ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

stochastic processes and their applications

www.elsevier.com/locate/spa

Tightness and duality of martingale transport on the Skorokhod space☆

Gaoyue Guo[a,*], Xiaolu Tan[b], Nizar Touzi[a]

[a] *CMAP, Ecole Polytechnique, France*

[b] *CEREMADE, University of Paris-Dauphine, PSL Research University, France*

Received 6 July 2015; received in revised form 24 May 2016; accepted 6 July 2016

Abstract

The *martingale optimal transport* aims to optimally transfer a probability measure to another along the class of martingales. This problem is mainly motivated by the robust superhedging of exotic derivatives in financial mathematics, which turns out to be the corresponding Kantorovich dual. In this paper we consider the continuous-time martingale transport on the Skorokhod space of càdlàg paths. Similar to the classical setting of optimal transport, we introduce different dual problems and establish the corresponding dualities by a crucial use of the S-topology and the dynamic programming principle.

MSC: 65B05; 60B10; 60G44; 91G20

Keywords: S-topology; Dynamic programming principle; Robust superhedging

1. Introduction

Initialed by the famous work of Monge and Kantorovich, the *optimal transport problem* optimizes the cost of the transfer of mass from one location to another. Namely, let $\mathcal{P}(\mathbb{R}^d)$ be the

☆ We 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). We are grateful to Adam Jakubowski and Alexander Cox for fruitful discussions and comments and two anonymous referees for helpful and useful comments and suggestions.

* Corresponding author.

E-mail addresses: guo@cmap.polytechnique.fr (G. Guo), tan@ceremade.dauphine.fr (X. Tan), nizar.touzi@polytechnique.edu (N. Touzi).

http://dx.doi.org/10.1016/j.spa.2016.07.005

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

space of probability measures on the Euclidean space $\mathbb{R}^d$. For any given measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, let

$$\mathcal{P}(\mu, \nu) := \left\{\mathbb{P} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : \mathbb{P} \circ X^{-1} = \mu \text{ and } \mathbb{P} \circ Y^{-1} = \nu\right\}, \tag{1.1}$$

where (X, Y) denotes the canonical process on $\mathbb{R}^d \times \mathbb{R}^d$, i.e. $X(x, y) = x$ and $Y(x, y) = y$ for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. Then the optimal transport problem consists in optimizing the expectation of some measurable function $\xi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ among all probability measures in $\mathcal{P}(\mu, \nu)$. Various related issues are studied, e.g. the general duality theory and optimality results, we refer to Rachev & Rüschendorf [37] and Villani [39] for a comprehensive account of the literature.

Recently, a martingale optimal transport problem was introduced in Beiglböck, Henry-Labordère & Penkner [2] in discrete-time (see Galichon, Henry-Labordère & Touzi [16] for the continuous-time case), where a maximization problem is considered over a subset $\mathcal{M}(\mu, \nu) := \left\{\mathbb{P} \in \mathcal{P}(\mu, \nu) : \mathbb{E}^{\mathbb{P}}[Y|X] = X, \ \mathbb{P}\text{-a.s.}\right\}$:

$$\mathsf{P}(\mu, \nu) := \sup_{\mathbb{P} \in \mathcal{M}(\mu, \nu)} \mathbb{E}^{\mathbb{P}}\big[\xi(X, Y)\big].$$

Each element of $\mathcal{M}(\mu, \nu)$ is called a *transport plan*. Similarly to the classical setting, the corresponding dual problem is defined by

$$\mathsf{D}(\mu, \nu) := \inf_{(\lambda, \varphi, H) \in \mathcal{D}(\mu, \nu)} \left\{\int \lambda d\mu + \int \varphi d\nu\right\},$$

with $\mathcal{D}(\mu, \nu)$ being the collection of triplets (λ, φ, H), where $\lambda, \varphi, H : \mathbb{R}^d \to \mathbb{R}$ are measurable functions such that $\lambda \in \mathbb{L}^1(\mu)$, $\varphi \in \mathbb{L}^1(\nu)$ and

$$\lambda(x) + \varphi(y) + H(x)(y - x) \geq \xi(x, y) \quad \text{for all } (x, y) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{1.2}$$

The last dual formulation has the interpretation of minimal robust superhedging cost of derivative security defined by the payoff ξ by trading the underlying security and any possible Vanilla option. When $d = 1$, as observed by Breeden & Litzenberger [4], the marginal distributions of the underlying asset are recovered by the market prices of calls for all strikes, and any Vanilla option has a non-ambiguous price as the integral of its payoff function with respect to the marginal. Therefore, the inequality (1.2) represents a super-replication of ξ, which consist of the trading of the underlying and Vanilla options at different maturities. Since there is no specific model imposed on the process (X, Y), the dual problem may be interpreted as the robust superhedging cost, i.e. the minimum cost to construct super-replications. Similar to the classical setting, the duality $\mathsf{P}(\mu, \nu) = \mathsf{D}(\mu, \nu)$ holds under quite general conditions.

The present paper considers the continuous-time martingale optimal transport problem. Let $\mathcal{X} := \left\{\omega = (\omega_t)_{0 \leq t \leq 1} : \omega_t \in \mathbb{R}^d \text{ for all } t \in [0, 1]\right\}$, where $\mathcal{X}$ is either the space of continuous functions or the Skorokhod space of càdlàg functions. Denote by $X = (X_t)_{0 \leq t \leq 1}$ the canonical process and by $\mathcal{M}$ the set of all martingale measures $\mathbb{P}$, i.e. X is a martingale under $\mathbb{P}$. For a given family of probability measures $\boldsymbol{\mu} = (\mu_t)_{t \in \mathbb{T}}$, where $\mathbb{T} \subseteq [0, 1]$ is a subset, define by $\mathcal{M}(\boldsymbol{\mu})$ the subset of transport plans $\mathbb{P}$, i.e. $\mathbb{P} \circ X_t^{-1} = \mu_t$ for all $t \in \mathbb{T}$. Then for a measurable function $\xi : \mathcal{X} \to \mathbb{R}$, the problem is defined by

$$\mathsf{P}(\boldsymbol{\mu}) := \sup_{\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}[\xi(X)]. \tag{1.3}$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

In contrast with the discrete-time case, the set $\mathcal{M}(\boldsymbol{\mu})$ is generally not tight with respect to the usual topologies. Without the crucial compactness, the arguments in the classical setting fail to be adapted to handle the related issues.

In the existing literature, there are two dual formulations for the problem (1.3), Galichon, Henry-Labordère & Touzi studied a class of transport plans defined by stochastic differential equations in [16] and introduced a quasi-sure dual problem. They applied a stochastic control approach and deduced the duality. Another important contribution is due to Dolinsky & Soner [14,13], see also Hou & Oblój [26], where the dual problem is still pathwisely formulated as in (1.2). By discretizing the paths and a technical construction of approximated martingale measures, they avoid the compactness issue and derive the duality.

In addition, the martingale optimal transport problem is studied by the approach of Skorokhod embedding problem. Following the seminal paper of Hobson [20], this methodology generated developments in many directions, see e.g. Brown, Hobson & Rogers [5], Cox & Oblój [8,7], Cox, Hobson & Oblój [6], Cox, Oblój & Touzi [9], Cox & Wang [10], Davis, Oblój & Raval [11], Gassiat, Oberhauser & dos Reis [17], Hobson & Klimmek [22,24,23], Hobson & Neuberger [25] and Madan & Yor [31]. A thorough literature is provided in the survey papers Hobson [21] and Oblój [36].

Our main contribution in the paper is to study systematically the tightness of the set $\mathcal{M}(\boldsymbol{\mu})$ by means of the S-topology introduced in Jakubowski [28]. Endowing properly the space of marginal laws with a Wasserstein kind topology, the tightness yields the upper semicontinuity of the map $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ and thus the first duality, obtained by penalizing the marginal constraints. Based on the first duality and using respectively the dynamic programming principle and the discretization argument of path-space, the dualities are established for both quasi-sure and pathwise dual formulations.

The above analysis immediately gives rise to a stability consequence. Denote $\overline{\mathsf{P}} := \mathsf{P}$ and $\underline{\mathsf{P}}(\boldsymbol{\mu}) := \inf_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}[\xi(X)]$, then it is shown that the map $\boldsymbol{\mu} \mapsto \overline{\mathsf{P}}(\boldsymbol{\mu})$ (resp. $\boldsymbol{\mu} \mapsto \underline{\mathsf{P}}(\boldsymbol{\mu})$) is upper (resp. lower) semicontinuous, which yields the stability, i.e. for any sequence $(\boldsymbol{\mu}^n)_{n\geq 1}$ convergent to $\boldsymbol{\mu}$, there exists a sequence $(\varepsilon_n)_{n\geq 1} \subseteq \mathbb{R}_+$ convergent to zero such that

$$\big[\underline{\mathsf{P}}(\boldsymbol{\mu}^n),\ \overline{\mathsf{P}}(\boldsymbol{\mu}^n)\big] \subseteq \big[\underline{\mathsf{P}}(\boldsymbol{\mu}) - \varepsilon_n,\ \overline{\mathsf{P}}(\boldsymbol{\mu}) + \varepsilon_n\big] \quad \text{for all } n \geq 1,$$

i.e. the interval of model-free prices is stable with respect to the market.

The paper is organized as follows. We formulate the martingale optimal transport problem and provide the dual problems in Section 2. In Section 3, the duality results are presented and we reduce the infinitely-many marginal constraints to the finitely-many marginal constraints. In Sections 4, 5 we focus on the finitely-many marginal case and provide all related proofs.

2. Martingale optimal transport

For all $0 \leq s < t$, denote by $\mathbb{D}([s,t],\mathbb{R}^d)$ the space of càdlàg functions defined on $[s,t]$ taking values in $\mathbb{R}^d$. Let $\Omega := \mathbb{D}([0,1],\mathbb{R}^d)$ with generic element denoted by ω. Denote further by $X := (X_t)_{0\leq t\leq 1}$ the canonical process, i.e. $X_t(\omega) = \omega_t$ and by $\mathbb{F} := (\mathcal{F}_t)_{0\leq t\leq 1}$ its natural filtration, i.e. $\mathcal{F}_t = \sigma(X_u, u \leq t)$. Let $\mathcal{P} := \mathcal{P}(\Omega, \mathcal{F}_1)$ be the set of probability measures on Ω. A probability measure $\mathbb{P} \in \mathcal{P}$ is called a martingale measure if the canonical process X is a martingale under $\mathbb{P}$. Denote by $\mathcal{M}$ the collection of all martingale measures.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

2.1. Peacock and martingale optimal transport

Let $\mathbf{P} := \mathbf{P}(\mathbb{R}^d)$ be the space of all probability measures μ on $\mathbb{R}^d$ with finite first moment. A pair $(\mu, \nu) \in \mathbf{P} \times \mathbf{P}$ is said to be increasing in convex ordering if

$$\int_{\mathbb{R}^d} \lambda(x)\mu(dx) := \mu(\lambda) \le \nu(\lambda) := \int_{\mathbb{R}^d} \lambda(x)\nu(dx)$$

holds for every convex function $\lambda : \mathbb{R}^d \to \mathbb{R}$. This relation is denoted by $\mu \preceq \nu$. Let $\mathbb{T} \subseteq [0, 1]$ be some subset containing 1 and define the $\mathbb{T}$-product of $\mathbf{P}$ by

$$\mathbf{P}^{\mathbb{T}} := \Big\{ \boldsymbol{\mu} := (\mu_t)_{t\in\mathbb{T}} : \mu_t \in \mathbf{P} \text{ for all } t \in \mathbb{T} \Big\}.$$

Definition 2.1. A family of probability measures $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}^{\mathbb{T}}$ is called a peacock ($\mathbb{T}$-peacock) if $\mu_s \preceq \mu_t$ holds for all $s, t \in \mathbb{T}$ such that $s \le t$. A peacock $\boldsymbol{\mu}$ is said to be càdlàg if the map $t \mapsto \mu_t$ is càdlàg on $\mathbb{T}$ with respect to the weak convergence. Denote by $\mathbf{P}^{\mathbb{T}}_{\preceq}$ the set of all càdlàg peacocks.

For each peacock $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$, define the set of transport plans

$$\mathcal{M}(\boldsymbol{\mu}) := \Big\{ \mathbb{P} \in \mathcal{M} : \mathbb{P} \circ X_t^{-1} = \mu_t \text{ for all } t \in \mathbb{T} \Big\}. \tag{2.1}$$

We may assume without loss of generality that $\mathbb{T}$ is closed under the *lower limit topology*, i.e. the topology generated by all half-open intervals $[s, t) \subseteq [0, 1]$, see e.g. Steen & Seebach [38]. Indeed, denote by $\bar{\mathbb{T}}$ the closure of $\mathbb{T}$ under the lower limit topology, then it follows that the law of X_t for $t \in \bar{\mathbb{T}}$ is uniquely determined by the right continuity of X. This implies that $\mathcal{M}(\bar{\boldsymbol{\mu}}) = \mathcal{M}(\boldsymbol{\mu})$, where $\bar{\boldsymbol{\mu}} := (\bar{\mu}_t)_{t\in\bar{\mathbb{T}}}$ is defined by

$$\bar{\mu}_t := \lim_{n\to\infty} \mu_{t_n} \text{ for any sequence } (t_n)_{n\ge1} \subseteq \mathbb{T} \text{ decreasing to } t. \tag{2.2}$$

Remark 2.2. (i) Since $\mu_{t_n} \preceq \mu_1$ for all n, we have

$$\mu_{t_n}\big((x_i - K)^+\big) \le \mu_1\big((x_i - K)^+\big) \quad \text{for all } i = 1, \dots, d,$$

thus showing that the sequence $(\mu_{t_n})_{n\ge1}$ is uniformly integrable. In particular, $(\mu_{t_n})_{n\ge1}$ is tight, and we may verify immediately by a direct density argument that any two possible accumulation points $\bar{\mu}_t$ and $\bar{\mu}'_t$ coincide, i.e. $\bar{\mu}_t = \bar{\mu}'_t$. Hence the sequence $(\mu_{t_n})_{n\ge1}$ converges weakly, justifying the convergence in (2.2) is well defined.

(ii) When $\mathbb{T} = [0, 1]$, $\mathcal{M}(\boldsymbol{\mu})$ is nonempty by Kellerer's theorem, see e.g. Hirsch & Roynette [18,19] and Kellerer [29]. For a general closed $\mathbb{T}$, we may extend $\boldsymbol{\mu}$ to some $\bar{\boldsymbol{\mu}} = (\bar{\mu}_t)_{0\le t\le1}$ by $\bar{\mu}_t := \bar{\mu}_{\bar{t}}$ with $\bar{t} := \inf\{s \ge t : s \in \mathbb{T}\}$. Clearly, $\bar{\boldsymbol{\mu}} \in \mathbf{P}^{[0,1]}_{\preceq}$ and $\bar{\mu}_t = \mu_t$ for all $t \in \mathbb{T}$. Hence $\mathcal{M}(\boldsymbol{\mu}) \supseteq \mathcal{M}(\bar{\boldsymbol{\mu}})$ is again nonempty.

Let $\xi : \Omega \to \mathbb{R}$ be a measurable function. For every peacock $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$, define the martingale optimal transport problem by

$$\mathsf{P}(\boldsymbol{\mu}) := \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\big[\xi(X)\big], \tag{2.3}$$

where $\mathbb{E}^{\mathbb{P}}[\xi] := \mathbb{E}^{\mathbb{P}}[\xi^+] - \mathbb{E}^{\mathbb{P}}[\xi^-]$ with the convention $+\infty - \infty = -\infty$.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

2.2. Dual problems

First dual problem. Let Λ be the set of continuous functions $\lambda : \mathbb{R}^d \to \mathbb{R}$ with linear growth, i.e. $\sup_{x\in\mathbb{R}^d}\big(|\lambda(x)|/(1+|x|)\big) < +\infty$. Define

$$\Lambda^{\mathbb{T}} := \Big\{\boldsymbol{\lambda} := (\lambda_{t_i})_{1\le i\le m} : t_i \in \mathbb{T},\ \lambda_{t_i} \in \Lambda \text{ for all } i = 1,\dots,m,\ m \in \mathbb{N}\Big\}.$$

For every $\boldsymbol{\lambda} = (\lambda_{t_i})_{1\le i\le m} \in \Lambda^{\mathbb{T}}$, $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$ and $\omega \in \Omega$, denote

$$\boldsymbol{\lambda}(\omega) := \sum_{i=1}^m \lambda_{t_i}(\omega_{t_i}) \quad \text{and} \quad \boldsymbol{\mu}(\boldsymbol{\lambda}) := \sum_{i=1}^m \mu_{t_i}(\lambda_{t_i}).$$

Next, we introduce three dual formulations. Roughly speaking, as X is required to be a martingale and has the given marginal laws in problem (2.3), then we dualize respectively these two constraints. The first dual problem is defined by

$$\mathsf{D}_1(\boldsymbol{\mu}) := \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}}\Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) + \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\big[\xi(X) - \boldsymbol{\lambda}(X)\big]\Big\}. \tag{2.4}$$

The dual problem D_1 is the Kuhn–Tucker formulation in convex optimization, where the marginal constraints $\boldsymbol{\mu}$ are penalized by the Lagrange multipliers $\boldsymbol{\lambda}$.

Second dual problem. The second dual problem dualizes further the martingale constraint and has close analogues in the mathematical finance literature in the context of a financial market with d risky assets, where the price process is modeled by the canonical process $X = (X_t)_{0\le t\le 1}$. For technical reasons, the underlying process X is assumed to be non-negative and start at some fixed price that may be normalized to be $\mathbf{1} := (1,\dots,1) \in \mathbb{R}^d$. Namely, define the set of market scenarios

$$\Omega_+ := \big\{\omega \in \Omega : \omega_0 = \mathbf{1} \text{ and } \omega_t \in \mathbb{R}^d_+ \text{ for all } t \in [0,1]\big\}$$

and the set of all possible models $\mathcal{M}_+ := \big\{\mathbb{P} \in \mathcal{M} : \mathrm{supp}(\mathbb{P}) \subseteq \Omega_+\big\}$. Consequently, the market calibration $\boldsymbol{\mu}$ should satisfy

$$\mu_0(dx) = \delta_{\mathbf{1}}(dx) \quad \text{and} \quad \mathrm{supp}(\mu_1) \subseteq \mathbb{R}^d_+. \tag{2.5}$$

Moreover, let us denote by $\mathbb{F}^U = (\mathcal{F}^U_t)_{0\le t\le 1}$ the universally completed filtration, i.e. $\mathcal{F}^U_t := \cap_{\mathbb{P}\in\mathcal{P}}\mathcal{F}^{\mathbb{P}}_t$, where $\mathcal{F}^{\mathbb{P}}_t$ is the completed σ-field of $\mathcal{F}_t$ under $\mathbb{P}$.

Definition 2.3. A process $S = (S_t)_{0\le t\le 1}$ is called a $\mathcal{M}_+$-supermartingale if it is $\mathbb{F}^U$-adapted and is a $\mathbb{P}$-supermartingale for all $\mathbb{P} \in \mathcal{M}_+$. Denote by $\mathcal{S}$ the collection of all $\mathcal{M}_+$-supermartingales and by $\mathcal{S}_0 \subseteq \mathcal{S}$ the subset of processes starting at 0. Denote further

$$\mathcal{D}_2(\xi) := \Big\{(\boldsymbol{\lambda}, S) \in \Lambda^{\mathbb{T}} \times \mathcal{S}_0 : \boldsymbol{\lambda}(\omega) + S_1(\omega) \ge \xi(\omega) \text{ for all } \omega \in \Omega_+\Big\}.$$

For a peacock $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$ satisfying (2.5), the second dual problem is defined by

$$\mathsf{D}_2(\boldsymbol{\mu}) := \inf_{(\boldsymbol{\lambda},S)\in\mathcal{D}_2(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{2.6}$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Remark 2.4. (i) Notice that the supermartingale $S \in \mathcal{S}$ is not required to have any regularity. If it were càdlàg then it would follow from Theorem 2.1 in Kramkov [30] that, for every $\mathbb{P} \in \mathcal{M}_+$ there exist a predictable process $H^{\mathbb{P}} = (H_t^{\mathbb{P}})_{0\le t\le 1}$ and an optional non-decreasing process $A^{\mathbb{P}} = (A_t^{\mathbb{P}})_{0\le t\le 1}$ such that

$$S_t = S_0 + \int_0^t H_s^{\mathbb{P}} dX_s - A_t^{\mathbb{P}} \quad \text{for all } t \in [0,1],\ \mathbb{P}\text{-a.s.}$$

However, it is not clear whether one can aggregate the last representation, i.e. find predictable processes H and A such that $(H, A) = (H^{\mathbb{P}}, A^{\mathbb{P}})$, $\mathbb{P}$-almost surely. See also Nutz [35] for a partial result of this direction.

(ii) In financial mathematics, the pair $(\boldsymbol{\lambda}, H^{\mathbb{P}})$ has the interpretation of a semi-static super-replicating strategy under the model $\mathbb{P}$. If the aggregation above were possible, then the dual problem D_2 turns to the quasi-sure formulation of the robust superhedging problem, see also Beiglböck, Nutz & Touzi [3], and the duality $\mathsf{P} = \mathsf{D}_2$ reduces to the quasi-sure pricing–hedging duality.

Third dual problem. Following the pioneering work [20] of Hobson, the martingale optimal transport approach is applied to study the robust hedging problems in finance. We do not postulate any specific model on the underlying assets and pursue here a robust approach. Assume further that all call/put options are liquid in the market for maturities $t \in \mathbb{T}$, thus yielding a family of marginal distributions $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}}$ that is considered to be exogenous, see e.g. Breeden & Litzenberger [4]. Then, the time 0 market price of any derivative $\lambda(X_t)$ is given by $\mu_t(\lambda)$. Hence, the cost of a static strategy $\boldsymbol{\lambda} \in \Lambda^{\mathbb{T}}$ is $\boldsymbol{\mu}(\boldsymbol{\lambda})$.

The return from a zero-initial cost dynamic trading, defined by a suitable process $H = (H_t)_{0\le t\le 1}$, is given by the stochastic integral $(H \cdot X)$ which we define similarly to Dolinsky & Soner [13]. We restrict $H : [0,1] \to \mathbb{R}^d$ to be left-continuous with bounded variation. Then, we may define the stochastic integral by integration by parts:

$$(H \cdot X)_t := H_t \cdot X_t - H_0 \cdot X_0 - \int_0^t X_u \cdot dH_u \quad \text{for all } t \in [0,1], \tag{2.7}$$

where $\int_0^t X_u \cdot dH_u$ refers to the scalar Lebesgue–Stieltjes integration.

Definition 2.5. An $\mathbb{F}$-adapted process $H : [0,1] \times \Omega_+ \to \mathbb{R}^d$ is called a dynamic strategy if $t \mapsto H_t(\omega)$ is left-continuous and of bounded variation for every $\omega \in \Omega_+$ and $(H \cdot X)$ is a supermartingale under every $\mathbb{P} \in \mathcal{M}_+$. Let $\mathcal{A}$ be the set of all dynamic strategies and define the set of robust super-replications

$$\mathcal{D}_3(\xi) := \Big\{(\boldsymbol{\lambda}, H) \in \Lambda^{\mathbb{T}} \times \mathcal{A} : \boldsymbol{\lambda}(\omega) + (H \cdot X)_1(\omega) \ge \xi(\omega) \text{ for all } \omega \in \Omega_+\Big\}.$$

For a peacock $\boldsymbol{\mu} \in \mathbf{P}_{\preceq}^{\mathbb{T}}$ satisfying (2.5), the third dual problem is defined by

$$\mathsf{D}_3(\boldsymbol{\mu}) := \inf_{(\boldsymbol{\lambda}, H)\in\mathcal{D}_3(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{2.8}$$

Remark 2.6. It is clear by definition that the weak duality $\mathsf{P}(\boldsymbol{\mu}) \le \mathsf{D}_1(\boldsymbol{\mu})$ holds. Moreover, if the peacock $\boldsymbol{\mu}$ satisfies (2.5), then

$$\mathsf{P}(\boldsymbol{\mu}) \le \mathsf{D}_1(\boldsymbol{\mu}) \le \mathsf{D}_2(\boldsymbol{\mu}) \le \mathsf{D}_3(\boldsymbol{\mu}).$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

3. Main results

We aim to study the existence of optimal transport plans and establish the dualities in a systematic way. Before providing these results in Sections 3.2 and 3.3, we first introduce some notions of topology on Ω and the associated space of probability measures in Section 3.1.

3.1. Preliminaries

In the classical optimal transport problem, the relevant results (existence of optimizers, duality, etc.) rely essentially on the compactness condition of $\mathcal{M}(\mu, \nu)$. However, when passing to the continuous-time case, as shown by Example 3.1 below, the set $\mathcal{M}(\boldsymbol{\mu})$ is in general not tight with respect to the topologies L^∞ (uniform topology) and J_1 (Skorokhod topology). For our purpose, we endow Ω with the S-topology introduced by Jakubowski [28] such that the Borel σ-field agrees with the projection σ-field $\mathcal{F}_1$, and more importantly, the S-topology facilitates the tightness issue and both Skorokhod representation theorem and Prohorov's theorem hold true. Before introducing the S-topology, we give an example which shows that the topologies L^∞ and J_1 are not convenient to handle the tightness of $\mathcal{M}(\boldsymbol{\mu})$.

Example 3.1. Let $M = (M_0, M_1, M_2)$ be a discrete-time martingale on some probability space such that $\mathbb{P}\big[M_0 \neq M_1 \text{ and } M_1 \neq M_2\big] > 0$. Define $\mathbb{P}_n := \mathbb{P} \circ (M^n)^{-1}$ for $n \geq 3$, where $M^n = (M_t^n)_{0\leq t\leq 1}$ is defined by

$$M_t^n := M_0 \mathbb{1}_{\left[0, \frac{1}{2}-\frac{1}{n}\right)}(t) + M_1 \mathbb{1}_{\left[\frac{1}{2}-\frac{1}{n}, \frac{1}{2}+\frac{1}{n}\right)}(t) + M_2 \mathbb{1}_{\left[\frac{1}{2}+\frac{1}{n}, 1\right]}(t).$$

Clearly, $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu})$ for all $n \geq 3$ with $\mathbb{T} = \{0, 1\}$ and $\boldsymbol{\mu} = \big(\mathbb{P} \circ M_0^{-1}, \mathbb{P} \circ M_2^{-1}\big)$. However, it follows from Theorem VI.3.21 in Jacod & Shiryaev [27] that, the sequence $(\mathbb{P}_n)_{n\geq 3}$ is not J_1-tight and thus not L^∞-tight.

Definition 3.2 (*S-topology*)**.** The S-topology on Ω is the sequential topology induced by the following S-convergence, i.e. a set $F \subseteq \Omega$ is closed under S-topology if it contains all limits of its S-convergent subsequences, where the S-convergence (denoted by $\xrightarrow{S}$) is defined as follows. Let $(\omega^n)_{n\geq 0} \subseteq \Omega$, we say that $\omega^n \xrightarrow{S} \omega^0$ as $n \to \infty$ if for each $\varepsilon > 0$, we may find a sequence $(v_\varepsilon^n)_{n\geq 0} \subseteq \Omega$ such that

$$v_\varepsilon^n \text{ has bounded variation}, \quad \|\omega^n - v_\varepsilon^n\| \leq \varepsilon \text{ for all } n \geq 0$$

and

$$\lim_{n\to\infty} \int_{[0,1]} f(t) \cdot dv_\varepsilon^n(t) = \int_{[0,1]} f(t) \cdot dv_\varepsilon^0(t) \quad \text{for all } f \in \mathcal{C}([0, 1], \mathbb{R}^d).$$

We denote by $\xrightarrow{S^*}$ the convergence induced by the S-topology.

Remark 3.3. (i) It is shown in Jakubowski [28] that the S-topology is not metrizable. However, its associated Borel σ-field coincides with $\mathcal{F}_1$. In a metric space, a subset is sequentially closed if and only if it is closed; but in a non metrizable space, a sequentially closed set may not be closed. In particular, a sequentially closed set under $\xrightarrow{S^*}$ may not be closed under S-topology (which is equivalent to be sequentially closed under $\xrightarrow{S}$). More precisely, it is

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

shown in Remark 2.6 of [28] that the convergence $\xrightarrow{S^*}$ is weaker than the original one $\xrightarrow{S}$. However, this is not a real problem for our case, since we know, from [28],

$$\omega^n \xrightarrow{S^*} \omega, \text{ if and only if, in every subsequence } (n_k)_{k\geq 1},$$

one may find a further subsequence $(n_{k_l})_{l\geq 1}$ such that $\omega^{n_{k_l}} \xrightarrow{S} \omega$.

In particular, a function $\xi : \Omega \to \mathbb{R}$ is S-continuous (semicontinuous) if and only if ξ is S^*-continuous (semicontinuous).

(ii) The functions $\omega \mapsto \omega_{i,1}$, $\omega \mapsto \int_0^1 \omega_{i,t}dt$ and $\omega \mapsto \int_0^1 |\omega_t|dt$ for $i = 1, \dots, d$ are S-continuous. The functions $\omega \mapsto \|\omega\|$ and $\omega \mapsto \sup_{0\leq t\leq 1} \omega_{i,t}$ for $i = 1, \dots, d$ are S-lower semicontinuous.

Notice that the S-topology is not metrizable, then instead of the usual weak convergence, we use another convergence of probability measures introduced in [28], which induces easy criteria for S-tightness and preserves the Prohorov's theorem, i.e. tightness yields sequential compactness.

Definition 3.4. Let $(\mathbb{P}_n)_{n\geq 1}$ be a sequence of probability measures on the space $(\Omega, \mathcal{F}_1)$. We say $\mathbb{P}_n \overset{*}{\Longrightarrow}_D \mathbb{P}$ if for each subsequence $(\mathbb{P}_{n_k})_{k\geq 1}$, one can find a further subsequence $(\mathbb{P}_{n_{k_l}})_{l\geq 1}$ and stochastic processes $(Y^l)_{l\geq 1}$ and Y defined on the probability space $\big([0,1], \mathcal{B}_{[0,1]}, \ell\big)$ endowed with the Lebesgue measure ℓ, such that $\mathcal{L}(Y^l) = \mathbb{P}_{n_{k_l}}$ for all $l \geq 1$, $\mathcal{L}(Y) = \mathbb{P}$,

$$Y^l(e) \xrightarrow{S^*} Y(e) \quad \text{for all } e \in [0,1],$$

and for each $\varepsilon > 0$, there exists an S^*-compact subset $K_\varepsilon \subseteq \Omega$ such that

$$\inf_{l\geq 1} \mathbb{P}_{n_{k_l}}\big[X \in K_\varepsilon\big] > 1 - \varepsilon.$$

It follows from Jakubowski [28] (see Theorem A.1) that the convergence $\overset{*}{\Longrightarrow}_D$ implies in some sense the convergence of finite dimensional distributions that is specified later, and more importantly, the limit of every convergent sequence of martingale measures is still a martingale measure.

Remark 3.5. Meyer & Zheng [32] have also introduced a topology on Ω (for the case $d = 1$), called pseudo-path topology, by considering the occupation measure induced by every path $\omega \in \Omega$ on $[0,1] \times \mathbb{R}$. We notice that $\omega_n \to \omega_0$ under S-topology induces $\omega_n \to \omega_0$ under the pseudo-path topology, and hence under the pseudo-path topology, it is easier to obtain the relative compactness of a sequence of martingale measures, but one has less continuous functionals defined on Ω. In particular, the simple maps $\omega \mapsto \|\omega\|$, $\omega \mapsto \omega_1$ are not upper semicontinuous, which makes it unsuitable to study the current martingale optimal transport.

We next introduce the Wasserstein distance for the purpose of deriving the duality $\mathsf{P} = \mathsf{D}_1$. Recall the set $\mathcal{P}(\mu, \nu)$ introduced in (1.1).

Definition 3.6. The Wasserstein distance of order 1 is defined by

$$\mathcal{W}_1(\mu, \nu) := \inf_{\mathbb{P}\in\mathcal{P}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}\big[\big|X - Y\big|\big] \quad \text{for all } \mu,\ \nu \in \mathbf{P}.$$

A sequence $(\mu^n)_{n\geq 1} \subseteq \mathbf{P}$ converges to $\mu \in \mathbf{P}$ in $\mathcal{W}_1$ if $\mathcal{W}_1(\mu^n, \mu) \to 0$ as $n \to \infty$ or, equivalently, $\lim_{n\to\infty} \mu^n(\lambda) = \mu(\lambda)$ for all $\lambda \in \Lambda$, see e.g. Theorem 6.9 in Villani [39].

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

For $\left(\boldsymbol{\mu}^n = (\mu_t^n)_{t\in\mathbb{T}}\right)_{n\geq 1} \subseteq \mathbf{P}_{\preceq}^{\mathbb{T}}$ and $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}_{\preceq}^{\mathbb{T}}$, we say that $\boldsymbol{\mu}^n$ converges to $\boldsymbol{\mu}$ if μ_t^n converges to μ_t in $\mathcal{W}_1$ for all $t \in \mathbb{T}$ and this convergence is denoted by $\xrightarrow{\mathcal{W}_1^{\mathbb{T}}}$. We now provide a crucial tightness result for the present paper which is a consequence from [28].

Let $\mathbb{T}_0 \subseteq \mathbb{T}$ be the collection of all condensation points under the lower limit topology, i.e. $t = 1$ or $[t, t+\varepsilon) \cap \mathbb{T}$ is uncountable for any $\varepsilon > 0$.

Lemma 3.7. *Let $(\mathbb{P}_n)_{n\geq 1}$ be a sequence of probability measures such that $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu}^n)$ for some $\boldsymbol{\mu}^n \in \mathbf{P}_{\preceq}^{\mathbb{T}}$, satisfying*

$$\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^{\mathbb{T}}} \boldsymbol{\mu} \in \mathbf{P}_{\preceq}^{\mathbb{T}}. \tag{3.1}$$

(i) *Then, $(\mathbb{P}_n)_{n\geq 1}$ is S-tight, i.e. any subsequence $(\mathbb{P}_{n_k})_{k\geq 1}$ admits a further convergent subsequence under $\overset{*}{\Longrightarrow}_D$. Moreover, any limit point $\mathbb{P}$ of $(\mathbb{P}_n)_{n\geq 1}$ is again a martingale measure.*

(ii) *Assume in addition that $\mathbb{T}_0 = \mathbb{T}$, then $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$.*

Proof. (i) By Theorem A.1, it is clear that $(\mathbb{P}_n)_{n\geq 1}$ is S-tight and there exists a convergent subsequence $(\mathbb{P}_{n_k})_{k\geq 1}$ with limit $\mathbb{P} \in \mathcal{P}$. Moreover, one has a countable subset $\mathcal{T} \subseteq [0,1)$ such that for any finite set $\{u_1, \ldots, u_r\} \subset [0,1] \setminus \mathcal{T}$,

$$\mathbb{P}_{n_k} \circ \left(X_{u_1}, \ldots, X_{u_r}\right)^{-1} \xrightarrow{\mathcal{L}} \mathbb{P} \circ \left(X_{u_1}, \ldots, X_{u_r}\right)^{-1} \quad \text{as } k \to \infty. \tag{3.2}$$

Let $s, t \in [0,1] \setminus \mathcal{T}$ such that $s < t$, and take a finite subset $\{u_1, \ldots, u_r\} \subseteq [0, s] \setminus \mathcal{T}$ and a sequence of bounded continuous functions $\{f_i\}_{1\leq i\leq r}$. Notice that for every $u \in [0,1]$, X_u is uniformly integrable with respect to $(\mathbb{P}_n)_{n\geq 1}$. Indeed,

$$\begin{aligned}
\lim_{R\to\infty} \sup_{n\geq 1} \mathbb{E}^{\mathbb{P}_n}\left[|X_u| 1_{|X_u|\geq R}\right] &\leq \lim_{R\to\infty} \sup_{n\geq 1} \mathbb{E}^{\mathbb{P}_n}\left[\left(|X_u| - R/2\right)_+\right] \\
&\leq \lim_{R\to\infty} \sup_{n\geq 1} \mathbb{E}^{\mathbb{P}_n}\left[\left(|X_1| - R/2\right)_+\right] \\
&= \lim_{R\to\infty} \sup_{n\geq 1} \mu_1^n\left((|x| - R/2)_+\right) = 0.
\end{aligned} \tag{3.3}$$

Combining (3.2) and (3.3), one has

$$\mathbb{E}^{\mathbb{P}}\left[f_1(X_{u_1}) \cdots f_r(X_{u_r})\left(X_t - X_s\right)\right] = \lim_{k\to\infty} \mathbb{E}^{\mathbb{P}_{n_k}}\left[f_1(X_{u_1}) \cdots f_r(X_{u_r})\left(X_t - X_s\right)\right] = 0.$$

Since $\mathcal{T}$ is at most countable, it follows that $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s] = X_s$ for any $s, t \in [0,1] \setminus \mathcal{T}$ such that $s < t$. It follows by the right continuity of X that $\mathbb{P} \in \mathcal{M}$.

(ii) To prove that $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$, it remains to show that $\mathbb{P} \circ X_t^{-1} = \mu_t$ for all $t \in \mathbb{T}$. When $t \in \mathbb{T} \setminus \mathcal{T}$, by the convergence (3.2) and the fact that $\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^{\mathbb{T}}} \boldsymbol{\mu}$, it follows that $\mathbb{P} \circ X_t^{-1} = \mu_t$. Further, notice that $\mathbb{T}_0 = \mathbb{T}$, then for every $t \in \mathbb{T}$, there exists a sequence $(t_i)_{i\geq 1} \subseteq \mathbb{T} \setminus \mathcal{T}$ decreasing to t. Using again the right continuity of X, we conclude $\mathbb{P} \circ X_t^{-1} = \lim_{i\to\infty} \mathbb{P} \circ X_{t_i}^{-1} = \mu_t$. □

As a consequence, the set $\mathcal{M}(\boldsymbol{\mu})$ is S-tight and it is closed if $\mathbb{T}_0 = \mathbb{T}$. The following example shows that the closeness may fail when $\mathbb{T}_0 \neq \mathbb{T}$.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Example 3.8. Let $\mathbb{T} = \{0, 1\}$ and consider a random variable Y such that $\mathbb{P}(Y = 1) = \mathbb{P}(Y = -1) = 1/2$. Define $\mathbb{P}_n := \mathbb{P} \circ (M^n)^{-1}$ for $n \geq 1$, where $M^n = (M^n_t)_{0\leq t\leq 1}$ is defined by

$$M^n_t := Y \mathbb{1}_{\left[\frac{1}{n},1\right]}(t).$$

Define a peacock $\boldsymbol{\mu} = (\mu_0, \mu_1)$ by $\mu_0 := \delta_{\{0\}}$ and $\mu_1 = \big(\delta_{\{-1\}} + \delta_{\{1\}}\big)/2$. Obviously, $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu})$ for all $n \geq 1$. However, the limit of $(\mathbb{P}_n)_{n\geq 1}$ is a martingale measure $\mathbb{P}$ such that $X_t = X_0$, $\mathbb{P}$-a.s. and $\mathbb{P} \circ X_0^{-1} = \mu_1$, which does not lie in $\mathcal{M}(\boldsymbol{\mu})$.

3.2. Finitely-many marginal constraints

We start by studying the finitely-many marginal case and assume throughout this subsection that $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$. Denote $\Delta t_i := t_i - t_{i-1}$ for all $i = 1, \ldots, m$ and $\Delta\mathbb{T} := \min_{1\leq i\leq m} \Delta t_i$. Let us formulate some conditions on the reward function ξ. We shall see later that the usual examples satisfy our conditions.

Assumption 3.9. $\limsup_{n\to\infty} \xi(\omega^n) \leq \xi(\omega)$ holds for all $(\omega^n)_{n\geq 1} \subseteq \Omega$ and $\omega \in \Omega$ such that

$$\omega^n \xrightarrow{S^*} \omega \quad \text{and} \quad \omega^n_{t_i} \longrightarrow \omega_{t_i} \quad \text{for all } i = 0, \ldots, m-1.$$

For $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_m) \in \mathbb{R}^m_+$ such that $|\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$, let $f_{\boldsymbol{\varepsilon}}$ (forward function) and $b_{\boldsymbol{\varepsilon}}$ (backward function) be two non-decreasing functions defined on $[0, 1]$:

$$f_{\boldsymbol{\varepsilon}}(t) := \sum_{i=1}^m \mathbb{1}_{(t_{i-1},t_i]}(t)\Big(t_{i-1} + \frac{\Delta t_i}{\Delta t_i - \varepsilon_i}\big(t - t_{i-1} - \varepsilon_i\big)^+\Big), \tag{3.4}$$

$$b_{\boldsymbol{\varepsilon}}(t) := \sum_{i=1}^m \mathbb{1}_{(t_{i-1},t_i]}(t)\Big(t_i - \Big(\Delta t_i - \frac{\Delta t_i}{\Delta t_i - \varepsilon_i}\big(t - t_{i-1}\big)\Big)^+\Big). \tag{3.5}$$

Assumption 3.10. There is a continuous function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ with $\alpha(0) = 0$ such that the following inequality holds for any $\boldsymbol{\varepsilon} \in \mathbb{R}^m_+$ satisfying $|\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$

$$\big|\xi(\omega) - \xi\big(\omega_{f_{\boldsymbol{\varepsilon}}}\big)\big| \leq \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + \sum_{i=0}^m |\omega_{t_i}| + \int_0^1 |\omega_t| dt\Big), \quad (1)$$

$$\big|\xi(\omega) - \xi\big(\omega_{b_{\boldsymbol{\varepsilon}}}\big)\big| \leq \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + \sum_{i=0}^m |\omega_{t_i}| + \int_0^1 |\omega_t| dt\Big), \quad (2)$$

where $\omega_{f_{\boldsymbol{\varepsilon}}}$ (resp. $\omega_{b_{\boldsymbol{\varepsilon}}}$) denotes the composition of ω and $f_{\boldsymbol{\varepsilon}}$ (resp. $b_{\boldsymbol{\varepsilon}}$).

Theorem 3.11. *Le ξ be bounded from above and satisfies Assumptions 3.9 and 3.10(1). Then for all $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$:*

(i) *The duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu})$ holds.*
(ii) *Assuming further that ξ is bounded, the duality $\mathsf{D}_1(\boldsymbol{\mu}) = \mathsf{D}_2(\boldsymbol{\mu})$ holds for all $\boldsymbol{\mu}$ satisfying (2.5).*

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

To establish the duality $\mathsf{D}_1(\boldsymbol{\mu}) = \mathsf{D}_3(\boldsymbol{\mu})$, we need more regularity conditions on ξ. Define a distance $\rho_{\mathbb{T}}$ on Ω by

$$\rho_{\mathbb{T}}(\omega, \omega') := \sum_{i=1}^{m} \rho_{[t_{i-1},t_i]}(\omega, \omega') + \left| \int_0^1 (\omega_u - \omega'_u) du \right| \quad \text{for all } \omega, \omega' \in \Omega, \tag{3.6}$$

where $\rho_{[s,t]} : \mathbb{D}([s,t], \mathbb{R}^d) \times \mathbb{D}([s,t], \mathbb{R}^d) \to \mathbb{R}_+$ denotes the Skorokhod metric on the space $\mathbb{D}([s,t], \mathbb{R}^d)$. Clearly, $|\omega_{t_i} - \omega'_{t_i}| \le \rho_{\mathbb{T}}(\omega, \omega')$ for all $\omega, \omega' \in \Omega$ and $i = 1, \dots, m$.

Assumption 3.12. ξ is locally $\rho_{\mathbb{T}}$-uniformly continuous, i.e. for every $R > 0$, there exists a continuous increasing function $h_R : \mathbb{R}_+ \to \mathbb{R}_+$ with $h(0) = 0$, such that

$$|\xi(\omega) - \xi(\omega')| \le h_R\left(\rho_{\mathbb{T}}(\omega, \omega')\right) \quad \text{for all } \|\omega\|,\ \|\omega'\| \le R.$$

Theorem 3.13. *Let ξ be bounded and $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$ satisfying (2.5). Then under Assumptions 3.9, 3.10 and 3.12, the duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_3(\boldsymbol{\mu})$ holds.*

Remark 3.14. Using the pathwise Doob's inequality in Acciaio, Beiglböck, Penkner, Schachermayer & Temme [1], the boundedness condition in Theorem 3.13 may be removed when $\mu_1(|x|^p) < +\infty$ for some $p > 1$, see also Dolinsky & Soner [13].

3.3. Infinitely-many marginal constraints

Using approximation techniques, we then obtain some results for the martingale transport problem under infinitely-many marginal constraints.

Proposition 3.15. *Let ξ be S^*-upper semicontinuous and bounded from above. For all $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$:*

(i) *Assume that there exists an increasing sequence of finite sets $\{\mathbb{T}_n\}_{n\ge1}$ such that $1 \in \mathbb{T}_n \subseteq \mathbb{T}$ for all $n \ge 1$ and $\cup_{n\ge1} \mathbb{T}_n$ is dense in $\mathbb{T}$ under the lower limit topology. Then*

$$\lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) = \mathsf{P}(\boldsymbol{\mu}) \quad \text{with } \boldsymbol{\mu}^n := (\mu_t)_{t\in\mathbb{T}_n}.$$

(ii) *Assume $\mathbb{T}_0 = \mathbb{T}$, then there exists an optimal transport plan $\mathbb{P}^* \in \mathcal{M}(\boldsymbol{\mu})$, i.e.*

$$\mathsf{P}(\boldsymbol{\mu}) = \mathbb{E}^{\mathbb{P}^*}[\xi(X)]. \tag{3.7}$$

Proof. (i) It follows by the definition of $\boldsymbol{\mu}^n$ that $\mathsf{P}(\boldsymbol{\mu}^n)$ is non-increasing with respect to n. Take a sequence $(\mathbb{P}_n)_{n\ge1}$ such that $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu}^n)$ and

$$\mathsf{P}(\boldsymbol{\mu}) \le \lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) = \lim_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}[\xi].$$

By Lemma 3.7(i), there is a convergent subsequence $(\mathbb{P}_{n_k})_{k\ge1}$ with some limit $\mathbb{P} \in \mathcal{M}$. It follows by the same arguments in the proof of Lemma 3.7 that $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$ and, the upper semicontinuity of ξ yields

$$\lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) = \lim_{k\to\infty} \mathbb{E}^{\mathbb{P}_{n_k}}[\xi] \le \mathbb{E}^{\mathbb{P}}[\xi] \le \mathsf{P}(\boldsymbol{\mu}).$$

(ii) Take a maximizing sequence $(\mathbb{P}_n)_{n\ge1} \subseteq \mathcal{M}(\boldsymbol{\mu})$, then we may get a limit point $\mathbb{P}^*$ and by Lemma 3.7(ii), $\mathbb{P}^*$ is the required optimal transport plan. □

Consequently, we obtain immediately the dualities for general $\mathbb{T}$ through Proposition 3.15.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Theorem 3.16. *Let ξ be S^*-upper semicontinuous and bounded from above and $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq}$, consider an increasing sequence of finite sets $\{\mathbb{T}_n\}_{n\geq 1}$ such that $\mathbb{T}_n \subseteq \mathbb{T}$, $\cup_{n\geq 1}\mathbb{T}_n$ is dense in $\mathbb{T}$, and set $\boldsymbol{\mu}^n := (\mu_t)_{t\in\mathbb{T}_n}$.*

(i) *Assume that $\mathsf{P}(\boldsymbol{\mu}^n) = \mathsf{D}_1(\boldsymbol{\mu}^n)$ for all $n \geq 1$. Then $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu})$.*
(ii) *Assume further that $\boldsymbol{\mu}$ satisfies (2.5) and $\mathsf{D}_1(\boldsymbol{\mu}^n) = \mathsf{D}_2(\boldsymbol{\mu}^n) = \mathsf{D}_3(\boldsymbol{\mu}^n)$ for all $n \geq 1$. Then $\mathsf{D}_1(\boldsymbol{\mu}) = \mathsf{D}_2(\boldsymbol{\mu}) = \mathsf{D}_3(\boldsymbol{\mu})$.*

Proof. It is enough to show (i). Notice by definition that $\mathsf{D}_1(\boldsymbol{\mu}^n) \geq \mathsf{D}_1(\boldsymbol{\mu})$ for all $n \geq 1$, then it follows by Proposition 3.15(i) that

$$\mathsf{P}(\boldsymbol{\mu}) = \lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) \geq \mathsf{D}_1(\boldsymbol{\mu}).$$

Then the proof is fulfilled by the weak duality $\mathsf{P}(\boldsymbol{\mu}) \leq \mathsf{D}_1(\boldsymbol{\mu})$. □

Remark 3.17. In the present setting, the marginal constraint $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}}$ is given by a family of joint distributions μ_t on $\mathbb{R}^d$. If we replace the probability distribution μ_t by, either d marginal distributions $(\mu_t^1, \ldots, \mu_t^d)$ on $\mathbb{R}$, or a joint distribution $\vec{\mu}_{\vec{t}}$ on $\mathbb{R}^{l\times d}$ for some $\vec{t} := (t_1, \ldots, t_l)$ with $0 \leq t_1 < \cdots < t_l \leq 1$, then all the arguments still hold true and we can obtain similar duality results as in Theorems 3.11, 3.13 and 3.16.

4. The dualities $\mathsf{P} = \mathsf{D}_1 = \mathsf{D}_2$

In the following, we focus on the finite-marginal case, i.e. $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$ and start by proving the first duality. To prove the equality $\mathsf{P} = \mathsf{D}_1$, we shall apply the following well-known result from convex analysis.

Theorem 4.1 (*Fenchel–Moreau*)**.** *Let (E, Σ) be a Hausdorff locally convex space and $F : E \to \mathbb{R}$ be a concave and upper semicontinuous function. Then F is equal to its biconjugate F^{**} which is defined by*

$$F^{**}(e) := \inf_{e^*\in E^*}\Big\{\langle e, e^*\rangle + \sup_{e'\in E}\big(F(e') - \langle e', e^*\rangle\big)\Big\}$$

and E^* denotes the dual space of E.

Next we show that the map $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ is $\mathcal{W}_1^{\mathbb{T}}$-upper semicontinuous and concave and then identify its dual space to be $\Lambda^{\mathbb{T}}$ by $\langle \boldsymbol{\mu}, \boldsymbol{\lambda}\rangle = \boldsymbol{\mu}(\boldsymbol{\lambda})$.

4.1. Space of signed measures on $\mathbb{R}^d$ and its dual space

Let $\mathbf{M}$ denote the space of all finite signed Borel measures μ on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d}\big(1 + |x|\big)|\mu|(dx) < +\infty.$$

It is clear that $\mathbf{M}$ is a linear vector space. We endow $\mathbf{M}$ with a topology (of Wasserstein kind) induced by the following convergence: Let $(\mu^n)_{n\geq 0} \subseteq \mathbf{M}$ be a sequence of bounded signed measures, we say μ^n converges to μ^0 if

$$\lim_{n\to\infty}\int_{\mathbb{R}^d}\lambda(x)\mu^n(dx) = \int_{\mathbb{R}^d}\lambda(x)\mu^0(dx) \quad \text{for all } \lambda \in \Lambda.$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Notice that the above topology restricted on the subspace $\mathbf{P} \subseteq \mathbf{M}$ of probability measures is exactly that induced by the Wasserstein distance. As for the space $\mathbf{M}_0$ of all finite signed Borel measures on $\mathbb{R}^d$ equipped with the weak convergence topology, it is well known that its dual space $\mathbf{M}_0^*$ can be identified as the space of all bounded continuous functions Λ_0, see e.g. Deuschel & Stroock [12]. The following lemma identifies the dual space of $\mathbf{M}$.

Lemma 4.2. *The space* $\mathbf{M}$ *is a Hausdorff locally convex space, and the duality relation*

$$(\lambda, \mu) \in \Lambda \times \mathbf{M} \longmapsto \mu(\lambda)$$

determines a representation of $\mathbf{M}^*$ *as* Λ.

The proof is almost the same as that of $\mathbf{M}_0^* = \Lambda_0$. For completeness, we provide a short proof in the Appendix. For the finite set $\mathbb{T}$, let us endow $\mathbf{M}^{\mathbb{T}}$ with the product topology and obviously, the dual space of $\mathbf{M}^{\mathbb{T}}$ is given by $\Lambda^{\mathbb{T}}$.

4.2. *Proof of the duality* $\mathsf{P} = \mathsf{D}_1$

In preparation for the first duality, we show first the upper semicontinuity of $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ in the context of Theorem 3.11(i). For $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m) \in \mathbb{R}_+^m$ such that $|\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$, we introduce

$$\mathcal{M}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}) := \Big\{\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu}) : X_t = X_{t_k} \text{ on } [t_k, t_k + \varepsilon_k) \text{ for all } k = 0, \dots, m-1,\ \mathbb{P}\text{-a.s.}\Big\}$$

and

$$\mathsf{P}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}) := \sup_{\mathbb{P} \in \mathcal{M}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\big[\xi(X)\big].$$

Proposition 4.3. *Let* ξ *be bounded from above and satisfying Assumptions* 3.9 *and* 3.10(1)*, then* $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ *is* $\mathcal{W}_1^{\mathbb{T}}$*-upper semicontinuous on* $\mathbf{P}_{\preceq}^{\mathbb{T}}$.

Proof. (i) First notice $\mathsf{P}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}) \le \mathsf{P}(\boldsymbol{\mu})$ since $\mathcal{M}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}) \subseteq \mathcal{M}(\boldsymbol{\mu})$. Next, for each $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$, define $\mathbb{P}^{\boldsymbol{\varepsilon}} := \mathbb{P} \circ X_{f_{\boldsymbol{\varepsilon}}}^{-1}$, where $f_{\boldsymbol{\varepsilon}}$ is defined in (3.4). It is clear that $\mathbb{P}^{\boldsymbol{\varepsilon}} \in \mathcal{M}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu})$ and $\mathbb{E}^{\mathbb{P}^{\boldsymbol{\varepsilon}}}[\xi(X)] = \mathbb{E}^{\mathbb{P}}[\xi(X_{f_{\boldsymbol{\varepsilon}}})]$. It follows by Assumption 3.10(1),

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\big[\xi(X)\big] &\le \mathbb{E}^{\mathbb{P}}\big[\xi(X_{f_{\boldsymbol{\varepsilon}}})\big] + \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + (m+2)\mathbb{E}^{\mathbb{P}}\big[|X_1|\big]\Big)\\ &= \mathbb{E}^{\mathbb{P}^{\boldsymbol{\varepsilon}}}\big[\xi(X)\big] + \alpha(|\boldsymbol{\varepsilon}|)\big(1 + (m+2)\mu_1(|x|)\big)\\ &\le \mathsf{P}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}) + \alpha(|\boldsymbol{\varepsilon}|)\big(1 + (m+2)\mu_1(|x|)\big),\end{aligned}$$

which implies that

$$\mathsf{P}(\boldsymbol{\mu}) = \inf_{0<|\boldsymbol{\varepsilon}|<\Delta\mathbb{T}}\Big(\mathsf{P}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}) + \alpha(|\boldsymbol{\varepsilon}|)\big(1 + (m+2)\mu_1(|x|)\big)\Big).$$

(ii) In order to prove that $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ is upper semicontinuous, it suffices to verify that $\boldsymbol{\mu} \mapsto \mathsf{P}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu})$ is upper semicontinuous. To see this, let $(\boldsymbol{\mu}^n)_{n\ge1} \subseteq \mathbf{P}_{\preceq}^{\mathbb{T}}$ be a sequence such that $\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^{\mathbb{T}}} \boldsymbol{\mu} \in \mathbf{P}_{\preceq}^{\mathbb{T}}$. By definition, we have a sequence $(\mathbb{P}_n)_{n\ge1}$ such that $\mathbb{P}_n \in \mathcal{M}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}^n)$ and

$$\limsup_{n\to\infty} \mathsf{P}^{\boldsymbol{\varepsilon}}(\boldsymbol{\mu}^n) = \limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}\big[\xi\big].$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Then one may find a convergent subsequence $(\mathbb{P}_{n_k})_{k\geq 1}$ with limit $\mathbb{P} \in \mathcal{M}$. It follows by exactly the same arguments as in Lemma 3.7 (ii) that $\mathbb{P} \in \mathcal{M}^{\varepsilon}(\mu)$. Since ξ is bounded from above, then it follows from Fatou's lemma that

$$\limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}[\xi] = \lim_{k\to\infty} \mathbb{E}^{\mathbb{P}_{n_k}}[\xi] \leq \mathbb{E}^{\mathbb{P}}[\xi] \leq \mathsf{P}^{\varepsilon}(\mu),$$

which concludes the proof. □

Now we are ready to provide the first duality $\mathsf{P}(\mu) = \mathsf{D}_1(\mu)$. To apply Fenchel–Moreau theorem, we need to embed $\mathbf{P}^{\mathbb{T}}_{\preceq}$ to a locally convex space. Recall that $\mathbf{M}$ is the space of all finite signed measures μ such that

$$\int_{\mathbb{R}^d} \big(1 + |x|\big)|\mu|(dx) < +\infty,$$

and $\mathbf{M}^{\mathbb{T}}$ is its $\mathbb{T}$-product. We then extend the map P from $\mathbf{P}^{\mathbb{T}}_{\preceq}$ to $\mathbf{M}^{\mathbb{T}}$ by

$$\widetilde{\mathsf{P}}(\mu) := \begin{cases} \mathsf{P}(\mu), & \text{if } \mu \in \mathbf{P}^{\mathbb{T}}_{\preceq}, \\ -\infty, & \text{otherwise.} \end{cases}$$

Proof of Theorem 3.11. (i) The concavity of the map $\mu \mapsto \mathsf{P}(\mu)$ is immediate from its definition. Together with the upper semicontinuity of Proposition 4.3, we may directly verify that the extended map $\widetilde{\mathsf{P}}$ is also $\mathcal{W}_1^{\mathbb{T}}$-upper semicontinuous and concave. Then, combining the Fenchel–Moreau theorem and Lemma 4.2, it follows that for all $\mu \in \mathbf{M}^{\mathbb{T}}$,

$$\widetilde{\mathsf{P}}(\mu) = \widetilde{\mathsf{P}}^{**}(\mu),$$

where $\widetilde{\mathsf{P}}^{**}$ denotes the biconjugate of $\widetilde{\mathsf{P}}$. In particular, for $\mu \in \mathbf{P}^{\mathbb{T}}_{\preceq}$ one has

$$\begin{aligned} \mathsf{P}(\mu) &= \widetilde{\mathsf{P}}(\mu) = \widetilde{\mathsf{P}}^{**}(\mu) \\ &= \inf_{\lambda\in\Lambda^{\mathbb{T}}} \big\{\mu(\lambda) - \widetilde{\mathsf{P}}^{*}(\lambda)\big\} = \inf_{\lambda\in\Lambda^{\mathbb{T}}} \Big\{\mu(\lambda) - \inf_{\nu\in\mathbf{M}^{\mathbb{T}}} \big\{\nu(\lambda) - \widetilde{\mathsf{P}}(\nu)\big\}\Big\} \\ &\geq \inf_{\lambda\in\Lambda^{\mathbb{T}}} \Big\{\mu(\lambda) + \sup_{\nu\in\mathbf{P}^{\mathbb{T}}_{\preceq}} \Big\{ \sup_{\mathbb{P}\in\mathcal{M}(\nu)} \mathbb{E}^{\mathbb{P}}\big[\xi - \lambda(X)\big]\Big\}\Big\} \\ &= \inf_{\lambda\in\Lambda^{\mathbb{T}}} \Big\{\mu(\lambda) + \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\big[\xi - \lambda(X)\big]\Big\} = \mathsf{D}_1(\mu) \geq \mathsf{P}(\mu), \end{aligned}$$

which yields $\mathsf{P}(\mu) = \mathsf{D}_1(\mu)$. □

4.3. *Proof of the duality* $\mathsf{D}_1 = \mathsf{D}_2$

For technical reasons, we need to restrict the static strategy λ to a smaller class of functions $\Lambda^{\mathbb{T}}_{lip}$ defined by

$$\Lambda^{\mathbb{T}}_{lip} := \Big\{\lambda = (\lambda_{t_i})_{1\leq i\leq m} \in \Lambda^{\mathbb{T}} : \text{each } \lambda_{t_i} \text{ is boundedly supported and Lipschitz}\Big\}.$$

Proposition 4.4. *Under the conditions of Theorem 3.11(ii) one has*

$$\mathsf{D}_1(\mu) = \inf_{\lambda\in\Lambda^{\mathbb{T}}_{lip}} \Big\{\mu(\lambda) + \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \lambda(X)\big]\Big\}. \tag{4.1}$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Proof. Clearly, by the definition of D_1 and the fact that $\mu_0 = \delta_{\mathbf{1}}(dx)$ and $\mathrm{supp}(\mu_1) \subseteq \mathbb{R}^d_+$, one obtains by interchanging inf and sup that

$$\begin{aligned}\mathsf{D}_1(\boldsymbol{\mu}) &\geq \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \sup_{\mathbb{P}\in\mathcal{M}_+} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) + \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}(X)\big]\Big\} \\ &\geq \sup_{\mathbb{P}\in\mathcal{M}_+} \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) + \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}(X)\big]\Big\} \\ &= \mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu}),\end{aligned}$$

by Theorem 3.11(i). Hence

$$\mathsf{D}_1(\boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) + \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}(X)\big]\Big\}.$$

Next for every $\boldsymbol{\lambda} = (\lambda_{t_i})_{1\leq i\leq m} \in \Lambda^{\mathbb{T}}$, there exists some constant $L > 0$ such that for every $1 \leq i \leq m$,

$$\lambda^L_{t_i}(x) := \lambda_{t_i}(x) - L(1 + \mathbf{1}\cdot x) \leq 0 \quad \text{for all } x \in \mathbb{R}^d_+.$$

Denote $\boldsymbol{\lambda}^L := (\lambda^L_{t_i})_{1\leq i\leq m}$, then for every martingale measure $\mathbb{P} \in \mathcal{M}_+$, we have

$$\boldsymbol{\mu}(\boldsymbol{\lambda}) + \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}(X)\big] = \boldsymbol{\mu}(\boldsymbol{\lambda}^L) + \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^L(X)\big].$$

Further, for each $R > 0$, let $\psi_R : \mathbb{R}^d \to [0, 1]$ be some continuous function such that

$$\psi_R(x) = 1 \quad \text{whenever } |x| \leq R \quad \text{and} \quad \psi_R(x) = 0 \quad \text{whenever } |x| > R + 1.$$

Let $\boldsymbol{\lambda}^{L,R} := \big(\lambda^{L,R}_{t_i}\big)_{1\leq i\leq m}$ with $\lambda^{L,R}_{t_i}(x) := \lambda^L_{t_i}(x)\psi_R(x) \geq \lambda^L_{t_i}(x)$, then

$$\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^{L,R}(X)\big] \leq \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^L(X)\big].$$

On the other hand, for all $\mathbb{P} \in \mathcal{M}_+$ we have by monotone convergence theorem

$$\lim_{R\to\infty} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^{L,R}(X)\big] = \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^L(X)\big].$$

Hence

$$\lim_{R\to\infty} \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^{L,R}(X)\big] = \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^L(X)\big].$$

It follows that

$$\begin{aligned}\lim_{R\to\infty}\Big(\boldsymbol{\mu}(\boldsymbol{\lambda}^{L,R}) + \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^{L,R}(X)\big]\Big) &= \boldsymbol{\mu}(\boldsymbol{\lambda}^L) + \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}^L(X)\big] \\ &= \boldsymbol{\mu}(\boldsymbol{\lambda}) + \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi - \boldsymbol{\lambda}(X)\big].\end{aligned}$$

Finally, by a convolution argument each $\lambda^{L,R}_{t_i}$ can be approximated uniformly by some Lipschitz function that is also boundedly supported, which yields the required result. □

For all $(\omega, t) \in \Omega_+ \times [0, 1]$, denote by $\mathcal{B}^{sem}_{\omega,t} \subseteq \mathcal{P}$ the set of probability measures $\mathbb{P}$ such that $\mathbb{P}\big[X_s = \omega_s \text{ for all } 0 \leq s \leq t\big] = 1$ and $(X_s)_{s\geq t}$ is a non-negative semimartingale under $\mathbb{P}$. Denote further

$$\mathcal{M}^{loc}_{\omega,t} := \Big\{\mathbb{P} \in \mathcal{B}^{sem}_{\omega,t} : (X_s)_{s\geq t} \text{ is a local martingale under } \mathbb{P}\Big\}.$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Write in particular $\mathcal{B}^{sem} = \mathcal{B}^{sem}_{\omega,0}$ and $\mathcal{M}^{loc} = \mathcal{M}^{loc}_{\omega,0}$. Let $\zeta : \Omega \to \mathbb{R}$ be a measurable function and put

$$V_t(\omega) := \sup_{\mathbb{P}\in\mathcal{M}^{loc}_{\omega,t}} \mathbb{E}^{\mathbb{P}}\big[\zeta(X)\big]. \tag{4.2}$$

Our objective now is to show that the process $(V_t)_{0\le t\le 1}$ is $\mathbb{F}^U$-adapted and that the dynamic programming principle holds. To achieve this, we use the related results in Neufeld & Nutz [33,34]. Let $\mathbb{P} \in \mathcal{B}^{sem}$ be a semimartingale measure with the triplet $(B^{\mathbb{P}}, C^{\mathbb{P}}, \nu^{\mathbb{P}})$ of predictable semimartingale characteristics, see e.g. Chapter II of Jacod & Shiryaev [27]. Notice that

$$\mathcal{M}^{loc} = \Big\{\mathbb{P} \in \mathcal{B}^{sem} : B^{\mathbb{P}}_t = 0 \text{ for all } t \in [0,1]\Big\}.$$

By Theorem 2.5 in [33], the map $\mathbb{P} \mapsto (B^{\mathbb{P}}, C^{\mathbb{P}}, \nu^{\mathbb{P}})$ is measurable, then it follows that $\mathcal{M}^{loc}$ is Borel. Moreover, by the same arguments we have the following lemma.

Lemma 4.5. *The set* $\big\{(\omega, t, \mathbb{P}) \in \Omega \times [0,1] \times \mathcal{P}(\Omega) : \mathbb{P} \in \mathcal{M}^{loc}_{\omega,t}\big\}$ *is Borel.*

By Theorem 2.1 in [34], we have the following lemma.

Lemma 4.6. *Let* $\mathbb{P} \in \mathcal{M}^{loc}_{\omega,t}$ *and* τ *be an* $\mathbb{F}$*-stopping time taking values in* $[t, 1]$.

(i) *There is a family of conditional probability* $(\mathbb{P}_\omega)_{\omega\in\Omega}$ *of* $\mathbb{P}$ *with respect to* $\mathcal{F}_\tau$ *such that* $\mathbb{P}_\omega \in \mathcal{M}^{loc}_{\omega,\tau(\omega)}$ *for* $\mathbb{P}$*-a.e.* $\omega \in \Omega$.

(ii) *Assume that there exists a family of probability measures* $(\mathbb{Q}_\omega)_{\omega\in\Omega_+}$ *such that*

$$\mathbb{Q}_\omega \in \mathcal{M}^{loc}_{\omega,\tau(\omega)} \text{ for } \mathbb{P}\text{-a.e. } \omega \in \Omega, \text{ and the map } \omega \mapsto \mathbb{Q}_\omega \text{ is } \mathcal{F}_\tau\text{-measurable},$$

then $\mathbb{P} \otimes \mathbb{Q} \in \mathcal{M}^{loc}_{\omega,t}$, *where*

$$\mathbb{P} \otimes \mathbb{Q}(\cdot) := \int_\Omega \mathbb{Q}_\omega(\cdot)\mathbb{P}(d\omega).$$

The dynamic programming principle follows by Lemmas 4.5 and 4.6, and as a consequence we have the following proposition.

Proposition 4.7. *Assume that* ζ *is bounded, then the process* $V = (V_t)_{0\le t\le 1}$ *defined in* (4.2) *is a* $\mathcal{M}_+$*-supermartingale, i.e.* $V \in \mathcal{S}$.

Proposition 4.8. *Let* ζ *be a measurable and bounded function, then one has*

$$\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}[\zeta(X)] = \inf\Big\{V_0 : (V_t)_{0\le t\le 1} \in \mathcal{S} \text{ such that } V_1(\omega) \ge \zeta(\omega) \text{ for all } \omega \in \Omega_+\Big\}.$$

Proof. By Proposition 4.7 with the process V defined in (4.2), it remains to show that

$$\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}[\zeta(X)] = \sup_{\mathbb{P}\in\mathcal{M}^{loc}} \mathbb{E}^{\mathbb{P}}[\zeta(X)].$$

It is clear that $\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}[\zeta(X)] \le \sup_{\mathbb{P}\in\mathcal{M}^{loc}} \mathbb{E}^{\mathbb{P}}[\zeta(X)]$ since $\mathcal{M}_+ \subseteq \mathcal{M}^{loc}$, then it suffices to prove the converse inequality. For each $\mathbb{P} \in \mathcal{M}^{loc}$, there exists an increasing sequence of

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

stopping times $(\sigma_n)_{n\geq 1}$ such that $\sigma_n \to +\infty$, $\mathbb{P}$-almost surely and $X_{\sigma_n\wedge\cdot}$ is a $\mathbb{P}$-martingale, where $X_{\sigma_n\wedge\cdot} := (X_{\sigma_n\wedge t})_{0\leq t\leq 1}$. Hence

$$\mathbb{E}^{\mathbb{P}}[\zeta(X_{\sigma_n\wedge\cdot})] \leq \sup_{\mathbb{Q}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{Q}}[\zeta(X)].$$

The required result follows from the dominated convergence theorem. □

Proof of Theorem 3.11. (ii) It remains to show $\mathsf{D}_1(\boldsymbol{\mu}) \geq \mathsf{D}_2(\boldsymbol{\mu})$. Indeed, one has by Proposition 4.4,

$$\mathsf{D}_1(\boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}_{lip}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) + \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\big[\xi(X) - \boldsymbol{\lambda}(X)\big]\Big\}.$$

For each $\varepsilon > 0$, by Proposition 4.8 there exist a vector $\boldsymbol{\lambda}^\varepsilon \in \Lambda^{\mathbb{T}}_{lip}$ and a process $V^\varepsilon = (V^\varepsilon_t)_{0\leq t\leq 1} \in \mathcal{S}$ such that

$$\mathsf{D}_1(\boldsymbol{\mu}) + \varepsilon \geq \boldsymbol{\mu}(\boldsymbol{\lambda}^\varepsilon) + V^\varepsilon_0 \quad \text{and} \quad V^\varepsilon_1(\omega) \geq \xi(\omega) - \boldsymbol{\lambda}^\varepsilon(\omega).$$

This implies that $\mathsf{D}_1(\boldsymbol{\mu}) + \varepsilon \geq \mathsf{D}_2(\boldsymbol{\mu})$, and the required result by the arbitrariness of ε. □

5. Proof of the duality $\mathsf{D}_1 = \mathsf{D}_3$

Now let us turn to prove the third duality $\mathsf{D}_1 = \mathsf{D}_3$ in Theorem 3.13. We will follow the idea in Dolinsky & Soner [13] to discretize the underlying paths and then use the classical constrained duality result of Föllmer & Kramkov [15]. The proof in [13] relies on the min–max theorem and the explicit approximation of a martingale measure. We emphasize that the present proof is less technically involved than [13] as the marginals constraints have already been reduced by the first duality.

5.1. Reduction of ξ to be boundedly supported

In this section we denote $\mathsf{P}(\boldsymbol{\mu}, \xi)$ and $\mathsf{D}_3(\boldsymbol{\mu}, \xi)$ in place of $\mathsf{P}(\boldsymbol{\mu})$ and $\mathsf{D}_3(\boldsymbol{\mu})$ to emphasize the dependence on ξ, then clearly for any $\xi, \xi' : \Omega \to \mathbb{R}$ and $c \in \mathbb{R}$, one has

$$\mathsf{D}_3(\boldsymbol{\mu}, \xi + \xi') \leq \mathsf{D}_3(\boldsymbol{\mu}, \xi) + \mathsf{D}_3(\boldsymbol{\mu}, \xi') \quad \text{and} \quad \mathsf{D}_3(\boldsymbol{\mu}, \xi + c) = \mathsf{D}_3(\boldsymbol{\mu}, \xi) + c.$$

In particular for $c > 0$ one has

$$\mathsf{D}_3(\boldsymbol{\mu}, c\xi) = c\mathsf{D}_3(\boldsymbol{\mu}, \xi).$$

Hence, under the conditions of Theorem 3.13, we may assume that $0 \leq \xi \leq 1$. Indeed, we show next that it suffices to establish the duality $\mathsf{P}(\boldsymbol{\mu}, \xi) = \mathsf{D}_3(\boldsymbol{\mu}, \xi)$ for ξ that is boundedly supported. For all $R > 0$, define the continuous function $\chi_R : \mathbb{R}_+ \to [0, 1]$ by

$$\chi_R(x) := \mathbb{1}_{[0,R]}(x) + (R + 1 - x)\mathbb{1}_{(R,R+1]}(x) \quad \text{for all } x \in \mathbb{R}_+.$$

Denote further for $R > 0$

$$\xi_R(\omega) := \xi(\omega)\chi_R(\|\omega\|) \quad \text{for all } \omega \in \Omega.$$

Notice that $0 \leq \xi \leq 1$ yields $\xi_R(\omega) \leq \xi(\omega) \leq \xi_R(\omega) + \mathbb{1}_{\{\|\omega\|\geq R\}}$, then it follows that

$$\mathsf{D}_3(\boldsymbol{\mu}, \xi_R) \leq \mathsf{D}_3(\boldsymbol{\mu}, \xi) \leq \mathsf{D}_3(\boldsymbol{\mu}, \xi_R) + \mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X\|\geq R\}}). \tag{5.1}$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Lemma 5.1. *Let ξ be bounded and $\boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T},+}_{\preceq}$. Then*

$$\mathsf{D}_3(\boldsymbol{\mu}, \xi) = \lim_{R\to\infty} \mathsf{D}_3(\boldsymbol{\mu}, \xi_R).$$

Proof. It is enough to prove by (5.1) that

$$\lim_{R\to+\infty} \mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X\|\geq R\}}) = 0.$$

This is indeed a direct consequence of the pathwise inequality, see e.g. Lemma 2.3 of Brown, Hobson and Rogers [5]

$$\mathbb{1}_{\{\|X_i\|\geq R\}} \leq \frac{(|X_{i,1}| - K)^+}{R-K} + \mathbb{1}_{\{\|X_i\|\geq R\}} \frac{R - X_{i,1}}{R-K} \quad \text{for all } i = 1, \dots, d$$

holds for every $0 < K < R$. It follows by taking $K = R/2d$ that

$$\mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X\|\geq R\}}) \leq \sum_{i=1}^{d} \mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X_i\|\geq R/d\}}) \leq \frac{2d}{R} \sum_{i=1}^{d} \mu_1\left(\left(x_i - \frac{R}{2d}\right)^+\right).$$

The proof is fulfilled by letting $R \to +\infty$. □

Next we show that ξ_R inherits almost the same properties as ξ.

Lemma 5.2. *For each $R > 0$:*

(i) *If ξ satisfies Assumptions 3.9 and 3.10, then so does ξ_R.*
(ii) *If ξ satisfies Assumption 3.12, then*

$$\begin{aligned} &\xi_R \text{ is } L^\infty\text{-uniformly continuous, and} \\ &\xi_R(\omega) - \xi_R(\omega') \leq \beta\big(\rho_{\mathbb{T}}(\omega, \omega')\big) \quad \text{for all } \omega, \omega' \in \Omega \text{ such that } \|\omega'\| \leq \|\omega\| \end{aligned} \tag{5.2}$$

for some continuous increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\beta(0) = 0$.

Proof. (i) follows by the fact that $\omega \mapsto \|\omega\|$ is S^*-lower semicontinuous and $\|\omega_{f_\varepsilon}\| = \|\omega_{b_\varepsilon}\| = \|\omega\|$. Let us turn to show (ii). Notice that ξ is $\rho_{\mathbb{T}}$-uniformly continuous on $\big\{\omega : \|\omega\| \leq R\big\}$, i.e. there exists a continuous increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\beta(0) = 0$ such that for all $\|\omega\|, \|\omega'\| \leq R$

$$|\xi(\omega) - \xi(\omega')| \leq \beta\big(\rho_{\mathbb{T}}(\omega, \omega')\big).$$

Hence, for any $\omega, \omega' \in \Omega$ such that $\|\omega'\| \leq \|\omega\|$, one has

$$\xi_R(\omega) - \xi_R(\omega') \leq \mathbb{1}_{\{\|\omega\|\leq R\}}\big(\xi(\omega) - \xi(\omega')\big) \leq \beta\big(\rho_{\mathbb{T}}(\omega, \omega')\big).$$

Moreover,

$$\begin{aligned} |\xi_R(\omega') - \xi_R(\omega)| &\leq |\xi(\omega') - \xi(\omega)|\chi_R(\|\omega\|) + |\chi_R(\|\omega\|) - \chi_R(\|\omega'\|)| \\ &\leq \beta\big(\rho_{\mathbb{T}}(\omega, \omega')\big) + \|\omega - \omega'\| \\ &\leq \beta\big(2\|\omega - \omega'\|\big) + \|\omega - \omega'\|, \end{aligned}$$

which yields the L^∞-uniform continuity of ξ_R. □

Therefore, in the following it suffices to consider the function ξ that is boundedly supported such that Assumptions 3.9 and 3.10 and Condition (5.2) hold. Similar to the proof of the duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_2(\boldsymbol{\mu})$, it remains to prove a duality without marginal constraints.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

5.2. Duality without marginal constraints

We consider in this section the optimization problem without marginal constraints. Let $\zeta : \Omega \to \mathbb{R}$ be bounded and define

$$\mathsf{P}(\zeta) := \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}[\zeta(X)] \quad \text{and} \quad \mathsf{D}(\zeta) := \inf_{(z,H)\in\mathcal{D}(\zeta)} z \tag{5.3}$$

where, with the same definition of integral in (2.7),

$$\mathcal{D}(\zeta) := \Big\{(z, H) \in \mathbb{R} \times \mathcal{A} : z + (H \cdot X)_1(\omega) \geq \zeta(\omega) \text{ for all } \omega \in \Omega_+\Big\}.$$

We provide immediately a duality result for the above optimization problems, and leave its proof in Section 5.3.

Theorem 5.3. *Suppose that ζ satisfies Assumptions 3.9, 3.10 and Condition (5.2), then*

$$\mathsf{P}(\zeta) = \mathsf{D}(\zeta). \tag{5.4}$$

By exactly the same arguments as in the proof of Theorem 3.11(ii), the duality $\mathsf{P}(\boldsymbol{\mu}, \xi_R) = \mathsf{D}_3(\boldsymbol{\mu}, \xi_R)$ follows immediately by taking $\zeta = \xi_R - \boldsymbol{\lambda}$ in Theorem 5.3.

Proof of Theorem 3.13. Using Lemma 5.1 as well as the first duality $\mathsf{P} = \mathsf{D}_1$ for ξ_R, one has

$$\mathsf{P}(\boldsymbol{\mu}, \xi) \geq \lim_{R\to\infty} \mathsf{P}(\boldsymbol{\mu}, \xi_R) = \lim_{R\to\infty} \mathsf{D}_3(\boldsymbol{\mu}, \xi_R) = \mathsf{D}_3(\boldsymbol{\mu}, \xi).$$

Hence we conclude the proof by the weak duality $\mathsf{P}(\boldsymbol{\mu}, \xi) \leq \mathsf{D}_3(\boldsymbol{\mu}, \xi)$. □

5.3. Proof of Theorem 5.3

Recall that $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$, $\Delta t_i = t_i - t_{i-1}$ for $i = 1, \ldots, m$ and $\Delta\mathbb{T} = \min_{1\leq i\leq m} \Delta t_i$. Let $\zeta : \Omega \to \mathbb{R}$ be measurable and boundedly supported. Then for each $0 \leq \delta < \Delta\mathbb{T}$, denote $\Omega^\delta := \mathbb{D}([0, 1 + \delta], \mathbb{R}^d)$ and all its elements by ω^δ. Put $\mathbb{T}^\delta := \{0 = t_0^\delta < \cdots < t_m^\delta = 1 + \delta\}$, where $t_i^\delta := k_\delta t_i$ for all $i = 0, \ldots, m$ with $k_\delta := 1 + \delta$. Define $\zeta^\delta : \Omega^\delta \to \mathbb{R}$ by

$$\zeta^\delta(\omega^\delta) := \zeta(\bar{\omega}^\delta), \quad \text{where } \bar{\omega}^\delta \in \Omega \text{ is defined by } \bar{\omega}_t^\delta := \omega_{k_\delta t}^\delta \text{ for all } t \in [0, 1]. \tag{5.5}$$

Proposition 5.4. *Assume that ζ satisfies Assumptions 3.9, 3.10 and Condition (5.2). Then:*

(i) *For all $0 \leq \delta < \Delta\mathbb{T}$, the ζ^δ defined by (5.5) satisfies Assumptions 3.9, 3.10 and Condition (5.2).*

(ii) *There is a continuous function $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\eta(0) = 0$ such that for all $0 \leq \delta < \delta' < \Delta\mathbb{T}$ the following inequality holds*

$$\Big|\zeta^\delta(\omega^\delta) - \zeta^{\delta'}(\omega^{\delta',\delta})\Big| \leq \eta\Big(\frac{\delta' - \delta}{1 + \delta'}\Big)\Big(1 + \sum_{i=0}^{m} |\omega^\delta_{t_i^\delta}| + \int_0^{1+\delta} |\omega_t^\delta| dt\Big) \quad \text{for all } \omega^\delta \in \Omega^\delta,$$

where $\omega^{\delta',\delta} \in \Omega^{\delta'}$ is defined by

$$\omega_t^{\delta',\delta} := \omega^\delta_{(t - t_i^{\delta'} + t_i^\delta)\wedge t_{i+1}^\delta} \quad \text{for all } t \in [t_i^{\delta'}, t_{i+1}^{\delta'}] \text{ and } i = 0, \ldots, m - 1.$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Proof. (i) will be proved in Lemmas 5.10–5.12 in Section 5.5.
(ii) Clearly, $\zeta^{\delta'}(\omega^{\delta',\delta}) = \zeta(\bar{\omega}^{\delta',\delta})$, where

$$\bar{\omega}_t^{\delta',\delta} := \omega^{\delta',\delta}_{k_{\delta' t}} \quad \text{for all } t \in [0,1].$$

Direct computation reveals that $\bar{\omega}^{\delta,'\delta} = \bar{\omega}^{\delta} \circ b_{\boldsymbol{\varepsilon}}$ with

$$\boldsymbol{\varepsilon} := \frac{\delta' - \delta}{1+\delta'}(\Delta t_1, \ldots, \Delta t_m).$$

Hence by Assumption 3.10 one obtains

$$\begin{aligned}\left|\zeta^{\delta}(\omega^{\delta}) - \zeta^{\delta'}(\omega^{\delta',\delta})\right| &= \left|\zeta(\bar{\omega}^{\delta}) - \zeta(\bar{\omega}^{\delta',\delta})\right| \\ &\le \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + \sum_{i=0}^{m} |\bar{\omega}^{\delta}_{t_i}| + \int_0^1 |\bar{\omega}^{\delta}_t| dt\Big) \\ &\le \alpha\Big(|\Delta\mathbb{T}|\frac{\delta'-\delta}{1+\delta'}\Big)\Big(1 + \sum_{i=0}^{m} |\omega^{\delta}_{t_i^{\delta}}| + \int_0^{1+\delta} |\omega^{\delta}_t| dt\Big).\end{aligned}$$

The proof is completed by taking $\eta(\cdot) = \alpha(\Delta\mathbb{T} \times \cdot)$. □

We are now ready to prove the required duality. Define

$$\Omega_+^{\delta} := \big\{\omega^{\delta} \in \Omega^{\delta} : \omega_0^{\delta} = \mathbf{1} \text{ and } \omega_t^{\delta} \in \mathbb{R}^d_+ \text{ for all } t \in [0, 1+\delta]\big\}$$

and the corresponding martingale optimal transport problem

$$\mathsf{P}_{\delta} := \sup_{\mathbb{P} \in \mathcal{M}_+^{\delta}} \mathbb{E}^{\mathbb{P}}[\zeta^{\delta}(X^{\delta})],$$

where similarly, $X^{\delta} = (X_t^{\delta})_{0 \le t \le k_{\delta}}$ denotes the canonical process and $\mathcal{M}_+^{\delta}$ denotes the set of martingale measures supported on Ω_+^{δ}. The dual problem is slightly different. Denote further

$$\Omega_+^{c,\delta} := \big\{\omega^{\delta} \in \Omega_+^{\delta} : \omega^{\delta}_{t_i^{\delta}-} = \omega^{\delta}_{t_i^{\delta}} \text{ for all } i = 1, \ldots, m\big\}$$

and define the dual problem by

$$\mathsf{D}_{\delta}^{c} := \inf_{(z,H) \in \mathcal{D}_{\delta}^{c}} z,$$

with $\mathcal{D}_{\delta}^{c}$ given by

$$\mathcal{D}_{\delta}^{c} := \Big\{(z^{\delta}, H^{\delta}) \in \mathbb{R} \times \mathcal{A}^{\delta} : z^{\delta} + (H^{\delta} \cdot \omega^{\delta})_1 \ge \zeta^{\delta}(\omega^{\delta}) \text{ for all } \omega^{\delta} \in \Omega_+^{c,\delta}\Big\},$$

where, similarly to Definition 2.5, $\mathcal{A}^{\delta}$ denotes the collection of all left-continuous adapted processes with bounded variation such that the stochastic integral $(H^{\delta} \cdot X^{\delta})$ is a supermartingale under all probability measures in $\mathcal{M}_+^{\delta}$.

The main technical step for our result is the following.

Lemma 5.5. *Suppose that ζ satisfies Assumptions 3.9, 3.10 and Condition (5.2). Then*

$$\mathsf{D}_{\delta}^{c} \le \mathsf{P}_{\delta} \quad \textit{for all } \delta \ge 0. \tag{5.6}$$

The proof of Lemma 5.5 is adapted from Dolinsky & Soner [13] and is reported in Section 5.4.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Lemma 5.6. *Suppose that ζ satisfies Assumptions 3.9, 3.10 and Condition (5.2). Then*

$$\liminf_{\delta\downarrow 0} \mathsf{D}^c_\delta \geq \mathsf{D}(\zeta) \quad \text{and} \quad \limsup_{\delta\downarrow 0} \mathsf{P}_\delta \leq \mathsf{P}(\zeta). \tag{5.7}$$

Proof. (i) For each $(z^\delta, H^\delta) \in \mathcal{D}^c_\delta$ with $\delta > 0$ let us construct a robust super-replication on Ω_+. For any $\omega \in \Omega_+$ define $H_0(\omega) = H^\delta_0(\omega^{\delta,0})$ and

$$H_t(\omega) := H^\delta_{t-t_i+t^\delta_i}(\omega^{\delta,0}) \quad \text{for all } t \in (t_i, t_{i+1}] \text{ and } i = 0, \dots, m-1,$$

where $\omega^{\delta,0} \in \Omega^{c,\delta}_+$ is defined as before by

$$\omega^{\delta,0}_t = \omega_{(t-t^\delta_i+t_i)\wedge t_{i+1}} \quad \text{for all } t \in [t^\delta_i, t^\delta_{i+1}] \text{ and } i = 0, \dots, m-1.$$

It is clear that H is $\mathbb{F}$-adapted, left-continuous, with bounded variation, and $(H \cdot X)$ is a supermartingale under every $\mathbb{P} \in \mathcal{M}_+$, hence $H \in \mathcal{A}$. Moreover,

$$z^\delta + \big(H^\delta \cdot \omega^{\delta,0}\big)_{1+\delta} \geq \zeta^\delta(\omega^{\delta,0}) \quad \text{for all } \omega \in \Omega_+. \tag{5.8}$$

Notice that $\big(H^\delta \cdot \omega^{\delta,0}\big)_{1+\delta} = (H \cdot \omega)_1$, thus we obtain by Assumption 3.10 and Condition (5.8)

$$z^\delta + (H \cdot \omega)_1 \geq \zeta(\omega) - \eta\Big(\frac{\delta}{1+\delta}\Big)\Big(1 + \sum_{i=0}^m |\omega_{t_i}| + \int_0^1 |\omega_t| dt\Big) \quad \text{for all } \omega \in \Omega_+,$$

which yields $\mathsf{D}^c_\delta + \big(1 + (m+2)d\big)\eta\big(\frac{\delta}{1+\delta}\big) \geq \mathsf{D}(\zeta)$ and therefore

$$\liminf_{\delta\downarrow 0} \mathsf{D}^c_\delta \geq \mathsf{D}(\zeta).$$

(ii) Let $(\delta_n)_{n\geq 1}$ be such that $\delta_n > 0$ and $\delta_n \downarrow 0$. Then there is a sequence $(\mathbb{P}_n)_{n\geq 1}$ such that

$$\limsup_{n\to\infty} \mathsf{P}_{\delta_n} = \limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}\big[\zeta^{\delta_n}(X^{\delta_n})\big].$$

For any fixed $\delta_0 > 0$, we assume without loss of generality that $\delta_n \leq \delta_0$ for all $n \geq 1$. Then for each $n \geq 1$, let us define $\tilde{\mathbb{P}}_n := \mathbb{P}_n \circ \big(\tilde{X}^{\delta_n}\big)^{-1}$ where $\tilde{X}^{\delta_n}(\omega^{\delta_n}) := X^{\delta_0}(\omega^{\delta_0,\delta_n})$ is the extended process from Ω^{δ_n} to Ω^{δ_0}. It follows by Proposition 5.4(ii) that

$$\begin{aligned}\mathbb{E}^{\mathbb{P}_n}\big[\zeta^{\delta_n}(X^{\delta_n})\big] &\leq \big(1 + (m+2)d\big)\eta\Big(\frac{\delta_0 - \delta_n}{1+\delta_0}\Big) + \mathbb{E}^{\mathbb{P}_n}\big[\zeta^{\delta_0}(\tilde{X}^{\delta_n})\big] \\ &= \big(1 + (m+2)d\big)\eta\Big(\frac{\delta_0 - \delta_n}{1+\delta_0}\Big) + \mathbb{E}^{\tilde{\mathbb{P}}_n}\big[\zeta^{\delta_0}(X^{\delta_0})\big].\end{aligned}$$

Again by the same argument in Proposition 4.3 we obtain

$$\limsup_{n\to\infty} \mathsf{P}_{\delta_n} \leq 2\big(1 + (m+2)d\big)\eta\Big(\frac{\delta_0}{1+\delta_0}\Big) + \mathsf{P}(\zeta)$$

which yields the required result since $\delta_0 > 0$ is arbitrary. □

Proof of Theorem 5.3. Let $(z, H) \in \mathcal{D}(\zeta)$, we know by definition $z + (H \cdot \omega)_1 \geq \zeta(\omega)$, $\forall \omega \in \Omega_+$. Taking expectation over each sides, it follows that

$$z \geq \mathbb{E}^{\mathbb{P}}[\zeta(X)] \quad \text{for all } \mathbb{P} \in \mathcal{M}_+.$$

Then we get the weak duality $\mathsf{P}(\zeta) \leq \mathsf{D}(\zeta)$. The reverse inequality follows by Lemmas 5.5 and 5.6. □

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

5.4. Proof of Lemma 5.5

The arguments are mainly adapted from Dolinsky & Soner [13] and the main idea is to discretize the paths on the Skorokhod space. By Proposition 5.4(i), the proof of $\mathsf{D}^c_\delta \le \mathsf{P}_\delta$ is not altered by the value of δ. We therefore consider $\delta = 0$ in this subsection.

5.4.1. A probabilistic hedging problem

For all $n \in \mathbb{N}$, put

$$A^{(n)} := \left\{2^{-n}q : q \in \mathbb{N}^d\right\} \quad \text{and} \quad B^{(n)} := \left\{i\sqrt{d}2^{-n} : i \in \mathbb{N}\right\} \cup \left\{\sqrt{d}2^{-n}/j : j \in \mathbb{N}^*\right\}.$$

We then define a subspace $\hat{\Omega} := \hat{\Omega}^{(n)} \subseteq \Omega_+$ as follows.

Definition 5.7. A path $\omega \in \Omega_+$ belongs to $\hat{\Omega}$ if there exist non-negative integers $0 = K_0 < K_1 + 1 < \cdots < K_m + m$ and a partition $\left\{0 = \hat{\tau}_0 < \hat{\tau}_1 < \cdots < \hat{\tau}_{K_m+m} = 1\right\}$ such that $\hat{\tau}_{K_i+i} = t_i$ for $1 \le i \le m$ and

$$\omega_t = \sum_{i=0}^{m-1}\Big(\sum_{k=K_i+i}^{K_{i+1}+i-1} \omega_{\hat{\tau}_k}\mathbb{1}_{[\hat{\tau}_k,\hat{\tau}_{k+1})}(t) + \omega_{t_{i+1}}\mathbb{1}_{[\hat{\tau}_{K_{i+1}+i},t_{i+1})}(t)\Big) + \omega_1\mathbb{1}_{\{t=1\}},$$

where $\omega_{t_i} \in A^{(n)}$ for $1 \le i \le m$ and for $0 \le i < m$

$$\omega_{\hat{\tau}_k} \in A^{(n+k-K_i-i)}, \quad K_i + i < k < K_{i+1} + i + 1,$$
$$\hat{\tau}_k - \hat{\tau}_{k-1} \in B^{(n+k-K_i-i)}, \quad K_i + i < k < K_{i+1} + i + 1.$$

Notice that $\hat{\Omega}$ is countable, then there exists a probability measure $\hat{\mathbb{P}} := \hat{\mathbb{P}}^{(n)}$ on Ω_+ supported on $\hat{\Omega}$ which gives positive weight to every element of $\hat{\Omega}$. In particular, the canonical process X has finitely many jumps $\hat{\mathbb{P}}$-almost surely. Denote by $\hat{\mathbb{F}}$ the completed filtration of $\mathbb{F}$ under $\hat{\mathbb{P}}$. Put

$$\hat{\mathcal{H}}^{(n)} := \left\{\hat{H} : [0,1] \times \Omega_+ \to \mathbb{R}^d \text{ is } \hat{\mathbb{F}}\text{-predictable such that } \|\hat{H}\| \le n\right\}$$

and

$$\hat{\mathcal{A}}^{(n)} := \left\{\hat{H} \in \hat{\mathcal{H}}^{(n)} : (\hat{H} \cdot X)_t \ge K \text{ for all } t \in [0,1],\ \hat{\mathbb{P}}\text{-a.s. for some } K \in \mathbb{R}\right\}.$$

Let

$$\hat{\mathcal{D}}^{(n)}(\zeta) := \left\{(z, \hat{H}) \in \mathbb{R} \times \hat{\mathcal{A}}^{(n)} : z + (\hat{H} \cdot X)_1 \ge \zeta(X),\ \hat{\mathbb{P}}\text{-a.s.}\right\}$$

and define the robust superhedging problem under the dominating measure $\hat{\mathbb{P}}$

$$\mathsf{D}^{(n)}(\zeta) := \inf_{(z,\hat{H})\in\hat{\mathcal{D}}^{(n)}(\zeta)} z.$$

Let $\hat{\mathcal{P}} \subseteq \mathcal{P}$ be the subset of probability measures supported on $\hat{\Omega}$, and $\hat{\mathcal{M}}_n \subseteq \hat{\mathcal{P}}$ be the subset of probability measures $\mathbb{Q}$ that have the following properties:

$$\mathbb{E}^{\mathbb{Q}}\Big[\sum_{k=1}^{K_m+m}\left|\mathbb{E}^{\mathbb{Q}}[X_{\hat{\tau}_k}|\mathcal{F}_{\hat{\tau}_k-}] - X_{\hat{\tau}_{k-1}}\right|\Big] \le \frac{1}{n},$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

where $0 < \hat{\tau}_1(\omega) < \cdots < \hat{\tau}_{K_m+m-1}(\omega) < 1$ are the jumps times of the piecewise constant process $X(\omega)$ with $\hat{\tau}_0(\omega) = 0$ and $\hat{\tau}_{K_m+m}(\omega) = 1$. Then the required result $\mathsf{D}^c(\zeta) \le \mathsf{P}(\zeta)$ follows from the following Propositions 5.8 and 5.9.

Proposition 5.8. *Assume that ζ satisfies Assumptions 3.9, 3.10 and is $\mathbb{L}^\infty$-uniformly continuous, then*

$$\limsup_{n\to\infty} \mathsf{D}^{(n)}(\zeta) \le \mathsf{P}(\zeta).$$

Proof. (i) From Example 2.3 and Proposition 4.1 in Föllmer & Kramkov [15], it follows that

$$\mathsf{D}^{(n)}(\zeta) = \sup_{\mathbb{Q}\in\hat{\mathcal{P}}} \mathbb{E}^{\mathbb{Q}}\Big[\zeta - n \sum_{k=1}^{K_m+m} \Big|X_{\hat{\tau}_{k-1}} - \mathbb{E}^{\mathbb{Q}}[X_{\hat{\tau}_k}|\mathcal{F}_{\hat{\tau}_k-}]\Big|\Big].$$

Since $0 \le \zeta \le 1$, we determine that $\mathsf{D}^{(n)}(\zeta) \ge 0$ and we have for every $\mathbb{Q} \in \hat{\mathcal{P}} \setminus \hat{\mathcal{M}}_n$,

$$\mathbb{E}^{\mathbb{Q}}\Big[\zeta - n \sum_{k=1}^{K_m+m} \Big|X_{\hat{\tau}_{k-1}} - \mathbb{E}^{\mathbb{Q}}[X_{\hat{\tau}_k}|\mathcal{F}_{\hat{\tau}_k-}]\Big|\Big] \le 0,$$

which yields

$$\mathsf{D}^{(n)}(\zeta) \le \sup_{\mathbb{Q}\in\hat{\mathcal{M}}_n} \mathbb{E}^{\mathbb{Q}}[\zeta(X)].$$

(ii) Let us take a sequence $(\mathbb{Q}_n)_{n\ge1}$ with $\mathbb{Q}_n \in \hat{\mathcal{M}}_n$ such that

$$\limsup_{n\to\infty}\Big\{\sup_{\mathbb{Q}\in\hat{\mathcal{M}}_n} \mathbb{E}^{\mathbb{Q}}[\zeta(X)]\Big\} = \limsup_{n\to\infty} \mathbb{E}^{\mathbb{Q}_n}[\zeta(X)].$$

Since under each $\mathbb{Q}_n$ the canonical process X is piecewise constant with jump times $0 < \hat{\tau}_1 < \cdots < \hat{\tau}_{K_m+m-1} < 1$, X is a $\mathbb{Q}_n$-semimartingale. Then we have the decomposition $X = M^{\mathbb{Q}_n} - A^{\mathbb{Q}_n}$, where $A^{\mathbb{Q}_n}$ is a predictable process of bounded variation and $M^{\mathbb{Q}_n}$ is a martingale under $\mathbb{Q}_n$. Moreover, $A^{\mathbb{Q}_n}$ is identified by

$$A_t^{\mathbb{Q}_n} = \sum_{k=1}^{K_m+m-1} \mathbb{1}_{[\hat{\tau}_k,\hat{\tau}_{k+1})}(t) \sum_{j=1}^{k}\Big[X_{\hat{\tau}_{j-1}} - \mathbb{E}^{\mathbb{Q}_n}[X_{\hat{\tau}_j}|\mathcal{F}_{\hat{\tau}_j-}]\Big] \quad \text{for all } t \in [0,1),$$

and $A_1^{\mathbb{Q}_n} = \lim_{t\to1} A_t^{\mathbb{Q}_n}$. It follows then $\mathbb{E}^{\mathbb{Q}_n}\big[|X_1 - M_1^{\mathbb{Q}_n}|\big] \le \mathbb{E}^{\mathbb{Q}_n}\big[|A_1^{\mathbb{Q}_n}|\big] \le 1/n$ and

$$\mathbb{Q}_n\Big[\|A^{\mathbb{Q}_n}\| \ge n^{-1/2}\Big] \le n^{1/2}\mathbb{E}^{\mathbb{Q}_n}\Big[\sum_{k=1}^{K_m+m-1} \Big|X_{\hat{\tau}_{k-1}} - \mathbb{E}^{\mathbb{Q}_n}[X_{\hat{\tau}_k}|\mathcal{F}_{\hat{\tau}_k-}]\Big|\Big] \le n^{-1/2}.$$

Since ζ is $\mathbb{L}^\infty$-uniformly continuous, one obtains

$$\limsup_{n\to\infty} \mathbb{E}^{\mathbb{Q}_n}[\zeta(X)] \le \limsup_{n\to\infty} \mathbb{E}^{\mathbb{Q}_n}[\zeta(M^{\mathbb{Q}_n})].$$

Let $\mathbb{P}_n = \mathbb{Q}_n \circ (M^{\mathbb{Q}_n})^{-1}$, then

$$\begin{aligned}\sup_{n\ge1} \mathbb{E}^{\mathbb{P}_n}[|X_1|] &= \sup_{n\ge1} \mathbb{E}^{\mathbb{Q}_n}[|M_1^{\mathbb{Q}_n}|] \\ &\le \sup_{n\ge1} \mathbb{E}^{\mathbb{Q}_n}[|M_1^{\mathbb{Q}_n} - X_1|] + \sup_{n\ge1} \mathbb{E}^{\mathbb{Q}_n}[X_1]\end{aligned}$$

$$\leq \sup_{n\geq 1}\mathbb{E}^{\mathbb{Q}_n}[|M_1^{\mathbb{Q}_n}-X_1|]+\sup_{n\geq 1}\mathbb{E}^{\mathbb{Q}_n}[X_1-M_1^{\mathbb{Q}_n}]+\sup_{n\geq 1}\mathbb{E}^{\mathbb{Q}_n}[M_1^{\mathbb{Q}_n}]$$
$$\leq 1+\frac{2}{n}\leq 3.$$

By Assumptions 3.9 and 3.10, it follows that for any $\boldsymbol{\varepsilon}\in\mathbb{R}_+^m$ such that $0<|\boldsymbol{\varepsilon}|<\Delta\mathbb{T}$

$$\limsup_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}[\zeta(X)]\leq\limsup_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}[\zeta(X_{f_{\boldsymbol{\varepsilon}}})]+\big(1+(m+2)d\big)\alpha(|\boldsymbol{\varepsilon}|).$$

Again with the same reasoning, we may prove

$$\limsup_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}[\zeta(X_{f_{\boldsymbol{\varepsilon}}})]\leq \mathsf{P}(\zeta).$$

Since $\boldsymbol{\varepsilon}$ is arbitrary we get

$$\limsup_{n\to\infty}\Big[\sup_{\mathbb{Q}\in\hat{\mathcal{M}}_n}\mathbb{E}^{\mathbb{Q}}[\zeta(X)]\Big]\leq \mathsf{P}(\zeta),$$

and hence the required result. □

5.4.2. Time–space discretization

Discretization: For each $\omega\in\Omega_+^c$ let us define $\tau_k:=\tau_k^{(n)}(\omega)$ and $K_i:=K_i^{(n)}(\omega)$ by

$$\tau_0:=0,\qquad K_0:=0,$$
$$\tau_1:=t_1\wedge\sqrt{d}2^{-n}\wedge\inf\left\{t>0:|\omega_t-\omega_0|\geq 2^{-n}\right\},$$
$$\tau_{k+1}:=t_1\wedge(\tau_k+\Delta\tau_k)\wedge\inf\left\{t>\tau_k:|\omega_t-\omega_{\tau_k}|\geq 2^{-n}\right\},\quad \Delta\tau_k=\tau_k-\tau_{k-1}\text{ for }k\geq 1.$$

Set further

$$K_1:=\min\{k\in\mathbb{N}:\tau_k=t_1\}.$$

Recursively, we define for $1\leq i\leq m-1$ and $k\geq K_i$,

$$\tau_{K_i+1}:=t_{i+1}\wedge(t_i+\sqrt{d}2^{-n})\wedge\inf\left\{t>t_i:|\omega_t-\omega_{t_i}|\geq 2^{-n}\right\},$$
$$\tau_{k+1}:=t_{i+1}\wedge(\tau_k+\Delta\tau_k)\wedge\inf\left\{t>\tau_k:|\omega_t-\omega_{\tau_k}|\geq 2^{-n}\right\}\quad\text{for }k\geq K_i+1$$

and

$$K_{i+1}:=\min\{k\in\mathbb{N}:\tau_k=t_{i+1}\}.$$

Notice that the above τ_k, $k\geq 0$ are all stopping times w.r.t. to the right-continuous filtration $\mathbb{F}^+=(\mathcal{F}_{t+})_{t\geq 0}$, and

$$0=\tau_0<\tau_1\cdots<\tau_{K_m}=1\quad\text{and}\quad\tau_{K_i}=t_i\quad\text{for all }i=1,\ldots,m.$$

Moreover, for $0\leq i\leq m-1$, $K_i<k\leq K_{i+1}$ and $t\in[\tau_{k-1},\tau_k)$,

$$|\omega_t-\omega_{\tau_{k-1}}|\leq 2^{-n}\quad\text{and}\quad\Delta\tau_{k+1}\leq\Delta\tau_k\leq 2^{-n}.$$

Also by the continuity of ω at $\tau_{K_i}=t_i$ for all $i=1,\ldots,m$

$$|\omega_t-\omega_{\tau_{K_i-1}}|\leq 2^{-n}\quad\text{for all }t\in[\tau_{K_i-1},t_i]\text{ and }i=1,\ldots,m.$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Lifting: Set $\hat{\tau}_0 := 0$ and for $0 \le i \le m-1$

$$\begin{aligned}\hat{\tau}_{K_i+1} &:= \hat{\tau}_{K_i} + \sqrt{d}2^{-n},\\ \hat{\tau}_k &:= \hat{\tau}_{k-1} + \big(1 - \sqrt{d}2^{-n}/\Delta t_{i+1}\big)\sup\big\{\Delta t > 0 : \Delta t \in B^{(n+k-K_i-i)},\\ &\qquad \Delta t < \Delta\tau_{k-1}\big\}, \quad \text{for all } K_i + i + 2 \le k \le K_{i+1} + i,\\ \hat{\tau}_{K_{i+1}+i+1} &:= t_{i+1}.\end{aligned}$$

Denote $\hat{\Pi}(\omega) = \big(\hat{\Pi}(\omega)\big)_{0\le t\le 1}$ by

$$\begin{aligned}\hat{\Pi}_t(\omega) := &\sum_{i=0}^{m-1}\Big\{\sum_{k=K_i+i}^{K_{i+1}+i-1}\pi^{(n+k-K_i-i)}(\omega_{\tau_k})\mathbb{1}_{[\hat{\tau}_k,\hat{\tau}_{k+1})}(t) + \pi^{(n)}(\omega_{t_i})\mathbb{1}_{[\hat{\tau}_{K_r},t_{i+1})}(t)\Big\}\\ &+ \pi^{(n)}(\omega_1)\mathbb{1}_{\{t=1\}},\end{aligned}$$

then $\hat{\Pi}(\omega) \in \hat{\Omega}$. For each $\hat{H} \in \hat{\mathcal{A}}$ we may define

$$H_t(\omega) := \sum_{k=0}^{K_m-1}\hat{H}_{\hat{\tau}_{k+1}(\omega)}(\hat{\Pi}(\omega))\mathbb{1}_{(\tau_k(\omega),\tau_{k+1}(\omega)]}(t) \quad \text{for all } (\omega,t) \in \Omega_+ \times [0,1]. \tag{5.9}$$

We observe that $(\omega,t) \mapsto H_t(\omega)$ is Borel measurable on $\Omega_+ \times [0,1]$, $t \mapsto H_t(\omega)$ is left-continuous, $\omega \mapsto H_t(\omega)$ is $\mathcal{F}_{t+}$-measurable. Hence H is $\mathbb{F}^+$-predictable, which is equivalent to be $\mathbb{F}$-predictable. Further, following the argument of Lemmas 3.5 and 3.6 of Dolinsky & Soner [13], we see that the process H defined by (5.9) belongs to $\mathcal{A}$, and more importantly, there exists some constant $C > 0$ independent of n such that for all $\omega \in \Omega_+$,

$$\begin{aligned}\rho_{\mathbb{T}}(\omega, \hat{\Pi}(\omega)) &\le C2^{-n}\big(1 + \|\omega\|\big) \quad \text{and} \quad \big|(H\cdot\omega)_1 - (\hat{H}(\hat{\Pi}(\omega))\cdot\hat{\Pi}(\omega))_1\big|\\ &\le Cn2^{-n}.\end{aligned} \tag{5.10}$$

Proposition 5.9. *Assume that ζ satisfies Condition (5.2), then one has*

$$\liminf_{n\to\infty}\mathsf{D}^{(n)}(\zeta) \ge \mathsf{D}^c(\zeta).$$

Proof. Take an arbitrary $(\hat{z},\hat{H}) \in \hat{\mathcal{D}}$. Then for any $\omega \in \Omega_+$ one has $\hat{\Pi}(\omega) \in \hat{\Omega}$ and thus

$$\hat{z} + (\hat{H}(\hat{\Pi}(\omega))\cdot\hat{\Pi}(\omega))_1 \ge \zeta(\hat{\Pi}(\omega)) \quad \text{for all } \omega \in \Omega_+.$$

Take H constructed as (5.9), then by (5.10), we have $H \in \mathcal{A}$ and

$$\hat{z} + (H\cdot\omega)_1 \ge \zeta(\hat{\Pi}(\omega)) - Cn2^{-n} \quad \text{for all } \omega \in \Omega_+.$$

Moreover, by the construction of $\hat{\Pi}(\omega)$ one has $\|\hat{\Pi}(\omega)\| \le \|\omega\|$. Notice that ζ is boundedly supported, saying by $\{\omega \in \Omega : \|\omega\| \le R\}$. Then by (5.2) one has a continuous increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\beta(0) = 0$ such that for all $\omega \in \Omega_+$,

$$\zeta(\hat{\Pi}(\omega)) \ge \zeta(\omega) - \mathbb{1}_{\{\|\omega\|\le R\}}\beta\big(\rho_{\mathbb{T}}\big(\omega,\hat{\Pi}(\omega)\big)\big) \ge \zeta(\omega) - \beta\big(C(1+R)2^{-n}\big),$$

which implies that $\big(\hat{z} + \beta\big(C(1+R)2^{-n}\big) + Cn2^{-n}, H\big) \in \mathcal{D}^c(\zeta)$. Hence

$$\mathsf{D}^c(\zeta) \le \mathsf{D}^{(n)}(\zeta) + \beta\big(C(1+R)2^{-n}\big) + Cn2^{-n},$$

which yields the required result. □

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

5.5. Proof of Proposition 5.4(i)

Recall that ξ satisfies Assumptions 3.9, 3.10 and Condition (5.2). The required statement follows from the three lemmas below.

Lemma 5.10. $\limsup_{n\to\infty} \xi^\delta(\omega^{\delta,n}) \le \xi(\omega^{\delta,0})$ *holds for any sequence* $(\omega^{\delta,n})_{n\ge0} \subseteq \Omega^\delta$ *such that*

$$\omega^{\delta,n} \xrightarrow{S^*} \omega^{\delta,0} \quad \text{and} \quad \omega^{\delta,n}_{t_i^\delta} \longrightarrow \omega^{\delta,0}_{t_i^\delta} \quad \text{for all } i = 0, \dots, m-1.$$

Proof. For the sake of simplicity we may assume that

$$\limsup_{n\to\infty} \xi^\delta(\omega^{\delta,n}) = \lim_{n\to\infty} \xi^\delta(\omega^{\delta,n}).$$

Since $(\omega^{\delta,n})_{n\ge1}$ is S-tight, then by the S-tightness criteria and the construction in (5.5) we determine that $(\bar\omega^{\delta,n})_{n\ge1}$ is again S-tight, which yields a convergent subsequence $(\bar\omega^{\delta,n_k})_{k\ge1}$ and a limit $\omega^0 \in \Omega$, i.e. $\bar\omega^{\delta,n_k} \xrightarrow{S^*} \omega^0$. Clearly, $\omega^{\delta,n}_{t_i^\delta} \longrightarrow \omega^{\delta,0}_{t_i^\delta}$ implies in particular that $\bar\omega^{\delta,n_k}_{t_i} \longrightarrow \bar\omega^{\delta,0}_{t_i}$ for all $i = 0, \dots, m-1$. Next, $\omega^{\delta,n_k} \xrightarrow{S^*} \omega^{\delta,0}$ yields a countable set $\mathcal{T} \subseteq [0, 1+\delta)$ such that

$$\omega^{\delta,n_k}_t \longrightarrow \omega^{\delta,0}_t \quad \text{for all } t \in [0, 1+\delta] \setminus \mathcal{T},$$

which yields another countable set $\mathcal{T}' \subseteq [0,1)$ such that

$$\bar\omega^{\delta,n_k}_t \longrightarrow \bar\omega^{\delta,0}_t \quad \text{for all } t \in [0,1] \setminus \mathcal{T}'.$$

Hence one has $\bar\omega^{\delta,0} = \omega^0$ and thus

$$\bar\omega^{\delta,n_k} \xrightarrow{S^*} \bar\omega^{\delta,0} \quad \text{and} \quad \bar\omega^{\delta,n_k}_{t_i} \longrightarrow \bar\omega^{\delta,0}_{t_i} \quad \text{for all } i = 0, \dots, m-1,$$

which implies that

$$\lim_{k\to\infty} \xi^\delta(\omega^{\delta,n_k}) = \lim_{k\to\infty} \xi(\bar\omega^{\delta,n_k}) \le \xi(\bar\omega^{\delta,0}) = \xi^\delta(\omega^{\delta,0}). \quad \square$$

Lemma 5.11. *There exists a continuous function* $\alpha_\delta : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $\alpha_\delta(0) = 0$ *such that for all* $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m) \in \mathbb{R}^m_+$ *sufficiently small one has*

$$\big|\xi^\delta(\omega^\delta) - \xi^\delta\big(\omega^\delta_{f^\delta_{\boldsymbol{\varepsilon}}}\big)\big|, \qquad \big|\xi^\delta(\omega^\delta) - \xi^\delta\big(\omega^\delta_{b^\delta_{\boldsymbol{\varepsilon}}}\big)\big| \le \alpha_\delta(|\boldsymbol{\varepsilon}|)\Big(1 + \sum_{i=0}^m |\omega^\delta_{t_i^\delta}| + \int_0^{1+\delta} |\omega^\delta_t| dt\Big),$$

where $f^\delta_{\boldsymbol{\varepsilon}}$, $b^\delta_{\boldsymbol{\varepsilon}} : [0, 1+\delta] \to [0, 1+\delta]$ *are two non-decreasing functions defined as in* (3.4) *and* (3.5).

Proof. We only prove the inequality on $f^\delta_{\boldsymbol{\varepsilon}}$, while the inequality on $b^\delta_{\boldsymbol{\varepsilon}}$ follows by the same arguments. Define $\bar f^\delta_{\boldsymbol{\varepsilon}} : [0, 1+\delta] \to [0, 1+\delta]$ by

$$\bar f^\delta_{\boldsymbol{\varepsilon}}(t) := \frac{1}{k_\delta} f^\delta_{\boldsymbol{\varepsilon}}(k_\delta t) \quad \text{for all } t \in [0,1].$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

Thus by the construction of $f^\delta_{\boldsymbol{\varepsilon}}$ we get

$$\begin{aligned}\bar f^\delta_{\boldsymbol{\varepsilon}}(t) &= \frac{1}{k_\delta}\sum_{i=1}^m \mathbb{1}_{(t^\delta_{i-1},t^\delta_i]}(k_\delta t)\Big(t^\delta_{i-1}+\frac{\Delta t^\delta_i}{\Delta t^\delta_i-\varepsilon_i}\big(k_\delta t - t^\delta_{i-1}-\varepsilon_i\big)^+\Big)\\ &= \sum_{i=1}^m \mathbb{1}_{(t_{i-1},t_i]}(t)\Big(t_{i-1}+\frac{\Delta t_i}{\Delta t_i-\varepsilon_i/k_\delta}\Big(t-t_{i-1}-\frac{\varepsilon_i}{k_\delta}\Big)^+\Big),\end{aligned}$$

which implies that

$$\bar\omega^\delta_{f^\delta_{\boldsymbol{\varepsilon}}(t)} = \omega^\delta_{k_\delta \bar f^\delta_{\boldsymbol{\varepsilon}}(t)} \quad \text{for all } t\in[0,1]$$

and thus $\bar\omega^\delta_{f^\delta_{\boldsymbol{\varepsilon}}} = \bar\omega^\delta \circ \bar f^\delta_{\boldsymbol{\varepsilon}}$. Hence

$$\begin{aligned}\big|\xi^\delta(\omega^\delta)-\xi^\delta\big(\bar\omega^\delta_{f^\delta_{\boldsymbol{\varepsilon}}}\big)\big| &= \big|\xi(\bar\omega^\delta)-\xi\big(\bar\omega^\delta\circ\bar f^\delta_{\boldsymbol{\varepsilon}}\big)\big|\\ &\le \alpha(|\boldsymbol{\varepsilon}|/k_\delta)\Big(1+\sum_{i=0}^m|\bar\omega^\delta_{t_i}|+\int_0^1|\bar\omega^\delta_t|dt\Big)\\ &= \alpha(|\boldsymbol{\varepsilon}|/k_\delta)\Big(1+\sum_{i=0}^m|\omega^\delta_{t^\delta_i}|+\frac{1}{k_\delta}\int_0^{1+\delta}|\omega^\delta_t|dt\Big)\\ &\le \alpha(|\boldsymbol{\varepsilon}|/k_\delta)\Big(1+\sum_{i=0}^m|\omega^\delta_{t^\delta_i}|+\int_0^{1+\delta}|\omega^\delta_t|dt\Big).\end{aligned}$$

The proof is completed by taking $\alpha_\delta(\cdot)=\alpha(\cdot/k_\delta)$. □

Lemma 5.12. *ξ^δ is L^∞-uniformly continuous and satisfies Condition (5.2) for $\rho_{\mathbb{T}^\delta}$.*

Proof. For any $\omega^\delta, v^\delta\in\Omega^\delta$ such that $\|v^\delta\|\le\|\omega^\delta\|$, one has

$$\begin{aligned}&\xi^\delta(\omega^\delta)-\xi^\delta(v^\delta)=\xi(\bar\omega^\delta)-\xi(\bar v^\delta)\le\beta\big(\rho_{\mathbf{T}}(\bar\omega^\delta,\bar v^\delta)\big),\\ &\big|\xi^\delta(\omega^\delta)-\xi^\delta(v^\delta)\big|=\big|\xi(\bar\omega^\delta)-\xi(\bar v^\delta)\big|.\end{aligned}$$

It is thus enough to show that

$$\rho_{[t_{i-1},t_i]}(\bar\omega^\delta,\bar v^\delta)\le\rho_{[t^\delta_{i-1},t^\delta_i]}(\omega^\delta,v^\delta)\quad\text{for all } i=1,\dots,m$$

and

$$\Big|\int_0^1(\bar\omega^\delta_t-\bar v^\delta_t)dt\Big|\le\Big|\int_0^{1+\delta}(\omega^\delta_t-v^\delta_t)dt\Big|.$$

Let $\Gamma_{[s,t]}$ denotes the collection of strictly increasing continuous functions γ defined on $[s,t]$ such that $\gamma(s)=s$ and $\gamma(t)=t$. For any $\gamma^\delta\in\Gamma_{[t^\delta_{i-1},t^\delta_i]}$, define $\gamma\in\Gamma_{[t_{i-1},t_i]}$ by

$$\gamma(t):=\frac{1}{k_\delta}\gamma^\delta\big(k_\delta t\big)\quad\text{for all } t\in[t_{i-1},t_i].$$

Hence

$$\begin{aligned}\sup_{t_{i-1}\le t\le t_i}\big|\bar\omega^\delta_{\gamma(t)}-\bar v^\delta_t\big| &= \sup_{t_{i-1}\le t\le t_i}\big|\omega^\delta_{k_\delta\gamma(t)}-v^\delta_{k_\delta t}\big| = \sup_{t_{i-1}\le t\le t_i}\big|\omega^\delta_{\gamma^\delta(k_\delta t)}-v^\delta_{k_\delta t}\big|\\ &= \sup_{t^\delta_{i-1}\le t\le t^\delta_i}\big|\omega^\delta_{\gamma^\delta(t)}-v^\delta_t\big|\end{aligned}$$

and

$$\sup_{t_{i-1}\le t\le t_i}\big|\gamma(t)-t\big| = \sup_{t_{i-1}\le t\le t_i}\Big|\frac{1}{k_\delta}\gamma^\delta\big(k_\delta t\big)-t\Big| = \frac{1}{k_\delta}\sup_{t_{i-1}\le t\le t_i}\big|\gamma^\delta\big(k_\delta t\big)-k_\delta t\big|$$
$$= \frac{1}{k_\delta}\sup_{t^\delta_{i-1}\le t\le t^\delta_i}\big|\gamma^\delta(t)-t\big| \le \sup_{t^\delta_{i-1}\le t\le t^\delta_i}\big|\gamma^\delta(t)-t\big|,$$

which implies that

$$\rho_{[t_{i-1},t_i]}(\bar\omega^\delta,\bar v^\delta)\le \rho_{[t^\delta_{i-1},t^\delta_i]}(\omega^\delta,v^\delta).$$

We may thus conclude by

$$\Big|\int_0^1(\bar\omega^\delta_t-\bar v^\delta_t)dt\Big| = \frac{1}{k_\delta}\Big|\int_0^{1+\delta}(\omega^\delta_t-v^\delta_t)dt\Big| \le \Big|\int_0^{1+\delta}(\omega^\delta_t-v^\delta_t)dt\Big|. \quad \square$$

Appendix

A.1. Tightness under S-topology

Recall that $\Omega=\mathbb{D}([0,1],\mathbb{R}^d)$ is the Skorokhod space of càdlàg paths on $[0,1]$, with canonical process $X=(X_t)_{0\le t\le 1}$ and canonical filtration $\mathbb{F}=(\mathcal{F}_t)_{0\le t\le 1}$, and $\mathcal{P}$ denotes the set of all probability measures on $(\Omega,\mathcal{F}_1)$. A sequence of probability measures $(\mathbb{P}_n)_{n\ge1}\subset\mathcal{P}$ is said to be S-tight if for any $\varepsilon>0$, there exists a S-compact set $K_\varepsilon\subset\Omega$ such that

$$\inf_{n\ge1}\mathbb{P}_n\big[X\in K_\varepsilon\big]\ge 1-\varepsilon.$$

The following result is recalled from Jakubowski [28] (see their Theorem 3.1 and the discussion at the beginning of Section 4 in [28]).

Theorem A.1 (*Jakubowski*)**.**

(i) *Let $(\mathbb{P}_n)_{n\ge1}\subseteq\mathcal{P}$ be a sequence of probability measures such that X is a $\mathbb{P}_n$-supermartingale for all $n\ge1$, then*

$$\sup_{n\ge1}\sup_{0\le t\le1}\mathbb{E}^{\mathbb{P}_n}\big[|X_t|\big]<+\infty \Longrightarrow (\mathbb{P}_n)_{n\ge1}\ \textit{is S-tight}.$$

(ii) *Let $(\mathbb{P}_n)_{n\ge1}\subseteq\mathcal{P}$ be a S-tight sequence of probability measures. Then there exist a subsequence $(\mathbb{P}_{n_k})_{k\ge1}$, a probability measure $\mathbb{P}\in\mathcal{P}$ and a countable subset $\mathcal{T}\subset[0,1)$ such that for all finite sets $\{u_1<u_2<\cdots<u_r\}\subset[0,1]\setminus\mathcal{T}$,*

$$\mathbb{P}_{n_k}\circ(X_{u_1},\dots,X_{u_r})^{-1}\longrightarrow\mathbb{P}\circ(X_{u_1},\dots,X_{u_r})^{-1}\quad \textit{as } k\to\infty. \tag{A.1}$$

In particular, $X^{n_k}\overset{*}{\Longrightarrow}_D X^0$ as $k\to\infty$.

A.2. Dual space of **M**

Recall that $\mathbf{M}$ denotes the space of all finite signed measures μ on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d}\big(1+|x|\big)|\mu|(dx)<+\infty,$$

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

and it is equipped with the topology induced by the convergence $\xrightarrow{\mathcal{W}_1}$. We would like to identify its dual space as Λ, where the arguments are mainly adapted from Lemma 3.2.3 of Deuschel & Stroock [12]. Notice that the topology on $\mathbf{M}$ is generated by all the following open balls

$$U_{\lambda^1,\dots,\lambda^m,c}(\mu) := \Big\{\nu \in \mathbf{M} : \big|\mu(\lambda^i) - \nu(\lambda^i)\big| < c \text{ for all } 1 \le i \le m\Big\},$$

where $\lambda^i \in \Lambda$ for $1 \le i \le m$ and $c > 0$. Let $\mathcal{O}$ be the collection of open sets generated by the open balls above, then clearly, every open set $U \in \mathcal{O}$ could be expressed as

$$U = \bigcup_\alpha U_{\lambda^1_\alpha,\dots,\lambda^{n_\alpha}_\alpha,c^\alpha}(\mu^\alpha) \quad \text{with } \lambda^i_\alpha \in \Lambda \text{ for } 1 \le i \le n_\alpha, n_\alpha \in \mathbb{N} \text{ and } c^\alpha > 0.$$

Theorem A.2. *The space* $(\mathbf{M}, \mathcal{O})$ *is a Hausdorff locally convex space, whose dual space can be identified by* $\mathbf{M}^* = \Lambda$.

Proof. (i) First, $(\mathbf{M}, \mathcal{O})$ is clearly a topological vector space. For every $\mu \in \mathbf{M}$, let

$$\mathcal{U}(\mu) := \big\{U_{\lambda^1_\alpha,\dots,\lambda^{n_\alpha}_\alpha,c^\alpha}(\mu) : \lambda^i_\alpha \in \Lambda_1 \text{ for } 1 \le i \le n_\alpha, n_\alpha \in \mathbb{N} \text{ and } c^\alpha > 0\big\}.$$

By definition, one can check that $\mathcal{U}(\mu)$ is a local basis of μ for every $\mu \in \mathbf{M}$. Moreover, by denoting $\mathbf{0} \in \mathbf{M}$ the null measure, $\mathcal{U}(\mathbf{0})$ is a local basis of absolutely convex absorbent sets and thus $\mathbf{M}$ is a locally convex space.
(ii) Now, let us identify the dual space of $\mathbf{M}$. First, for every $\lambda \in \Lambda$, the map $F_\lambda : \mathbf{M} \to \mathbb{R}$ defined by $F_\lambda(\mu) := \mu(\lambda)$ gives a unique element in $\mathbf{M}^*$, and hence $\Lambda \subseteq \mathbf{M}^*$. On the other hand, for any $F \in \mathbf{M}^*$, we define a function λ^F by

$$\lambda^F(x) := F(\delta_{\{x\}}) \quad \text{for all } x \in \mathbb{R}^d.$$

Clearly one has the following implication

$$x_n \to x_0 \Longrightarrow \delta_{\{x_n\}} \xrightarrow{\mathcal{O}} \delta_{\{x_0\}} \Longrightarrow \lambda^F(x_n) \to \lambda^F(x_0),$$

which implies that λ^F is continuous. It follows that the set $F^{-1}\big((-1,1)\big)$ is open and thus there exists some $U_{\lambda^1,\dots,\lambda^m,c}(\mathbf{0})$ such that

$$U_{\lambda^1,\dots,\lambda^m,c}(\mathbf{0}) \subseteq F^{-1}\big((-1,1)\big),$$

where $\lambda^i \in \Lambda$ for all $i = 1, \dots, m$ and $c > 0$. Now for any $\mu \in \mathbf{M}$ such that $\sum_{i=1}^m \big|\mu(\lambda^i)\big| > 0$, we define

$$\bar{\mu} := \frac{c\mu}{\sum_{i=1}^m \big|\mu(\lambda^i)\big|}.$$

Then $\bar{\mu} \in U_{\lambda^1,\dots,\lambda^m,c}(\mathbf{0})$ and thus $|F(\bar{\mu})| < 1$. It follows that

$$|F(\mu)| \le c \sum_{i=1}^m \big|\mu(\lambda^i)\big| \quad \text{for all } \mu \in \mathbf{M},$$

and hence $\lambda^F \in \Lambda$. When μ is a linear combination of Dirac measures, it is obvious that $F(\mu) = \mu(\lambda^F)$. Moreover, since such μ are dense in $\mathbf{M}$, it follows that $F(\mu) = \mu(\lambda^F)$ holds for all $\mu \in \mathbf{M}$. □

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005

References

[1] B. Acciaio, M. Beiglböck, F. Penkner, W. Schachermayer, J. Temme, A trajectorial interpretation of Doob's martingale inequalities, Ann. Appl. Probab. 23 (4) (2013) 1494–1505.
[2] M. Beiglböck, P. Henry-Labordère, F. Penkner, Model-independent bounds for option prices: a mass transport approach.
[3] M. Beiglböck, M. Nutz, N. Touzi, Robust super-replication for time-invariant derivatives, Preprint, 2015.
[4] D.T. Breeden, R.H. Litzenberger, Prices of state-contingent claims implicit in options prices, J. Bus. 51 (1978).
[5] H. Brown, D. Hobson, L.C.G. Roggers, Robust hedging of barrier options, Math. Finance 11 (2001) 285–314.
[6] A. Cox, D. Hobson, J. Oblój, Pathwise inequalities for local time: applications to Skorokhod embeddings and optimal stopping, Ann. Appl. Probab. 18 (5) (2008) 1870–1896.
[7] A. Cox, J. Oblój, Robust hedging of double touch barrier options, SIAM J. Financ. Math. 2 (2011) 141–182.
[8] A. Cox, J. Oblój, Robust pricing and hedging of double no-touch options, Finance Stoch. 15 (3) (2011) 573–605.
[9] A. Cox, J. Oblój, N. Touzi, The Root solution to the multi-marginal embedding problem: an optimal stopping and time-reveral approach, preprint, 2015.
[10] A. Cox, J. Wang, Root's barrier: Construction, optimality and applications to variance options, Ann. Appl. Probab. 23 (3) (2013) 859–894.
[11] M. Davis, J. Oblój, V. Raval, Arbitrage bounds for weighted variance swap prices, Math. Finance 24 (4) (2014) 821–854.
[12] J.D. Deuschel, D.M. Stroock, Large Deviations, American Mathematical Society, 1991.
[13] Y. Dolinsky, H.M. Soner, Martingale optimal transport in the Skorokhod space, Preprint, 2014.
[14] Y. Dolinsky, H.M. Soner, Robust hedging and martingale optimal transport in continuous time, Probab. Theory Related Fields 160 (1-2) (2014) 391–427.
[15] H. Föllmer, D. Kramkov, Optional decompositions under constraints, Probab. Theory Related Fields 109 (1997) 1–25.
[16] A. Galichon, P. Henry-Labordère, N. Touzi, A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options, Ann. Appl. Probab. 24 (1) (2014) 312–336.
[17] P. Gassiat, H. Oberhauser, G. dos Reis, Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs, Preprint, 2015.
[18] F. Hirsch, B. Roynette, A new proof of Kellerer's theorem, ESAIM Probab. Stat. 16 (2012) 48–60.
[19] F. Hirsch, B. Roynette, On $\mathbb{R}^d$-valued peacocks, ESAIM Probab. Stat. 17 (2013) 444–454.
[20] D. Hobson, Robust hedging of the lookback option, Finance Stoch. 2 (1998) 329–347.
[21] D. Hobson, The Skorokhod embedding problem and model-independent bounds for option prices, in: Paris-Princeton Lectures on Mathematical Finance, in: LNM, vol. 2003, Springer, 2010.
[22] D. Hobson, M. Klimmek, Model independent hedging strategies for variance swaps, Finance Stoch. 16 (2012) 611–649. http://dx.doi.org/10.1007/s00780-012-0190-3.
[23] D. Hobson, M. Klimmek, Maximising functionals of the joint law of the maximum and terminal value in the Skorokhod embedding problem, Ann. Appl. Probab. 23 (5) (2013) 2020–2052.
[24] D. Hobson, M. Klimmek, Robust price bounds for the forward starting straddle, Finance Stoch. 19 (1) (2015) 189–214.
[25] D. Hobson, A. Neuberger, Robust bounds for forward start options, Math. Finance 22 (2012) 31–56.
[26] Z. Hou, J. Oblój, On robust pricing-hedging duality in continuous time, Preprint, 2015.
[27] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, in: Grundlehren der Mathematischen Wissenschaften ([Fundamental Principles of Mathematical Sciences]), vol. 228, Springer-Verlag, Berlin, 1987.
[28] A. Jakubowski, A non-Skorohod topology on the Skorohod space, Electron. J. Probab. 2 (4) (1997) 1–21.
[29] Hans G. Kellerer, Markov-Komposition und eine Anwendung auf Martingale, Math. Ann. (1972).
[30] D.O. Kramkov, Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets, Probab. Theory Related Fields (1995).
[31] D. Madan, M. Yor, Making Markov martingales meet marginals: with explicit constructions, Bernoulli 8 (4) (2002) 509–536.
[32] P.A. Meyer, W. Zheng, Tightness criteria for laws of semimartingales, Ann. Inst. Henri Poincare B 20 (4) (1984) 353–372.
[33] A. Neufeld, M. Nutz, Measurability of semimartingale characteristics with respect to the probability law, Stochastic Process. Appl. 124 (11) (2014) 3819–3845.
[34] A. Neufeld, M. Nutz, Nonlinear Lévy processes and their characteristics, Trans. Amer. Math. Soc. (2016).
[35] M. Nutz, Robust Superhedging with Jumps and Diffusion, Preprint, 2014.
[36] J. Oblój, The Skorokhod embedding problem and its offspring, Probab. Surv. 1 (2004) 321–392.
[37] S.T. Rachev, L. Rüschendorf, Mass Transportation Problems Volume 1: Theory, Vol. XXVI, 1998, p. 508.
[38] J. Arthur Seebach, Lynn Arthur Steen, Counterexamples in Topology, Springer-Verlag, New York, 1978.
[39] C. Villani, Optimal Transport: Old and New, in: Grundlehren der Mathematischen Wissenschaften, 2009.

Please cite this article in press as: G. Guo, et al., Tightness and duality of martingale transport on the Skorokhod space, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2016.07.005