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

Pierre Henry-Labordère † Xiaolu Tan ‡ Nizar Touzi§

February 13, 2014

Abstract

We extend the martingale version of the one-dimensional Brenier's theorem (Fréchet-Hoeffding coupling, see [16, 30]) established in Henry-Labordère and Touzi [21] to the infinitely-many marginals case. In short, their results give an explicit characterization of the optimal martingale transference plans as well as the optimal dual components of a two marginals discrete-time martingale transportation (MT) problem for a large class of reward functions. We consider here the limiting continuous-time case, which leads to an infinitely-many marginals MT problem. By approximation technique, we show that for a class of reward functions, the optimal martingale transference plan is provided by a pure downward jump local Lévy model. In particular, it provides a new construction of the martingale peacock process (PCOC "Processus Croissant pour l'Ordre Convexe," see Hirsch, Profeta, Roynette and Yor [24]), and a new remarkable example of discontinuous fake Brownian motions. Further, as in [21], we also provide a duality result together with dual optimizer in explicit form. Finally, 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 transportation (OT) problem was initially formulated by Monge in his treatise "Théorie des déblais et des remblais". 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 transportation problem consists in minimizing the transportation 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 problem was later relaxed by Kantorovich to minimizing the value $\mathbb{E}^{\mathbb{P}}[c(X_0, X_1)]$ among all probability measures $\mathbb{P}$ such that $X_0 \sim^{\mathbb{P}} \mu_0$ and $X_1 \sim^{\mathbb{P}} \mu_1$. The last formulation is easily seen to admit optimal solutions and has the same value as that of Monge's original problem when the

*X. Tan and N. Touzi gratefully acknowledge the financial support of the ERC 321111 Rofirm, the ANR Isotace, and the Chairs Financial Risks (Risk Foundation, sponsored by Société Générale) and Finance and Sustainable Development (IEF sponsored by EDF and CA).

†Société Générale, Global Market Quantitative Research, pierre.henry-labordere@sgcib.com

‡Université Paris Dauphine, Ceremade, tan@ceremade.dauphine.fr

§Ecole Polytechnique Paris, Centre de Mathématiques Appliquées, nizar.touzi@polytechnique.edu

probability measure μ_0 has no atoms. As for Monge's original problem, the existence of an optimal solution and its characterization has been obtained later when the cost function satisfies the so-called Spence-Mirrless condition, known as Brenier's theorem. We refer to Rachev and Ruschendorf [47] and Villani [50] for a detailed presentation.

The theory has been extended to the multiple marginals case by Gangbo and Święch [18], Carlier [8], Olkin and S.T. Rachev [41], Knott and Smith [36], Rüschendorf and Uckelmann [48], Heinich [20], and Pass [43, 44, 45], etc. We also refer to the full-marginals case addressed by Pass [46].

Recently, a martingale transportation (MT) problem was introduced in Beiglböck, Henry-Labordère and Penkner [3] and in Galichon, Henry-Labordère and Touzi [17]. Given two probability measures μ_0 and μ_1, one considers all (discrete- or continuous-time) martingales (or general stochastic dynamics) X such that $X_0 \sim \mu_0$ and $X_1 \sim \mu_1$, which minimizes a path-dependent cost function. With later developments (see e.g. [2, 5, 12, 13, 49] etc.), some duality results are established, numerical resolutions of the dual problem are proposed and the applications in probability theory and financial mathematics are developed.

This new transportation problem is in fact motivated by the computation of arbitrage-free model-independent bounds of exotic options consistent with market prices of vanilla options. As explained in Breeden and Litzenberger [6], marginal distributions of the underlying stock can be recovered from the market values of vanilla options for all strikes. Then by considering all martingales satisfying the marginal distribution constraints, one can obtain the arbitrage-free model-independent bound for exotic options. Based on the fact that any martingale can be represented as a time-changed Brownian motion, this problem was initially studied through the Skorokhod Embedding Problem (SEP) approach, see e.g. [10, 25, 26, 39] etc., which consists in finding a stopping time τ of Brownian motion B such that B_τ follows a given distribution.

In view of the extension to the multi-marginals case, the MT approach is more natural than the SEP. Moreover, its dual formulation can be naturally interpreted as the minimum superhedging (or maximum subhedging) price for exotic options, and is better suited for possible numerical resolutions.

In the one-dimensional case $d = 1$, for a discrete-time MT problem, Beiglböck and Juillet [4] introduced a left/right monotone martingale transference plan induced by a binomial model, and show that it is in fact unique and is the optimal transference plan for a large class of cost/reward functions. Henry-Labordère and Touzi [21] provide 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 large 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. Concretely, we are given a family of probability measures $(\mu_t)_{0\le t\le 1}$ on $\mathbb{R}$ which is non-decreasing in convex ordering, i.e. $t \mapsto \mu_t(\phi)$ is non-decreasing for every convex function ϕ. Then for every time discretization of the interval $[0, 1]$, we obtain a finite number of marginal distributions along the discretization. Following the construction in [21], there is a binomial model fitted 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 the limit continuous-time process is optimal for a corresponding MT problem with infinitely-many marginals, when the limit exists.

Given a continuous family of marginal distributions which is non-decreasing in convex ordering, a stochastic process fitting all the marginals is called a peacock process (or PCOC "Processus Croissant pour l'Ordre Convexe" in French) in Hirsch, Profeta, Roynette and Yor [24]. 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 [1], Fan, Hamza and Klebaner [14], Hamza and Klebaner [19], Hirsch et al. [23], Hobson [27], Oleszkiewicz [40], Pagès [42] etc.

Our procedure gives a new construction of martingales associated with peacock processes, and in particular a discontinuous fake Brownian motion. Moreover, our constructed martingale is optimal among all martingales with given marginal distributions for a large class of cost/reward functions, i.e. it solves a martingale transportation problem.

The rest of the paper is organized as follows. In Section 2, we recall the martingale version of Brenier's theorem for a discrete-time MT problem in two marginals case as well as its extension in multi-marginals case, established in Henry-Labordère and Touzi [21]. In Section 3, we formulate a continuous-time MT problem under full marginals constraints. The problem is next solved in Section 4. Namely, by taking the limit of the optimal martingale measure for the multi-marginals MT problem, we obtain a continuous-time martingale fitted to the given marginals, or equivalently, a martingale associated with peacock processes. From a point of view of the forward Kolmogorov-Fokker-Planck (KFP) equation, this martingale peacock is related to a local Lévy process. Under additional conditions, 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. As an application in finance, we provide an optimal robust hedging strategy for the variance swap option. In Section 5, we discuss some examples of extremal peacock processes following our construction, including a discontinuous fake Brownian motion and a self-similar martingale. 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 Discrete-time martingale transportation

This section recalls from [21] the martingale version of the one-dimensional Brenier theorem. The corresponding result in the standard optimal transport theory is known as the Fréchet-Hoeffding coupling [16, 30], see also Rachev and Ruschendorf [47].

2.1 The two marginals case

Let μ_0, μ_1 be two probability measures on $\mathbb{R}$ with finite first moments, such that $\mu_0 \preceq \mu_1$ in convex ordering, i.e. $\mu_0(\phi) \le \mu_1(\phi)$ for every convex function, where $\mu(\phi) := \int_{\mathbb{R}} \phi(x)\mu(dx)$ for every probability measure μ and one-sided integrable function ϕ. Denote by F_0 (resp. F_1) the cumulative distribution function of μ_0 (resp. μ_1), and $\delta F := F_1 - F_0$, we assume in addition that μ_0 and μ_1 have no atoms (i.e. F_0 and F_1 are continuous on $\mathbb{R}$) and the function $x \mapsto \delta F(x)$ has at most countable local maximizers, where the set of all local maximum points admits no accumulation points in $[-\infty, \infty)$.

Denote the set of all martingale measures with marginals μ_0 and μ_1 by

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

The two-marginals MT problem is then given by

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

where $c : \mathbb{R}^2 \to \mathbb{R}$ is a reward function such that $c(x,y) \le a(x) + b(y)$ for some $a \in \mathbb{L}^1(\mu_0)$ and $b \in \mathbb{L}^1(\mu_1)$.

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

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

where the collection $\mathcal{D}_2$ of dual components is defined, with notations $(\varphi \oplus \psi)(x,y) := \varphi(x) + \psi(y)$ and $h^{\otimes}(x,y) := h(x)(y-x)$, by

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

As a financial interpretation, $\mathbf{D}_2(\mu_0, \mu_1)$ is the minimum superhedging cost for the path-dependent derivative payoff defined by $c(X_0, X_1)$ by static and dynamic trading strategies (φ, ψ, h). Moreover, under mild conditions, a strong duality (i.e. $\mathbf{P}_2(\mu_0, \mu_1) = \mathbf{D}_2(\mu_0, \mu_1)$) is proved in Beiglböck, Henry-Labordère and Penkner [3].

An explicit construction of the solution to the last MT problem (2.1) and the corresponding dual problem (2.2) is provided in [21], under the condition $c_{xyy} > 0$. Let m_1 be the smallest local maximizer of δF, $x_0 := \inf\{x \in \mathbb{R} \;: \delta F \text{ strictly increasing at } x\}$, we define the two right-continuous functions $T_u, T_d : \mathbb{R} \to \mathbb{R}$ and a sequence $(x_k, m_{k+1})_{k \ge 0}$ by the following procedure: Let

$$A_k \;:=\; (x_0, m_k] \setminus \big(\cup_{i<k} \big\{ T_d([m_i, x_i)) \cup [m_i, x_i) \big\} \big) \;=\; (x_0, m_k] \setminus \big\{ \cup_{i<k} \big(T_d(x_i), x_i\big] \big\};$$

$g(x,y) := F_1^{-1}\big(F_0(x)+\delta F(y)\big)$, for every $x > m_k$, $t^{A_k}(x,m_k)$ is the unique point in A_k such that

$$\int_{-\infty}^{x} \big[F_1^{-1}(F_0(\xi)) - \xi\big] dF_0(\xi) \;+\; \int_{A\cap(-\infty,t^A(x,m)]} \big(g(x,\xi)-\xi\big) d\delta F(\xi) \;=\; 0; \quad (2.3)$$

$x_k := \inf\big\{x > m_k \;:\, g\big(x, t^{A_k}(x,m_k)\big) \le x\big\}$, $m_{k+1} := \inf\big\{m \ge x_k \;:\, m \text{ is a local maximizer of } \delta F\big\}$ and

$$T_d(x) := t^{A_k}(x,m_k), \quad T_u(x) := g(x,T_d(x)) \quad \text{for} \quad m_k \le x < x_k.$$

Denote $D_0 := \cup_{k\ge1}(x_{k-1}, m_k]$, it turns out that, under the condition that F_0 and F_1 are continuous, $T_u(x) \in D_0^c$ and $T_d(x) \in D_0$ for every $x \in D_0^c$. Moreover, both functions are continuous except on points $(x_k)_{k\ge1}$ and $\big(T_d^{-1}(x_k-)\big)_{k\ge1}$, where T_d^{-1} denotes the right-continuous version of the inverse function of T_d.

Remark 2.1. *In the case δF has only one local maximizer m_1, we have $D_0 = (-\infty, m_1]$ and $D_0^c = (m_1,\infty)$, T_d maps from D_0^c to D_0 and T_u maps from D_0^c to D_0^c.*

Remark 2.2. *In [21], the functions T_u and T_d are obtained by solving the ODE*

$$d(\delta F\circ T_d) = -(1-q)dF_0, \quad d(F_1\circ T_u) = q dF_0, \quad \textit{where} \quad q(x) \;:=\; \frac{x - T_d(x)}{T_u(x)-T_d(x)}, \quad (2.4)$$

on the continuity domain of T_d.

With the two functions T_u and T_d, one can then construct a discrete-time martingale (X_0^*, X_1^*) satisfying the marginal constraints (μ_0,μ_1) as follows: (i) X_0^* is a random variable of distribution μ_0; (ii) conditioned on $X_0^* \in D_0$, we have $X_1^* := X_0^*$; (iii) conditioned on $X_0^* \in D_0^c$, X_1^* takes the value $T_u(X_0^*)$ with probability $q(X_0^*)$ and the value $T_d(X_0^*)$ with probability $1-q(X_0^*)$. In other words, the above construction gives a probability kernel T_* from $\mathbb{R}$ to $\mathbb{R}$,

$$T_*(x,dy) \;:=\; \mathbf{1}_{D_0}(x)\delta_x(dy) + \mathbf{1}_{D_0^c}(x)\big[q(x)\delta_{T_u(x)}(dy) + (1-q(x))\delta_{T_d(x)}(dy)\big]. \quad (2.5)$$

Further, they construct a dual component (a superhedging strategy) $(\varphi_*,\psi_*,h_*) \in \mathcal{D}_2$. The dynamic strategy h_* and static strategy ψ_* are defined, up to a constant, by

$$\begin{aligned} h_*'(x) &= \frac{c_x(x,T_u(x)) - c_x(x,T_d(x))}{T_u(x)-T_d(x)}, && \text{for} \quad x \in D_0^c, \quad (2.6)\\ h_*(x) &:= h_*\big(T_d^{-1}(x)\big) + c_y(x,x) - c_y\big(T_d^{-1}(x),x\big), && \text{for} \quad x \in D_0, \end{aligned}$$

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

and

$$\varphi_*(x) \;:=\; q(x)\big(c(x,T_u(x)) - \psi_*(T_u(x))\big) + (1-q(x))\big(c(x,T_d(x)) - \psi_*(T_d(x))\big), \;\forall x \in \mathbb{R},$$

where we set $q(x) := 1$ for $x \in D_0$. Moreover, h_*, ψ_* are chosen such that

$$c(\cdot,T_u(\cdot)) - \psi_*(T_u(\cdot)) - c(\cdot,T_d(\cdot)) + \psi_*(T_d(\cdot)) - \big(T_u(\cdot)-T_d(\cdot)\big)h_*(\cdot) \quad (2.7)$$

is a continuous function.

Theorem 2.3. *[21, Theorem 3.13] Suppose that the partial derivative c_{xyy} exists and $c_{xyy} > 0$ on $\mathbb{R} \times \mathbb{R}$. Then,*
(i) *the probability $\mathbb{P}_*(dx, dy) := \mu_0(dx)T_*(x, dy) \in \mathcal{M}_2(\mu_0, \mu_1)$ and $(\varphi_*, \psi_*, h_*) \in \mathcal{D}_2$;*
(ii) *the martingale transference plan $\mathbb{P}_*$ solves the primal problem (2.1) and (φ_*, ψ_*, h_*) 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 2.4. *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 3.14 of [21]).*

2.2 The multi-marginals case

The above result are easily extended, in Section 4 of [21], to the multi-marginals case when the reward function is given by $c(x) := \sum_{i=1}^n c^i(x_{i-1}, x_i)$, $\forall x \in \mathbb{R}^{n+1}$. 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

$$\mathbb{E}\big[c(X_0, \cdots, X_n)\big] \;=\; \mathbb{E}\Big[\sum_{i=1}^n c^i(X_{i-1}, X_i)\Big] \tag{2.8}$$

among all martingales $(X_0, \cdots, X_n)$ satisfying the marginal distribution constraints ($X_i \sim \mu_i, i = 0, \cdots, n$). For every (μ_{i-1}, μ_i), we construct the corresponding functions (T^i_d, T^i_u) as well as T^i_* and $(\varphi^i_*, \psi^i_*, h^i_*)$ as in (2.5, 2.6). Assume $c^i_{xyy} > 0$ for every $1 \le i \le n$, it follows that the optimal martingale measure is given by $\mathbb{P}^*_n(dx) = \mu_0(dx_0)\Pi_{i=1}^n T^i_*(x_{i-1}, dx_i)$, and $(\varphi^i_*, \psi^i_*, h^i_*)_{1 \le i \le n}$ is an optimal superhedging strategy, i.e. for all $(x_0, \cdots, x_n) \in \mathbb{R}^{n+1}$,

$$c(x_0, \cdots, x_n) \;\le\; \sum_{i=1}^n \big(\varphi^i_*(x_{i-1}) + \psi^i_*(x_i)\big) \;+\; \sum_{i=1}^n h^i_*(x_{i-1})\big(x_i - x_{i-1}\big).$$

3 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 2.2 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 $\Omega := D([0, 1], \mathbb{R})$ 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. 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. By Karandikar [34], there is an 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 $[X]^c_\cdot$ the continuous part of the non-decreasing process $[X]_\cdot$.

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 [35] (see also Hirsch and Roynette [22]), $\mathcal{M}_\infty(\mu)$ is nonempty if and only if the family $(\mu_t)_{0\le t\le 1}$ admits finite first order moment, is non-decreasing in convex ordering, and $t \mapsto \mu_t$ is right-continuous. Suppose that every μ_t is supported on the smallest interval $[l(t), r(t)]$, where $(l,r) : [0,1] \to (\mathbb{R}\cup\{-\infty\}) \times (\mathbb{R}\cup\{+\infty\})$ satisfy $-\infty \le l(t) < r(t) \le \infty$. Denote also

$$l_0 := l(0), \quad l_1 := l(1), \quad r_0 := r(0), \quad r_1 := r(1).$$

Similar to Hobson and Klimmek [28], 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 (2.8), 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 3.1. *The reward function $c : \mathbb{R}^2 \to \mathbb{R}$ is in $C^3(\mathbb{R}^2)$ and satisfies*

$$c(x,x) = c_y(x,x) = 0 \quad \text{and} \quad c_{xyy}(x,y) > 0, \quad \forall (x,y) \in (l_1,r_1)\times(l_1,r_1). \tag{3.1}$$

Further, to obtain the convergence, we also need to use the technique of pathwise Itô's analysis introduced in Föllmer [15], which is also used in Hobson and Klimmek [28] and Davis, Oblój and Raval [11] (see in particular their Appendix B).

Definition 3.2 (Föllmer [15])**.** *Let $(\pi_n)_{n\ge1}$ be a sequence of partitions of $[0,1]$, i.e. $\pi_n = (0 = t_0^n < \cdots < t_n^n = 1)$, such that $|\pi_n| := \max_{1\le k\le n} |t_k^n - t_{k-1}^n| \to 0$ as $n \to \infty$. A càdlàg path $\mathbf{x} : [0,1] \to \mathbb{R}$ has a finite quadratic variation along $(\pi_n)_{n\ge1}$ if the sequence of measures on $[0,1]$,*

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

converges weakly to a measure $[\mathbf{x}]^F$ on $[0,1]$. We then denote $[\mathbf{x}]^F_t := [\mathbf{x}]^F([0,t])$ which is clearly a non-decreasing process, and by $[\mathbf{x}]^{F,c}_\cdot$ its continuous part.

The following convergence result follows the same line of proof as in Lemma 7.4 of Hobson and Klimmek [28].

Lemma 3.3. *Let Assumption 3.1 hold true. Then for every path $\mathbf{x} \in \Omega$ with finite quadratic variation $[\mathbf{x}]^F$ along a sequence of partition $(\pi_n)_{n\ge1}$, we have*

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

Notice that $[\mathbf{x}]^F$ depends on the sequence of partitions $(\pi_n)_{n\geq 1}$ and it is not defined for every path $\mathbf{x} \in \Omega$. Therefore, we also use the non-decreasing process $[\mathbf{x}]$ introduced in Karandikar [34] which is defined for every $\mathbf{x} \in \Omega$ and coincides "almost surely" with the "quadratic variation" in the martingale case.

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

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

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

$$\mathbf{P}_\infty(\mu) \;:=\; \sup_{\mathbb{P}\in\mathcal{M}_\infty(\mu)} \mathbb{E}^{\mathbb{P}}\big[C(X_\cdot)\big]. \tag{3.2}$$

Remark 3.4. (i) Under the condition $c(x,x) = c_y(x,x) = 0$ in Assumption 3.1, we have $|c(x, x+\Delta x)| \leq K(x)\Delta x^2$ for all $\Delta x \in [-1,1]$ with some positive function $K(x)$ which is locally bounded. Therefore, the sum $\sum_{0\leq t\leq 1} c(X_{t^-}, X_t)$ is in fact finite for almost every path, under every martingale measure $\mathbb{P} \in \mathcal{M}_\infty$.
(ii) Let us fix a martingale probability $\mathbb{P} \in \mathcal{M}_\infty$, under which the canonical process X is a martingale and hence $\sum_{t_k\in\pi_n} \big(X_{t_k} - X_{t_{k-1}}\big)^2$ converges in probability to its quadratic variation. By considering a sub-sequence of $(\pi_n)_{n\geq 1}$, it follows that $\mathbb{P}$-almost every path admits finite quadratic variation, denoted by $[X]^F$, along this sub-sequence in sense of Definition 3.2. It follows that $[X] = [X]^F$, $\mathbb{P}$-a.s.

Now, let us introduce the dual formulation of the above MT problem (3.2). We first introduce the class of admissible dynamic and static strategies. Denote by $\mathbb{H}_0$ the class of all locally bounded processes $H : [0,1] \times \Omega \to \mathbb{R}$ which are predictable w.r.t. the canonical filtration $\mathbb{F}$. 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 and Shiryaev [32] Chapter I.4). Define

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

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 admits a representation $\lambda(x, dt) = \lambda_0(t,x)\gamma(dt)$ for some locally bounded, measurable function $\lambda_0 : [0,1] \times \mathbb{R} \to \mathbb{R}$ and finite non-negative measure γ on $[0,1]$. We then denote

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

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

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

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

$$\mathbf{D}_\infty(\mu) \ := \inf_{(H,\lambda)\in\mathcal{D}_\infty(\mu)} \mu(\lambda). \tag{3.3}$$

4 Main results

We will first provide, in Section 4.1, an approximation construction of $\mathbb{P} \in \mathcal{M}_\infty(\mu)$ for a given family of $\mu = (\mu_t)_{0\le t\le 1}$, which leads to a local Lévy process characterization. Under further conditions, we show that the local Lévy process solves the continuous-time MT problem (3.2) and we give also the optimal solution to the dual problem (3.3) together with a strong duality in Section 4.2. The corresponding proofs are postponed to Section 6. Finally, as an application in finance, we compute the optimal arbitrage-free lower and upper bounds of a variance swap in Section 4.3.

4.1 Construction of martingales with given marginals

4.1.1 On the left-monotone martingale transport

For every $t \in [0,1]$, let us denote by $F(t,\cdot)$ the cumulative distribution function of the probability measure μ_t on $\mathbb{R}$. We first make the following assumption on the marginals μ.

Assumption 4.1. (i) *The marginal distributions $\mu = (\mu_t)_{t\in[0,1]}$ are non-decreasing in convex ordering and have finite first order.*
(ii) *The distribution function $F \in C_b^4(E)$ with $E := \{(t,x) \ : t \in [0,1], x \in (l(t), r(t))\}$, and its density function $f(t,x) := \partial_x F(t,x)$ satisfies $\inf_{x\in[-K,K]\cap(l(t),r(t))} f(t,x) > 0$ for all $K > 0$ large enough.*

Notice that under Assumption 4.1, the function l (resp. r): $[0,1] \to \mathbb{R}$ defined before Assumption 3.1 is clearly continuous and non-increasing (resp. non-decreasing) on $[0,1]$ whenever it is finite. Moreover, the continuity of the density function f implies that $t \mapsto \mu_t$ is continuous, and hence $\mathcal{M}_\infty(\mu)$ is nonempty under the above Assumption 4.1.

Before introducing our continuous-time martingale transference plan, we consider the discrete case. For $t \in [0,1)$ and $\varepsilon \in (0, 1-t)$, denote

$$\delta^\varepsilon F(t,x) := F(t+\varepsilon, x) - F(t,x) \quad \text{and} \quad \delta^\varepsilon f(t,x) := f(t+\varepsilon,x) - f(t,x).$$

Our Assumption 4.2 below states that $x \mapsto \delta^\varepsilon F(t,x)$ has at most countably-many local maximizers and the collection of all local maximizers admits no accumulation points in $[-\infty,\infty)$. Then by considering the two marginals μ_t and $\mu_{t+\varepsilon}$, one obtains a unique left-monotone martingale measure following the construction recalled in Section 2. Suppose that the construction gives a sequence of points $\big(x_k^\varepsilon(t), m_{k+1}^\varepsilon(t)\big)_{k\ge 0}$ as well as the functions $T_u^\varepsilon(t,\cdot)$ and $T_d^\varepsilon(t,\cdot)$ following the procedure below equation (2.3). Similarly, denote $D^\varepsilon(t) :=$

$\cup_{k\geq 1}\big(x^\varepsilon_{k-1}(t), m^\varepsilon_k(t)\big]$ and $A^\varepsilon_k(t) := \big(x^\varepsilon_0(t), m^\varepsilon_k(t)\big)\setminus\big\{\cup_{i<k}\big(T^\varepsilon_d(t, x^\varepsilon_i(t)), x^\varepsilon_i(t)\big]\big\}$. In particular, we notice that for every $x \in \big[m^\varepsilon_k(t), x^\varepsilon_k(t)\big)$, $T^\varepsilon_d(t,x) \in A^\varepsilon_k(t)$ is uniquely determined by

$$\int_{-\infty}^{x}\big[F^{-1}\big(t+\varepsilon,\ F(t,\xi)\big)-\xi\big]f(t,\xi)d\xi+\int_{\left(-\infty,T^\varepsilon_d(t,x)\right]\cap A^\varepsilon_k(t)}\big[g^\varepsilon_t(x,\xi)-\xi\big]\delta^\varepsilon f(t,\xi)d\xi = 0, \quad (4.1)$$

and $T^\varepsilon_u(t,x) := g^\varepsilon_t\big(x, T^\varepsilon_d(t,x)\big)$, where F^{-1} denotes the inverse function of $x \mapsto F(t,x)$ and $g^\varepsilon_t(x,y) := F^{-1}\big(t+\varepsilon, F(t,x) + \delta^\varepsilon F(t,y)\big)$. Similarly, $(T^\varepsilon_d)^{-1}$ denotes the right-continuous version of the inverse function of $x \mapsto T^\varepsilon_d(t,x)$. We also introduce the jump size function

$$J^\varepsilon_u(t,x) := T^\varepsilon_u(t,x) - x, \qquad J^\varepsilon_d(t,x) := x - T^\varepsilon_d(t,x),$$

and

$$y^\varepsilon_k(t) := \big(T^\varepsilon_d\big)^{-1}\big(t, x^\varepsilon_{k-1}(t)-\big), \qquad z^\varepsilon_k(t) := T^\varepsilon_d\big(t, x^\varepsilon_k(t)-\big).$$

Notice that we have $m^\varepsilon_k(t) < x^\varepsilon_k(t)$, and $T^\varepsilon_d(t,\cdot)$ and $T^\varepsilon_u(t,\cdot)$ are both right-continuous, and continuous except at the points $\big(x^\varepsilon_k(t)\big)_{k\geq 1}$ and $\big(y^\varepsilon_k(t)\big)_{k\geq 1}$.

Our main result below shows a convergence result to a continuous-time limit which has a similar structure. We suppose that $x \mapsto \partial_t F(t,x)$ has at most countably-many local maximizers with no accumulation points in $[-\infty,\infty)$. Let $x_0(t) := l(t)$ and $m_1(t)$ be the first local maximizer, we then define a sequence of points $\big(m_k(t), x_k(t), y_k(t), z_k(t)\big)_{k\geq 1}$ by the following procedure:

$$m_k(t) \;:=\; \inf\big\{m > x_{k-1}(t)\ : m \text{ is a local maximizer of } x \mapsto \partial_t F(t,x)\big\}, \qquad (4.2)$$
$$x_k(t) := \inf\big\{x > m_k(t)\ : \partial_t F(t,x) \geq \partial_t F(t, T_d(t, x-))\big\},\ y_k(t) := T_d^{-1}\big(t, x_{k-1}(t)-\big), (4.3)$$

and $z_k(t) := T_d\big(t, x_k(t)-\big)$, where for every $x \in [m_k(t), x_k(t))$, $T_d(t,x)$ is the unique solution (see Lemma 4.3 below) of

$$\int_{T_d(t,x)}^{x}(x-\xi)\partial_t f(t,\xi)d\xi = 0 \text{ in } A_k(t) := \big(x_0(t), m_k(t)\big] \setminus \big\{\cup_{i<k}\big(T_d(t,x_i(t)), x_i(t)\big]\ \big\}. \quad (4.4)$$

Similarly, we denote

$$D(t) \;:=\; \cup_{k\geq 1}\big(x_{k-1}(t), m_k(t)\big], \qquad D \;:=\; \big\{(t,x) : t\in[0,1], x \in D(t)\big\},$$
$$\text{and}\quad D^{c,\circ}(t) \;:=\; \mathrm{int}\big(D(t)^c\big) \setminus \{y_k(t), k\geq 1\}.$$

We also define $j_d(t,x)$ and $j_u(t,x)$, for every $t \in [0,1]$ and $x \in D^c(t)$, by

$$j_d(t,x) \;:=\; x - T_d(t,x) \quad\text{and}\quad j_u(t,x) \;:=\; \frac{\partial_t F\big(t, T_d(t,x)\big) - \partial_t F(t,x)}{f(t,x)}. \qquad (4.5)$$

In particular, $j_u(t,\cdot)$ and $j_d(t,\cdot)$ are both positive and both continuous on $D^{c,\circ}(t)$.

In order to guarantee the convergence of T^ε_d and T^ε_u towards their natural limits (described by j_d and j_u), we need an additional assumption on the marginal distributions.

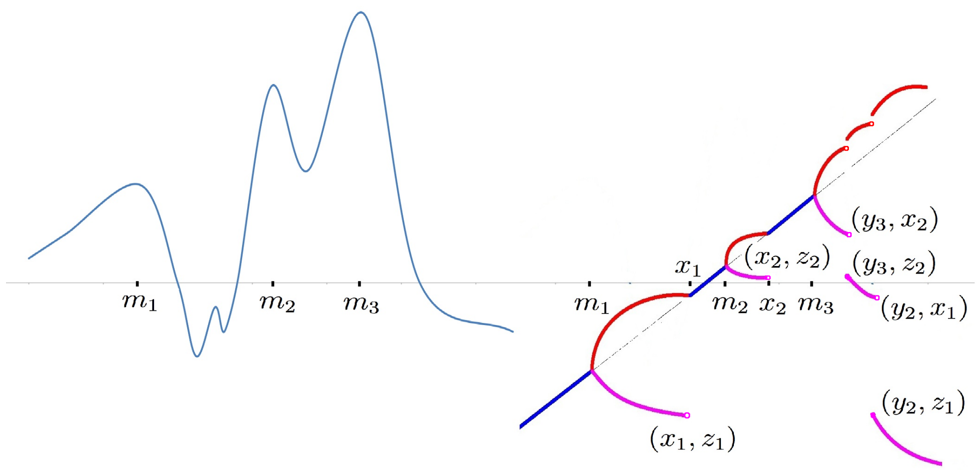

Figure 1: *An illustration of* (m_k, x_k, y_k, z_k) *for some fixed* $t \in [0,1]$. *The figure on the left is the function* $x \mapsto \partial_t F(t,x)$. *The figure on the right is the functions* $x \mapsto T_d(t,x)$ *and* $x \mapsto x + j_u(t,x)$, *which are constructed to be right-continuous. We notice that there is a local maximizer of* $\partial_t F(t,x)$ *between* m_1 *and* m_2, *which is not involved in the construction of* j_d *and* j_u.

Assumption 4.2. (i) *For every* $t \in [0,1)$ *and* $\varepsilon \in (0, 1-t)$, *the maps* $x \mapsto \delta^\varepsilon F(t,x)$ *and* $x \mapsto \partial_t F(t,x)$ *have at most countably-many local maximizers, with no accumulation points in* $[-\infty, \infty)$.
(ii) *The functions* $\big(m_k^\varepsilon, x_k^\varepsilon, y_k^\varepsilon, z_k^\varepsilon\big)_{k\geq 1}$ *and* $\big(m_k, x_k, y_k, z_k\big)_{k \geq 1}$ *are all uniformly continuous in* t *with the same modulus* ρ_0 *and* $\big(m_k^\varepsilon, x_k^\varepsilon, y_k^\varepsilon, z_k^\varepsilon\big) \to \big(m_k, x_k, y_k, z_k\big)$ *pointwise (and hence uniformly) as* $\varepsilon \to 0$ *for all* $k \geq 1$.

Lemma 4.3. *Let Assumptions 4.1 and 4.2 (i) hold true. Then*
(i) *for all* $k \geq 1$, $t \in [0,1]$, $x_k(t) > m_k(t)$ *is well-defined by* (4.3), *and for all* $x \in (m_k(t), x_k(t))$, *the equation* (4.4) *has a unique solution* $T_d(t,x)$ *in* $A_k(t)$.
(ii) *on every interval* $(m_k(t), x_k(t))$, *the map* $x \mapsto T_d(t,x)$ *is strictly decreasing.*

4.1.2 Construction of martingales with given marginals

In this subsection, we provide a first part of the main results of the paper, which gives an approximation construction of continuous-time martingales satisfying full marginals constraints, as well as a local Lévy process characterization.

We recall that $\Omega := D([0,1], \mathbb{R})$ 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 a $(n+1)$-marginals MT problem as recalled in Section

2.2, which consists in maximizing

$$\mathbb{E}\Big[\sum_{0\le k\le n-1} c(X_k^n, X_{k+1}^n)\Big]$$

among all discrete-time martingales satisfying the marginal distribution constraints. To simplify the notation, we suppose without loss of generality that the partition is uniform and $|\pi_n| = \varepsilon$ with $\varepsilon = \frac{1}{n}$. It is clear that under Assumptions 3.1, 4.1 and 4.2, the optimal martingale measure, denoted by $\mathbb{P}^{*,n}$, is provided by the left-monotone transference plan recalled in Section 4.1.1. More precisely, let $\Omega^{*,n} := \mathbb{R}^{n+1}$ be the canonical space of discrete-time process, $X^{*,n} = (X_k^{*,n})_{0\le k\le n}$ be the canonical process, under $\mathbb{P}^{*,n}$, $X^{*,n}$ is a discrete-time martingale and at the same time a Markov chain, characterized by $T_u^\varepsilon(t_k^n,\cdot)$ and $T_d^\varepsilon(t_k^n,\cdot)$ in (4.1). By abuse of notation, we extend the Markov chain $X^{*,n}$ to a continuous-time càdlàg process, denoted by $X^{*,n} = (X_t^{*,n})_{0\le t\le 1}$, by setting $X_t^{*,n} := X_k^{*,n}$ whenever $t \in [t_k^n, t_{k+1}^n)$. Let $\mathbb{P}^n := \mathbb{P}^{*,n} \circ (X^{*,n})^{-1}$ denote the probability measure on Ω, induced by $X^{*,n}$ under $\mathbb{P}^{*,n}$.

Our first result is the convergence of $(\mathbb{P}^n)_{n\ge 1}$ as $|\pi_n| \to 0$.

Proposition 4.4. *Let Assumptions 4.1 and 4.2 (i) hold true. Then the sequence $\big(\mathbb{P}^n\big)_{n\ge1}$ is tight (w.r.t. the Skorokhod topology on Ω); moreover, every limit probability measure $\mathbb{P}^0$ satisfies $\mathbb{P}^0 \in \mathcal{M}_\infty(\mu)$, i.e. the canonical process X is a martingale fitted to the marginal distributions $\mu = (\mu_t)_{0\le t\le 1}$ under $\mathbb{P}^0$.*

Remark 4.5. (i) *The result in Proposition 4.4 provides an approximation construction of martingales fitted to infinitely-many marginals, or equivalently the martingales associated with some peacock process.*
(ii) *Jakubowski [33] introduced the so-called S-topology on the canonical space Ω, which is coarser than the classical Skorokhod topology, and under which the associated tightness of probability measures on Ω is easier to be verified. Using the technique of S-topology, we can also obtain a martingale peacock by convergence without the technical condition in Assumption 4.1* (ii). *However, to characterize the limit martingale as in Theorem 4.11 below, we need to use the standard localization technique as in Jacod and Shiryaev [32], which is our main reason to keep on using the Skorokhod topology on Ω. We shall investigate the use of S-topology in martingale transport in our future works.*

We next provide a point of view from the forward Kolmogorov-Fokker-Planck (KFP) equation. Recall that $D^{c,\circ}(t)$ is defined below (4.4), and denote $D^\circ(t) := T_d(t, D^{c,\circ}(t))$.

Lemma 4.6. *Under Assumptions 4.1 and 4.2, the density function $f(t,x)$ satisfies*

$$\partial_t f(t,x) = -\mathbf{1}_{\{x\in D(t)\}} \frac{j_u f}{j_d(1-\partial_x j_d)}\big(t, T_d^{-1}(t,x)\big) - \mathbf{1}_{\{x\in D^c(t)\}}\Big(\frac{j_u f}{j_d} - \partial_x(j_u f)\Big)(t,x), \quad (4.6)$$

for all $t\in[0,1)$ and $x \in D^\circ(t)\cup D^{c,\circ}(t)$.

The first order PDE (4.6)can be viewed as a KFP forward equation of the following SDE:

$$dX_t \;=\; -\mathbf{1}_{\{X_{t-}\in D^c(t)\}} j_d(t,X_{t-})(dN_t - \nu_t dt), \quad \nu_t \;:=\; \tfrac{j_u}{j_d}(t,X_{t-})\mathbf{1}_{\{X_{t-}\in D^c(t)\}}, \quad (4.7)$$

where $(N_t)_{0\le t\le 1}$ is a jump process with unit jump size and with predictable compensated process $(\nu_t)_{0\le t\le 1}$. Notice that this pure jump process is in the spirit of the local Lévy models introduced in Carr, Geman, Madan and Yor [9]. However, the intensity process $(\nu_t)_{0\le t\le 1}$ in our context is state-dependent.

Proposition 4.7. *Let Assumptions 4.1 and 4.2 hold true. Suppose that the SDE* (4.7) *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*

$$\mathbb{E}\Big[\int_t^{t+\varepsilon_0} j_u(s,\widehat{X}_s)\mathbf{1}_{\widehat{X}_s\in D^c(s)}ds\Big] \;\; < \;\; \infty. \tag{4.8}$$

(i) *Then, the density function $f^{\widehat{X}}$ of $\widehat{X}$ defined in* (4.7) *satisfies the KFP equation* (4.6)*.*
(ii) *Consequently, if uniqueness holds for the KFP equation* (4.6)*, the pure jump process $\widehat{X}$ is a martingale with marginals $\widehat{X}_t \sim \mu_t$ for all $t \in [0,1]$.*

4.2 Optimality of the local Lévy process

For the optimality of the local Lévy process (4.7), we formulate another more restrictive assumption on the marginal distributions. The main technical reason for this condition is that j_u and j_d have generally discontinuous points under the multiple local maximizer conditions in Assumption 4.2.

Assumption 4.8. (i) *There is some constant $\varepsilon_0 > 0$ such that, for all $t \in [0,1]$ and $0 < \varepsilon \le \varepsilon_0 \wedge (1-t)$, $x \mapsto \delta^\varepsilon F(t,x)$ (resp. $x \mapsto \partial_t F(t,x)$) has only one local maximizer (which is hence the global maximizer) on $\big(l(t+\varepsilon), r(t+\varepsilon)\big)$ (resp. $\big(l(t), r(t)\big)$), denoted by $m^\varepsilon(t)$ (resp. $m(t)$).*
(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 $\partial_{tx} f(t, m(t)) < 0$.*

Remark 4.9. Part (i) of the above assumption is in fact very similar to the Dispersion Assumption 2.1 in Hobson and Klimmek [29], i.e. it implies that $\{x\ : \delta^\varepsilon f(t,x) \ge 0\}$ lies in an interval on $\mathbb{R}$. An example which satisfies both Assumptions 4.1 and 4.8 is the marginals $(\mu_t)_{t\in[\delta,1+\delta]}$ of the Brownian motion for some $\delta > 0$, where $f(t,x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}$, $m^\varepsilon(t) = -\sqrt{\frac{t(t+\varepsilon)}{\varepsilon}\log(1+\varepsilon/t)}$ and $m(t) = -\sqrt{t}$. See also Section 5 for more discussions.

4.2.1 On the optimal dual component

In preparation, let us introduce the candidates of the optimal dual components for the dual problem (3.3), under Assumption 4.8. For ease of presentation, let us suppose that $l(t) = -\infty$ and $r(t) = \infty$. First, let us consider the discrete-time two marginals MT problem associated with initial distribution μ_t, terminal distribution $\mu_{t+\varepsilon}$ and reward function $c : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$. As recalled in Section 2, under Assumption 4.8, the optimal superhedging strategy $(\varphi^\varepsilon, \psi^\varepsilon, h^\varepsilon)$ is explicitly given as follows:

$$
\begin{aligned}
\partial_x h^\varepsilon(t,x) &:= \frac{c_x(x, T_u^\varepsilon(t,x)) - c_x(x, T_d^\varepsilon(t,x))}{T_u^\varepsilon(t,x) - T_d^\varepsilon(t,x)}, && x \geq m^\varepsilon(t),\\
h^\varepsilon(t,x) &:= h^\varepsilon\big(t, (T_d^\varepsilon)^{-1}(t,x)\big) - c_y\big((T_d^\varepsilon)^{-1}(t,x), x\big), && x < m^\varepsilon(t).
\end{aligned}
$$

Denote $(T^\varepsilon)^{-1}(t,x) := (T_u^\varepsilon)^{-1}(t,x)\mathbf{1}_{x\geq m^\varepsilon(t)} + (T_d^\varepsilon)^{-1}(t,x)\mathbf{1}_{x<m^\varepsilon(t)}$, ψ^ε is defined 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

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

Clearly, h^ε and ψ^ε are unique up to a constant. More importantly, h^ε and ψ^ε can be chosen continuous on $[0,1)\times\mathbb{R}$ so that (2.7) holds true, since T_u^ε and T_d^ε are both continuous under Assumption 4.8.

We shall see later that Assumption 3.1 on the reward function c implies that the continuous-time limit of the optimal dual components is given as follows. The function $h^* : [0,1]\times\mathbb{R}$ is defined, up to a constant, by

$$
\partial_x h^*(t,x) \;:=\; \frac{c_x(x,x) - c_x(x, T_d(t,x))}{j_d(t,x)}, \qquad \text{when } x \geq m(t), \tag{4.9}
$$

$$
h^*(t,x) \;:=\; h^*\big(t, T_d^{-1}(t,x)\big) - c_y\big(T_d^{-1}(t,x), x\big), \qquad \text{when } x < m(t). \tag{4.10}
$$

Finally, $\psi^* : [0,1]\times\mathbb{R}\to\mathbb{R}$ is defined, up to a constant, by

$$
\partial_x \psi^*(t,x) \;:=\; -\,h^*(t,x), \qquad (t,x)\in[0,1]\times\mathbb{R}.
$$

Lemma 4.10. *Let Assumptions 4.1 and 4.8 hold true. Then $j_d(t,x)\mathbf{1}_{x>m(t)}$, $j_u(t,x)\mathbf{1}_{x>m(t)}$ and $\frac{j_u}{j_d}(t,x)\mathbf{1}_{x>m(t)}$ are all locally Lipschitz in x. Consequently, suppose in addition that Assumption 3.1 holds true, then $\psi^*\in C^{1,1}([0,1]\times\mathbb{R})$, i.e. $\partial_t\psi^*$ and $\partial_x\psi^*$ are both continuous.*

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 $(t,x)\in(0,1)\times\mathbb{R}$,

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

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

4.2.2 Optimality of the local Lévy process

Under the one maximizer condition of Assumption 4.8, we now strengthen the result of Propositions 4.4 and 4.7. We recall that the family $(\mathbb{P}^n)_{n\geq 1}$ is induced by the discrete-time left-monotone martingale transferences (see the beginning of Section 4.1.2).

Theorem 4.11. *Suppose that Assumptions 4.1 and 4.8 hold true, then $\mathbb{P}^n \to \mathbb{P}^0$, where $\mathbb{P}^0$ is the unique weak solution of the SDE*

$$X_t \;=\; X_0 - \int_0^t \mathbf{1}_{\{X_{s-}>m(s)\}} j_d(s, X_{s-})(dN_s - \nu_s ds), \quad \nu_s \;:=\; \frac{j_u}{j_d}(s, X_{s-})\mathbf{1}_{X_{s-}>m(s)}, \tag{4.11}$$

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

Our next result is the optimality of the above local Lévy process (4.11), as well as that of the dual component introduced in and below (4.9). Similar to [21] and [29], we obtain in addition a strong duality for the MT problem (3.2) and (3.3). Let H^* be the $\mathbb{F}$-predictable process on Ω defined by $H^*_t := h^*(t, X_{t-})$.

Theorem 4.12. *Let Assumptions 3.1, 4.1 and 4.8 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 (3.2) is solved by the local Lévy process (4.11). Moreover, $(H^*, \lambda^*) \in \mathcal{D}_\infty(\mu)$ and we have the duality*

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

where the optimal value is given by

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

Remark 4.13. The proofs of Theorems 4.11 and 4.12 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 4.10. This is also the main reason for which we restrict to the one maximizer case under Assumption 4.8. See also Remark 6.5 for more discussions.

Remark 4.14. *By symmetry, we can consider the right monotone martingale transference plan as discussed in Remark 3.14 of [21]. This leads to a upward pure jump process with explicit characterizations, assuming that $x \mapsto \partial_t F(t,x)$ has only one local minimizer $\tilde{m}(t)$. More precisely, we define*

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

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:

$$dX_t = \mathbf{1}_{X_{t^-}<\tilde{m}(t)} \tilde{j}_u(t, X_{t^-})(d\tilde{N}_t - \tilde{\nu}_t dt) \qquad \tilde{\nu}_t \ := \ \frac{\tilde{j}_d}{\tilde{j}_u}(t, X_{t^-})\mathbf{1}_{X_{t^-}<\tilde{m}(t)}, \tag{4.12}$$

where $(\tilde{N}_t)_{0\le t\le 1}$ is a upward jump process with unit jump size and predictable compensated process $(\tilde{\nu}_t)_{0\le t\le 1}$. Moreover, under Assumption 3.1 together with regularity conditions on the density function f, this martingale solves a corresponding minimization MT problem with optimal value

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

4.3 Application: Robust hedging of variance swap

As an application, let us finally consider the reward function $c_0(x, y) := (\ln x - \ln y)^2$, corresponding to the payoff of a so-called "variance swap". More precisely, the payoff of "variance swap" is given by $\sum_{k=1}^{n-1} \ln^2 \frac{X_{t_{k+1}}}{X_k}$ in the discrete-time case, and by

$$\int_0^1 \frac{1}{X_t^2} d[X]_t^{\mathrm{c}} + \sum_{0<t\le 1} \ln^2 \frac{X_t}{X_{t-}}$$

in the continuous-time case, following the convergence result in Lemma 3.3.

We can easily verify that c_0 satisfies Assumption 3.1. Therefore, given a continuous-time family of marginals $(\mu_t)_{0\le t\le 1}$ which are all supported on $(0, \infty)$ and satisfy Assumptions 4.1 and 4.8, we can then construct a left-monotone (resp. right-monotone) martingale with characteristics m, j_u and j_d (resp. $\tilde{m}$, $\tilde{j}_u$ and $\tilde{j}_d$). In addition, suppose that the constructed optimal static strategy λ^* satisfies the integrability conditions in Theorem 4.12, we then get the following result:

Proposition 4.15. *Under the above conditions and with the same notations, the optimal upper bound of the variance swap is given by*

$$\int_0^1 dt \int_{m(t)}^{\infty} dx \ \frac{j_u(t,x)}{j_d(t,x)} \ \ln^2 \frac{x}{x - j_d(t,x)} \ f(t,x),$$

and the optimal lower bound is given by

$$\int_0^1 dt \int_0^{\tilde{m}(t)} dx \ \frac{\tilde{j}_d(t,x)}{\tilde{j}_u(t,x)} \ \ln^2 \frac{x + \tilde{j}_u(t,x)}{x} \ f(t,x),$$

where the optimal martingale measures are given by the local Lévy processes (4.11) *and* (4.12)*.*

5 Examples of extremal peacock processes

With the introduction of peacock (or PCOC "Processus Croissant pour l'Ordre Convexe" in French) by Hirsch, Profeta, Roynette and Yor [24], the construction of martingales with given marginal distributions becomes an interesting subject. When the marginals are given by those of the Brownian motion, such a martingale is called a fake Brownian motion. The above two jump processes provide two new constructions of martingale peacocks and in particular two discontinuous fake Brownian motions if we take for $f(t,x)$ the density of a Brownian motion. We also refer to Albin [1], Fan, Hamza and Klebaner [14], Hamza and Klebaner [19], Hirsch et al. [23], Hobson [27], Oleszkiewicz [40], Pagès [42] etc. for other solutions and related results. Moreover, our two fake Brownian motions are remarkable since they are optimal for a large class of reward functions. Let us provide here some explicit characterizations of the first discontinuous fake Brownian motion as well as that of a self-similar martingale induced by Theorem 4.11.

5.1 A remarkable fake Brownian motion

Let $\mu_t := \mathcal{N}(0,t)$ with $t \in [\delta, 1]$ for some $\delta > 0$, for which Assumptions 4.1 and 4.8 are satisfied. In particular, by direct computation, we have $m^\varepsilon(t) = -\sqrt{\frac{t(t+\varepsilon)}{\varepsilon}\log(1+\varepsilon/t)}$ and $m(t) = -\sqrt{t}$ for all $t \in [\delta, 1]$. In this case, it follows that $T_d(t,x)$ is defined by the equation:

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

By direct change of variables, this provides the scaled solution $T_d(t,x) := t^{1/2}\widehat{T}_d\big(t^{-1/2}x\big)$, where:

$$\widehat{T}_d(x) \leq -1 \quad \text{is defined for all } x \geq -1 \text{ by } \quad \int_{\widehat{T}_d(x)}^{x} (x-\xi)(\xi^2-1)e^{-\xi^2/2}d\xi = 0.$$

i.e.

$$e^{-\widehat{T}_d(x)^2/2}\Big(1 + \widehat{T}_d(x)^2 - x\widehat{T}_d(x)\Big) = e^{-x^2/2}.$$

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

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

We also plot the maps $\widehat{T}_d(x)$ and $\widehat{T}_u(x) := x + \widehat{j}_u(x)$ in Fig. 2.

5.2 A new construction of self-similar martingales

In Hirsch, Profeta, Roynette and Yor [23], the authors construct martingales M_t which enjoy the (inhomogeneous) Markov property and the Brownian scaling property:

$$\forall\, c > 0\,, (M_{c^2 t}, t \geq 0) \sim (cM_t, t \geq 0).$$

When the marginals of M admit a density, this property means that the density function $f(t,x)$ scales as $f\left(c^2 t, x\right) = \frac{1}{c} f\left(\frac{x}{c}, t\right)$, i.e. $f(t,x) = f(1, x/\sqrt{t})/\sqrt{t}$. A first methodology for

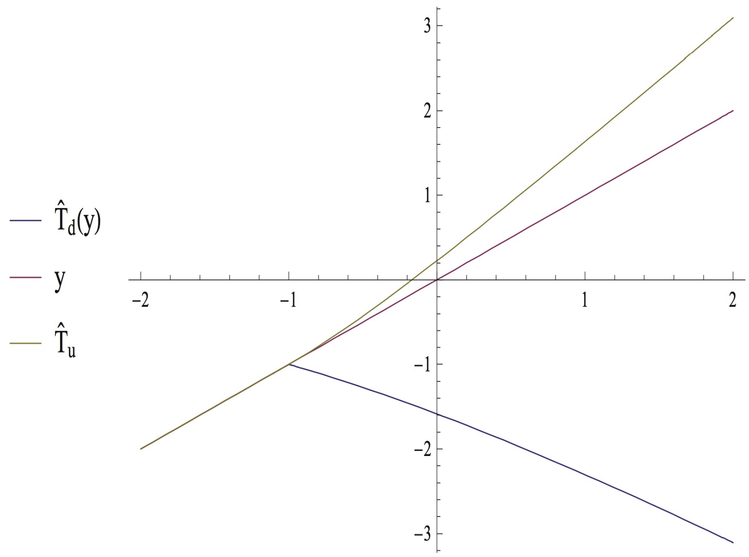

Figure 2: *Fake Brownian motion: Maps* $\hat{T}_d$ *and* $\hat{T}_u$

constructing such martingales, initiated in [38], uses the Azéma-Yor embedding algorithm under a condition on $\mu = \mathrm{Law}(M_1)$, which is equivalent to

$$x \mapsto \frac{x}{b_1(x)} \quad \text{is increasing on } \mathbb{R}_+,$$

with b_1 the Hardy-Littlewood barycenter function

$$b_1(x) \quad := \quad \frac{1}{\mu([x,\infty))} \int_{[x,\infty)} y\mu(dy).$$

Their second method uses randomization techniques, and allows to reach any centered law with finite moment of order 1.

Following our approximation approach, one can construct a self-similar martingale that can reach marginals without using randomization. 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 property of j_d and j_u, observed in the previous subsection, still applies; and $\widehat{T}_d$ as well as $\widehat{j}_u$ can be computed by

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

6 Proofs

In this section, we complete the proofs of the main results (Theorems 4.11 and 4.12) stated in Section 4. We first provide some asymptotic estimates for the left-monotone transference plan of Section 6.1 as a preparation. Then we complete the proofs for the convergence results in Section 6.2, and finally the proof of the optimality result in Section 6.3.

6.1 Asymptotic estimates of the left-monotone transference plan

We recall that the left-monotone transference plan is described by T_u^ε and T_d^ε, which are defined in and below (4.1). Moreover, $J_u^\varepsilon(t,x) := T_u^\varepsilon(t,x) - x$ denotes the upward jump size, $J_d^\varepsilon(t,x) := x - T_d^\varepsilon(t,x)$ the downward jump size and

$$q^\varepsilon(t,x) \;\;:=\;\; \frac{J_u^\varepsilon(t,x)}{J_u^\varepsilon(t,x) + J_d^\varepsilon(t,x)} \;\;=\;\; \frac{T_u^\varepsilon(t,x) - x}{T_u^\varepsilon(t,x) - T_d^\varepsilon(t,x)}$$

the probability of a downward jump.

Lemma 6.1. *Let Assumptions 4.1 and 4.2 (i) hold true. Then for every $K > 0$, there are some 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 (4.1)), 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 4.1, then it follows by the definition of T_u^ε below (4.1) 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 (4.1), we have

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

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. □

Proof of Lemma 4.3. To simplify the notation, let us omit the argument t and write (m_k, x_k, y_k, z_k) in place of $(m_k(t), x_k(t), y_k(t), z_k(t))$, etc. We shall prove the lemma by induction.

(i) Let $k \ge 1$, suppose that x_i and $T_d(t,x)$ is well defined for all $x \in [m_i, x_i)$ and every $i < k$. Denote

$$G_k(t,x,y) := \int_{[m_k,x) \cup \{(y,m_k) \cap A_k\}} (x-\xi)\partial_t f(t,\xi) d\xi, \tag{6.1}$$

where we recall that $A_k := \big(\ell(t), m_k\big) \setminus \big\{ \cup_{i<k} \big(z_i, x_i\big] \big\}$ and $z_i := T_d(t, x_i-)$. Further, using the equality $\partial_t F(t, x_i) = \partial_t F(t, z_k)$, it is easy to show by induction that $\int_{[T_d(t,x_i),x_i)} (x_i - \xi)\partial_t f(t,\xi) d\xi = 0$ and hence

$$\int_{\cup_{i<k}(z_i,x_i]} (x-\xi)\partial_t f(t,\xi)d\xi \;\;=\;\; 0. \tag{6.2}$$

It follows then the equation (4.4) is equivalent to

$$G_k\big(t, x, T_d(t,x)\big) \;\;=\;\; 0. \tag{6.3}$$

(ii) Using again (6.2), we have

$$G_k(t,x,-\infty) \;\;=\;\; \textstyle\int_{[m_k,x)\cup\{(-\infty,m_k)\cap A_k\}} (x-\xi)\partial_t f(t,\xi)d\xi \;\;=\;\; \int_{-\infty}^x (x-\xi)\partial_t f(t,\xi)d\xi.$$

Therefore, $G_k(t,x,-\infty) \geq 0$ using Jensen's inequality and the condition that $(\mu_t)_{t\in[0,1]}$ is non-decreasing in the convex ordering in Assumption 4.1. In fact, since $\partial_t f(t,x) > 0$ for $x \in A_k \subset (-\infty, m_k]$, the inequality is strict, i.e.

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

Next, since $\partial_t f(t,\xi) > 0$ for every $\xi \in A_k$ by Assumption 4.2 (i), it follows that $y \mapsto G_k(t,x,y)$ is strictly decreasing on A_k.

Further, denote $c_k := \inf\{y > m_k \;:\; \partial_t F(t,y) \geq \partial_t F(t,m_k)\}$. Then it is clear that $G_k(t,m_k,m_k) = 0$ and $x \mapsto G_k(t,x,m_k)$ is decreasing on interval (m_k, c_k) since

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

In summary, for every $x \in (m_k, c_k)$, we have $G_k(t,x,m_k) < 0$, $G_k(t,x,-\infty) > 0$ and $y \mapsto G_k(t,x,y)$ is continuous, strictly decreasing on A_k. It follows that the equation (6.3) has a unique solution $T_d(t,x)$ and it takes values in A_k, which implies that the equation (4.4) has a unique solution in A_k.

(iii) Since m_k is a local maximizer of $x \mapsto \partial_t F(t,x)$, then for $x > m_k$ close enough to m_k, we have $T_d(t,x) \in (x_{k-1}, m_k)$, i.e.

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

It follows that

$$\int_{T_d(t,x)}^{m_k} (x-m_k)\partial_t f(t,\xi)d\xi \;+\; \int_{m_k}^x (x-m_k)\partial_t f(t,\xi)d\xi \;\;<\;\; 0,$$

which implies that $\partial_t F(t,x) < \partial_t F(t, T_d(t,x))$ for $x > m_k$ close enough to m_k. Therefore, x_k is well defined in (4.3) and satisfies that $x_k > m_k$.

(iv) Differentiating both sides of equation (4.4) w.r.t. x for $x \notin D^{c,\circ}(t)$, 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 $x \in (m_k(t), x_k(t))$,

$$\partial_x T_d(t,x) \;=\; \frac{\partial_t F(t,x) - \partial_t F(t, T_d(t,x))}{\big(x - T_d(t,x)\big)\; \partial_t f\big(t, T_d(t,x)\big)} \;<\; 0, \tag{6.4}$$

and hence $x \mapsto T_d(t,x)$ is strictly decreasing in x on interval $(m_k(t), x_k(t))$. $\square$

Under Assumption 4.2, we provide an asymptotic estimate for J_u^ε and J_d^ε. First, Assumption 4.2 states that $(m_k^\varepsilon(t), x_k^\varepsilon(t), y_k^\varepsilon(t), z_k^\varepsilon(t))$ converges uniformly to $(m_k(t), x_k(t), y_k(t), z_k(t))$ for $t \in [0,1]$. Let us denote the uniform convergence rate of $(m_k^\varepsilon, x_k^\varepsilon, y_k^\varepsilon, z_k^\varepsilon)$ to (m_k, x_k, y_k, z_k) by $\rho_k(\varepsilon)$.

Lemma 6.2. *Under Assumptions 4.1 and 4.2, we have for all $t_0 \in [0,1)$ and $x_0 \in D^{c,\circ}(t_0)$:*

$$J_u^\varepsilon(t,x) = \varepsilon j_u^\varepsilon(t,x) + \varepsilon^2 e_u^\varepsilon(t,x), \quad \textit{and} \quad J_d^\varepsilon(t,x) = j_d(t,x) + \big(\varepsilon \vee \rho_k(\varepsilon)\big)\; e_d^\varepsilon(t,x), \tag{6.5}$$

where $(\varepsilon, t, x) \mapsto e_u^\varepsilon(t,x), e_d^\varepsilon(t,x)$ are locally bounded near $(t_0, x_0, 0)$, j_d is defined in (4.5), and

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

Proof. For $\delta > 0$ small enough and $0 \le \varepsilon \le \delta$, the function $g_t^\varepsilon(x,\xi)$ is uniformly bounded (let us say bounded by $K > 0$) for $x \in B_\delta(t_0, x_0, 0)$ and $\xi \in \mathbb{R}$. Let $||_\infty$ be the $\mathbb{L}^\infty-$norm in the corresponding space of variables. For the density function f, we denote $m_f := \inf_{t \in [t_0-\delta, t_0+\delta], y \in [-K,K] \cap (l(t), r(t))} f(t,y)$ which is strictly positive by Assumption 4.1.

Notice that $T_d(t,x)$ is clearly continuous on $B_\delta(t_0, x_0)$ for $\delta > 0$ small enough, since $x_0 \in D^{c,\circ}(t_0)$. We first claim that the family $(|T_d^\varepsilon(t,x)|)_{(t,x,\varepsilon) \in B_\delta(t_0,x_0,0)}$ is uniformly bounded by some constant $K > 0$, such that $K > |T_d(t,x)| + 1$ for every $(t,x) \in B_\delta(t_0, x_0)$.

(i) By the definition of T_u^ε below (4.1), we write:

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

By direct expansion, we see that the first equality in (6.5) holds true with

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

(ii) Let us now consider the second equality in (6.5). First,

$$\begin{aligned}
\int_{-\infty}^{x} [F^{-1}(t+\varepsilon, F(t,\xi)) - \xi] f(t,\xi) d\xi &= \int_{-\infty}^{x} \xi \; \delta^\varepsilon f(t,\xi) d\xi + \int_{x}^{F^{-1}(t+\varepsilon, F(t,x))} \xi f(t+\varepsilon, \xi) d\xi \\
&= \int_{-\infty}^{x} \xi \; \delta^\varepsilon f(t,\xi) d\xi - \delta^\varepsilon F(t,x)\big(x + C_1(t,x)\varepsilon\big) \\
&= \int_{-\infty}^{x} (\xi - x) \; \delta^\varepsilon f(t,\xi) d\xi \;+\; C_2(t,x)\varepsilon^2,
\end{aligned}$$

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

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

$$\int_{(-\infty, T^\varepsilon_d(t,x)]\cap A^\varepsilon_k} \big(g^\varepsilon_t(x,\xi) - \xi\big)\delta f(t,\xi)d\xi \;=\; \int_{(-\infty, T^\varepsilon_d(t,x)]\cap A^\varepsilon_k} (x-\xi)\ \delta^\varepsilon f(t,\xi)d\xi + C_4(t,x)\varepsilon^2,$$

where $|C_4(t,x)| \le 2\frac{|\partial_t F|_\infty}{m_f}|\partial_t F|_\infty$. Combining the above estimates with (4.1), it follows that

$$\int_{(T^\varepsilon_d(t,x),x]\cup\{\cup_{i<k}[z^\varepsilon_i, x^\varepsilon_i)\}} (x-\xi)\frac{1}{\varepsilon}\delta^\varepsilon f(t,\xi)\ d\xi \;=\; \big(C_2(t,x)\vee C_4(t,x)\big)\varepsilon.$$

It follows then

$$\int_{(T^\varepsilon_d(t,x),x]\cup\{\cup_{i<k}[z_i, x_i)\}} (x-\xi)\partial_t f(t,\xi)\ d\xi \;=\; C_5(t,x)\big(\varepsilon\vee\rho_k(\varepsilon)\big),$$

where $|C_5(t,x)| \le (x+K)(|\partial_t f|_\infty + |\partial^2_{tt} f|_\infty)$. Recall that $\partial_t F(t,x_i) = \partial_t F(t,z_i)$ and $\int_{z_i}^{x_i}(x_i-\xi)\partial_t f(t,\xi)d\xi = 0$, we get

$$\int_{(m_k,x]\cup\{(T^\varepsilon_d(t,x),m_k]\cap A_k\}} (x-\xi)\partial_t f(t,\xi)\ d\xi \;=\; C_5(t,x)\big(\varepsilon\vee\rho_k(\varepsilon)\big). \tag{6.6}$$

This implies the first estimation in (6.5) since $\partial_t f(t,x) > 0$ on A_k.
(iii) To complete the proof, it is enough the prove the claim that $(|T^\varepsilon_d(t,x)|)_{(t,x,\varepsilon)\in B_\delta(t_0,x_0,0)}$ is uniformly bounded. We notice that for $\varepsilon > 0$ small enough, there is some $T^\varepsilon_d(t,x)$ bounded by $K > |T_d(t,x)| + 1$, which solves equation (6.6). Then this $T^\varepsilon_d(t,x)$ also solves (4.1) since (6.6) is equivalent to (4.1) by the definition of $C_5(t,x)$. Finally, by the uniqueness of solution of T^ε in (4.1), it follows that the family $(|T^\varepsilon_d(t,x)|)_{(t,x,\varepsilon)\in B_\delta(t_0,x_0,0)}$ is indeed uniformly bounded by some constant K, and we hence conclude the proof. □

Our next result will be proved in the one local maximizer context of Assumption 4.8. Then, $D := \{x \le m(t)\}$, and $T_d : D^c \to D$ is defined by

$$\int_{T_d(t,x)}^x (x-\xi)\partial_t f(t,\xi)d\xi = 0, \quad \text{for all} \quad (t,x)\in D^c. \tag{6.7}$$

We also introduce the following subsets of $D^c := \{x > m(t)\}$:

$$E_\delta := \{(t,x)\in D^c : m(t) < x < m(t)+\delta\},\quad E_{\delta,K} := \{(t,x)\in D^c : m(t)+\delta \le x \le (m(t)+K)\wedge r(t)\}.$$

Remark 6.3. *Under the additional Assumption 4.8 and by the same arguments, the functions e^ε_u and e^ε_d in Lemma 6.2 are uniformly bounded on every $E_{\delta,K}$, for every $0<\delta<K$. Consequently, there is constant $C_{\delta,K}$ such that q^ε admits the asymptotic expansion:*

$$q^\varepsilon(t,x) \;=\; \varepsilon\frac{j_u(t,x)}{j_d(t,x)} \;+\; C_{\delta,K}\,\varepsilon\big(\varepsilon\vee\rho_0(\varepsilon)\big), \quad \textit{for } (t,x)\in E^+_{\delta,K}. \tag{6.8}$$

Proof of Lemma 4.10. (i) To see that $j_d \mathbf{1}_{D^c}$is locally Lipschitz in x on $[m(t), \infty)$, we shall verify that $\partial_x T_d \mathbf{1}_{D^c}$ is locally bounded. From (6.4), 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 4.8, we have

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

where $C_1(t,\xi)$ is uniformly bounded for $|\xi - m(t)| \le \delta$. Inserting the above expression into (6.7), it follows that

$$\int_{T_d(t,x)}^{x} \big(x-\xi\big)\big(\xi - m(t)\big) d\xi \quad = \quad C_2(t,x)\big(x - T_d(t,x)\big)^4,$$

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 4.8. By direct computation, it follows that

$$\big(x - T_d(t,x)\big)^2\big(x - m(t) + 2\big(m(t) - T_d(t,x)\big)\big) = C_2(t,x)\big(x - T_d(t,x)\big)^4,$$

which implies that

$$T_d(t,x) = m(t) - \frac{1}{2}\big(x - m(t)\big) + C_2(t,x)\big(x - T_d(t,x)\big)^2, \tag{6.9}$$

Using again the expression (6.4), we have

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

where C_3 is also uniformly bounded on E_δ. Finally, by the uniqueness of solution T_d of (6.7), we get

$$\partial_x T_d\big(t,x\big) \quad = \quad -\frac{1}{2} \ + \ C_4(t,x)\big(x - m(t)\big), \tag{6.10}$$

for some C_4 uniformly bounded on E_δ, implying that $T_d \mathbf{1}_{D^c}$ is locally Lipschitz in x.
(ii) By the expression of j_u in (4.5), i.e.

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

By (6.10), 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.
(iii) We now consider $\partial_t T_d(t,x)$. By direct computation,

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

which is clearly continuous in (t,x) on D^c, and hence uniformly bounded on $E_{\delta,K}$, for $K>\delta>0$. Using again (6.9), 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. In particular, we have

$$\partial_t T_d(t,m(t)+\delta) \;\to\; -\frac{3}{2}\frac{\partial^2_{tt}f(t,m(t))}{\partial^2_{tx}f(t,m(t))} \;\text{ uniformly for } t\in[0,1], \quad \text{as } \delta\searrow 0.$$

(iv) To prove $\psi^*\in C^{1,1}([0,1]\times\mathbb{R})$, we need to check that $\partial_x\psi^*=h^*$ and $\partial_t\psi^*$ are both continuous on $[0,1]\times\mathbb{R}$. First, By its definition below (4.9), h^* is clearly 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 can also easily check the continuity of h^* at the point $(t,m(t))$ by (4.9), since $T_d^{-1}(t,x)\to m(t)$ as $x\to m(t)$ and $c_y(x,x)=0$. Finally, by (4.9) with direct computation, we get

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

which is also locally bounded on D^c by (6.9). It follows then that $\partial_t\psi^*(t,x)=\int_0^x\partial_t h^*(t,\xi)d\xi$ is continuous in (t,x). □

Lemma 6.4. *Under Assumptions 3.1, 4.1 and 4.8, we have*

$$T_d^\varepsilon\mathbf{1}_{D_\varepsilon^c}\to T_d\mathbf{1}_{D^c},\; h^\varepsilon\to h^*,\; \partial_t\psi^\varepsilon\to\partial_t\psi^*,\; \textit{and } \psi^\varepsilon\to\psi^*,$$

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 4.8, the definition of $T_d^\varepsilon(t,x)$ in (4.1) is reduced to be

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

or equivalently

$$\int_{T_d^\varepsilon(t,x)}^x\xi\;\delta^\varepsilon f(t,\xi)d\xi \;+\; \int_x^{T_u^\varepsilon(t,x)}\xi f(t+\varepsilon,\xi)d\xi \;=\; 0, \tag{6.11}$$

with $T_u^\varepsilon(t,x):=g_t^\varepsilon\big(x,T_d^\varepsilon(t,x)\big)$. Differentiating (6.11), it follows that

$$\partial_t T_d^\varepsilon(t,x) \;:=\; -\frac{A^\varepsilon(t,x)}{\big(T_u^\varepsilon-T_d^\varepsilon\big)\delta^\varepsilon f\big(t,T_d^\varepsilon(\cdot)\big)}\,(t,x), \tag{6.12}$$

with

$$\begin{aligned}
A^\varepsilon(t,x) \;&:=\; \int_{T_d^\varepsilon(t,x)}^x\xi\partial_t\delta^\varepsilon f(t,\xi)d\xi \;+\; \int_x^{T_u^\varepsilon(t,x)}\xi\partial_t f(t+\varepsilon,\xi)d\xi\\
&\quad + T_u^\varepsilon(t,x)\Big(\partial_t F(t,x)-\partial_t F\big(t+\varepsilon,T_u^\varepsilon(t,x)\big)+\partial_t\delta^\varepsilon F\big(t,T_d^\varepsilon(t,x)\big)\Big),\\
&=\; -\big(T_u^\varepsilon(t,x)-x\big)\big(\partial_t\delta^\varepsilon F(t,x)-\partial_t\delta^\varepsilon F(t,T_d^\varepsilon(t,x)\big)\\
&\quad -\int_x^{T_u^\varepsilon(t,x)}\big(T_u^\varepsilon(t,x)-\xi\big)\partial_t f(t+\varepsilon,\xi)d\xi \;-\; \int_{T_d^\varepsilon(t,x)}^x(x-\xi)\partial_t\delta^\varepsilon f(t,\xi)d\xi
\end{aligned}$$

and

$$\partial_x T_d^\varepsilon(t,x) \quad := \quad - \frac{T_u^\varepsilon(t,x) - x}{\big(T_u^\varepsilon(t,x) - T_d^\varepsilon(t,x)\big)\ \delta^\varepsilon f\big(t, T_d^\varepsilon(t,x)\big)}\ f(t,x), \tag{6.13}$$

where the last term is exactly the same as that induced by ODE (2.4).

(ii) Taking the limit $\varepsilon \to 0$, it follows by direct computation and the convergence $T_d^\varepsilon(t,x) \to T_d(t,x)$ in Lemma 6.2 that $\partial_x T_d^\varepsilon(t,x) \to \partial_x T_d(t,x)$ and $\partial_t T_d^\varepsilon(t,x) \to \partial_t T_d(t,x)$ for every $(t,x) \in D^c$. Moreover, by the local uniform convergence result in Lemma 6.2, we deduce that $\partial_x T_d^\varepsilon$ and $\partial_t T_d^\varepsilon$ also converge locally uniformly. Denote $T_d^0 := T_d$, it follows that the mapping $(t,x,\varepsilon) \to \big(\partial_t T_d^\varepsilon(t,x), \partial_x T_d^\varepsilon(t,x)\big)$ is continuous on

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

where $m^0(t) := m(t)$ and $r^0(t) := r(t)$.

(iii) By exactly the same computation as in Proposition 3.12 of [21], we have

$$\partial_x T_d^\varepsilon(t,x) \quad = \quad \Big(1 + O(\varepsilon) + O\big(x - T_d^\varepsilon\big)\Big) \frac{\big(x - m^\varepsilon(t)\big) - \frac{1}{2}\big(x - T_d^\varepsilon\big) + O\big(\big(x - T_d^\varepsilon\big)^2\big)}{\big(x - m^\varepsilon(t)\big) - \big(x - T_d^\varepsilon\big) + O\big(\big(x - T_d^\varepsilon\big)^2\big)}(t,x),$$

and it follows by similar arguments as in [21] that

$$T_d^\varepsilon(t,x) - m^\varepsilon(t) \quad = \quad -\frac{1}{2}\big(x - m^\varepsilon(t)\big)\ +\ O\big((x - m^\varepsilon)^2\big),$$

and hence

$$\partial_x T_d^\varepsilon(t, m^\varepsilon(t) + \delta) \to -\frac{1}{2} \text{ uniformly for } t \in [0,1) \text{ and } \varepsilon \in [0, \varepsilon_0 \wedge 1 - t],\ \text{ as } \delta \searrow 0. \tag{6.14}$$

Next, using the estimation (6.14) 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,

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

It follows by the uniform convergence in (6.14) that

$$\partial_t T_d^\varepsilon(t,x) \quad = \quad -\frac{3}{2}\ \frac{\partial_t \delta^\varepsilon f(t, m^\varepsilon(t))}{\partial_x \delta^\varepsilon f(t, m^\varepsilon(t))}\ +\ C_4(\varepsilon,t,x)\big(x - m^\varepsilon(t)\big), \tag{6.15}$$

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 (6.14) and (6.15) 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 (6.10) and (6.14), 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 $\varepsilon \to 0$. □

6.2 Weak convergence to the Peacock process

Proof of Proposition 4.4. We recall that $\mathbb{P}^n$ is a martingale measure on the canonical space Ω, induced by the discrete-time martingale $X^{*,n}$ under the probability $\mathbb{P}^{*,n}$, whose upward jump size is $J_u^\varepsilon(t,x)$ and downward jump size is $J_d^\varepsilon(t,x)$ with $\varepsilon := 1/n$ (see the beginning of Section 4.1.2).
(i) To prove the tightness of $(\mathbb{P}^n)_{n\geq 1}$, we shall use Theorem VI.4.5 of Jacod and Shiryaev [32, P. 356].
First, Doob's martingale inequality implies that

$$\mathbb{P}^n\Big[\sup_{0\leq t\leq 1}|X_t|\geq K\Big] \;\leq\; \frac{\mathbb{E}^{\mathbb{P}^n}[|X_1|]}{K} \;=\; \frac{1}{K}\int_{\mathbb{R}}|x|\mu_1(dx) \;=:\; \frac{L_1}{K}, \qquad \forall K>0. \tag{6.16}$$

It follows that, to prove the tightness of $(\mathbb{P}^n)_{n\geq 1}$, we can suppose, without loss of generality, that the canonical process X under every $\mathbb{P}^n$ is a bounded martingale. Let $\delta>0$, it follows by Lemma 6.1 that the upward jump size $J_u^\varepsilon(t,x)$ is uniformly bounded by $C\varepsilon$ for some constant C on $D_\delta := \{(t,x)\;: m(t)\leq x\leq r(t)-\delta/2\}$. We then consider $\theta>0$ small enough such that $\theta\leq\frac{\delta}{2C}$ and $|l(t+\theta)-l(t)|+|r(t+\theta)-r(t)|\leq\delta/2$ 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 $X^{*,n}$ only increases between S and $S+\theta$, then clearly $|X_T^{*,n}-X_S^{*,n}|<\delta$. Therefore

$$\begin{aligned}\mathbb{P}^{*,n}\big[\big|X_T^{*,n}-X_S^{*,n}\big|\geq\delta\big] &\leq \mathbb{P}^{*,n}\big[\text{There is a down jump of } X^{*,n} \text{ on } [S,S+\theta]\big]\\ &\leq 1-(1-C\varepsilon)^{\theta/\varepsilon+1},\end{aligned}$$

where the last inequality follows by the estimate of q^ε in Lemma 6.1. It is clear that

$$\lim_{\theta\to 0}\;\limsup_{n\to\infty}\;\big(1-(1-C\varepsilon)^{\theta/\varepsilon+1}\big) \;=\; 0.$$

We then deduce from Theorem VI.4.5 of Jacod and Shiryaev [32] that the sequence $(X^{*,n})_{n\geq 1}$ 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 the Markov chain $X^{*,n}$, 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

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

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. $\square$

To show that a limit of $(\mathbb{P}^n)_{n\geq 1}$ provides a weak solution of (4.7), we shall consider the associated martingale problem. Let

$$\begin{aligned} M_t(\varphi,\mathbf{x}) &:= \varphi(\mathbf{x}_t) - \int_0^t j_u(s,\mathbf{x}_{s^-})D\varphi(\mathbf{x}_{s^-})\mathbf{1}_{\mathbf{x}_{s^-}>m(s)}ds \\ &\quad + \int_0^t \Big[\big[\varphi(\mathbf{x}_{s^-} - j_d(s,\mathbf{x}_{s^-})) - \varphi(\mathbf{x}_{s^-})\big]\frac{j_u}{j_d}(s,\mathbf{x}_{s^-})\Big]\mathbf{1}_{\mathbf{x}_{s^-}>m(s)}ds, \quad (6.17) \end{aligned}$$

for all $\mathbf{x} \in \Omega := D([0,1],\mathbb{R})$ 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 and Shiryaev [32]. In preparation, let us introduce, for every constant $p > 0$, an $\mathbb{F}$-stopping time and the corresponding stopped canonical process

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

Following [32], denote also $J(\mathbf{x}) := \{t > 0\ : \Delta\mathbf{x}(t) \neq 0\}$,

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

Proof of Theorem 4.11. 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 (4.11), 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 4.10, we need to adapt the localization technique in Jacod and Shiryaev [32], by using the stopping time τ_p. Our proof will be very similar to that of Theorem IX.3.39 in [32].

First, for every $n \geq 1$, $\mathbb{P}^n$ is induced by the Markov chain $(X^{*,n},\mathbb{P}^{*,n})$, then

$$\begin{aligned} \mathbb{E}_{t_k}^{\mathbb{P}^n}\big[\varphi(X_{t_{k+1}}) - \varphi(X_{t_k})\big] &= \mathbb{E}_{t_k}^{\mathbb{P}^n}\Big[\big\{\varphi\big(X_{t_k} + J_u^\varepsilon(t_k,X_{t_k})\big) - \varphi(X_{t_k})\big\}\Big(1 - \frac{J_d}{J_d+J_u}\Big)\mathbf{1}_{X_{t_k}\geq m^\varepsilon(t_k)}\Big] \\ &\quad + \mathbb{E}_{t_k}^{\mathbb{P}^n}\Big[\big\{\varphi\big(X_{t_k} - J_d^\varepsilon(t_k,X_{t_k})\big) - \varphi(X_{t_k})\big\}\frac{J_d}{J_d+J_u}\mathbf{1}_{X_{t_k}\geq m^\varepsilon(t_k)}\Big] \\ &=: \quad \alpha \ + \ \beta. \end{aligned}$$

By (6.8) in Remark 6.3 and the uniform continuity of $m^\varepsilon(t)$, we see that

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

where ρ_0 is the continuity modulus of $(t,\varepsilon) \mapsto m^\varepsilon(t)$ in Assumption 4.8. We also estimate similarly that

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

Therefore, let $0 \leq s < t \leq 1$, $p \in \mathbb{N}$, $\phi_s(X.)$ 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

$$\mathbb{E}^{\mathbb{P}^n}\Big[\phi_s(X.)\big(M_{t\wedge\tau_p}(\varphi, X) - M_{s\wedge\tau_p}(\varphi, X)\big)\Big] \quad \leq \quad C_p\ \varepsilon. \tag{6.18}$$

To proceed, we follow the same localization arguments as in the proof of Theorem IX.3.39 of Jacod and Shiryaev [32]. 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 $\mathcal{L}^{\mathbb{P}^n}(X^p_.) \to \mathbb{P}^{0,p}$ as $n \to \infty$. Due to the proof of Proposition IX.1.17 of [32], 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 [32] that $\omega \mapsto \tau_{a_p}(\omega)$ is $\mathbb{P}^{0,p}$-a.s. continuous and the law $\mathcal{L}^{\mathbb{P}^n}\big(X^p_., X^{a_p}_.\big)$ converges to $\mathcal{L}^{\mathbb{P}^{0,p}}\big(X, X^{\tau_{a_p}}\big)$.

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 $\mathcal{L}^{\mathbb{P}^n}\big(X^{a_p}\big) \to \tilde{\mathbb{P}}^{0,p}$. In particular, since there is a countable set $\mathbb{T}^* \subset [0,1]$ such that

$$\mathbf{x} \quad \mapsto \quad M_{t\wedge\tau_{a_p}}(\varphi, \mathbf{x}) - M_{s\wedge\tau_{a_p}}(\varphi, \mathbf{x}) \tag{6.19}$$

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 (6.18), we obtain

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

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 (6.17) 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 (6.17) and hence a weak solution to SDE (4.11).

Finally, for uniqueness of solutions to SDE (4.11), it is enough to use Theorem III-4 of Lepeltier and Marchal [37] (see also Theorem 14.18 of Jacod [31, P. 453]) together with localization technique to conclude the proof. □

Remark 6.5. *In the multiple local maximizers case under Assumption 4.2, the functions j_u and j_d are no more continuous, then the mapping* (6.19) *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 4.8 in Theorem 4.11.*

Proof of Lemma 4.6. We recall that by Theorem 3.8 in [21] (ii), the corresponding maps $T^\varepsilon_u(t,.)$ and $T^\varepsilon_d(t,.)$ solve the following ODEs:

$$\frac{d}{dx}\ \delta^\varepsilon F\big(t+\varepsilon, T^\varepsilon_d(t,x)\big) \quad = \quad (1-q)(t,x)f(t,x), \tag{6.20}$$

$$\frac{d}{dx}F\big(t+\varepsilon, T^\varepsilon_u(t,x)\big) \quad = \quad q(t,x)f(t,x) \ \text{ for all } \ x \in (D^\varepsilon)^c(t), \tag{6.21}$$

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 6.2. By direct substitution of this expression in the system of ODEs (6.20-6.21), 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 (4.6) holds true. $\square$

Proof of Proposition 4.7. By Lemma 4.6, item (ii) of Proposition 4.7 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

$$\begin{aligned} d\big(\widehat{X}_t - x\big)^+ &= dM_t + L_t \\ &:= \mathbf{1}_{\{\widehat{X}_{t-} > x\}} d\widehat{X}_t + \Big(\big(\widehat{X}_t - x\big)^+ - \big(\widehat{X}_{t-} - x\big)^+ - \mathbf{1}_{\{\widehat{X}_{t-} > x\}} \Delta \widehat{X}_t\Big), \end{aligned} \tag{6.22}$$

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

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

by direct computation, it follows by (4.8) together with dominated convergence theorem that

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

for every $\varepsilon \le \varepsilon_0$, where $\varepsilon_0 \in (0, 1-t)$. Then, integrating (6.22) between t and $t+\varepsilon$, and taking expectations, it follows that

$$\begin{aligned} &\mathbb{E}\Big[\Big(\widehat{X}_{t+\varepsilon} - x\Big)^+\Big] - \mathbb{E}\Big[\Big(\widehat{X}_t - x\Big)^+\Big] \\ &= \int_t^{t+\varepsilon} \int_{\mathbb{R}} \big(x - T_d(s,y)\big) \frac{j_u}{j_d}(s,y) \mathbf{1}_{\left\{T_d(s,y) \le x < y,\ y \in D^c(s)\right\}} f^{\widehat{X}}(s,y)\, dy\, ds. \end{aligned} \tag{6.23}$$

Let us now differentiate both sides of (6.23). For the left hand side, since the density function $f^{\widehat{X}}(t,.)$ of $\widehat{X}_t$ is continuous, the function $x \mapsto \mathbb{E}\big[(\widehat{X}_t - x)^+\big] = \int_x^\infty (y-x) f^{\widehat{X}}(t,y) dy$ is differentiable and

$$\partial_x \mathbb{E}\Big[\big(\widehat{X}_t - x\big)^+\Big] = \int_x^\infty -f^{\widehat{X}}(t,y) dy, \qquad \partial^2_{xx} \mathbb{E}\big[\big(\widehat{X}_t - x\big)^+\big] = f^{\widehat{X}}(t,x). \tag{6.24}$$

We now consider the rhs of (6.23) and denote

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

Let us fix $s \in [0,1)$ and $x \in D^{c,\circ}(s)$, then it is clear that

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

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

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

and

$$\partial^2_{xx} l(s,x) \quad = \quad \partial_x\big(j_u f^{\widehat{X}}\big)(s,x) \ - \ \frac{j_u}{j_d} f^{\widehat{X}}(s,x). \tag{6.25}$$

We now consider the case $x \in D^\circ$. Notice that $T_d(s,\cdot) : D^c(s) \to D(s)$ 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 $D(s)$, given by

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

Then by the expression that

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

we get

$$\partial^2_{xx} l(x) \quad = \quad \tfrac{j_u}{j_d}\big(s, T_d^{-1}(s,x)\big) f^{\widehat{T}}(s,z) \quad = \quad \tfrac{j_u}{j_d}\big(s, T_d^{-1}(s,x)\big) \tfrac{f^{\widehat{X}}}{\partial_x T_d}(s,x), \quad \forall x \in D^\circ(s). \tag{6.26}$$

Finally, differentiating both sides of (6.23) (with (6.24), (6.25) and (6.26)), then dividing them by ε and sending $\varepsilon \searrow 0$, it follows that

$$\partial_t f^{\widehat{X}}(t,x) \quad = \quad \mathbf{1}_{\{x\in D^c(t)\}} \Big(\partial_x\big(f^{\widehat{X}} j_u\big) - \frac{j_u f^{\widehat{X}}}{j_d}\Big)(t,x) - \mathbf{1}_{\{x\in D(t)\}} \frac{j_u f^{\widehat{X}}}{j_d(1-\partial_x j_d)}\big(t, T_d^{-1}(t,x)\big),$$

for every $t \in [0,1)$ and $x \in D^\circ(t) \cup D^{c,\circ}(t)$. □

6.3 Convergence of the robust superhedging strategy

To prove Theorem 4.12, we recall that under every $\mathbb{P}^n$, we have $\mathbb{P}^n$-a.s. that

$$\sum_{k=0}^{n-1} \big(\varphi^\varepsilon(t_k, X_{t_k}) + \psi^\varepsilon(t_k, X_{t_{k+1}})\big) + \sum_{k=0}^{n-1} h^\varepsilon(t_k, X_{t_k}) \left(X_{t_{k+1}} - X_{t_k}\right) \ge \sum_{k=0}^{n-1} c\big(X_{t_k}, X_{t_{k+1}}\big) \tag{6.27}$$

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 (4.11).

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

$$\begin{aligned}\Psi^*(\mathbf{x}) \;&:=\; \psi^*(1,\mathbf{x}_1)-\psi^*(0,\mathbf{x}_0)-\int_0^1\Big(\partial_t\psi^*(t,\mathbf{x}_t)+j_u(t,\mathbf{x}_t)\mathbf{1}_{\mathbf{x}_t>m(t)}\partial_x\psi^*(t,\mathbf{x}_t)\Big)dt \qquad (6.28)\\ &\quad+\int_0^1\frac{j_u(t,\mathbf{x}_t)}{j_d(t,\mathbf{x}_t)}\mathbf{1}_{\mathbf{x}_t>m(t)}\Big(\psi^*(t,\mathbf{x}_t)-\psi^*(t,\mathbf{x}_t-j_d(t,\mathbf{x}_t))+c(\mathbf{x}_t,\mathbf{x}_t-j_d(t,\mathbf{x}_t))\Big)dt.\end{aligned}$$

Lemma 6.6. *Let Assumptions 4.1 and 4.8 hold true. Then for every càdlàg path $\mathbf{x}\in D([0,1])$ taking value in (ℓ_1,r_1), we have*

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

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

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

First, we have $\psi^\varepsilon(t_{n-1},\mathbf{x}_1)\to\psi^*(1,\mathbf{x}_1)$ and by Lemma 6.4,

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

Further, when $x>m(t)$,

$$\begin{aligned}\varphi^\varepsilon+\psi^\varepsilon\;&=\;\psi^\varepsilon-\psi^\varepsilon(.,T_u^\varepsilon)+\frac{J_u^\varepsilon}{J_u^\varepsilon+J_d^\varepsilon}\Big(\psi^\varepsilon(.,T_u^\varepsilon)+c(.,T_d^\varepsilon)-\psi^\varepsilon(.,T_d^\varepsilon)\Big)+o(\varepsilon)\\ &=\;-\varepsilon j_u\partial_x\psi^\varepsilon+\varepsilon\frac{j_u}{j_d}\Big(\psi^\varepsilon-\psi^\varepsilon(.,T_d)+c(.,T_d)\Big)+o(\varepsilon).\end{aligned}$$

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

$$\begin{aligned}&\int_0^1-\partial_x\psi^*(t,\mathbf{x}_t)j_u(t,\mathbf{x}_t)dt\\ +\;&\int_0^1\frac{j_u(t,\mathbf{x}_t)}{j_d(t,\mathbf{x}_t)}1_{\mathbf{x}_t>m(t)}\Big(\psi^*(t,\mathbf{x}_t)-\psi^*(t,\mathbf{x}_t-j_d(t,\mathbf{x}_t))+c(\mathbf{x}_t,\mathbf{x}_t-j_d(t,\mathbf{x}_t))\Big)dt,\end{aligned}$$

which concludes the proof. □

Lemma 6.7. *Let Assumptions 4.1 and 4.8 hold true, and $\mu(\overline{\lambda}^*)<\infty$. Then for the limit probability measure $\mathbb{P}^0$ given in Theorem 4.11, we have*

$$\mathbb{E}^{\mathbb{P}^0}\big[C(X.)\big]\;=\;\mathbb{E}^{\mathbb{P}^0}\big[\Psi^*(X.)\big]\;=\;\mu(\lambda^*)\;=\;\int_0^1\int_{m(t)}^{r(t)}\frac{j_u(t,x)}{j_d(t,x)}\;c\big(x,x-j_d(t,x)\big)f(t,x)dxdt.$$

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

$$
\begin{aligned}
&\psi^*(t, X_t) - \psi^*(0, X_0) - \int_0^t \partial_t \psi^*(t, X_s)dt \\
- &\int_0^t \Big[j_u(s, X_s)\partial_x \psi^*(s, X_s) + \frac{j_u}{j_d}(s, X_s)\big[\psi^*(s, X_s - j_d(s, X_s)) - \psi^*(s, X_s)\big]\Big] \mathbf{1}_{X_s > m(s)} ds.
\end{aligned}
$$

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

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

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

To computer $\mathbb{E}^{\mathbb{P}^0}[C(X_\cdot)]$, we notice that $[X]^c_t = 0$, $\mathbb{P}^0 - a.s.$, 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

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

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. □

Next, let us consider the limit of the second term on the left hand side of (6.27).

Lemma 6.8. *Let Assumptions 4.1 and 4.8 hold true. Then we have the following convergence in probability under every martingale measure $\mathbb{P} \in \mathcal{M}_\infty$:*

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

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 6.4. By the right continuity of martingale X, the above lemma is then a direct application of Theorem I.4.31 of Jacod and Shiryaev [32]. □

Proof of Theorem 4.12. Using (6.27), together with Lemmas 3.3, 6.6 and 6.8, 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

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

Further, by weak duality, we have

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

Since $\mathbb{E}^{\mathbb{P}^0}[C(X.)] = \mu(\lambda^*)$ by Lemma 6.7, this implies the strong duality as well as the optimality of the local Lévy process (4.11) and the semi-static superhedging strategy described by (h^*, ψ^*). $\square$

References

[1] Albin, J. M. P.: *A continuous non-Brownian motion martingale with Brownian motion martingale distributions*, Statistics and Probability Letters, 78, 6, 682686. 2008.

[2] Beiglböck, and Huesmann, M. Optimal Transport and Skorokhod Embedding. Preprint 2013.

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

[4] Beiglböck, M. , Juillet, N. : *On a problem of optimal transport under marginal martingale constraints*, preprint, 2012, arXiv:1208.1509.

[5] 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.

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

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

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

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

[10] 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.

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

[12] Dolinsky, Y., Soner, H.M. : *Robust hedging and martingale optimal transport in continuous time*, 2012.

[13] Dolinsky, Y., Soner, H.M. : *Robust Hedging under Proportional Transaction Costs*, Preprint, 2013.

[14] Fan, J.Y. Hamza, K., Klebaner, F.C. : *Mimicking self-similar processes.* Preprint, 2012.

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

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

[17] 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, to appear.

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

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

[20] H. Heinich. *Problème de Monge pour n probabilités.* C.R. Math. Acad. Sci. Paris, 334(9):793795, 2002.

[21] Henry-Labordère, P., and Touzi, N. : *An Explicit Martingale Version of Brenier's Theorem.* Preprint 2013.

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

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

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

[25] Hobson, D. : *Robust hedging of the lookback option*, Finance and Stochastics, 2, pp329-347, 1998.

[26] 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.

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

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

[29] Hobson, D., Klimmek, M. : *Robust price bounds for the forward starting straddle*, preprint, 2013.

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

[31] Jacod, J. : *Calcul stochastique et problemes de martingales*, Spinrger, 1979.

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

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

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

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

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

[37] 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.

[38] Madan, D.B., Yor, M. : *Making Markov martingales meet marginals: with explicit constructions*, Bernoulli, 8(4):509536, 2002.

[39] Obloj, J. : *The Skorokhod embedding problem and its offspring*, Probab. Surv., 1:321390, 2004.

[40] Oleszkiewicz, K. : *On fake Brownian motions*, Statistics and Probability Letters, 78, 12511254, 2008.

[41] I. Olkin and S.T. Rachev. *Maximum submatrix traces for positive denite matrices.* SIAM J. Matrix Ana. Appl., 14:39039, 1993.

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

[43] Pass, B.: *Uniqueness and the Monge solutions in the multi marginal optimal transport problem.* SIAM Journal on Mathematical Analysis, 43(6):27582775, 2011.

[44] 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.

[45] 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.")

[46] Pass, B.: *On a class of optimal transportation problems with infinitely many marginals*, preprint (Jun. 2012), arXiv:1206.5543v1.

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

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

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

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