



HAL
open science

Second order in time schemes for gradient flows in Wasserstein and geodesic metric spaces

Guillaume Legendre, Gabriel Turinici

► **To cite this version:**

Guillaume Legendre, Gabriel Turinici. Second order in time schemes for gradient flows in Wasserstein and geodesic metric spaces. 2016. hal-01317769v1

HAL Id: hal-01317769

<https://hal.science/hal-01317769v1>

Preprint submitted on 18 May 2016 (v1), last revised 24 Feb 2017 (v3)

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

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

Second order in time schemes for gradient flows in Wasserstein and geodesic metric spaces

Guillaume Legendre^a, Gabriel Turinici^{a,b}

^aUniversité Paris-Dauphine, PSL Research University, CNRS, UMR 7534, CEREMADE, 75016 Paris, France

^bInstitut Universitaire de France

Received *****; accepted after revision +++++

Presented by \mathbb{R}^n

Abstract

The time discretization of gradient flows in metric spaces uses variants of the celebrated implicit Euler-type scheme of Jordan, Kinderlehrer and Otto [8]. We propose in this Note a different approach which allows to construct two second order in time numerical schemes. We show that these schemes are well defined and prove (theoretically and numerically) the convergence for a gradient flow in the Wasserstein space corresponding to a Fokker–Planck equation.

To cite this article: G. Legendre, G. Turinici, C. R. Acad. Sci. Paris, Ser. I 340 (2016).

Résumé

Schémas d’ordre deux en temps pour des flots de gradient dans des espaces métriques géodésiques et de Wasserstein. La discrétisation temporelle des flots de gradient dans des espaces métriques utilise des variantes du schéma d’Euler implicite issu du travail séminal de Jordan, Kinderlehrer et Otto [8]. Nous proposons dans cette Note une approche différente permettant de construire deux schémas numériques d’ordre deux en temps. Nous montrons que ces schémas sont bien définis et que leur convergence est (théoriquement et numériquement) d’ordre deux pour un flot de gradient dans l’espace de Wasserstein correspondant à une équation de Fokker–Planck.

Pour citer cet article : G. Legendre, G. Turinici, C. R. Acad. Sci. Paris, Ser. I 340 (2016).

Version française abrégée

Soit (\mathcal{X}, d) un espace métrique et $F : (\mathcal{X}, d) \rightarrow \mathbb{R} \cup \{\infty\}$ une fonctionnelle. L’étude théorique et numérique du flot de gradient de F dans \mathcal{X} utilise souvent une version variationnelle du schéma d’Euler

Email addresses: Guillaume.Legendre@dauphine.fr (Guillaume Legendre), Gabriel.Turinici@dauphine.fr (Gabriel Turinici).

implicite, introduit par Jordan, Kinderlehrer et Otto [8] (noté dorénavant JKO), consistant à avancer de τ unités du temps à partir du point courant X_n par le choix d'un point X_{n+1} minimisant (1). Ceci, couplé avec une interpolation adéquate, permet de démontrer l'existence du flot de gradient, mais aussi de calculer numériquement une approximation d'ordre $O(\tau)$ de ce flot. Nous proposons dans cette Note deux schémas d'ordre deux en temps.

Étant donné \bar{X} dans un espace géodésique \mathcal{X} , notons par $\frac{X+Y}{2}$ l'ensemble des points milieux du couple (X, Y) (voir [7, Chapitre 2] et [1]). Le schéma variationnel implicite du point milieu (VIM) est défini de la manière suivante : pour avancer en temps τ unités de temps à partir de \bar{X} , nous cherchons X et $Z \in \frac{\bar{X}+X}{2}$ qui vérifient (2). Nous montrons le résultat suivant.

Theorem 0.1 *Soit (\mathcal{X}, d) un espace polonais tel que, pour tout $U \in \mathcal{X}$, l'ensemble $\cup_{Y \in \mathcal{X}} \frac{U+Y}{2}$ est fermé. Supposons aussi que la fonctionnelle F est semi-continue inférieurement, bornée inférieurement et telle que, pour tous $r > 0$ et $c \in \mathbb{R}$ l'ensemble $\{Z \in \mathcal{X} \mid F(Z) \leq c, d(\bar{X}, Z) \leq r\}$ est compact. Alors il existe un couple $(X, Z) \in \mathcal{X}^2$ tels que (2) ait lieu.*

Supposons à présent que \mathcal{X} est non-branching (voir [12]). On observe alors que, si l'on connaît X et $Z \in \frac{X+Y}{2}$, le point Y est déterminé de manière unique et appelé le 2-extrapolé de X par rapport à Z . Si pour (X, Z) donné dans \mathcal{X}^2 , il n'existe pas de Y tel que $Z \in \frac{X+Y}{2}$, le 2-extrapolé est défini comme étant le point le plus éloigné de X sur toute géodésique partant de X et contenant Z . Avec ces définitions, nous constatons qu'une autre possibilité pour définir un schéma d'ordre deux arrive de l'identité (3). Le schéma variationnel d'Euler implicite extrapolé (EVIE) est ainsi défini en construisant Z à partir de \bar{X} avec le schéma JKO utilisant un pas de temps $\tau/2$, puis en prenant X comme le 2-extrapolé de \bar{X} par rapport à Z .

Dans le cas particulier du flot de gradient de F définie par (4) et d'un potentiel de la forme $V(x) = \theta \frac{(x-\mu)^2}{2}$ placé dans l'espace de Wasserstein sur \mathbb{R} , nous montrons que ces deux schémas convergent à l'ordre deux en temps, aussi bien en version discrète en temps et exacte en espace qu'en version discrète en temps et espace.

1. Background on gradient flows and relationship with the implicit Euler scheme

Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be a smooth convex function; its gradient flow is the curve solution to $X'(t) = -\nabla f(X(t))$. A subject that received considerable attention recently is the extension of this definition to a Polish metric space (\mathcal{X}, d) and a functional $F : (\mathcal{X}, d) \rightarrow \mathbb{R} \cup \{+\infty\}$ (see [3,13,2,12]). An important advance was in particular the work Jordan, Kinderlehrer and Otto [8] that introduced the following numerical scheme: given X_n , select X_{n+1} as a minimizer of

$$X \mapsto P_F^{JKO}(X; X_n, \tau) := \frac{1}{2} d(X_n, X)^2 + \tau F(X). \quad (1)$$

When the metric space is a Hilbert space and the functional F is smooth enough, the above scheme becomes the implicit Euler scheme, i.e., $\frac{X_{n+1}-X_n}{\tau} = -F'(X_{n+1})$. As a consequence, the Jordan–Kinderlehrer–Otto method (named JKO from now on) can be seen as a variational generalization of the implicit Euler scheme in a metric space. Coupled with a suitable interpolation, the sequence constructed by this scheme defines, as the time step τ tends to zero, a (metric) gradient flow, the convergence being of order one in τ . As a natural numerical counterpart of the theoretical procedure, the JKO scheme was thus used to compute the gradient flow in many applications (see [9,5,4] for instance).

However, while the JKO scheme has theoretical advantages, its first order convergence may be judged insufficient in practice. When the time dependence is smooth enough, one may seek to replace it with higher order alternatives. The purpose of this Note is to propose two second order schemes, which are

both easy to implement (one of them uses directly a step of the JKO method). We also give several preliminary theoretical and numerical results pertaining to these schemes.

2. Second order numerical schemes on metric spaces

Let (\mathcal{X}, d) be a Polish metric space. A curve $\gamma : [0, 1] \rightarrow \mathcal{X}$ is called a (constant speed) geodesic (see also [2, page 37]) provided that $d(\gamma(t), \gamma(s)) = |t - s| \cdot d(\gamma(0), \gamma(1))$, and the space is called geodesic if, for any couple of points $(X, Y) \in \mathcal{X}^2$, there exists a geodesic γ connecting them, that is, such that $\gamma(0) = X$, $\gamma(1) = Y$. The set of all geodesics connecting X and Y is denoted $Geod_{X,Y}$. Moreover $\{\gamma(1/2) \mid \gamma \in Geod_{X,Y}\}$ is called the set of midpoints of the couple (X, Y) and will be denoted $\frac{X+Y}{2}$ (see [7, Chapter 2] for general metric spaces and [1] for Wasserstein spaces).

Given $\bar{X} \in \mathcal{X}$, we define the *Variational Implicit Midpoint (VIM) scheme* starting from \bar{X} for the (gradient flow of the) functional F by selecting a point X such that, for some $Z \in \frac{\bar{X}+X}{2}$,

$$\forall Y \in \mathcal{X}, \forall U \in \frac{\bar{X}+Y}{2}, \frac{1}{2} d(\bar{X}, X)^2 + 2\tau F(Z) \leq \frac{1}{2} d(\bar{X}, Y)^2 + 2\tau F(U). \quad (2)$$

Intuitively, this means that X is a minimizer of $\frac{1}{2} d(\bar{X}, X)^2 + 2\tau F\left(\frac{\bar{X}+X}{2}\right)$. In the particular case of a Hilbert space and a smooth functional F , the point X is effectively a solution of the equation $\frac{X-\bar{X}}{\tau} + F'\left(\frac{\bar{X}+X}{2}\right) = 0$, which characterizes the implicit midpoint rule method, a one-stage implicit Runge–Kutta method of second order [11, Chapter 16, equation (16.6.29)].

We now prove that the VIM scheme is well-defined by establishing the existence of a solution to (2).

Theorem 2.1 *Let (\mathcal{X}, d) be a Polish metric space such that, for any $U \in \mathcal{X}$, the set $\cup_{Y \in \mathcal{X}} \frac{U+Y}{2}$ is closed. Assume that F is lower semicontinuous, bounded from below, and such that, for all $r > 0$ and $c \in \mathbb{R}$, the set $\{U \in \mathcal{X} \mid F(U) \leq c, d(\bar{X}, U) \leq r\}$ is compact. Then, there exists $(X, Z) \in \mathcal{X}^2$ such that (2) holds.*

Proof Let $(X_n, Z_n)_{n \geq 1}$ be a minimizing sequence for (2). In particular, the sequence $(\frac{1}{2} d(\bar{X}, X_n)^2)_{n \geq 1}$ is bounded and so is $(d(\bar{X}, Z_n))_{n \geq 1}$. On the other hand, the sequence $(F(Z_n))_{n \geq 1}$ is also bounded, so that, by hypothesis, the sequence $(Z_n)_{n \geq 1}$ lives in a compact set and thus converges, up to a subsequence, to some Z . Since $\cup_{Y \in \mathcal{X}} \frac{\bar{X}+Y}{2}$ is closed, there exists $X \in \mathcal{X}$ such that $Z \in \frac{\bar{X}+X}{2}$. In addition, recalling that $Z_n \in \frac{\bar{X}+X_n}{2}$, we obtain that $d(\bar{X}, X) = 2 d(\bar{X}, Z) = 2 \lim_{n \rightarrow \infty} d(\bar{X}, Z_n) = \lim_{n \rightarrow \infty} d(\bar{X}, X_n)$. The lower semicontinuity property of F then allows to conclude. ■

Remark 1 *The assumptions on the functional F in the theorem are classical (see [2, Section 4.2.2] for instance) but can be weakened. In particular, we do not need F to be bounded from below.*

Note that, for any $Z \in \frac{\bar{X}+X}{2}$, one has

$$\frac{1}{2} d(\bar{X}, X)^2 + 2\tau F(Z) = 2 d(\bar{X}, Z)^2 + 2\tau F(Z) = 4 P_F^{JKO}(Z; \bar{X}, \tau/2). \quad (3)$$

Another way to interpret the VIM numerical scheme is thus to view the midpoint Z as computed by a JKO scheme with time step $\tau/2$ and to *extrapolate* it as described below.

Recall that a geodesic space (\mathcal{X}, d) is called non-branching (see [2, Definition 3.15] or [3, Theorem 11.2.10]) if, for any $t \in]0, 1[$, a constant speed geodesic γ is uniquely determined by its initial point $\gamma(0)$ and by the point $\gamma(t)$. In particular, if $U \in \frac{X+Y_1}{2} \cap \frac{X+Y_2}{2}$ then $Y_1 = Y_2$. By definition, the 2-extrapolate of the couple (X, U) (denoted $2U - X$ hereafter) is the unique point Y such that $U \in \frac{X+Y}{2}$. On the contrary, when there exists no such point, the 2-extrapolate is defined as the point Y with largest distance to X such that U is on a geodesic from X to Y . Since the space is non-branching and complete, using the properties of the distance between two points on a geodesic and the fact that any Cauchy sequence converges, we

obtain that for fixed X and U , the set $\{d(X, Y) \mid \exists \gamma \in \text{Geod}_{X, Y}, U \in \gamma\}$ is closed. The 2-extrapolate is thus well defined and unique. It moreover satisfies $d(\bar{X}, U) \leq d(\bar{X}, 2U - \bar{X}) \leq 2d(\bar{X}, U)$.

These considerations lead us to define the *Extrapolated Variational Implicit Euler (EVIE) scheme*, as the 2-extrapolate $2Z - \bar{X}$ of the solution Z of the JKO scheme with time step $\tau/2$.

From the implementation point of view, the EVIE scheme has the obvious advantage of using the JKO scheme. From a theoretical point of view, the existence and uniqueness of a solution in a non-branching space is assured under the same hypothesis as those required for the JKO scheme. While adaptations certainly allow to define it in branching spaces, uniqueness may be lost, even when the JKO scheme possesses a unique solution at each step.

Remark 2 Combining first and second order schemes allow to build adaptive stepsize schemes cf. [11, Section 16.2].

3. Second order schemes for the Fokker–Planck equation as a gradient flow in the Wasserstein space

In this section the space \mathcal{X} is the 2-Wasserstein space $\mathcal{P}_2(\mathbb{R})$ endowed with the 2-Wasserstein metric (see [12, Chapter 5] or [8,13]). Let $\nu \in \mathcal{P}_2(\mathbb{R})$. When ν is absolutely continuous with respect to the Lebesgue measure, with density ρ , the functional F reads:

$$F(\nu) = E(\rho) + S(\rho) \text{ with } S(\rho) = \frac{\sigma^2}{2} \int_{\mathbb{R}} \rho(x) \log(\rho(x)) dx \text{ and } E(\rho) = \int_{\mathbb{R}} V(x)\rho(x) dx. \quad (4)$$

When ν is not absolutely continuous with respect to the Lebesgue measure, we set $F(\nu) = +\infty$. For a smooth potential V the gradient flow in $\mathcal{P}_2(\mathbb{R})$ of the functional F is a curve in $t \mapsto \nu(t) \in \mathcal{P}_2(\mathbb{R})$, each $\nu(t)$ being absolutely continuous with respect to the Lebesgue measure (see [8]), $\nu(t) = \rho(t, x) dx$ where $\rho(t, x)$ is solution of the Fokker-Planck equation:

$$\frac{\partial \rho}{\partial t}(t, x) = \frac{\partial}{\partial x} [V'(x)\rho(t, x)] + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}(t, x), \quad (5)$$

associated with the Stochastic Differential Equation (SDE) $dX(t) = -V'(X(t))dt + \sigma dW_t$, where W is the one dimensional Brownian motion.

We consider quadratic potentials of the form $V(x) = \theta \frac{(x-\mu)^2}{2}$ and denote by $\mathcal{N}(a, b^2)$ the probability law of a normal variable with mean a and variance b^2 and also, by abusing the notation, its density (with respect to the Lebesgue measure).

Proposition 3.1 *If $\rho(0, \cdot) = \mathcal{N}(\mu_0, \sigma_0^2)$, any intermediary state of the semi-discrete VIM and EVIE schemes is a Gaussian and the error of both schemes has order $O(\tau^2)$.*

Proof We will prove the assertion for the VIM scheme, the proof for the EVIE one being similar. When the potential V is quadratic, the solution of the SDE is a Ornstein–Uhlenbeck process $X(t) = X(0)e^{-\theta t} + \mu(1 - e^{-\theta t}) + \frac{\sigma}{\sqrt{2\theta}}e^{-\theta t}W_{e^{2\theta t}-1}$. In particular, the exact evolution from $\rho_n = \mathcal{N}(\mu_n, \sigma_n^2)$ after τ time units is

$$\rho_{n+1}^{exact} = \mathcal{N}\left(e^{-\theta\tau}\mu_n + \mu(1 - e^{-\theta\tau}), e^{-2\theta\tau}\sigma_n^2 + \frac{\sigma^2}{2\theta}(1 - e^{-2\theta\tau})\right).$$

On the other hand, ρ_n being absolutely continuous with respect to the Lebesgue measure, the set $\text{Geod}_{\rho_n, \rho}$ is a singleton for any $\rho \in \mathcal{P}_2(\mathbb{R})$ (and so is the set $\frac{\rho_n + \rho}{2}$). The unique geodesic $\gamma_\rho \in \text{Geod}_{\rho_n, \rho}$ is determined by a map $T : \mathbb{R} \rightarrow \mathbb{R}$ (which is the gradient of a convex function, see [6,13]), $\gamma_\rho(t)$ being the push-forward of ρ_n through $x \mapsto (1-t)x + tT(x)$. The minimization in (2) can thus be expressed in terms of the map T , for instance $\frac{1}{2}d(\rho_n, \rho)^2 = \int_{\mathbb{R}} \frac{(T(x)-x)^2}{2}\rho_n(x) dx$. After some tedious but rather straightforward computations, one obtains that the unique element ρ_{n+1}^{VIM} minimizing $\frac{1}{2}d(\rho_n, \rho)^2 + 2\tau F\left(\frac{\rho_n + \rho}{2}\right)$

is also a Gaussian $\rho_{n+1}^{VIM} = \mathcal{N}(\mu_{n+1}, \sigma_{n+1}^2)$, with $\mu_{n+1} = (1 + \frac{\theta\tau}{2})^{-1} [\mu_n (1 - \frac{\theta\tau}{2}) + \frac{\mu\theta\tau}{2}]$, $\sigma_{n+1} = \left[\frac{\sqrt{\sigma_n^2 + \sigma^2\tau(1 + \theta\tau/2)} - \sigma_n\theta\tau/2}{[1 + \theta\tau/2]} \right]$. Consequently, ρ_{n+1}^{VIM} is an approximation of ρ_{n+1}^{exact} exact up to second order and, by usual Gronwall inequalities, the error of the VIM scheme is globally of order $O(\tau^2)$. \blacksquare

3.1. Fully discrete schemes

To illustrate the performances of the proposed schemes, we consider the numerical solution of a one-dimensional Fokker–Planck equation (see [4,10] for possible approaches in higher dimensions).

Roughly speaking, we employ a continuous piecewise affine (P1) finite element discretization for the cumulative distribution functions of elements in $\mathcal{P}_2(\mathbb{R})$, which amounts to replace the space $\mathcal{P}_2(\mathbb{R})$ by a space $\mathcal{P}_2(\mathbb{R})_m$ defined as follows. Let $M \in \mathbb{N}$ and $m = (m_1, \dots, m_M)$ be a (ordered) mass distribution with $m_k > 0$, $\sum_{k=1}^M m_k = 1$, fixed once for all. An element in $\mathcal{P}_2(\mathbb{R})_m$ with density $\rho_X(\cdot)$ is defined by a vector $X = (x_0, \dots, x_M) \in \mathbb{R}^{M+1}$, with $x_0 < x_1 < \dots < x_{M-1} < x_M$, by setting $\rho_X(\cdot)$ constant and equal to $\frac{m_k}{x_{k+1} - x_k}$ on the segment $[x_k, x_{k+1})$.

The potential energy $E(\rho_X) = \int_{\mathbb{R}} \rho_X(x)V(x)dx$ is discretized by an operator E_m where the integral is approximated by the Simpson rule:

$$E_m(X) = \sum_{k=0}^M \frac{m_k}{6} \left[V(x_k) + 4V\left(\frac{x_k + x_{k+1}}{2}\right) + V(x_{k+1}) \right]. \quad (6)$$

The internal energy $S(\rho_X)$ is computed exactly by the operator S_m and has the simple form:

$$S_m(X) = \sum_{k=0}^M \frac{m_k}{x_{k+1} - x_k} (x_{k+1} - x_k) \log \left[\frac{m_k}{x_{k+1} - x_k} \right] = \sum_{k=0}^M m_k \log(m_k) - \sum_{k=0}^M m_k \log(x_{k+1} - x_k). \quad (7)$$

To compute the W_2 distance between, say ρ_{X_1} and ρ_{X_2} , note that the map T is the unique piecewise linear function on any interval $[x_k^1, x_{k+1}^1]$ which associates x_k^1 to x_k^2 and x_{k+1}^1 to x_{k+1}^2 . Straightforward computations show that the distance $\text{dist}_{W_2}(\rho_{X_1}, \rho_{X_2})^2 = \int (T(x) - x)^2 \rho_{X_1}(x)dx$ reduces to

$$\text{dist}_{W_2}(\rho_{X_1}, \rho_{X_2})^2 = \sum_{k=0}^M m_k [(x_k^2 - x_k^1)^2 + (x_{k+1}^2 - x_{k+1}^1) \cdot (x_k^2 - x_k^1) + (x_{k+1}^2 - x_{k+1}^1)^2] / 3. \quad (8)$$

We use final time equal to 1, $\sigma = 1$, $V(x) = (x - 5)^2/2$, $M = 32$, $m_k = 1/M$, and the initial datum is a centered reduced Gaussian (projected on the discrete space). The results presented in Figure 1 confirm numerically that both schemes are of second order in time.

The internal energy $S(\rho)$ constitutes a barrier enforcing the constraint $x_k < x_{k+1}$. As such the (continuous in time) dynamics can also be viewed as a dynamics of the vector X which follows a, cumbersome but regular, system of ordinary differential equations. The proposed schemes provide a consistent, second order in time, discretization of this equation. We have thus obtained the following result.

Corollary 3.1 *The scheme (2) with space discretization (6), (7), (8) converges at second order in time to the exact solution of the semi-discrete Fokker–Planck gradient flow.*

Acknowledgements

G.T. acknowledges support from the Agence Nationale de la Recherche (ANR), projects EMAQS ANR-2011-BS01-017-01, CINE-PARA 2015-19 and ISOTACE (ANR-12-MONU-0013).

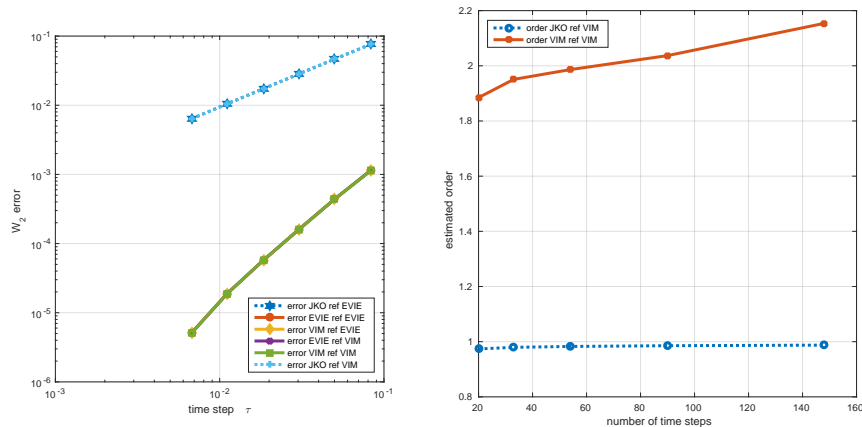


Figure 1. Results for the Section 3.1. **Left:** The error of the JKO (dotted) and second order (VIM, EVIE) (solid) schemes. The time steps are $1/4, 1/7, 1/12, 1/20, 1/33, 1/54, 1/90, 1/148$. The error shown is the W_2 distance between the numerical solution and a fine (VIM and EVIE) numerical solution obtained with 244 time steps. Four time steps with any second order scheme are as accurate as 90 steps with standard JKO scheme; similarly, 7 VIM / EVIE steps are as good as 144 JKO steps. The points of the VIM and EVIE schemes are indistinguishable to machine precision. **Right:** the numerical estimation of the order of convergence: dotted line the JKO scheme, solid line the second order schemes.

References

- [1] Martial Agueh and Guillaume Carlier. Barycenters in the Wasserstein space. *SIAM J. Math. Anal.*, 43(2):904–924, 2011.
- [2] Luigi Ambrosio and Nicola Gigli. *Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009*, Editors: Benedetto Piccoli, Michel Rascle, chapter A User’s Guide to Optimal Transport, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures. 2nd ed.* Basel: Birkhäuser, 2nd ed. edition, 2008.
- [4] Jean-David Benamou, Guillaume Carlier, Quentin Mérigot, and Édouard Oudet. Discretization of functionals involving the Monge–Ampère operator. *Numer. Math.*, pages 1–26, 2015.
- [5] Adriën Blanchet, Vincent Calvez, and José A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.
- [6] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.*, 44(4):375–417, 1991.
- [7] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [8] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [9] David Kinderlehrer and Noel J. Walkington. Approximation of parabolic equations using the Wasserstein metric. *M2AN Math. Model. Numer. Anal.*, 33(4):837–852, 1999.
- [10] Jean-Marie Mirebeau. Discretization of the 3D Monge-Ampère operator, between wide stencils and power diagrams. *ESAIM Math. Model. Numer. Anal.*, 49(5):1511–1523, 2015.
- [11] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C++*. Cambridge University Press, Cambridge, 2002. The art of scientific computing, Second edition, updated for C++.
- [12] Filippo Santambrogio. *Optimal transport for applied mathematicians*. Progress in Nonlinear Differential Equations and their Applications, 87. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
- [13] Cédric Villani. *Optimal transport, Old and new*, volume 338. Springer-Verlag, Berlin, 2009.