Applied Mathematics & Optimization (2023) 87:37
https://doi.org/10.1007/s00245-022-09948-y

A Note on the Differentiability of the Hellinger–Kantorovich Distances

Florentine Catharina Fleißner[1]

Accepted: 24 November 2022 / Published online: 13 March 2023

Abstract

This paper will deal with differentiability properties of the class of Hellinger–Kantorovich distances $\mathsf{HK}_{\Lambda,\Sigma}$ $(\Lambda, \Sigma > 0)$ which was recently introduced on the space $\mathcal{M}(\mathbb{R}^d)$ of finite nonnegative Radon measures. The derivatives of $t \mapsto \mathsf{HK}_{\Lambda,\Sigma}(\mu_t, \nu_t)^2$, for absolutely continuous curves $(\mu_t)_t, (\nu_t)_t$ in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$, will be computed $\mathscr{L}^1$-a.e.. The characterization of absolutely continuous curves in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$ will be refined.

1 Introduction

Recently, a new class of distances on the space $\mathcal{M}(\mathbb{R}^d)$ of finite nonnegative Radon measures was established by three independent teams [3, 4, 7–9]. We will follow the presentation of these distances by Liero, Mielke and Savaré [8, 9] who named it *Hellinger–Kantorovich distances*. The class of Hellinger–Kantorovich distances $\mathsf{HK}_{\Lambda,\Sigma}$ $(\Lambda, \Sigma > 0)$ is based on the conversion of one measure into another one (possibly having different total mass) by means of transport and creation / annihilation of mass. The parameters Λ and Σ serve as weightings of the transport part and the mass creation/annihilation part respectively. To be more precise, the square $\mathsf{HK}_{\Lambda,\Sigma}(\mu_1, \mu_2)^2$ of the Hellinger–Kantorovich distance $\mathsf{HK}_{\Lambda,\Sigma}$ between two measures $\mu_1, \mu_2 \in \mathcal{M}(\mathbb{R}^d)$ on $\mathbb{R}^d$ corresponds to

$$\min\Big\{\sum_{i=1}^{2}\frac{4}{\Sigma}\int_{\mathbb{R}^d}(\sigma_i\log\sigma_i - \sigma_i + 1)\,\mathrm{d}\mu_i \\ + \int_{\mathbb{R}^d\times\mathbb{R}^d}\mathsf{c}_{\Lambda,\Sigma}(|x_1 - x_2|)\,\mathrm{d}\gamma \,:\, \gamma \in \mathcal{M}(\mathbb{R}^d\times\mathbb{R}^d),\ \gamma_i \ll \mu_i\Big\}, \tag{1.1}$$

✉ Florentine Catharina Fleißner
fleissne@ma.tum.de

1 Technische Universität München, Munich, Germany

with entropy cost functions $\frac{4}{\Sigma}(\sigma_i \log \sigma_i - \sigma_i + 1)$,

$$\sigma_i := \frac{\mathrm{d}\gamma_i}{\mathrm{d}\mu_i} \quad (\gamma_i \text{ i-th marginal of } \gamma), \tag{1.2}$$

and transportation cost function

$$\mathsf{c}_{\Lambda,\Sigma}(\mathsf{d}) := \begin{cases} -\frac{8}{\Sigma}\log(\cos(\sqrt{\Sigma/(4\Lambda)}\mathsf{d})) & \text{if } \mathsf{d} < \pi\sqrt{\Lambda/\Sigma}, \\ +\infty & \text{if } \mathsf{d} \geq \pi\sqrt{\Lambda/\Sigma}. \end{cases} \tag{1.3}$$

There exists an optimal plan γ for the Logarithmic Entropy-Transport problem (1.1) (cf. Thm. 3.3 in [9]), and if μ_1 is absolutely continuous with respect to the Lebesgue measure and γ is such optimal plan, then there exists a Borel optimal transport mapping $t : \mathbb{R}^d \to \mathbb{R}^d$ so that γ takes the form

$$\gamma \ = \ (I \times t)_\# \gamma_1 \ = \ (I \times t)_\# \sigma_1 \mu_1$$

(cf. Thm. 4.5 in [6] and Thm. 6.6 in [9]). We refer the reader to ([9], Cor. 7.14, Thms. 7.17 and 7.20) for the proofs that $\mathsf{HK}_{\Lambda,\Sigma}$ defined via the Logarithmic Entropy-Transport problem (1.1) indeed represents a distance on the space of finite nonnegative Radon measures and that $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$ is a complete metric space. Furthermore, the Hellinger–Kantorovich distance $\mathsf{HK}_{\Lambda,\Sigma}$ metrizes the weak topology on $\mathcal{M}(\mathbb{R}^d)$ in duality with continuous and bounded functions (cf. Thm. 7.15 in [9]) and can be interpreted as weighted infimal convolution of the Kantorovich-Wasserstein distance and the Hellinger-Kakutani distance. A representation formula à la Benamou-Brenier which can be proved for $\mathsf{HK}_{\Lambda,\Sigma}$ (cf. ([9], Thm. 8.18; [8], Thm. 3.6(v))) justifies this interpretation:

$$\mathsf{HK}_{\Lambda,\Sigma}(\mu_1,\mu_2)^2 = \min\left\{ \int_0^1 \int_{\mathbb{R}^d} (\Lambda|v_t|^2 + \Sigma|w_t|^2)\,\mathrm{d}\mu_t\,\mathrm{d}t \,:\, \mu_1 \overset{(\mu,v,w)}{\rightsquigarrow} \mu_2 \right\} \tag{1.4}$$

where $\mu_1 \overset{(\mu,v,w)}{\rightsquigarrow} \mu_2$ means that $\mu : [0,1] \to \mathcal{M}(\mathbb{R}^d)$ is a continuous curve connecting $\mu(0) = \mu_1$ and $\mu(1) = \mu_2$ and satisfying the continuity equation with reaction

$$\partial_t \mu_t = -\Lambda \mathrm{div}(v_t \mu_t) + \Sigma w_t \mu_t, \tag{1.5}$$

governed by Borel functions $v : (0,1) \times \mathbb{R}^d \to \mathbb{R}^d$ and $w : (0,1) \times \mathbb{R}^d \to \mathbb{R}$ with

$$\int_0^1 \int_{\mathbb{R}^d} (\Lambda|v_t|^2 + \Sigma|w_t|^2)\,\mathrm{d}\mu_t\,\mathrm{d}t < +\infty, \tag{1.6}$$

in duality with C^∞-functions with compact support in $(0.1) \times \mathbb{R}^d$, i.e.

$$\int_0^1 \int_{\mathbb{R}^d} (\partial_t \psi(t,x) + \Lambda \langle \nabla \psi(t,x), v(t,x) \rangle \\ + \Sigma \psi(t,x) w(t,x)) \, \mathrm{d}\mu_t(x) \, \mathrm{d}t \; = \; 0 \tag{1.7}$$

for all $\psi \in C_c^\infty((0,1) \times \mathbb{R}^d)$.

The class of such continuous curves μ satisfying (1.5), (1.6) for some Borel vector field (v, w) coincides with the class of *absolutely continuous* curves $(\mu_t)_{t\in[0,1]}$ in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$ *with square-integrable metric derivatives* (cf. Thms. 8.16 and 8.17 in [9], see Sect. 3 in this paper).

In order to deepen our understanding of a distance, it is always worth studying its differentiability along absolutely continuous curves (e.g. see Chap. 8 in [1] for the corresponding analysis of the Kantorovich-Wasserstein distance on the space of Borel probability measures with finite second order moments). The present paper addresses this issue for the class of Hellinger–Kantorovich distances on the space of finite nonnegative Radon measures. Clearly, if $(\mu_t)_{t\in[0,1]}, (\nu_t)_{t\in[0,1]}$ are absolutely continuous curves in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$, then the mapping

$$t \mapsto \mathsf{HK}_{\Lambda,\Sigma}(\mu_t, \nu_t)^2 \tag{1.8}$$

is absolutely continuous and therefore $\mathscr{L}^1$-a.e. differentiable. A natural question that arises is the one of the concrete form of the corresponding derivatives. We will answer this question for absolutely continuous curves with square-integrable metric derivatives (for which such characterization (1.5) is available), refine that characterization by providing more information on (v, w) (see Prop. 3.1), establish a linearization result (see Thm. 3.4), and determine

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathsf{HK}_{\Lambda,\Sigma}(\mu_t, \nu_t)^2 \tag{1.9}$$

at $\mathscr{L}^1$-a.e. $t \in [0,1]$ (see Thm. 4.1). This piece of work can be viewed as continuation of Sect. 3 in the author's paper [5] constituting a starting point for the study of differentiability properties of the Hellinger–Kantorovich distances. Therein, we identified elements of the Fréchet subdifferential of mappings

$$t \mapsto -\mathsf{HK}_{\Lambda,\Sigma}((I + tv)_\#(1 + tR)^2 \mu_0, \nu)^2$$

at $t = 0$, for $\mu_0, \nu \in \mathcal{M}(\mathbb{R}^d)$ and bounded Borel functions $v : \mathbb{R}^d \to \mathbb{R}^d$ and $R : \mathbb{R}^d \to \mathbb{R}$. That subdifferential calculus was an essential ingredient for our Minimizing Movement approach to a class of scalar reaction-diffusion equations [5] substantiating their gradient-flow-like structure in the space of finite nonnegative Radon measures endowed with the Hellinger–Kantorovich distance $\mathsf{HK}_{\Lambda,\Sigma}$.

The proof in [9] that absolutely continuous curves in $(\mathcal{M}(\mathbb{H}), \mathsf{HK}_{\Lambda,\Sigma})$ with square-integrable metric derivatives are characterized via (1.5), (1.6) was carried out only for $\mathbb{H} = \mathbb{R}^d$, endowed with usual scalar product $\langle \cdot, \cdot \rangle$ and norm $|\cdot| := \sqrt{\langle \cdot, \cdot \rangle}$, but according to a comment at the beginning of Sect. 8.5 in [9], it should be possible to prove such characterization result in a more general setting. We would like to remark that also our computation of the derivatives (1.9) may be adapted for general separable Hilbert spaces $\mathbb{H}$.

Our plan for the paper is to give an equivalent characterization of the Hellinger–Kantorovich distances in Sect. 2, to state and prove new results on absolutely continuous curves in Sect. 3 and to perform the computation of the derivatives (1.9) in Sect. 4.

2 Optimal Transportation on the Cone

According to ([8], Sect. 4) and ([9], Sect. 7), the Logarithmic Entropy-Transport problem (1.1) translates into a problem of optimal transportation on the geometric cone $\mathfrak{C}$ on $\mathbb{R}^d$, see (2.16), (2.17) below. The fact that all the information on transport of mass and creation / annihilation of mass according to (1.1) lies in a pure transportation problem has proved extremely useful for the analysis of $\mathsf{HK}_{\Lambda,\Sigma}$ in [9] and for our subdifferential calculus in [5].

Geometric cone $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$. The geometric cone is defined as the quotient space

$$\mathfrak{C} := \mathbb{R}^d \times [0, +\infty)/\sim \tag{2.1}$$

with

$$(x_1, r_1) \sim (x_2, r_2) \quad \Leftrightarrow \quad r_1 = r_2 = 0 \text{ or } r_1 = r_2,\ x_1 = x_2 \tag{2.2}$$

and is endowed with a class of distances $\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}$ $(\Lambda, \Sigma > 0)$. The vertex $\mathfrak{o}$ (for $r = 0$) and $[x, r]$ (for $x \in \mathbb{R}^d$ and $r > 0$) denote the corresponding equivalence classes and

$$\begin{aligned}&\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1, r_1], [x_2, r_2])^2\\ &\quad := \frac{4}{\Sigma}\Big(r_1^2 + r_2^2 - 2r_1 r_2 \cos\Big(\Big(\sqrt{\Sigma/4\Lambda}\,|x_1 - x_2|\Big) \wedge \pi\Big)\Big)\end{aligned} \tag{2.3}$$

(where $\mathfrak{o}$ is identified with $[\bar{x}, 0]$ for some $\bar{x} \in \mathbb{R}^d$). It can be proved that

$$\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(y_0, y_1)^2 = \min\left\{ \int_0^1 \Big(\frac{4}{\Sigma}(\dot{r}(s))^2 + \frac{1}{\Lambda} r(s)^2 |\dot{x}(s)|^2\Big)\,\mathrm{d}s \,\Big|\, y_0 \overset{[x,r]}{\rightsquigarrow} y_1 \right\} \tag{2.4}$$

for $y_i = [x_i, r_i] \in \mathfrak{C}$, where $y_0 \overset{[x,r]}{\rightsquigarrow} y_1$ means that $x \in \mathrm{C}^1([0,1]; \mathbb{R}^d)$, $r \in \mathrm{C}^1([0,1]; [0, +\infty))$ and $[x(i), r(i)] = y_i$, so that the cone distance may be interpreted as dissipation distance generated by the metric tensor

$$\mathfrak{G}^{\Lambda,\Sigma}_{[x,r]}((\dot{x}_1, \dot{r}_1), (\dot{x}_2, \dot{r}_2)) := \frac{4}{\Sigma}\dot{r}_1\dot{r}_2 + \frac{1}{\Lambda} r^2 \langle \dot{x}_1, \dot{x}_2 \rangle \tag{2.5}$$

(cf. Sect. 8.1 in [9]). This metric tensor (2.5) will appear in the formulas in our differential calculus of $\mathsf{HK}_{\Lambda,\Sigma}$.

We show how to construct geodesics in $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$ (cf. Sect. 8.1 in [9]) as they will play an important role in our analysis of (1.9), too. Let $y_i := [x_i, r_i] \in \mathfrak{C}$, $i = 1, 2$, and suppose that $|x_1 - x_2| \leq \pi\sqrt{\Lambda/\Sigma}$, $r_1, r_2 > 0$. We search for functions $\mathcal{R}_{y_1,y_2} : [0, 1] \to [0, +\infty)$ and $\theta_{y_1,y_2} : [0, 1] \to [0, 1]$ so that the curve $\eta : [0, 1] \to \mathfrak{C}$ defined as $\eta(s) := [x_1 + \theta_{y_1,y_2}(s)(x_2 - x_1), \mathcal{R}_{y_1,y_2}(s)]$ is a (constant speed) geodesic connecting $[x_1, r_1]$ and $[x_2, r_2]$, which means $\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\eta(s), \eta(t)) = |s - t|\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1, r_1], [x_2, r_2])$ for all $s, t \in [0, 1]$. If $x_1 = x_2$, we set $\theta_{y_1,y_2} \equiv 0$. We note that

$$\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\eta(s), \eta(t))^2 = |z(s) - z(t)|^2_{\mathbb{C}}, \tag{2.6}$$

where $z : [0, 1] \to \mathbb{C}$ is the curve in the complex plane $\mathbb{C}$ defined as

$$z(s) := \frac{2}{\sqrt{\Sigma}}\mathcal{R}_{y_1,y_2}(s) \exp\Big(i\theta_{y_1,y_2}(s)\sqrt{\Sigma/4\Lambda}\,|x_1 - x_2|\Big), \tag{2.7}$$

and $|\cdot|_{\mathbb{C}}$ denotes the absolute value for complex numbers. Thus, if z is a geodesic in the complex plane between $z_1 := \frac{2}{\sqrt{\Sigma}}r_1$ and $z_2 := \frac{2}{\sqrt{\Sigma}}r_2 \exp\Big(i\sqrt{\Sigma/4\Lambda}\,|x_1 - x_2|\Big)$, i.e.

$$z(s) = z_1 + s(z_2 - z_1) \quad \text{for all } s \in [0, 1], \tag{2.8}$$

then η is a geodesic in $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$ between $[x_1, r_1]$ and $[x_2, r_2]$. This condition yields an appropriate choice for $\mathcal{R}_{y_1,y_2} : [0, 1] \to [0, +\infty)$ and $\theta_{y_1,y_2} : [0, 1] \to [0, 1]$, and it is not difficult to see that they are both smooth functions, their first derivatives satisfy

$$\begin{aligned}&\frac{4}{\Sigma}(\mathcal{R}'_{y_1,y_2}(s))^2 + \frac{1}{\Lambda}\mathcal{R}_{y_1,y_2}(s)^2(\theta'_{y_1,y_2}(s))^2|x_1 - x_2|^2 \\ &\quad = \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1, r_1], [x_2, r_2])^2 \quad \text{for all } s \in (0, 1),\end{aligned} \tag{2.9}$$

and they are right differentiable at $s = 0$ with right derivatives

$$\begin{aligned}\theta'_{y_1,y_2,+}(0) &= \frac{r_2}{r_1}\frac{\sin(\sqrt{\Sigma/4\Lambda}\,\mathsf{d}(x_1, x_2))}{\sqrt{\Sigma/4\Lambda}\,\mathsf{d}(x_1, x_2)}, \\ \mathcal{R}'_{y_1,y_2,+}(0) &= r_2 \cos\Big(\sqrt{\Sigma/4\Lambda}\,\mathsf{d}(x_1, x_2)\Big) - r_1.\end{aligned} \tag{2.10}$$

It is noteworthy that

$$\mathsf{t}_{y_1,y_2}(s) := \Big(\theta'_{y_1,y_2}(s)(x_2 - x_1), \mathcal{R}'_{y_1,y_2}(s)\Big) \tag{2.11}$$

represents the tangent vector to the geodesic at $\eta(s)$, $s \in (0, 1)$, with

$$\mathsf{t}_{y_1,y_2}(0) := \lim_{s\downarrow 0} \mathsf{t}_{y_1,y_2}(s) = \Big(\theta'_{y_1,y_2,+}(0)(x_2 - x_1), \mathcal{R}'_{y_1,y_2,+}(0)\Big), \tag{2.12}$$

and the left-hand side of (2.9) equals the metric tensor $\mathfrak{G}^{\Lambda,\Sigma}_{\eta(s)}(\mathsf{t}_{y_1,y_2}(s), \mathsf{t}_{y_1,y_2}(s))$ (cf. (2.5)).

We obtain a geodesic from $[x_1, r_1]$ to the vertex $\mathfrak{o}$ by setting $\theta_{y_1,\mathfrak{o}} \equiv 0$ and $\mathcal{R}_{y_1,\mathfrak{o}}(s) := (1 - s)r_1$ and identifying $\mathfrak{o}$ with $[x_1, 0]$. Also in this case, (2.9) and the second part of (2.10) hold good.

Optimal transportation problem. The distance $\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}$ gives rise to an optimal transport problem on the cone and therefore to an extended quadratic Kantorovich-Wasserstein distance $\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}$ on the space $\mathcal{M}_2(\mathfrak{C})$ of finite nonnegative Radon measures on $\mathfrak{C}$ with finite second order moments, i.e. $\int_{\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x, r], \mathfrak{o})^2 \, \mathrm{d}\alpha([x, r]) < +\infty$. The extended Kantorovich-Wasserstein distance $\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_1, \alpha_2)$ between two measures $\alpha_1, \alpha_2 \in \mathcal{M}_2(\mathfrak{C})$ is equal to $+\infty$ if $\alpha_1(\mathfrak{C}) \neq \alpha_2(\mathfrak{C})$ and is given by

$$\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_1, \alpha_2)^2 := \min\Big\{ \int_{\mathfrak{C}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1, r_1], [x_2, r_2])^2 \, \mathrm{d}\beta \mid \beta \in \Gamma(\alpha_1, \alpha_2)\Big\} \tag{2.13}$$

if $\alpha_1(\mathfrak{C}) = \alpha_2(\mathfrak{C})$, with $\Gamma(\alpha_1, \alpha_2)$ being the set of finite nonnegative Radon measures on $\mathfrak{C} \times \mathfrak{C}$ whose first and second marginals coincide with α_1 and α_2. Every measure $\alpha \in \mathcal{M}_2(\mathfrak{C})$ on the cone is assigned a measure $\mathfrak{h}\alpha \in \mathcal{M}(\mathbb{R}^d)$ on $\mathbb{R}^d$,

$$\mathfrak{h}\alpha := \mathsf{x}_\#(\mathsf{r}^2\alpha), \tag{2.14}$$

with $(\mathsf{x}, \mathsf{r}) : \mathfrak{C} \to \mathbb{R}^d \times [0, +\infty)$ defined as

$$(\mathsf{x}, \mathsf{r})([x, r]) := (x, r) \text{ for } [x, r] \in \mathfrak{C},\ r > 0,\ (\mathsf{x}, \mathsf{r})(\mathfrak{o}) := (\bar{x}, 0), \tag{2.15}$$

which means $\int_{\mathbb{R}^d} \phi(x) \, \mathrm{d}(\mathfrak{h}\alpha) = \int_{\mathfrak{C}} \mathsf{r}^2\phi(\mathsf{x}) \, \mathrm{d}\alpha$ for all continuous and bounded functions $\phi : \mathbb{R}^d \to \mathbb{R}$ (short $\phi \in \mathrm{C}^0_b(\mathbb{R}^d)$). Please note that the mapping $\mathfrak{h} : \mathcal{M}_2(\mathfrak{C}) \to \mathcal{M}(\mathbb{R}^d)$ is not injective.

Now, an equivalent characterization of the Hellinger–Kantorovich distance $\mathsf{HK}_{\Lambda,\Sigma}$ is given by the transportation problems

$$\mathsf{HK}_{\Lambda,\Sigma}(\mu_1, \mu_2)^2 = \min\Big\{\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_1, \alpha_2)^2 \,\Big|\, \alpha_i \in \mathcal{M}_2(\mathfrak{C}),\ \mathfrak{h}\alpha_i = \mu_i\Big\} \tag{2.16}$$

$$= \min\Big\{\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_1, \alpha_2)^2 + \frac{4}{\Sigma}\sum_{i=1}^{2}(\mu_i - \mathfrak{h}\alpha_i)(\mathbb{R}^d) \Big|\, \alpha_i \in \mathcal{M}_2(\mathfrak{C}), \ \mathfrak{h}\alpha_i \le \mu_i\Big\}, \tag{2.17}$$

cf. Probl. 7.4, Thm. 7.6, Lem. 7.9, Thm. 7.20 in [9]. Every solution $\gamma \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ to the Logarithmic Entropy-Transport problem (1.1) induces a solution $\beta \in \mathcal{M}(\mathfrak{C} \times \mathfrak{C})$ to ((2.17), (2.13)): if γ is an optimal plan for (1.1) with Lebesgue decompositions [1]

$$\mu_i = \rho_i \gamma_i + \mu_i^\perp, \tag{2.18}$$

then

$$\beta := ([x_1, \sqrt{\rho_1(x_1)}], [x_2, \sqrt{\rho_2(x_2)}])_{\#}\gamma \ \in \mathcal{M}(\mathfrak{C} \times \mathfrak{C}) \tag{2.19}$$

is an optimal plan for the transport problem (2.17), (2.13) (cf. ([9], Thm. 7.20(iii))). Furthermore, if $\beta \in \mathcal{M}(\mathfrak{C} \times \mathfrak{C})$ is a solution to (2.17), (2.13) or a solution to (2.16), (2.13) (which exists by ([9], Thm. 7.6)), then

$$\beta\left(\left\{([x_1, r_1], [x_2, r_2]) \in \mathfrak{C} \times \mathfrak{C} : \ r_1, r_2 > 0, \ |x_1 - x_2| > \pi\sqrt{\Lambda/\Sigma}\right\}\right) = 0, \tag{2.20}$$

(cf. ([9], Lem. 7.19)).

3 Absolutely Continuous Curves

We fix $\Lambda, \Sigma > 0$ and examine the behaviour of absolutely continuous curves in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$.

Let $(\mu_t)_{t\in[0,1]}$ be an absolutely continuous curve in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK}_{\Lambda,\Sigma})$ with square-integrable metric derivative, i.e. the limit

$$|\mu_t'| := \lim_{h\to 0} \frac{\mathsf{HK}_{\Lambda,\Sigma}(\mu_{t+h}, \mu_t)}{|h|} \tag{3.1}$$

exists for $\mathscr{L}^1$-a.e. $t \in (0, 1)$, the function $t \mapsto |\mu_t'|$ which is called *metric derivative* of $(\mu_t)_t$ belongs to $\mathrm{L}^2((0, 1))$ and

$$\mathsf{HK}_{\Lambda,\Sigma}(\mu_s, \mu_t) \le \int_s^t |\mu_r'|\,\mathrm{d}r \qquad \text{for all } 0 \le s \le t \le 1 \tag{3.2}$$

(cf. Def. 1.1.1 and Thm. 1.1.2 in [1]). According to Thms. 8.16 and 8.17 in [9], there exists an essentially unique Borel vector field $(v, w) : (0, 1) \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}$ so that the continuity equation with reaction

$$\partial_t \mu_t = -\Lambda \mathrm{div}(v_t \mu_t) + \Sigma w_t \mu_t \tag{3.3}$$

[1] according to Lem. 2.3 in [9], there exist Borel functions $\rho_i : \mathbb{R}^d \to [0, +\infty)$ and nonnegative finite Radon measures $\mu_i^\perp \in \mathcal{M}(\mathbb{R}^d)$, $\mu_i^\perp \perp \gamma_i$, so that (2.18) holds good.

$(v_t := v(t,\cdot),\ w_t := w(t,\cdot))$ holds good, in duality with C^∞-functions with compact support in $(0,1)\times\mathbb{R}^d$ (see (1.7)), and

$$\int_{\mathbb{R}^d} (\Lambda|v_t|^2 + \Sigma|w_t|^2)\,\mathrm{d}\mu_t \;=\; |\mu_t'|^2 \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,1). \tag{3.4}$$

For every $t \in (0,1)$ and $h \in (-t, 1-t)$, there exists a plan $\beta_{t,t+h} \in \mathcal{M}(\mathfrak{C}\times\mathfrak{C})$ which is optimal in the definition of $\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h})^2$ according to (2.16), (2.13), i.e.

$$\begin{aligned}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h})^2 &= \int_{\mathfrak{C}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2\,\mathrm{d}\beta_{t,t+h},\\ \mathfrak{h}(\pi^1_\#\beta_{t,t+h}) &= \mu_t,\ \mathfrak{h}(\pi^2_\#\beta_{t,t+h}) = \mu_{t+h},\end{aligned}$$

and whose first marginal $\pi^1_\#\beta_{t,t+h}$ satisfies

$$\int_{\mathfrak{C}} \phi([x,r])\,\mathrm{d}(\pi^1_\#\beta_{t,t+h}) = \int_{\mathbb{R}^d} \phi([x,1])\,\mathrm{d}\mu_t + h^2\phi(\mathfrak{o}) \tag{3.5}$$

for all $\phi \in \mathrm{C}^0_b(\mathfrak{C})$ (cf. Thm. 7.6 and Lem. 7.10 in [9]).

This notation holds good throughout the rest of the paper.

As a first result of our analysis of absolutely continuous curves, Prop. 3.1 will identify (v_t, w_t) as belonging to a particular class of functions.

Proposition 3.1 *For $\mathscr{L}^1$-a.e. $t \in (0,1)$, the Borel function (v_t, w_t) belongs to the closure in $\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})$ of the subspace $\{(\nabla\zeta,\zeta) :\ \zeta \in \mathrm{C}^\infty_c(\mathbb{R}^d)\}$.*

Here $(\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R}), ||\cdot||_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})})$ denotes the normed space of all μ_t-measurable functions $(\bar v,\bar w)$ from $\mathbb{R}^d$ to $\mathbb{R}^d\times\mathbb{R}$ satisfying

$$||(\bar v,\bar w)||_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})} := \Big(\int_{\mathbb{R}^d} (\Lambda|\bar v|^2 + \Sigma|\bar w|^2)\,\mathrm{d}\mu_t\Big)^{1/2} < +\infty. \tag{3.6}$$

Proof We construct a Borel vector field $(\tilde v,\tilde w) : (0,1)\times\mathbb{R}^d \to \mathbb{R}^d\times\mathbb{R}$ satisfying (3.3) so that, for $\mathscr{L}^1$-a.e. $t \in (0,1)$, the function $(\tilde v_t,\tilde w_t)$ belongs to the closure in $\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})$ of the subspace $\{(\nabla\zeta,\zeta) :\ \zeta \in \mathrm{C}^\infty_c(\mathbb{R}^d)\}$ and

$$||(\tilde v_t,\tilde w_t)||^2_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})} = \int_{\mathbb{R}^d} (\Lambda|\tilde v_t|^2 + \Sigma|\tilde w_t|^2)\,\mathrm{d}\mu_t \;\le\; |\mu_t'|^2. \tag{3.7}$$

We begin the proof with some estimations. Let $\phi \in C_c^\infty(\mathbb{R}^d)$. It follows from the construction of $\mathcal{R}_{[x_1,r_1],[x_2,r_2]}$ and $\theta_{[x_1,r_1],[x_2,r_2]}$ according to (2.6)-(2.9) that

$$
\begin{aligned}
&\frac{2}{\Sigma}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s)^2 = \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2,\\
&\Big|\theta''_{[x_1,r_1],[x_2,r_2]}(s)\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s)^2(x_2-x_1)\Big| \le C_{\Sigma,\Lambda}\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2,\\
&\Big|2\theta'_{[x_1,r_1],[x_2,r_2]}(s)\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s)\mathcal{R}'_{[x_1,r_1],[x_2,r_2]}(s)(x_2-x_1)\Big|\\
&\le C_{\Sigma,\Lambda}\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2,\\
&\left|\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big[\phi(x_1+\theta_{[x_1,r_1],[x_2,r_2]}(s)(x_2-x_1))\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s)^2\Big]\right|\\
&\le C_\phi C_{\Sigma,\Lambda}\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2,
\end{aligned}
$$

for $s \in (0,1)$, with $C_\phi > 0$ only depending on ϕ and $C_{\Sigma,\Lambda} := 2\Sigma + 4\Lambda$; we refer the reader to the proof of Prop. 2.5 in [5] for details. With (2.9) and these estimations on hand, it is straightforward to prove that there exists a constant $C_{\phi,\Lambda,\Sigma} > 0$ only depending on ϕ, Λ and Σ so that

$$
|\varphi'_{y_1,y_2}(\bar{s}) - \varphi'_{y_1,y_2}(s)| \le C_{\phi,\Lambda,\Sigma}\,\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(y_1,y_2)^2, \tag{3.8}
$$

$$
\begin{aligned}
&\Big|\varphi'_{y_1,y_2}(s) - \langle\nabla\phi(x_1),\theta'_{y_1,y_2}(s)(x_2-x_1)\rangle\mathcal{R}_{y_1,y_2}(s)^2 + 2\phi(x_1)\mathcal{R}'_{y_1,y_2}(s)\mathcal{R}_{y_1,y_2}(s)\Big|\\
&\le C_{\phi,\Lambda,\Sigma}\,\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(y_1,y_2)^2
\end{aligned} \tag{3.9}
$$

and

$$
\begin{aligned}
&\Big|\Big(\langle\nabla\phi(x_1),\theta'_{y_1,y_2}(s)(x_2-x_1)\rangle\mathcal{R}_{y_1,y_2}(s) + 2\phi(x_1)\mathcal{R}'_{y_1,y_2}(s)\Big)\Big(\mathcal{R}_{y_1,y_2}(s) - r_1\Big)\Big|\\
&\le C_{\phi,\Lambda,\Sigma}\,\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(y_1,y_2)^2
\end{aligned} \tag{3.10}
$$

for all $s,\bar{s} \in (0,1)$, with $y_i := [x_i,r_i]$, $\varphi_{y_1,y_2}(s) := \phi(x_1 + \theta_{[x_1,r_1],[x_2,r_2]}(s)(x_2 - x_1))\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s)^2$.

Now, let $t \in (0,1)$ so that the limit (3.1) exists and $\mathfrak{C}_\mathfrak{o} := \mathfrak{C}\setminus\{\mathfrak{o}\}$. By applying (2.20), (3.9), (3.10), (3.5), Hölder's inequality and (2.9), we obtain

$$
\begin{aligned}
&\Big|\int_{\mathbb{R}^d}\phi\,\mathrm{d}\mu_{t+h} - \int_{\mathbb{R}^d}\phi\,\mathrm{d}\mu_t\Big| = \Big|\int_{\mathfrak{C}\times\mathfrak{C}}(\phi(x_2)r_2^2 - \phi(x_1)r_1^2)\,\mathrm{d}\beta_{t,t+h}\Big|\\
&\le \int_{\mathfrak{C}\times\mathfrak{C}}\int_0^1|\varphi'_{y_1,y_2}(s)|\,\mathrm{d}s\,\mathrm{d}\beta_{t,t+h} \le\\
&\int_{\mathfrak{C}_\mathfrak{o}\times\mathfrak{C}}\int_0^1\Big|\langle\nabla\phi(x_1),\theta'_{[x_1,r_1],[x_2,r_2]}(s)(x_2-x_1)\rangle\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s) + 2\phi(x_1)\mathcal{R}'_{[x_1,r_1],[x_2,r_2]}(s)\Big|\,\mathrm{d}s\,\mathrm{d}\beta_{t,t+h}\\
&\quad + 2C_{\phi,\Lambda,\Sigma}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h})^2\\
&\le \Big(\int_{\mathfrak{C}_\mathfrak{o}}\Big(\Lambda|\nabla\phi|^2 + \Sigma\phi^2\Big)\,\mathrm{d}(\pi^1_\#\beta_{t,t+h})\Big)^{1/2} Big(\int_{\mathfrak{C}_\mathfrak{o}\times\mathfrak{C}}\int_0^1\Big(\frac{1}{\Lambda}\mathcal{R}^2(\theta')^2|x_2-x_1|^2 + \frac{4}{\Sigma}(\mathcal{R}')^2\Big)\,\mathrm{d}s\,\mathrm{d}\beta_{t,t+h}\Big)^{1/2}\\
&\quad + 2C_{\phi,\Lambda,\Sigma}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h})^2 \;\le\; ||(\nabla\phi,\phi)||_{L^2(\mu_t,\mathbb{R}^d\times\mathbb{R})}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h}) + 2C_{\phi,\Lambda,\Sigma}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h})^2
\end{aligned}
$$

and thus,

$$\limsup_{h\to 0} \frac{1}{|h|}\Big| \int_{\mathbb{R}^d} \phi \, \mathrm{d}\mu_{t+h} - \int_{\mathbb{R}^d} \phi \, \mathrm{d}\mu_t \Big| \ \le \ ||(\nabla\phi, \phi)||_{\mathrm{L}^2(\mu_t, \mathbb{R}^d\times\mathbb{R})}|\mu_t'|. \quad (3.11)$$

At this point, we may follow the proof of Thm. 8.3.1 in [1]. Therein, a similar characterization of absolutely continuous curves in the space of Borel probability measures with finite second order moments, endowed with the Kantorovich-Wasserstein distance, was given by solving a suitable minimum problem. We adapt that approach. Let $\mu \in \mathcal{M}((0, 1) \times \mathbb{R}^d)$ be defined by

$$\int_{(0,1)\times\mathbb{R}^d} \psi(t, x) \, \mathrm{d}\mu(t, x) \ = \ \int_0^1 \int_{\mathbb{R}^d} \psi(t, x) \, \mathrm{d}\mu_t(x) \, \mathrm{d}t$$

for all $\psi \in \mathrm{C}_b^0((0, 1) \times \mathbb{R}^d)$, and let $(\mathrm{L}^2(\mu, \mathbb{R}^d \times \mathbb{R}), || \cdot ||_{\mathrm{L}^2(\mu,\mathbb{R}^d\times\mathbb{R})})$ denote the normed space of all μ-measurable vector fields $(\hat{v}, \hat{w})$ from $(0, 1) \times \mathbb{R}^d$ to $\mathbb{R}^d \times \mathbb{R}$ satisfying

$$||(\hat{v}, \hat{w})||_{\mathrm{L}^2(\mu,\mathbb{R}^d\times\mathbb{R})} := \Big(\int_0^1 \int_{\mathbb{R}^d} (\Lambda|\hat{v}_t|^2 + \Sigma|\hat{w}_t|^2) \, \mathrm{d}\mu_t \, \mathrm{d}t \Big)^{1/2} < +\infty. \quad (3.12)$$

An application of (3.11), Fatou's Lemma, Hölder's inequality and Hahn-Banach Theorem shows that there exists a unique bounded linear functional L defined on the closure $\mathcal{V}$ in $\mathrm{L}^2(\mu, \mathbb{R}^d \times \mathbb{R})$ of the subspace $\{(\nabla\zeta, \zeta) \ : \ \zeta \in \mathrm{C}_c^\infty((0, 1) \times \mathbb{R}^d)\}$, satisfying

$$L((\nabla\zeta, \zeta)) := -\int_0^1 \int_{\mathbb{R}^d} \partial_t\zeta(t, x) \, \mathrm{d}\mu_t \, \mathrm{d}t \quad \text{for all } \zeta \in \mathrm{C}_c^\infty((0, 1) \times \mathbb{R}^d). \quad (3.13)$$

We consider the minimum problem

$$\min \Big\{ \frac{1}{2}||(\hat{v}, \hat{w})||^2_{\mathrm{L}^2(\mu,\mathbb{R}^d\times\mathbb{R})} - L((\hat{v}, \hat{w})) : \ (\hat{v}, \hat{w}) \in \mathcal{V} \Big\}. \quad (3.14)$$

The same argument as in the proof of Thm. 8.3.1 in [1] proves that the unique solution $(\tilde{v}, \tilde{w})$ to (3.14) (which clearly exists) satisfies (3.3) and, for $\mathscr{L}^1$-a.e. $t \in (0, 1)$, the function $(\tilde{v}_t, \tilde{w}_t)$ belongs to the closure in $\mathrm{L}^2(\mu_t, \mathbb{R}^d \times \mathbb{R})$ of the subspace $\{(\nabla\zeta, \zeta) : \zeta \in \mathrm{C}_c^\infty(\mathbb{R}^d)\}$ and (3.7) holds good. By Thm. 8.17 in [9], for every Borel vector field $(\hat{v}, \hat{w}) \in \mathrm{L}^2(\mu, \mathbb{R}^d \times \mathbb{R})$ satisfying the continuity equation with reaction (3.3) the opposite inequality holds good, i.e.

$$\int_{\mathbb{R}^d} (\Lambda|\hat{v}_t|^2 + \Sigma|\hat{w}_t|^2)\,\mathrm{d}\mu_t \;\geq\; |\mu_t'|^2 \quad \text{for } \mathscr{L}^1\text{-a.e.} t \in (0,1).$$

It follows from this and from the strict convexity of $||\cdot||^2_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})}$ that the Borel vector field $(\tilde{v}, \tilde{w})$ solves (3.3), (3.4) and that it coincides $\mathscr{L}^1$-a.e. with any other vector field solving (3.3), (3.4). This completes the proof of Prop. 3.1. □

Definition 3.2 Let $\mathcal{C}(\mathbb{R}^d)$ be a countable subset of $\mathrm{C}_c^\infty(\mathbb{R}^d)$ so that every function in $\mathrm{C}_c^\infty(\mathbb{R}^d)$ can be approximated in the C^1-norm by a sequence of functions in $\mathcal{C}(\mathbb{R}^d)$.
We define $\mathcal{N}_\mu$ as the set of points $t \in (0,1)$ at which the following holds good:

(i) The limit (3.1) exists,
(ii) (v_t, w_t) belongs to the closure in $\mathrm{L}^2(\mu_t, \mathbb{R}^d \times \mathbb{R})$ of the subspace $\{(\nabla\zeta, \zeta) : \zeta \in \mathrm{C}_c^\infty(\mathbb{R}^d)\}$ and satisfies (3.4),
(iii) and, for all $\psi \in \mathcal{C}(\mathbb{R}^d)$,

$$\lim_{h\to 0} \frac{1}{h}\Big(\int_{\mathbb{R}^d} \psi\,\mathrm{d}\mu_{t+h} - \int_{\mathbb{R}^d} \psi\,\mathrm{d}\mu_t\Big) = \int_{\mathbb{R}^d} (\Lambda\langle\nabla\psi, v_t\rangle + \Sigma\psi w_t)\,\mathrm{d}\mu_t. \tag{3.15}$$

Please note that $(0,1)\setminus\mathcal{N}_\mu$ is an $\mathscr{L}^1$-negligible set; it follows from (1.7) that, for fixed $\psi \in \mathrm{C}_c^\infty(\mathbb{R}^d)$, the mapping $t \mapsto \int_{\mathbb{R}^d} \psi\,\mathrm{d}\mu_t$ is absolutely continuous from [0, 1] to $\mathbb{R}$ and (3.15) holds good at $\mathscr{L}^1$-a.e. $t \in (0,1)$.

The second step in our analysis is to establish a connection between the "tangent vector" (v_t, w_t) to μ_t and tangent vectors to geodesics in $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$, measured by $\beta_{t,t+h}$ for $|h|$ small. For $t \in \mathcal{N}_\mu$, $h \in (-t, 1-t)$ and $s \in (0,1)$, the mappings

$$\begin{aligned}\mathfrak{D}_{t,h,s} : (y_1, y_2) \;\mapsto\; \Big((\mathsf{x}(y_1), \mathsf{r}(y_1)), \Big(&\frac{1}{h\Lambda}\mathcal{R}_{y_1,y_2}(s)\theta'_{y_1,y_2}(s)(\mathsf{x}(y_2)\\ &-\mathsf{x}(y_1)), \frac{2}{h\Sigma}\mathcal{R}'_{y_1,y_2}(s)\Big)\Big)\end{aligned} \tag{3.16}$$

from $\Big(\mathfrak{C}\times\mathfrak{C}\Big)\setminus\Big\{([x_1,r_1],[x_2,r_2]) \in \mathfrak{C}\times\mathfrak{C} :\; r_1, r_2 > 0,\; |x_1 - x_2| > \pi\sqrt{\Lambda/\Sigma}\Big\}$ to $(\mathbb{R}^d\times\mathbb{R})\times(\mathbb{R}^d\times\mathbb{R})$ will be considered, with x, r as in (2.15), and $\mathcal{R}_{y_1,y_2}$, θ_{y_1,y_2} being constructed according to (2.6)–(2.9). Their second components may be interpreted as blow-ups of tangent vectors to geodesics in $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$; in fact, the transition from (x, r) to the local chart $(1/\Lambda\,\mathcal{R}_{y_1,y_2}(s)\,\mathsf{x}, 2/\Sigma\,\mathsf{r})$ transforms the tangent vector $\mathfrak{t}_{y_1,y_2}(s)$ from (2.11) into the tangent vector

$$\tilde{\mathfrak{t}}_{y_1,y_2}(s) := \Big(\frac{1}{\Lambda}\mathcal{R}_{y_1,y_2}(s)\theta'_{y_1,y_2}(s)(\mathsf{x}(y_2) - \mathsf{x}(y_1)), \frac{2}{\Sigma}\mathcal{R}'_{y_1,y_2}(s)\Big). \tag{3.17}$$

We will take advantage of the fact that this chart transition transforms the metric tensor $\mathfrak{G}^{\Lambda,\Sigma}$ from (2.5) into a metric tensor which is equal to $\Lambda < \mathfrak{v}_1, \mathfrak{w}_1 > +\Sigma\, \mathfrak{v}_2\mathfrak{w}_2$ for tangent vectors $\mathfrak{v} := (\mathfrak{v}_1, \mathfrak{v}_2)$, $\mathfrak{w} := (\mathfrak{w}_1, \mathfrak{w}_2) \in \mathbb{R}^d \times \mathbb{R}$ at $[x, \mathcal{R}_{y_1,y_2}(s)] \in \mathfrak{C}$.

We turn to the push-forward $\Delta_{t,h,s} \in \mathcal{M}((\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}))$ of $\beta_{t,t+h}$ through (3.16), defined by

$$\int_{(\mathbb{R}^d\times\mathbb{R})\times(\mathbb{R}^d\times\mathbb{R})} \Phi(y)\,\mathrm{d}\Delta_{t,h,s} \quad = \quad \int_{\mathfrak{C}\times\mathfrak{C}} \Phi(\mathfrak{D}_{t,h,s}(y_1, y_2))\,\mathrm{d}\beta_{t,t+h}$$

for all $\Phi \in \mathrm{C}_b^0((\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}))$. Please recall (2.20) in this context and note that, by (2.9), the mappings (3.16) are Borel measurable. The following proposition will provide information on the limits of $\Delta_{t,h,s}$ as $h \to 0$, linking them to (v_t, w_t).

Proposition 3.3 *Let $t \in \mathcal{N}_\mu$ and $s \in (0, 1)$. Then*

$$\lim_{h\to 0} \int_{(\mathbb{R}^d\times\mathbb{R})\times(\mathbb{R}^d\times\mathbb{R})} \Phi(y)\,\mathrm{d}\Delta_{t,h,s} \;=\; \int_{\mathbb{R}^d} \Phi((x, 1), (v_t(x), w_t(x)))\,\mathrm{d}\mu_t \tag{3.18}$$

for all continuous functions $\Phi : (\mathbb{R}^d \times \mathbb{R}) \times (\mathbb{R}^d \times \mathbb{R}) \to \mathbb{R}$ satisfying the growth condition

$$|\Phi((x_1, r_1), (x_2, r_2))| \;\le\; C\Big(1 + |x_2|^2 + |r_2|^2\Big) \tag{3.19}$$

for some $C > 0$.

Proof We set $Y := \mathbb{R}^d \times \mathbb{R}$.

Let $t \in \mathcal{N}_\mu$ and $s \in (0, 1)$. We note that, by (2.9) and Def. 3.2(i),

$$\begin{aligned}&\int_{Y\times Y} (\Lambda|x_2|^2 + \Sigma|r_2|^2)\,\mathrm{d}\Delta_{t,h,s}((x_1, r_1), (x_2, r_2)) \\ &\quad = \frac{\mathsf{HK}_{\Lambda,\Sigma}(\mu_t, \mu_{t+h})^2}{h^2} \to |\mu_t'|^2 \quad \text{as } h \to 0.\end{aligned} \tag{3.20}$$

We may apply Prokhorov's Theorem to any sequence $(\Delta_{t,h_k,s})_{k\in\mathbb{N}}$, $h_k \to 0$, of measures from the family $(\Delta_{t,h,s})_{h\in(-t,1-t)} \subset \mathcal{M}(Y \times Y)$, since such sequence is bounded and equally tight by (3.5) and (3.20), and we obtain a subsequence $h_{k_l} \to 0$ and a measure $\Delta \in \mathcal{M}(Y \times Y)$ so that $(\Delta_{t,h_{k_l},s})_{l\in\mathbb{N}}$ converges to Δ in the weak topology on $\mathcal{M}(Y \times Y)$, in duality with continuous and bounded functions. So let $(\Delta_{t,h_l,s})_{l\in\mathbb{N}}$ ($h_l \to 0$) be a convergent sequence with limit measure $\Delta \in \mathcal{M}(Y \times Y)$, i.e.

$$\lim_{l\to\infty} \int_{Y\times Y} \Phi(y)\,\mathrm{d}\Delta_{t,h_l,s} \;=\; \int_{Y\times Y} \Phi(y)\,\mathrm{d}\Delta \tag{3.21}$$

for all $\Phi \in \mathrm{C}_b^0(Y \times Y)$. We want to identify Δ as $((x, 1), (v_t(x), w_t(x)))_\#\mu_t$. It is not difficult to infer from (3.5) that the first marginal $\pi^1_\#\Delta$ of Δ coincides with $(x, 1)_\#\mu_t$, i.e.

$$\int_Y \phi((x, r))\,\mathrm{d}(\pi^1_\#\Delta) = \int_{\mathbb{R}^d} \phi((x, 1))\,\mathrm{d}\mu_t \tag{3.22}$$

for all $\phi \in \mathrm{C}_b^0(Y)$. Let $\psi \in \mathcal{C}(\mathbb{R}^d)$. Then (3.21) also holds good for $\Phi((x_1, r_1), (x_2, r_2)) := \Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big]r_1$: Indeed, we have

$$\lim_{l\to\infty}\int_{Y\times Y}(\Phi_N)\,\mathrm{d}\Delta_{t,h_l,s} = \int_{Y\times Y}(\Phi_N)\,\mathrm{d}\Delta$$

for all $N > 0$, with $\Phi_N := (\Phi\wedge N)\vee(-N)$. Setting $Y_N := \{(x, r) \in Y : |x| + |r| > N\}$, $C_\psi := \sup_{x\in\mathbb{R}^d}\{|\nabla\psi(x)| + |\psi(x)|\}$, and applying (3.5), (3.20) and (3.22), we conclude that for every $\epsilon > 0$ there exists $N_\epsilon > 0$ so that

$$\int_{Y\times Y_N}(|x_2| + |r_2|)\,\mathrm{d}\Delta_{t,h_l,s} + \int_{Y\times Y_N}(|x_2| + |r_2|)\,\mathrm{d}\Delta \;\le \epsilon \quad \text{for all } N \ge N_\epsilon,\ l \in \mathbb{N},$$

and

$$\begin{aligned}
&\limsup_{l\to\infty}\Big|\int_{Y\times Y}\Phi\,\mathrm{d}\Delta_{t,h_l,s} - \int_{Y\times Y}\Phi\,\mathrm{d}\Delta\Big| \\
&\le \limsup_{l\to\infty}\Big|\int_{Y\times Y}(\Phi_{C_\psi(\Lambda+\Sigma)N_\epsilon})\,\mathrm{d}\Delta_{t,h_l,s} - \int_{Y\times Y}\Phi_{C_\psi(\Lambda+\Sigma)N_\epsilon}\,\mathrm{d}\Delta\Big| \\
&\quad + C_\psi(\Lambda+\Sigma)\limsup_{l\to\infty}\int_{Y\times Y_{N_\epsilon}}(|x_2| + |r_2|)\,\mathrm{d}(\Delta_{t,h_l,s} + \Delta) \\
&\le C_\psi(\Lambda+\Sigma)\epsilon.
\end{aligned}$$

Hence, taking (3.22) into account, we obtain

$$\begin{aligned}
&\lim_{l\to\infty}\int_{Y\times Y}\Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big]r_1\,\mathrm{d}\Delta_{t,h_l,s} \\
&\quad = \int_{Y\times Y}\Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big]\,\mathrm{d}\Delta.
\end{aligned}\tag{3.23}$$

It holds that

$$\begin{aligned}
&\int_{\mathbb{R}^d}\psi\,\mathrm{d}\mu_{t+h_l} - \int_{\mathbb{R}^d}\psi\,\mathrm{d}\mu_t = \int_{\mathfrak{C}\times\mathfrak{C}}(\psi(x_2)r_2^2 - \psi(x_1)r_1^2)\,\mathrm{d}\beta_{t,t+h} \\
&\quad = \int_{\mathfrak{C}\times\mathfrak{C}}\int_0^1\frac{\mathrm{d}}{\mathrm{d}s}\Big[\psi(x_1 + \theta_{[x_1,r_1],[x_2,r_2]}(s)(x_2 - x_1))\mathcal{R}_{[x_1,r_1],[x_2,r_2]}(s)^2\Big]\,\mathrm{d}s\,\mathrm{d}\beta_{t,t+h_l}
\end{aligned}$$

so that (3.15), (3.8), (3.9), (3.10), Def. 3.2(i) and (3.23) yield

$$\begin{aligned}
&\int_{\mathbb{R}^d}(\Lambda\langle\nabla\psi, v_t\rangle + \Sigma\psi w_t)\,\mathrm{d}\mu_t = \lim_{l\to\infty}\frac{1}{h_l}\Big(\int_{\mathbb{R}^d}\psi\,\mathrm{d}\mu_{t+h_l} - \int_{\mathbb{R}^d}\psi\,\mathrm{d}\mu_t\Big) \\
&\quad = \lim_{l\to\infty}\int_{Y\times Y}\Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big]r_1\,\mathrm{d}\Delta_{t,h_l,s} \\
&\quad = \int_{Y\times Y}\Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big]\,\mathrm{d}\Delta.
\end{aligned}$$

According to the Disintegration Theorem (see e.g. Thm. 5.3.1 in [1]) and (3.22), there exists a Borel family of probability measures $(\Delta_{x_1})_{x_1\in\mathbb{R}^d} \subset \mathcal{M}(Y),\ \Delta_{x_1}(Y) = 1$, so that

$$\int_{Y\times Y} \Phi \,\mathrm{d}\Delta = \int_{\mathbb{R}^d} \Big(\int_Y \Phi((x_1, 1), (x_2, r_2)) \,\mathrm{d}\Delta_{x_1}((x_2, r_2))\Big) \,\mathrm{d}\mu_t(x_1)$$

for all Δ-integrable maps $\Phi : Y \times Y \to \mathbb{R}$. We infer from (3.20) that, for μ_t-a.e. $x_1 \in \mathbb{R}^d$, the measure Δ_{x_1} has finite second order moment and we define the function $(v_\Delta, w_\Delta) : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}$ by

$$\begin{aligned} v_\Delta(x_1) \ &:= \int_Y x_2 \,\mathrm{d}\Delta_{x_1}((x_2, r_2)),\ \ w_\Delta(x_1) := \int_Y r_2 \,\mathrm{d}\Delta_{x_1}((x_2, r_2)) \\ &\text{for } \mu_t\text{-a.e. } x_1 \in \mathbb{R}^d. \end{aligned} \tag{3.24}$$

The function (v_Δ, w_Δ) is Borel measurable (cf. (5.3.1) and Def. 5.4.2 in [1]), and

$$\begin{aligned} &\int_{Y\times Y} \Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big] \,\mathrm{d}\Delta \\ &= \int_{\mathbb{R}^d} \Big(\int_Y \Big[\Lambda\langle\nabla\psi(x_1), x_2\rangle + \Sigma\psi(x_1)r_2\Big] \,\mathrm{d}\Delta_{x_1}((x_2, r_2))\Big) \,\mathrm{d}\mu_t(x_1) \\ &= \int_{\mathbb{R}^d} (\Lambda\langle\nabla\psi, v_\Delta\rangle + \Sigma\psi w_\Delta) \,\mathrm{d}\mu_t. \end{aligned}$$

All in all, we have found that

$$\int_{\mathbb{R}^d} (\Lambda\langle\nabla\psi, v_t\rangle + \Sigma\psi w_t) \,\mathrm{d}\mu_t = \int_{\mathbb{R}^d} (\Lambda\langle\nabla\psi, v_\Delta\rangle + \Sigma\psi w_\Delta) \,\mathrm{d}\mu_t \tag{3.25}$$

for all $\psi \in \mathcal{C}(\mathbb{R}^d)$. Since every function in $\mathrm{C}_c^\infty(\mathbb{R}^d)$ can be approximated in the C^1-norm by a sequence of functions in $\mathcal{C}(\mathbb{R}^d)$ (cf. Def. 3.2) and, by (3.20) and Def. 3.2(ii), the functions $v_\Delta, w_\Delta, v_t, w_t$ are square-integrable w.r.t. μ_t, (3.25) holds good for all $\psi \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ and for all pairs in the $\mathrm{L}^2(\mu_t, \mathbb{R}^d \times \mathbb{R})$-closure of $\{(\nabla\zeta, \zeta) : \zeta \in \mathrm{C}_c^\infty(\mathbb{R}^d)\}$. It follows from this and from Def. 3.2(ii) that

$$||(v_t, w_t)||^2_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})} = \int_{\mathbb{R}^d} (\Lambda\langle v_t, v_\Delta\rangle + \Sigma w_t w_\Delta) \,\mathrm{d}\mu_t. \tag{3.26}$$

Applying Hölder's inequality to (3.26), taking the definition (3.24) of v_Δ, w_Δ, Jensen's inequality, (3.21), (3.20) and Def. 3.2(ii) into account, we obtain

$$\begin{aligned}||(v_t, w_t)||_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})} &\le ||(v_\Delta, w_\Delta)||_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})}\\ &\le \Big(\int_{Y\times Y}(\Lambda|x_2|^2+\Sigma|r_2|^2)\,\mathrm{d}\Delta\Big)^{1/2} \le \qquad (3.27)\\ &\le \lim_{l\to\infty}\Big(\int_{Y\times Y}(\Lambda|x_2|^2+\Sigma|r_2|^2)\,\mathrm{d}\Delta_{t,h_l,s}\Big)^{1/2}\\ &= ||(v_t, w_t)||_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})} \qquad (3.28)\end{aligned}$$

so that, in fact, equality holds good everywhere in (3.27) and (3.28). We infer from this and from (3.26) that

$$||(v_t, w_t) - (v_\Delta, w_\Delta)||_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})} = 0$$

which means

$$v_t(x) = v_\Delta(x) \text{ and } w_t(x) = w_\Delta(x) \qquad \text{for } \mu_t\text{-a.e. } x \in \mathbb{R}^d. \tag{3.29}$$

Moreover, the fact that the second inequality in (3.27), resulting from Jensen's inequality, is in fact an equality and (3.29) yield $\Delta_{x_1} = \delta_{v_t(x_1)} \otimes \delta_{w_t(x_1)}$ for μ_t-a.e. $x_1 \in \mathbb{R}^d$ (cf. a canonical proof of Jensen's inequality), i.e.

$$\int_Y \phi((x, r))\,\mathrm{d}\Delta_{x_1} = \phi(v_t(x_1), w_t(x_1)) \tag{3.30}$$

for all $\phi \in \mathrm{C}_b^0(Y)$, for μ_t-a.e. $x_1 \in \mathbb{R}^d$.

Altogether, we may conclude that $\Delta = ((x, 1), (v_t(x), w_t(x)))_\#\mu_t$,

$$\begin{aligned}\int_{Y\times Y}(\Lambda|x_2|^2+\Sigma|r_2|^2)\,\mathrm{d}\Delta &= |\mu_t'|^2\\ &= \lim_{l\to\infty}\int_{Y\times Y}(\Lambda|x_2|^2+\Sigma|r_2|^2)\,\mathrm{d}\Delta_{t,h_l,s} \qquad (3.31)\end{aligned}$$

and that (3.18) holds good for all $\Phi \in \mathrm{C}_b^0(Y \times Y)$. A similar argument as in the proof of (3.23), making use of (3.31), will show (3.18) for all continuous functions $\Phi : Y \times Y \to \mathbb{R}$ satisfying the growth condition (3.19) (cf. Thm. 7.12 in [10] where the space of Borel probability measures with finite second order moments is considered and the equivalence between convergence in the Kantorovich-Wasserstein distance and convergence in duality with continuous functions satisfying a suitable growth condition is proved). This completes the proof of Prop. 3.3. □

Now, Theorem 3.4 yields a linearization result for absolutely continuous curves.

Theorem 3.4 *Let $t \in \mathcal{N}_\mu$.*

Define $\mathfrak{C}_{t,h} := \left\{[x,r] \in \mathfrak{C} \setminus \{\mathfrak{o}\} : \; |v_t(x)| < \frac{1}{\sqrt{|h|}} \text{ and } |w_t(x)| < \frac{2}{\sqrt{|h|}\Sigma}\right\}$ *and* $\Xi_{t,h} : \mathfrak{C} \to \mathfrak{C}$,

$$\Xi_{t,h}([x,r]) := \begin{cases} [x + \Lambda h v_t(x), r(1 + \frac{\Sigma}{2} h w_t(x))] \text{ if } [x,r] \in \mathfrak{C}_{t,h}, \\ [x,r] \text{ else }. \end{cases} \tag{3.32}$$

Let $\chi_{t,h} := (\Xi_{t,h})_\#(\pi^1_\# \beta_{t,t+h})$ *be the push-forward of the first marginal of* $\beta_{t,t+h}$ *through* $\Xi_{t,h}$, *i.e.*

$$\int_{\mathfrak{C}} \phi([x,r]) \, \mathrm{d}\chi_{t,h} = \int_{\mathfrak{C}} \phi(\Xi_{t,h}([x,r])) \, \mathrm{d}(\pi^1_\# \beta_{t,t+h})$$

for all $\phi \in \mathrm{C}^0_b(\mathfrak{C})$. *Then*

$$\lim_{h \to 0} \frac{\mathsf{HK}_{\Lambda,\Sigma}(\mu_{t+h}, \mathfrak{h}\chi_{t,h})^2}{h^2} \; = \; 0. \tag{3.33}$$

Remark 3.5 The technical role of $\mathfrak{C}_{t,h}$ will be visible in the proof. First, the restriction to $\mathfrak{C}_{t,h}$ ensures that $[x + \Lambda h v_t(x), r(1 + \frac{\Sigma}{2} h w_t(x))] \in \mathfrak{C}$ is well-defined, and second, we will take advantage thereof in order to suitably estimate $\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}(y_1), y_2)^2/h^2$ for $(y_1, y_2) \in \operatorname{supp} \beta_{t,t+h}$.

Proof We set $Y := \mathbb{R}^d \times \mathbb{R}$.

Let $t \in \mathcal{N}_\mu$. According to (2.13), (2.16), we have

$$\frac{\mathsf{HK}_{\Lambda,\Sigma}(\mu_{t+h}, \mathfrak{h}\chi_{t,h})^2}{h^2} \; \le \; \frac{1}{h^2} \int_{\mathfrak{C}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}([x_1,r_1]), [x_2,r_2])^2 \, \mathrm{d}\beta_{t,t+h}. \tag{3.34}$$

We will prove that the right-hand side of (3.34) converges to 0 as $h \to 0$.

First we note that, by Prokhorov's Theorem, Def. 3.2(ii) and the proof of Prop. 3.3, every sequence $\left(((v_t(x_1), w_t(x_1)), (x_2, r_2))_\# \Delta_{t,h_l,s}\right)_{l \in \mathbb{N}}$, $h_l \to 0$, is relatively compact w.r.t. the weak topology in $\mathcal{M}(Y \times Y)$ and in duality with continuous functions $\Phi : Y \times Y \to \mathbb{R}$ satisfying (3.19), and the second marginals of the corresponding limit measures coincide with $(v_t(x), w_t(x))_\# \mu_t$. It follows therefrom that for $N \in \mathbb{N}$, $\bar{s} \in (0,1)$,

$$\begin{aligned} &\limsup_{h \to 0} \frac{1}{h^2} \int_{(\mathfrak{C} \setminus \mathfrak{C}_{t,1/N}) \times \mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2 \, \mathrm{d}\beta_{t,t+h} \\ &= \limsup_{h \to 0} \int_{Y \times Y} (\Lambda |x_2|^2 + \Sigma |r_2|^2) \mathbb{1}_{\{x : |v_t(x)| \ge \sqrt{N} \text{ or } |w_t(x)| \ge 2\sqrt{N}/\Sigma\}}(x_1) \, \mathrm{d}\Delta_{t,h,\bar{s}} \\ &\le \int_{Y \times Y} (\Lambda |x_2|^2 + \Sigma |r_2|^2) \mathbb{1}_{\{(x,r) : |x| \ge \sqrt{N} \text{ or } |r| \ge 2\sqrt{N}/\Sigma\}}(x_1, r_1) \, \mathrm{d}\tilde{\Delta}, \end{aligned}$$

(where $\tilde{\Delta}$ denotes a suitable limit measure of $((v_t(x_1), w_t(x_1)), (x_2, r_2))_\# \Delta_{t,h,\bar{s}})$ and an application of the Dominated Convergence Theorem then yields

$$\lim_{N\to\infty} \limsup_{h\to 0} \frac{1}{h^2} \int_{(\mathfrak{C}\setminus\mathfrak{C}_{t,1/N})\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1, r_1], [x_2, r_2])^2 \, \mathrm{d}\beta_{t,t+h} \;=\; 0,$$

which implies

$$\lim_{h\to 0} \frac{1}{h^2} \int_{(\mathfrak{C}\setminus\mathfrak{C}_{t,h})\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}([x_1, r_1], [x_2, r_2])^2 \, \mathrm{d}\beta_{t,t+h} \;=\; 0. \tag{3.35}$$

Next we consider $\frac{1}{h^2}\int_{\mathfrak{C}_{t,h}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}([x_1, r_1]), [x_2, r_2])^2 \, \mathrm{d}\beta_{t,t+h}$. According to ([2], Sect. 3.6) and ([9], Sect. 8.1), the geometric cone $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$ is a length space and it holds that any curve $\eta := [x, r] : [0, 1] \to \mathfrak{C}$ for C^1-functions $x : [0, 1] \to \mathbb{R}^d$ and $r : [0, 1] \to [0, +\infty)$ is absolutely continuous in $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma})$ and

$$\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\eta(1), \eta(0))^2 \le \int_0^1 \Big(\frac{4}{\Sigma}(r'(s))^2 + \frac{1}{\Lambda} r(s)^2 |x'(s)|^2\Big) \, \mathrm{d}s$$

(cf. ([9], Lem. 8.1)). We define, for $y_1 := [x_1, r_1] \in \mathfrak{C}_{t,h}$, $y_2 := [x_2, r_2] \in \mathfrak{C}$, with $|x_1 - x_2| \le \pi\sqrt{\Lambda/\Sigma}$ if $r_2 > 0$, an absolutely continuous curve $\mathcal{A}_{h,\Xi(y_1),y_2} : [0, 1] \to \mathfrak{C}$ connecting $\Xi(y_1) = [x_1 + \Lambda h v_t(x_1), r_1(1 + \Sigma h w_t(x_1)/2)]$ and y_2 by setting $\mathcal{A}_{h,\Xi(y_1),y_2} := [\mathcal{X}_{h,\Xi(y_1),y_2}, \mathcal{R}_{h,\Xi(y_1),y_2}]$,

$$\mathcal{X}_{h,\Xi(y_1),y_2}(s) := x_1 + \theta_{y_1,y_2}(s)(x_2 - x_1) + \Lambda(1-s)h v_t(x_1), \tag{3.36}$$

$$\mathcal{R}_{h,\Xi(y_1),y_2}(s) := \mathcal{R}_{y_1,y_2}(s)\Big(1 + \Sigma(1-s)h w_t(x_1)/2\Big) \tag{3.37}$$

(cf. (2.6)-(2.9), (2.20)). The functions $\mathcal{X}_{h,\Xi(y_1),y_2} : [0, 1] \to \mathbb{R}^d$ and $\mathcal{R}_{h,\Xi(y_1),y_2} : [0, 1] \to [0, +\infty)$ are continuously differentiable with

$$\begin{aligned}
&(\mathcal{R}'_{h,\Xi(y_1),y_2}(s))^2 \\
&\quad = \Big(\Sigma\mathcal{R}'_{y_1,y_2}(s)(1-s)h w_t(x_1)/2 + \mathcal{R}'_{y_1,y_2}(s) - \Sigma\mathcal{R}_{y_1,y_2}(s) h w_t(x_1)/2\Big)^2 \\
&\quad \le 2|h|\Sigma \, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(y_1, y_2)^2 + 2\Big(\mathcal{R}'_{y_1,y_2}(s) - \Sigma r_1 h w_t(x_1)/2\Big)^2
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{R}_{h,\Xi(y_1),y_2}(s)^2 |\mathcal{X}'_{h,\Xi(y_1),y_2}(s)|^2 \\
&\quad \le 4\mathcal{R}_{y_1,y_2}(s)^2 |\theta'_{y_1,y_2}(s)(x_2 - x_1) - \Lambda h v_t(x_1)|^2 \\
&\quad \le 8\Big(|\mathcal{R}_{y_1,y_2}(s)\theta'_{y_1,y_2}(s)(x_2 - x_1) - \Lambda r_1 h v_t(x_1)|^2 + \Lambda^2 |h| |\mathcal{R}_{y_1,y_2}(s) - r_1|^2\Big) \\
&\quad \le 8\Big(|\mathcal{R}_{y_1,y_2}(s)\theta'_{y_1,y_2}(s)(x_2 - x_1) - \Lambda r_1 h v_t(x_1)|^2 + \Lambda^2\Sigma|h|/4 \, \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(y_1, y_2)^2\Big),
\end{aligned}$$

where we have made use of (2.9) and the fact that $y_1 = [x_1, r_1] \in \mathfrak{C}_{t,h}$. It follows from the above estimations and an application of Fubini's Theorem that

$$
\begin{aligned}
&\frac{1}{h^2}\int_{\mathfrak{C}_{t,h}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}([x_1,r_1]),[x_2,r_2])^2 \,\mathrm{d}\beta_{t,t+h}\\
&\le \frac{1}{h^2}\int_{\mathfrak{C}_{t,h}\times\mathfrak{C}}\int_0^1 \Big(\frac{4}{\Sigma}(\mathcal{R}'_{h,\Xi(y_1),y_2}(s))^2 + \frac{1}{\Lambda}\mathcal{R}_{h,\Xi(y_1),y_2}(s)^2|\mathcal{X}'_{h,\Xi(y_1),y_2}(s)|^2\Big)\,\mathrm{d}s\,\mathrm{d}\beta_{t,t+h}\\
&\le \int_0^1\int_{Y\times Y}\Big(2\Sigma(r_2 - r_1 w_t(x_1))^2 + 8\Lambda|x_2 - r_1 v_t(x_1)|^2\Big)\,\mathrm{d}\Delta_{t,h,s}((x_1,r_1),(x_2,r_2))\,\mathrm{d}s\\
&\quad + C_{\Lambda,\Sigma}\frac{\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\mu_{t+h})^2}{|h|}
\end{aligned}
$$

with $C_{\Lambda,\Sigma}$ only depending on Λ and Σ. According to Def. 3.2(ii), there exists a sequence of functions $\zeta_n \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ $(n \in \mathbb{N})$ so that $((\nabla\zeta_n,\zeta_n))_{n\in\mathbb{N}}$ converges to (v_t, w_t) in $\mathrm{L}^2(\mu_t, \mathbb{R}^d \times \mathbb{R})$, which means

$$
\begin{aligned}
\lim_{n\to\infty}\int_{Y\times Y}&\Big(r_1^2(\zeta_n(x_1) - w_t(x_1))^2\\
&+r_1^2|\nabla\zeta_n(x_1) - v_t(x_1)|^2\Big)\,\mathrm{d}\Delta_{t,h,s}((x_1,r_1),(x_2,r_2)) = 0
\end{aligned}
\tag{3.38}
$$

uniformly in $h \in (-t, 1-t)$ and $s \in (0, 1)$. Moreover, Prop. 3.3 and (3.5) yield

$$
\begin{aligned}
\lim_{h\to 0}\int_{Y\times Y}&\Big(\Sigma(r_2 - r_1\zeta_n(x_1))^2 + \Lambda|x_2 - r_1\nabla\zeta_n(x_1)|^2\Big)\,\mathrm{d}\Delta_{t,h,s}\\
&= ||(v_t, w_t) - (\nabla\zeta_n,\zeta_n)||^2_{\mathrm{L}^2(\mu_t,\mathbb{R}^d\times\mathbb{R})}
\end{aligned}
\tag{3.39}
$$

for all $n \in \mathbb{N}$ and $s \in (0, 1)$. Combining (3.38) and (3.39) and the fact that the right-hand side of (3.39) converges to 0 as $n \to \infty$, we obtain

$$
\begin{aligned}
\limsup_{h\to 0}\int_{Y\times Y}&\Big(2\Sigma(r_2 - r_1 w_t(x_1))^2\\
&+8\Lambda|x_2 - r_1 v_t(x_1)|^2\Big)\,\mathrm{d}\Delta_{t,h,s}((x_1,r_1),(x_2,r_2)) \;=\; 0,
\end{aligned}
$$

for every $s \in (0, 1)$, and thus, by Fatou's lemma,

$$
\begin{aligned}
\limsup_{h\to 0}\int_0^1\int_{Y\times Y}&\Big(2\Sigma(r_2 - r_1 w_t(x_1))^2\\
&+8\Lambda|x_2 - r_1 v_t(x_1)|^2\Big)\,\mathrm{d}\Delta_{t,h,s}((x_1,r_1),(x_2,r_2))\,\mathrm{d}s \;=\; 0.
\end{aligned}
\tag{3.40}
$$

Finally, applying the above estimation of $\frac{1}{h^2}\int_{\mathfrak{C}_{t,h}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}([x_1,r_1]),[x_2,r_2])^2\,\mathrm{d}\beta_{t,t+h}$, (3.40) and Def. 3.2(i), we obtain

$$\lim_{h\to 0}\frac{1}{h^2}\int_{\mathfrak{C}_{t,h}\times\mathfrak{C}} \mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}([x_1,r_1]),[x_2,r_2])^2\,\mathrm{d}\beta_{t,t+h} \;=\; 0, \tag{3.41}$$

which completes the proof of Thm. 3.4. $\square$

4 Differentiability Results

This section finally treats the differentiability of the Hellinger–Kantorovich distance $\mathsf{HK}_{\Lambda,\Sigma}$ along absolutely continuous curves; the linearization result of Thm. 3.4 puts us in a position to precisely compute the corresponding derivatives.
We fix another absolutely continuous curve $(\nu_t)_{t\in[0,1]}$ in $(\mathcal{M}(\mathbb{R}^d),\mathsf{HK}_{\Lambda,\Sigma})$ with square-integrable metric derivative $t\mapsto|\nu_t'|$. It follows from (3.2) that

$$t\mapsto\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2 \tag{4.1}$$

is an absolutely continuous mapping from [0, 1] to $[0,+\infty)$ and thus $\mathscr{L}^1$-a.e. differentiable.

Let $(\bar{v},\bar{w}):(0,1)\times\mathbb{R}^d\to\mathbb{R}^d\times\mathbb{R}$ be the essentially unique Borel vector field associated with $(\nu_t)_t$ so that the continuity equation with reaction

$$\partial_t\nu_t=-\Lambda\mathrm{div}(\bar{v}_t\nu_t)+\Sigma\bar{w}_t\nu_t$$

holds good and

$$\int_{\mathbb{R}^d}(\Lambda|\bar{v}_t|^2+\Sigma|\bar{w}_t|^2)\,\mathrm{d}\nu_t\;=\;|\nu_t'|^2\quad\text{for }\mathscr{L}^1\text{-a.e. }t\in(0,1),$$

let $\mathcal{N}_\nu$ be the associated set of times defined according to Def. 3.2 and let $\mathcal{N}$ denote the set of times $t\in\mathcal{N}_\mu\cap\mathcal{N}_\nu$ at which (4.1) is differentiable. Clearly, $(0,1)\setminus\mathcal{N}$ is an $\mathscr{L}^1$-negligible set.

Theorem 4.1 *If $t\in\mathcal{N}$ and $\beta_t\in\mathcal{M}(\mathfrak{C}\times\mathfrak{C})$ is optimal in the definition of $\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2$ according to ((2.17), (2.13)), i.e.*

$$\begin{aligned}&\hat{\mu}_t:=\mu_t-\mathfrak{h}(\pi^1_\#\beta_t)\ge 0,\quad \hat{\nu}_t:=\nu_t-\mathfrak{h}(\pi^2_\#\beta_t)\ge 0,\\ &\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2\;=\;\int_{\mathfrak{C}\times\mathfrak{C}} d_{\mathfrak{C},\Lambda,\Sigma}([x_1,r_1],[x_2,r_2])^2\,\mathrm{d}\beta_t\\ &\qquad\qquad+4/\Sigma\;\hat{\mu}_t(\mathbb{R}^d)\;+\;4/\Sigma\;\hat{\nu}_t(\mathbb{R}^d),\end{aligned}$$

then the derivative $\frac{\mathrm{d}}{\mathrm{d}t}[\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2]$ *of* (4.1) *at* t *coincides with*

$$-\int_{\mathfrak{C}\times\mathfrak{C}}[\mathfrak{G}^{\Lambda,\Sigma}_{y_1}(\mathfrak{t}_{y_1,y_2}(0),\mathfrak{s}^{\mu}_{t,y_1})+\mathfrak{G}^{\Lambda,\Sigma}_{y_2}(\mathfrak{t}_{y_2,y_1}(0),\mathfrak{s}^{\nu}_{t,y_2})]\,\mathrm{d}\beta_t$$
$$+2\Big(\int_{\mathbb{R}^d} w_t\,\mathrm{d}\hat{\mu}_t+\int_{\mathbb{R}^d}\bar{w}_t\,\mathrm{d}\hat{\nu}_t\Big) \tag{4.2}$$

where $\mathfrak{s}^{\mu}_{t,y}:=(\Lambda v_t(x(y)),\,\Sigma/2\,r(y)w_t(x(y)))$ *and* $\mathfrak{s}^{\nu}_{t,y}:=(\Lambda\bar{v}_t(x(y)),\,\Sigma/2r(y)\bar{w}_t(x(y)))$.

Before proving Thm. 4.1, let us try to gain an insight into the above formula (4.2).

Remark 4.2 Suppose that $\nu_s\equiv\nu\in\mathcal{M}(\mathbb{R}^d)$. There exists an optimal plan β_t associated with μ_t and ν whose marginals satisfy $\mu_t=\mathfrak{h}(\pi^1_{\#}\beta_t)$ and $\nu=\mathfrak{h}(\pi^2_{\#}\beta_t)$ (cf. Thm. 7.6 in [9]). The derivative $\frac{\mathrm{d}}{\mathrm{d}t}[\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu)^2]$ at $t\in\mathcal{N}$ then takes the form

$$-\int_{\mathfrak{C}\times\mathfrak{C}}\mathfrak{G}^{\Lambda,\Sigma}_{y_1}(\mathfrak{t}_{y_1,y_2}(0),\mathfrak{s}^{\mu}_{t,y_1})\,\mathrm{d}\beta_t. \tag{4.3}$$

The tangent vectors $\mathfrak{t}_{y_1,y_2}(0)$ (see (2.12)) and $\mathfrak{s}^{\mu}_{t,y_1}$ to the geometric cone $\mathfrak{C}$, for $(y_1,y_2)\in\operatorname{supp}\beta_t$, represent the directions $\mu_t\rightsquigarrow\nu$ and $\mu_t\rightsquigarrow\mu_{t+h}$ (for $h>0$ small) respectively on an infinitesimal level (cf. Thm. 3.4). It is noteworthy that the metric tensor $\mathfrak{G}^{\Lambda,\Sigma}$ (see (2.5)) at $y_1\in\mathfrak{C}$ between such tangent vectors $\mathfrak{t}_{y_1,y_2}(0)$ and $\mathfrak{s}^{\mu}_{t,y_1}$ is equal to the derivative at $h=0$ of $h\mapsto -1/2\;\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}(y_1),y_2)^2$ (see (3.32)), i.e.

$$-\mathfrak{G}^{\Lambda,\Sigma}_{y_1}(\mathfrak{t}_{y_1,y_2}(0),\mathfrak{s}^{\mu}_{t,y_1}) \;=\; \frac{1}{2}\frac{\mathsf{d}}{\mathsf{d}h}\Big|_{h=0}\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{t,h}(y_1),y_2)^2 \tag{4.4}$$

for a simple computation shows that both terms in (4.4) equal

$$2r_1^2w_t(x_1)-2r_1r_2w_t(x_1)\cos(\sqrt{\Sigma/4\Lambda}|x_1-x_2|)$$
$$-2r_1r_2\sqrt{\Lambda/\Sigma}\left\langle\frac{\sin(\sqrt{\Sigma/4\Lambda}|x_1-x_2|)}{|x_1-x_2|}(x_2-x_1),v_t(x_1)\right\rangle$$

$(y_i=[x_i,r_i]\in\mathfrak{C})$.

Also, we would like to remark that the derivatives of (4.1) at $t\in\mathcal{N}$ can be expressed equally in terms of the Logarithmic Entropy-Transport characterization (1.1) of the Hellinger–Kantorovich distance $\mathsf{HK}_{\Lambda,\Sigma}$, by applying (2.19) to the above representation (4.2) of the derivatives.

Proof Let $t\in\mathcal{N}$. Then $t\in\mathcal{N}_\mu\cap\mathcal{N}_\nu$ and (4.1) is differentiable at t. We apply Thm. 3.4 to both curves $(\mu_s)_s$ and $(\nu_s)_s$ defining $\Xi_{\mu,t,h},\ \chi_{\mu,t,h}$ and $\Xi_{\nu,t,h},\ \chi_{\nu,t,h}$ respectively according thereto so that, by the corresponding linearization results,

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}\left[\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mu_s,\nu_s)^2\right]\bigg|_{s=t} \\
&= \lim_{h\to 0}\frac{\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mathfrak{h}\chi_{\mu,t,h},\mathfrak{h}\chi_{\nu,t,h})^2-\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2}{h}
\end{aligned}
\tag{4.5}
$$

(cf. (3.32), (3.33)). Let $\bar\chi_{\mu,t,h} := (\Xi_{\mu,t,h})_\#\alpha_{\mu,t}$ and $\bar\chi_{\nu,t,h} := (\Xi_{\nu,t,h})_\#\alpha_{\nu,t}$ be the push-forwards of the marginals $\alpha_{\mu,t} := \pi^1_\#\beta_t$ and $\alpha_{\nu,t} := \pi^2_\#\beta_t$ of β_t through the mappings $\Xi_{\mu,t,h}$ and $\Xi_{\nu,t,h}$ respectively. We have

$$
\begin{aligned}
\int_{\mathbb{R}^d}\phi\,\mathrm{d}(\mathfrak{h}\bar\chi_{\mu,t,h}) &= \int_{\mathfrak{C}_{\mu,t,h}}\mathsf{r}^2(1+\Sigma h w_t(\mathsf{x})/2)^2\phi(\mathsf{x}+\Lambda h v_t(\mathsf{x}))\,\mathrm{d}\alpha_{\mu,t} \\
&\quad+\int_{\mathfrak{C}\setminus\mathfrak{C}_{\mu,t,h}}\mathsf{r}^2\phi(\mathsf{x})\,\mathrm{d}\alpha_{\mu,t} \\
&= \int_{\mathsf{x}(\mathfrak{C}_{\mu,t,h})}(1+\Sigma h w_t(x)/2)^2\phi(x+\Lambda h v_t(x))\,\mathrm{d}\mathfrak{h}\alpha_{\mu,t} \\
&\quad+\int_{\mathsf{x}(\mathfrak{C}\setminus\mathfrak{C}_{\mu,t,h})}\phi(x)\,\mathrm{d}\mathfrak{h}\alpha_{\mu,t} \\
&\le \int_{\mathsf{x}(\mathfrak{C}_{\mu,t,h})}(1+\Sigma h w_t(x)/2)^2\phi(x+\Lambda h v_t(x))\,\mathrm{d}\mu_t \\
&\quad+\int_{\mathsf{x}(\mathfrak{C}\setminus\mathfrak{C}_{\mu,t,h})}\phi(x)\,\mathrm{d}\mu_t \;=\int_{\mathbb{R}^d}\phi\,\mathrm{d}(\mathfrak{h}\chi_{\mu,t,h})
\end{aligned}
$$

for all nonnegative bounded Borel functions $\phi:\mathbb{R}^d\to\mathbb{R}$ (cf. (2.14), (2.15)), from which we infer that

$$
\mathfrak{h}\bar\chi_{\mu,t,h}\;\le\mathfrak{h}\chi_{\mu,t,h},\,(\mathfrak{h}\chi_{\mu,t,h}-\mathfrak{h}\bar\chi_{\mu,t,h})(\mathbb{R}^d)=\hat\mu_t(\mathbb{R}^d)\;+\int_{\mathsf{x}(\mathfrak{C}_{\mu,t,h})}\Big(\Sigma h w_t(x)+\frac{\Sigma^2}{4}h^2 w_t(x)^2\Big)\,\mathrm{d}\hat\mu_t.
$$

Similarly,

$$
\begin{aligned}
&\mathfrak{h}\bar\chi_{\nu,t,h}\le\mathfrak{h}\chi_{\nu,t,h}, \\
(\mathfrak{h}\chi_{\nu,t,h}-\mathfrak{h}\bar\chi_{\nu,t,h})(\mathbb{R}^d)&=\hat\nu_t(\mathbb{R}^d)\;+\int_{\mathsf{x}(\mathfrak{C}_{\nu,t,h})}\Big(\Sigma h\bar w_t(x)+\frac{\Sigma^2}{4}h^2\bar w_t(x)^2\Big)\,\mathrm{d}\hat\nu_t.
\end{aligned}
$$

We obtain

$$
\begin{aligned}
&\frac{1}{2}\Big(\mathsf{HK}_{\Lambda,\Sigma}(\mathfrak{h}\chi_{\mu,t,h},\mathfrak{h}\chi_{\nu,t,h})^2-\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2\Big) \\
&\le\frac{1}{2}\Big(\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\bar\chi_{\mu,t,h},\bar\chi_{\nu,t,h})^2-\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_{\mu,t},\alpha_{\nu,t})^2\Big) \\
&\quad+2\int_{\mathsf{x}(\mathfrak{C}_{\mu,t,h})}\Big(h w_t(x)+\frac{\Sigma}{4}h^2 w_t(x)^2\Big)\,\mathrm{d}\hat\mu_t
\end{aligned}
$$

$$+2\int_{\mathsf{x}(\mathfrak{C}_{\nu,t,h})}\Big(h\bar{w}_t(x)+\frac{\Sigma}{4}h^2\bar{w}_t(x)^2\Big)\,\mathrm{d}\hat{\nu}_t,$$

and

$$\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\bar{\chi}_{\mu,t,h},\bar{\chi}_{\nu,t,h})^2\ \le\ \int_{\mathfrak{C}\times\mathfrak{C}}\mathsf{d}_{\mathfrak{C},\Lambda,\Sigma}(\Xi_{\mu,t,h}([x_1,r_1]),\Xi_{\nu,t,h}([x_2,r_2]))^2\,\mathrm{d}\beta_t.$$

The same argument as in the proof of Lem. 2.2 in [5] then yields

$$\begin{aligned}
&\limsup_{h\downarrow 0}\frac{\frac{1}{2}\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\bar{\chi}_{\mu,t,h},\bar{\chi}_{\nu,t,h})^2-\frac{1}{2}\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_{\mu,t},\alpha_{\nu,t})^2}{h}\\
&\le 2\int_{\mathfrak{C}\times\mathfrak{C}}\Big[r_1^2 w_t(x_1)-r_1r_2w_t(x_1)\cos(\sqrt{\Sigma/4\Lambda}|x_1-x_2|)\\
&-r_1r_2\sqrt{\Lambda/\Sigma}\,\langle S_{\Lambda,\Sigma}(x_1,x_2),v_t(x_1)\rangle\Big]\,\mathrm{d}\beta_t\\
&+2\int_{\mathfrak{C}\times\mathfrak{C}}\Big[r_2^2\bar{w}_t(x_2)-r_1r_2\bar{w}_t(x_2)\cos(\sqrt{\Sigma/4\Lambda}|x_1-x_2|)\\
&+r_1r_2\sqrt{\Lambda/\Sigma}\,\langle S_{\Lambda,\Sigma}(x_1,x_2),\bar{v}_t(x_2)\rangle\Big]\,\mathrm{d}\beta_t\\
&\le\liminf_{h\uparrow 0}\frac{\frac{1}{2}\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\bar{\chi}_{\mu,t,h},\bar{\chi}_{\nu,t,h})^2-\frac{1}{2}\mathcal{W}_{\mathfrak{C},\Lambda,\Sigma}(\alpha_{\mu,t},\alpha_{\nu,t})^2}{h},
\end{aligned}$$

with $S_{\Lambda,\Sigma}$ defined as

$$S_{\Lambda,\Sigma}(x_1,x_2):=\begin{cases}\frac{\sin(\sqrt{\Sigma/4\Lambda}|x_1-x_2|)}{|x_1-x_2|}(x_2-x_1) & \text{if } x_1\neq x_2,\\ 0 & \text{if } x_1=x_2.\end{cases}$$

Since the limit (4.5) exists, the sum of the above integrands is identical with

$$-\mathfrak{G}^{\Lambda,\Sigma}_{y_1}(\mathfrak{t}_{y_1,y_2}(0),\mathfrak{s}^{\mu}_{t,y_1})-\mathfrak{G}^{\Lambda,\Sigma}_{y_2}(\mathfrak{t}_{y_2,y_1}(0),\mathfrak{s}^{\nu}_{t,y_2})\qquad(y_i:=[x_i,r_i])$$

(cf. Rem. 4.2), and

$$\begin{aligned}
\lim_{h\to 0}\int_{\mathsf{x}(\mathfrak{C}_{\mu,t,h})}\Big(w_t(x)+\frac{\Sigma}{4}hw_t(x)^2\Big)\,\mathrm{d}\hat{\mu}_t&=\int_{\mathbb{R}^d}w_t(x)\,\mathrm{d}\hat{\mu}_t,\\
\lim_{h\to 0}\int_{\mathsf{x}(\mathfrak{C}_{\nu,t,h})}\Big(\bar{w}_t(x)+\frac{\Sigma}{4}h\bar{w}_t(x)^2\Big)\,\mathrm{d}\hat{\nu}_t&=\int_{\mathbb{R}^d}\bar{w}_t(x)\,\mathrm{d}\hat{\nu}_t,
\end{aligned}$$

it follows from the above computations that

$$
\begin{aligned}
\lim_{h\to 0} & \frac{\frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mathfrak{h}\chi_{\mu,t,h},\mathfrak{h}\chi_{\nu,t,h})^2 - \frac{1}{2}\mathsf{HK}_{\Lambda,\Sigma}(\mu_t,\nu_t)^2}{h} \\
&= -\int_{\mathfrak{C}\times\mathfrak{C}} [\mathfrak{G}^{\Lambda,\Sigma}_{y_1}(\mathfrak{t}_{y_1,y_2}(0),\mathfrak{s}^{\mu}_{t,y_1}) + \mathfrak{G}^{\Lambda,\Sigma}_{y_2}(\mathfrak{t}_{y_2,y_1}(0),\mathfrak{s}^{\nu}_{t,y_2})]\,\mathrm{d}\beta_t \\
&\quad + 2\Big(\int_{\mathbb{R}^d} w_t\,\mathrm{d}\hat{\mu}_t + \int_{\mathbb{R}^d} \bar{w}_t\,\mathrm{d}\hat{\nu}_t\Big).
\end{aligned}
$$

The proof of Thm. 4.1 is complete. □

Acknowledgements I gratefully acknowledge support from the Erwin Schrödinger International Institute for Mathematics and Physics (Vienna) during my participation in the programme “Optimal Transport”.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics. ETH Zürich, Birkhäuser (2005)
2. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry. Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence, RI (2001)
3. Chizat, L., Peyré, G., Schmitzer, B., Vialard, F.-X.: An interpolating distance between optimal transport and Fisher-Rao metrics. Found. Comput. Math. **18**, 1–44 (2018)
4. Chizat, L., Peyré, G., Schmitzer, B., Vialard, F.-X.: Unbalanced optimal transport: dynamic and Kantorovich formulations. J. Funct. Anal. **274**, 3090–3123 (2018)
5. Fleißner, F.: A Minimizing Movement approach to a class of scalar reaction-diffusion equations, ESAIM:COCV 27 (2021) 18. https://arXiv.org/2002.04496 (2020)
6. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**, 113–161 (1996)
7. Kondratyev, S., Monsaingeon, L., Vorotnikov, D., et al.: A new optimal transport distance on the space of finite Radon measures. Adv. Differ. Equ. **21**, 1117–1164 (2016)
8. Liero, M., Mielke, A., Savaré, G.: Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. SIAM J. Math. Anal. **48**, 2869–2911 (2016)
9. Liero, M., Mielke, A., Savaré, G.: Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. Invent. Math. **211**, 969–1117 (2018)
10. Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence, RI (2003)