An Explicit Martingale Version of the One-dimensional Brenier's Theorem with Full Marginals Constraint *

Pierre Henry-Labordère [†]   Xiaolu Tan [‡]   Nizar Touzi[§]

February 29, 2016

Abstract

We provide an extension of the martingale version of the Fréchet-Hoeffding coupling to the infinitely-many marginals constraints setting. In the two-marginal context, this extension was obtained by Beiglböck & Juillet [7], and further developed by Henry-Labordère & Touzi [40], see also [6].

Our main result applies to a special class of reward functions and requires some restrictions on the marginal distributions. We show that the optimal martingale transference plan is induced by a pure downward jump local Lévy model. In particular, this provides a new martingale peacock process (PCOC "Processus Croissant pour l'Ordre Convexe," see Hirsch, Profeta, Roynette & Yor [43]), and a new remarkable example of discontinuous fake Brownian motions. Further, as in [40], we also provide a duality result together with the corresponding dual optimizer in explicit form.

As an application to financial mathematics, our results give the model-independent optimal lower and upper bounds for variance swaps.

1 Introduction

The classical optimal transport (OT) problem was initially formulated by Monge in his treatise "Théorie des déblais et des remblais" as follows. Let μ_0, μ_1 be two probability measures on $\mathbb{R}^d$, $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a cost function, then the optimal transport problem consists in minimizing the cost $\int_{\mathbb{R}^d} c(x, T(x)) \mu_0(dx)$ among all transference plans, i.e. all measurable functions $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $\mu_1 = \mu_0 \circ T^{-1}$. The relaxed formulation of the problem, as introduced by Kantorovich, consists in minimizing 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 so-called Spence-Mirrlees or Twist condition, the optimal Monge transference plan is characterized by the Brenier Theorem, and explicitly given by the Fréchet-Hoefding in the one-dimensioanl setting. We refer to Rachev & Ruschendorf [71] and Villani [75] for a detailed presentation.

The theory has been extended to the multiple marginals case by Gangbo & Święch [33], Carlier [15], Olkin & Rachev [65], Knott & Smith [60], Rüschendorf & Uckelmann [72], Heinich [38], and Pass [67, 68, 69], etc. We also refer to the full-marginals case addressed by Pass [70].

Recently, a martingale transportation (MT) problem was introduced in Beiglböck, Henry-Labordère & Penkner [5] and in Galichon, Henry-Labordère & Touzi [32]. Given two probability measures μ_0 and μ_1, the problem consists in minimizing some expected cost among all

probability measures $\mathbb{P}$ with fixed marginals $\mathbb{P} \circ X_0^{-1} = \mu_0$, $\mathbb{P} \circ X_1^{-1} = \mu_1$, and such that the canonical process X is a $\mathbb{P}-$martingale.

This new optimal transport problem is motivated by the problem of robust subhedging exotic options in a frictionless financial market allowing for trading the underlying asset and the corresponding vanilla options for the maturities 0 and 1. As observed by Breeden & Litzenberger [12], the market values of vanilla options for all strikes allows to recover the marginal distributions of the underlying asset price. This suggests a dual formulation of the robust superhedging problem defined as the minimization of the $\mathbb{P}-$expected payoff of the exotic option over all martingale measures $\mathbb{P}$ satisfying the marginal distribution constraint.

Based on the fact that any martingale can be represented as a time-changed Brownian motion, this problem was initially studied in the seminal paper of Hobson [44] by means of the Skorokhod Embedding Problem (SEP) approach, which consists in finding a stopping time τ of Brownian motion B such that B_τ has some given distribution. This methodology generated developments in many directions, namely for different derivative contracts and/or multiple-marginals constraints, see e.g. Brown, Hobson & Rogers [13], Madan & Yor [62], Cox, Hobson & Oblój [18], Cox & Oblój [19, 20], Davis, Oblój & Raval [23], Cox & Wang [22], Gassiat, Oberhauser & dos Reis [34], Cox, Oblój & Touzi [21], Hobson & Neuberger [51], and Hobson & Klimmek [47, 48, 49, 50]. We also refer to the survey papers by Oblój [63] and Hobson [45] for more details.

Recently, a rich literature has emerged around the martingale optimal transport approach to robust hedging. For models in discrete-time, we refer to Acciaio, Beiglböck, Penkner & Schachermayer[1], Beiglböck & Nutz [8], Beiglböck, Henry-Labordère & Touzi [6], Bouchard & Nutz [11], Campi, Laachir & Martini [14], Fahim & Huang [28], De Marco & Henry-Labordère [24]. For models in continuous-time, we refer to Biagini, Bouchard, Kardaras & Nutz [9], Dolinsky & Soner [25, 26, 27], Juillet [56], Henry-Labordère, Obloj, Spoida & Touzi [39], Källblad, Tan & Touzi [57], Stebegg [73], Bonnans & Tan [10], and Tan & Touzi [74]. We finally mention the work by Beiglböck, Cox & Huesmann [3] which derives new results on the Skorohod embedding problem by using the martingale transport approach, see also Beiglböck, Cox, Huesmann, Perkovsky & Promel [4], and Guo, Tan & Touzi [35, 36].

In the context of a one-period and one-dimensional martingale transport problem, Beiglböck & Juillet [7] introduced the left/right monotone martingale transference plan by formulating a martingale version of the so-called cyclic monotonicity in optimal transport theory. When the starting measure μ_0 has no atoms, the left/right monotone martingale transference is induced by a binomial model, called left/right curtain. More importantly, it is proved in [7] that such a left/right monotone transference plan exists and is unique, see also Beiglböck, Henry-Labordère & Touzi [6] for an alternative argument.

Under the additional conditions that the measures μ_0 and μ_1 are atomless, and the corresponding difference of c.d.f. has essentially finitely many local maximizers, Henry-Labordère & Touzi [40] provided an explicit construction of this left/right-monotone martingale transference plan, which extends the Fréchet-Hoeffding coupling in standard one-dimensional optimal transport. Moreover, they obtained an explicit expression of the solution of the dual problem, and hence by the duality result, they showed the optimality of their constructed transference plan for a class of cost/reward functions. An immediate extension to the multiple marginals case follows for a family of cost/reward functions.

In this paper, we are interested in the continuous-time case, as the limit of the multiple marginals MT problem. Let $(\mu_t)_{0 \le t \le 1}$ be a given family of probability measures on $\mathbb{R}$ which is non-decreasing in convex ordering, and assume that any pair $(\mu_t, \mu_{t+\varepsilon})$ of measures satisfies the technical conditions in [40]. Then every discretization of the time interval $[0, 1]$ induces a finite number of marginal constraints. Following the construction in [40], there is a binomial model calibrated to the corresponding multiple marginal distributions, which is of course optimal for

a class of cost/reward functions. Two natural questions can then be addressed. The first is whether the discrete binomial process converges when the time step converges to zero, and the second is whether such a limiting continuous-time process is optimal for a corresponding MT problem with full marginals.

Given a continuous family of marginal distributions which is non-decreasing in convex ordering, a stochastic process calibrated to all the marginals is called a peacock (or PCOC "Processus Croissant pour l'Ordre Convexe" in French) in Hirsch, Profeta, Roynette & Yor [43]. It follows by Kellerer's theorem that a process is a peacock if and only if there is a martingale with the same marginal distributions at each time, it is then interesting to construct such martingales associated with a given peacock (or equivalently with a given family of marginal distributions). In particular, when the marginal distributions are given by those of a Brownian motion, such a martingale is called a fake Brownian motion. Some examples of martingale peacock (or fake Brownian motion) have been provided by Albin [2], Fan, Hamza & Klebaner [29], Hamza & Klebaner [37], Hirsch et al. [42], Hobson [46], Oleszkiewicz [64], Pagès [66].

Our procedure gives a new construction of martingales associated with some peacock processes, under technical conditions (see Assumption 3.1 below), and in particular a discontinuous fake Brownian motion. Moreover, assuming that the difference of c.d.f. has exactly one local maximizer and some additional conditions (see Assumption 3.4 below), our constructed martingale is optimal among all martingales with given marginal distributions for a class of cost/reward functions, i.e. it solves a martingale transportation problem.

We would like to observe that our procedure depends on the explicit construction of the left-monotone martingale transference plan in [40] under restrictive conditions. This is in contrast with the recent parallel work of Juillet [56], where the convergence of the left-monotone martingale transference plan is investigated under general conditions and/or in various specific situations.

This paper is organized as follows. In Section 2, we recall the discrete-time martingale transport problem under finitely many marginal constraints. Taking the limit as the time step goes to 0, this leads naturally to a continuous-time MT problem under full marginals constraints. which is solved in Section 3 under technical conditions. Namely, by taking the limit of the optimal martingale measure for the multi-marginals MT problem, we obtain a continuous-time martingale calibrated to the given marginals, or equivalently, a martingale associated with the corresponding peacock. Under additional conditions, including the unique local maximizer of the c.d.f. difference, we prove that this limit martingale is a local Lévy process and solves the infinitely-many marginals MT problem for a class of cost/reward functions. In particular, we provide an explicit characterization of this optimal solution as well as the dual optimizer. In Section 4, we discuss some examples of extremal peacock following our construction, including a discontinuous fake Brownian motion and a self-similar martingale. As an application in finance, we provide an optimal robust hedging strategy for the variance swap option in Section 5. Finally, we complete the proofs of our main results in Section 6, where the main idea is to approximate the infinitely-many marginals case by the multi marginals case.

2 A martingale optimal transport problem

We first consider the martingale optimal transport problem in finite discrete time, and recall some results from [7, 40, 6]. By taking the limit, it leads to a continuous-time martingale transport problem.

2.1 Discrete-time martingale transport

Let μ_0, μ_1 be two atomless Borel probability measures on $\mathbb{R}$ with extreme left-point and right-point of the support $-\infty \le l_{\mu_i} < r_{\mu_i} \le +\infty$. We assume that μ_0, μ_1 have finite first moments and $\mu_0 \preceq \mu_1$ in the convex order, i.e. $\mu_0(\phi) \le \mu_1(\phi)$ for every convex function ϕ, where $\mu_i(\phi) := \int_{\mathbb{R}} \phi(x)\mu_i(dx)$ for $i = 0, 1$.[1]

Let $\mathcal{P}_{\mathbb{R}^2}$ be the collection of of all Borel probability measures on $\mathbb{R}^2$. The corresponding canonical process is denoted by (X_0, X_1), i.e. $X_i(x_0, x_1) = x_i$ for $i = 0, 1$ and $(x_0, x_1) \in \mathbb{R}^2$. We consider the set

$$\mathcal{M}_2(\mu_0, \mu_1) \quad := \quad \big\{ \mathbb{P} \in \mathcal{P}_{\mathbb{R}^2} : X_0 \sim^{\mathbb{P}} \mu_0, \ X_1 \sim^{\mathbb{P}} \mu_1 \text{ and } \mathbb{E}^{\mathbb{P}}[X_1 | X_0] = X_0 \big\}$$

of all martingale measures with marginals μ_0 and μ_1.

Let $c : \mathbb{R}^2 \to \mathbb{R}$ be a reward function such that $c(x, y) \le a(x) + b(y)$ for some $a \in \mathbb{L}^1(\mu_0)$ and $b \in \mathbb{L}^1(\mu_1)$, a two-marginals MT problem is defined by

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

The dual formulation of the MT problem (2.1) turns out to be

$$\mathbf{D}_2(\mu_0, \mu_1) \quad := \quad \inf_{(\varphi, \psi, h) \in \mathcal{D}_2} \big\{ \mu_0(\varphi) + \mu_1(\psi) \big\}, \tag{2.2}$$

where the dual variables set is:

$$\mathcal{D}_2 \quad := \quad \big\{ (\varphi, \psi, h) \ : \varphi^+ \in \mathbb{L}^1(\mu_0), \ \psi^+ \in \mathbb{L}^1(\mu_1), \ h \in \mathbb{L}^0 \text{ and } \varphi \oplus \psi + h^\otimes \ge c \big\}, \tag{2.3}$$

with notations $\varphi^+ := \varphi \vee 0$, $\psi^+ := \psi \vee 0$, $(\varphi \oplus \psi)(x, y) := \varphi(x) + \psi(y)$ and $h^\otimes(x, y) := h(x)(y - x)$.

Under mild conditions, the existence of the optimal martingale as well as the strong duality (i.e. $\mathbf{P}_2(\mu_0, \mu_1) = \mathbf{D}_2(\mu_0, \mu_1)$) are proved in Beiglböck, Henry-Labordère & Penkner [5]. More recently, under additional conditions on the reward function c, the optimal martingale has been characterized as left-monotone transference map in [7], see more extensions in [40, 6]. Let us recall it in a simpler context.

Theorem 2.1. *Suppose that the partial derivative c_{xyy} exists and $c_{xyy} > 0$ on $(l_{\mu_0}, r_{\mu_0}) \times (l_{\mu_1}, r_{\mu_1})$. Then the optimal martingale measure $\mathbb{P}_*(dx, dy)$ of problem (2.1) is unique, and can be characterized by $\mathbb{P}_*(dx, dy) = \mu_0(dx)T_*(x, dy)$, where the probability kernel T_* is of the form*

$$T_*(x, dy) \quad := \quad q(x)\delta_{T_u(x)}(dy) + (1 - q(x))\delta_{T_d(x)}(dy),$$

for some functions $T_u, T_d : \mathbb{R} \to \mathbb{R}$ satisfying $T_d(x) \le x \le T_u(x)$ and $q(x) \in [0, 1]$.

More precise properties of T_d and T_u have been obtained in [7, 40], which will be recalled and used in Section 6.

The above problem and results can be easily extended to the finitely-many marginals case, when the cost function is given by $\sum_{i=1}^{n} c(x_{i-1}, x_i)$. More precisely, with $n+1$ given probability measures $(\mu_0, \cdots, \mu_n) \in (\mathcal{P}_{\mathbb{R}})^{n+1}$ such that $\mu_0 \preceq \cdots \preceq \mu_n$ in the convex ordering, the problem consists in maximizing

$$\mathbb{E}^{\mathbb{P}}\big[c(X_0, \cdots, X_n)\big] \quad = \quad \mathbb{E}^{\mathbb{P}}\Big[\sum_{i=1}^{n} c^i(X_{i-1}, X_i) \Big] \tag{2.4}$$

among all martingales probability measures on $\mathbb{R}^{n+1}$ satisfying the marginal distribution constraints ($X_i \sim^{\mathbb{P}} \mu_i, i = 0, \cdots, n$). Under the same technical condition on c, the optimal martingale measure is obtained by a Markovian iteration of the above left-monotone transference map, thus inducing a binomial tree.

---

[1] The last integrals are well-defined and valued in $\mathbb{R} \cup \{\infty\}$, as convex functions are bounded from below by an affine function, which is integrable w.r.t. μ by the finite first moment condition.

2.2 Continuous-time martingale transport under full marginals constraints

We now introduce a continuous-time martingale transportation (MT) problem under full marginals constraints, as the limit of the multi-marginals MT recalled in Section 2.1 above. Namely, given a family of probability measures $\mu = (\mu_t)_{t \in [0,1]}$, we consider all continuous-time martingales satisfying the marginal constraints, and optimize w.r.t. a class of reward functions. To avoid the problem of integration, we define, for every random variable ξ, the expectation $\mathbb{E}[\xi] := \mathbb{E}[\xi^+] - \mathbb{E}[\xi^-]$ with the convention $\infty - \infty = -\infty$.

Let $\Omega := \mathbb{D}([0,1], \mathbb{R})$ denote the canonical space of all càdlàg paths on $[0,1]$, X the canonical process and $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le 1}$ the canonical filtration generated by X, i.e. $\mathcal{F}_t := \sigma\{X_s, 0 \le s \le t\}$. We denote by $\mathcal{M}_\infty$ the collection of all martingale measures on Ω, i.e. the collection of all probability measures on Ω under which the canonical process X is a martingale. The set $\mathcal{M}_\infty$ is equipped with the weak convergence topology throughout the paper. By Karandikar [58], there is a non-decreasing process $([X]_t)_{t \in [0,1]}$ defined on Ω which coincides with the $\mathbb{P}$-quadratic variation of X, $\mathbb{P}$-a.s. for every martingale measure $\mathbb{P} \in \mathcal{M}_\infty$. Denote also by $[X]^c$ the continuous part of the non-decreasing process $[X]$.

Given a family of probability measures $\mu = (\mu_t)_{0 \le t \le 1}$, denote by $\mathcal{M}_\infty(\mu) \subset \mathcal{M}_\infty$ the collection of all martingale measures on Ω such that $X_t \sim^{\mathbb{P}} \mu_t$ for all $t \in [0,1]$. In particular, following Kellerer [59] (see also Hirsch and Roynette [41]), $\mathcal{M}_\infty(\mu)$ is nonempty if and only if the family $(\mu_t)_{0 \le t \le 1}$ admits a finite first order moment, is non-decreasing in convex ordering, and $t \mapsto \mu_t$ is right-continuous.

Finally, for all $t \in [0,1]$, we denote by $-\infty \le l_t \le r_t \le \infty$ the left and right extreme boundaries of the support [2] of μ_t. Similar to Hobson & Klimmek [47], our continuous-time MT problem is obtained as a continuous-time limit of the multi-marginals MT problem, by considering the limit of the reward function $\sum_{i=1}^n c(\mathbf{x}_{t_{i-1}}, \mathbf{x}_{t_i})$ as in (2.4), where $(t_i)_{1 \le i \le n}$ is a partition of the interval $[0,1]$ with mesh size vanishing to zero. For this purpose, we formulate the following assumption on the reward function.

Assumption 2.2. *The reward function $c : \mathbb{R}^2 \to \mathbb{R}$ is in $C^3((l_1, r_1) \times (l_1, r_1))$ and satisfies*

$$c(x,x) = c_y(x,x) = 0 \quad and \quad c_{xyy}(x,y) > 0, \quad \forall (x,y) \in (l_1, r_1) \times (l_1, r_1).$$

In order to obtain the convergence, we need to use the pathwise Itô calculus introduced in Föllmer [30], which is also used in Hobson & Klimmek [47] and Davis, Obłój & Raval [23] (see in particular their Appendix B).

Definition 2.3 (Föllmer [30]). *Let $\pi_n = (0 = t_0^n < \cdots < t_n^n = 1), n \ge 1$ be partitions of $[0,1]$ with $|\pi_n| := \max_{1 \le k \le n} |t_k^n - t_{k-1}^n| \to 0$ as $n \to \infty$. A càdlàg path $\mathbf{x} : [0,1] \to \mathbb{R}$ has a finite quadratic variation along $(\pi_n)_{n \ge 1}$ if the sequence of measures on $[0,1]$,*

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

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

The next result follows the same line of proof as in Lemma 7.4 of [47].

---

[2] (i) The functions l and r are non-increasing and non-decreasing, respectively. We only report the justification for the right boundary of the support r, a similar argument applies to l. For $0 \le s \le t \le 1$, it follows from the increase in convex order that for all constant $c \in \mathbb{R}$, we have $\int_c^\infty (x-c)^+ \mu_t(dx) \ge \int_c^\infty (x-c)^+ \mu_s(dx)$, so that $\mu_s((c,\infty)) > 0$ implies that $\mu_t((c,\infty)) > 0$, and therefore r is non-decreasing.

(ii) Assume that μ_t has a density function for all $t \in [0,1]$ and $t \mapsto \mu_t$ is continuous w.r.t the weak convergence topology, then the functions $\mathbf{1}_{(-\infty,c)}$ and $\mathbf{1}_{(c,\infty)}(x)$ are μ_t−a.s. continuous for all $t \in [0,1]$, and it follows that l and r are continuous.

Lemma 2.4. *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*

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

We next introduce the non-decreasing process $[\mathbf{x}]$ as in Karandikar [58], which is defined for every $\mathbf{x} \in \Omega$, and coincides $\mathbb{P}$−a.s. with the quadratic variation, and also with $[x]^F$, for all $\mathbb{P} \in \mathcal{M}_\infty$.

Motivated by the last convergence result, we introduce a reward function

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

where the integral and the sum are defined as the difference of the positive and negative parts, under the convention $\infty - \infty = -\infty$. We then formulate a continuous-time MT problem under full marginals constraints by

$$\mathbf{P}_\infty(\mu) \quad := \quad \sup_{\mathbb{P} \in \mathcal{M}_\infty(\mu)} \mathbb{E}^{\mathbb{P}}\big[C(X)\big]. \tag{2.5}$$

Remark 2.5. Under the condition $c(x,x) = c_y(x,x) = 0$ in Assumption 2.2, we have $|c(x, x + h)| \leq K_1(x)h^2$ for all $h \in [-1, 1]$ with some locally bounded positive function K_1. Since we have, $\mathbb{P}$−a.s. for all $\mathbb{P} \in \mathcal{M}_\infty$, that $\sum_{s \leq t}(\Delta X_s)^2 < \infty$, then we also have $\sum_{0 \leq t \leq 1} c(X_{t-}, X_t) < \infty$.

Now, let us introduce the dual formulation of the above MT problem (2.5). We first introduce the class of admissible dynamic and static strategies. Denote by $\mathbb{H}_0$ the class of all $\mathbb{F}$−predictable and locally bounded processes $H : [0, 1] \times \Omega \to \mathbb{R}$, i.e. there is an increasing family of $\mathbb{F}$−stopping times $(\tau_n)_{n \geq 1}$ taking value in $[0, 1] \cup \{\infty\}$ such that the process $H_{. \wedge \tau_n}$ is bounded for all $n \geq 1$ and $\tau_n \to \infty$ as $n \to \infty$. Then for every $H \in \mathbb{H}_0$ and under every martingale measure $\mathbb{P} \in \mathcal{M}_\infty$, one can define the integral, denoted by $H \cdot X$, of H w.r.t. the martingale X (see e.g. Jacod & Shiryaev [54] Chapter I.4). Define

$$\mathcal{H} \quad := \quad \big\{ H \in \mathbb{H}_0 \; : H \cdot X \text{ is a } \mathbb{P}\text{-supermartingale for every } \mathbb{P} \in \mathcal{M}_\infty \big\}.$$

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

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

We also introduce a family of random measures $\delta^X = (\delta_t^X)_{0 \leq t \leq 1}$ on $\mathbb{R}$, induced by the canonical process X, by $\delta_t^X(dx) := \delta_{X_t}(dx)$. In particular, we have

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

Then the collection of all superhedging strategies is given by

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

and our dual problem is defined by

$$\mathbf{D}_\infty(\mu) \quad := \quad \inf_{(H, \lambda) \in \mathcal{D}_\infty(\mu)} \mu(\lambda). \tag{2.6}$$

3 Main results

We first construct in Subsection 3.1 a martingale transport peacock corresponding to the full-marginals $(\mu_t)_{t\in[0,1]}$ under technical conditions. This is obtained in Proposition 3.2 as an accumulation point of a sequence $(\mathbb{P}^n)_n$ of solutions of $n-$periods discrete martingale transport problems. In order to further characterize such a martingale peacock, we next restrict our analysis to the one-maximizer context of Assumption 3.4 below. Then, our second main result in Subsection 3.2 shows that the sequence $(\mathbb{P}^n)_n$ converges to the distribution of a local Lévy process. Finally our third main result is reported in Subsection 3.3. We show that this limit indeed solves the continuous-time MT problem (2.5), and that the duality result holds with explicit characterization of the solution of the dual problem under Assumption 3.4 below.

3.1 A martingale transport plan under full marginals constraint

For every $t \in [0,1]$, we denote by $F(t, \cdot)$ the cumulative distribution function of the probability measure μ_t on $\mathbb{R}$, and $F^{-1}(t, \cdot)$ the corresponding right-continuous inverse with respect to the $x-$variable. We also denote for $t \in [0,1)$, $\varepsilon \in (0, 1-t]$:

$$\delta^\varepsilon F(t,x) := F(t+\varepsilon, x) - F(t,x), \quad g_t^\varepsilon(x,y) := F^{-1}\big(t+\varepsilon, F(t,x) + \delta^\varepsilon F(t,y)\big), \qquad (3.1)$$

and

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

Assumption 3.1. (i) *The marginal distributions $\mu = (\mu_t)_{t\in[0,1]}$ are non-decreasing in convex ordering and have finite first order moment.*
(ii) *For all $t \in [0,1]$, the measure μ_t has a density function $f(t, \cdot)$, and $t \mapsto \mu_t$ is continuous w.r.t. the weak convergence topology. Moreover,*

$$F \in C_b^4(E) \quad and \quad \inf_{x\in[-K,K]\cap(l_t,r_t)} f(t,x) > 0 \;\; for\;all\;\; K > 0.$$

(iii) *For every $t \in [0,1)$ and $\varepsilon \in (0, 1-t]$, the set of local maximizers $\mathbf{M}\big(\delta^\varepsilon F(t,x)\big)$ of function $x \mapsto \delta^\varepsilon F(t,x)$ is finite; and moreover, the following irreducibility condition [3] holds true,*

$$\int_x^\infty (\xi - x)\delta^\varepsilon F(d\xi) = -\int_x^\infty \delta^\varepsilon F(t,y)dy > 0, \quad for\;all\;x \in (l_t, r_t). \qquad (3.2)$$

Under the convex ordering condition, by considering the two marginals $(\mu_t, \mu_{t+\varepsilon})$, the corresponding unique left-monotone martingale measure can be characterized by two functions $T_u^\varepsilon(t,x)$ and $T_d^\varepsilon(t,x)$ as in Theorem 2.1.

We recall that $\Omega := \mathbb{D}([0,1], \mathbb{R})$ is the canonical space of càdlàg paths, which is a Polish space (separable, complete metric space) equipped with the Skorokhod topology; and X is the canonical process. Let $(\pi_n)_{n\geq 1}$ be a sequence, where every $\pi_n = (t_k^n)_{0\leq k\leq n}$ is a partition of the interval $[0,1]$, i.e. $0 = t_0^n < \cdots < t_n^n = 1$. Suppose in addition that $|\pi_n| := \max_{1\leq k\leq n}(t_k^n - t_{k-1}^n) \to 0$. Then for every partition π_n, by considering the marginal distributions $(\mu_{t_k^n})_{0\leq k\leq n}$, one obtains an $(n+1)$-marginals MT problem, which consists in maximizing

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

among all discrete-time martingales $\tilde{X}^n = (\tilde{X}_k^n)_{0\leq k\leq n}$ satisfying the marginal distribution constraints. Under Assumptions 2.2 and 3.1, the left-monotone transference plan denoted by $\mathbb{P}^{*,n}$ is a solution of the last martingale transport problem. Let $\Omega^{*,n} := \mathbb{R}^{n+1}$ be the canonical space

---

[3]We refer to Section 8.1 of [7] for the decomposition of any pair (μ_0, μ_1) in irreducible components.

of discrete-time process, $X^n = (X^n_k)_{0 \le k \le n}$ be the canonical process. Then under the optimal martingale measure $\mathbb{P}^{*,n}$, X^n is a discrete-time martingale and at the same time a Markov chain, characterized by $T_u^{\Delta t^n_{k+1}}(t^n_k, \cdot)$ and $T_d^{\Delta t^n_{k+1}}(t^n_k, \cdot)$ with $\Delta t^n_{k+1} := t^n_{k+1} - t^n_k$, induced by the two marginals $(\mu_{t^n_k}, \mu_{t^n_{k+1}})$ as in Theorem 2.1. We then extend the Markov chain X^n to a continuous-time càdlàg process $X^{*,n} = (X^{*,n}_t)_{0 \le t \le 1}$ defined by

$$X^{*,n}_t := \sum_{k=1}^{n} X^n_{k-1} \mathbf{1}_{[t^n_{k-1}, t^n_k)}(t), \quad t \in [0,1],$$

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

Proposition 3.2. *Under Assumption 3.1, the sequence $(\mathbb{P}^n)_{n \ge 1}$ is tight w.r.t. the Skorokhod topology on Ω. Moreover, every limit point $\mathbb{P}^0$ satisfies $\mathbb{P}^0 \in \mathcal{M}_\infty(\mu)$.*

Remark 3.3. *In a recent work, Juillet [56] analyzes the convergence of the left-monotone martingale transport plan under more general conditions and/or in some other specific context.*

3.2 A Local Lévy process characterization

We next seek for a further characterization of the limiting peacocks obtained in Proposition 3.2. The remaining part of our results is established under the following unique local maximum condition. Let $t \in [0,1)$ and $\varepsilon \in (0, 1-t]$, we denote by $\mathbf{M}(\delta^\varepsilon F(t, \cdot))$ (resp. $\mathbf{M}(\partial_t F(t, \cdot))$) the set of all local maximizer of function $\delta^\varepsilon F(t, \cdot)$ (resp. $\partial_t F(t, \cdot)$).

Assumption 3.4. (i) *There is a constant $\varepsilon_0 > 0$ such that, for all $t \in [0,1]$ and $0 < \varepsilon \le \varepsilon_0 \wedge (1-t)$, we have $\mathbf{M}(\delta^\varepsilon F(t, \cdot)) = \{m^\varepsilon(t)\}$ and $\mathbf{M}(\partial_t F(t, \cdot)) = \{m_t\}$.*
(ii) *Denote $m^0(t) := m_t$, then $(t, \varepsilon) \mapsto m^\varepsilon(t)$ is continuous (hence uniformly continuous with continuity modulus ρ_0) on $\{(t, \varepsilon) : 0 \le \varepsilon \le \varepsilon_0, \ 0 \le t \le 1 - \varepsilon\}$.*
(iii) *For every $t \in [0,1]$, we have $\partial_{tx} f(t, m_t) < 0$.*

Example 3.13 below provides some concrete peacocks satisfying the last conditions.

Remark 3.5. *Under the one local maximizer condition in Assumption 3.4 (i), the irreducibility condition (3.2) holds true. Indeed, assume that (3.2) is not true, the for some $t \in [0,t)$, $\varepsilon \in (0, 1-t]$, and some $K \in (l_t, r_t)$, one has $\int_K^\infty (\xi - x)\delta^\varepsilon F(d\xi) = 0 = \int_{-\infty}^K (x - \xi)\delta^\varepsilon F(d\xi)$. It follows by Theorem 8.4 of Beiglböck and Juillet [7], we have the decomposition $\mu_t = \rho \mu_t^1 + (1-\rho)\mu_t^2$ and $\mu_{t+\varepsilon} = \rho \mu_{t+\varepsilon}^1 + (1-\rho)\mu_{t+\varepsilon}^2$ for some $\rho \in (0,1)$, where μ_t^1 and $\mu_{t+\varepsilon}^1$ (resp. μ_t^2 and $\mu_{t+\varepsilon}^2$) are probability measures supported in $(-\infty, K)$ (resp. $[K, \infty)$), and $\mu_t^i \preceq \mu_{t+\varepsilon}^i$ in convex ordering for $i = 1, 2$. Notice that the corresponding differences of c.d.f. $\delta^\varepsilon F^i$ satisfies*

$$\delta^\varepsilon F = \rho \mathbf{1}_{(-\infty, K)} \delta^\varepsilon F^1 + (1-\rho)\mathbf{1}_{(K, \infty)} \delta^\varepsilon F^2.$$

Since $\mu_t^i \preceq \mu_{t+\varepsilon}^i$ in convex ordering for $i = 1, 2$, $\delta^\varepsilon F^1(x)$ (resp. $\delta^\varepsilon F^2(x)$) has at least one local maximizer in $(-\infty, K) \cap (l_t, r_t)$ (resp. $(K, \infty) \cap (l_t, r_t)$). It follows that $\delta^\varepsilon F$ has at least two local maximizer on (l_t, r_t), which is a contradiction to Assumption 3.4 (i).

Recall that the left-monotone transference plan can be characterized by $(T_u^\varepsilon, T_d^\varepsilon)$, whose explicit construction are provided by [40]. Direct calculation suggests that the sequence $(T_d^\varepsilon)_\varepsilon$ is expected to converge to the function T_d obtained via the limiting equation

$$\int_{T_d(t,x)}^{x} (x - \xi)\partial_t f(t, \xi)d\xi = 0. \tag{3.3}$$

We also define $j_d(t, x)$ and $j_u(t, x)$ by

$$j_d(t, x) := x - T_d(t, x) \text{ and } j_u(t, x) := \frac{\partial_t F(t, T_d(t,x)) - \partial_t F(t, x)}{f(t, x)}, \quad t \in [0,1], x > m_t. \tag{3.4}$$

We notice that $j_u(t, \cdot)$ and $j_d(t, \cdot)$ are both positive and continuous on $D^c(t) := (m_t, r_t)$. Let us also introduce $D(t) := (l_t, m_t]$ and

$$D^c := \{(t, x) : t \in [0, 1], x \in (m_t, r_t)\} \quad \text{and} \quad D := \{(t, x) : t \in [0, 1], x \in D(t)\}. \quad (3.5)$$

Lemma 3.6. *Let Assumptions 3.1 and 3.4 hold true. Then for all $x \in (m_t, r_t)$, the equation (3.3) has a unique solution $T_d(t, x) \in (l_t, m_t)$. Moreover*
(i) T_d *is strictly decreasing in x*
(ii) $j_d(t, x)\mathbf{1}_{x > m_t}$, $j_u(t, x)\mathbf{1}_{x > m_t}$ and $\frac{j_u}{j_d}(t, x)\mathbf{1}_{x > m_t}$ *are all locally Lipschitz in (t, x).*

Our second main result is the following convergence result for the sequence $(\mathbb{P}^n)_n$ with an explicit characterization of the limit as the law of a local Lévy process.

Theorem 3.7. *Suppose that Assumptions 3.1 and 3.4 hold true, then $\mathbb{P}^n \to \mathbb{P}^0$, where $\mathbb{P}^0$ is the unique weak solution of the SDE*

$$X_t = X_0 - \int_0^t \mathbf{1}_{\{X_{s-} > m(s)\}} j_d(s, X_{s-})(dN_s - \nu_s ds), \quad \nu_s := \frac{j_u}{j_d}(s, X_{s-})\mathbf{1}_{X_{s-} > m(s)}, \quad (3.6)$$

and $(N_s)_{0 \le s \le 1}$ is a unit size jump process with predictable compensated process $(\nu_s)_{0 \le s \le 1}$.

The pure jump process (3.6) is in the spirit of the local Lévy models introduced in Carr, Geman, Madan & Yor [16]. Notice however that the intensity process $(\nu_t)_{0 \le t \le 1}$ in our context is state-dependent.

We conclude this subsection by providing a point of view from the perspective of the forward Kolmogorov-Fokker-Planck (KFP) equation.

Lemma 3.8. *Under Assumptions 3.1 and 3.4, the density function $f(t, x)$ satisfies*

$$\partial_t f(t, x) = -\mathbf{1}_{\{x < m_t\}} \frac{j_u f}{j_d(1 - \partial_x j_d)}\big(t, T_d^{-1}(t, x)\big) - \mathbf{1}_{\{x > m_t\}} \Big(\frac{j_u f}{j_d} - \partial_x(j_u f)\Big)(t, x), \quad (3.7)$$

for all $t \in [0, 1)$ and $x \in (l_t, r_t) \setminus \{m_t\}$.

The first order PDE (3.7) can be viewed as a KFP forward equation of SDE (3.6).

Proposition 3.9. *Let Assumptions 3.1 and 3.4 hold true. Suppose that the SDE (3.6) has a weak solution $\widehat{X}$ which is a martingale whose marginal distribution admits a density function $f^{\widehat{X}}(t, x) \in C^1([0, 1] \times \mathbb{R})$. Suppose in addition that $\mathbb{E}\big[\big|\widehat{X}_1\big|^p\big] < \infty$ for some $p > 1$, and for every $t \in [0, 1)$, there is some $\varepsilon_0 \in (0, 1 - t]$ such that*

$$\mathbb{E}\Big[\int_t^{t + \varepsilon_0} j_u(s, \widehat{X}_s) \mathbf{1}_{\widehat{X}_s \in D^c(s)} ds\Big] < \infty. \quad (3.8)$$

Then, the density function $f^{\widehat{X}}$ of $\widehat{X}$ defined in (3.6) satisfies the KFP equation (3.7)

3.3 Optimality of the local Lévy process

The optimality results of this subsection are also obtained in the one-maximizer context of Assumption 3.4. We first introduce the candidates of the optimal dual components for the dual problem (2.6). Following Section 5 of [40], (see also the recalling in Section 6.1.1 below), the optimal superhedging strategy $(\varphi^\varepsilon, \psi^\varepsilon, h^\varepsilon)$ for the two marginals MT problem associated with initial distribution μ_t, terminal distribution $\mu_{t+\varepsilon}$, and reward function $c : (l_1, r_1) \times (l_1, r_1) \to \mathbb{R}$, is explicitly given by:

$$\partial_x h^\varepsilon(t, x) := \frac{c_x(x, T_u^\varepsilon(t, x)) - c_x(x, T_d^\varepsilon(t, x))}{T_u^\varepsilon(t, x) - T_d^\varepsilon(t, x)}, \qquad x \ge m^\varepsilon(t),$$

$$h^\varepsilon(t, x) := h^\varepsilon\big(t, (T_d^\varepsilon)^{-1}(t, x)\big) - c_y\big((T_d^\varepsilon)^{-1}(t, x), x\big), \qquad x < m^\varepsilon(t);$$

9

denoting $(T^\varepsilon)^{-1}(t,x) := (T_u^\varepsilon)^{-1}(t,x)\mathbf{1}_{x \geq m^\varepsilon(t)} + (T_d^\varepsilon)^{-1}(t,x)\mathbf{1}_{x < m^\varepsilon(t)}$, the function ψ^ε is given by

$$\partial_x \psi^\varepsilon(t,x) = c_y\big((T^\varepsilon)^{-1}(t,x),x\big) - h^\varepsilon\big(t,(T^\varepsilon)^{-1}(t,x)\big),$$

and

$$\varphi^\varepsilon(t,x) := \frac{x - T_d^\varepsilon(t,x)}{T_u^\varepsilon(t,x) - T_d^\varepsilon(t,x)} \Big(c\big(x,T_u^\varepsilon(t,x)\big) - \psi^\varepsilon\big(t,T_u^\varepsilon(t,x)\big)\Big)$$
$$+ \frac{T_u^\varepsilon(t,x) - x}{T_u^\varepsilon(t,x) - T_d^\varepsilon(t,x)} \Big(c\big(x,T_d^\varepsilon(t,x)\big) - \psi^\varepsilon\big(t,T_d^\varepsilon(t,x)\big)\Big).$$

Clearly, h^ε and ψ^ε are unique up to a constant. More importantly, h^ε and ψ^ε can be chosen continuous on D so that

$$c(\cdot,T_u^\varepsilon(t,\cdot)) - \psi^\varepsilon(T_u^\varepsilon(\cdot)) - c(\cdot,T_d^\varepsilon(\cdot)) + \psi^\varepsilon(T_d^\varepsilon(\cdot)) - \big(T_u^\varepsilon(\cdot) - T_d^\varepsilon(\cdot)\big)h^\varepsilon(\cdot)$$

is continuous, since T_u^ε and T_d^ε are both continuous under Assumption 3.4.

We shall see later that Assumption 2.2 on the reward function c implies that the continuous-time limit of the optimal dual components is given as follows. The function $h^* : D \to \mathbb{R}$ is defined, up to a constant, by

$$\partial_x h^*(t,x) := \frac{c_x(x,x) - c_x(x,T_d(t,x))}{j_d(t,x)}, \qquad \text{when } x \geq m_t, \tag{3.9}$$

$$h^*(t,x) := h^*\big(t,T_d^{-1}(t,x)\big) - c_y\big(T_d^{-1}(t,x),x\big), \qquad \text{when } x < m_t, \tag{3.10}$$

and $\psi^* : D \to \mathbb{R}$ is defined, up to a constant, by

$$\partial_x \psi^*(t,x) := -h^*(t,x), \quad \text{for all } (t,x) \in D. \tag{3.11}$$

As an immediate consequence of Lemma 3.6 (ii), we obtain the following regularity results

Corollary 3.10. *Under Assumptions 2.2, 3.1, and 3.4, we have $\psi^* \in C^{1,1}(D)$.*

Proof. Clearly, h^* is continuous in (t,x) for $x \neq m_t$, since the function $T_d(t,x)\mathbf{1}_{x \geq m_t}$ is continuous. We also easily check the continuity of h^* at the point (t,m_t) by (3.9), since $T_d^{-1}(t,x) \to m_t$ as $x \to m_t$ and $c_y(x,x) = 0$. By its definition in (3.11), $\partial_x \psi^*$ is also continuous. We next compute, by (3.9), that

$$\partial_{t,x} h^*(t,x) = \partial_t j_d(t,x) \frac{c_{xy}\big(x,T_d(t,x)\big)j_d(t,x) - \big(c_x(x,x) - c_x(x,T_d(t,x))\big)}{\big(j_d(t,x)\big)^2},$$

which is also locally bounded on D^c since $T_d(t,x)$ is locally bounded by Lemma 3.6. Then the function $\partial_t \psi^*(t,x) = C - \int_0^x \partial_t h^*(t,\xi)d\xi$, for some fixed constant C, is also continuous in (t,x). $\square$

In order to introduce a dual static strategy in Λ, we let $\gamma^*(dt) := \delta_{\{0\}}(dt) + \delta_{\{1\}}(dt) + Leb(dt)$ be a finite measure on $[0,1]$, where $Leb(dt)$ denotes the Lebesgue measure on $[0,1]$; we define λ_0^* and $\overline{\lambda}_0^*$ by $\lambda_0^*(0,x) := \psi^*(0,x)$, $\lambda_0^*(1,x) := \psi^*(1,x)$, $\overline{\lambda}_0^*(0,x) := |\psi^*(0,x)|$, $\overline{\lambda}_0^*(1,x) := \sup_{t \in [0,1]} |\psi^*(t,x)|$; and for all $(t,x) \in (0,1) \times \mathbb{R}$,

$$\lambda_0^* := \partial_t \psi^* + \mathbf{1}_{D^c}\big(\partial_x \psi^* j_u + \nu[\psi^* - \psi^*(.,T_d) + c(.,T_d)]\big),$$
$$\overline{\lambda}_0^* := \big|\partial_t \psi^* + \mathbf{1}_{D^c}\big(\partial_x \psi^* j_u + \nu[\psi^* - \psi^*(.,T_d)]\big)\big| + \mathbf{1}_{D^c}\nu[|c(.,T_d)|],$$

where we recall that $D^c = \{(t,x) : x > m_t\}$. Finally, we denote $\lambda^*(x,dt) := \lambda_0^*(t,x)\gamma^*(dt)$ and $\overline{\lambda}^*(x,dt) := \overline{\lambda}_0^*(t,x)\gamma^*(dt)$.

We are now ready for our third main result which states the optimality of the local Lévy process (3.6), as well as that of the dual component introduced above. Similar to [40] and [47], we obtain in addition a strong duality for the MT problem (2.5) and (2.6). Let H^* be the $\mathbb{F}$-predictable process on Ω defined by

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

Theorem 3.11. *Let Assumptions 2.2, 3.1 and 3.4 hold true, suppose in addition that $\mu(\overline{\lambda}^*) = \int_0^1 \int_{\mathbb{R}} \overline{\lambda}_0^*(t, x) \mu_t(dx) \gamma^*(dt) < \infty$. Then the martingale transport problem (2.5) is solved by the local Lévy process (3.6). Moreover, $(H^*, \lambda^*) \in \mathcal{D}_\infty(\mu)$ and we have the duality*

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

where the optimal value is given by

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

The proofs of Theorems 3.7 and 3.11 are reported later in Section 6, the main idea is to use the approximation technique, where we need in particular the continuity property of the characteristic functions in Lemma 3.10.

Remark 3.12. By symmetry, we may consider the right monotone martingale transference plan as in Remark 3.14 of [40]. This leads to an upward pure jump process with explicit characteristics, assuming that $x \mapsto \partial_t F(t, x)$ has only one local minimizer $\tilde{m}_t$. Indeed, let

$$\tilde{j}_u(t, x) := \tilde{T}_u(t, x) - x \quad \text{and} \quad \tilde{j}_d(t, x) := \frac{\partial_t F(t, x) - \partial_t F(t, \tilde{T}_u(t, x))}{f(t, x)},$$

where $\tilde{T}_u(t, x) : (l_t, \tilde{m}_t] \to [\tilde{m}_t, r_t)$ is defined as the unique solution to

$$\int_x^{\tilde{T}_u(t,x)} (\xi - x) \partial_t f(t, \xi) d\xi \;=\; 0.$$

The limit process solves SDE:

$$dX_t = \mathbf{1}_{X_{t-} < \tilde{m}_t} \tilde{j}_u(t, X_{t-})(d\tilde{N}_t - \tilde{\nu}_t dt) \qquad \tilde{\nu}_t := \frac{\tilde{j}_d}{\tilde{j}_u}(t, X_{t-}) \mathbf{1}_{X_{t-} < \tilde{m}_t}, \tag{3.12}$$

where $(\tilde{N}_t)_{0 \le t \le 1}$ is an upward jump process with unit jump size and predictable compensated process $(\tilde{\nu}_t)_{0 \le t \le 1}$. Moreover, under Assumption 2.2 together with further technical conditions on μ, this martingale solves a corresponding minimization MT problem with optimal value

$$\int_0^1 \int_{l_t}^{\tilde{m}_t} \frac{\tilde{j}_d(t, x)}{\tilde{j}_u(t, x)} \, c\big(x, x + \tilde{j}_u(t, x)\big) \, f(t, x) \, dx \, dt.$$

The corresponding dual optimizer is also explicitly given by Theorem 3.11 once $(\tilde{j}_d, \tilde{j}_u)$ is substituted to (j_d, j_u).

We conclude this section by some examples of peacocks satisfying Assumption 3.4.

Example 3.13. (i) Let $F_0 : \mathbb{R} \to [0, 1]$ be a distribution function of random variable such that its density function $f_0(x) := F_0'(x)$ is strictly positive on $\mathbb{R}$. Define $F(t, x) := F_0(x/t)$ and $f(t, x) := \frac{1}{t} f_0(x/t)$, then the associated marginal distribution is clearly a peacock. Denote further $\hat{f}_0(x) := \log f_0(x)$. Assume that $\hat{f}_0'(x) > 0$ for $x \in (-\infty, 0)$ and $\hat{f}_0'(x) < 0$ for $x \in (0, \infty)$,

then for every $t \in (0,1)$ and $\varepsilon \in (0, 1-t)$, the map $y \mapsto \hat{f}_0(y) - \hat{f}_0(y + \varepsilon y)$ is strictly increasing on $(-\infty, 0)$ and strictly decreasing on $(0, \infty)$. In this case, the set

$$\{x \ : f(t,x) = f(t+\varepsilon, x)\} \quad = \quad \Big\{x \ : \log(1 + \varepsilon/t) = \hat{f}_0\Big(\frac{x}{t+\varepsilon}\Big) - \hat{f}_0\Big(\frac{x}{t+\varepsilon}(1 + \varepsilon/t)\Big)\Big\}$$

has exactly two points, and one can check that the smaller one is the maximizer $m^\varepsilon(t)$ of $x \mapsto \delta^\varepsilon F(t,x)$, and the second one is the minimizer of $x \mapsto \delta^\varepsilon F(t,x)$. Denote by m_t the maximizer of $x \mapsto \partial_t F(t,x)$, which is also is the smaller solution of

$$-\hat{f}_0'\Big(\frac{x}{t}\Big)\frac{x}{t} \ = \ t \quad \Longleftrightarrow \quad \partial_t f(t,x) = 0, \quad \text{for a fixed } t > 0. \tag{3.13}$$

Assume in addition that that $\hat{f}_0'(m_t) \neq 0$, then by the fact that

$$\hat{f}_0\Big(\frac{x}{t+\varepsilon}\Big) - \hat{f}_0\Big(\frac{x}{t+\varepsilon}(1 + \varepsilon/t)\Big) \ = \ -\hat{f}_0'\Big(\frac{x}{t+\varepsilon}\Big)\frac{x}{t+\varepsilon}\frac{\varepsilon}{t} \ + \ o(\varepsilon/t),$$

we can also prove the convergence of $m^\varepsilon(t) \to m_t$. In this case, Assumption 3.4 holds true.

(ii) In particular, when the marginals $(\mu_t)_{t \in [\delta, 1+\delta]}$ are those of the Brownian motion for some $\delta > 0$, then both Assumptions 3.1 and 3.4 hold true with $f(t,x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}$, $m^\varepsilon(t) = -\sqrt{\frac{t(t+\varepsilon)}{\varepsilon}}\log(1 + \varepsilon/t)$ and $m_t = -\sqrt{t}$. See also Section 4 for more discussions.

(iii) For general function $\hat{f}_0$, it is clear that (3.13) may have more than two solutions, then Assumption 3.4 is no more true.

4 Examples of extremal peacock processes

4.1 A remarkable fake Brownian motion

In this subsection, we specialize the discussion to the special case $\mu_t :=:= \mathcal{N}(0, t)$, $t \in [0, 1]$. This is the first appearance of peacocks in the literature under the name of fake Brownian motion, see Albin [2], Fan, Hamza & Klebaner [29], Hamza & Klebaner [37], Hirsch et al. [42], Hobson [46], Oleszkiewicz [64], Pagès [66].

Our construction in Theorem 3.7 provides a new example of fake Brownian motion which is remarkable by the corresponding optimality property.

It is easily checked that our Assumptions 3.1 and 3.4 are satisfied. By direct computation, we have $m^\varepsilon(t) = -\sqrt{\frac{t(t+\varepsilon)}{\varepsilon}}\log(1 + \varepsilon/t)$, $m_t = -\sqrt{t}$ for all $t \in [\delta, 1]$, and it follows that $T_d(t,x)$ is defined by the equation:

$$\int_{T_d(t,x)}^x (x - \xi)(\xi^2 - t)e^{-\xi^2/2t}d\xi = 0 \quad \text{for all} \quad x \geq m_t.$$

Direct change of variables provides the scaled solution:

$$t^{1/2}\widehat{T}_d\big(t^{-1/2}x\big) := T_d(t,x), \quad x \geq -1, \tag{4.1}$$

where $\widehat{T}_d(x)$ is the unique solution in $(-\infty, -1]$ of $\int_{\widehat{T}_d(x)}^x (x - \xi)(\xi^2 - 1)e^{-\xi^2/2}d\xi = 0$, for all $x \geq -1$. Equivalently, $\widehat{T}_d$ is characterized by:

$$\widehat{T}_d(x) \leq -1 \quad \text{and} \quad e^{-\widehat{T}_d(x)^2/2}\big(1 + \widehat{T}_d(x)^2 - x\widehat{T}_d(x)\big) = e^{-x^2/2}; \quad x \geq -1.$$

Similarly, we see that , where
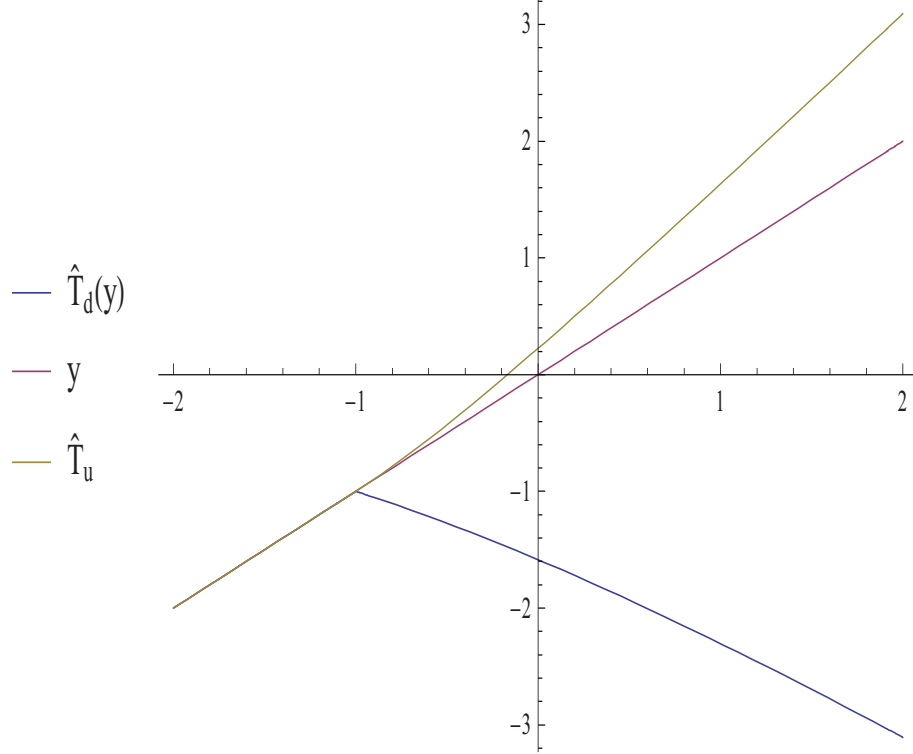
$$j_u(t,x) := t^{-1/2}\widehat{j}_u\big(t^{-1/2}x\big) \tag{4.2}$$

12

Figure 1: *Fake Brownian motion: Maps $\hat{T}_d$ and $\hat{T}_u$*

where the scaled function $\widehat{j}$ is given by:

$$\widehat{j}_u(x) := \frac{1}{2}\Big[x - \widehat{T}_d(x)e^{-(\widehat{T}_d(x)^2 - x^2)/2}\Big] = \frac{1}{2}\Big[x - \frac{\widehat{T}_d(x)}{1 + \widehat{T}_d(x)^2 - x\widehat{T}_d(x)}\Big]; \quad x \geq -1.$$

The maps $\widehat{T}_d(x)$ and $\widehat{T}_u(x) := x + \widehat{j}_u(x)$ are plotted in Figure 1.

4.2   A new construction of self-similar martingales

In Hirsch, Profeta, Roynette & Yor [42], the authors construct martingales M_t which enjoy the (inhomogeneous) Markov property and the Brownian scaling property $\{M_{c^2 t}, t \geq 0\} \sim \{cM_t, t \geq 0\}$ for all $c > 0$. When the marginals of M admit a density $f(t, x)$, this property translates to:

$$f(t, x) = t^{-1/2}f\big(1, t^{-1/2}x\big), \quad t > 0, \ x \in \mathbb{R}.$$

In order to apply our construction result of Theorem 3.7, we assume that $\partial_t F(t, x)$ has a unique maximizer which is given by $m_t = \sqrt{t}\widehat{m}$, where $\widehat{m}$ is the smallest solution of

$$f(1, \widehat{m}) + \widehat{m}f_x(1, \widehat{m}) \quad = \quad 0.$$

The scaling properties (4.1)-(4.2) apply again in this case, and we compute that

$$\int_{\widehat{T}_d(x)}^{x} (x - \zeta)\big(f(1, \zeta) + \zeta f_x(1, \zeta)\big)d\zeta = 0, \ \text{and} \ \widehat{j}_u(x) := \frac{1}{2}\Big[x - \frac{\widehat{T}_d(x)f(1, \widehat{T}_d(x))}{f(1, x)}\Big]; \ x \geq \widehat{m}.$$

13

5   Application: Robust sub/superhedging of variance swap

As a final application, we consider the reward function $c_0(x, y) := (\ln x - \ln y)^2$, which clearly satisfies Assumption 2.2. The so-called "variance swap" is defined by the payoff $\sum_{k=0}^{n-1} c_0(X_{t_k}, X_{t_{k+1}})$, and is usually approximated by the continuous-time limit of Lemma 2.4:

$$C_0(X) := \int_0^1 \frac{d[X]_t^c}{X_t^2} + \sum_{0 < t \le 1} \ln^2 \left(\frac{X_t}{X_{t-}} \right).$$

We recall that, under any pathwise continuous martingale model $\mathbb{P}^c \in \mathcal{M}_\infty$, the variance swap can be perfectly replicated by a semi-static hedging strategy. This follows from a direct application of Itô's formula:

$$C_0(X) = \int_0^1 \frac{d[X]_t}{X_t^2} = 2 \int_0^1 \frac{dX_t}{X_t} - 2 \ln \left(\frac{X_1}{X_0} \right), \quad \mathbb{P}^c - \text{a.s.}$$

However, since our set of admissible models $\mathcal{M}_\infty$ also allows for models with jumps, it turns out that the upper and lower optimal bounds for the variance swap define an interval containing $\int \ln(x/X_0) \mu_1(dx)$ in its interior.

In the following statement, $(\mu_t)_{0 \le t \le 1}$ is a non-decreasing family of marginals with support in $(0, \infty)$, and satisfying Assumptions 3.1 and 3.4. The corresponding left-monotone and right-monotone martingales are defined by the characteristics (m, j_u, j_d) and $(\tilde{m}, \tilde{j}_u, \tilde{j}_d)$, respectively. In addition, we assume that the constructed optimal static strategy λ^* satisfies the integrability conditions in Theorem 3.11,

Proposition 5.1. *The optimal martingale measures are given by the local Lévy processes* (3.6) *and* (3.12), *and the corresponding optimal upper and lower bounds for the variance swap are given by*

$$\int_0^1 \!\!\! \int_{m_t}^\infty \frac{j_u(t,x)}{j_d(t,x)} \ln^2 \left[\frac{x}{x - j_d(t,x)} \right] \mu_t(dx) dt \quad and \quad \int_0^1 \!\!\! \int_0^{\tilde{m}_t} \frac{\tilde{j}_d(t,x)}{\tilde{j}_u(t,x)} \ln^2 \left[\frac{x + \tilde{j}_u(t,x)}{x} \right] \mu_t(dx) dt$$

We have compared these bounds to market values denoted VS_{mkt} for the DAX index (2-Feb-2013) for different maturities (see Table 1). The market marginals $(\mu_t)_{0 \le t \le 1}$ are induced from market prices $C_t(K)$ of call options, with strike K and maturity t, by $\mu_t(dx) = C_t''(dx)$.[4]

In Figure 2, we have plotted market marginals $\partial_t F(t, K) = -\partial_t \partial_K C(t, K)$ for different maturities t and checked that $\partial_t F$ admits only one local maximizer.

The prices in Table 1 are quoted in volatility $\times 100$. Note that for maturities less than 1.5 years, our upper bound is below the market price, highlighting an arbitrage opportunity. In practice, this arbitrage disappears if we include transaction costs for trading vanilla options with low/high strikes. Recall also that we have assumed that vanilla options with all maturities are traded.

---

[4]Let $\sigma_t^{\text{BS}}(K)$ be the Black-Scholes implied volatility of strike K and maturity t, and denote by $V_t^{\text{BS}}(K) := t\sigma_t^{\text{BS}}(K)^2$ the corresponding total variance. In our numerical experiments, we have used the following interpolation from the observed maturities t_1, t_2:

$$V_t^{\text{BS}}(K)^2 = \left(\frac{t - t_1}{t_2 - t_1} \right) \left[V_{t_2}^{\text{BS}}(K) - V_{t_1}^{\text{BS}}(K) \right] + V_{t_1}^{\text{BS}}(K).$$

This linear interpolation guarantees that $C(t_1, K) \le C(t, K) \le C(t_2, K)$ for all K, and thus preserves the consistency of the interpolation with the no-arbitrage condition.
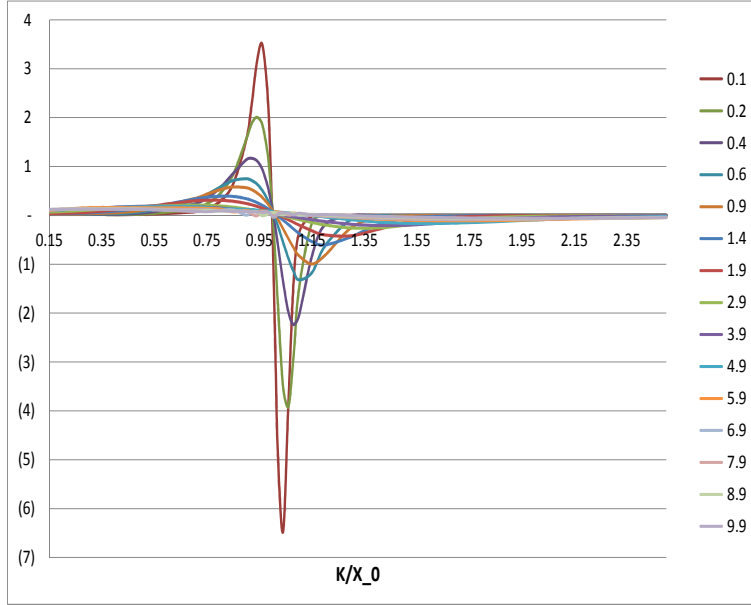
Figure 2: Market marginals $\partial_t F(t, K) = \partial_t \partial_K C(t, K)$ for different (liquid) maturities t inferred from DAX index (2-Feb-2013). For each t, $\partial_t F(t, \cdot)$ admits only one local maximizer.

Maturity (years)	VS_{mkt}	Upper	Lower
0.4	18.47	18.45	16.73
0.6	19.14	18.70	17.23
0.9	20.03	19.63	17.89
1.4	21.77	21.62	19.03
1.9	22.89	23.06	19.63

Table 1: Implied volatility for variance swap as a function of the maturity - DAX index (2-Feb-2013). Lower/upper bounds versus market prices (quoted in volatility $\times 100$).

6 Proofs

6.1 Tightness of the monotone martingale plan under the finitely-many maxima condition of the difference of cdf

In order to prove Proposition 3.2, we start by recalling in Subsection 6.1.1 the explicit characterization of the left-monotone transference plan, as derived in [40], and we collect the required properties for our convergence results in Subsection 6.1.2.

6.1.1 Reminders on the left-monotone martingale transport plan

Let (μ_0, μ_1) be a pair of scalar atomless probability measures with $\mu_0 \preceq \mu_1$ in convex order. For the purpose of the present paper, the corresponding cdf F_0 and F_1 are assumed to be continuous. We also assume that (μ_0, μ_1) is irreducible which translates as a condition on the difference of cdf $\delta F := F_1 - F_0$:

$$\int_x^\infty \delta F(y) dy \; < \; 0, \quad \text{for all} \quad x \in (l_{\mu_1}, r_{\mu_1}),$$

and we denote $x_0 := \inf \{x \in \mathbb{R} \; : \delta F \text{ increasing on a right neighborhood of } x\}$.

We further assume that δF has a finite number of local maximizers

$$\mathbf{M}(\delta F) := \{m_1^0, \ldots, m_n^0\}, \quad l_{\mu_1} < m_1^0 < \ldots < m_n^0 < r_{\mu_1}.$$

Denote by F_1^{-1} the right-continuous inverse of F_1, and

$$g(x, y) \quad := \quad F_1^{-1}\big(F_0(x) + \delta F(y)\big).$$

and we set $m_1 := m_1^0$, and we introduce the right-continuous functions $T_u, T_d : \mathbb{R} \to \mathbb{R}$ by

$$T_d(x) := t^{A_k}(x, m_k), \quad T_u(x) := g(x, T_d(x)); \quad x \in [m_k, x_k)$$
$$\text{and} \; T_d(x) = T_u(x) = x; \quad x \in (x_{k-1}, m_k],$$

where:

- $A_k := (x_0, m_k] \setminus \big(\cup_{i<k} \{T_d([m_i, x_i)) \cup [m_i, x_i)\}\big) \;=\; (x_0, m_k] \setminus \{\cup_{i<k} \big(T_d(x_i), x_i\big]\};$
- for $x > m_k$, $t^{A_k}(x, m_k)$ is the unique point in A_k such that

$$\int_{-\infty}^x \big[F_1^{-1}(F_0(\xi)) - \xi\big] dF_0(\xi) + \int_{-\infty}^{t^{A_k}(x,m_k)} \mathbf{1}_{A_k}(\xi)\big(g(x, \xi) - \xi\big) d\delta F(\xi) = 0; \tag{6.1}$$

- for $k \geq 2$, define

$$x_k := \inf \{x > m_k \; : g\big(x, t^{A_k}(x, m_k)\big) \leq x\} \quad \text{and} \quad m_{k+1} := \inf \big(\mathbf{M}(\delta F) \cap [x_k, \infty)\big).$$

Let $D_0 := \cup_{k \geq 0}(x_k, m_{k+1}]$ and $D_0^c := (l_{\mu_1}, r_{\mu_1}) \setminus D_0$, by the continuity of F_0 and F_1, it follows from the last construction that T_d and T_u take values in D_0 and D_0^c, respectively. Both functions are continuous except on points $(x_k)_{k \geq 1}$ and $\big(T_d^{-1}(x_k-)\big)_{k \geq 1}$, where T_d^{-1} denotes the right-continuous version of the inverse function of T_d. Moreover, one has $T_d(x) \leq x \leq T_u(x)$ for all $x \in \mathbb{R}$.

Remark 6.1. (i) *In the case δF has only one local maximizer m_1^0, we have $D_0 = (l_{\mu_1}, m_1^0]$ and $D_0^c = (m_1^0, r_{\mu_1})$, T_d maps D_0^c to D_0 and T_u maps D_0^c to D_0^c.*
(ii) *In [40], the functions T_u and T_d are obtained by solving the ODE*

$$d(\delta F \circ T_d) = -(1 - q)dF_0, \quad d(F_1 \circ T_u) = qdF_0, \tag{6.2}$$

on the continuity domain of T_d, where

$$q(x) \quad := \quad \mathbf{1}_{\{T_u(x) - T_d(x) = 0\}} + \frac{x - T_d(x)}{T_u(x) - T_d(x)}\mathbf{1}_{\{T_u(x) - T_d(x) \neq 0\}}. \tag{6.3}$$

Given the two functions T_d and T_u, we introduce a probability kernel T_* from $\mathbb{R}$ to $\mathbb{R}$:

$$T_*(x, dy) \quad := \quad \mathbf{1}_{D_0}(x)\delta_x(dy) + \mathbf{1}_{D_0^c}(x)\big[q(x)\delta_{T_u(x)}(dy) + (1 - q(x))\delta_{T_d(x)}(dy)\big]. \qquad (6.4)$$

We next introduce the optimal dual semi-static hedging strategy. The dynamic hedging function $h_* : \mathbb{R} \to \mathbb{R}$ is defined up to a constant by:

$$h'_*(x) := \frac{c_x(x, T_u(x)) - c_x(x, T_d(x))}{T_u(x) - T_d(x)}, \quad \text{for} \quad x \in D_0^c, \qquad (6.5)$$

$$h_*(x) := h_*\big(T_d^{-1}(x)\big) + c_y(x, x) - c_y\big(T_d^{-1}(x), x\big), \quad \text{for} \quad x \in D_0.$$

The stating hedging functions ψ_* and φ_* are defined, also up to a constant, by

$$\psi'_* := c_y\big(T_u^{-1}, \cdot\big) - h_* \circ T_u^{-1} \quad \text{on} \quad D_0^c, \quad \psi'_* := c_y\big(T_d^{-1}, \cdot\big) - h_* \circ T_d^{-1} \quad \text{on} \quad D_0,$$

$$\varphi_*(x) := q(x)\big(c(x, T_u(x)) - \psi_*(T_u(x))\big) + (1 - q(x))\big(c(x, T_d(x)) - \psi_*(T_d(x))\big); \quad x \in (l_{\mu_0}, r_{\mu_0}).$$

Finally, the free constants are fixed by requiring that:

$$c(\cdot, T_u(\cdot)) - \psi_*(T_u(\cdot)) - c(\cdot, T_d(\cdot)) + \psi_*(T_d(\cdot)) - \big(T_u(\cdot) - T_d(\cdot)\big)h_*(\cdot) \quad \text{is continuous.}$$

Theorem 6.2. *[40, Theorems 4.5 and 5.1] Suppose that the partial derivative c_{xyy} exists and $c_{xyy} > 0$ on $(l_{\mu_0}, r_{\mu_0}) \times (l_{\mu_1}, r_{\mu_1})$. Then,*
(i) the probability $\mathbb{P}_*(dx, dy) := \mu_0(dx)T_*(x, dy) \in \mathcal{M}_2(\mu_0, \mu_1)$ is the unique left-monotone martingale transport plan, which solves the primal problem (2.1).
(ii) Assume in addition that the positive parts $\varphi_*^+ \in \mathbb{L}^1(\mu_0)$ and $\psi_*^+ \in \mathbb{L}^1(\mu_1)$. Then $(\varphi_*, \psi_*, h_*) \in \mathcal{D}_2$ (see (2.3)) and it solves the dual problem (2.2); moreover, we have the duality

$$\mathbb{E}^{\mathbb{P}_*}\big[c(X_0, X_1)\big] \;=\; \mathbf{P}_2(\mu_0, \mu_1) \;=\; \mathbf{D}_2(\mu_0, \mu_1) \;=\; \mu_0(\varphi_*) + \mu_1(\psi_*).$$

Remark 6.3. By symmetry, one can also consider the c.d.f. $\tilde{F}_i(x) := 1 - F_i(-x)$, $x \in \mathbb{R}$, $i = 0, 1$, and construct a right monotone martingale transference plan which solves the minimization transportation problem (see more discussions in Remark 5.2 of [40]).

6.1.2 Asymptotic estimates of the left-monotone transference plan

Let $t \in [0, 1)$ and $\varepsilon \in (0, 1 - t]$, we consider the pair $(\mu_t, \mu_{t+\varepsilon})$ of the two marginal measures. Under Assumption 3.1 (iii), an explicit construction of the left-monotone martingale measure corresponding to $(\mu_t, \mu_{t+\varepsilon})$ can be obtained by following the procedure recalled in Section 6.1.1. We denote the corresponding characteristics by $\big(x_k^\varepsilon(t), m_{k+1}^\varepsilon(t)\big)_{k \geq 0}$, $T_u^\varepsilon(t, \cdot)$ and $T_d^\varepsilon(t, \cdot)$. Similarly, denote $D^\varepsilon(t) := \cup_{k \geq 1}\big(x_{k-1}^\varepsilon(t), m_k^\varepsilon(t)\big]$, $(D^\varepsilon)^c(t) := (l_{t+\varepsilon}, m_{t+\varepsilon}) \setminus D^\varepsilon(t)$ and $A_k^\varepsilon(t) := \big(x_0^\varepsilon(t), m_k^\varepsilon(t)\big) \setminus \big\{ \cup_{i<k} \big(T_d^\varepsilon(t, x_i^\varepsilon(t)), x_i^\varepsilon(t)\big]\big\}$. Recall that $\delta^\varepsilon F$ and g_t^ε are introduced in (3.1), and let us define

$$\delta^\varepsilon f(t, x) := f(t + \varepsilon, x) - f(t, x), \quad t \in [0, 1), \ \varepsilon \in (0, 1 - t].$$

Then for every $x \in \big[m_k^\varepsilon(t), x_k^\varepsilon(t)\big)$,
• the function $T_d^\varepsilon(t, x) \in A_k^\varepsilon(t)$ is uniquely determined by

$$\int_{-\infty}^x \big[F^{-1}\big(t + \varepsilon, \ F(t, \xi)\big) - \xi\big]f(t, \xi)d\xi + \int_{-\infty}^{T_d^\varepsilon(t,x)} \mathbf{1}_{A_k^\varepsilon(t)}(\xi)\big[g_t^\varepsilon(x, \xi) - \xi\big]\delta^\varepsilon f(t, \xi)d\xi = 0, \quad (6.6)$$

• the function T_u^ε is given by $T_u^\varepsilon(t, x) := g_t^\varepsilon\big(x, T_d^\varepsilon(t, x)\big)$.

Since the functions T_d^ε and T_u^ε are (piecewise) monotone, we may introduce $(T_d^\varepsilon)^{-1}$ and $(T_u^\varepsilon)^{-1}$ as their right-continuous inverse function in x denotes the of $x \mapsto T_d^\varepsilon(t, x)$. We also introduce the jump size function by

$$J_u^\varepsilon(t, x) := T_u^\varepsilon(t, x) - x, \qquad J_d^\varepsilon(t, x) := x - T_d^\varepsilon(t, x).$$

17

and the probability of a downward jump by

$$q^\varepsilon(t, x) \ := \ \frac{J_u^\varepsilon(t, x)}{J_u^\varepsilon(t, x) + J_d^\varepsilon(t, x)} \mathbf{1}_{\{J_u^\varepsilon(t,x)>0\}} \ = \ \frac{T_u^\varepsilon(t, x) - x}{T_u^\varepsilon(t, x) - T_d^\varepsilon(t, x)} \mathbf{1}_{\{T_u^\varepsilon(t,x)>x\}}.$$

Lemma 6.4. *Let Assumption 3.1 hold true. Then for every $K > 0$, there is a constant C independent of (t, x, ε) such that*

$$J_u^\varepsilon(t, x) \ + \ q^\varepsilon(t, x) \ \le \ C\varepsilon, \qquad \forall x \in [-K, K] \cap (l_t, r_t).$$

Proof. Differentiating g_t^ε (defined below (6.6)), we have

$$\partial_y \, g_t^\varepsilon(x, y) \ = \ \frac{\delta^\varepsilon f(t, y)}{f\big(t + \varepsilon, g_t^\varepsilon(x, y)\big)}.$$

Notice that $|\delta^\varepsilon F(t, x)| + |\delta^\varepsilon f(t, x)| \le C_1 \varepsilon$ for some constant C_1 independent of (t, x, ε). Then for $\varepsilon > 0$ small enough, the value of $g_t^\varepsilon(x, y)$ is uniformly bounded for all $t \in [0, 1]$ and all $x \in [-K, K] \cap (l_t, r_t)$ and $y \in \mathbb{R}$. Further, the density function satisfies $\inf_{x \in [-\bar{K}, \bar{K}] \cap (l_t, r_t)} f(t, x) > 0$ for every $\bar{K} > 0$ large enough, by Assumption 3.1, then it follows by the definition of T_u^ε below (6.6) that

$$q^\varepsilon(t, x) \ \le \ \frac{T_u^\varepsilon(t, x) - x}{x - T_d^\varepsilon(t, x)} \ = \ \frac{g_t^\varepsilon(x, T_d^\varepsilon(t, x)) - g_t^\varepsilon(x, x)}{x - T_d^\varepsilon(t, x)} \ \le \ C\varepsilon.$$

Finally, by the definition of T_u^ε below (6.6), we have

$$
\begin{aligned}
J_u^\varepsilon(t, x) \ &= \ F^{-1}\big(t + \varepsilon, F(t, x) + \delta^\varepsilon F(t, T_d^\varepsilon(t, x))\big) - F^{-1}\big(t + \varepsilon, F(t, x) + \delta^\varepsilon F(t, x)\big) \\
&\le \ \frac{\big|\delta^\varepsilon F\big(t, T_d^\varepsilon(t, x)\big)\big| + |\delta^\varepsilon F(t, x)|}{f\big(t + \varepsilon, F^{-1}\big(t + \varepsilon, F(t, x) + \xi\big)\big)},
\end{aligned}
$$

for some ξ between $\delta^\varepsilon F\big(t, T_d^\varepsilon(t, x)\big)$ and $\delta^\varepsilon F(t, x)$. We can then conclude the proof by the fact that $|\delta^\varepsilon F| \le C_1 \varepsilon$ for some constant C_1. $\qquad \square$

6.1.3   Proof of Proposition 3.2

We recall that $\mathbb{P}^n$ is a martingale measure on the canonical space Ω, induced by the continuous-time martingale $X^{*,n}$ under the probability $\mathbb{P}^{*,n}$. The martingale $X^{*,n}$ jumps only on discrete time grid $\pi_n = (t_k^n)_{1 \le k \le n}$. Moreover, at time t_{k+1}^n, the upward jump size is $J_u^\varepsilon(t_k^n, X_{t_k^n})$ and downward jump size is $J_d^\varepsilon(t_k^n, X_{t_k^n})$ with $\varepsilon := t_{k+1}^n - t_k^n$ (see Section 3.1). Let C, θ be some positive constants, we define

$$\mathcal{E}_n(C, \theta) \ := \ \inf\Big\{ \Pi_{i=j}^{k-1}\big(1 - C(t_{i+1}^n - t_i^n)\big) \ : \ \text{for some } s \in [0, 1) \text{ and } 0 \le j \le k \le n$$

$$\text{such that } s \le t_j^n \le t_{k+1}^n \le s + \theta \Big\}.$$

Since $|\pi_n| := \max_{1 \le k \le n}(t_k^n - t_{k-1}^n) \longrightarrow 0$, it follows that $\mathcal{E}_n(C, \theta) \longrightarrow e^{-C\theta}$ as $n \longrightarrow \infty$.

(i) To prove the tightness of $(\mathbb{P}^n)_{n \ge 1}$, we shall use Theorem VI.4.5 of Jacod & Shiryaev [54, P. 356].

First, Doob's martingale inequality implies that

$$\mathbb{P}^n\Big[\sup_{0 \le t \le 1} |X_t| \ge K\Big] \ \le \ \frac{\mathbb{E}^{\mathbb{P}^n}[|X_1|]}{K} \ = \ \frac{1}{K}\int_{\mathbb{R}} |x| \mu_1(dx) \ =: \ \frac{L_1}{K}, \qquad \forall K > 0. \qquad (6.7)$$

Let $\eta > 0$ be an arbitrary small real number, then there is some $K > 0$ such that

$$\mathbb{P}^n\Big[\sup_{0 \le t \le 1} |X_t| \ge K\Big] \ \le \ \eta, \quad \text{for all } n \ge 1.$$

18

We can assume K large enough so that $-K < m_t < K$ for all $t \in [0,1]$. Denote then $r^K(t) := r_t \wedge K$ and $l^K(t) := l_t \vee (-K)$.

Let $\delta > 0$, it follows by Lemma 6.4 that the upward jump size $J_u^\varepsilon(t, x)$ is uniformly bounded by $C\varepsilon$ for some constant C on $D_\delta^K := \{(t, x) : m_t \leq x \leq r^K(t) - \delta/2\}$. We then consider $\theta > 0$ small enough such that $\theta \leq \frac{\delta}{2C}$ and $|l^K(t + \theta) - l^K(t)| + |r^K(t + \theta) - r^K(t)| \leq \delta/2$ for all $t \in [0, 1 - \theta]$. Let S, T be two stopping times w.r.t to the filtration generated by $X^{*,n}$ such that $0 \leq S \leq T \leq S + \theta \leq 1$. When $\sup_{0 \leq t \leq 1} |X_t^{*,n}| \leq K$ and $X^{*,n}$ only increases between S and $S + \theta$, then clearly $|X_T^{*,n} - X_S^{*,n}| < \delta$. Therefore

$$\mathbb{P}^{*,n}\Big[\sup_{0 \leq t \leq 1} |X_t^{*,n}| \leq K, \; \big|X_T^{*,n} - X_S^{*,n}\big| \geq \delta\Big]$$

$$\leq \quad \mathbb{P}^{*,n}\Big[\sup_{0 \leq t \leq 1} |X_t^{*,n}| \leq K, \text{ and there is a down jump of } X^{*,n} \text{ on } [S, S + \theta]\Big]$$

$$\leq \quad 1 - \mathcal{E}_n(C, \theta),$$

where the last inequality follows by the estimate of q^ε in Lemma 6.4. Then it follows that

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

$$\leq \quad \limsup_{\theta \to 0} \limsup_{n \to \infty} \left(\mathbb{P}^{*,n}\Big[\sup_{0 \leq t \leq 1} |X_t^{*,n}| \leq K, \; \big|X_T^{*,n} - X_S^{*,n}\big| \geq \delta\Big] + \mathbb{P}^{*,n}\Big[\sup_{0 \leq t \leq 1} |X_t^{*,n}| \geq K\Big]\right)$$

$$\leq \quad \limsup_{\theta \to 0} \limsup_{n \to \infty} \; \big(1 - \mathcal{E}_n(C, \theta)\big) \; + \; \eta \; = \; \eta.$$

Since $\eta > 0$ is an arbitrary small real number, we then obtain that

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

Then it follows by Theorem VI.4.5 of [54] that the sequence $(X^{*,n}, \mathbb{P}^{*,n})_{n \geq 1}$ is tight, and hence $\big(\mathbb{P}^n\big)_{n \geq 1}$ is tight.

(ii) Let $\mathbb{P}^0$ be a limit of $(\mathbb{P}^n)_{n \geq 1}$, let us now check that $\mathbb{P}^0 \circ X_t^{-1} = \mu_t$ for every $t \in [0, 1]$. By extracting the sub-sequence, we suppose that $\mathbb{P}^n \to \mathbb{P}^0$, then $\mathbb{P}^{*,n} \circ (X_t^{*,n})^{-1} = \mathbb{P}^n \circ X_t^{-1} \to \mathbb{P}^0 \circ X_t^{-1}$. By the construction of $X^{*,n}$, there is a sequence $(s_n)_{n \geq 1}$ in $[0, 1]$ such that $s_n \to t$ and $X_t^{*,n} = X_{s_n}^{*,n} \sim \mu_{s_n}$ under $\mathbb{P}^{*,n}$. It follows by the continuity of the distribution function $F(t, x)$ that $\mu_{s_n} \to \mu_t$, and hence $\mathbb{P}^0 \circ X^{-1} = \mu_t$.

(iii) Finally, let us show that X is still a martingale under $\mathbb{P}^0$. For every $K > 0$, denote $X_t^K := (-K) \vee X_t \wedge K$. Let $s < t$ and $\varphi(s, X_.)$ be a bounded continuous, $\mathcal{F}_s$-measurable function, by weak convergence, we have

$$\mathbb{E}^{\mathbb{P}^n}\big[\varphi(s, X_.)(X_t^K - X_s^K)\big] \longrightarrow \mathbb{E}^{\mathbb{P}^0}\big[\varphi(s, X_.)(X_t^K - X_s^K)\big].$$

Moreover, since the marginals $(\mu_t)_{t \in [0,1]}$ form a peacock, and hence are uniformly integrable, it follows that

$$\big|\mathbb{E}^{\mathbb{P}^n}\big[\varphi(s, X_.)(X_t^K - X_s^K)\big]\big| \quad \leq \quad 2|\varphi|_\infty \sup_{r \leq 1} \int |x| \mathbf{1}_{\{|x| \geq K\}} \mu_r(dx) \; \longrightarrow 0, \quad \text{as} \quad K \to \infty,$$

uniformly in n. Then, by the fact that X is a $\mathbb{P}^n$−martingale, we have $\mathbb{E}^{\mathbb{P}^0}\big[\varphi(s, X_.)(X_t - X_s)\big] = 0$. By the arbitrariness of φ, this proves that X is a $\mathbb{P}^0$−martingale. $\square$

6.2 Convergence to the peacock under the one-maximizer condition on the cdf difference

We recall that under Assumption 3.4, the function $x \mapsto \partial_t F(t, x)$ has one unique local maximizer on (l_t, r_t). Recall also in this context, one has $D(t) := (l_t, m_t]$, $D^c(t) := (m_t, r_t)$ and

$$D^c := \{(t, x) : t \in [0, 1], x \in D^c(t)\}, \quad D := \{(t, x) : t \in [0, 1], x \in D(t)\}.$$

6.2.1 Full-marginals left-monotone martingale transport

This section is dedicated to the

Proof of Lemma 3.6. We first rewrite (3.3) into:

$$G\big(t, x, T_d(t,x)\big) = 0, \quad \text{where} \quad G(t,x,y) := \int_y^x (x - \xi)\partial_t f(t,\xi)d\xi. \tag{6.8}$$

Step 1: We first prove the existence and uniqueness of T_d. By the non-decrease of $(\mu_t)_{t\in[0,1]}$ in convex order in Assumption 3.1, it follows that

$$G(t, x, -\infty) = \int_{-\infty}^x (x - \xi)\partial_t f(t,\xi)d\xi \geq 0.$$

Since $\mathbf{M}(\partial_t F(t,.)) = \{m_t\}$, the function $x \longmapsto \partial_t F(t,x)$ is strictly increasing on the interval $(l_t, m_t]$, implying that the last inequality is strict, i.e.

$$G(t, x, -\infty) > 0.$$

The same argument also implies that $y \longmapsto G(t,x,y)$ is strictly decreasing on (l_t, m_t).

Next, notice that $G(t, m_t, m_t) = 0$ and $x \longmapsto G(t, x, m_t)$ is decreasing on interval (m_t, r_t):

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

In summary, for every $x \in (m_t, r_t)$, we have $G(t, x, m_t) < 0$, $G(t, x, -\infty) > 0$ and $y \mapsto G(t,x,y)$ is continuous, strictly decreasing on (l_t, m_t). It follows that the equation (6.8) has a unique solution $T_d(t,x)$ and it takes values in (l_t, m_t), which implies that the equation (3.3) has a unique solution in (l_t, m_t).

Step 2: We next prove (i). Differentiating both sides of equation (3.3) w.r.t. $x \in (m_t, r_t)$, it follows that

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

Therefore, for every $x \in (m_t, r_t)$,

$$\partial_x T_d(t,x) = \frac{\partial_t F(t,x) - \partial_t F(t, T_d(t,x))}{\big(x - T_d(t,x)\big)\,\partial_t f\big(t, T_d(t,x)\big)} < 0, \tag{6.9}$$

and hence $x \mapsto T_d(t,x)$ is strictly decreasing in x on interval (m_t, r_t).

Step 3: It remains to prove (ii). We define, for $0 < \delta < K < \infty$,

$$E_\delta := \{(t,x) \in D^c : m_t < x < m_t + \delta\}, \tag{6.10}$$

$$E_{\delta,K} := \{(t,x) \in D^c : m_t + \delta \leq x \leq (m_t + K) \wedge r_t\}. \tag{6.11}$$

Let us first prove that $j_d \mathbf{1}_{D^c}$ is locally Lipschitz in x on $[m_t, r_t)$, it is enough to verify that $\partial_x T_d \mathbf{1}_{D^c}$ is locally bounded. From (6.9), we have

$$\partial_x T_d(t,x) = \frac{\partial_t F(t,x) - \partial_t F(t, T_d(t,x))}{\big(x - T_d(t,x)\big)\,\partial_t f\big(t, T_d(t,x)\big)}, \quad (t,x) \in D^c.$$

It is clear that $\partial_x T_d(t,x)$ is continuous on D^c and hence bounded on $E_{\delta,K}$ for every $0 < \delta < K$. We then focus on the case $(t,x) \in E_\delta$. Since $\partial_t f(t, m_t) = 0$ and $\partial_{tx} f(t, m_t) < 0$ by Assumption 3.4, we have

$$\partial_t f(t,\xi) = \partial_{tx} f\big(t, m_t\big)\big(\xi - m_t\big) + C_1(t,\xi)\big(\xi - m_t\big)^2,$$

where $C_1(t, \xi)$ is uniformly bounded for $|\xi - m_t| \le \delta$. Inserting the above expression into (6.8), it follows that

$$\int_{T_d(t,x)}^{x} (x - \xi)(\xi - m_t)d\xi \;=\; C_2(t,x)(x - T_d(t,x))^4,$$

where C_2 is also uniformly bounded on E_δ since $\min_{0 \le t \le 1} \partial_{t,x} f(t, m_t) < 0$ by Assumption 3.4. By direct computation, it follows that

$$(x - T_d(t,x))^2 (x - m_t + 2(m_t - T_d(t,x))) = C_2(t,x)(x - T_d(t,x))^4,$$

which implies that

$$T_d(t,x) = m_t - \frac{1}{2}(x - m_t) + C_2(t,x)(x - T_d(t,x))^2, \tag{6.12}$$

Using again the expression (6.9), we have

$$\partial_x T_d(t,x) \;=\; -\frac{1}{2} \;+\; C_3(t,x)(x - T_d(t,x)),$$

where C_3 is also uniformly bounded on E_δ. Finally, by the uniqueness of solution T_d of (6.8), we get

$$\partial_x T_d(t,x) \;=\; -\frac{1}{2} \;+\; C_4(t,x)(x - m_t), \tag{6.13}$$

for some C_4 uniformly bounded on E_δ, implying that $T_d \mathbf{1}_{D^c}$ is locally Lipschitz in x. Moreover, by the expression of j_u in (3.4), i.e.

$$j_u(t,x) \;:=\; \frac{\partial_t F(t, T_d(t,x)) - \partial_t F(t,x)}{f(t,x)},$$

together with (6.13), it is easy to check hat $j_u \mathbf{1}_{D^c}$ and $(j_u/j_d)\mathbf{1}_{D^c}$ are also locally Lipschitz in x.

To prove that these functions are also locally Lipschitz in t, we consider $\partial_t T_d(t,x)$. By direct computation, we obtain

$$\partial_t T_d(t,x) \;=\; \frac{\int_{T_d(t,x)}^{x} (x - \xi)\partial_{tt}^2 f(t,\xi)d\xi}{(x - T_d(t,x))\partial_t f(t, T_d(t,x))},$$

which is clearly continuous in (t,x) on D^c, and hence uniformly bounded on $E_{\delta,K}$, for $K > \delta > 0$. Using again (6.12), it is easy to verify that $\partial_t T_d$ is also uniformly bounded on $E_{0,\delta}$, and hence $\partial_t T_d(t,x)$ is also locally bounded on D^c. Therefore, $j_d(t,x)\mathbf{1}_{x > m_t}$, $j_u(t,x)\mathbf{1}_{x > m_t}$ and $\frac{j_u}{j_d}(t,x)\mathbf{1}_{x > m_t}$ are all locally Lipschitz. $\qquad\square$

Remark 6.5. *From the previous proof, we have in particular that*

$$\partial_t T_d(t, m_t + \delta) \;\xrightarrow{\delta \searrow 0}\; -\frac{3}{2} \frac{\partial_{tt}^2 f(t, m_t)}{\partial_{tx}^2 f(t, m_t)} \quad \text{uniformly in } t \in [0,1].$$

6.2.2 Additional asymptotic estimates of the left-monotone transference plan

We shall use the notations E_δ and $E_{\delta,K}$ as defined in (6.10)-(6.11).

Lemma 6.6. *Let Assumptions 3.1 and 3.4 hold true. Then J_u^ε and J_d^ε admit the expansion*

$$J_u^\varepsilon(t,x) = \varepsilon j_u^\varepsilon(t,x) + \varepsilon^2 e_u^\varepsilon(t,x), \quad \text{and} \quad J_d^\varepsilon(t,x) = j_d(t,x) + (\varepsilon \vee \rho(\varepsilon)) \, e_d^\varepsilon(t,x), \tag{6.14}$$

where j_d is defined in (3.4), and

$$j_u^\varepsilon(t,x) \;\; := \;\; \frac{\partial_t F\big(t, x - J_d^\varepsilon(t,x)\big) - \partial_t F(t,x)}{f(t+\varepsilon, x)}.$$

Moreover, for all $0 < \delta < K < \infty$, $e_u^\varepsilon(t,x), e_d^\varepsilon(t,x)$ are uniformly bounded; and consequently, there is constant $C_{\delta,K}$ such that q^ε admits the asymptotic expansion:

$$q^\varepsilon(t,x) \;\; = \;\; \varepsilon \frac{j_u(t,x)}{j_d(t,x)} \;+\; C_{\delta,K}\, \varepsilon\big(\varepsilon \vee \rho_0(\varepsilon)\big), \quad for\; (t,x) \in E_{\delta,K}. \tag{6.15}$$

Proof. Let $\delta < K$ be fixed. Notice that by its definition, the function $T_d(t,x)$ is continuous on compact set $E_{\delta,K}$ and hence it is uniformly bounded.

(i) By the definition of T_u^ε below (6.6), we have

$$T_u^\varepsilon(t,x) \;\; = \;\; F^{-1}\big(t+\varepsilon,\; F(t+\varepsilon,x) + \delta^\varepsilon F(t, T_d^\varepsilon(t,x)) - \delta^\varepsilon F(t,x)\big).$$

By direct expansion, we see that the first equality in (6.14) holds true with

$$|e_u^\varepsilon(t,x)| \;\; \leq \;\; \sup_{t \leq s \leq t+\varepsilon,\; T_d^\varepsilon(s,x) \leq \xi \leq x} \frac{2\partial_{tt} F(s,\xi) \partial_x f(s,\xi)}{f^3(s,\xi)}.$$

(ii) Let us now consider the second equality in (6.14). First,

$$\int_{-\infty}^x [F^{-1}(t+\varepsilon, F(t,\xi)) - \xi] f(t,\xi)d\xi = \int_{-\infty}^x \xi\, \delta^\varepsilon f(t,\xi)d\xi + \int_x^{F^{-1}(t+\varepsilon, F(t,x))} \xi f(t+\varepsilon,\xi)d\xi$$

$$= \int_{-\infty}^x \xi\, \delta^\varepsilon f(t,\xi)d\xi - \delta^\varepsilon F(t,x)\big(x + C_1(t,x)\varepsilon\big)$$

$$= \int_{-\infty}^x (\xi - x)\, \delta^\varepsilon f(t,\xi)d\xi \;+\; C_2(t,x)\varepsilon^2,$$

where $|C_1(t,x)| \leq |F^{-1}(t+\varepsilon, F(t,x)) - x|^2 |f|_\infty$ and $|C_2(t,x)| \leq |C_2(t,x)||\partial_t F|_\infty$.

We next note that $g_t^\varepsilon(x,\xi) = x + C_3(t,x,\xi)\varepsilon$, where $|C_3(t,x,\xi)| \leq 2\frac{|\partial_t F|_\infty}{m_f}$. Then it follows by direct computation that Further, for every $k \geq 1$,

$$\int_{-\infty}^{T_d^\varepsilon(t,x)} \big(g_t^\varepsilon(x,\xi) - \xi\big)\delta f(t,\xi)d\xi \;\; = \;\; \int_{-\infty}^{T_d^\varepsilon(t,x)} (x - \xi)\, \delta^\varepsilon f(t,\xi)d\xi \;+\; C_4(t,x)\varepsilon^2,$$

where $|C_4(t,x)| \leq 2\frac{|\partial_t F|_\infty}{m_f}|\partial_t F|_\infty$. Combining the above estimates with (6.6), it follows that

$$\int_{T_d^\varepsilon(t,x)}^x (x - \xi)\frac{1}{\varepsilon}\delta^\varepsilon f(t,\xi)\, d\xi \;\; = \;\; \big(C_2(t,x) \vee C_4(t,x)\big)\varepsilon.$$

It follows then

$$\int_{T_d^\varepsilon(t,x)}^x (x - \xi)\partial_t f(t,\xi)\, d\xi \;\; = \;\; C_5(t,x)\big(\varepsilon \vee \rho(\varepsilon)\big),$$

where $|C_5(t,x)| \leq (x + K)(|\partial_t f|_\infty + |\partial_{tt}^2 f|_\infty)$. This implies the first estimation in (6.14) since $\partial_t f(t,x) > 0$ for $x \in (l_t, m_t)$. $\qquad\square$

Lemma 6.7. *Under Assumptions 2.2, 3.1 and 3.4, we have*

$$T_d^\varepsilon \mathbf{1}_{\{x > m^\varepsilon(t)\}} \to T_d \mathbf{1}_{\{x > m(t)\}}, \;\; h^\varepsilon \to h^*, \;\; \partial_t \psi^\varepsilon \to \partial_t \psi^*, \;\; and \;\; \psi^\varepsilon \to \psi^*,$$

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

Proof. (i) In the one local maximizer case under Assumption 3.4, the definition of $T_{\hat{d}}^{\varepsilon}(t,x)$ in (6.6) is reduced to be

$$\int_{-\infty}^{x}\left[F^{-1}\big(t+\varepsilon,F(t,\xi)\big)-\xi\right]f(t,\xi)d\xi \;+\; \int_{-\infty}^{T_{\hat{d}}^{\varepsilon}(t,x)}\left[g_{t}^{\varepsilon}(x,\xi)-\xi\right]\delta^{\varepsilon}f(t,\xi)d\xi \;=\; 0,$$

or equivalently

$$\int_{T_{\hat{d}}^{\varepsilon}(t,x)}^{x}\xi\,\delta^{\varepsilon}f(t,\xi)d\xi \;+\; \int_{x}^{T_{u}^{\varepsilon}(t,x)}\xi f(t+\varepsilon,\xi)d\xi \;=\; 0, \tag{6.16}$$

with $T_{u}^{\varepsilon}(t,x):=g_{t}^{\varepsilon}\big(x,T_{\hat{d}}^{\varepsilon}(t,x)\big)$. Differentiating (6.16), it follows that

$$\partial_{t}T_{\hat{d}}^{\varepsilon}(t,x) \;:=\; -\,\frac{A^{\varepsilon}(t,x)}{\big(T_{u}^{\varepsilon}-T_{\hat{d}}^{\varepsilon}\big)\delta^{\varepsilon}f\big(t,T_{\hat{d}}^{\varepsilon}(\cdot)\big)}\,(t,x), \tag{6.17}$$

with

$$\begin{aligned}
A^{\varepsilon}(t,x) \;:=&\; \int_{T_{\hat{d}}^{\varepsilon}(t,x)}^{x}\xi\partial_{t}\delta^{\varepsilon}f(t,\xi)d\xi \;+\; \int_{x}^{T_{u}^{\varepsilon}(t,x)}\xi\partial_{t}f(t+\varepsilon,\xi)d\xi \\
&+\; T_{u}^{\varepsilon}(t,x)\Big(\partial_{t}F(t,x)-\partial_{t}F\big(t+\varepsilon,T_{u}^{\varepsilon}(t,x)\big)+\partial_{t}\delta^{\varepsilon}F\big(t,T_{\hat{d}}^{\varepsilon}(t,x)\big)\Big), \\
=&\; -\big(T_{u}^{\varepsilon}(t,x)-x\big)\big(\partial_{t}\delta^{\varepsilon}F(t,x)-\partial_{t}\delta^{\varepsilon}F(t,T_{\hat{d}}^{\varepsilon}(t,x))\big) \\
&-\; \int_{x}^{T_{u}^{\varepsilon}(t,x)}\big(T_{u}^{\varepsilon}(t,x)-\xi\big)\partial_{t}f(t+\varepsilon,\xi)d\xi \;-\; \int_{T_{\hat{d}}^{\varepsilon}(t,x)}^{x}(x-\xi)\partial_{t}\delta^{\varepsilon}f(t,\xi)d\xi
\end{aligned}$$

and

$$\partial_{x}T_{\hat{d}}^{\varepsilon}(t,x) \;:=\; -\,\frac{T_{u}^{\varepsilon}(t,x)-x}{\big(T_{u}^{\varepsilon}(t,x)-T_{\hat{d}}^{\varepsilon}(t,x)\big)\,\delta^{\varepsilon}f\big(t,T_{\hat{d}}^{\varepsilon}(t,x)\big)}\,f(t,x), \tag{6.18}$$

where the last term is exactly the same as that induced by ODE (6.2).

(ii) Taking the limit $\varepsilon\to 0$, it follows by direct computation and the convergence $T_{\hat{d}}^{\varepsilon}(t,x)\to T_{d}(t,x)$ in Lemma 6.6 that $\partial_{x}T_{\hat{d}}^{\varepsilon}(t,x)\to\partial_{x}T_{d}(t,x)$ and $\partial_{t}T_{\hat{d}}^{\varepsilon}(t,x)\to\partial_{t}T_{d}(t,x)$ for every $(t,x)\in D^{c}$. Moreover, by the local uniform convergence result in Lemma 6.6, we deduce that $\partial_{x}T_{\hat{d}}^{\varepsilon}$ and $\partial_{t}T_{\hat{d}}^{\varepsilon}$ also converge locally uniformly. Denote $T_{\hat{d}}^{0}:=T_{d}$, it follows that the mapping $(t,x,\varepsilon)\to\big(\partial_{t}T_{\hat{d}}^{\varepsilon}(t,x),\partial_{x}T_{\hat{d}}^{\varepsilon}(t,x)\big)$ is continuous on

$$\overline{E} \;:=\; \big\{(t,x,\varepsilon) \;:\; t\in[0,1],\; \varepsilon\in[0,1-t],\; m^{\varepsilon}(t)<x<r^{\varepsilon}(t)\big\},$$

where $m^{0}(t):=m_{t}$ and $r^{0}(t):=r_{t}$.

(iii) By exactly the same computation as in Proposition 3.12 of [40], we have

$$\partial_{x}T_{\hat{d}}^{\varepsilon}(t,x) \;=\; \Big(1+O(\varepsilon)+O\big(x-T_{\hat{d}}^{\varepsilon}\big)\Big)\frac{\big(x-m^{\varepsilon}(t)\big)-\frac{1}{2}\big(x-T_{\hat{d}}^{\varepsilon}\big)+O\big(\big(x-T_{\hat{d}}^{\varepsilon}\big)^{2}\big)}{\big(x-m^{\varepsilon}(t)\big)-\big(x-T_{\hat{d}}^{\varepsilon}\big)+O\big(\big(x-T_{\hat{d}}^{\varepsilon}\big)^{2}\big)}(t,x),$$

and it follows by similar arguments as in [40] that

$$T_{\hat{d}}^{\varepsilon}(t,x)-m^{\varepsilon}(t) \;=\; -\frac{1}{2}\big(x-m^{\varepsilon}(t)\big) \;+\; O\big((x-m^{\varepsilon})^{2}\big),$$

and hence

$$\partial_{x}T_{\hat{d}}^{\varepsilon}(t,m^{\varepsilon}(t)+\delta) \to -\frac{1}{2}\;\text{ uniformly for } t\in[0,1)\text{ and }\varepsilon\in[0,\varepsilon_{0}\wedge(1-t)],\text{ as }\delta\searrow 0. \tag{6.19}$$

23

Next, using the estimation (6.19) and the definition of T_u^ε, we have

$$T_u^\varepsilon(t,x) - x = C_1(\varepsilon,t,x)\big(x - T_d^\varepsilon(t,x)\big)^2 \quad \text{and} \quad \frac{T_u^\varepsilon(t,x) - x}{\varepsilon} = C_2(\varepsilon,t,x)\big(x - T_d^\varepsilon(t,x)\big).$$

Therefore, by direct computation,

$$\frac{1}{\varepsilon}A^\varepsilon(t,x) = -\frac{1}{2}\big(x - T_d^\varepsilon(t,x)\big)^2 \partial_t \frac{1}{\varepsilon}\delta^\varepsilon f(t, m^\varepsilon(t)) + C_3(\varepsilon,t,x)\big(x - m^\varepsilon(t)\big)^3.$$

It follows by the uniform convergence in (6.19) that

$$\partial_t T_d^\varepsilon(t,x) = -\frac{3}{2}\frac{\partial_t \delta^\varepsilon f(t, m^\varepsilon(t))}{\partial_x \delta^\varepsilon f(t, m^\varepsilon(t))} + C_4(\varepsilon,t,x)\big(x - m^\varepsilon(t)\big), \tag{6.20}$$

where we notice that C_4 is uniformly bounded for $\varepsilon > 0$ and $x - T_d^\varepsilon(m^\varepsilon(t))$ small enough. Finally, the two uniform convergence results in (6.19) and (6.20) together with the continuity of $(t,x,\varepsilon) \to \big(\partial_t T_d^\varepsilon(t,x), \partial_x T_d^\varepsilon(t,x)\big)$ implies that $\partial_t T_d^\varepsilon(t,x)$ and $\partial_x T_d^\varepsilon(t,x)$ are uniformly bounded on $\overline{E} \cap \{(t,x,\varepsilon) : |x| \le m^\varepsilon(t) + K\}$ for every $K > 0$.

(iv) Therefore, it follows by Arzelà-Ascoli's theorem that T_d^ε converges to T_d locally uniformly. Finally, by the local uniform convergence of $T_d^\varepsilon \to T_d$, together with the estimations in (6.13) and (6.19), it is easy to deduce the local uniform convergence of $h^\varepsilon \to h$, $\partial_t \psi^\varepsilon \to \partial_t \psi^*$ and $\psi^\varepsilon \to \psi^*$ as $\varepsilon \to 0$. $\qquad\square$

6.2.3 Convergence to the peacock

Recall that the sequence of martingale measures $(\mathbb{P}^n)_{n\ge 1}$ induced by the left-monotone transference plan is tight by Proposition 3.2, we now show that any limit of $(\mathbb{P}^n)_{n\ge 1}$ provides a weak solution of (3.6). Let

$$
\begin{aligned}
M_t(\varphi, \mathbf{x}) := &\ \varphi(\mathbf{x}_t) - \int_0^t j_u(s, \mathbf{x}_{s-}) D\varphi(\mathbf{x}_{s-}) \mathbf{1}_{\mathbf{x}_{s-} > m(s)} ds \\
&+ \int_0^t \left[\big[\varphi(\mathbf{x}_{s-} - j_d(s,\mathbf{x}_{s-})) - \varphi(\mathbf{x}_{s-})\big] \frac{\dot{j}_u}{\dot{j}_d}(s,\mathbf{x}_{s-}) \right] \mathbf{1}_{\mathbf{x}_{s-} > m(s)} ds, \quad (6.21)
\end{aligned}
$$

for all $\mathbf{x} \in \Omega := \mathbb{D}([0,1], \mathbb{R})$ and $\varphi \in C^1(\mathbb{R})$. Then the process $M(\varphi, X)$ is clearly progressively measurable w.r.t. the canonical filtration $\mathbb{F}$. For the martingale problem, we also need to use the standard localization technique in Jacod & Shiryaev [54]. In preparation, let us introduce, for every constant $p > 0$, an $\mathbb{F}$-stopping time and the corresponding stopped canonical process

$$\tau_p := \inf\big\{t \ge 0 : |X_t| \ge p \ \text{or} \ |X_{t-}| \ge p\big\}, \quad X_t^p := X_{t \wedge \tau_p}. \tag{6.22}$$

Following [54], denote also $J(\mathbf{x}) := \{t > 0 : \Delta\mathbf{x}(t) \neq 0\}$,

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

Proof of Theorem 3.7. By extracting subsequences, we can suppose without loss of generality that $\mathbb{P}^n \to \mathbb{P}^0$ weakly. To prove that $\mathbb{P}^0$ is a weak solution of SDE (3.6), it is sufficient to show that $\big(M_t(\varphi, X)\big)_{t\in[0,1]}$ is a local martingale under $\mathbb{P}_0$ for every $\varphi \in C_b^1(\mathbb{R})$. Since the functions j_u and j_d are only locally Lipschitz (not uniformly bounded) by Lemma 3.10, we need to adapt the localization technique in Jacod & Shiryaev [54], by using the stopping time τ_p defined by (6.22). Our proof will be very similar to that of Theorem IX.3.39 in [54].

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

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

where, denoting $\varepsilon_k^n := t_{k+1}^n - t_k^n$,

$$\alpha_i := \mathbb{E}_{t_k^n}^{\mathbb{P}^n}\left[\left\{\varphi\big(X_{t_k^n} + J_i^{\varepsilon_k^n}(t_k^n, X_{t_k^n})\big) - \varphi(X_{t_k^n})\right\}\frac{J_i}{J_d + J_u}\mathbf{1}_{X_{t_k^n} \geq m^{\varepsilon_k^n}(t_k^n)}\right], \quad i = d, u.$$

By (6.15) in Lemma 6.6 and the uniform continuity of $m^{\varepsilon_k^n}(t)$, we have

$$\alpha_u = \mathbb{E}_{t_k^n}^{\mathbb{P}^n}\left[\int_{t_k^n}^{t_{k+1}^n} D\varphi(X_s)j_u(s, X_s)\mathbf{1}_{X_s \geq m(s)}ds\right] + O\big(\varepsilon_k^n(\varepsilon_k^n \vee \rho_0(\varepsilon_k^n))\big),$$

where ρ_0 is the continuity modulus of $(t, \varepsilon_k^n) \mapsto m^{\varepsilon_k^n}(t)$ in Assumption 3.4. Similarly,

$$\alpha_d = \mathbb{E}_{t_k^n}^{\mathbb{P}^n}\left[\int_{t_k^n}^{t_{k+1}^n}\big(\varphi(X_s - j_d(s, X_s)) - \varphi(X_s)\big)\frac{j_u}{j_d}(s, X_s)\mathbf{1}_{X_s \geq m(s)}ds\right] + O\big(\varepsilon_k^n(\varepsilon_k^n \vee \rho_0(\varepsilon_k^n))\big).$$

Therefore, 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 Ω such that $\phi : \Omega \to \mathbb{R}$ is continuous under the Skorokhod topology, we have

$$\mathbb{E}^{\mathbb{P}^n}\left[\phi_s(X_\cdot)\big(M_{t \wedge \tau_p}(\varphi, X) - M_{s \wedge \tau_p}(\varphi, X)\big)\right] \leq C_p\big(|\pi_n| \vee \rho_0(|\pi_n|)\big), \qquad (6.23)$$

for some constant $C_p > 0$.

To proceed, we follow the same localization arguments as in the proof of Theorem IX.3.39 of Jacod & Shiryaev [54]. Since $\mathbb{P}^n \to \mathbb{P}^0$ as $n \to \infty$, then for every $p \in \mathbb{N}$, the distribution of the stopped process X^p under $\mathbb{P}^n$ also converges, i.e. there is $\mathbb{P}^{0,p}$ such that $\mathcal{L}^{\mathbb{P}^n}(X_\cdot^p) \to \mathbb{P}^{0,p}$ as $n \to \infty$. Due to the proof of Proposition IX.1.17 of [54], there are at most countably-many $a > 0$ such that

$$\mathbb{P}^{0,p}\big(\omega : a \in V(\omega) \cup V'(\omega)\big) > 0.$$

So we can choose $a_p \in [p-1, p]$ such that

$$\mathbb{P}^{0,p}\big[\omega : a_p \in V(\omega) \cup V'(\omega)\big] = 0.$$

It follows by Theorem 2.11 of [54] that $\omega \mapsto \tau_{a_p}(\omega)$ is $\mathbb{P}^{0,p}$-a.s. continuous and the law $\mathcal{L}^{\mathbb{P}^n}(X_\cdot^p, X_\cdot^{a_p})$ converges to $\mathcal{L}^{\mathbb{P}^{0,p}}(X, X^{\tau_{a_p}})$.

Denote by $\tilde{\mathbb{P}}^{0,p}$ the law of $X^{\tau_{a_p}}$ on $(\Omega, \mathcal{F}, \mathbb{P}^{0,p})$, we then have $\omega \mapsto \tau_{a_p}(\omega)$ is $\tilde{\mathbb{P}}^{0,p}$-a.s. continuous and $\mathcal{L}^{\mathbb{P}^n}(X^{a_p}) \to \tilde{\mathbb{P}}^{0,p}$. In particular, since there is a countable set $\mathbb{T}^* \subset [0,1]$ such that

$$\mathbf{x} \mapsto M_{t \wedge \tau_{a_p}}(\varphi, \mathbf{x}) - M_{s \wedge \tau_{a_p}}(\varphi, \mathbf{x}) \qquad (6.24)$$

is $\tilde{\mathbb{P}}^{0,p}$-almost surely continuous for all $s < t$ such that $s, t \notin \mathbb{T}^*$. Therefore, by taking the limit of (6.23), we obtain

$$\mathbb{E}^{\tilde{\mathbb{P}}^{0,p}}\left[\phi_s(X_\cdot)\big(M_t(\varphi, X) - M_s(\varphi, X)\big)\right] = 0,$$

whenever $s \leq t$ and $t \notin \mathbb{T}^*$. Combining with the right-continuity of $M_t(\varphi, \mathbf{x})$, we know $\tilde{\mathbb{P}}^{0,p}$ is a solution of the martingale problem (6.21) between 0 and τ_{a_p}, i.e. $\big(M_{t \wedge \tau_{a_p}}(\varphi, X)\big)_{0 \leq t \leq 1}$ is a martingale under $\tilde{\mathbb{P}}^{0,p}$. Moreover, since $\tilde{\mathbb{P}}^{0,p} = \mathbb{P}^0$ in restriction to $(\Omega, \mathcal{F}_{\tau_{a_p}})$ and $\tau_{a_p} \to \infty$ as $p \to \infty$, it follows by taking the limit $p \to \infty$ that $\big(M_t(\varphi, X)\big)_{0 \leq t \leq 1}$ is a local martingale under $\mathbb{P}^0$, i.e. $\mathbb{P}^0$ is a solution to the martingale problem (6.21) and hence a weak solution to SDE (3.6).

Finally, for uniqueness of solutions to SDE (3.6), it is enough to use Theorem III-4 of Lepeltier & Marchal [61] (see also Theorem 14.18 of Jacod [53, P. 453]) together with localization technique to conclude the proof. $\qquad\square$

Remark 6.8. *In the multiple local maximizers case under Assumption 3.1, the functions j_u and j_d are no more continuous, then the mapping (6.24) may not be a.s. continuous and the limiting argument thereafter does not hold true. This is the main reason for which we restrict to the one maximizer case under Assumption 3.4 in Theorem 3.7.*

Proof of Lemma 3.8. We recall that by Theorem 3.8 in [40] (ii), the corresponding maps $T_u^\varepsilon(t,.)$ and $T_d^\varepsilon(t,.)$ solve the following ODEs:

$$\frac{d}{dx}\,\delta^\varepsilon F\big(t+\varepsilon, T_d^\varepsilon(t,x)\big) \;=\; (1-q)(t,x)f(t,x), \qquad (6.25)$$

$$\frac{d}{dx}F\big(t+\varepsilon, T_u^\varepsilon(t,x)\big) \;=\; q(t,x)f(t,x) \quad \text{for all } x \in (D^\varepsilon)^c(t), \qquad (6.26)$$

where $\delta^\varepsilon F(t+\varepsilon,.) := F(t+\varepsilon,.) - F(t,.)$. With the asymptotic estimates

$$T_d^\varepsilon(t,x) - x = -j_d(t,x) + \circ(\varepsilon) \quad \text{and} \quad T_u^\varepsilon(t,x) - x = \varepsilon j_u(t,x) + O(\varepsilon),$$

which is locally uniform by Lemma 6.6. By direct substitution of this expression in the system of ODEs (6.25-6.26), we see that the limiting maps (j_d, j_u) of $(T_u^\varepsilon, T_d^\varepsilon)$, as $\varepsilon \searrow 0$, satisfy the following system of first order partial differential equations (PDEs):

$$\partial_x j_d(t,x) = 1 + \frac{j_u(t,x)}{j_d(t,x)}\frac{f(t,x)}{\partial_t f(t,x-j_d(t,x))}, \qquad \partial_x\{j_u f\}(t,x) = -\partial_t f(t,x) - \frac{j_u(t,x)}{j_d(t,x)}f(t,x).$$

Since $x \in D^c(t)$ and $x - j_d(t,x) \in D(t)$, it follows directly that (3.7) holds true. $\qquad \square$

Proof of Proposition 3.9. By Lemma 3.8, item (ii) of Proposition 3.9 is a direct consequence of item (i), then we only need to prove (i).

Let $x \in \mathbb{R}$, the function $y \mapsto (y-x)^+$ is continuous and smooth on both $(-\infty, x]$ and $[x, \infty)$, then it follows by Itô's lemma that

$$d\big(\widehat{X}_t - x\big)^+ \;=\; dM_t \;+\; L_t \qquad (6.27)$$
$$:= \; \mathbf{1}_{\{\widehat{X}_{t-}>x\}}d\widehat{X}_t \;+\; \Big(\big(\widehat{X}_t - x\big)^+ - \big(\widehat{X}_{t-} - x\big)^+ - \mathbf{1}_{\{\widehat{X}_{t-}>x\}}\Delta\widehat{X}_t\Big),$$

where $(M_t)_{0\le t\le 1}$ is a local martingale. Notice that $\mathbf{1}_{\{\widehat{X}_{t-}>x\}}$ is bounded and $\widehat{X}_1 \in L^p$ for some $p > 1$. Using BDG inequality and then Doob's inequality, it is a standard result that $(M_t)_{0\le t\le 1}$ is a real martingale. Further, the local Lévy process $\widehat{X}$ is clearly quasi left continuous. Moreover, since

$$L_s \;=\; \Big(x - T_d\big(s, \widehat{X}_{s-}\big)\Big)\mathbf{1}_{\left\{T_d\big(s,\widehat{X}_{s-}\big) \,\le\, x \,<\, \widehat{X}_{s-},\ \widehat{X}_{s-}\in D^c(s)\right\}} \;\le\; j_d\big(s,\widehat{X}_{s-}\big)\mathbf{1}_{\left\{\widehat{X}_{s-}\in D^c(s)\right\}}$$

by direct computation, it follows by (3.8) together with dominated convergence theorem that

$$\mathbb{E}\Big[\sum_{t\le s\le t+\varepsilon} L_s\Big] = \mathbb{E}\Big[\int_t^{t+\varepsilon}\Big(x - T_d\big(s,\widehat{X}_{s-}\big)\Big)\frac{j_u}{j_d}\big(s,\widehat{X}_{s-}\big)\mathbf{1}_{\left\{T_d\big(s,\widehat{X}_{s-}\big)\le x<\widehat{X}_{s-},\ \widehat{X}_{s-}\in D^c(s)\right\}}ds\Big],$$

for every $\varepsilon \le \varepsilon_0$, where $\varepsilon_0 \in (0, 1-t]$. Then, integrating (6.27) between t and $t+\varepsilon$, and taking expectations, it follows that

$$\mathbb{E}\Big[\big(\widehat{X}_{t+\varepsilon} - x\big)^+\Big] - \mathbb{E}\Big[\big(\widehat{X}_t - x\big)^+\Big]$$
$$= \int_t^{t+\varepsilon}\int_{\mathbb{R}}\big(x - T_d(s,y)\big)\frac{j_u}{j_d}(s,y)\mathbf{1}_{\left\{T_d(s,y)\le x<y,\ y\in D^c(s)\right\}}f^{\widehat{X}}(s,y)\,dy\,ds. \qquad (6.28)$$

Let us now differentiate both sides of (6.28). For the left hand side, since the density function $f^{\widehat{X}}(t,.)$ of $\widehat{X}_t$ is continuous, the function $x \mapsto \mathbb{E}[(\widehat{X}_t - x)^+] = \int_x^\infty (y-x)f^{\widehat{X}}(t,y)dy$ is differentiable and

$$\partial_x \mathbb{E}\left[(\widehat{X}_t - x)^+\right] = \int_x^\infty -f^{\widehat{X}}(t,y)dy, \qquad \partial_{xx}^2 \mathbb{E}\left[(\widehat{X}_t - x)^+\right] = f^{\widehat{X}}(t,x). \qquad (6.29)$$

We now consider the rhs of (6.28) and denote

$$l(s,x) := \int_{\mathbb{R}} (x - T_d(s,y))\frac{j_u}{j_d}(s,y)\mathbf{1}_{\left\{T_d(s,y)\leq x < y,\ y\in D^c(s)\right\}}f^{\widehat{X}}(s,y)\ dy.$$

Let us fix $s \in [0,1)$ and $x \in D^c(s) := (m_s, r_s)$, then it is clear that

$$l(s,x) = \int_x^\infty (x - T_d(s,y))\frac{j_u}{j_d}(s,y)\mathbf{1}_{\left\{T_d(s,y)\leq x,\ y\in D^c(s)\right\}}f^{\widehat{X}}(s,y)dy,$$

where the integrand is smooth in x for every $y \in \mathbb{R}$. Hence for every $x \in D^c(s)$,

$$\partial_x l(s,x) = j_u(s,x)f^{\widehat{X}}(s,x) + \int_x^\infty \frac{j_u}{j_d}(s,y)\mathbf{1}_{\left\{T_d(s,y)\leq x,\ y\in D^c(s)\right\}}f^{\widehat{X}}(s,y)dy,$$

and

$$\partial_{xx}^2 l(s,x) = \partial_x(j_u f^{\widehat{X}})(s,x) - \frac{j_u}{j_d}f^{\widehat{X}}(s,x). \qquad (6.30)$$

We now consider the case $x \in (l_t, m_t)$. Notice that $T_d(s,\cdot) : D^c(s) \to (l_s, m_s)$ is a bijection and $\widehat{X}_s$ admits a density function. It follows that the random variable $T_d(s,\widehat{X}_s)$ also admits a density function on (m_s, r_s), given by

$$f^{\widehat{T}}(s,y) = \frac{f^{\widehat{X}}}{\partial_x T_d}(s,y), \quad \forall y \in (l_s, m_s).$$

Then by the expression that

$$l(s,x) = \int_{-\infty}^x (x - z)\frac{j_u}{j_d}\left(s, T_d^{-1}(s,z)\right)\mathbf{1}_{\{x \leq T_d^{-1}(z)\}}f^{\widehat{T}}(s,z)dz,$$

we get

$$\partial_{xx}^2 l(x) = \frac{j_u}{j_d}\left(s, T_d^{-1}(s,x)\right)f^{\widehat{T}}(s,z) = \frac{j_u}{j_d}\left(s, T_d^{-1}(s,x)\right)\frac{f^{\widehat{X}}}{\partial_x T_d}(s,x), \quad \forall x \in (l_s, m_s). \quad (6.31)$$

Finally, differentiating both sides of (6.28) (with (6.29), (6.30) and (6.31)), then dividing them by ε and sending $\varepsilon \searrow 0$, it follows that

$$\partial_t f^{\widehat{X}}(t,x) = \mathbf{1}_{\{x > m_t\}}\left(\partial_x(f^{\widehat{X}}j_u) - \frac{j_u f^{\widehat{X}}}{j_d}\right)(t,x) - \mathbf{1}_{\{x < m_t\}}\frac{j_u f^{\widehat{X}}}{j_d(1 - \partial_x j_d)}\left(t, T_d^{-1}(t,x)\right),$$

for every $t \in [0,1)$ and $x \in (l_t, r_t) \setminus \{m_t\}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

6.3 Convergence of the robust superhedging strategy

To prove Theorem 3.11, we will consider a special sequence of partitions of $[0,1]$, $\pi_n = (t_k^n)_{0 \leq k \leq n}$, where $t_k^n := k\varepsilon$ with $\varepsilon = \frac{1}{n}$. To avoid heavy notation, we will omit the superscript and simplify t_k^n to t_k. We also recall that under every $\mathbb{P}^n$, we have $\mathbb{P}^n$-a.s. that

$$\sum_{k=0}^{n-1}\left(\varphi^\varepsilon(t_k, X_{t_k}) + \psi^\varepsilon(t_k, X_{t_{k+1}})\right) + \sum_{k=0}^{n-1}h^\varepsilon(t_k, X_{t_k})\left(X_{t_{k+1}} - X_{t_k}\right) \geq \sum_{k=0}^{n-1}c(X_{t_k}, X_{t_{k+1}}). \quad (6.32)$$

By taking the limit of every term, we obtain a superhedging strategy for the continuous-time reward function, and we can then check that this superhedging strategy induces a duality of the transportation problem as well as the optimality of the local Lévy process (3.6).

Let us first introduce $\Psi^* : \Omega \to \mathbb{R}$ by

$$\Psi^*(\mathbf{x}) \quad := \quad \psi^*(1, \mathbf{x}_1) - \psi^*(0, \mathbf{x}_0) - \int_0^1 \Big(\partial_t \psi^*(t, \mathbf{x}_t) + j_u(t, \mathbf{x}_t) \mathbf{1}_{\mathbf{x}_t > m_t} \partial_x \psi^*(t, \mathbf{x}_t) \Big) dt \qquad (6.33)$$

$$+ \int_0^1 \frac{j_u(t, \mathbf{x}_t)}{j_d(t, \mathbf{x}_t)} \mathbf{1}_{\mathbf{x}_t > m_t} \Big(\psi^*(t, \mathbf{x}_t) - \psi^*(t, \mathbf{x}_t - j_d(t, \mathbf{x}_t)) + c(\mathbf{x}_t, \mathbf{x}_t - j_d(t, \mathbf{x}_t)) \Big) dt.$$

Lemma 6.9. *Let Assumptions 3.1 and 3.4 hold true. Then for every càdlàg path $\mathbf{x} \in \mathbb{D}([0,1], \mathbb{R})$ taking value in (l_1, r_1), we have*

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

Proof. By direct computation, we have for every $n \ge 1$,

$$\sum_{k=0}^{n-1} \big(\varphi^\varepsilon(t_k, \mathbf{x}_{t_k}) + \psi^\varepsilon(t_k, \mathbf{x}_{t_{k+1}}) \big) \quad = \quad \sum_{k=1}^{n-1} \big(\psi^\varepsilon(t_{k-1}, \mathbf{x}_{t_k}) - \psi^\varepsilon(t_k, \mathbf{x}_{t_k}) \big) + \psi^\varepsilon(t_{n-1}, \mathbf{x}_1)$$

$$+ \sum_{k=0}^{n-1} \big(\varphi^\varepsilon(t_k, \mathbf{x}_{t_k}) + \psi^\varepsilon(t_k, \mathbf{x}_{t_k}) \big) - \psi^\varepsilon(0, \mathbf{x}_0).$$

First, we have $\psi^\varepsilon(t_{n-1}, \mathbf{x}_1) \to \psi^*(1, \mathbf{x}_1)$ and by Lemma 6.7,

$$\sum_{k=1}^{n-1} \big(\psi^\varepsilon(t_{k-1}, \mathbf{x}_{t_k}) - \psi^\varepsilon(t_k, \mathbf{x}_{t_k}) \big) = - \int_0^1 \sum_{k=1}^{n-1} \partial_t \psi^\varepsilon(s, \mathbf{x}_{t_k}) \mathbf{1}_{s \in [t_k, t_{k+1})} ds \longrightarrow - \int_0^1 \partial_t \psi^*(s, \mathbf{x}_s) ds.$$

Further, when $x > m_t$,

$$\varphi^\varepsilon + \psi^\varepsilon \quad = \quad \psi^\varepsilon - \psi^\varepsilon(., T_u^\varepsilon) + \frac{J_u^\varepsilon}{J_u^\varepsilon + J_d^\varepsilon} \Big(\psi^\varepsilon(., T_u^\varepsilon) + c(., T_d^\varepsilon) - \psi^\varepsilon(., T_d^\varepsilon) \Big) + o(\varepsilon)$$

$$= \quad -\varepsilon j_u \partial_x \psi^\varepsilon + \varepsilon \frac{j_u}{j_d} \Big(\psi^\varepsilon - \psi^\varepsilon(., T_d) + c(., T_d) \Big) + o(\varepsilon).$$

It follows that $\sum_{k=0}^{n-1} \big(\varphi^\varepsilon(t_k, \mathbf{x}_{t_k}) + \psi^\varepsilon(t_k, \mathbf{x}_{t_k}) \big)$ converges to

$$\int_0^1 -\partial_x \psi^*(t, \mathbf{x}_t) j_u(t, \mathbf{x}_t) dt$$

$$+ \quad \int_0^1 \frac{j_u(t, \mathbf{x}_t)}{j_d(t, \mathbf{x}_t)} \mathbf{1}_{\mathbf{x}_t > m_t} \Big(\psi^*(t, \mathbf{x}_t) - \psi^*(t, \mathbf{x}_t - j_d(t, \mathbf{x}_t)) + c(\mathbf{x}_t, \mathbf{x}_t - j_d(t, \mathbf{x}_t)) \Big) dt,$$

which concludes the proof. $\qquad \square$

Lemma 6.10. *Let Assumptions 3.1 and 3.4 hold true, and $\mu(\overline{\lambda}^*) < \infty$. Then for the limit probability measure $\mathbb{P}^0$ given in Theorem 3.7, we have*

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

Proof. We notice that under the limit probability measure $\mathbb{P}^0$, X is a pure jump martingale with intensity $\frac{j_u}{j_d}(s, X_{s-})$. Then by Itô's formula, the following process is a local martingale

$$\psi^*(t, X_t) - \psi^*(0, X_0) - \int_0^t \partial_t \psi^*(t, X_s) dt$$

$$- \quad \int_0^t \Big[j_u(s, X_s) \partial_x \psi^*(s, X_s) + \frac{j_u}{j_d}(s, X_s) \big[\psi^*(s, X_s - j_d(s, X_s)) - \psi^*(s, X_s) \big] \Big] \mathbf{1}_{X_s > m(s)} ds.$$

Moreover, since $\mu(\overline{\lambda}^*) < \infty$, it follows by dominated convergence theorem that

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}^0}\big[\Psi^*(X.)\big] &= \mathbb{E}^{\mathbb{P}^0}\Big[\int_0^1 \frac{j_u}{j_d}(s, X_s)\mathbf{1}_{X_s > m(s)} c(X_s, X_s - j_d(s, X_s))ds\Big] \\
&= \int_0^1 \int_{m_t}^{r_t} \frac{j_u(t, x)}{j_d(t, x)} \, c\big(x, x - j_d(t, x)\big) f(t, x)dxdt,
\end{aligned}
$$

since the marginals of X under $\mathbb{P}^0$ are $(\mu_t)_{0 \le t \le 1}$.

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

$$
Y_t := \sum_{s \le t} |c(X_{s-}, X_s)| - \int_0^t |c(X_{s-}, X_{s-} - j_d(s, X_{s-}))| \frac{j_u(s, X_{s-})}{j_d(s, X_{s-})} \mathbf{1}_{X_{s-} \ge m_t} ds,
$$

is a local martingale. Since $\mu(\overline{\lambda}^*) < \infty$, we have

$$
\begin{aligned}
&\int_0^1 |c(X_{s-}, X_{s-} - j_d(s, X_{s-}))| \frac{j_u(s, X_{s-})}{j_d(s, X_{s-})} \mathbf{1}_{X_{s-} \ge m_t} ds \\
&= \int_0^1 \int_{m_t}^{r_t} \frac{j_u(t, x)}{j_d(t, x)} |c\big(x, x - j_d(t, x)\big)| f(t, x)dxdt \quad < \quad \infty,
\end{aligned}
$$

which implies that Y is a martingale and hence $\mathbb{E}[Y_1] = 0$. Finally, using similar arguments together with dominated convergence theorem, we get that

$$
\mathbb{E}\Big[\sum_{s \le t} c(X_{s-}, X_s)\Big] = \int_0^1 \int_{m_t}^{r_t} \frac{j_u(t, x)}{j_d(t, x)} \, c\big(x, x - j_d(t, x)\big) f(t, x)dxdt,
$$

which concludes the proof. $\qquad\square$

Next, let us consider the limit of the second term on the left hand side of (6.32).

Lemma 6.11. *Let Assumptions 3.1 and 3.4 hold true. Then we have the following convergence in probability under every martingale measure $\mathbb{P} \in \mathcal{M}_\infty$:*

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

Proof. The functions h^ε are all locally Lipschitz uniformly in ε and $h^\varepsilon \to h^*$ locally uniformly, as $\varepsilon \to 0$, by Lemma 6.7. By the right continuity of martingale X, the above lemma is then a direct application of Theorem I.4.31 of Jacod & Shiryaev [54]. $\qquad\square$

Proof of Theorem 3.11. Using (6.32), together with Lemmas 2.4, 6.9 and 6.11, it follows that under every $\mathbb{P} \in \mathcal{M}_\infty$ (i.e. the canonical process X is a martingale under $\mathbb{P}$), we have the superhedging property

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

Further, by weak duality, we have

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

Since $\mathbb{E}^{\mathbb{P}^0}[C(X.)] = \mu(\lambda^*)$ by Lemma 6.10, this implies the strong duality as well as the optimality of the local Lévy process (3.6) and the semi-static superhedging strategy described by (h^*, ψ^*). $\qquad\square$

29

References

[1] Acciaio, B., Beiglböck, M., Penkner, F. and Schachermayer, W. : *A Model-free Version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem*, Math. Finance, 2013.

[2] Albin, J. M. P.: *A continuous non-Brownian motion martingale with Brownian motion martingale distributions*, Statistics and Probability Letters, 78(6):682-686, 2008.

[3] Beiglböck, M., Cox, A. and Huesmann, M. : *Optimal Transport and Skorokhod Embedding*, Preprint, 2013.

[4] Beiglböck, M., Cox, A., Huesmann, M., Perkowsky, N. and Promel, J. : *Pathwise Super-Hedging via Vovks Outer Measure,* Preprint, 2015.

[5] Beiglböck, M., Henry-Labordère, P., Penkner, F. : *Model-independent Bounds for Option Prices: A Mass-Transport Approach*, Finance and Stochastics, 3:477-501, 2013.

[6] Beiglböck, M., Henry-Labordère, P., Touzi, N. : *Monotone Martingale Transport Plans and Skorohod Embedding*, Preprint.

[7] Beiglböck, M. , Juillet, N. : *On a problem of optimal transport under marginal martingale constraints*, Ann. Probab., to appear.

[8] Beiglböck, M. , Nutz, M. : *Martingale Inequalities and Deterministic Counterparts*, Electron. J. Probab, 2014.

[9] Biagini, S., Bouchard, B., Kardaras, C. and Nutz, M. : *Robust Fundamental Theorem for Continuous Processes*, Mathematical Finance, 2015.

[10] Bonnans J.F. and Tan X. : *A model-free no-arbitrage price bound for variance options*, Applied Mathematics & Optimization, Vol. 68, Issue 1, 43-73, 2013.

[11] Bouchard, B. and Nutz, M. : *Arbitrage and duality in nondominated discrete-time models*, Annals of Applied Probability, 25(2), 823-859, 2015.

[12] Breeden, D.T., Litzenberger, R.H. : *Prices of state-contingent claims implicit in options prices*, J. Business, 51, 1978.

[13] Brown, H., Hobson, D., Rogers, L.C.G. : *Robust hedging of barrier options*, Math. Finance, 11(3):285314, 2001.

[14] Campi, L., Laachir, I. and Martini, C. : *Change of numeraire in the two-marginals martingale transport problem*, preprint, 2014.

[15] Carlier, G. : *On a class of multidimensional optimal transportation problems*. J. Convex Anal., 10(2):517529, 2003.

[16] Carr, P. , Geman, H., Madan, D.B., Yor, M. : *From local volatility to local Lévy models*, Quantitative Finance, Oct. 2004.

[17] Cox, A.M.G and Hobson, D.G. : *Skorokhod embeddings, minimality and non-centred target distributions*, Probability Theory and Related Fields, 135(3):395-414, 2006.

[18] Cox, A., Hobson, D. and Obloj, J. : *Pathwise inequalities for local time: Applications to Skorokhod embeddings and optimal stopping*, Annals of Applied Probability, 18 (5), 1870-1896, 2008.

[19] Cox, A. and Obloj, J. : *Robust pricing and hedging of double no-touch options*, Finance and Stochastics, 15(3):573-605, 2011.

[20] Cox, A. and Obloj, J. : *On joint distributions of the maximum, minimum and terminal value of a continuous uniformly integrable martingale*, Stochastic Processes and their Applications, 2015.

[21] Cox, A., Obloj, J. and Touzi, N. : *The Root solution to the multi-marginal embedding problem: an optimal stopping and time-reversal approach*, preprint, 2015.

[22] Cox, A. and Wang, J. : *Root's Barrier: Construction, Optimality and Applications to Variance Options*, Annals of Applied Probability, 23 (3), 859-894, 2013.

[23] Davis, M., Oblój, J., Raval, V. : *Arbitrage Bounds for Weighted Variance Swap Prices*, Mathematical Finance, 2012.

[24] De Marco, S. and Henry-Labordère, P. : *Linking Vanillas and VIX options: A constrained martingale optimal transport problem*, SIAM Journal on Financial Mathematics, 6(1):1171-1194, 2015.

[25] Dolinsky, Y., Soner, H.M. : *Robust hedging and martingale optimal transport in continuous time*, Probability Theory and Related Fields, 160(1-2):391-427, 2014.

[26] Dolinsky, Y., Soner, H.M. : *Robust Hedging under Proportional Transaction Costs*, Finance and Stochastics, 18(2), 327-347, 2014.

[27] Dolinsky, Y., Soner, H.M. : *Convex duality with transaction costs,* arXiv:1502.01735, preprint, 2015.

[28] Fahim, A. and Huang, Y. : *Model-independent Superhedging under Portfolio Constraints*, Finance and Stochastics, 20(1):51-81, 2016.

[29] Fan, J.Y. Hamza, K., Klebaner, F.C. : *Mimicking self-similar processes.* Bernoulli, 21(3):1341-1360, 2015.

[30] Föllmer, H : *calcul d'Itô sans probabilités*, Séminaire de Probabilités, XV: 09, 143-150, LNM 850, 1981.

[31] Fréchet, H : *Sur la distance de deux lois de probabilité.* C.R. Acad. Sci. Paris, 244:689-692, 1957.

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

[33] Gangbo, W. and Święch, A. : *Optimal maps for the multidimensional Monge-Kantorovich problem.* Comm. Pure Appl. Math., 51(1):23–45, 1998

[34] Gassiat, P., Oberhauser, H. and dos Reis, G. : *Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs*, Stochastic Processes and their Applications, 125(12):4601-4631, 2015.

[35] Guo, G., Tan, X. and Touzi, N. *Optimal Skorokhod embedding under finitely-many marginal constraints*, preprint, 2015.

[36] Guo, G., Tan, X. and Touzi, N. *On the monotonicity principle of optimal Skorokhod embedding problem*, preprint, 2015.

[37] Hamza, K., Klebaner, F.C. : *A family of non-Gaussian martingales with Gaussian marginals*, J. Appl. Math. and Stochastic Analysis, Article Id 92723, 2007.

[38] Heinich, H. : *Problème de Monge pour n probabilités.* C.R. Math. Acad. Sci. Paris, 334(9):793-795, 2002.

[39] Henry-Labordère, P., Obloj, J., Spoida, P. and Touzi, N. : *Maximum Maximum of Martingales given Marginals*, Annals of Applied Probability, to appear.

[40] Henry-Labordère, P., and Touzi, N. : *An Explicit Martingale Version of Brenier's Theorem.* Finance and Stochastics, to appear.

[41] Hirsch, F., and Roynette, B. *A new proof of Kellerer's theorem*, ESAIM: Probability and Statistics 16:48-60, 2012.

[42] Hirsch, F., Profeta, C., Roynette, B., Yor, M.: *Constructing self-similar martingales via two Skorokhod embeddings*, Séminaire de Probabilités XLIII Lecture Notes in Mathematics, 451-503, 2011.

[43] Hirsch, F., Profeta, C., Roynette, B., Yor, M. : *Peacocks and Associated Martingales, with Explicit Constructions*, Springer (2011).

[44] Hobson, D. : *Robust hedging of the lookback option*, Finance and Stochastics, 2:329-347, 1998.

[45] Hobson, D. : *The Skorokhod embedding problem and model-independent bounds for option prices*, In Paris-Princeton Lectures on Mathematical Finance 2010, volume 2003 of Lecture Notes in Math., pages 267318. Springer, Berlin, 2011.

[46] Hobson, D. : *Fake Exponential Brownian Motion*, Statistics and Probability Letters, 83(10):2386-2390, 2013.

[47] Hobson, D., Klimmek, M. : *Model independent hedging strategies for variance swaps*, Finance and Stochastics 16, 611-649, 2012.

[48] Hobson, D., Klimmek, M. : *Robust price bounds for the forward starting straddle*, Finance and Stochastics 19 (1), 189-214, 2015.

[49] Hobson, D., Klimmek, M. : *Maximizing functionals of the maximum in the Skorokhod embedding problem and an application to variance swaps*, The Annals of Applied Probability 23 (5), 2020-2052, 2013.

[50] Hobson, D., Klimmek, M. : *Constructing time-homogeneous generalized diffusions consistent with optimal stopping values*, Stochastics: An International Journal of Probability and Stochastic Processes, Vol.83 (No.4-6). 477-503. ISSN 1744-2508, 2011.

[51] Hobson, D., Neuberger, A. : *Robust bounds for forward start options*, Mathematical Finance 22, 31-56, 2012.

[52] Hoeffding, W. : *Mastabinvariante Korrelationstheorie.* Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin, 5:181-233, 1940.

[53] Jacod, J. : *Calcul stochastique et problemes de martingales*, Springer, 1979.

[54] Jacod, J., Shiryaev A.N. : *Limit theorems for stochastic processes*, Springer, 2003.

[55] Jakubowski, A. : *A non-Skorohod topology on the Skorohod space*, Electronic journal of probability, 2(4):1-21, 1997.

[56] Juillet, N. : *Martingales associated to peacocks using the curtain coupling*, preprint, 2014.

[57] Kallblad, S., Tan X. and Touzi, N. : *Optimal Skorokhod embedding given full marginals and Azma-Yor peacocks*, Preprint, 2015.

[58] Karandikar, R.L. : *On pathwise stochastic integration*, Stoch. Proc. and their App. 57(1):11–18, 1995.

[59] Kellerer, H.G. : *Markov-Komposition und eine anwendung auf martingale*, Math. Ann. 198:99-122, 1972.

[60] Knott, M. and Smith, C. : *On a generalization of cyclic monotonicity and distances among random vectors.* Linear Algebra Appl., 199:363371, 1994.

[61] Lepeltier J.-P., Marchal B. : *Problème des martingales et équations différentielles stochastiques associées a un opérateur intégro-différentiel*, Annales de l'I.H.P., section B, 12(1):43-103, 1976.

[62] Madan, D.B., Yor, M. : *Making Markov martingales meet marginals: with explicit constructions*, Bernoulli, 8(4):509-536, 2002.

[63] Obloj, J. : *The Skorokhod embedding problem and its offspring*, Probab. Surv., 1:321-390, 2004.

[64] Oleszkiewicz, K. : *On fake Brownian motions*, Statistics and Probability Letters, 78, 1251-1254, 2008.

[65] Olkin, I. and Rachev, S.T. : *Maximum submatrix traces for positive definite matrices.* SIAM J. Matrix Ana. Appl., 14:390-397, 1993.

[66] Pagès, G. : *A functional co-monotony principle with an application to peacocks and barrier options*, Séminaire de Probabilités, 2012.

[67] Pass, B.: *Uniqueness and the Monge solutions in the multi marginal optimal transport problem.* SIAM Journal on Mathematical Analysis, 43(6):2758-2775, 2011.

[68] Pass, B.: *On the local structure of optimal measures in the multi-marginal optimal transportation problem.* Calculus of Variations and Partial Differential Equations, 43:529–536, 2012.

[69] Pass, B.: *Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions*, Discrete Contin. Dyn. Syst. 34 (2014) 1623-1639. (Special issue on "Optimal transport and applications.")

[70] Pass, B.: *On a class of optimal transportation problems with infinitely many marginals*, SIAM Journal on Mathematical Analysis, 45(4), 2557-2575, 2013.

[71] Rachev, S. T. and Ruschendorf, L. : *Mass Transportation Problems.* In Vol. 1: Theory. Vol. 2: Applications. Springer, Berlin, 1998.

[72] Ruschendorf, L. and Uckelmann, L. : *On the $n-$coupling problem.* J. Multivariate Anal. 81:242–258, 2002.

[73] Stebegg, F. : *Model-Independent Pricing of Asian Options via Optimal Martingale Transport*, Preprint, 2015.

[74] Tan, X. and Touzi, N. : *Optimal Transportation under Controlled Stochastic Dynamics*, Annals of Probability, Vol. 41, No. 5, 3201-3240, 2013.

[75] Villani, C. : *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften, Vol 338, Springer, 2008.