Electronic Journal of Probability

Electron. J. Probab. **25** (2020), article no. 136, 1–50.
ISSN: 1083-6489 https://doi.org/10.1214/20-EJP543

A new family of one dimensional martingale couplings

B. Jourdain* W. Margheriti*

Abstract

In this paper, we exhibit a new family of martingale couplings between two one-dimensional probability measures μ and ν in the convex order. This family is parametrised by two dimensional probability measures on the unit square with respective marginal densities proportional to the positive and negative parts of the difference between the quantile functions of μ and ν. It contains the inverse transform martingale coupling which is explicit in terms of the quantile functions of these marginal densities. The integral of $|x-y|$ with respect to each of these couplings is smaller than twice the $\mathcal{W}_1$ distance between μ and ν. When the comonotonous coupling between μ and ν is given by a map T, the elements of the family minimise $\int_{\mathbb{R}} |y-T(x)|\, M(dx,dy)$ among all martingale couplings between μ and ν. When μ and ν are in the decreasing (resp. increasing) convex order, the construction is generalised to exhibit super (resp. sub) martingale couplings.

Keywords: convex order; martingale optimal transport; Wasserstein distance; martingale couplings.
MSC2020 subject classifications: 60G42; 60E15; 91G80.
Submitted to EJP on March 7, 2019, final version accepted on October 22, 2020.
Supersedes arXiv:1808.01390v2.
Supersedes HAL:hal-01876809.

1 Introduction

For all $d \in \mathbb{N}^*$, $\rho \geq 1$ and μ,ν in the set $\mathcal{P}_\rho(\mathbb{R}^d)$ of probability measures on $\mathbb{R}^d$ with finite ρ-th moment, we define the Wasserstein distance with index ρ by $\mathcal{W}_\rho(\mu,\nu) = (\inf_{P\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^\rho\, P(dx,dy))^{1/\rho}$, where $\Pi(\mu,\nu)$ denotes the set of couplings between μ and ν, that is $\Pi(\mu,\nu) = \{P \in \mathcal{P}_1(\mathbb{R}^d\times\mathbb{R}^d) \mid \forall A \in \mathcal{B}(\mathbb{R}^d),\ P(A\times\mathbb{R}^d) = \mu(A) \text{ and } P(\mathbb{R}^d\times A) = \nu(A)\}$. Let $\Pi^{\mathrm{M}}(\mu,\nu)$ be the set of martingale couplings between μ and ν, that is

$$\Pi^{\mathrm{M}}(\mu,\nu) = \left\{M \in \Pi(\mu,\nu) \mid \mu(dx)\text{-a.e., } \int_{\mathbb{R}^d} |y|\, m(x,dy) < +\infty \text{ and } \int_{\mathbb{R}^d} y\, m(x,dy) = x\right\}.$$

*Université Paris-Est, Cermics (ENPC), INRIA F-77455 Marne-la-Vallée, France. E-mail: benjamin.jourdain@enpc.fr,william.margheriti@enpc.fr – This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque.

The celebrated Strassen theorem [22] ensures that if $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$, then $\Pi^{\mathrm{M}}(\mu, \nu) \neq \emptyset$ iff μ and ν are in the convex order. We recall that two probability measures $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ are in the convex order, and denote $\mu \leq_{cx} \nu$, if $\int_{\mathbb{R}^d} f(x)\,\mu(dx) \leq \int_{\mathbb{R}^d} f(y)\,\nu(dy)$ for any convex function $f : \mathbb{R}^d \to \mathbb{R}$. We denote $\mu <_{cx} \nu$ if $\mu \leq_{cx} \nu$ and $\mu \neq \nu$. For all $\rho \geq 1$ and $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$, we define $\mathcal{M}_\rho(\mu, \nu)$ by

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

Our main result is the following stability inequality which shows that if μ and ν are in the convex order and close to each other, then there exists a martingale coupling which expresses this proximity:

$$\forall \mu, \nu \in \mathcal{P}_1(\mathbb{R}) \text{ such that } \mu \leq_{cx} \nu, \quad \mathcal{M}_1(\mu, \nu) \leq 2\mathcal{W}_1(\mu, \nu). \tag{1.1}$$

It is well known (see for instance [23, Remark 2.19 (ii) Chapter 2]) that for all $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R})$,

$$\mathcal{W}_\rho(\mu, \nu) = \left(\int_0^1 \left| F_\mu^{-1}(u) - F_\nu^{-1}(u) \right|^\rho \, du \right)^{1/\rho}, \tag{1.2}$$

where we denote by $F_\eta(x) = \eta((-\infty, x])$ and $F_\eta^{-1}(u) = \inf\{x \in \mathbb{R} \mid F_\eta(x) \geq u\}, u \in (0, 1)$, the cumulative distribution function and the quantile function of a probability measure η on $\mathbb{R}$. We prove the inequality (1.1) by exhibiting a new family of martingale couplings M such that $\int_{\mathbb{R}\times\mathbb{R}} |x - y| \, M(dx, dy) \leq 2\mathcal{W}_1(\mu, \nu)$. We will show (see the proof of Theorem 2.12) that the constant 2 is sharp in (1.1). We will also see that (1.1) cannot be generalised with $\mathcal{M}_1(\mu, \nu)$ and $\mathcal{W}_1(\mu, \nu)$ replaced with $\mathcal{M}_\rho(\mu, \nu)$ and $\mathcal{W}_\rho(\mu, \nu)$ for $\rho > 1$. The case $\rho = 2$ is easy, since for all $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$ and $M \in \Pi^{\mathrm{M}}(\mu, \nu)$, $\int_{\mathbb{R}\times\mathbb{R}} |x - y|^2 \, M(dx, dy) = \int_{\mathbb{R}} y^2 \, \nu(dy) - \int_{\mathbb{R}} x^2 \, \mu(dx)$, which is independent from M. For all $n \in \mathbb{N}^*$, let μ_n be the centred Gaussian distribution with variance n^2. Then we get that $\mathcal{M}_2(\mu_n, \mu_{n+1}) = \sqrt{2n + 1} \underset{n\to+\infty}{\longrightarrow} +\infty$, whereas $\mathcal{W}_\rho(\mu_n, \mu_{n+1}) = (\int_0^1 |nF_{\mu_1}^{-1}(u) - (n + 1)F_{\mu_1}^{-1}(u)|^\rho \, du)^{1/\rho} = \mathbb{E}[|G|^\rho]^{1/\rho} < +\infty$ for $G \sim \mathcal{N}_1(0, 1)$, which makes the equivalent of (1.1) impossible to hold. Extension to the case $\rho > 2$ is immediate with the same example thanks to Jensen's inequality which provides $\mathcal{M}_\rho(\mu_n, \mu_{n+1}) \geq \mathcal{M}_2(\mu_n, \mu_{n+1}) = \sqrt{2n + 1}$, whereas $\mathcal{W}_\rho(\mu_n, \mu_{n+1})$ is still bounded.

This problem is motivated by the resolution of the Martingale Optimal Transport (MOT) problem introduced by Beiglböck, Henry-Labordère and Penkner [3] in a discrete time setting, and Galichon, Henry-Labordère and Touzi [12] in a continuous time setting. For adaptations of celebrated results on classical optimal transport theory to the MOT problem, we refer to Henry-Labordère, Tan and Touzi [14] and Henry-Labordère and Touzi [15]. To tackle numerically the MOT problem, we refer to Alfonsi, Corbetta and Jourdain [1], Alfonsi, Corbetta and Jourdain [2], De March [7] and Guo and Obłój [13]. On duality, we refer to Beiglböck, Nutz and Touzi [6], Beiglböck, Lim and Obłój [5] and De March [9]. We also refer to De March [8] and De March and Touzi [10] for the multi-dimensional case. Once the martingale optimal transport problem is discretised by approximating μ and ν by probability measures with finite support and in the convex order, one can raise the question of the convergence of the discrete optimal cost towards the continuous one. The present paper is a step forward in proving the stability of the martingale optimal transport problem with respect to the marginals.

We develop in Section 2 an abstract construction of a new family of martingale couplings between two probability measures μ and ν on the real line with finite first moments and comparable in the convex order. This family is parametrised by two dimensional probability measures on the unit square with respective marginal densities proportional to the positive and the negative parts of the difference $F_\mu^{-1} - F_\nu^{-1}$ between the quantile functions of μ and ν. Moreover, each martingale coupling in the family is obtained as the image of $\mathbb{1}_{(0,1)}(u)\,du\,\widetilde{m}^Q(u,dy)$ by $(u,y) \mapsto (F_\mu^{-1}(u), y)$ where $\widetilde{m}^Q$ is a Markov kernel on $(0,1)\times\mathbb{R}$ such that $\int_{(0,1)} \widetilde{m}^Q(u, \{y \in \mathbb{R} \mid |y - F_\nu^{-1}(u)| = (y - F_\nu^{-1}(u))\mathrm{sg}(F_\mu^{-1}(u) - F_\nu^{-1}(u))\})du = 1$, where $\mathrm{sg}(x) = \mathbb{1}_{\{x>0\}} - \mathbb{1}_{\{x<0\}}$ for $x \in \mathbb{R}$. Therefore, for (U,Y) distributed according to $\mathbb{1}_{(0,1)}(u)\,du\,\widetilde{m}^Q(u,dy)$, $(F_\mu^{-1}(U), Y)$ is a martingale coupling and

$$\begin{aligned}\mathbb{E}[|Y - F_\nu^{-1}(U)|] &= \mathbb{E}[\mathrm{sg}(F_\mu^{-1}(U) - F_\nu^{-1}(U))\mathbb{E}[Y - F_\nu^{-1}(U)|U]] \\ &= \mathbb{E}[|F_\mu^{-1}(U) - F_\nu^{-1}(U)|] = \mathcal{W}_1(\mu,\nu).\end{aligned} \tag{1.3}$$

When the comonotonous coupling between μ and ν, that is the probability distribution of $(F_\mu^{-1}(U), F_\nu^{-1}(U))$, is given by a map T, the elements of the family minimise $\int_\mathbb{R} |y - T(x)|\,M(dx,dy)$ among all martingale couplings between μ and ν. We deduce from (1.3) that $\mathbb{E}[|Y - F_\mu^{-1}(U)|] \le \mathbb{E}[|Y - F_\nu^{-1}(U)|] + \mathbb{E}[|F_\nu^{-1}(U) - F_\mu^{-1}(U)|] = 2\mathcal{W}_1(\mu,\nu)$ which implies (1.1) as soon as the parameter set of probability measures on the unit square is non empty.

In Section 3, we give an explicit example of such a probability measure on the unit square. We call the associated martingale coupling the inverse transform martingale coupling. This coupling is explicit in terms of the cumulative distribution functions of the above-mentioned densities and their left-continuous generalised inverses. It is therefore more explicit than the left-curtain (and right-curtain) coupling introduced by Beiglböck and Juillet [4] which under the condition that ν has no atoms and the set of local maximal values of $F_\nu - F_\mu$ is finite can be made explicit according to Henry-Labordère and Touzi [15] by solving two coupled ordinary differential equations starting from each right-most local maximiser. We also check that the inverse transform martingale coupling is stable with respect to its marginals μ and ν for the Wasserstein distance. The building brick of the inverse transform martingale coupling is a martingale coupling between $\mu_{u,v} = p\delta_{F_\mu^{-1}(u)} + (1-p)\delta_{F_\mu^{-1}(v)}$ and $\nu_{u,v} = p\delta_{F_\nu^{-1}(u)} + (1-p)\delta_{F_\nu^{-1}(v)}$ with $0 < u < v < 1$ such that

$$F_\nu^{-1}(u) < F_\mu^{-1}(u) < F_\mu^{-1}(v) < F_\nu^{-1}(v), \tag{1.4}$$

where we choose a common weight p (resp. $1-p$) for $F_\mu^{-1}(u)$ and $F_\nu^{-1}(u)$ (resp. $F_\mu^{-1}(v)$ and $F_\nu^{-1}(v)$) to help ensuring that the second marginal is equal to ν when the first is equal to μ. Then p is given by the equality of the means which in view of the condition (1.4) on the supports is equivalent to the convex order between $\mu_{u,v}$ and $\nu_{u,v}$: $\frac{1-p}{p} = \frac{F_\mu^{-1}(u) - F_\nu^{-1}(u)}{F_\nu^{-1}(v) - F_\mu^{-1}(v)}$. We rely on the necessary condition of [21, Theorem 3.A.5 Chapter 3]: $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ are such that $\mu \le_{cx} \nu$ iff for all $u \in [0,1]$, $\int_0^u F_\mu^{-1}(v)\,dv \ge \int_0^u F_\nu^{-1}(v)dv$ with equality for $u = 1$. This implies that for all $u \in [0,1]$, $\Psi_+(u) := \int_0^u (F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \ge \int_0^u (F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv := \Psi_-(u)$ where $x^+ := \max(x,0)$ and $x^- := \max(-x,0)$ respectively denote the positive and negative parts of a real number x. We now choose $v = \Psi_-^{-1}(\Psi_+(u))$ where Ψ_-^{-1} is the left-continuous generalised inverse of Ψ_-. Then $d\Psi_+(u)$ a.e. $u < v$ (consequence of $\Psi_- \le \Psi_+$) and $F_\nu^{-1}(u) < F_\mu^{-1}(u) < F_\mu^{-1}(v) < F_\nu^{-1}(v)$ (consequence of the definitions of Ψ_+ and Ψ_-, see Section 3.1). Moreover the key equality $\frac{dv}{du} = \frac{(F_\mu^{-1} - F_\nu^{-1})^+(u)}{(F_\mu^{-1} - F_\nu^{-1})^-(v)} = \frac{1-p}{p}$ explains why the construction succeeds. More details are given in Section 3.

The cardinality of this new family of martingale couplings between μ and ν is discussed in Section 4. This family is shown to be convex and is therefore either a singleton like when ν only weighs two points, or uncountably infinite like when $\mu(\{x\}) = \nu(\{x\}) = 0$ for all $x \in \mathbb{R}$.

The construction is finally generalised in Section 5 to exhibit super (resp. sub) martingale couplings as soon as μ is smaller than ν in the decreasing (resp. increasing) convex order. We recall that two probability measures $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ are in the decreasing (resp. increasing) convex order and denote $\mu\le_{dcx}\nu$ (resp. $\mu\le_{icx}\nu$) if $\int_{\mathbb{R}} f(x)\,\mu(dx)\le\int_{\mathbb{R}} f(x)\,\nu(dx)$ for any decreasing (resp. increasing) convex function $f:\mathbb{R}\to\mathbb{R}$. In particular, we generalise the stability inequality to the super (resp. sub) martingale case.

Throughout the present article, a capital letter M which denotes a coupling between μ and ν is associated to its small letter m which denotes the regular conditional probability distribution of M with respect to μ, that is the (μ-a.e.) unique Markov kernel such that $M(dx,dy)=\mu(dx)\,m(x,dy)$.

2 A new family of martingale couplings

2.1 A simple example

Let us construct a coupling in dimension 1 which shows that (1.1) holds true in a simple case. We say that a centred probability measure $\mu\in\mathcal{P}_1(\mathbb{R})$ is symmetric if $\mu=\bar\mu$, where $\bar\mu$ denotes the image of μ by $x\mapsto -x$. Let then μ and ν be centred and symmetric probability measures on $\mathbb{R}$ such that $F_\mu^{-1}(u)\ge F_\nu^{-1}(u)$ for all $u\in(0,1/2]$ and $F_\mu^{-1}(u)\le F_\nu^{-1}(u)$ for all $u\in(1/2,1)$. Let U be a random variable uniformly distributed on $(0,1)$. According to the inverse transform sampling, the probability distributions of $F_\mu^{-1}(U)$ and $F_\nu^{-1}(U)$ are respectively μ and ν. Let Y be the random variable defined by

$$Y=F_\nu^{-1}(U)\mathbb{1}_{\left\{F_\nu^{-1}(U)\neq 0,\,V\le\frac{F_\mu^{-1}(U)+F_\nu^{-1}(U)}{2F_\nu^{-1}(U)}\right\}}-F_\nu^{-1}(U)\mathbb{1}_{\left\{F_\nu^{-1}(U)\neq 0,\,V>\frac{F_\mu^{-1}(U)+F_\nu^{-1}(U)}{2F_\nu^{-1}(U)}\right\}},\qquad(2.1)$$

where V is a random variable uniformly distributed on $(0,1)$ independent from U. It is clear by symmetry of μ that $F_\mu(0)\ge 1/2$, so $F_\mu^{-1}(1/2)\le 0$. Moreover, for all $x\in\mathbb{R}$ and $u>1/2$, $F_\mu(x)\ge u$ implies $x\ge 0$, so $F_\mu^{-1}(u)\ge 0$. Therefore, we have

$$\forall u\in(0,1/2],\quad F_\nu^{-1}(u)\le F_\mu^{-1}(u)\le 0\quad\text{and}\quad\forall u\in(1/2,1),\quad 0\le F_\mu^{-1}(u)\le F_\nu^{-1}(u).\qquad(2.2)$$

In particular, when $F_\nu^{-1}(U)=0$, then $F_\mu^{-1}(U)=0$ and $Y=0$. Let us check that Y is distributed according to ν. Using that $(F_\mu^{-1}(U),F_\nu^{-1}(U))$ and $(-F_\mu^{-1}(U),-F_\nu^{-1}(U))$ are identically distributed (see Lemma 6.5 below in Section 6), we have for all measurable and bounded functions $h:\mathbb{R}\to\mathbb{R}$,

$$\begin{aligned}\mathbb{E}[h(Y)]&=\mathbb{E}\left[h(F_\nu^{-1}(U))\mathbb{1}_{\left\{F_\nu^{-1}(U)\neq 0,\,V\le\frac{F_\mu^{-1}(U)+F_\nu^{-1}(U)}{2F_\nu^{-1}(U)}\right\}}\right]\\&\quad+\mathbb{E}\left[h(F_\nu^{-1}(U))\mathbb{1}_{\left\{F_\nu^{-1}(U)\neq 0,\,V>\frac{F_\mu^{-1}(U)+F_\nu^{-1}(U)}{2F_\nu^{-1}(U)}\right\}}\right]+h(0)\mathbb{P}(F_\nu^{-1}(U)=0)\\&=\mathbb{E}[h(F_\nu^{-1}(U))].\end{aligned}$$

Moreover, according to (2.2), we have $\frac{F_\mu^{-1}(u)+F_\nu^{-1}(u)}{2F_\nu^{-1}(u)}\in[0,1]$ for all $u\in(0,1)$ such that $F_\nu^{-1}(u)\neq 0$. In addition to that, we have

$$F_\nu^{-1}(u)\frac{F_\mu^{-1}(u)+F_\nu^{-1}(u)}{2F_\nu^{-1}(u)}-F_\nu^{-1}(u)\left(1-\frac{F_\mu^{-1}(u)+F_\nu^{-1}(u)}{2F_\nu^{-1}(u)}\right)=F_\mu^{-1}(u),$$

for all $u\in(0,1)$ such that $F_\nu^{-1}(u)\neq 0$. So $\mathbb{E}[Y|U]=F_\mu^{-1}(U)\mathbb{1}_{\{F_\nu^{-1}(U)\neq 0\}}=F_\mu^{-1}(U)$ since $F_\nu^{-1}(U)=0$ implies $F_\mu^{-1}(U)=0$. So we deduce that $\mathbb{E}[Y|F_\mu^{-1}(U)]=F_\mu^{-1}(U)$. Therefore, the law of $(F_\mu^{-1}(U),Y)$ is an explicit martingale coupling between μ and ν.

Furthermore, remarking that $|Y-F_\nu^{-1}(U)|=(Y-F_\nu^{-1}(U))\mathrm{sg}(F_\mu^{-1}(U)-F_\nu^{-1}(U))$, we deduce from the equality (1.3) that $\mathbb{E}[|Y-F_\mu^{-1}(U)|]\le\mathbb{E}[|Y-F_\nu^{-1}(U)|]+\mathbb{E}[|F_\nu^{-1}(U)-F_\mu^{-1}(U)|]=2\mathcal{W}_1(\mu,\nu)$, so (1.1) holds.

2.2 Definition

Let μ and ν be two probability measures on $\mathbb{R}$ with finite first moment such that $\int_{\mathbb{R}} x\,\mu(dx) = \int_{\mathbb{R}} y\,\nu(dy)$ and $\mu \neq \nu$. We recall that Ψ_+ and Ψ_- are defined for all $u \in [0,1]$ by $\Psi_+(u) = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv$ and $\Psi_-(u) = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv$. Let $\mathcal{U}_+$, $\mathcal{U}_-$ and $\mathcal{U}_0$ be defined by

$$\mathcal{U}_+ = \{u \in (0,1) \mid F_\mu^{-1}(u) > F_\nu^{-1}(u)\}, \quad \mathcal{U}_- = \{u \in (0,1) \mid F_\mu^{-1}(u) < F_\nu^{-1}(u)\}$$
$$\text{and} \quad \mathcal{U}_0 = \{u \in (0,1) \mid F_\mu^{-1}(u) = F_\nu^{-1}(u)\}. \tag{2.3}$$

Notice that $d\Psi_+(u)$-a.e. (resp. $d\Psi_-(u)$-a.e.), we have $u \in \mathcal{U}_+$ (resp. $u \in \mathcal{U}_-$). Since μ and ν have equal means, we can set $\gamma = \int_0^1 (F_\mu^{-1} - F_\nu^{-1})^+(u)\,du = \int_0^1 (F_\mu^{-1} - F_\nu^{-1})^-(u)\,du \in (0,+\infty)$. We note $\mathcal{Q}$ the set of probability measures $Q(du,dv)$ on $(0,1)^2$ such that

(i) Q has first marginal $\frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du = \frac{1}{\gamma}\,d\Psi_+(u)$;

(ii) Q has second marginal $\frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv = \frac{1}{\gamma}\,d\Psi_-(v)$;

(iii) $Q\left(\{(u,v) \in (0,1)^2 \mid u < v\}\right) = 1$.

Example 2.1. Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Suppose that the difference of the quantile functions changes sign only once, that is there exists $p \in (0,1)$ such that $u \mapsto \int_0^u (F_\mu^{-1}(v) - F_\nu^{-1}(v))\,dv$ is nondecreasing on $[0,p]$ and nonincreasing on $[p,1]$. Then one can easily see that any probability measure Q defined on $(0,1)$ satisfying properties (i) and (ii) of the definition of $\mathcal{Q}$ is concentrated on $(0,p) \times (p,1)$ and therefore satisfies (iii). In particular, the probability measure Q_1 defined on $(0,1)^2$ by

$$Q_1(du,dv) = \frac{1}{\gamma^2}(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du\,(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv \tag{2.4}$$

is an element of $\mathcal{Q}$.

In view of (i) and (ii), one could rewrite (iii) as $Q\left(\{(u,v) \in \mathcal{U}_+ \times \mathcal{U}_- \mid u < v\}\right) = 1$. A characterisation of the support of Q in terms of the irreducible components of μ and ν is given by Proposition 2.8 below. In the general case, the construction of a probability measure $Q \in \mathcal{Q}$ is not straightforward, but a direct consequence of Proposition 3.1 below is that $\mathcal{Q}$ is non-empty as long as $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ are such that $\mu <_{cx} \nu$. Moreover, the convexity of $\mathcal{Q}$ is clear.

Proposition 2.2. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Then $\mathcal{Q}$ is a non-empty convex set.*

Let Q be an element of $\mathcal{Q}$. Let π_-^Q and π_+^Q be two sub-Markov kernels on $(0,1)$ such that for du-almost all $u \in \mathcal{U}_+$ and dv-almost all $v \in \mathcal{U}_-$, $\pi_+^Q(u,(0,1)) = 1$, $\pi_-^Q(v,(0,1)) = 1$ and

$$Q(du,dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du\,\pi_+^Q(u,dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv\,\pi_-^Q(v,du).$$

Let $(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$ be the Markov kernel defined by

$$\begin{cases} \displaystyle\int_{v\in(0,1)} \frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\,\delta_{F_\nu^{-1}(v)}(dy)\,\pi_+^Q(u,dv) \\ \quad + \displaystyle\int_{v\in(0,1)} \frac{F_\nu^{-1}(v)-F_\mu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\,\pi_+^Q(u,dv)\,\delta_{F_\nu^{-1}(u)}(dy) \\ \quad \text{for } u\in\mathcal{U}_+ \text{ such that } \pi_+^Q(u,\{v\in(0,1)\mid F_\nu^{-1}(v)>F_\mu^{-1}(u)\})=1; \\ \displaystyle\int_{v\in(0,1)} \frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\,\delta_{F_\nu^{-1}(v)}(dy)\,\pi_-^Q(u,dv) \\ \quad + \displaystyle\int_{v\in(0,1)} \frac{F_\nu^{-1}(v)-F_\mu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\,\pi_-^Q(u,dv)\,\delta_{F_\nu^{-1}(u)}(dy) \\ \quad \text{for } u\in\mathcal{U}_- \text{ such that } \pi_-^Q(u,\{v\in(0,1)\mid F_\nu^{-1}(v)<F_\mu^{-1}(u)\})=1; \\ \delta_{F_\nu^{-1}(u)}(dy) \qquad \text{otherwise.} \end{cases} \tag{2.5}$$

For any Markov kernel $(\widetilde{m}(u,dy))_{u\in(0,1)}$, we denote by $(m(x,dy))_{x\in\mathbb{R}}$ the Markov kernel defined by

$$\begin{cases} \delta_x(dy) & \text{if } F_\mu(x)=0 \text{ or } F_\mu(x_-)=1; \\ \displaystyle\frac{1}{\mu(\{x\})}\int_{u=F_\mu(x_-)}^{F_\mu(x)} \widetilde{m}(u,dy)\,du & \text{if } \mu(\{x\})>0; \\ \widetilde{m}(F_\mu(x),dy) & \text{otherwise.} \end{cases} \tag{2.6}$$

For all $x\in\mathbb{R}$ such that $F_\mu(x)>0$ and $F_\mu(x_-)<1$, $m(x,dy)$ can be rewritten as

$$m(x,dy)=\int_{v=0}^1 \widetilde{m}(F_\mu(x_-)+v(F_\mu(x)-F_\mu(x_-)),dy)\,dv. \tag{2.7}$$

Conversely, let $(p(x,dy))_{x\in\mathbb{R}}$ be a Markov kernel. Let then $(\widetilde{m}(u,dy))_{u\in(0,1)}$ be the Markov kernel defined for all $u\in(0,1)$ by $\widetilde{m}(u,dy)=p(F_\mu^{-1}(u),dy)$. Let $(m(x,dy))_{x\in\mathbb{R}}$ be the Markov kernel defined by (2.6). Let $x\in\mathbb{R}$ be such that $F_\mu(x_-)>0$ and $F_\mu(x)<1$. If $\mu(\{x\})>0$, then for all $u\in(F_\mu(x_-),F_\mu(x)]$, $F_\mu^{-1}(u)=x$. Hence $m(x,dy)=\frac{1}{\mu(\{x\})}\int_{u=F_\mu(x_-)}^{F_\mu(x)}\widetilde{m}(u,dy)\,du=\frac{1}{\mu(\{x\})}\int_{u=F_\mu(x_-)}^{F_\mu(x)}p(x,dy)\,du=p(x,dy)$. By Lemma 6.3 below, $F_\mu^{-1}(F_\mu(x))=x$, $\mu(dx)$-almost everywhere. So for $\mu(dx)$-almost all $x\in\mathbb{R}$ such that $F_\mu(x_-)>0$, $F_\mu(x)<1$ and $\mu(\{x\})=0$, $m(x,dy)=p(F_\mu^{-1}(F_\mu(x)),dy)=p(x,dy)$. Therefore, for $\mu(dx)$-almost all $x\in\mathbb{R}$, $p(x,dy)=m(x,dy)$.

Throughout the present article, for any $Q\in\mathcal{Q}$, $(m^Q(x,dy))_{x\in\mathbb{R}}$ and M^Q will respectively denote the Markov kernel given by (2.6) when $(\widetilde{m}(u,dy))_{u\in(0,1)}=(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$ and the probability measure on $\mathbb{R}^2$ defined by $M^Q(dx,dy)=\mu(dx)\,m^Q(x,dy)$.

Proposition 2.3. *Let μ and ν be two distinct probability measures on $\mathbb{R}$ with finite first moment and equal means such that $\mathcal{Q}$ is non-empty. Let $Q\in\mathcal{Q}$. Then the probability measure M^Q is a martingale coupling between μ and ν.*

One can easily check thanks to Jensen's inequality that the existence of a martingale coupling between μ and ν implies that $\mu\le_{cx}\nu$ (see Remark 3.2 for a proof). A direct consequence of the latter fact and the last two propositions is an easy characterisation of the emptiness of $\mathcal{Q}$.

Corollary 2.4. *Let μ and ν be two distinct probability measures on $\mathbb{R}$ with finite first moment and equal means. Then $\mathcal{Q} \neq \emptyset$ iff $\mu \leq_{cx} \nu$.*

The proof of Proposition 2.3 relies on the two following lemmas.

Lemma 2.5. *Let $Q \in \mathcal{Q}$. For du-almost all $u \in (0,1)$,*

$$\begin{cases} u \in \mathcal{U}_+ & \Longrightarrow \quad F_\nu^{-1}(v) > F_\mu^{-1}(u), \ \pi_+^Q(u,dv)\text{-a.e;} \\ u \in \mathcal{U}_- & \Longrightarrow \quad F_\nu^{-1}(v) < F_\mu^{-1}(u), \ \pi_-^Q(u,dv)\text{-a.e.} \end{cases}$$

Proof of Lemma 2.5. We have

$$\begin{aligned} &\int_{(0,1)} \left(\int_{(0,1)} \mathbb{1}_{\{F_\nu^{-1}(v) \leq F_\mu^{-1}(u)\}} \pi_+^Q(u,dv) \right) (F_\mu^{-1} - F_\nu^{-1})^+(u)\, du \\ &= \gamma \int_{(0,1)^2} \mathbb{1}_{\{F_\nu^{-1}(v) \leq F_\mu^{-1}(u)\}}\, Q(du,dv) \leq \gamma \int_{(0,1)^2} \mathbb{1}_{\{F_\nu^{-1}(v) \leq F_\mu^{-1}(v)\}}\, Q(du,dv) \\ &= \int_{(0,1)^2} \mathbb{1}_{\{F_\mu^{-1}(v) - F_\nu^{-1}(v) \geq 0\}} (F_\mu^{-1} - F_\nu^{-1})^-(v)\, dv\, \pi_-^Q(v,du) = 0, \end{aligned}$$

where we used for the inequality that $u < v$, $Q(du,dv)$-almost everywhere and that F_μ^{-1} is nondecreasing. So for du-almost all $u \in \mathcal{U}_+$, $\pi_+^Q(u,dv)$-a.e., $F_\nu^{-1}(v) > F_\mu^{-1}(u)$. With a symmetric reasoning, we obtain that for du-almost all $u \in \mathcal{U}_-$, $\pi_-^Q(u,dv)$-a.e., $F_\nu^{-1}(v) < F_\mu^{-1}(u)$. □

Lemma 2.6. *Let $(\widetilde{m}(u,dy))_{u\in(0,1)}$ be a Markov kernel and let $(m(x,dy))_{x\in\mathbb{R}}$ be given by* (2.6)*. Then*

$$\mu(dx)\, m(x,dy) = (F_\mu^{-1}(u), y)_\sharp \left(\mathbb{1}_{(0,1)}(u)\, du\, \widetilde{m}(u,dy)\right),$$

where $\sharp$ denotes the pushforward operation.

Proof of Lemma 2.6. Let $h : \mathbb{R}^2 \to \mathbb{R}$ be a measurable and nonnegative function. By Lemma 6.4 below, $F_\mu(x) > 0$ and $F_\mu(x_-) < 1$, $\mu(dx)$-almost everywhere. So using (2.7), we have

$$\begin{aligned} &\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\, \mu(dx)\, m(x,dy) \\ &= \int_{\mathbb{R}\times\mathbb{R}\times(0,1)} h(x,y) \mathbb{1}_{\{0<F_\mu(x), F_\mu(x_-)<1\}}\, \mu(dx)\, \widetilde{m}(F_\mu(x_-) + v(F_\mu(x) - F_\mu(x_-)), dy)\, dv. \end{aligned}$$

Let $\theta : (x,v) \mapsto F_\mu(x_-) + v(F_\mu(x) - F_\mu(x_-))$. By Lemma 6.6 below, $x = F_\mu^{-1}(\theta(x,v))$, $\mu(dx) \otimes dv$-almost everywhere on $\mathbb{R} \times (0,1)$ and $\theta(x,v)_\sharp(\mu(dx) \otimes \mathbb{1}_{(0,1)}(v)\, dv) = \mathbb{1}_{(0,1)}(u)\, du$. So

$$\begin{aligned} &\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\, \mu(dx)\, m(x,dy) \\ &= \int_{\mathbb{R}\times\mathbb{R}\times(0,1)} h(F_\mu^{-1}(\theta(x,v)), y) \mathbb{1}_{\{0<F_\mu(F_\mu^{-1}(\theta(x,v))), F_\mu(F_\mu^{-1}(\theta(x,v)_-))<1\}}\, \mu(dx)\, \widetilde{m}(\theta(x,v), dy)\, dv \\ &= \int_{\mathbb{R}\times(0,1)} h(F_\mu^{-1}(u), y) \mathbb{1}_{\{0<F_\mu(F_\mu^{-1}(u)), F_\mu(F_\mu^{-1}(u)_-)<1\}}\, \widetilde{m}(u,dy)\, du. \end{aligned}$$

By Lemma 6.4 below and the inverse transform sampling, $F_\mu(F_\mu^{-1}(u)) > 0$ and $F_\mu(F_\mu^{-1}(u)_-) < 1$, du-almost everywhere on $(0,1)$, hence

$$\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\, \mu(dx)\, m(x,dy) = \int_{\mathbb{R}\times(0,1)} h(F_\mu^{-1}(u), y)\, \widetilde{m}(u,dy)\, du. \qquad \square$$

Proof of Proposition 2.3. Let us show that M^Q defines a coupling between μ and ν. Let $h:\mathbb{R}\to\mathbb{R}$ be a measurable and nonnegative (or bounded) function. We want to show that

$$\int_{\mathbb{R}\times\mathbb{R}} h(y)\,\mu(dx)m^Q(x,dy)=\int_{\mathbb{R}} h(y)\,\nu(dy),$$

which by Lemma 2.6 and the inverse transform sampling is equivalent to

$$\int_0^1\int_{\mathbb{R}} h(y)\,\widetilde{m}^Q(u,dy)\,du=\int_0^1 h(F_\nu^{-1}(u))\,du. \tag{2.8}$$

Thanks to Lemma 2.5, we get for du-almost all $u\in(0,1)$,

$$\begin{aligned}
&\int_{\mathbb{R}} h(y)\,\widetilde{m}^Q(u,dy)\\
&=\int_{(0,1)}\left(1-\frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\right)h(F_\nu^{-1}(u))\left(\pi_+^Q(u,dv)\mathbb{1}_{\{F_\mu^{-1}(u)>F_\nu^{-1}(u)\}}\right.\\
&\qquad\qquad\qquad\left.+\pi_-^Q(u,dv)\mathbb{1}_{\{F_\mu^{-1}(u)<F_\nu^{-1}(u)\}}\right)\\
&\quad+\int_{(0,1)}\left(\frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\right)h(F_\nu^{-1}(v))\left(\pi_+^Q(u,dv)\mathbb{1}_{\{F_\mu^{-1}(u)>F_\nu^{-1}(u)\}}\right.\\
&\qquad\qquad\qquad\left.+\pi_-^Q(u,dv)\mathbb{1}_{\{F_\mu^{-1}(u)<F_\nu^{-1}(u)\}}\right)\\
&\quad+h(F_\nu^{-1}(u))\mathbb{1}_{\{F_\mu^{-1}(u)=F_\nu^{-1}(u)\}}\\
&=h(F_\nu^{-1}(u))+\int_{(0,1)}\frac{(F_\mu^{-1}-F_\nu^{-1})^+(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}(h(F_\nu^{-1}(v))-h(F_\nu^{-1}(u)))\,\pi_+^Q(u,dv)\\
&\quad+\int_{(0,1)}\frac{(F_\mu^{-1}-F_\nu^{-1})^-(u)}{F_\nu^{-1}(u)-F_\nu^{-1}(v)}(h(F_\nu^{-1}(v))-h(F_\nu^{-1}(u)))\,\pi_-^Q(u,dv).
\end{aligned} \tag{2.9}$$

Since

$$\begin{aligned}
&\int_{(0,1)^2}\frac{(F_\mu^{-1}-F_\nu^{-1})^+(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}(h(F_\nu^{-1}(v)-h(F_\nu^{-1}(u)))\,\pi_+^Q(u,dv)\,du\\
&=\gamma\int_{(0,1)^2}\frac{h(F_\nu^{-1}(v))-h(F_\nu^{-1}(u))}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}Q(du,dv)\\
&=\int_{(0,1)^2}\frac{h(F_\nu^{-1}(v))-h(F_\nu^{-1}(u))}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}(F_\mu^{-1}-F_\nu^{-1})^-(v)\,\pi_-^Q(v,du)\,dv\\
&=-\int_{(0,1)^2}\frac{(F_\mu^{-1}-F_\nu^{-1})^-(u)}{F_\nu^{-1}(u)-F_\nu^{-1}(v)}(h(F_\nu^{-1}(v))-h(F_\nu^{-1}(u)))\,\pi_-^Q(u,dv)\,du,
\end{aligned}$$

we deduce that $\int_0^1\int_{\mathbb{R}} h(y)\,\widetilde{m}^Q(u,dy)\,du=\int_0^1 h(F_\nu^{-1}(u))\,du$. We conclude that M^Q is a coupling between μ and ν. In particular for $h:y\mapsto|y|$, using the inverse transform sampling, we have

$$\int_0^1\int_{\mathbb{R}}|y|\,\widetilde{m}^Q(u,dy)\,du=\int_0^1|F_\nu^{-1}(u)|\,du=\int_{\mathbb{R}}|y|\,\nu(dy)<+\infty.$$

So $\int_{\mathbb{R}} y\,\widetilde{m}^Q(u,dy)$ is well defined du-almost everywhere on $(0,1)$.

Let us show now that M^Q defines a martingale coupling between μ and ν. By Lemma 2.5, for du-almost all $u \in \mathcal{U}_+$,

$$\begin{aligned}\int_{\mathbb{R}} y\,\widetilde{m}^Q(u,dy) &= \int_{(0,1)} \left(1 - \frac{F_\mu^{-1}(u) - F_\nu^{-1}(u)}{F_\nu^{-1}(v) - F_\nu^{-1}(u)}\right) F_\nu^{-1}(u)\,\pi_+^Q(u,dv) \\ &\quad + \int_{(0,1)} \left(\frac{F_\mu^{-1}(u) - F_\nu^{-1}(u)}{F_\nu^{-1}(v) - F_\nu^{-1}(u)}\right) F_\nu^{-1}(v)\,\pi_+^Q(u,dv) \\ &= \int_{(0,1)} (F_\nu^{-1}(u) + F_\mu^{-1}(u) - F_\nu^{-1}(u))\,\pi_+^Q(u,dv) \\ &= F_\mu^{-1}(u). \end{aligned} \tag{2.10}$$

In the same way, for du-almost all $u \in \mathcal{U}_-$,

$$\int_{\mathbb{R}} y\,\widetilde{m}^Q(u,dy) = F_\mu^{-1}(u). \tag{2.11}$$

Else if $u \in \mathcal{U}_0$, then by definition of $\widetilde{m}^Q(u,dy)$,

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

so for du-almost all $u \in (0,1)$, $\int_{\mathbb{R}} y\,\widetilde{m}^Q(u,dy) = F_\mu^{-1}(u)$.

Let $h : \mathbb{R} \to \mathbb{R}$ be a measurable and bounded function. By Lemma 2.6,

$$\int_{\mathbb{R}\times\mathbb{R}} h(x)(y-x)\,\mu(dx)\,m^Q(x,dy) = \int_0^1 h(F_\mu^{-1}(u)) \left(\int_{\mathbb{R}} (y - F_\mu^{-1}(u))\,\widetilde{m}^Q(u,dy)\right) du = 0.$$

So $\mu(dx)\,m^Q(x,dy)$ is a martingale coupling between μ and ν. □

Let $H : \mathbb{R}^2 \to \mathbb{R}$ be a measurable and bounded function. Lemma 2.6 and (2.9) written with $h : y \mapsto H(F_\mu^{-1}(u), y)$ yield the following formula, which illustrates well how the martingale coupling M^Q differs from the comonotous coupling between μ and ν:

$$\begin{aligned}&\int_{\mathbb{R}\times\mathbb{R}} H(x,y)\,M^Q(dx,dy) - \int_0^1 H(F_\mu^{-1}(u), F_\nu^{-1}(u))\,du \\ &= \gamma \int_{(0,1)^2} \frac{H(F_\mu^{-1}(u), F_\nu^{-1}(v)) - H(F_\mu^{-1}(u), F_\nu^{-1}(u)) + H(F_\mu^{-1}(v), F_\nu^{-1}(u)) - H(F_\mu^{-1}(v), F_\nu^{-1}(v))}{F_\nu^{-1}(v) - F_\nu^{-1}(u)}\,Q(du,dv). \end{aligned} \tag{2.12}$$

Notice that the last integral is well defined since we have $Q(du,dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du\,\pi_+^Q(u,dv)$ and according to Lemma 2.5, there holds $Q(du,dv)$-almost everywhere $F_\nu^{-1}(v) > F_\mu^{-1}(u) > F_\nu^{-1}(u)$. Moreover, the fact that μ and ν have finite first moment along with the inverse transform sampling show that (2.12) also holds for any measurable map $H : \mathbb{R}^2 \to \mathbb{R}$ with at most linear growth. As shown in the next proposition, we can easily deduce from this formula that the map $\mathcal{Q} \ni Q \mapsto M^Q$ is one-to-one as soon as F_μ and F_ν are continuous.

Proposition 2.7. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. If F_μ and F_ν are continuous, then the map $\mathcal{Q} \ni Q \mapsto M^Q$ is one-to-one.*

Proof. Let $Q, Q' \in \mathcal{Q}$ be such that $Q \neq Q'$. Then there exists a borel set $A \subset (0,1)^2$ such that $Q(A) \neq Q'(A)$. Let

$$H : (x,y) \mapsto (y - F_\nu^{-1}(F_\mu(x)))^+ \mathbb{1}_{\{F_\mu(x)\in(0,1)\}} \mathbb{1}_A(F_\mu(x), F_\nu(y)) \mathbb{1}_{\{F_\mu(x) < F_\mu(y)\}}.$$

Since F_μ and F_ν are continuous, for all $u,v\in(0,1)$, we have $F_\mu(F_\mu^{-1}(u))=u$ and $F_\nu(F_\nu^{-1}(v))=v$, so $H(F_\mu^{-1}(u),F_\nu^{-1}(v))=(F_\nu^{-1}(v)-F_\nu^{-1}(u))^+\mathbb{1}_A(u,v)\mathbb{1}_{\{u<v\}}$. We deduce that for all $u,v\in(0,1)$, $H(F_\mu^{-1}(u),F_\nu^{-1}(u))=H(F_\mu^{-1}(v),F_\nu^{-1}(v))=0$ and since $(Q+Q')(du,dv)$-almost everywhere on $(0,1)^2$, $u<v$, we have that $(Q+Q')(du,dv)$-almost everywhere on $(0,1)^2$, $H(F_\mu^{-1}(v),F_\nu^{-1}(u))=0$. Since $H(x,y)$ grows at most linearly in $F_\nu^{-1}(F_\mu(x))$ and y, one can easily deduce from the integrability of μ and ν and the inverse transform sampling that (2.12) holds. Using that $(Q+Q')(du,dv)$ almost everywhere on $(0,1)^2$, $F_\nu^{-1}(u)<F_\mu^{-1}(u)<F_\nu^{-1}(v)$, which is a consequence of Lemma 2.5, we obtain

$$
\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}} H(x,y)\,M^Q(dx,dy)-\int_{\mathbb{R}\times\mathbb{R}} H(x,y)\,M^{Q'}(dx,dy)\\
&=\gamma\int_{(0,1)^2}\frac{H(F_\mu^{-1}(u),F_\nu^{-1}(v))}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\,Q(du,dv)-\gamma\int_{(0,1)^2}\frac{H(F_\mu^{-1}(u),F_\nu^{-1}(v))}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\,Q'(du,dv)\\
&=\gamma(Q(A)-Q'(A))\neq 0,
\end{aligned}
$$

hence $M^Q\neq M^{Q'}$ and the map $\mathcal{Q}\ni Q\mapsto M^Q$ is one-to-one. □

According to [4, Theorem A.4], there exist $N\in\mathbb{N}^*\cup\{+\infty\}$ and a sequence of disjoint open intervals $((\underline{t}_n,\overline{t}_n))_{1\le n\le N}$ such that

$$\left\{t\in\mathbb{R}\mid\int_{-\infty}^t F_\mu(x)\,dx<\int_{-\infty}^t F_\nu(x)\,dx\right\}=\bigcup_{n=1}^N(\underline{t}_n,\overline{t}_n). \tag{2.13}$$

These intervals are called the irreducible components of the pair (μ,ν). Moreover, there exists a unique decomposition of probability measures $(\mu_n,\nu_n)_{1\le n\le N}$, such that the choice of any martingale coupling M between μ and ν reduces to the choice of a sequence of martingale couplings $(M_n)_{1\le n\le N}$. More precisely, for all $1\le n\le N$,

$$F_\mu(\underline{t}_n)\le F_\nu(\underline{t}_n)\le F_\nu((\overline{t}_n)_-)\le F_\mu((\overline{t}_n)_-),\quad F_\mu(\underline{t}_n)<F_\mu((\overline{t}_n)_-), \tag{2.14}$$

and μ_n and ν_n are given by

$$\left\{\begin{array}{rcl}\mu_n(dx)&=&\frac{1}{F_\mu((\overline{t}_n)_-)-F_\mu(\underline{t}_n)}\mathbb{1}_{(\underline{t}_n,\overline{t}_n)}(x)\,\mu(dx);\\ \nu_n(dy)&=&\frac{1}{F_\mu((\overline{t}_n)_-)-F_\mu(\underline{t}_n)}\Big(\mathbb{1}_{(\underline{t}_n,\overline{t}_n)}(y)\,\nu(dy)+(F_\nu(\underline{t}_n)-F_\mu(\underline{t}_n))\,\delta_{\underline{t}_n}(dy)\\ &&\qquad+(F_\mu((\overline{t}_n)_-)-F_\nu((\overline{t}_n)_-))\,\delta_{\overline{t}_n}(dy)\Big).\end{array}\right. \tag{2.15}$$

Then a probability measure M on $\mathbb{R}^2$ is a martingale coupling between μ and ν if and only if there exists a sequence $(M_n)_{1\le n\le N}$ such that for all $1\le n\le N$, M_n is a martingale coupling between μ_n and ν_n and

$$M(dx,dy)=\mathbb{1}_{\mathbb{R}\setminus\bigcup_{n=1}^N(\underline{t}_n,\overline{t}_n)}(x)\,\mu(dx)\,\delta_x(dy)+\sum_{n=1}^N\mu((\underline{t}_n,\overline{t}_n))\,M_n(dx,dy).$$

We can establish a strong connection between the support of any probability measure $Q\in\mathcal{Q}$ and the irreducible components of (μ,ν).

Proposition 2.8. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu<_{cx}\nu$. Let $(\underline{t}_n,\overline{t}_n)_{1\le n\le N}$ denote the irreducible components of (μ,ν). Then for all $Q\in\mathcal{Q}$, we have*

$$Q\left(\bigcup_{1\le n\le N}\big(F_\mu(\underline{t}_n),F_\mu((\overline{t}_n)_-)\big)^2\right)=1.$$

Proof. Let $Q \in \mathcal{Q}$. By [2, Lemma A.8], we have

$$\mathcal{W} := \bigcup_{n=1}^{N} \left(F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-)\right) = \left\{u \in (0,1) \mid \int_0^u F_\mu^{-1}(v)\,dv > \int_0^u F_\nu^{-1}(v)\,dv\right\}.$$

Let $u \in (0,1)$ be such that $F_\mu^{-1}(u) > F_\nu^{-1}(u)$, that is $u \in \mathcal{U}_+$. Since $\mu \leq_{cx} \nu$, according to the necessary condition of [21, Theorem 3.A.5 Chapter 3] (see also Remark 3.2 for a proof), for all $q \in [0,1]$, $\int_0^q F_\mu^{-1}(v)\,dv \geq \int_0^q F_\nu^{-1}(v)\,dv$. By left-continuity of F_μ^{-1} and F_ν^{-1}, we deduce that $\int_0^u F_\mu^{-1}(v)\,dv > \int_0^u F_\nu^{-1}(v)\,dv$, that is $u \in \mathcal{W}$. So $\mathcal{U}_+ \subset \mathcal{W}$.

Let $1 \leq n \leq N$. Then M^Q transports $(\underline{t}_n, \bar{t}_n)$ to $[\underline{t}_n, \bar{t}_n]$, namely for $\mu(dx)$-almost all $x \in (\underline{t}_n, \bar{t}_n)$, $m^Q(x, [\underline{t}_n, \bar{t}_n]) = 1$. So using Lemma 2.6 for the last equality, we have

$$\begin{aligned}\int_{F_\mu(\underline{t}_n)}^{F_\mu((\bar{t}_n)_-)} du = \mu((\underline{t}_n, \bar{t}_n)) &= \int_{\mathbb{R}} \mathbb{1}_{\{\underline{t}_n < x < \bar{t}_n\}} \mu(dx) \\ &= \int_{\mathbb{R}^2} \mathbb{1}_{\{\underline{t}_n < x < \bar{t}_n\}} \mathbb{1}_{\{\underline{t}_n \leq y \leq \bar{t}_n\}}\, \mu(dx)\, m^Q(x, dy) \\ &= \int_{(0,1)\times\mathbb{R}} \mathbb{1}_{\{\underline{t}_n < F_\mu^{-1}(u) < \bar{t}_n\}} \mathbb{1}_{\{\underline{t}_n \leq y \leq \bar{t}_n\}}\, du\, \widetilde{m}^Q(u, dy).\end{aligned}$$

Using Lemma 6.3 below, one can easily see that for all $u \in (0,1)$,

$$\mathbb{1}_{\{F_\mu(\underline{t}_n) < u < F_\mu((\bar{t}_n)_-)\}} \leq \mathbb{1}_{\{\underline{t}_n < F_\mu^{-1}(u) < \bar{t}_n\}} \leq \mathbb{1}_{\{F_\mu(\underline{t}_n) < u \leq F_\mu((\bar{t}_n)_-)\}},$$

so

$$\begin{aligned}\int_{F_\mu(\underline{t}_n)}^{F_\mu((\bar{t}_n)_-)} du &= \int_{(0,1)\times\mathbb{R}} \mathbb{1}_{\{F_\mu(\underline{t}_n) < u < F_\mu((\bar{t}_n)_-)\}} \mathbb{1}_{\{\underline{t}_n \leq y \leq \bar{t}_n\}}\, du\, \widetilde{m}^Q(u, dy) \\ &= \int_{F_\mu(\underline{t}_n)}^{F_\mu((\bar{t}_n)_-)} \widetilde{m}^Q(u, [\underline{t}_n, \bar{t}_n])\, du.\end{aligned}$$

So for du-almost all $u \in (F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-))$, $\widetilde{m}^Q(u, [\underline{t}_n, \bar{t}_n]) = 1$. By Lemma 2.5, $d\Psi_+(u)$-almost everywhere on $(F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-))$,

$$\begin{aligned}1 &= \pi_+^Q(u, \{v \in (0,1) \mid F_\nu^{-1}(v) \in [\underline{t}_n, \bar{t}_n]\}) \\ &= \pi_+^Q(u, \mathcal{U}_- \cap (u,1) \cap \{v \in (0,1) \mid F_\nu^{-1}(v) \in [\underline{t}_n, \bar{t}_n]\}),\end{aligned}$$

where the last equality derives from conditions (ii) and (iii) satisfied by Q. Let $u \in (F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-))$. Let us check that

$$\mathcal{U}_- \cap (u,1) \cap \{v \in (0,1) \mid F_\nu^{-1}(v) \in [\underline{t}_n, \bar{t}_n]\} \subset \mathcal{U}_- \cap (u,1) \cap (F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-)]. \quad (2.16)$$

Let $v \in (0,1)$ be such that $F_\nu^{-1}(v) \in [\underline{t}_n, \bar{t}_n]$. First of all, if $v > u$ then $v > F_\mu(\underline{t}_n)$. Second, if $v > F_\mu((\bar{t}_n)_-)$, then according to (2.14) and Lemma 6.3 below, we have $F_\nu((\bar{t}_n)_-) \leq F_\mu((\bar{t}_n)_-) < v \leq F_\nu(\bar{t}_n)$. In that case, if $v \leq F_\mu(\bar{t}_n)$, then $v \in (F_\nu((\bar{t}_n)_-), F_\nu(\bar{t}_n)] \cap (F_\mu((\bar{t}_n)_-), F_\mu(\bar{t}_n)]$, so $F_\nu^{-1}(v) = F_\mu^{-1}(v) = \bar{t}_n$ and $v \in \mathcal{U}_0$. Else if $v > F_\mu(\bar{t}_n)$, then $v \in (F_\mu(\bar{t}_n), F_\nu(\bar{t}_n)]$ so $F_\nu^{-1}(v) \leq \bar{t}_n < F_\mu^{-1}(v)$ and $v \in \mathcal{U}_+$. This proves (2.16).

Using conditions (ii) and (iii) satisfied by Q again and the fact that the second marginal of Q has a density, we get that $d\Psi_+(u)$-almost everywhere on $(F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-))$,

$$\begin{aligned}1 &= \pi_+^Q(u, \mathcal{U}_- \cap (u,1) \cap \{v \in (0,1) \mid F_\nu^{-1}(v) \in [\underline{t}_n, \bar{t}_n]\}) \\ &\leq \pi_+^Q(u, \mathcal{U}_- \cap (u,1) \cap (F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-)]) \\ &= \pi_+^Q(u, (F_\mu(\underline{t}_n), F_\mu((\bar{t}_n)_-))).\end{aligned}$$

We deduce that

$$
\begin{aligned}
Q\left(\bigcup_{1\le n\le N}\left(F_\mu(\underline{t}_n),F_\mu((\bar{t}_n)_-)\right)^2\right) &= \sum_{n=1}^N Q\left(\left(F_\mu(\underline{t}_n),F_\mu((\bar{t}_n)_-)\right)^2\right)\\
&= \frac{1}{\gamma}\sum_{n=1}^N\int_{F_\mu(\underline{t}_n)}^{F_\mu((\bar{t}_n)_-)} d\Psi_+(u)\,\pi_+^Q(u,(F_\mu(\underline{t}_n),F_\mu((\bar{t}_n)_-)))\\
&= \frac{1}{\gamma}\sum_{n=1}^N\int_{F_\mu(\underline{t}_n)}^{F_\mu((\bar{t}_n)_-)} d\Psi_+(u) = \frac{1}{\gamma}\,d\Psi_+(\mathcal{W})\\
&\ge \frac{1}{\gamma}\,d\Psi_+(\mathcal{U}_+)\\
&= 1,
\end{aligned}
$$

where we used the fact that $\mathcal{U}_+\subset\mathcal{W}$ for the inequality. □

The next proposition clarifies the structure of the set of martingale couplings deriving from $\mathcal{Q}$ and states a linearity property of the map $Q\in\mathcal{Q}\mapsto M^Q$. In particular, it ensures that the set of martingale couplings deriving from $\mathcal{Q}$ is either a singleton, or uncountably infinite.

Proposition 2.9. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu<_{cx}\nu$. Then for all $Q,Q'\in\mathcal{Q}$ and $\lambda\in[0,1]$,*

$$M^{\lambda Q+(1-\lambda)Q'}=\lambda M^Q+(1-\lambda)M^{Q'}.$$

In particular, the set $\{M^Q\mid Q\in\mathcal{Q}\}$ is convex.

Proof. Let $Q,Q'\in\mathcal{Q}$ and let $\lambda\in[0,1]$. It is straightforward that for du-almost all $u\in\mathcal{U}_+$ and dv-almost all $v\in\mathcal{U}_-$,

$$
\begin{aligned}
\pi_+^{\lambda Q+(1-\lambda)Q'}(u,dy)&=\lambda\pi_+^Q(u,dy)+(1-\lambda)\pi_+^{Q'}(u,dy);\\
\pi_-^{\lambda Q+(1-\lambda)Q'}(v,dy)&=\lambda\pi_-^Q(v,dy)+(1-\lambda)\pi_-^{Q'}(v,dy).
\end{aligned}
$$

Using Lemma 2.5, we get that for du-almost all $u\in(0,1)$,

$$\widetilde{m}^{\lambda Q+(1-\lambda)Q'}(u,dy)=\lambda\widetilde{m}^Q(u,dy)+(1-\lambda)\widetilde{m}^{Q'}(u,dy).$$

Let $h:\mathbb{R}^2\to\mathbb{R}$ be a measurable and bounded function. By Lemma 2.6,

$$
\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\,M^{\lambda Q+(1-\lambda)Q'}(dx,dy)\\
&=\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\,\mu(dx)\,m^{\lambda Q+(1-\lambda)Q'}(x,dy)=\int_0^1\left(\int_{\mathbb{R}} h(F_\mu^{-1}(u),y)\,\widetilde{m}^{\lambda Q+(1-\lambda)Q'}(u,dy)\right)du\\
&=\lambda\int_0^1\left(\int_{\mathbb{R}} h(F_\mu^{-1}(u),y)\,\widetilde{m}^Q(u,dy)\right)du+(1-\lambda)\int_0^1\left(\int_{\mathbb{R}} h(F_\mu^{-1}(u),y)\,\widetilde{m}^{Q'}(u,dy)\right)du\\
&=\lambda\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\,\mu(dx)\,m^Q(x,dy)+(1-\lambda)\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\,\mu(dx)\,m^{Q'}(x,dy)\\
&=\int_{\mathbb{R}\times\mathbb{R}} h(x,y)\,(\lambda M^Q+(1-\lambda)M^{Q'})(dx,dy).
\end{aligned}
$$

So $M^{\lambda Q+(1-\lambda)Q'}=\lambda M^Q+(1-\lambda)M^{Q'}$. □

We deduce that if $Q, Q' \in \mathcal{Q}$ are such that $M^Q \neq M^{Q'}$, then there exists a whole segment of martingale couplings between μ and ν, all parametrised by $\mathcal{Q}$. More details are given in Section 4. Let us complete this section by revisiting the example given in Section 2.1.

Example 2.10. Suppose now that $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ are symmetric with common mean $\alpha \in \mathbb{R}$, that is $(x-\alpha)_\sharp\mu(dx) = (\alpha - x)_\sharp\mu(dx)$ and $(y-\alpha)_\sharp\nu(dy) = (\alpha-y)_\sharp\nu(dy)$ where $\sharp$ denotes the pushforward operation. Suppose in addition that their respective quantile functions satisfy $F_\mu^{-1} \geq F_\nu^{-1}$ on $(0, 1/2]$ and $F_\mu^{-1} \leq F_\nu^{-1}$ on $(1/2, 1)$. We saw in Section 2.1 that when U is a random variable uniformly distributed on $[0, 1]$ and Z is given by (2.1), $(F_\mu^{-1}(U), Z)$ is an explicit coupling between μ and ν in the case $\alpha = 0$. Let us show here that this coupling is in fact associated to a particular element of $\mathcal{Q}$. According to Lemma 6.5 below, we have $F_\mu^{-1}(u) = 2\alpha - F_\mu^{-1}(1-u)$ and $F_\nu^{-1}(u) = 2\alpha - F_\nu^{-1}(1-u)$ for du-almost all $u \in (0, 1)$, which is helpful in order to see that the probability measure Q_2 defined on $(0, 1)^2$ by

$$Q_2(du, dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\, du\, \delta_{1-u}(dv) \tag{2.17}$$

is an element of $\mathcal{Q}$ (in particular to check that it satisfies (ii)). For that element Q_2, using (2.5), Lemma 2.5 and Lemma 6.5 below, we have for du-almost all $u \in \mathcal{U}_+ \cup \mathcal{U}_-$,

$$\widetilde{m}^{Q_2}(u, dy) = \frac{F_\mu^{-1}(u) + F_\nu^{-1}(u) - 2\alpha}{2(F_\nu^{-1}(u) - \alpha)}\delta_{F_\nu^{-1}(u)}(dy) + \frac{F_\nu^{-1}(u) - F_\mu^{-1}(u)}{2(F_\nu^{-1}(u) - \alpha)}\delta_{2\alpha - F_\nu^{-1}(u)}(dy), \tag{2.18}$$

and $\widetilde{m}^{Q_2}(u, dy) = \delta_{F_\nu^{-1}(u)}(dy)$ if $u \in \mathcal{U}_0$. Let $u \in (0, 1)$. If $F_\mu^{-1}(u) = F_\nu^{-1}(u) \neq \alpha$, then $\delta_{F_\nu^{-1}(u)}(dy)$ coincides with the right-hand side of (2.18). Furthermore if $F_\nu^{-1}(u) = \alpha$, since $\alpha \geq F_\mu^{-1}(u) \geq F_\nu^{-1}(u)$ or $\alpha \leq F_\mu^{-1}(u) \leq F_\nu^{-1}(u)$ by an easy generalisation of (2.2), then $F_\mu^{-1}(u) = \alpha$. Therefore (2.18) holds for du-almost all $u \in (0, 1)$ such that $F_\nu^{-1}(u) \neq \alpha$ and $\widetilde{m}^{Q_2}(u, dy) = \delta_{F_\nu^{-1}(u)}(dy)$ for du-almost all $u \in (0, 1)$ such that $F_\nu^{-1}(u) = \alpha$.

Let U and V be two independent random variables uniformly distributed on $(0, 1)$ and let Y be defined as in (2.1) but with the mean α taken into account, that is

$$\begin{aligned} Y = F_\nu^{-1}(U)\mathbb{1}&_{\left\{F_\nu^{-1}(U) \neq \alpha, V \leq \frac{F_\mu^{-1}(U) + F_\nu^{-1}(U) - 2\alpha}{2(F_\nu^{-1}(U) - \alpha)}\right\}} \\ &+ (2\alpha - F_\nu^{-1}(U))\mathbb{1}_{\left\{F_\nu^{-1}(U) \neq \alpha, V > \frac{F_\mu^{-1}(U) + F_\nu^{-1}(U) - 2\alpha}{2(F_\nu^{-1}(U) - \alpha)}\right\}} + \alpha\mathbb{1}_{\{F_\nu^{-1}(U) = \alpha\}}. \end{aligned}$$

Then (U, Y) is distributed according to $\mathbb{1}_{(0,1)}(u)\, du\, \widetilde{m}^{Q_2}(u, dy)$. By Lemma 2.6, the random vector $(F_\mu^{-1}(U), Y)$ is distributed according to $\mu(dx)\, m(x, dy)$.

2.3 Optimality property

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$. It is well known that F_ν^{-1} is constant on the jumps of F_μ, that is F_ν^{-1} is constant on the intervals of the form $(F_\mu(x_-), F_\mu(x)]$, iff the comonotonous coupling between μ and ν is concentrated on the graph of a map $T : \mathbb{R} \to \mathbb{R}$, and then

$$T = F_\nu^{-1} \circ F_\mu. \tag{2.19}$$

We will refer to T as the Monge transport map.

Proposition 2.11. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Suppose in addition that F_ν^{-1} is constant on the intervals of the form $(F_\mu(x_-), F_\mu(x)]$. Let T be the Monge transport map. Let $Q \in \mathcal{Q}$. Then*

$$\inf_{M \in \Pi^{\mathrm{M}}(\mu, \nu)} \int_{\mathbb{R}\times\mathbb{R}} |y - T(x)|\, M(dx, dy) = \int_{\mathbb{R}\times\mathbb{R}} |y - T(x)|\, M^Q(dx, dy) = \mathcal{W}_1(\mu, \nu).$$

Proof. This is a particular case of Proposition 2.18 below. Indeed, let $M(dx,dy) = \mu(dx)\,m(x,dy)$ be a martingale coupling between μ and ν. Let $(\widetilde{m}(u,dy))_{u\in(0,1)}$ be the kernel defined for all $u\in(0,1)$ by $\widetilde{m}(u,dy) = m(F_\mu^{-1}(u),dy)$. Using the inverse transform sampling, we have

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|\,\mu(dx)\,m(x,dy) &= \int_{(0,1)\times\mathbb{R}} |y-T(F_\mu^{-1}(u))|\,du\,m(F_\mu^{-1}(u),dy)\\ &= \int_{(0,1)\times\mathbb{R}} |y-F_\nu^{-1}(F_\mu((F_\mu^{-1}(u))))|\,du\,\widetilde{m}(u,dy),\end{aligned}$$

where we used for the last equality that $T = F_\nu^{-1}\circ F_\mu$. Let $u\in(0,1)$. If there exists $x\in\mathbb{R}$ such that $u = F_\mu(x)$, then $F_\mu(F_\mu^{-1}(u)) = F_\mu(F_\mu^{-1}(F_\mu(x))) = F_\mu(x) = u$, so $F_\nu^{-1}(F_\mu(F_\mu^{-1}(u))) = F_\nu^{-1}(u)$. Else there exists x in the set of discontinuities of F_μ such that $F_\mu(x_-)\le u < F_\mu(x)$. In that case, if $u > F_\mu(x_-)$ then $x = F_\mu^{-1}(u)$, so $F_\nu^{-1}(F_\mu(F_\mu^{-1}(u))) = F_\nu^{-1}(F_\mu(x)) = F_\nu^{-1}(u)$ since F_ν^{-1} is constant on the jumps of F_μ. Hence

$$du\text{-a.e. on } (0,1),\quad F_\nu^{-1}(F_\mu(F_\mu^{-1}(u))) = F_\nu^{-1}(u). \tag{2.20}$$

We deduce that

$$\int_{(0,1)\times\mathbb{R}} |y-F_\nu^{-1}(F_\mu(F_\mu^{-1}(u)))|\,du\,\widetilde{m}(u,dy) = \int_{(0,1)\times\mathbb{R}} |y-F_\nu^{-1}(u)|\,du\,\widetilde{m}(u,dy).$$

With a similar reasoning, we have $\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|\,\mu(dx)\,m^Q(x,dy) = \int_{(0,1)\times\mathbb{R}} |y-F_\nu^{-1}(u)|\,du\,\widetilde{m}^Q(u,dy)$. Therefore, using Proposition 2.18 combined with Remark 2.19 below, we get that $\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|\,M(dx,dy)$ is minimised when $M = M^Q$, for which we have $\int_{\mathbb{R}\times\mathbb{R}} |y-T(x)|\,M^Q(dx,dy) = \mathcal{W}_1(\mu,\nu)$. □

2.4 Stability inequality

We can now state our main result. In the minimisation of the cost function $(x,y)\mapsto |x-y|$ with respect to the couplings between μ and ν, the addition of the martingale constraint does not cost more than a factor 2.

Theorem 2.12. *For all $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ such that $\mu <_{cx} \nu$ and for all Q in the non-empty set $\mathcal{Q}$,*

$$\int_{\mathbb{R}\times\mathbb{R}} |x-y|\,M^Q(dx,dy) \le 2\mathcal{W}_1(\mu,\nu). \tag{2.21}$$

Consequently,

$$\mathcal{M}_1(\mu,\nu) \le 2\mathcal{W}_1(\mu,\nu). \tag{2.22}$$

Moreover, the constant 2 is sharp.

The proof of Theorem 2.12 relies on Proposition 2.18 below. Note that since $\Pi^{\mathrm{M}}(\mu,\nu)\subset\Pi(\mu,\nu)$, we always have $\mathcal{W}_1(\mu,\nu)\le\mathcal{M}_1(\mu,\nu)$. Moreover, the stability inequality (2.22) can be tensorised: it holds in greater dimension when the marginals are independent, as the next corollary states.

Corollary 2.13. *Let $d\in\mathbb{N}^*$ and $\mu_1,\cdots,\mu_d,\nu_1,\cdots,\nu_d\in\mathcal{P}_1(\mathbb{R})$ be such that for all $1\le i\le d$, $\mu_i\le_{cx}\nu_i$. Let $\mu = \mu_1\otimes\cdots\otimes\mu_d$ and $\nu = \nu_1\otimes\cdots\otimes\nu_d$. Then $\mu\le_{cx}\nu$ and*

$$\mathcal{M}_1(\mu,\nu) \le 2\mathcal{W}_1(\mu,\nu),$$

when $\mathbb{R}^d$ is endowed with the L^1-norm.

Proof of Corollary 2.13. For all $1 \le i \le d$, since $\mu_i \le_{cx} \nu_i$, Strassen's theorem or Proposition 2.3 and Corollary 2.4 ensure the existence of a martingale coupling $M_i(dx_i, dy_i) = \mu_i(dx_i)\, m_i(x_i, dy_i)$ between μ_i and ν_i. Let then M be the probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ defined by $M(dx, dy) = \mu(dx)\, m_1(x_1, dy_1) \cdots m_d(x_d, dy_d)$. Then it is clear that M is a martingale coupling between μ and ν, which shows that $\mu \le_{cx} \nu$, and

$$\begin{aligned}\mathcal{M}_1(\mu,\nu) &\le \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, M(dx,dy) = \sum_{i=1}^d \int_{\mathbb{R}^d\times\mathbb{R}^d} |x_i - y_i|\, M(dx,dy)\\ &= \sum_{i=1}^d \int_{\mathbb{R}\times\mathbb{R}} |x_i-y_i|\, M_i(dx_i,dy_i).\end{aligned}$$

For all $1 \le i \le d$, let $\mathcal{Q}_i$ denote the set $\mathcal{Q}$ with respect to $\mu = \mu_i$ and $\nu = \nu_i$ and let $Q_i \in \mathcal{Q}_i$. Then for $M_1 = M^{Q_1}, \cdots, M_d = M^{Q_d}$, we deduce from Theorem 2.12 that

$$\mathcal{M}_1(\mu,\nu) \le \sum_{i=1}^d \int_{\mathbb{R}\times\mathbb{R}} |x_i-y_i|\, M^{Q_i}(dx_i,dy_i) \le 2\sum_{i=1}^d \mathcal{W}_1(\mu_i,\nu_i).$$

Let $P \in \Pi(\mu,\nu)$ be a coupling between μ and ν. For all $1 \le i \le d$, let P_i be the marginals of P with respect to the coordinates i and $i+d$, so that P_i is a coupling between μ_i and ν_i. Then

$$\begin{aligned}\sum_{i=1}^d \mathcal{W}_1(\mu_i,\nu_i) &\le \sum_{i=1}^d \int_{\mathbb{R}\times\mathbb{R}} |x_i-y_i|\, P_i(dx_i,dy_i) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \sum_{i=1}^d |x_i-y_i|\, P(dx,dy)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|\, P(dx,dy).\end{aligned}$$

Since the inequality above is true for any coupling P between μ and ν, we deduce that $\sum_{i=1}^d \mathcal{W}_1(\mu_i,\nu_i) \le \mathcal{W}_1(\mu,\nu)$, which proves the assertion. □

In the following remarks, we first look in which case the minimiser of (2.22), studied by Hobson and Klimmek [16], derives from $\mathcal{Q}$. Second, we see that the left-curtain martingale coupling introduced by Beiglböck and Juillet [4] does not always satisfy (2.22).

Remark 2.14. The optimal martingale coupling $M \in \Pi^{\mathrm{M}}(\mu,\nu)$ which minimises $\int_{\mathbb{R}\times\mathbb{R}} |x-y|\, M(dx,dy)$ was actually characterised by Hobson and Klimmek [16] under the dispersion assumption that there exists a bounded interval E of positive length such that $(\mu-\nu)^+(E^{\complement}) = (\nu-\mu)^+(E) = 0$. They show that the optimal coupling M^{HK} is unique. Moreover, in the simpler case where $\mu \wedge \nu = 0$, if $a < b$ denote the endpoints of E, then there exist two nonincreasing functions $R : (0,1) \to (-\infty, a]$ and $S : (0,1) \to [b, +\infty)$ such that for all $u \in (0,1)$, denoting $\widetilde{m}^{HK}(u,dy) = m^{HK}(F_\mu^{-1}(u), dy)$ where $m^{HK}(x,dy)\,\mu(dx) = M^{HK}(dx,dy)$, one has

$$\widetilde{m}^{HK}(u,dy) = \frac{S(u) - F_\mu^{-1}(u)}{S(u)-R(u)}\delta_{R(u)}(dy) + \frac{F_\mu^{-1}(u) - R(u)}{S(u)-R(u)}\delta_{S(u)}(dy).$$

We can discuss in which case M^{HK} derives from $\mathcal{Q}$. Suppose first that F_ν^{-1} takes at least three different values, that is there exist $u, v, w \in (0,1)$ such that $F_\nu^{-1}(u) < F_\nu^{-1}(v) < F_\nu^{-1}(w)$. By left-continuity of F_ν^{-1}, there exists $\varepsilon > 0$ such that $F_\nu^{-1}(u) < F_\nu^{-1}(v-\varepsilon)$ and $F_\nu^{-1}(v) < F_\nu^{-1}(w-\varepsilon)$. Let $I_1 = (0,u]$, $I_2 = (v-\varepsilon, v]$ and $I_3 = (w-\varepsilon, 1]$. Those three intervals are such that for all $s \in I_1$ (resp. $s \in I_2$) and $t \in I_2$ (resp. $t \in I_3$), we have $F_\nu^{-1}(s) < F_\nu^{-1}(t)$. Since R is nonincreasing, if the graph of R meets the graph of F_ν^{-1} on one of those three intervals, then they cannot meet on the two others. We can assert the same with the graph of S since S is nonincreasing as well. Therefore, there exists $k \in \{1,2,3\}$ such that the intersection of $F_\nu^{-1}(I_k)$ and $R(I_k) \cup S(I_k)$ is empty. In particular, for all $t \in I_k$, $\widetilde{m}^{HK}(t, \{F_\nu^{-1}(t)\}) = 0$. However, thanks to Lemma 2.5, we can see that for all $Q \in \mathcal{Q}$, the Markov kernel $\widetilde{m}^Q$ is such that $\widetilde{m}^Q(u, \{F_\nu^{-1}(u)\}) > 0$ for du-almost all $u \in (0,1)$. Therefore, M^{HK} does not derive from $\mathcal{Q}$.

If F_ν^{-1} does not take more than two different values, that is if ν is reduced to two atoms at most, then there exists a unique martingale coupling between μ and ν, so M^{HK} derives of course from $\mathcal{Q}$.

Note that the maximisation problem $\sup_{M \in \Pi^{\mathrm{M}}(\mu,\nu)} \int_{\mathbb{R}\times\mathbb{R}} |x-y|\, M(dx,dy)$ is discussed by Hobson and Neuberger [17].

Example 2.15. For instance, if μ and ν are defined by $\mu(dx) = \frac{1}{2}\mathbb{1}_{[-1,1]}(x)\,dx$ and $\nu(dy) = \frac{1}{2}(\mathbb{1}_{[-2,-1)} + \mathbb{1}_{(1,2]})(y)\,dy$, then (see [16, Example 6.1] for an equivalent calculation)

$$\begin{aligned} m^{HK}(x,dy) = &\left(\frac{1}{2} - \frac{3x}{2\sqrt{12-3x^2}}\right)\delta_{-\frac{1}{2}(x+\sqrt{12-3x^2})}(dy) \\ &+ \left(\frac{1}{2} + \frac{3x}{2\sqrt{12-3x^2}}\right)\delta_{\frac{1}{2}(-x+\sqrt{12-3x^2})}(dy), \end{aligned}$$

which satisfies $m^{HK}(x, \{F_\nu^{-1}(F_\mu(x))\}) > 0$ iff $x \in \{(3-\sqrt{33})/6, (\sqrt{33}-3)/6\}$. On the other hand, for all $Q \in \mathcal{Q}$, the Markov kernel m^Q is such that $m^Q(x, \{F_\nu^{-1}(F_\mu(x))\}) > 0$ for dx-almost all $x \in (-1,1)$.

Remark 2.16. We investigate an example where the left-curtain martingale coupling introduced by Beiglböck and Juillet [4] does not satisfy (2.21). Let $\mu \in \mathcal{P}_1(\mathbb{R})$ be with density f_μ and let $u > 1$ and $d > 0$. Let M^{LC} be defined by

$$M^{LC}(dx,dy) = \mu(dx)\,\left(\mathbb{1}_{\{x>0\}}\,(q\,\delta_{ux}(dy) + (1-q)\,\delta_{-dx}(dy)) + \mathbb{1}_{\{x\le 0\}}\delta_x(dy)\right),$$

where $q = \frac{1+d}{u+d}$. Let ν denote the second marginal of M^{LC}. So ν has density f_ν defined by $f_\nu(x) = \frac{q}{u} f_\mu(\frac{x}{u})$ for all $x > 0$ and $f_\nu(x) = f_\mu(x) + \frac{1-q}{d} f_\mu(-\frac{x}{d})$ for all $x \le 0$. Then M^{LC} is the left-curtain martingale coupling between μ and ν. One can easily compute $\int_{\mathbb{R}^d} |y - x|\, M^{LC}(dx,dy) = 2\frac{(u-1)(1+d)}{u+d}\int_{\mathbb{R}_+} x f_\mu(x)\,dx$. On the other hand, $\mathcal{W}_1(\mu,\nu) = \int_{\mathbb{R}} |F_\mu(t) - F_\nu(t)|\,dt$ (see for instance [23, Remark 2.19 (iii) Chapter 2]). From the relation between f_ν and f_μ, one can deduce that for all $x \ge 0$, $F_\nu(x) = 1 - q + qF_\mu(x/u)$, and for all $x \le 0$, $F_\nu(x) = F_\mu(x) + (1-q)\overline{F}_\mu(-x/d)$, where $\overline{F}_\mu : x \mapsto \mu((x,+\infty)) = 1 - F_\mu(x)$. Using $|x| = x + 2x^-$, we have

$$\begin{aligned} &\mathcal{W}_1(\mu,\nu) \\ &= \int_{\mathbb{R}_-} (1-q)\overline{F}_\mu(-x/d)\,dx + \int_{\mathbb{R}_+} |\overline{F}_\mu(x) - q\overline{F}_\mu(x/u)|\,dx \\ &= \int_{\mathbb{R}_-} (1-q)\overline{F}_\mu(-x/d)\,dx + \int_{\mathbb{R}_+} (\overline{F}_\mu(x) - q\overline{F}_\mu(x/u))\,dx + 2\int_{\mathbb{R}_+} (\overline{F}_\mu(x) - q\overline{F}_\mu(x/u))^-\,dx \\ &= d(1-q)\int_{\mathbb{R}+} x f_\mu(x)\,dx + (1-qu)\int_{\mathbb{R}_+} x f_\mu(x)\,dx + 2\int_{\mathbb{R}_+} (\overline{F}_\mu(x) - q\overline{F}_\mu(x/u))^-\,dx \\ &= 2\int_{\mathbb{R}_+} (\overline{F}_\mu(x) - q\overline{F}_\mu(x/u))^-\,dx. \end{aligned}$$

Then M^{LC} satisfies (2.21) iff

$$\frac{(u-1)(1+d)}{u+d}\int_{\mathbb{R}_+} x f_\mu(x)\,dx \le 2\int_{\mathbb{R}_+} \left(\bar{F}_\mu(x) - q\bar{F}_\mu(x/u)\right)^-\,dx. \tag{2.23}$$

The next example illustrates a contradiction of (2.23) and therefore (2.21) for M^{LC}.

Example 2.17. Let $\mu(dx) = \lambda\exp(-\lambda x)\mathbb{1}_{\{x>0\}}\,dx$, where $\lambda > 0$, and let ν be the probability distribution with density f_ν given by $f_\nu(x) = \frac{q}{u} f_\mu(x/u)$ for $x > 0$ and $f_\nu(x) = \frac{1-q}{d} f_\mu(-x/d)$ for $x \le 0$. Then for all $x \in \mathbb{R}$, $\overline{F}_\mu(x) = \exp(-\lambda x)$, and (2.23) is equivalent to

$$\begin{aligned}
\frac{(u-1)(1+d)}{u+d} \times \frac{1}{\lambda} &> 2\int_{\mathbb{R}_+} \left(\exp(-\lambda x) - q\exp(-\lambda x/u)\right)^-\,dx \\
&= 2\int_{\frac{\ln q}{\lambda(\frac{1}{u}-1)}}^{+\infty} \left(q\exp(-\lambda x/u) - \exp(-\lambda x)\right)dx \\
\iff \frac{(u-1)q}{\lambda} &> 2\left(\frac{qu}{\lambda}\exp\left(-\frac{\ln q}{1-u}\right) - \frac{1}{\lambda}\exp\left(-\frac{\ln q}{\frac{1}{u}-1}\right)\right) = 2\frac{q}{\lambda}(u-1)q^{-1/(1-u)} \\
\iff 2^{1-u} &> q = \frac{1+d}{u+d},
\end{aligned}$$

which can be satisfied for example with $u = \frac{5}{4}$ and $d = \frac{1}{4}$. Note that this condition does not depend on the value of λ. Therefore, the left-curtain martingale coupling

$$M^{LC}(dx,dy) = \lambda\exp(-\lambda x)\mathbb{1}_{\{x>0\}}\,dx\left(\frac{5}{6}\,\delta_{\frac{5x}{4}}(dy) + \frac{1}{6}\,\delta_{-\frac{x}{4}}(dy)\right)$$

does not satisfy (2.21), for any $\lambda > 0$.

Proposition 2.18. *Let $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Let $Q \in \mathcal{Q}$. Then the Markov kernel $(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$ minimises*

$$\int_0^1\int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\,\widetilde{m}(u,dy)\,du$$

among all Markov kernels $(\widetilde{m}(u,dy))_{u\in(0,1)}$ such that

$$\begin{aligned}
&\int_{u\in(0,1)} \widetilde{m}(u,dy)\,du = \nu(dy), \quad \int_{\mathbb{R}} |y|\,\widetilde{m}(u,dy) < +\infty, \\
\text{and} \quad &\int_{\mathbb{R}} y\,\widetilde{m}(u,dy) = F_\mu^{-1}(u),\ du\text{-almost everywhere on } (0,1).
\end{aligned} \tag{2.24}$$

Moreover, $\int_0^1\int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\,\widetilde{m}^Q(u,dy)\,du = \mathcal{W}_1(\mu,\nu)$.

Remark 2.19. If $(\widetilde{m}(u,dy))_{u\in(0,1)}$ is a Markov kernel satisfying (2.24), then using Lemma 2.6, we get that $\mu(dx)\,m(x,dy)$ with $(m(x,dy))_{x\in\mathbb{R}}$ denoting the Markov kernel given by (2.6) is a martingale coupling between μ and ν.

Conversely, if $\mu(dx)\,m(x,dy)$ is a martingale coupling between μ and ν, then using the inverse transform sampling, we get that the Markov kernel $(m(F_\mu^{-1}(u),dy))_{u\in(0,1)}$ satisfies (2.24).

Remark 2.20. The martingale couplings parametrised by $Q \in \mathcal{Q}$ are not the only ones to minimise $\int_0^1\int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\,\widetilde{m}(u,dy)\,du$ among all Markov kernels $(\widetilde{m}(u,dy))_{u\in(0,1)}$ which satisfy (2.24). Indeed, let $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, $\nu = \frac{1}{8}\delta_{-8} + \frac{1}{4}\delta_{-6} + \frac{5}{8}\delta_4$ and

$$M = \frac{1}{8}\left(2\delta_{(-1,-6)} + 2\delta_{(-1,4)} + \delta_{(1,-8)} + 3\delta_{(1,4)}\right).$$

For $m(-1,dy) = \frac{1}{2}\delta_{-6} + \frac{1}{2}\delta_4$ and $m(1,dy) = \frac{1}{4}\delta_{-8} + \frac{3}{4}\delta_4$, we have $M(dx,dy) = \mu(dx)\,m(x,dy)$. Let $(\widetilde{m}(u,dy))_{u\in(0,1)}$ be defined by $\widetilde{m}(u,dy) = m(F_\mu^{-1}(u),dy)$ for all $u \in (0,1)$. It is easy to see that M is a martingale coupling between μ and ν, so $(\widetilde{m}(u,dy))_{u\in(0,1)}$ satisfies (2.24). For all $u \in (0,1)$, we have $F_\mu^{-1}(u) = \mathbb{1}_{\{u\le 1/2\}}(-1) +$

$\mathbb{1}_{\{u>1/2\}}$ and $F_\nu^{-1}(u) = \mathbb{1}_{\{u\le 1/8\}}(-8)+\mathbb{1}_{\{1/8<u\le 3/8\}}(-6)+\mathbb{1}_{\{u>3/8\}}\times 4$. So for all $u\in(0,1)$, we have

$$\widetilde{m}(u,dy) = \mathbb{1}_{\{u\le\frac{1}{2}\}}\left(\frac{1}{2}\delta_{-6}+\frac{1}{2}\delta_4\right)+\mathbb{1}_{\{u>\frac{1}{2}\}}\left(\frac{1}{4}\delta_{-8}+\frac{3}{4}\delta_4\right).$$

We can compute $\int_0^1\int_{\mathbb{R}}|F_\nu^{-1}(u)-y|\,\widetilde{m}(u,dy)\,du = \frac{17}{4} = \int_0^1|F_\mu^{-1}(u)-F_\nu^{-1}(u)|\,du = \mathcal{W}_1(\mu,\nu)$, so $(\widetilde{m}(u,dy))_{u\in(0,1)}$ is optimal.

Thanks to Lemma 2.5, we can see that for all $Q\in\mathcal{Q}$, the Markov kernel $\widetilde{m}^Q$ is such that $\widetilde{m}^Q(u,\{F_\nu^{-1}(u)\})>0$ for du-almost all $u\in(0,1)$. However for all $u\in(0,1/8]$, we have $\widetilde{m}(u,\{F_\nu^{-1}(u)\})=\widetilde{m}(u,\{-8\})=0$. Therefore, $\widetilde{m}$ does not derive from $\mathcal{Q}$.

Proof of Proposition 2.18. Let $\widetilde{m}$ be a Markov kernel satisfying (2.24). By Jensen's inequality, for du-almost every $u\in(0,1)$,

$$|F_\nu^{-1}(u)-F_\mu^{-1}(u)| = \left|\int_{\mathbb{R}}(F_\nu^{-1}(u)-y)\,\widetilde{m}(u,dy)\right| \le \int_{\mathbb{R}}|F_\nu^{-1}(u)-y|\,\widetilde{m}(u,dy).$$

So $\int_0^1|F_\nu^{-1}(u)-F_\mu^{-1}(u)|\,du \le \int_0^1\int_{\mathbb{R}}|F_\nu^{-1}(u)-y|\,\widetilde{m}(u,dy)\,du$.

Therefore, to conclude, it is sufficient to prove that $\int_{\mathbb{R}}|F_\nu^{-1}(u)-y|\,\widetilde{m}^Q(u,dy) = |F_\nu^{-1}(u)-F_\mu^{-1}(u)|$, du-almost everywhere on $(0,1)$.

Applying (2.9) to the measurable and nonnegative function $h : y\mapsto |F_\nu^{-1}(u)-y|$ yields for du-almost all $u\in(0,1)$

$$\begin{aligned}\int_{\mathbb{R}}|F_\nu^{-1}(u)-y|\,\widetilde{m}^Q(u,dy) &= \int_{(0,1)}\frac{(F_\mu^{-1}-F_\nu^{-1})^+(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}|F_\nu^{-1}(u)-F_\nu^{-1}(v)|\pi_+^Q(u,dv)\\ &\quad+\int_{(0,1)}\frac{(F_\mu^{-1}-F_\nu^{-1})^-(u)}{F_\nu^{-1}(u)-F_\nu^{-1}(v)}|F_\nu^{-1}(u)-F_\nu^{-1}(v)|\pi_-^Q(u,dv).\end{aligned}$$

Using Lemma 2.5, we deduce that for du-almost all $u\in(0,1)$

$$\begin{aligned}&\int_{\mathbb{R}}|F_\nu^{-1}(u)-y|\,\widetilde{m}^Q(u,dy)\\ &=\int_{(0,1)}(F_\mu^{-1}-F_\nu^{-1})^+(u)\pi_+^Q(u,dv)+\int_{(0,1)}(F_\mu^{-1}-F_\nu^{-1})^-(u)\pi_-^Q(u,dv)\\ &=|F_\nu^{-1}(u)-F_\mu^{-1}(u)|.\end{aligned}$$

□

Proof of Theorem 2.12. Let $Q\in\mathcal{Q}$ and let $\widetilde{m}^Q$ be the Markov kernel defined by (2.5). By Lemma 2.6 and Proposition 2.18,

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}}|y-x|\,\mu(dx)\,m^Q(x,dy) &= \int_0^1\int_{\mathbb{R}}|y-F_\mu^{-1}(u)|\,\widetilde{m}^Q(u,dy)\,du\\ &\le\int_0^1\int_{\mathbb{R}}|y-F_\nu^{-1}(u)|\,\widetilde{m}^Q(u,dy)\,du\\ &\quad+\int_0^1\int_{\mathbb{R}}|F_\nu^{-1}(u)-F_\mu^{-1}(u)|\,\widetilde{m}^Q(u,dy)\,du\\ &=2\mathcal{W}_1(\mu,\nu).\end{aligned}$$

Since $M^Q(dx,dy)=\mu(dx)\,m^Q(x,dy)$ is a martingale coupling between μ and ν (Proposition 2.3), we get (2.22).

Let us show now that the constant 2 is sharp, that is

$$\sup_{\substack{\mu,\nu\in\mathcal{P}_1(\mathbb{R})\\ \mu<_{cx}\nu}} \frac{\mathcal{M}_1(\mu,\nu)}{\mathcal{W}_1(\mu,\nu)} = 2.$$

Let $a,b\in\mathbb{R}$ be such that $0<a<b$. Let $\mu=\frac{1}{2}\delta_{-a}+\frac{1}{2}\delta_a$ and $\nu=\frac{1}{2}\delta_{-b}+\frac{1}{2}\delta_b$. Since μ and ν are two probability measures with equal means such that μ is concentrated on $[-a,a]$ and ν on $\mathbb{R}\backslash[-a,a]$, then $\mu<_{cx}\nu$. Any coupling H between μ and ν is of the form

$$H = r\delta_{(-a,-b)} + r'\delta_{(-a,b)} + p\delta_{(a,b)} + p'\delta_{(a,-b)},$$

where $r,r',p,p'\geq 0$ and $p+p'=r+r'=p+r'=p'+r=1/2$. One can easily see that H is a martingale coupling iff $b(p-p')=a/2$ and $b(r'-r)=-a/2$, that is

$$H = \frac{(b+a)}{4b}\delta_{(-a,-b)} + \frac{(b-a)}{4b}\delta_{(-a,b)} + \frac{(b+a)}{4b}\delta_{(a,b)} + \frac{(b-a)}{4b}\delta_{(a,-b)}. \tag{2.25}$$

Since there is only one martingale coupling, we trivially have

$$\mathcal{M}_1(\mu,\nu) = \int_{\mathbb{R}\times\mathbb{R}} |x-y|\, H(dx,dy) = \frac{b^2-a^2}{b}\cdot$$

On the other hand, since $\mathcal{W}_1(\mu,\nu)=\int_{\mathbb{R}}|F_\mu(t)-F_\nu(t)|\,dt$ (see for instance [23, Remark 2.19 (iii) Chapter 2]),

$$\mathcal{W}_1(\mu,\nu) = \int_{-\infty}^{-b} 0\,dt + \int_{-b}^{-a}\frac{1}{2}\,dt + \int_{-a}^{a} 0\,dt + \int_a^b \frac{1}{2}\,dt + \int_b^{+\infty} 0\,dt = b-a.$$

So, we have

$$\frac{\mathcal{M}_1(\mu,\nu)}{\mathcal{W}_1(\mu,\nu)} = 1+\frac{a}{b},$$

which tends to 2 as b tends to a. □

Also, the stability inequality (2.22) does not generalise with $\mathcal{M}_1(\mu,\nu)$ and $\mathcal{W}_1(\mu,\nu)$ replaced with $\mathcal{M}_\rho(\mu,\nu)$ and $\mathcal{W}_\rho(\mu,\nu)$ for $\rho>1$, as shown in the next proposition in general dimension.

Proposition 2.21. *Let $d\geq 1$ and $\rho>1$. Then*

$$\sup_{\substack{\mu,\nu\in\mathcal{P}_\rho(\mathbb{R}^d)\\ \mu<_{cx}\nu}} \frac{\mathcal{M}_\rho(\mu,\nu)}{\mathcal{W}_\rho(\mu,\nu)} = +\infty.$$

The proof of Proposition 2.21 will use the following lemma for the case $1<\rho<2$.

Lemma 2.22. *Let $d\geq 1$ and $\rho\in(1,2)$. Let $|\cdot|$ denote the Euclidean norm on $\mathbb{R}^d$. Then there exists $C_\rho>0$ such that*

$$\forall (x,y)\in\mathbb{R}^d\times\mathbb{R}^d,\quad |x-y|^\rho \geq C_\rho\left(|x|^\rho - \frac{\rho}{\rho-1}|x|^{\rho-2}\langle x,y\rangle_{\mathbb{R}^d} + \frac{1}{\rho-1}|y|^\rho\right), \tag{2.26}$$

where, by convention, for all $y\in\mathbb{R}^d$ and for $x=0$ we choose $|x|^{\rho-2}\langle x,y\rangle_{\mathbb{R}^d}$ equal to its limit 0 as $x\to 0$.

When $\rho=2$, both sides of the inequality are equal with $C_2=1$.

Proof of Lemma 2.22. If $x = 0$, any $C_\rho \le \rho - 1$ suits. Else, dividing by $|x|^\rho$ and using that $y/|x|$ explores $\mathbb{R}^d$ when y explores $\mathbb{R}^d$, we see that the statement reduces to show that for all $x, y \in \mathbb{R}^d$ such that $|x| = 1$,

$$|x-y|^\rho \ge C_\rho \left(1 - \frac{\rho}{\rho-1}\langle x, y\rangle_{\mathbb{R}^d} + \frac{1}{\rho-1}|y|^\rho\right).$$

For all $x, y \in \mathbb{R}^d$ such that $|x| = 1$, there exist $y_1, y_2 \in \mathbb{R}$ such that $y = y_1 x + y_2 x^\perp$, where $x^\perp$ is an element of $\mathrm{span}(x)^\perp$ such that $|x^\perp| = 1$. The inequality to prove becomes

$$\forall (y_1, y_2) \in \mathbb{R}^2, \quad \left((1-y_1)^2 + y_2^2\right)^{\rho/2} \ge C_\rho \left(1 - \frac{\rho}{\rho-1}y_1 + \frac{1}{\rho-1}(y_1^2+y_2^2)^{\rho/2}\right). \tag{2.27}$$

Let $L : (y_1, y_2) \mapsto \left((1-y_1)^2 + y_2^2\right)^{\rho/2}$ and $R : (y_1, y_2) \mapsto 1 - \frac{\rho}{\rho-1}y_1 + \frac{1}{\rho-1}(y_1^2+y_2^2)^{\rho/2}$. When $(y_1, y_2) \to (1, 0)$, we have

$$\begin{aligned}
R(y_1, y_2) &= \frac{1}{\rho-1}\left(\rho - 1 - \rho(y_1 - 1 + 1) + (1 + 2(y_1-1) + (y_1-1)^2 + y_2^2)^{\rho/2}\right) \\
&= \frac{1}{\rho-1}\Big(-1 - \rho(y_1-1) + 1 + \rho(y_1-1) + \frac{\rho}{2}(y_1-1)^2 + \frac{\rho}{2}y_2^2 \\
&\qquad + \rho(\frac{\rho}{2}-1)(y_1-1)^2 + o((y_1-1)^2+y_2^2)\Big) \\
&= \frac{1}{\rho-1}\left(\frac{\rho}{2}(y_1-1)^2 + \frac{\rho}{2}y_2^2 - \rho(1-\frac{\rho}{2})(y_1-1)^2 + o((y_1-1)^2+y_2^2)\right).
\end{aligned}$$

Since $\rho < 2$, $L(y_1, y_2) \ge (1-y_1)^2 + y_2^2$ for any (y_1, y_2) in the ball centred at $(1, 0)$ with radius 1. So

$$\limsup_{\substack{(y_1,y_2)\to(1,0)\\(y_1,y_2)\neq(1,0)}} \frac{R(y_1,y_2)}{L(y_1,y_2)} \le \frac{\rho}{2(\rho-1)},$$

On the other hand, when $y_1^2 + y_2^2 \to +\infty$,

$$\frac{R(y_1,y_2)}{L(y_1,y_2)} \sim \frac{(y_1^2+y_2^2)^{\rho/2}}{(\rho-1)(y_1^2+y_2^2)^{\rho/2}} = \frac{1}{\rho-1}.$$

So $(y_1, y_2) \mapsto R(y_1,y_2)/L(y_1,y_2)$ is defined and continuous on $(\mathbb{R}^d \times \mathbb{R}^d)\backslash\{(1,0)\}$, bounded from above in the ball centred at $(1, 0)$ with radius 1 and has a finite limit when the norm of (y_1, y_2) tends to $+\infty$. Therefore this function is bounded from above on $(\mathbb{R}^d \times \mathbb{R}^d)\backslash\{(1,0)\}$ by a certain constant $K \ge \frac{1}{\rho-1}$. Since both sides of (2.27) vanish for $(y_1, y_2) = (1, 0)$, we conclude that this inequality holds with constant $C_\rho = \frac{1}{K}$ and (2.26) with constant $C_\rho = \frac{1}{K}$. □

Proof of Proposition 2.21. Since all norms on $\mathbb{R}^d$ are equivalent, we can suppose that $\mathbb{R}^d$ is endowed with the Euclidean norm. The case $\rho \ge 2$ was addressed in the introduction in the one dimensional case. Its extension to dimension d is immediate. Indeed, for all $n \in \mathbb{N}^*$, let $\mu_n = \mathcal{N}_1(0, n^2)$ and $\mu'_n(dx_1, \cdots, dx_d) = (x_1, 0, \cdots, 0)_\sharp \mu_n(dx_1)$ where $\sharp$ denotes the pushforward operation. By reduction to the one dimensional case, we have

$$\frac{\mathcal{M}_\rho(\mu'_n, \mu'_{n+1})}{\mathcal{W}_\rho(\mu'_n, \mu'_{n+1})} = \frac{\mathcal{M}_\rho(\mu_n, \mu_{n+1})}{\mathcal{W}_\rho(\mu_n, \mu_{n+1})} \underset{n\to+\infty}{\longrightarrow} +\infty.$$

We now consider the case $1 < \rho < 2$. Let $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R}^d)$ be such that $\mu <_{cx} \nu$, and let M be a martingale coupling between μ and ν, which exists according to Strassen's theorem or Proposition 2.3 and Corollary 2.4. Thanks to Lemma 2.22, there exists $C_\rho > 0$ such that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^\rho\, M(dx,dy) \geq C_\rho \left(\int_{\mathbb{R}^d} |x|^\rho\, \mu(dx) - \frac{\rho}{\rho-1} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x|^{\rho-2}\langle x,y\rangle_{\mathbb{R}^d}\, M(dx,dy) \right.$$
$$\left. + \frac{1}{\rho-1}\int_{\mathbb{R}^d} |y|^\rho\, \nu(dx)\right).$$

Since $M(dx,dy) = \mu(dx)\, m(x,dy)$ is a martingale coupling, we have for $\mu(dx)$-almost all $x \in \mathbb{R}^d$, $\int_{\mathbb{R}^d} |x|^{\rho-2}\langle x,y\rangle_{\mathbb{R}^d}\, m(x,dy) = |x|^\rho$, where both sides are equal to 0 when $x=0$. So we get

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^\rho\, M(dx,dy) \geq \frac{C_\rho}{\rho-1}\left(\int_{\mathbb{R}^d} |y|^\rho\, \nu(dx) - \int_{\mathbb{R}^d} |x|^\rho\, \mu(dx)\right).$$

For all $n \in \mathbb{N}$, let $\mu_n = \mathcal{N}_d(0, n^2 I_d)$. Let $G \sim \mathcal{N}_d(0, I_d)$. Then for all $n \in \mathbb{N}$, $\mathcal{W}_\rho^\rho(\mu_n, \mu_{n+1}) \leq \mathbb{E}[|G|^\rho]$ and

$$\begin{aligned}\frac{\mathcal{M}_\rho^\rho(\mu_n,\mu_{n+1})}{\mathcal{W}_\rho^\rho(\mu_n,\mu_{n+1})} &\geq \frac{C_\rho}{\rho-1}\frac{(\mathbb{E}[|(n+1)G|^\rho] - \mathbb{E}[|nG|^\rho])}{\mathbb{E}[|G|^\rho]} \\ &= \frac{((n+1)^\rho - n^\rho)C_\rho}{\rho-1} \\ &\sim_{n\to+\infty} \frac{\rho C_\rho}{\rho-1} n^{\rho-1} \underset{n\to+\infty}{\longrightarrow} +\infty.\end{aligned}$$

□

3 The inverse transform martingale coupling

3.1 Definition and stability of the inverse transform martingale coupling

Let $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. We recall that Ψ_+ and Ψ_- are defined for all $u \in [0,1]$ by $\Psi_+(u) = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^+(v)\, dv$ and $\Psi_-(u) = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^-(v)\, dv$. Let Ψ_-^{-1} (resp. Ψ_+^{-1}) denote the left continuous generalised inverse of Ψ_- (resp. Ψ_+). Let $\varphi : [0,1] \to [0,1]$ and $\widetilde{\varphi} : [0,1] \to [0,1]$ be defined for all $u \in [0,1]$ by

$$\begin{aligned}\varphi(u) &= \Psi_-^{-1}(\Psi_+(u)) = \inf\{r \in [0,1] \mid \Psi_-(r) \geq \Psi_+(u)\};\\ \widetilde{\varphi}(u) &= \Psi_+^{-1}(\Psi_-(u)) = \inf\{r \in [0,1] \mid \Psi_+(r) \geq \Psi_-(u)\},\end{aligned}$$

which are well defined thanks to the equality $\Psi_-(1) = \Psi_+(1)$, consequence of the equality of the means.

Let Q^{IT} be the measure defined on $(0,1)^2$ by

$$Q^{IT}(du,dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\, du\, \pi_+^{IT}(u,dv) \quad \text{where } \pi_+^{IT}(u,dv) = \mathbb{1}_{\{0<\varphi(u)<1\}}\, \delta_{\varphi(u)}(dv), \tag{3.1}$$

with $\gamma = \Psi_-(1) = \Psi_+(1)$. According to the next proposition, this measure belongs to $\mathcal{Q}$.

Proposition 3.1. *Let $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. The measure Q^{IT} is an element of $\mathcal{Q}$ as defined in Section 2. Moreover,*

$$Q^{IT}(du,dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\, dv\, \pi_-^{IT}(v,du) \text{ where } \pi_-^{IT}(v,du) = \mathbb{1}_{\{0<\widetilde{\varphi}(v)<1\}}\, \delta_{\widetilde{\varphi}(v)}(du).$$

From now on we write $(\widetilde{m}^{IT}(u,dy))_{u\in(0,1)}$ instead of $(\widetilde{m}^{Q^{IT}}(u,dy))_{u\in(0,1)}$ and write $(m^{IT}(x,dy))_{x\in\mathbb{R}}$ instead of $(m^{Q^{IT}}(x,dy))_{x\in\mathbb{R}}$. Then Proposition 2.3 implies that the probability measure $M^{IT}(dx,dy) = \mu(dx)\, m^{IT}(x,dy)$ is a martingale coupling between μ and ν, which we call the inverse transform martingale coupling.

We deduce from the expression of π_-^{IT} given in Proposition 3.1 that the definition of $(\widetilde{m}^{IT}(u,dy))_{u\in(0,1)}$ reduces to

$$\begin{cases} \dfrac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(\varphi(u))-F_\nu^{-1}(u)}\delta_{F_\nu^{-1}(\varphi(u))}(dy)+\left(1-\dfrac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(\varphi(u))-F_\nu^{-1}(u)}\right)\delta_{F_\nu^{-1}(u)}(dy) \\ \qquad\qquad \text{if } F_\nu^{-1}(\varphi(u))>F_\mu^{-1}(u)>F_\nu^{-1}(u) \text{ and } \varphi(u)<1; \\ \dfrac{F_\nu^{-1}(u)-F_\mu^{-1}(u)}{F_\nu^{-1}(u)-F_\nu^{-1}(\widetilde{\varphi}(u))}\delta_{F_\nu^{-1}(\widetilde{\varphi}(u))}(dy)+\left(1-\dfrac{F_\nu^{-1}(u)-F_\mu^{-1}(u)}{F_\nu^{-1}(u)-F_\nu^{-1}(\widetilde{\varphi}(u))}\right)\delta_{F_\nu^{-1}(u)}(dy) \\ \qquad\qquad \text{if } F_\nu^{-1}(\widetilde{\varphi}(u))<F_\mu^{-1}(u)<F_\nu^{-1}(u) \text{ and } \widetilde{\varphi}(u)<1; \\ \delta_{F_\nu^{-1}(u)}(dy) \qquad \text{otherwise.} \end{cases} \tag{3.2}$$

Note that if $F_\mu^{-1}(u)>F_\nu^{-1}(u)$, then by left-continuity of F_μ^{-1} and F_ν^{-1}, $\Psi_+(u)>0$, which implies $\varphi(u)>0$. Therefore $F_\mu^{-1}(u)>F_\nu^{-1}(u)$ implies $\varphi(u)>0$ so that with the condition $\varphi(u)<1$, $F_\nu^{-1}(\varphi(u))$ makes sense. For similar reasons, if $F_\mu^{-1}(u)<F_\nu^{-1}(u)$ and $\widetilde{\varphi}(u)<1$ then $F_\nu^{-1}(\widetilde{\varphi}(u))$ makes sense.

Remark 3.2. We recall the celebrated Strassen theorem: if $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$, then $\mu\le_{cx}\nu$ iff there exists a martingale coupling between μ and ν. The sufficient condition is a straightforward consequence of Jensen's inequality. Indeed, if $M(dx,dy)=\mu(dx)\,m(x,dy)$ is a martingale coupling between $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$, then for all convex functions $f:\mathbb{R}\to\mathbb{R}$,

$$\int_{\mathbb{R}} f(x)\,\mu(dx)=\int_{\mathbb{R}} f\left(\int_{\mathbb{R}} y\,m(x,dy)\right)\mu(dx)\le\int_{\mathbb{R}^2} f(y)\,m(x,dy)\,\mu(dx)=\int_{\mathbb{R}} f(y)\,\nu(dy).$$

Conversely, suppose that $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ are such that $\mu\le_{cx}\nu$. For $t\in\mathbb{R}$, $\int_{\mathbb{R}}(t-x)^+\,\mu(dx)\le\int_{\mathbb{R}}(t-x)^+\,\nu(dx)$ by convexity of $x\in\mathbb{R}\mapsto(t-x)^+$. By the Fubini-Tonelli theorem, $\int_{\mathbb{R}}(t-x)^+\,\mu(dx)=\int_{-\infty}^t F_\mu(x)\,dx$. Hence $\varphi_\mu(t)=\int_{-\infty}^t F_\mu(x)\,dx\le\varphi_\nu(t)=\int_{-\infty}^t F_\nu(x)\,dx$ for all $t\in\mathbb{R}$. Hence the respective Fenchel-Legendre transforms φ_μ^* and φ_ν^* of φ_μ and φ_ν satisfy $\varphi_\mu^*\ge\varphi_\nu^*$. For all $u\in(0,1)$ and for all $t\in\mathbb{R}$, $F_\mu^{-1}(u)\le t\iff u\le F_\mu(t)$, so

$$\sup_{q\in[0,1]}\left(qt-\int_0^q F_\mu^{-1}(u)\,du\right)=\int_0^{F_\mu(t)}(t-F_\mu^{-1}(u))\,du=\int_0^1(t-F_\mu^{-1}(u))^+\,du=\varphi_\mu(t).$$

Since $q\mapsto(\int_0^q F_\mu^{-1}(u)\,du)$ is convex on $[0,1]$, we get the well known fact (see for instance [11, Lemma A.22]) that for all $q\in\mathbb{R}$, $\varphi_\mu^*(q)=(\int_0^q F_\mu^{-1}(u)\,du)\mathbb{1}_{[0,1]}(q)+(+\infty)\mathbb{1}_{[0,1]^\complement}(q)$. Hence

$$\int_0^q F_\mu^{-1}(u)\,du\ge\int_0^q F_\nu^{-1}(u)\,du \quad \text{for all } q\in[0,1]\text{, with equality for } q=1. \tag{3.3}$$

We will see in the proof of Proposition 3.1 that if $\mu\neq\nu$, then (3.3) implies that Q^{IT} belongs to $\mathcal{Q}$, which ensures that the inverse transform martingale coupling M^{IT} exists. If $\mu=\nu$, the existence of a martingale coupling is straightforward. Therefore, the construction of the inverse transform martingale coupling gives a constructive proof of the necessary condition in Strassen's theorem in dimension 1.

Proof of Proposition 3.1. By Lemma 6.1 below,

$$\begin{aligned} Q^{IT}((0,1)^2)&=\frac{1}{\gamma}\int_0^1(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\varphi(u)<1\}}\,du \\ &=\frac{1}{\gamma}\int_0^1(F_\mu^{-1}-F_\nu^{-1})^-(u)\mathbb{1}_{\{0<u<1\}}\,du=1, \end{aligned}$$

so Q^{IT} is a probability measure on $(0,1)^2$. Let $h:(0,1)^2\to\mathbb{R}$ be a measurable and bounded function. We have

$$\int_{(0,1)^2} h(u,v)\,Q^{IT}(du,dv)=\frac{1}{\gamma}\int_{(0,1)} h(u,\varphi(u))(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\varphi(u)<1\}}\,du. \tag{3.4}$$

Since Ψ_- is continuous, one has $\Psi_-(\Psi_-^{-1}(u))=u$ for all $u\in(0,1)$. By Lemma 6.3 below, we deduce that $\widetilde{\varphi}(\varphi(u))=u$, $(F_\mu^{-1}-F_\nu^{-1})^+(u)\,du$-almost everywhere on $(0,1)$. Therefore, by Lemma 6.1 below,

$$\begin{aligned}
&\int_{(0,1)} h(u,\varphi(u))(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\varphi(u)<1\}}\,du\\
&=\int_{(0,1)} h(\widetilde{\varphi}(\varphi(u)),\varphi(u))(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\varphi(u)<1\}}\mathbb{1}_{\{0<\widetilde{\varphi}(\varphi(u))<1\}}\,du\\
&=\int_{(0,1)} h(\widetilde{\varphi}(v),v)(F_\mu^{-1}-F_\nu^{-1})^-(v)\mathbb{1}_{\{0<v<1\}}\mathbb{1}_{\{0<\widetilde{\varphi}(v)<1\}}\,dv\\
&=\int_{(0,1)^2} h(u,v)(F_\mu^{-1}-F_\nu^{-1})^-(v)\mathbb{1}_{\{0<\widetilde{\varphi}(v)<1\}}\,\delta_{\widetilde{\varphi}(v)}(du)\,dv.
\end{aligned} \tag{3.5}$$

So

$$\int_{(0,1)^2} h(u,v)\,Q^{IT}(du,dv)=\frac{1}{\gamma}\int_{(0,1)^2} h(u,v)(F_\mu^{-1}-F_\nu^{-1})^-(v)\mathbb{1}_{\{0<\widetilde{\varphi}(v)<1\}}\,\delta_{\widetilde{\varphi}(v)}(du)\,dv.$$

Hence we have $Q^{IT}(du,dv)=\frac{1}{\gamma}(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv\,\pi_-^{IT}(v,du)$, where $\pi_-^{IT}(v,du)=\mathbb{1}_{\{0<\widetilde{\varphi}(v)<1\}}\,\delta_{\widetilde{\varphi}(v)}(du)$. Moreover, since Q^{IT} is a probability measure on $(0,1)^2$, it proves that

$$d\Psi_+(u)\text{-a.e. (resp. } d\Psi_-(u)\text{-a.e.)},\quad 0<\varphi(u)<1 \text{ (resp. } 0<\widetilde{\varphi}(u)<1). \tag{3.6}$$

Therefore, it is clear that Q^{IT} has first marginal $\frac{1}{\gamma}(F_\mu^{-1}-F_\nu^{-1})^+(u)\,du$ and second marginal $\frac{1}{\gamma}(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv$. For $h:(u,v)\mapsto\mathbb{1}_{\{u<v\}}$, (3.4) writes

$$Q^{IT}\left(\{(u,v)\in(0,1)^2\mid u<v\}\right)=\frac{1}{\gamma}\int_0^1 \mathbb{1}_{\{u<\varphi(u)\}}(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\varphi(u)<1\}}\,du.$$

Let us show that $u<\varphi(u)$, $(F_\mu^{-1}-F_\nu^{-1})^+(u)\,du$-almost everywhere on $(0,1)$. By the definition of φ and Lemma 6.3 below, for all $u\in(0,1)$, $\varphi(u)\le u\iff\Psi_-^{-1}(\Psi_+(u))\le u\iff\Psi_+(u)\le\Psi_-(u)$. Recall that since $\mu\le_{cx}\nu$, according (3.3), for all $u\in(0,1)$, $\int_0^u F_\mu^{-1}(v)\,dv\ge\int_0^u F_\nu^{-1}(v)\,dv$, so $\Psi_+(u)\ge\Psi_-(u)$. Therefore, we get that

$$\forall u\in(0,1),\quad \varphi(u)\le u\iff\Psi_+(u)=\Psi_-(u). \tag{3.7}$$

Suppose $F_\mu^{-1}(u)>F_\nu^{-1}(u)$. Since F_μ^{-1} and F_ν^{-1} are left continuous, this implies $F_\mu^{-1}(u-\varepsilon)>F_\nu^{-1}(u-\varepsilon)$ for $\varepsilon>0$ small enough. So, for $\varepsilon>0$ small enough, $\Psi_-(u)=\Psi_-(u-\varepsilon)\le\Psi_+(u-\varepsilon)<\Psi_+(u)$, which implies

$$u<\varphi(u),\quad (F_\mu^{-1}-F_\nu^{-1})^+(u)\,du\text{-almost everywhere on }(0,1). \tag{3.8}$$

So

$$\begin{aligned}
Q^{IT}\left(\{(u,v)\in(0,1)^2\mid u<v\}\right)&=\frac{1}{\gamma}\int_0^1(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\varphi(u)<1\}}\,du=Q^{IT}((0,1)^2)\\
&=1,
\end{aligned}$$

since Q^{IT} is a probability measure on $(0,1)^2$. □

We end this section with the stability of the inverse transform martingale coupling with respect to its marginals μ and ν for the Wasserstein distance topology. The following proposition is a direct consequence of Proposition 5.10, whose proof is given in the supermartingale setting. For the sake of generality, the only martingale coupling between a probability measure $\mu \in \mathcal{P}_1(\mathbb{R})$ and itself, namely $\mu(dx)\,\delta_x(dy)$, is still called inverse transform martingale coupling.

Proposition 3.3. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{cx} \nu$. Let $(\mu_n)_{n\in\mathbb{N}}$ and $(\nu_n)_{n\in\mathbb{N}}$ be two sequences of probability measures on $\mathbb{R}$ with finite first moments such that for all $n \in \mathbb{N}$, $\mu_n \leq_{cx} \nu_n$. For all $n \in \mathbb{N}$, let M_n^{IT} (resp. M^{IT}) be the inverse transform martingale coupling between μ_n and ν_n (resp. between μ and ν).*

If $\mathcal{W}_1(\mu_n, \mu) \underset{n\to+\infty}{\longrightarrow} 0$ and $\mathcal{W}_1(\nu_n, \nu) \underset{n\to+\infty}{\longrightarrow} 0$, then

$$\mathcal{W}_1(M_n^{IT}, M^{IT}) \underset{n\to+\infty}{\longrightarrow} 0.$$

3.2 Optimality properties

Let us now suppose that $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ are such that $\mu <_{cx} \nu$ and there exists $p \in (0,1)$ such that $u \mapsto \int_0^u (F_\mu^{-1}(v) - F_\nu^{-1}(v))\,dv$ is nondecreasing on $[0,p]$ and nonincreasing on $[p,1]$. We saw in Example 2.1 a concrete example of an element $Q_1 \in \mathcal{Q}$. Any probability measure Q defined on $(0,1)$ satisfying properties (i) and (ii) of the definition of $\mathcal{Q}$ is concentrated on $(0,p)\times(p,1)$ and therefore satisfies (iii). The probability measure Q_1 is a simple example that comes to mind. The inverse transform martingale coupling presented in this section is a valid example as well and inspires another coupling which is sort of the nonincreasing twin of the inverse transform martingale coupling.

Let $\chi_- : u \in [0,1] \mapsto \int_{1-u}^1 (F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv = \int_0^u (F_\mu^{-1} - F_\nu^{-1})^-(1-v)\,dv$, $\chi_+ : u \in [0,1] \mapsto \int_0^u (F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv$ and $\Gamma = \chi_-^{-1} \circ \chi_+$ where χ_-^{-1} denotes the left continuous generalised inverse of χ_-, that is

$$\Gamma : u \in [0,1] \mapsto \inf\{r \in [0,1] \mid \chi_-(r) \geq \chi_+(u)\},$$

which is well defined since $\chi_+(1) = \chi_+(p) = \chi_-(1-p) = \gamma$, consequence of the equality of the means. Let Q^{NIT} be the probability measure defined on $(0,1)^2$ by

$$\begin{aligned} Q^{NIT}(du, dv) &= \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du\,\pi_+^{NIT}(u, dv), \\ \text{where} \quad \pi_+^{NIT}(u, dv) &= \mathbb{1}_{\{\Gamma(u)>0\}}\,\delta_{1-\Gamma(u)}(dv). \end{aligned} \tag{3.9}$$

Proposition 3.4. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Assume that there exists $p \in (0,1)$ such that $u \mapsto \int_0^u (F_\mu^{-1}(v) - F_\nu^{-1}(v))\,dv$ is nondecreasing on $[0,p]$ and nonincreasing on $[p,1]$. Then $Q^{NIT} \in \mathcal{Q}$.*

In the symmetric case, that is when μ and ν are symmetric and $p = 1/2$, we have $\Gamma(u) = u$ and therefore $Q^{NIT} = Q_2$ (see (2.17)). Hence Q^{NIT} is a generalisation of the symmetric coupling.

Proof of Proposition 3.4. Note that $\Gamma(1) \leq 1-p$, hence $\Gamma(u) < 1$ for all $u \in (0,1)$. It is clear that Q^{NIT} satisfies property (i) of the definition of $\mathcal{Q}$. By Lemma 6.1 below applied with the functions $f_1 : u \in (0,1) \mapsto (F_\mu^{-1} - F_\nu^{-1})^+(u)$ and $f_2 : u \in (0,1) \mapsto (F_\mu^{-1} - F_\nu^{-1})^-(1-u)$, we have

$$\begin{aligned} \frac{1}{\gamma}\int_0^1 (F_\mu^{-1} - F_\nu^{-1})^+(u) h(1-\Gamma(u)) \mathbb{1}_{\{\Gamma(u)>0\}}\,du &= \frac{1}{\gamma}\int_0^1 (F_\mu^{-1} - F_\nu^{-1})^-(1-v) h(1-v)\,dv \\ &= \frac{1}{\gamma}\int_0^1 (F_\mu^{-1} - F_\nu^{-1})^{-1}(v) h(v)\,dv, \end{aligned}$$

for any measurable and bounded function $h : (0,1) \to \mathbb{R}$. So Q^{NIT} satisfies (ii) as well, and therefore (iii). $\square$

We saw with Proposition 2.18 that for all $Q \in \mathcal{Q}$, $\int_0^1 \int_{\mathbb{R}} |F_\nu^{-1}(u) - y| \, \widetilde{m}^Q(u, dy) \, du = \mathcal{W}_1(\mu, \nu)$. The next proposition shows that the inverse transform martingale coupling and its nonincreasing twin, when it exists, play particular roles among the martingale couplings which derive from $\mathcal{Q}$ when $|F_\nu^{-1}(u) - y|$ is replaced with $|F_\nu^{-1}(u) - y|^\rho$ with $\rho \in \mathbb{R}$.

Proposition 3.5. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. For all $\rho \in \mathbb{R}$ and for any Markov kernel $(\widetilde{m}(u, dy))_{u \in (0,1)}$, let $\mathcal{C}_\rho(\widetilde{m})$ be defined by*

$$\mathcal{C}_\rho(\widetilde{m}) = \int_{\mathbb{R} \times (0,1)} |F_\nu^{-1}(u) - y|^\rho \mathbb{1}_{\{y \neq F_\nu^{-1}(u)\}} \, \widetilde{m}(u, dy) \, du. \tag{3.10}$$

Then, for all $Q \in \mathcal{Q}$,

$$\begin{aligned} \forall \rho \in (-\infty, 1] \cup [2, +\infty), \quad & \mathcal{C}_\rho(\widetilde{m}^{IT}) \leq \mathcal{C}_\rho(\widetilde{m}^Q); \\ \forall \rho \in [1, 2], \quad & \mathcal{C}_\rho(\widetilde{m}^Q) \leq \mathcal{C}_\rho(\widetilde{m}^{IT}); \\ \forall \rho \in \{1, 2\}, \quad & \mathcal{C}_\rho(\widetilde{m}^{IT}) = \mathcal{C}_\rho(\widetilde{m}^Q). \end{aligned} \tag{3.11}$$

Let us now assume that there exists $p \in (0,1)$ such that $u \mapsto \int_0^u (F_\mu^{-1}(v) - F_\nu^{-1}(v)) \, dv$ is nondecreasing on $[0, p]$ and nonincreasing on $[p, 1]$ and denote $(\widetilde{m}^{NIT}(u, dy))_{u \in (0,1)}$ for $(\widetilde{m}^{Q^{NIT}}(u, dy))_{u \in (0,1)}$. Then, for all $Q \in \mathcal{Q}$,

$$\begin{aligned} \forall \rho \in (-\infty, 1] \cup [2, +\infty), \quad & \mathcal{C}_\rho(\widetilde{m}^Q) \leq \mathcal{C}_\rho(\widetilde{m}^{NIT}); \\ \forall \rho \in [1, 2], \quad & \mathcal{C}_\rho(\widetilde{m}^{NIT}) \leq \mathcal{C}_\rho(\widetilde{m}^Q); \\ \forall \rho \in \{1, 2\}, \quad & \mathcal{C}_\rho(\widetilde{m}^{NIT}) = \mathcal{C}_\rho(\widetilde{m}^Q). \end{aligned} \tag{3.12}$$

Remark 3.6. Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. By Proposition 3.5 for $\rho = 0$, we deduce that

$$\sup_{Q \in \mathcal{Q}} \{\mathbb{P}(Y = F_\nu^{-1}(U)) \mid (U, Y) \sim \mathbb{1}_{(0,1)}(u) \, du \, \widetilde{m}^Q(u, dy)\}$$

is attained for the inverse transform martingale coupling.

Suppose in addition that F_ν^{-1} is constant on the intervals of the form $(F_\mu(x_-), F_\mu(x)]$, $x \in \mathbb{R}$. Let $M(dx, dy) = \mu(dx) \, m(x, dy)$ be a martingale coupling between μ and ν. Let $(\widetilde{m}(u, dy))_{u \in (0,1)}$ be the kernel defined for all $u \in (0,1)$ by $\widetilde{m}(u, dy) = m(F_\mu^{-1}(u), dy)$. Let T be the Monge transport map. According to (2.20), $F_\nu^{-1}(u) = F_\nu^{-1}(F_\mu(F_\mu^{-1}(u)))$ for du-almost all $u \in (0,1)$. So by Lemma 2.6, for all $\rho \in \mathbb{R}$,

$$\begin{aligned} &\int_{\mathbb{R} \times \mathbb{R}} |y - T(x)|^\rho \mathbb{1}_{\{y \neq T(x)\}} \, \mu(dx) \, m(x, dy) \\ &= \int_0^1 \int_{\mathbb{R}} |y - T(F_\mu^{-1}(u))|^\rho \mathbb{1}_{\{y \neq T(F_\mu^{-1}(u))\}} \, \widetilde{m}(u, dy) \, du \\ &= \int_0^1 \int_{\mathbb{R}} |y - F_\nu^{-1}(u)|^\rho \mathbb{1}_{\{y \neq F_\nu^{-1}(u)\}} \, \widetilde{m}(u, dy) \, du. \end{aligned}$$

We deduce that the supremum of $\mathbb{P}(Y = T(X))$ among all random variables X and Y such that $(X, Y) \sim M^Q$ for $Q \in \mathcal{Q}$ is attained for the inverse transform martingale coupling.

Proof of Proposition 3.5. Let $\rho \in \mathbb{R}$ and $Q \in \mathcal{Q}$. Let $\varepsilon > 0$ and $f_\varepsilon : \mathbb{R} \to \mathbb{R}$ be defined for all $x \in \mathbb{R}$ by

$$f_\varepsilon(x) = \varepsilon^{\rho - 2} \left((\rho - 1)x + (2 - \rho)\varepsilon \right) \mathbb{1}_{\{x \leq \varepsilon\}} + x^{\rho - 1} \mathbb{1}_{\{x > \varepsilon\}}.$$

It is clear that f_ε is convex for $\rho \in (-\infty, 1] \cup [2, +\infty)$ and concave for $\rho \in [1, 2]$. Let $c_\varepsilon : (0,1)^2 \to \mathbb{R}$ be the right-continuous function defined for all $(u, v) \in (0,1)^2$ by $c_\varepsilon(u, v) = f_\varepsilon(|F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|)$.

If $\rho \in (-\infty, 1] \cup [2, +\infty)$, then c_ε satisfies the Monge condition, that is for all $u, u', v, v' \in (0,1)$ such that $u \leq u'$ and $v \leq v'$,

$$c_\varepsilon(u', v') - c_\varepsilon(u, v') - c_\varepsilon(u', v) + c_\varepsilon(u, v) \leq 0,$$

which follows from the monotonicity of F_ν^{-1} and the fact that $(x, y) \mapsto f_\varepsilon(|x - y|)$ is convex and therefore satisfies the Monge condition. Since Q has marginals $d\Psi_+/\gamma$ and $d\Psi_-/\gamma$, by [19, Theorem 3.1.2 Chapter 3], we have

$$\begin{aligned}\int_0^1 c_\varepsilon(\Psi_+^{-1}(\gamma u), \Psi_-^{-1}(\gamma u))\,du &\leq \int_{(0,1)^2} c_\varepsilon(u, v)\,Q(du, dv)\\ &\leq \int_0^1 c_\varepsilon(\Psi_+^{-1}(\gamma u), \Psi_-^{-1}(\gamma(1-u)))\,du.\end{aligned}$$

It is easy to check that for all $u, v \in (0,1)$, the map $(0,1) \ni \varepsilon \mapsto c_\varepsilon(u, v)$ is nonincreasing, bounded from below by $2 - \rho$ and converges to $|F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}$ when $\varepsilon \to 0$ where by convention, we choose $0^0 = 1$ and for all $\alpha < 0$ and $x = 0$, we choose x^α equal to its limit $+\infty$ as $x \to 0_+$. Therefore, by the monotone convergence theorem for $\varepsilon \to 0$, we have

$$\begin{aligned}&\forall \rho \in (-\infty, 1] \cup [2, +\infty),\\ &\quad\int_0^1 |F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+)|^{\rho-1}\,du\\ &\qquad\leq \int_{(0,1)^2} |F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}\,Q(du, dv)\\ &\qquad\leq \int_0^1 |F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma(1-u))_+)|^{\rho-1}\,du. \end{aligned} \tag{3.13}$$

If $1 \leq \rho \leq 2$, then $(x, y) \mapsto f_\varepsilon(|x - y|)$ is concave so $-c_\varepsilon$ satisfies the Monge condition and a symmetric reasoning shows that

$$\begin{aligned}\int_0^1 c_\varepsilon(\Psi_+^{-1}(\gamma u), \Psi_-^{-1}(\gamma(1-u)))\,du &\leq \int_{(0,1)^2} c_\varepsilon(u, v)\,Q(du, dv)\\ &\leq \int_0^1 c_\varepsilon(\Psi_+^{-1}(\gamma u), \Psi_-^{-1}(\gamma u))\,du.\end{aligned} \tag{3.14}$$

It is easy to check that for all $u, v \in (0,1)$, the map $(0,1) \ni \varepsilon \mapsto c_\varepsilon(u, v)$ is bounded from above by $1 + |F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}$ and converges to its lower bound $|F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}$ when $\varepsilon \to 0$. Consider one of the three integrals in (3.14). If the pointwise limit for $\varepsilon \to 0$ of its integrand is integrable, then we can apply the dominated convergence theorem. Otherwise, the integral is infinite for all $\varepsilon \in (0,1)$. Therefore, for $\varepsilon \to 0$, we have

$$\begin{aligned}\forall 1 \leq \rho \leq 2,\quad &\int_0^1 |F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma(1-u))_+)|^{\rho-1}\,du\\ &\leq \int_{(0,1)^2} |F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}\,Q(du, dv)\\ &\leq \int_0^1 |F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+)|^{\rho-1}\,du. \end{aligned} \tag{3.15}$$

For all $\rho\in\mathbb{R}$, applying (2.9) to the measurable and nonnegative function $h:y\mapsto |F_\nu^{-1}(u)-y|^\rho\mathbb{1}_{\{y\neq F_\nu^{-1}(u)\}}$ yields du-almost everywhere on $(0,1)$,

$$\begin{aligned}
&\int_{\mathbb{R}}|F_\nu^{-1}(u)-y|^\rho\mathbb{1}_{\{y\neq F_\nu^{-1}(u)\}}\,\widetilde{m}^Q(u,dy)\\
&=\int_{(0,1)}\frac{(F_\mu^{-1}-F_\nu^{-1})^+(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}|F_\nu^{-1}(u)-F_\nu^{-1}(v)|^\rho\mathbb{1}_{\{F_\nu^{-1}(v)\neq F_\nu^{-1}(u)\}}\,\pi_+^Q(u,dv)\\
&\quad+\int_{(0,1)}\frac{(F_\mu^{-1}-F_\nu^{-1})^-(u)}{F_\nu^{-1}(u)-F_\nu^{-1}(v)}|F_\nu^{-1}(u)-F_\nu^{-1}(v)|^\rho\mathbb{1}_{\{F_\nu^{-1}(v)\neq F_\nu^{-1}(u)\}}\,\pi_-^Q(u,dv),
\end{aligned}$$

where according to Lemma 2.5, for $(F_\mu^{-1}-F_\nu^{-1})^+(u)\,du$-almost all $u\in(0,1)$, $\pi_+^Q(u,dv)$-a.e., $F_\nu^{-1}(v)>F_\nu^{-1}(u)$ and for $(F_\mu^{-1}-F_\nu^{-1})^-(u)\,du$-almost all $u\in(0,1)$, $\pi_-^Q(u,dv)$-a.e, $F_\nu^{-1}(v)<F_\nu^{-1}(u)$. We deduce that

$$\begin{aligned}
C_\rho(\widetilde{m}^Q)&=\int_{(0,1)^2}(F_\mu^{-1}-F_\nu^{-1})^+(u)|F_\nu^{-1}(u)-F_\nu^{-1}(v)|^{\rho-1}\,du\,\pi_+^Q(u,dv)\\
&\quad+\int_{(0,1)^2}(F_\mu^{-1}-F_\nu^{-1})^-(u)|F_\nu^{-1}(u)-F_\nu^{-1}(v)|^{\rho-1}\,du\,\pi_-^Q(u,dv)\\
&=2\gamma\int_{(0,1)^2}|F_\nu^{-1}(u)-F_\nu^{-1}(v)|^{\rho-1}\,Q(du,dv).
\end{aligned}\tag{3.16}$$

Since the set of discontinuities of F_ν^{-1} is at most countable and since the marginals of Q have densities, we have

$$C_\rho(\widetilde{m}^Q)=2\gamma\int_{(0,1)^2}|F_\nu^{-1}(u_+)-F_\nu^{-1}(v_+)|^{\rho-1}\,Q(du,dv).\tag{3.17}$$

Let us show that

$$\mathcal{C}_\rho(\widetilde{m}^{IT})=2\gamma\int_0^1|F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+)-F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+)|^{\rho-1}\,du.\tag{3.18}$$

By Lemma 6.3 below, $\Psi_+^{-1}(\Psi_+(u))=u$, $d\Psi_+(u)$-almost everywhere on $(0,1)$, so using (3.16), Proposition 6.2 below and the fact that $0<\Psi_\pm^{-1}(u)<1$ for all $u\in(0,\gamma)$, we have

$$\begin{aligned}
\mathcal{C}_\rho(\widetilde{m}^{IT})&=2\int_0^1(F_\mu^{-1}-F_\nu^{-1})^+(u)|F_\nu^{-1}(u)-F_\nu^{-1}(\varphi(u))|^{\rho-1}\mathbb{1}_{\{0<\varphi(u)<1\}}\,du\\
&=2\int_0^1|F_\nu^{-1}(\Psi_+^{-1}(\Psi_+(u)))-F_\nu^{-1}(\Psi_-^{-1}(\Psi_+(u)))|^{\rho-1}\mathbb{1}_{\{0<\Psi_-^{-1}(\Psi_+(u))<1\}}\,d\Psi_+(u)\\
&=2\int_0^\gamma|F_\nu^{-1}(\Psi_+^{-1}(u))-F_\nu^{-1}(\Psi_-^{-1}(u))|^{\rho-1}\mathbb{1}_{\{0<\Psi_-^{-1}(u)<1\}}\,du\\
&=2\gamma\int_0^1|F_\nu^{-1}(\Psi_+^{-1}(\gamma u))-F_\nu^{-1}(\Psi_-^{-1}(\gamma u))|^{\rho-1}\,du.
\end{aligned}$$

Since the set of discontinuities of Ψ_+^{-1}, Ψ_-^{-1}, $(\Psi_+\circ F_\nu)^{-1}=F_\nu^{-1}\circ\Psi_+^{-1}$ and $(\Psi_-\circ F_\nu)^{-1}=F_\nu^{-1}\circ\Psi_-^{-1}$ are at most countable, we get that for du-almost all $u\in(0,1)$, $F_\nu^{-1}(\Psi_+^{-1}(\gamma u))=F_\nu^{-1}\circ\Psi_+^{-1}(\gamma u_+)=F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+)$ and $F_\nu^{-1}(\Psi_-^{-1}(\gamma u))=F_\nu^{-1}(\Psi^{-1}(\gamma u)_+)$, which proves (3.18). Then (3.11) is deduced from (3.13), (3.15), (3.17) and (3.18).

Assume now that there exists $p\in(0,1)$ such that $u\mapsto\int_0^u(F_\mu^{-1}(v)-F_\nu^{-1}(v))\,dv$ is nondecreasing on $[0,p]$ and nonincreasing on $[p,1]$. For all $u\in(0,1)$, $\chi_+(u)=\Psi_+(u)$ and $\chi_-(u)=\gamma-\Psi_-(1-u)$. If U is a random variable uniformly distributed on $(0,1)$, one can easily check that $1-\Psi_-^{-1}(\gamma(1-U))$ has distribution $d\chi_-/\gamma$. Since $u\mapsto 1-\Psi_-^{-1}(\gamma(1-u))$ is nondecreasing, it is shown in [2, Lemma A.3] that $1-\Psi_-^{-1}(\gamma(1-u))=\chi_-^{-1}(\gamma u)$, du-almost everywhere on $(0,1)$. So we show with similar arguments as above that

$$\mathcal{C}_\rho(\widetilde{m}^{NIT})=2\gamma\int_0^1|F_\nu^{-1}(\Psi_+^{-1}(\gamma u)_+)-F_\nu^{-1}(\Psi_-^{-1}(\gamma(1-u))_+)|^{\rho-1}\,du. \tag{3.19}$$

Then (3.12) is deduced from (3.13), (3.15), (3.17) and (3.19). □

4 On the uniqueness of martingale couplings parametrised by $\mathcal{Q}$

Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu<_{cx}\nu$. A direct consequence of Proposition 2.9 is that the set of martingale couplings between μ and ν parametrised by $\mathcal{Q}$ is either a singleton, or uncountably infinite. Since $\mathcal{Q}$ is convex, we deduce from Proposition 4.2 below that $\mathcal{Q}$ is infinite as soon as $\mu<_{cx}\nu$. When μ and ν are such that F_μ and F_ν are continuous, Corollary 4.5 below ensures that there exist uncountably many martingale couplings between μ and ν parametrised by $\mathcal{Q}$. However this does not necessarily hold in the general case. We saw that when ν is reduced to two atoms only, there exists a unique martingale coupling between μ and ν. Suppose now that the comonotonous coupling is a martingale coupling between μ and ν, and $\mu,\nu\in\mathcal{P}_2(\mathbb{R})$. For any martingale coupling $M\in\Pi^{\mathrm{M}}(\mu,\nu)$, we have $\int_{\mathbb{R}\times\mathbb{R}}|x-y|^2\,M(dx,dy)=\int_{\mathbb{R}}y^2\,\nu(dy)-\int_{\mathbb{R}}x^2\,\mu(dx)$. So all the martingale couplings between μ and ν yield the same quadratic cost. In particular, they yield the same quadratic cost as the comonotonous coupling, which is the only minimiser of the quadratic cost among $\Pi(\mu,\nu)$. So the comonotonous coupling is the only martingale coupling between μ and ν. The next proposition states that this conclusion still holds when μ and ν only have finite first order moments.

Proposition 4.1. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu<_{cx}\nu$. If the comonotonous coupling between μ and ν is a martingale coupling, that is for U a random variable uniformly distributed on $(0,1)$,*

$$\mathbb{E}[F_\nu^{-1}(U)|F_\mu^{-1}(U)]=F_\mu^{-1}(U)\quad\textit{almost surely},$$

then it is the only martingale coupling between μ and ν.

Proof. Let U be a random variable uniformly distributed on $(0,1)$. The couple $(U,F_\nu^{-1}(U))$ is distributed according to $\mathbb{1}_{(0,1)}(u)\,du\,\delta_{F_\nu^{-1}(u)}(dy)$. By Lemma 2.6 applied with the Markov kernel $(\widetilde{m}(u,dy))_{u\in(0,1)}=(\delta_{F_\nu^{-1}(u)}(dy))_{u\in(0,1)}$, we get that $(F_\mu^{-1}(U),F_\nu^{-1}(U))$ is distributed according to $\mu(dx)\,m(x,dy)$ where $(m(x,dy))_{x\in\mathbb{R}}$ is given by (2.6). By Lemma 6.4 below combined with the inverse transform sampling and (2.7), $(F_\mu^{-1}(U),F_\nu^{-1}(U))$ is distributed according to $\mu(dx)\int_{v=0}^1\delta_{F_\nu^{-1}(F_\mu(x_-)+v(F_\mu(x)-F_\mu(x_-)))}(dy)\,dv$. So almost surely,

$$\begin{aligned}F_\mu^{-1}(U)&=\mathbb{E}\left[F_\nu^{-1}(U)|F_\mu^{-1}(U)\right]\\&=\int_{v=0}^1\left(\int_{y\in\mathbb{R}}y\,\delta_{F_\nu^{-1}(F_\mu(F_\mu^{-1}(U)_-)+v(F_\mu(F_\mu^{-1}(U))-F_\mu(F_\mu^{-1}(U)_-)))}(dy)\right)dv\\&=\int_0^1F_\nu^{-1}(F_\mu(F_\mu^{-1}(U)_-)+v(F_\mu(F_\mu^{-1}(U))-F_\mu(F_\mu^{-1}(U)_-)))\,dv.\end{aligned}$$

By the inverse transform sampling, we deduce that for $\mu(dx)$-almost all $x\in\mathbb{R}$,

$$\int_0^1F_\nu^{-1}(F_\mu(x_-)+v(F_\mu(x)-F_\mu(x_-)))\,dv=x. \tag{4.1}$$

Let $(\underline{t}_n,\bar{t}_n)_{1\le n\le N}$ denote the irreducible components of (μ,ν), whose definition is given by (2.13). We recall that the choice of any martingale coupling M between μ and

ν reduces to the choice of a sequence of martingale couplings $(M_n)_{1\le n\le N}$ such that for all $1\le n\le N$, M_n is a martingale coupling between the probability measures μ_n and ν_n defined by (2.15). If for each n, μ_n reduces to a single atom, then we necessarily have $M_n(dx,dy)=\mu_n(dx)\,\nu_n(dy)$, so there is a unique choice of the sequence $(M_n)_{1\le n\le N}$ and therefore M, which is the comonotonous coupling. Let us then prove that μ_n reduces to a single atom.

Let I be the at most countable set of $x\in\mathbb{R}$ such that $\mu(\{x\})>0$ and F_ν^{-1} is nonconstant on $(F_\mu(x_-),F_\mu(x)]$. Let us show that

$$\bigcup_{x\in I}(F_\mu(x_-),F_\mu(x))=\bigcup_{n=1}^N\left(F_\mu(\underline{t}_n),F_\mu((\bar t_n)_-)\right). \tag{4.2}$$

By [2, Lemma A.8], we have

$$\bigcup_{n=1}^N\left(F_\mu(\underline{t}_n),F_\mu((\bar t_n)_-)\right)=\left\{u\in(0,1)\mid\int_0^u F_\mu^{-1}(v)\,dv>\int_0^u F_\nu^{-1}(v)\,dv\right\}. \tag{4.3}$$

Let $u\in(0,1)$. Suppose first that there exists $t\in\mathbb{R}$ such that $u=F_\mu(t)$. We recall that $(F_\mu^{-1}(U),F_\nu^{-1}(U))$ follows $\mu(dx)\int_{v=0}^1\delta_{F_\nu^{-1}(F_\mu(x_-)+v(F_\mu(x)-F_\mu(x_-)))}(dy)\,dv$. So

$$\begin{aligned}
&\int_0^{F_\mu(t)}F_\nu^{-1}(v)\,dv\\
&=\int_0^1\mathbb{1}_{\{v\le F_\mu(t)\}}F_\nu^{-1}(v)\,dv=\int_0^1\mathbb{1}_{\{F_\mu^{-1}(v)\le t\}}F_\nu^{-1}(v)\,dv\\
&=\int_{x\in\mathbb{R}}\left(\int_{v=0}^1\mathbb{1}_{\{x\le t\}}\left(\int_{y\in\mathbb{R}}y\,\delta_{F_\nu^{-1}(F_\mu(x_-)+v(F_\mu(x)-F_\mu(x_-)))}(dy)\right)dv\right)\mu(dx)\\
&=\int_{x\in\mathbb{R}}\mathbb{1}_{\{x\le t\}}\left(\int_{v=0}^1F_\nu^{-1}(F_\mu(x_-)+v(F_\mu(x)-F_\mu(x_-)))\,dv\right)\mu(dx)\\
&=\int_{x\in\mathbb{R}}\mathbb{1}_{\{x\le t\}}x\,\mu(dx)=\int_0^1\mathbb{1}_{\{F_\mu^{-1}(v)\le t\}}F_\mu^{-1}(v)\,dv=\int_0^{F_\mu(t)}F_\mu^{-1}(v)\,dv,
\end{aligned}$$

where we used (4.1) for the fifth equality and the inverse transform sampling for the sixth equality. By continuity, we also deduce that for all $t\in\mathbb{R}$, $\int_0^{F_\mu(t_-)}F_\nu^{-1}(v)\,dv=\int_0^{F_\mu(t_-)}F_\mu^{-1}(v)\,dv$.

Suppose now that there exists $x\in\mathbb{R}$ in the set of discontinuities of F_μ such that $F_\mu(x_-)<u<F_\mu(x)$. According to (4.1), we have $\int_{F_\mu(x_-)}^{F_\mu(x)}F_\nu^{-1}(v)\,dv=\mu(\{x\})x=\int_{F_\mu(x_-)}^{F_\mu(x)}x\,dv=\int_{F_\mu(x_-)}^{F_\mu(x)}F_\mu^{-1}(v)\,dv$.

If F_ν^{-1} is constant on $(F_\mu(x_-),F_\mu(x)]$, then for all $v\in(F_\mu(x_-),F_\mu(x)]$, $F_\nu^{-1}(v)=x=F_\mu^{-1}(v)$, so

$$\begin{aligned}
&\int_0^u F_\nu^{-1}(v)\,dv\\
&=\int_0^{F_\mu(x_-)}F_\nu^{-1}(v)\,dv+\int_{F_\mu(x_-)}^u F_\nu^{-1}(v)\,dv=\int_0^{F_\mu(x_-)}F_\mu^{-1}(v)\,dv+\int_{F_\mu(x_-)}^u F_\mu^{-1}(v)\,dv\\
&=\int_0^u F_\mu^{-1}(v)\,dv.
\end{aligned}$$

If F_ν^{-1} is nonconstant on $(F_\mu(x_-),F_\mu(x)]$, then using the monotonicity of F_ν^{-1}, one can easily show that for all $u\in(F_\mu(x_-),F_\mu(x))$,

$$\frac{1}{u-F_\mu(x_-)}\int_{F_\mu(x_-)}^u F_\nu^{-1}(v)\,dv<\frac{1}{F_\mu(x)-F_\mu(x_-)}\int_{F_\mu(x_-)}^{F_\mu(x)}F_\nu^{-1}(v)\,dv.$$

We deduce that for all $u \in (F_\mu(x_-), F_\mu(x))$,

$$\int_{F_\mu(x_-)}^u F_\nu^{-1}(v)\,dv < \frac{u - F_\mu(x_-)}{F_\mu(x) - F_\mu(x_-)} x\mu(\{x\}) = (u - F_\mu(x_-))x$$
$$= \int_{F_\mu(x_-)}^u x\,dv = \int_{F_\mu(x_-)}^u F_\mu^{-1}(v)\,dv,$$

and $\int_0^u F_\mu^{-1}(v)\,dv > \int_0^u F_\nu^{-1}(v)\,dv$. With (4.3), we deduce (4.2). Since the intervals $((\underline{t}_n, \overline{t}_n))_{1\le n\le N}$ are disjoint, the intervals $((F_\mu(\underline{t}_n), F_\mu((\overline{t}_n)_-)))_{1\le n\le N}$ are disjoint as well. By equality of unions of disjoint intervals, we proved that for all $1 \le n \le N$, there exists $x \in I$ such that $(F_\mu(\underline{t}_n), F_\mu((\overline{t}_n)_-)) = (F_\mu(x_-), F_\mu(x))$. So $x \in (\underline{t}_n, \overline{t}_n)$ and $\mu((\underline{t}_n, \overline{t}_n)) = F_\mu((\overline{t}_n)_-) - F_\mu(\underline{t}_n) = F_\mu(x) - F_\mu(x_-) = \mu(\{x\})$. So $\mu_n = \delta_x$, and the discussion above concludes that there exists only one martingale coupling between μ and ν, namely the comonotonous coupling. □

We saw in Section 3.1 that we can build a nonincreasing twin of the inverse transform martingale coupling (see (3.9)) as soon as the two marginals satisfy the assumption in Proposition 3.4. This corresponds to a general inversion of the monoticity of φ on $(0,1)$. In the general case, such an inversion is not possible on $(0,1)$, but can be made locally.

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Since $\mu \ne \nu$, there exists $u \in (0,1)$ such that $\Psi_+(u) > \Psi_-(u)$. Let $v = \Psi_+^{-1}(\Psi_+(u))$. Then $\Psi_+(v) = \Psi_+(\Psi_+^{-1}(\Psi_+(u)) = \Psi_+(u)$ so that $v > 0$ and $\Psi_+(v) > \Psi_-(u) \ge \Psi_-(v)$. By left-continuity of Ψ_- and Ψ_+, there exists $\eta \in (0, v)$ such that $\Psi_+(w) > \Psi_-(w)$ for all $w \in [v-\eta, v]$. By definition of v, we have $\Psi_+(v-\eta) < \Psi_+(v)$, so there exists $u_0 \in (v-\eta, v)$ such that $(F_\mu^{-1} - F_\nu^{-1})^+(u_0) > 0$. Since $u_0 \in (v-\eta, v)$, we have $\Psi_+(v) > \Psi_+(u_0) > \Psi_-(u_0)$ so $1 > \varphi(u_0) > u_0$ according to (3.7). By left-continuity of F_μ^{-1}, F_ν^{-1} and φ, there exists $\varepsilon \in (0, u_0)$ such that

$$\forall u \in [u_0 - \varepsilon, u_0], \quad 1 > \varphi(u) > u_0 \quad \text{and} \quad F_\mu^{-1}(u) > F_\nu^{-1}(u). \tag{4.4}$$

Since $(u_0 - \varepsilon, u_0] \subset \mathcal{U}_+$, Ψ_+ is increasing and is therefore one-to-one onto from $(u_0-\varepsilon, u_0]$ to $(\Psi_+(u_0-\varepsilon), \Psi_+(u_0)]$. Since the set of discontinuities of Ψ_-^{-1} is at most countable, up to choosing ε smaller, we may also suppose that in addition to (4.4), ε is such that Ψ_-^{-1} is continuous at $\Psi_+(u_0-\varepsilon)$. Let then $\zeta : [0,1] \to [0,1]$ and $\widetilde{\zeta} : [0,1] \to [0,1]$ be defined for all $u \in (0,1)$ by

$$\zeta(u) = \Psi_-^{-1}(G(\Psi_+(u))) \quad \text{and} \quad \widetilde{\zeta}(u) = \Psi_+^{-1}(G(\Psi_-(u))), \tag{4.5}$$

where $G : u \mapsto u\mathbb{1}_{(u_0-\varepsilon,u_0]^\complement}(\Psi_+^{-1}(u)) + (\Psi_+(u_0) - u + \Psi_+(u_0-\varepsilon))\,\mathbb{1}_{(u_0-\varepsilon,u_0]}(\Psi_+^{-1}(u))$. Let Q^ζ be the measure defined on $(0,1)^2$ by

$$Q^\zeta(du,dv) = \frac{1}{\gamma}\left(F_\mu^{-1} - F_\nu^{-1}\right)^+(u)\,du\,\pi_+^\zeta(u,dv), \text{ where } \pi_+^\zeta(u,dv) = \mathbb{1}_{\{0<\zeta(u)<1\}}\,\delta_{\zeta(u)}(dv), \tag{4.6}$$

with $\gamma = \Psi_-(1) = \Psi_+(1)$.

Proposition 4.2. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. The measure Q^ζ defined by* (4.6) *is an element of $\mathcal{Q}$. Moreover,*

$$Q^\zeta(du,dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv\,\pi_-^\zeta(v,du), \text{ where } \pi_-^\zeta(v,du) = \mathbb{1}_{\{0<\widetilde{\zeta}(v)<1\}}\,\delta_{\widetilde{\zeta}(v)}(du).$$

As said above, Ψ_+ is one-to-one onto from $(u_0-\varepsilon, u_0]$ to $(\Psi_+(u_0-\varepsilon), \Psi_+(u_0)]$. So, for all $u \in (u_0-\varepsilon, u_0]$, $\Psi_+^{-1}(\Psi_+(u)) = u$ and $G(\Psi_+(u)) = \Psi_+(u_0) - \Psi_+(u) + \Psi_+(u_0-\varepsilon)$. So

$$\forall u \in (u_0-\varepsilon, u_0], \quad \zeta(u) = \Psi_-^{-1}(\Psi_+(u_0) - \Psi_+(u) + \Psi_+(u_0-\varepsilon)). \tag{4.7}$$

Since Ψ_- is continuous, Ψ_-^{-1} is one-to-one. Moreover, Ψ_+ is increasing on $(u_0-\varepsilon, u_0]$, so for all $u \in (u_0-\varepsilon, u_0]\backslash\{\Psi_+^{-1}(\frac{\Psi_+(u_0)+\Psi_+(u_0-\varepsilon)}{2})\}$, $\zeta(u) \neq \varphi(u)$. Since $(u_0-\varepsilon, u_0] \subset \mathcal{U}_+$, considering the first marginal of Q^ζ and Q^{IT}, we deduce that $Q^\zeta \neq Q^{IT}$. As a direct consequence of the convexity of $\mathcal{Q}$, we deduce that $\mathcal{Q}$ is uncountably infinite.

Corollary 4.3. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Then $\mathcal{Q}$ is uncountably infinite.*

Proof of Proposition 4.2. Let $h : (0,1)^2 \to \mathbb{R}$ be a measurable and bounded function. We have

$$\begin{aligned}\int_{(0,1)^2} h(u,v)\, Q^\zeta(du,dv) &= \frac{1}{\gamma}\int_{(0,1)^2} h(u,v)(F_\mu^{-1}-F_\nu^{-1})^+(u)\mathbb{1}_{\{0<\zeta(u)<1\}}\,\delta_{\zeta(u)}(dv)\,du \\ &= \frac{1}{\gamma}\int_0^1 h(u,\zeta(u))\mathbb{1}_{\{0<\zeta(u)<1\}}\,d\Psi_+(u) \\ &= \frac{1}{\gamma}\int_0^1 h(\Psi_+^{-1}(\Psi_+(u)),\zeta(u))\mathbb{1}_{\{0<\zeta(u)<1\}}\mathbb{1}_{\{0<u<1\}}\,d\Psi_+(u),\end{aligned} \tag{4.8}$$

where the last equality is a consequence of Lemma 6.3 below. By Proposition 6.2 below,

$$\begin{aligned}&\int_{(0,1)^2} h(u,v)\, Q^\zeta(du,dv) \\ &= \frac{1}{\gamma}\int_0^{\Psi_+(1)} h(\Psi_+^{-1}(u),\Psi_-^{-1}(G(u)))\mathbb{1}_{\{0<\Psi_-^{-1}(G(u))<1\}}\mathbb{1}_{\{0<\Psi_+^{-1}(u)<1\}}\,du.\end{aligned}$$

By Lemma 6.3 below, for all $u \in (0,\Psi_+(1))$, $u_0-\varepsilon < \Psi_+^{-1}(u) \le u_0 \iff \Psi_+(u_0-\varepsilon) < u \le \Psi_+(u_0)$. Hence G is a piecewise affine function which satisfies $G(G(u)) = u$ for all $u \in (0,\Psi_+(1))\backslash\{\Psi_+(u_0)\}$ and $G(G(\Psi_+(u_0))) = \Psi_+(u_0-\varepsilon)$. So by the change of variables $w = G(u)$, we have

$$\begin{aligned}&\int_{(0,1)^2} h(u,v)\, Q^\zeta(du,dv) \\ &= \frac{1}{\gamma}\int_0^{\Psi_+(1)} h(\Psi_+^{-1}(G(w)),\Psi_-^{-1}(w))\mathbb{1}_{\{0<\Psi_-^{-1}(w)<1\}}\mathbb{1}_{\{0<\Psi_+^{-1}(G(w))<1\}}\,dw.\end{aligned} \tag{4.9}$$

By continuity of Ψ_- and Proposition 6.2 below, using that $\Psi_+(1) = \Psi_-(1)$, we have

$$\begin{aligned}&\int_{(0,1)^2} h(u,v)\, Q^\zeta(du,dv) \\ &= \frac{1}{\gamma}\int_0^1 h(\Psi_+^{-1}(G(\Psi_-(u))),\Psi_-^{-1}(\Psi_-(u)))\mathbb{1}_{\{0<\Psi_-^{-1}(\Psi_-(u))<1\}}\mathbb{1}_{\{0<\Psi_+^{-1}(G(\Psi_-(u)))<1\}}\,d\Psi_-(u) \\ &= \frac{1}{\gamma}\int_0^1 h(\widetilde{\zeta}(u),u)\mathbb{1}_{\{0<\widetilde{\zeta}(u)<1\}}\,d\Psi_-(u),\end{aligned}$$

where we used for the last equality that $\Psi_-^{-1}(\Psi_-(u)) = u$, $d\Psi_-(u)$-almost everywhere on $(0,1)$ according to Lemma 6.3 below.

Hence we have that $Q(du,dv) = \frac{1}{\gamma}(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv\,\pi_-^\zeta(v,du)$ where $\pi_-^\zeta(v,du) = \mathbb{1}_{\{0<\widetilde{\zeta}(v)<1\}}\,\delta_{\widetilde{\zeta}(v)}(du)$.

For $h : (u,v) \mapsto 1$, (4.9) writes

$$Q^\zeta((0,1)^2) = \frac{1}{\gamma}\int_0^{\Psi_+(1)} \mathbb{1}_{\{0<\Psi_-^{-1}(w)<1\}}\mathbb{1}_{\{0<\Psi_+^{-1}(G(w))<1\}}\,dw.$$

By continuity of Ψ_-, Proposition 6.2 and Lemma 6.3 below, $\int_0^{\Psi_+(1)} \mathbb{1}_{\{0<\Psi_-^{-1}(w)<1\}}\,dw = \int_0^1 \mathbb{1}_{\{0<\Psi_-^{-1}(\Psi_-(w))<1\}}\,d\Psi_-(w) = \int_0^1 d\,\Psi_-(u) = \Psi_-(1) = \Psi_+(1)$. So $0 < \Psi_-^{-1}(w) < 1$, dw-almost everywhere on $(0,\Psi_+(1))$. By a similar reasoning, $0 < \Psi_+^{-1}(w) < 1$ for dw-almost all $w \in (0, \Psi_+(1))$. Since G is piecewise affine and bijective from $(0, \Psi_+(1)) \backslash \{\Psi_+(u_0)\}$ to itself, $0 < \Psi_+^{-1}(G(w)) < 1$ for dw-almost all $w \in (0, \Psi_+(1))$. Hence

$$Q^\zeta((0,1)^2) = \frac{1}{\gamma} \int_0^{\Psi_+(1)} dw = 1,$$

so Q^ζ is a probability measure, with first marginal $\frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du$ and second marginal $\frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv$.

We have

$$Q^\zeta\left(\{(u,v) \in (0,1)^2 \mid u < v\}\right) = \frac{1}{\gamma} \int_0^1 \mathbb{1}_{\{u < \zeta(u)\}} \mathbb{1}_{\{0 < \zeta(u) < 1\}}\, d\Psi_+(u).$$

According to (3.8), $u < \varphi(u)$, $d\Psi_+(u)$-almost everywhere on $(0,1)$. According to (4.7) and (4.4), for all $u \in (u_0 - \varepsilon, u_0]$, $\zeta(u) \geq \zeta(u_0)$ and

$$\zeta(u_0) = \varphi(u_0 - \varepsilon) > u_0. \tag{4.10}$$

So for all $u \in (u_0 - \varepsilon, u_0]$, $\zeta(u) > u_0 \geq u$. Moreover, by Lemma 6.3 below, $\Psi_+^{-1}(\Psi_+(u)) = u$, $d\Psi_+(u)$-almost everywhere on $(0,1)$. So ζ coincides with φ, $d\Psi_+$-almost everywhere on $(u_0 - \varepsilon, u_0]^\complement$, hence $u < \zeta(u)$, $d\Psi_+(u)$-almost everywhere on $(0,1)$. So using (4.8) for $h = 1$, we get that

$$Q^\zeta\left(\{(u,v) \in (0,1)^2 \mid u < v\}\right) = \frac{1}{\gamma} \int_0^1 \mathbb{1}_{\{0 < \zeta(u) < 1\}}\, d\Psi_+(u) = Q^\zeta((0,1)^2),$$

which is equal to 1 since Q^ζ is a probability measure on $(0,1)^2$. □

Corollary 4.4. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$. Let $u_0 \in (0,1)$ and $\varepsilon \in (0, u_0)$ be such that (4.4) is satisfied and Ψ_-^{-1} is continuous at $\Psi_+(u_0 - \varepsilon)$. If F_ν^{-1} is nonconstant on $(\varphi(u_0 - \varepsilon), \varphi(u_0)]$ and if F_μ^{-1} is such that for all $\varepsilon' \in (0, \varepsilon)$, the set $\{u \in (0,1) \mid F_\mu^{-1}(u_0 - \varepsilon') < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)\}$ has positive Lebesgue measure, then there exist uncountably many martingale couplings parametrised by $\mathcal{Q}$ between μ and ν.*

Notice that by left-continuity, the condition on F_μ^{-1} in the statement of Corollary 4.4 is satisfied if for all $\varepsilon' \in (0, \varepsilon)$, F_μ^{-1} takes at least three different values on $[u_0 - \varepsilon', u_0]$. A direct consequence of Corollary 4.4 is the infinite amount of martingale couplings between μ and ν when F_μ^{-1} and F_ν^{-1} are increasing, or equivalently when F_μ and F_ν are continuous.

Corollary 4.5. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu <_{cx} \nu$ and $\mu(\{x\}) = \nu(\{x\}) = 0$ for all $x \in \mathbb{R}$. Then there exist uncountably many martingale couplings parametrised by $\mathcal{Q}$ between μ and ν.*

Remark 4.6. Corollary 4.5 is also a consequence of Corollary 4.3 together with Proposition 2.7.

Proof of Corollary 4.4. Let ζ be defined by (4.5) and let Q^ζ be the probability measure defined by (4.6). By Proposition 3.1, Proposition 4.2 and Lemma 2.5, for du-almost all $u \in (u_0 - \varepsilon, u_0)$, $F_\nu^{-1}(\zeta(u)) > F_\mu^{-1}(u) > F_\nu^{-1}(u)$ and $F_\nu^{-1}(\varphi(u)) > F_\mu^{-1}(u) > F_\nu^{-1}(u)$.

Since F_ν^{-1} is left-continuous and nonconstant on $(\varphi(u_0 - \varepsilon), \varphi(u_0)]$, F_ν^{-1} is nonconstant on $(\varphi(u_0 - \varepsilon), \varphi(u_0))$. So there exist $a, b \in (\varphi(u_0 - \varepsilon), \varphi(u_0))$ such that $F_\nu^{-1}(a) < F_\nu^{-1}(b)$. Let then $c = \inf\{u \in (a,b) \mid F_\nu^{-1}(u) = F_\nu^{-1}(b)\}$. Let $u \in [a,b]$. If $F_\nu^{-1}(u) = F_\nu^{-1}(b)$, then $c \leq u$. Else if $F_\nu^{-1}(u) < F_\nu^{-1}(b)$, then $c \geq u$. We deduce that $a \leq c \leq b$ and for all $u, v \in (\varphi(u_0 - \varepsilon), \varphi(u_0))$ such that $u < c < v$, we have $F_\nu^{-1}(u) < F_\nu^{-1}(b) \leq F_\nu^{-1}(v)$.

Using (4.10), we have $F_\nu^{-1}(\varphi(u_0)) > F_\nu^{-1}(\varphi(u_0 - \varepsilon)) = F_\nu^{-1}(\zeta(u_0))$. The map φ is left-continuous, and since ε is such that Ψ_-^{-1} is continuous at $\Psi_+(u_0 - \varepsilon)$, ζ is left-continuous at

u_0. So there exists $\tau \in (0, \varepsilon)$ such that for all $u \in [u_0 - \tau, u_0]$, $\zeta(u) < c < \varphi(u)$. We deduce that for all $u \in [u_0 - \tau, u_0]$, $F_\nu^{-1}(\varphi(u)) > F_\nu^{-1}(\zeta(u))$. So for du-almost all $u \in [u_0 - \tau, u_0]$, we have

$$F_\nu^{-1}(\varphi(u)) > F_\nu^{-1}(\zeta(u)) > F_\mu^{-1}(u) > F_\nu^{-1}(u). \tag{4.11}$$

Let $a, b, c, d \in \mathbb{R}$ be such that $a > b > c > d$. Then

$$\begin{aligned}\left(\frac{c-d}{a-d}\right) a^2 + \left(\frac{a-c}{a-d}\right) d^2 &= \frac{ca^2 - da^2 + ad^2 - cd^2}{a-d} = \frac{(a-d)(ac - ad + dc)}{(a-d)} \\ &= a(c-d) + dc > b(c-d) + dc \\ &= \left(\frac{c-d}{b-d}\right) b^2 + \left(\frac{b-c}{b-d}\right) d^2.\end{aligned}$$

Thanks to (4.11) and this inequality applied with

$$(a, b, c, d) = (F_\nu^{-1}(\varphi(u)), F_\nu^{-1}(\zeta(u)), F_\mu^{-1}(u), F_\nu^{-1}(u)),$$

we deduce that for du-almost all $u \in [u_0 - \tau, u_0]$,

$$\int_{\mathbb{R}} y^2\, \widetilde{m}^{IT}(u, dy) > \int_{\mathbb{R}} y^2\, \widetilde{m}^{Q^\zeta}(u, dy). \tag{4.12}$$

By Lemma 2.6, we have

$$\begin{aligned}&\int_{\mathbb{R}^2} \mathbb{1}_{\{F_\mu^{-1}(u_0-\tau) < x < F_\mu^{-1}(u_0)\}} y^2\, M^{IT}(dx, dy) \\ &= \int_{u=0}^1 \mathbb{1}_{\{F_\mu^{-1}(u_0-\tau) < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)\}} \int_{y \in \mathbb{R}} y^2\, \widetilde{m}^{IT}(u, dy)\, du.\end{aligned}$$

For all $u \in (0, 1)$ such that $F_\mu^{-1}(u_0 - \tau) < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)$, we have $u \in [u_0 - \tau, u_0]$. So by (4.11), for $Q \in \{Q^{IT}, Q^\zeta\}$ and for du-almost all $u \in (0, 1)$ such that $F_\mu^{-1}(u_0 - \tau) < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)$, $y^2 \le \max(F_\nu^{-1}(\varphi(u_0))^2, F_\nu^{-1}(u_0 - \varepsilon)^2)$, $\widetilde{m}^Q(u, dy)$-almost everywhere. Therefore, for $Q \in \{Q^{IT}, Q^\zeta\}$, we have

$$\begin{aligned}\int_{u=0}^1 \mathbb{1}_{\{F_\mu^{-1}(u_0-\tau) < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)\}} \int_{y \in \mathbb{R}} y^2\, \widetilde{m}^Q(u, dy)\, du &\le \max(F_\nu^{-1}(\varphi(u_0))^2, F_\nu^{-1}(u_0 - \varepsilon)^2) \\ &< +\infty.\end{aligned}$$

Since by assumption the Lebesgue measure of $\{u \in (0, 1) \mid F_\mu^{-1}(u_0 - \tau) < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)\}$ is positive, according to (4.12), we get that

$$\begin{aligned}&\int_{\mathbb{R}^2} \mathbb{1}_{\{F_\mu^{-1}(u_0-\tau) < x < F_\mu^{-1}(u_0)\}} y^2\, M^{IT}(dx, dy) \\ &> \int_{u=0}^1 \mathbb{1}_{\{F_\mu^{-1}(u_0-\tau) < F_\mu^{-1}(u) < F_\mu^{-1}(u_0)\}} \int_{y \in \mathbb{R}} y^2\, \widetilde{m}^{Q^\zeta}(u, dy)\, du \\ &= \int_{\mathbb{R}^2} \mathbb{1}_{\{F_\mu^{-1}(u_0-\tau) < x < F_\mu^{-1}(u_0)\}} y^2\, M^{Q^\zeta}(dx, dy).\end{aligned}$$

So $M^{IT} \neq M^{Q^\zeta}$. By Proposition 2.9, we deduce that $(M^{\lambda Q^{IT} + (1-\lambda) Q^\zeta})_{\lambda \in [0,1]}$ is a family of distinct martingale couplings between μ and ν. ☐

5 Corresponding super and submartingale couplings

We recall that two probability measures $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ are in the decreasing (resp. increasing) convex order and denote $\mu\le_{dcx}\nu$ (resp. $\mu\le_{icx}\nu$) if $\int_\mathbb{R} f(x)\,\mu(dx)\le\int_\mathbb{R} f(y)\,\nu(dy)$ for any decreasing (resp. increasing) convex function $f:\mathbb{R}\to\mathbb{R}$. For two probability measures $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ such that $\mu\le_{dcx}\nu$ (resp. $\mu\le_{icx}\nu$), let

$$\mathcal{S}_1(\mu,\nu)=\inf\int_{\mathbb{R}\times\mathbb{R}}|x-y|\,M(dx,dy),$$

where the infimum is taken over all supermartingale (resp. submartingale) couplings M between μ and ν. Our main result, namely Theorem 2.12, can be generalised for the decreasing and increasing convex orders. We use the definitions of $\mathcal{U}_+$, $\mathcal{U}_-$, $\mathcal{U}_0$ given by (2.3) and the definitions of Ψ_+ and Ψ_- given at the beginning of Section 3.1.

Theorem 5.1. *For all $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ such that $\mu\le_{dcx}\nu$,*

$$\mathcal{S}_1(\mu,\nu)\le 2\Psi_-(1)+\mathcal{W}_1(\mu,\nu). \tag{5.1}$$

For all $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ such that $\mu\le_{icx}\nu$,

$$\mathcal{S}_1(\mu,\nu)\le 2\Psi_+(1)+\mathcal{W}_1(\mu,\nu). \tag{5.2}$$

Remark 5.2. In the martingale case, that is $\mu\le_{cx}\nu$, we have that $2\Psi_-(1)=2\Psi_+(1)=\mathcal{W}_1(\mu,\nu)$, consequence of the equality of the means and (1.2), so that we find Theorem 2.12 again.

The statement (5.2) for the increasing convex order can easily be deduced from (5.1) for the decreasing convex order. Indeed, let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu\le_{icx}\nu$. For any probability measure τ on $\mathbb{R}$ or $\mathbb{R}^2$, let $\overline{\tau}$ denote the image of τ by $x\mapsto -x$, so that $\overline{\mu}\le_{dcx}\overline{\nu}$. By the inverse transform sampling, $-F_\mu^{-1}(1-U)$ is distributing according to $\overline{\mu}$ for U a random variable uniformly distributed on $(0,1)$. Since $u\mapsto -F_\mu^{-1}(1-u)$ is nondecreasing, we have $F_{\overline{\mu}}^{-1}(u)=-F_\mu^{-1}(1-u)$, du-almost everywhere on $(0,1)$ (see for instance [2, Lemma A.3] and Lemma 6.5 below for an idea of the proof). The fact that quantile functions are left-continuous and have at most countable sets of discontinuities then yields $F_{\overline{\mu}}^{-1}(u)=-F_\mu^{-1}((1-u)_+)$ for all $u\in(0,1)$. Since the map $M\mapsto\overline{M}$ is a one-to-one correspondence between the set of supermartingale couplings between $\overline{\mu}$ and $\overline{\nu}$ and the set of submartingale couplings between μ and ν, we have $\mathcal{S}_1(\overline{\mu},\overline{\nu})=\mathcal{S}_1(\mu,\nu)$. So if (5.1) is true for $\overline{\mu}\le_{dcx}\overline{\nu}$, then

$$\begin{aligned}\mathcal{S}_1(\mu,\nu)=\mathcal{S}_1(\overline{\mu},\overline{\nu})&\le 2\int_0^1(F_{\overline{\mu}}^{-1}-F_{\overline{\nu}}^{-1})^-(u)\,du+\mathcal{W}_1(\overline{\mu},\overline{\nu})\\&=2\int_0^1(F_\nu^{-1}-F_\mu^{-1})^-(u)\,du+\mathcal{W}_1(\mu,\nu)=2\int_0^1(F_\mu^{-1}-F_\nu^{-1})^+(u)\,du+\mathcal{W}_1(\mu,\nu)\\&=2\Psi_+(1)+\mathcal{W}_1(\mu,\nu),\end{aligned}$$

hence (5.2) holds.

Frow now on, we suppose $\mu\le_{dcx}\nu$. We recall that two probability measures $\eta,\tau\in\mathcal{P}(\mathbb{R})$ are in the stochastic order, denoted $\eta\le_{st}\tau$, iff for all $u\in(0,1)$, $F_\eta^{-1}(u)\le F_\tau^{-1}(u)$, and in that case $\tau\le_{dcx}\eta$. If $\nu\le_{st}\mu$, then for U a random variable uniformly distributed on $(0,1)$, by the inverse transform sampling, $(F_\mu^{-1}(U),F_\nu^{-1}(U))$ is a supermartingale coupling between μ and ν, that is $\mathbb{E}[F_\nu^{-1}(U)|F_\mu^{-1}(U)]\le F_\mu^{-1}(U)$ almost surely. In that case,

$$\mathcal{S}_1(\mu,\nu)\le\mathbb{E}[F_\nu^{-1}(U)-F_\mu^{-1}(U)]=\mathcal{W}_1(\mu,\nu), \tag{5.3}$$

so (5.1) is satisfied as soon as $\nu \leq_{st} \mu$, which is equivalent to $\Psi_-(1) = 0$. If $\nu \not\leq_{st} \mu$, then Inequality (5.1) is a direct consequence of Proposition 5.5 and Proposition 5.7 below. As mentioned above, this concludes the proof of Theorem 5.1.

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{dcx} \nu$ and $\nu \not\leq_{st} \mu$, so that $\Psi_-(1) > 0$. According to [21, Theorem 4.A.3 Chapter 4], for all $u \in [0,1]$, $\int_0^u F_\mu^{-1}(v)\,dv \geq \int_0^u F_\nu^{-1}(v)\,dv$. This implies that for all $u \in [0,1]$, $\Psi_+(u) \geq \Psi_-(u)$. Let $\widetilde{\mathcal{U}}_+$ be a measurable subset of $\mathcal{U}_+$ which satisfies

$$\forall u \in (0,1), \quad \int_0^u \mathbb{1}_{\widetilde{\mathcal{U}}_+(v)}(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \geq \int_0^u (F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv, \quad \text{with equality for } u = 1. \tag{5.4}$$

Let $u_d = \Psi_+^{-1}(\Psi_-(1))$. Since Ψ_+ is continuous, $\Psi_+(u_d) = \Psi_-(1)$. If $u_d = 0$, then $\Psi_-(1) = 0$, which implies that for all $u \in (0,1)$, $F_\mu^{-1}(u) \geq F_\nu^{-1}(u)$. We deduce that the condition $\nu \not\leq_{st} \mu$ is equivalent to $u_d > 0$. One readily sees that (5.4) is satisfied for $\widetilde{\mathcal{U}}_+ = (0, u_d)$. Let

$$\overline{u} = \sup\{u \in [0,1] \mid \Psi_+(u) = \Psi_-(u)\}. \tag{5.5}$$

We deduce from the definition of $\overline{u}$ and (5.4) that $\Psi_+(\overline{u}) = \int_0^{\overline{u}} \mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du$, so for du-almost all $u \in \mathcal{U}_+ \cap [0, \overline{u}]$, $u \in \widetilde{\mathcal{U}}_+$. Therefore, the only room for manoeuvre of $\widetilde{\mathcal{U}}_+$ is $[\overline{u}, 1]$.

Let $\gamma = \int_0^1 (F_\mu^{-1} - F_\nu^{-1})^-(u)\,du \in (0, +\infty)$. We note $\mathcal{Q}$ the set of probability measures Q on $(0,1)^2$ such that there exists a measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ which satisfies (5.4) and

(i) Q has first marginal $\frac{1}{\gamma}\mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du$;

(ii) Q has second marginal $\frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv$;

(iii) $Q\left(\{(u,v) \in (0,1)^2 \mid u < v\}\right) = 1$.

For $\widetilde{\mathcal{U}}_+$ a measurable subset of $\mathcal{U}_+$ satisfying (5.4), let $\widetilde{\Psi}_+ : [0,1] \to \mathbb{R}_+$ be defined for all $u \in [0,1]$ by $\widetilde{\Psi}_+(u) = \int_0^u \mathbb{1}_{\widetilde{\mathcal{U}}_+}(v)(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv$. Let $\varphi : [0,1] \to [0,1]$ and $\widetilde{\varphi} : [0,1] \to [0,1]$ be defined for all $u \in [0,1]$ by

$$\varphi(u) = \Psi_-^{-1}(\widetilde{\Psi}_+(u)) = \inf\{r \in [0,1] \mid \Psi_-(r) \geq \widetilde{\Psi}_+(u)\};$$
$$\widetilde{\varphi}(u) = \widetilde{\Psi}_+^{-1}(\Psi_-(v)) = \inf\{r \in [0,1] \mid \widetilde{\Psi}_+(r) \geq \Psi_-(u)\},$$

which are well defined thanks to the equality $\Psi_-(1) = \widetilde{\Psi}_+(1)$, consequence of (5.4). Let then $Q^{IT}_{\widetilde{\mathcal{U}}_+}$ be the measure defined on $(0,1)^2$ by

$$Q^{IT}_{\widetilde{\mathcal{U}}_+}(du, dv) = \frac{1}{\gamma}\mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du\,\pi_+^Q(u, dv), \quad \text{where} \quad \pi_+^Q(u, dv) = \mathbb{1}_{\{0<\varphi(u)<1\}}\,\delta_{\varphi(u)}(dv). \tag{5.6}$$

Proposition 5.3. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{dcx} \nu$ and $\nu \not\leq_{st} \mu$. Let $\widetilde{\mathcal{U}}_+$ be a measurable subset of $\mathcal{U}_+$ such that* (5.4) *holds. The measure $Q^{IT}_{\widetilde{\mathcal{U}}_+}$ defined by* (5.6) *is an element of $\mathcal{Q}$. Moreover,*

$$Q^{IT}_{\widetilde{\mathcal{U}}_+}(du, dv) = \frac{1}{\gamma}(F_\mu^{-1} - F_\nu^{-1})^-(v)\,dv\,\pi_-^Q(v, du), \text{ where } \pi_-^Q(v, du) = \mathbb{1}_{\{0<\widetilde{\varphi}(v)<1\}}\,\delta_{\widetilde{\varphi}(v)}(du). \tag{5.7}$$

Proof of Proposition 5.3. A mild adaptation of the proof of Proposition 3.1 is conclusive. In particular, (3.5) is replaced with

$$\int_0^1 h(u,\varphi(u))\,d\widetilde{\Psi}_+(u)=\int_0^1 h(\widetilde{\varphi}(v),v)\,d\Psi_-(v),$$

for any measurable and bounded function $h:[0,1]^2\to\mathbb{R}$, consequence of Lemma 6.1 below with $f_1:u\mapsto \mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1}-F_\nu^{-1})^+(u)$, $f_2:v\mapsto(F_\mu^{-1}-F_\nu^{-1})^-(v)$ and $u_0=1$, which gives the key property to show that $Q^{IT}_{\widetilde{\mathcal{U}}_+}\in\mathcal{Q}$. $\square$

The existence of the inverse transform supermartingale coupling introduced below for the choice $\widetilde{\mathcal{U}}_+=(0,u_d)$ implies that $\mathcal{Q}$ is non-empty. More generally, for any measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ satisfying (5.4), there exists $Q\in\mathcal{Q}$ with first marginal $\frac{1}{\gamma}\mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1}-F_\nu^{-1})^+(u)\,du$. For Q an element of $\mathcal{Q}$, let π_+^Q and π_-^Q be two sub-Markov kernels such that

$$Q(du,dv)=\frac{1}{\gamma}\mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1}(u)-F_\nu^{-1})^+(u)\,du\,\pi_+^Q(u,dv)=\frac{1}{\gamma}(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv\,\pi_-^Q(v,du).$$

Let $(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$ be the Markov kernel defined by

$$\begin{cases}\displaystyle\int_{(0,1)}\frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\delta_{F_\nu^{-1}(v)}(dy)\pi_+^Q(u,dv)+\int_{(0,1)}\frac{F_\nu^{-1}(v)-F_\mu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\pi_+^Q(u,dv)\delta_{F_\nu^{-1}(u)}(dy)\\ \qquad\text{for } u\in\widetilde{\mathcal{U}}_+ \text{ such that } \pi_+^Q(u,\{v\in(0,1)\mid F_\nu^{-1}(v)>F_\mu^{-1}(u)\})=1;\\ \displaystyle\int_{\widetilde{\mathcal{U}}_+}\frac{F_\mu^{-1}(u)-F_\nu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\delta_{F_\nu^{-1}(v)}(dy)\pi_-^Q(u,dv)+\int_{\widetilde{\mathcal{U}}_+}\frac{F_\nu^{-1}(v)-F_\mu^{-1}(u)}{F_\nu^{-1}(v)-F_\nu^{-1}(u)}\pi_-^Q(u,dv)\delta_{F_\nu^{-1}(u)}(dy)\\ \qquad\text{for } u\in\mathcal{U}_- \text{ such that } \pi_-^Q(u,\{v\in(0,1)\mid F_\nu^{-1}(v)<F_\mu^{-1}(u)\})=1;\\ \delta_{F_\nu^{-1}(u)}(dy)\qquad\text{otherwise.}\end{cases}\tag{5.8}$$

The idea of this construction is as follows: for $u\in\mathcal{U}_-$, we can associate to $F_\mu^{-1}(u)$ a martingale contribution with $F_\nu^{-1}(u)$ and $F_\nu^{-1}(v)$ as in Section 2. If $F_\mu^{-1}(u)=F_\nu^{-1}(u)$, we associate $F_\nu^{-1}(u)$ to $F_\mu^{-1}(u)$. For $u\in\mathcal{U}_+\backslash\widetilde{\mathcal{U}}_+$, we only associate $F_\nu^{-1}(u)<F_\mu^{-1}(u)$ to $F_\mu^{-1}(u)$ since there is no partner $v\in\mathcal{U}_-\cap(u,1)$ available to construct a martingale contribution: all such possible partners have already been associated to values in $\widetilde{\mathcal{U}}_+$. Since du-almost all u in $\mathcal{U}_+\cap[0,\overline{u}]$ belong to $\widetilde{\mathcal{U}}_+$, our construction is such that we associate to $F_\mu^{-1}(u)$ a martingale contribution at least for du-almost all $u\in\mathcal{U}_+\cap[0,\overline{u}]$, which is actually not a particularity of our construction but a common property satisfied by all supermartingale couplings, as shown in the next proposition.

Proposition 5.4. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu\le_{dcx}\nu$, $M(dx,dy)=\mu(dx)\,m(x,dy)$ be a supermartingale coupling between μ and ν and $\overline{u}$ be defined by* (5.5)*. Then $\int_\mathbb{R} y\,m(x,dy)=x$ for $\mu(dx)$-almost all $x\le F_\mu^{-1}(\overline{u})$, or equivalently,*

$$\int_\mathbb{R}\left(x-\int_\mathbb{R} y\,m(x,dy)\right)\mathbb{1}_{\{x\le F_\mu^{-1}(\overline{u})\}}\,\mu(dx)=0.\tag{5.9}$$

Proof. Let $(\widetilde{m}(u,dy))_{u\in(0,1)}=(m(F_\mu^{-1}(u),dy))_{u\in(0,1)}$ and η be the image of μ by the map $x\mapsto\int_\mathbb{R} y\,m(x,dy)$. Since M is a supermartingale coupling, we deduce by the inverse transform sampling that $F_\mu^{-1}(u)\ge\int_\mathbb{R} y\,\widetilde{m}(u,dy)$ for du-almost all $u\in(0,1)$. Therefore,

$$\int_0^{\overline{u}}F_\mu^{-1}(u)\,du\ge\int_0^{\overline{u}}\int_\mathbb{R} y\,\widetilde{m}(u,dy)\,du=\int_0^{\overline{u}}\int_\mathbb{R} y\,m(F_\mu^{-1}(u),dy)\,du.\tag{5.10}$$

Let U and V be two independent random variables uniformly distributed on $(0,1)$. By the inverse transform sampling, $\int_{\mathbb{R}} y\, m(F_\mu^{-1}(U), dy)$ is distributed according to η, so by Lemma 6.6 below, the map $f : (0,1)^2 \to \mathbb{R}$ defined for all $u, v \in (0,1)$ by

$$f(u,v) = F_\eta\left(\left(\int_{\mathbb{R}} y\, m(F_\mu^{-1}(u), dy)\right)_-\right) + v\eta\left(\left\{\int_{\mathbb{R}} y\, m(F_\mu^{-1}(u), dy)\right\}\right)$$

is such that the random variable $f(U,V)$ is uniformly distributed on $(0,1)$ and satisfies $F_\eta^{-1}(f(U,V)) = \int_{\mathbb{R}} y\, m(F_\mu^{-1}(U), dy)$ almost surely. For $d \in \{1,2\}$, let λ_d denote the Lebesgue measure on $[0,1]^d$ and $A = f((0,\overline{u}) \times (0,1))$. Then $\overline{u} = \lambda_2((0,\overline{u}) \times (0,1)) \le \lambda_2(f^{-1}(f((0,\overline{u})\times(0,1)))) = \lambda_1(A)$. We deduce that $\lambda_1(A\cap(0,\overline{u})^\complement) = \lambda_1(A) - \lambda_1(A\cap(0,\overline{u})) \ge \lambda_1((0,\overline{u})) - \lambda_1(A\cap(0,\overline{u})) = \lambda_1(A^\complement \cap (0,\overline{u}))$ and

$$\begin{aligned}
\int_0^{\overline{u}} \int_{\mathbb{R}} y\, m(F_\mu^{-1}(u), dy)\, du &= \int_{(0,\overline{u})\times(0,1)} F_\eta^{-1}(f(u,v))\, du\, dv \\
&= \int_{A\cap(0,\overline{u})^\complement} F_\eta^{-1}(u)\, du + \int_{A\cap(0,\overline{u})} F_\eta^{-1}(u)\, du \\
&\ge \lambda_1(A\cap(0,\overline{u})^\complement) F_\eta^{-1}(\overline{u}) + \int_{A\cap(0,\overline{u})} F_\eta^{-1}(u)\, du \\
&\ge \lambda_1(A^\complement\cap(0,\overline{u})) F_\eta^{-1}(\overline{u}) + \int_{A\cap(0,\overline{u})} F_\eta^{-1}(u)\, du \\
&\ge \int_{A^\complement\cap(0,\overline{u})} F_\eta^{-1}(u)\, du + \int_{A\cap(0,\overline{u})} F_\eta^{-1}(u)\, du = \int_0^{\overline{u}} F_\eta^{-1}(u)\, du.
\end{aligned}$$

For any convex function $f : \mathbb{R} \to \mathbb{R}$, Jensen's inequality yields

$$\int_{\mathbb{R}} f(y)\, \eta(dy) = \int_{\mathbb{R}} f\left(\int_{\mathbb{R}} y\, m(x, dy)\right) \mu(dx) \le \int_{\mathbb{R}}\int_{\mathbb{R}} f(y)\, m(x,dy)\, \mu(dx) = \int_{\mathbb{R}} f(y)\, \nu(dy),$$

hence $\eta \le_{cx} \nu$. We then deduce from (3.3) that $\int_0^{\overline{u}} F_\eta^{-1}(u)\, du \ge \int_0^{\overline{u}} F_\nu^{-1}(v)\, dv$. Finally, we showed that

$$\begin{aligned}
0 = \Psi_+(\overline{u}) - \Psi_-(\overline{u}) &= \int_0^{\overline{u}} F_\mu^{-1}(u)\, du - \int_0^{\overline{u}} F_\nu^{-1}(u)\, du \\
&\ge \int_0^{\overline{u}} F_\mu^{-1}(u)\, du - \int_0^{\overline{u}} \int_{\mathbb{R}} y\, m(F_\mu^{-1}(u), dy)\, du \ge 0,
\end{aligned}$$

where the last inequality comes from (5.10). Therefore we have $\int_0^{\overline{u}} (F_\mu^{-1}(u) - \int_{\mathbb{R}} y\, m(F_\mu^{-1}(u), dy))\, du = 0$. Let $u \in (0,1)$ be such that $\overline{u} \le u$ and $F_\mu^{-1}(\overline{u}) = F_\mu^{-1}(u)$. Then $\Psi_+(u) = \Psi_+(\overline{u}) = \Psi_-(\overline{u}) \le \Psi_-(u) \le \Psi_+(u)$, so these inequalities are equalities and $u = \overline{u}$ by definition of $\overline{u}$. Therefore, $u \le \overline{u} \iff F_\mu^{-1}(u) \le F_\mu^{-1}(\overline{u})$ and by the inverse transform sampling,

$$\begin{aligned}
0 &= \int_0^1 \left(F_\mu^{-1}(u) - \int_{\mathbb{R}} y\, m(F_\mu^{-1}(u), dy)\right) \mathbb{1}_{F_\mu^{-1}(u) \le F_\mu^{-1}(\overline{u})}\, du \\
&= \int_0^1 \left(x - \int_{\mathbb{R}} y\, m(x, dy)\right) \mathbb{1}_{x \le F_\mu^{-1}(\overline{u})}\, \mu(dx),
\end{aligned}$$

which proves (5.9). □

Let $(m^Q(x,dy))_{x\in\mathbb{R}}$ be the Markov kernel defined as in (2.6) with $(\widetilde{m}(u,dy))_{u\in(0,1)}$ replaced with $(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$. Then $\mu(dx)\, m^Q(x,dy)$ is expected to be a supermartingale coupling between μ and ν, as the next proposition states.

Proposition 5.5. *Let $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \le_{dcx} \nu$ and $\nu \not\le_{st} \mu$. Then for all Q in the non-empty set $\mathcal{Q}$, the probability measure $M^Q(dx,dy) = \mu(dx)\, m^Q(x,dy)$ is a supermartingale coupling between μ and ν.*

Notice that M^Q is a martingale coupling between μ and ν iff μ and ν have equal means, which is equivalent to $\Psi_+(1) = \Psi_-(1)$.

Proof of Proposition 5.5. With the very same arguments as in Section 2, we show that $M^Q(dx,dy)$ is a coupling between μ and ν (see Proposition 2.3). The same calculation as (2.10) for du-almost all $u \in \widetilde{\mathcal{U}}_+ \cup \mathcal{U}_-$ and the definition of $\widetilde{m}^Q$ for $u \in \mathcal{U}_0$ and $u \in \mathcal{U}_+ \backslash \widetilde{\mathcal{U}}_+$ yield

$$\int_{\mathbb{R}} |y| \widetilde{m}^Q(u,dy) < +\infty,$$

$$\text{and} \quad \int_{\mathbb{R}} y\, \widetilde{m}^Q(u,dy) = \begin{cases} F_\mu^{-1}(u) & \text{for } du\text{-almost all } u \in \widetilde{\mathcal{U}}_+ \cup \mathcal{U}_-; \\ F_\nu^{-1}(u) = F_\mu^{-1}(u) & \text{for } u \in \mathcal{U}_0; \\ F_\nu^{-1}(u) < F_\mu^{-1}(u) & \text{for } u \in \mathcal{U}_+ \backslash \widetilde{\mathcal{U}}_+. \end{cases} \tag{5.11}$$

Therefore, for du-almost all $u \in (0,1)$, $\int_{\mathbb{R}} y\, \widetilde{m}^Q(u,dy) \le F_\mu^{-1}(u)$. Let $h : \mathbb{R} \to \mathbb{R}$ be a measurable nonnegative and bounded function. By Lemma 2.6,

$$\int_{\mathbb{R}\times\mathbb{R}} h(x)(y-x)\, \mu(dx)\, m^Q(x,dy) = \int_0^1 h(F_\mu^{-1}(u)) \left(\int_{\mathbb{R}} (y - F_\mu^{-1}(u))\, \widetilde{m}^Q(u,dy) \right) du \le 0. \tag{5.12}$$

Therefore, for all $Q \in \mathcal{Q}$, $M^Q(dx,dy)$ is a supermartingale coupling between μ and ν. □

For $\widetilde{\mathcal{U}}_+$ a measurable subset of $\mathcal{U}_+$ which satifies (5.4), let us write $(m^{IT}_{\widetilde{\mathcal{U}}_+}(x,dy))_{x\in\mathbb{R}}$ instead of $(m^{Q^{IT}_{\widetilde{\mathcal{U}}_+}}(x,dy))_{x\in\mathbb{R}}$ and $(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}(u,dy))_{u\in(0,1)}$ instead of $(\widetilde{m}^{Q^{IT}_{\widetilde{\mathcal{U}}_+}}(u,dy))_{u\in(0,1)}$, whose definition, given by (5.8), reduces to

$$\begin{cases} \dfrac{F_\mu^{-1}(u) - F_\nu^{-1}(u)}{F_\nu^{-1}(\varphi(u)) - F_\nu^{-1}(u)} \delta_{F_\nu^{-1}(\varphi(u))}(dy) + \left(1 - \dfrac{F_\mu^{-1}(u) - F_\nu^{-1}(u)}{F_\nu^{-1}(\varphi(u)) - F_\nu^{-1}(u)}\right) \delta_{F_\nu^{-1}(u)}(dy) \\ \qquad\qquad \text{if } u \in \widetilde{\mathcal{U}}_+,\ F_\nu^{-1}(\varphi(u)) > F_\mu^{-1}(u) > F_\nu^{-1}(u) \text{ and } \varphi(u) < 1; \\ \dfrac{F_\nu^{-1}(u) - F_\mu^{-1}(u)}{F_\nu^{-1}(u) - F_\nu^{-1}(\widetilde{\varphi}(u))} \delta_{F_\nu^{-1}(\widetilde{\varphi}(u))}(dy) + \left(1 - \dfrac{F_\nu^{-1}(u) - F_\mu^{-1}(u)}{F_\nu^{-1}(u) - F_\nu^{-1}(\widetilde{\varphi}(u))}\right) \delta_{F_\nu^{-1}(u)}(dy) \\ \qquad\qquad \text{if } F_\nu^{-1}(\widetilde{\varphi}(u)) < F_\mu^{-1}(u) < F_\nu^{-1}(u) \text{ and } \widetilde{\varphi}(u) < 1; \\ \delta_{F_\nu^{-1}(u)}(dy) \qquad \text{otherwise.} \end{cases} \tag{5.13}$$

Then $M^{IT}_{\widetilde{\mathcal{U}}_+}(dx,dy) = \mu(dx)\, m^{IT}_{\widetilde{\mathcal{U}}_+}(x,dy)$ is a supermartingale coupling. Let $Q^{ITS} = Q^{IT}_{(0,u_d)}$, that is the element of $\mathcal{Q}$ defined by (5.6) for $\widetilde{\mathcal{U}}_+ = (0,u_d)$. From now on, we write $(\widetilde{m}^{ITS}(u,dy))_{u\in(0,1)}$ and $(m^{ITS}(x,dy))_{x\in\mathbb{R}}$ instead of $(\widetilde{m}^{Q^{ITS}}(u,dy))_{u\in(0,1)} = (\widetilde{m}^{IT}_{(0,u_d)}(u,dy))_{u\in(0,1)}$ and $(m^{Q^{ITS}}(x,dy))_{x\in\mathbb{R}} = (m^{IT}_{(0,u_d)}(x,dy))_{x\in\mathbb{R}}$ respectively, and call inverse transform supermartingale coupling the probability measure $M^{ITS}(dx,dy) = \mu(dx)\, m^{ITS}(x,dy)$.

The next statement generalises Proposition 2.18.

Proposition 5.6. *Let $\mu,\nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{dcx} \nu$ and $\nu \not\leq_{st} \mu$. Let $Q \in \mathcal{Q}$. Then the Markov kernel $(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$ minimises*

$$\int_0^1 \int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\, \widetilde{m}(u,dy)\, du$$

among all Markov kernels $(\widetilde{m}(u,dy))_{u\in(0,1)}$ such that

$$\begin{aligned} &\int_{u\in(0,1)} \widetilde{m}(u,dy)\, du = \nu(dy), \quad \int_{\mathbb{R}} |y|\, \widetilde{m}(u,dy) < +\infty, \\ \text{and} \quad &\int_{\mathbb{R}} y\, \widetilde{m}(u,dy) \leq F_\mu^{-1}(u),\ du\text{-almost everywhere on } (0,1). \end{aligned} \tag{5.14}$$

Moreover, $\int_0^1 \int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\, \widetilde{m}^Q(u,dy)\, du = 2\Psi_-(1)$.

Proof. Let $\widetilde{m}$ be a Markov kernel satisfying (5.14). By monotonicity of the negative part and Jensen's inequality, for du-almost every $u \in (0,1)$,

$$(F_\mu^{-1} - F_\nu^{-1})^-(u) \leq \left(\int_{\mathbb{R}} (y - F_\nu^{-1}(u))\, \widetilde{m}(u,dy) \right)^- \leq \int_{\mathbb{R}} (y - F_\nu^{-1}(u))^-\, \widetilde{m}(u,dy).$$

Using the equality $2x^- = |x| - x$ valid for $x \in \mathbb{R}$ and the inverse transform sampling, we deduce that

$$\begin{aligned} 2\Psi_-(1) &\leq 2\int_0^1 \int_{\mathbb{R}} (y - F_\nu^{-1}(u))^-\, \widetilde{m}(u,dy)\, du \\ &= \int_0^1 \int_{\mathbb{R}} |y - F_\nu^{-1}(u)|\, \widetilde{m}(u,dy)\, du - \int_0^1 \int_{\mathbb{R}} y\, \widetilde{m}(u,dy)\, du + \int_0^1 F_\nu^{-1}(u)\, du \\ &= \int_0^1 \int_{\mathbb{R}} |y - F_\nu^{-1}(u)|\, \widetilde{m}(u,dy)\, du - \int_{\mathbb{R}} y\, \nu(dy) + \int_{\mathbb{R}} y\, \nu(dy) \\ &= \int_0^1 \int_{\mathbb{R}} |y - F_\nu^{-1}(u)|\, \widetilde{m}(u,dy)\, du. \end{aligned}$$

According to Proposition 5.5, $\mu(dx)\, m^Q(x,dy)$ is a coupling between μ and ν, so by Lemma 2.6, $\int_{u\in(0,1)} \widetilde{m}^Q(u,dy)\, dy = \nu(dy)$. Moreover, we deduce from (5.11) that $(\widetilde{m}^Q(u,dy))_{u\in(0,1)}$ satisfies (5.14). Therefore, to conclude, it is sufficient to prove that $\int_0^1 \int_{\mathbb{R}} |y - F_\nu^{-1}(u)|\, \widetilde{m}^Q(u,dy)\, du = 2\Psi_-(1)$.

Using the definition (5.8) of $\widetilde{m}^Q$, we get for du-almost all $u \in (0,1)$

$$\begin{aligned} \int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\, \widetilde{m}^Q(u,dy) = &\int_{(0,1)} \mathbb{1}_{\widetilde{\mathcal{U}}_+}(u) \frac{(F_\mu^{-1} - F_\nu^{-1})^+(u)}{F_\nu^{-1}(v) - F_\nu^{-1}(u)} |F_\nu^{-1}(u) - F_\nu^{-1}(v)| \pi_+^Q(u,dv) \\ &+ \int_{(0,1)} \frac{(F_\mu^{-1} - F_\nu^{-1})^-(u)}{F_\nu^{-1}(u) - F_\nu^{-1}(v)} |F_\nu^{-1}(u) - F_\nu^{-1}(v)| \pi_-^Q(u,dv). \end{aligned}$$

A mild adaptation of the proof of Lemma 2.5 yields for du-almost all $u \in (0,1)$,

$$\begin{cases} u \in \widetilde{\mathcal{U}}_+ \implies F_\nu^{-1}(v) > F_\mu^{-1}(u),\ \pi_+^Q(u,dv)\text{-a.e;} \\ u \in \mathcal{U}_- \implies F_\nu^{-1}(v) < F_\mu^{-1}(u),\ \pi_-^Q(u,dv)\text{-a.e.} \end{cases} \tag{5.15}$$

We deduce that $\int_{\mathbb{R}} |F_\nu^{-1}(u) - y|\, \widetilde{m}^Q(u,dy) \leq \mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1} - F_\nu^{-1})^+(u) + (F_\nu^{-1} - F_\mu^{-1})^-(u)$ for du-almost all $u \in (0,1)$. Using (5.4) for $u = 1$, we conclude that $\int_0^1 \int_{\mathbb{R}} |y - F_\nu^{-1}(u)|\, \widetilde{m}^Q(u,dy)\, du \leq 2\Psi_-(1)$. □

The next statement generalises the first statement in Theorem 2.12.

Proposition 5.7. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu\le_{dcx}\nu$ and $\nu\not\le_{st}\mu$. For all Q in the non-empty set $\mathcal{Q}$,*

$$\int_{\mathbb{R}\times\mathbb{R}}|x-y|\,M^Q(dx,dy)\le 2\Psi_-(1)+\mathcal{W}_1(\mu,\nu). \tag{5.16}$$

Proof. Let $Q\in\mathcal{Q}$ and let $\widetilde{m}^Q$ be the Markov kernel defined by (5.8). By Lemma 2.6 and Proposition 5.6,

$$\begin{aligned}&\int_{\mathbb{R}\times\mathbb{R}}|y-x|\,M^Q(dx,dy)\\&=\int_{\mathbb{R}\times\mathbb{R}}|y-x|\,\mu(dx)\,m^Q(x,dy)=\int_0^1\int_{\mathbb{R}}|y-F_\mu^{-1}(u)|\,\widetilde{m}^Q(u,dy)\,du\\&\le\int_0^1\int_{\mathbb{R}}|y-F_\nu^{-1}(u)|\,\widetilde{m}^Q(u,dy)\,du+\int_0^1\int_{\mathbb{R}}|F_\nu^{-1}(u)-F_\mu^{-1}(u)|\,\widetilde{m}^Q(u,dy)\,du\\&=2\Psi_-(1)+\mathcal{W}_1(\mu,\nu).\end{aligned}$$

□

Among all the measurable subsets $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ which satisfy (5.4), $(0,u_d)$ is the leftmost one. This is one of the reasons for which the inverse transform supermartingale coupling plays a particular role among the supermartingale couplings which derive from $\mathcal{Q}$, as stated in the next Proposition. It is also natural to investigate the rightmost measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ which satisfies (5.4), that is such that $\widetilde{\Psi}_+$ is as small as possible. Notice that a measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ satisfies (5.4) iff it satisfies

$$\forall u\in(0,1),\quad \int_{1-u}^1 \mathbb{1}_{\widetilde{\mathcal{U}}_+(v)}(F_\mu^{-1}-F_\nu^{-1})^+(v)\,dv\le\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv, \tag{5.17}$$

with equality for $u=1$.

Therefore, we look for a measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ such that for $u\in[0,1]$, $\int_{1-u}^1\mathbb{1}_{\widetilde{\mathcal{U}}_+}(v)\,(F_\mu^{-1}-F_\nu^{-1})^+(v)\,dv$ is as big as possible while still being smaller than $\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv$ with equality for $u=1$. This is equivalent to have

$$\begin{aligned}&\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})(v)\,dv-\int_{1-u}^1\mathbb{1}_{\mathcal{U}_+\setminus\widetilde{\mathcal{U}}_+}(v)(F_\mu^{-1}-F_\nu^{-1})(v)\,dv\\&=\int_{1-u}^1\mathbb{1}_{\widetilde{\mathcal{U}}_+}(v)\,(F_\mu^{-1}-F_\nu^{-1})^+(v)\,dv-\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})^-(v)\,dv\le 0,\end{aligned}$$

with equality for $u=1$. Therefore, we look for a measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ such that $\int_{1-u}^1\mathbb{1}_{\mathcal{U}_+\setminus\widetilde{\mathcal{U}}_+}(v)(F_\mu^{-1}-F_\nu^{-1})(v)\,dv$ is as small as possible while still being greater than $\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})(v)\,dv$. Let then $R:[0,1]\to\mathbb{R}$ be defined for all $u\in[0,1]$ by

$$R(u)=\sup_{v\in[0,u]}\int_{1-v}^1(F_\mu^{-1}-F_\nu^{-1})(w)\,dw, \tag{5.18}$$

which can easily be proved to be the minimum of the set of nonnegative and nondecreasing functions $f:[0,1]\to\mathbb{R}$ which satisfy $f(u)\ge\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})(v)\,dv$ for all $u\in[0,1]$. The following proposition makes the connection between R and the rightmost measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ which satisfies (5.4).

Proposition 5.8. *Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R})$ be such that $\mu\le_{dcx}\nu$, R be defined by* (5.18) *and* $B:[0,1]\ni u\mapsto\int_{1-u}^1(F_\mu^{-1}-F_\nu^{-1})(v)\,dv$. *Let*

$$\widetilde{\mathcal{U}}_+^R=\{u\in\mathcal{U}_+\mid R(1-u)>B(1-u)\}\quad\textit{and}\quad\widetilde{\Psi}_+^R:u\mapsto\int_0^u\mathbb{1}_{\widetilde{\mathcal{U}}_+^R}(v)(F_\mu^{-1}-F_\nu^{-1})(v)\,dv. \tag{5.19}$$

Then $\widetilde{\mathcal{U}}_+^R$ is a measurable subset of $\mathcal{U}_+$ which satisfies (5.4) *and for any measurable subset $\widetilde{\mathcal{U}}_+$ of $\mathcal{U}_+$ satisfying* (5.4)*, we have that*

$$\forall u\in[0,1],\quad\widetilde{\Psi}_+^R(u)\le\widetilde{\Psi}_+(u).$$

Proof. For $\varepsilon>0$, let $\varphi_\varepsilon:\mathbb{R}\to\mathbb{R}$ be a continuously differentiable map such that $\varphi_\varepsilon(x)=0$ for $x\le-\varepsilon$, $\varphi_\varepsilon(x)=x$ for $x\ge\varepsilon$, $\varphi_\varepsilon'(x)\in[0,1]$ for $x\in\mathbb{R}$ and $\varphi_\varepsilon'(0)=1$. One could choose for instance

$$\varphi_\varepsilon:x\mapsto\left(\frac{\varepsilon}{2}+x+\frac{1}{2\varepsilon}x^2\right)\mathbb{1}_{\{-\varepsilon<x\le0\}}+\left(\frac{\varepsilon}{2}+x-\frac{3}{2\varepsilon}x^2+\frac{1}{\varepsilon^2}x^3\right)\mathbb{1}_{\{0<x\le\varepsilon\}}+x\mathbb{1}_{\{x>\varepsilon\}}.$$

Since φ_ε is continuously differentiable, the chain rule formula (see for instance [20, Proposition 4.6 Chapter 0]) yields for all $0\le u<v\le1$,

$$\varphi_\varepsilon((B-R)(v))-\varphi_\varepsilon((B-R)(u))=\int_{(u,v]}\varphi_\varepsilon'((B-R)(w))\,d(B-R)(w).$$

We deduce from the dominated convergence theorem for $\varepsilon\to0$ that for all $0\le u<v\le1$,

$$(B-R)^+(v)-(B-R)^+(u)=\int_{(u,v]}\mathbb{1}_{\{(B-R)(w)\ge0\}}\,d(B-R)(w).$$

Since $R(u)\ge B(u)$ for all $u\in[0,1]$, we get that

$$0=d(B-R)^+(u)=\mathbb{1}_{\{R(u)=B(u)\}}\,dB(u)-\mathbb{1}_{\{R(u)=B(u)\}}\,dR(u). \tag{5.20}$$

According to [18, Theorem 1.1.1], the map R solves a Skorokhod problem and may increase only at points $u\in(0,1)$ such that $R(u)=B(u)$, that is $dR(u)=\mathbb{1}_{\{R(u)=B(u)\}}\,dR(u)$. With (5.20), we deduce that

$$dR(u)=\mathbb{1}_{\{R(u)=B(u)\}}(F_\mu^{-1}-F_\nu^{-1})(1-u)\,du.$$

By monotonicity of R, we have

$$\begin{aligned}0&\le\int_{(0,1)}\mathbb{1}_{\{F_\mu^{-1}(1-u)\le F_\nu^{-1}(1-u)\}}\,dR(u)\\&=\int_{(0,1)}\mathbb{1}_{\{R(u)=B(u)\}}\mathbb{1}_{\{F_\mu^{-1}(1-u)\le F_\nu^{-1}(1-u)\}}(F_\mu^{-1}-F_\nu^{-1})(1-u)\,du\le0,\end{aligned}$$

so those inequalities are equalities and for $dR(u)$-almost all $u\in(0,1)$, $1-u\in\mathcal{U}_+$. Therefore, $dR(u)=\mathbb{1}_{\{R(u)=B(u)\}}\mathbb{1}_{\{(1-u)\in\mathcal{U}_+\}}(F_\mu^{-1}-F_\nu^{-1})(1-u)\,du$, so that the set $\widetilde{\mathcal{U}}_+^R:=\{u\in\mathcal{U}_+\mid R(1-u)>B(1-u)\}$ is such that for all $u\in[0,1]$, $R(u)=\int_{1-u}^1\mathbb{1}_{\mathcal{U}_+\setminus\widetilde{\mathcal{U}}_+^R}(v)(F_\mu^{-1}-F_\nu^{-1})(v)\,dv$.

Let us now prove that $\widetilde{\mathcal{U}}_+^R$ satisfies (5.4), which will end the proof. Let $\widetilde{\Psi}_+^R:u\mapsto\int_0^u\mathbb{1}_{\widetilde{\mathcal{U}}_+^R}(v)(F_\mu^{-1}-F_\nu^{-1})(v)\,dv$. On the one hand, using that $\Psi_+(u)\ge\Psi_-(u)$ for all $u\in[0,1]$, we have

$$\begin{aligned}B(1)\le R(1)&=\sup_{v\in[0,1]}B(v)=\sup_{v\in[0,1]}(\Psi_+(1)-\Psi_-(1)-\Psi_+(1-v)+\Psi_-(1-v))\\&\le\Psi_+(1)-\Psi_-(1)=B(1),\end{aligned}$$

so those inequalities are equalities and $R(1) = \Psi_+(1) - \Psi_-(1)$. We deduce that $\widetilde{\Psi}^R_+(1) = \Psi_+(1) - R(1) = \Psi_-(1)$. On the other hand, for $u \in [0,1]$,

$$\widetilde{\Psi}^R_+(u) = \widetilde{\Psi}^R_+(1) + R(1-u) - \Psi_+(1) + \Psi_+(u) \geq \Psi_-(1) + B(1-u) - \Psi_+(1) + \Psi_+(u) = \Psi_-(u),$$

so $\widetilde{\mathcal{U}}^R_+$ satisfies (5.4). □

Proposition 5.9. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \leq_{dcx} \nu$ and $\nu \not\leq_{st} \mu$. For all $\rho \in \mathbb{R}$ and for any Markov kernel $(\widetilde{m}(u,dy))_{u\in(0,1)}$, let $\mathcal{C}_\rho(\widetilde{m})$ be defined by*

$$\mathcal{C}_\rho(\widetilde{m}) = \int_{\mathbb{R}\times(0,1)} |F_\nu^{-1}(u) - y|^\rho \mathbb{1}_{\{y \neq F_\nu^{-1}(u)\}} \widetilde{m}(u,dy)\,du. \tag{5.21}$$

Let $(\widetilde{m}^R(u,dy))_{u\in(0,1)} = (\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}^R_+}(u,dy))_{u\in(0,1)}$, where $\widetilde{\mathcal{U}}^R_+$ is defined by (5.19). *Then, for all $Q \in \mathcal{Q}$,*

$$\begin{aligned} \forall \rho \in (-\infty,1], &\quad \mathcal{C}_\rho(\widetilde{m}^{ITS}) \leq \mathcal{C}_\rho(\widetilde{m}^Q); \\ \forall \rho \in [1,2], &\quad \mathcal{C}_\rho(\widetilde{m}^Q) \leq \mathcal{C}_\rho(\widetilde{m}^{ITS}); \\ \forall \rho \in [2,+\infty), &\quad \mathcal{C}_\rho(\widetilde{m}^R) \leq \mathcal{C}_\rho(\widetilde{m}^Q). \end{aligned} \tag{5.22}$$

Proof. Let $\widetilde{\mathcal{U}}_+$ be a subset of $\mathcal{U}_+$ which satisfies (5.4). Let Q be any element of $\mathcal{Q}$ with first marginal $\frac{1}{\gamma}\mathbb{1}_{\widetilde{\mathcal{U}}_+}(u)(F_\mu^{-1} - F_\nu^{-1})^+(u)\,du$, and $Q^{IT}_{\widetilde{\mathcal{U}}_+}$ be defined by (5.6). Reasoning like in the derivation of (3.13) and (3.15), we obtain

$$\begin{aligned} \forall \rho \in (-\infty,1] \cup [2,+\infty), \quad & \int_0^1 |F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+)|^{\rho-1}\,du \\ & \leq \int_{(0,1)^2} |F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}\,Q(du,dv), \end{aligned} \tag{5.23}$$

and

$$\begin{aligned} \forall 1 \leq \rho \leq 2, \quad & \int_{(0,1)^2} |F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}\,Q(du,dv) \\ & \leq \int_0^1 |F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+)|^{\rho-1}\,du. \end{aligned} \tag{5.24}$$

Moreover, (3.17) and (3.18) generalise into $C_\rho(\widetilde{m}^Q) = 2\gamma\int_{(0,1)^2}|F_\nu^{-1}(u_+) - F_\nu^{-1}(v_+)|^{\rho-1}\,Q(du,dv)$ and $\mathcal{C}_\rho(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}) = 2\gamma\int_0^1 |F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\gamma u)_+) - F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+)|^{\rho-1}\,du$. We deduce that

$$\forall \rho \in (-\infty,1] \cup [2,+\infty), \quad \mathcal{C}_\rho(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}) \leq \mathcal{C}_\rho(\widetilde{m}^Q) \quad \text{and} \quad \forall \rho \in [1,2], \quad \mathcal{C}_\rho(\widetilde{m}^Q) \leq \mathcal{C}_\rho(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}). \tag{5.25}$$

Notice that since $\Psi_-(u) \leq \widetilde{\Psi}_+(u)$ for $u \in [0,1]$, we have $\Psi_-^{-1}(v) \geq \widetilde{\Psi}_+^{-1}(v)$ for $v \in (0,\gamma)$, so by monotonicity of F_ν^{-1}, we also have

$$C_\rho(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}) = 2\gamma\int_0^1 \left(F_\nu^{-1}(\Psi_-^{-1}(\gamma u)_+ - F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\gamma u)_+))\right)^{\rho-1}\,du. \tag{5.26}$$

Let $\widetilde{\Psi}^{ITS}_+ : [0,1] \to \mathbb{R}$ and $\widetilde{\Psi}^R_+ : [0,1] \to \mathbb{R}$ be defined for all $u \in [0,1]$ by $\widetilde{\Psi}^{ITS}_+(u) = \int_0^{u\wedge u_d}(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv$ and $\widetilde{\Psi}^R_+(u) = \int_0^u \mathbb{1}_{\widetilde{\mathcal{U}}^R_+}(v)(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv$. For all $u \in [0,u_d]$, $\widetilde{\Psi}_+(u) = \int_0^u \mathbb{1}_{\widetilde{\mathcal{U}}_+}(v)(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \leq \int_0^u (F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \leq \widetilde{\Psi}^{ITS}_+(u)$ and for all $u \in$

$[u_d, 1]$, $\widetilde{\Psi}_+(u) \le \widetilde{\Psi}_+(1) = \Psi_-(1) = \widetilde{\Psi}^{ITS}(u)$. Moreover, let $\alpha : u \mapsto \int_{1-u}^1 \mathbb{1}_{\mathcal{U}_+ \setminus \widetilde{\mathcal{U}}_+}(v)(F_\mu^{-1} - F_\nu^{-1})(v)$. The map α is nonnegative, nondecreasing and satisfies

$$\begin{aligned}
\alpha(u) &= \int_{1-u}^1 \mathbb{1}_{\mathcal{U}_+}(v)(F_\mu^{-1} - F_\nu^{-1})(v)\,dv - \widetilde{\Psi}_+(1) + \widetilde{\Psi}_+(1-u) \\
&\ge \int_{1-u}^1 (F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv - \Psi_-(1) + \Psi_-(1-u) \\
&= \int_{1-u}^1 (F_\mu^{-1} - F_\nu^{-1})(v)\,dv,
\end{aligned}$$

where we used (5.4) for the inequality. By definition of R, we deduce that for all $u \in [0,1]$, $\alpha(u) \ge R(u)$, hence

$$\begin{aligned}
\widetilde{\Psi}_+^R(u) &= \widetilde{\Psi}_+^R(1) - \int_u^1 \mathbb{1}_{\widetilde{\mathcal{U}}_+^R}(v)(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \\
&= \widetilde{\Psi}_+(1) - \int_u^1 \mathbb{1}_{\widetilde{\mathcal{U}}_+^R}(v)(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \\
&= \widetilde{\Psi}_+(1) - \Psi_+(1) + \Psi_+(u) + \int_u^1 \mathbb{1}_{\mathcal{U}_+ \setminus \widetilde{\mathcal{U}}_+^R}(v)(F_\mu^{-1} - F_\nu^{-1})^+(v)\,dv \\
&= \widetilde{\Psi}_+(1) - \Psi_+(1) + \Psi_+(u) + R(1-u) \\
&\le \widetilde{\Psi}_+(1) - \Psi_+(1) + \Psi_+(u) + \alpha(1-u) = \widetilde{\Psi}_+(u).
\end{aligned}$$

Since $\widetilde{\Psi}_+^R(u) \le \widetilde{\Psi}_+(u) \le \widetilde{\Psi}_+^{ITS}(u)$ for all $u \in [0,1]$, we deduce that

$$\forall u \in (0,\gamma), \quad (\widetilde{\Psi}_+^{ITS})^{-1}(u) \le \widetilde{\Psi}_+^{-1}(u) \le (\widetilde{\Psi}_+^R)^{-1}(u). \tag{5.27}$$

By (5.26), (5.27) and monotonicity of the maps $\mathbb{R}_+ \ni x \mapsto x^{\rho-1}$ and F_ν^{-1}, we have

$$\begin{aligned}
&\forall \rho \in (-\infty, 1], \quad C_\rho(\widetilde{m}^{ITS}) \le C_\rho(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}) \le C_\rho(\widetilde{m}^R), \\
\text{and} \quad &\forall \rho \in [1, +\infty), \quad C_\rho(\widetilde{m}^R) \le C_\rho(\widetilde{m}^{IT}_{\widetilde{\mathcal{U}}_+}) \le C_\rho(\widetilde{m}^{ITS}).
\end{aligned} \tag{5.28}$$

Then (5.22) is deduced from (5.25) and (5.28). □

We now show the stability of the inverse transform supermartingale coupling with respect to its marginals for the Wasserstein distance topology. So far, the inverse transform supermartingale coupling has been defined just before Proposition 5.6 for all $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ such that $\mu \le_{dcx} \nu$ and $\nu \not\le_{st} \mu$. When $\nu \le_{st} \mu$, we simply define the inverse transform supermartingale coupling as the comonotonous coupling between μ and ν.

Proposition 5.10. *Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ be such that $\mu \le_{dcx} \nu$. Let $(\mu_n)_{n\in\mathbb{N}}$ and $(\nu_n)_{n\in\mathbb{N}}$ be two sequences of probability measures on $\mathbb{R}$ with finite first moments such that for all $n \in \mathbb{N}$, $\mu_n \le_{dcx} \nu_n$. For all $n \in \mathbb{N}$, let M_n^{ITS} (resp. M^{ITS}) be the inverse transform supermartingale coupling between μ_n and ν_n (resp. μ and ν).*

If $\mathcal{W}_1(\mu_n, \mu) \underset{n\to+\infty}{\longrightarrow} 0$ and $\mathcal{W}_1(\nu_n, \nu) \underset{n\to+\infty}{\longrightarrow} 0$, then

$$\mathcal{W}_1(M_n^{ITS}, M^{ITS}) \underset{n\to+\infty}{\longrightarrow} 0.$$

Proof. For all $n \in \mathbb{N}$, let $\Psi_{n+} : u \in [0,1] \mapsto \int_0^u (F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^+(v)\,dv$, $\Psi_{n-} : u \in [0,1] \mapsto \int_0^u (F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^-(v)\,dv$, $(u_d)_n = \Psi_{n+}^{-1}(\Psi_{n-}(1))$ if $\nu_n \not\le_{st} \mu_n$ and $(u_d)_n = 0$ otherwise, $\widetilde{\mathcal{U}}_{n+} = (0, (u_d)_n)$, $\widetilde{\Psi}_{n+} : u \in [0,1] \mapsto \int_0^u \mathbb{1}_{\widetilde{\mathcal{U}}_{n+}}(v)(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^+(v)\,dv$, $\varphi_n = \Psi_{n-}^{-1} \circ \widetilde{\Psi}_{n+}$ and $\widetilde{\varphi}_n = \widetilde{\Psi}_{n+}^{-1} \circ \Psi_{n-}$. Let $H : \mathbb{R}^2 \to \mathbb{R}$ be a bounded and continuous function such that h is

Lipschitz continuous with respect to its second variable. One can easily prove that (2.12) still holds in the supermartingale case, which writes, for $Q = Q^{IT}_{\widetilde{\mathcal{U}}_{n+}}$,

$$\begin{aligned}
&\int_{\mathbb{R}^2} H(x,y)\, M_n^{ITS}(dx,dy)\\
&= \int_{(0,1)} H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u))\, du\\
&+ \int_{(0,1)} \mathbb{1}_{\widetilde{\mathcal{U}}_{n+}}(u) \frac{(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^+(u)}{F_{\nu_n}^{-1}(\varphi_n(u)) - F_{\nu_n}^{-1}(u)} (H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(\varphi_n(u))) - H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u)))\, du\\
&+ \int_{(0,1)} \frac{(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^-(u)}{F_{\nu_n}^{-1}(u) - F_{\nu_n}^{-1}(\widetilde{\varphi}_n(u))} (H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(\widetilde{\varphi}_n(u))) - H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u)))\, du,
\end{aligned} \tag{5.29}$$

where the last two integrands are zero when $\nu_n \leq_{st} \mu_n$. Since μ_n converges weakly towards μ, then $F_{\mu_n}^{-1}(u)$ (resp. $F_{\nu_n}^{-1}(u)$) converges towards $F_\mu^{-1}(u)$ (resp. $F_\nu^{-1}(u)$) du-almost everywhere on $(0,1)$. Since H is continuous and bounded, by the dominated convergence theorem,

$$\int_{(0,1)} H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u))\, du \underset{n\to+\infty}{\longrightarrow} \int_{(0,1)} H(F_\mu^{-1}(u), F_\nu^{-1}(u))\, du. \tag{5.30}$$

Since for all $u \in [0,1]$, $x \mapsto x^+$ is Lipschitz continuous with constant 1,

$$\begin{aligned}
|\Psi_{n-}(u) - \Psi_-(u)| &\leq \int_0^u |(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^-(v) - (F_\mu^{-1} - F_\nu^{-1})^-(v)|\, dv\\
&\leq \int_0^u |F_{\mu_n}^{-1}(v) - F_\mu^{-1}(v)|\, dv + \int_0^u |F_{\nu_n}^{-1}(v) - F_\nu^{-1}(v)|\, dv\\
&\leq \mathcal{W}_1(\mu_n,\mu) + \mathcal{W}_1(\nu_n,\nu),
\end{aligned}$$

so Ψ_{n-} converges uniformly to Ψ_- on $[0,1]$. We deduce with the same reasoning that Ψ_{n+} converges uniformly to Ψ_+ on $[0,1]$. Since $\widetilde{\mathcal{U}}_{n+} = (0,(u_d)_n)$, we deduce from the definition of $(u_d)_n$ that for all $u \in [0,1]$, $\widetilde{\Psi}_{n+}(u) = \Psi_{n+}(u \wedge (u_d)_n) = \Psi_{n+}(u) \wedge \Psi_{n-}(1)$. Let $(a,b,c,d) \in \mathbb{R}^4$. Then $((a-b)^+ - (c-d)^+)((b-a)^+ - (d-c)^+) = -(a-b)^+(d-c)^+ - (c-d)^+(b-a)^+ \leq 0$, so $(a-b)^+ - (c-d)^+$ and $(b-a)^+ - (d-c)^+$ have opposite signs. Therefore, we can apply the inequality $|x| \leq |x+\alpha| \vee |x+\beta|$ valid for $(x,\alpha,\beta) \in \mathbb{R}^3$ such that α and β have opposite signs with $(x,\alpha,\beta) = (a\wedge b - c\wedge d, (a-b)^+ - (c-d)^+, (b-a)^+ - (d-c)^+)$, which yields

$$\begin{aligned}
|a\wedge b - c\wedge d| &\leq |a\wedge b - c\wedge d + (a-b)^+ - (c-d)^+| \vee |a\wedge b - c\wedge d + (b-a)^+ - (d-c)^+|\\
&= |a-c| \vee |b-d|.
\end{aligned} \tag{5.31}$$

Using (5.31) with $(a,b,c,d) = (\Psi_{n+}(u), \Psi_{n-}(1), \Psi_+(u), \Psi_-(1))$, we deduce that

$$\begin{aligned}
|\widetilde{\Psi}_{n+}(u) - \widetilde{\Psi}_+(u)| &= |\Psi_{n+}(u) \wedge \Psi_{n-}(1) - \Psi_+(u) \wedge \Psi_-(1)|\\
&\leq |\Psi_{n+}(u) - \Psi_+(u)| \vee |\Psi_{n-}(1) - \Psi_-(1)|,
\end{aligned}$$

hence $\widetilde{\Psi}_{n+}$ converges uniformly to $\widetilde{\Psi}_+$ on $[0,1]$. If $\nu \leq_{st} \mu$, then we deduce from the Lipschitz continuity of H with respect to its second variable, (5.29) and (5.30) that there exists $K \in \mathbb{R}_+$ such that

$$\begin{aligned}
&\left|\int_{\mathbb{R}\times\mathbb{R}} H(x,y)\, M_n^{ITS}(dx,dy) - \int_{\mathbb{R}\times\mathbb{R}} H(x,y)\, M^{ITS}(dx,dy)\right| \\
&\leq \left|\int_{(0,1)} H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u))\, du - \int_{(0,1)} H(F_\mu^{-1}(u), F_\nu^{-1}(u))\, du\right| + K(\widetilde{\Psi}_{n+}(1) + \Psi_{n-}(1)) \\
&\underset{n\to+\infty}{\longrightarrow} K(\widetilde{\Psi}_+(1) + \Psi_-(1)) = 0.
\end{aligned}$$

We conclude that $M_n^{ITS} \underset{n\to+\infty}{\longrightarrow} M^{ITS}$ for the weak convergence topology as soon as $\nu \leq_{st} \mu$. From now on, we suppose $\nu \not\leq_{st} \mu$. Since $\Psi_{n-}(1) \underset{n\to+\infty}{\longrightarrow} \Psi_-(1) > 0$, $\nu_n \not\leq_{st} \mu_n$ for n large enough, so we can suppose without loss of generality that $\nu_n \not\leq_{st} \mu_n$ for all $n \in \mathbb{N}$. Using Lemma 6.3 below for the first equality, then Proposition 6.2 below for the second equality and the change of variables $u = \widetilde{\Psi}_{n+}(1)v$ with the equality $\widetilde{\Psi}_{n+}(1) = \Psi_{n-}(1)$ for the last equality, we have

$$\begin{aligned}
&\int_{(0,1)} \frac{(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^+(u)}{F_{\nu_n}^{-1}(\varphi_n(u)) - F_{\nu_n}^{-1}(u)} (H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(\varphi_n(u))) - H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u)))\, du \\
&= \int_{(0,1)} \frac{H(F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(u))), F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\widetilde{\Psi}_{n+}(u)))) - H(F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(u))), F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(u))))}{F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\widetilde{\Psi}_{n+}(u))) - F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(u)))}\, d\widetilde{\Psi}_{n+}(u) \\
&= \int_{(0,\widetilde{\Psi}_{n+}(1))} \frac{H(F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(u)), F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(u))) - H(F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(u)), F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(u)))}{F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(u)) - F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(u))}\, du \\
&= \widetilde{\Psi}_{n+}(1) \int_{(0,1)} \frac{H(F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(1)v)), F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\Psi_{n-}(1)v))) - H(F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(1)v)), F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(1)v)))}{(F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\Psi_{n-}(1)v)) - F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(1)v)))}\, dv.
\end{aligned}$$

Since H is Lipschitz continuous with respect to its second variable, then the integrand above is bounded. Moreover, for all $x \in \mathbb{R}$, $|\widetilde{\Psi}_{n+}(F_{\mu_n}(x)) - \widetilde{\Psi}_+(F_\mu(x))| \leq \sup_{[0,1]} |\widetilde{\Psi}_{n+} - \widetilde{\Psi}_+| + |\widetilde{\Psi}_+(F_{\mu_n}(x)) - \widetilde{\Psi}_+(F_\mu(x))|$, so $\widetilde{\Psi}_{n+}(F_{\mu_n}(x))/\widetilde{\Psi}_{n+}(1) \underset{n\to+\infty}{\to} \widetilde{\Psi}_+(F_\mu(x))/\widetilde{\Psi}_+(1)$ for all $x \in \mathbb{R}$ outside the at most countable set of discontinuities of F_μ. This implies that $d(\widetilde{\Psi}_{n+}(F_{\mu_n}(x))/\widetilde{\Psi}_{n+}(1))$ converges to $d(\widetilde{\Psi}_+(F_\mu(x))/\widetilde{\Psi}_+(1))$ for the weak convergence topology. We deduce the pointwise convergence of the left continuous pseudo-inverses du-almost everywhere on $(0,1)$, that is $F_{\mu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(1)u)) \underset{n\to+\infty}{\longrightarrow} F_\mu^{-1}(\widetilde{\Psi}_+^{-1}(\widetilde{\Psi}_+(1)u))$ for du-almost all $u \in (0,1)$. In the same way, $F_{\nu_n}^{-1}(\widetilde{\Psi}_{n+}^{-1}(\widetilde{\Psi}_{n+}(1)u)) \underset{n\to+\infty}{\longrightarrow} F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\widetilde{\Psi}_+(1)u))$ and $F_{\nu_n}^{-1}(\Psi_{n-}^{-1}(\Psi_{n-}(1)u)) \underset{n\to+\infty}{\longrightarrow} F_\nu^{-1}(\Psi_-^{-1}(\Psi_-(1)u))$ for du-almost all $u \in (0,1)$. Therefore, by the dominated convergence theorem,

$$\begin{aligned}
&\int_{(0,1)} \frac{(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^+(u)}{F_{\nu_n}^{-1}(\varphi_n(u)) - F_{\nu_n}^{-1}(u)} (H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(\varphi_n(u))) - H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u)))\, du \\
&\underset{n\to+\infty}{\longrightarrow} \widetilde{\Psi}_+(1) \int_{(0,1)} \frac{H(F_\mu^{-1}(\widetilde{\Psi}_+^{-1}(\widetilde{\Psi}_+(1)v)), F_\nu^{-1}(\Psi_-^{-1}(\Psi_-(1)v))) - H(F_\mu^{-1}(\widetilde{\Psi}_+^{-1}(\widetilde{\Psi}_+(1)v)), F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\widetilde{\Psi}_+(1)v)))}{(F_\nu^{-1}(\Psi_-^{-1}(\Psi_-(1)v)) - F_\nu^{-1}(\widetilde{\Psi}_+^{-1}(\widetilde{\Psi}_+(1)v)))}\, dv \\
&= \int_{(0,1)} \frac{(F_\mu^{-1} - F_\nu^{-1})^+(u)}{F_\nu^{-1}(\varphi(u)) - F_\nu^{-1}(u)} (H(F_\mu^{-1}(u), F_\nu^{-1}(\varphi(u))) - H(F_\mu^{-1}(u), F_\nu^{-1}(u)))\, du.
\end{aligned}$$

We can show in the same way that

$$\begin{aligned}
&\int_{(0,1)} \frac{(F_{\mu_n}^{-1} - F_{\nu_n}^{-1})^-(u)}{F_{\nu_n}^{-1}(u) - F_{\nu_n}^{-1}(\widetilde{\varphi}_n(u))} (H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(\widetilde{\varphi}_n(u))) - H(F_{\mu_n}^{-1}(u), F_{\nu_n}^{-1}(u)))\, du \\
&\underset{n\to+\infty}{\longrightarrow} \int_{(0,1)} \frac{(F_\mu^{-1} - F_\nu^{-1})^-(u)}{F_\nu^{-1}(u) - F_\nu^{-1}(\widetilde{\varphi}(u))} (H(F_\mu^{-1}(u), F_\nu^{-1}(\widetilde{\varphi}(u))) - H(F_\mu^{-1}(u), F_\nu^{-1}(u)))\, du.
\end{aligned}$$

Finally, we showed that

$$\int_{\mathbb{R}\times\mathbb{R}} H(x,y)\, M_n^{ITS}(x,dy) \underset{n\to+\infty}{\longrightarrow} \int_{\mathbb{R}\times\mathbb{R}} H(x,y)\, M^{ITS}(x,dy),$$

for any bounded and continuous function $H : \mathbb{R}^2 \to \mathbb{R}$ which is Lipschitz continuous with respect to its second variable, that is $M_n^{ITS} \underset{n\to+\infty}{\longrightarrow} M^{ITS}$ for the weak convergence topology.

Since the convergence for the Wasserstein distance topology is equivalent to the convergence for the weak convergence topology and the convergence of the first order moments (see for instance [24, Theorem 6.9 Chapter 6]), $\int_{\mathbb{R}} |x|\, \mu_n(dx) \underset{n\to+\infty}{\longrightarrow} \int_{\mathbb{R}} |x|\, \mu(dx)$ and $\int_{\mathbb{R}} |y|\, \nu_n(dy) \underset{n\to+\infty}{\longrightarrow} \int_{\mathbb{R}} |y|\, \nu(dy)$. Therefore, $\mathcal{W}_1(M_n^{ITS}, M^{ITS}) \underset{n\to+\infty}{\longrightarrow} 0$ when $\mathbb{R}^2$ is endowed with the L^1-norm. Since all norms on $\mathbb{R}^2$ are equivalent, $\mathcal{W}_1(M_n^{ITS}, M^{ITS}) \underset{n\to+\infty}{\longrightarrow} 0$ when $\mathbb{R}^2$ is endowed with any norm. □

6 Appendix

We begin with a key result for the construction of the inverse transform martingale coupling.

Lemma 6.1. *Let $f_1, f_2 : (0,1) \to \mathbb{R}$ be two measurable nonnegative and integrable functions and $u_0 \in [0,1]$ be such that $\int_0^{u_0} f_1(u)\, du = \int_0^1 f_2(u)\, du$. Let $\Psi_1 : [0,1] \ni u \mapsto \int_0^u f_1(v)\, dv$, $\Psi_2 : [0,1] \ni u \mapsto \int_0^u f_2(v)\, dv$ and $\Gamma = \Psi_2^{-1} \circ \Psi_1$ where Ψ_2^{-1} denotes the càg pseudo-inverse of Ψ_2. Then Γ is well defined on $[0, u_0]$ and for any measurable and bounded function $h : [0,1] \to \mathbb{R}$,*

$$\int_0^{u_0} h(\Gamma(u)) f_1(u)\, du = \int_0^1 h(v) f_2(v)\, dv.$$

The proof of Lemma 6.1 relies on the next proposition, which is a well known result of integration by continuous and nondecreasing substitution, whose proof can be found for instance in [20, Proposition 4.10 Chapter 0].

Proposition 6.2. *Let $a, b \in \mathbb{R}$ be such that $a < b$. Let $\Psi : [a,b] \to \mathbb{R}$ be a continuous and nondecreasing function. Then for any Borel function $f : [\Psi(a), \Psi(b)] \to \mathbb{R}$,*

$$\int_a^b f(\Psi(s))\, d\Psi(s) = \int_{\Psi(a)}^{\Psi(b)} f(t)\, dt.$$

Proof of Lemma 6.1. Let $h : [0,1] \to \mathbb{R}$ be a measurable and bounded function. Since Ψ_1 is nondecreasing and continuous, using Proposition 6.2, we have

$$\int_0^{u_0} h(\Gamma(u)) f_1(u)\, du = \int_0^{u_0} h(\Psi_2^{-1}(\Psi_1(u)))\, d\Psi_1(u) = \int_0^{\Psi_1(u_0)} h(\Psi_2^{-1}(w))\, dw.$$

Since $\int_0^{u_0} f_1(u)\, du = \int_0^1 f_2(u)\, du$, we have $\Psi_1(u_0) = \Psi_2(1)$, and since Ψ_2 is nondecreasing and continuous, using once again Proposition 6.2, we have

$$\int_0^{\Psi_1(u_0)} h(\Psi_2^{-1}(w))\, dw = \int_0^{\Psi_2(1)} h(\Psi_2^{-1}(w))\, dw = \int_0^1 h(\Psi_2^{-1}(\Psi_2(v)))\, d\Psi_2(v).$$

Since by Lemma 6.3 below, $\Psi_2^{-1}(\Psi_2(v)) = v$, $d\Psi_2(v)$-almost everywhere on $(0,1)$, we conclude that

$$\int_0^1 h(\Psi_2^{-1}(\Psi_2(v)))\, d\Psi_2(v) = \int_0^1 h(v) f_2(v)\, dv.$$ □

We complete this section with standard lemmas with their proofs, so that the present article is self-contained.

Lemma 6.3. *Let $I \subset \mathbb{R}$ be an interval, $F : I \to \mathbb{R}$ be a bounded and nondecreasing càdlàg function, $F(I)$ be the image of I by F and F^{-1} be the left continuous pseudo-inverse of F, that is*

$$F^{-1} : u \in F(I) \mapsto \inf\{r \in I \mid F(r) \geq u\}$$

Then for all $(x, u) \in I \times F(I)$, $F(x) \geq u \iff x \geq F^{-1}(u)$. Moreover, $F^{-1}(F(x)) = x$, $dF(x)$-almost everywhere on I.

Proof. Let $(x, u) \in I \times F(I)$. If $F(x) \geq u$, then by definition of the infimum, $x \geq F^{-1}(u)$. Conversely, if $x \geq F^{-1}(u)$, then let $(r_n)_{n\in\mathbb{N}} \in I^{\mathbb{N}}$ be a decreasing sequence converging to $F^{-1}(u)$. For all $n \in \mathbb{N}$, $F(r_n) \geq u$. By right-continuity of F, we get $F(F^{-1}(u)) \geq u$ for $n \to +\infty$. By monotonicity of F, we have $F(x) \geq F(F^{-1}(u)) \geq u$.

Let us now prove the second statement. Let $a = \inf F(I)$ and $b = \sup F(I)$. If $a = b$, then $dF(x)$ is the trivial measure on I so the statement is straightforward. Else, let $G : I \to [0, 1]$ be defined for all $x \in I$ by $G(x) = (F(x) - a)/(b - a)$ and let G^{-1} be its left-continuous pseudo-inverse. It is well known that for all $u \in (0, 1)$, $G^{-1}(G(G^{-1}(u))) = G^{-1}(u)$. So $G^{-1}(G(G^{-1}(U))) = G^{-1}(U)$, where U is a random variable uniformly distributed on $[0, 1]$. By the inverse transform sampling, it implies that $G^{-1}(G(x)) = x$, $dG(x)$-almost everywhere on I. For all $u \in F(I)$, we have $F^{-1}(u) = G^{-1}((u - a)/(b - a))$, hence $F^{-1}(F(x)) = G^{-1}(G(x)) = x$, $dG(x)$-almost everywhere on I. Since $dG(x) = \frac{1}{b-a} dF(x)$, $dG(x)$ and $dF(x)$ are equivalent, so $F^{-1}(F(x)) = x$, $dF(x)$-almost everywhere on I. □

Lemma 6.4. *Let $\mu \in \mathcal{P}(\mathbb{R})$. Then $F_\mu(x) > 0$ and $F_\mu(x_-) < 1$, $\mu(dx)$-almost everywhere on $\mathbb{R}$.*

Proof. If $\{x \in \mathbb{R} \mid F_\mu(x) = 0\}$ is nonempty, then it is an interval of the form $(-\infty, a]$ or $(-\infty, a)$, depending on whether $F_\mu(a) = 0$ or not. If $F_\mu(a) = 0$, then $\mu(\{x \in \mathbb{R} \mid F_\mu(x) = 0\}) = \mu((-\infty, a]) = F_\mu(a) = 0$. Else, since for all $x < a$, $F_\mu(x) = 0$, then $\mu(\{x \in \mathbb{R} \mid F_\mu(x) = 0\}) = \mu((-\infty, a)) = F_\mu(a_-) = 0$.

If $\{x \in \mathbb{R} \mid F_\mu(x_-) = 1\}$ is nonempty, then it is an interval of the form $[a, +\infty)$ or $(a, +\infty)$, depending whether $F_\mu(a_-) = 1$ or not. If $F_\mu(a_-) = 1$, then $\mu(\{x \in \mathbb{R} \mid F_\mu(x_-) = 1\}) = \mu([a, +\infty)) = 1 - F_\mu(a_-) = 0$. Else, since for all $x > a$, $F_\mu(x_-) = 1$, then $\mu(\{x \in \mathbb{R} \mid F_\mu(x_-) = 1\}) = \mu((a, +\infty)) = 1 - F_\mu(a) = 1 - \lim_{x\to a, x>a} F_\mu(x_-) = 0$, by right continuity of F_μ. □

Lemma 6.5. *Let $\mu \in \mathcal{P}_1(\mathbb{R})$. Then μ is symmetric with mean $\alpha \in \mathbb{R}$, that is $(x-\alpha)_\sharp\mu(dx) = (\alpha - x)_\sharp\mu(dx)$ where $\sharp$ denotes the pushforward operation, iff*

$$F_\mu^{-1}(u_+) = 2\alpha - F_\mu^{-1}(1 - u),$$

for all $u \in (0, 1)$. In that case, $F_\mu^{-1}(u) = 2\alpha - F_\mu^{-1}(1 - u)$ for $u \in (0, 1)$ up to the at most countable set of discontinuities of F_μ^{-1}.

Proof. Let U be a random variable uniformly distributed on $[0, 1]$. Then, by the inverse transform sampling, $F_\mu^{-1}(1 - U) \sim \mu$, so $2\alpha - F_\mu^{-1}(1 - U) \sim \mu$ since μ is symmetric with mean α. Since $u \mapsto 2\alpha - F_\mu^{-1}(1 - u)$ is nondecreasing, then one can show that $2\alpha - F_\mu^{-1}(1 - u) = F_\mu^{-1}(u)$, du-almost everywhere on $(0, 1)$. Indeed, as shown in [2, Lemma

A.3], for all $u, q \in (0,1)$ such that $q < u$, if $F_\mu^{-1}(u) < 2\alpha - F_\mu^{-1}(1-q)$, then

$$\begin{aligned}\mathbb{P}(2\alpha - F_\mu^{-1}(1-U) \le F_\mu^{-1}(u)) &\le \mathbb{P}(2\alpha - F_\mu^{-1}(1-U) < 2\alpha - F_\mu^{-1}(1-q)) \le q \\ &< u \le \mathbb{P}(F_\mu^{-1}(U) \le F_\mu^{-1}(u)) \\ &= \mathbb{P}(2\alpha - F_\mu^{-1}(1-U) \le F_\mu^{-1}(u)),\end{aligned}$$

which is contradictory, so $F_\mu^{-1}(u) \ge \sup_{q\in(0,u)}(2\alpha - F_\mu^{-1}(1-q))$. By symmetry, $2\alpha - F_\mu^{-1}(1-u) \ge \sup_{q\in(0,u)} F_\mu^{-1}(q) = F_\mu^{-1}(u)$ by left-continuity and monotonicity of F_μ^{-1}. Since F_μ^{-1} has an at most countable set of discontinuities, then for du-almost all $u \in (0,1)$, $2\alpha - F_\mu^{-1}(1-u) = \sup_{q\in(0,u)}(2\alpha - F_\mu^{-1}(1-q)) \le F_\mu^{-1}(u) \le 2\alpha - F_\mu^{-1}(1-u)$. Therefore, $2\alpha - F_\mu^{-1}(1-u) = F_\mu^{-1}(u_+)$, du-almost everywhere on $(0,1)$ and even everywhere on $(0,1)$ since both sides are right-continuous. □

Lemma 6.6. *Let $\mu \in \mathcal{P}(\mathbb{R})$, let $X : \Omega \to \mathbb{R}$ be a random variable with distribution μ and let V be a random variable independent from X and uniformly distributed on $(0,1)$. Let $W : \Omega \to \mathbb{R}$ be the random variable defined by*

$$W = F_\mu(X_-) + V(F_\mu(X) - F_\mu(X_-)).$$

Then W is uniformly distributed on $(0,1)$, and $F_\mu^{-1}(W) = X$ almost surely.

Proof. Let $h : \mathbb{R} \to \mathbb{R}$ be a measurable and bounded function. Then

$$\begin{aligned}\mathbb{E}[h(W)] &= \mathbb{E}[h(F_\mu(X_-) + V(F_\mu(X) - F_\mu(X_-))] \\ &= \int_0^1 \int_{\mathbb{R}} h(F_\mu(x_-) + v(F_\mu(x) - F_\mu(x_-)))\,\mu(dx)\,dv \\ &= \int_0^1 \int_{\mathbb{R}} \mathbb{1}_{\{\mu(\{x\})=0\}} h(F_\mu(x_-) + v(F_\mu(x) - F_\mu(x_-)))\,\mu(dx)\,dv \\ &\quad + \int_0^1 \int_{\mathbb{R}} \mathbb{1}_{\{\mu(\{x\})>0\}} h(F_\mu(x_-) + v(F_\mu(x) - F_\mu(x_-)))\,\mu(dx)\,dv \\ &= \int_{\mathbb{R}} \mathbb{1}_{\{\mu(\{x\})=0\}} h(F_\mu(x))\,\mu(dx) + \sum_{x\in\mathbb{R}:\mu(\{x\})>0} \int_{F_\mu(x_-)}^{F_\mu(x)} h(v)\,dv \\ &= \int_0^1 \mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})=0\}} h(F_\mu(F_\mu^{-1}(u)))\,du + \sum_{x\in\mathbb{R}:\mu(\{x\})>0} \int_{F_\mu(x_-)}^{F_\mu(x)} h(v)\,dv \\ &= \int_0^1 \mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})=0\}} h(u)\,du + \sum_{x\in\mathbb{R}:\mu(\{x\})>0} \int_{F_\mu(x_-)}^{F_\mu(x)} h(v)\,dv,\end{aligned}$$

where we used for the last but one equality the inverse transform sampling, and for the last equality the fact that $F_\mu(F_\mu^{-1}(u)) = u$ if F_μ is continuous at $F_\mu^{-1}(u)$. One can easily see that for all $x \in \mathbb{R}$ and $u \in (0,1)$,

$$F_\mu(x_-) < u \le F_\mu(x) \implies x = F_\mu^{-1}(u) \implies F_\mu(x_-) \le u \le F_\mu(x),$$

which implies

$$\begin{aligned}\bigcup_{x\in\mathbb{R}:\mu(\{x\})>0} (F_\mu(x_-), F_\mu(x)] &\subset \{u \in (0,1) \mid \mu(\{F_\mu^{-1}(u)\}) > 0\} \\ &\subset \bigcup_{x\in\mathbb{R}:\mu(\{x\})>0} [F_\mu(x_-), F_\mu(x)],\end{aligned}$$

so

$$\sum_{x\in\mathbb{R}:\mu(\{x\})>0}\int_{F_\mu(x_-)}^{F_\mu(x)} h(v)\,dv = \int_0^1 \mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})>0\}} h(u)\,du.$$

Therefore,

$$\mathbb{E}[h(W)] = \int_0^1 \mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})=0\}} h(u)\,du + \int_0^1 \mathbb{1}_{\{\mu(\{F_\mu^{-1}(u)\})>0\}} h(u)\,du = \int_0^1 h(u)\,du,$$

so W is uniformly distributed on $(0,1)$. Moreover, on $\{F_\mu(X_-) = F_\mu(X)\}$, $W = F_\mu(X)$ and by Lemma 6.3, $F_\mu^{-1}(W) = X$ almost surely. Since $V > 0$ a.s., on $\{F_\mu(X_-) < F_\mu(X)\}$, a.s., $F_\mu(X_-) < W \leq F_\mu(X)$ so $F_\mu^{-1}(W) = X$. □

References

[1] Alfonsi, A., Corbetta, J. and Jourdain, B.: Sampling of one-dimensional probability measures in the convex order and computation of robust option price bounds. *International Journal of Theoretical and Applied Finance* **22**, (2019). MR-3951948

[2] Alfonsi, A., Corbetta, J. and Jourdain, B.: Sampling of probability measures in the convex order by Wasserstein projection. *Annales de l'Institut Henri Poincaré B, Probability and Statistics* **56**, (2020), 1706–1729. MR-4116705

[3] Beiglböck, M., Henry-Labordère, P. and Penkner, F.: Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics* **17**, (2013), 477–501. MR-3066985

[4] Beiglböck, M. and Juillet, N.: On a problem of optimal transport under marginal martingale constraints. *Annals of Probability* **44**, (2016), 42–106. MR-3456332

[5] Beiglböck, M., Lim, T. and Obłój, J.: Dual attainment for the martingale transport problem. *Bernoulli* **25**, (2019), 1640–1658. MR-3961225

[6] Beiglböck, M., Nutz, M. and Touzi, N.: Complete Duality for Martingale Optimal Transport on the Line. *Annals of Probability* **45**, (2017), 3038–3074. MR-3706738

[7] De March, H.: Entropic approximation for multi-dimensional martingale optimal transport, arXiv:1812.11104

[8] De March, H.: Local structure of multi-dimensional martingale optimal transport, arXiv:1805.09469

[9] De March, H.: Quasi-sure duality for multi-dimensional martingale optimal transport, arXiv:1805.01757

[10] De March, H. and Touzi, N.: Irreducible convex paving for decomposition of multi-dimensional martingale transport plans. *Annals of Probability* **47**, (2019), 1726–1774. MR-3945758

[11] Föllmer, H. and Schied, A.: Stochastic finance. An introduction in discrete time. Third revised and extended edition. *Walter de Gruyter & Co., Berlin*, 2011. xii+544 pp. MR-2779313

[12] Galichon, A., Henry-Labordère, P. and Touzi., N.: A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Annals of Applied Probability* **24**, (2014), 312–336. MR-3161649

[13] Guo, G. and Obłój, J.: Computational methods for martingale optimal transport problems. *Annals of Applied Probability* **29**, (2019), 3311–3347. MR-4047982

[14] Henry-Labordère, P., Tan, X. and Touzi, N.: An explicit version of the one-dimensional brenier's theorem with full marginals constraint. *Stochastic Processes and their Applications* **126**, (2016), 2800–2834. MR-3522302

[15] Henry-Labordère, P. and Touzi, N.: An explicit martingale version of the one-dimensional Brenier theorem. *Finance and Stochastics* **20**, (2016), 635–668. MR-3519164

[16] Hobson, D. and Klimmek, M.: Robust price bounds for the forward starting straddle. *Finance and Stochastics* **19**, (2015), 189–214. MR-3292129

[17] Hobson, D. and Neuberger, A.: Robust Bounds for Forward Start Options. *Mathematical Finance* **22**, (2012), 31–56. MR-2881879

[18] Pilipenko, A.: An introduction to stochastic differential equations with reflection. *Lectures in pure and applied mathematics. Universitätsverlag Potsdam, Potsdam*, 2014. ix+75 pp.

[19] Rachev, S. T. and Rüschendorf, L.: Mass Transportation Problems. Volume 1: Theory. *Probability and Its Applications. Springer-Verlag, New York*, 1998. xxvi+508 pp. MR-1619170

[20] Revuz, D. and Yor, M.: Continuous Martingales and Brownian Motion. Third edition. *Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg*, 1999. xiv+602 pp. MR-1725357

[21] Shaked, M. and Shanthikumar, G.: Stochastic Orders. *Springer Series in Statistics. Springer-Verlag, New York*, 2007. xvi+473 pp. MR-2265633

[22] Strassen, V.: The Existence of Probability Measures with Given Marginals. *Annals of Mathematical Statistics* **36**, (1965), 423–439. MR-0177430

[23] Villani, C.: Topics in Optimal Transportation. *American Mathematical Society, Providence, RI*, 2003. xvi+370 pp. MR-1964483

[24] Villani, C.: Optimal Transport: Old and New. *Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg*, 2009. xxii+973 pp. MR-2459454

Acknowledgments. We thank Jean-François Delmas (CERMICS) for numerous and fruitful discussions.