [4] defined hereafter. Let (X, γ) denote an arbitrary probability space. Given a function $p : X \to \Omega$, the empirical measure associated with p is denoted $p_*\gamma$. It is defined for any measurable set B by

$$p_*\gamma(B) \coloneqq \gamma(p^{-1}(B)),$$

where γ denotes the Lebesgue measure on the interval $[0, 1]$. Intuitively, the quantity $p_*\gamma(B)$ represents the "time" spent by the function p in B. Note that $p_*\gamma$ is a probability measure since it is positive and $p_*\gamma(\Omega) = 1$. Given a measure μ of kind $\mu = p_*\gamma$, the function p is called *parameterization* of μ.

Let $\mathcal{P}$ denote a set of parameterizations $p : X \to \Omega$ and $\mathcal{M}(\mathcal{P})$ denote the associated set of pushforward-measures:

$$\mathcal{M}(\mathcal{P}) \coloneqq \{\mu = p_*\gamma, p \in \mathcal{P}\}.$$

In the rest of this paragraph we give sufficient conditions so that a projection on $\mathcal{M}(\mathcal{P})$ exists. We first need the following proposition.

Proposition 5. *Let $(p_n)_{n \in \mathbb{N}}$ denote a sequence in $\mathcal{P}$ that converges to p pointwise. Then $(p_{n*}\gamma)_{n \in \mathbb{N}}$ converges weakly to $p_*\gamma$.*

Proof. Let $f \in \mathcal{C}(\Omega)$. Since Ω is compact, f is bounded. Hence dominated convergence yields $\int_X f(p_n(x)) - f(p(x))d\gamma(x) \to 0$. $\square$

Proposition 6. *Assume that $\mathcal{P}$ is compact for the topology of pointwise convergence. Then there exists a minimizer to Problem (6) with $\mathcal{M}_N = \mathcal{M}(\mathcal{P})$.*

Proof. By Proposition 4 it is enough to show that $\mathcal{M}(\mathcal{P})$ is weakly compact. First, $\mathcal{M}(\mathcal{P})$ is bounded in TV-norm since it is a subspace of probability measures. Consider a sequence $(p_n)_{n \in \mathbb{N}}$ in $\mathcal{P}$ such that the sequence $(p_{n*}\gamma)_{n \in \mathbb{N}}$ weakly converges to a measure μ. Since $\mathcal{P}$ is compact for the topology of pointwise convergence, there is a subsequence $(p_{n_k})_{k \in \mathbb{N}}$ converging pointwise to $p \in \mathcal{P}$. By Proposition 5, the pushforward-measure $p_*\gamma = \mu$ so that $\mu \in \mathcal{M}(\mathcal{P})$ and $\mathcal{P}$ is weakly closed. $\square$

3.3 Consistency

In this paragraph, we consider a sequence $(\mathcal{M}_N)_{N \in \mathbb{N}}$ of weakly compact subsets of $\mathcal{M}_\Delta$. By Proposition 4 there exists a minimizer $\mu_N^* \in \mathcal{M}_N$ to Problem (6) for every N. We study conditions on $(\mathcal{M}_N)_{N \in \mathbb{N}}$ that ensure consistency, i.e. $\mu_N^* \underset{N \to \infty}{\rightharpoonup} \pi$. In the case of image rendering, it basically means that if N is taken sufficiently large, the projection μ_N^* and the target image π will be indistinguishable from a perceptual point of view.

In order to evaluate distances between μ_N^* and π, the most natural metric is the minimized norm $\mathcal{N}_h(\mu_N^* - \pi)$. However, its analysis is easy in the Fourier domain, whereas all measures sets in this paper are defined in the space domain. We therefore prefer to use another metrization of weak convergence, given by the transportation distance. Moreover we will see in Theorem 1 that the transportation distance defined below dominates $\mathcal{N}_h$.

Definition 2. *The L^1 transportation distance, also known as Kantorovitch or Wasserstein distance, between two measures with same TV norm is given by:*

$$W_1(\mu, \nu) := \inf_c \int \|x - y\|_1 \, \mathrm{d}c(x, y)$$

where the infimum runs over all couplings of μ and ν, that is the measures c on $\Omega \times \Omega$ with marginals satisfying $c(A, \Omega) = \mu(A)$ and $c(\Omega, A) = \nu(A)$ for all Borelians A.

Equivalently, we may define the distance through the dual, that is the action on Lipschitz functions:

$$W_1(\mu, \nu) = \sup_{f : Lip(f) \leq 1} \mu(f) - \nu(f). \tag{7}$$

We define the point-to-set distance as

$$W_1(\mathcal{M}_N, \pi) := \inf_{\mu \in \mathcal{M}_N} W_1(\mu, \pi).$$

Obviously this distance satisfies:

$$W_1(\mathcal{M}_N, \pi) \leq \delta_N := \sup_{\pi \in \mathcal{M}_\Delta} \inf_{\mu \in \mathcal{M}_N} W_1(\mu, \pi). \tag{8}$$

Theorem 1. *Assume that $h \in \mathcal{C}(\Omega)$ denote a Lipschitz continuous function with Lipschitz constant L. Then*

$$\mathcal{N}_h(\mu - \pi) \leq L W_1(\mu, \pi) \tag{9}$$

and

$$\mathcal{N}_h(\mu_N^* - \pi) \leq L W_1(\mathcal{M}_N, \pi) \leq L \delta_N. \tag{10}$$

Proof. Let $\tau_x : h(\cdot) \mapsto h(x - \cdot)$ denote the symmetrization and shift operator. Let us first prove inequality (9):

$$\begin{aligned} \|h \star (\mu - \pi)\|_2^2 &= \int_\Omega [h \star (\mu - \pi)(x)]^2 \, dx \\ &= \int_\Omega |\mu(\tau_x h) - \pi(\tau_x h)|^2 \, dx \\ &\leq |\Omega| L^2 W_1^2(\mu, \pi), \end{aligned}$$

where we used the dual definition (7) of the Wasserstein distance to obtain the last inequality.

Let μ_N denote a minimizer of $\inf_{\mu \in \mathcal{M}_N} W_1(\mu, \pi)$. If no minimizer exists we may take an ϵ-solution with arbitrary small ϵ instead. By definition of the projection μ_N^*, we have:

$$\mathcal{N}_h(\mu_N^* - \pi) \leq \mathcal{N}_h(\mu_N - \pi) \leq W(\mu_N, \pi) \leq \delta_N. \tag{11}$$

$\square$

Even though the bound (10) is pessimistic in general, it provides some insight on which sequences of measures spaces allow a fast weak convergence.

3.4   Application to image stippling

In order to illustrate the proposed theory, we first focus on the case of N-point measures $\mathcal{M}(\Omega^N)$ defined in Eq. 2. This setting is the standard one considered for probability quantization (see [13, 18] for similar results). As mentioned earlier, it has many applications including image stippling. Our main results read as follows.

Theorem 2. *Let h denote an L-Lipschitz kernel. The set of N-point measures $\mathcal{M}(\Omega^N)$ satisfies the following inequalities:*

$$\delta_N = \sup_{\pi \in \mathcal{M}_\Delta} \inf_{\mu \in \mathcal{M}(\Omega^N)} W_1(\mu, \pi) \leq \left(\frac{\sqrt{d}}{2} + 1 \right) \frac{1}{N^{1/d} - 1} \tag{12}$$

and

$$\sup_{\pi \in \mathcal{M}_\Delta} \inf_{\mu \in \mathcal{M}(\Omega^N)} \mathcal{N}_h(\mu - \pi) \leq L \left(\frac{\sqrt{d}}{2} + 1 \right) \frac{1}{N^{1/d} - 1}. \tag{13}$$

As a direct consequence, we get the following corollary.

Corollary 1. *Let $\mathcal{M}_N = \mathcal{M}(\Omega^N)$ denote the set of N-point measures. Then there exist solutions μ_N^* to the projection problem (6). Moreover, for any L-Lipschitz kernel $h \in \mathcal{C}(\Omega)$:*

i) $\mu_N^* \underset{N \to \infty}{\rightharpoonup} \pi$.

ii) $\mathcal{N}_h(\mu_N^* - \pi) = \mathcal{O}\left(L N^{-\frac{1}{d}} \right)$.

Proof. We first evaluate the bound δ_N defined in (8). To this end, for any given π, we construct an explicit sequence of measures $\mu_0, \ldots, \mu_N$, the last of which is an N-point measure approximating π.

Note that $\mathbb{T}^d$ can be thought of as the unit cube $[0,1)^d$. It may therefore be partitioned in C^d smaller cubes of edge length $1/C$ with $C = \lfloor N^{1/d} \rfloor$. We let $(\omega_i)_{1 \leq i \leq C^d}$ denote the small cubes and x_i denote their center. We assume that the cubes are ordered in such a way that ω_i and ω_{i+1} are contiguous.

We define $\mu_0 = \sum_{i=1}^{C^d} \pi(\omega_i) \delta_{x_i}$. The measure μ_0 satisfies

$$W_1(\pi, \mu_0) \leqslant \frac{1}{2} \sup_i \text{Diameter}(\omega_i)$$

$$\leqslant \frac{\sqrt{d}}{2} \lfloor N^{1/d} \rfloor^{-1}$$

$$\leqslant \frac{\sqrt{d}}{2} \frac{1}{N^{1/d} - 1},$$

but is not an N-point measure since $N\pi(\omega_i)$ is not an integer.

To obtain an N-point measure, we recursively build μ_l as follows:

$$\mu_l(\{x_l\}) = \frac{1}{N} \lfloor N\mu_{l-1}(\{x_l\}) \rfloor,$$

$$\mu_l(\{x_{l+1}\}) = \mu_{l-1}(\{x_{l+1}, x_l\}) - \frac{1}{N} \lfloor N\mu_{l-1}(\{x_l\}) \rfloor$$

$$\text{if } l \leq (1/C)^d - 1,$$

$$\mu_l(\{x_i\}) = \mu_{l-1}(\{x_i\}) \quad \text{if } i \notin \{l, l+1\}.$$

We stop the process for $l = (1/C)^d$ and let $\tilde{\mu} = \mu_{(1/C)^d}$. Notice that $N\mu_l(x_i)$ is an integer for all $i \leqslant l$ and that μ_l is a probability measure for all l. Therefore $\tilde{\mu}$ is an N-point measure. Moreover:

$$W_1(\mu_l, \mu_{l+1}) \leqslant \frac{1}{N} \|x_l - x_{l+1}\|_2$$

$$\leqslant \frac{1}{N(N^{1/d} - 1)}.$$

Since the transportation distance is a distance, we have the triangle inequality. Therefore:

$$W_1(\pi, \tilde{\mu}) \leq W_1(\pi, \mu_0) + \sum_{l=1}^{N} W_1(\mu_{l-1}, \mu_l),$$

$$= \frac{\sqrt{d}}{2} \frac{1}{N^{1/d} - 1} + N \frac{1}{N(N^{1/d} - 1)}$$

$$= \left(\frac{\sqrt{d}}{2} + 1 \right) \frac{1}{N^{1/d} - 1}.$$

The inequality (13) is a direct consequence of this result and Proposition 1.

We now turn to the proof of Corollary 1. To prove the existence, first notice that the projection problem (6) can be recast as (1). Let $p = (p_1, \cdots, p_N) \in \Omega^N$. The mapping $p \mapsto \left\| h \star \left(\pi - \frac{1}{N} \sum_{i=1}^{N} \delta_{p_i} \right) \right\|_2^2$ is continuous. Problem (1) therefore consists of minimizing a finite dimensional continuous function over a compact set. The existence of a solution follows. Point ii) is a direct consequence of Theorem 1 and bound (13). Point i) is due to the fact that $\mathcal{N}_h$ metrizes weak convergence, see Proposition 3. □

4   Numerical resolution

In this section, we propose a generic numerical algorithm to solve the projection problem (6). We first draw a connection with the recent works on electrostatic halftoning [26, 28] in subsection 4.1. We then recall the algorithm proposed in [26, 28] when $\mathcal{M}_N$ is the set of N-point measures. Finally, we extend this principle to arbitrary measures spaces and provide some results on their theoretical performance in section 4.3.

4.1   Relationship to electrostatic-halftoning

In a recent series of papers [26, 28, 12, 14], it was suggested to use electrostatic principles to perform image halftoning. This technique was shown to produce

results having a number of nice properties such as few visual artifacts. Motivated by preliminary results in [26], the authors of [28] proposed to choose the N points locations $p = (p_i)_{1 \le i \le N} \in \Omega^N$ as a solution of the following variational problem:

$$\min_{p \in \Omega^N} \underbrace{\frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} H(p_i - p_j)}_{\text{Repulsion potential}} - \underbrace{\frac{1}{N} \sum_{i=1}^{N} \int_{\Omega} H(x - p_i) \, d\pi(x)}_{\text{Attraction potential}}, \qquad (14)$$

where H was initially defined as $H(x) = \|x\|_2$ in [26, 28] and then extended to a few other functions in [12]. The attraction potential tends to attract points towards the bright regions of the image (regions where the measure π has a large mass) whereas the repulsion potential can be regarded as a counter-balancing term that tends to maximize the distance between all pairs of points. Deriving an algorithm to solve problem (14) with good precision can be seen as a generalization of Thomson's problem [29], which belongs to Smale's list of mathematical questions to solve for the XXIst century [27].

Proposition 7 below shows that this attraction-repulsion problem is actually equivalent to the projection problem (6) on the set of N-point measures defined in (2). We let $\mathcal{P}^*$ denote the set of solutions of (14) and $\mathcal{M}(\mathcal{P}^*) = \{\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{p_i^*}, \ p^* \in \mathcal{P}^*\}$. We also let $\mathcal{M}^*$ denote the set of solutions to problem (6).

Proposition 7. *Let $h \in \mathcal{C}(\Omega)$ denote a kernel such that $|\hat{h}|(\xi) > 0, \ \forall \xi \in \mathbb{Z}^d$. Define H through its Fourier series by $\hat{H}(\xi) = |\hat{h}|^2(\xi)$. Then problems (6) and (14) yield the same solutions set:*

$$\mathcal{M}^* = \mathcal{M}(\mathcal{P}^*).$$

Proof. First, note that since H and h are continuous both problems are well defined and admit at least one solution. Let us first expand the L^2-norm in (6):

$$\frac{1}{2} \|h \star (\mu - \pi)\|_2^2 = \frac{1}{2} \langle h \star (\mu - \pi), h \star (\mu - \pi) \rangle$$

$$= \frac{1}{2} \langle H \star (\mu - \pi), \mu - \pi \rangle$$

$$= \frac{1}{2} \left(\langle H \star \mu, \mu \rangle - 2 \langle H \star \mu, \pi \rangle + \langle H \star \pi, \pi \rangle \right).$$

Therefore

$$\operatorname*{Arg\,min}_{\mu \in \mathcal{M}_N} \frac{1}{2} \|h \star (\mu - \pi)\|_2^2 = \operatorname*{Arg\,min}_{\mu \in \mathcal{M}_N} \frac{1}{2} \left(\langle H \star \mu, \mu \rangle - 2 \langle H \star \mu, \pi \rangle \right).$$

To conclude, it suffices to remark that for a measure μ of kind $\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{p_i}$,

$$\frac{1}{2} \left(\langle H \star \mu, \mu \rangle - 2 \langle H \star \mu, \pi \rangle \right)$$

$$= \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} H(p_i - p_j) - \frac{1}{N} \sum_{i=1}^{N} \int_{\Omega} H(x - p_i) \, d\pi(x).$$

$\square$

Remark 2. *It is rather easy to show that a sufficient condition for h to be continuous is that $H \in \mathcal{C}^3(\Omega)$ or H be Hölder continuous with exponent $\alpha > 2$. These conditions are however strong and exclude kernels such as $H(x) = \|x\|_2$.*

From Remark 1, it is actually sufficient that $h \in \mathcal{L}^\infty(\Omega)$ for $\mathcal{N}_h$ to be well defined. This leads to less stringent conditions on H. We do not discuss this possibility further to keep the arguments simple.

Remark 3. *Corollary 1 sheds light on the approximation quality of the minimizers of attraction-repulsion functionals. Let us mention that consistency of problem* (14) *was already studied in the recent papers [12, 10, 11]. To the best of our knowledge, Corollary 1 is stronger than existing results since it yields a convergence rate and holds true under more general assumptions.*

Though formulations (6) and (14) are equivalent, we believe that the proposed one (6) has some advantages: it is probably more intuitive, shows that the convolution kernel h should be chosen depending on physical considerations and simplifies some parts of the mathematical analysis such as consistency. However, the set of admissible measures $\mathcal{M}(\Omega^N)$ has a complex geometry and this formulation as such is hardly amenable to numerical implementation. For instance, $\mathcal{M}(\Omega^N)$ is not a vector space, since adding two N-point measures usually leads to $(2N)$-point measures. On the other hand, the attraction-repulsion formulation (14) is an optimization problem of a continuous function over the set Ω^N. It therefore looks easier to handle numerically using non-linear programming techniques. This is what we will implement in the next paragraphs following previous works [26, 28].

4.2   The case of N-point measures

In this section, we develop an algorithm specific to the projection on the set of N-point measures defined in (2). This algorithm generates stippling results such as in Fig. 1. In stippling, the measure is supported by a union of discs, i.e., a sum of diracs convoluted with a disc indicator. We simply have to consider the image deconvoluted with this disc indicator as π to include stippling in the framework of N-point measures. We will generalize this algorithm to arbitrary sets of measures in the next section. We assume without further mention that $\hat{H}(\xi)$ is real and positive for all ξ. This implies that H is real and even. Moreover, Proposition 7 implies that problems (6) and (14) yield the same solutions sets. We let $p = (p_1, \ldots, p_N)$ and set

$$\tilde{J}(p) := \underbrace{\frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N H(p_i - p_j)}_{F(p)} - \underbrace{\frac{1}{N} \sum_{i=1}^N \int_\Omega H(x - p_i) \, d\pi(x)}_{\tilde{G}(p)}. \qquad (15)$$

The projection problem therefore rereads as:

$$\min_{p \in \Omega^N} \tilde{J}(p). \qquad (16)$$

For practical purposes, the integrals in $\tilde{G}(p)$ first have to be replaced by numerical quadratures. We let $G(p) \simeq \tilde{G}(p)$ denote the numerical approximation of

$\tilde{G}(p)$. This approximation can be written as

$$G(p) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{n} w_j H(x_j - p_i)\pi_j,$$

where n is the number of discretization points x_j and w_j are weights that depend on the integration rule. In particular, since we want to approximate integration with respect to a probability measure, we require that

$$\sum_{j=1}^{n} w_j \pi_j = 1.$$

In our numerical experiments we use the rectangle rule. We may then take π_j as the integral of π over the corresponding rectangle. After discretization, the projection problem therefore rereads as:

$$\min_{p \in \Omega^N} J(p) := F(p) - G(p). \tag{17}$$

The following result [1, Theorem 5.3] will be useful to design a convergent algorithm. We refer to [1] for a comprehensive introduction to the definition of Kurdyka-Łojasiewicz functions and to its applications to algorithmic analysis. In particular, we recall that semi-algebraic functions are Kurdyka-Łojasiewicz [19].

Theorem 3. *Let $K : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$ function whose gradient is L-Lipschitz continuous and let C be a nonempty closed subset of $\mathbb{R}^n$. Being given $\varepsilon \in \left(0, \frac{1}{2L}\right)$ and a sequence of stepsizes $\gamma^{(k)}$ such that $\varepsilon < \gamma^{(k)} < \frac{1}{L} - \varepsilon$, we consider a sequence $(x^{(k)})_{k \in \mathbb{N}}$ that complies with*

$$x^{(k+1)} \in P_C \left(x^{(k)} - \gamma^{(k)} \nabla K(x^{(k)}) \right), \text{ with } x^{(0)} \in C \tag{18}$$

If the function $K + i_C$ is a Kurdyka-Łojasiewicz function and if $(x^{(k)})_{k \in \mathbb{N}}$ is bounded, then the sequence $(x^{(k)})_{k \in \mathbb{N}}$ converges to a critical point x^* in C.

A consequence of this important result is the following.

Corollary 2. *Assume that H is a $\mathcal{C}^1$ semi-algebraic function with L-Lipschitz continuous gradient. Set $0 < \gamma < \frac{N}{3L}$. Then the following sequence converges to a critical point of problem* (17)

$$p^{(k+1)} \in P_{\Omega^N} \left(p^{(k)} - \gamma \nabla J(p^{(k)}) \right), \text{ with } p^{(0)} \in \Omega^N. \tag{19}$$

If H is convex, $0 < \gamma < \frac{N}{2L}$ ensures convergence to a critical point.

Remark 4. *The semi-algebraicity is useful to obtain convergence to a critical point. In some cases it might however not be needed. For instance, in the case where C is convex and closed, it is straightforward to establish the decrease of the cost function assuming only that ∇J is Lipschitz. Nesterov in [23, Theorem 3] also provides a convergence rate in $\mathcal{O}\left(\frac{1}{\sqrt{k+1}}\right)$ in terms of objective function values.*

Proof. First notice that J is semi-algebraic as a finite sum of semi-algebraic functions.

Function J is $\mathcal{C}^1$ by Leibniz integral rule. Let ∂_k denote the derivative with respect to p_k. Then, since H is even

$$\partial_k F(p) = \frac{1}{N^2} \sum_{i=1}^{N} \nabla H(p_k - p_i) \tag{20}$$

and

$$\partial_k G(p) = -\frac{1}{N} \sum_{j=1}^{n} w_j \nabla H(x_j - p_k) \pi_j. \tag{21}$$

For any two sets of N points $p^{(1)} = (p_k^{(1)})_{1 \leqslant k \leqslant N}$, $p^{(2)} = (p_k^{(2)})_{1 \leqslant k \leqslant N}$:

$$\|\nabla F(p^{(1)}) - \nabla F(p^{(2)})\|_2^2 = \sum_{k=1}^{N} \left\| \partial_k F(p^{(1)}) - \partial_k F(p^{(2)}) \right\|_2^2$$

$$= \frac{1}{N^4} \sum_{k=1}^{N} \left\| \sum_{i=1}^{N} \nabla H(p_k^{(1)} - p_i^{(1)}) - \nabla H(p_k^{(2)} - p_i^{(2)}) \right\|_2^2$$

$$\leqslant \frac{1}{N^4} \sum_{k=1}^{N} \left(\sum_{i=1}^{N} L \| p_k^{(1)} - p_i^{(1)} - (p_k^{(2)} - p_i^{(2)}) \|_2 \right)^2$$

$$\leqslant \frac{L^2}{N^4} \sum_{k=1}^{N} \left(\sum_{i=1}^{N} \| p_k^{(1)} - p_k^{(2)} \|_2 + \| p_i^{(1)} - p_i^{(2)} \|_2 \right)^2$$

$$\leqslant \frac{L^2}{N^4} \sum_{k=1}^{N} N \left(\sum_{i=1}^{N} \left(\| p_k^{(1)} - p_k^{(2)} \|_2 + \| p_i^{(1)} - p_i^{(2)} \|_2 \right)^2 \right)$$

$$\leqslant \frac{2L^2}{N^3} \sum_{k=1}^{N} \sum_{i=1}^{N} \| p_k^{(1)} - p_k^{(2)} \|_2^2 + \| p_i^{(1)} - p_i^{(2)} \|_2^2$$

$$= \frac{4L^2}{N^2} \| p^{(1)} - p^{(2)} \|_2^2,$$

and

$$\|\nabla G(p^{(1)}) - \nabla G(p^{(2)})\|_2^2 = \sum_{k=1}^{N} \left\| \partial_k G(p^{(1)}) - \partial_k G(p^{(2)}) \right\|_2^2$$

$$= \frac{1}{N^2} \sum_{k=1}^{N} \left\| \sum_{j=1}^{n} w_j \pi_j \left(\nabla H(p_k^{(1)} - x) - \nabla H(p_k^{(2)} - x) \right) \right\|_2^2$$

$$\leqslant \frac{1}{N^2} \sum_{k=1}^{N} \left(\sum_{j=1}^{n} w_j \pi_j L \| p_k^{(1)} - p_k^{(2)} \| \right)^2$$

$$= \frac{L^2}{N^2} \left(\sum_{j=1}^{n} w_j \pi_j \right) \| p^{(1)} - p^{(2)} \|_2^2$$

$$= \frac{L^2}{N^2} \| p^{(1)} - p^{(2)} \|_2^2.$$

Finally,

$$\|\nabla J(p^{(1)}) - \nabla J(p^{(2)})\|_2$$
$$\leqslant \|\nabla F(p^{(1)}) - \nabla F(p^{(2)})\|_2 + \|\nabla G(p^{(1)}) - \nabla G(p^{(2)})\|_2$$
$$\leqslant \left(\frac{2L}{N} + \frac{L}{N}\right)\|p^{(1)} - p^{(2)}\|_2 = \frac{3L}{N}\|p^{(1)} - p^{(2)}\|_2.$$

Now, if we assume that H is convex and $\mathcal{C}^2$ (this hypothesis is not necessary, but simplifies the proof). Then F and G are also convex and $\mathcal{C}^2$. We let $\nabla^2 F$ denote the Hessian matrix of F. Given the previous inequalities, we have $0 \preccurlyeq \nabla^2 F \preccurlyeq \frac{2L}{N}\text{Id}$ and $0 \preccurlyeq \nabla^2 G \preccurlyeq \frac{L}{N}\text{Id}$. Hence, the largest eigenvalue in magnitude of $\nabla^2(F - G)$ is bounded above by $\frac{2L}{N}$.

Moreover, the sequence $(x^{(k)})_{k \in \mathbb{N}}$ is bounded since Ω^N is bounded. $\qquad\square$

4.3   A generic projection algorithm

We now turn to the problem of finding a solution of (6), where $\mathcal{M}_N$ denotes our arbitrary measures set. In the previous paragraph, it was shown that critical points of $J + i_{\Omega^N}$ could be obtained with a simple projected gradient algorithm under mild assumtpions. Although this algorithm only yields critical points, they usually correspond to point configurations that are visually pleasing after only a few hundreds of iterations. For instance, the lion in Figure 1b was obtained after 500 iterations. Motivated by this appealing numerical behavior, we propose to extend this algorithm to the following abstract construction:

1. Approximate $\mathcal{M}_N$ by a subset $\mathcal{A}_n$ of n-point measures.

2. Use the generic Algorithm (18) to obtain an approximate projection μ_n^* on $\mathcal{A}_n$.

3. When possible, reconstruct an approximation $\mu_N \in \mathcal{M}_N$ of a projection μ_N^* using μ_n^*.

To formalize the approximation step, we need the definition of Hausdorff distance:

Definition 3. *The Hausdorff distance between two subsets X and Y of a metric space (M, d) is:*

$$\mathcal{H}_d(X, Y) := \max\left\{\sup_{x \in X}\inf_{y \in Y} d(x, y), \sup_{y \in Y}\inf_{x \in X} d(y, x)\right\}.$$

In words, two sets are close if any point in one set is close to at least a point in the other set. In this paper, the relevant metric space is the space of signed measures $\mathcal{M}$ with the norm $\mathcal{N}_h$. The corresponding Hausdorff distance is denoted $\mathcal{H}_{\mathcal{N}_h}$.

The following proposition clarifies why controlling the Hausdorff distance is relevant to design approximation sets $\mathcal{A}_n$.

Proposition 8. *Let $\mathcal{A}_n$ and $\mathcal{M}_N$ be two TV-bounded weakly closed sets of measures such that $\mathcal{H}_{\mathcal{N}_h}(\mathcal{A}_n, \mathcal{M}_N) \leq \varepsilon$. Let μ_n^* be a projection on $\mathcal{A}_n$. Then there is a point $\mu_N \in \mathcal{M}_N$ such that $\mathcal{N}_h(\mu_n^* - \mu_N) \leq \varepsilon$ and $\mathcal{N}_h(\pi - \mu_N) \leq \inf_{\mu \in \mathcal{M}_N} \mathcal{N}_h(\pi - \mu) + 2\varepsilon$.*

15

Corollary 3. *If* $\lim_{n\to\infty} \mathcal{H}_{\mathcal{N}_h}(\mathcal{A}_n, \mathcal{M}_N) = 0$, *then* $(\mu_n^*)_{n\in\mathbb{N}}$ *converges weakly along a subsequence to a solution* μ_N^* *of Problem* (6).

Proof. We first prove Proposition 8. Since $\mathcal{A}_n$ and $\mathcal{M}_N$ are bounded weakly closed, by Proposition 4, there exists at least one projection μ_n^* on $\mathcal{A}_n$ and one projection μ_N^* on $\mathcal{M}_N$.

Moreover since $\mathcal{A}_n$ and $\mathcal{M}_N$ are bounded weakly closed, they are also closed for $\mathcal{N}_h$, so that the infimum in the Hausdorff distances are attained. Hence there exists $\mu_n \in \mathcal{A}_n$ such that $\mathcal{N}_h(\mu_n - \mu_N^*) \leq \mathcal{H}_{\mathcal{N}_h}(\mathcal{A}_n, \mathcal{M}_N) \leq \varepsilon$ and $\mu_N \in \mathcal{M}_N$ such that $\mathcal{N}_h(\mu_N - \mu_n^*) \leq \varepsilon$. The proposition follows from the triangle inequality:

$$
\begin{aligned}
\mathcal{N}_h(\mu_N - \pi) &\leq \mathcal{N}_h(\mu_N - \mu_n^*) + \mathcal{N}_h(\mu_n^* - \pi) \\
&\leq \varepsilon + \mathcal{N}_h(\mu_n - \pi) \\
&\leq \varepsilon + \mathcal{N}_h(\mu_n - \mu_N^*) + \mathcal{N}_h(\mu_N^* - \pi) \\
&\leq \mathcal{N}_h(\mu_N^* - \pi) + 2\varepsilon.
\end{aligned}
$$

For the corollary, let us consider the sequence $(\mu_n^*)_{n\in\mathbb{N}}$ as n tends to infinity. Since all μ_n are in $\mathcal{M}_\Delta$, which is weakly compact, we have a subsequence that converges to μ_∞^*. Since $\mathcal{N}_h$ is a metrization of weak convergence on $\mathcal{M}_N$, this μ_∞^* is indeed a solution to Problem (6):

$$
\begin{aligned}
\mathcal{N}_h(\mu_\infty^* - \pi) &= \lim_{n\to\infty} \mathcal{N}_h(\mu_n^* - \pi) \\
&= \inf_{\mu\in\mathcal{M}_N} \mathcal{N}_h(\pi - \mu).
\end{aligned}
$$

$\square$

To conclude this section, we show that it is always possible to construct an approximation set $\mathcal{A}_n \subseteq \mathcal{M}(\Omega^n)$ with a control on the Hausdorff distance to $\mathcal{M}_N$. Let $\mathcal{M}_N^\epsilon$ denote an ϵ-enlargement of $\mathcal{M}_N$ w.r.t. the $\mathcal{N}_h$-norm, i.e.:

$$
\mathcal{M}_N^\epsilon = \cup_{\mu_N\in\mathcal{M}_N}\{\mu\in\mathcal{M}_\Delta, \mathcal{N}_h(\mu - \mu_N) \leq \epsilon\}. \tag{22}
$$

We may define an approximation set $\mathcal{A}_n^\epsilon$ as follows:

$$
\mathcal{A}_n^\epsilon = \mathcal{M}(\Omega^n) \cap \mathcal{M}_N^\epsilon. \tag{23}
$$

For sufficient large n, this set is non-empty and can be rewritten as

$$
\mathcal{A}_n^\epsilon = \left\{\mu = \frac{1}{n}\sum_{i=1}^n \delta_{p_i}, \quad \text{with } p = (p_i)_{1\leq i\leq n} \in \mathcal{P}_n^\epsilon\right\}, \tag{24}
$$

where the parameterization set $\mathcal{P}_n^\epsilon$ depends on $\mathcal{M}_N$ and ϵ. With this discretization of $\mathcal{M}_N$ at hand, one can then apply (at least formally) the following projected gradient descent algorithm:

$$
p^{(k+1)} \in P_{\mathcal{P}_n^\epsilon}\left(p^{(k)} - \gamma\nabla J(p^{(k)})\right), \quad \text{with } p^{(0)} \in \mathcal{P}_n^\epsilon. \tag{25}
$$

The following proposition summarizes the main approximation result:

Proposition 9. *Assume that h is L-Lipschitz. Set $\epsilon = \left(\frac{\sqrt{d}}{2} + 1\right)\frac{L}{n^{1/d}-1}$ and $\mathcal{A}_n = \mathcal{A}_n^\epsilon$, then*

$$
\mathcal{H}_{\mathcal{N}_h}(\mathcal{A}_n, \mathcal{M}_N) = \mathcal{O}\left(Ln^{-1/d}\right).
$$

Proof. By construction, $\mathcal{A}_n$ satisfies

$$\sup_{\mu_n \in \mathcal{A}_n} \inf_{\mu_N \in \mathcal{M}_N} \mathcal{N}_h(\mu_n - \mu_N) \leq \epsilon.$$

Let μ_N be an arbitrary measure in $\mathcal{M}_N$. By inequality (12), there exists $\mu_n \in \mathcal{M}(\Omega^n)$ such that $\mathcal{N}_h(\mu_n - \mu_N) \leq \epsilon$. Therefore μ_n also belongs to $\mathcal{A}_n^\epsilon$. This shows that

$$\sup_{\mu_N \in \mathcal{M}_N} \inf_{\mu_n \in \mathcal{A}_n} \mathcal{N}_h(\mu_n - \mu_N) \leq \epsilon.$$

$\square$

The approximation process proposed (23) is non-constructive since it does not provide any explicit formula for $\mathcal{P}_n^\epsilon$. Moreover, $\mathcal{P}_n^\epsilon$ can be an arbitrary set and the projection on $\mathcal{P}_n^\epsilon$ might not be implementable. We will provide constructive approximations for specific measures spaces in Section 5.

5 Application to continuous line drawing

In this section, we concentrate on the continuous line drawing problem described in the introduction. We first construct a set of admissible measures $\mathcal{M}_T$ that is a natural representative of artistic continuous line drawings. The index T represents the time spent to draw the picture. We then show that using this set in problem (6) ensures existence of a solution and weak convergence of the minimizers μ_T^* to any $\pi \in \mathcal{M}_\Delta$. We finish by designing a numerical algorithm to solve the problem and analyze its theoretical guarantees.

5.1 Problem formalization

Let us assume that an artist draws a picture with a pencil. The trajectory of the pencil tip can be defined as a parameterized curve $p : [0, T] \to \Omega$. The body, elbow, arm and hand are subject to non-trivial constraints [21]. The curve p should therefore belong to some admissible parameterized curves set denoted $\mathcal{P}_T$. In this paper, we simply assume that $\mathcal{P}_T$ contains curves with bounded first and second order derivatives in $L^q([0, T])$. More precisely, we consider the following sets of admissible curves:

1. Curves with bounded speed:

$$\mathcal{P}_T^{1,\infty} = \left\{ p \in (W^{1,\infty}([0, T]))^d, \ p([0, T]) \subset \Omega, \|\dot{p}\|_\infty \leq \alpha_1 \right\},$$

where α_1 is a positive real.

2. Curves with bounded speed and acceleration:

$$\mathcal{P}_T^{2,\infty} = \left\{ p \in (W^{2,\infty}([0, T]))^d, \ p([0, T]) \subset \Omega, \|\dot{p}\|_\infty \leq \alpha_1, \right.$$
$$\left. \|\ddot{p}\|_\infty \leq \alpha_2 \right\},$$

where α_1 and α_2 are positive reals. This set models rather accurately kinematic constraints that are met in vehicles. It is obviously a rough approximation of arm constraints.

3. The proposed theory and algorithm apply to a more general setting. For instance they cover the case of curves with derivatives up to an arbitrary order bounded in L^q with $q \in [1, \infty]$. We let

$$\mathcal{P}_T^{m,q} = \Big\{ p \in (W^{m,q}([0,T]))^d,\ p([0,T]) \subset \Omega,$$

$$\forall i \in \{1, \ldots, m\},\ \|p^{(i)}\|_q \leq \alpha_i \Big\}.$$

where $(\alpha_i)_{i=1\ldots m}$ are positive reals. This case will be treated only in the numerical experiments to illustrate the variety of results that can be obtained in applications.

Note that all above mentionned sets are convex. The convexity property will help deriving efficient numerical procedures.

In the rest of this section, we consider the following projection problem:

$$\inf_{\mu \in \mathcal{M}\left(\mathcal{P}_T^{m,q}\right)} \mathcal{N}_h(\mu - \pi), \tag{26}$$

with a special emphasis on the set $\mathcal{M}\left(\mathcal{P}_T^{m,\infty}\right)$ since it best describes standard kinematic constraints. This problem basically consists of finding the "best" way to represent a picture in a given amount of time T.

5.2 Existence and consistency

We first provide existence results using the results derived in Section 3 for $q = \infty$.

Theorem 4. *For any $m \in \mathbb{N}^*$, Problem (26) admits at least one solution in $\mathcal{M}\left(\mathcal{P}_T^{m,\infty}\right)$.*

Proof. From Proposition 6, it suffices to show that $\mathcal{P}_T^{m,\infty}$ is compact for the topology of pointwise convergence.

Let $(p_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_T^{m,\infty}$ that converges pointwise to p. Since p_n is in $W^{m,\infty}$, its $(m-1)$-th derivative is Lipschitz continuous. By definition of $\mathcal{P}_T^{m,\infty}$, the $p_n^{(m-1)}$ are both uniformly bounded by α_{m-1} and α_m-Lipschitz, hence equicontinuous. Next, by Ascoli's theorem, up to taking a subsequence, $p_n^{(m-1)}$ uniformly converges to a continuous $p^{(m-1)}$. Integrating yields that $p_n^{(i)} \to p^{(i)}$ uniformly for all $i \leq m-1$, so that $\|p^{(i)}\|_\infty \leq \alpha_i$ for $i \leq m-1$. Finally, a limit of L-Lipschitz functions is also L-Lipschitz, so that $\|p^{(m)}\|_\infty \leq \alpha_m$. Hence $p \in \mathcal{P}_T^{m,\infty}$, ending the proof. $\qquad\square$

Let us now turn to weak convergence.

Theorem 5. *Let T be an arbitrary positive real. Let $\mu_T^* \in \mathcal{M}\left(\mathcal{P}_T^{m,\infty}\right)$ denote any solution of Problem (26). Then, for any Lipschitz kernel $h \in \mathcal{C}(\Omega)$:*

i) $\mu_T^* \underset{T\to\infty}{\rightharpoonup} \pi$,

ii) $\mathcal{N}_h(\mu_T^* - \pi) = \mathcal{O}\left(T^{-\frac{m}{m(d+1)-1}}\right)$.

Proof. Let us consider a function $u : [0,1] \to \mathbb{R}$ such that:

- The m-th derivative is bounded by α_m, that is $\|u^{(m)}\|_\infty \leq \alpha_m$.

18

- For all integers $i \in \{1, \ldots, m - 1\}$, endpoint values are zero, that is $u^{(i)}(0) = u^{(i)}(1) = 0$.

- Start point is zero, that is $u(0) = 0$.

- Endpoint is positive, that is $u(1) = C > 0$.

Let x and y in Ω, such that $\|x - y\|_2 = Cr^m$, and let τ_{xy} be the unit vector from x to y. Then, for r small enough, the function $s[x, y] : t \mapsto x + \tau_{xy} u(\frac{t}{r})$ belongs to $\mathcal{P}_T^{m,\infty}$, with all its first $(m - 1)$ derivatives zero at its endpoints. The condition r small enough is for controlling the norm of the i-th derivatives for $i \leq m - 1$, which scale as r^{m-i}.

Now, let us split $\Omega = [0, 1]^d$ in N^d small cubes ω_i. We may order them such that each ω_i is adjacent to the next cube ω_{i+1}. We write x_i for the center of ω_i. We now build functions $s \in \mathcal{P}_T^{m,\infty}$ by concatenating paths from x_i to x_{i+1} and waiting times in x_i:

$$0 = t_1^1 \leq \cdots \leq t_{i-1}^2 \leq t_i^1 \leq t_i^2 \leq t_{i+1}^1 \leq \cdots \leq t_{N^d}^2 = T,$$

$$t_i^2 - t_i^1 = \left(\frac{1}{NC}\right)^{\frac{1}{m}},$$

$$s(t) = \begin{cases} x_i & \text{if} \quad t_i^1 \leq t \leq t_i^2, \\ s[x_i, x_{i+1}](t - t_i^2) & \text{if} \quad t_i^2 \leq t \leq t_{i+1}^1, \end{cases}$$

under the condition $T \geq T_N := (N^d - 1) \left(\frac{1}{NC}\right)^{\frac{1}{m}}$, that is to say that we have enough time to loop through all the cube centers.

Let now $\pi \in \mathcal{M}_\Delta$. We may choose $t_i^2 - t_i^1 \leq T\pi(\omega_i)$ for all i. Then, we may couple π and $s_* \gamma_T$ with $c(x_i, \omega_i) = \frac{t_i^2 - t_i^1}{T}$. Since the small cubes have radius $\sqrt{d}/N$ and the big one has radius $\sqrt{d}$, we obtain:

$$W_1(\pi, s_* \gamma_T) \leq \frac{\sqrt{d}}{2N} \sum_i \frac{t_i^2 - t_i^1}{T} + \sqrt{d} \sum_{i < N^d} \frac{t_{i+1}^1 - t_i^2}{T}$$

$$= \frac{\sqrt{d}}{2N} \frac{T - T_N}{T} + \sqrt{d} \frac{T_N}{T}.$$

In particular, taking $N = T^{\frac{m}{m(d+1)-1}}$, we find that $W_1\left(\mathcal{M}\left(\mathcal{P}_T^{m,\infty}\right), \pi\right) = O\left(T^{-\frac{m}{m(d+1)-1}}\right)$, hence $\bigcup_T \mathcal{M}\left(\mathcal{P}_T^{m,\infty}\right)$ is weakly dense in $\mathcal{M}_\Delta$.

$\square$

5.3 Numerical resolution

We now turn to the numerical resolution of problem (26). We first discretize the problem. We set $\Delta t := \frac{T}{N}$ and define discrete curves s as vectors of $\mathbb{R}^{N \cdot d}$. We let $s(i) \in \mathbb{R}^d$ denote the curve location at discrete time i, corresponding to the continuous time $i\Delta t$.

We define $D_1 : \mathbb{R}^{N \cdot d} \to \mathbb{R}^{N \cdot d}$, the discrete first order derivative operator, as follows:

$$(D_1 s)(i) = \frac{1}{\Delta t} \begin{cases} 0 & \text{if } i = 1, \\ s(i) - s(i - 1) & \text{if } i \in \{2, \ldots, N\}. \end{cases}$$

In what follows, D_i denotes a discretization of the derivative operator of order i. In the numerical experiments, we set $D_2 = -D_1^* D_1$.

We define $P_N^{m,q}$, a discretized version of $\mathcal{P}_T^{m,q}$, as follows:

$$P_N^{m,q} = \left\{ s \in \mathbb{R}^{N \cdot d}, \text{such that } \forall i \in \{1, \ldots N\}, \ s(i) \in \Omega, \right. \tag{27}$$

$$\left. \text{and } \forall j \in \{1, \ldots, m\}, \ \|D_j s\|_q \leqslant \alpha_j \right\}. \tag{28}$$

Here, $\|\cdot\|_q$ is defined by: $\|x\|_q = \left(\sum_{i=1}^{N \cdot d} \|x_i\|_2^q \right)^{\frac{1}{q}}$ for $q \in [1, +\infty)$ and $\|x\|_\infty = \max_{1 \leqslant i \leqslant N \cdot d} \|x_i\|_2$.

The measures set $\mathcal{M}(\mathcal{P}_T^{m,q})$ can be approximated by the set of N-point measures $\mathcal{M}(P_N^{m,q})$. From Corollary 3, it suffices to control the Hausdorff distance $\mathcal{H}_{W_1}(\mathcal{M}(\mathcal{P}_T^{m,q}), \mathcal{M}(P_N^{m,q}))$, to ensure that the solution of the discrete problem (6) with $\mathcal{M}_N = \mathcal{M}(P_N^{m,q})$ is a good approximation of problem (26). Unfortunately, the control of this distance is rather technical and falls beyond the scope of this paper for general m and q. In the following proposition, we therefore limit ourselves to the case $m = 1, q = \infty$.

Proposition 10. $\mathcal{H}_{W_1}(\mathcal{M}(\mathcal{P}_T^{1,\infty}), \mathcal{M}(P_N^{1,\infty})) \leqslant \alpha_1 \frac{T}{N}$.

Proof. 1. Let us show that $\sup_{\mu \in \mathcal{M}(\mathcal{P}_T^{1,\infty})} \inf_{\tilde{\mu} \in \mathcal{M}(P_N^{1,\infty})} W_1(\mu, \tilde{\mu}) \leqslant \frac{\alpha_1 T}{N}$.

Let $\mu \in \mathcal{M}(\mathcal{P}_T^{1,\infty})$ and denote by $p \in \mathcal{P}_T^{1,\infty}$ a parameterization such that $\mu = p_* \gamma$. Define $\tilde{\mu} = \frac{1}{N} \sum_{i=0}^{N-1} \delta_{p\left(\frac{iT}{N}\right)}$. Then a parameterization of $\tilde{\mu}$ is defined by $s(i) = p\left(\frac{iT}{N}\right)$. Moreover, for $i \in \{2, \ldots N\}$, $|(D_1 s)(i)| = \frac{1}{\Delta t}\left| p\left(\frac{iT}{N}\right) - p\left(\frac{(i-1)T}{N}\right) \right| = \frac{1}{\Delta t}\left| \int_{\frac{(i-1)T}{N}}^{\frac{iT}{N}} \dot{p}(t)\, dt \right| \leqslant \frac{1}{\Delta t} \int_{\frac{(i-1)T}{N}}^{\frac{iT}{N}} |\dot{p}(t)|\, dt \leqslant \alpha_1$. Therefore $s \in P_N^{1,\infty}$.

Let us consider the transportation map coupling the curve arcs between times $(i-1)\frac{T}{N}$ and $i\frac{T}{N}$ and the Diracs at $p\left(i\frac{T}{N}\right)$. Then $W_1(p_* \gamma, s_* \gamma) \leqslant$

$$\sum_{i=1}^{N} \frac{1}{N} \sup_{(i-1)\frac{T}{N} \leqslant t \leqslant i\frac{T}{N}} \left\| s(t) - s\left((i-1)\frac{T}{N}\right) \right\| \leqslant \alpha_1 \frac{T}{N}.$$

2. Let us fix $\mu \in \mathcal{M}\left(P_N^{1,\infty}\right)$ and let $s \in P_N^{1,\infty}$ such that $s_* \gamma = \mu$. We set $p(0) = s(1)$, and:

$$p(t) = \begin{cases} s(1) & \text{for } t \in \left]0, \frac{T}{N}\right], \\ s(i) + \left(\frac{t}{\Delta t} - i\right)(s(i+1) - s(i)) & \text{for } t \in \left]\frac{iT}{N}, \frac{(i+1)T}{N}\right] \\ & , i \in \{1, \ldots N-1\}. \end{cases}$$

Since $s \in \Omega^N$ and Ω is convex, $p([0,T]) \subset \Omega$. Moreover, p is continuous and piecewise differentiable. Finally, for $i \in \{1, \ldots, N-1\}$ and $t \in \left]\frac{iT}{N}, \frac{(i+1)T}{N}\right]$, $\dot{p}(t) = \frac{1}{\Delta t}(s(i+1) - s(i)) = D_1(s)(i)$. Therefore, $\|\dot{p}\|_\infty \leqslant$

α_1, ensuring that $p \in \mathcal{P}_T^{1,\infty}$. With the same coupling as above, we have $W_1(p_*\gamma, s_*\gamma) \leqslant \alpha_1 \frac{T}{N}$, which ends the proof.

$\square$

To end up, let us describe precisely a solver for the following variational problem:

$$\inf_{\mu \in \mathcal{M}\left(\mathcal{P}_T^{1,\infty}\right)} \mathcal{N}_h(\mu - \pi). \tag{29}$$

We let $\mathcal{M}^*$ denote the set of minimizers and $\mathcal{P}^*$ denote the associated set of parameterizations.

Algorithm 1 A projection algorithm on $\mathcal{M}\left(\mathcal{P}_T^{1,\infty}\right)$.

Input:

- π: target measure.

- N: a number of discretization points.

- $s^{(0)} \in P_N^{1,\infty}$: initial parameterized curve.

- H: a semi-algebraic function with Lipschitz continuous gradient.

- nit: number of iterations.

Output:

- $s^{(nit)}$: an approximation of a curve in $\mathcal{P}^*$.

- $\mu^{(nit)} = (s^{(nit)})_* \gamma_T$: an approximation of an element of $\mathcal{M}^*$.

for $0 \leq k \leq nit$ **do**

- Compute $\eta^{(k)} = \nabla J(s^{(k)})$
- Set $s^{(k+1)} = P_{P_N^{1,\infty}}\left(s^{(k)} - \tau\eta^{(k)}\right)$

Remark 5. *The implementation of Algorithm 1 requires computing the gradients (20) and (21) and computing a projection on $P_N^{1,\infty}$. Both problems are actually non trivial.*

The naive approach to compute the gradient of F consists of using the explicit formula (20). This approach is feasible only for a small amount of points N (less than 1000) since its complexity is $\mathcal{O}\left(N^2\right)$. In our numerical experiments, we therefore resort to fast summation algorithms [25, 17] commonly used in particles simulation. This part of the numerical analysis is described in [28] and we do not discuss it in this paper.

The set $P_N^{1,\infty}$ and more generally the sets $P_N^{m,q}$ are convex for $q \in [1,\infty]$. Projections can be computed using first-order iterative algorithms for convex

functions. In our numerical experiments, we use accelerated proximal gradient descents on the dual problem [3, 23, 31]. A precise description is given in [7].

6   Results

To illustrate the results, we focus on the continuous line drawing problem discussed throughout the paper. It is performed using Algorithm 1. In the following experiments, we set H as a smoothed L^1-norm. This is similar to what was proposed in the original halftoning papers in [26, 28].

We first concentrate on the projection onto $P_N^{1,\infty}$. In Figure 3, we show the evolution of the curve $s^{(k)}$ across iterations, for different choices of $s^{(0)}$. After $30,000$ iterations, the evolution seems to be stabilized. The cost function during the 400 first iterations is depicted in Figure 4 for the three different initializations. As can be seen, the curve evolves toward a satisfactoryt representation of the lion, whatever the initialization. This is a very nice feature that is somehow surprising since our algorithm simply consists of minimizing a highly nonconvex function with a first order method.

In Figure 5, we show the projection onto $P_N^{1,\infty}$ of the famous Meisje met de Parel painting (Girl with a Pearl Earring), after $10,000$ iterations. To really see the precision of the algorithm, we advise the reader to blink the eyes or to take a printed version of the paper away. From a close distance, the curves or points are visible. From a long distance, only the painting appears.

To finish, we consider projections onto more general measure spaces, such as $\mathcal{M}(\mathcal{P}_T^{m,q})$. In Fig. 6, we show different behaviours for different $m \in \{1, 2\}$ and $q \in \{1, 2, \infty\}$. We also show a large scale example with a picture of Marylin Monroe in Figure 7.

7   Conclusion

We analyzed the basic properties of a variational problem to project a target Radon measure π on arbitrary measures sets $\mathcal{M}_N$. We then proposed a numerical algorithm to find approximate solutions of this problem and gave several guarantees. An important application covered by this algorithm is the projection on the set of N-point measures, which is often called quantization and appears in many different areas such as finance, imaging, biology,... To the best of our knowledge, the extension to *arbitrary* measures set is new, and opens many interesting application perspectives. As examples in imaging, let us mention open topics such as the detection of singularities [2] (e.g. curves in 3D images) and sparse spike deconvolution in dimension d [9].

To finish, let us mention an important open question. We provided necessary and sufficient conditions on the sequence $(\mathcal{M}_N)_{N \in \mathbb{N}}$ for the sequence of *global* minimizers $(\mu_N^*)_{N \in \mathbb{N}}$ to weakly converge to π. In practice, finding the global minimizer is impossible and we can only expect finding critical points. One may therefore wonder whether all sequences of critical points weakly converge to π. An interesting perspective to answer this question is the use of mean-field limits [10].
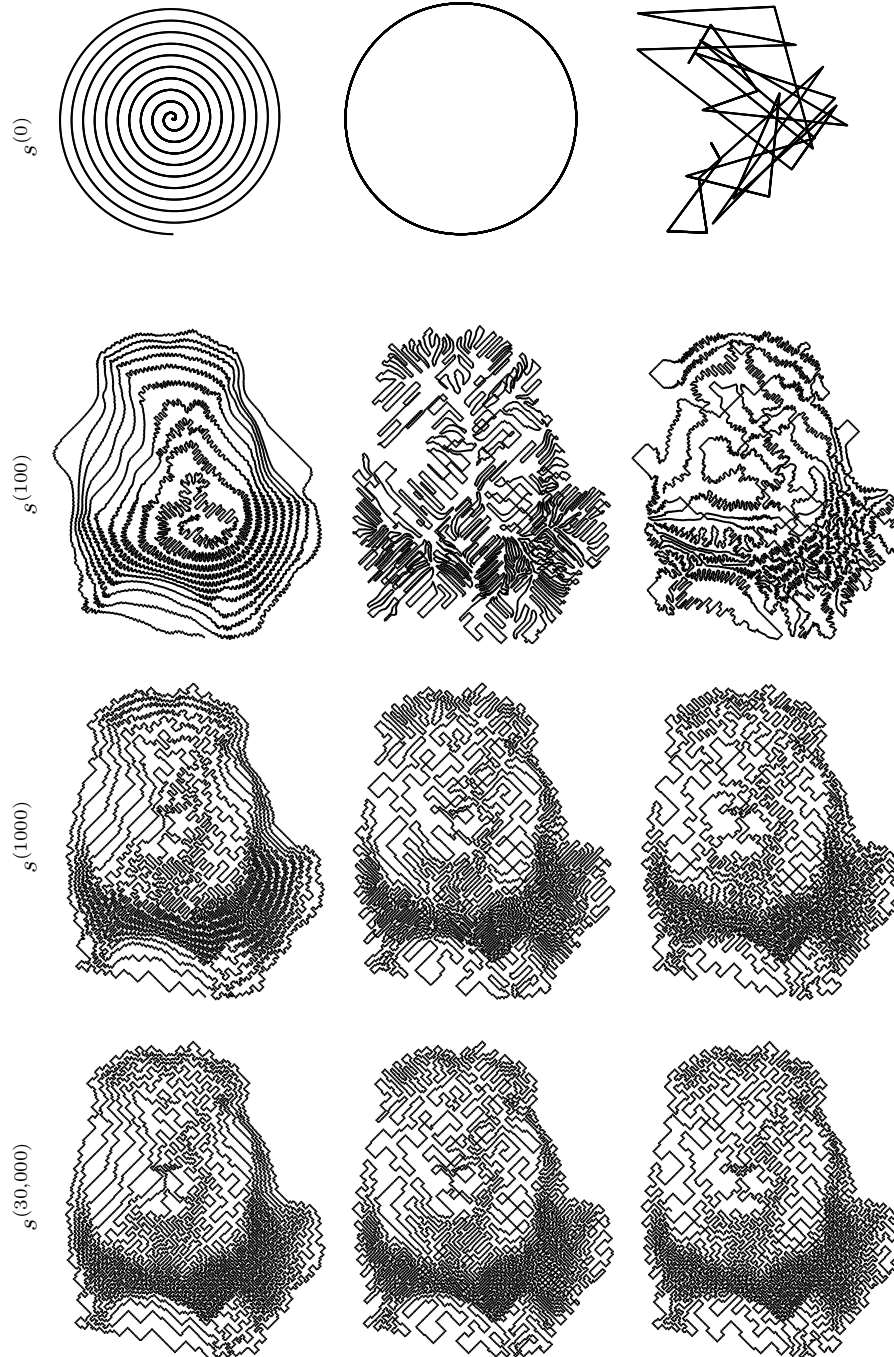
Figure 3: Projection of the lion image onto $P_N^{1,\infty}$ with $N = 8,000$. The figure depicts $s^{(k)}$ with several values of the iterate k in Algorithm 1.
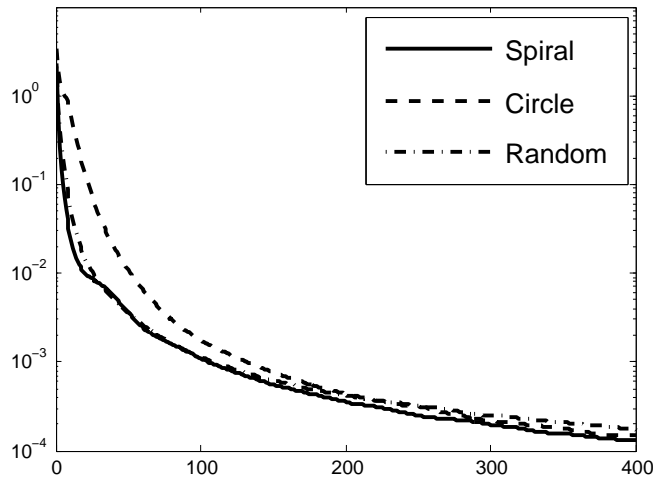
Figure 4: Decay of the cost function J for the three experiments depicted in Fig. 3. We represent $\log_{10}(J(k) - m)$ for $k \leq 400$ where m is the mimimal value of J during the first $30,000$ iterations.

Acknowledgements

References

[1] Hedy Attouch, Jérôme Bolte, and Benar Fux Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized gauss–seidel methods. *Mathematical Programming*, 137(1-2):91–129, 2013.

[2] Gilles Aubert, Jean-François Aujol, and Laure Blanc-Féraud. Detecting codimension-two objects in an image with Ginzburg-Landau models. *International Journal of Computer Vision*, 65(1-2):29–42, 2005.

[3] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[4] Vladimir Igorevich Bogachev and Maria Aparecida Soares Ruas. *Measure theory I*, volume 1. Springer, 2007.

[5] Robert Bosch and Adrianne Herman. Continuous line drawings via the traveling salesman problem. *Operations Research Letters*, 32(4):302–303, 2004.

[6] Claire Boyer, Nicolas Chauffert, Philippe Ciuciu, Jonas Kahn, and Pierre Weiss. On the generation of sampling schemes for magnetic resonance imaging. *arXiv preprint*, 2016.

[7] Nicolas Chauffert, Pierre Weiss, Jonas Kahn, and Philippe Ciuciu. Gradient waveform design for variable density sampling in Magnetic Resonance Imaging. *arXiv preprint arXiv:1412.4621*, 2014.

[8] John G Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision research*, 20(10):847–856, 1980.

[9] Vincent Duval and Gabriel Peyré. Exact support recovery for sparse spikes deconvolution. *arXiv preprint arXiv:1306.6909*, 2013.

[10] Massimo Fornasier, Jan Haškovec, and Gabriele Steidl. Consistency of variational continuous-domain quantization via kinetic theory. *Applicable Analysis*, 92(6):1283–1298, 2013.

[11] Massimo Fornasier and Jan-Christian Hütter. Consistency of probability measure quantization by means of power repulsion-attraction potentials. *arXiv preprint arXiv:1310.1120*, 2013.

[12] Manuel Gräf, Daniel Potts, and Gabriele Steidl. Quadrature errors, discrepancies, and their relations to halftoning on the torus and the sphere. *SIAM Journal on Scientific Computing*, 34(5):A2760–A2791, 2012.

[13] Peter M Gruber. Optimum quantization and its applications. *Advances in Mathematics*, 186(2):456–497, 2004.

[14] Pascal Gwosdek, Christian Schmaltz, Joachim Weickert, and Tanja Teuber. Fast electrostatic halftoning. *Journal of real-time image processing*, 9(2):379–392, 2014.

[15] Craig S Kaplan, Robert Bosch, et al. Tsp art. In *Renaissance Banff: Mathematics, Music, Art, Culture*, pages 301–308. Canadian Mathematical Society, 2005.

[16] Yitzhak Katznelson. An introduction to harmonic analysis. *New York*, 1968.

[17] Jens Keiner, Stefan Kunis, and Daniel Potts. Using nfft 3—a software library for various nonequispaced fast fourier transforms. *ACM Transactions on Mathematical Software (TOMS)*, 36(4):19, 2009.

[18] Benoit Kloeckner. Approximation by finitely supported measures. *ESAIM: Control, Optimisation and Calculus of Variations*, 18(02):343–359, 2012.

[19] Krzysztof Kurdyka. On gradients of functions definable in o-minimal structures. In *Annales de l'institut Fourier*, volume 48, pages 769–783. Institut Fourier, 1998.

[20] Hua Li and David Mould. Continuous line drawings and designs. *International Journal of Creative Interfaces and Computer Graphics*, 2014.

[21] RG Marteniuk, CL MacKenzie, M Jeannerod, S Athenes, and C Dugas. Constraints on human arm movement trajectories. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 41(3):365, 1987.

[22] Boris S Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory*, volume 330. Springer, 2006.

[23] Yu Nesterov. Gradient methods for minimizing composite functions. *Mathematical Programming*, 140(1):125–161, 2013.

[24] Thrasyvoulos N Pappas and David L Neuhoff. Least-squares model-based halftoning. *Image Processing, IEEE Transactions on*, 8(8):1102–1116, 1999.

[25] Daniel Potts and Gabriele Steidl. Fast summation at nonequispaced knots by NFFT. *SIAM Journal on Scientific Computing*, 24(6):2013–2037, 2003.

[26] Christian Schmaltz, Pascal Gwosdek, Andrés Bruhn, and Joachim Weickert. Electrostatic halftoning. In *Computer Graphics Forum*, volume 29, pages 2313–2327. Wiley Online Library, 2010.

[27] Steve Smale. Mathematical problems for the next century. *The Mathematical Intelligencer*, 20(2):7–15, 1998.

[28] Tanja Teuber, Gabriele Steidl, Pascal Gwosdek, Christian Schmaltz, and Joachim Weickert. Dithering by differences of convex functions. *SIAM Journal on Imaging Sciences*, 4(1):79–108, 2011.

[29] Joseph John Thomson. On the structure of the atom. *Philos. Mag., Ser. 6*, 7:237–265, 1904.

[30] Robert Ulichney. *Digital halftoning*. MIT press, 1987.

[31] Pierre Weiss, Laure Blanc-Féraud, and Gilles Aubert. Efficient schemes for total variation minimization under constraints in image processing. *SIAM journal on Scientific Computing*, 31(3):2047–2080, 2009.

[32] Fernando J Wong and Shigeo Takahashi. A graph-based approach to continuous line illustrations with variable levels of detail. In *Computer Graphics Forum*, volume 30, pages 1931–1939. Wiley Online Library, 2011.

[33] Jie Xu and Craig S Kaplan. Image-guided maze construction. In *ACM Transactions on Graphics (TOG)*, volume 26, page 29. ACM, 2007.

Figure 5: Projection of Meisje met de Parel, Vermeer 1665, onto $P_N^{1,\infty}$ with $N = 150,000$. The figure depicts $s^{(10,000)}$ obtained with Algorithm 1.
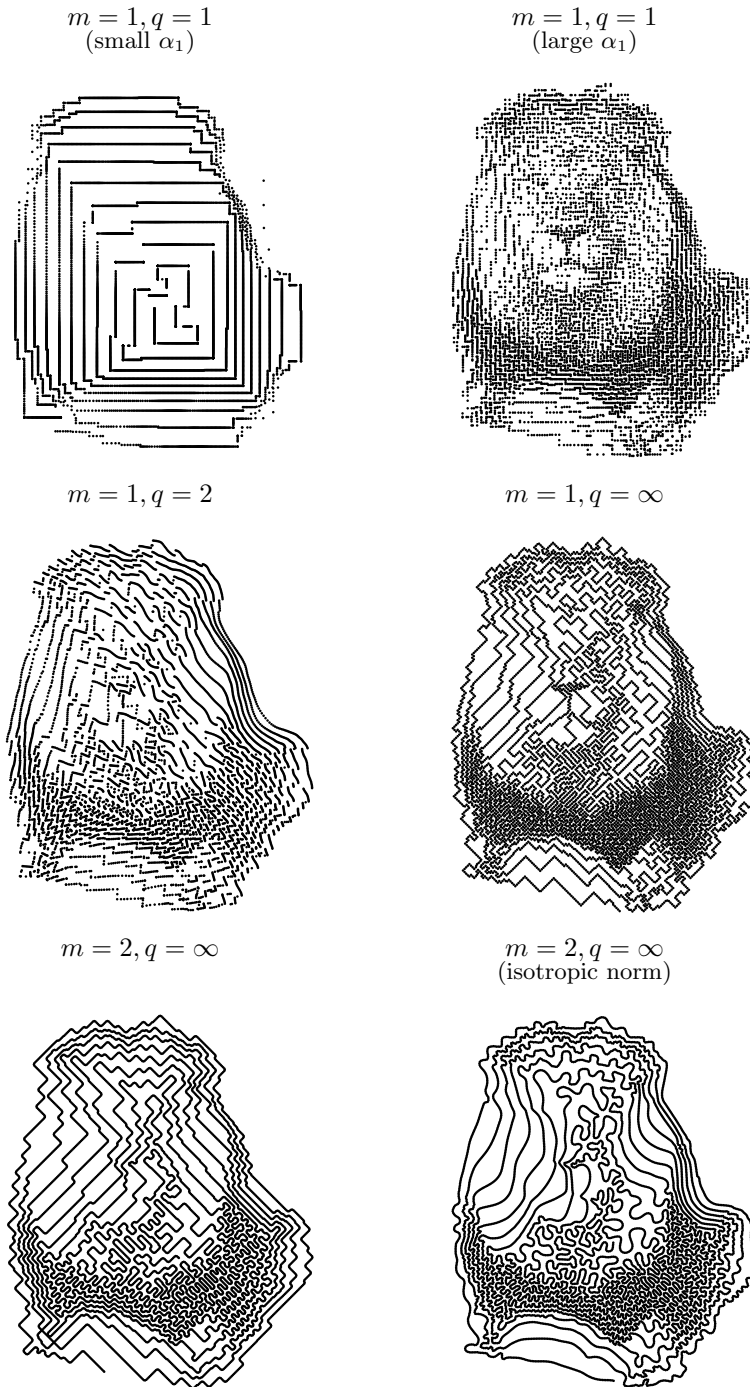
$m = 1, q = 1$
(small α_1)

$m = 1, q = 1$
(large α_1)

$m = 1, q = 2$

$m = 1, q = \infty$

$m = 2, q = \infty$

$m = 2, q = \infty$
(isotropic norm)

Figure 6: Projection of the lion image onto $P_N^{m,q}$ with $N = 8,000$, and $m \in \{1, 2\}$ and $q \in \{1, 2, \infty\}$.

Figure 7: Projection of Marylin image, onto the set:
$\mathcal{C} = \{p \in (W^{2,\infty}([0,T]))^2, \sup_{i\in[1,N]} (\|D_1 p(i)\|_2) \leq \alpha_1, \sup_{i\in[1,N]} (\|D_2 p(i)\|_2) \leq \alpha_2\}$,
with $N = 100,000$. The figure depicts $s^{(10,000)}$ obtained with Algorithm 1.