Fixed Point Theory, 26(2025), No. 1, 37-56
DOI: 10.24193/fpt-ro.2025.1.03
http://www.math.ubbcluj.ro/~nodeacj/sfptcj.html

QUASI α-FIRMLY NONEXPANSIVE MAPPINGS IN WASSERSTEIN SPACES

ARIAN BËRDËLLIMA* AND GABRIELE STEIDL**

*Faculty of Engineering, German International University in Berlin,
Am Borsigturm 162, 13507 Berlin, Germany*
E-mail: arian.berdellima@giu-berlin.de

**Institute of Mathematics, Technische Universität Berlin,
Straße des 17. Juni 136, D-10623 Berlin, Germany**
E-mail: steidl@math.tu-berlin.de

Abstract. This paper introduces the concept of quasi α-firmly nonexpansive mappings in Wasserstein spaces over $\mathbb{R}^d$ and analyzes properties of these mappings. We prove that for quasi α-firmly nonexpansive mappings satisfying a certain quadratic growth condition, the fixed point iterations converge in the narrow topology. As a byproduct, we will get the known convergence of the proximal point algorithm in Wasserstein spaces. We apply our results to show for the first time that cyclic proximal point algorithms for minimizing the sum of certain functionals on Wasserstein spaces converge under appropriate assumptions.
Key Words and Phrases: Quasi α-firmly nonexpansive mappings, Wasserstein spaces, fixed point, cyclic proximal algorithm, push-forward operator, optimal transport.
2020 Mathematics Subject Classification: 46T99, 47H10, 47J25, 28A33.

1. Introduction

Splitting algorithms which include proximal operators have recently found broad interest both in Hilbert spaces [8] and nonlinear CAT(0) spaces [6]. For certain applications in finite dimensional linear spaces we refer to the overview papers [16], and in Hadamard manifolds to [7, 12]. On the other hand, Wasserstein spaces and Wasserstein proximal mappings are popular in connection with gradient and neural gradient flows [22, 33].

In this paper, we introduce the concept of quasi α-firmly nonexpansive mappings in Wasserstein-2 spaces over $\mathbb{R}^d$. For linear spaces such operators were examined in various papers, see, e.g. [8, 11, 26]. The main motivation for studying (quasi) α-firmly nonexpansive operators in linear spaces, particularly in Hilbert spaces, is their connection with the so-called averaged operators which are essential in fixed point theory, see, e.g. the classical works [17, 18, 24, 27]. In the context of nonlinear CAT(0) spaces (quasi) α-firmly nonexpansive mappings were introduced in [9] and later extended in [10] to more general settings. For $d = 1$ the Wasserstein space is CAT(0) and the theory about (quasi) α-firmly nonexpansive operators follows from [9]. Therefore our

theory is a new contribution in the case $d \geq 2$. We will see that quasi α-firmly nonexpansive mappings in Wasserstein spaces are closed under compositions of operators, whenever they share a common fixed point. Prominent examples of such mappings are the Wasserstein proximal mappings of proper, lower semicontinuous, coercive functions that are convex along generalized geodesics. Also the push-forward mappings of measures by an α-firmly nonexpansive operator in $\mathbb{R}^d$ constitute an example of its own interest. As an application of such mappings, we consider the cyclic proximal point algorithm. In contrast to CAT(0) spaces, Wasserstein spaces have a positive Alexandrov curvature [1] for $d \geq 2$, which makes the analysis of algorithms including these operators in general quite tricky. Under appropriate conditions we show that the iterations of this algorithm converge in the narrow topology to a minimizer of a given finite sum of proper, lower semicontinuous, coercive functions that are convex along generalized geodesics. Both situations when these functions share or don't share a common minimizer are treated. In the latter case, Lipschitz continuity of each constituent function is needed. These results have direct relations to finding the minimum of certain energy and relative entropy functionals, see [3, §9.3, §9.4]. It is known that the minima of such functionals are the stationary solutions of corresponding stochastic differential equations and that the corresponding density functions appear as solutions of partial differential equations as, e.g. the well-examined Fokker–Planck equation, see, e.g., [23].

The outline of this paper is as follows: Section 2 contains the basic notation required for our analysis in Wasserstein spaces. In Section 3, we study proximal mappings of functions that are proper, lsc, coercive and convex along generalized geodesics. Then, in Section 4, we introduce the concept of quasi α-firmly nonexpansive mappings in Wasserstein spaces. We show that proximal mappings of certain functions are quasi $\frac{1}{2}$-firmly nonexpansive. Further, we examine push-forward operators of measures for operators on $\mathbb{R}^d$ which are themselves (quasi) α-firmly nonexpansive. Since α-firmly nonexpansive operators in Hilbert spaces are important due to the related fixed point theory, we examine the fixed point properties of such operators in Wasserstein spaces in Section 5. As in Hilbert spaces, the path to go is via Opial's property and Fejér's monotonicity. In Section 6, we apply our results to prove the convergence of the cyclic proximal point algorithm.

2. Preliminaries

The following section provides the necessary facts and notation on Wasserstein spaces as they can be found in several textbooks as [3, 2, 32, 34]. For applications we also refer to [19] and for novel developments about the sliced-Wasserstein distances see [15, 14].

Let $\mathbb{R}^d$, $d \geq 1$ be equipped with the Euclidean norm $\|\cdot\|$, and let $\mathcal{B}(\mathbb{R}^d)$ be its Borel σ-algebra. By $\mathcal{P}_2(\mathbb{R}^d)$, we denote the set of probability measures on $\mathcal{B}(\mathbb{R}^d)$ with finite second moments. With the *L^2-Wasserstein metric*

$$W_2(\mu,\nu) := \Big(\min_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x-y\|^2 \, d\pi(x,y) \Big)^{1/2}, \tag{2.1}$$

where $\Pi(\mu,\nu)$ denotes the set of all transport plans between μ and ν, i.e.,

$$\pi(A\times\mathbb{R}^d)=\mu(A) \text{ and } \pi(\mathbb{R}^d\times B)=\nu(B) \quad \text{for all } A,B\in\mathcal{B}(\mathbb{R}^d), \tag{2.2}$$

the space $\mathcal{P}_2(\mathbb{R}^d)$ becomes a separable, complete metric space, called Wasserstein space or briefly Wasserstein space. Let $\Pi_{opt}(\mu,\nu)$ denote the set of optimal transport plans, i.e. the set of all elements in $\Pi(\mu,\nu)$ that attain (2.1).

Remark 2.1. *If μ is absolutely continuous with respect to the Lebesgue measure, then the optimal transport plan π_μ^ν is unique and is induced by the unique minimizer T_μ^ν of the so-called Monge problem*

$$\tilde{W}_2(\mu,\nu)=\inf_T\int_{\mathbb{R}^d}\|x-T(x)\|^2\,d\mu(x) \quad \textit{subject to } \nu=T_\#\mu:=\mu\circ T^{-1} \tag{2.3}$$

by $\pi_\mu^\nu=(\mathrm{Id},T_\mu^\nu)_\#\mu$. In this case $\tilde{W}_2(\mu,\nu)$ coincides with $W_2(\mu,\nu)$. The situation changes if μ is not absolutely continuous. Then, in contrast to the minimization problem (2.1)*, which is also known as Kantorovich problem, the Monge problem* (2.3) *may fail to have a minimizer. Further, if an optimal transport map T_μ^ν in* (2.3) *exists, then $\pi:=(\mathrm{Id},T_\mu^\nu)_\#\mu\in\Pi(\mu,\nu)$, i.e. this plan π fulfills the marginal conditions. However, it need not be optimal as the example $\mu:=\frac{1}{4}\delta_0+\frac{3}{4}\delta_1$ and $\nu:=\frac{3}{4}\delta_0+\frac{1}{4}\delta_1$ shows.*

A sequence $(\mu_n)_{n\in\mathbb{N}}\subset\mathcal{P}_2(\mathbb{R}^d)$ converges to $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, denoted by $\mu_n\to\mu$, if

$$\lim_{n\to\infty}W_2(\mu_n,\mu)=0.$$

A sequence $(\mu_n)_{n\in\mathbb{N}}\subset\mathcal{P}_2(\mathbb{R}^d)$ *converges narrowly* to $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, denoted by $\mu_n\overset{\mathcal{N}}{\to}\mu$, if

$$\int_{\mathbb{R}^d}\varphi(x)\,d\mu_n(x)\to\int_{\mathbb{R}^d}\varphi(x)\,d\mu(x) \quad \text{for all } \varphi\in C_b(\mathbb{R}^d). \tag{2.4}$$

The relation between both topologies is given by the following theorem.

Theorem 2.2. [34, Theorem 6.9] *For $(\mu_n)_{n\in\mathbb{N}}\subset\mathcal{P}_2(\mathbb{R}^d)$, we have $\mu_n\to\mu$ if and only if $\mu_n\overset{\mathcal{N}}{\to}\mu$ and*

$$\int_{\mathbb{R}^d}\|x\|^2\,d\mu_n(x)\to\int_{\mathbb{R}^d}\|x\|^2\,d\mu(x) \quad \textit{as } n\to+\infty.$$

For all $\nu\in\mathcal{P}_2(\mathbb{R}^d)$, the Wasserstein metric $W_2(\cdot,\nu)$ is lower semicontinuous in the narrow topology, i.e. $W_2(\mu,\nu)\le\liminf_{n\to+\infty}W_2(\mu_n,\nu)$ whenever $\mu_n\overset{\mathcal{N}}{\to}\mu$, see [34, Lemma 4.3]. An important concept is the tightness of a set in $\mathcal{P}_2(\mathbb{R}^d)$. A set $S\subseteq\mathcal{P}_2(\mathbb{R}^d)$ is *tight*, if for every $\varepsilon>0$ there exists a compact set $K_\varepsilon\subseteq\mathbb{R}^d$ such that $\mu(\mathbb{R}^d\setminus K_\varepsilon)\le\varepsilon$ for all $\mu\in S$.

Theorem 2.3 (Prokhorov's Theorem [30])**.** *A set $S\subseteq\mathcal{P}_2(\mathbb{R}^d)$ is tight if and only if S is relatively compact in the topology of narrow convergence.*

In particular, we have the following lemma.

Lemma 2.4. [35, Theorem 1] *Closed balls in $\big(\mathcal{P}_2(\mathbb{R}^d),W_2\big)$ are tight.*

The Wasserstein spaces are so-called *geodesic spaces*, meaning, that for every $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a curve $\gamma : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ with $\gamma(0) = \mu$, $\gamma(1) = \nu$ and

$$W_2(\gamma(t), \gamma(s)) = |t - s| W_2(\gamma(0), \gamma(1)) \quad \text{for every } t, s \in [0,1]. \tag{2.5}$$

A curve $\gamma : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ with property (2.5) is called constant speed *geodesic*. If $\pi \in \Pi_{opt}(\mu_1, \mu_2)$, then the curve

$$\mu_t^{1\to 2} := g(t, \cdot)_{\#}\pi, \quad t \in [0,1], \tag{2.6}$$

with $g : [0,1] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $(t, x_1, x_2) \mapsto (1-t)x_1 + tx_2$ is a constant speed geodesic connecting μ_1 and μ_2. Conversely, every constant speed geodesic connecting μ_1 and μ_2 has a representation (2.6) for a suitable $\pi \in \Pi_{opt}(\mu_1, \mu_2)$, see [3, Theorem 7.2.2]. In particular, if μ_1 is absolutely continuous with respect to the d-dimensional Lebesgue measure, then, by Remark 2.1, there exists exactly one such constant speed geodesic.

We will need a more general definition of geodesics in order to make the Wasserstein proximal mappings in the next section well-defined. For $\mu_0, \mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$, let $\Pi(\mu_0, \mu_1, \mu_2)$ denote the set of measures $\boldsymbol{\pi} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ with marginals μ_i, $i = 0, 1, 2$, and let $\Pi^{j,k} : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$, $(x_0, x_1, x_2) \mapsto (x_j, x_k)$ for $j, k = 0, 1, 2$. A *generalized geodesic* connecting μ_1 and μ_2 with base μ_0 is any curve of type

$$\mu_{0,t}^{1\to 2} := h(t, \cdot)_{\#}\boldsymbol{\pi},$$

with $h(t, \cdot) : [0,1] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $(t, x_0, x_1, x_2) \mapsto (1-t)x_1 + tx_2$, where

$$\boldsymbol{\pi} \in \Pi(\mu_0, \mu_1, \mu_2), \qquad \Pi^{0,1}_{\#}\boldsymbol{\pi} = \Pi_{opt}(\mu_0, \mu_1), \quad \Pi^{0,2}_{\#}\boldsymbol{\pi} = \Pi_{opt}(\mu_0, \mu_2). \tag{2.7}$$

Choosing the base $\mu_0 = \mu_1$, we have again the definition of a geodesic. Moreover, for an absolutely continuous base measure μ_0, the generalized geodesic connecting μ_1 and μ_2 is uniquely determined and the plan in (2.7) is given via the optimal transport maps by $\boldsymbol{\pi} = (\mathrm{Id}, T_{\mu_0}^{\mu_1}, T_{\mu_0}^{\mu_2})_{\#}\mu_0$, e.g. see [3, Remark 9.2.3].

We consider functions $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ with effective domain

$$D(\mathscr{F}) := \{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \mathscr{F}(\mu) < \infty\}$$

and call a function *proper* if $D(\mathscr{F}) \neq \emptyset$. A function $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ is said to be *convex along generalized geodesics*, if for every $\mu_0, \mu_1, \mu_2 \in D(\mathscr{F})$, there exists a generalized geodesic $\mu_{0,t}^{1\to 2}$ with base μ_0 such that

$$\mathscr{F}(\mu_{0,t}^{1\to 2}) \le (1-t)\mathscr{F}(\mu_1) + t\mathscr{F}(\mu_2) \quad \text{for all } t \in [0,1]. \tag{2.8}$$

Typical examples of functions defined on $\mathcal{P}_2(\mathbb{R}^d)$ that are convex along generalized geodesics are the potential and interaction energy and the relative entropy discussed, e.g., in [3, §9.3, §9.4].

3. Proximal mappings

In this section, we consider proximal mappings in Wasserstein spaces, which play an important role in Wasserstein gradient flow methods. Let $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ be proper, lower semicontinuous (lsc), coercive (in the sense of [3, (2.4.10)]) and convex

along generalized geodesics. Then the proximal mapping $\mathscr{J}_\tau : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ given by

$$\mathscr{J}_\tau(\mu) := \operatorname{argmin}_{\nu\in\mathcal{P}_2(\mathbb{R}^d)} \Big\{\mathscr{F}(\nu) + \frac{1}{2\,\tau} W_2^2(\nu,\mu)\Big\}, \quad \tau > 0 \tag{3.1}$$

is well-defined, i.e., for every $\mu \in \overline{D(\mathscr{F})}$, the minimizer in (3.1) exits and is unique, see [3, Theorem 4.1.2, Lemma 9.2.7]. Moreover, by [3, Theorem 4.1.2] (with $\lambda = 0$), for all $\mu \in \overline{D(\mathscr{F})}$ and all $\nu \in D(\mathscr{F})$, the following inequality is satisfied

$$\frac{1}{2\,\tau} W_2^2(\mathscr{J}_\tau(\mu),\nu) - \frac{1}{2\,\tau} W_2^2(\mu,\nu) \le \mathscr{F}(\nu) - \mathscr{F}(\mathscr{J}_\tau(\mu)) - \frac{1}{2\,\tau} W_2^2(\mathscr{J}_\tau(\mu),\mu). \tag{3.2}$$

Replacing ν with $\mathscr{J}_\tau(\nu)$ and changing the roles of μ with ν in (3.2), we obtain for all $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^d)$ that

$$W_2^2(\mathscr{J}_\tau(\mu), \mathscr{J}_\tau(\nu)) \le \frac{1}{2}\Big(W_2^2(\mu, \mathscr{J}_\tau(\nu)) + W_2^2(\mathscr{J}_\tau(\mu),\nu) \tag{3.3}$$

$$- W_2^2(\mathscr{J}_\tau(\mu),\mu) - W_2^2(\mathscr{J}_\tau(\nu),\nu)\Big). \tag{3.4}$$

As in Hilbert spaces, minimizers of $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty,+\infty]$ and fixed points of its proximal mapping are closely related.

Proposition 3.1. *Let $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty,+\infty]$ be proper, lsc, coercive and convex along generalized geodesics. Then it holds*

$$\operatorname{argmin}_{\mu\in\mathcal{P}(\mathbb{R}^d)} \mathscr{F}(\mu) = \operatorname{Fix} \mathscr{J}_\tau.$$

Proof. Let $\hat\mu$ be a minimizer of $\mathscr{F}$. Then, for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\mathscr{F}(\hat\mu) + \frac{1}{2\tau}\, W_2^2(\hat\mu,\hat\mu) = \mathscr{F}(\hat\mu) \le \mathscr{F}(\mu) \le \mathscr{F}(\mu) + \frac{1}{2\tau}\, W_2^2(\mu,\hat\mu)$$

implying $\hat\mu = \mathscr{J}_\tau(\hat\mu)$. The converse follows immediately from inequality (3.2) with μ replaced by $\hat\mu$ and using $\hat\mu = \mathscr{J}_\tau(\hat\mu)$,

$$\begin{aligned} 0 &= \frac{1}{2\,\tau} W_2^2(\mathscr{J}_\tau(\hat\mu),\nu) - \frac{1}{2\,\tau} W_2^2(\hat\mu,\nu) \\ &\le \mathscr{F}(\nu) - \mathscr{F}(\mathscr{J}_{\hat\tau}(\mu)) - \frac{1}{2\,\tau} W_2^2(\mathscr{J}_\tau(\hat\mu),\hat\mu) = \mathscr{F}(\nu) - \mathscr{F}(\hat\mu), \end{aligned}$$

i.e. $\mathscr{F}(\hat\mu) \le \mathscr{F}(\nu)$ for all $\nu \in D(\mathscr{F})$. □

4. Quasi α-firmly nonexpansive mappings

Recall that for $\alpha \in (0,1)$, an operator $T : \mathbb{R}^d \to \mathbb{R}^d$ is *α-firmly nonexpansive*, if we have

$$\|Tx - Ty\|^2 \le \|x-y\|^2 - \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 \quad \text{for all } x,y \in \mathbb{R}^d. \tag{4.1}$$

If the fixed point set $\operatorname{Fix} T := \{x \in \mathbb{R}^d : T(x) = x\}$ is nonempty and (4.1) is restricted to $y \in \operatorname{Fix} F$, i.e.,

$$\|Tx - y\|^2 \le \|x-y\|^2 - \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-T)x\|^2 \quad \text{for all } x \in \mathbb{R}^d,\ y \in \operatorname{Fix} T, \tag{4.2}$$

then T is called *quasi α-firmly nonexpansive*. We do not know how to translate the definition of α-firmly nonexpansive operators on $\mathbb{R}^d$ to Wasserstein spaces. However, for quasi α-firmly nonexpansive operators this is possible. First, we say as usual that a mapping $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is *nonexpansive*, if

$$W_2(\mathscr{T}(\mu), \mathscr{T}(\nu)) \leq W_2(\mu, \nu) \quad \text{for all } \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.3}$$

If the fixed point set $\operatorname{Fix} \mathscr{T} := \{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \mathscr{T}(\mu) = \mu\}$ is nonempty and (4.3) holds for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and all $\nu \in \operatorname{Fix} \mathscr{T}$, then $\mathscr{T}$ is said to be a *quasi nonexpansive mapping*. Finally, for $\alpha \in (0,1)$, a mapping $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is *quasi α-firmly nonexpansive*, if $\operatorname{Fix} \mathscr{T} \neq \emptyset$ and for all $\mu \in \mathcal{P}(\mathbb{R}^d)$, $\nu \in \operatorname{Fix} \mathscr{T}$, the following inequality holds true:

$$W_2^2(\mathscr{T}(\mu), \nu) \leq W_2^2(\mu, \nu) - \frac{1-\alpha}{\alpha} W_2^2(\mu, \mathscr{T}(\mu)). \tag{4.4}$$

By the next proposition, proximal mappings are quasi 1/2-firmly nonexpansive.

Proposition 4.1. *Let $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ be proper, lsc, coercive and convex along generalized geodesics. Then, for every $\tau > 0$, the proximal mapping $\mathscr{J}_\tau : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is quasi α-firmly nonexpansive on $\overline{D(\mathscr{F})}$ with $\alpha = 1/2$.*

Proof. The claim follows immediately from (3.3) using $\mathscr{J}_\tau(\nu) = \nu$. □

Next, we are interested in the behavior of push-forward operators of (quasi) α-firmly nonexpansive operators $T : \mathbb{R}^d \to \mathbb{R}^d$. In other words, we consider $\mathscr{T}_T : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ defined by

$$\mathscr{T}_T(\mu) := T_\# \mu = \mu \circ T^{-1}. \tag{4.5}$$

Proposition 4.2. *Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be an α-firmly nonexpansive operator for some $\alpha \in (0,1)$. Then $\mathscr{T}_T : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ in (4.5) with $\widetilde{T} := \operatorname{Id} - T$ satisfies*

$$W_2^2(\mathscr{T}_T(\mu), \mathscr{T}_T(\nu)) \leq W_2^2(\mu, \nu) - \frac{1-\alpha}{\alpha} W_2^2(\mathscr{T}_{\widetilde{T}}(\mu), \mathscr{T}_{\widetilde{T}}(\nu)) \quad \text{for all } \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.6}$$

In particular $\mathscr{T}_T$ is nonexpansive.

Proof. Let $\pi \in \Pi_{opt}(\mu, \nu)$. Then

$$\begin{aligned} W_2^2(\mathscr{T}_T(\mu), \mathscr{T}_T(\nu)) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 \, d\left((T,T)_\# \pi\right)(x,y) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \|Tx - Ty\|^2 \, d\pi(x,y). \end{aligned}$$

By assumption $T : \mathbb{R}^d \to \mathbb{R}^d$ is α-firmly nonexpansive for some $\alpha \in (0,1)$, so that

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\mathbb{R}^d} \|Tx - Ty\|^2 \, d\pi(x,y) \\
&\qquad \le \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2 \, d\pi(x,y) - \frac{1-\alpha}{\alpha} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 \, d\pi(x,y) \\
&\qquad = W_2^2(\mu,\nu) - \frac{1-\alpha}{\alpha} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2 \, d\left((\widetilde{T},\widetilde{T})_{\#}\pi\right)(x,y) \\
&\qquad \le W_2^2(\mu,\nu) - \frac{1-\alpha}{\alpha} W_2^2(\mathscr{T}_{\widetilde{T}}(\mu), \mathscr{T}_{\widetilde{T}}(\nu)).
\end{aligned}$$

This proves (4.6) and in particular that $\mathscr{T}_T$ is a nonexpansive mapping. □

The next result describes the relationships between $\operatorname{Fix} T$ and $\operatorname{Fix} \mathscr{T}_T$.

Proposition 4.3. *Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be a quasi α-firmly nonexpansive operator. Then $\operatorname{Fix} \mathscr{T}_T \neq \emptyset$ and in particular $\nu \in \operatorname{Fix} \mathscr{T}_T$ if $\operatorname{supp}(\nu) \subseteq \operatorname{Fix} T$.*

Proof. Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be quasi α-firmly nonexpansive mapping. By [11, Lemma 4.1] it follows that $\operatorname{Fix} T$ is nonempty, closed and convex. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ be arbitrary. Let $f(x) := P_{\operatorname{Fix} T}(x)$ for $x \in \mathbb{R}^d$, where $P_{\operatorname{Fix} T}$ denotes the metric projection onto $\operatorname{Fix} T$. Then we have for $\nu := f_{\#}\mu$ that $\operatorname{supp}(\nu) \subseteq \operatorname{Fix} T$. If a measure ν fulfills the later condition, then we obtain for every $B \in \mathcal{B}(\mathbb{R}^d)$ that

$$\mathscr{T}_T(\nu)(B) = (T_{\#}\nu)(B) = \nu(T^{-1}(B)) = \nu(T^{-1}(B) \cap \operatorname{Fix} T) = \nu(B \cap \operatorname{Fix} T) = \nu(B),$$

i.e. $\nu \in \operatorname{Fix} \mathscr{T}_T$. It remains to show that $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. By definition of f, we have for any $B \in \mathcal{B}(\mathbb{R}^d)$ that $\nu(B) = \mu(f^{-1}(B)) = \mu(f^{-1}(B \cap \operatorname{Fix} T))$. It is evident that $\nu(B) \ge 0$ and that $\nu(\mathbb{R}^d) = \mu(f^{-1}(\operatorname{Fix} T)) = \mu(\mathbb{R}^d) = 1$. Moreover, if $(B_n)_{n\in\mathbb{N}}$ is a countable family of disjoint Borel sets, then so is the family $(B_n \cap \operatorname{Fix} T)_{n\in\mathbb{N}}$ and consequently

$$\begin{aligned}
\nu\Big(\bigcup_{n\in\mathbb{N}} B_n\Big) &= \mu\Big(f^{-1}\Big(\bigcup_{n\in\mathbb{N}} B_n \cap \operatorname{Fix} T\Big)\Big) \\
&= \mu\Big(\bigcup_{n\in\mathbb{N}} f^{-1}(B_n \cap \operatorname{Fix} T)\Big) = \sum_{n\in\mathbb{N}} \mu(f^{-1}(B_n \cap \operatorname{Fix} T)) = \sum_{n\in\mathbb{N}} \nu(B_n).
\end{aligned}$$

Therefore ν is indeed a probability measure. Next, consider

$$\int_{\mathbb{R}^d} \|x\|^2 \, d\nu(x) = \int_{\operatorname{Fix} T} \|x\|^2 \, d\mu(x) + \int_{\mathbb{R}^d \setminus \operatorname{Fix} T} \|P_{\operatorname{Fix} T} x\|^2 \, d\mu(x).$$

The second integral can be estimated with an arbitrary fixed $x_0 \in \operatorname{Fix} T$ as follows:

$$\begin{aligned}
\int_{\mathbb{R}^d\setminus\operatorname{Fix} T} \|P_{\operatorname{Fix} T} x\|^2 \, d\mu(x) &\le \int_{\mathbb{R}^d\setminus\operatorname{Fix} T} (\|P_{\operatorname{Fix} T} x - x_0\| + \|x_0\|)^2 \, d\mu(x) \\
&\le 2\Big(\int_{\mathbb{R}^d\setminus\operatorname{Fix} T} \|P_{\operatorname{Fix} T} x - x_0\|^2 \, d\mu(x) + \int_{\mathbb{R}^d\setminus\operatorname{Fix} T} \|x_0\|^2 \, d\mu(x)\Big) \\
&\le 2\Big(\int_{\mathbb{R}^d\setminus\operatorname{Fix} T} \|x - x_0\|^2 \, d\mu(x) + \int_{\mathbb{R}^d\setminus\operatorname{Fix} T} \|x_0\|^2 \, d\mu(x)\Big).
\end{aligned}$$

Since $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ this completes the proof. □

In particular we have showed the following relation.

Corollary 4.4. *If $T : \mathbb{R}^d \to \mathbb{R}^d$ is a quasi α-firmly nonexpansive mapping, then it holds $(P_{\operatorname{Fix} T})_{\#}\mathcal{P}_2(\mathbb{R}^d) \subseteq \operatorname{Fix} \mathscr{T}_T$.*

Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\pi \in \Pi(\mu, \nu)$. Let $\{\nu_x\}_{x\in\mathbb{R}^d}$ be the family of disintegrations of π with respect to μ , i.e.

$$\pi(A \times B) = \int_A \Big(\int_B \nu_x(y)\, dy \Big)\, d\mu(x) \quad \text{for all } A, B \in \mathcal{B}(\mathbb{R}^d), \tag{4.7}$$

Proposition 4.5. *Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be a quasi α-firmly nonexpansive operator. Suppose that for every $\nu \in \operatorname{Fix} \mathscr{T}_T$ there is a family of disintegrating measures $\{\nu_x\}_{x\in\mathbb{R}^d}$ satisfying $\nu_x(\operatorname{Fix} T) \geq C_{\operatorname{Fix} T}$ for some positive uniform constant $C_{\operatorname{Fix} T}$ possibly depending on $\operatorname{Fix} T$. Then $\mathscr{T}_T : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is a quasi β-firmly nonexpansive operator for some β depending on α.*

Proof. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\nu \in \operatorname{Fix} \mathscr{T}_T$ satisfying $\nu_x(\operatorname{Fix} T) \geq C_{\operatorname{Fix} T}$ uniformly for some positive constant $C_{\operatorname{Fix} T}$. Let $\pi \in \Pi_{opt}(\mu, \nu)$. From Proposition 4.3 we have that $\operatorname{supp}(\nu) \subseteq \operatorname{Fix} T$, therefore

$$\begin{aligned} W_2^2(\mathscr{T}_T(\mu), \mathscr{T}_T(\nu)) &\leq \int_{\mathbb{R}^d \times \operatorname{Fix} T} \|Tx - Ty\|^2 \, d\pi(x, y) \\ &\leq \int_{\mathbb{R}^d \times \operatorname{Fix} T} \|x - y\|^2 \, d\pi(x, y) \quad - \frac{1-\alpha}{\alpha} \int_{\mathbb{R}^d \times \operatorname{Fix} T} \|x - Tx\|^2 \, d\pi(x, y). \end{aligned}$$

From the assumption on the disintegration we estimate the second integral from below as

$$\begin{aligned} \int_{\mathbb{R}^d \times \operatorname{Fix} T} \|x - Tx\|^2 \, d\pi(x, y) &= \int_{\mathbb{R}^d} \|x - Tx\|^2 \Big(\int_{\operatorname{Fix} T} \nu_x(y)\, dy \Big)\, d\mu(x) \\ &= \int_{\mathbb{R}^d} \|x - Tx\|^2 \nu_x(\operatorname{Fix} T)\, d\mu(x) \geq C_{\operatorname{Fix} T} \int_{\mathbb{R}^d} \|x - Tx\|^2 \, d\mu(x). \end{aligned}$$

Together with $\nu \in \operatorname{Fix} \mathscr{T}_T$, this implies

$$\begin{aligned} W_2^2(\mathscr{T}_T(\mu), \mathscr{T}_T(\nu)) = W_2^2(\mathscr{T}_T(\mu), \nu) &\leq W_2^2(\mu, \nu) - \frac{1-\alpha}{\alpha} C_{\operatorname{Fix} T} \int_{\mathbb{R}^d} \|x - Tx\|^2 \, d\mu(x) \\ &\leq W_2^2(\mu, \nu) - \frac{1-\alpha}{\alpha} C_{\operatorname{Fix} T}\, W_2^2(\mu, \mathscr{T}_T(\mu)). \end{aligned}$$

Taking $\beta = (1 + C_{\operatorname{Fix} T}(1-\alpha)/\alpha)^{-1}$, this completes the proof. □

In the particular case when $\operatorname{Fix} T$ consists of a unique element $x_0 \in \mathbb{R}^d$, the previous result holds without any disintegration condition as shown below.

Corollary 4.6. *Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be a quasi α-firmly nonexpansive operator such that $\operatorname{Fix} T = \{x_0\}$ for some $x_0 \in \mathbb{R}^d$. Then $\nu_0 := \delta_{x_0} \in \mathcal{P}_2(\mathbb{R}^d)$ is a fixed point of the push-forward operator $\mathscr{T}_T$ and fulfills*

$$W_2^2(\mathscr{T}_T(\mu), \mathscr{T}_T(\nu_0)) \leq W_2^2(\mu, \nu_0) - \frac{1-\alpha}{\alpha} W_2^2(\mu, \mathscr{T}_T(\mu)).$$

Proof. Let $\operatorname{Fix} T = \{x_0\}$ for some $x_0 \in \mathbb{R}^d$. Then, by the proof of Proposition 4.3, there is $\nu_0 \in \operatorname{Fix} \mathscr{T}_T$ with $\operatorname{supp}(\nu_0) \subseteq \{x_0\}$, i.e. $\nu_0 = \delta_{x_0}$. For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the only transport plan between μ and ν_0 is $\pi = \mu \otimes \delta_{x_0}$. By similar calculations as in Proposition 4.5 we obtain

$$\begin{aligned} W_2^2(\mathscr{T}_T(\mu), \mathscr{T}_T(\nu_0)) &\leq \int_{\mathbb{R}^d \times \{x_0\}} \|Tx - Tx_0\|^2 \, d\pi(x, x_0) = \int_{\mathbb{R}^d} \|Tx - x_0\|^2 \, d\mu(x) \\ &\leq \int_{\mathbb{R}^d} \|x - x_0\|^2 \, d\mu(x) - \frac{1-\alpha}{\alpha} \int_{\mathbb{R}^d} \|Tx - x\|^2 \, d\mu(x) \\ &\leq W_2^2(\mu, \nu_0) - \frac{1-\alpha}{\alpha} W_2^2(\mu, \mathscr{T}_T(\mu)). \end{aligned}$$

□

We conclude this section by showing that quasi α-firmly nonexpansiveness is well behaved under the composition of mappings that share at least a common fixed point. The proofs are modifications of arguments from [9] for our setting.

Lemma 4.7. *Let $\mathscr{S}, \mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ satisfy $\operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S} \neq \emptyset$. If $\mathscr{S}$ is quasi α-firmly nonexpansive and $\mathscr{T}$ is quasi nonexpansive, then $\operatorname{Fix}(\mathscr{T} \circ \mathscr{S}) = \operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S}$.*

Proof. The inclusion $\operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S} \subseteq \operatorname{Fix}(\mathscr{T} \circ \mathscr{S})$ is obvious. Now let $\mu \in \operatorname{Fix}(\mathscr{T} \circ \mathscr{S})$. First, suppose that $\mathscr{S}(\mu) \in \operatorname{Fix} \mathscr{T}$. Then $\mathscr{S}(\mu) = \mathscr{T}(\mathscr{S}(\mu)) = \mu$ implies $\mu \in \operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S}$. Next, let $\mathscr{S}(\mu) \notin \operatorname{Fix} \mathscr{T}$. Then we distinguish two subcases $\mu \in \operatorname{Fix} \mathscr{S}$ and $\mu \notin \operatorname{Fix} \mathscr{S}$. If $\mu \in \operatorname{Fix} \mathscr{S}$, then $\mu = \mathscr{T}(\mathscr{S}(\mu)) = \mathscr{T}(\mu)$ implies $\mu \in \operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S}$. Finally, let $\mu \notin \operatorname{Fix} \mathscr{S}$ and take $\nu \in \operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S}$. This yields

$$W_2^2(\mu, \nu) = W_2^2(\mathscr{T}(\mathscr{S}(\mu)), \mathscr{T}(\nu)) \leq W_2^2(\mathscr{S}(\mu), \nu) \leq W_2^2(\mu, \nu) - \frac{1-\alpha}{\alpha} W_2^2(\mu, \mathscr{S}(\mu)),$$

which implies $\mathscr{S}(\mu) = \mu$, a contradiction. □

Proposition 4.8. *Let $\mathscr{S} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be quasi α-firmly nonexpansive and let $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be quasi β-firmly nonexpansive. Suppose that $\operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S} \neq \emptyset$. Then $\mathscr{T} \circ \mathscr{S}$ is quasi γ-firmly nonexpansive with*

$$\gamma := \frac{\alpha + \beta - 2\alpha\beta}{1 - \alpha\beta}. \tag{4.8}$$

Proof. By Lemma 4.7, assumption $\operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S} \neq \emptyset$ implies that $\operatorname{Fix}(\mathscr{T} \circ \mathscr{S}) = \operatorname{Fix} \mathscr{T} \cap \operatorname{Fix} \mathscr{S}$. Let $\nu \in \operatorname{Fix}(\mathscr{T} \circ \mathscr{S})$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then an application of (4.4) yields

$$\begin{aligned} W_2^2(\mathscr{T}(\mathscr{S}(\mu)), \nu) &\leq W_2^2(\mathscr{S}(\mu), \nu) - \frac{1-\beta}{\beta} W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu))) \\ &\leq W_2^2(\mu, \nu) - \frac{1-\alpha}{\alpha} W_2^2(\mu, \mathscr{S}(\mu)) - \frac{1-\beta}{\beta} W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)). \end{aligned}$$

It suffices to show that

$$\frac{1-\gamma}{\gamma} W_2^2(\mu, \mathscr{T}(\mathscr{S}(\mu)) \leq \frac{1-\alpha}{\alpha} W_2^2(\mu, \mathscr{S}(\mu)) + \frac{1-\beta}{\beta} W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)). \tag{4.9}$$

With $\tau := (1-\alpha)/\alpha + (1-\beta)/\beta$ inequality (4.9) is equivalent to

$$\Big(\frac{1-\alpha}{\tau\alpha}\Big)^2 W_2^2(\mu, \mathscr{S}(\mu)) + \Big(\frac{1-\beta}{\tau\beta}\Big)^2 W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu))$$
$$+ \Big(\frac{1-\alpha}{\tau\alpha}\Big)\Big(\frac{1-\beta}{\tau\beta}\Big)\Big(W_2^2(\mu, \mathscr{S}(\mu)) + W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)) - W_2^2(\mu, \mathscr{T}(\mathscr{S}(\mu))\Big) \geq 0.$$

Setting $\kappa := \dfrac{1-\alpha}{\alpha} / \dfrac{1-\beta}{\beta}$, this is equivalent to

$$(\kappa+1) W_2^2(\mu, \mathscr{S}(\mu)) + \frac{\kappa+1}{\kappa} W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)) - W_2^2(\mu, \mathscr{T}(\mathscr{S}(\mu)) \geq 0.$$

The elementary inequality

$$\kappa W_2^2(\mu, \mathscr{S}(\mu)) + \frac{1}{\kappa} W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)) \geq 2\, W_2(\mu, \mathscr{S}(\mu)) W_2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu))$$

for all $\kappa > 0$ together with the triangle inequality

$$W_2(\mu, \mathscr{S}(\mu)) + W_2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)) \geq W_2(\mu, \mathscr{T}(\mathscr{S}(\mu))$$

implies

$$(\kappa+1) W_2^2(\mu, \mathscr{S}(\mu)) + \frac{\kappa+1}{\kappa} W_2^2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu))$$
$$\geq (W_2(\mu, \mathscr{S}(\mu)) + W_2(\mathscr{S}(\mu), \mathscr{T}(\mathscr{S}(\mu)))^2 \geq W_2^2(\mu, \mathscr{T}(\mathscr{S}(\mu)).$$

□

The following corollary is an immediate consequence of the above proposition.

Corollary 4.9. *Let $(\mathscr{T}_i)_{i=1}^N$ be a finite family of quasi α_i-firmly nonexpansive mappings. Suppose that their fixed point sets have a nonempty intersection. Then $\mathscr{T} = \mathscr{T}_{i_n} \circ \mathscr{T}_{i_{n-1}} \circ ... \circ \mathscr{T}_{i_1}$, where $i_j \in \{1, 2, ..., n\}$ are distinct, is also quasi α-firmly nonexpansive for some $\alpha \in (0,1)$ dependent on α_i.*

5. Fixed point theory of quasi α-firmly nonexpansive mappings

5.1. **Opial's property and Fejér monotonicity.** A mapping $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is *asymptotic regular at* $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, if

$$\lim_{n \to +\infty} W_2(\mathscr{T}^{n+1}(\mu), \mathscr{T}^n(\mu)) = 0. \tag{5.1}$$

Here $\mathscr{T}^n := \underbrace{\mathscr{T} \circ \cdots \circ \mathscr{T}}_{n-\text{times}}$. If the limit in (5.1) holds for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, then $\mathscr{T}$ is said to be *asymptotic regular* on $\mathcal{P}_2(\mathbb{R}^d)$. An immediate consequence of quasi α-firmly nonexpansiveness is the following evident lemma.

Lemma 5.1. *Let $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be a quasi α-firmly nonexpansive mapping. Then $\mathscr{T}$ is asymptotic regular on $\mathcal{P}_2(\mathbb{R}^d)$.*

Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_2(\mathbb{R}^d)$. An element $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ is a *narrow cluster point* of $(\mu_n)_{n\in\mathbb{N}}$ if and only if there exists a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ such that $\mu_{n_k} \overset{\mathcal{N}}{\to} \mu$. Recently, it has been shown [28, Theorem 5.1] that if $\mu_n \overset{\mathcal{N}}{\to} \mu$, then the following inequality holds true

$$\liminf_{n\to+\infty} W_2(\mu_n, \mu) < \liminf_{n\to+\infty} W_2(\mu_n, \nu), \quad \text{for all } \nu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \{\mu\}. \tag{5.2}$$

This is known as the *Opial's property*. It implies, for all $\nu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \{\mu\}$, that

$$\limsup_{n\to+\infty} W_2(\mu_n, \mu) = \lim_{k\to+\infty} W_2(\mu_{n_k}, \mu) < \liminf_{k\to+\infty} W_2(\mu_{n_k}, \nu) \leq \limsup_{n\to+\infty} W_2(\mu_n, \nu). \tag{5.3}$$

A sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}_2(\mathbb{R}^d)$ is *Fejér monotone with respect to a set* $S \subseteq \mathcal{P}_2(\mathbb{R}^d)$ if $W_2(\mu_{n+1}, \nu) \leq W_2(\mu_n, \nu)$ for all $\nu \in S$ and for all $n \in \mathbb{N}$.

Lemma 5.2. *Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}_2(\mathbb{R}^d)$ be Fejér monotone with respect to a set $S \subseteq \mathcal{P}_2(\mathbb{R}^d)$. If all narrow cluster points of $(\mu_n)_{n\in\mathbb{N}}$ belong to S, then $\mu_n \overset{\mathcal{N}}{\to} \mu$ for some $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.*

Proof. The proof follows similar lines as in [11, Lemma 5.3]. □

5.2. **Opial's Theorem.** A well-known result of Opial [29] for uniformly convex Banach spaces X satisfying Opial's property states that the iterations $x_{n+1} = Tx_n$ of a nonexpansive and asymptotic regular operator $T : X \to X$ with $\operatorname{Fix} T \neq \emptyset$ always converge weakly to an element in $\operatorname{Fix} T$. Recently, in [28, Theorem 6.9], it was shown that such a result holds true as well in the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$ for mappings that are nonexpansive, asymptotic regular and have a nonempty fixed point set. This is because the space $\mathcal{P}_2(\mathbb{R}^d)$ satisfies Opial's property with respect to the narrow convergence.

Theorem 5.3. [28, Theorem 6.9] *Let $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be a nonexpansive mapping that is asymptotic regular on $\mathcal{P}_2(\mathbb{R}^d)$. Then $\operatorname{Fix} \mathscr{T} \neq \emptyset$ if and only if for some $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ (hence any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$) the iterates $\mu_{n+1} = \mathscr{T}(\mu_n)$ are bounded in $\mathcal{P}_2(\mathbb{R}^d)$, in which case they narrowly converge to some $\mu \in \operatorname{Fix} \mathscr{T}$.*

As a consequence of this theorem we have the following corollary.

Corollary 5.4. *Let $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be a nonexpansive, quasi α-firmly nonexpansive mapping. Then, for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, the iterates $\mu_{n+1} = \mathscr{T}(\mu_n)$ converge narrowly to some element $\mu \in \operatorname{Fix} \mathscr{T}$.*

Proof. By Lemma 5.1 the operator $\mathscr{T}$ is asymptotic regular, whenever it is quasi α-firmly nonexpansive. Moreover, by Definition (4.4), the fixed point set $\operatorname{Fix} \mathscr{T}$ is nonempty. Hence the result follows directly from Theorem 5.3. □

A function $\phi : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ is a *characteristic function* of a mapping $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$, if

$$\operatorname{argmin}\{\phi(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\} = \operatorname{Fix} \mathscr{T}$$

whenever the latter is nonempty. Let $\mathscr{C}_{\mathscr{T}}$ denote the set of all characteristic functions associated to the mapping $\mathscr{T}$. Note that $\mathscr{C}_{\mathscr{T}} \neq \emptyset$ since $\phi(\mu) = W_2(\mu, \mathscr{T}(\mu))$ is a

characteristic function for any mapping $\mathscr{T}$ satisfying $\operatorname{Fix}\mathscr{T} \neq \emptyset$. A mapping $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is said to satisfy the *quadratic growth condition*, if there exist a constant $C > 0$ and a proper, narrow lower semicontinuous function $\phi \in \mathscr{C}_{\mathscr{T}}$ satisfying

$$W_2^2(\mathscr{T}(\mu), \nu) - W_2^2(\mu, \nu) \leq C\,(\phi(\nu) - \phi(\mathscr{T}(\mu))) \quad \text{for all } \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.4}$$

Theorem 5.5. *Let $\mathscr{T} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be a quasi α-firmly nonexpansive mapping satisfying the quadratic growth condition* (5.4)*. Then, for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, the iterates $\mu_{n+1} = \mathscr{T}(\mu_n)$ converge narrowly to some element $\mu \in \operatorname{Fix}\mathscr{T}$.*

Proof. Since $\mathscr{T}$ is quasi α-firmly nonexpansive, then, by similar arguments as in Proposition 5.4, it follows that, for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, the sequence $(\mu_n)_{n\in\mathbb{N}}$ defined by $\mu_{n+1} = \mathscr{T}(\mu_n)$ is bounded and therefore it contains a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ narrowly converging to a certain element $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. The assumption that $\mathscr{T}$ satisfies the quadratic growth condition implies that there is a constant $C > 0$ and a proper narrow lsc function $\phi \in \mathscr{C}_{\mathscr{T}}$ such that inequality (5.4) is satisfied. In particular, it follows that $\phi(\mathscr{T}(\mu)) \leq \phi(\mu)$ for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Consequently, we obtain $\phi(\mu_{n+1}) \leq \phi(\mu_n)$ for every $n \in \mathbb{N}$. Again, from condition (5.4) for every $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ it follows

$$\sum_{n=0}^{N}(W_2^2(\mathscr{T}(\mu_n), \nu) - W_2^2(\mu_n, \nu)) \leq C \sum_{n=0}^{N}(\phi(\nu) - \phi(\mathscr{T}(\mu_n))), \tag{5.5}$$

$$W_2^2(\mathscr{T}(\mu_N), \nu) - W_2^2(\mu_0, \nu) \leq C\,(N+1)\,\phi(\nu) - C\sum_{n=0}^{N}\phi(\mathscr{T}(\mu_n)). \tag{5.6}$$

Rearranging the terms and using the monotonicity of $(\phi(\mu_n))_{n\in\mathbb{N}}$ yields

$$\phi(\mu_{N+1}) + \frac{1}{C\,(N+1)}(W_2^2(\mu_{N+1}, \nu) - W_2^2(\mu_0, \nu)) \leq \phi(\nu) \quad \text{for all } \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

Consequently, by narrow lower semicontinuity of ϕ, we get

$$\phi(\nu) \geq \limsup_{N\to+\infty} \phi(\mu_{N+1}) \geq \liminf_{k\to+\infty} \phi(\mu_{n_k}) \geq \phi(\mu) \quad \text{for all } \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

Therefore, $\mu \in \operatorname{argmin}\{\phi(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$.

Since ϕ is a characteristic function for the mapping $\mathscr{T}$ we have $\mu \in \operatorname{Fix}\mathscr{T}$. If $\widetilde{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$ is another narrow cluster point of $(\mu_n)_{n\in\mathbb{N}}$, then by the same arguments we conclude that $\tilde{\mu} \in \operatorname{Fix}\mathscr{T}$. By quasi α-firmly nonexpansiveness of the mapping $\mathscr{T}$, it holds that $(\mu_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix}\mathscr{T}$. By Lemma 5.2 it follows that the whole sequence $(\mu_n)_{n\in\mathbb{N}}$ narrowly converges to $\mu \in \operatorname{Fix}\mathscr{T}$. □

As a corollary, we obtain a result on the convergence of the so-called *proximal point algorithm* which is already known from the literature, see, e.g., [28, Theorem 6.7] .

Corollary 5.6. *Let $\mathscr{F} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty + \infty]$ be proper, lsc, coercive and convex along generalized geodesics. For $\tau > 0$, let $\mathscr{J}_\tau : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be the proximal mapping defined in* (3.1)*. Then, for any $\mu_0 \in \overline{D(\mathscr{F})}$, the iterates $\mu_{n+1} = \mathscr{J}_\tau(\mu_n)$ converge narrowly to some $\mu \in \operatorname{argmin}\{\mathscr{F}(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$.*

Proof. By Proposition 4.1, the proximal mapping $\mathscr{J}_\tau$ is quasi α-firmly nonexpansive with $\alpha = 1/2$. Moreover, from inequality (3.2) we know that $\mathscr{J}_\tau$ satisfies the quadratic growth condition with $C = 2\tau$ and $\phi = \mathscr{F}$. Then the result follows from Theorem 5.5. □

6. Cyclic proximal point algorithm

Let $\mathscr{F}_i : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ be proper, lsc, coercive and convex along generalized geodesics for $i = 1, 2, \cdots, N$. Consider the minimization problem

$$\inf_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \sum_{i=1}^{N} \mathscr{F}_i(\mu). \tag{6.1}$$

The function $\mathscr{F} = \sum_{i=1}^N \mathscr{F}_i$ is itself proper, lsc, coercive and convex along generalized geodesics, since it is the sum of finitely many such functions. Suppose further that $D(\mathscr{F}) \subseteq D(\mathscr{F}_i)$. A popular method to solve a problem of this kind is the proximal point algorithm from the previous section. However, computing the proximal mapping $\mathscr{J}_\tau$ might be complicated because it could happen that the function $\mathscr{F}$ is difficult to handle, both theoretically and computationally. One way out consists in considering the functions $\mathscr{F}_i$ separately, that is one computes the proximal mapping $\mathscr{J}_{\tau_i}$ for each function, where $\tau_i > 0$ is the corresponding step size for $i = 1, 2, \cdots, N$. Then we consider the iterates

$$\mu_{n+1} = \mathscr{J}_{\tau_{[n]}}(\mu_n), \quad \text{where } [n] = n \,(\mathrm{mod}\, N) + 1 \in \{1, 2, \cdots, N\}. \tag{6.2}$$

Such a method is known as the *cyclic proximal point method*. For two operators, it is also called backward-backward splitting method. Splitting methods in convex analysis date back to papers of Lions, Mercier [25, Lions–Mercier, 1979] who studied splitting algorithms for stationary and evolution problems involving the sum of two monotone (multivalued) operators defined on a Hilbert space. In finite dimensional, linear spaces cyclic proximal point algorithms go back to [13]. Since then splitting algorithms have been applied to more general problems in the setting of both linear and non-linear spaces. For example, in the context of complete CAT(0) spaces, this proximal point algorithms were studied in [4], see also [21] and their cyclic version in [5]. In this paper, we introduce the cyclic proximal point algorithm in $\mathcal{P}_2(\mathbb{R}^d)$. For recent papers on related algorithms, see e.g. [20, 31].

Theorem 6.1. *Let $\mathscr{F}_i : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ be proper, lsc, coercive and convex along generalized geodesics. Denote by $\mathscr{F} = \sum_{i=1}^N \mathscr{F}_i$ and suppose that $\emptyset \neq D(\mathscr{F}) \subseteq D(\mathscr{F}_i)$ for $i = 1, 2, \cdots, N$. Let $\mathscr{J}_{\tau_i} : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ be the proximal mapping of $\mathscr{F}_i$ for $i = 1, 2, \cdots, N$. Assume that $\bigcap_{i=1}^N \mathrm{Fix}\, \mathscr{J}_{\tau_i} \neq \emptyset$. Then, for any $\mu_0 \in \overline{D(\mathscr{F})}$, the iterates $\mu_{n+1} = \mathscr{J}_{[n]}(\mu_n)$ converge narrowly to a solution of* (6.1).

Proof. Let $\mathscr{J} := \mathscr{J}_{\tau_N} \circ \mathscr{J}_{\tau_{N-1}} \circ \cdots \circ \mathscr{J}_{\tau_1}$. Assumption $\bigcap_{i=1}^N \mathrm{Fix}\, \mathscr{J}_{\tau_i} \neq \emptyset$ implies by Lemma 4.7 that $\mathrm{Fix}\, \mathscr{J} = \bigcap_{i=1}^N \mathrm{Fix}\, \mathscr{J}_{\tau_i}$. Given $\mu_0 \in \overline{D(\mathscr{F})}$, we define $\mu_{n+1} = \mathscr{J}_{\tau_{[n]}}(\mu_n)$. By Proposition 4.1, the mapping $\mathscr{J}_{\tau_i}$ is quasi 1/2-firmly nonexpansive

for every $i = 1, 2, \cdots, N$ and in particular $\mathscr{J}_{\tau_i}$ is quasi nonexpansive for every $i = 1, 2, \cdots, N$. Therefore, we get for every $\nu \in \operatorname{Fix} \mathscr{J}$ that

$$W_2(\mu_{n+1}, \nu) = W_2(\mathscr{J}_{\tau_{[n]}}(\mu_n), \nu) \leq W_2(\mu_n, \nu).$$

Consequently, the sequence $(\mu_n)_{n\in\mathbb{N}}$ is bounded. Hence, by Lemma 2.4 and Prokhorov's Theorem 2.3 it has a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ narrowly converging to some measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Since there are finitely many indices $i \in \{1, 2, \cdots, N\}$, we get by the pigeonhole principle that $\mu_{n_k} = \mathscr{J}_{\tau_j}(\mu_{n_k-1})$ for infinitely many $k \in \mathbb{N}$, for some $j \in \{1, 2, \cdots, N\}$. Moreover, by inequality (3.2), we have for all $l \in \mathbb{N}$ and all $\nu \in D(\mathscr{F})$ that

$$\frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l)}}, \nu) - \frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l)}-1}, \nu) \leq \mathscr{F}_j(\nu) - \mathscr{F}_j(\mu_{n_{k(l)}}).$$

From Fejér monotonicity, we get $W_2(\mu_{n_{k(l)}}, \nu) \geq W_2(\mu_{n_{k(l+1)}-1}, \nu)$ for any $\nu \in \operatorname{Fix} \mathscr{J}$ and every $l \in \mathbb{N}$. Then, rearranging terms in the last inequality, yields for all $l \in \mathbb{N}$ and all $\nu \in \operatorname{Fix} \mathscr{J}$ that

$$\mathscr{F}_j(\mu_{n_{k(l)}}) \leq \mathscr{F}_j(\nu) + \frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l-1)}}, \nu) - \frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l)}}, \nu).$$

Passing to the limit as $l \to +\infty$ and from Fejér monotonicity of $(\mu_n)_{n\in\mathbb{N}}$ with respect to $\operatorname{Fix} \mathscr{J}$, we obtain for all $\nu \in \operatorname{Fix} \mathscr{J}$ that

$$\liminf_{l\to+\infty} \mathscr{F}_j(\mu_{n_{k(l)}}) \leq \mathscr{F}_j(\nu) + \frac{1}{2\tau_j} \lim_{l\to+\infty} \Big(W_2^2(\mu_{n_{k(l-1)}}, \nu) - W_2^2(\mu_{n_{k(l)}}, \nu) \Big) = \mathscr{F}_j(\nu).$$

On the other hand, $\operatorname{Fix} \mathscr{J} \subseteq \operatorname{Fix} \mathscr{J}_{\tau_j} = \operatorname{argmin}\{\mathscr{F}_j(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$ implies that the last inequality holds for all $\nu \in \operatorname{argmin}\{\mathscr{F}_j(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$ and so for all $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Narrow lsc of $\mathscr{F}_j$ gives

$$\mathscr{F}_j(\mu) \leq \liminf_{l\to+\infty} \mathscr{F}_j(\mu_{n_{k(l)}}) \leq \mathscr{F}_j(\nu) \quad \text{for all } \nu \in \mathcal{P}_2(\mathbb{R}^d),$$

and therefore $\mu \in \operatorname{argmin}\{\mathscr{F}_j(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$. Now consider the sequence $(\mu_{n_{k(l)-1}})_{l\in\mathbb{N}}$ that by construction satisfies $\mu_{n_{k(l)-1}} = \mathscr{J}_{\tau_{j-1}}(\mu_{n_{k(l)-2}})$. Since $(\mu_{n_{k(l)-1}})_{l\in\mathbb{N}}$ is bounded, let $\mu_{n_{k(l)-1}} \overset{\mathcal{N}}{\to} \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, else by Lemma 2.4 and Prokhorov's Theorem 2.3, we can always pass to a subsequence of $(\mu_{n_{k(l)-1}})_{l\in\mathbb{N}}$ with this property. By similar arguments as above, we find that the limit $\mu' \in \operatorname{argmin}\{\mathscr{F}_{j-1}(\nu) : \nu \in \mathcal{P}_2(R^d)\}$. By inequality (3.2) we have

$$\frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l)}}, \mu_{n_{k(l)-1}}) + \frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l)}}, \nu) - \frac{1}{2\tau_j} W_2^2(\mu_{n_{k(l)-1}}, \nu) \leq \mathscr{F}_j(\nu) - \mathscr{F}_j(\mu_{n_{k(l)}})$$

for all $\nu \in D(\mathscr{F})$ and in particular for all $\nu \in \operatorname{Fix} \mathscr{J}$. From narrow lsc of $W(\cdot, \cdot)$ and Fejér monotonicity of $(\mu_n)_{n\in\mathbb{N}}$ with respect to $\operatorname{Fix} \mathscr{J}$, passing to the limit as $l \to +\infty$, we obtain

$$\frac{1}{2\tau_j} W_2^2(\mu, \mu') \leq \mathscr{F}_j(\nu) - \mathscr{F}_j(\mu) \quad \text{for all } \nu \in \operatorname{Fix} \mathscr{J}.$$

Since $\operatorname{Fix} \mathscr{J} \subseteq \operatorname{argmin}\{\mathscr{F}_j(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$, the last inequality holds in particular for all $\nu \in \operatorname{argmin}\{\mathscr{F}_j(\nu) : \nu \in \mathcal{P}_2(R^d)\}$. Therefore $W_2(\mu, \mu') \leq 0$ implies that $\mu = \mu'$.

This means that $\mu \in \operatorname{argmin}\{\mathscr{F}_{j-1}(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$. Repeating the same argument for every index $i \in \{1, 2, \cdots, N\}$ yields that $\mu \in \operatorname{argmin}\{\mathscr{F}_i(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$ for all $i = 1, 2, \cdots, N$, so that

$$\mu \in \bigcap_{i=1}^{N} \operatorname{argmin}\{\mathscr{F}_i(\nu) \,:\, \nu \in \mathcal{P}_2(\mathbb{R}^d)\} \subseteq \operatorname{argmin}\{\mathscr{F}(\nu) \,:\, \nu \in \mathcal{P}_2(\mathbb{R}^d)\}.$$

Hence we obtain for the original subsequence $\mu_{n_k} \overset{\mathcal{N}}{\to} \mu \in \bigcap_{i=1}^{N} \operatorname{argmin}\{\mathscr{F}_i(\nu) : \nu \in \mathcal{P}_2(\mathbb{R}^d)\}$. If μ' is another narrow cluster point of $(\mu_n)_{n\in\mathbb{N}}$, then, by same arguments, we obtain

$$\mu' \in \bigcap_{i=1}^{N} \operatorname{argmin}\{\mathscr{F}_i(\nu) \,:\, \nu \in \mathcal{P}_2(\mathbb{R}^d)\}.$$

We have that $\bigcap_{i=1}^{N} \operatorname{argmin}\{\mathscr{F}_i(\nu) \,:\, \nu \in \mathcal{P}_2(\mathbb{R}^d)\} = \operatorname{Fix} \mathscr{J}$ and that the sequence $(\mu_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix} \mathscr{J}$. Consequently, by Lemma 5.2, the whole sequence converges $\mu_n \overset{\mathcal{N}}{\to} \mu \in \operatorname{Fix} \mathscr{J}$. This completes the proof. □

Since the our theory relies on the assumption that the intersection of the fixed points sets of a finite collection of quasi α-firmly nonexpansive operators is nonempty, the last results cannot be applied to a situation when this intersection is empty. However, inspired by a result of Bačak [6, Theorem 6.3.7], we can provide a convergence theorem, when the functions $\mathscr{F}_i$ do not have a common minimizer, which essentially is the case when the corresponding proximal mappings $\mathscr{J}_{\tau_i}$ have no common fixed point. However, we need to add two conditions. First, each function $\mathscr{F}_i$ is Lipschitz continuous on $D(\mathscr{F}_i)$. This means that there exists $L_i > 0$ such that $|\mathscr{F}_i(\mu) - \mathscr{F}_i(\nu)| \leq L_i\, W_2(\mu, \nu)$ for all $\mu, \nu \in D(\mathscr{F}_i)$. Second, if $\mathscr{J}_{i,\tau_k}$ is the proximal mapping of $\mathscr{F}_i$ with step size τ_k, we require that $(\tau_k)_{k\in\mathbb{N}_0}$ satisfies $\sum_{k\in\mathbb{N}} \tau_k = +\infty$ and $\sum_{k\in\mathbb{N}_0} \tau_k^2 < +\infty$. Then we consider the iterations

$$\mu_{kN+n+1} = \mathscr{J}_{[n],\tau_k}(\mu_{kN+n}), \quad [n] = n \,(\operatorname{mod} N) + 1 \in \{1, 2, \cdots, N\},\ k = 0, 1, 2, \cdots. \tag{6.3}$$

Theorem 6.2. *Let $\mathscr{F}_i : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ be proper, lsc, coercive and convex along generalized geodesics. Denote by $\mathscr{F} = \sum_{i=1}^{N} \mathscr{F}_i$ and suppose that $\emptyset \neq D(\mathscr{F}) \subseteq D(\mathscr{F}_i)$ for $i = 1, 2, \cdots, N$. Assume that $\mathscr{F}_i$ is Lipschitz continuous on $D(\mathscr{F}_i)$ for every $i \in \{1, 2, \cdots, N\}$. Denote by $\mathscr{J}_{i,\tau_k}$ the proximal mapping of $\mathscr{F}_i$ with step size $\tau_k > 0$ satisfying $\sum_{k\in\mathbb{N}_0} \tau_k = +\infty$ and $\sum_{k\in\mathbb{N}_0} \tau_k^2 < +\infty$. Then, for any initial measure $\mu_0 \in \overline{D(\mathscr{F})}$, the iterates $\mu_{kN+n+1} = \mathscr{J}_{[n],\tau_k}(\mu_{kN+n})$ converge narrowly to a solution of problem* (6.1).

The proof follows similar steps as in [6, Theorem 6.3.7].

Proof. First, we get from inequality (3.2) for each $i \in \{1, 2, \cdots, N\}$ that

$$W_2^2(\mu_{kN+i}, \nu) \leq W_2^2(\mu_{kN+i-1}, \nu) - 2\tau_k(\mathscr{F}_i(\mu_{kN+i}) - \mathscr{F}_i(\nu)) \quad \text{for all } \nu \in D(\mathscr{F}).$$

Summing on the both sides of this inequality yields

$$W_2^2(\mu_{kN+N},\nu) \le W_2^2(\mu_{kN},\nu) - 2\tau_k \sum_{i=1}^{N} \mathscr{F}_i(\mu_{kN+i}) + 2\tau_k \mathscr{F}(\nu) \tag{6.4}$$

$$= W_2^2(\mu_{kN},\nu) - 2\tau_k(\mathscr{F}(\mu_{kN}) - \mathscr{F}(\nu)) + 2\tau_k \mathscr{F}(\mu_{kN}) - 2\tau_k \sum_{i=1}^{N} \mathscr{F}_i(\mu_{kN+i}). \tag{6.5}$$

The assumption that $\mathscr{F}_i$ is Lipschitz continuous on $D(\mathscr{F}_i)$ and hence on $D(\mathscr{F})$ for every $i \in \{1,2,\cdots,N\}$ implies for all $k \in \mathbb{N}_0$ that

$$\mathscr{F}(\mu_{kN}) - \sum_{i=1}^{N} \mathscr{F}_i(\mu_{kN+i}) = \sum_{i=1}^{N} (\mathscr{F}_i(\mu_{kN}) - \mathscr{F}_i(\mu_{kN+i})) \le \sum_{i=1}^{N} L_i\, W_2(\mu_{kN},\mu_{kN+i}).$$

By definition of the proximum we have for all $i \in \{1,2,\cdots,N\}$ and all $k \in \mathbb{N}_0$ that

$$\mathscr{F}_i(\mu_{kN+i}) + \frac{1}{2\tau_k} W_2^2(\mu_{kN+i-1},\mu_{kN+i}) \le \mathscr{F}_i(\mu_{kN+i-1}).$$

This implies

$$W_2(\mu_{kN+i-1},\mu_{kN+i}) \le 2\tau_k \frac{\mathscr{F}_i(\mu_{kN+i-1}) - \mathscr{F}_i(\mu_{kN+i})}{W_2(\mu_{kN+i-1},\mu_{kN+i})} \le 2\tau_k\, L_i.$$

This upper estimate together with an iterative application of the triangle inequality to the expression $W_2(\mu_{kN},\mu_{kN+i}) \le W_2(\mu_{kN},\mu_{kN+1}) + \cdots + W_2(\mu_{kN+i-1},\mu_{kN+i})$ yields

$$\sum_{i=1}^{N} \mathscr{F}_i(\mu_{kN}) - \mathscr{F}_i(\mu_{kN+i}) \le 2\tau_k \sum_{i=1}^{N} L_i \sum_{j=1}^{i} L_j \le \tau_k L_{\max}^2 N(N+1),$$

where $L_{\max} := \max\{L_i : i = 1,2,\cdots,N\}$. Therefore, we obtain for all $\nu \in D(\mathscr{F})$ the inequality

$$W_2^2(\mu_{kN+N},\nu) \le W_2^2(\mu_{kN},\nu) - 2\tau_k(\mathscr{F}(\mu_{kN}) - \mathscr{F}(\nu)) + 2\tau_k^2\, L_{\max}^2 N(N+1). \tag{6.6}$$

In particular, (6.6) holds if $\nu \in \operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$. Applying [6, Exercise 6.5] with $a_k := W_2^2(\mu_{kN},\nu)$, $b_k := \mathscr{F}(\mu_{kN}) - \mathscr{F}(\nu)$ and $c_k := 2\tau_k^2\, L_{\max}^2 N(N+1)$ yields that the sequence $(W_2(\mu_{kN},\nu))_{k\in\mathbb{N}_0}$ converges to a certain number $d(\nu) \ge 0$. In particular, the sequence $(\mu_{kN})_{k\in\mathbb{N}_0}$ is bounded. By Lemma 2.4 and Prokhorov's Theorem 2.3, there is a subsequence $\mu_{k_jN} \xrightarrow{\mathcal{N}} \mu$ for some $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. On the other hand, again by [6, Exercise 6.5], it holds that

$$\sum_{k\in\mathbb{N}_0} \tau_k(\mathscr{F}(\mu_{kN}) - \mathscr{F}(\nu)) < +\infty \quad \text{for all } \nu \in \operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}.$$

Therefore $\lim_{k\to+\infty} \tau_k(\mathscr{F}(\mu_{kN}) - \mathscr{F}(\nu)) = 0$ implies that $\lim_{k\to+\infty} \mathscr{F}(\mu_{kN}) = \mathscr{F}(\nu)$, else we can always pass to a subsequence of $(\mu_{kN})_{k\in\mathbb{N}_0}$ having this property. By narrow lsc of $\mathscr{F}$ we get that

$$\mathscr{F}(\mu) \le \liminf_{j\to+\infty} \mathscr{F}(\mu_{k_jN}) \le \limsup_{k\to+\infty} \mathscr{F}(\mu_{kN}) = \mathscr{F}(\nu)$$

for all $\nu \in \operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$.

Thus, $\mu \in \operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$. Let $(\mu_{k_m N})_{m\in\mathbb{N}}$ be another narrowly convergent subsequence of $(\mu_{kN})_{k\in\mathbb{N}_0}$. Let $\mu_{k_m N} \overset{\mathcal{N}}{\to} \mu' \in \mathcal{P}_2(\mathbb{R}^d)$. Note that (6.6) acts as a substitute in the argument of Lemma 5.2 for Fejér monotonicity of $(\mu_{kN})_{k\in\mathbb{N}_0}$ with respect to $\operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$. Indeed, let $r_1 = \limsup_{j\to+\infty} W_2(\mu_{k_j N}, \mu)$ and $r_2 = \limsup_{m\to+\infty} W_2(\mu_{k_m N}, \mu')$. Suppose w.l.o.g. that $r_1 \leq r_2$. From Opial's property (5.3) it follows that $r_2 < \limsup_{m\to+\infty} W_2(\mu_{k_m N}, \mu)$. For every $\varepsilon > 0$, there is $j_0 \in \mathbb{N}$ such that $W_2(\mu_{k_j N}, \mu) < r_1 + \varepsilon$, whenever $j \geq j_0$. In (6.6), let $\varepsilon_k^2 := 2\tau_k^2 L_{\max}^2 N(N+1)$. Then we have $W_2^2(\mu_{k_m N}, \mu) < (r_1+\varepsilon)^2 + \sum_{l=k_{j_0}}^{k_m} \varepsilon_l^2$ whenever $k_m \geq k_{j_0}$. For a fixed difference $\Delta(m, j_0) = k_m - k_{j_0}$, we let $j_0 \to +\infty$, i.e., also $m \to +\infty$. Since $\varepsilon_l \to 0$ and the sum $\sum_{l=k_{j_0}}^{k_m} \varepsilon_l^2$ is finite, we get for sufficiently large j_0 and sufficiently large m that $W_2(\mu_{k_m N}, \mu) < r_1 + 2\varepsilon$. Therefore there exists $m_1 \in \mathbb{N}$ such that $W_2(\mu_{k_m N}, \mu) < r_2 + 2\varepsilon$, whenever $m \geq m_1$, implying $\limsup_{m\to+\infty} W_2(\mu_{k_m N}, \mu) \leq r_2$. This would raise a contradiction. Therefore the sequence $(\mu_{kN})_{k\in\mathbb{N}_0}$ must have a unique narrow cluster point. Following exactly the same arguments as in Lemma 5.2, we get that the whole sequence converges $\mu_{kN} \overset{\mathcal{N}}{\to} \mu$. Now consider $(\mu_{kN+i})_{k\in\mathbb{N}_0}$ for $i = 1, 2, \cdots, N-1$. Repeating the same reasoning as for $(\mu_{kN})_{k\in\mathbb{N}_0}$, we conclude that $\mu_{kN+i} \overset{\mathcal{N}}{\to} \mu_i \in \operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$ for each $i = 1, 2, \cdots, N-1$. From the estimate $W_2(\mu_{kN}, \mu_{kN+i}) \leq 2\tau_k \sum_{j=1}^{i} L_j$ and narrow lsc of $W_2(\cdot,\cdot)$, we obtain that

$$0 \leq W_2(\mu, \mu_i) \leq \liminf_{k\to+\infty} W_2(\mu_{kN}, \mu_{kN+i}) \leq \lim_{k\to+\infty} (2\tau_k \sum_{j=1}^{i} L_j) = 0.$$

Hence $\mu = \mu_i$ for every $i = 1, 2, \cdots, N-1$. This means that the whole sequence of iterates $\mu_{kN+n+1} = \mathscr{J}_{[n],\tau_k}(\mu_{kN+n})$ converges narrowly to $\mu \in \operatorname{argmin}\{\mathscr{F}(\sigma) : \sigma \in \mathcal{P}_2(\mathbb{R}^d)\}$. □

Acknowledgment: This research was supported by DFG under Germany's Excellence Strategy The Berlin Mathematics Research Center MATH+ (EXC-2046/1, Projektnummer: 390685689).

References

[1] A.D. Aleksandrov, *A theorem on triangles in a metric space and some of its applications*, Trudy Mat. Inst. Steklov., **38**(1951), 5–23.

[2] L. Ambrosio, E. Brué, D. Semola, *Lectures on Optimal Transport*, Unitext, Springer, no. 130, 2021.

[3] L. Ambrosio, N. Gigli, G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics ETH Zürich, Birkhäuser, 2005.

[4] M. Bačak, *The proximal point algorithm in metric spaces*, Isr. J. Math., **194**(2013), no. 2, 689–701.

[5] M. Bačak, *Computing medians and means in Hadamard spaces*, SIAM J. Optim., **24**(2014), no. 3, 1542–1566.

[6] M. Bačak, *Convex Analysis and Optimization in Hadamard Spaces*, 22 of De Gruyter Series in Nonlinear Analysis and Applications, De Gruyter, Berlin, 2014.

[7] M. Bačak, R. Bergmann, G. Steidl, A. Weinmann, *A second order non-smooth variational model for restoring manifold-valued images*, SIAM J. Sci. Comput., **38**(2016), no. 1, 567–597.

[8] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
[9] A. Bërdëllima, *On a notion of averaged mappings in $CAT(0)$ spaces*, Funct. Anal. its Appl., **56**(2022), no. 1, 37–50.
[10] A. Bërdëllima, F. Lauster, D.R. Luke, *α-Firmly nonexpansive operators on metric spaces*, J. Fixed Point Theory Appl., **, 24**(2022), no. 1.
[11] A. Bërdëllima, G. Steidl, *On α-firmly nonexpansive operators in r-uniformly convex spaces*, Results Math., **76**(2021).
[12] R. Bergmann, J. Persch, G. Steidl, *A parallel Douglas-Rachford algorithm for minimizing ROF-like functionals on images with values in symmetric Hadamard manifolds*, SIAM J. Imag. Sci., bf 9(2016), no. 3, 901–937.
[13] D.P. Bertsekas, *Incremental proximal methods for large scale convex optimization*, Math. Program., Ser. B, **129**(2011), no. 2, 163–195.
[14] C. Bonet, L. Chapel, L. Drumetz, N. Courty, *Hyperbolic sliced-Wasserstein via geodesic and horospherical projections*, Topological, Algebraic and Geometric Learning Workshops, PMLR, 2023, 334–370.
[15] C. Bonet, B. Malazieux, A. Rakotomamonjy, L. Drumetz, T. Moreau, M. Kowalski, N. Courty, *Sliced-Wasserstein on symmetric positive definite matrices for M/EEG signals*, International Conference on Machine Learning, PMLR, 2023, 2777–2805.
[16] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn., **3**(2011), no. 1, 101–122.
[17] F.E. Browder, *Convergence theorems for sequences of nonlinear operators in Banach spaces*, Math. Zeitschr., **100**(1967), 201–225.
[18] R.E. Bruck Jr., *Nonexpansive projections on subsets of Banach spaces*, Pacific J. Math., **47**(1973), 341–355.
[19] M. Cuturi, G. Peyré, *Computational optimal transport*, Found. Trends Mach. Learn., **11**(2019), no. 5-6, 355–607.
[20] J. Fan, A. Taghvaei, Y. Chen, *Variational Wasserstein gradient flow*, arXiv: 2112.02424, 2021.
[21] O.P. Ferreira, P.R. Oliveira, *Proximal point algorithm on Riemannian manifolds*, Optim., **51**(2002), no. 2, 257–270.
[22] J. Hertrich, R. Beinert, M. Gräf, G. Steidl, *Wasserstein steepest descent flow of discrepancies with Riezs kernels*, J. Math. Anal. Appl, **531**(2024), no. 1, part 1.
[23] R. Jordan, D. Kinderlehrer, F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal., **29**(1998), 1–17.
[24] M.A. Krasnoselskij, *Two remarks on the method of successive approximations*, Uspehi. Mat. Nauk (N.S.), **10**(1955), 123–127.
[25] P.L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., **16**(1979), no. 6, 964–979.
[26] D.R. Luke, N.H. Thao, M.K. Tam, *Quantitative convergence analysis of iterated expansive, set-valued mappings*, Math. Oper. Res., **43**(2018), no. 4, 1143–1176.
[27] W.R. Mann, *Mean value methods in iteration*, Proc. Amer. Math. Soc., **4**(1953), 506–510.
[28] E. Naldi, G. Savaré, *Weak topology and Opial property in Wasserstein spaces, with applications to gradient flows and proximal point algorithms of geodesically convex functionals*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur., **32**(2022), no. 4, 725–750.
[29] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc., **73**(1967), 591–597.
[30] Y.V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, (Russian), Theory Probab. Its Appl., **1**(1956), no. 2, 157–214.
[31] A. Salim, A. Korba, G. Luise, *The Wasserstein proximal gradient algorithm*, arXiv: 2002.03035, 2021.
[32] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Springer, 2015.
[33] F. Santambrogio, *Euclidean, metric, and Wasserstein gradient flows: An overview*, Bull. Math. Sci., **7**(2017), 87-154.

[34] C. Villani, *Optimal Transport: Old and New*, Grundlehren der Mathematischen Wissenschaften 338, Springer-Verlag Berlin Heidelberg, 2009.
[35] M.C. Yue, D. Kuhn, W. Wiesemann, *On linear optimization over Wasserstein balls*, Math. Program., 2021.

Received: September 19, 2022; Accepted: March 18, 2024.