One dimensional martingale rearrangement couplings

B. Jourdain* W. Margheriti*

Abstract

We are interested in martingale rearrangement couplings. As introduced by Wiesel [39] in order to prove the stability of Martingale Optimal Transport problems, these are projections in adapted Wasserstein distance of couplings between two probability measures on the real line in the convex order onto the set of martingale couplings between these two marginals. In reason of the lack of relative compactness of the set of couplings with given marginals for the adapted Wasserstein topology, the existence of such a projection is not clear at all. Under a barycentre dispersion assumption on the original coupling which is in particular satisfied by the Hoeffding-Fréchet or comonotone coupling, Wiesel gives a clear algorithmic construction of a martingale rearrangement when the marginals are finitely supported and then gets rid of the finite support assumption by relying on a rather messy limiting procedure to overcome the lack of relative compactness. Here, we give a direct general construction of a martingale rearrangement coupling under the barycentre dispersion assumption. This martingale rearrangement is obtained from the original coupling by an approach similar to the construction we gave in [26] of the inverse transform martingale coupling, a member of a family of martingale couplings close to the Hoeffding-Fréchet coupling, but for a slightly different injection in the set of extended couplings introduced by Beiglböck and Juillet [11] and which involve the uniform distribution on $[0,1]$ in addition to the two marginals. We last discuss the stability in adapted Wassertein distance of the inverse transform martingale coupling with respect to the marginal distributions.

Keywords: Martingale couplings, Martingale Optimal Transport, Adapted Wasserstein distance, Robust finance, Convex order.

1 Introduction

Let $\rho \geq 1$ and μ, ν be in the set $\mathcal{P}_\rho(\mathbb{R})$ of probability measures on the real line with finite order ρ moment. We denote by $\Pi(\mu,\nu)$ the set of couplings between μ and ν, that is $\pi \in \Pi(\mu,\nu)$ iff π is a measure on $\mathbb{R}\times\mathbb{R}$ with first marginal μ and second marginal ν. We denote by $\Pi^{\mathrm{M}}(\mu,\nu)$ the set of martingale couplings between μ and ν:

$$\Pi^{\mathrm{M}}(\mu,\nu) = \left\{ M \in \Pi(\mu,\nu) \mid \mu(dx)\text{-a.e., } \int_{\mathbb{R}} y\, M_x(dy) = x \right\}, \tag{1.1}$$

where for any coupling $\pi \in \Pi(\mu,\nu)$ we denote by $(\pi_x)_{x\in\mathbb{R}}$ its disintegration with respect to its first marginal, that is $\pi(dx,dy) = \mu(dx)\,\pi_x(dy)$, or with a slight abuse of notation $\pi = \mu \times \pi_x$. The celebrated Strassen theorem [38] ensures that $\Pi^{\mathrm{M}}(\mu,\nu) \neq \emptyset$ iff μ and ν are in the convex order, which we denote $\mu \leq_{cx} \nu$, that is iff $\int_{\mathbb{R}} f(x)\,\mu(dx) \leq \int_{\mathbb{R}} f(y)\,\nu(dy)$ for any convex function $f : \mathbb{R}\to\mathbb{R}$, which implies that $\int_{\mathbb{R}} x\mu(dx) = \int_{\mathbb{R}} y\nu(dy)$.

Fix $\pi \in \Pi(\mu,\nu)$ with $\mu \leq_{cx} \nu$. We are interested in finding a projection of π on the set $\Pi^{\mathrm{M}}(\mu,\nu)$ for the adapted Wasserstein distance $\mathcal{AW}_\rho$ (defined in (1.5) below), that is finding a martingale coupling M between μ and ν such that

$$\mathcal{AW}_\rho(\pi, M) = \inf_{M' \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_\rho(\pi, M'). \tag{1.2}$$

*CERMICS, Ecole des Ponts, INRIA, Marne-la-Vallée, France. E-mails: benjamin.jourdain@enpc.fr, william.margheriti@enpc.fr - This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque.

This problem arose our interest when Wiesel [39] highlighted its connection for $\rho = 1$ with the stability of the Martingale Optimal Transport (MOT) problem. The MOT problem was introduced in discrete time by Beiglböck, Henry-Labordère and Penkner [7] and in continuous time by Galichon, Henry- Labordère and Touzi [19] in order to get model-free bounds of an option price. It consists in the classical Optimal Transport problem, which was formulated by Gaspard Monge [30] in 1781 and modernised by Kantorovich [28] in 1942, to which an additional martingale constraint is added in order to reflect the arbitrage-free condition of the market. In our setting the MOT problem consists in the minimisation

$$\mathrm{MOT}(\mu,\nu) := \inf_{M\in\Pi^{\mathrm{M}}(\mu,\nu)} \int_{\mathbb{R}\times\mathbb{R}} C(x,y)\, M(dx,dy), \tag{MOT}$$

where $C : \mathbb{R}\times\mathbb{R} \to \mathbb{R}_+$ is a nonnegative measurable payoff function. The study of its stability, that is the continuity of the map $(\mu,\nu) \mapsto \mathrm{MOT}(\mu,\nu)$, represents a major stake, since it confirms the robustness of model-free bounds of an option price. Backhoff-Veraguas and Pammer [6] gave a positive answer under mild regularity assumptions by showing the stability of the so called martingale C-monotonicity property, which is proved sufficient for optimality. Independently, Wiesel [39] also gave a positive answer. More recently, Beiglböck, Pammer and the two authors generalised those stability results to the weak MOT problem [8]. For adaptations of celebrated results on classical optimal transport theory to the MOT problem, we refer to Beiglböck and Juillet [10], Henry-Labordère, Tan and Touzi [24] and Henry- Labordère and Touzi [25]. On duality, we refer to Beiglböck, Nutz and Touzi [13], Beiglböck, Lim and Obłój [12] and De March [17]. We also refer to Ghoussoub, Kim and Lim [22], De March [16] and De March and Touzi [18] for the multi-dimensional case, where stability fails according to a nice counter-example by Brückerhoff and Juillet [15].

We recall that the Wasserstein distance with index ρ between μ and ν is defined by

$$\mathcal{W}_\rho(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{\mathbb{R}\times\mathbb{R}} |x-y|^\rho \, \pi(dx,dy) \right)^{1/\rho}. \tag{1.3}$$

The infimum is attained by the comonotonic or Hoeffding-Fréchet coupling π^{HF} between μ and ν, that is the image of the Lebesgue measure on $(0,1)$ by $u \mapsto (F_\mu^{-1}(u), F_\nu^{-1}(u))$, where $F_\eta^{-1}(u) = \inf\{x \in \mathbb{R} : \eta((-\infty,x]) \geq u\}$ denotes the quantile function of a probability measure η on $\mathbb{R}$. As a consequence,

$$\mathcal{W}_\rho(\mu,\nu) = \left(\int_{(0,1)} |F_\mu^{-1}(u) - F_\nu^{-1}(u)|^\rho du \right)^{1/\rho}. \tag{1.4}$$

The topology induced by the Wasserstein distance is not always well suited for any setting, especially in mathematical finance. Indeed, the symmetry of this distance does not take into account the temporal structure of martingales. One can easily get convinced that two stochastic processes very close in Wasserstein distance can yield radically unalike information, as [3, Figure 1] illustrates very well. Therefore, one needs to strengthen, or adapt this usual topology. This can be done in many different ways, such as the adapted weak topology (see below), Hellwig's information topology [23], Aldous's extended weak topology [1] or the optimal stopping topology [4]. Strikingly, all those apparently independent topologies are actually equal, at least in discrete time [4, Theorem 1.1].

Hence it induces no loss of generality to focus on the so called adapted Wasserstein distance. For an extensive background, we refer to [31, 32, 33, 34, 29, 14]. For all $\mu',\nu' \in \mathcal{P}_\rho(\mathbb{R})$ and $\pi' \in \Pi(\mu',\nu')$, the adapted Wasserstein distance with index ρ between π and π' is defined by

$$\mathcal{AW}_\rho(\pi,\pi') = \inf_{\chi\in\Pi(\mu,\mu')} \left(\int_{\mathbb{R}\times\mathbb{R}} \left(|x-x'|^\rho + \mathcal{W}_\rho^\rho(\pi_x,\pi'_{x'}) \right) \chi(dx,dx') \right)^{1/\rho}. \tag{1.5}$$

Note that by Lemma 5.1 below there always exists a coupling $\chi \in \Pi(\mu,\mu)$ optimal for $\mathcal{AW}_\rho(\pi,\pi')$. Moreover it is easy to check that $\mathcal{W}_\rho \leq \mathcal{AW}_\rho$, so that $\mathcal{AW}_\rho$ induces a finer topology than $\mathcal{W}_\rho$. For $\pi \in \Pi(\mu,\nu)$ with $\mu \leq_{cx} \nu$, Wiesel [39] studies Problem (1.2) for $\rho = 1$ and introduces the notion of martingale rearrangement: a martingale coupling $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ is called a martingale rearrangement coupling of π if

$$\mathcal{AW}_1(\pi, M) = \inf_{M' \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_1(\pi, M'). \tag{1.6}$$

Actually, he works with the nested Wasserstein distance, which according to [5, (1.3)] is equal to the adapted Wasserstein distance. In the present paper, even if we mainly concentrate on martingale rearrangements, we will also consider a slight extension of the latter definition: a martingale coupling $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ is called an $\mathcal{AW}_\rho$-minimal martingale rearrangement coupling of π if

$$\mathcal{AW}_\rho(\pi, M) = \inf_{M' \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_\rho(\pi, M'). \tag{1.7}$$

Note that the existence of an $\mathcal{AW}_\rho$-minimal martingale rearrangement coupling is not clear in the general case. Indeed, let $(M_n)_{n\in\mathbb{N}}$ be a sequence of martingale couplings between μ and ν such that $(\mathcal{AW}_\rho(\pi, M_n))_{n\in\mathbb{N}}$ converges to $\mathcal{AW}_\rho(\pi, M)$. The tightness of the marginals μ and ν guarantees tightness and therefore relative compactness of $(M_n)_{n\in\mathbb{N}}$ for the $\mathcal{W}_\rho$-distance, but not necessarily for the $\mathcal{AW}_\rho$-distance. In order to compensate this lack of relative compactness, Wiesel [39] introduces a new assumption: the coupling $\pi \in \Pi(\mu,\nu)$ is said to satisfy the barycentre dispersion assumption iff

$$\forall a \in \mathbb{R}, \quad \int_{\mathbb{R}} \mathbb{1}_{[a,+\infty)}(x) \left(x - \int_{\mathbb{R}} y \, \pi_x(dy) \right) \mu(dx) \leq 0. \tag{1.8}$$

The latter assumption is important in this context since it provides a sufficient condition for a coupling π between μ and ν to admit a martingale rearrangement coupling. More precisely, Wiesel shows [39, Lemma 2.1] that in the general case,

$$\inf_{M' \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_1(\pi, M') \geq \int_{\mathbb{R}} \left| \int_{\mathbb{R}} y \, \pi_x(dy) - x \right| \mu(dx), \tag{1.9}$$

and there exists $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ such that $\mathcal{AW}_1(\pi, M) = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} y \, \pi_x(dy) - x \right| \mu(dx)$ when π satisfies the barycentre dispersion assumption (1.8) [39, Proposition 2.4].

The problem (1.2) was in a certain way already considered by Rüschendorf [37], who looked for a projection of a probability measure on a set of probability measures with given linear constraints. Since the martingale constraint is linear, his study encompasses our problem. Yet he considered the projection with respect to the Kullback-Leibler distance, also known as relative entropy, in place of $\mathcal{AW}_\rho$ and this does not suit our purpose. More recently, Gerhold and Gülüm looked at a very similar problem [20, Problem 2.4] for the infinity Wasserstein distance, the Prokhorov distance, the stop-loss distance, the Lévy distance or modified versions of them. Once again, despite being of great interest in their setting, in particular for their application to the existence of a market model which is consistent with a finite set of European call options prices on a single underlying asset [21], their choice of distance is still inadequate for a connection with the stability of (MOT).

In Section 2 we briefly recall Wiesel's construction [39] of a martingale rearrangement coupling of any coupling π which satisfies the barycentre dispersion assumption (1.8). Then we design our own construction of a martingale rearrangement coupling of π. This construction is actually done through lifted couplings, in the sense of Beiglböck and Juillet [11], that is probability measures on the enlarged space $(0,1) \times \mathbb{R} \times \mathbb{R}$ in which the spatial domain $\mathbb{R} \times \mathbb{R}$ of regular couplings is embedded.

Our construction in Section 2 is highly inspired of the one we did in [26], where we designed the martingale inverse transform coupling M^{IT} between μ and ν as a special element of a family $(M^Q)_{Q\in\mathcal{Q}}$ of martingale couplings between μ and ν such that $\mu \leq_{cx} \nu$ parametrised by a set $\mathcal{Q}$ of probability measures on $(0,1)^2$. This family was meant to be as close as possible to the Hoeffding-Fréchet coupling π^{HF} between μ and ν. As proved by Wiesel [39, Lemma 2.3], π^{HF} satisfies the barycentre dispersion assumption (1.8). Section 3 is specialised to martingale rearrangements of the Hoeffding-Fréchet coupling π^{HF}. After presenting the family $(M^Q)_{Q\in\mathcal{Q}}$ and the martingale inverse transform coupling M^{IT}, we show that the lifted coupling associated with any element of $(M^Q)_{Q\in\mathcal{Q}}$ is, in a very natural sense, a lifted martingale rearrangement coupling of a lift of π^{HF}. At the level of regular couplings on $\mathbb{R}\times\mathbb{R}$, we can conclude the same as soon as the sign of $F_\nu^{-1}-F_\mu^{-1}$ is constant on the jumps of F_μ, which holds when $F_\nu^{-1}-F_\mu^{-1}$ is constant on these jumps and π^{HF} is concentrated on the graph of the Monge transport map $T=F_\nu^{-1}\circ F_\mu$. When this condition is not met, the inverse transform martingale coupling M^{IT} may fail to be a martingale rearrangement of π^{HF} as we show in Example 3.5.

We finally show in Section 4 the stability of the inverse transform martingale coupling for the $\mathcal{AW}_\rho$-distance with respect to its marginals. The latter stability holds in full generality at the lifted level but a condition on the first marginals is needed at the level of regular couplings.

Let us now recall some standard results about cumulative distribution functions and quantile functions since they will prove very handy one-dimensional tools. Proofs can be found for instance in [26, Appendix]. For any probability measure η on $\mathbb{R}$, denoting by $F_\eta(x)=\eta((-\infty,x])$ and $F_\eta^{-1}(u)=\inf\{x\in\mathbb{R}:F_\eta(x)\geq u\}$ the cumulative distribution function and the quantile function of η, we have

(1) F_η, resp. F_η^{-1}, is right continuous, resp. left continuous, and nondecreasing;

(2) For all $(x,u)\in\mathbb{R}\times(0,1)$,
$$F_\eta^{-1}(u)\leq x \iff u\leq F_\eta(x), \tag{1.10}$$
which implies
$$F_\eta(x-)<u\leq F_\eta(x) \implies x=F_\eta^{-1}(u), \tag{1.11}$$
$$\text{and}\quad F_\eta(F_\eta^{-1}(u)-)\leq u\leq F_\eta(F_\eta^{-1}(u)); \tag{1.12}$$

(3) For $\eta(dx)$-almost every $x\in\mathbb{R}$,
$$0<F_\eta(x),\quad F_\eta(x-)<1\quad\text{and}\quad F_\eta^{-1}(F_\eta(x))=x; \tag{1.13}$$

(4) The image of the Lebesgue measure on $(0,1)$ by F_η^{-1} is η. This property is referred to as inverse transform sampling.

(5) Denoting by $\lambda_{(0,1)}$, resp. $\lambda_{(0,1)^2}$, the Lebesgue measure on $(0,1)$, resp. $(0,1)^2$ and setting
$$\theta(x,v)=F_\mu(x-)+v\mu(\{x\})\quad\text{for}\quad(x,v)\in\mathbb{R}\times[0,1], \tag{1.14}$$
we have
$$\left((u,v)\mapsto\theta(F_\mu^{-1}(u),v)\right)_\sharp\lambda_{(0,1)^2}=\lambda_{(0,1)}, \tag{1.15}$$
where $\sharp$ denotes the pushforward operation. Coupled with the inverse transform sampling we also have the equivalent formulation
$$\theta_\sharp(\mu\times\lambda_{(0,1)})=\lambda_{(0,1)}. \tag{1.16}$$

2 Martingale rearrangements of couplings which satisfy the barycentre dispersion assumption

2.1 Regular and lifted martingale rearrangement couplings

By (1.4) for the first equality and the inverse transform sampling for the second one, we have for $\eta,\eta'\in\mathcal{P}_1(\mathbb{R})$,

$$\mathcal{W}_1(\eta,\eta')=\int_{(0,1)}\left|F_\eta^{-1}(u)-F_{\eta'}^{-1}(u)\right|du\geq\left|\int_{(0,1)}F_\eta^{-1}(u)du-\int_{(0,1)}F_{\eta'}^{-1}(u)du\right|=\left|\int_\mathbb{R}x\,\eta(dx)-\int_\mathbb{R}x\,\eta'(dx)\right|. \tag{2.1}$$

The inequality is an equality iff either $\forall u \in (0,1)$, $F_\eta^{-1}(u) \le F_{\eta'}^{-1}(u)$ i.e. η is smaller than η' for the stochastic order which we denote $\eta \le_{st} \eta'$ or $\forall u \in (0,1)$, $F_\eta^{-1}(u) \ge F_{\eta'}^{-1}(u)$ i.e. $\eta \ge_{st} \eta'$.

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ such that $\mu \le_{cx} \nu$. We are now ready to reproduce the proof of [39, Lemma 2.1] to check (1.9). For $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ and $\chi \in \Pi(\mu,\mu)$ we have, using (2.1) then the triangle inequality,

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}} (|x-x'| + \mathcal{W}_1(\pi_x, M_{x'}))\, \chi(dx,dx') &\ge \int_{\mathbb{R}\times\mathbb{R}} \left(|x-x'| + \left|\int_{\mathbb{R}} y\, \pi_x(dy) - x'\right|\right) \chi(dx,dx') \\ &\ge \int_{\mathbb{R}} \left|\int_{\mathbb{R}} y\, \pi_x(dy) - x\right| \mu(dx).\end{aligned} \tag{2.2}$$

When π satisfies the barycentre dispersion assumption (1.8), finding a martingale rearrangement coupling of π amounts to find a martingale coupling such that the inequalities in (2.2) are equalities. This observation leads to the following lemma.

Lemma 2.1. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \le_{cx} \nu$ and $\pi \in \Pi(\mu,\nu)$ satisfy the barycentre dispersion assumption* (1.8). *Then $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ is a martingale rearrangement coupling of π iff there exists $\chi \in \Pi(\mu,\mu)$ such that $\chi(dx,dx')$-almost everywhere,*

$$x < x' \implies \pi_x \ge_{st} M_{x'}, \quad x > x' \implies \pi_x \le_{st} M_{x'} \quad \text{and} \quad x = x' \implies \pi_x \le_{st} M_x \text{ or } \pi_x \ge_{st} M_x, \tag{2.3}$$

in which case χ is optimal for $\mathcal{AW}_1(\pi, M)$.

Proof. Suppose that M is a martingale rearrangement coupling of π and χ is optimal for $\mathcal{AW}_1(\pi, M)$. Since π satisfies the barycentre dispersion assumption, we know by [39, Proposition 2.4] that $\mathcal{AW}_1(M,\pi) = \int_{\mathbb{R}} \left|\int_{\mathbb{R}} y\, \pi_x(dy) - x\right| \mu(dx)$. Then the first inequality in (2.2) is an equality, hence $\chi(dx,dx')$-almost everywhere, $\mathcal{W}_1(\pi_x, M_{x'}) = \left|\int_{\mathbb{R}} y\, \pi_x(dy) - x'\right|$, or equivalently π_x and $M_{x'}$ are comparable in the stochastic order. Morever the second inequality in (2.2) is an equality as well, hence $\chi(dx,dx')$-almost everywhere, x' lies between x and $\int_{\mathbb{R}} y\, \pi_x(dy)$. We deduce that $\chi(dx,dx')$-almost everywhere,

$$(x-x')\left(x' - \int_{\mathbb{R}} y\, \pi_x(dy)\right) \ge 0, \tag{2.4}$$

and $\pi_x \le_{st} M_{x'}$ or $\pi_x \ge_{st} M_{x'}$. Then (2.3) is easily deduced from the fact that the map $\eta \mapsto \int_{\mathbb{R}} z\, \eta(dz)$ is increasing for the stochastic order.

Conversely, suppose that (2.3) and therefore (2.4) holds for some $\chi \in \Pi(\mu,\mu)$. Then the inequalities in (2.2) are equalities, hence χ is optimal for $\mathcal{AW}_1(\pi, M)$ and M is a martingale rearrangement coupling of π. □

To construct a martingale rearrangement coupling of π satisfying the barycenter dispersion assumption (1.8), we will define a probability kernel $(m_u)_{u\in(0,1)}$ such that $\int_{\mathbb{R}} y\, m_u(dy) = F_\mu^{-1}(u)$ du-a.e. and deduce that the probability measure

$$M(dx,dy) = \int_0^1 \delta_{F_\mu^{-1}(u)}(dx)\, m_u(dy)\, du \tag{2.5}$$

is a martingale coupling between μ and ν. Yet the probability kernel $(m_u)_{u\in(0,1)}$ is not uniquely determined from the knowledge of M. Hence the definition (2.5) induces a loss of information. In order to keep this information, one can consider like Beiglböck and Juillet [11] instead of M its lifted martingale coupling

$$\widehat{M}(du,dx,dy) = \lambda_{(0,1)}(du)\, \delta_{F_\mu^{-1}(u)}(dx)\, m_u(dy) \in \Pi(\lambda_{(0,1)}, \mu, \nu), \tag{2.6}$$

where $\lambda_{(0,1)}$ denotes the Lebesgue measure on $(0,1)$. In the present paper, we only use the quantile coupling $\lambda_{(0,1)}(du)\, \delta_{F_\mu^{-1}(u)}(dx)$ between $\lambda_{(0,1)}(du)$ and $\mu(dx)$ whereas other couplings and in particular the independent one are also considered in [11]. More generally, for any $\pi \in \Pi(\mu,\nu)$, we call lifted coupling of π any coupling $\widehat{\pi} \in \Pi(\lambda_{(0,1)}, \mu, \nu)$ such that there exists a probability kernel $(p_u)_{u\in(0,1)}$ which satisfies

$$\widehat{\pi}(du, dx, dy) = \lambda_{(0,1)}(du)\, \delta_{F_\mu^{-1}(u)}(dx)\, p_u(dy) \quad \text{and} \quad \int_{u\in(0,1)} \widehat{\pi}(du, dx, dy) = \pi(dx, dy).$$

We denote by $\widehat{\Pi}(\mu, \nu)$ the set of all lifted couplings between μ and ν. Notice that there exists an easy embedding

$$\iota : \Pi(\mu, \nu) \to \widehat{\Pi}(\mu, \nu), \quad \pi \mapsto \lambda_{(0,1)}(du)\, \delta_{F_\mu^{-1}(u)}(dx)\, \pi_{F_\mu^{-1}(u)}(dy). \tag{2.7}$$

For $\widehat{\pi} = \lambda_{(0,1)} \times \widehat{\pi}_u = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times p_u$ and $\widehat{\pi}' = \lambda_{(0,1)} \times \widehat{\pi}'_u = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times p'_u$ two lifted couplings of $\pi \in \Pi(\mu, \nu)$ and $\pi' \in \Pi(\mu', \nu')$, we define their lifted adapted Wasserstein distance of order ρ by

$$\begin{aligned}
\widehat{\mathcal{AW}}_\rho(\pi, \pi') &= \inf_{\chi\in\Pi(\lambda_{(0,1)},\lambda_{(0,1)})} \left(\int_{(0,1)\times(0,1)} \left(|u-u'|^\rho + \mathcal{AW}_\rho^\rho(\widehat{\pi}_u, \widehat{\pi}'_{u'}) \right) \chi(du, du') \right)^{1/\rho} \\
&= \inf_{\chi\in\Pi(\lambda_{(0,1)},\lambda_{(0,1)})} \left(\int_{(0,1)\times(0,1)} \left(|u-u'|^\rho + |F_\mu^{-1}(u) - F_{\mu'}^{-1}(u')|^\rho + \mathcal{W}_\rho^\rho(p_u, p'_{u'}) \right) \chi(du, du') \right)^{1/\rho}.
\end{aligned}$$

Note that by Remark 5.2 below there always exists a coupling $\chi \in \Pi(\lambda_{(0,1)}, \lambda_{(0,1)})$ optimal for $\widehat{\mathcal{AW}}_\rho(\widehat{\pi}, \widehat{\pi}')$. We denote by $\widehat{\Pi}^{\mathrm{M}}(\mu, \nu)$ the set of all lifted martingale couplings between μ and ν, that is the set of all lifted couplings $\lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times m_u \in \widehat{\Pi}(\mu, \nu)$ such that $\int_{\mathbb{R}} y\, m_u(dy) = F_\mu^{-1}(u)$ for du-almost all $u \in (0,1)$. For $\rho \geq 1$, we then call lifted $\widehat{\mathcal{AW}}_\rho$-minimal martingale rearrangement coupling (or simply lifted martingale rearrangement coupling when $\rho = 1$) of $\widehat{\pi} \in \widehat{\Pi}(\mu, \nu)$ any lifted martingale coupling $\widehat{M} \in \widehat{\Pi}^{\mathrm{M}}(\mu, \nu)$ such that

$$\widehat{\mathcal{AW}}_\rho(\widehat{\pi}, \widehat{M}) = \inf_{\widehat{M}'\in\widehat{\Pi}^{\mathrm{M}}(\mu,\nu)} \widehat{\mathcal{AW}}_\rho(\widehat{\pi}, \widehat{M}').$$

Ignoring the non-negative contribution of $|u - u'|$ in the definition of $\widehat{\mathcal{AW}}_1$ and reasoning like in (2.2), we easily check the following lower bound analogous, at the lifted level, to (1.9).

Lemma 2.2. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$. Then for all $\widehat{\pi} = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times p_u \in \widehat{\Pi}(\mu, \nu)$,*

$$\inf_{\widehat{M}\in\widehat{\Pi}^{\mathrm{M}}(\mu,\nu)} \widehat{\mathcal{AW}}_1(\widehat{\pi}, \widehat{M}) \geq \int_{(0,1)} \left| \int_{\mathbb{R}} y\, p_u(dy) - F_\mu^{-1}(u) \right| du.$$

The next proposition gives a sufficient condition for the collapse through (2.5) of a lifted martingale coupling to be a martingale rearrangement.

Proposition 2.3. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$. Let $\widehat{\pi} = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times p_u \in \widehat{\Pi}(\mu, \nu)$ be such that $u \mapsto p_u$ is constant on the jumps of F_μ, that is constant on the intervals $(F_\mu(x-), F_\mu(x)]$, $x \in \mathbb{R}$, which is trivially satisfied when μ is atomless. Suppose that $\widehat{M} = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times m_u \in \widehat{\Pi}^{\mathrm{M}}(\mu, \nu)$ is such that*

$$\int_{(0,1)} \mathcal{W}_1(p_u, m_u)\, du \leq \int_{(0,1)} \left| \int_{\mathbb{R}} y\, p_u(dy) - F_\mu^{-1}(u) \right| du.$$

Then the martingale coupling $M(dx, dy) = \int_{u\in(0,1)} \delta_{F_\mu^{-1}(u)}(dx)\, m_u(dy)\, du$ is a martingale rearrangement coupling of $\pi = \int_{u\in(0,1)} \delta_{F_\mu^{-1}(u)}(dx)\, p_u(dy)\, du$ which satisfies

$$\mathcal{AW}_1(\pi, M) = \int_{\mathbb{R}} \mathcal{W}_1(\pi_x, M_x)\, \mu(dx) = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} y\, \pi_x(dy) - x \right| \mu(dx).$$

Of course, under the hypotheses, $\widehat{\mathcal{AW}}_1(\widehat{\pi},\widehat{M}) \le \int_{(0,1)} \mathcal{W}_1(p_u, m_u)\,du \le \int_{(0,1)} \left|\int_{\mathbb{R}} y\,p_u(dy) - F_\mu^{-1}(u)\right|\,du$ so that, by Lemma 2.2, these inequalities are equalities and $\widehat{M}$ is a lifted martingale rearrangement of $\widehat{\pi}$.

Proof. By (1.9) it suffices to show that

$$\int_{\mathbb{R}} \mathcal{W}_1(\pi_x, M_x)\,\mu(dx) \le \int_{\mathbb{R}} \left|\int_{\mathbb{R}} y\,\pi_x(dy) - x\right|\,\mu(dx). \tag{2.8}$$

For $(x,v) \in \mathbb{R}\times(0,1)$, let $\theta(x,v) = F_\mu(x-)+v\mu(\{x\})$. Using (1.16) and the fact that $F_\mu^{-1}(\theta(x',v)) = x'$ for all $(x',v) \in \mathbb{R}\times(0,1)$, we get

$$\begin{aligned}\pi(dx,dy) &= \int_{u\in(0,1)} \delta_{F_\mu^{-1}(u)}(dx)\,p_u(dy)\,du = \int_{(x',v)\in\mathbb{R}\times(0,1)} \delta_{x'}(dx)\,p_{\theta(x',v)}(dy)\,\mu(dx')\,dv \\ &= \int_{v\in(0,1)} \mu(dx)\,p_{\theta(x,v)}(dy)\,dv.\end{aligned} \tag{2.9}$$

Hence we have $\mu(dx)$-almost everywhere $\pi_x(dy) = \int_0^1 p_{\theta(x,v)}(dy)\,dv$, and similarly we find $M_x(dy) = \int_0^1 m_{\theta(x,v)}(dy)\,dv$. Using (1.16) for the first and last equality, we deduce that

$$\begin{aligned}\int_{\mathbb{R}} \mathcal{W}_1(\pi_x, M_x)\,\mu(dx) &\le \int_{\mathbb{R}\times(0,1)} \mathcal{W}_1(m_{\theta(x,v)}, p_{\theta(x,v)})\,\mu(dx)\,dv \\ &= \int_{(0,1)} \mathcal{W}_1(m_u, p_u)\,du \\ &\le \int_{(0,1)} \left|\int_{\mathbb{R}} y\,p_u(dy) - F_\mu^{-1}(u)\right|\,du \\ &= \int_{\mathbb{R}\times(0,1)} \left|\int_{\mathbb{R}} y\,p_{\theta(x,v)}(dy) - F_\mu^{-1}(\theta(x,v))\right|\,\mu(dx)\,dv.\end{aligned}$$

For $(x,v) \in \mathbb{R}\times(0,1)$, $F_\mu^{-1}(\theta(x,v)) = x$, and since $u \mapsto p_u$ is constant on the jumps of F_μ, the map $v \mapsto p_{\theta(x,v)}$ is constant on $(0,1)$, hence

$$\int_{(0,1)} \left|\int_{\mathbb{R}} y\,p_{\theta(x,v)}(dy) - F_\mu^{-1}(\theta(x,v))\right|\,dv = \left|\int_{\mathbb{R}\times(0,1)} y\,p_{\theta(x,v)}(dy)\,dv - x\right|.$$

We deduce that

$$\int_{\mathbb{R}} \mathcal{W}_1(\pi_x, M_x)\,\mu(dx) \le \int_{\mathbb{R}} \left|\int_{\mathbb{R}\times(0,1)} y\,p_{\theta(x,v)}(dy)\,dv - x\right|\,\mu(dx) = \int_{\mathbb{R}} \left|\int_{\mathbb{R}} y\,\pi_x(dy) - x\right|\,\mu(dx),$$

which proves (2.8) and concludes the proof. □

2.2 Construction of an explicit martingale rearrangement coupling

We recall that a coupling $\pi \in \Pi(\mu,\nu)$ between two probability measures $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ in the convex order satisfies the barycentre dispersion assumption formulated by Wiesel [39] iff

$$\forall a \in \mathbb{R}, \quad \int_{\mathbb{R}} \mathbb{1}_{[a,+\infty)}(x)\left(x - \int_{\mathbb{R}} y\,\pi_x(dy)\right)\,\mu(dx) \le 0. \tag{2.10}$$

First we briefly recall Wiesel's construction [39] of a martingale rearrangement coupling of a coupling π which satisfies (2.10), which is well perceivable as soon as π has finite support but becomes rather implicit in the general case. Then we design our own construction of such a martingale rearrangement coupling, whose intelligibility does not depend on the finiteness of the support of π. Since the Hoeffding-Fréchet satisfies (2.10) [39, Lemma 2.3], this construction extends the study made in Section 3.

Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu\le_{cx}\nu$ and $\mu\neq\nu$ and $\pi\in\Pi(\mu,\nu)\setminus\Pi^M(\mu,\nu)$ be a coupling between μ and ν which satisfies the barycentre dispersion assumption (2.10). Suppose first that π has finite support. As Wiesel [39] points out, the barycentre dispersion assumption (2.10) and the convex order between distinct μ and ν imply that

$$x^- := \max\left\{x : \mu(\{x\})>0 \text{ and } \int_{\mathbb{R}} y\,\pi_x(dy) < x\right\} < \max\left\{x : \mu(\{x\})>0 \text{ and } \int_{\mathbb{R}} y\,\pi_x(dy) > x\right\} =: x^+.$$

He then switches as much as possible the mass at $y^- := \min\{y : \pi_{x^-}(\{y\})>0\}$ and $y^+ := \max\{y : \pi_{x^+}(\{y\})>0\}$ of π_{x^-} and π_{x^+} in order to rectify the barycentres. More precisely, he defines for all $x\in S$

$$\pi^{(1)}_x = \mathbb{1}_{\{x\notin\{x^-,x^+\}\}}\,\pi_x + \mathbb{1}_{\{x=x^-\}}\left(\pi_{x^-} + \frac{\lambda}{\mu(\{x^-\})}(\delta_{y^+}-\delta_{y^-})\right) + \mathbb{1}_{\{x=x^+\}}\left(\pi_{x^+} + \frac{\lambda}{\mu(\{x^+\})}(\delta_{y^-}-\delta_{y^+})\right),$$

where $\lambda\ge 0$ is taken as large as possible, so that

$$\text{either}\quad \pi^{(1)}_{x^-}(\{y^-\})=0,\quad \pi^{(1)}_{x^+}(\{y^+\})=0,\quad \int_{\mathbb{R}} y\,\pi^{(1)}_{x^+}(dy)=x^+ \quad\text{or}\quad \int_{\mathbb{R}} y\,\pi^{(1)}_{x^-}(dy)=x^-.$$

Then the measure $\pi^{(1)}(dx,dy)=\mu(dx)\,\pi^{(1)}_x(dy)$ is a coupling between μ and ν which satisfies the barycentre dispersion assumption (2.10). After finitely many (in reason of the finite support of π) repetitions of this process, the obtained coupling is a martingale coupling and even a martingale rearrangement coupling of π.

In the general case, there exists by [39, Lemma 4.1] a sequence $(\pi^n)_{n\in\mathbb{N}^*}$ of finitely supported measures such that $\mathcal{W}^{nd}_1(\pi^n,\pi)\le 1/n$ for all $n\in\mathbb{N}^*$. The marginals μ_n and ν_n of π^n are not in the convex order, but a mere adaptation of the previous reasoning yields the existence of a coupling π^n_{mr} between μ_n and ν_n which is almost a martingale rearrangement coupling of π^n, in the sense that

$$\int_{\mathbb{R}}\left|x-\int_{\mathbb{R}} y\,(\pi^n_{mr})_x(dy)\right|\,\mu^n(dx)\le\frac{1}{n}\quad\text{and}\quad \mathcal{AW}_1(\pi^n_{mr},\pi^n)\le\int_{\mathbb{R}}\left|x-\int_{\mathbb{R}} y\,\pi^n_x(dy)\right|\,\mu^n(dx). \tag{2.11}$$

Then Wiesel shows the existence of a coupling π_{mr} between μ and ν such that $\mathcal{AW}_1\left(\frac{1}{n}\sum_{k=1}^n\pi^k_{mr},\pi_{mr}\right)$ vanishes as n goes to $+\infty$. By (2.11) taken to the limit $n\to+\infty$ and (1.9) he deduces that π_{mr} is a martingale rearrangement coupling of π.

We now propose an alternate construction of a martingale rearrangement coupling of π, regardless of the finiteness of its support, deduced from a lifted martingale coupling. Let us first give an intuitive description of the construction. For $u\in(0,1)$ we set $G(u)=\int_{\mathbb{R}} y\,\pi_{F_\mu^{-1}(u)}(dy)$ so that, locally, the lack of martingale property writes $G(u)\neq F_\mu^{-1}(u)$. We want for each $u\in(0,1)$ such that $G(u)<F^{-1}(u)$ to find a partner $v\in(u,1)$ such that $G(v)>F_\mu^{-1}(v)$ (and conversely for each $v\in(0,1)$ such that $G(v)>F_\mu^{-1}(v)$ to find a partner $u\in(0,v)$ such that $G(u)<F^{-1}(u)$) and to mix $\pi_{F_\mu^{-1}(u)}$ and $\pi_{F_\mu^{-1}(v)}$ in order to construct probability measures m_u and m_v with respective means $F_\mu^{-1}(u)$ and $F_\mu^{-1}(v)$ in order to restaure the martingale constraint. By taking expectations, the only possible $p\in[0,1]$ for the equality $p\pi_{F_\mu^{-1}(u)}+(1-p)\pi_{F_\mu^{-1}(v)}=pm_u+(1-p)m_v$ to hold is

$$p(u,v)=\frac{G(v)-F_\mu^{-1}(v)}{G(v)-F_\mu^{-1}(v)+F_\mu^{-1}(u)-G(u)}.$$

According to Lemma 2.4 below, for this choice, it is possible to find m_u and m_v as desired with the additional property $\pi_{F_\mu^{-1}(u)}\le_{st} m_u$ and $m_v\le_{st}\pi_{F_\mu^{-1}(v)}$, which, in view of Lemma 2.1, is a desirable feature in order to obtain a martingale rearrangement coupling when replacing $(\pi_{F_\mu^{-1}(u)}, \pi_{F_\mu^{-1}(v)})$ by (m_u, m_v). Introducing

$$\Delta_+(u) = \int_0^u (F_\mu^{-1} - G)^+(v)\,dv \quad \text{and} \quad \Delta_-(u) = \int_0^u (F_\mu^{-1} - G)^-(v)\,dv, \tag{2.12}$$

we will show that the barycenter dispersion assumption (2.10) is equivalent to

$$\forall u \in [0,1], \quad \Delta_+(u) \geq \Delta_-(u), \tag{2.13}$$

so that the choice $\Delta_+(u) = \Delta_-(v)$ ensures both that $u < v$ and $(1-p(u,v))du = p(u,v)dv$. The latter equality ensures that the "rates" of consumption of $\pi_{F_\mu^{-1}(u)}$, of consumption of $\pi_{F_\mu^{-1}(v)}$, of production of m_u and of production of m_v through the above mixing procedure are equal which is the key reason why the second marginal ν is preserved.

To now make the construction precise, we first show that (2.10) is equivalent to (2.13). Using (1.10) we see that for all $u \in (0,1)$ and $a \in \mathbb{R}$, $u > F_\mu(a-) \implies F_\mu^{-1}(u) \geq a \implies u \geq F_\mu(a-)$. By the latter implications and the inverse transform sampling we deduce that (2.10) is equivalent to

$$\forall a \in \mathbb{R}, \quad \int_{F_\mu(a-)}^1 (F_\mu^{-1}(u) - G(u))\,du \leq 0.$$

Since $\Delta_+(1) = \Delta_-(1)$, consequence of the equality of the respective means of μ and ν, we deduce that it is equivalent to

$$\forall a \in \mathbb{R}, \quad \Delta_+(F_\mu(a-)) \geq \Delta_-(F_\mu(a-)).$$

By right continuity of F_μ, for all $a \in \mathbb{R}$ we have $F_\mu(a) = \lim_{h\to 0, h>0} F_\mu((a+h)-)$, so by continuity of Δ_+ and Δ_- we also have $\Delta_+(F_\mu(a)) \geq \Delta_-(F_\mu(a))$ for all $a \in \mathbb{R}$. Moreover, for all $a \in \mathbb{R}$ such that $\mu(\{a\}) > 0$ and $u \in (F_\mu(a-), F_\mu(a)]$, we have by (1.11) that $F_\mu^{-1}(u) = a$, so Δ_+ and Δ_- are affine on $(F_\mu(a-), F_\mu(a)]$. We deduce that we also have $\Delta_+ \geq \Delta_-$ on $(F_\mu(a-), F_\mu(a)]$, hence the equivalence with (2.13).

We define

$$\begin{aligned}
&\mathcal{U}^+ = \{u \in (0,1) \mid F_\mu^{-1}(u) > G(u)\}, \quad \mathcal{U}^- = \{u \in (0,1) \mid F_\mu^{-1}(u) < G(u)\}, \\
&\text{and} \quad \mathcal{U}^0 = \{u \in (0,1) \mid F_\mu^{-1}(u) = G(u)\},
\end{aligned}$$

and thanks to the equality $\Delta_+(1) = \Delta_-(1)$ we can set for all $u \in [0,1]$

$$\phi(u) = \begin{cases} \Delta_-^{-1}(\Delta_+(u)) & \text{if} \quad u \in \mathcal{U}^+; \\ \Delta_+^{-1}(\Delta_-(u)) & \text{if} \quad u \in \mathcal{U}^-; \\ u & \text{if} \quad u \in \mathcal{U}^0. \end{cases}$$

Applying [26, Lemma 6.1] again with $f_1 = (F_\mu^{-1} - G)^+$, $f_2 = (F_\mu^{-1} - G)^-$, $u_0 = 1$ and $h : u \mapsto \mathbb{1}_{\{G(\phi(u)) \leq F_\mu^{-1}(\phi(u))\}}$ yields

$$\int_0^1 \mathbb{1}_{\{G(\phi(u)) \leq F_\mu^{-1}(\phi(u))\}}\,d\Delta_+(u) = \int_0^1 \mathbb{1}_{\{G(v) \leq F_\mu^{-1}(v)\}}\,d\Delta_-(u) = 0.$$

Similarly, we get $\int_0^1 \mathbb{1}_{\{G(\phi(u)) \geq F_\mu^{-1}(\phi(u))\}}\,d\Delta_-(u) = 0$. We deduce that

$$\phi(u) \in \mathcal{U}^-, \quad \text{resp. } \phi(u) \in \mathcal{U}^+, \quad \text{for } du\text{-almost all } u \in \mathcal{U}^+, \quad \text{resp. } \mathcal{U}^-. \tag{2.14}$$

This allows us to define for du-almost all $u \in \mathcal{U}^+ \cup \mathcal{U}^-$

$$p(u) = \frac{G(\phi(u)) - F_\mu^{-1}(\phi(u))}{F_\mu^{-1}(u) - G(u) + G(\phi(u)) - F_\mu^{-1}(\phi(u))}. \tag{2.15}$$

Notice that (2.14) implies that for du-almost all $u \in \mathcal{U}^+$, $\phi(\phi(u)) = \Delta_+^{-1}(\Delta_-(\Delta_-^{-1}(\Delta_+(u))))$. Since Δ_- is continuous we have $\Delta_-(\Delta_-^{-1}(v)) = v$ for all $v \in [0, \Delta_-(1)]$, and using (1.13) after an appropriate normalisation we get $\Delta_+^{-1}(\Delta_+(v)) = v$ for dv-almost all $v \in \mathcal{U}^+$. We deduce that

$$u = \phi(\phi(u)), \quad du\text{-almost everywhere on } \mathcal{U}^+. \tag{2.16}$$

Similarly, $\phi(\phi(u)) = u$ for all du-almost all $u \in \mathcal{U}^-$. We deduce that

$$\text{for } du\text{-almost all } u \in \mathcal{U}^+ \cup \mathcal{U}^-, \quad \phi(\phi(u)) = u, \tag{2.17}$$

and

$$\text{for } du\text{-almost all } u \in \mathcal{U}^+ \cup \mathcal{U}^-, \quad p(\phi(u)) = \frac{G(u) - F_\mu^{-1}(u)}{F_\mu^{-1}(\phi(u)) - G(\phi(u)) + G(u) - F_\mu^{-1}(u)} = 1 - p(u). \tag{2.18}$$

In order to define the appropriate martingale kernel, we rely on the following lemma which allows us to inject some stochastic order in the construction, a convenient tool for the computation of Wasserstein distances. We recall that two probability measures μ and ν on the real line are said to be in the stochastic order, denoted $\mu \le_{st} \nu$, iff $F_\mu^{-1}(u) \le F_\nu^{-1}(u)$ for all $u \in [0,1]$. Since the Hoeffding-Fréchet coupling between μ and ν is optimal for $\mathcal{W}_1(\mu,\nu)$, this implies by the inverse transform sampling that $\mathcal{W}_1(\mu,\nu) = \int_\mathbb{R} y\,\nu(dy) - \int_\mathbb{R} x\,\mu(dx)$.

Lemma 2.4. *Let $\mathfrak{B}$ be the set of all quadruples $(y, \widetilde{y}, \mu, \widetilde{\mu}) \in \mathbb{R} \times \mathbb{R} \times \mathcal{P}_1(\mathbb{R}) \times \mathcal{P}_1(\mathbb{R})$ such that μ and $\widetilde{\mu}$ have respective means x and $\widetilde{x}$ and $x < y \le \widetilde{y} < \widetilde{x}$. Endow $\mathcal{P}_1(\mathbb{R})$ with the Borel σ-algebra of the weak convergence topology and $\mathfrak{B}$ with the trace of the product σ-algebra on $\mathbb{R} \times \mathbb{R} \times \mathcal{P}_1(\mathbb{R}) \times \mathcal{P}_1(\mathbb{R})$.*

Then there exist two measurable maps $\beta, \widetilde{\beta} : \mathfrak{B} \to \mathcal{P}_1(\mathbb{R})$ such that for all $(y, \widetilde{y}, \mu, \widetilde{\mu})$, denoting $\nu = \beta(y, \widetilde{y}, \mu, \widetilde{\mu})$, $\widetilde{\nu} = \widetilde{\beta}(y, \widetilde{y}, \mu, \widetilde{\mu})$ and $p = \frac{\widetilde{x} - \widetilde{y}}{y - x + \widetilde{x} - \widetilde{y}}$ where x and $\widetilde{x}$ are the respective means of μ and $\widetilde{\mu}$, we have

$$\int_\mathbb{R} w\,\nu(dw) = y, \quad \int_\mathbb{R} w\,\widetilde{\nu}(dw) = \widetilde{y}, \quad \mu \le_{st} \nu, \quad \widetilde{\nu} \le_{st} \widetilde{\mu} \quad \textit{and} \quad p\nu + (1-p)\widetilde{\nu} = p\mu + (1-p)\widetilde{\mu}. \tag{2.19}$$

In particular, $p\,\delta_y(dz)\,\nu(dw) + (1-p)\,\delta_{\widetilde{y}}(dz)\,\widetilde{\nu}(dw)$ is a martingale coupling between $p\delta_y(dz) + (1-p)\delta_{\widetilde{y}}(dz)$ and $p\mu(dw) + (1-p)\widetilde{\mu}(dw)$, and $\mathcal{W}_1(\mu,\nu) = y - x$, $\mathcal{W}_1(\widetilde{\mu}, \widetilde{\nu}) = \widetilde{x} - \widetilde{y}$. The proof, which consists in exhibiting particular maps β and $\widetilde{\beta}$, is moved to the end of the present section.

In order to use Lemma 2.4 we need to compare ϕ to the identity function. The inequality (2.13) is equivalent by appropriate normalisation of (1.10) to $u \ge \Delta_+^{-1}(\Delta_-(u))$ for all $u \in [0,1]$, hence

$$\forall u \in \mathcal{U}^-, \quad \phi(u) \le u. \tag{2.20}$$

Moreover, by (2.17), [26, Lemma 6.1] applied with $f_1 = (F_\mu^{-1} - G)^+$, $f_2 = (F_\mu^{-1} - G)^-$, $u_0 = 1$ and $h : u \mapsto \mathbb{1}_{\{u < \phi(u)\}}$ we have

$$\int_0^1 \mathbb{1}_{\{\phi(u) < u\}}\, d\Delta_+(u) = \int_0^1 \mathbb{1}_{\{\phi(u) < \phi(\phi(u))\}}\, d\Delta_+(u) = \int_0^1 \mathbb{1}_{\{u < \phi(u)\}}\, d\Delta_-(u).$$

By (2.20) the right-hand side is 0, hence

$$\text{for } du\text{-almost all } u \in \mathcal{U}^+, \quad \phi(u) \ge u. \tag{2.21}$$

Let

$$A^+ = \{u \in \mathcal{U}^+ \mid F_\mu^{-1}(\phi(u)) < G(\phi(u)),\ \phi(\phi(u)) = u \text{ and } p(\phi(u)) = 1 - p(u)\}$$
$$\text{and} \quad A^- = \{u \in \mathcal{U}^- \mid F_\mu^{-1}(\phi(u)) > G(\phi(u)),\ \phi(\phi(u)) = u \text{ and } p(\phi(u)) = 1 - p(u)\}.$$

For all $u \in A^+$, we have by definition

$$\phi(u) \in \mathcal{U}^-, \quad F_\mu^{-1}(\phi(\phi(u))) = F_\mu^{-1}(u) > G(u) = G(\phi(\phi(u))),$$
$$\phi(\phi(\phi(u))) = \phi(u) \quad \text{and} \quad p(\phi(\phi(u))) = p(u) = 1 - p(\phi(u)),$$

hence $\phi(u) \in A^-$. Similarly, for all $u \in A^-$, $\phi(u) \in A^+$. By (2.14), (2.17), (2.18), (2.21) and the monotonicity of F_μ^{-1}, we deduce that A^+ and A^- are two disjoint Borel sets such that the Lebesgue measure of $(\mathcal{U}^+ \backslash A^+) \cup (\mathcal{U}^- \backslash A^-)$ is 0 and

$$\forall u \in A^+, \quad G(u) < F_\mu^{-1}(u) \leq F_\mu^{-1}(\phi(u)) < G(\phi(u)). \tag{2.22}$$

For all $u \in A^+$, $\pi_{F_\mu^{-1}(u)}$ and $\pi_{F_\mu^{-1}(\phi(u))}$ have by definition respective means $G(u)$ and $G(\phi(u))$, so by (2.22) we can apply Lemma 2.4 with

$$(y, \widetilde{y}, \mu, \widetilde{\mu}) = (F_\mu^{-1}(u), F_\mu^{-1}(\phi(u)), \pi_{F_\mu^{-1}(u)}, \pi_{F_\mu^{-1}(\phi(u))}).$$

Hence there exist two probability measures $m_u, \widetilde{m}_u \in \mathcal{P}_1(\mathbb{R})$ with respective means $F_\mu^{-1}(u)$, $F_\mu^{-1}(\phi(u))$ and such that

$$\begin{aligned} &\pi_{F_\mu^{-1}(u)} \leq_{st} m_u, \quad \widetilde{m}_u \leq_{st} \pi_{F_\mu^{-1}(\phi(u))}, \\ \text{and} \quad &p(u)m_u + (1-p(u))\widetilde{m}_u = p(u)\pi_{F_\mu^{-1}(u)} + (1-p(u))\pi_{F_\mu^{-1}(\phi(u))}. \end{aligned} \tag{2.23}$$

Since $A^+ = \phi(A^-)$ and $A^- = \phi(A^+)$, for all $u \in A^-$ we can set $m_u = \widetilde{m}_{\phi(u)}$, so that

$$\forall u \in A^+, \quad \pi_{F_\mu^{-1}(u)} \leq_{st} m_u \quad \text{and} \quad \forall u \in A^-, \quad m_u \leq_{st} \pi_{F_\mu^{-1}(u)}, \tag{2.24}$$

and ,

$$\forall u \in A^+ \cup A^-, \ p(u)m_u + p(\phi(u))m_{\phi(u)} = p(u)\pi_{F_\mu^{-1}(u)} + p(\phi(u))\pi_{F_\mu^{-1}(\phi(u))}. \tag{2.25}$$

Finally, for all $u \in \mathcal{U}^0 \cup (\mathcal{U}^+ \backslash A^+) \cup (\mathcal{U}^- \backslash A^-)$ set $m_u = \pi_{F_\mu^{-1}(u)}$. By composition of the measurable map $u \mapsto (F_\mu^{-1}(u), F_\mu^{-1}(\phi(u)), \pi_{F_\mu^{-1}(u)}, \pi_{F_\mu^{-1}(\phi(u))})$ and the measurable map β defined in Lemma 2.4, the map $u \mapsto m_u$ is measurable. By [2, Theorem 19.12] it is equivalent to say that $(m_u)_{u \in (0,1)}$ is a probability kernel, hence we can define

$$\widehat{M}(du, dx, dy) = \lambda_{(0,1)}(du)\, \delta_{F_\mu^{-1}(u)}(dx)\, m_u(dy), \tag{2.26}$$

and

$$M(dx, dy) = \int_0^1 \delta_{F_\mu^{-1}(u)}(dx)\, m_u(dy)\, du. \tag{2.27}$$

We now state that $\widehat{M}$ is a lifted martingale rearrangement coupling of $\widehat{\pi} = \iota(\pi)$.

Proposition 2.5. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$ and $\mu \neq \nu$ and $\pi \in \Pi(\mu, \nu)$ be a coupling between μ and ν which satisfies the barycentre dispersion assumption* (2.10)*. Then the measure $\widehat{M}$ defined by* (2.26) *is a lifted martingale rearrangement coupling of the lifted coupling $\widehat{\pi} = \iota(\pi)$:*

$$\inf_{\widehat{M}' \in \widehat{\Pi}^{\mathrm{M}}(\mu,\nu)} \widehat{\mathcal{AW}}_1(\widehat{\pi}, \widehat{M}') = \widehat{\mathcal{AW}}_1(\widehat{\pi}, \widehat{M}) = \int_{(0,1)} \mathcal{W}_1(\pi_{F_\mu^{-1}(u)}, m_u)\, du = \int_{(0,1)} \left| G(u) - F_\mu^{-1}(u) \right| du.$$

Since $u \mapsto \pi_{F_\mu^{-1}(u)}$ is constant on the jumps on F_μ by (1.11), we immediately deduce by Proposition 2.3 that M is a martingale rearrangement coupling of π.

Corollary 2.6. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$ and $\mu \neq \nu$ and $\pi \in \Pi(\mu, \nu)$ be a coupling between μ and ν which satisfies the barycentre dispersion assumption* (2.10)*. Then the measure M defined by* (2.27) *is a martingale rearrangement coupling of π:*

$$\inf_{M' \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_1(\pi, M') = \mathcal{AW}_1(\pi, M) = \int_{\mathbb{R}} \mathcal{W}_1(\pi_x, M_x)\, \mu(dx) = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} y\, \pi_x(dy) - x \right| \mu(dx).$$

Remark 2.7. As seen from the proof of Proposition 2.5 just below, for $\widehat{M}$ defined by (2.26) to be a lifted martingale rearrangement coupling of the lifted coupling $\widehat{\pi} = \iota(\pi)$ and therefore M defined by (2.27) to be a martingale rearrangement coupling of π, it is enough that $u \mapsto m_u$ is measurable, satisfies (2.24), (2.25) and $m_u = \pi_{F_\mu^{-1}(u)}$ for all $u \in \mathcal{U}^0 \cup (\mathcal{U}^+ \backslash A^+) \cup (\mathcal{U}^- \backslash A^-)$.

Proof of Proposition 2.5. Assume for a moment that $\widehat{M} \in \widehat{\Pi}^{\mathrm{M}}(\mu, \nu)$. Then we have by (2.24) that for all $u \in (0,1)$, $\pi_{F_\mu^{-1}(u)} \leq_{st} m_u$ or $m_u \leq_{st} \pi_{F_\mu^{-1}(u)}$, hence $\mathcal{W}_1(\pi_{F_\mu^{-1}(u)}, m_u) = |G(u) - F_\mu^{-1}(u)|$ and

$$\widehat{\mathcal{AW}}_1(\widehat{\pi}, \widehat{M}) \leq \int_{(0,1)} \mathcal{W}_1(\pi_{F_\mu^{-1}(u)}, m_u)\, du = \int_{(0,1)} |G(u) - F_\mu^{-1}(u)|\, du,$$

which proves the claim by Lemma 2.2.

It remains to show that $\widehat{M} \in \widehat{\Pi}^{\mathrm{M}}(\mu, \nu)$. By the inverse transform sampling and the fact that m_u has mean $F_\mu^{-1}(u)$ for all $u \in (0,1)$, it is clear that $\widehat{M}$ is a lifted martingale coupling between μ and $\int_{u\in(0,1)} m_u(dy)\, du$. To conclude it is therefore sufficient to check that

$$\int_{u\in(0,1)} m_u(dy)\, du = \nu. \tag{2.28}$$

To this end, let $H : [0,1] \to \mathbb{R}$ be measurable and bounded. Using (2.15), (2.17) and [26, Lemma 6.1] applied with $f_1 = (F_\mu^{-1} - G)^+$, $f_2 = (F_\mu^{-1} - G)^-$, $u_0 = 1$ and $h : u \mapsto \frac{H(\phi(u))}{F_\mu^{-1}(\phi(u)) - G(\phi(u)) + G(u) - F_\mu^{-1}(u)}$ for the third equality, we get

$$\begin{aligned}
\int_{\mathcal{U}^+} (1 - p(u)) H(u)\, du &= \int_0^1 \frac{(F_\mu^{-1} - G)^+(u)}{F_\mu^{-1}(u) - G(u) + G(\phi(u)) - F_\mu^{-1}(\phi(u))} H(u)\, du \\
&= \int_0^1 h(\phi(u))\, d\Delta_+(u) \\
&= \int_0^1 h(v)\, d\Delta_-(v) \\
&= \int_0^1 \frac{(F_\mu^{-1} - G)^-(v)}{F_\mu^{-1}(\phi(v)) - G(\phi(v)) + G(v) - F_\mu^{-1}(v)} H(\phi(v))\, dv \\
&= \int_{\mathcal{U}^-} p(\phi(v)) H(\phi(v))\, dv.
\end{aligned}$$

Similarly, we have $\int_{\mathcal{U}^-} (1 - p(u)) H(u)\, du = \int_{\mathcal{U}^+} p(\phi(u)) H(\phi(u))\, du$. We deduce that

$$\begin{aligned}
\int_0^1 H(u)\, du &= \int_{\mathcal{U}^0} H(u)\, du + \int_{\mathcal{U}^+} p(u) H(u)\, du + \int_{\mathcal{U}^+} (1 - p(u)) H(u)\, du \\
&\quad + \int_{\mathcal{U}^-} p(u) H(u)\, du + \int_{\mathcal{U}^-} (1 - p(u)) H(u)\, du \\
&= \int_{\mathcal{U}^0} H(u)\, du + \int_{\mathcal{U}^+ \cup \mathcal{U}^-} (p(u) H(u) + p(\phi(u)) H(\phi(u)))\, du.
\end{aligned} \tag{2.29}$$

Let $f : \mathbb{R} \to \mathbb{R}$ be measurable and bounded. Using (2.29) applied with $H : u \mapsto \int_{\mathbb{R}} f(y)\, m_u(dy)$ for the first equality, the fact that $m_u = \pi_{F_\mu^{-1}(u)}$ for all $u \in \mathcal{U}^0$ and (2.25) for the second equality, (2.29) again applied with $H : u \mapsto \int_{\mathbb{R}} f(y)\, \pi_{F_\mu^{-1}(u)}(dy)$ for the third equality and the inverse transform sampling for the last equality, we get

$$\begin{aligned}
&\int_0^1 \int_{\mathbb{R}} f(y)\, m_u(dy)\, du \\
&= \int_{\mathcal{U}^0} \int_{\mathbb{R}} f(y)\, m_u(dy)\, du + \int_{\mathcal{U}^+ \cup \mathcal{U}^-} \int_{\mathbb{R}} f(y)\, (p(u)\, m_u(dy) + p(\phi(u))\, m_{\phi(u)}(dy))\, du
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathcal{U}^0} \int_{\mathbb{R}} f(y)\, \pi_{F_\mu^{-1}(u)}(dy)\, du \\
&\quad + \int_{\mathcal{U}^+ \cup \mathcal{U}^-} \int_{\mathbb{R}} f(y)\, (p(u)\, \pi_{F_\mu^{-1}(u)}(dy) + p(\phi(u))\, \pi_{F_\mu^{-1}(\phi(u))}(dy))\, du \\
&= \int_0^1 \int_{\mathbb{R}} f(y)\, \pi_{F_\mu^{-1}(u)}(dy)\, du \\
&= \int_{\mathbb{R}} f(y)\, \nu(dy),
\end{aligned}
$$

which shows (2.28) and concludes the proof. □

Proof of Lemma 2.4. Let $(y, \widetilde{y}, \mu, \widetilde{\mu}) \in \mathfrak{B}$, x and $\widetilde{x}$ be the respective means of μ and $\widetilde{\mu}$ and $p = \frac{\widetilde{x}-\widetilde{y}}{y-x+\widetilde{x}-\widetilde{y}}$. First we construct two measures $\nu, \widetilde{\nu} \in \mathcal{P}_1(\mathbb{R})$ which satisfy (2.19). Then we show that ν and $\widetilde{\nu}$ are measurable in $(y, \widetilde{y}, \mu, \widetilde{\mu})$.

Let $\mu \vee \widetilde{\mu}$, resp. $\mu \wedge \widetilde{\mu}$ be the image of the Lebesgue measure on $(0,1)$ by $F_\mu^{-1} \vee F_{\widetilde{\mu}}^{-1}$, resp. $F_\mu^{-1} \wedge F_{\widetilde{\mu}}^{-1}$. Let z, resp. $\widetilde{z}$ be the mean of $\mu \vee \widetilde{\mu}$, resp. $\mu \wedge \widetilde{\mu}$, that is

$$
z = \int_0^1 (F_\mu^{-1} \vee F_{\widetilde{\mu}}^{-1})(u)\, du \quad \text{and} \quad \widetilde{z} = \int_0^1 (F_\mu^{-1} \wedge F_{\widetilde{\mu}}^{-1})(u)\, du.
$$

With the inverse transform sampling in mind, we have

$$
\int_0^1 F_\mu^{-1}(u)\, du = x < y \leq \widetilde{y} < \widetilde{x} = \int_0^1 F_{\widetilde{\mu}}^{-1}(u)\, du,
$$

from which we readily deduce that $z > y > x$ and $\widetilde{z} < \widetilde{y} < \widetilde{x}$. Therefore we can set $a = \frac{z-y}{z-x} \in (0,1)$, $\widetilde{a} = \frac{\widetilde{y}-\widetilde{z}}{\widetilde{x}-\widetilde{z}} \in (0,1)$ and define the probability measures

$$
\nu = a\mu + (1-a)(\mu \vee \widetilde{\mu}) \quad \text{and} \quad \widetilde{\nu} = \widetilde{a}\widetilde{\mu} + (1-\widetilde{a})(\mu \wedge \widetilde{\mu}).
$$

Since $\nu \leq \mu + \widetilde{\mu}$ and $\widetilde{\nu} \leq \mu + \widetilde{\mu}$, we have that $\nu, \widetilde{\nu} \in \mathcal{P}_1(\mathbb{R})$.

We can easily check that $ax + (1-a)z = y$ and $\widetilde{a}\widetilde{x} + (1-\widetilde{a})\widetilde{z} = \widetilde{y}$, hence ν and $\widetilde{\nu}$ have respective means y and $\widetilde{y}$.

By definition of the stochastic order, it is clear that $\mu \leq_{st} \mu \vee \widetilde{\mu}$ and $\mu \wedge \widetilde{\mu} \leq_{st} \widetilde{\mu}$, which directly implies that $\mu \leq_{st} \nu$ and $\widetilde{\nu} \leq_{st} \widetilde{\mu}$.

Since $\mu \vee \widetilde{\mu} + \mu \wedge \widetilde{\mu} = \mu + \widetilde{\mu}$, by taking the means we have that $z + \widetilde{z} = x + \widetilde{x}$, or equivalently $z - x = \widetilde{x} - \widetilde{z}$. This helps us to see that

$$
p(1-a) = \frac{\widetilde{x}-\widetilde{y}}{y-x+\widetilde{x}-\widetilde{y}} \times \frac{y-x}{z-x} = \frac{y-x}{y-x+\widetilde{x}-\widetilde{y}} \times \frac{\widetilde{x}-\widetilde{y}}{\widetilde{x}-\widetilde{z}} = (1-p)(1-\widetilde{a}).
$$

Then we derive

$$
\begin{aligned}
p\nu + (1-p)\widetilde{\nu} &= pa\mu + p(1-a)(\mu \vee \widetilde{\mu}) + (1-p)\widetilde{a}\widetilde{\mu} + (1-p)(1-\widetilde{a})(\mu \wedge \widetilde{\mu}) \\
&= pa\mu + (1-p)\widetilde{a}\widetilde{\mu} + p(1-a)(\mu \vee \widetilde{\mu} + \mu \wedge \widetilde{\mu}) \\
&= pa\mu + (1-p)\widetilde{a}\widetilde{\mu} + p(1-a)(\mu + \widetilde{\mu}) \\
&= pa\mu + p(1-a)\mu + (1-p)\widetilde{a}\widetilde{\mu} + (1-p)(1-\widetilde{a})\widetilde{\mu} \\
&= p\mu + (1-p)\widetilde{\mu}.
\end{aligned}
$$

It remains to show that ν and $\widetilde{\nu}$ are measurable in $(y, \widetilde{y}, \mu, \widetilde{\mu})$. From their definition it is clear that we must show that a, $\widetilde{a}$, $\mu \vee \widetilde{\mu}$ and $\mu \wedge \widetilde{\mu}$ are measurable in $(y, \widetilde{y}, \mu, \widetilde{\mu})$. Since a and $\widetilde{a}$ clearly are measurable functions of y, $\widetilde{y}$ and the means of μ, $\widetilde{\mu}$, $\mu \vee \widetilde{\mu}$ and $\mu \wedge \widetilde{\mu}$, the only non-straightforward measurability properties to prove are that of the maps

$$
\mathcal{P}_1(\mathbb{R}) \ni \eta \mapsto \int_{\mathbb{R}} x\, \eta(dx) \quad \text{and} \quad \mathcal{P}_1(\mathbb{R}) \times \mathcal{P}_1(\mathbb{R}) \ni (\mu, \widetilde{\mu}) \mapsto (\mu \vee \widetilde{\mu}, \mu \wedge \widetilde{\mu}).
$$

First of all, the functions $x \mapsto x^+$ and $x \mapsto x^-$ being nonnegative and continuous, the maps $\mathcal{P}_1(\mathbb{R}) \ni \eta \mapsto \int_\mathbb{R} x^+ \, \eta(dx)$ and $\mathcal{P}_1(\mathbb{R}) \ni \eta \mapsto \int_\mathbb{R} x^- \, \eta(dx)$ are lower semicontinuous and therefore measurable with respect to the weak convergence topology. Hence their difference $\mathcal{P}_1(\mathbb{R}) \ni \eta \mapsto \int_\mathbb{R} x \, \eta(dx)$ is measurable.

Second of all, let $f : \mathbb{R} \to \mathbb{R}$ be continuous and bounded, and $(\mu_n)_{n\in\mathbb{N}}, (\widetilde{\mu}_n)_{n\in\mathbb{N}} \in \mathcal{P}_1(\mathbb{R})^\mathbb{N}$ converge weakly to μ and $\widetilde{\mu}$ respectively. Then for all u outside the at most countable sets of discontinuities of F_μ^{-1} and $F_{\widetilde{\mu}}^{-1}$, the sequences $(F_{\mu_n}^{-1}(u))_{n\in\mathbb{N}}$ and $(F_{\widetilde{\mu}_n}^{-1}(u))_{n\in\mathbb{N}}$ converge to $F_\mu^{-1}(u)$ and $F_{\widetilde{\mu}}^{-1}(u)$ respectively. We then deduce by the dominated convergence theorem that

$$\begin{aligned}\int_\mathbb{R} f(x) \, (\mu_n \vee \widetilde{\mu}_n)(dx) &= \int_0^1 f\left(F_{\mu_n}^{-1}(u) \vee F_{\widetilde{\mu}_n}^{-1}(u)\right) du \\ &\underset{n\to+\infty}{\longrightarrow} \int_0^1 f\left(F_\mu^{-1}(u) \vee F_{\widetilde{\mu}}^{-1}(u)\right) du = \int_\mathbb{R} f(x) \, (\mu \vee \widetilde{\mu})(dx),\end{aligned}$$

hence $(\mu_n \vee \widetilde{\mu}_n)_{n\in\mathbb{N}}$ converges weakly to $\mu \vee \widetilde{\mu}$. Similarly, $(\mu_n \wedge \widetilde{\mu}_n)_{n\in\mathbb{N}}$ converges weakly to $\mu \wedge \widetilde{\mu}$. We deduce the continuity and therefore the measurability of $\mathcal{P}_1(\mathbb{R}) \times \mathcal{P}_1(\mathbb{R}) \ni (\mu, \widetilde{\mu}) \mapsto (\mu \vee \widetilde{\mu}, \mu \wedge \widetilde{\mu})$, which ends the proof. □

3 Martingale rearrangement couplings of the Hoeffding-Fréchet coupling

3.1 The inverse transform martingale coupling

We come back on the inverse transform martingale coupling and the family parametrised by $\mathcal{Q}$ introduced in [26] since they will have particular significance in the remaining of the present paper. We briefly recall the construction and main properties and refer to [26] for an extensive study. Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$ and $\mu \neq \nu$. For $u \in [0, 1]$ we define

$$\Psi_+(u) = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^+(v) \, dv \quad \text{and} \quad \Psi_-(u) = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^-(v) \, dv, \tag{3.1}$$

with respective left continuous generalised inverses Ψ_+^{-1} and Ψ_-^{-1}. We then define $\mathcal{Q}$ as the set of probability measures on $(0,1)^2$ with first marginal $\frac{1}{\Psi_+(1)} d\Psi_+$, second marginal $\frac{1}{\Psi_+(1)} d\Psi_-$ and such that $u < v$ for $Q(du, dv)$-almost every $(u, v) \in (0,1)^2$. Since $d\,\Psi_+$ and $d\,\Psi_-$ are concentrated on two disjoint Borel sets, there exists for each $Q \in \mathcal{Q}$ a probability kernel $(\pi_u^Q)_{u\in(0,1)}$ such that

$$Q(du, dv) = \frac{1}{\Psi_+(1)} d\Psi_+(u) \, \pi_u^Q(dv) = \frac{1}{\Psi_+(1)} d\Psi_-(v) \, \pi_v^Q(du), \tag{3.2}$$

and we exhibit a probability kernel $(\widetilde{m}_u^Q)_{u\in(0,1)}$ which satisfies for du-almost all $u \in (0,1)$ such that $F_\mu^{-1}(u) \neq F_\nu^{-1}(u)$

$$\widetilde{m}_u^Q(dy) = \int_{v\in(0,1)} \left(\frac{F_\mu^{-1}(u) - F_\nu^{-1}(u)}{F_\nu^{-1}(v) - F_\nu^{-1}(u)} \delta_{F_\nu^{-1}(v)}(dy) + \frac{F_\nu^{-1}(v) - F_\mu^{-1}(u)}{F_\nu^{-1}(v) - F_\nu^{-1}(u)} \delta_{F_\nu^{-1}(u)}(dy) \right) \pi_u^Q(dv), \tag{3.3}$$

and $\widetilde{m}_u^Q(dy) = \delta_{F_\nu^{-1}(u)}(dy)$ for all $u \in (0,1)$ such that $F_\mu^{-1}(u) = F_\nu^{-1}(u)$. Then the measure

$$\widehat{M}^Q(du, dx, dy) = \lambda_{(0,1)}(du) \, \delta_{F_\mu^{-1}(u)}(dx) \, \widetilde{m}_u^Q(dy) \tag{3.4}$$

is a lifted martingale coupling between μ and ν. Moreover it was shown by [26, Proposition 2.18] and its proof that for du-almost all $u \in (0,1)$,

$$\int_\mathbb{R} |y - F_\nu^{-1}(u)| \, \widetilde{m}_u^Q(dy) = |F_\mu^{-1}(u) - F_\nu^{-1}(u)|, \tag{3.5}$$

from which we deduce that the measure

$$M^Q(dx,dy) = \int_0^1 \delta_{F_\mu^{-1}(u)}(dx)\, \widetilde{m}_u^Q(dy)\, du \tag{3.6}$$

is a martingale coupling between μ and ν which satisfies $\int_{\mathbb{R}\times\mathbb{R}} |y-x|\, M^Q(dx,dy) \le 2\mathcal{W}_1(\mu,\nu)$. Let also

$$\mathcal{U}^+ = \{u\in(0,1) \mid F_\mu^{-1}(u) > F_\nu^{-1}(u)\},\quad \mathcal{U}^- = \{u\in(0,1) \mid F_\mu^{-1}(u) < F_\nu^{-1}(u)\}, \tag{3.7}$$

$$\text{and}\quad \mathcal{U}^0 = \{u\in(0,1) \mid F_\mu^{-1}(u) = F_\nu^{-1}(u)\}. \tag{3.8}$$

Thanks to the equality $\Psi_+(1) = \Psi_-(1)$, consequence of the equality of the respective means of μ and ν, we can set for all $u\in[0,1]$

$$\varphi(u) = \begin{cases} \Psi_-^{-1}(\Psi_+(u)) & \text{if} \quad u\in\mathcal{U}^+; \\ \Psi_+^{-1}(\Psi_-(u)) & \text{if} \quad u\in\mathcal{U}^-; \\ u & \text{if} \quad u\in\mathcal{U}^0. \end{cases} \tag{3.9}$$

Then the measure $Q^{IT}(du,dv) = \frac{1}{\Psi_+(1)} d\Psi_+(u) \mathbb{1}_{\{0<\varphi(u)<1\}}\, \delta_{\varphi(u)}(dv)$ belongs to $\mathcal{Q}$. The martingale coupling $M^{IT} = M^{Q^{IT}}$ is the so called inverse transform martingale coupling, associated to the probability kernel $\widetilde{m}^{IT} = \widetilde{m}^{Q^{IT}}$ which satisfies for du-almost all $u\in(0,1)$

$$\widetilde{m}^{IT}(u,dy) = p(u)\, \delta_{F_\nu^{-1}(\varphi(u))}(dy) + (1-p(u))\, \delta_{F_\nu^{-1}(u)}(dy), \tag{3.10}$$

where $p(u) = \mathbb{1}_{\{F_\mu^{-1}(u)\neq F_\nu^{-1}(u)\}} \frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(\varphi(u))-F_\nu^{-1}(u)}$.

3.2 The Hoeffding-Fréchet coupling

Let μ and ν be two probability measures on the real line with finite first moment. We recall that the Hoeffding-Fréchet coupling between μ and ν, denoted π^{HF}, is by definition the comonotonic coupling between μ and ν, that is the image of the Lebesgue measure on $(0,1)$ by the map $u\mapsto(F_\mu^{-1}(u), F_\nu^{-1}(u))$. Equivalently, we can write

$$\pi^{HF}(dx,dy) = \int_{(0,1)} \delta_{(F_\mu^{-1}(u),F_\nu^{-1}(u))}(dx,dy)\, du.$$

This coupling is of paramount importance in the classical optimal transport theory in dimension 1 since it attains the infimum in the minimisation problem

$$\inf_{P\in\Pi(\mu,\nu)} \int_{\mathbb{R}\times\mathbb{R}} c(x,y)\, P(dx,dy)$$

as soon as c satisfies the so called Monge condition, see [35, Theorem 3.1.2]. The latter condition being satisfied for any function $(x,y)\mapsto h(|y-x|)$ where $h:\mathbb{R}_+\to\mathbb{R}$ is convex and non-decreasing, we deduce that π^{HF} is optimal for $\mathcal{W}_\rho(\mu,\nu)$ for all $\rho\ge1$. By strict convexity, it is even the only coupling optimal for $\mathcal{W}_\rho(\mu,\nu)$ for $\rho>1$. Reasoning like in (2.9), we get that for $\mu(dx)$-almost all $x\in\mathbb{R}$,

$$\pi_x^{HF}(dy) = \int_{(0,1)} \delta_{F_\nu^{-1}(\theta(x,v))}(dy)\, dv. \tag{3.11}$$

By (3.11) and monotonicity and left continuity of F_ν^{-1} we recover the well known fact that π^{HF} is given by a measurable map, i.e. is the image of μ by $x\mapsto(x,T(x))$ where $T:\mathbb{R}\to\mathbb{R}$ is measurable, iff for all $x\in\mathbb{R}$ such that $\mu(\{x\})>0$, F_ν^{-1} is constant on $(F_\mu(x-),F_\mu(x)]$. In that case, we have $T = F_\nu^{-1}\circ F_\mu$, referred to as the Monge transport map.

3.3 Martingale rearrangement couplings

Our family $(M^Q)_{Q\in\mathcal{Q}}$ of martingale couplings mentioned above was meant to contain the closest martingale couplings from the Hoeffding-Fréchet coupling, the latter being well known for minimising the Wasserstein distance. Thanks to Wiesel's definition of martingale rearrangement couplings we can now rephrase the latter sentence in a more formal way. Let π^{HF} be the Hoeffding-Fréchet coupling between μ and ν. We will consider the following lifted coupling of π^{HF}:

$$\widehat{\pi}^{HF}(du,dx,dy) = \lambda_{(0,1)}(du)\,\delta_{F_\mu^{-1}(u)}(dx)\,\delta_{F_\nu^{-1}(u)}(dy). \tag{3.12}$$

Recall the embedding ι defined by (2.7) and the definition of the map θ given by (1.14). Then

$$\iota(\pi^{HF})(du,dx,dy) = \lambda_{(0,1)}(du)\,\delta_{F_\mu^{-1}(u)}(dx)\int_0^1 \delta_{F_\nu^{-1}(\theta(F_\mu^{-1}(u),v))}(dy)\,dv,$$

which is different from $\widehat{\pi}^{HF}$ when F_ν^{-1} is not constant on the jumps of F_μ. We can actually see that $\widehat{\pi}^{HF} = \iota'(\pi^{HF})$, where ι' is another embedding $\Pi(\mu,\nu)$ to $\widehat{\Pi}(\mu,\nu)$, such that for all $\pi\in\Pi(\mu,\nu)$, $\iota'(\pi)$ is defined by

$$\lambda_{(0,1)}(du)\,\delta_{F_\mu^{-1}(u)}(dx)\left(\mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})>0\}}\delta_{\left(F_{\pi_{F_\mu^{-1}(u)}}\right)^{-1}\left(\frac{u-F_\mu(F_\mu^{-1}(u)-)}{\mu(\{F_\mu^{-1}(u)\})}\right)}(dy) + \mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})=0\}}\pi_{F_\mu^{-1}(u)}(dy)\right).$$

Although $\widehat{\pi}^{HF}$ is a very natural lifted coupling of π^{HF}, the embedding ι used in Section 2.1 appears to be in general simpler than ι'.

Proposition 3.1. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu\le_{cx}\nu$. Then for all $Q\in\mathcal{Q}$, the lifted martingale coupling $\widehat{M}^Q$ defined by* (3.4) *is a lifted martingale rearrangement coupling of the lifted Hoeffding-Fréchet coupling $\widehat{\pi}^{HF}$ defined by* (3.12)*:*

$$\forall Q\in\mathcal{Q},\quad \widehat{\mathcal{AW}}_1(\widehat{\pi}^{HF},\widehat{M}^Q) = \inf_{\widehat{M}\in\widehat{\Pi}^{\mathrm{M}}(\mu,\nu)} \widehat{\mathcal{AW}}_1(\widehat{\pi}^{HF},\widehat{M}).$$

Proof. Let $Q\in\mathcal{Q}$. The fact that $\widehat{M}^Q\in\widehat{\Pi}^{\mathrm{M}}(\mu,\nu)$ is clear. By (3.5) we have

$$\widehat{\mathcal{AW}}_1(\widehat{\pi}^{HF},\widehat{M}^Q) \le \int_{(0,1)} \mathcal{W}_1(\delta_{F_\nu^{-1}(u)},\widetilde{m}_u^Q)\,du = \int_{(0,1)}\int_{\mathbb{R}} |y-F_\nu^{-1}(u)|\,\widetilde{m}_u^Q(dy)du = \int_{(0,1)} |F_\mu^{-1}(u)-F_\nu^{-1}(u)|\,du,$$

which proves the claim by Lemma 2.2. □

We can also easily show that any lifted martingale coupling is a lifted quadratic martingale rearrangement coupling of the lifted Hoeffding-Fréchet coupling.

Proposition 3.2. *Let $\mu,\nu\in\mathcal{P}_2(\mathbb{R})$ be such that $\mu\le_{cx}\nu$. Then any lifted martingale coupling between μ and ν is a $\widehat{\mathcal{AW}}_2$-minimal lifted martingale rearrangement coupling of the lifted Hoeffding-Fréchet coupling $\widehat{\pi}^{HF}$ defined by* (3.12)*:*

$$\forall M,M'\in\widehat{\Pi}^{\mathrm{M}}(\mu,\nu),\quad \widehat{\mathcal{AW}}_2(\widehat{M},\widehat{\pi}^{HF}) = \widehat{\mathcal{AW}}_2(\widehat{M}',\widehat{\pi}^{HF}).$$

Proof. Let $\widehat{M} = \lambda_{(0,1)}\times\delta_{F_\mu^{-1}(u)}\times m_u\in\widehat{\Pi}^{\mathrm{M}}(\mu,\nu)$ and $\chi\in\Pi(\lambda_{(0,1)},\lambda_{(0,1)})$ be optimal for $\widehat{\mathcal{AW}}_2(\widehat{M},\widehat{\pi}^{HF})$, so that

$$\begin{aligned}\widehat{\mathcal{AW}}_2^2(\widehat{M},\widehat{\pi}^{HF}) &= \int_{(0,1)\times(0,1)} \left(|u-u'|^2 + |F_\mu^{-1}(u)-F_\mu^{-1}(u')|^2 + \mathcal{W}_2^2(m_u,\delta_{F_\nu^{-1}(u')})\right)\chi(du,du')\\ &\ge \int_{(0,1)\times(0,1)} \mathcal{W}_2^2(m_u,\delta_{F_\nu^{-1}(u')})\,\chi(du,du').\end{aligned}$$

By bias-variance decomposition for the first equality, the fact that the image of $\lambda_{(0,1)}$ by $u \mapsto (F_\mu^{-1}(u), F_\nu^{-1}(u))$ is optimal for $\mathcal{W}_2^2(\mu,\nu)$ for the inequality, and by bias-variance decomposition again for the second equality, we have that

$$\begin{aligned}
&\int_{(0,1)\times(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u')})\,\chi(du,du') \\
&= \int_{(0,1)\times(0,1)} \left(|F_\nu^{-1}(u') - F_\mu^{-1}(u)|^2 + \int_{\mathbb{R}} |F_\mu^{-1}(u) - y|^2\, m_u(dy) \right) \chi(du,du') \\
&\geq \int_{(0,1)} \left(|F_\nu^{-1}(u) - F_\mu^{-1}(u)|^2 + \int_{\mathbb{R}} |F_\mu^{-1}(u) - y|^2\, m_u(dy) \right) du \qquad (3.13)\\
&= \int_{(0,1)} \int_{\mathbb{R}} |F_\nu^{-1}(u) - y|^2\, m_u(dy)\, du \\
&= \int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\, du \geq \widehat{\mathcal{AW}}_2^2(\widehat{M}, \widehat{\pi}^{HF}).
\end{aligned}$$

Using the fact that $\int_{(0,1)} \int_{\mathbb{R}} |F_\mu^{-1}(u) - y|^2\, m_u(dy)\, du = \int_{\mathbb{R}} |y|^2\, \nu(dy) - \int_{\mathbb{R}} |x|^2\, \mu(dx)$, we deduce that

$$\widehat{\mathcal{AW}}_2^2(\widehat{M}, \widehat{\pi}^{HF}) = \int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\, du = \mathcal{W}_2^2(\mu,\nu) + \int_{\mathbb{R}} |y|^2\, \nu(dy) - \int_{\mathbb{R}} |x|^2\, \mu(dx),$$

hence $\widehat{\mathcal{AW}}_2^2(\widehat{M}, \widehat{\pi}^{HF})$ does not depend on the choice of M. □

A similar conclusion holds for regular couplings. Just this once, we provide a proof valid in any dimension. In the following statement, $d \in \mathbb{N}^*$. The definitions (1.1), (1.3), (1.5) (1.7) given in $\mathbb{R}$ have straightfoward extensions to $\mathbb{R}^d$ endowed with the Euclidean norm $|\cdot|$.

Proposition 3.3. *Let $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^d)$ be such that $\mu \leq_{cx} \nu$ and $\pi \in \Pi(\mu,\nu)$ be optimal for $\mathcal{W}_2(\mu,\nu)$ and concentrated on the graph of a measurable map $T : \mathbb{R}^d \to \mathbb{R}^d$. Then any $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ is an $\mathcal{AW}_2$-minimal martingale rearrangement coupling of π.*

Proof. Let $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ and $\chi \in \Pi(\mu,\mu)$ be optimal for $\mathcal{AW}_2(M,\pi)$, so that

$$\mathcal{AW}_2^2(M,\pi) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \left(|x-x'|^2 + \mathcal{W}_2^2(M_x, \delta_{T(x')}) \right) \chi(dx,dx') \geq \int_{\mathbb{R}^d\times\mathbb{R}^d} \mathcal{W}_2^2(M_x, \delta_{T(x')})\,\chi(dx,dx').$$

By bias-variance decomposition for the first equality and the fact that the image of χ by $(x,x') \mapsto (x, T(x'))$ is a coupling between μ and ν for the first inequality, and by bias-variance decomposition again for the second equality, we have that

$$\begin{aligned}
\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathcal{W}_2^2(M_x, \delta_{T(x')})\,\chi(dx,dx') &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \left(|T(x')-x|^2 + \int_{\mathbb{R}^d} |x-y|^2\, M_x(dy) \right) \chi(dx,dx') \\
&\geq \mathcal{W}_2^2(\mu,\nu) + \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\, M(dx,dy) \\
&= \int_{\mathbb{R}^d} \left(|x-T(x)|^2 + \int_{\mathbb{R}^d} |x-y|^2\, M_x(dy) \right) \mu(dx) \qquad (3.14)\\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d} |y-T(x)|^2\, M(dx,dy) \\
&= \int_{\mathbb{R}^d} \mathcal{W}_2^2(M_x, \delta_{T(x)})\,\mu(dx) \geq \mathcal{AW}_2^2(M,\pi).
\end{aligned}$$

Using the fact that $\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\, M(dx,dy) = \int_{\mathbb{R}^d} |y|^2\, \nu(dy) - \int_{\mathbb{R}^d} |x|^2\, \mu(dx)$, we deduce that

$$\mathcal{AW}_2^2(M,\pi) = \int_{\mathbb{R}^d} \mathcal{W}_2^2(M_x, \delta_{T(x)})\,\mu(dx) = \mathcal{W}_2^2(\mu,\nu) + \int_{\mathbb{R}^d} |y|^2\, \nu(dy) - \int_{\mathbb{R}^d} |x|^2\, \mu(dx),$$

hence any martingale coupling $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ is an $\mathcal{AW}_2$-minimal martingale rearrrangement coupling of π. □

The use of Lemma 2.1 allows us to easily prove that the analogue of Proposition 3.1 holds for regular couplings as soon as on each interval $(F_\mu(x-), F_\mu(x)]$, where $x \in \mathbb{R}$, the sign of $u \mapsto F_\mu^{-1}(u) - F_\nu^{-1}(u)$ is constant. Of course this includes the case where F_ν^{-1} is constant on the intervals of the form $(F_\mu(x-), F_\mu(x)]$ for $x \in \mathbb{R}$, or equivalently the Hoeffding-Fréchet coupling π^{HF} between μ and ν is concentrated on the graph of the Monge transport map $T = F_\nu^{-1} \circ F_\mu$. In the latter case, the conclusion of Proposition 3.4 below can also be seen as an immediate consequence of Proposition 2.3 and the proof of Proposition 3.1.

Proposition 3.4. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$ and on each interval $(F_\mu(x-), F_\mu(x)]$, where $x \in \mathbb{R}$, the sign of $u \mapsto F_\mu^{-1}(u) - F_\nu^{-1}(u)$ is constant. Then for all $Q \in \mathcal{Q}$, the martingale coupling M^Q defined by* (3.6) *is a martingale rearrangement coupling of the Hoeffding-Fréchet coupling π^{HF}:*

$$\forall Q \in \mathcal{Q}, \quad \mathcal{AW}_1(\pi^{HF}, M^Q) = \inf_{M \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_1(\pi^{HF}, M).$$

Proof. By Lemma 2.1 applied with $\chi = (x \mapsto (x,x))_\sharp \mu$, it suffices to show that $\mu(dx)$-almost everywhere,

$$\pi_x^{HF} \leq_{st} M_x^Q \quad \text{or} \quad \pi_x^{HF} \geq_{st} M_x^Q. \tag{3.15}$$

Reasoning like in (2.9) we get that $\mu(dx)$-almost everywhere,

$$\pi_x^{HF}(dy) = \int_{(0,1)} \delta_{F_\nu^{-1}(\theta(x,v))}(dy)\, dv \quad \text{and} \quad M_x^Q(dy) = \int_{(0,1)} \widetilde{m}^Q_{\theta(x,v)}(dy)\, dv.$$

We deduce from [26, Lemma 2.5] that for du-almost all $u \in (0,1)$ such that $F_\mu^{-1}(u) \geq F_\nu^{-1}(u)$ (resp $F_\mu^{-1}(u) \leq F_\nu^{-1}(u)$), $\delta_{F_\nu^{-1}(u)} \leq_{st} \widetilde{m}_u^Q$ (resp. $\delta_{F_\nu^{-1}(u)} \geq_{st} \widetilde{m}_u^Q$). This implies that for du-almost all $u \in (0,1)$ such that $\mu(\{F_\mu^{-1}(u)\}) = 0$, $\pi^{HF}_{F_\mu^{-1}(u)} = \delta_{F_\nu^{-1}(u)}$ and $M^Q_{F_\mu^{-1}(u)} = \widetilde{m}_u^Q$ are comparable under the stochastic order. Moreover, the assumption made on the sign of the map $F_\mu^{-1} - F_\nu^{-1}$ on the jumps of F_μ implies that for du-almost all $u \in (0,1)$ such that $\mu(\{F_\mu^{-1}(u)\}) > 0$, we have either $(\theta(F_\mu^{-1}(u), 0), \theta(F_\mu^{-1}(u), 1)] \subset \{F_\mu^{-1} \geq F_\nu^{-1}\}$ so that, using the characterization of the stochastic order in terms of the cumulative disribution functions,

$$\pi^{HF}_{F_\mu^{-1}(u)}(dy) = \int_{(0,1)} \delta_{F_\nu^{-1}(\theta(F_\mu^{-1}(u),v))}(dy)\, dv \leq_{st} \int_{(0,1)} \widetilde{m}^Q_{\theta(F_\mu^{-1}(u),v)}(dy)\, dv = M^Q_{F_\mu^{-1}(u)}(dy),$$

or $(\theta(F_\mu^{-1}(u), 0), \theta(F_\mu^{-1}(u), 1)] \subset \{F_\mu^{-1} \leq F_\nu^{-1}\}$ so that $\pi^{HF}_{F_\mu^{-1}(u)} \geq_{st} M^Q_{F_\mu^{-1}(u)}$. By the inverse transform sampling, this shows (3.15) and completes the proof. □

In the next example where the above constant sign condition fails, the inverse transform martingale coupling between μ and ν is not a martingale rearrangement coupling of π^{HF}. Therefore, in general, we cannot say that every element of our family $(M^Q)_{Q \in \mathcal{Q}}$ is a martingale rearrangement coupling of the Hoeffding-Fréchet coupling. However, we can always find a specific parameter $Q \in \mathcal{Q}$ such that the martingale coupling M^Q is a martingale rearrangement coupling of π^{HF} (see [27, Proposition 3.6]).

Example 3.5. Let $\mu = \frac{1}{4}(\delta_{-1} + 2\delta_0 + \delta_1)$ and $\nu = \frac{1}{4}(\delta_{-2} + \delta_{-1} + \delta_1 + \delta_2)$. The Hoeffding-Fréchet coupling π^{HF} between μ and ν is given by

$$\pi^{HF} = \frac{1}{4}\left(\delta_{(-1,-2)} + \delta_{(0,-1)} + \delta_{(0,1)} + \delta_{(1,2)}\right).$$

To see that the inverse transform martingale coupling

$$M^{IT} = \frac{1}{6}\delta_{(-1,-2)} + \frac{1}{12}\delta_{(-1,1)} + \frac{1}{12}\delta_{(0,-2)} + \frac{1}{6}\delta_{(0,-1)} + \frac{1}{6}\delta_{(0,1)} + \frac{1}{12}\delta_{(0,2)} + \frac{1}{12}\delta_{(1,-1)} + \frac{1}{6}\delta_{(1,2)}$$

is not a martingale rearrangement coupling of π^{HF}, we rely on the equivalent condition provided by Lemma 2.1. One can readily compute $M_0^{IT} = \frac{1}{6}\delta_{-2} + \frac{1}{3}\delta_{-1} + \frac{1}{3}\delta_1 + \frac{1}{6}\delta_2$, $\pi_{-1}^{HF} = \delta_{-2}$, $\pi_0^{HF} = \frac{1}{2}(\delta_{-1} + \delta_1)$ and $\pi_1^{HF} = \delta_2$. Then $-1 < 0$, $1 > 0$ and $0 = 0$, but we have neither $\pi_{-1}^{HF} \geq_{st} M_0^{IT}$, $\pi_1^{HF} \leq_{st} M_0^{IT}$, $\pi_0^{HF} \leq_{st} M_0^{IT}$ nor $\pi_0^{HF} \geq_{st} M_0^{IT}$. We deduce by Lemma 2.1 that M^{IT} is not a martingale rearrangement coupling of π^{HF}.

Note that the martingale rearrangement constructed in Section 2.2 is

$$\frac{3}{16}\delta_{(-1,-2)} + \frac{1}{16}\delta_{(-1,2)} + \frac{1}{4}\delta_{(0,-1)} + \frac{1}{4}\delta_{(0,1)} + \frac{1}{16}\delta_{(1,-2)} + \frac{3}{16}\delta_{(1,2)}.$$

3.4 An example of $\mathcal{AW}_\rho$-minimal martingale rearrangement for $\rho > 2$

Let $f : \mathbb{R} \to \mathbb{R}$ and $q : \mathbb{R} \to [0,1]$ be defined for all $y \in \mathbb{R}$ by

$$f(y) = \frac{1+\mathrm{e}}{6}\left(\mathrm{e}^{-|y|}\mathbb{1}_{\{|y|\geq 1\}} + \frac{\mathrm{e}^{-|y|}+1}{1+\mathrm{e}}\mathbb{1}_{\{|y|<1\}}\right);$$
$$q(y) = \frac{\mathrm{e}}{1+\mathrm{e}}\mathbb{1}_{\{y\leq -1\}} + \frac{1}{1+\mathrm{e}^y}\mathbb{1}_{\{-1<y<1\}} + \frac{1}{1+\mathrm{e}}\mathbb{1}_{\{y\geq 1\}}.$$

Let $T : \mathbb{R} \to \mathbb{R}$ be the inverse of the continuous increasing map $y \mapsto y + 2q(y) - 1$, so that for all $y \in \mathbb{R}$, $q(y) = \frac{1+T^{-1}(y)-y}{2}$. Let $\nu(dy) = f(y)\,dy$ and $\mu = (T^{-1})_\sharp\nu$. We can easily compute

$$\sup_{x\in\mathbb{R}} |x - T(x)| = \sup_{y\in\mathbb{R}} |T^{-1}(y) - y| = \sup_{y\in\mathbb{R}} |2q(y) - 1| = \frac{\mathrm{e}-1}{\mathrm{e}+1} < 1.$$

By considering the cases $y \leq -2$, $-2 < y \leq -1$, $-1 < y \leq 0$, $0 < y \leq 1$, $1 < y \leq 2$ and $2 \leq y$, it is easy to check that

$$\forall y \in \mathbb{R}, \quad q(y-1)f(y-1) + (1 - q(y+1))f(y+1) = f(y). \tag{3.16}$$

Let

$$m_u^0 = q(F_\nu^{-1}(u))\,\delta_{F_\nu^{-1}(u)+1} + (1 - q(F_\nu^{-1}(u)))\,\delta_{F_\nu^{-1}(u)-1} \tag{3.17}$$

and $h : \mathbb{R} \to \mathbb{R}$ be measurable and bounded. Then

$$\begin{aligned}\int_{(0,1)\times\mathbb{R}} h(y)\,du\,m_u^0(dy) &= \int_{(0,1)} \big(q(F_\nu^{-1}(u))h(F_\nu^{-1}(u)+1) + (1 - q(F_\nu^{-1}(u)))h(F_\nu^{-1}(u)-1)\big)\,du \\ &= \int_{\mathbb{R}} (q(y)h(y+1) + (1-q(y))h(y-1))\,\nu(dy) \\ &= \int_{\mathbb{R}} q(y)h(y+1)f(y)\,dy + \int_{\mathbb{R}} (1-q(y))h(y-1)f(y)\,dy \\ &= \int_{\mathbb{R}} (q(y-1)f(y-1) + (1-q(y+1))f(y+1))\,h(y)\,dy \\ &= \int_{\mathbb{R}} f(y)h(y)\,dy,\end{aligned}$$

where we used (3.16) for the last equality. We deduce that $\int_{u\in(0,1)} m_u^0(dy)\,du = \nu(dy)$. Hence $\widehat{M}^0 = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times m_u^0 \in \widehat{\Pi}^{\mathrm{M}}(\mu,\nu)$.

Let us now show that $\widehat{M}^0$ is the only $\widehat{\mathcal{AW}}_\rho$-minimal martingale rearrangement coupling of $\widehat{\pi}^{HF}$ for $\rho > 2$. Since $|y - F_\nu^{-1}(u)|$ is $du\,m_u^0(dy)$-almost everywhere constant, we have

$$\begin{aligned}\left(\int_{(0,1)} \mathcal{W}_2^2(m_u^0, \delta_{F_\nu^{-1}(u)})\,du\right)^{\rho/2} &= \left(\int_{(0,1)}\int_{\mathbb{R}} |F_\nu^{-1}(u) - y|^2\,m_u^0(dy)\,du\right)^{\rho/2} \\ &= \int_{(0,1)}\int_{\mathbb{R}} |F_\nu^{-1}(u) - y|^\rho\,m_u^0(dy)\,du \geq \widehat{\mathcal{AW}}_\rho^\rho(\widehat{M}^0, \widehat{\pi}^{HF}).\end{aligned} \tag{3.18}$$

Since by Proposition 3.2 and its proof, $\int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\,du$ does not depend on $\widehat{M} = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times m_u \in \widehat{\Pi}^{\mathrm{M}}(\mu,\nu)$, to conclude it is enough to show that for $\widehat{M} \neq \widehat{M}^0$,

$$\widehat{\mathcal{AW}}_\rho^\rho(\widehat{M}, \widehat{\pi}^{HF}) > \left(\int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\,du\right)^{\rho/2}. \tag{3.19}$$

Let $\chi \in \Pi(\lambda_{(0,1)}, \lambda_{(0,1)})$ be optimal for $\widehat{\mathcal{AW}}_\rho(\widehat{M}, \widehat{\pi}^{HF})$. Suppose first that $\chi(du, du') = \lambda_{(0,1)}(du)\,\delta_u(du')$. Since

$$\int_{(0,1)}\int_{\mathbb{R}} |y - F_\nu^{-1}(u)|^2\, m_u(dy)\,du = \int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\,du = \int_{(0,1)} \mathcal{W}_2^2(m_u^0, \delta_{F_\nu^{-1}(u)})\,du = 1,$$

and $\widehat{M} \neq \widehat{M}^0$, $|y - F_\nu^{-1}(u)|$ is not $du\, m_u(dy)$-almost everywhere constant, so by Jensen's strict inequality we have

$$\begin{aligned}\widehat{\mathcal{AW}}_\rho^\rho(\widehat{M}, \widehat{\pi}^{HF}) &= \int_{(0,1)} \mathcal{W}_\rho^\rho(m_u, \delta_{F_\nu^{-1}(u)})\,du = \int_{\mathbb{R}\times(0,1)} |y - F_\nu^{-1}(u)|^\rho\, m_u(dy)\,du \\ &> \left(\int_{\mathbb{R}\times(0,1)} |y - F_\nu^{-1}(u)|^2\, m_u(dy)\,du\right)^{\rho/2} = \left(\int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\,du\right)^{\rho/2}.\end{aligned}$$

Else if $\chi(du, du') \neq \lambda_{(0,1)}(du)\,\delta_u(du')$, then using Jensen's inequality for the third inequality and (3.13) for the fourth, we have

$$\begin{aligned}\widehat{\mathcal{AW}}_\rho^\rho(\widehat{M}, \widehat{\pi}^{HF}) &> \int_{(0,1)\times(0,1)} \mathcal{W}_\rho^\rho(m_u, \delta_{F_\nu^{-1}(u')})\,\chi(du, du') \geq \int_{(0,1)\times(0,1)} \mathcal{W}_2^\rho(m_u, \delta_{F_\nu^{-1}(u')})\,\chi(du, du') \\ &\geq \left(\int_{(0,1)\times(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u')})\,\chi(du, du')\right)^{\rho/2} \geq \left(\int_{(0,1)} \mathcal{W}_2^2(m_u, \delta_{F_\nu^{-1}(u)})\,du\right)^{\rho/2},\end{aligned} \tag{3.20}$$

which proves (3.19) and therefore that $\widehat{M}^0$ is the only $\widehat{\mathcal{AW}}_\rho$-minimal martingale rearrangement coupling of $\widehat{\pi}^{HF}$. Note that (3.20) is valid for $\widehat{M} = \widehat{M}^0$, which in view of (3.18) shows that $\lambda_{(0,1)}(du)\,\delta_u(du')$ is the only coupling between $\lambda_{(0,1)}$ and $\lambda_{(0,1)}$ optimal for $\widehat{\mathcal{AW}}_\rho(\widehat{M}^0, \widehat{\pi}^{HF})$. With similar arguments we prove that

$$M^0(dx, dy) = \mu(dx)\left(q(T(x))\,\delta_{T(x)+1}(dy) + (1 - q(T(x)))\,\delta_{T(x)-1}(dy)\right)$$

is the only $\mathcal{AW}_\rho$-minimal martingale rearrangement coupling of π^{HF}, and $\mu(dx)\,\delta_x(dx')$ is the only coupling between μ and μ optimal for $\mathcal{AW}_\rho(M^0, \pi^{HF})$.

Remark 3.6. Since F_μ is continuous, by Proposition 3.4, for each $Q \in \mathcal{Q}$, M^Q defined by (3.6) is a martingale rearrangement coupling of π^{HF} : $\mathcal{AW}_1(\pi^{HF}, M^Q) = \inf_{M \in \Pi^{\mathrm{M}}(\mu,\nu)} \mathcal{AW}_1(\pi^{HF}, M)$. A related but slightly different minimisation problem is considered in [26], where, according to Proposition 2.11, any element of the family $(M^Q)_{Q\in\mathcal{Q}}$ of martingale couplings minimises

$$\int_{\mathbb{R}\times\mathbb{R}} |y - T(x)|\, M(dx, dy) = \int_{\mathbb{R}} \mathcal{W}_1(\delta_{T(x)}, M_x)\mu(dx) = \int_{\mathbb{R}} \mathcal{W}_1(\pi_x^{HF}, M_x)\mu(dx)$$

among all martingale couplings M between μ and ν and satisfies $\int_{\mathbb{R}\times\mathbb{R}} |y - T(x)|\, M^Q(dx, dy) = \mathcal{W}_1(\mu, \nu)$. According to [26, Proposition 3.5], since $\rho > 2$, the inverse transform martingale coupling M^{IT} minimises $\int_{\mathbb{R}\times\mathbb{R}} |y - T(x)|^\rho\, M^Q(dx, dy)$ among all martingale couplings M^Q parametrised by $Q \in \mathcal{Q}$. Yet the minimiser over the whole set of martingale couplings between μ and ν is not M^{IT} but M^0.

Indeed, by construction we have $M_x^{IT}(\{T(x)\}) > 0$, $\mu(dx)$-almost everywhere, hence $M^{IT} \neq M^0$ and $|y-T(x)|$ is not $M^{IT}(dx,dy)$-almost everywhere constant. Then by Jensen's strict inequality and the fact that $\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|^2 \, M(dx,dy)$ does not depend on the choice of $M \in \Pi^{\mathrm{M}}(\mu,\nu)$, we get

$$\begin{aligned}\left(\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|^\rho \, M^{IT}(dx,dy)\right)^{1/\rho} &> \left(\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|^2 \, M^0(dx,dy)\right)^{1/2} \\ &= \left(\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|^\rho \, M^0(dx,dy)\right)^{1/\rho}.\end{aligned} \tag{3.21}$$

Note that in [26, Proposition 3.5] (resp. [26, Proposition 5.9]), (3.11) and (3.12) (resp. (5.22)) are only valid in the case $\rho \in \{1\} \cup [2,+\infty)$ since the function c_ε defined in the proof lacks the claimed convexity property when $\rho < 2$ (resp. since (5.25) is only valid for $\rho \in \{1\} \cup [2,+\infty)$).

4 Stability of the inverse transform martingale coupling

In the next proposition we prove the stability in $\widehat{\mathcal{AW}}_\rho$, for $\rho \geq 1$, of the lifted inverse transform martingale coupling, defined for all $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ in the convex order by

$$\widehat{M}^{IT}(du,dx,dy) = \lambda_{(0,1)}(du)\, \delta_{F_\mu^{-1}(u)}(dx)\, \widetilde{m}_u^{IT}(dy),$$

where $(\widetilde{m}_u^{IT})_{u\in(0,1)}$ is defined by (3.10). In another proposition, we give a condition on the first marginals ensuring that the inverse transform martingale coupling is stable in $\mathcal{AW}_\rho$.

Proposition 4.1. *Let $\rho \geq 1$ and $\mu_n, \nu_n \in \mathcal{P}_\rho(\mathbb{R})$, $n \in \mathbb{N}$, be in convex order and respectively converge to μ and ν in $\mathcal{W}_\rho$ as $n \to +\infty$. Then*

$$\widehat{\mathcal{AW}}_\rho^\rho(\widehat{M}_n^{IT}, \widehat{M}^{IT}) \leq \int_{(0,1)} \mathcal{W}_\rho^\rho((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT})\, du \underset{n\to+\infty}{\longrightarrow} 0, \tag{4.1}$$

where $\widehat{M}_n^{IT} = \lambda_{(0,1)} \times \delta_{F_{\mu_n}^{-1}(u)} \times (\widetilde{m}_n^{IT})_u$, resp. $M^{IT} = \lambda_{(0,1)} \times \delta_{F_\mu^{-1}(u)} \times \widetilde{m}_u^{IT}$, denotes the lifted inverse transform martingale coupling between μ_n and ν_n, resp. μ and ν.

Proof. By Remark 5.4, it is enough to prove that $\widehat{\mathcal{AW}}_1(\widehat{M}_n^{IT}, \widehat{M}^{IT})$ goes to 0 as $n \to \infty$. Since

$$\widehat{\mathcal{AW}}_1(\widehat{M}_n^{IT}, \widehat{M}^{IT}) \leq \int_{(0,1)} \mathcal{W}_1((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT})\, du,$$

it suffices to show that the right-hand side vanishes as n goes to $+\infty$. This is achieved in two steps. First, we prove that, on the probability space $(0,1)$ endowed with the Lebesgue measure, the family of random variables $\left(\mathcal{W}_\rho^\rho\left((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT}\right)\right)_{n\in\mathbb{N}}$ is uniformly integrable, which, with the inequality $\mathcal{W}_\rho \geq \mathcal{W}_1$, implies the uniform integrability of $\left(\mathcal{W}_1\left((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT}\right)\right)_{n\in\mathbb{N}}$. Second we show for du-almost all $u \in (0,1)$ that

$$\mathcal{W}_1\left((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT}\right) \underset{n\to+\infty}{\longrightarrow} 0 \tag{4.2}$$

Let us begin with the uniform integrability. For $u \in (0,1)$, we can estimate

$$\mathcal{W}_\rho^\rho\left((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT}\right) \leq 2^{\rho-1} \int_{\mathbb{R}} |y|^\rho \left((\widetilde{m}_n^{IT})_u(dy) + \widetilde{m}_u^{IT}(dy)\right). \tag{4.3}$$

According to [26, Lemma 2.6], M^{IT} is the image of $\mathbb{1}_{(0,1)}(u)\, du\, \widetilde{m}_u^{IT}(dy)$ by $(u,y) \mapsto (F_\mu^{-1}(u), y)$ so that the second marginal of this measure is $\nu(dy)$. Therefore

$$\int_{(0,1)} \int_{\mathbb{R}} |y|^\rho\, \widetilde{m}_u^{IT}(dy)\, du = \int_{\mathbb{R}} |y|^\rho\, \nu(dy) < +\infty.$$

Hence it is enough to check the uniform integrability of $\left(\int_{\mathbb{R}} |y|^\rho\, (\widetilde{m}^{IT}_n)_u(dy)\right)_{n\in\mathbb{N}}$ to ensure that of $\left(\mathcal{W}_\rho^\rho\left((\widetilde{m}^{IT}_n)_u, \widetilde{m}^{IT}_u\right)\right)_{n\in\mathbb{N}}$. Since the second marginal of the measure $\mathbb{1}_{(0,1)}(u)\,du\,(\widetilde{m}^{IT}_n)_u(dy)$ is $\nu_n(dy)$, this measure also writes $\nu_n(dy)k^n_y(du)$ for some probability kernel k^n on $\mathbb{R}\times(0,1)$. Let $\varepsilon>0$ and A be a measurable subset of $(0,1)$ such that $\lambda(A)<\varepsilon$. For all $n\in\mathbb{N}$, we have

$$J_n(A) := \int_A \int_{\mathbb{R}} |y|^\rho\, (\widetilde{m}^{IT}_n)_u(dy)\, du = \int_{\mathbb{R}} |y|^\rho\, \tau_n(dy),$$

where $\tau_n(dy) = \int_{u=0}^1 \mathbb{1}_A(u)\, k^n_y(du)\, \nu_n(dy)$ is such that $\tau_n \le \nu_n$ and $\tau_n(\mathbb{R}) = \lambda(A)$. Hence

$$\sup_{A\in\mathcal{B}((0,1)),\ \lambda(A)\le\varepsilon} J_n(A) \le I^\rho_\varepsilon(\nu_n),$$

where $I^\rho_\varepsilon(\zeta)$ is defined for all $\zeta\in\mathcal{P}_\rho(\mathbb{R})$ as the supremum of $\int_{\mathbb{R}}|y|^\rho\,\tau(dy)$ over all finite measures τ on $\mathbb{R}$ such that $\tau\le\zeta$ and $\tau(\mathbb{R})\le\varepsilon$. Let $\eta>0$. By [9, Lemma 3.1 (b)], since $\nu\in\mathcal{P}_\rho(\mathbb{R})$, there exists $\varepsilon'>0$ such that $I^\rho_{\varepsilon'}(\nu)<\eta$. Let then $N\in\mathbb{N}$ be such that for all $n>N$, $\mathcal{W}^\rho_\rho(\nu_n,\nu)<\eta$, so that by [9, Lemma 3.1 (c)], $I^\rho_{\varepsilon'}(\nu_n) \le 2^{\rho-1}(\mathcal{W}^\rho_\rho(\nu_n,\nu) + I^\rho_{\varepsilon'}(\nu)) < 2^\rho\eta$. By [9, Lemma 3.1 (b)] again there exists $\varepsilon''>0$ such that for all $n\le N$, $I^\rho_{\varepsilon''}(\nu_n)<2^\rho\eta$. We deduce that for all $\varepsilon\in(0,\varepsilon'\wedge\varepsilon'')$,

$$\sup_{n\in\mathbb{N}}\ \sup_{A\in\mathcal{B}((0,1)),\ \lambda(A)\le\varepsilon} J_n(A) \le 2^\rho\eta,$$

which yields uniform integrability of $\left(\int_{\mathbb{R}}|y|^\rho\,(\widetilde{m}^{IT}_n)_u(dy)\right)_{n\in\mathbb{N}}$.

Next, we show the du-almost everywhere pointwise convergence of (4.2). Since, by monotonicity, $u\mapsto(F_\mu^{-1}(u),F_\nu^{-1}(u))$ is continuous du-almost everywhere on $(0,1)$ and, then, the weak convergence implies that

$$(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u)) \underset{n\to+\infty}{\longrightarrow} (F_\mu^{-1}(u), F_\nu^{-1}(u)), \tag{4.4}$$

we suppose without loss of generality that this convergence holds. Let $n\in\mathbb{N}$. Let Ψ_{n+}, resp. Ψ_{n-}, be the map defined by the left-hand, resp. right-hand side of (3.1), with (μ_n,ν_n) replacing (μ,ν). By (3.10),

$$(\widetilde{m}^{IT}_n)_u = p_n(u)\delta_{F_{\nu_n}^{-1}(\varphi_n(u))} + (1-p_n(u))\delta_{F_{\nu_n}^{-1}(u)} \quad \text{with} \quad p_n(u) = \mathbb{1}_{\{F_{\mu_n}^{-1}(u)\neq F_{\nu_n}^{-1}(u)\}}\frac{F_{\mu_n}^{-1}(u)-F_{\nu_n}^{-1}(u)}{F_{\nu_n}^{-1}(\varphi_n(u))-F_{\nu_n}^{-1}(u)} \in [0,1]$$

and $\varphi_n(u) = \Psi_{n-}^{-1}(\Psi_{n+}(u))$.

Suppose first that $u\in\mathcal{U}_0$ i.e. $F_\mu^{-1}(u)=F_\nu^{-1}(u)$, so that $\widetilde{m}^{IT}_u = \delta_{F_\nu^{-1}(u)}$. We have

$$\begin{aligned}
\mathcal{W}_1((\widetilde{m}^{IT}_n)_u, \widetilde{m}^{IT}_u) &= p_n(u)|F_{\nu_n}^{-1}(\varphi_n(u)) - F_\nu^{-1}(u)| + (1-p_n(u))|F_{\nu_n}^{-1}(u)-F_\nu^{-1}(u)| \\
&\le p_n(u)|F_{\nu_n}^{-1}(\varphi_n(u)) - F_{\nu_n}^{-1}(u)| + |F_{\nu_n}^{-1}(u)-F_\nu^{-1}(u)| \\
&= |F_{\mu_n}^{-1}(u)-F_{\nu_n}^{-1}(u)| + |F_{\nu_n}^{-1}(u)-F_\nu^{-1}(u)| \\
&\le |F_{\mu_n}^{-1}(u)-F_\mu^{-1}(u)| + 2|F_{\nu_n}^{-1}(u)-F_\nu^{-1}(u)|,
\end{aligned} \tag{4.5}$$

where the right-hand side goes to 0 as $n\to\infty$ by (4.4).

Suppose next that $u\in\mathcal{U}_+$ i.e. $F_\mu^{-1}(u)>F_\nu^{-1}(u)$, the case $u\in\mathcal{U}_-$ being treated in a similar way. Then without loss of generality

$$\widetilde{m}^{IT}_u = p(u)\delta_{F_\nu^{-1}(\varphi(u))} + (1-p(u))\delta_{F_\nu^{-1}(u)} \text{ with } p(u) = \frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(\varphi(u))-F_\nu^{-1}(u)}$$

and $\varphi(u)=\Psi_-^{-1}(\Psi_+(u))$. By (4.4), for n large enough, $u\in\mathcal{U}_{n+}$ so that without loss of generality, $p_n(u) = \frac{F_{\mu_n}^{-1}(u)-F_{\nu_n}^{-1}(u)}{F_{\nu_n}^{-1}(\varphi_n(u))-F_{\nu_n}^{-1}(u)}$ and checking (4.2) amounts to show that

$$F_{\nu_n}^{-1}(\varphi_n(u)) \underset{n\to+\infty}{\longrightarrow} F_\nu^{-1}(\varphi(u)). \tag{4.6}$$

It was shown in the proof of [26, Proposition 5.10] that Ψ_{n+} converges uniformly to Ψ_+ on $[0,1]$ and for dv-almost every $v \in (0,1)$,

$$F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\Psi_{n+}(1)v)) \underset{n\to+\infty}{\longrightarrow} F_\nu^{-1}(\Psi_-^{-1}(\Psi_+(1)v)). \tag{4.7}$$

Let $\mathcal{D}$ be the set of discontinuities of $F_\nu^{-1}\circ\Psi_-^{-1}$, which is at most countable by monotonicity. Then [36, Proposition 4.10, Chapter 0] yields

$$0 = \int_{\Psi_+(0)}^{\Psi_+(1)} \mathbb{1}_{\mathcal{D}}(v)\,dv = \int_0^1 \mathbb{1}_{\{\Psi_+(u)\in\mathcal{D}\}}\,d\Psi_+(u).$$

We deduce that for du-almost all $u \in \mathcal{U}_+$, $F_\nu^{-1}\circ\Psi_-^{-1}$ is continuous at $\Psi_+(u)$, which we suppose from now. According to (4.7), there exists $\varepsilon > 0$ arbitrarily small such that

$$\begin{aligned} &F_{\nu_n}^{-1}\left(\Psi_{n-}^{-1}\left(\Psi_{n+}(1)\frac{\Psi_+(u)-\varepsilon}{\Psi_+(1)}\right)\right) \underset{n\to+\infty}{\longrightarrow} F_\nu^{-1}\left(\Psi_-^{-1}\left(\Psi_+(u)-\varepsilon\right)\right) \\ \text{and}\quad &F_{\nu_n}^{-1}\left(\Psi_{n-}^{-1}\left(\Psi_{n+}(1)\frac{\Psi_+(u)+\varepsilon}{\Psi_+(1)}\right)\right) \underset{n\to+\infty}{\longrightarrow} F_\nu^{-1}\left(\Psi_-^{-1}\left(\Psi_+(u)+\varepsilon\right)\right). \end{aligned}$$

For n large enough, we have $\Psi_+(u) \in \left[\Psi_{n+}(1)\frac{\Psi_+(u)-\varepsilon}{\Psi_+(1)}, \Psi_{n+}(1)\frac{\Psi_+(u)+\varepsilon}{\Psi_+(1)}\right]$. Therefore, by monotonicity, we have

$$\begin{aligned} F_\nu^{-1}\left(\Psi_-^{-1}\left(\Psi_+(u)-\varepsilon\right)\right) &= \liminf_{n\to+\infty} F_{\nu_n}^{-1}\left(\Psi_{n-}^{-1}\left(\Psi_{n+}(1)\frac{\Psi_+(u)-\varepsilon}{\Psi_+(1)}\right)\right) \\ &\le \liminf_{n\to+\infty} F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\Psi_{n+}(u))) \\ &\le \limsup_{n\to+\infty} F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\Psi_{n+}(u))) \\ &\le \limsup_{n\to+\infty} F_{\nu_n}^{-1}\left(\Psi_{n-}^{-1}\left(\Psi_{n+}(1)\frac{\Psi_+(u)+\varepsilon}{\Psi_+(1)}\right)\right) \\ &= F_\nu^{-1}\left(\Psi_-^{-1}\left(\Psi_+(u)+\varepsilon\right)\right). \end{aligned}$$

Since $F_\nu^{-1}\circ\Psi_-^{-1}$ is continuous at $\Psi_+(u)$, we get when ε vanishes the convergence (4.6), which concludes the proof of (4.2) and therefore (4.1) □

Proposition 4.2. *Let $\rho \ge 1$ and $\mu_n, \nu_n \in \mathcal{P}_\rho(\mathbb{R})$, $n \in \mathbb{N}$, be in convex order and respectively converge to μ and ν in $\mathcal{W}_\rho$ as $n \to +\infty$. Suppose that asymptotically, any jump of F_μ is included in a jump of F_{μ_n}, that is*

$$\forall x\in\mathbb{R},\quad \mu(\{x\})>0 \implies \exists (x_n)_{n\in\mathbb{N}}\in\mathbb{R}^{\mathbb{N}},\quad F_{\mu_n}(x_n)\wedge F_\mu(x) - F_{\mu_n}(x_n-)\vee F_\mu(x-) \underset{n\to+\infty}{\longrightarrow} \mu(\{x\}), \tag{4.8}$$

which is for instance satisfied if μ is non-atomic. Then

$$\mathcal{AW}_\rho(M_n^{IT}, M^{IT}) \underset{n\to+\infty}{\longrightarrow} 0, \tag{4.9}$$

where M_n^{IT}, resp. M^{IT}, denotes the inverse transform martingale coupling between μ_n and ν_n, resp. μ and ν.

Remark 4.3. If (4.8) is not satisfied, then (4.9) may not hold. Indeed, for $n \in \mathbb{N}^*$, let $\mu_n = \mathcal{U}((-1/n, 1/n))$, $\mu = \delta_0$ and $\nu_n = \nu = \mathcal{U}((-1,1))$. We trivially have $M^{IT}(dx,dy) = \mu(dx)\,\nu(dy)$, so $\mathcal{AW}_1(M_n^{IT}, M^{IT}) \ge \int_{x\in\mathbb{R}} \mathcal{W}_1((M_n^{IT})_x, \nu)\,\mu_n(dx)$. However, for $n \in \mathbb{N}^*$, since F_{μ_n} is continuous, we have that for all $x \in \mathbb{R}$, $(M_n^{IT})_x = (\widetilde{m}_n^{IT})_{F_\mu(x)}$, where according to (3.10), $((\widetilde{m}_n^{IT})_u(dy))_{u\in(0,1)}$ is a probability kernel such that for all $u \in (0,1)$, there exist $a, b \in [-1,1]$ and $p \in [0,1]$ which satisfy $\widetilde{m}_n^{IT}(u,dy) = p\delta_a + (1-p)\delta_b$. Using the fact that the comonotonic coupling is optimal for the $\mathcal{W}_1$-distance, we get

$$\mathcal{W}_1(p\delta_a + (1-p)\delta_b, \nu) = \int_0^p |a+1-2u|\,du + \int_p^1 |b+1-2u|\,du.$$

It is easy to show that $\int_0^p |a+1-2u|\,du$ is equal to $p(a+1-p) \geq p^2$ if $(a+1)/2 > p$, and equal to $(a+1)^2/2 - p(a+1) + p^2 \leq p^2$ if $(a+1)/2 \leq p$. Therefore, one can readily show that $\int_0^p |a+1-2u|\,du \geq p^2/2$, attained for $a = p-1$. Similarly, we have $\int_p^1 |b+1-2u|\,du \geq (1-p)^2/2$, attained for $b = p$. We deduce that for all $(a,b,p) \in \mathbb{R}^2 \times [0,1]$, $\mathcal{W}_1(p\delta_a + (1-p)\delta_b, \nu) \geq (p^2 + (1-p)^2)/2 \geq 1/4$, attained for $p = 1/2$, hence $\int_{x\in\mathbb{R}} \mathcal{W}_1((M_n^{IT})_x, \nu)\,\mu_n(dx) \geq 1/4$, which proves that (4.9) is not satisfied.

Proof of Proposition 4.2. By Lemma 5.3 below we may suppose without loss of generality that $\rho = 1$. We have

$$\begin{aligned}\mathcal{AW}_1(M_n^{IT}, M^{IT}) &\leq \int_0^1 \left(|F_{\mu_n}^{-1}(u) - F_\mu^{-1}(u)| + \mathcal{W}_1\left((M_n^{IT})_{F_{\mu_n}^{-1}(u)}, M^{IT}_{F_\mu^{-1}(u)}\right)\right) du \\ &= \mathcal{W}_1(\mu_n, \mu) + \int_0^1 \mathcal{W}_1\left((M_n^{IT})_{F_{\mu_n}^{-1}(u)}, M^{IT}_{F_\mu^{-1}(u)}\right) du.\end{aligned}$$

For $(x,v) \in \mathbb{R} \times [0,1]$ and $n \in \mathbb{N}$, let $\theta(x,v) = F_\mu(x-) + v\mu(\{x\})$, $\theta_n(x,v) = F_{\mu_n}(x-) + v\mu_n(\{x\})$ and

$$(M_n)_x(dy) = \int_{v=0}^1 \widetilde{m}^{IT}_{\theta_n(x,v)}(dy)\,dv.$$

Then (1.15) and the triangle inequality yield

$$\begin{aligned}&\int_{(0,1)} \mathcal{W}_1\left((M_n^{IT})_{F_{\mu_n}^{-1}(u)}, M^{IT}_{F_\mu^{-1}(u)}\right) du \\ &\leq \int_{(0,1)} \left(\mathcal{W}_1\left((M_n^{IT})_{F_{\mu_n}^{-1}(u)}, (M_n)_{F_{\mu_n}^{-1}(u)}\right) + \mathcal{W}_1\left((M_n)_{F_{\mu_n}^{-1}(u)}, M^{IT}_{F_\mu^{-1}(u)}\right)\right) du \\ &\leq \int_{(0,1)^2} \left(\mathcal{W}_1\left((\widetilde{m}_n^{IT})_{\theta_n(F_{\mu_n}^{-1}(u),v)}, \widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}\right) + \mathcal{W}_1\left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}, \widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}\right)\right) du\,dv \\ &= \int_{(0,1)} \mathcal{W}_1\left((\widetilde{m}_n^{IT})_u, \widetilde{m}_u^{IT}\right) du + \int_{(0,1)^2} \mathcal{W}_1\left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}, \widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}\right) du\,dv.\end{aligned}$$

In order to show (4.9), it is therefore sufficient by (4.1) to prove that the second summand in right-hand side vanishes when n goes to $+\infty$. This is achieved in two steps. First, we prove that, on the probability space $(0,1)^2$ endowed with the Lebesgue measure, the family of random variables $\left(\mathcal{W}_1\left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}, \widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}\right)\right)_{n\in\mathbb{N}}$ is uniformly integrable. Second, we show for $du\,dv$-almost every $(u,v) \in (0,1)^2$ that

$$\mathcal{W}_1\left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}, \widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}\right) \underset{n\to+\infty}{\longrightarrow} 0. \tag{4.10}$$

Let us begin with the uniform integrability. For $(u,v) \in (0,1)^2$, we can estimate

$$\mathcal{W}_1\left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}, \widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}\right) \leq \int_{\mathbb{R}} |y| \left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}(dy) + \widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}(dy)\right).$$

For each nonnegative measurable function $f : \mathbb{R} \to \mathbb{R}$, we have by (1.15)

$$\begin{aligned}\int_{(0,1)^2} f\left(\int_{\mathbb{R}} |y|\,\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)}(dy)\right) du\,dv &= \int_{(0,1)^2} f\left(\int_{\mathbb{R}} |y|\,\widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}(dy)\right) du\,dv \\ &= \int_{(0,1)} f\left(\int_{\mathbb{R}} |y|\,\widetilde{m}^{IT}_u(dy)\right) du.\end{aligned}$$

According to [26, Lemma 2.6], M^{IT} is the image of $\mathbb{1}_{(0,1)}(u)\,du\,\widetilde{m}^{IT}_u(dy)$ by $(u,y)\mapsto(F_\mu^{-1}(u),y)$ so that the second marginal of this measure is $\nu(dy)$, hence the random variables $\left(\mathcal{W}_1\left(\widetilde{m}^{IT}_{\theta_n(F_{\mu_n}^{-1}(u),v)},\widetilde{m}^{IT}_{\theta(F_\mu^{-1}(u),v)}\right)\right)_{n\in\mathbb{N}}$ are uniformly integrable.

Next, we show the $du\,dv$-almost everywhere pointwise convergence of (4.10). Let $w\in(0,1)$ be in the set of continuity points of F_μ^{-1}, F_ν^{-1}, $F_\nu^{-1}\circ\Psi_-^{-1}\circ\Psi_+$ and $F_\nu^{-1}\circ\Psi_+^{-1}\circ\Psi_-$. Recall that we have

$$\widetilde{m}^{IT}_w=p(w)\delta_{F_\nu^{-1}(\varphi(w))}+(1-p(w))\delta_{F_\nu^{-1}(w)}\text{ with }p(w)=\mathbb{1}_{\{F_\mu^{-1}(w)\neq F_\nu^{-1}(w)\}}\frac{F_\mu^{-1}(w)-F_\nu^{-1}(w)}{F_\nu^{-1}(\varphi(w))-F_\nu^{-1}(w)}\in[0,1].$$

Let $(w_n)_{n\in\mathbb{N}}$ be a sequence with values in $(0,1)$ converging to w and let us show that

$$\mathcal{W}_1(\widetilde{m}^{IT}_{w_n},\widetilde{m}^{IT}_w)\underset{n\to+\infty}{\longrightarrow}0. \tag{4.11}$$

Suppose first that $w\in\mathcal{U}_0$ i.e. $F_\mu^{-1}(w)=F_\nu^{-1}(w)$. Then a computation similar to (4.5) yields

$$\mathcal{W}_1(\widetilde{m}^{IT}_{w_n},\widetilde{m}^{IT}_w)\le|F_\mu^{-1}(w_n)-F_\mu^{-1}(w)|+2|F_\nu^{-1}(w_n)-F_\nu^{-1}(w)|,$$

where the right-hand side goes to 0 as $n\to+\infty$ by continuity of F_μ^{-1} and F_ν^{-1} at w.

Suppose next that $w\in\mathcal{U}_+$ i.e. $F_\mu^{-1}(w)>F_\nu^{-1}(w)$, the case $w\in\mathcal{U}_-$ being treated in a similar way. Then by continuity of F_μ^{-1} and F_ν^{-1} at w, $w_n\in\mathcal{U}_+$ for n large enough so that without loss of generality

$$p(w)=\frac{F_\mu^{-1}(w)-F_\nu^{-1}(w)}{F_\nu^{-1}(\varphi(w))-F_\nu^{-1}(w)},\quad p(w_n)=\frac{F_\mu^{-1}(w_n)-F_\nu^{-1}(w_n)}{F_\nu^{-1}(\varphi(w_n))-F_\nu^{-1}(w_n)},$$

$\varphi(w)=\Psi_-^{-1}(\Psi_+(w))$, and $\varphi(w_n)=\Psi_-^{-1}(\Psi_+(w_n))$, hence (4.11) follows from the continuity at w of F_μ^{-1}, F_ν^{-1} and $F_\nu^{-1}\circ\Psi_-^{-1}\circ\Psi_+$. Since the set of discontinuity points of the non-decreasing functions F_μ^{-1}, F_ν^{-1}, $F_\nu^{-1}\circ\Psi_-^{-1}\circ\Psi_+$ and $F_\nu^{-1}\circ\Psi_+^{-1}\circ\Psi_-$ are at most countable, we deduce by (1.15) and (4.11) that it is sufficient to show for $du\,dv$-almost every $(u,v)\in(0,1)^2$

$$\theta_n(F_{\mu_n}^{-1}(u),v)\underset{n\to+\infty}{\longrightarrow}\theta(F_\mu^{-1}(u),v),$$

or equivalently

$$(F_{\mu_n}(x_u^n-),F_{\mu_n}(x_u^n))\underset{n\to+\infty}{\longrightarrow}(F_\mu(x_u-),F_\mu(x_u)) \tag{4.12}$$

for du-almost every $u\in(0,1)$, where $x_u:=F_\mu^{-1}(u)$ and $x_u^n:=F_{\mu_n}^{-1}(u)$.

Let then $u\in(0,1)$. Since, by monotonicity, $u\mapsto(F_\mu^{-1}(u),F_\nu^{-1}(u))$ is continuous du-almost everywhere on $(0,1)$ and, then, the weak convergence implies that

$$(F_{\mu_n}^{-1}(u),F_{\nu_n}^{-1}(u))\underset{n\to+\infty}{\longrightarrow}(F_\mu^{-1}(u),F_\nu^{-1}(u)), \tag{4.13}$$

we suppose without loss of generality that this convergence holds. For $n\in\mathbb{N}$, define $l_n=\inf_{k\ge n}x_u^k$ and $r_n=\sup_{k\ge n}x_u^k$. Since (4.13) holds, we find that $(l_n)_{n\in\mathbb{N}}$, resp. $(r_n)_{n\in\mathbb{N}}$, is a nondecreasing, resp. nonincreasing, sequence converging to x_u. Due to right continuity of F_μ and left continuity of $x\mapsto F_\mu(x-)$ we have

$$F_\mu(x_u-)=\lim_{p\to+\infty}F_\mu(l_p-)\quad\text{and}\quad\lim_{p\to+\infty}F_\mu(r_p)=F_\mu(x_u).$$

By Portmanteau's theorem and monotonicity of cumulative distribution functions we have

$$F_\mu(l_p-)\le\liminf_{n\to+\infty}F_{\mu_n}(l_p-)\le\liminf_{n\to+\infty}F_{\mu_n}(x_u^n-)\le\limsup_{n\to+\infty}F_{\mu_n}(x_u^n)\le\limsup_{n\to+\infty}F_{\mu_n}(r_p)\le F_\mu(r_p).$$

By taking the limit $p\to+\infty$, we find

$$F_\mu(x_u-)\le\liminf_{n\to+\infty}F_{\mu_n}(x_u^n-)\le\limsup_{n\to+\infty}F_{\mu_n}(x_u^n)\le F_\mu(x_u).$$

This implies (4.12) as soon as F_μ is continuous at x_u. Suppose now that F_μ is discontinuous at x_u. Since μ has countably many atoms, we may suppose without loss of generality that $u \in (F_\mu(x_u-), F_\mu(x_u))$. Let $(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^\mathbb{N}$ be the sequence associated to $x = x_u$ by (4.8). For n large enough, we have $u \in (F_{\mu_n}(x_n-), F_{\mu_n}(x_n))$, hence $x_n = x_u^n$. Using the assumption made in (4.8), we get

$$\begin{aligned}\liminf_{n\to+\infty} F_{\mu_n}(x_u^n) &= \liminf_{n\to+\infty}(F_{\mu_n}(x_u^n) \wedge F_\mu(x_u))\\ &= \liminf_{n\to+\infty}(F_{\mu_n}(x_u^n) \wedge F_\mu(x_u) - F_{\mu_n}(x_u^n-) \vee F_\mu(x_u-) + F_{\mu_n}(x_u^n-) \vee F_\mu(x_u-))\\ &= \mu(\{x_u\}) + \liminf_{n\to+\infty}(F_{\mu_n}(x_u^n-) \vee F_\mu(x_u-)) \geq F_\mu(x_u),\end{aligned}$$

hence $F_{\mu_n}(x_u^n) \underset{n\to+\infty}{\longrightarrow} F_\mu(x_u)$. Similarly, $F_{\mu_n}(x_u^n-) \underset{n\to+\infty}{\longrightarrow} F_\mu(x_u-)$, which shows (4.12) and concludes the proof. □

5 Appendix: adapted Wasserstein distances

A useful point of view is the following: for all $\mu,\nu \in \mathcal{P}_\rho(\mathbb{R})$ and $\pi \in \Pi(\mu,\nu)$, let $J(\pi)$ be the probability measure on $\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})$ defined by

$$J(\pi)(dx, dp) = \mu(dx)\,\delta_{\pi_x}(dp). \tag{5.1}$$

Then one can readily show that for any $\mu',\nu' \in \mathcal{P}_\rho(\mathbb{R})$ and $\pi' \in \Pi(\mu,\nu)$,

$$\mathcal{AW}_\rho(\pi,\pi') = \mathcal{W}_\rho(J(\pi), J(\pi')), \tag{5.2}$$

where $\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})$ is of course endowed with the product metric $((x,p),(x',p')) \mapsto (|x-x'|^\rho + \mathcal{W}_\rho^\rho(p,p'))^{1/\rho}$. Therefore, the topology induced by $\mathcal{AW}_\rho$ coincides with the initial topology with respect to J. This allows us to easily derive the two following lemmas.

Lemma 5.1. *Let $\rho \geq 1$, $\mu,\nu,\mu',\nu' \in \mathcal{P}_\rho(\mathbb{R})$ and $\pi \in \Pi(\mu,\nu), \pi' \in \Pi(\mu',\nu')$. Then there exists a coupling $\chi \in \Pi(\mu,\mu')$ optimal for $\mathcal{AW}_\rho(\pi,\pi')$, i.e. such that*

$$\mathcal{AW}_\rho^\rho(\pi,\pi') = \int_{\mathbb{R}\times\mathbb{R}} \left(|x-x'|^\rho + \mathcal{W}_\rho^\rho(\pi_x,\pi'_{x'})\right)\,\chi(dx,dx').$$

Remark 5.2. A similar statement holds when π,π' have three marginals. In that case, writing $\pi(dx,dy,dz) = \mu(dx)\,\pi_x(dy,dz)$ and $\pi'(dx',dy',dz') = \mu'(dx')\,\pi'_{x'}(dy',dz')$ we define

$$\mathcal{AW}_\rho^\rho(\pi,\pi') = \inf_{\chi\in\Pi(\mu,\mu')} \int_{\mathbb{R}\times\mathbb{R}} \left(|x-x'|^\rho + \mathcal{AW}_\rho^\rho(\pi_x,\pi'_{x'})\right)\,\chi(dx,dx').$$

Let $K(\pi)$ be the probability measure on $\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R}))$ defined by

$$K(\pi)(dx,dp) = \mu(dx)\,\delta_{J(\pi_x)}(dp).$$

Then one can readily show that

$$\mathcal{AW}_\rho(\pi,\pi') = \mathcal{W}_\rho(K(\pi), K(\pi')),$$

where $\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R}))$ is of course endowed with the product metric $((x,p),(x,p')) \mapsto \left(|x-x'|^\rho + \mathcal{W}_\rho^\rho(p,p')\right)^{1/\rho}$. Similarly to Lemma 5.1, the latter characterisation allows us to easily see that there exists a coupling $\chi \in \Pi(\mu,\mu')$ optimal for $\mathcal{AW}_\rho(\pi,\pi')$.

Proof of Lemma 5.1. Since $\mathbb{R}$ is Polish, so are the set $\mathcal{P}_\rho(\mathbb{R})$ and the set of probability measures on $\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})$. Hence there exists a coupling $P\in\Pi(J(\pi),J(\pi'))$ optimal for $\mathcal{W}_\rho(J(\pi),J(\pi'))$, i.e.

$$\mathcal{W}_\rho^\rho(J(\pi),J(\pi'))=\int_{\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})\times\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})}|(x,p)-(x',p')|^\rho\,P(dx,dp,dx',dp').$$

Since the $J(\pi)$ and $J(\pi')$ are concentrated on graphs of measurable maps, it is clear that $P(dx,dp,dx',dp')=\chi(dx,dx')\,\delta_{\pi_x}(dp)\,\delta_{\pi'_{x'}}(dp')$ for $\chi(dx,dx')=\int_{(p,p')\in\mathcal{P}_\rho(\mathbb{R})\times\mathcal{P}_\rho(\mathbb{R})}P(dx,dp,dx',dp')\in\Pi(\mu,\mu')$. Then

$$\begin{aligned}\mathcal{AW}_\rho^\rho(\pi,\pi')&=\mathcal{W}_\rho^\rho(J(\pi),J(\pi'))\\&=\int_{\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})\times\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})}\left(|x-x'|^\rho+\mathcal{W}_\rho^\rho(p,p')\right)\,\chi(dx,dx')\,\delta_{\pi_x}(dp)\,\delta_{\pi'_{x'}}(dp')\\&=\int_{\mathbb{R}\times\mathbb{R}}\left(|x-x'|^\rho+\mathcal{W}_\rho^\rho(\pi_x,\pi'_{x'})\right)\,\chi(dx,dx'),\end{aligned}$$

hence χ is optimal for $\mathcal{AW}_\rho(\pi,\pi')$. ☐

Lemma 5.3. *Let $\rho\geq 1$, $\mu,\nu\in\mathcal{P}_\rho(\mathbb{R})$, $(\mu_n)_{n\in\mathbb{N}},(\nu_n)_{n\in\mathbb{N}}\in\mathcal{P}_\rho(\mathbb{R})^{\mathbb{N}}$, $\pi\in\Pi(\mu,\nu)$ and $(\pi_n)_{n\in\mathbb{N}}\in\prod_{n\in\mathbb{N}}\Pi(\mu_n,\nu_n)$. Then*

$$\mathcal{AW}_\rho(\pi_n,\pi)\underset{n\to+\infty}{\longrightarrow}0\iff\mathcal{AW}_1(\pi_n,\pi)+\mathcal{W}_\rho(\mu_n,\mu)+\mathcal{W}_\rho(\nu_n,\nu)\underset{n\to+\infty}{\longrightarrow}0.\tag{5.3}$$

Proof. Clearly,

$$\begin{aligned}\int_{\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R})}\left(|x|^\rho+\mathcal{W}_\rho^\rho(p,\delta_0)\right)\,J(\pi)(dx,dp)&=\int_{\mathbb{R}}\left(|x|^\rho+\int_{\mathbb{R}}|y|^\rho\,\pi_x(dy)\right)\mu(dx)\\&=\int_{\mathbb{R}}|x|^\rho\,\mu(dx)+\int_{\mathbb{R}}|y|^\rho\,\nu(dy)\end{aligned}$$

so that π and $J(\pi)$ have equal ρ-th moments. Since convergence in $\mathcal{W}_\rho$ is equivalent to convergence in $\mathcal{W}_1$ coupled with convergence of the ρ-th moments, we deduce from (5.2) that

$$\begin{aligned}&\mathcal{AW}_\rho(\pi_n,\pi)\underset{n\to+\infty}{\longrightarrow}0\\&\iff\mathcal{AW}_1(\pi_n,\pi)+\left|\int_{\mathbb{R}}|x|^\rho\,\mu_n(dx)+\int_{\mathbb{R}}|y|^\rho\,\nu_n(dy)-\int_{\mathbb{R}}|x|^\rho\,\mu(dx)-\int_{\mathbb{R}}|y|^\rho\,\nu(dy)\right|\underset{n\to+\infty}{\longrightarrow}0.\end{aligned}$$

Since $\mathcal{W}_1\leq\mathcal{AW}_1$ and $\mathcal{W}_1$-convergence of the couplings implies that of their respective marginals and $\liminf_{n\to\infty}\int_{\mathbb{R}}|x|^\rho\,\mu_n(dx)\geq\int_{\mathbb{R}}|x|^\rho\,\mu(dx)$, $\liminf_{n\to\infty}\int_{\mathbb{R}}|y|^\rho\,\nu_n(dy)\geq\int_{\mathbb{R}}|y|^\rho\,\nu(dy)$, using the fact that convergence in $\mathcal{W}_\rho$ is equivalent to convergence in $\mathcal{W}_1$ coupled with convergence of the ρ-th moments again, we conclude that the right-hand side is clearly equivalent to

$$\mathcal{AW}_1(\pi_n,\pi)+\mathcal{W}_\rho(\mu_n,\mu)+\mathcal{W}_\rho(\nu_n,\nu)\underset{n\to+\infty}{\longrightarrow}0,$$

which proves (5.4). ☐

Remark 5.4. For $\rho\geq 1$, let $\lambda,\mu,\nu\in\mathcal{P}_\rho(\mathbb{R})$, $(\lambda_n)_{n\in\mathbb{N}},(\mu_n)_{n\in\mathbb{N}},(\nu_n)_{n\in\mathbb{N}}\in\mathcal{P}_\rho(\mathbb{R})^{\mathbb{N}}$, $\pi\in\mathcal{P}(\mathbb{R}\times\mathbb{R}\times\mathbb{R})$ with marginals λ,μ,ν and $\pi_n\in\mathcal{P}(\mathbb{R}\times\mathbb{R}\times\mathbb{R})$ with marginals λ_n,μ_n,ν_n for $n\in\mathbb{N}$. Then

$$\mathcal{AW}_\rho(\pi_n,\pi)\underset{n\to+\infty}{\longrightarrow}0\iff\mathcal{AW}_1(\pi_n,\pi)+\mathcal{W}_\rho(\lambda_n,\lambda)+\mathcal{W}_\rho(\mu_n,\mu)+\mathcal{W}_\rho(\nu_n,\nu)\underset{n\to+\infty}{\longrightarrow}0.\tag{5.4}$$

This can be proved by the same argument as in the previous proof since, for the mapping K introduced in Remark 5.2, one has

$$\begin{aligned}\int_{\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R}\times\mathcal{P}_\rho(\mathbb{R}))}\left(|x|^\rho+\mathcal{W}_\rho^\rho(p,\delta_{(0,\delta_0)})\right)K(\pi)(dx,dp)&=\int_{\mathbb{R}}\left(|x|^\rho+\mathcal{W}_\rho^\rho(J(\pi_x),\delta_{(0,\delta_0)})\right)\lambda(dx)\\&=\int_{\mathbb{R}}|x|^\rho\lambda(dx)+\int_{\mathbb{R}\times\mathbb{R}}\left(|y|^\rho+\mathcal{W}_\rho^\rho(\pi_{x,y},\delta_0)\right)\pi_x(dy,\mathbb{R})\lambda(dx)\\&=\int_{\mathbb{R}}|x|^\rho\lambda(dx)+\int_{\mathbb{R}}|y|^\rho\mu(dy)+\int_{\mathbb{R}}|z|^\rho\nu(dz).\end{aligned}$$

References

[1] D. J. Aldous. Weak convergence and general theory of processes. Unpublished incomplete draft of monograph; Department of Statistics, University of California, Berkeley, CA 94720, 1981.

[2] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis: A hitchhiker's guide*. Springer, 3rd edition, 2006.

[3] J. Backhoff-Veraguas, D. Bartl, M. Beiglböck, and M. Eder. Adapted Wasserstein distances and stability in mathematical finance. *Finance and Stochastics*, 24(3):601–632, 2020.

[4] J. Backhoff-Veraguas, D. Bartl, M. Beiglböck, and M. Eder. All adapted topologies are equal. *Probab. Theory Related Fields*, 178(3-4):1125–1172, 2020.

[5] J. Backhoff-Veraguas, D. Bartl, M. Beiglböck, and J. Wiesel. Estimating processes in adapted Wasserstein distance. *arXiv e-prints:2002.07261*, 2020.

[6] J. Backhoff-Veraguas and G. Pammer. Stability of martingale optimal transport and weak optimal transport. *Annals of Applied probability*, to appear.

[7] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: A mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.

[8] M. Beiglböck, B. Jourdain, W. Margheriti, and G. Pammer. Stability of the weak martingale optimal transport problem. *arXiv e-prints:2109.06322*, 2021.

[9] M. Beiglböck, B. Jourdain, W. Margheriti, and G. Pammer. Approximation of martingale couplings on the line in the weak adapted topology. *Probability Theory and related Fields*, to appear.

[10] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Annals of Probability*, 44(1):42–106, 2016.

[11] M. Beiglböck and N. Juillet. Shadow couplings. *Trans. Amer. Math. Soc.*, 374(7):4973–5002, 2021.

[12] M. Beiglböck, T. Lim, and J. Obłój. Dual attainment for the martingale transport problem. *Bernoulli*, 25(3):1640–1658, 2019.

[13] M. Beiglböck, M. Nutz, and N. Touzi. Complete Duality for Martingale Optimal Transport on the Line. *Annals of Probability*, 45(5):3038–3074, 2017.

[14] J. Bion-Nadal and D. Talay. On a Wasserstein-type distance between solutions to stochastic differential equations. *Annals of Applied Probability*, 29(3):1609–1639, 2019.

[15] M. Brückerhoff and N. Juillet. Instability of martingale optimal transport in dimension d $\geq$ 2. *arXiv e-print:2101.06964*, 2021.

[16] H. De March. Local structure of multi-dimensional martingale optimal transport. *arXiv e-prints:1805.09469*, 2018.

[17] H. De March. Quasi-sure duality for multi-dimensional martingale optimal transport. *arXiv e-prints:1805.01757*, 2018.

[18] H. De March and N. Touzi. Irreducible convex paving for decomposition of multi-dimensional martingale transport plans. *Annals of Probability*, 47(3):1726–1774, 2019.

[19] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Annals of Applied Probability*, 24(1):312–336, 2014.

[20] S. Gerhold and I. C. Gülüm. Peacocks nearby: Approximating sequences of measures. *Stochastic Processes and their Applications*, 129(7):2406–2436, 2019.

[21] S. Gerhold and I. C. Gülüm. Consistency of option prices under bid–ask spreads. *Mathematical Finance*, 30(2):377–402, 2020.

[22] N. Ghoussoub, Y.-H. Kim, and T. Lim. Structure of optimal martingale transport plans in general dimensions. *Ann. Probab.*, 47(1):109–164, 2019.

[23] M. F. Hellwig. Sequential decisions under uncertainty and the maximum theorem. *Journal of Mathematical Economics*, 25(4):443–464, 1996.

[24] P. Henry-Labordère, X. Tan, and N. Touzi. An explicit version of the one-dimensional Brenier's theorem with full marginals constraint. *Stochastic Processes and their Applications*, 126(9):2800–2834, 2016.

[25] P. Henry-Labordère and N. Touzi. An explicit martingale version of the one-dimensional Brenier theorem. *Finance and Stochastics*, 20(3):635–668, 2016.

[26] B. Jourdain and W. Margheriti. A new family of one dimensional martingale couplings. *Electronic Journal of Probability*, 25, 2020.

[27] B. Jourdain and W. Margheriti. One dimensional martingale rearrangement couplings. *arXiv e-prints:2101.12651*, 2021.

[28] L. V. Kantorovich. On the translocation of masses. *Doklady Akademii Nauk SSSR*, 37(7-8):199–201, 1942.

[29] R. Lassalle. Causal transference plans and their Monge-Kantorovich problems. *Stochastic Analysis and Applications*, 36(3):452–484, 2018.

[30] G. Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'académie Royale des Sciences de Paris*, 1781.

[31] G. C. Pflug and A. Pichler. A distance for multistage stochastic optimization models. *SIAM Journal on Optimization*, 22(1):1–23, 2012.

[32] G. C. Pflug and A. Pichler. *Multistage stochastic optimization*. Springer Series in Operations Research and Financial Engineering. Springer, Cham, 2014.

[33] G. C. Pflug and A. Pichler. Dynamic generation of scenario trees. *Computational Optimization and Applications*, 62(3):641–668, 2015.

[34] G. C. Pflug and A. Pichler. From empirical observations to tree models for stochastic optimization: convergence properties. *SIAM Journal on Optimization*, 26(3):1715–1740, 2016.

[35] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume I: Theory*. Springer Science & Business Media, 1998.

[36] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg, 3 edition, 1999.

[37] L. Rüschendorf. On the minimum discrimination information theorem. *Statistics & Decisions Supplement Issue*, 1:263–283, 1984.

[38] V. Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36(2):423–439, 1965.

[39] J. Wiesel. Continuity of the martingale optimal transport problem on the real line. *arXiv e-prints:1905.04574*, 2020.