

# Consistency of Variational Continuous-Domain Quantization via Kinetic Theory

Massimo Fornasier\*, Jan Haškovec† and Gabriele Steidl‡

January 18, 2012

## Abstract

We study the kinetic mean-field limits of the discrete systems of interacting particles used for halftoning of images in the sense of continuous-domain quantization. Under mild assumptions on the regularity of the interacting kernels we provide a rigorous derivation of the mean-field kinetic equation. Moreover, we study the energy of the system, show that it is a Lyapunov functional and prove that in the long time limit the solution tends to an equilibrium given by a local minimum of the energy. In a special case we prove that the equilibrium is unique and is identical to the prescribed image profile. This proves the consistency of the particle halftoning method when the number of particles tends to infinity.

**AMS subject classification (MSC 2010):** 82C10, 82C22, 68U10

**Key Words:** Image processing, dithering, halftoning, mean-field limit, interacting particles.

## 1 Introduction

A halftoning method places black dots in an image in such a way that their density gives the impression of tone. For an illustration see Fig. 1. Due to its various applications, halftoning is an active field of research and we refer to the recent papers [12, 4] for deterministic and, resp., stochastic point distributions.

In this paper, we consider a continuous-domain quantization method based on electrostatic-like principles studied in [14], where the basic idea goes back to [12]. The method consists in considering a system of  $N$  particles with electrostatic-like interaction (repulsion) and exposed to an attractive external potential  $w$  with compact support in  $\mathbb{R}^d$  which represents the image to be approximated by points. In [14] the authors considered the discrete energy functional

$$E(p) := \sum_{i=1}^N \int_{\mathbb{R}^d} w(x) |p_i - x| dx - \frac{\lambda}{2} \sum_{i,j=1}^N |p_i - p_j|, \quad (1)$$

---

\*Faculty of Mathematics, Technical University of Munich, Boltzmannstrasse 3, D-85748 Garching, Germany, email: [massimo.fornasier@ma.tum.de](mailto:massimo.fornasier@ma.tum.de).

†Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria, email: [jan.haskovec@oeaw.ac.at](mailto:jan.haskovec@oeaw.ac.at).

‡Department of Mathematics, Technical University of Kaiserslautern, Germany email: [steidl@mathematik.uni-kl.de](mailto:steidl@mathematik.uni-kl.de).



Figure 1: Left: Original  $256 \times 256$  image. Right: Halftoning result with  $m = 30150$  points using the technique from [14].

where  $p := (p_1, \dots, p_N)^T \in \mathbb{R}^{N \times d}$ , the Euclidean distance is denoted by  $|\cdot|$ ,

$$\lambda := \frac{1}{N} \int_{\mathbb{R}^d} w(x) dx, \quad (2)$$

and defined the evolution of the particle system as the corresponding gradient flow  $\partial_t p \in -\partial_p E$ . They showed that for  $d = 1$  the energy functional is continuous and coercive, and calculated explicitly its minimizers. In the two-dimensional setting it is not possible to obtain explicit expressions for the minimizers anymore. Instead, the authors employed a difference of convex functions (DC) algorithm together with fast summation methods for non-equispaced knots to obtain a local minimum of the variational problem numerically. In [8] the function  $w$  was also considered on other sets such as  $\mathbb{T}^d$  or  $\mathbb{S}^2$  and kernels other than the Euclidean distance were used. This generalized approach stems from the study of optimal quadrature error functionals on reproducing kernel Hilbert spaces with respect to the quadrature knots and is closely related to so-called *discrepancy functionals* [7].

In this paper we study the kinetic mean-field limits of discrete systems like (1), that are obtained as the number of particles tends to infinity. First, in Section 2, we specify the interaction kernels which we will consider and provide some particular examples. Moreover, we introduce a generalized setting with different kernels for the attractive and repulsive interactions, and show that this new setting can be reduced to the previous one with appropriately modified data. In Section 3 we provide, under mild regularity conditions on the interaction kernels, a rigorous derivation of the mean field kinetic equation obtained in the limit as the number of particles tends to infinity. We show that the corresponding energy functional, obtained as a formal limit of the discrete energy, is a Lyapunov functional for the kinetic equation and that in the long-time limit the solution tends to an equilibrium, which is a local minimum of the energy. For a special choice of the interaction kernel, we are able to show that the equilibrium coincides with the prescribed image profile. This proves the consistency of the discrete halftoning method when the number of particles tends to infinity. Finally, in Section 4, we provide numerical examples confirming our theoretical results, and showing the behavior of the model

subject to different competitive attraction and repulsion terms as well as the consistency of the particle system with respect to its continuous kinetic limit.

## 2 Energy Functionals with Kernels

In the following, let  $\Omega$  be a domain, which we will be either the Euclidean space  $\Omega = \mathbb{R}^d$  or the torus  $\Omega = \mathbb{T}^d$ ,  $d \geq 1$ . A symmetric function  $K : \Omega \times \Omega \rightarrow \mathbb{R}$  is said to be *positive semi-definite* if for any  $N \in \mathbb{N}$  points  $x_1, \dots, x_N \in \Omega$  and any  $a \in \mathbb{R}^N \setminus \{0\}$  the relationship  $a^T (K(x_i, x_j))_{i,j=1}^N a \geq 0$  holds true, and *positive definite* if we have strict inequality. Let  $K : \Omega \times \Omega \rightarrow \mathbb{R}$  be a symmetric, positive semi-definite function, and  $H_K$  be the reproducing kernel Hilbert space (RKHS) associated with  $K$ , see [1]. Then we are interested in the functional

$$E_K(p) := \frac{\lambda}{2} \sum_{i,j=1}^N K(p_i, p_j) - \sum_{i=1}^N \int_{\Omega} w(x) K(p_i, x) dx. \quad (3)$$

In the case  $\Omega = \mathbb{R}^d$  we suppose that the function  $w : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is compactly supported. In [8] it was shown that this functional is related to the optimality of a certain quadrature rule for functions in  $H_K$  depending on the knots  $p_i$ ,  $i = 1, \dots, N$ . By the following remark, which was proved in [8], slight modifications of the kernel do not change the minimizers of the functional  $E_K$ .

**Remark 1** *Let  $K : \Omega \times \Omega \rightarrow \mathbb{R}$  be a symmetric function and  $\tilde{K}(x, y) := aK(x, y) + b(K(x, 0) + K(0, y)) + c$  with  $a > 0$  and  $b, c \in \mathbb{R}$ . Then the minimizers of  $E_K$  and  $E_{\tilde{K}}$  coincide.*

In this paper, we consider *radial kernels* on  $\Omega = \mathbb{R}^d$ , i.e.,

$$K(x, y) = \varphi(x - y) = \Phi(|x - y|) \quad (4)$$

with  $\Phi : [0, \infty) \rightarrow \mathbb{R}$ . For the 1-periodic setting  $\Omega = \mathbb{T}^1$  we use the same notation, where  $|x - y|$  has to be replaced by the ‘‘periodic’’ distance  $\min\{|x - y|, 1 - |x - y|\}$ . In the case  $\Omega = \mathbb{T}^d$  we consider tensor products of radial kernels. We call  $\varphi$  positive semi-definite (resp. positive definite) if the corresponding kernel is positive semi-definite (resp. positive definite). The functional of interest becomes then

$$E_{\varphi}(p) := \frac{\lambda}{2} \sum_{i,j=1}^N \varphi(p_i - p_j) - \sum_{i=1}^N \int_{\Omega} w(x) \varphi(p_i - x) dx. \quad (5)$$

**Example.** We give some interesting examples of positive semi-definite kernels, see [16] and [15].

1. Let  $\Omega := \mathbb{R}^d$ . Then the functions

$$\begin{aligned} \varphi(x) &= (1 - |x|)_{+}^{\tau}, \quad \tau \geq \left\lfloor \frac{d}{2} \right\rfloor + 1, \\ \varphi(x) &= (\varepsilon^2 + |x|^2)^{-\beta}, \quad \beta > \frac{d}{2} \quad (\text{inverse multiquadrics}) \end{aligned}$$

are positive definite. Next, consider the *conditionally positive definite radial kernels of order 1* defined by

$$\begin{aligned}\varphi(x) &:= -|x|^\tau, \quad 0 < \tau < 2, \\ \varphi(x) &:= -(\varepsilon^2 + |x|^2)^\tau, \quad 0 < \tau < 1, \quad (\text{multiquadrics}).\end{aligned}$$

The kernels  $K(x, y) = \varphi(x - y)$  are not positive semi-definite. However, their slight modifications given by

$$\tilde{K}(x, y) := \varphi(x - y) - \varphi(y) - \varphi(x) + \varphi(0)$$

define positive semi-definite kernels, and the corresponding RKHSs were characterized in [16, Theorem 10.18]. By Remark 1,  $E_K$  and  $E_{\tilde{K}}$  have the same minimizers, so that we can work with the original kernel  $K$  in the energy functional.

2. Let  $\Omega := \mathbb{T}^1$ . Up to an additive constant, Wahba's spline kernels are given by

$$K(x, y) := \frac{(-1)^{m-1}}{(2m)!} B_{2m}(|x - y|) = 2 \sum_{k=1}^{\infty} \frac{1}{(2\pi k)^{2m}} \cos(2\pi k(x - y))$$

where  $B_{2m}$  denotes the *Bernoulli polynomial* of degree  $2m$ . Note that  $B_{2m}(1 - t) = B_{2m}(t)$ . For example, we have

$$B_2(t) = t^2 - t + \frac{1}{6}, \quad B_4(t) = t^4 - 2t^3 + t^2 - \frac{1}{30}.$$

Let us now introduce a slight generalization of (5), where we consider different kernels for the attractive interaction (attraction of the particles by the image profile  $w$ ) and the repulsive interaction (particle-particle repulsion). In particular, we introduce the functions  $\varphi$  and  $\psi$  related to different radial kernels (4), and the generalized energy functional

$$E_{\psi, \varphi}(p) := \frac{\lambda}{2} \sum_{i, j=1}^N \psi(p_i - p_j) - \sum_{i=1}^N \int_{\Omega} w(x) \varphi(p_i - x) \, dx. \quad (6)$$

The following remark gives an intuition on the behavior of the corresponding quantization process.

**Remark 2** *Assume that the kernel  $K$  in (3) is in addition continuous and an element of  $L^2(\Omega \times \Omega)$  (Mercer kernel). Then it can be expanded into an absolutely and uniformly convergent series,*

$$K(x, y) = K_\lambda(x, y) := \sum_{\ell=1}^{\infty} \lambda_\ell \eta_\ell(x) \eta_\ell(y)$$

*of orthonormal eigenfunctions  $\eta_\ell \in L^2(\Omega)$  and associated eigenvalues  $\lambda_\ell > 0$  of the compact, self-adjoint integral operator  $T_K$  on  $L^2(\Omega)$  given by*

$$T_K f(x) := \int_{\Omega} K(x, y) f(y) \, dy.$$

Assume further that  $w$  can be also expanded into an absolutely convergent series  $w(x) := \sum_{k=1}^{\infty} w_k \eta_k(x)$ . Then the functional (3) becomes

$$\begin{aligned}
E_K(p) &= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \int_{\Omega} w(x) K_\lambda(p_i, x) \, dx \\
&= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \int_{\Omega} \sum_{k=1}^{\infty} w_k \eta_k(x) \sum_{\ell=1}^{\infty} \lambda_\ell \eta_\ell(p_i) \eta_\ell(x) \, dx \\
&= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \sum_{k,\ell=1}^{\infty} w_k \lambda_\ell \eta_\ell(p_i) \int_{\Omega} \eta_k(x) \eta_\ell(x) \, dx \\
&= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \sum_{k=1}^{\infty} w_k \lambda_k \eta_k(p_i).
\end{aligned}$$

On the other hand, if we consider  $E_K$  for another function  $\tilde{w}(x) := \sum_{k=1}^{\infty} w_k v_k \eta_k(x)$ , we have

$$\begin{aligned}
E_K(p) &= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \int_{\Omega} \tilde{w}(x) K_\lambda(p_i, x) \, dx \\
&= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \sum_{k=1}^{\infty} w_k v_k \lambda_k \eta_k(p_i) \\
&= \frac{\lambda}{2} \sum_{i,j=1}^N K_\lambda(p_i, p_j) - \sum_{i=1}^N \int_{\Omega} w(x) K_\mu(p_i, x) \, dx,
\end{aligned}$$

with  $K_\mu(x, y) := \sum_{\ell=1}^{\infty} \mu_\ell \eta_\ell(x) \eta_\ell(y)$  and  $\mu_k := v_k \lambda_k$ , where we assume absolute convergence of the involved series. Hence, using a smoother kernel  $K_\mu$  for the interaction with the datum  $w$  than  $K_\lambda$  (i.e.,  $\mu_k$  decays faster than  $\lambda_k$ ) leads to the approximation of a smoother function  $\tilde{w}$  ( $w_k v_k$  decays faster than  $w_k$ ), and vice versa.

### 3 Mean-Field Limit

We are interested in the passage to the limit when the number of particles  $N$  tends to infinity. For simplicity, we restrict our attention to radial kernels and  $\Omega = \mathbb{R}^d$  although the analysis works as well for the periodic settings  $\Omega = \mathbb{T}$  and  $\Omega = \mathbb{T}^d$ ,  $d \geq 2$  with the tensor product of radial kernels. Moreover, without loss of generality, we prescribe the normalization

$$\int_{\Omega} w(x) \, dx = 1, \quad N\lambda = 1$$

and suppose that  $w \geq 0$  is compactly supported.

### 3.1 Passage to the Mean-Field Limit

The evolution of the  $N$ -particle system according to the gradient flow of the discrete generalized energy functional

$$E(p) = E_{\varphi,\psi}(p) := \frac{1}{2N} \sum_{k,\ell=1}^N \psi(p_k - p_\ell) - \sum_{k=1}^N \int_{\Omega} w(x) \varphi(p_k - x) dx \quad (7)$$

is given, under the assumption that  $\varphi, \psi \in C^1(\Omega)$ , by

$$\begin{aligned} \frac{d}{dt} p_i(t) &= -\nabla_{p_i} E(p(t)) \\ &= \int_{\Omega} w(x) \nabla \varphi(p_i(t) - x) dx - \frac{1}{N} \sum_{\ell=1}^N \nabla \psi(p_i(t) - p_\ell(t)), \quad i = 1, \dots, N, \end{aligned} \quad (8)$$

subject to the initial condition

$$p_i(0) = p_i^0, \quad i = 1, \dots, N. \quad (9)$$

The mean field limit is obtained as the number of particles  $N$  tends to infinity. Then, the vector of time-dependent particle positions  $p(t) \in \Omega^N$  is replaced by the time-dependent probability measure  $f(x, t)$ , where, roughly speaking,  $f(x, t) dx$  can be understood as the probability that a particle is located in the space element  $dx$  around the position  $x \in \Omega$  at time  $t \geq 0$ .

In the following, let  $\mathcal{M}(\Omega)$  denote the space of Radon measures on  $\Omega$  and  $C_c(\Omega)$  the space of continuous, compactly supported functions on  $\Omega$ . Further, let  $L^\infty(\mathbb{R}_+, \mathcal{M}(\Omega))$  denote the space of functions from  $\mathbb{R}_+$  to  $\mathcal{M}(\Omega)$  which are essentially bounded, i.e.,  $f : t \rightarrow f(\cdot, t) = f_t$  with  $\text{ess sup}_{t \in \mathbb{R}_+} \int_{\Omega} d|f(\cdot, t)| < \infty$ ,  $|f|$  denoting the total variation of the measure  $f$ . Note that  $L^\infty(\mathbb{R}_+, \mathcal{M}(\Omega))$  is the dual space of  $L^1(\mathbb{R}_+, C_c(\Omega))$ , the space of functions from  $\mathbb{R}_+$  to  $C_c(\Omega)$  such that  $\int_0^\infty \|g(\cdot, t)\|_\infty dt < \infty$ , see, e.g., [2]. Moreover, let us denote by  $\mathcal{M}^1(\Omega)$  the set of probability measures on  $\Omega$ , i.e.,  $f \in \mathcal{M}^1(\Omega)$  if and only if  $f$  is a nonnegative Radon measure such that  $\int_{\Omega} df = 1$ .

For any  $N \in \mathbb{N}$ , let us denote by  $f^N$  the empirical measures

$$f^N(\cdot, t) = \frac{1}{N} \sum_{i=1}^N \delta(\cdot - p_i(t)) \quad (10)$$

corresponding to the evolution of the  $N$ -particle system (8). Then, each  $f^N$  is a time-dependent probability measure, such that  $f^N \in L^\infty(\mathbb{R}_+, \mathcal{M}(\Omega))$ . In the following theorem we carry out the rigorous mean field limit passage  $N \rightarrow \infty$ .

**Theorem 1** *Let  $\varphi, \psi \in C^1(\Omega)$ , where  $\nabla \psi$  is in addition bounded. Let  $(f^N)_{N \in \mathbb{N}}$  be given by (10), corresponding to the system (8) with the initial datum (9). Moreover, assume that there exists a probability measure  $f_0 \in \mathcal{M}^1(\Omega)$  such that  $f^N(\cdot, 0) \rightarrow f_0(\cdot)$  weakly- $*$  in  $\mathcal{M}(\Omega)$  as  $N \rightarrow \infty$ .*

*Then there exists a subsequence  $(f^{N_k})_{k \in \mathbb{N}}$  which converges weakly- $*$  in  $L^\infty(\mathbb{R}_+, \mathcal{M}(\Omega))$  to a time-dependent probability measure  $f \in L^\infty(\mathbb{R}_+, \mathcal{M}^1(\Omega))$  which solves, in the sense of*

distributions, the mean-field equation

$$\begin{aligned}\partial_t f &= \nabla_y \cdot \left( \int_{\Omega} (f(x, t) \nabla \psi(y - x) - w(x) \nabla \varphi(y - x)) f(y, t) dx \right) \\ &= \nabla \cdot (\nabla \mathcal{K}[f]f),\end{aligned}\tag{11}$$

where

$$\mathcal{K}[f](y, t) := \int_{\Omega} (f(x, t) \psi(y - x) - w(x) \varphi(y - x)) dx,\tag{12}$$

subject to the initial condition

$$f(\cdot, 0) = f_0.\tag{13}$$

**Proof:** First we show that for all  $N \in \mathbb{N}$  the empirical measures (10) are distributional solutions of (11). Indeed, considering a smooth, compactly supported test function  $\xi \in C_c^\infty(\Omega \times [0, \infty))$ , we obtain

$$\begin{aligned}& \int_0^\infty \int_{\Omega} f^N(y, t) \partial_t \xi(y, t) dy dt + \int_{\Omega} f^N(y, 0) \xi(y, 0) dy \\ &= \frac{1}{N} \int_0^\infty \sum_{i=1}^N \partial_t \xi(p_i(t), t) dt + \frac{1}{N} \sum_{i=1}^N \xi(p_i(0), 0) \\ &= \frac{1}{N} \int_0^\infty \sum_{i=1}^N \left[ \frac{d}{dt} \xi(p_i(t), t) - \nabla_{p_i} \xi(p_i(t), t) \cdot \frac{d}{dt} p_i(t) \right] dt + \frac{1}{N} \sum_{i=1}^N \xi(p_i(0), 0) \\ &= \frac{1}{N} \int_0^\infty \sum_{i=1}^N \nabla_{p_i} \xi(p_i(t), t) \cdot \left[ \frac{1}{N} \sum_{\ell=1}^N \nabla \psi(p_i(t) - p_\ell(t)) - \int_{\Omega} w(x) \nabla \varphi(p_i(t) - x) dx \right] dt \\ &= \int_0^\infty \int_{\Omega} f^N(y, t) \nabla_y \xi(y, t) \cdot \left[ \int_{\Omega} f^N(x, t) \nabla \psi(y - x) dx - \int_{\Omega} w(x) \nabla \varphi(y - x) dx \right] dy dt,\end{aligned}$$

Therefore, we have the identity

$$\begin{aligned}& \int_0^\infty \int_{\Omega} f^N(y, t) \partial_t \xi(y, t) dy dt + \int_{\Omega} f^N(y, 0) \xi(y, 0) dy \\ &= \int_0^\infty \int_{\Omega} f^N(y, t) \nabla_y \xi(y, t) \cdot \left[ \int_{\Omega} f^N(x, t) \nabla \psi(y - x) dx - \int_{\Omega} w(x) \nabla \varphi(y - x) dx \right] dy dt,\end{aligned}\tag{14}$$

for all test functions  $\xi \in C_c^\infty(\Omega \times [0, \infty))$ , that is the distributional formulation of (11) with  $f^N$  in place of  $f$  and the initial condition  $f^N(\cdot, 0) = \frac{1}{N} \sum_{i=1}^N \delta(\cdot - p_i^0)$ .

Now, since  $(f^N)_{N \in \mathbb{N}}$  is a sequence of time-dependent probability measures, it is uniformly bounded in  $L^\infty(\mathbb{R}_+, \mathcal{M}(\Omega))$ , so that there exists a subsequence  $(f^{N_k})_{k \in \mathbb{N}}$  which converges weakly-\* to some  $f$  in  $L^\infty(\mathbb{R}_+, \mathcal{M}^1(\Omega))$ . We show that  $f$  is a distributional solution of (11). The limit passage in the linear terms of (14) follows immediately. Moreover, since  $\nabla \psi$  is assumed to be continuous and bounded, the sequence  $\int_{\Omega} f^N(x, t) \nabla \psi(y - x) dx$  is uniformly equicontinuous and uniformly bounded for  $y \in \text{supp } \xi(\cdot, t)$  and for almost every  $t \in \mathbb{R}_+$ . Therefore, due to the Arzelà-Ascoli theorem, a subsequence converges strongly in  $L^\infty(\text{supp } \xi(\cdot, t))$  for almost all  $t \in \mathbb{R}_+$ . Finally, the bounded convergence theorem ensures the strong convergence of the sequence in  $L^1(\mathbb{R}_+, L^\infty(\text{supp } \xi))$  and this justifies the limit passage in the nonlinear term.

Therefore, in the limit  $N \rightarrow \infty$ , we have obtained

$$\begin{aligned} & \int_0^\infty \int_\Omega f(y, t) \partial_t \xi(y, t) \, dy \, dt + \int_\Omega f_0(y) \xi(y, 0) \, dy \\ &= \int_0^\infty \int_\Omega f(y, t) \nabla_y \xi(y, t) \cdot \left[ \int_\Omega f(x, t) \nabla \psi(y - x) \, dx - \int_\Omega w(x) \nabla \varphi(y - x) \, dx \right] \, dy \, dt, \end{aligned} \quad (15)$$

which is the distributional formulation of (11) subject to the initial condition (13).  $\blacksquare$

Equation (11) describes the evolution of the time-dependent probability measure  $f(\cdot, t)$  due to the mutual repulsive interaction between the particles and the attractive interaction with the datum  $w$ . In the following lemma we show that, under mild regularity assumptions on  $\varphi$ ,  $\psi$  and  $f_0$ , the solution  $f$  of (11) is in fact classical.

**Lemma 1** *Let  $\nabla \varphi$  and  $\nabla \psi$  be globally Lipschitz continuous on  $\mathbb{R}^d$ , i.e., there exist constants  $L_1, L_2$  such that*

$$\begin{aligned} |\nabla \psi(x) - \nabla \psi(y)| &\leq L_1 |x - y|, \\ |\nabla \varphi(x) - \nabla \varphi(y)| &\leq L_2 |x - y|, \end{aligned}$$

for all  $x, y \in \Omega$ . Let  $f_0 \in C_c^1(\Omega)$  be nonnegative, compactly supported and fulfill  $\int_\Omega f_0(x) \, dx = 1$ . Then the corresponding distributional solution  $f$  of (11) subject to the initial condition (13) is in fact the unique classical solution with  $f \in C^1(\Omega \times \mathbb{R}_+)$  and  $f(\cdot, t) \geq 0$  for all  $t \geq 0$ . Moreover,  $f(\cdot, t)$  is compactly supported on  $\Omega$  for any  $t \in \mathbb{R}_+$ .

**Proof:** Since the distributional solution  $f \in L^\infty(\mathbb{R}_+, \mathcal{M}^1(\Omega))$  constructed in Theorem 1 is a time-dependent probability measure, we have  $f(\cdot, t) \geq 0$  and

$$\int_\Omega df(\cdot, t) \equiv 1 \quad \text{for all } t \geq 0.$$

The essential point is to observe that due to the assumptions on  $\varphi$  and  $\psi$ , the transport field  $\nabla \mathcal{K}[f]$  for (11) is Lipschitz continuous. Indeed, we have

$$\begin{aligned} |\nabla \mathcal{K}[f](p, t) - \nabla \mathcal{K}[f](q, t)| &\leq \int_\Omega f(x, t) |\nabla \psi(p - x) - \nabla \psi(q - x)| \, dx \\ &\quad + \int_\Omega w(x) |\nabla \varphi(p - x) - \nabla \varphi(q - x)| \, dx \\ &\leq L_1 \int_\Omega f(x, t) |p - q| \, dx + L_2 \int_\Omega w(x) |p - q| \, dx \\ &\leq (L_1 + L_2) |p - q|. \end{aligned}$$

Therefore,  $f$  is a solution of a hyperbolic transport equation with globally Lipschitz-continuous transport field  $\nabla \mathcal{K}[f]$ . As such, the values of the initial condition  $f_0$  propagate along the characteristics

$$\dot{x}(t) = \nabla \mathcal{K}[f](x(t), t) \quad (16)$$

with finite speeds. Due to the assumption  $f_0 \in C_c^1(\Omega)$ , from the standard theory of hyperbolic transport equations (method of characteristics and Cauchy-Lipschitz theorem, see [3]) it follows



that  $f \in C^1(\Omega \times \mathbb{R}_+)$  is the unique classical solution. The compactness of the support of  $f$  for all times follows from the assumed compactness of the support of  $f_0$  and the finite characteristic speeds (16). ■

**Remark 3** *Similarly as in Remark 2, we observe that the evolution induced by (11) with two different interaction kernels  $\varphi$  and  $\psi$  is in fact equivalent to an evolution produced by using same interaction kernels, but with a modified data  $w$ . Indeed, assuming that  $\varphi$  and  $\psi$  are continuously differentiable and that the Fourier transform of  $\psi$  is nonzero almost everywhere (as is the case for positive definite kernels), the Fourier transform of (11) with respect to  $y$  reads*

$$\begin{aligned} \partial_t \hat{f}(\xi, t) &= i\xi \cdot \left( (\widehat{f\nabla\psi} - \hat{w}\widehat{\nabla\varphi}) * \hat{f}(\xi, t) \right) \\ &= -|\xi|^2 \left( (\hat{f}\hat{\psi} - \hat{w}\hat{\varphi}) * \hat{f}(\xi, t) \right) \\ &= -|\xi|^2 \left( \hat{\psi} \left( \hat{f} - \hat{w} \frac{\hat{\varphi}}{\hat{\psi}} \right) * \hat{f}(\xi, t) \right). \end{aligned}$$

Applying the inverse Fourier transform, we get

$$\partial_t f = \nabla \cdot ((f - \tilde{w}) * \nabla\psi) f,$$

where  $\tilde{w}$  is the inverse Fourier transform of  $(\hat{w}\hat{\varphi}/\hat{\psi})$ . Therefore, taking  $\varphi$  smoother than  $\psi$  corresponds to a smoothing of  $w$ , while  $\psi$  smoother than  $\varphi$  corresponds to a “sharpening” (anti-smoothing) of  $w$ .

### 3.2 Energy dissipation, long time behavior and equilibria

Let us observe that the formal limit of the discrete energy (7) as  $N \rightarrow \infty$  is given by the continuous energy functional

$$\mathcal{E}[f] = \frac{1}{2} \int_{\Omega} \int_{\Omega} f(x)\psi(p-x)f(p) dp dx - \int_{\Omega} \int_{\Omega} w(x)\varphi(p-x)f(p) dx dp, \quad (17)$$

defined for all  $f \in \mathcal{M}^1(\Omega)$ . Let us mention that the corresponding formal gradient flow with respect to the topology of 2-Wasserstein distance on the space of probability measures (see [10] or [11] for details) is given by the hyperbolic transport equation (11). This suggests that the energy  $\mathcal{E}[f(x, \cdot)]$  actually is a Lyapunov functional, thus nonincreasing along the solutions of (11):

$$\frac{d}{dt} \mathcal{E}[f(x, t)] \leq 0. \quad (18)$$

Indeed, at least for classical solutions, this inequality can be proven rigorously:

**Lemma 2** *Let  $\varphi, \psi$  be of the form (4). Let  $f \in C^1(\Omega \times \mathbb{R}_+)$  be a classical solution of (11)–(13) in the sense of Lemma 1. Then (18) holds true.*

**Proof:** The proof follows from the direct calculation

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}[f(x, t)] &= -\int_{\Omega}\int_{\Omega}w(x)\varphi(p-x)\nabla_p\cdot(\nabla\mathcal{K}[f](p, t)f(p, t))\,dp\,dx \\
&\quad +\frac{1}{2}\int_{\Omega}\int_{\Omega}\psi(p-x)f(x, t)\nabla_p\cdot(\nabla\mathcal{K}[f](p, t)f(p, t))\,dp\,dx \\
&\quad +\frac{1}{2}\int_{\Omega}\int_{\Omega}\psi(p-x)f(p, t)\nabla_x\cdot(\nabla\mathcal{K}[f](x, t)f(x, t))\,dp\,dx \\
&= \int_{\Omega}\int_{\Omega}w(x)\nabla\varphi(p-x)\cdot(\nabla\mathcal{K}[f](p, t)f(p, t))\,dp\,dx \\
&\quad -\frac{1}{2}\int_{\Omega}\int_{\Omega}\nabla\psi(p-x)f(x, t)\cdot(\nabla\mathcal{K}[f](p, t)f(p, t))\,dp\,dx \\
&\quad +\frac{1}{2}\int_{\Omega}\int_{\Omega}\nabla\psi(p-x)f(p, t)\cdot(\nabla\mathcal{K}[f](x, t)f(x, t))\,dp\,dx \\
&= \int_{\Omega}\int_{\Omega}(w(x)\nabla\varphi(p-x)-\nabla\psi(p-x)f(x, t))\cdot(\nabla\mathcal{K}[f](p, t)f(p, t))\,dp\,dx \\
&= -\int_{\Omega}|\nabla\mathcal{K}[f](p, t)f(p, t)|^2f(p, t)\,dp\leq 0, \tag{19}
\end{aligned}$$

where we have used the symmetry of  $\psi$ , i.e.,  $\nabla\psi(p-x)=-\nabla\psi(x-p)$ , and integration by parts, where the boundary terms vanish due to the compact support of  $f$ . (In the case  $\Omega=\mathbb{T}^d$  due to the periodicity of the interaction kernels.) ■

In the rest of this section, we study the question whether the classical solution  $f$  of (11) tends for  $t\rightarrow\infty$  to an equilibrium  $f^*$ , characterized by the condition  $|\nabla\mathcal{K}[f^*](p)f^*(p)|^2f^*(p)=0$  a.e. on  $\Omega$ , which stems from setting the right-hand side of (19) equal to zero. Without loss of generality (see Remark 3), we restrict our attention to the case  $\varphi=\psi$ ; then we have

$$\nabla\mathcal{K}[f](p, t)=\int_{\Omega}(f(x, t)-w(x))\nabla\psi(p-x)\,dx=((f(\cdot, t)-w)*\nabla\psi)(p). \tag{20}$$

Moreover, we adopt the assumption that  $f$  is a classical solution of (11) in the sense of Lemma 1.

First we show that the energy functional (17) is bounded from below.

**Lemma 3** *Let  $\Phi:[0, \infty)\rightarrow\mathbb{R}$  be concave and monotone increasing and  $\Phi(0)$  be finite. Then  $\mathcal{E}[f]$  with  $\varphi=\psi=\Phi(|\cdot|)$  is bounded from below for all  $f\in\mathcal{M}^1(\Omega)$ .*

**Proof:** Without loss of generality, we can assume that  $\Phi(0)=0$ . Concavity and monotonicity of  $\Phi$  imply its subadditivity [14], which implies further

$$\Phi(|p-x|)\leq\Phi(|p|)+\Phi(|x|), \quad \Phi(|p-x|)\geq\Phi(|p|)-\Phi(|x|).$$

Therefore, since  $w, f\geq 0$  and  $\int_{\Omega}w(p)\,dp=\int_{\Omega}f(p)\,dp=1$ , we obtain

$$\begin{aligned}
\mathcal{E}[f] &\geq \int_{\Omega}\int_{\Omega}\left[w(x)(\Phi(|p|)-\Phi(|x|))-\frac{1}{2}f(x)(\Phi(|p|)+\Phi(|x|))\right]f(p)\,dp\,dx \\
&= \int_{\Omega}\Phi(|p|)f(p)\,dp-\int_{\Omega}w(x)\Phi(|x|)\,dx-\int_{\Omega}\Phi(|p|)f(p)\,dp \\
&= -\int_{\Omega}w(x)\Phi(|x|)\,dx,
\end{aligned}$$

which is the announced boundedness from below. ■

**Remark 4** Note that if  $\Phi(r)$  grows as  $r^\tau$  with  $\tau > 2$  as  $r \rightarrow \infty$  and  $w$  is compactly supported, then  $\mathcal{E}$  is not bounded from below. This can be seen as follows: Define  $f_q(x) := \frac{1}{2}(\delta(x - q) + \delta(x + q))$  for  $q \in \mathbb{R}^d$ . Then, due to the compact support of  $w$ , we get

$$\begin{aligned} \mathcal{E}[f_q] &= \frac{1}{2} \int w(x) (\Phi(|q - x|) + \Phi(|q + x|_2)) \, dx - \frac{1}{4} \Phi(2|q|) \\ &\sim |q|^\tau - \frac{1}{4} (2|q|)^\tau = (1 - \frac{1}{4} 2^\tau) |q|^\tau \rightarrow -\infty \text{ as } |q| \rightarrow \infty. \end{aligned}$$

**Lemma 4** Let  $\varphi, \psi$  be of the form (4) and let  $f \in C^1(\Omega \times \mathbb{R}_+)$  be a classical solution of (11)–(13) in the sense of Lemma 1. Moreover, either let  $\Omega$  be bounded (for instance,  $\Omega = \mathbb{T}^d$ ), or let  $\psi$  be a uniformly bounded function. Then  $f(\cdot, t) \rightarrow f^*$  weakly-\* in  $\mathcal{M}^1(\Omega)$  as  $t \rightarrow \infty$ , where  $f^*$  satisfies

$$|\nabla \mathcal{K}[f^*](p) f^*(p)|^2 f^*(p) = 0 \quad \text{a.e. on } \Omega. \quad (21)$$

**Proof:** Since by Lemma 2 the energy  $\mathcal{E}[f(x, t)]$  is nonincreasing as a function of time and by Lemma 3 bounded below, the limit of  $\mathcal{E}[f(x, t)]$  as  $t \rightarrow \infty$  exists and is finite.

Due to the boundedness of  $f(\cdot, t)$  in the space of Radon measures, there exists a sequence  $t_j \rightarrow \infty$  and a Radon measure  $f^* \in \mathcal{M}^1(\Omega)$  such that  $f(\cdot, t_j) \rightarrow f^*$  weakly-\* as  $t_j \rightarrow \infty$ . Since, as assumed,  $f$  is the unique classical solution with  $f \in C^1(\Omega \times \mathbb{R}_+)$ , we have  $f(\cdot, t_j) \rightarrow f^*$  weakly-\*, and due to the boundedness assumptions on  $\Omega$  or  $\psi$ ,  $\mathcal{E}[f(\cdot, t_j)] \rightarrow \mathcal{E}[f^*]$  as  $t_j \rightarrow \infty$ . Consequently,  $\mathcal{E}[f(\cdot, t)] \rightarrow \mathcal{E}[f^*]$  as  $t \rightarrow \infty$ . By (19), the equilibrium  $f^*$  is characterized by the condition (21). ■

By (20) the choice  $f^* \equiv w$  is always a solution of (21). This corresponds to the intuitive expectation that, as we let the number of particles  $N$  tend to infinity, we should recover the profile  $w$  in the long-time limit, regardless of the initial distribution of particles. However, the question whether the choice  $f^* \equiv w$  is the unique solution of (21) in the class of probability measures seems to be rather nontrivial. Although we believe that the affirmative indeed holds for a broad class of interaction potentials, we are so far only able to provide a proof for the special case  $\Omega = \mathbb{R}$  and  $\varphi(\cdot) = \psi(\cdot) = -|\cdot|$ . Unfortunately, this case does not match our assumptions on the smoothness of  $\varphi$  and  $\psi$  made in Theorem 1 and Lemmata 1 and 2. However, the weak formulation (15) of equation (11) perfectly makes sense if we insert the distributional derivative  $\varphi'(\cdot) = \psi'(\cdot) = \text{sign}(\cdot)$  into (12), as long as  $w, f \in L^1(\Omega)$ , since then the integrals

$$\int_{\Omega} w(x) \varphi'(p - x) \, dx = \int_{\Omega} w(x) \text{sign}(p - x) \, dx$$

and

$$\int_{\Omega} f(x) \psi'(p - x) \, dx = \int_{\Omega} f(x) \text{sign}(p - x) \, dx$$

are well defined for all  $p \in \mathbb{R}$  and uniformly bounded. Consequently, we can formulate the following Lemma:

**Lemma 5** *Let  $\Omega = \mathbb{R}$  and  $\varphi(\cdot) = \psi(\cdot) = -|\cdot|$ . Let  $w \geq 0$  be compactly supported in  $\Omega$  and such that  $\int_{\Omega} w(x) dx = 1$ . Then the solution  $f^* \equiv w$  of (21) is unique in the class  $\mathcal{X} := \{f \in L^1(\Omega) : f \geq 0 \text{ a.e. on } \Omega, \int_{\Omega} f(x) dx = 1\}$ .*

**Proof:** Let  $f \in \mathcal{X}$  fulfill (21), which can be recast as

$$\left| \int_{\Omega} (w(x) - f(x)) \text{sign}(p - x) dx \right|^2 f(p) = 0 \quad \text{a.e. on } \Omega. \quad (22)$$

Due to the normalization  $\int_{\Omega} w(x) dx = \int_{\Omega} f(x) dx = 1$ , we have

$$\begin{aligned} G(p) &:= \int_{\Omega} (w(x) - f(x)) \text{sign}(p - x) dx \\ &= \int_{-\infty}^p w(x) dx - \int_p^{\infty} w(x) dx - \int_{-\infty}^p f(x) dx + \int_p^{\infty} f(x) dx \\ &= 2W(p) - 2F(p), \end{aligned}$$

where we denoted

$$W(p) := \int_{-\infty}^p w(x) dx \quad \text{and} \quad F(p) := \int_{-\infty}^p f(x) dx.$$

The condition (22) can then be rewritten as  $|G(p)|^2 f(p) = 0$  for almost all  $p \in \Omega$ ; let us note that  $G$  is a continuous function. By assumption, there exist real numbers  $\alpha$  and  $\omega$  such that  $\text{supp } w \subset [\alpha, \omega]$ . We prove that  $G \equiv 0$  in three steps:

- First we show that  $G(p) \equiv 0$  for all  $p \leq \alpha$ . For a contradiction, let us assume that there exists a  $p_0 \leq \alpha$  such that  $G(p_0) \neq 0$ . Then we have  $G(p_0) = 2W(p_0) - 2F(p_0) = -2F(p_0) < 0$  and, therefore, there exists a set  $S \subset (-\infty, p_0)$  of positive Lebesgue measure, such that  $f > 0$  almost everywhere on  $S$ . But then  $F > 0$  almost everywhere on  $S$ , and, consequently,  $G = -2F < 0$  almost everywhere on  $S$ , a contradiction to  $|G(p)|^2 f(p) = 0$ .
- Using a symmetry argument, we can prove that

$$G(p) = -2 \int_p^{\infty} w(x) dx + 2 \int_p^{\infty} f(x) dx \equiv 0$$

for all  $p \geq \omega$ .

- Finally, we prove that  $G(p) \equiv 0$  also for all  $p \in (\alpha, \omega)$ . By the continuity of  $G$ , the set  $\mathcal{S} := \{p \in (\alpha, \omega) : G(p) \neq 0\}$  is open and  $G$  vanishes on its boundary  $\partial\mathcal{S}$ . Since (22) dictates that  $f(p) = 0$  almost everywhere on  $\mathcal{S}$ , we have for every  $p_0 \in \mathcal{S}$  and  $\delta > 0$  small enough,

$$G(p_0 + \delta) = G(p_0) + 2 \int_{p_0}^{p_0 + \delta} w(x) dx \geq G(p_0).$$

Therefore,  $G$  is nondecreasing on each subinterval of  $\mathcal{S}$ , and, again by the continuity of  $G$ , we conclude that  $G \equiv 0$ .

We finish the proof by observing that

$$\int_{-\infty}^p (w(x) - f(x)) dx = 0 \quad \text{for all } p \in \mathbb{R}$$

implies  $f = w$  almost everywhere on  $\mathbb{R}$ . ■

## 4 Numerical Examples

In this section we present few numerical results for the discrete particle system (6) and the mean-field limit (11). We consider two cases:

- Smoothing case:  $\varphi(s) = -|s|^{1.1}$  and  $\psi(s) = -|s|$
- Sharpening case:  $\varphi(s) = -|s|$  and  $\psi(s) = -|s|^{1.1}$

In both cases, we consider the 1D full-space setting  $\Omega = \mathbb{R}$  with the datum  $w = 4\chi_{[0.25, 0.5]}$ . For the discrete particle system, we use  $N = 20, 50$  and  $100$  particles. We integrate the ODE system (8) in time using the explicit Euler method until the steady state (which does not depend on the initial condition). The results for the smoothing and sharpening cases are shown in Fig. 2.

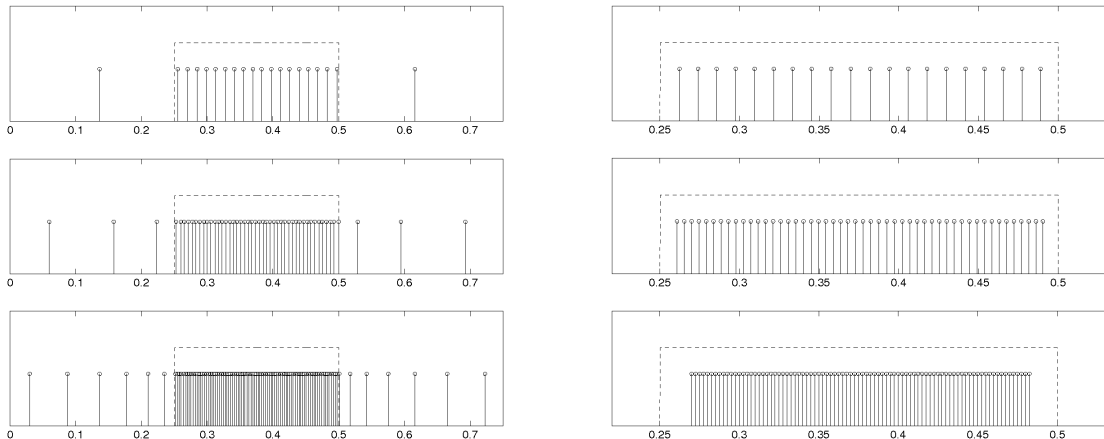


Figure 2: Steady states of the discrete particle system (8) with  $N = 20, 50, 100$  particles (first, second and, resp., third row) in the smoothing (left panels) and sharpening case (right panels). The datum  $w$  is visualized with the dashed line. Note the different horizontal axis limits in the smoothing and sharpening case.

For the mean-field limit (11), we impose the initial condition  $f_0 = 4\chi_{[0.65, 0.9]}$ . We discretize (11) using the semi-implicit finite difference method with upwinding in space and explicit Euler method in time. Snapshots of the solutions are shown in Fig. 3 for the smoothing case and in Fig. 4 for the sharpening case.

As we expect, the solutions converge to some steady states as  $t \rightarrow \infty$  in both the discrete and mean-field cases. In the smoothing case, the equilibrium profile is a smoothed version of

the data  $w$ , while in the sharpening case the equilibrium is an anti-smoothed version of  $w$ . It is interesting to compare these results with the discrete particle calculation from the previous example. Indeed, the steady state particle distributions shown on the left panels of Fig. 2 can be regarded as approximations of the steady state on Fig. 3, and the same hold for the right panels of Fig. 2 and the steady state of Fig. 4.

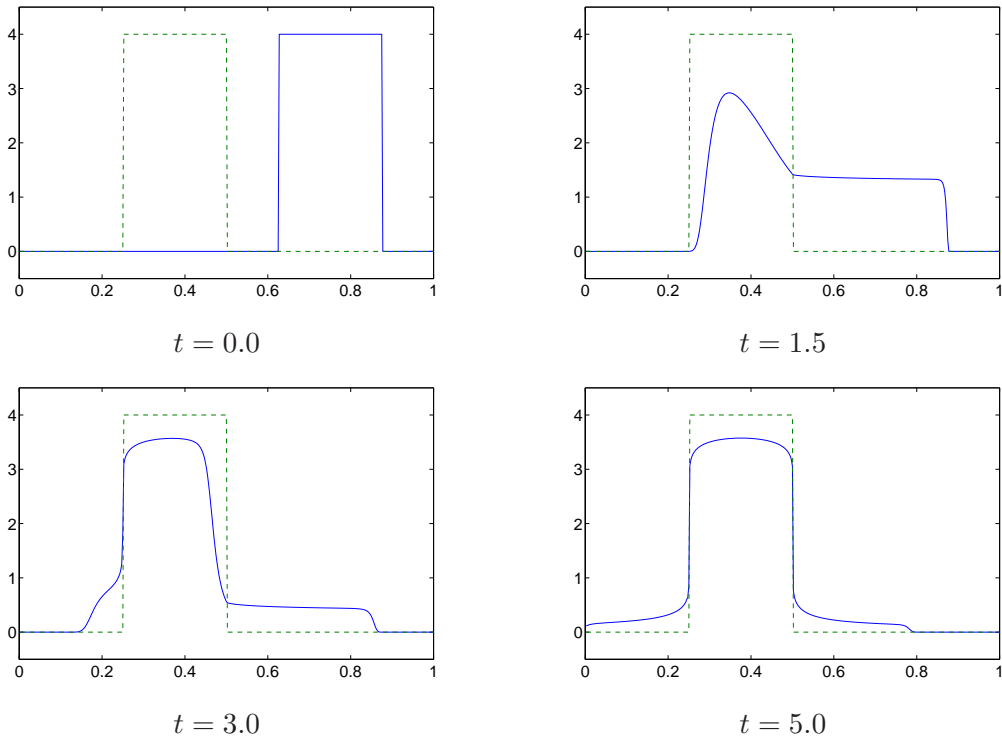


Figure 3: Smoothing case of the mean-field limit (11). The solid line represents the solution  $f$ , the dashed line the data  $w$ . The upper left panel shows the initial condition, the lower right panel is the steady state.

In our last example, we show how the kinetic equation (11) can be used for deblurring of images in the case when the blurring kernel is known. As our “image” we take again the datum  $w = 4\chi_{[0.25,5]}$ , and blur it by applying 200 time-steps of (11) with  $\varphi(s) = -|s|^{1.1}$  and  $\psi(s) = -|s|$  (smoothing case), the time-step length is 0.01. For simplicity, we use the initial datum  $f_0 := w$ . The result of this blurring process,  $g = f(\cdot, t = 2)$ , is shown on Fig. 5, left panel. Then, we perform deblurring of the “image” by applying (11) with reversed interaction potentials, i.e.,  $\varphi(s) = -|s|$  and  $\psi(s) = -|s|^{1.1}$ , and with  $w := g$ . Again, for simplicity, we take  $f_0 := w = g$  as the initial condition for  $f$ . We let (11) evolve until steady state, which is reached around  $t = 30$ , and plot it in the right panel of Fig. 5. We see that  $f(\cdot, t = 30)$  is indistinguishable from the original, un-blurred image  $w$ ; the relative difference between them in the  $L^1$ -norm is approximately 0.6%.

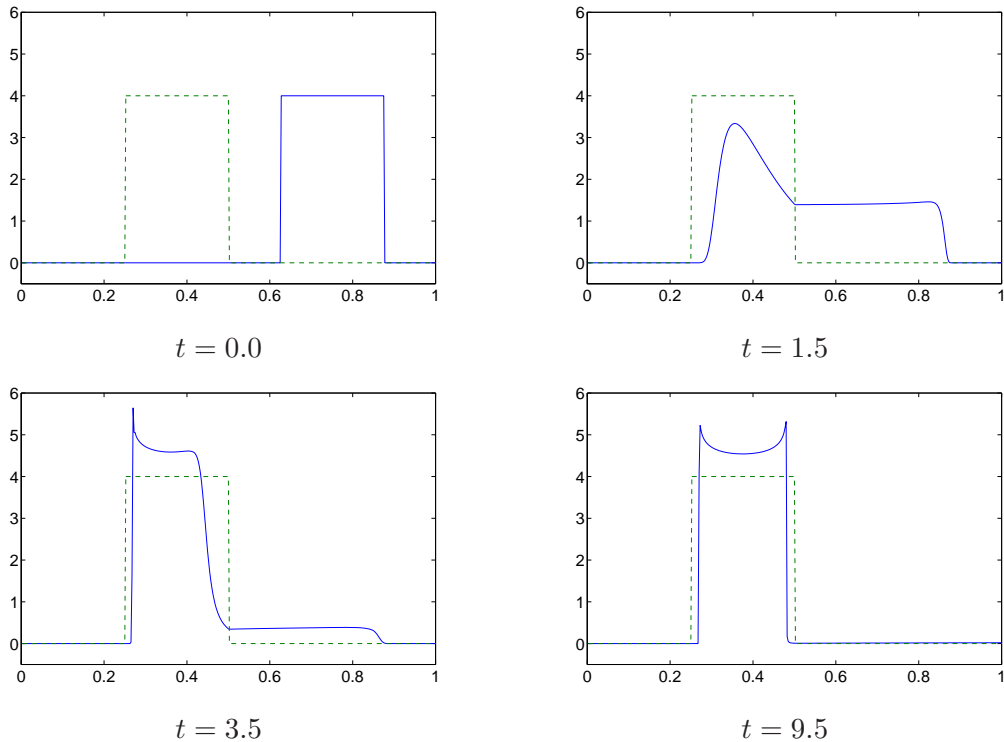


Figure 4: Sharpening case of the mean-field limit (11). The solid line represents the solution  $f$ , the dashed line the data  $w$ . The upper left panel shows the initial condition, the lower right panel is the steady state.

Let us mention that (11) in the spatially 2D setting potentially might be of interest for application in image processing, in particular for deblurring of images. However, it is well known that the process of deconvolution without further regularizations is not robust to noise. Therefore, a logical next step would be to introduce a regularization term into the expression for the energy (17), for instance the total variation term  $\int_{\Omega} |\nabla f(x, t)| dx$ . Moreover, let us remark that the numerical implementation of (11) in the spatially 2D setting is quite a demanding task due to the high numerical cost. This is caused by the necessity of evaluation of the convolution  $f * \nabla \psi$  in each time step, which in general takes  $\mathcal{O}(N^4)$  multiplications if the grid consists of  $N^2$  points. A possible speed-up of this operation can be achieved by using fast multipole expansion or FFT-based methods, see for instance [5]. We postpone these tasks for a future work.

### Acknowledgments

Massimo Fornasier and Jan Haškovec acknowledge the financial support provided by the START award “Sparse Approximation and Optimization in High Dimensions” no. FWF Y 432-N15 of the Fonds zur Förderung der wissenschaftlichen Forschung (Austrian Science Foundation).

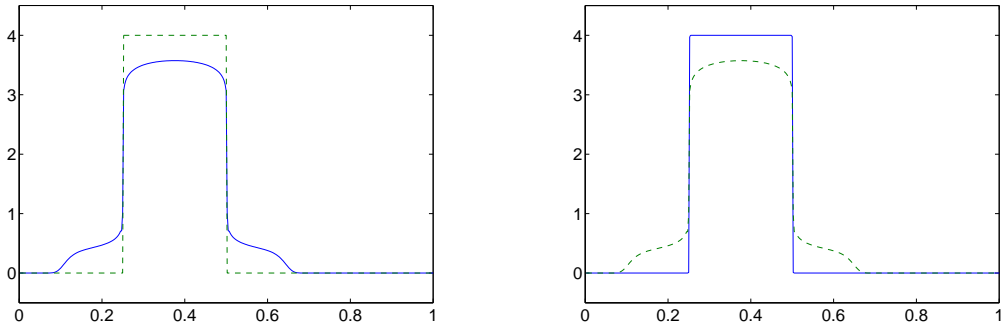


Figure 5: Left panel: The original image  $w$  (dashed line) and its blurred version  $\tilde{w} = f(t = 2)$  (solid line). Right panel: The result of deblurring of  $\tilde{w}$  (dashed line), given by the steady state  $f(\cdot, t = 30)$  of (11), solid line.

## References

- [1] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.
- [2] P. Malliavin. *Integration and Probability*. Graduate Texts in Mathematics 157, Springer Verlag, 1995, pp. 96-97.
- [3] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, 1998.
- [4] R. Fattal. Blue-noise point sampling using kernel density model. *ACM SIGGRAPH 2011*, 28(3):1–10, 2011.
- [5] D. Potts, G. Steidl and A. Nieslony Fast convolution with radial kernels at nonequispaced knots, *Numer. Math.*,98:329–351, 2004.
- [6] F. Golse. The mean-field limit for the dynamics of large particle systems. Journées “Équations aux Dérivées Partielles”, Exp. No. IX, 47 pp., Univ. Nantes, Nantes, 2003.
- [7] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Lecture Notes in Mathematics 1730, Springer, Berlin - Heidelberg - New York, 2000.
- [8] M. Gräf, D. Potts and G. Steidl. Quadrature rules, discrepancies and their relations to halftoning on the torus and the sphere, *SIAM J. Sci. Comput.*, to appear.
- [9] L. Hörmander. *The Analysis of Linear Partial Differential Operators I*. Springer, Berlin-Heidelberg-New York, 2003.
- [10] R. Jordan, D. Kinderlehrer and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* 29(1):1–17, 1998.
- [11] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* 26(1-2):101–174, 2001.



- [12] C. Schmaltz, P. Gwosdek, A. Bruhn and J. Weickert. Electrostatic halftoning. *Computer Graphics Forum* 29:2313–2327, 2010.
- [13] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970.
- [14] T. Teuber, G. Steidl, P. Gwosdek, C. Schmalz and J. Weickert. Dithering by differences of convex functions. *SIAM Journal on Mathematical Imaging* 4(1):79–108, 2011.
- [15] G. Wahba. *Spline Models for Observational Data*. SIAM, Philadelphia, 1990.
- [16] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, Cambridge, 2005.