

# Supermartingale Brenier's Theorem with full-marginals constraint\*

Erhan Bayraktar <sup>†</sup>

Shuoqing Deng <sup>‡</sup>

Dominykas Norgilas <sup>§</sup>

## Abstract

We explicitly construct the supermartingale version of the Fréchet-Hoeffding coupling in the setting with infinitely many marginal constraints. This extends the results of Henry-Labordère et al. [41] obtained in the martingale setting. Our construction is based on the Markovian iteration of one-period optimal supermartingale couplings. In the limit, as the number of iterations goes to infinity, we obtain a pure jump process that belongs to a family of local Lévy models introduced by Carr et al. [22]. We show that the constructed processes solve the continuous-time supermartingale optimal transport problem for a particular family of (path-dependent) cost functions. The explicit computations are provided in the following three cases: the uniform case, the Bachelier model and the Geometric Brownian Motion case.

**Key words.** Supermartingale optimal transport, Brenier's Theorem, PCOCD.

**MSC (2010).** Primary: 60G40, 60G05; Secondary: 49M29.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Supermartingale optimal transport problem</b>	<b>4</b>
2.1	Discrete-time supermartingale optimal transport . . . . .	4
2.2	Continuous-time supermartingale optimal transport . . . . .	5
<b>3</b>	<b>Main results</b>	<b>7</b>
3.1	Convergence of the sequence $(\mathbb{P}^n)_{n \geq 1}$ . . . . .	7
3.2	Characterization of the limiting process . . . . .	12
3.3	Optimality of the local Lévy process . . . . .	17
<b>4</b>	<b>Examples</b>	<b>20</b>
4.1	Uniform distribution with bounded support . . . . .	20
4.2	Bachelier dynamics with negative drift . . . . .	25
4.3	Geometric Brownian motion with decreasing average . . . . .	29
<b>A</b>	<b>Brenier's Theorem</b>	<b>32</b>
A.1	Brenier's Theorem in Optimal Transport (OT) . . . . .	32
A.2	Brenier's Theorem in Martingale Optimal Transport (MOT) . . . . .	33

---

\*The authors are grateful to Xiaolu Tan for fruitful discussions.

<sup>†</sup>Department of Mathematics, University of Michigan, Ann Arbor, Email: erhan@umich.edu. E.B. is partially supported by the National Science Foundation under grant DMS-2106556 and by the Susan M. Smith chair.

<sup>‡</sup>Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. Email: masdeng@ust.hk. S.D. is partially supported by the Start-up Grant of HKUST.

<sup>§</sup>Department of Mathematics, University of Michigan, Ann Arbor. Email: dnorgila@umich.edu.

<b>B One-period Supermartingale Optimal Transport</b>	<b>34</b>
B.1 Solution to the primal SMOT problem and related properties . . . . .	34
B.2 Optimal dual strategy for decreasing SMOT . . . . .	37

# 1 Introduction

The classical optimal transport (OT) problem, first introduced by Monge, and then relaxed by Kantorovich, has the following formulation: given two probability measures  $\mu_0, \mu_1$  on  $\mathbb{R}$  and a cost (or payoff) function  $c : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , the goal is to minimize (or maximize) the value  $\mathbb{E}^{\mathbb{P}}[c(X_0, X_1)]$  among all probability measures  $\mathbb{P}$  such that  $\mathbb{P} \circ X_0^{-1} = \mu_0$  and  $\mathbb{P} \circ X_1^{-1} = \mu_1$ . Under the Spence-Mirrlees condition  $c_{xy} > 0$ , the maximizing transport plan is characterized by the Brenier's Theorem (see Brenier [18] and Rachev and Rüschendorf [61]) and corresponds to the Fréchet-Hoeffding (or quantile) coupling  $\pi^{FH}$ .

Motivated by the applications in financial mathematics, the Martingale Optimal Transport (MOT) problem was introduced by Beiglböck et al. [8] (in the discrete-time setting) and Galichon et al. [36] (in the continuous-time setting), and has been widely studied since then. Given two probability measures  $\mu_0, \mu_1$  which are increasing in convex order, the problem consists in minimizing (or maximizing)  $\mathbb{E}^{\mathbb{P}}[c(X_0, X_1)]$  among all martingale measures  $\mathbb{P}$  such that  $\mathbb{P} \circ X_0^{-1} = \mu_0$  and  $\mathbb{P} \circ X_1^{-1} = \mu_1$ . Under the so-called martingale Spence-Mirrlees condition  $c_{xyy} > 0$ , the maximizer corresponds to the so-called left-curtain martingale coupling  $\pi^{lc}$ , introduced by Beiglböck and Juillet [12].  $\pi^{lc}$  was explicitly constructed by Henry-Labordère and Touzi [42] using ODE arguments (and under some technical assumptions). The general construction was later obtained by Hobson and Norgilas [48] using geometric arguments and properties of the potential functions of marginals.

The financial motivation of MOT comes from the robust sub/super-replication of financial derivatives in a market where the underlying asset and the corresponding vanilla options with certain maturities are available for trading. From the well-known Breeden-Litzenberger formula (see [17]), the marginal distributions of the underlying asset can be then obtained using the prices of vanilla options. Each sensible pricing model should be calibrated to the given market data, and thus should produce the same marginal distributions at fixed times. The search of a model that produces the highest no-arbitrage price of an exotic claim, among all calibrated models, then naturally corresponds to the MOT problem. For more recent developments of MOT problems, see for example Acciaio et al. [1], Backhoff-Veraguas et al. [2, 3], Beiglböck et al. [6, 16, 7, 9, 10, 11], Beiglböck and Juillet [12, 13], Brücknerhoff et al. [20], Campi et al. [21], De Marco and Henry-Labordère [30], Dolinsky and Soner [31], Fahim nad Huang [34], Gaoyue et al. [39], Hobson and Neuberger [46], Hobson and Klimmek [45], Hobson and Norgilas [47], Nutz et al. [58], Wiesel [63].

The problem of finding the models that give robust no-arbitrage bounds for the prices of exotic derivatives was initially studied in the seminal work of Hobson [49] by means of the Skorokhod Embedding Problem (SEP). SEP consists in finding a stopping time  $\tau$  of a Brownian motion  $B$ , such that  $B_\tau$  has a prescribed law. This approach generated developments in many (probabilistic and financially motivated) directions, see, for example, Brown et. al [19], Madan and Yor [56], Cox et al. [23, 27], Cox and Oblój [25, 26], Davis et al. [29]. See also the survey papers on SEP by Oblój [59] and Hobson [44].

More recently, the Supermartingale Optimal Transport(SOT) problem was introduced by Nutz and Stebegg [57]. Given two probability measures  $\mu_0, \mu_1$  which are increasing in convex-decreasing order, the problem consists in minimizing (or maximizing)  $\mathbb{E}^{\mathbb{P}}[c(X_0, X_1)]$  among all supermartingale measures  $\mathbb{P}$  such that  $\mathbb{P} \circ X_0^{-1} = \mu_0$  and  $\mathbb{P} \circ X_1^{-1} = \mu_1$ . In particular, the authors of [57] introduced two canonical supermartingale couplings, namely the increasing supermartingale coupling  $\pi^I$  and the decreasing supermartingale coupling  $\pi^D$ . These transport plans are canonical in the sense that they can be equivalently characterised by any, and then all

of the following properties: the optimality (in terms of minimization or maximization problem) for a large class of cost functions, the monotonicity of the support, and the order-theoretic minimality. (The optimal couplings in the OT and MOT settings have similar characterizations.)

In this paper, we shall focus on the continuous-time case, which corresponds to a suitable limit of the multi-marginal SOT problem. Given a continuous family of marginal distributions  $(\mu_t)_{t \in [0,1]}$ , which is non-decreasing in convex-decreasing order, we call a stochastic process that is calibrated to all marginals  $(\mu_t)_{t \in [0,1]}$  a *PCOCD* (“Processus Croissant pour l’Ordre Convexe Décroissant” in French). Notice that in the martingale case such process is baptized as a peacock (*PCOC* for short) and is extensively studied in the book of Hirsch et al. [43].

We will closely follow Henri-Labordère et al. [41]: in the martingale setting, by considering a Markovian iteration of the one-period left-curtain martingale couplings  $\pi^{lc}$ , the authors obtained a pure jump process in the spirit of the local Lévy models introduced by Carr et al. [22] and showed that it solves a particular continuous-time MOT problem. Our main goal is to extend the results of Henri-Labordère et al. [41] to the supermartingale setting.

In the supermartingale case, the construction is different in the following aspects. First, in the martingale case, the left-curtain and the right-curtain transport plans are symmetric, see Henry-Labordère et al. [41, Remark 3.12]. More precisely, the limiting process, associated to the left-curtain coupling, is driven by downward jumps and upward drift, while in the right-curtain case it is the upward jumps and downward drift that drive the limiting process. On the other hand, in the supermartingale case, the increasing supermartingale transport plan and the decreasing supermartingale transport plan demonstrate different behaviour see Bayraktar et al. [4], [5]. In the increasing case, the limiting process is a martingale in the interior of the support of marginals, and possesses strict supermartingale characteristics when it escapes to the upper boundary of the support (it jumps from the upper all the way down to the lower boundary). In the case of the decreasing supermartingale coupling, there are martingale and supermartingale regions in the interior of the support. Once the boundary curve (that separates these two regions) is hit, the process becomes deterministic and follows a decreasing (w.r.t. time) curve.

Second, in case of the the decreasing supermartingale coupling, there might exist multiple *phase transition curves* for the limiting process, that divide the martingale and strict supermartingale regions. To obtain explicit solutions, in this article we will impose certain conditions that will guarantee the uniqueness of the phase transition curve. On the other hand, if the increasing supermartingale coupling is used for construction, then the regime switching boundary is unique, and is always given by the upper boundary of the support of the marginal distributions.

In financial terms, the dual SOT problem corresponds to the super-replication of an exotic payoff using semi-static strategies with no short-selling constraints. In order to prove the optimality of the limiting process, we will modify the arguments of Henry-Labordère et al. [42] and explicitly construct a candidate optimal dual strategy, for a large (but very particular) class of cost functions. The choice of dual variables is closely related to the martingale and strict supermartingale regions of a candidate optimal transport plan. In the case of the decreasing supermartingale coupling, the dual methods are related to those of the dual MOT problem when the right-curtain martingale coupling is optimal, and also the classical OT duality when the quantile coupling is an optimizer. The strategies must be modified accordingly, if the construction of the limiting process is based on the increasing supermartingale coupling.

The rest of the paper is organized as follows. In Section 2, we recall the formulation of the discrete-time supermartingale transport problem and introduce the continuous-time SOT problem. In Section 3, we provide the convergence theorem which gives the characterization of the limiting process, and the optimality theorem which shows that the limiting process solves a continuous-time SOT problem for a class of costs functions. In Section 4, we explicitly obtain the limiting process when the marginals correspond to uniform measures on bounded support,

Normal distributions with decreasing means (Brownian motion with drift) and log-Normal distributions (geometric Brownian motion). In Appendix A and B we collect some important results regarding one-period OT, MOT and SOT problems.

## 2 Supermartingale optimal transport problem

### 2.1 Discrete-time supermartingale optimal transport

Let  $\mathcal{M}$  (resp.  $\mathcal{P}$ ) be the set of finite (resp. probability) Borel measures on  $\mathbb{R}$  with finite first moments. The support of  $\eta \in \mathcal{M}$ , denoted by  $\text{supp}(\eta)$ , is the smallest closed set  $I \subseteq \mathbb{R}$  with  $\eta(I) = \eta(\mathbb{R})$ . We set  $\ell_\eta := \inf\{k \in \text{supp}(\eta)\} \in [-\infty, \infty)$  and  $r_\eta := \sup\{k \in \text{supp}(\eta)\} \in (-\infty, \infty]$ . The mean of  $\eta \in \mathcal{M}$  is denoted by  $\bar{\eta}$ , so that the barycenter of  $\eta$  is given by  $\bar{\eta}/\eta(\mathbb{R})$ .

Throughout the paper, for any random variable  $\xi$ , the expectation of  $\xi$  is defined as  $\mathbb{E}[\xi] := \mathbb{E}[\xi^+] - \mathbb{E}[\xi^-]$  with the convention  $\infty - \infty = -\infty$ .

The canonical process on  $\mathbb{R}^2$  is denoted by  $(X_0, X_1)$ , so that, for  $i = 0, 1$ ,  $X_i(x_0, x_1) = x_i$  for all  $(x_0, x_1) \in \mathbb{R}^2$ . For  $\mu_0, \mu_1 \in \mathcal{P}$ , let  $\mathcal{P}(\mu_0, \mu_1)$  be the set of (Borel) probability measures on  $\mathbb{R}^2$  with first and second marginals  $\mu_0$  and  $\mu_1$ , respectively. (Each  $\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)$  is often called a transport plan or coupling of  $\mu_0$  and  $\mu_1$ .) Then  $X_0 \sim^\mathbb{P} \mu_0$ ,  $X_1 \sim^\mathbb{P} \mu_1$  for all  $\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)$ . We also introduce the set of supermartingale couplings of  $\mu_0$  and  $\mu_1$ :

$$\mathcal{S}_2(\mu_0, \mu_1) := \{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1) : \mathbb{E}^\mathbb{P}[X_1|X_0] \leq X_0 \text{ } \mathbb{P}\text{-a.s.}\}.$$

By the classical result of Strassen [62],  $\mathcal{S}_2(\mu_0, \mu_1)$  is non-empty if and only if  $\mu_0 \leq_{cd} \mu_1$ , i.e.,  $\mu_0$  is smaller than  $\mu_1$  in convex-decreasing order. Recall that  $\mu_0 \leq_{cd} \mu_1$  if  $\mu_0(\phi) \leq \mu_1(\phi)$  for all convex and non-increasing  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , where for a measurable  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $\eta \in \mathcal{M}$  we write  $\eta(f) = \int_{\mathbb{R}} f d\eta$ . Note that  $\bar{\mu}_0 \leq \bar{\mu}_1$  whenever  $\mu_0 \leq_{cd} \mu_1$ . On the other hand, if  $\mu_0 \leq_{cd} \mu_1$  and  $\bar{\mu}_0 = \bar{\mu}_1$ , then  $\mathcal{S}_2(\mu_0, \mu_1)$  reduces to the set of martingale couplings of  $\mu_0$  and  $\mu_1$ . Indeed, every supermartingale with constant mean is a martingale.

For a (Borel) measurable reward (or cost) function  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ , a two-marginal supermartingale transport problem is defined by

$$\mathbf{P}_2(\mu_0, \mu_1) := \sup_{\mathbb{P} \in \mathcal{S}_2(\mu_0, \mu_1)} \mathbb{E}^\mathbb{P}[c(X_0, X_1)]. \quad (2.1)$$

Suppose that there exists  $a_0, a_1 : \mathbb{R} \rightarrow \mathbb{R}$  that are  $\mu_0$  and  $\mu_1$ -integrable, respectively, and such that  $|c(x, y)| \leq a_0(x) + a_1(y)$ ,  $x, y \in \mathbb{R}$ . Then in the case  $c$  satisfies the following *supermartingale Spence-Mirrlees* condition

$$c(x', \cdot) - c(x, \cdot) \text{ is decreasing and convex for all } x < x',$$

Nutz and Stebegg [57] proved that the supremum in (2.1) is attained by the so-called *increasing* supermartingale coupling  $\pi^I$ . Note that, in the case  $c$  is smooth, supermartingale Spence-Mirrlees condition can be equivalently stated in terms of cross-derivatives:  $c_{xy} \leq 0$  and  $c_{xyy} \geq 0$ . On the other hand, if  $-c$  is supermartingale Spence-Mirrlees, then the optimizer is given by the *decreasing* supermartingale coupling  $\pi^D$ . The main focus of this paper is on  $\pi^D$  (see Appendix B.1).

We now describe the dual problem that is associated to (2.1). We write  $(\phi, \psi, h) \in \mathcal{D}_2(\mu_0, \mu_1)$  if  $\phi, \psi, h : \mathbb{R} \rightarrow \mathbb{R}$  are (Borel) measurable functions such that  $(\int_{\mathbb{R}} (\phi \vee 0) d\mu_0) \vee (\int_{\mathbb{R}} (\psi \vee 0) d\mu_1) < \infty$ ,  $h \geq 0$  and

$$c(x, y) \leq \phi(x) + \psi(y) + h(x)(y - x), \quad x, y \in \mathbb{R}.$$

Then the weak duality holds

$$\mathbf{P}_2(\mu_0, \mu_1) \leq \mathbf{D}_2(\mu_0, \mu_1) := \inf_{(\phi, \psi, h) \in \mathcal{D}_2(\mu_0, \mu_1)} \{\mu_0(\phi) + \mu_1(\psi)\}. \quad (2.2)$$

Nutz and Stebegg [57] showed that the strong duality (i.e., equality in (2.2)) holds under some additional assumptions (for example, if  $\mu_0 \leq_{cd} \mu_1$  are *irreducible*). In general (and without the irreducibility condition), the strong duality and the existence of dual optimizers can still be obtained, provided one agrees to switch from a point-wise to a quasi-sure formulation of the dual problem. An explicit construction of optimal dual variables is (under some further assumptions) provided in Appendix B.2.

The problem (2.1) can be easily extended to a  $T$ -period (or  $(T + 1)$ -marginal) setting if we restrict to cost functions  $c : \mathbb{R}^{T+1} \rightarrow \mathbb{R}$  of the form  $c(x_0, \dots, x_T) = \sum_{i=1}^T c_i(x_{i-1}, x_i)$ ,  $(x_0, \dots, x_T) \in \mathbb{R}^{T+1}$ , for a collection of one-period (Borel) cost functions  $c_i : \mathbb{R} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, T$ . More precisely, we are given a sequence of probability measures  $\mu_0, \dots, \mu_T \in \mathcal{P}$  satisfying  $\mu_0 \leq_{cd} \dots \leq_{cd} \mu_T$ , and the goal is to maximize

$$\mathbb{E}^{\mathbb{P}}[c(X_0, \dots, X_T)] = \sum_{i=1}^T \mathbb{E}^{\mathbb{P}}[c_i(X_{i-1}, X_i)] \quad (2.3)$$

over all (Borel) probability measures  $\mathbb{P}$  on  $\mathbb{R}^{T+1}$  such that  $X_i \sim^{\mathbb{P}} \mu_i$ ,  $i = 0, \dots, T$ , and satisfying  $\mathbb{E}^{\mathbb{P}}[X_i | X_0, \dots, X_{i-1}] \leq X_{i-1}$   $\mathbb{P}$ -a.s., for all  $i = 1, \dots, T$ . (For the martingale version with general cost functions  $c : \mathbb{R}^{T+1} \rightarrow \mathbb{R}$ , see Nutz et al. [58], while a corresponding continuous-time extension is given by Juillet et al. [20].) Then, if each  $c_i$  (or  $-c_i$ ) is supermartingale Spence-Mirrlees, the optimal supermartingale coupling (with fixed  $T + 1$  marginals) can be obtained by a Markovian iteration of the increasing (or decreasing) one-period supermartingale transport plans.

## 2.2 Continuous-time supermartingale optimal transport

In this section we introduce a continuous-time supermartingale optimal transport problem under full marginals constraint, as the limit of the multi-period supermartingale optimal transport introduced at the end of Section 2.1.

Let  $\Omega := \mathbb{D}([0, 1], \mathbb{R})$  be the canonical space of all càdlàg paths on  $[0, 1]$ ,  $X$  the canonical process and  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq 1}$  the canonical filtration generated by  $X$ , i.e.  $\mathcal{F}_t := \sigma(\{X_s, 0 \leq s \leq t\})$ . We denote by  $\mathcal{S}_{\infty}$  the collection of all supermartingale measures on  $\Omega$ , i.e., under each  $\mathbb{P} \in \mathcal{S}_{\infty}$  the canonical process  $X$  is a supermartingale. We equip  $\mathcal{S}_{\infty}$  with the weak convergence topology throughout the paper. By the classical results (see, for example, Karandikar [54]), there is a non-decreasing process  $([X]_t)_{t \in [0, 1]}$ , defined on  $\Omega$ , and which coincides with the  $\mathbb{P}$ -quadratic variation of  $X$ ,  $\mathbb{P}$ -a.s. for every supermartingale measure  $\mathbb{P} \in \mathcal{S}_{\infty}$ . Denote by  $[X]_t^c$  the continuous part of  $[X]_t$ .

Given a family of (integrable) probability measures  $\mu = (\mu_t)_{0 \leq t \leq 1}$ , we denote by  $\mathcal{S}_{\infty}(\mu) \subset \mathcal{S}_{\infty}$  the collection of all supermartingale measures on  $\Omega$  such that  $X_t \sim^{\mathbb{P}} \mu_t$  for all  $t \in [0, 1]$ . In particular, from Ewald and Yor [33], we know that  $\mathcal{S}_{\infty}(\mu)$  is non-empty if and only if the family  $(\mu_t)_{0 \leq t \leq 1}$  is non-decreasing in convex-decreasing order and  $t \mapsto \mu_t$  is right-continuous.

For all  $t \in [0, 1]$ , we denote by  $-\infty \leq \ell_t := \ell_{\mu_t} \leq r_{\mu_t} \leq \infty$  the left and right extreme boundaries of  $\text{supp}(\mu_t)$ . Suppose that  $\mu_s \leq_{cd} \mu_t$  for all  $0 \leq s \leq t \leq 1$ . Then since  $x \mapsto p_k(x) := (k - x)^+$  is convex and non-increasing for all  $k \in \mathbb{R}$ , we have that  $\mu_s(p_k) \leq \mu_t(p_k)$  for  $k \in \mathbb{R}$ . In particular,  $\mu_s((-\infty, k)) > 0$  implies that  $\mu_t((-\infty, k)) > 0$ , from which it follows that  $t \mapsto \ell_t$  is non-increasing. Note that, in general,  $t \mapsto r_t$  may fail to be non-decreasing. In addition, if  $\mu_t$  admits a density function for all  $t \in [0, 1]$  and  $t \mapsto \mu_t$  is continuous w.r.t. the weak convergence topology, then  $t \mapsto \ell_t$  and  $t \mapsto r_t$  are continuous (see Henry-Labordère and Touzi [42, p. 2805]).

As in Henry-Labordère et al. [41] (see also Hobson and Klimmek [45]), our continuous-time SOT problem arises as a limit of the multi-period SOT problems, by considering the limit of the reward function  $\sum_{i=1}^T c_i(x_{t_{i-1}}, x_{t_i})$  as in (2.3), where  $(t_i)_{0 \leq i \leq T}$  is a partition of  $[0, 1]$  with  $\max_{1 \leq i \leq T} |t_i - t_{i-1}| \rightarrow 0$  as  $T \rightarrow \infty$ . To obtain the convergence we need the pathwise Itô calculus.

**Definition 2.1** (Föllmer [35]). Let  $\pi_n = (0 = t_0^n < \dots < t_n^n = 1)$ ,  $n \geq 1$  be partitions of  $[0, 1]$  with  $|\pi_n| := \max_{1 \leq k \leq n} |t_k^n - t_{k-1}^n| \rightarrow 0$  as  $n \rightarrow \infty$ . A càdlàg path  $\mathbf{x} : [0, 1] \rightarrow \mathbb{R}$  has a finite quadratic variation along  $(\pi_n)_{n \geq 1}$  if the sequence of measures on  $[0, 1]$ ,

$$\sum_{1 \leq k \leq n} (\mathbf{x}_{t_k^n} - \mathbf{x}_{t_{k-1}^n})^2 \delta_{\{t_{k-1}^n\}}(dt),$$

converges weakly to a measure  $[\mathbf{x}]^F$  on  $[0, 1]$ . For each  $t \in [0, 1]$ , set  $[\mathbf{x}]_t^F := [\mathbf{x}]^F([0, t])$ , and let  $[\mathbf{x}]^{F,c}([0, t])$  be the continuous part of this non-decreasing path.

The following will be one of our standing assumptions.

**Assumption 2.2.** The cost function  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$  is in  $C^3((l_1, \max_{0 \leq t \leq 1} r_t) \times (l_1, \max_{0 \leq t \leq 1} r_t))$  and satisfies

$$c(x, x) = c_y(x, x) = 0, \quad c_{xyy} < 0, \quad c_{xy} > 0, \quad (x, y) \in (l_1, \max_{0 \leq t \leq 1} r_t) \times (l_1, \max_{0 \leq t \leq 1} r_t).$$

Following Hobson and Klimmek [45] we then have (in fact, a weaker assumption on  $c$  would suffice; see Touzi et al. [42])

**Lemma 2.3.** Let Assumption 2.2 hold true. Then for every path  $\mathbf{x} \in \Omega$  with finite quadratic variation  $[\mathbf{x}]^F$  along a sequence of partitions  $(\pi_n)_{n \geq 1}$ , we have

$$\sum_{k=0}^{n-1} c(\mathbf{x}_{t_k^n}, \mathbf{x}_{t_{k-1}^n}) \rightarrow \frac{1}{2} \int_0^1 c_{yy}(\mathbf{x}_t, \mathbf{x}_t) d[\mathbf{x}]_t^{F,c} + \sum_{0 \leq t \leq 1} c(\mathbf{x}_{t-}, \mathbf{x}_t).$$

The non-decreasing process  $[X]_\cdot$ , as in Karandikar [54], is defined for every càdlàg path  $\mathbf{x}$  and coincides  $\mathbb{P}$ -a.s. with both, the quadratic variation and  $[\mathbf{x}]^F$ , for every  $\mathbb{P} \in \mathcal{S}_\infty$ . This and Lemma 2.3 motivates us to define the following continuous-time reward function:

$$C(\mathbf{x}) := \frac{1}{2} \int_0^1 c_{yy}(\mathbf{x}_t, \mathbf{x}_t) d[\mathbf{x}]_t^c + \sum_{0 \leq t \leq 1} c(\mathbf{x}_{t-}, \mathbf{x}_t), \quad \text{for all } \mathbf{x} \in \Omega,$$

where the integral and the sum are defined using positive and negative parts with a convention  $\infty - \infty = -\infty$ .

The continuous-time supermartingale optimal transport problem is now given by

$$\mathbf{P}_\infty(\mu) := \sup_{\mathbb{P} \in \mathcal{S}_\infty(\mu)} \mathbb{E}^\mathbb{P} [C(X)]. \quad (2.4)$$

Note that  $\sum_{0 \leq t \leq 1} c(X_{t-}, X_t) < \infty$   $\mathbb{P}$ -a.s. for all  $\mathbb{P} \in \mathcal{S}_\infty$ ; see Henry-Labordère et al. [41, Remark 2.5].

We now follow Henry-Labordère et al. [41] and introduce the dual formulation of the problem (2.4). First we introduce the set of admissible semi-static trading strategies. Let  $\mathbb{H}_0$  be the class of  $\mathbb{F}$ -predictable and locally bounded processes  $H : [0, 1] \times \Omega \rightarrow \mathbb{R}$ , i.e., there exists an increasing family of  $\mathbb{F}$ -stopping times  $(\tau_n)_{n \geq 1}$  taking values in  $[0, 1] \cup \{\infty\}$  such that  $H_{\cdot \wedge \tau_n}$  is bounded for all  $n \geq 1$  and  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Then for all  $H \in \mathbb{H}_0$  and under every supermartingale measure  $\mathbb{P} \in \mathcal{S}_\infty$ , one can define the integral, denoted by  $H \cdot X$ , see Jacod and Shiryaev [51, Chapter I.4]. Define

$$\mathcal{H} := \{H \in \mathbb{H}_0 : H \geq 0 \text{ and } H \cdot X \text{ is a } \mathbb{P}\text{-supermartingale for every } \mathbb{P} \in \mathcal{S}_\infty\}.$$

For the static strategy, we denote by  $M([0, 1])$  the space of all finite signed measures on  $[0, 1]$  which is a Polish space under the weak convergence topology, and by  $\Lambda$  the class of all measurable



maps  $\lambda : \mathbb{R} \rightarrow M([0, 1])$  which admit a representation  $\lambda(x, dt) = \lambda_0(t, x)\gamma(dt)$  for some finite non-negative measure  $\gamma$  on  $[0, 1]$  and some measurable function  $\lambda_0 : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$  which is bounded on  $[0, 1] \times K$  for all compact  $K$  of  $\mathbb{R}$ . We then denote

$$\Lambda(\mu) := \{\lambda \in \Lambda : \mu(|\lambda|) < \infty\}, \quad \text{where } \mu(|\lambda|) := \int_0^1 \int_{\mathbb{R}} |\lambda_0(t, x)| \mu_t(dx) \gamma(dt).$$

We also introduce a family of random measures  $\delta^X = (\delta_t^X)_{0 \leq t \leq 1}$ , associated to the canonical process  $X$ , by setting  $\delta_t^X(dx) := \delta_{X_t}(dx)$ . In particular, one has

$$\delta^X(\lambda) = \int_0^1 \lambda(X_t, dt) = \int_0^1 \lambda_0(t, X_t) \gamma(dt).$$

The set of admissible superhedging strategies with no short-selling constraint is defined by

$$\mathcal{D}_\infty := \{(H, \lambda) \in \mathcal{H} \times \Lambda(\mu) : \delta^X(\lambda) + (H \cdot X)_1 \geq C(X), \mathbb{P}\text{-a.s.}, \forall \mathbb{P} \in \mathcal{S}_\infty\},$$

and then the dual problem is given by

$$\mathbf{D}_\infty(\mu) := \inf_{(H, \lambda) \in \mathcal{D}_\infty(\mu)} \mu(\lambda). \quad (2.5)$$

Furthermore, the weak duality holds:  $\mathbf{P}_\infty(\mu) \leq \mathbf{D}_\infty(\mu)$ .

### 3 Main results

Our contribution splits into three parts. First we construct a continuous-time supermartingale with given marginals  $(\mu_t)_{t \in [0, 1]}$ . This is achieved as an accumulation point of a sequence  $(\mathbb{P}^n)_n$  of solutions of  $n$ -period supermartingale transport problems. We then characterize the limiting law further and show that it corresponds to the distribution of a local Lévy process. Finally we prove that this limit solves the continuous-time supermartingale optimal transport problem (2.4). The last point is achieved by proving the strong duality  $\mathbf{P}_\infty(\mu) = \mathbf{D}_\infty(\mu)$ , and explicitly constructing an optimal superhedging strategy.

#### 3.1 Convergence of the sequence $(\mathbb{P}^n)_{n \geq 1}$

For every  $t \in [0, 1]$ , we denote by  $\mathbb{R} \ni x \mapsto F(t, x)$  and  $[0, 1] \ni u \mapsto F^{-1}(t, u)$  the cumulative distribution function and the corresponding right-continuous inverse (w.r.t. the  $x$ -variable) of the probability measure  $\mu_t$ . For  $t \in [0, 1]$ ,  $\epsilon \in (0, 1 - t]$  and  $x, y \in \mathbb{R}$  let

$$\delta^\epsilon F(t, x) := F(t + \epsilon, x) - F(t, x)$$

and define

$$E := \{(t, x) : t \in [0, 1], x \in (\ell_t, r_t)\}.$$

**Assumption 3.1.** (i) *The marginal distributions  $\mu = (\mu_t)_{t \in [0, 1]}$  belong to  $\mathcal{P}$ , are non-decreasing in convex-decreasing and  $t \mapsto \mu_t$  is continuous w.r.t. the weak convergence topology.*

(ii)  *$F \in C_b^4(E)$ . In particular, for all  $t \in [0, 1]$ , the measure  $\mu_t$  has a density function  $f(t, \cdot)$  such that  $f(t, \cdot) > 0$  on  $(\ell_t, r_t)$ .*

(iii) *There exists  $\epsilon_1 \in (0, 1]$  such that for all  $t \in [0, 1]$  and  $0 < \epsilon \leq \epsilon_1 \wedge (1 - t)$ ,  $\mu_t$  and  $\mu_{t+\epsilon}$  satisfy the Dispersion Assumption:*

- *There exists  $m_\epsilon(t) < m^\epsilon(t)$  such that  $\ell_{t+\epsilon} \leq \ell_t \leq m_\epsilon(t) < m^\epsilon(t) \leq r_t \leq r_{t+\epsilon}$ , with  $\ell_{t+\epsilon} < \ell_t \wedge m_\epsilon(t)$  and  $m^\epsilon(t) \vee r_t < r_{t+\epsilon}$ .*

- Furthermore,  $f(t + \epsilon, \cdot) < f(t, \cdot)$  on  $(m_\epsilon(t), m^\epsilon(t))$ ,  $f(t + \epsilon, x) = f(t, x)$  for  $x \in \{m_\epsilon(t), m^\epsilon(t)\}$ , and  $f(t + \epsilon, \cdot) > f(t, \cdot)$  on  $(\ell_{t+\epsilon}, m_\epsilon(t)) \cup (m^\epsilon(t), r_{t+\epsilon})$ .

Note that Assumption 3.1(iii) is precisely Assumption A.3, and thus all the results of Appendix B are valid.

**Remark 3.2.** Under Assumption 3.1 several further implications follow. First, since, for all  $t \in [0, 1]$ ,  $f(t, \cdot) > 0$  on  $(\ell_t, r_t)$ , we have that  $\inf_{x \in [-K, K] \cap (\ell_t, r_t)} f(t, x) > 0$  for all  $K > 0$ . Furthermore, due to the Dispersion Assumption (see Assumption 3.1(iii)), for every  $t \in [0, 1)$  and sufficiently small  $\epsilon \in (0, 1 - t]$ , the function  $x \mapsto \delta^\epsilon F(t, x)$  attains the (unique) maximum and minimum at  $m_\epsilon(t)$  and  $m^\epsilon(t)$ , respectively. It is also easy to see that, for all  $x \in (\ell_t, r_t)$ ,  $\int_x^\infty (y - x) \delta^\epsilon F(t, dy) > 0$ . In particular, Assumption 3.1 implies Henry-Labordère et al. [41, Assumption 3.1].

In the rest of this paper we will always work with  $(\mu_t)_{t \in [0, 1]}$  satisfying Assumption 3.1.

Fix  $t \in [0, 1)$  and  $\epsilon \in (0, 1 - t]$ , and consider a pair of marginals  $(\mu_t, \mu_{t+\epsilon})$ . Then the corresponding decreasing supermartingale coupling can be determined by a pair of functions  $(T_d^\epsilon(t, \cdot) = \hat{T}_d^{\mu_t, \mu_{t+\epsilon}}(\cdot), T_u^\epsilon(t, \cdot) = \hat{T}_u^{\mu_t, \mu_{t+\epsilon}}(\cdot))$ , where  $(\hat{T}_d^{\mu_t, \mu_{t+\epsilon}}(\cdot), \hat{T}_u^{\mu_t, \mu_{t+\epsilon}}(\cdot))$  are as in Appendix B.1. Furthermore, set  $x_1^\epsilon(t) := x_1^{\mu_t, \mu_{t+\epsilon}}$  and  $y_1^\epsilon(t) := y_1^{\mu_t, \mu_{t+\epsilon}}$ , where  $x_1^{\mu_t, \mu_{t+\epsilon}}$  is defined by (B.50), while  $y_1^{\mu_t, \mu_{t+\epsilon}}$  is given by (B.51).

Recall that the canonical space of càdlàg paths  $\Omega := \mathbb{D}([0, 1]; \mathbb{R})$ , is a Polish space equipped with the Skorokhod topology, while  $X$  denotes the canonical process. Let  $(\pi_n)_{n \geq 1}$  be a sequence of partitions of  $[0, 1]$ , so that each  $\pi_n = (t_k^n)_{0 \leq k \leq n}$  is such that  $0 = t_0^n < \dots < t_n^n = 1$ . Suppose in addition that  $|\pi_n| := \max_{1 \leq k \leq n} (t_k^n - t_{k-1}^n) \rightarrow 0$  as  $n \rightarrow \infty$ . Then for every partition  $\pi_n$ , by considering the marginal distributions  $(\mu_{t_k^n})_{0 \leq k \leq n}$ , one obtains an  $(n+1)$ -marginal (or  $n$ -period) SOT problem, where the goal is to maximize

$$\mathbb{E} \left[ \sum_{0 \leq k \leq n-1} c(\tilde{X}_k^n, \tilde{X}_{k+1}^n) \right]$$

among all discrete-time supermartingales  $\tilde{X}^n = (\tilde{X}_k^n)_{0 \leq k \leq n}$  satisfying marginal constraints. Under Assumptions 2.2 and 3.1, the iterated  $n$ -period decreasing supermartingale coupling is the solution to the above supermartingale transport problem. Let  $\Omega^{*,n} := \mathbb{R}^{n+1}$  be the canonical space of discrete-time process, and  $X^n = (X_k^n)_{0 \leq k \leq n}$  be the canonical process on  $\Omega^{*,n}$ . Then under the optimal supermartingale measure  $\mathbb{P}^{*,n}$ ,  $X^n$  is a discrete-time supermartingale and at the same time a Markov chain, characterized by  $T_u^{\Delta t_{k+1}^n}(t_k^n, \cdot)$ ,  $T_d^{\Delta t_{k+1}^n}(t_k^n, \cdot)$  with  $\Delta t_{k+1}^n := t_{k+1}^n - t_k^n$ , induced by the two marginals  $(\mu_{t_k^n}, \mu_{t_{k+1}^n})$ ; see Appendix B.1. We then extend the Markov chain  $X^n$  to a continuous-time càdlàg process  $X^{*,n} = (X_t^{*,n})_{0 \leq t \leq 1}$  defined by

$$X_t^{*,n} := \sum_{k=1}^n X_{k-1}^n \mathbf{1}_{[t_{k-1}^n, t_k^n)}(t), \quad t \in [0, 1],$$

and define the probability measure  $\mathbb{P}^n := \mathbb{P}^{*,n} \circ (X^{*,n})^{-1}$  on  $\Omega$ .

We further denote by  $\mathbf{m}(\delta^\epsilon F(t, \cdot))$  (resp.  $\mathbf{m}(\partial_t F(t, \cdot))$ ) the set of all local minimizers of functions  $\delta^\epsilon F(t, \cdot)$  (resp.  $\partial F(t, \cdot)$ ). Let  $\mathbf{M}(\partial_t F(t, \cdot))$  be the set of local maximizers of  $\partial F(t, \cdot)$ . Note that, by Assumption 3.1,  $\mathbf{m}(\delta^\epsilon F(t, \cdot)) = \{m^\epsilon(t)\}$  and  $\mathbf{M}(\delta^\epsilon F(t, \cdot)) = \{m_\epsilon(t)\}$  for all  $0 < \epsilon \leq \epsilon_1 \wedge (1 - t)$ .

- Assumption 3.3.** (i) There is a constant  $\epsilon_2 > 0$  such that, for all  $t \in [0, 1]$  and  $0 < \epsilon \leq \epsilon_2 \wedge (1 - t)$ , we have  $\mathbf{m}(\partial_t F(t, \cdot)) = \{m_t\}$  and  $\mathbf{M}(\partial_t F(t, \cdot)) = \{\tilde{m}_t\}$ .
- (ii) Let  $m^0(t) = m_t$ . Then the map  $(t, \epsilon) \mapsto m^\epsilon(t)$  is continuous (and hence uniformly continuous with continuity modulus  $\rho_0$ ) on  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_2, 0 \leq t \leq 1 - \epsilon\}$ .



**Proposition 3.4.** Suppose that Assumptions 3.1 and 3.3 are valid. Then the sequence  $(\mathbb{P}^n)_{n \geq 1}$  is tight w.r.t. the Skorokhod topology on  $\Omega$ . Moreover, every limit point  $\mathbb{P}^0$  satisfies  $\mathbb{P}^0 \in \mathcal{S}_\infty(\mu)$ .

The proof of Proposition 3.4 requires two additional lemmas.

**Lemma 3.5.** Let  $(\mu_t)_{t \in [0,1]}$  be a family of probability measures in  $\mathcal{P}$  which are increasing in convex-decreasing order. Then  $(\mu_t)_{t \in [0,1]}$  is uniformly integrable, i.e.

$$\lim_{K \rightarrow +\infty} \sup_{0 \leq t \leq 1} \int |x| \mathbf{1}_{\{|x| \geq K\}} \mu_t(dx) = 0.$$

*Proof.* Using the inequality  $|x| \mathbf{1}_{\{|x| \geq K\}} \leq 2(|x| - \frac{K}{2}) \mathbf{1}_{\{|x| \geq \frac{K}{2}\}}$ , we have that

$$\begin{aligned} \sup_{0 \leq t \leq 1} \int |x| \mathbf{1}_{\{|x| \geq K\}} \mu_t(dx) &= \sup_{0 \leq t \leq 1} \int 2(|x| - \frac{K}{2}) \mathbf{1}_{\{|x| \geq \frac{K}{2}\}} \mu_t(dx) \\ &\leq 2 \sup_{0 \leq t \leq 1} \int (x - \frac{K}{2}) \mathbf{1}_{\{x \geq \frac{K}{2}\}} \mu_t(dx) + 2 \sup_{0 \leq t \leq 1} \int (-x - \frac{K}{2}) \mathbf{1}_{\{x \leq -\frac{K}{2}\}} \mu_t(dx) \\ &= 2 \sup_{0 \leq t \leq 1} C_{\mu_t}(\frac{K}{2}) + 2 \sup_{0 \leq t \leq 1} P_{\mu_t}(-\frac{K}{2}) \\ &= 2C_{\tilde{\mu}}(\frac{K}{2}) + 2P_{\mu_1}(-\frac{K}{2}) \rightarrow 0, \text{ when } K \rightarrow \infty, \end{aligned}$$

where  $\tilde{\mu}$  is some probability measure on  $\mathbb{R}$  with the same mass and mean with  $\mu_0$ . We now explain the last equality. First,  $x \mapsto (-x - \frac{K}{2})^+$  is convex and non-increasing, and thus using the convex-decreasing ordering of  $(\mu_t)_{t \in [0,1]}$  we obtain that  $\sup_{0 \leq t \leq 1} P_{\mu_t}(-\frac{K}{2}) = P_{\mu_1}(-\frac{K}{2})$ . On the other hand,  $\sup_{0 \leq t \leq 1} C_{\mu_t}(\frac{K}{2}) = C_{\tilde{\mu}}(\frac{K}{2})$  follows from the fact that the point-wise supremum of a family of convex functions is convex, and  $\lim_{|x| \rightarrow \infty} \{\sup_{0 \leq t \leq 1} C_{\mu_t}(x) - \{\bar{\mu}_0 - \mu_0(\mathbb{R})x\}\} = 0$ . Therefore, there exists some probability measure  $\tilde{\mu}$ , such that  $\tilde{\mu}(\mathbb{R}) = \mu_0(\mathbb{R})$ ,  $\bar{\mu}_0 = \bar{\tilde{\mu}}$  and  $C_{\tilde{\mu}} = \sup_{0 \leq t \leq 1} C_{\mu_t}$ .  $\square$

Recall that  $(\mu_t)_{t \in [0,1]}$  satisfies Assumption 3.1. Now for fixed  $t \in [0,1]$ ,  $\epsilon \in (0, 1-t]$  and  $x \in (r_t, m^\epsilon(t))$  denote by  $q^\epsilon(t, x)$  the conditional probability of the (martingale) upward jump under the decreasing supermartingale coupling, so that  $q^\epsilon(t, x) = \frac{x - T_d^\epsilon(t, x)}{T_u^\epsilon(t, x) - T_d^\epsilon(t, x)}$ . Note that, for  $x \leq x_1^\epsilon(t)$ ,  $T_u^\epsilon(t, x) = \infty$  and therefore  $q^\epsilon(t, x) = 0$ .

Furthermore, define  $J_d^\epsilon(t, \cdot), J_u^\epsilon(t, \cdot)$  by

$$J_d^\epsilon(t, x) = x - T_d^\epsilon(t, x), \quad J_u^\epsilon(t, x) = T_u^\epsilon(t, x) - x, \quad x \in (\ell_t, r_t),$$

so that  $J_d^\epsilon(t, \cdot), J_u^\epsilon(t, \cdot)$  correspond to the downward and upward distance each particle travels (between times  $t$  and  $t+\epsilon$ ) under the decreasing supermartingale coupling, conditioned it started at  $x$ . (Note that  $J_u^\epsilon(t, x) = \infty$  for  $x \leq x_1^\epsilon(t)$ , but this happens with zero probability.)

**Lemma 3.6.** Suppose Assumptions 3.1 and 3.3 hold. Then

(i) For every  $K > 0$ , there is a constant  $C_1$ , independent of  $(t, x, \epsilon)$  such that

$$J_d^\epsilon(t, x) + q^\epsilon(t, x) \leq C_1 \epsilon, \quad \forall x \in [-K, K] \cap (\ell_t, r_t) \cap (x_1^\epsilon(t), m_t^\epsilon).$$

(ii) For every  $K > 0$ , there is a constant  $C_2$ , independent of  $(t, x, \epsilon)$  such that

$$J_d^\epsilon(t, x) \leq C_2 \epsilon, \quad \forall x \in [-K, K] \cap (\ell_t, x_1^\epsilon(t)).$$

*Proof.* (i) Since  $\mu_t|_{[x_1^\epsilon(t), r_t]} \leq_c \mu_{t+\epsilon}|_{[y_1^\epsilon(t), r_{t+\epsilon}]}$  and the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  coincides with the right-curtain coupling of  $\mu_t|_{[x_1^\epsilon(t), r_t]} \leq_c \mu_{t+\epsilon}|_{[y_1^\epsilon(t), r_{t+\epsilon}]}$  (see Appendix B.1), the result follows immediately from Henry-Labordère and Touzi [42, Lemma 6.4] applied to the right-curtain coupling.

(ii) For  $x \leq x_1^\epsilon(t)$ ,  $T_d^\epsilon(t, x) = F^{-1}(t + \epsilon, F(t, x))$  and therefore

$$\begin{aligned} J_d^\epsilon(t, x) &= F^{-1}(t + \epsilon, F(t + \epsilon, x)) - F^{-1}(t + \epsilon, F(t, x)) \\ &= \frac{1}{f(t + \epsilon, F^{-1}(t + \epsilon, \xi))} \delta^\epsilon F(t, x), \end{aligned}$$

for some  $\xi$  between  $F(t + \epsilon, x)$  and  $F(t, x)$ , by the mean value theorem. Now one draws the desired conclusion from the fact that  $|\delta^\epsilon F| \leq C\epsilon$ , for some constant  $C > 0$ .  $\square$

We are now ready to prove Proposition 3.4.

*Proof of Proposition 3.4.* Recall that  $\mathbb{P}^n$  is a supermartingale measure on the canonical space  $\Omega$ , induced by the continuous-time supermartingale  $X^{*,n}$  under the probability  $\mathbb{P}^{*,n}$ . The supermartingale  $X^{*,n}$  jumps only at discrete time points relative to the partition  $\pi_n = (t_k^n)_{0 \leq k \leq n}$ . Moreover, at time  $t_{k+1}^n$ , on  $\{x_1^\epsilon(t_k^n) \leq X_{t_k^n}^k\}$ , the upward jump size is  $J_u^\epsilon(t_k^n, X_{t_k^n}^k)$  and downward jump size is  $J_d^\epsilon(t_k^n, X_{t_k^n}^k)$ , where  $\epsilon := t_{k+1}^n - t_k^n$ . In addition, at time  $t_{k+1}^n$  and on  $\{x_1^\epsilon(t_k^n) \geq X_{t_k^n}^k\}$ , the downward jump size is  $J_d^\epsilon(t_k^n, X_{t_k^n}^k)$  with probability 1.

For given positive constants  $C, \theta$ , we introduce

$$\begin{aligned} \mathcal{E}_n(C, \theta) &:= \inf \{ \Pi_{i=j}^{k-1} (1 - C(t_{i+1}^n - t_i^n)) : \text{for some } s \in [0, 1) \text{ and} \\ &\quad 0 \leq j \leq k \leq n \text{ such that } s \leq t_j^n \leq t_{k+1}^n \leq s + \theta \}. \end{aligned}$$

(i) First, using Jacod and Shiryaev [51, Theorem VI.4.5], we show that  $(\mathbb{P}^n)_{n \geq 1}$  is tight. Let  $\tau := 1 \wedge \inf \{s : |X_s| \geq K\}$ , and  $\tau$  is a stopping time w.r.t. the canonical filtration generated by  $(X_s)_{0 \leq s \leq 1}$ . In addition,

$$\begin{aligned} K \mathbb{P}^n \left[ \sup_{0 \leq t \leq 1} |X_t| \geq K \right] &\leq \mathbb{E}^{\mathbb{P}^n} [|X_\tau|] \\ &= \mathbb{E}^{\mathbb{P}^n} [X_\tau + 2X_\tau^-] \leq \mathbb{E}^{\mathbb{P}^n} [X_0 + 2X_\tau^-] \\ &\leq \mathbb{E}^{\mathbb{P}^n} [X_0 + 2X_1^-] \leq \mathbb{E}^{\mathbb{P}^n} [|X_0|] + 2\mathbb{E}^{\mathbb{P}^n} [|X_1|] < +\infty, \end{aligned}$$

where on the above, the first inequality comes from Markov inequality,  $\mathbb{P}^{*,n} \circ (X^{*,n})^{-1} = \mathbb{P}^n \circ X^{-1}$ , and that fact that  $X_t^{*,n}$  is piecewise constant. The third inequality comes from optional stopping theorem and the fact that  $X^-$  is a submartingale. Let  $\eta > 0$  be an arbitrary small real number, then there is some  $K > 0$  such that

$$\mathbb{P}^n \left[ \sup_{0 \leq t \leq 1} |X_t| \geq K \right] \leq \eta, \quad \text{for all } n \geq 1.$$

We can assume w.l.o.g. that  $K$  satisfies  $-K < m_t < K$  for all  $t \in [0, 1]$ ; recall Assumption 3.3.

The remaining proof is almost identical to the proof of [41, Proposition 3.2]. However, in few places the supermartingale transitions appear (which are not present in [41]) and thus we sketch the arguments.

Denote in addition that  $r^K(t) := r_t \wedge K$  and  $\ell^K(t) := \ell_t \vee (-K)$ . Let  $\delta > 0$ , then it follows from Lemma 3.6 that  $J_d^\epsilon(t_k^n, X_{t_k^n}^k)$  is uniformly bounded by  $C\epsilon$  for some constant  $C$  on  $D_\delta^K := \{(t, x) : \ell^K(t) + \delta/2 \leq x \leq m_t\}$ . Let  $\theta > 0$  satisfy  $\theta \leq \frac{\delta}{2C}$  and  $|\ell^K(t + \theta) - \ell^K(t)| + |r^K(t + \theta) - r^K(t)| \leq \delta/2$ . Let  $S, T$  be two stopping times w.r.t. the filtration

generated by  $X^{*,n}$  such that  $0 \leq S \leq T \leq S + \theta \leq 1$ . Since, by Lemma 3.6, the big jumps of  $X^{*,n}$  correspond to  $J_u^\epsilon(t_n^k, X_{t_n^k}^k)$ , and not to  $J_d^\epsilon(t_n^k, X_{t_n^k}^k)$ , using the same arguments as in the proof of [41, Proposition 3.2], we have that

$$\mathbb{P}^{*,n} \left[ \sup_{0 \leq t \leq 1} |X_t^{*,n}| \leq K, |X_T^{*,n} - X_S^{*,n}| \geq \delta \right] \leq 1 - \mathcal{E}_n(C, \theta).$$

It follows that

$$\begin{aligned} & \limsup_{\theta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^{*,n} [|X_T^{*,n} - X_S^{*,n}| \geq \delta] \\ & \leq \limsup_{\theta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^{*,n} \left[ \sup_{0 \leq t \leq 1} |X_t^{*,n}| \leq K, |X_T^{*,n} - X_S^{*,n}| \geq \delta \right] + \mathbb{P}^{*,n} \left[ \sup_{0 \leq t \leq 1} |X_t^{*,n}| \geq K \right] \\ & \leq \limsup_{\theta \rightarrow 0} \limsup_{n \rightarrow \infty} (1 - \mathcal{E}_n(C, \theta)) + \eta = \eta. \end{aligned}$$

Since  $\eta > 0$  can be arbitrarily small, we conclude that

$$\limsup_{\theta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^{*,n} [|X_T^{*,n} - X_S^{*,n}| \geq \delta] = 0.$$

Finally, Jacod and Shiryaev [51, Theorem VI.4.5] shows that  $(\mathbb{P}^n)_{n \geq 1}$  is tight.

- (ii) Let  $\mathbb{P}^0$  be a limit of  $(\mathbb{P}^n)_{n \geq 1}$ . Then repeating the arguments of the second part of the proof of [41, Proposition 3.2] (which relies on the continuity of  $F(t, x)$ ), we obtain that  $\mathbb{P}^0 \circ X_t^{-1} = \mu_t$ .
- (iii) In the last step, we show that  $X$  is still a supermartingale under  $\mathbb{P}^0$ . For every  $K > 0$ , we first define the auxiliary process  $X_t^K := (-K) \vee X_t \wedge K$ . Given  $s < t$  and  $\varphi(s, X.)$  a bounded continuous,  $\mathcal{F}_s$ -measurable function, it follows from weak convergence that

$$\lim_{n \rightarrow +\infty} \mathbb{E}^{\mathbb{P}^n} [\varphi(s, X.) (X_t^K - X_s^K)] = \mathbb{E}^{\mathbb{P}^0} [\varphi(s, X.) (X_t^K - X_s^K)].$$

As  $X$  is a  $\mathbb{P}^n$ -supermartingale, we have that

$$\begin{aligned} \mathbb{E}^{\mathbb{P}^n} [\varphi(s, X.) (X_t^K - X_s^K)] & \leq -\mathbb{E}^{\mathbb{P}^n} [\varphi(s, X.) (X_t \mathbf{1}_{\{|X_t| \geq K\}} - X_s \mathbf{1}_{\{|X_s| \geq K\}})] \\ & \leq |\mathbb{E}^{\mathbb{P}^n} [\varphi(s, X.) (X_t \mathbf{1}_{\{|X_t| \geq K\}} - X_s \mathbf{1}_{\{|X_s| \geq K\}})]| \\ & \leq 2|\varphi|_\infty \sup_{0 \leq t \leq 1} \int |x| \mathbf{1}_{\{|x| \geq K\}} \mu_t(dx) \\ & \rightarrow 0, \text{ as } K \rightarrow \infty, \end{aligned}$$

where the last inequality follows from Lemma 3.5 and the above convergence is uniformly in  $n$ . Finally, by the dominated convergence theorem, we have that

$$\begin{aligned} \mathbb{E}^{\mathbb{P}^0} [\varphi(s, X.) (X_t - X_s)] & = \lim_{K \rightarrow \infty} \mathbb{E}^{\mathbb{P}^0} [\varphi(s, X.) (X_t^K - X_s^K)] \\ & = \lim_{K \rightarrow \infty} \lim_{n \rightarrow +\infty} \mathbb{E}^{\mathbb{P}^n} [\varphi(s, X.) (X_t^K - X_s^K)] \\ & \leq 0. \end{aligned}$$

As  $\varphi$  is arbitrary, we conclude that  $X$  is a  $\mathbb{P}^0$ -supermartingale. □

### 3.2 Characterization of the limiting process

To further analyse the limiting process of the previous section, we consider the decreasing supermartingale coupling (see Appendix B) of two marginals  $\mu_t$  and  $\mu_{t+\epsilon}$ . Recall that  $x_1^\epsilon(t)$  denotes the unique phase transition point which divides the martingale region (to the right of  $x_1^\epsilon(t)$ ) and the supermartingale region (to the left of  $x_1^\epsilon(t)$ ). In the supermartingale region, the mass at  $x \leq x_1^\epsilon(t)$  is transported (through the quantile coupling) to the destination  $T_d^\epsilon(t, x) = F^{-1}(t + \epsilon, F(t, x))$ . In the martingale region, the two supporting maps  $T_u^\epsilon$  and  $T_d^\epsilon$  are characterized in Proposition B.4.

Set  $\epsilon_0 := \epsilon_1 \wedge \epsilon_2$ , where  $\epsilon_1, \epsilon_2$  are as in Assumptions 3.1 and 3.3, respectively.

**Assumption 3.7.** (i) *The map  $(t, \epsilon) \mapsto x_1^\epsilon(t)$  is uniformly continuous (with continuity modulus  $\rho_1$ ) and also uniformly bounded on  $\{(t, \epsilon) : 0 < \epsilon \leq \epsilon_0, 0 \leq t \leq 1 - \epsilon\}$ .*

(ii) *For every  $t \in [0, 1]$ ,  $r_t = r \in (-\infty, \infty]$ .*

(iii) *For every  $t \in [0, 1]$ , we have  $\partial_{tx}f(t, m_t) > 0$ .*

**Lemma 3.8.** *Suppose that Assumptions 3.1, 3.3, 3.7(i) and (ii) hold. Then the limit  $\lim_{\epsilon \rightarrow 0} x_1^\epsilon(t)$  exists. Moreover, if we denote the limit by  $x_1(t)$ , then  $t \mapsto x_1(t)$  is continuous and uniquely determined by the equation*

$$\int_{x_1(t)}^{+\infty} (x_1(t) - \xi) \partial_t f(t, \xi) d\xi = 0, \quad t \in [0, 1].$$

*Proof.* It follows from Lemma B.3 that  $x_1^\epsilon(t)$  is unique for all  $t \in [0, 1)$ ,  $\epsilon \in (0, \epsilon_0]$ . From Assumption 3.7 (i), using the Arzelà-Ascoli theorem, we have that  $x_1^\epsilon(t)$  converges when  $\epsilon \rightarrow 0$ , and the limit remains a continuous function.

Using the properties of the decreasing supermartingale coupling (see Appendix B), we have that  $x_1^\epsilon(t)$  and  $y_1^\epsilon(t)$  satisfy the mean and mass preservation conditions:

$$\int_{x_1^\epsilon(t)}^{r_t} f(t, x) dx = \int_{y_1^\epsilon(t)}^{r_{t+\epsilon}} f(t + \epsilon, x) dx, \quad \int_{x_1^\epsilon(t)}^{r_t} x f(t, x) dx = \int_{y_1^\epsilon(t)}^{r_{t+\epsilon}} x f(t + \epsilon, x) dx.$$

Rearranging, and using that  $r_t = r_{t+\epsilon}$ , gives

$$\int_{x_1^\epsilon(t)}^{r_t} (f(t, x) - f(t + \epsilon, x)) dx = \int_{y_1^\epsilon(t)}^{x_1^\epsilon(t)} f(t + \epsilon, x) dx, \quad \int_{x_1^\epsilon(t)}^{+\infty} x(f(t, x) - f(t + \epsilon, x)) dx = \int_{y_1^\epsilon(t)}^{x_1^\epsilon(t)} x f(t + \epsilon, x) dx.$$

Using Taylor's expansion  $f(t + \epsilon, x) - f(t, x) = \partial_t f(t, x) \epsilon + \frac{1}{2} \partial_{tt} f(\xi_1, x) \epsilon^2$  (with  $\xi_1 \in (t, t + \epsilon)$ ) in the first equation above and applying the mean value theorem for definite integrals, we get that

$$-\epsilon \int_{x_1^\epsilon(t)}^{+\infty} \partial_t f(t, x) dx - \frac{1}{2} \epsilon^2 \int_{x_1^\epsilon(t)}^{+\infty} \partial_{tt} f(\xi_1, x) dx = (x_1^\epsilon(t) - y_1^\epsilon(t)) f(t, \xi_2),$$

where  $\xi_2$  lies in  $[y_1^\epsilon(t), x_1^\epsilon(t)]$ . Letting  $\epsilon$  tend to 0 on both sides, the l.h.s. tends to 0, and consequently r.h.s. also tends to 0. It follows that  $\lim_{\epsilon \rightarrow 0} x_1^\epsilon(t) = \lim_{\epsilon \rightarrow 0} y_1^\epsilon(t)$ . Dividing both sides by  $\epsilon$  we get

$$-\int_{x_1^\epsilon(t)}^{+\infty} \partial_t f(t, x) dx = \frac{x_1^\epsilon(t) - y_1^\epsilon(t)}{\epsilon} f(t, \xi_2) + O(\epsilon), \quad (3.6)$$

Similarly,

$$-\int_{x_1^\epsilon(t)}^{+\infty} x \partial_t f(t, x) dx = \frac{x_1^\epsilon(t) - y_1^\epsilon(t)}{\epsilon} \xi_2 f(t, \xi_2) + O(\epsilon). \quad (3.7)$$

Note that Assumption 3.7(ii) ensures that  $x \mapsto \partial_t f(t, x)$  is not a constant.

Combining (3.6) and (3.7), and letting  $\epsilon$  tend to 0, we get that

$$\int_{x_1(t)}^{+\infty} (x - x_1(t)) \partial_t f(t, x) dx = 0. \quad \square$$

**Remark 3.9.** The Assumption 3.7(ii) can be relaxed. For example, if  $t \mapsto r_t$  is differentiable, then the equation that characterizes  $x_1(t)$ , as in Lemma 3.8, can be shown to be of the following form:

$$\int_{x_1(t)}^{r_t} (x_1(t) - \xi) \partial_t f(t, \xi) d\xi = -f(t, r_t)(x_1(t) - r_t) \frac{dr_t}{dt}, \quad t \in [0, 1].$$

Consequently, the subsequent arguments would have to be adjusted. Since our main motivation is the examples that satisfy Assumption 3.7(ii), we do not include the details.

We also introduce  $T_u$  through the following integral equation:

$$\int_x^{T_u(t, x)} (x - \xi) \partial_t f(t, \xi) d\xi = 0, \quad (3.8)$$

**Proposition 3.10.** Suppose that Assumptions 3.1, 3.3, 3.7(i) and (ii) hold. For  $x \in (x_1(t), m_t)$ , (3.8) has the unique solution  $T_u(t, x)$  on  $(m_t, r_t)$ .

*Proof.* Define  $G(t, x, y) := \int_x^y (x - \xi) \partial_t f(t, \xi) d\xi$ , as  $y \mapsto G(t, x, y)$  is continuous, it is enough to show that, for  $x \in [x_1(t), m_t]$ ,  $y \mapsto G(t, x, y)$  is decreasing on  $y \in [m_t, \infty)$ ,  $G(t, x, m_t) > 0$  and  $G(t, x, +\infty) < 0$ .

We first show that for  $x \in [x_1(t), m_t]$ ,  $y \mapsto G(t, x, y)$  is decreasing on  $y \in [m_t, \infty)$ . Indeed, for  $m_t \leq y_1 < y_2$ , we have

$$G(t, x, y_2) - G(t, x, y_1) = \int_{y_1}^{y_2} (x - \xi) \partial_t f(t, \xi) d\xi < 0,$$

where the last inequality follows from the fact that  $\xi > y_1 \geq m_t \geq x$  and  $\partial_t f(t, \xi) > 0$  on  $[y_1, y_2]$  (as  $x \mapsto \partial_t F(t, x)$  is increasing on  $(m_t, \infty)$ ).

Notice that the equation  $\partial_t F(t, x) = 0$  has a unique solution under Assumption 3.3, and we denote it by  $\bar{x}(t)$ . Now let us show that  $x \mapsto G(t, x, +\infty)$  is decreasing on  $(x_1(t), \bar{x}(t))$ , and increasing on  $(\bar{x}(t), m_t)$ . Indeed,

$$\partial_x G(t, x, +\infty) = \int_x^{+\infty} \partial_t f(t, \xi) d\xi = \partial_t F(t, +\infty) - \partial_t F(t, x) = -\partial_t F(t, x),$$

which is negative on  $(x_1(t), \bar{x}(t))$ , and positive on  $(\bar{x}(t), m_t)$ . To show  $G(t, x, +\infty) < 0$  for all  $x_1(t) < x < m_t$ , it is enough to show that  $G(t, x_1(t), +\infty) \leq 0$  and  $G(t, m_t, +\infty) \leq 0$ . Notice that from Lemma 3.8,  $x_1(t)$  satisfies that

$$\int_{x_1(t)}^{+\infty} (x_1(t) - \xi) \partial_t f(t, \xi) d\xi = 0,$$

i.e.,  $G(t, x_1(t), +\infty) = 0$ . Moreover, as  $G(t, m_t, m_t) = 0$  and  $y \mapsto G(t, m_t, y)$  is strictly decreasing on  $y \in [m_t, \infty]$  which is shown above, it follows that  $G(t, m_t, +\infty) < 0$ .

In addition,  $x \mapsto G(t, x, m_t)$  is decreasing on the interval  $(x_1(t), m_t)$ :

$$\partial_x G(t, x, m_t) = \int_x^{m_t} \partial_t f(t, \xi) d\xi = \partial_t F(t, m_t) - \partial_t F(t, x) < 0.$$

It follows that  $G(t, x, m_t) > G(t, m_t, m_t) = 0$ . □

We further define  $j_u$  and  $j_d$  for  $t \in [0, 1], x \in (x_1(t), m_t)$  by  $j_u(t, x) := T_u(t, x) - x$  and  $j_d(t, x) := \frac{\partial_t F(t, x) - \partial_t F(t, T_u(t, x))}{f(t, x)}$ . The following lemma analyzes the asymptotic behaviour of  $T_d^\epsilon$ ,  $T_u^\epsilon$  and  $q^\epsilon(t, x)$  in the martingale region, and when  $\epsilon$  is small. The proof is analogous to the proof of Henry-Labordère et al. [41, Lemma 6.6] when applied to the right-curtain coupling, and thus omitted.

**Lemma 3.11.** *Suppose that Assumptions 3.1, 3.3, 3.7(i) and (ii) hold.*

*For  $0 < \delta < K < \infty$ , let  $E_{\delta,K} := \{(t, x) : t \in [0, 1], (m_t - K) \vee x_1(t) < x \leq m_t - \delta\}$ . Define  $j_d^\epsilon(t, x) := \frac{\partial_t F(t, x) - \partial_t F(t, T_u^\epsilon(t, x))}{f(t + \epsilon, x)}$ . Then  $T_u^\epsilon(t, x)$  and  $T_d^\epsilon(t, x)$  admit the following expansions:*

$$T_u^\epsilon(t, x) - x := J_u^\epsilon(t, x) = j_u(t, x) + (\epsilon \vee \rho_0(\epsilon)) e_u^\epsilon(t, x),$$

and

$$x - T_d^\epsilon(t, x) := J_d^\epsilon(t, x) = \epsilon j_d^\epsilon(t, x) + \epsilon^2 e_d^\epsilon(t, x),$$

with  $e_d^\epsilon(t, x)$  and  $e_u^\epsilon(t, x)$  bounded on  $E_{\delta,K}$ . Consequently, there exists a constant  $C_{\delta,K}$  such that the probability of the upward jump  $q^\epsilon$  admits the asymptotic expansion:

$$q^\epsilon(t, x) = \frac{J_d^\epsilon(t, x)}{J_u^\epsilon(t, x) + J_d^\epsilon(t, x)} = \epsilon \frac{j_d(t, x)}{j_u(t, x)} + C_{\delta,K} \epsilon (\epsilon \vee \rho_0(\epsilon)), \quad \text{for } (t, x) \in E_{\delta,K}. \quad (3.9)$$

The next lemma considers the asymptotic expansion of  $J_d^\epsilon$  in the supermartingale region.

**Lemma 3.12.** *Suppose that Assumptions 3.1, 3.3, 3.7(i) and (ii) hold.*

*For  $0 < K < \infty$ , let  $E_K := \{(t, x) : t \in [0, 1], (x_1(t) - K) \vee l_t \leq x \leq x_1(t)\}$ . Then,  $x - T_d^\epsilon(t, x) = J_d^\epsilon(t, x) = \epsilon j_d^\epsilon(t, x) + \epsilon^2 e_d^\epsilon(t, x)$ , with  $j_d^\epsilon(t, x) := \frac{1}{f(t + \epsilon, x)} \partial_t F(t, x)$ , and  $e_d^\epsilon(t, x)$  bounded on  $E_K$ .*

*Proof.* For  $J_d^\epsilon(t, x)$ , applying Taylor's expansion to the first order, we have that

$$\begin{aligned} J_d^\epsilon(t, x) &= x - T_d^\epsilon(t, x) = F^{-1}(t + \epsilon, F(t + \epsilon, x)) - F^{-1}(t + \epsilon, F(t, x)) \\ &= \frac{\partial F^{-1}}{\partial x}(t + \epsilon, F(t + \epsilon, x)) [F(t + \epsilon, x) - F(t, x)] + \frac{1}{2} \frac{\partial^2 F^{-1}}{\partial x^2}(t + \epsilon, c) [F(t + \epsilon, x) - F(t, x)]^2 \\ &= \frac{1}{f(t + \epsilon, x)} [F(t + \epsilon, x) - F(t, x)] + \frac{1}{2} (-1) \frac{f'(t + \epsilon, \xi)}{f^3(t + \epsilon, \xi)} [F(t + \epsilon, x) - F(t, x)]^2, \end{aligned}$$

where on the above,  $c$  is between  $F(t, x)$  and  $F(t + \epsilon, x)$ , and  $\xi = F^{-1}(t + \epsilon, c)$ . Applying further second-order Taylor's expansion to the first  $F(t + \epsilon, x) - F(t, x)$  and first-order Taylor's expansion to the second  $F(t + \epsilon, x) - F(t, x)$ , we have

$$\begin{aligned} J_d^\epsilon(t, x) &= \frac{1}{f(t + \epsilon, x)} \left[ \partial_t F(t, x) \epsilon + \frac{1}{2} \partial_{tt} F(\xi_1, x) \epsilon^2 \right] - \frac{1}{2} \frac{f'(t + \epsilon, \xi)}{f^3(t + \epsilon, \xi)} [\partial_t F(\xi_2, x) \epsilon]^2 \\ &= \frac{\partial_t F(t, x)}{f(t + \epsilon, x)} \epsilon + \left[ \frac{\partial_{tt} F(\xi_1, x)}{2f(t + \epsilon, x)} - \frac{f'(t + \epsilon, \xi)}{2f^3(t + \epsilon, \xi)} (\partial_t F(\xi_2, x))^2 \right] \epsilon^2, \end{aligned}$$

with  $\xi_1, \xi_2$  between  $t$  and  $t + \epsilon$ . Hence the expansion of  $J_d^\epsilon$  is valid, with

$$|e_d^\epsilon(t, x)| \leq \sup_{t \leq s \leq t + \epsilon, F^{-1}(t + \epsilon, F(t, x)) \leq \xi \leq x} \left| \frac{\partial_{tt} F(s, \xi)}{2f(s, \xi)} - \frac{f'(s, \xi)}{2f^3(s, \xi)} (\partial_t F(s, \xi))^2 \right|.$$

□

For  $t \in [0, 1]$ , extend the definition of  $j_d$  to  $(\ell_t, x_1(t))$ , by setting  $j_d(t, x) := \frac{1}{f(t, x)} \partial_t F(t, x)$ . For the main result of this section we need one additional result.

**Lemma 3.13.** *Suppose that Assumptions 3.1, 3.3 and 3.7 hold.*

- (i)  $T_u$  is strictly decreasing in  $x$  on  $(x_1(t), m_t)$ .
- (ii)  $j_d(t, x) \mathbf{1}_{\{x_1(t) < x < m_t\}}, j_u(t, x) \mathbf{1}_{\{x_1(t) < x < m_t\}}, \frac{j_d}{j_u}(t, x) \mathbf{1}_{\{x_1(t) < x < m_t\}}$  and  $j_d(t, x) \mathbf{1}_{\{x \leq x_1(t)\}}$  are all locally Lipschitz in  $(t, x) \in \{0 \leq t \leq 1, (-K \vee l_t) \leq x \leq (K \wedge r_t)\}$ .



*Proof.* (i) Differentiating both sides of (3.8) with respect to  $x \in (x_1(t), m_t)$ , we get that

$$\partial_x T_u(t, x) (x - T_u(t, x)) \partial_t f(t, T_u(t, x)) + \int_x^{T_u(t, x)} \partial_t f(t, \xi) d\xi = 0.$$

Hence

$$\partial_x T_u(t, x) = \frac{\partial_t F(t, T_u(t, x)) - \partial_t F(t, x)}{(x - T_u(t, x)) \partial_t f(t, T_u(t, x))} < 0, \quad (3.10)$$

where the last inequality follows from  $\partial_t F(t, T_u(t, x)) > \partial_t F(t, x)$ .

- (ii) Using Assumption 3.3 and repeating the arguments of the proof of [41, Lemma 3.6] (applied to the right-curtain coupling case) we get that all  $j_d(t, x) \mathbf{1}_{\{x_1(t) < x < m_t\}}$ ,  $j_u(t, x) \mathbf{1}_{\{x_1(t) < x < m_t\}}$ ,  $\frac{j_d}{j_u}(t, x) \mathbf{1}_{\{x_1(t) < x < m_t\}}$  are locally Lipschitz. It is left to verify the regularity of the function  $j_d(t, x) \mathbf{1}_{\{x \leq x_1(t)\}}$ . But Assumption 3.1 and the definition  $j_d(t, x) = \frac{1}{f(t, x)} \partial_t F(t, x)$  immediately imply the local Lipschitz property.  $\square$

In the following, we provide a characterization of the limit of the sequence  $(\mathbb{P}^n)_{n \geq 1}$ . The limiting process is a pure-jump process that is similar in spirit to the local Lévy models introduced by Carr, Geman, Madan and Yor [22]. (See also [41] for the related martingale case.)

**Theorem 3.14.** *Suppose that Assumptions 3.1, 3.3, 3.7 hold. Then  $\mathbb{P}^n \rightarrow \mathbb{P}^0$ , where  $\mathbb{P}^0$  is the unique weak solution to the SDE:*

$$X_t = X_0 + \int_0^t j_u(t, X_{s-}) (dN_s - \nu_s ds) \mathbf{1}_{\{x_1(s) < X_{s-} < m_s\}} - \int_0^t \mathbf{1}_{\{X_{s-} \leq x_1(s)\}} j_d(t, X_{s-}) ds, \quad (3.11)$$

and  $(N_s)_{0 \leq s \leq 1}$  is a unit size jump process with predictable compensated process  $(\nu_s)_{0 \leq s \leq 1}$  given by:

$$\nu_s := \frac{j_d}{j_u}(s, X_{s-}) \mathbf{1}_{\{x_1(s) < X_{s-} < m_s\}};$$

here  $x_1(t)$  is given by Lemma 3.8,  $j_u(t, x) := T_u(t, x) - x$  and  $j_d(t, x) := \frac{\partial_t F(t, x) - \partial_t F(t, T_u(t, x))}{f(t, x)}$  for  $x \in (x_1(t), m_t)$ ,  $j_d(t, x) := \frac{1}{f(t, x)} \partial_t F(t, x)$  for  $x \in (\ell_t, x_1(t)]$ , and  $T_u$  is characterized by (3.8) on  $x \in (x_1(t), m_t)$ .

*Proof.* By Proposition 3.4, the sequence of supermartingale measures  $(\mathbb{P}^n)_{n \geq 1}$  induced by the decreasing supermartingale transport plan is tight. We shall prove that any limit of  $(\mathbb{P}^n)_{n \geq 1}$  provides a weak solution to (3.11). Our strategy is identical to that of Henry-Labordère and Touzi [42, Theorem 3.7], and thus we only present the steps that are affected by the supermartingale feature of our construction.

For all  $\mathbf{x} \in \Omega := \mathbb{D}([0, 1], \mathbb{R})$  and  $\varphi \in C_b^1(\mathbb{R})$ , we define

$$\begin{aligned} M_t(\varphi, \mathbf{x}) &:= \varphi(\mathbf{x}_t) - \int_0^t j_d(s, \mathbf{x}_{s-}) D\varphi(\mathbf{x}_{s-}) \mathbf{1}_{\{x_1(s) < \mathbf{x}_{s-} < m_s\}} ds - \int_0^t [[\varphi(\mathbf{x}_{s-} + j_u(s, \mathbf{x}_{s-})) \\ &\quad - \varphi(\mathbf{x}_{s-})] \frac{j_d}{j_u}(s, \mathbf{x}_{s-})] \mathbf{1}_{\{x_1(s) < \mathbf{x}_{s-} < m_s\}} ds - \int_0^t j_d(s, \mathbf{x}_{s-}) D\varphi(\mathbf{x}_{s-}) \mathbf{1}_{\{\mathbf{x}_{s-} \leq x_1(s)\}} ds. \end{aligned}$$

The process  $M(\varphi, X)$  is progressively measurable w.r.t. the canonical filtration  $\mathbb{F}$ . For every constant  $p > 0$ , we introduce an  $\mathbb{F}$ -stopping time and the corresponding stopped canonical process

$$\tau_p := \inf\{t \geq 0 : |X_t| \geq p \text{ or } |X_{t-}| \geq p\}, \quad X_t^p := X_{t \wedge \tau_p}.$$

Denote further  $J(\mathbf{x}) := \{t > 0 : \Delta \mathbf{x}(t) \neq 0\}$ ,  $V(\mathbf{x}) := \{a > 0 : \tau_a(\mathbf{x}) < \tau_{a+}(\mathbf{x})\}$  and

$$V'(\mathbf{x}) := \{a > 0 : \tau_a(\mathbf{x}) \in J(\mathbf{x}) \text{ and } |\mathbf{x}(\tau_a(\mathbf{x}))| = a\}.$$

By extracting a subsequence, we can suppose w.l.o.g. that  $\mathbb{P}^n \rightarrow \mathbb{P}^0$  weakly. To prove that  $\mathbb{P}^0$  is a weak solution to SDE (3.11), it is enough to show that  $(M_t(\varphi, X))_{t \in [0,1]}$  is a local martingale under  $\mathbb{P}^0$  for every  $\varphi \in C_b^1(\mathbb{R})$ . Since the functions  $j_d$  and  $j_u$  are locally Lipchitz, but not necessarily uniformly bounded, as in the proof of Henry-Labordère and Touzi [42, Theorem 3.7], we need to adapt the localization technique in Jacod and Shiryaev [51].

First, since  $\mathbb{P}^n$  is induced by the Markov chain  $(X^n, \mathbb{P}^{*,n})$  for all  $n \geq 1$ , we have

$$\mathbb{E}_{t_k^n}^{\mathbb{P}^n} [\varphi(X_{t_{k+1}^n}) - \varphi(X_{t_k^n})] = \alpha_u + \alpha_d + \alpha_{\bar{d}},$$

where

$$\begin{aligned} \alpha_u &:= \mathbb{E}_{t_k^n}^{\mathbb{P}^n} [\{\varphi(X_{t_k^n} + J_u^{\epsilon_k^n}(t_k^n, X_{t_k^n})) - \varphi(X_{t_k^n})\} \frac{J_d^{\epsilon_k^n}}{J_d^{\epsilon_k^n} + J_u^{\epsilon_k^n}} \mathbf{1}_{\{x_1^{\epsilon_k^n}(t_k^n) < X_{t_k^n} < m^{\epsilon_k^n}(t_k^n)\}}], \\ \alpha_d &:= \mathbb{E}_{t_k^n}^{\mathbb{P}^n} [\{\varphi(X_{t_k^n} - J_d^{\epsilon_k^n}(t_k^n, X_{t_k^n})) - \varphi(X_{t_k^n})\} \frac{J_u^{\epsilon_k^n}}{J_d^{\epsilon_k^n} + J_u^{\epsilon_k^n}} \mathbf{1}_{\{x_1^{\epsilon_k^n}(t_k^n) < X_{t_k^n} < m^{\epsilon_k^n}(t_k^n)\}}], \\ \alpha_{\bar{d}} &:= \mathbb{E}_{t_k^n}^{\mathbb{P}^n} [\{\varphi(X_{t_k^n} - J_d^{\epsilon_k^n}(t_k^n, X_{t_k^n})) - \varphi(X_{t_k^n})\} \mathbf{1}_{\{X_{t_k^n} \leq x_1^{\epsilon_k^n}(t_k^n)\}}], \end{aligned}$$

with  $\epsilon_k^n := t_{k+1}^n - t_k^n$  (also recall that, under the present assumptions (Dispersion Assumption in particular),  $x_1^{\epsilon_k^n}(t_k^n)$  is the unique regime-switching boundary).

By the continuity of  $(\epsilon, t) \mapsto (m^\epsilon(t), x_1^\epsilon(t))$  and noticing that the density  $f(t, x) \in C_b^3(E)$ , we have that for  $\epsilon_k^n$  small and for all  $s \in [t_k^n, t_{k+1}^n]$ ,

$$\begin{aligned} &|\mathbb{E}_{t_k^n}^{\mathbb{P}^n} [\mathbf{1}_{\{x_1^{\epsilon_k^n}(t_k^n) < X_{t_k^n} < m^{\epsilon_k^n}(t_k^n)\}} - \mathbf{1}_{\{x_1(s) < X_s < m_s\}}]| \\ &\leq |\int_{x_1^{\epsilon_k^n}(t_k^n)}^{x_1(s)} f(t, x) dx| + |\int_{m^{\epsilon_k^n}(t_k^n)}^{m_s} f(t, x) dx| = O(\rho_0(\epsilon_k^n) \vee \rho_1(\epsilon_k^n)). \end{aligned}$$

Then using Lemma 3.11 and the arguments of [41, Theorem 3.7] we have that

$$\begin{aligned} \alpha_u &= \mathbb{E}_{t_k^n}^{\mathbb{P}^n} \left[ \int_{t_k^n}^{t_{k+1}^n} [\varphi(X_s + j_u(s, X_s)) - \varphi(X_s)] \frac{j_d}{j_u}(s, X_s) \mathbf{1}_{\{x_1(s) < X_s < m_s\}} ds \right], \\ \alpha_d &= -\mathbb{E}_{t_k^n}^{\mathbb{P}^n} \left[ \int_{t_k^n}^{t_{k+1}^n} j_d(s, X_s) D\varphi(X_s) \mathbf{1}_{\{x_1(s) < X_s < m_s\}} ds \right] + O(\epsilon_k^n(\epsilon_k^n \vee \rho_0(\epsilon_k^n) \vee \rho_1(\epsilon_k^n))). \end{aligned}$$

Similarly, using Lemma 3.12 we have that

$$\alpha_{\bar{d}} = -\mathbb{E}_{t_k^n}^{\mathbb{P}^n} \left[ \int_{t_k^n}^{t_{k+1}^n} j_d(s, X_s) D\varphi(X_s) \mathbf{1}_{\{X_s \leq x_1(s)\}} ds \right] + O(\epsilon_k^n(\epsilon_k^n \vee \rho_0(\epsilon_k^n) \vee \rho_1(\epsilon_k^n))).$$

Now let  $0 \leq s < t \leq 1$ ,  $p \in \mathbb{N}$ ,  $\phi_s(X_\cdot)$  be a  $\mathcal{F}_s$ -measurable bounded random variable on  $\Omega$  such that  $\phi : \Omega \rightarrow \mathbb{R}$  is continuous under the Skorokhod topology. Then

$$\mathbb{E}^{\mathbb{P}^n} [\phi_s(X_\cdot) (M_{t \wedge \tau_p}(\varphi, X) - M_{s \wedge \tau_p}(\varphi, X))] \leq C_p (|\pi_n| \vee \rho_0(|\pi_n|) \vee \rho_1(|\pi_n|)).$$

for some constant  $C_p > 0$ . The remaining arguments are those of [41, Theorem 3.7], and thus omitted.  $\square$

### 3.3 Optimality of the local Lévy process

In this section, we establish the optimality of the previously obtained limiting process to the continuous-time primal problem  $\mathbf{P}_\infty$ .

We will need the optimal dual strategy (in discrete time one-period case) which is given in Lemma B.5; in the following, we denote  $h^\epsilon := \hat{h}^{\mu_t, \mu_{t+\epsilon}}, \varphi^\epsilon := \hat{\varphi}^{\mu_t, \mu_{t+\epsilon}}, \psi^\epsilon := \hat{\psi}^{\mu_t, \mu_{t+\epsilon}}$ .

Now let us introduce the optimal dual strategies for the continuous-time transport problem. First, we define the dynamic strategy  $h^*$  by

$$\begin{aligned} h^*(t, x) &= 0, \quad \text{when } x \leq x_1(t), \\ \partial_x h^*(t, x) &= \frac{c_x(x, T_u(t, x)) - c_x(x, x)}{j_u(t, x)}, \quad x_1(t) < x < m_t, \\ \lim_{x \downarrow x_1(t)} h^*(t, x) &= 0, \\ h^*(t, x) &= h^*(t, T_u^{-1}(t, x)) - c_y(T_u^{-1}(t, x), x), \quad \text{when } x \geq m_t. \end{aligned}$$

We remark that from the above definition, for all  $t \in [0, 1]$ ,  $x \mapsto h^*(t, x)$  is continuous both at  $x_1(t)$  and  $m_t$ . In addition,  $\psi^*$  is defined, up to a constant, by

$$\partial_x \psi^*(t, x) := -h^*(t, x), \quad (t, x) \in E = \{(t, x) : t \in [0, 1], x \in (l_t, r_t)\}.$$

Now using the same arguments as [41, Corollary 3.10] together with the fact that  $h^*(t, x) = 0$  for  $t \in [0, 1]$  and  $x \leq x_1(t)$ , we have that  $\psi^* \in C^{1,1}(E)$ .

For the (static part of the) dual limiting strategy, let  $\gamma^*(dt) := \delta_{\{0\}}(dt) + \delta_{\{1\}}(dt) + \text{Leb}(dt)$  be a finite measure on  $[0, 1]$ , where  $\text{Leb}$  denotes the Lebesgue measure on  $[0, 1]$ .

One further defines  $\lambda_0^*$  and  $\bar{\lambda}_0^*$  by  $\lambda_0^*(0, x) := \psi^*(0, x)$ ,  $\lambda_0^*(1, x) := \psi^*(1, x)$ ,  $\bar{\lambda}_0^*(0, x) := |\psi^*(0, x)|$ , and  $\bar{\lambda}_0^*(1, x) := \sup_{t \in [0, 1]} |\psi^*(t, x)|$ ; and for all  $(t, x) \in (0, 1) \times \mathbb{R}$ ,

$$\begin{aligned} \lambda_0^*(t, x) &:= \partial_t \psi^*(t, x) + \mathbf{1}_{\{x_1(t) < x \leq m_t\}} (\partial_x \psi^* j_d + \nu[\psi^* - \psi^*(\cdot, T_u) + c(\cdot, T_u)]) \\ &\quad + \mathbf{1}_{\{x \leq x_1(t)\}} \partial_x \psi^* j_d, \end{aligned}$$

$$\begin{aligned} \bar{\lambda}_0^* &:= |\partial_t \psi^*(t, x) + \mathbf{1}_{\{x_1(t) < x \leq m_t\}} (\partial_x \psi^* j_d + \nu[\psi^* - \psi^*(\cdot, T_u)])| + \mathbf{1}_{\{x_1(t) < x < m_t\}} \nu[|c(\cdot, T_u)|] \\ &\quad + \mathbf{1}_{\{x \leq x_1(t)\}} |\partial_x \psi^* j_d|. \end{aligned}$$

Finally we define  $\lambda^*(x, dt) := \lambda_0^*(t, x) \gamma^*(dt)$  and  $\bar{\lambda}^*(x, dt) := \bar{\lambda}_0^*(t, x) \gamma^*(dt)$ .

The next result shows that the limiting process of the previous sections solves the continuous-time supermartingale optimal transport problem. Let  $H^*$  be the  $\mathbb{F}$ -predictable process on  $\Omega$  defined by

$$H_t^* := h^*(t, X_{t-}), \quad t \in [0, 1].$$

**Theorem 3.15.** *Suppose that Assumptions 2.2, 3.1, 3.3, 3.7 hold, and that*

$$\mu(\bar{\lambda}^*) = \int_0^1 \int_{\mathbb{R}} \bar{\lambda}_0^*(t, x) \mu_t(dx) \gamma^*(dt) < \infty. \quad (3.12)$$

*Then the supermartingale transport problem (2.4) is solved by the local Lévy process (3.11). Moreover, we have that  $(H^*, \lambda^*) \in \mathcal{D}_\infty(\mu)$  and the following duality is valid:*

$$\mathbb{E}^{\mathbb{P}^0}[C(X_\cdot)] = \mathbf{P}_\infty(\mu) = \mathbf{D}_\infty(\mu) = \mu(\lambda^*),$$

*and the optimal value is given by*

$$\mu(\lambda^*) = \int_0^1 \int_{x_1(t)}^{m_t} \frac{j_d(t, x)}{j_u(t, x)} c(x, x + j_u(t, x)) f(t, x) dx dt.$$

In order to prove Theorem 3.15, we consider the following partition of  $[0, 1]$ :  $\pi_n = (t_k^n)_{0 \leq k \leq n}$ , where  $t_k^n := k\epsilon$  with  $\epsilon = \frac{1}{n}$ . In the following, we shall simplify the notation  $t_k^n$  to  $t_k$ . Remind that under every  $\mathbb{P}^n$ , we have the following super-replication inequality:

$$\sum_{k=0}^{n-1} (\varphi^\epsilon(t_k, X_{t_k}) + \psi^\epsilon(t_k, X_{t_{k+1}})) + \sum_{k=0}^{n-1} h^\epsilon(t_k, X_{t_k})(X_{t_{k+1}} - X_{t_k}) \geq \sum_{k=0}^{n-1} c(X_{t_k}, X_{t_{k+1}}). \quad (3.13)$$

We further define  $\Psi^* : \Omega \rightarrow \mathbb{R}$  by

$$\begin{aligned} \Psi^*(\mathbf{x}) &:= \psi^*(1, \mathbf{x}_1) - \psi^*(0, \mathbf{x}_0) - \int_0^1 (\partial_t \psi^*(t, \mathbf{x}_t) + j_d(t, \mathbf{x}_t) \mathbf{1}_{\{\mathbf{x}_t < m_t\}} \partial_x \psi^*(t, \mathbf{x}_t)) dt \\ &+ \int_0^1 \frac{j_d(t, \mathbf{x}_t)}{j_u(t, \mathbf{x}_t)} \mathbf{1}_{\{x_1(t) < \mathbf{x}_t < m_t\}} (\psi^*(t, \mathbf{x}_t) - \psi^*(t, \mathbf{x}_t + j_u(t, \mathbf{x}_t)) + c(\mathbf{x}_t, \mathbf{x}_t + j_u(t, \mathbf{x}_t))) dt. \end{aligned}$$

To prove Theorem 3.15, we need several auxiliary results. The proof of the following lemma relies on Lemmas 3.11 and 3.12, but otherwise is the same as the proof of [41, Lemma 6.7], and thus we omit the details.

**Lemma 3.16.** *Under Assumptions 2.2, 3.1, 3.3, we have that*

$$T_u^\epsilon \mathbf{1}_{\{x_1^\epsilon(t) < x < m_t^\epsilon\}} \rightarrow T_u \mathbf{1}_{\{x_1(t) < x < m_t\}}, \quad h^\epsilon \rightarrow h^*, \quad \partial_t \psi^\epsilon \rightarrow \partial_t \psi^*, \quad \psi^\epsilon \rightarrow \psi^*$$

locally uniformly on  $E = \{(t, x) : t \in [0, 1], x \in (l_t, r_t)\}$ .

**Lemma 3.17.** *Suppose Assumptions 2.2, 3.1, 3.3 hold. Then for every càdlàg path  $\mathbf{x} \in \mathbb{D}([0, 1], \mathbb{R})$  taking value in  $(l_1, r_1)$ , we have*

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} (\varphi^\epsilon(t_k, \mathbf{x}_{t_k}) + \psi^\epsilon(t_k, \mathbf{x}_{t_{k+1}})) \rightarrow \Psi^*(\mathbf{x}), \quad \text{as } \epsilon \rightarrow 0.$$

*Proof.* The arguments are almost identical to those of [41, Lemma 6.9], and thus we only briefly sketch them. First note that

$$\begin{aligned} \sum_{k=0}^{n-1} (\varphi^\epsilon(t_k, \mathbf{x}_{t_k}) + \psi^\epsilon(t_k, \mathbf{x}_{t_{k+1}})) &= \sum_{k=0}^{n-1} (\varphi^\epsilon(t_k, \mathbf{x}_{t_k}) + \psi^\epsilon(t_k, \mathbf{x}_{t_k})) + \psi^\epsilon(t_{n-1}, \mathbf{x}_1) \\ &+ \sum_{k=1}^{n-1} (\psi^\epsilon(t_{k-1}, \mathbf{x}_{t_k}) - \psi^\epsilon(t_k, \mathbf{x}_{t_k})) - \psi^\epsilon(0, \mathbf{x}_0). \end{aligned}$$

Then by Lemma 3.16 and the arguments of [41, Lemma 6.9], as  $\epsilon \rightarrow 0$ , we have that  $\psi^\epsilon(t_{n-1}, \mathbf{x}_1) \rightarrow \psi^*(1, \mathbf{x}_1)$ ,

$$\sum_{k=1}^{n-1} (\psi^\epsilon(t_{k-1}, \mathbf{x}_{t_k}) - \psi^\epsilon(t_k, \mathbf{x}_{t_k})) \rightarrow - \int_0^1 \partial_t \psi^*(t, \mathbf{x}_t) dt,$$

and  $\sum_{k=0}^{n-1} (\varphi^\epsilon(t_k, \mathbf{x}_{t_k}) + \psi^\epsilon(t_k, \mathbf{x}_{t_k})) \mathbf{1}_{\{x_1^\epsilon(t_k) < \mathbf{x}_{t_k} < m^\epsilon(t_k)\}}$  converges to

$$\begin{aligned} &\int_0^1 -\partial_x \psi^*(t, \mathbf{x}_t) j_d(t, \mathbf{x}_t) dt \\ &+ \int_0^1 \frac{j_d(t, \mathbf{x}_t)}{j_u(t, \mathbf{x}_t)} \mathbf{1}_{\{x_1(t) < \mathbf{x}_t < m_t\}} (\psi^*(t, \mathbf{x}_t) - \psi^*(t, \mathbf{x}_t + j_u(t, \mathbf{x}_t)) + c(\mathbf{x}_t, \mathbf{x}_t + j_u(t, \mathbf{x}_t))) dt. \end{aligned}$$

Finally, using similar arguments, and noting that  $c(x, x) = 0$  for all  $x \in \mathbb{R}$ , we have that

$$\sum_{k=0}^{n-1} (\varphi^\epsilon(t_k, \mathbf{x}_{t_k}) + \psi^\epsilon(t_k, \mathbf{x}_{t_k})) \mathbf{1}_{\{\mathbf{x}_{t_k} \leq x_1^\epsilon(t_k)\}} \rightarrow - \int_0^1 j_d(t, \mathbf{x}_t) \mathbf{1}_{\{\mathbf{x}_t \leq x_1(t)\}} \partial_x \psi^*(t, \mathbf{x}_t) dt,$$

which concludes the proof.  $\square$

**Lemma 3.18.** *Suppose that Assumptions 2.2, 3.1, 3.3, 3.7 hold and  $\mu(\bar{\lambda}^*) < \infty$ . Then for the probability measure  $\mathbb{P}^0$  given in Theorem 3.14, we have that*

$$\mathbb{E}^{\mathbb{P}^0}[C(X)] = \mathbb{E}^{\mathbb{P}^0}[\Psi^*(X)] = \mu(\lambda^*) = \int_0^1 \int_{x_1(t)}^{m_t} \frac{j_d(t, x)}{j_u(t, x)} c(x, x + j_u(t, x)) f(t, x) dx dt.$$

*Proof.* The arguments are almost identical to those of [41, Lemma 6.10] (adapted to the right-curtain coupling situation). We only need to adapt various definitions to include the supermartingale regions governed by  $t \mapsto x_1(t)$ .

By Itô's formula, the following process is a local martingale:

$$\begin{aligned} & \psi^*(t, X_t) - \psi^*(0, X_0) - \int_0^1 (\partial_t \psi^*(t, X_t) + (j_d(t, X_t) \mathbf{1}_{\{x_1(t) < X_t < m_t\}} + j_d(t, X_t) \mathbf{1}_{\{X_t \leq x_1(t)\}}) \\ & \partial_x \psi^*(t, X_t)) dt + \int_0^1 \frac{j_d(t, x_t)}{j_u(t, x_t)} \mathbf{1}_{\{x_1(t) < X_t < m_t\}} (\psi^*(t, X_t) - \psi^*(t, X_t + j_u(t, X_t))) dt. \end{aligned}$$

Moreover, using the fact that  $\mu(\bar{\lambda}^*) < \infty$  together with the dominated convergence theorem, we have that

$$\mathbb{E}^{\mathbb{P}^0}[\Psi^*(X)] = \int_0^1 \int_{x_1(t)}^{m_t} \frac{j_d(t, x)}{j_u(t, x)} c(x, x + j_u(t, x)) f(t, x) dx dt.$$

To compute  $\mathbb{E}^{\mathbb{P}^0}[C(X)]$ , noticing that  $[X]_t^c = 0$ ,  $\mathbb{P}^0$ -a.s., and the process

$$Y_t := \sum_{s \leq t} |c(X_{s-}, X_s)| - \int_0^t |c(X_{s-}, X_{s-} + j_u(t, X_{s-}))| \frac{j_d(t, X_{s-})}{j_u(t, X_{s-})} \mathbf{1}_{\{x_1(t) < X_t < m_t\}} dt$$

is a local martingale. The remainder of the proof follows the arguments of [42, Lemma 6.10].  $\square$

In the following lemma, we establish the continuous limit of the dynamic part. The proof relies on Lemma 3.16, but otherwise is identical to the proof of [41, Lemma 6.11], and thus omitted.

**Lemma 3.19.** *Suppose Assumptions 3.1, 3.3, 3.7 hold. Then the following convergence in probability (as  $\epsilon \rightarrow 0$ ) holds under every supermartingale measure  $\mathbb{P} \in \mathcal{S}_\infty$ :*

$$\sum_{k=1}^{n-1} h^\epsilon(t_k, X)(X_{t_{k+1}} - X_{t_k}) \rightarrow \int_0^1 h^*(t, X_{t-}) dX_t.$$

Finally, we provide the proof of Theorem 3.15.

*Proof of Theorem 3.15.* Using the discrete superhedging duality (3.13), together with the convergence results Lemma 2.3, 3.17, 3.19, we have that under every  $\mathbb{P} \in \mathcal{S}_\infty$ ,  $(\Psi^*, h^*)$  superhedges the continuous-time cost  $C(X)$ :

$$\Psi^*(X) + \int_0^1 h^*(t, X_{t-}) dX_t \geq \int_0^1 \frac{1}{2} c_{yy}(X_t, X_t) d[X]_t^c + \sum_{0 \leq t \leq 1} c(X_{t-}, X_t), \quad \mathbb{P}\text{-a.s.}$$

By the weak duality,

$$\mathbb{E}^{\mathbb{P}^0}[C(X)] \leq \mathbf{P}_\infty(\mu) \leq \mathbf{D}_\infty(\mu) \leq \mu(\lambda^*).$$

Then Lemma 3.18 concludes the proof.  $\square$

## 4 Examples

In this section, we will consider three examples of supermartingales matching a given PCOCD. In contrary to the martingale case of [52, 41] (that covers left-curtain and right-curtain couplings), in the supermartingale context, the increasing and decreasing couplings provide significantly different dynamics in continuous-time.

### 4.1 Uniform distribution with bounded support

In this section the family of marginal distributions  $(\mu_t)_{t \in [0,1]}$  satisfies

$$\mu_t = \frac{1}{e^t + e^{2t}} \lambda_{[-e^{2t}, e^t]}, \quad t \in [0, 1]. \quad (4.14)$$

(This example was also considered in the martingale case by Juillet [52].) Our goal is to explicitly determine the quantities that define the SDE given in Theorem 3.14.

Note that  $\ell_t = -e^{2t}$  and  $r_t = e^t$  for all  $t \in [0, 1]$ . Furthermore,  $(\mu_t)_{t \in [0,1]}$  satisfies Assumption 3.1 with  $m_\epsilon(t) = \ell_t$  and  $m^\epsilon(t) = r_t$ , for each  $t \in [0, 1)$  and  $\epsilon \in (0, 1 - t]$ . Recall that  $m_\epsilon, m^\epsilon$  maximizes and minimizes, respectively, the function  $x \mapsto F(t + \epsilon, x) - F(t, x)$ . Observe that  $m_\epsilon, m^\epsilon$  are independent of  $\epsilon > 0$ . By direct computation we have that, for all  $t \in (0, 1)$ ,  $\partial_t F(t, \cdot)$  is minimized and maximized at  $m_t = e^t$  and  $\tilde{m}_t = -e^{2t}$ . Then  $m^0(t) := m_t = \lim_{\epsilon \downarrow 0} m^\epsilon(t) = e^t$  for all  $t \in [0, 1]$ . It is now easy to see that  $(\mu_t)_{t \in [0,1]}$  satisfies Assumption 3.3. However, since  $t \mapsto r_t$  is not constant and  $m_t = r_t$ , Assumption 3.7 does not hold. For this reason, most of the results of Section 3.2 (that lead to Theorem 3.14) cannot be applied, and thus we will prove the relevant statements by direct calculations.

Now fix  $t \in [0, 1)$  and  $\epsilon \in (0, 1 - t]$ , and consider the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  of  $\mu_t$  and  $\mu_{t+\epsilon}$  (see Appendix B.1). Since, Assumption 3.1 is satisfied, the phase transition point, that separates the martingale and supermartingale regions, is unique; see Lemma B.3. The transition threshold is denoted by

$$x_1^\epsilon(t) := x_1^{\mu_t, \mu_{t+\epsilon}},$$

see (B.50).

**Lemma 4.1.** *If  $(\mu_t)_{t \in [0,1]}$  is given by (4.14), then the phase transition  $x_1^\epsilon(t)$  is given by*

$$x_1^\epsilon(t) = \frac{e^t(e^{2(t+\epsilon)} - e^{t+\epsilon} + e^t) + e^{2t}(e^t - 2e^{t+\epsilon})}{e^{t+\epsilon} + e^{2(t+\epsilon)} - e^t - e^{2t}}. \quad (4.15)$$

Furthermore, the limit  $x_1(t) := \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$  exists and is given by

$$x_1(t) = \frac{-e^t}{1 + 2e^t} \in (-e^{2t}, e^t), \quad t \in [0, 1]. \quad (4.16)$$

*Proof.* To determine  $x_1^\epsilon(t) \in (\ell_t, r_t)$  we use the fact that under the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}, \mu|_{[x_1^\epsilon(t), r_t]}$  is mapped to  $\nu|_{[y_1^\epsilon(t), r_{t+\epsilon}]}$  in a martingale way, where  $y_1^\epsilon(t) = y_1^{\mu_t, \mu_{t+\epsilon}}$  (see (B.51)). In particular, the pair  $(x_1^\epsilon(t), y_1^\epsilon(t))$  is uniquely determined by the mass and mean preservation condition

$$\int_{x_1^\epsilon(t)}^{r_t} z^i \mu_t(dz) = \int_{y_1^\epsilon(t)}^{r_{t+\epsilon}} z^i \mu_{t+\epsilon}(dz), \quad i = 0, 1. \quad (4.17)$$

Direct computation leads to

$$\frac{e^t - x_1^\epsilon(t)}{e^t + e^{2t}} = \frac{e^{t+\epsilon} - y_1^\epsilon(t)}{e^{t+\epsilon} + e^{2(t+\epsilon)}}, \quad \frac{e^{2t} - (x_1^\epsilon(t))^2}{e^t + e^{2t}} = \frac{e^{2(t+\epsilon)} - (y_1^\epsilon(t))^2}{e^{t+\epsilon} + e^{2(t+\epsilon)}}.$$



Solving the above equations gives

$$x_1^\epsilon(t) = \frac{e^t(e^{2(t+\epsilon)} - e^{t+\epsilon} + e^t) + e^{2t}(e^t - 2e^{t+\epsilon})}{e^{t+\epsilon} + e^{2(t+\epsilon)} - e^t - e^{2t}}, \quad y_1^\epsilon(t) = x_1^\epsilon(t) - e^{t+\epsilon} + e^t.$$

By letting  $\epsilon \downarrow 0$ , direct calculation (using the L'Hôpital's rule) gives that  $x_1(t) = \frac{-e^t}{1+2e^t}$  (note that  $\lim_{\epsilon \downarrow 0} y_1^\epsilon(t) = x_1(t)$ ), which concludes the proof.  $\square$

Even though not all the Assumptions 3.1, 3.3, 3.7 are satisfied in the case of (4.14), the following result shows that the statement of Theorem 3.14 remains true.

**Proposition 4.2.** *If  $(\mu_t)_{t \in [0,1]}$  is given by (4.14), then the statement of Theorem 3.14 holds, and the corresponding SDE is explicitly given by*

$$X_t = X_0 + \int_0^t \mathbf{1}_{\{X_{s-} > x_1(s)\}} j_u(s, X_{s-}) (dN_s - \nu_s ds) - \int_0^t \mathbf{1}_{\{X_{s-} \leq x_1(s)\}} \tilde{\nu}_s ds,$$

where  $(N_s)_{0 \leq s \leq 1}$  is a unit size jump process with predictable compensated process  $(\nu_s)_{0 \leq s \leq 1}$  given by  $\nu_s := \frac{j_d}{j_u}(s, X_{s-})$ , with  $j_d(s, x) = \frac{1}{2} \frac{e^s - x}{1+e^s} (1 + 2e^s)$ ,  $j_u(s, x) = e^s - x$ , and  $\tilde{\nu}_s := \frac{e^{2s} - x(1+2e^s)}{1+e^s}$ .

In addition, if Assumption 2.2 and the integrability condition (3.12) are satisfied, then the above process is the optimal solution for the primal problem (2.4) and the conclusion of Theorem 3.15 holds.

*Proof.* Since, for  $t \in [0, 1]$ , we have that  $m_t = r_t$ , it follows that  $\mathbf{1}_{\{x_1(t) < X_{t-} < m_t\}} = \mathbf{1}_{\{X_{t-} > x_1(t)\}}$  a.s. We now determine  $(j_d, j_u)$ .

For  $t \in [0, 1)$  and  $\epsilon \in (0, 1 - t]$ , we consider the one-period decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$ , which is supported on the graphs of two functions  $T_u^\epsilon(t, \cdot) = T_u^{\mu_t, \mu_{t+\epsilon}}(\cdot)$ ,  $T_d^\epsilon(t, \cdot) = T_d^{\mu_t, \mu_{t+\epsilon}}(\cdot)$ ; see Appendix B.1. On  $(x_1^\epsilon(t), m^\epsilon(t) = r_t)$ ,  $(T_d^\epsilon(t, \cdot), T_u^\epsilon(t, \cdot))$  can be determined from the mass and mean preservation condition

$$\int_x^{T_u^\epsilon(t, x)} z^i \mu_t(dz) = \int_{T_d^\epsilon(t, x)}^{T_u^\epsilon(t, x)} z^i \mu_{t+\epsilon}(dz), \quad i = 0, 1.$$

Note that, since  $T_u^\epsilon(t, x) \geq m^\epsilon(t) = r_t$  for  $x \leq r_t$ , we have that  $\int_x^{T_u^\epsilon(t, x)} z^i \mu_t(dz) = \int_x^{r_t} z^i \mu_t(dz)$  for  $i = 0, 1$ . Then direct calculation gives that

$$\frac{e^t - x}{e^t + e^{2t}} = \frac{T_u^\epsilon(t, x) - T_d^\epsilon(t, x)}{e^{t+\epsilon} + e^{2(t+\epsilon)}}, \quad T_u^\epsilon(t, x) = x + e^t - T_d^\epsilon(t, x), \quad x \in (x_1^\epsilon(t), e^t).$$

It follows that

$$x - T_d^\epsilon(t, x) = \frac{1}{2} \frac{e^t - x}{e^t + e^{2t}} \left[ e^{t+\epsilon} - e^t + e^{2(t+\epsilon)} - e^{2t} \right], \quad x \in (x_1^\epsilon(t), e^t).$$

This further leads to

$$j_d(t, x) := \lim_{\epsilon \downarrow 0} j_d^\epsilon(t, x) := \lim_{\epsilon \downarrow 0} \frac{x - T_d^\epsilon(t, x)}{\epsilon} = \frac{1}{2} \frac{e^t - x}{1 + e^t} (1 + 2e^t), \quad x \in (x_1(t), e^t),$$

$$j_u(t, x) := \lim_{\epsilon \downarrow 0} (T_u^\epsilon(t, x) - x) = e^t - x, \quad x \in (x_1(t), e^t).$$

Then also

$$\frac{j_d}{j_u}(t, x) = \frac{1}{2} \frac{1 + 2e^t}{1 + e^t}, \quad x \in (x_1(t), e^t).$$

If we define  $T_u$  by  $T_u(t, x) := \lim_{\epsilon \downarrow 0} T_u^\epsilon(t, x)$ , then  $T_u(t, x) = e^t$  for  $x \in (x_1(t), e^t)$ , which is clearly non-decreasing and continuous.

The definition of  $j_d$  on  $[0, 1) \times (\ell_t, x_1(t))$  remains the same:

$$j_d(t, x) := \lim_{\epsilon \downarrow 0} j_d^\epsilon(t, x) = \lim_{\epsilon \downarrow 0} \frac{x - T_d^\epsilon(t, x)}{\epsilon}, \quad x \in (-e^{2t}, x_1(t)).$$

The explicit form is recovered by using the fact that  $T_d^\epsilon(t, x) = F^{-1}(t + \epsilon, F(t, x))$  for  $x \leq x_1^\epsilon(t)$ . Direct computation gives that

$$j_d(t, x) = \frac{e^{2t} - x(1 + 2e^t)}{1 + e^t}, \quad x \in (-e^{2t}, x_1(t)).$$

It follows that the statements of Lemmas 3.11, 3.12, 3.13(ii) (and consequently Theorem 3.14) hold, which proves the first part of the present proposition.

The second claim is established by following the arguments of the proof of Theorem 3.15. In particular we need continuity of  $T_u$ , Lemmas 3.13(ii) and 3.16. However, these results (by direct computation) also hold in the present setting,  $\square$

**Increasing supermartingale coupling** Let  $(\mu_t)_{t \in [0, 1]}$  be as in (4.14). Fix  $t \in [0, 1)$  and  $\epsilon \in (0, 1 - t]$ , and denote by  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  the increasing supermartingale coupling of  $\mu_t$  and  $\mu_{t+\epsilon}$  (see Nutz and Stebegg [57] and Bayraktar et al. [4]).

The main difference between  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  and the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  is that  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  is constructed by working from left to right and mapping  $\mu_t|_{(-\infty, x]}$  to the shadow measure  $S^{\mu_{t+\epsilon}}(\mu_t|_{(-\infty, x]})$ , for each  $x \in \mathbb{R}$ . The construction of  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  is similar, but works from right to left and thus considers the measures  $(\mu_t|_{[x, \infty)})_{x \in \mathbb{R}}$  and their corresponding ‘shadows’ in  $\mu_{t+\epsilon}$ .

The main feature of  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  is that there exists the unique point  $x_1^\epsilon(t) \in (\ell_t, r_t)$  such that  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  is a martingale on  $(-\infty, x_1^\epsilon(t)] \times \mathbb{R}$  and strict supermartingale elsewhere. In particular,

- on  $(-\infty, x_1^\epsilon(t)] \times \mathbb{R}$ ,  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  coincides with the left-curtain martingale coupling of  $\mu_t|_{(-\infty, x_1^\epsilon(t)]}$  and  $S^{\mu_{t+\epsilon}}(\mu_t|_{(-\infty, x_1^\epsilon(t)]})$ ;
- the support of  $(\mu_t - \mu_t|_{(-\infty, x_1^\epsilon(t)]})$  is strictly to the right of the support of  $(\mu_{t+\epsilon} - S^{\mu_{t+\epsilon}}(\mu_t|_{(-\infty, x_1^\epsilon(t)]}))$ ;
- on  $(x_1^\epsilon(t), \infty) \times \mathbb{R}$ ,  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  coincides with the *antitone* coupling of  $(\mu_t - \mu_t|_{(-\infty, x_1^\epsilon(t)]})$  and  $(\mu_{t+\epsilon} - S^{\mu_{t+\epsilon}}(\mu_t|_{(-\infty, x_1^\epsilon(t)]}))$ ;

see Bayraktar et al. [4] for details. (In fact, these properties hold for any measures  $\mu \leq_{cd} \nu$ , and not necessarily  $\mu_t \leq_{cd} \mu_{t+\epsilon}$  as in (4.14).) The left-curtain martingale coupling was introduced by Beiglböck and Juillet [12] (it is the symmetric counterpart of the right-curtain martingale coupling; see Appendix A.2). The aforementioned antitone coupling of two measures  $\eta, \chi$  (with  $\eta(\mathbb{R}) = \chi(\mathbb{R})$  and atom-less  $\eta$ ) is denoted by  $\pi^{a, \eta, \chi}$  and given by

$$\pi^{a, \eta, \chi}(dx, dy) = \eta(dx) \delta_{F_\chi^{-1}(\chi(\mathbb{R}) - F_\eta(x))}(dy).$$

Note that  $\pi^{a, \eta, \chi}$  is supported on a graph of a decreasing function  $x \mapsto F_\chi^{-1}(\chi(\mathbb{R}) - F_\eta(x))$ .

**Lemma 4.3.** *If  $(\mu_t)_{t \in [0, 1]}$  is defined by (4.14), then the phase transition  $x_1^\epsilon(t)$  is given by*

$$x_1^\epsilon(t) = \frac{e^{3t}(1 - e^{2\epsilon}) + e^t(2e^\epsilon + e^t(1 + e^\epsilon))}{1 + e^\epsilon + e^t(1 + e^{2\epsilon})}, \quad t \in [0, 1), \quad \epsilon \in (0, 1 - t]. \quad (4.18)$$

*In particular, under  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$ ,  $\mu_t|_{(\ell_t, x_1^\epsilon(t))}$  is mapped to  $\mu_{t+\epsilon}|_{[y_1^\epsilon(t), r_t]}$  in a martingale way, where*

$$y_1^\epsilon(t) = x_1^\epsilon(t) - e^{t+\epsilon} - e^{2t},$$

while  $\mu_t|_{(x_1^\epsilon(t), \infty)}$  is mapped to  $\mu_{t+\epsilon}|_{(-\infty, y_1^\epsilon(t))}$  in a strict supermartingale way.

Furthermore, the limit  $x_1(t) := \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$  exists and is given by

$$x_1(t) = e^t = r_t, \quad t \in [0, 1]. \quad (4.19)$$

*Proof.* To determine  $x_1^\epsilon(t) \in (\ell_t, r_t)$  we use the fact that  $x_1^\epsilon(t)$  is the largest  $x \in (\ell_t, r_t)$  for which  $\mu_t|_{(\ell_t, x]} \leq_c S^{\mu_{t+\epsilon}}(\mu_t|_{(\ell_t, x]})$  and that  $r_{S^{\mu_{t+\epsilon}}(\mu_t|_{(\ell_t, x_1^\epsilon(t))})} = \sup\{k \in \text{supp}(S^{\mu_{t+\epsilon}}(\mu_t|_{(\ell_t, x_1^\epsilon(t))}))\} = r_t$ . Using that  $(\mu_t)_{t \in [0, 1]}$  are given by (4.14) and to properties of the shadow measures, it follows that  $\mu_t|_{(\ell_t, x_1^\epsilon(t))} \leq_c S^{\mu_{t+\epsilon}}(\mu_t|_{(\ell_t, x_1^\epsilon(t))}) = \mu_{t+\epsilon}|_{(y_1^\epsilon(t), r_{t+\epsilon})}$  for some (unique)  $y_1^\epsilon(t) \in (\ell_{t+\epsilon}, \ell_t)$ . In particular, the pair  $(x_1^\epsilon(t), y_1^\epsilon(t))$  is uniquely characterized by the mass and mean preservation condition

$$\int_{\ell_t}^{x_1^\epsilon(t)} z^i \mu_t(dz) = \int_{y_1^\epsilon(t)}^{r_{t+\epsilon}} z^i \mu_{t+\epsilon}(dz), \quad i = 0, 1. \quad (4.20)$$

Direct computation leads to

$$\frac{x_1^\epsilon(t) + e^{2t}}{e^t + e^{2t}} = \frac{e^{t+\epsilon} - y_1^\epsilon(t)}{e^{t+\epsilon} + e^{2(t+\epsilon)}}, \quad \frac{(x_1^\epsilon(t))^2 - e^{4t}}{e^t + e^{2t}} = \frac{e^{2(t+\epsilon)} - (y_1^\epsilon(t))^2}{e^{t+\epsilon} + e^{2(t+\epsilon)}}.$$

Solving the above equations gives

$$x_1^\epsilon(t) = \frac{e^{3t}(1 - e^{2\epsilon}) + e^t(2e^\epsilon + e^t(1 + e^\epsilon))}{1 + e^\epsilon + e^t(1 + e^{2\epsilon})}, \quad y_1^\epsilon(t) = x_1^\epsilon(t) - e^{t+\epsilon} - e^{2t}.$$

Letting  $\epsilon \downarrow 0$  gives that  $x_1(t) = e^t$  and  $y_1^\epsilon(t) = -e^{2t}$ , which concludes the proof.  $\square$

It is well-known (see Henry-Labordère and Touzi [42]) that the left-curtain martingale coupling (of two measures in convex order) is supported on the graphs of two functions that can be computed explicitly. The following gives an explicit representation of  $\mathbb{P}^{\mu_t, \mu_{t+\epsilon}}$  in terms of the supporting functions:

**Lemma 4.4.** *Let  $(\mu_t)_{t \in [0, 1]}$  be given by (4.14). Fix  $t \in [0, 1)$  and  $\epsilon \in (0, 1 - t]$ . Let  $x_1^\epsilon(t)$  be as in Lemma 4.3.*

*Then  $\mathbb{P}^{\mu_t, \mu_{t+\epsilon}}$  is given by*

$$\begin{aligned} \bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}(dx, dy) &= \mu(dx) \mathbf{1}_{[x_1^\epsilon(t), e^t)}(x) \delta_{\bar{T}_d^\epsilon(t, x)}(dy) \\ &\quad + \mu(dx) \mathbf{1}_{(-e^{2t}, x_1^\epsilon(t))}(x) \left\{ \frac{\bar{T}_u^\epsilon(t, x) - x}{\bar{T}_u^\epsilon(t, x) - \bar{T}_d^\epsilon(t, x)} \delta_{\bar{T}_d^\epsilon(t, x)}(dy) + \frac{x - \bar{T}_d^\epsilon(t, x)}{\bar{T}_u^\epsilon(t, x) - \bar{T}_d^\epsilon(t, x)} \delta_{\bar{T}_u^\epsilon(t, x)}(dy) \right\}, \end{aligned}$$

where

$$\begin{aligned} T_u^\epsilon(t, x) &= \frac{1}{2} \left\{ \frac{e^\epsilon(1 + e^{t+\epsilon})(x + e^{2t})}{1 + e^t} + x - e^{2t} \right\}, \quad x \in (-e^{2t}, x_1^\epsilon(t)), \\ T_d^\epsilon(t, x) &= x - e^{2t} - T_u^\epsilon(t, x), \quad x \in (-e^{2t}, x_1^\epsilon(t)), \\ T_d^\epsilon(t, x) &= \left( e^{t+\epsilon} + e^{2(t+\epsilon)} \right) \left( 1 - \frac{x + e^{2t}}{e^t + e^{2t}} \right) - e^{2t} \quad x \in [x_1^\epsilon(t), e^t]. \end{aligned}$$

*Proof.* Since, on  $(\ell_t, x_1^\epsilon(t)) \times (y_1^\epsilon(t), r_{t+\epsilon})$ ,  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  coincides with the left-curtain martingale coupling of  $\mu_t|_{(\ell_t, x_1^\epsilon(t))}$  and  $\mu_{t+\epsilon}|_{(y_1^\epsilon(t), r_{t+\epsilon})}$ , the supporting functions  $(T_d^\epsilon(t, \cdot), T_u^\epsilon(t, \cdot))$  can be determined from the mass and mean preservation condition

$$\int_{-e^{2t}}^x z^i \mu_t(dz) = \int_{T_d^\epsilon(t, x)}^{T_u^\epsilon(t, x)} z^i \mu_{t+\epsilon}(dz), \quad i = 0, 1.$$

The direct computation gives that

$$\frac{x + e^{2t}}{e^t + e^{2t}} = \frac{T_u^\epsilon(t, x) - T_d^\epsilon(t, x)}{e^{t+\epsilon} + e^{2(t+\epsilon)}}, \quad T_d^\epsilon(t, x) = x - e^{2t} - T_u^\epsilon(t, x), \quad x \in (-e^{2t}, x_1^\epsilon(t)).$$

Solving for  $T_u^\epsilon(t, x)$  (which is easy) gives the required representations of the supporting functions on  $(-e^{2t}, x_1^\epsilon(t))$ .

To obtain  $T_d^\epsilon(t, \cdot)$  on  $[x_1^\epsilon(t), r_t]$ , we use that  $\bar{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}$  coincides with the antitone coupling  $\pi^{a, \mu_t | [x_1^\epsilon(t), r_t], \mu_{t+\epsilon} | (\ell_{t+\epsilon}, y_1^\epsilon(t))}$  of  $\mu_t | [x_1^\epsilon(t), r_t]$  and  $\mu_{t+\epsilon} | (\ell_{t+\epsilon}, y_1^\epsilon(t), r_{t+\epsilon})$ . It follows that

$$T_d^\epsilon(t, x) = F^{-1}(t + \epsilon, 1 - F(t, x)), \quad x \in [x_1^\epsilon(t), r_t],$$

with  $F(t, x) = (x + e^{2t})/(e^t + e^{2t})$ , which proves the claim.  $\square$

Now define  $J_d^\epsilon(t, x) = x - T_d^\epsilon(t, x)$  for  $x \in (\ell_t, r_t)$  and  $J_u^\epsilon(t, x) = T_u^\epsilon(t, x) - x$  for  $x \in (\ell_t, x_1^\epsilon(t))$ . Then direct computation gives that

$$j_d(t, x) := \lim_{\epsilon \downarrow 0} J_d^\epsilon(t, x) = e^{2t} + \lim_{\epsilon \downarrow 0} T_u^\epsilon(t, x) = e^{2t} + x, \quad x \in (\ell_t - e^{2t}, x_1(t) = r_t = e^t) \quad (4.21)$$

and

$$j_u(t, x) := \lim_{\epsilon \downarrow 0} \frac{J_u^\epsilon(t, x)}{\epsilon} = \frac{1 + 2e^{2t}}{2} \frac{x + e^{2t}}{1 + e^t}, \quad x \in (-e^{2t}, e^t). \quad (4.22)$$

**Proposition 4.5.** *Let  $(\mu_t)_{t \in [0,1]}$  be as in (4.14). For each  $n \geq 1$ , let  $\mathbb{P}^n := \mathbb{P}^{*,n} \circ (X^{*,n})^{-1}$ , where  $\mathbb{P}^{*,n}$  is the  $n$ -period supermartingale coupling (w.r.t. partition  $\pi_n$ ), obtained through the Markovian iteration of one-period increasing supermartingale couplings  $\bar{\mathbb{P}}^{\mu_{t_k^n}, \mu_{t_{k+1}^n}}$ . (Here  $X^{*,n}$  is the piece-wise constant canonical process; see Section 3.)*

*Then the sequence  $(\mathbb{P}^n)_{n \geq 1}$  is tight w.r.t. the Skorokhod topology on  $\Omega$ .*

*Moreover, any accumulation point of  $(\mathbb{P}^n)_{n \geq 1}$ , denoted by  $\mathbb{P}^0$ , is the law of the following SDE:*

$$X_t = X_0 - \int_0^t j_d(t, X_{s-})(dN_s - \nu_s ds) - \sum_{s \leq t} \mathbf{1}_{\{X_{s-} = e^s\}}(e^s + e^{2s}),$$

where  $(N_s)_{0 \leq s \leq 1}$  is a unit size jump process, with predictable compensator  $\nu_s := \frac{j_u}{j_d}(s, X_{s-})$ , with  $j_d, j_u$  as in (4.21), (4.22).

Finally, if Assumption 2.2 holds with  $c_{xyy} > 0$  and  $c_{xy} < 0$ , then  $\mathbb{P}^0$  solves  $\mathbf{P}_\infty(\mu)$  and the strong duality  $\mathbf{P}_\infty(\mu) = \mathbf{D}_\infty(\mu)$  holds.

*Proof.* The tightness of  $(\mathbb{P}^n)_{n \geq 1}$  is guaranteed by Proposition 3.4.

The proof of the SDE characterization of  $\mathbb{P}^0$  and its optimality use similar arguments as in Theorems 3.14 and 3.15. Some notable difference arise due to the fact that the construction is based on the one-period increasing supermartingale coupling (and not on the decreasing one). For example, the drift part of  $X$  is determined by  $j_u$  (and thus by the upward transitions of the increasing supermartingale coupling), while the jumps of  $X$  are governed by  $j_d$  (and thus arise from the downward transitions of the increasing coupling). In particular, the terms  $\alpha_u, \alpha_d$  (which are defined as in the proof of Theorem 3.14) will determine the integral part of the SDE. This does not introduce new arguments and thus we omit the details.

In order to explain the summation term that appears in the SDE, we turn to the term  $\alpha_{\bar{d}}$  (see the proof of Theorem 3.14):

$$\alpha_{\bar{d}} = \mathbb{E}_{t_k^n}^{\mathbb{P}^n} [\{\varphi(X_{t_k^n} - J_d^{\epsilon_k^n}(t_k^n, X_{t_k^n})) - \varphi(X_{t_k^n})\} \mathbf{1}_{\{x_1^{\epsilon_k^n}(t_k^n) \leq X_{t_k^n} < e^{\frac{t_k^n}{k}}\}}],$$

where  $\epsilon_k^n = t_{k+1}^n - t_k^n$ . Note that in the present setting  $x_1^\epsilon(t) \rightarrow e^t$ , and thus in the limit, the supermartingale region  $[x_1^\epsilon(t), r_t)$  becomes a single point. Using this together with the fact that  $J_d^\epsilon(t, x) \rightarrow e^{2t} + x$ , we obtain that

$$\alpha_{\tilde{d}} = \mathbb{E}_{t_k^n}^{\mathbb{P}^n} \left[ \sum_{t_k^n \leq s \leq t_{k+1}^n} [\varphi(-e^{2s}) - \varphi(e^s)] \mathbf{1}_{\{X_s = e^s\}} \right] + O(\epsilon_k^n (\epsilon_k^n \vee \rho_1(\epsilon_k^n))),$$

which, in the limit, determines the summation term  $\sum_{s \leq t} \mathbf{1}_{\{X_s = e^s\}} (e^s + e^{2s})$  of the SDE.

For the optimality, the arguments are similar to those of Theorem 3.15. The candidate optimal dual strategies should be redefined to reflect the fact that we are working with the increasing supermartingale coupling (or the left-curtain martingale coupling in the martingale region). In the limit, we have that  $X$  is a martingale on  $(\ell_t, r_t)$ ,  $t \in [0, 1]$ , and has a strict supermartingale transition (a jump down to the lower boundary  $\ell_t$ ) only if it escapes to the upper boundary  $r_t$  at time  $t$ . Hence the dual strategies (see Section 3.3) in fact can be defined as in Henry-Labordère et al. [41, Eq. (3.9), (3.10), (3.11)]:

$$\partial_x h^*(t, x) = \frac{c_x(x, x) - c_x(x, T_d(t, x))}{j_d(t, x)}, \quad x \in (\ell_t, r_t),$$

$$\lim_{x \uparrow x_1(t) = e^t} h^*(t, x) = 0 = h^*(t, e^t).$$

We remark that from the above definition, for all  $t \in [0, 1]$ ,  $x \mapsto h^*(t, x)$  is continuous. In addition,  $\psi^*$  is defined, up to a constant, by

$$\partial_x \psi^*(t, x) := -h^*(t, x), \quad (t, x) \in E = \{(t, x) : t \in [0, 1], x \in (\ell_t, r_t)\}.$$

The limiting arguments that prove optimality of  $(h^*, \psi^*)$  are as in the decreasing case, and thus we omit the details.  $\square$

## 4.2 Bachelier dynamics with negative drift

In this section we consider the marginals  $(\mu_t)_{t \in (0, 1]}$  such that each  $\mu_t$  has the density

$$f(t, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x+t)^2}{2t}}, \quad t \in (0, 1]. \quad (4.23)$$

Note that Assumption 3.1 is almost satisfied; indeed  $\mu_0$  is left undefined. We could overcome this by introducing  $\delta > 0$  and then working with  $t \in [\delta, 1]$ . Alternatively, we could specify  $\tilde{\mu}_t$  to have density  $f(t + \delta, \cdot)$  for each  $t \in [0, 1]$ . For the convenience of notation we choose to work on  $(0, 1]$ .

Fix  $t \in (0, 1)$  and  $\epsilon \in (0, 1 - t]$ . Since  $m_\epsilon(t)$  and  $m^\epsilon(t)$ , as in Assumption 3.1, correspond to two unique crossing points of the densities  $f(t, \cdot), f(t + \epsilon, \cdot)$ , by direct computation one easily obtains that

$$m^2 = \frac{t(t + \epsilon)}{\epsilon} \ln \frac{t + \epsilon}{t} + (t + \epsilon)t, \quad m \in \{m_\epsilon(t), m^\epsilon(t)\}.$$

On the other hand, straightforward calculations show that  $m_t$  (resp.  $\tilde{m}_t$ ), the minimizer (resp. maximizer) of  $\partial_t F(t, \cdot)$ , is given by  $m_t = -\tilde{m}_t = \sqrt{t(t + 1)}$ . Note that  $m_t = \lim_{\epsilon \downarrow 0} m^\epsilon(t)$ ,  $\tilde{m}_t = \lim_{\epsilon \downarrow 0} m_\epsilon(t)$ , for each  $t \in (0, 1)$ . It follows that Assumption 3.3 holds.

Our next goal is to show that  $(\mu_t)_{t \in (0, 1]}$  also satisfies Assumption 3.7. Since,  $r_t = \infty$  for all  $t \in (0, 1]$ , Assumption 3.7(ii) is immediate. For Assumption 3.7 (iii), simple calculations lead to

$$\partial_{tx} f(t, x) = t^{-\frac{5}{2}} \Phi' \left( \frac{x + t}{\sqrt{t}} \right) \left[ x + t + \frac{1}{2}(t - x) \left( \frac{(x + t)^2}{t} - 1 \right) \right],$$

where  $\Phi(x) := \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$  is the c.d.f. of the standard normal random variable. Now using that  $m_t = \sqrt{t(t+1)}$ , we consequently have that  $\partial_{tx} f(t, m_t) = t^{-2} \sqrt{t+1} \Phi'(\sqrt{t} + \sqrt{t+1}) > 0$ .

It is left to verify Assumption 3.7(i). Let  $x_1^\epsilon(t)$  be the unique phase transition point of the decreasing supermartingale coupling of  $\mu_t$  and  $\mu_{t+\epsilon}$ ; see Section B.1.

**Lemma 4.6.** *Let  $(\mu_t)_{t \in (0,1]}$  be specified by (4.23). Then,  $x_1^\epsilon(t)$  is uniquely determined by the equation*

$$1 - \Phi\left(\frac{x_1^\epsilon(t) + t}{\sqrt{t}}\right) = \frac{1}{\epsilon}(\sqrt{t+\epsilon} - \sqrt{t})\Phi'\left(\frac{x_1^\epsilon(t) + t}{\sqrt{t}}\right), \quad t \in (0,1), \quad \epsilon \in (0, t-1].$$

Furthermore, the limit  $x_1(t) := \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$  exists, and is uniquely determined by the equation

$$2\sqrt{t}\left(1 - \Phi\left(\frac{x_1(t) + t}{\sqrt{t}}\right)\right) = \Phi'\left(\frac{x_1(t) + t}{\sqrt{t}}\right). \quad (4.24)$$

The proof of Lemma 4.6 requires the following auxiliary result.

**Lemma 4.7.** *For each  $t \in (0,1]$ , the equation  $2\sqrt{t}(1 - \Phi(x)) = \Phi'(x)$  admits a unique (real-valued) solution  $x_t^* \in (-\infty, 2\sqrt{t})$ .*

*Proof.* The proof is an application of the intermediate value theorem. Let us consider the function  $F(x) := 2\sqrt{t}(1 - \Phi(x)) - \Phi'(x)$ . First, as  $F'(x) = \Phi'(x)(x - 2\sqrt{t})$ , it is clear that  $x \mapsto F(x)$  is continuous on  $\mathbb{R}$ , decreasing on  $(-\infty, 2\sqrt{t}]$ , increasing on  $[2\sqrt{t}, \infty)$ . Furthermore,  $\lim_{x \rightarrow -\infty} F(x) = 2\sqrt{t} > 0$  and  $\lim_{x \rightarrow \infty} F(x) = 0$ . It follows that the equation  $F(x) = 0$  admits a unique solution in the interval  $(-\infty, 2\sqrt{t})$ .  $\square$

*Proof of Lemma 4.6.* To determine  $x_1^\epsilon(t) \in (\ell_t, r_t)$  we use the fact that under the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}, \mu|_{[x_1^\epsilon(t), r_t]}$  is mapped to  $\nu|_{[y_1^\epsilon(t), r_{t+\epsilon}]}$  in a martingale way, where  $y_1^\epsilon(t) = y_1^{\mu_t, \mu_{t+\epsilon}}$  (see (B.51)). In particular, the pair  $(x_1^\epsilon(t), y_1^\epsilon(t))$  is uniquely determined by the mass and mean preservation condition

$$\int_{x_1^\epsilon(t)}^{r_t} z^i \mu_t(dz) = \int_{y_1^\epsilon(t)}^{r_{t+\epsilon}} z^i \mu_{t+\epsilon}(dz), \quad i = 0, 1. \quad (4.25)$$

Note that due to the Dispersion Assumption (see Assumption 3.1), we must have that  $y_1^\epsilon(t) \leq x_1^\epsilon(t) < m^\epsilon(t)$ .

Now apply the change of variables:  $\hat{x}_1^\epsilon(t) = \frac{x_1^\epsilon(t) + t}{\sqrt{t}}$ ,  $\hat{y}_1^\epsilon(t) = \frac{y_1^\epsilon(t) + t + \epsilon}{\sqrt{t+\epsilon}}$ . Then (4.25) with  $i = 0$  reads

$$\Phi(\hat{x}_1^\epsilon(t)) = \Phi(\hat{y}_1^\epsilon(t)). \quad (4.26)$$

For (4.25) with  $i = 1$ , we first re-write it as

$$\int_{x_1^\epsilon(t)}^{\infty} (x+t)f(t, x)dx + \epsilon \int_{x_1^\epsilon(t)}^{\infty} f(t, x)dx = \int_{y_1^\epsilon(t)}^{\infty} (x+t+\epsilon)f(t+\epsilon, x)dx,$$

by adding  $(t+\epsilon)$  times the equation (4.25) with  $i = 0$  on both sides. Then direct calculation leads to

$$\sqrt{t}\Phi'(\hat{x}_1^\epsilon(t)) + \epsilon(1 - \Phi(\hat{x}_1^\epsilon(t))) = \sqrt{t+\epsilon}\Phi'(\hat{y}_1^\epsilon(t)). \quad (4.27)$$

From (4.26), we get  $\hat{x}_1^\epsilon(t) = \hat{y}_1^\epsilon(t)$ , or equivalently  $\frac{x_1^\epsilon(t) + t}{\sqrt{t}} = \frac{y_1^\epsilon(t) + t + \epsilon}{\sqrt{t+\epsilon}}$ . By plugging this relation  $\hat{x}_1^\epsilon(t) = \hat{x}_2^\epsilon(t)$  into (4.27), we obtain that

$$1 - \Phi(\hat{x}_1^\epsilon(t)) = \frac{1}{\epsilon}(\sqrt{t+\epsilon} - \sqrt{t})\Phi'(\hat{x}_1^\epsilon(t)), \quad (4.28)$$



which is precisely the required algebraic equation that characterizes  $x_1^\epsilon(t)$ .

We now deal with the limit  $x_1(t) := \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$ .

Recall that  $\epsilon_0 := \epsilon_1 \wedge \epsilon_2$ , where  $\epsilon_1, \epsilon_2$  are as in Assumptions 3.1 and 3.3, respectively.

For each (small)  $\delta > 0$ , define  $F = F^\delta : \mathbb{R}^3 \rightarrow \mathbb{R}$  on  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_0, \delta \leq t \leq 1 - \epsilon\} \times \mathbb{R}$  by

$$F(t, \epsilon, x) := (1 - \Phi(x)) - \frac{\sqrt{t+\epsilon} - \sqrt{t}}{\epsilon} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad (t, \epsilon, x) \in [\delta, 1 - \epsilon] \times (0, \epsilon_0] \times \mathbb{R} \quad (4.29)$$

and

$$F(t, 0, x) := (1 - \Phi(x)) - \frac{1}{2\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad (t, x) \in [\delta, 1] \times \mathbb{R}. \quad (4.30)$$

Note that, by (4.28),  $F(t, \epsilon, \hat{x}_1^\epsilon(t)) = 0$  for all  $\epsilon \in (0, \epsilon_0]$  and  $t \in [\delta, 1 - \epsilon]$ .

Using Lemma 4.7, we define  $\hat{x}_1^0(t)$ , for all  $t \in (0, 1]$ , to be the unique solution to  $F(t, 0, x) = 0$ . Then  $x_1^0(t)$  is defined as  $x_1^0(t) := \sqrt{t} \hat{x}_1^0(t) - t$ , for each  $t \in (0, 1]$ . Note that  $x_1^0(\cdot)$  is continuous if and only if  $\hat{x}_1^0(\cdot)$  is. Now since  $\hat{x}_1^0(t)$  uniquely satisfies  $F(t, 0, x) = 0$ , we immediately have that  $\lim_{\epsilon \downarrow 0} \hat{x}_1^\epsilon(t)$  exists, for all  $t \in (0, 1]$ , and uniquely satisfies (4.24).  $\square$

**Lemma 4.8.** *For  $\epsilon \in [0, \epsilon_0]$  and  $t \in (0, 1]$ , let  $x_1^\epsilon(t)$  be given by Lemma 4.6 with  $x_1^0(t) := x_1(t) = \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$ .*

*Then, for each (small)  $\delta > 0$ , the map  $(t, \epsilon) \mapsto x_1^\epsilon(t)$  is continuous on  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_0, \delta \leq t \leq 1 - \epsilon\} \times \mathbb{R}$ .*

*Proof.* Let  $F$  be defined as in (4.29) and (4.30). To prove the continuity of  $(t, \epsilon) \mapsto x_1^\epsilon(t) := \sqrt{t} \hat{x}_1^\epsilon(t) - t$ , we will use the implicit function theorem and show that  $(t, \epsilon) \mapsto \hat{x}_1^\epsilon(t)$  is continuous, where  $F(t, \epsilon, \hat{x}_1^\epsilon(t)) = 0$ .

We claim that  $(t, \epsilon, x) \mapsto F(t, \epsilon, x)$  is continuously differentiable.

First, for  $0 < \epsilon \leq \epsilon_0$ , we have that

$$\frac{\partial F}{\partial \epsilon}(t, \epsilon, x) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{\frac{1}{2}(t+\epsilon)^{-\frac{1}{2}}\epsilon - (\sqrt{t+\epsilon} - \sqrt{t})}{\epsilon^2},$$

and it follows that  $\lim_{\epsilon \rightarrow 0} \frac{\partial F}{\partial \epsilon}(t, \epsilon, x) = -\Phi'(x) \frac{1}{4} t^{-\frac{3}{2}}$ . On the other hand,

$$\frac{\partial F}{\partial \epsilon}(t, 0, x) = \lim_{\epsilon \rightarrow 0} \frac{F(t, \epsilon, x) - F(t, 0, x)}{\epsilon} = \lim_{\epsilon \rightarrow 0} -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{\frac{1}{2} t^{-\frac{1}{2}} \epsilon - (\sqrt{t+\epsilon} - \sqrt{t})}{\epsilon^2} = -\Phi'(x) \frac{1}{4} t^{-\frac{3}{2}},$$

and hence  $\frac{\partial F}{\partial \epsilon}$  is continuous. The continuity of  $\frac{\partial F}{\partial t}$  and  $\frac{\partial F}{\partial x}$  follow similarly by the direct computations.

Now we check that  $\frac{\partial F}{\partial x}(t, \epsilon, x)|_{F(t, \epsilon, x)=0} \neq 0$ . For  $\epsilon \in [0, \epsilon_0]$ , direct calculations show that

$$\frac{\partial F}{\partial x}(t, \epsilon, x)|_{F(t, \epsilon, x)=0} = \Phi'(x) \left[ x \frac{1 - \Phi(x)}{\Phi'(x)} - 1 \right] < 0,$$

where for the inequality we use the fact that  $\frac{1 - \Phi(x)}{\Phi'(x)} < \frac{1}{x}$  when  $x > 0$  (for  $x \leq 0$  the inequality is trivially satisfied).

By the implicit function theorem, we get that  $(t, \epsilon) \mapsto \hat{x}_1^\epsilon(t)$  is continuously differentiable on a compact set  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_0, \delta \leq t \leq 1 - \epsilon\}$ , and consequently uniformly bounded and continuous. The regularity of  $(t, \epsilon) \mapsto x_1^\epsilon(t)$  follows from the regularity of  $(t, x) \mapsto \hat{x}_1^\epsilon(t)$ .  $\square$

Since all the assumption required for Theorems 3.14 and 3.15 hold we have the following.

**Proposition 4.9.** Let  $(\mu_t)_{t \in (0,1]}$  be specified by (4.23) and consider  $t \mapsto x_1(t)$  as in Lemma 4.6. Then the SDE as in Theorem 3.14 is explicitly given by

$$X_t = X_0 + \int_0^t j_u(s, X_{s-})(dN_s - \nu_s ds) \mathbf{1}_{\{x_1(s) < X_{s-} < m_s\}} - \int_0^t \mathbf{1}_{\{X_{s-} \leq x_1(s)\}} j_d(s, X_{s-}) ds,$$

where  $m_s = \sqrt{s(s+1)}$ ,  $\nu_s(s, X_{s-}) := \frac{j_d}{j_u}(s, X_{s-})$  with  $j_u(t, x) = T_u(t, x) - x$ , and  $j_d(t, x) = \frac{1}{2\sqrt{t}}(\tilde{T}_u(t, x) - 2\sqrt{t})e^{\frac{\hat{x}^2}{2} - \frac{\tilde{T}_u(t, x)^2}{2}} - \frac{\hat{x}}{2\sqrt{t}} + 1$ , for  $x \in (x_1(s), m_s)$ , and  $j_d(s, X_{s-}) := \frac{s-x}{2s}$  for  $x \in (\ell_s, x_1(s)]$ ; here  $\hat{x} := \frac{x+t}{\sqrt{t}}$  and  $\tilde{T}_u(t, x) := \sqrt{t}T_u(t, x) - t$  is uniquely determined by the equation

$$(\tilde{T}_u(t, x) - \hat{x})(\tilde{T}_u(t, x) - 2\sqrt{t})\Phi'(\tilde{T}_u(t, x)) + 2\sqrt{t}(\Phi(\tilde{T}_u(t, x)) - \Phi(\hat{x})) = \Phi'(\hat{x}) - \Phi'(\tilde{T}_u(t, x)).$$

In addition, if Assumption 2.2(ii) and the integrability condition (3.12) are satisfied, then the law  $\mathbb{P}^0$  of the above process is the optimal solution for the primal problem (2.4), and the strong duality

$$\mathbb{E}^{\mathbb{P}^0}[C(X)] = \mathbf{P}_\infty(\mu) = \mathbf{D}_\infty(\mu)$$

holds.

*Proof.* As it has been verified that Assumptions 3.1, 3.3, 3.7 are valid (up to  $t = 0$ ), using Theorems 3.14 and 3.15 we obtain the validity of the SDE and the optimality of  $\mathbb{P}^0$ . It is left to derive the explicit expressions of the terms  $T_u, j_u, j_d$  that define the SDE.

In the supermartingale region, i.e., for  $x \in (\ell_t, x_1(t)]$ , we have that

$$j_d(t, x) = \frac{1}{f(t, x)} \partial_t F(t, x) = \frac{\sqrt{t}}{\Phi'(\frac{x+t}{\sqrt{t}})} \partial_t \Phi(\frac{x+t}{\sqrt{t}}) = \frac{t-x}{2t}.$$

In the martingale region, i.e., for  $x \in (x_1(t), m_t)$ , using (3.8) we have that  $T_u(t, x)$  is uniquely determined by

$$\int_x^{T_u(t, x)} (x - \xi) \frac{\partial}{\partial t} \left( \frac{1}{\sqrt{2\pi t}} e^{-\frac{(\xi+t)^2}{2t}} \right) d\xi = 0.$$

By a change of variables  $\tilde{T}_u(t, x) := \sqrt{t}T_u(t, x) - t$ , we obtain that  $\tilde{T}_u(t, x)$  satisfies

$$(\tilde{T}_u(t, x) - \hat{x})(\tilde{T}_u(t, x) - 2\sqrt{t})\Phi'(\tilde{T}_u(t, x)) + 2\sqrt{t}(\Phi(\tilde{T}_u(t, x)) - \Phi(\hat{x})) = \Phi'(\hat{x}) - \Phi'(\tilde{T}_u(t, x)).$$

Finally, for  $x \in (x_1(t), m_t)$ ,

$$j_d(t, x) = \frac{\partial_t F(t, x) - \partial_t F(t, T_u(t, x))}{f(t, x)} = \frac{1}{2\sqrt{t}}(\tilde{T}_u(t, x) - 2\sqrt{t})e^{\frac{\hat{x}^2}{2} - \frac{\tilde{T}_u(t, x)^2}{2}} - \frac{\hat{x}}{2\sqrt{t}} + 1,$$

as required.  $\square$

**Remark 4.10.** Let  $(\mu_t)_{t \in (0,1]}$  be specified by (4.23) and for  $t \in (0, 1)$  and  $\epsilon \in (0, 1-t]$ , consider the increasing supermartingale coupling of  $\mu_t$  and  $\mu_{t+\epsilon}$ . Then there exists the unique threshold  $x_1^\epsilon(t)$ , such that  $(\ell_t, x_1^\epsilon(t)]$  is a martingale region, while  $(x_1^\epsilon(t), r_t)$  corresponds to the supermartingale region.

By changing variables  $\hat{x}_1^\epsilon(t) = \frac{x_1^\epsilon(t)+t}{\sqrt{t}}$  and using similar arguments as in Lemma 4.6, we have that  $x_1^\epsilon(t)$  uniquely solves

$$\epsilon \Phi(\hat{x}_1^\epsilon(t)) = \sqrt{t} \Phi'(\hat{x}_1^\epsilon(t)) + \sqrt{t+\epsilon} \Phi'(-\hat{x}_1^\epsilon(t)).$$

Then letting  $\epsilon \downarrow 0$  we have that  $\hat{x}_1^\epsilon(t) \rightarrow +\infty$  (provided the limit exists). It follows that  $x_1^\epsilon(t) \rightarrow +\infty$ , and therefore one could expect that the supermartingale transitions of the limiting SDE correspond to jumps from  $\infty$  to  $-\infty$ .

This is in-line with our observations in Proposition 4.5 obtained in the uniform case. In particular, when the increasing supermartingale coupling is used in construction, and in the case of unbounded support of marginals, the convergence results may fail in general. We will consider these issues in our future research.

### 4.3 Geometric Brownian motion with decreasing average

In this section we consider the marginals  $(\mu_t)_{t \in (0,1]}$  such that each  $\mu_t$  has the density

$$f(t, x) = \frac{1}{x\sqrt{2\pi t}} e^{-\frac{(\ln x + t)^2}{2t}}, \quad t \in (0, 1]. \quad (4.31)$$

Similarly to the previous Bachelier case (see Section 4.2), we can introduce  $\delta > 0$  and then work with  $t \in [\delta, 1]$  (or  $t \in [\delta, 1 + \delta]$ ). For the convenience of notation we shall work on  $(0, 1]$ .

Assumption 3.1 is clearly satisfied. Fix  $t \in (0, 1)$  and  $\epsilon \in (0, 1 - t]$ . For  $m_\epsilon(t)$  and  $m^\epsilon(t)$  (as in Assumption 3.1), direct computation leads to

$$(\ln m)^2 = \frac{t(t + \epsilon)}{\epsilon} \ln \frac{t + \epsilon}{t} + (t + \epsilon)t, \quad m \in \{m_\epsilon(t), m^\epsilon(t)\}.$$

On the other hand, straightforward calculations show that  $m_t$  (resp.  $\tilde{m}_t$ ), the minimizer (resp. maximizer) of  $\partial_t F(t, \cdot)$ , is given by  $\ln m_t = -\ln \tilde{m}_t = \sqrt{t(t+1)}$ . It follows that Assumption 3.3 is also valid.

We now show that  $(\mu_t)_{t \in (0,1]}$  also satisfies Assumption 3.7. First, Assumption 3.7(ii) is immediate, as  $r_t = \infty$  for all  $t \in (0, 1]$ . For Assumption 3.7 (iii), simple calculations lead to

$$\partial_{tx} f(t, m_t) = t^{-\frac{3}{2}} \sqrt{t+1} \Phi'(\sqrt{t} + \sqrt{t+1}) e^{-2\sqrt{t(t+1)}} > 0.$$

It is left to verify Assumption 3.7(i). Let  $x_1^\epsilon(t)$  be the unique phase transition point of the decreasing supermartingale coupling of  $\mu_t$  and  $\mu_{t+\epsilon}$ ; see Section B.1.

**Lemma 4.11.** *Let  $(\mu_t)_{t \in (0,1]}$  be specified by (4.31). Then,  $x_1^\epsilon(t)$  is uniquely determined by the equation*

$$1 - \Phi\left(\frac{\ln x_1^\epsilon(t)}{\sqrt{t}}\right) = e^{-\frac{1}{2}\epsilon} [1 - \Phi\left(\frac{\ln x_1^\epsilon(t) + t}{\sqrt{t}} - \sqrt{t + \epsilon}\right)], \quad t \in (0, 1), \quad \epsilon \in (0, t - 1].$$

Furthermore, the limit  $x_1(t) := \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$  exists, and is uniquely determined by the equation

$$\frac{1}{\sqrt{t}} \Phi'\left(\frac{\ln x_1(t)}{\sqrt{t}}\right) = 1 - \Phi\left(\frac{\ln x_1(t)}{\sqrt{t}}\right). \quad (4.32)$$

The proof of Lemma 4.11 requires the following auxiliary result.

**Lemma 4.12.** *For each  $t \in (0, 1]$ , the equation  $\sqrt{t}(1 - \Phi(x)) = \Phi'(x)$  admits the unique (real-valued) solution  $x_t^* \in (-\infty, \sqrt{t})$ .*

*Proof.* The proof is similar to Lemma 4.7 and is an application of the intermediate value theorem. Hence we omit the details here.  $\square$

*Proof of Lemma 4.11.* To determine  $x_1^\epsilon(t) \in (\ell_t, r_t)$  we use the fact that under the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu_t, \mu_{t+\epsilon}}, \mu|_{[x_1^\epsilon(t), r_t]}$  is mapped to  $\nu|_{[y_1^\epsilon(t), r_{t+\epsilon}]}$  in a martingale way, where  $y_1^\epsilon(t) = y_1^{\mu_t, \mu_{t+\epsilon}}$  (see (B.51)). In particular, the pair  $(x_1^\epsilon(t), y_1^\epsilon(t))$  is uniquely determined by the mass and mean preservation condition

$$\int_{x_1^\epsilon(t)}^{r_t=\infty} z^i \mu_t(dz) = \int_{y_1^\epsilon(t)}^{r_{t+\epsilon}=\infty} z^i \mu_{t+\epsilon}(dz), \quad i = 0, 1. \quad (4.33)$$

Note that due to the Dispersion Assumption (see Assumption 3.1), we must have that  $y_1^\epsilon(t) \leq x_1^\epsilon(t) < m^\epsilon(t)$ .

Now apply the change of variables:  $\hat{x}_1^\epsilon(t) = \frac{\ln x_1^\epsilon(t) + t}{\sqrt{t}}$ ,  $\hat{y}_1^\epsilon(t) = \frac{\ln y_1^\epsilon(t) + t + \epsilon}{\sqrt{t + \epsilon}}$ . Then (4.33) with  $i = 0$  reads

$$\Phi(\hat{x}_1^\epsilon(t)) = \Phi(\hat{y}_1^\epsilon(t)). \quad (4.34)$$

For (4.33) with  $i = 1$ , it is equivalent to

$$\int_{\hat{x}_1^\epsilon(t)}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y - \sqrt{t})^2} dy = e^{-\frac{1}{2}\epsilon} \int_{\hat{y}_1^\epsilon(t)}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y - \sqrt{t + \epsilon})^2} dy,$$

which leads to

$$1 - \Phi(\hat{x}_1^\epsilon(t) - \sqrt{t}) = e^{-\frac{1}{2}\epsilon} [1 - \Phi(\hat{y}_1^\epsilon(t) - \sqrt{t + \epsilon})]. \quad (4.35)$$

From (4.26), we get  $\hat{x}_1^\epsilon(t) = \hat{y}_1^\epsilon(t)$ , or equivalently  $\frac{\ln x_1^\epsilon(t) + t}{\sqrt{t}} = \frac{\ln y_1^\epsilon(t) + t + \epsilon}{\sqrt{t + \epsilon}}$ . By plugging  $\hat{x}_1^\epsilon(t) = \hat{y}_1^\epsilon(t)$  into (4.35), we obtain that

$$1 - \Phi\left(\frac{\ln x_1^\epsilon(t)}{\sqrt{t}}\right) = e^{-\frac{1}{2}\epsilon} [1 - \Phi\left(\frac{\ln x_1^\epsilon(t) + t}{\sqrt{t}} - \sqrt{t + \epsilon}\right)], \quad (4.36)$$

which characterizes  $x_1^\epsilon(t)$ .

We now deal with the limit  $x_1(t) := \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$ .

Recall that  $\epsilon_0 := \epsilon_1 \wedge \epsilon_2$ , where  $\epsilon_1, \epsilon_2$  are as in Assumptions 3.1 and 3.3, respectively.

For each (small)  $\delta > 0$ , define  $G = G^\delta : \mathbb{R}^3 \rightarrow \mathbb{R}$  on  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_0, \delta \leq t \leq 1 - \epsilon\} \times \mathbb{R}$  by

$$G(t, \epsilon, x) := \frac{1}{\epsilon} \left\{ (1 - \Phi(x)) - e^{-\frac{1}{2}\epsilon} [1 - \Phi(x + \sqrt{t} - \sqrt{t + \epsilon})] \right\}, \quad (t, \epsilon, x) \in [\delta, 1 - \epsilon] \times (0, \epsilon_0] \times \mathbb{R} \quad (4.37)$$

and

$$G(t, 0, x) := (1 - \Phi(x)) - \frac{1}{\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (4.38)$$

Note that, by (4.36),  $G(t, \epsilon, \hat{x}_1^\epsilon(t)) = 0$  for all  $\epsilon \in (0, \epsilon_0]$  and  $t \in [\delta, 1 - \epsilon]$ .

Using Lemma 4.12, we define  $\hat{x}_1^0(t)$ , for all  $t \in (0, 1]$ , to be the unique solution to  $F(t, 0, x) = 0$ . Then  $x_1^0(t)$  is defined as  $x_1^0(t) := \exp(\sqrt{t} \hat{x}_1^0(t) - t)$ , for each  $t \in (0, 1]$ . Note that  $x_1^0(\cdot)$  is continuous if and only if  $\hat{x}_1^0(\cdot)$  is. Now since  $\hat{x}_1^\epsilon(t)$  uniquely satisfies  $F(t, 0, x) = 0$ , we immediately have that  $\lim_{\epsilon \downarrow 0} \hat{x}_1^\epsilon(t)$  exists, for all  $t \in (0, 1]$ , and uniquely satisfies (4.24).  $\square$

**Lemma 4.13.** For  $\epsilon \in [0, \epsilon_0]$  and  $t \in (0, 1]$ , let  $x_1^\epsilon(t)$  be given by Lemma 4.11 with  $x_1^0(t) := x_1(t) = \lim_{\epsilon \downarrow 0} x_1^\epsilon(t)$ .

Then, for each (small)  $\delta > 0$ , the map  $(t, \epsilon) \mapsto x_1^\epsilon(t)$  is continuous on  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_0, \delta \leq t \leq 1 - \epsilon\} \times \mathbb{R}$ .

*Proof.* Let  $G$  be defined as in (4.37) and (4.38). To prove the continuity of  $(t, \epsilon) \mapsto x_1^\epsilon(t) := \exp(\sqrt{t} \hat{x}_1^\epsilon(t) - t)$ , we will use the implicit function theorem and show that  $(t, \epsilon) \mapsto \hat{x}_1^\epsilon(t)$  is continuous, where  $G(t, \epsilon, \hat{x}_1^\epsilon(t)) = 0$ .

We claim that  $(t, \epsilon, x) \mapsto G(t, \epsilon, x)$  is continuously differentiable. Indeed,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{\partial G}{\partial \epsilon}(t, \epsilon, x) &= -\frac{1}{8}(1 - \Phi(x)) + \frac{1}{8}\Phi'(x)t^{-\frac{3}{2}}(2t + 1) + \frac{1}{8}\Phi''(x)t^{-1} \\ &= \lim_{\epsilon \rightarrow 0} \frac{G(t, \epsilon, x) - G(t, 0, x)}{\epsilon}, \end{aligned}$$

and hence  $\frac{\partial G}{\partial \epsilon}$  is continuous. The continuity of  $\frac{\partial G}{\partial t}$  and  $\frac{\partial G}{\partial x}$  follow similarly by the direct computations.

Now we check that  $\frac{\partial G}{\partial x}(t, \epsilon, x)|_{G(t, \epsilon, x)=0} \neq 0$ . For  $\epsilon \in [0, \epsilon_0]$ , direct calculations show that

$$\frac{\partial G}{\partial x}(t, \epsilon, x)|_{G(t, \epsilon, x)=0} = -\Phi'(x) + e^{-\frac{1}{2}\epsilon}\Phi'(x + \sqrt{t} - \sqrt{t + \epsilon}) = -\frac{\Phi'(x)}{1 - \Phi(x)} + \frac{\Phi'(x + \sqrt{t} - \sqrt{t + \epsilon})}{1 - \Phi(x + \sqrt{t} - \sqrt{t + \epsilon})},$$

and the desired conclusion follows.

We now study the monotonicity of the map  $x \mapsto \frac{\Phi'(x)}{1 - \Phi(x)}$ :

$$\left( \frac{\Phi'(x)}{1 - \Phi(x)} \right)' = \frac{\Phi''(x)(1 - \Phi(x)) + (\Phi'(x))^2}{(1 - \Phi(x))^2} = \Phi'(x) \frac{\Phi'(x) - x(1 - \Phi(x))}{(1 - \Phi(x))^2} > 0,$$

where in the last step, in the case  $x > 0$ , we use the inequality  $\frac{1 - \Phi(x)}{\Phi'(x)} < \frac{1}{x}$  (the case  $x \geq 0$  is trivially satisfied). Hence  $x \mapsto \frac{\Phi'(x)}{1 - \Phi(x)}$  is strictly increasing. It follows that  $\frac{\partial G}{\partial x}|_{G(t, \epsilon, x)=0} < 0$  for  $\epsilon \in (0, \epsilon_0]$ . Similarly, for  $\epsilon = 0$ , we have that

$$\frac{\partial G}{\partial x}(t, 0, x)|_{G(t, 0, x)=0} = -\Phi'(x) + \frac{1}{\sqrt{t}}\Phi'(x)x = -\Phi'(x) + x(1 - \Phi(x)) < 0.$$

Finally, from the implicit function theorem, we have that  $(t, \epsilon) \mapsto \ln x_1^\epsilon(t)$  is continuously differentiable on a compact set  $\{(t, \epsilon) : 0 \leq \epsilon \leq \epsilon_0, \delta \leq t \leq 1 - \epsilon\}$ , and consequently uniformly bounded and continuous.  $\square$

Finally we have the following characterization and the optimality of the limiting process.

**Proposition 4.14.** *Let  $(\mu_t)_{t \in (0, 1]}$  be specified by (4.31) and consider  $t \mapsto x_1(t)$  as in Lemma 4.11. Then the SDE as in Theorem 3.14 is explicitly given by*

$$X_t = X_0 + \int_0^t j_u(s, X_{s-})(dN_s - \nu_s ds) \mathbf{1}_{\{x_1(s) < X_{s-} < m_s\}} - \int_0^t \mathbf{1}_{\{X_{s-} \leq x_1(s)\}} j_d(s, X_{s-}) ds,$$

where  $m_s = e^{\sqrt{s(s+1)}}$ ,  $\nu_s(s, X_{s-}) := \frac{j_d}{j_u}(s, X_{s-})$  with  $j_u(t, x) = T_u(t, x) - x$ , and

$$j_d(t, x) = \frac{x}{2} \left[ e^{\frac{(\ln x + t)^2}{2t}} - \frac{(\ln(T_u(t, x)) + t)^2}{2t} \left( \frac{\ln(T_u(t, x))}{\sqrt{t}} - \sqrt{t} \right) - \left( \frac{\ln x}{\sqrt{t}} - \sqrt{t} \right) \right],$$

for  $x \in (x_1(s), m_s)$ , and  $j_d(s, X_{s-}) := x \frac{s - \ln x}{2s}$  for  $x \in (\ell_s, x_1(s)]$ ; here  $T_u(t, x)$  is uniquely determined by the following equation:

$$\left(1 - \frac{x}{T_u(t, x)}\right) e^{-\frac{(\ln(T_u(t, x)))^2}{2t}} \ln(T_u(t, x)) = t \left(1 - \frac{x}{T_u(t, x)}\right) e^{-\frac{(\ln x)^2}{2t}} - t^{\frac{3}{2}} \left[ \Phi\left(\frac{\ln(T_u(t, x))}{\sqrt{t}}\right) - \Phi\left(\frac{\ln x}{\sqrt{t}}\right) \right].$$

In addition, if Assumption 2.2(ii) and the integrability condition (3.12) are satisfied, then the law  $\mathbb{P}^0$  of the above process is the optimal solution for the primal problem (2.4), and the strong duality

$$\mathbb{E}^{\mathbb{P}^0}[C(X)] = \mathbf{P}_\infty(\mu) = \mathbf{D}_\infty(\mu)$$

holds.

*Proof.* As it has been verified that Assumptions 3.1, 3.3, 3.7 are valid (up to  $t = 0$ ), using Theorems 3.14 and 3.15 we obtain the validity of the SDE and the optimality of  $\mathbb{P}^0$ . It is left to derive the explicit expressions of the terms  $T_u, j_u, j_d$  that define the SDE.

The calculations are similar to those of Proposition 4.9, and we omit the details.  $\square$

## A Brenier's Theorem

### A.1 Brenier's Theorem in Optimal Transport (OT)

Let  $(X, Y)$  be a random vector in  $\mathbb{R}^2$  and let  $\mathcal{P}_2$  be the set of all (Borel) probability measures on  $\mathbb{R}^2$ . We denote respectively  $\mu$  and  $\nu$  the (integrable) marginal distributions of  $X$  and  $Y$ .

Let  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a (measurable) cost, or payoff, function. The primal Optimal Transport problem corresponds to

$$P_{OT}(\mu, \nu) := \sup_{\mathbb{P} \in \mathcal{P}_2(\mu, \nu)} \mathbb{E}^{\mathbb{P}}[c(X, Y)], \quad \text{where } \mathcal{P}_2(\mu, \nu) := \{\mathbb{P} \in \mathcal{P}_2 : X \sim_{\mathbb{P}} \mu, Y \sim_{\mathbb{P}} \nu\}.$$

The above optimisation problem has a dual problem, which is defined by

$$D_{OT}(\mu, \nu) := \inf_{(\varphi, \psi) \in \mathcal{D}_s} \{\mu(\varphi) + \nu(\psi)\}, \quad \text{where } \mathcal{D}_s := \{(\varphi, \psi) : \mu(\varphi \vee 0) + \nu(\psi \vee 0) < \infty, \varphi \oplus \psi \geq c\}.$$

In the above,  $\varphi \oplus \psi(x, y) := \varphi(x) + \psi(y)$  and  $\chi(f) := \int f d\chi$ , for a (Borel) measure  $\chi$  on  $\mathbb{R}$  and a (Borel) measurable  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Under mild conditions, it can be proved that the duality  $P_{OT}(\mu, \nu) = D_{OT}(\mu, \nu)$  holds. (The notation  $\mathcal{D}_s$  is motivated by the fact that, in financial terms,  $(\varphi, \psi) \in \mathcal{D}_s$  corresponds to a *static* hedging strategy.)

The Brenier's theorem gives the explicit expressions of the optimizers of both the primal and dual problems. In the case  $\mu$  is atomless, a candidate primal optimizer is given by the Fréchet-Hoeffding (or quantile) coupling  $\mathbb{P}_*$  defined by

$$\mathbb{P}_*(dx, dy) := \mu(dx) \delta_{T_*(x)}(dy), \quad (\text{A.39})$$

push-forward map from  $\mu$  to  $\nu$  is where  $T_* := F_\nu^{-1} \circ F_\mu$ . Here  $F_\nu^{-1}$  is the right-continuous inverse of (the c.d.f. of  $\nu$ )  $F_\nu$ :  $F_\nu^{-1}(t) := \inf\{y : F_\nu > t\}$ .

We further introduce candidate dual optimizers  $\varphi_*$  and  $\psi_*$  (up to a constant) by setting

$$\varphi_*(x) := c(x, T_*(x)) - \psi_* \circ T_*(x), \quad \psi'_*(y) := c_y(T_*^{-1}(y), y), \quad \forall x, y \in \mathbb{R}. \quad (\text{A.40})$$

**Theorem A.1** (Brenier [18], Rachev and Rüschendorf [61]). *Suppose  $\mu$  has no atoms. Let  $c$  be an upper semi-continuous function with (at most) linear growth. Assume that the partial derivative  $c_{xy}$  exists and satisfies the Spence-Mirrlees condition  $c_{xy} > 0$ .*

*Let  $\mathbb{P}_*$  be as in (A.39),  $(\varphi_*, \psi_*)$  as in (A.40), and assume that  $(\mu(\varphi_* \vee 0) + \nu(\psi_* \vee 0)) < \infty$ . Then  $\mathbb{P}_* \in \mathcal{P}_2(\mu, \nu)$ ,  $(\varphi_*, \psi_*) \in \mathcal{D}_s$  and*

$$\int c(x, T_*(x)) \mu(dx) = P_{OT}(\mu, \nu) = D_{OT}(\mu, \nu) = \mu(\varphi_*) + \nu(\psi_*).$$

**Remark A.2.** *In the case  $c_{xy} < 0$ , then  $P_{OT} = D_{OT}$  and the optimal coupling is the antitone coupling which is supported on  $x \mapsto T^*(x) := F_\nu^{-1} \circ (1 - F_\mu(x))$ . The dual optimizers are still defined through (A.40), but with  $T^*$  in place of  $T_*$ .*

## A.2 Brenier's Theorem in Martingale Optimal Transport (MOT)

Two (integrable) probability measures  $\mu, \nu$  on  $\mathbb{R}$  are increasing in convex order, denoted by  $\mu \leq_c \nu$ , if for all convex functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ , we have that  $\mu(f) \leq \nu(f)$ . By the Strassen's classical result [62],  $\mu \leq_c \nu$  is equivalent to non-emptiness of the set  $\mathcal{M}_2(\mu, \nu) := \{\mathbb{P} \in \mathcal{P}_2(\mu, \nu) : \mathbb{E}^\mathbb{P}[Y|X] = X\}$ .

In this section we suppose that  $\mu$  and  $\nu$  have finite first moments,  $\mu \leq_c \nu$ , and the following Dispersion Assumption holds:

**Assumption A.3** (Dispersion).  $\mu, \nu$  admit densities  $f_\mu, f_\nu$  that satisfy

1.  $f_\mu > 0$  on  $(\ell_\mu, r_\mu)$ ,  $f_\nu > 0$  on  $(\ell_\nu, r_\nu)$ , and  $\ell_\nu \leq \ell_\mu < r_\mu \leq r_\nu$ .
2. There exists  $\ell_\mu < \underline{m} := \underline{m}^{\mu, \nu} < \overline{m} := \overline{m}^{\mu, \nu} < r_\mu$  such that

$$f_\mu > f_\nu, \text{ on } (\underline{m}, \overline{m}), \quad f_\nu > f_\mu \text{ on } (\ell_\nu, \underline{m}) \cup (\overline{m}, r_\nu), \quad f_\mu = f_\nu, \text{ on } \{\underline{m}, \overline{m}\}.$$

**Remark A.4.** Note that if  $\overline{\mu} = \overline{\nu}$  and Assumption A.3 holds, then  $\mu \leq_c \nu$ . Furthermore, the difference of distribution functions  $y \mapsto F_\nu(y) - F_\mu(y)$  attains the (unique) maximum and minimum at  $\underline{m}^{\mu, \nu}$  and  $\overline{m}^{\mu, \nu}$ , respectively.

Given a (measurable) payoff function  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ , the primal problem of MOT is defined by

$$P_{MOT}(\mu, \nu) := \sup_{\mathbb{P} \in \mathcal{M}_2(\mu, \nu)} \mathbb{E}^\mathbb{P}[c(X, Y)].$$

The MOT problem also has a dual problem:

$$D_{MOT}(\mu, \nu) := \inf_{(\varphi, \psi, h) \in \mathcal{D}_{ss}} \{\mu(\varphi) + \nu(\psi)\},$$

where the set of dual variables is defined as:

$$\mathcal{D}_{ss} := \{(\varphi, \psi, h) : \mu(\varphi \vee 0) + \nu(\psi \vee 0) < \infty, h : \mathbb{R} \rightarrow \mathbb{R}, \varphi \oplus \psi + h^\otimes \geq c\}$$

with  $\varphi \oplus \psi(x, y) := \varphi(x) + \psi(y)$ , and  $h^\otimes(x, y) := h(x)(y - x)$ . (The notation  $\mathcal{D}_{ss}$  is motivated by the fact that, in financial terms,  $(\varphi, \psi, h) \in \mathcal{D}_{ss}$  corresponds to a *semi-static* hedging strategy, where the static part is  $(\varphi, \psi)$  and the dynamic part corresponds to  $h$ .)

We now introduce candidate optimal quantities. We start with a pair of functions  $(T_d^{\mu, \nu}, T_u^{\mu, \nu})$  that will support the optimal coupling in the case  $\mu \leq_c \nu$  satisfy Assumption A.3.

Set

$$g^{\mu, \nu}(x, y) := F_\nu^{-1}(F_\mu(x) + F_\nu(y) - F_\mu(y)), \quad x \in (\ell_\mu, r_\mu), y \in (\ell_\nu, r_\nu). \quad (\text{A.41})$$

Define

$$T_d^{\mu, \nu}(x) = T_u^{\mu, \nu}(x) = x, \quad x \in [\overline{m}^{\mu, \nu}, r_\mu]. \quad (\text{A.42})$$

For  $x \in (\ell_\mu, \overline{m}^{\mu, \nu})$ , let  $T_d^{\mu, \nu}(x)$  be the (unique) scalar such that

$$\int_{-\infty}^x [F_\nu^{-1}(F_\mu(\xi)) - \xi] f_\mu(\xi) d\xi + \int_{-\infty}^{T_d(x)} \mathbf{1}_{(-\infty, \overline{m}^{\mu, \nu}]}(\xi) [g^{\mu, \nu}(x, \xi) - \xi] (f_\nu(\xi) - f_\mu(\xi)) d\xi = 0 \quad (\text{A.43})$$

and set

$$T_u^{\mu, \nu}(x) = g^{\mu, \nu}(x, T_d^{\mu, \nu}(x)), \quad x \in (\ell_\mu, \overline{m}^{\mu, \nu}). \quad (\text{A.44})$$

**Remark A.5.** The existence and uniqueness of (A.43) is given by Henry-Labordère and Touzi [42]. Furthermore,  $T_d(x) \leq x \leq T_u(x)$  for all  $x \in (\ell_\mu, r_\mu)$ , and, under Assumption A.3,  $T_d, T_u$  are both continuous,  $T_d$  (resp.  $T_u$ ) is strictly increasing (resp. decreasing) on  $(\ell_\mu, \overline{m}^{\mu, \nu}]$ . See Henry-Labordère and Touzi [42, Section 4] (note, however, that [42] constructed the left-curtain martingale coupling, while in the present paper we are working with the right-curtain coupling.)



Using the pair  $(T_d^{\mu,\nu}, T_u^{\mu,\nu})$  we now define a candidate martingale coupling through its disintegration w.r.t. the first marginal  $\mu$ . Define a measure  $\mathbb{P}^{\mu,\nu}$  on  $\mathbb{R}^2$  by

$$\begin{aligned} \mathbb{P}^{\mu,\nu}(dx, dy) &= \mu(dx) \mathbf{1}_{[\overline{m}^{\mu,\nu}, r_\mu)}(x) \delta_x(dy) \\ &\quad + \mu(dx) \mathbf{1}_{(\ell_\mu, \overline{m}^{\mu,\nu})}(x) \left[ q^{\mu,\nu}(x) \delta_{T_u^{\mu,\nu}(x)}(dy) + (1 - q^{\mu,\nu}(x)) \delta_{T_d^{\mu,\nu}(x)}(dy) \right], \end{aligned} \quad (\text{A.45})$$

where

$$q^{\mu,\nu}(x) := \frac{x - T_d^{\mu,\nu}(x)}{T_u^{\mu,\nu}(x) - T_d^{\mu,\nu}(x)}, \quad x \in (\ell_\mu, \overline{m}^{\mu,\nu}), \quad (\text{A.46})$$

corresponds to the (conditional) martingale probability of jumping to  $T_u^{\mu,\nu}(x)$  from  $x$ .

We now introduce a candidate triple  $(\varphi_*, \psi_*, h_*) = (\varphi_*^{\mu,\nu}, \psi_*^{\mu,\nu}, h_*^{\mu,\nu})$  for the dual problem. To ease the notation we will write  $(T_d, T_u) = (T_d^{\mu,\nu}, T_u^{\mu,\nu})$ .

Define  $h_*$ , up to a constant, by

$$\begin{aligned} h_* &= h_* \circ T_d^{-1} + c_y(\cdot, \cdot) - c_y(T_d^{-1}, \cdot) \quad \text{on } [\overline{m}^{\mu,\nu}, r_\mu), \\ h'_* &= \frac{c_x(\cdot, T_u) - c_x(\cdot, T_d)}{T_u - T_d} \quad \text{on } (\ell_\mu, \overline{m}^{\mu,\nu}). \end{aligned} \quad (\text{A.47})$$

(Note that the constant can be chosen such that  $h_*$  is continuous.) The (continuous) function  $\psi_*$  is then given (also up to a constant) by

$$\begin{aligned} \psi'_* &= c_y(T_d^{-1}, \cdot) - h_* \circ T_d^{-1} \quad \text{on } [\overline{m}^{\mu,\nu}, r_\mu), \\ \psi'_* &= c_y(T_u^{-1}, \cdot) - h_* \circ T_u^{-1} \quad \text{on } (\ell_\mu, \overline{m}^{\mu,\nu}). \end{aligned} \quad (\text{A.48})$$

Finally the function  $\varphi_*$  is defined by

$$\varphi_*(x) = \mathbb{E}^{\mathbb{P}^{\mu,\nu}}[c(X, Y) - \psi_*(Y) | X = x], \quad x \in (\ell_\mu, r_\mu). \quad (\text{A.49})$$

**Theorem A.6** (Henry-Labordère and Touzi [42, Theorems 4.5, 5.1]). *Suppose  $\mu \leq_c \nu$  satisfies Assumption A.3. Assume further that  $\int \varphi_* \vee 0 d\mu + \int \psi_* \vee 0 d\nu < \infty$ , and that the partial derivative  $c_{xyy}$  exists and  $c_{xyy} < 0$  on  $\mathbb{R}^2$ . Let  $\mathbb{P}_* = \mathbb{P}^{\mu,\nu}$  be as in (A.45) and  $(\varphi_*, \psi_*, h_*)$  as in (A.49), (A.48), (A.47).*

*Then  $\mathbb{P}_* \in \mathcal{M}_2(\mu, \nu)$ ,  $(\varphi_*, \psi_*, h_*) \in \mathcal{D}_{ss}$  and the strong duality holds:*

$$\mathbb{E}^{\mathbb{P}_*}[c(X, Y)] = P_{MOT}(\mu, \nu) = D_{MOT}(\mu, \nu) = \mu(\varphi_*) + \nu(\psi_*).$$

## B One-period Supermartingale Optimal Transport

### B.1 Solution to the primal SMOT problem and related properties

Let  $\mu, \nu \in \mathcal{P}$  be in convex-decreasing order, i.e.,  $\mu \leq_{cd} \nu$  (for two measures  $\eta, \chi$  on  $\mathbb{R}$  we also write  $\eta \leq \chi$  if  $\eta(A) \leq \chi(A)$  for all Borel  $A \subseteq \mathbb{R}$ ). Let  $\mathcal{S}_2(\mu, \nu) := \{\mathbb{P} \in \mathcal{P}_2(\mu, \nu) : \mathbb{E}^{\mathbb{P}}[Y|X] \leq X\}$  be the set of supermartingale couplings with given marginals  $\mu$  and  $\nu$ . The following theorem defines the decreasing supermartingale coupling  $\hat{\mathbb{P}} = \hat{\mathbb{P}}^{\mu,\nu} \in \mathcal{S}_2(\mu, \nu)$ .

**Theorem B.1** (Nutz and Stebegg [57]). *Suppose  $\mu \leq_{cd} \nu$  and let  $\hat{\mathbb{P}} \in \mathcal{S}_2(\mu, \nu)$  be the decreasing supermartingale coupling. It satisfies any, and then all, of the following properties*

1.  $\hat{\mathbb{P}}$  solves (2.1) whenever  $c$  satisfies  $c_{xyy} < 0$  and  $c_{xy} > 0$ ;
2. For each  $x \in \mathbb{R}$ ,  $\hat{\mathbb{P}}|_{[x, \infty) \times \mathbb{R}}$  is the smallest element (w.r.t.  $\leq_{cd}$ ) of  $\{\theta : \mu|_{[x, \infty)} \leq_{cd} \theta \leq \nu\}$ ;
3. There exists  $\Gamma \subseteq \mathbb{R}^2$  and  $M \subset \mathbb{R}$  such that  $\hat{\mathbb{P}}_*$  is second-order right-monotone and first-order left-monotone w.r.t.  $(\Gamma, M)$  in the sense of Definition B.2.

**Definition B.2.** Let  $(\Gamma, M) \in B(\mathbb{R}^2) \times B(\mathbb{R})$ . We say

1.  $\Gamma$  is second-order right-monotone if for all  $(x, y_1), (x, y_2), (x', y') \in \Gamma$  with  $x' < x$  we have that  $y' \notin (y_1, y_2)$ ;
2.  $(\Gamma, M)$  is first-order left-monotone if for all  $(x_1, y_1), (x_2, y_2) \in \Gamma$  with  $x_1 < x_2$  and  $x_2 \notin M$  we have that  $y_1 \leq y_2$ .

The second property of  $\hat{\mathbb{P}}$  in Theorem B.1 implies that, for each  $x \in \mathbb{R}$ ,  $\hat{\mathbb{P}}|_{[x, \infty) \times \mathbb{R}}$  is the shadow of  $\mu|_{[x, \infty)}$  in  $\nu$ : given  $\mu_0 \leq \mu \leq_{cd} \nu$ , the *shadow measure*, denoted by  $S^\nu(\mu_0)$ , is the smallest element (w.r.t.  $\leq_{cd}$ ) of  $\{\theta : \mu_0 \leq_{cd} \theta \leq \nu\}$ .

Furthermore, the set  $M$ , that appears in the third property of  $\hat{\mathbb{P}}$  in Theorem B.1, corresponds to the martingale points of  $\hat{\mathbb{P}}$  in the sense that  $\hat{\mathbb{P}}|_{M \times \mathbb{R}}$  is a martingale. The explicit constructions of  $\hat{\mathbb{P}}$  and  $M$  were provided by Bayraktar et. al. [5, Theorem 6.1, Proposition 5.1]. In particular, there exists  $(x_n)_{n \geq 1}$  in  $\mathbb{R}$  such that  $M = \bigcup_{n \geq 1} [x_{n+1}, x_n]$ , so that  $\hat{\mathbb{P}}$  alternates between being a martingale and supermartingale at most countably many times. For each  $n \geq 1$ , the transitions of  $\hat{\mathbb{P}}$  on  $[x_{n+1}, x_n]$  are of the right-curtain type (see (A.45)), while on  $\mathbb{R} \setminus M$ ,  $\hat{\mathbb{P}}$  mimics the transitions of the quantile coupling (see (A.39)).

For a measure  $\eta$  on  $\mathbb{R}$  let  $C_\eta : \mathbb{R} \rightarrow \mathbb{R}$  be given by  $C_\eta(k) := \int_{\mathbb{R}} (x - k)^+ \eta(dx)$ ,  $k \in \mathbb{R}$ . Define  $c^{\mu, \nu} : \mathbb{R} \rightarrow \mathbb{R}$  by

$$c^{\mu, \nu}(x) = \sup_{k \in \mathbb{R}} \{C_{\mu|_{[x, \infty)}}(k) - C_\nu(k)\}, \quad x \in \mathbb{R}.$$

Then, provided  $\mu$  is atom-less,  $c$  is non-negative non-decreasing, continuous and

$$c(x) = \overline{\mu|_{[x, \infty)}} - \overline{S^\nu(\mu|_{[x, \infty)})}.$$

See Bayraktar et. al. [5, Lemmas 3.1, 5.1]. Then the first point (starting from the right, i.e.,  $r_\mu$ , and moving to the left towards  $\ell_\mu$ ) where  $\hat{\mathbb{P}}$  transitions from being a martingale to supermartingale is given by

$$x_1^{\mu, \nu} = \inf\{x \in \mathbb{R} : c^{\mu, \nu}(x) = 0\} \quad \text{with} \quad \inf \emptyset = r_\mu. \quad (\text{B.50})$$

Furthermore, define

$$y_1^{\mu, \nu} = \ell_{S^\nu(\mu|_{[x_1^{\mu, \nu}, \infty)})}. \quad (\text{B.51})$$

Note that, since  $\mu|_{[x_1^{\mu, \nu}, \infty)} \leq_{cd} S^\nu(\mu|_{[x_1^{\mu, \nu}, \infty)})$ , we have that  $y_1^{\mu, \nu} \leq x_1^{\mu, \nu}$ .

**Lemma B.3.** Suppose that  $\mu \leq_{cd} \nu$  satisfy Assumption A.3. Let  $\hat{\mathbb{P}}$  be the decreasing supermartingale coupling of Theorem B.1.

Then the ( $\mu$ -a.s. unique) set of martingale points of  $\hat{\mathbb{P}}$  is given by  $M = [x_1^{\mu, \nu}, r_\mu]$ . Furthermore, the second marginal of  $\hat{\mathbb{P}}|_{M \times \mathbb{R}}$  is  $\nu|_{[y_1^{\mu, \nu}, r_\nu]}$ .

*Proof of Lemma B.3.* We assume that  $\bar{\mu} > \bar{\nu}$ , since otherwise  $\mu \leq_c \nu$ , and there is nothing to prove. Indeed, if  $\bar{\mu} = \bar{\nu}$ , then  $x_1^{\mu, \nu} = \ell_\mu$  and  $\hat{\mathbb{P}} = \mathbb{P}^{\mu, \nu}$  is the right-curtain martingale coupling; see (A.45).

Suppose that  $x_1^{\mu, \nu} \in (\ell_\mu, r_\mu]$ . Let  $\bar{m} := \bar{m}^{\mu, \nu}$ . Note that  $\mu|_{[\bar{m}, r_\mu]} \leq \nu$  and thus  $\mu|_{[\bar{m}, r_\mu]} = S^\nu(\mu|_{[\bar{m}, r_\mu]})$ . It follows that  $x_1^{\mu, \nu} \leq \bar{m}$ . In fact, since  $\mu$  and  $\nu - \mu|_{[\bar{m}, r_\mu]}$  both have strictly positive densities in the neighbourhood of  $\bar{m}$  and  $f_\mu > f_\nu$  to the left of  $\bar{m}$ , we must have that  $x_1^{\mu, \nu} < \bar{m}$ .

We now argue that  $S^\nu(\mu|_{[x_1^{\mu, \nu}, r_\mu]}) = \nu|_{[y_1^{\mu, \nu}, r_\nu]}$ . First, using the associativity of the shadow measure (see Bayraktar et al. [4]), we have that

$$S^\nu(\mu|_{[x_1^{\mu, \nu}, r_\mu]}) = \mu|_{[\bar{m}, r_\mu]} + S^{\nu - \mu|_{[\bar{m}, r_\mu]}}(\mu|_{[x_1^{\mu, \nu}, \bar{m})}).$$

Thanks to the properties of the shadow measure  $(S^{\nu-\mu|\overline{m},r_\mu})(\mu|_{[x_1^{\mu,\nu},\overline{m}]})$  has the smallest variance among all possible target laws of  $\mu|_{[x_1^{\mu,\nu},\overline{m}]}$  within  $\nu - \mu|_{[\overline{m},r_\mu]}$ , we have that

$$S^{\nu-\mu|\overline{m},r_\mu}(\mu|_{[x_1^{\mu,\nu},\overline{m}]}) = \nu|_{[y_1(\mu,\nu),\overline{m}]} + (\nu - \mu|_{[\overline{m},r_\mu]})(\overline{m},r_0)$$

for some  $r_0 \in (\overline{m}, r_\nu]$ . But if  $r_0 < r_\nu$ , then  $\nu - S^\nu(\mu|_{[x_1^{\mu,\nu},r_\mu]})$  charges  $[r_0, r_\nu]$ . Then, since  $x_1^{\mu,\nu} < \overline{m} < r_0 < r_\nu$ , we can find a small enough  $x < x_1^{\mu,\nu}$ , such that  $\mu|_{[x,x_1^{\mu,\nu}]} \leq_c S^{\nu-S^\nu(\mu|_{[x_1^{\mu,\nu},r_\mu]})}(\mu|_{[x,x_1^{\mu,\nu}]})$ , contradicting the minimality of  $x_1^{\mu,\nu}$ . It follows that  $r_0 = r_\nu$ .

It is left to show that  $x_1^{\mu,\nu}$  is the  $(\mu$ -a.s.) unique regime-switching point. For this we use the following observations. The transitions of the decreasing supermartingale coupling  $\hat{\mathbb{P}}$  is either those of the right-curtain (see (A.45)) or the quantile coupling (see (A.39)); see Nutz and Stebegg [57] or Bayraktar et al. [5]. In particular, by the definition of  $x_1^{\mu,\nu}$ , we must have that, (locally) to the left of  $x_1^{\mu,\nu}$ ,  $\hat{\mathbb{P}}$  corresponds to the quantile coupling of  $\mu|_{(-\infty, x_1^{\mu,\nu})}$  and  $\nu|_{(-\infty, y_1^{\mu,\nu})}$ , denoted by  $\pi^q$  (which is just a restriction of the quantile coupling of  $\mu$  and  $\nu$  to  $(-\infty, x_1^{\mu,\nu}) \times (-\infty, y_1^{\mu,\nu})$ ).

Under  $\pi^q$ , each  $x \in (\ell_\mu, x_1^{\mu,\nu})$  is mapped to  $F_\nu^{-1}(F_\mu(x))$ . Note that  $x \mapsto F_\nu^{-1}(F_\mu(x))$  is non-decreasing, continuous, and by construction (and the supermartingale requirement) we have that  $F_\nu^{-1}(F_\mu(x_1^{\mu,\nu})) = y_1^{\mu,\nu} \leq x_1^{\mu,\nu}$ . Let  $\bar{x}$  be defined by  $F_\mu(\bar{x}) = F_\nu(\bar{x})$ . Due to Assumption A.3,  $\bar{x}$  is unique and satisfies  $\bar{x} \in (\underline{m}, \overline{m})$ . In particular,  $F_\nu > F_\mu$  on  $(-\infty, \bar{x})$ , while  $F_\nu < F_\mu$  on  $(\bar{x}, \infty)$ .

If  $x_1^{\mu,\nu} \leq \bar{x}$ , then  $F_\nu^{-1}(F_\mu(x)) < x$  for all  $x < x_1^{\mu,\nu}$ , and the quantile coupling provides strict supermartingale transitions on  $(-\infty, x_1^{\mu,\nu})$ . Then using characterization of  $\hat{P}_*$  in terms of the shadow measure, it follows that  $\hat{\mathbb{P}} = \pi^q$  on  $(-\infty, x_1^{\mu,\nu})$ , which proves our claim. On the other hand,  $x_1^{\mu,\nu} > \bar{x}$  cannot happen, since then  $y_1^{\mu,\nu} = F_\nu^{-1}(F_\mu(x_1^{\mu,\nu})) > x_1^{\mu,\nu}$ , a contradiction.  $\square$

Finally, provided that  $\mu$  is atomless and using the construction of Bayraktar et al. [5], we have that the decreasing supermartingale coupling is supported on the graph of two functions  $\hat{T}_d = \hat{T}_d^{\mu,\nu}, \hat{T}_u = \hat{T}_u^{\mu,\nu} : (\ell_\mu, r_\mu) \rightarrow (\ell_\nu, r_\nu)$ . If, in addition, the Dispersion Assumption holds, then using Lemma B.3 and utilising the construction and properties of the right-curtain martingale coupling  $P^{\mu,\nu}$  (as in (A.45)) we have that  $(\hat{T}_d, \hat{T}_u)$  satisfies some useful properties.

First,

$$\hat{T}_d(x) \leq x \leq \hat{T}_u(x) \text{ for all } x \in (\ell_\mu, r_\mu), \quad \text{and} \quad \hat{T}_d = \hat{T}_u \text{ on } [\overline{m}, r_\mu]. \quad (\text{B.52})$$

Furthermore,

$$\begin{aligned} \hat{T}_d &\text{ is continuous and strictly increasing on } (\ell_\mu, r_\mu), \\ \hat{T}_d(x) &= F_\nu^{-1} \circ F_\mu \text{ on } (\ell_\mu, x_1^{\mu,\nu}) \end{aligned} \quad (\text{B.53})$$

and

$$\begin{aligned} \hat{T}_u &\text{ is continuous on } (x_1^{\mu,\nu}, r_\mu), \text{ strictly decreasing on } (x_1^{\mu,\nu}, \overline{m}), \\ \hat{T}_u &\equiv \infty \text{ on } (\ell_\mu, x_1^{\mu,\nu}]. \end{aligned} \quad (\text{B.54})$$

In particular, in terms of  $(\hat{T}_d, \hat{T}_u)$ , the decreasing supermartingale coupling  $\mathbb{P}^{\mu,\nu}$  is given by

$$\begin{aligned} \hat{\mathbb{P}}(dx, dy) &= \mu(dx) \mathbf{1}_{(\ell_\mu, x_1^{\mu,\nu}] \cup [\overline{m}, r_\mu)}(x) \delta_{\hat{T}_d(x)}(dy) \\ &\quad + \mu(dx) \mathbf{1}_{(x_1^{\mu,\nu}, \overline{m})} \left\{ \frac{\hat{T}_u(x) - x}{\hat{T}_u(x) - \hat{T}_d(x)} \delta_{\hat{T}_d(x)}(dy) + \frac{x - \hat{T}_d(x)}{\hat{T}_u(x) - \hat{T}_d(x)} \delta_{\hat{T}_u(x)}(dy) \right\}. \end{aligned} \quad (\text{B.55})$$

Equivalently, we can express  $\hat{\mathbb{P}}$  in terms of the right-curtain coupling  $\mathbb{P}^{\mu|_{[x_1^{\mu,\nu}, r_\mu]}, \nu|_{[y_1^{\mu,\nu}, r_\nu]}}$  of  $\mu|_{[x_1^{\mu,\nu}, r_\mu]}$  and  $\nu|_{[y_1^{\mu,\nu}, r_\nu]}$  (see (A.45)):

$$\hat{\mathbb{P}}(dx, dy) = \mu(dx) \mathbf{1}_{(\ell_\mu, x_1^{\mu,\nu}]}(x) \delta_{F_\nu^{-1}(F_\mu(x))}(dy) + \mathbf{1}_{(x_1^{\mu,\nu}, r_\mu)} \mathbb{P}^{\mu|_{[x_1^{\mu,\nu}, r_\mu]}, \nu|_{[y_1^{\mu,\nu}, r_\nu]}}(dx, dy). \quad (\text{B.56})$$

While the pair  $(\hat{T}_d, \hat{T}_u)$  can be explicitly constructed using potential functions of the underlying measures (see Bayraktar et al. [5]), the following representation is more useful in the present context.

Recall the definition of  $g = g_{\mu, \nu}$  (see (A.41)).

**Proposition B.4.** *For each  $x \in [x_1^{\mu, \nu}, \bar{m}^{\mu, \nu})$ ,  $\hat{T}_u(x)$  is uniquely characterised by the integral equation:*

$$\int_x^\infty [F_\nu^{-1}(F_\mu(\xi)) - \xi] f_\mu(\xi) d\xi + \int_{\hat{T}_u(x)}^\infty \mathbf{1}_{(x_1^{\mu, \nu}, \infty)}(\xi) [g(x, \xi) - \xi] (f_\nu(\xi) - f_\mu(\xi)) d\xi = 0,$$

and  $\hat{T}_d(x) = g(x, \hat{T}_u(x))$ .

*Proof.* Since  $\mu|_{[x_1^{\mu, \nu}, r_\mu] \leq c} \nu|_{[y_1(\mu, \nu), r_\nu]}$  and the decreasing supermartingale coupling  $\hat{\mathbb{P}}^{\mu, \nu}$  of  $\mu \leq_{cd} \nu$  coincides with the right-curtain coupling of  $\mu|_{[x_1^{\mu, \nu}, r_\mu] \leq c} \nu|_{[y_1(\mu, \nu), r_\nu]}$  on  $[x_1^{\mu, \nu}, r_\mu)$ , the result follows immediately from the construction of the latter, which is given in Section A.2 (see (A.43) and (A.44)).  $\square$

## B.2 Optimal dual strategy for decreasing SMOT

We introduce now the optimal dual strategies to the dual problem (2.5) in the case the decreasing supermartingale coupling  $\hat{\mathbb{P}} = \hat{\mathbb{P}}^{\mu, \nu}$  (see Theorem B.1) is optimal. Recall that the pair  $(\hat{T}_d, \hat{T}_u)$  represents the supporting functions of  $\hat{\mathbb{P}}$ , while  $x_1 = x_1^{\mu, \nu}$  is the unique threshold that separates martingale and supermartingale regions.

In the following we suppose that  $\mu \leq_{cd} \nu$  satisfy Assumption A.3.

We introduce the following triple  $(\hat{\varphi}, \hat{\psi}, \hat{h}) = (\hat{\varphi}(\mu, \nu), \hat{\psi}(\mu, \nu), \hat{h}(\mu, \nu))$ , as the candidate optimal dual strategies for  $\mathbf{D}_2(\mu, \nu)$  (with  $\mathbf{D}_2$  defined in (2.2)):

$$\begin{aligned} \hat{h}'(x) &:= \frac{c_x(x, \hat{T}_u(x)) - c_x(x, \hat{T}_d(x))}{\hat{T}_u(x) - \hat{T}_d(x)}, \quad x_1 < x < \bar{m}; \\ \lim_{x \downarrow x_1} \hat{h}(x) &= 0; \\ \hat{h}(x) &:= \hat{h}((\hat{T}_u)^{-1}(x)) - c_y((\hat{T}_u)^{-1}(x), x), \quad x \geq \bar{m}; \\ \hat{h}(x) &:= 0, \quad x \leq x_1. \end{aligned} \tag{B.57}$$

Note that, when  $x = x_1$ ,  $\hat{T}_u(x) = \infty$  and therefore  $\hat{h}'(x) = 0$ . Furthermore,  $\hat{h}$  is continuous at  $x \in \{x_1, \bar{m}\}$  (note that  $c_y(y, y) = 0$  and  $\lim_{x \rightarrow \bar{m}} (\hat{T}_u)^{-1}(x) = \bar{m}$ ).

Denoting  $L^{-1}(x) := (\hat{T}_u)^{-1}(x) \mathbf{1}_{\{x \geq \bar{m}\}} + (\hat{T}_d)^{-1}(x) \mathbf{1}_{\{x < \bar{m}\}}$ , we introduce  $\hat{\psi}$ , up to a constant, by

$$\hat{\psi}'(x) := c_y(L^{-1}(x), x) - \hat{h}(L^{-1}(x)).$$

Note that  $\hat{\psi}$  can be chosen to be continuous, which then also makes

$$c(\cdot, \hat{T}_u(\cdot)) - \hat{\psi}(\hat{T}_u(\cdot)) - c(\cdot, \hat{T}_d(\cdot)) + \hat{\psi}(\hat{T}_d(\cdot)) - (\hat{T}_u(\cdot) - \hat{T}_d(\cdot)) \hat{h}(\cdot)$$

continuous.

Finally, we define

$$\hat{\varphi}(x) := \mathbb{E}^{\hat{\mathbb{P}}} [c(X, Y) - \psi(Y) | X = x], \quad x \in \mathbb{R}.$$

Note that, for  $x_1 \leq x \leq \bar{m}$

$$\hat{\varphi}(x) = \frac{x - \hat{T}_d(x)}{\hat{T}_u(x) - \hat{T}_d(x)} \left( c(x, \hat{T}_u(x)) - \hat{\psi}(\hat{T}_u(x)) \right) + \frac{\hat{T}_u(x) - x}{\hat{T}_u(x) - \hat{T}_d(x)} \left( c(x, \hat{T}_d(x)) - \hat{\psi}(\hat{T}_d(x)) \right),$$

$$\hat{\varphi}(x) := c(x, x) - \hat{\psi}(x) = -\hat{\psi}(x), \quad x > \overline{m}$$

and

$$\hat{\varphi}(x) := c(x, \hat{T}_d(x)) - \hat{\psi}(\hat{T}_d(x)), \quad x < x_1.$$

From our definitions, we immediately have that  $\mathbb{E}^{\hat{\mathbb{P}}}[c(X, Y)] = \mu(\hat{\varphi}(X)) + \nu(\hat{\psi}(Y))$ . It is left to show that the triple  $(\hat{\varphi}, \hat{\psi}, \hat{h})$  defines a superhedge.

**Lemma B.5.** *Let  $\mu \leq_{cd} \nu$  be such that Assumption A.3 holds. Suppose that the reward function  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$  satisfy Assumption 2.2. Assume further that  $\int \hat{\varphi} \vee 0 d\mu + \int \hat{\psi} \vee 0 d\nu < \infty$ .*

*We have that  $(\hat{\varphi}, \hat{\psi}, \hat{h}) \in \mathcal{D}_2(\mu, \nu)$ .*

*Proof.* To obtain the proof we need to show that  $\hat{h} \geq 0$  and that  $c(x, y) \leq \hat{\varphi}(x) + \hat{\psi}(y) + \hat{h}(x)(y - x)$  for all  $x, y \in \mathbb{R}$ . The second point, however, follows immediately from the arguments of the classical and martingale versions of Brenier's theorem (see Theorems A.1 and A.6). We now verify that  $\hat{h} \geq 0$ ; we will use that  $c(\cdot, \cdot) = c_y(\cdot, \cdot) = 0$  and  $c_{xy} > 0$ .

First, as  $c_{xy} > 0$ , we have that when  $x_1 \leq x < \overline{m}$ ,

$$\hat{h}'(x) := \frac{c_x(x, \hat{T}_u(x)) - c_x(x, \hat{T}_d(x))}{\hat{T}_u(x) - \hat{T}_d(x)} > 0.$$

Hence, since  $\hat{h}(x_1) = 0$ , we have that  $\hat{h}(x) \geq 0$  for all  $x_1 \leq x < \overline{m}$ .

For  $x \geq \overline{m}$ , as  $x_1 \leq (\hat{T}_u)^{-1}(x) \leq \overline{m}$ , we have that  $\hat{h}((\hat{T}_u)^{-1}(x)) \geq 0$ . In addition, as  $c_y(x, x) = 0$  and  $c_{xy} > 0$ , we have that  $c_y((\hat{T}_u)^{-1}(x), x) \leq 0$ . It follows that

$$\hat{h}(x) = \hat{h}((\hat{T}_u)^{-1}(x)) - c_y((\hat{T}_u)^{-1}(x), x) \geq 0.$$

Since  $\hat{h} \equiv 0$  on  $(-\infty, x_1)$ , the no short-selling constraint is satisfied.  $\square$

## References

- [1] Beatrice Acciaio, Mathias Beiglböck, and Gudmund Pammer. Weak transport for non-convex costs and model-independence in a fixed-income market. *Mathematical Finance*, 31(4):1423–1453, 2021.
- [2] Julio Backhoff-Veraguas, Mathias Beiglböck, Martin Huesmann, and Sigrid Källblad. Martingale benamou–brenier: a probabilistic perspective. *The Annals of Probability*, 48(5):2258–2289, 2020.
- [3] Julio Backhoff-Veraguas and Gudmund Pammer. Stability of martingale optimal transport and weak optimal transport. *The Annals of Applied Probability*, 32(1):721–752, 2022.
- [4] Erhan Bayraktar, Shuoqing Deng, and Dominykas Norgilas. A potential-based construction of the increasing supermartingale coupling. *arXiv preprint arXiv:2108.03450 (to appear in the Annals of Applied Probability)*, 2021.
- [5] Erhan Bayraktar, Shuoqing Deng, and Dominykas Norgilas. Supermartingale shadow couplings: the decreasing case. *arXiv preprint arXiv:2207.11732*, 2022.
- [6] Mathias Beiglböck, Alexander MG Cox, and Martin Huesmann. Optimal transport and skorokhod embedding. *Inventiones mathematicae*, 208(2):327–400, 2017.
- [7] Mathias Beiglböck, Alexander MG Cox, and Martin Huesmann. The geometry of multi-marginal skorokhod embedding. *Probability Theory and Related Fields*, 176(3):1045–1096, 2020.

- [8] Mathias Beiglböck, Pierre Henry-Labordere, and Friedrich Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.
- [9] Mathias Beiglböck, Pierre Henry-Labordère, and Nizar Touzi. Monotone martingale transport plans and skorokhod embedding. *Stochastic Processes and their Applications*, 127(9):3005–3013, 2017.
- [10] Mathias Beiglböck, David Hobson, and Dominykas Norgilas. The potential of the shadow measure. *Electronic Communications in Probability*, 27:1–12, 2022.
- [11] Mathias Beiglböck, Benjamin Jourdain, William Margheriti, and Gudmund Pammer. Approximation of martingale couplings on the line in the weak adapted topology. *arXiv preprint arXiv:2101.02517*, 2021.
- [12] Mathias Beiglböck and Nicolas Juillet. On a problem of optimal transport under marginal martingale constraints. *The Annals of Probability*, 44(1):42–106, 2016.
- [13] Mathias Beiglböck and Nicolas Juillet. Shadow couplings. *Transactions of the American Mathematical Society*, 374(7):4973–5002, 2021.
- [14] Mathias Beiglböck and Marcel Nutz. Martingale inequalities and deterministic counterparts. *Electronic Journal of Probability*, 19:1–15, 2014.
- [15] Mathias Beiglböck, Marcel Nutz, and Florian Stebegg. Fine properties of the optimal skorokhod embedding problem. *Journal of the European Mathematical Society*, 24(4):1389–1429, 2021.
- [16] Mathias Beiglböck, Marcel Nutz, and Nizar Touzi. Complete duality for martingale optimal transport on the line. *The Annals of Probability*, 45(5):3038–3074, 2017.
- [17] Douglas T Breeden and Robert H Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651, 1978.
- [18] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.
- [19] Haydyn Brown, David Hobson, and Leonard CG Rogers. Robust hedging of barrier options. *Mathematical Finance*, 11(3):285–314, 2001.
- [20] Martin Brücknerhoff, Martin Huesmann, and Nicolas Juillet. Shadow martingales—a stochastic mass transport approach to the peacock problem. *Electronic Journal of Probability*, 27:1–62, 2022.
- [21] Luciano Campi, Ismail Laachir, and Claude Martini. Change of numeraire in the two-marginals martingale transport problem. *Finance and Stochastics*, 21(2):471–486, 2017.
- [22] Peter Carr, Hélyette Geman, Dilip B Madan 5, and Marc Yor. From local volatility to local lévy models. *Quantitative Finance*, 4(5):581–588, 2004.
- [23] Alexander MG Cox, David Hobson, and Jan Oblój. Pathwise inequalities for local time: applications to skorokhod embeddings and optimal stopping. *The Annals of Applied Probability*, 18(5):1870–1896, 2008.
- [24] Alexander MG Cox and David G Hobson. Skorokhod embeddings, minimality and non-centred target distributions. *Probability Theory and Related Fields*, 135(3):395–414, 2006.
- [25] Alexander MG Cox and Jan Oblój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 15(3):573–605, 2011.
- [26] Alexander MG Cox and Jan Oblój. On joint distributions of the maximum, minimum and terminal value of a continuous uniformly integrable martingale. *Stochastic Processes and their Applications*, 125(8):3280–3300, 2015.

- [27] Alexander MG Cox, Jan Oblój, and Nizar Touzi. The root solution to the multi-marginal embedding problem: an optimal stopping and time-reversal approach. *Probability Theory and Related Fields*, 173(1):211–259, 2019.
- [28] Alexander MG Cox and Jiajie Wang. Root’s barrier: Construction, optimality and applications to variance options. *The Annals of Applied Probability*, 23(3):859–894, 2013.
- [29] Mark Davis, Jan Oblój, and Vimal Raval. Arbitrage bounds for prices of weighted variance swaps. *Mathematical Finance*, 24(4):821–854, 2014.
- [30] Stefano De Marco and Pierre Henry-Labordere. Linking vanillas and vix options: a constrained martingale optimal transport problem. *SIAM Journal on Financial Mathematics*, 6(1):1171–1194, 2015.
- [31] Yan Dolinsky and H Mete Soner. Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields*, 160(1):391–427, 2014.
- [32] Yan Dolinsky and H Mete Soner. Martingale optimal transport in the skorokhod space. *Stochastic Processes and their Applications*, 125(10):3893–3931, 2015.
- [33] Christian-Oliver Ewald and Marc Yor. On peacocks and lyrebirds: Australian options, brownian bridges, and the average of submartingales. *Mathematical Finance*, 28(2):536–549, 2018.
- [34] Arash Fahim and Yu-Jui Huang. Model-independent superhedging under portfolio constraints. *Finance and Stochastics*, 20(1):51–81, 2016.
- [35] Hans Föllmer. Calcul d’itô sans probabilités. In *Séminaire de Probabilités XV 1979/80*, pages 143–150. Springer, 1981.
- [36] Alfred Galichon, Pierre Henry-Labordere, and Nizar Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312–336, 2014.
- [37] Paul Gassiat, Harald Oberhauser, and Gonçalo Dos Reis. Root’s barrier, viscosity solutions of obstacle problems and reflected fbsdes. *Stochastic Processes and their Applications*, 125(12):4601–4631, 2015.
- [38] Gaoyue Guo, Xiaolu Tan, and Nizar Touzi. On the monotonicity principle of optimal skorokhod embedding problem. *SIAM Journal on Control and Optimization*, 54(5):2478–2489, 2016.
- [39] Gaoyue Guo, Xiaolu Tan, and Nizar Touzi. Optimal skorokhod embedding under finitely many marginal constraints. *SIAM Journal on Control and Optimization*, 54(4):2174–2201, 2016.
- [40] Gaoyue Guo, Xiaolu Tan, and Nizar Touzi. Tightness and duality of martingale transport on the skorokhod space. *Stochastic Processes and their Applications*, 127(3):927–956, 2017.
- [41] Pierre Henry-Labordère, Xiaolu Tan, and Nizar Touzi. An explicit martingale version of the one-dimensional brenier’s theorem with full marginals constraint. *Stochastic Processes and their Applications*, 126(9):2800–2834, 2016.
- [42] Pierre Henry-Labordère and Nizar Touzi. An explicit martingale version of the one-dimensional brenier theorem. *Finance and Stochastics*, 20(3):635–668, 2016.
- [43] Francis Hirsch, Christophe Profeta, Bernard Roynette, and Marc Yor. *Peacocks and associated martingales, with explicit constructions*. Springer Science & Business Media, 2011.
- [44] David Hobson. The skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton lectures on mathematical finance 2010*, pages 267–318. Springer, 2011.



- [45] David Hobson and Martin Klimmek. Model-independent hedging strategies for variance swaps. *Finance and Stochastics*, 16(4):611–649, 2012.
- [46] David Hobson and Anthony Neuberger. Robust bounds for forward start options. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 22(1):31–56, 2012.
- [47] David Hobson and Dominykas Norgilas. Robust bounds for the american put. *Finance and Stochastics*, 23(2):359–395, 2019.
- [48] David Hobson and Dominykas Norgilas. A construction of the left-curtain coupling. *Electronic Journal of Probability*, 27:1–46, 2022.
- [49] David G Hobson. Robust hedging of the lookback option. *Finance and Stochastics*, 2(4):329–347, 1998.
- [50] David G Hobson and Dominykas Norgilas. The left-curtain martingale coupling in the presence of atoms. *The Annals of Applied Probability*, 29(3):1904–1928, 2019.
- [51] Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- [52] Nicolas Juillet. Martingales associated to peacocks using the curtain coupling. *Electronic Journal of Probability*, 23:1–29, 2018.
- [53] Sigrid Källblad, Xiaolu Tan, and Nizar Touzi. Optimal skorokhod embedding given full marginals and azéma–yor peacocks. *The Annals of Applied Probability*, 27(2):686–719, 2017.
- [54] Rajeeva L Karandikar. On pathwise stochastic integration. *Stochastic Processes and their applications*, 57(1):11–18, 1995.
- [55] Hans G Kellerer. Markov-komposition und eine anwendung auf martingale. *Mathematische Annalen*, 198(3):99–122, 1972.
- [56] Dilip B Madan and Marc Yor. Making markov martingales meet marginals: with explicit constructions. *Bernoulli*, pages 509–536, 2002.
- [57] Marcel Nutz and Florian Stebegg. Canonical supermartingale couplings. *The Annals of Probability*, 46(6):3351–3398, 2018.
- [58] Marcel Nutz, Florian Stebegg, and Xiaowei Tan. Multiperiod martingale transport. *Stochastic Processes and their Applications*, 130(3):1568–1615, 2020.
- [59] Jan Oblój. The skorokhod embedding problem and its offspring. *Probability Surveys*, 1:321–392, 2004.
- [60] Jan Oblój and Peter Spoida. An iterated azéma–yor type embedding for finitely many marginals. *arXiv preprint arXiv:1304.0368*, 2013.
- [61] Svetlozar T Rachev and Ludger Rüschendorf. *Mass Transportation Problems: Volume I: Theory*, volume 1. Springer Science & Business Media, 1998.
- [62] Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, 1965.
- [63] Johannes Wiesel. Continuity of the martingale optimal transport problem on the real line. *arXiv preprint arXiv:1905.04574*, 2019.