The Annals of Probability
2017, Vol. 45, No. 4, 2210–2247
DOI: 10.1214/16-AOP1110

AN ITERATED AZÉMA–YOR TYPE EMBEDDING FOR FINITELY MANY MARGINALS

BY JAN OBŁÓJ[1] AND PETER SPOIDA[2]

University of Oxford

We solve the n-marginal Skorokhod embedding problem for a continuous local martingale and a sequence of probability measures $\mu_1, \dots, \mu_n$ which are in convex order and satisfy an additional technical assumption. Our construction is explicit and is a multiple marginal generalization of the Azéma and Yor [In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)* (1979) 90–115 Springer] solution. In particular, we recover the stopping boundaries obtained by Brown, Hobson and Rogers [*Probab. Theory Related Fields* **119** (2001) 558–578] and Madan and Yor [*Bernoulli* **8** (2002) 509–536]. Our technical assumption is necessary for the explicit embedding, as demonstrated with a counterexample. We discuss extensions to the general case giving details when $n = 3$.

In our analysis we compute the law of the maximum at each of the n stopping times. This is used in Henry-Labordère et al. [*Ann. Appl. Probab.* **26** (2016) 1–44] to show that the construction maximizes the distribution of the maximum among all solutions to the n-marginal Skorokhod embedding problem. The result has direct implications for robust pricing and hedging of Lookback options.

1. Introduction. We consider here an n-marginal Skorokhod embedding problem (SEP). We construct an explicit solution which has desirable optimal properties. The classical (one-marginal) SEP consists in finding a stopping time τ such that a given stochastic process (X_t) stopped at τ has a given distribution μ. For the solution to be useful (and nontrivial), one further requires τ to be *minimal* (cf. Obłój [20], Section 8). When X is a continuous local martingale and μ is centred in X_0, this is equivalent to $(X_{t\wedge\tau} : t \geq 0)$ being a uniformly integrable martingale. The problem dates back to the original work in Skorokhod [24] and has remained an active field of research since. New solutions often either considered new classes of processes X or focused on finding stopping times τ with additional optimal properties. This paper contributes to the latter category. We are motivated,

Received February 2014; revised October 2015.

[1]Supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement No. 335421 and Oxford-Man Institute of Quantitative Finance and St. John's College in Oxford.

[2]Supported by the Oxford-Man Institute of Quantitative Finance and the DAAD.

MSC2010 subject classifications. 60G40, 60G44.

Key words and phrases. Skorokhod embedding problem, Azéma–Yor embedding, maximum process, martingale optimal transport, continuous martingale, marginal constraints.

as was the case for several earlier works in the field, by questions arising in mathematical finance which we highlight below.

The problem and main results. To describe the problem of interest, consider a standard Brownian motion B and a sequence of probability measures $\mu_1, \dots, \mu_n$. A solution to the n-marginal SEP is a sequence of stopping times $\tau_1 \le \cdots \le \tau_n$ such that $B_{\tau_i} \sim \mu_i$, $1 \le i \le n$, and $(B_{t\wedge\tau_n})_{t\ge 0}$ is a uniformly integrable martingale. It follows from Jensen's inequality that a solution may exist only if all μ_i are centred and the sequence is in convex order. And then it is easy to see how to solve the problem: it suffices to iterate a solution to the classical case $n = 1$ developed for a nontrivial initial distribution of B_0, of which several exist.

In contrast, the question of optimality is much more involved. In general, there is no guarantee that a simple iteration of optimal embeddings would be globally optimal. Indeed, this is usually not the case. Consider the embedding of Azéma and Yor [1] which consists of a first exit time for the joint process $(B_t, \bar{B}_t)_{t\ge 0}$, where $\bar{B}_t = \sup_{s\le t} B_s$. More precisely, their solution $\tau^{\mathrm{AY}} = \inf\{t \ge 0 : B_t \le \xi_\mu(\bar{B}_t)\}$ leads to a functional relation $B_{\tau^{\mathrm{AY}}} = \xi_\mu(\bar{B}_{\tau^{\mathrm{AY}}})$. This then translates into the optimal property that the distribution of $\bar{B}_{\tau^{\mathrm{AY}}}$ is maximized in stochastic order amongst all solutions to SEP for μ, that is, for all y,

$$\mathbb{P}[\bar{B}_{\tau^{\mathrm{AY}}} \ge y] = \sup\{\mathbb{P}[\bar{B}_\rho \ge y] : \rho \text{ s.t. } B_\rho \sim \mu, (B_{t\wedge\rho}) \text{ is UI}\}.$$

It is not hard to generalize the Azéma–Yor embedding to a nontrivial starting law; see Obłój [20], Section 5. Consequently, we can find η_i such that $\tau_i = \inf\{t \ge \tau_{i-1} : B_t \le \eta_i(\sup_{\tau_{i-1}\le s\le t} B_s)\}$ solve the n-marginal SEP. However, this construction will maximize stochastically the distributions of $\sup_{\tau_{i-1}\le t\le\tau_i} B_t$, for each $1 \le i \le n$, but not of the global maximum $\bar{B}_{\tau_n}$. The latter is achieved with a new solution which we develop here.

Our construction involves an interplay between all n-marginals, and hence is not an iteration of a one-marginal solution. However, it preserves the spirit of the Azéma–Yor embedding in the following sense. Each τ_i is still a first exit for $(B_t, \bar{B}_t)_{t\ge\tau_{i-1}}$ which is designed in such a way as to obtain a "strong relation" between B_{τ_i} and $\bar{B}_{\tau_i}$, ideally a functional relation. Under our technical assumption about the measures $\mu_1, \dots, \mu_n$, Assumption $\circledast$, we describe this relation in detail in Lemma 3.1.

For $n = 2$, we recover the results of Brown, Hobson and Rogers [5]. We also recover the trivial case $\tau_i = \tau^{AY}_{\mu_i}$ which happens when $\xi_{\mu_i} \le \xi_{\mu_{i+1}}$, we refer to Madan and Yor [18] who in particular then investigate properties of the arising time-changed process. However, as a counterexample shows, our construction does not work for all laws $\mu_1, \dots, \mu_n$ which are in convex order. Assumption $\circledast$ fails when a special interdependence between the marginals is present and the analysis then becomes more technical and the resulting quantities are, in a way, less explicit. We only sketch the appropriate arguments for the case $n = 3$.

We stress that the problem considered in this paper is significantly more complex that the special case $n = 1$. For $n = 1$, several solutions to SEP exist with different optimal properties. For $n = 2$, only one such construction, the generalization of the Azéma–Yor embedding obtained by Brown, Hobson and Rogers [4], seems to be known. To the best of our knowledge, the solution we present here is the first one to deal with the general n-marginal SEP.[3]

Motivation and applications. Our results have direct implications for, and were motivated by, robust pricing and hedging of lookback options. In mathematical finance, one models the price process S as a martingale and specifying prices of call options at maturity T is equivalent to fixing the distribution μ of S_T. Understanding no-arbitrage price bounds for a functional O, which time-changes appropriately, is then equivalent to finding the range of $\mathbb{E}[O(B)_\tau]$ among all solutions to the Skorokhod embedding problem for μ. This link between SEP and robust pricing and hedging was pioneered by Hobson [16] who considered Lookback options. Barrier options were subsequently dealt with by Brown, Hobson and Rogers [5]. More recently, Cox and Obłój [7, 8] considered the case of double touch/no-touch barrier options, Hobson and Neuberger [15] looked at forward starting straddles and analysis for variance options was undertaken by Cox and Wang [10]. We refer to Hobson [14] and Obłój [21] for an exposition of the main ideas and more references. However, all the previous works considered essentially the case of call options with one maturity, that is, a one-marginal SEP, while in practice prices for many intermediate maturities may also be available. This motivated our investigation.

We started our quest for a general n-marginal optimal embedding by computing the value function $\sup \mathbb{E}[\phi(\sup_{t\le\tau_n} B_t)]$ among all solutions to the n-marginal SEP. This was achieved using stochastic control methods, developed first for $n = 1$ by Galichon, Henry-Labordère and Touzi [11], and is reported in a companion paper Henry-Labordère et al. [13]. Knowing the value function, we could start guessing the form of the optimizer and this led to the present paper. Consequently, the optimal properties of our embedding, namely that it indeed achieves the value function in question, are shown in Henry-Labordère et al. [13]. In fact, we give two proofs in that paper, one via stochastic control methods and another one by constructing appropriate pathwise inequalities and exploiting the key Lemma 3.1 below; cf. Henry-Labordère et al. [13], Section 4.

Organization of the paper. The remainder of the paper is organized as follows. In Section 2, we explain the main quantities for the embedding and state the main result. We also present the restriction on the measures $\mu_1, \dots, \mu_n$ which we require for our construction to work (Assumption $\circledast$). In Section 3, we prove the main result and Section 4 provides a discussion of extensions together with comments on Assumption $\circledast$. The proof of an important but technical lemma is relegated to the Appendix.

[3]We note that recently an n-fold version of the Root solution was announced in Cox, Obłój and Touzi [9].

2. Main results. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)$, be a filtered probability space satisfying the usual hypotheses and B a continuous $\mathbb{F}$-local martingale, $B_0 = 0$, $\langle B\rangle_\infty = \infty$ a.s. and B has no intervals of constancy a.s. We denote $\bar{B}_t := \sup_{s\le t} B_t$. We are primarily interested in the case when B is a standard Brownian motion and it is convenient to keep this example in mind, hence the notation. We allow for more generality as this introduces no changes to the statements or the proofs.

2.1. *Definitions*. We introduce below the fundamental objects of our study: the stopping boundaries $\xi_1, \dots, \xi_n$ for our iterated Azéma–Yor type embedding together with quantities $K_1, \dots, K_n$ which will be later linked to the law of the maximum at subsequent stopping times. We define various quantities assuming that a family of probability measures $(\mu_i)_{1\le i\le n}$ is given. Later, appropriate assumptions on (μ_i) will be made to ensure all objects are well defined. We think of n as a parameter: definitions below are recursive in n and the proofs will be mostly done by induction on n. We denote the left and right endpoints of the support of the measure μ_i by

$$l_{\mu_i} := \sup\{x : \mu_i([x,\infty)) = 1\}, \qquad r_{\mu_i} := \inf\{x : \mu_i((x,\infty)) = 0\}, \tag{2.1}$$

respectively, and, for $1 \le i \le n$, we let

$$c_i(\zeta) := \int_{\mathbb{R}} (x-\zeta)^+ \mu_i(\mathrm{d}x), \qquad \zeta \in \mathbb{R} \quad \text{and} \tag{2.2}$$

$$\bar{K}_n(\zeta_1, \dots, \zeta_n, y) := \sum_{i=1}^n \frac{c_i(\zeta_i) - c_{i-1}(\zeta_i)}{y - \zeta_i}, \qquad y \ge 0, \zeta_1, \dots, \zeta_n \in (-\infty, y], \tag{2.3}$$

where $c_0 \equiv 0$ and the values of $\bar{K}_n$ for $\zeta_i = y$ are understood as limits for $\zeta_i \nearrow y$. We sometimes refer to c_i as "call prices," a nomenclature borrowed from mathematical finance.

DEFINITION 2.1 (Stopping boundaries). Set the initial values as

$$c_0 \equiv 0, \qquad K_0 \equiv 0, \qquad \xi_0 \equiv -\infty. \tag{2.4}$$

For $n \in \mathbb{N}$, having previously defined $\xi_1(y), \dots, \xi_{n-1}(y)$, we write

$$\zeta_i^k(y) := \min_{i\le j\le k} \xi_j(y), \qquad y \ge 0, 1 \le i \le k, \tag{2.5}$$

for $k \le n-1$, and define the subsequent stopping boundary by $\xi_n(0) = l_{\mu_n}$,

$$\xi_n(y) := \sup\Big\{\arg\min_{\zeta\le y} \bar{K}_n\big(\zeta_1^{n-1}(y)\wedge\zeta, \dots, \zeta_{n-1}^{n-1}(y)\wedge\zeta, \zeta, y\big)\Big\}, \qquad y > 0. \tag{2.6}$$

Finally, we let ζ_i^n be given by (2.5) with $k = n$ and define

$$K_n(y) := \bar{K}_n\big(\zeta_1^n(y), \dots, \zeta_n^n(y), y\big), \qquad y \ge 0. \tag{2.7}$$

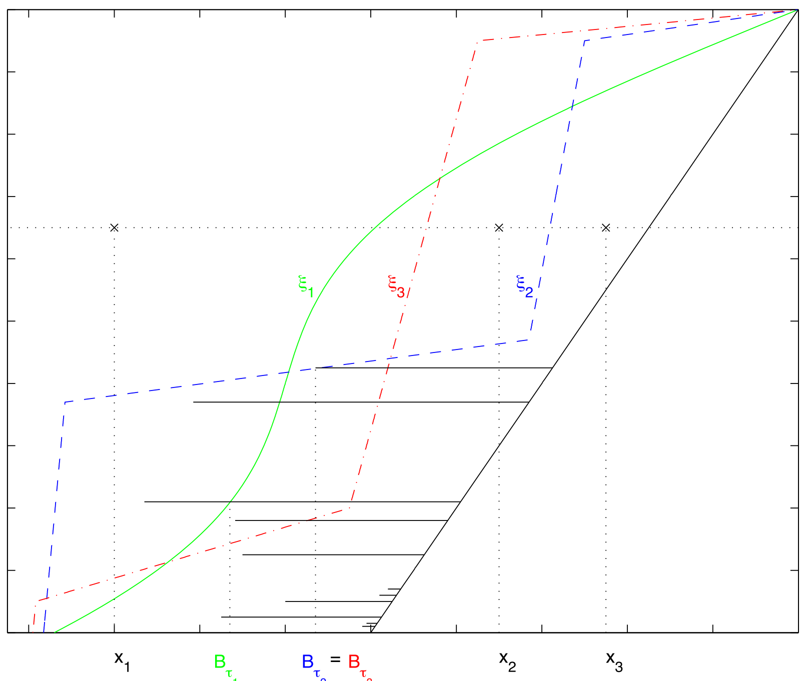

FIG. 1. *We illustrate possible stopping boundaries ξ_1, ξ_2, ξ_3. The horizontal lines represent a sample path of the process $(B_t, \bar{B}_t)$ where the x-axis is the value of B and the y-axis the value of $\bar{B}$. Each horizontal segment is an excursion of B away from its maximum $\bar{B}$. According to the definition of the embedding, the first stopping time τ_1 is found when the process first hits ξ_1. Since $\xi_1(\bar{B}_{\tau_1}) > \xi_2(\bar{B}_{\tau_1})$ the process continues and targets ξ_2. The stopping time τ_2 is found when the process first hits ξ_2. Since $\xi_2(\bar{B}_{\tau_2}) \le \xi_3(\bar{B}_{\tau_2})$, we get $\tau_3 = \tau_2$. For the y we fixed, we have $\iota_3(x_1, y) = 0, \iota_3(x_2, y) = 1, \iota_3(x_3, y) = 2$; see* (2.12).

We show below in Remark 2.8 that when $c_i \ge c_{i-1}$ then the optimization in (2.6) is well posed and setting $\xi_n(0) = l_{\mu_n}$ is consistent.

DEFINITION 2.2 (Embedding). Given stopping boundaries $\xi_1, \dots, \xi_n$ define the associated stopping times by setting $\tau_0 \equiv 0$ and

$$\tau_i := \inf\{t \ge \tau_{i-1} : B_t \le \xi_i(\bar{B}_t)\}, \qquad i = 1, \dots, n. \tag{2.8}$$

Figure 1 illustrates a set of possible stopping boundaries ξ_1, ξ_2, ξ_3 in the case of $n = 3$. If Assumption $\circledast$ is in place (see Sections 2.2 and 2.4), we will show that the stopping boundaries are continuous (except possibly for ξ_1) and nondecreasing; cf. Section 2.6. Note that the nth stopping boundary ξ_n is obtained from an optimization problem which features $\xi_1, \dots, \xi_{n-1}$ and $c_1, \dots, c_n$. $K_n(y)$ is the value of the objective function at the optimal value $\xi_n(y)$. The previously defined stopping boundaries $\xi_1, \dots, \xi_{n-1}$ and the quantities $K_1, \dots, K_{n-1}$ remain unchanged. This gives an iterative structure allowing to "add one marginal at a time" and enables us to prove our results by induction on n.

2.2. *The main result.* We start by imposing an important restriction on the given measures $(\mu_1, \dots, \mu_n)$.

ASSUMPTION $\circledast$ (Restriction on measures). We assume that $\mu_1, \dots, \mu_n$ are probability measures which satisfy:

(i) for all $1 \le i \le n$, $\int |x|\mu_i(\mathrm{d}x) < \infty$ with $\int x\mu_i(\mathrm{d}x) = 0$ and $c_{i-1} \le c_i$ with strict inequality $c_{i-1} < c_i$ on (l_{μ_i}, r_{μ_i});
(ii) for all $2 \le i \le n$ and all $0 < y < r_{\mu_i}$ the mapping

$$\zeta \mapsto \bar{K}_n\big(\zeta_1^{n-1}(y) \wedge \zeta, \dots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y\big) \tag{2.9}$$

admits a unique minimizer $\zeta^\star$ on (l_{μ_i}, y).

We discuss below in detail the significance of the above assumption. However, first we state our main result giving an n-fold embedding of $(\mu_1, \dots, \mu_n)$ in the spirit of Azéma and Yor [1] and Brown, Hobson and Rogers [4].

THEOREM 2.3. *Recall Definitions 2.1 and 2.2. Let $n \in \mathbb{N}$ and assume $\mu_1, \dots, \mu_n$ are given and satisfy Assumption $\circledast$ above.*

Then $\tau_i < \infty$ a.s., $B_{\tau_i} \sim \mu_i$ for all $i = 1, \dots, n$ and $(B_{\tau_n \wedge t})_{t \ge 0}$ is a uniformly integrable martingale. In addition, we have for $y \ge 0$ and $i = 1, \dots, n$,

$$\mathbb{P}[\bar{B}_{\tau_i} \ge y] = K_i(y), \tag{2.10}$$

where K_i is defined in (2.7).

REMARK 2.4 (Minimality). Since $\tau_i \le \tau_{i+1}$, it follows that $(B_{t\wedge\tau_i})_{t\ge 0}$ is a uniformly integrable martingale for any $1 \le i \le n$ and all τ_i are *minimal* (in the sense of Monroe [19]).

REMARK 2.5 (Uniqueness). In general, two sets of nondecreasing boundaries (ξ_i) could give the same distributions for B_{τ_i}, with τ_i in (2.8). This is due to the fact that parts of ξ_i may be never "seen by the embedding." For example, suppose $\xi_1(y) = \cdots = \xi_n(y)$ for $y \in [0, y_0]$. The embedding is then not affected by any change of boundaries ξ_i, $2 \le i \le n$, on $[0, y_0]$ which satisfies $\xi_i \ge \xi_1$ on $[0, y_0]$ and preserves global monotonicity. To obtain unicity in all generality, one would have to define "regular" stopping barriers, analogously to the way Loynes [17] defined regular Root's stopping barriers.

However, for measures satisfying Assumption $\circledast$, there is essentially a one-to-one correspondence between measures and (suitably nice) stopping boundaries. This is due to the strict inequality between potentials. Let us argue this briefly.

Clearly, because of the embedding property, different sets of target measures cannot generate the same stopping boundaries. Reversely, consider a set of continuous nondecreasing stopping boundaries $(\tilde{\xi}_i)$ with the associated stopping times $(\tilde{\tau}_i)$ in (2.8) such that $(B_{\tilde{\tau}_n\wedge t} : t \geq 0)$ is uniformly integrable. Denote (μ_i) the embedded measures, $\mu_i \sim B_{\tilde{\tau}_i}$, and suppose (μ_i) satisfy Assumption $\circledast$. For unicity, we need to assume that $\tilde{\xi}_i(y) = y$ for $y \geq r_{\mu_i}$ as these values are never seen by the embedding. Suppose now that (ξ_i), as obtained in Theorem 2.3, are different from $(\tilde{\xi}_i)$. Then, using continuity and monotonicity of (ξ_i) (see Section 2.6 below), we may assume that $\xi_j = \tilde{\xi}_j$ for $j < i$ and that on some interval $[y_0, y_1)$ we have $\xi_i < \tilde{\xi}_i$ (or the reverse inequality), for some $i \leq n$. Note that

$$\mathbb{P}(B_{\tilde{\tau}_{i-1}} \leq x, B_{\tilde{\tau}_i} > x) > 0, \qquad \forall x \in (l_{\mu_{i-1}}, r_{\mu_{i-1}}),$$

and likewise for τ_{i-1}, τ_i, as otherwise we would have $c_{i-1}(x) = c_i(x)$. In consequence, B_{τ_i} has a positive probability of stopping by hitting the boundary ξ_i on any interval. Together with $(B_{\tilde{\tau}_i}, \bar{B}_{\tilde{\tau}_i}) = (B_{\tau_i}, \bar{B}_{\tau_i})$, an explicit excursion theoretical computation shows that the distributions of $B_{\tilde{\tau}_i}$ and B_{τ_i} have to differ which establishes a contradiction.

2.3. *Alternative characterization of the stopping boundaries.* Let us investigate in more detail the optimization problem in (2.6). First, note that it is a constrained version of the global minimization problem

$$C_n(y) := \inf_{\zeta_1(y)\leq\cdots\leq\zeta_n(y)<y} \bar{K}_n\big(\zeta_1(y), \dots, \zeta_n(y), y\big) \tag{2.11}$$

in that the minimization is over ζ_n and the first $n-1$ coordinates are taken to be $\zeta_i^{n-1}(y) \wedge \zeta_n(y)$. In particular, $C_n(y) \leq K_n(y)$. The above global minimization appears as (3.5) in [13]. Theorem 3.3 therein, or more directly the simple pathwise inequality in Proposition 3.1, together with our Theorem 2.3 above, show that

$$K_n(y) = \mathbb{P}(\bar{B}_{\tau_n} \geq y) \leq C(y)$$

from which it follows that $K_n(y) = C_n(y)$ and $\zeta_1^n, \dots, \zeta_n^n$ solve (2.11). We exploit this further in Example 2.14 below.

It is useful and insightful to derive a representation of (2.6) by exhibiting explicitly the terms which collapse in the telescoping sum in (2.3). Let

$$\begin{aligned} &\iota_n(\cdot, y) : (-\infty, y] \to \{0, 1, \dots, n-1\}, \qquad y \geq 0, \text{ be given by} \\ &\iota_n(\zeta, y) := \max\{k \leq n-1 : \xi_k(y) < \zeta\} = \max\{k : \zeta_k^{n-1}(y) < \zeta\}. \end{aligned} \tag{2.12}$$

Note that $\zeta_i^{n-1} = \zeta_i^{\iota_n(\zeta,y)}$ for $i \leq \iota_n(\zeta, y)$ and $\zeta_i^{n-1} \wedge \zeta = \zeta$ for $\iota_n(\zeta, y) < i \leq n$. By considering the summands for $i \leq \iota_n(\zeta, y)$ and $i > \iota_n(\zeta, y)$ in (2.3), it follows that

$$\begin{aligned} &\bar{K}_n\big(\zeta_1^{n-1}(y) \wedge \zeta, \dots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y\big) \\ &\quad = K_{\iota_n(\zeta,y)}(y) + \frac{c_n(\zeta) - c_{\iota_n(\zeta,y)}(\zeta)}{y - \zeta} \end{aligned} \tag{2.13}$$

and for future reference it will be useful to define

$$\begin{aligned} &c^n(\cdot, y) :(-\infty, y] \to \mathbb{R} \cup \{\infty\}, \qquad y \geq 0, \\ &c^n(\zeta, y) := (y-\zeta) K_{\iota_n(\zeta,y)}(y) + c_n(\zeta) - c_{\iota_n(\zeta,y)}(\zeta) \\ &\qquad = \bar{K}_n\big(\zeta_1^{n-1}(y) \wedge \zeta, \ldots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y\big)(y-\zeta). \end{aligned} \tag{2.14}$$

It follows that the minimization problem in (2.6) is equivalent to the following minimization problem:

$$\xi_n(y) = \sup\left\{\arg\min_{\zeta \leq y} \frac{c^n(\zeta, y)}{y-\zeta}\right\}, \qquad y > 0. \tag{2.15}$$

Finally, we let

$$\jmath_n(y) := \iota_n\big(\xi_n(y), y\big), \tag{2.16}$$

which is the index of the last marginal $\mu_i, i < n$, which represents, locally at the level of the maximum y, a *binding constraint* for the embedding. Observe that

$$K_n(y) = K_{\jmath_n(y)}(y) + \frac{c_n(\xi_n(y)) - c_{\jmath_n(y)}(\xi_n(y))}{y - \xi_n(y)}. \tag{2.17}$$

To close this section, let us explain the relation of the above construction to the (one-marginal) classical Azéma–Yor embedding. Denote the barycentre function of μ_i by

$$b_i(x) := \frac{\int_{[x,\infty)} u\mu_i(\mathrm{d}u)}{\mu_i([x,\infty))} \mathbb{1}_{\{l_{\mu_i} < x < r_{\mu_i}\}} + x\mathbb{1}_{\{x \geq r_{\mu_i}\}}, \qquad x \in \mathbb{R}. \tag{2.18}$$

As shown by Brown, Hobson and Rogers [5], the right-continuous inverse of b_i, denoted by b_i^{-1}, can be represented as

$$b_i^{-1}(y) = \sup\left\{\arg\min_{\zeta \leq y} \frac{c_i(\zeta)}{y-\zeta}\right\}. \tag{2.19}$$

In particular, from Definition 2.1, we have $\xi_1 = b_1^{-1}$. It is clear, and was studied in more detail by Madan and Yor [18], that if the sequence of barycentre functions is increasing in i, then the intermediate law constraints do not have an impact on the corresponding iterated Azéma–Yor embedding. However, in general the barycentre functions will not be increasing in i (cf. Brown, Hobson and Rogers [4]), and hence will affect the embedding. As compared to the optimization from which b_n^{-1} is obtained [cf. (2.19)], the optimization from which ξ_n is obtained [cf. (2.15)] has a penalty term.

2.4. *Restrictions on measures*. We turn to the discussion of our key assumption on the marginals (μ_i). We proceed via a series of remarks.

REMARK 2.6 [Assumption $\circledast$(i)]. The condition that the call prices are nondecreasing in maturity

$$c_i \le c_{i+1}, \qquad i = 1, \dots, n-1, \tag{2.20}$$

can be rephrased by saying that $\mu_1, \dots, \mu_n$ are nondecreasing in the convex order. It is the necessary and sufficient condition for a uniformly integrable martingale with these marginals to exist, as shown by Strassen [25], Theorem 2. Condition (i) in Assumption $\circledast$ is stronger in that we require a strict inequality inside the support.

REMARK 2.7 (Discontinuity of ξ_1). Note that Assumption $\circledast$(ii) does not require that the mapping

$$\zeta \mapsto \frac{c^1(\zeta, y)}{y - \zeta} = \frac{c_1(\zeta)}{y - \zeta} \tag{2.21}$$

has a unique minimizer. It may happen that there is an interval of minimizers and then ξ_1 is discontinuous at such y.

REMARK 2.8 [Well-posedness of optimization (2.6)]. First, observe that

$$\bar{K}_n\big(\zeta_1^{n-1}(y) \wedge \zeta, \dots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y\big) = \frac{c^n(\zeta, y)}{y - \zeta} = \frac{c_n(\zeta)}{y - \zeta} \qquad \text{for } \zeta \le \zeta_1^{n-1}(y).$$

With a fixed $y > 0$, for ζ small enough, this is smaller than one. Also, under (2.20), $\bar{K}_n \ge 0$ so in consequence

$$0 \le K_n(y) \le 1, \qquad y \ge 0. \tag{2.22}$$

Further, the function $\frac{c_n(\zeta)}{y-\zeta}$ is nonincreasing on $(-\infty, b_n^{-1}(y)]$, where b_n^{-1} denotes the right-continuous inverse of the barycentre function b_n; cf. (2.19). It follows by induction that

$$\min_{1 \le i \le n} b_i^{-1}(y) \le \xi_n(y) \le y, \qquad y \ge 0, \tag{2.23}$$

so minimization in (2.6) is over a compact interval. Also, we see that under (2.20), the limit of $\frac{c_i(\zeta) - c_{i-1}(\zeta)}{y - \zeta}$, as $\zeta \nearrow y$, is well defined and under the strict inequalities in Assumption $\circledast$(i) it is equal to $+\infty$. Existence of a minimizer $\zeta^\star$ in (2.6) now follows from the continuity of $\bar{K}_n$ and by the above $\zeta^\star < y$.

Finally, we also note that for $y = 0$, and assuming $\xi_i(0) = l_{\mu_i}$ for $1 \le i \le n-1$, we have

$$\bar{K}_n\big(\zeta_1^{n-1}(y) \wedge \zeta, \dots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y\big) = 1 + \frac{c_n(\zeta) - c_{n-1}(\zeta)}{-\zeta},$$

which on $(-\infty, 0]$ is minimized for $\zeta = l_{\mu_n}$ (where it is understood as an asymptotic statement for $l_{\mu_i} = -\infty$) showing our convention $\xi_n(0) = l_{\mu_n}$ is consistent. We then have $K_n(0) = 1$.

REMARK 2.9 (Asymptotic behavior of ξ_n and K_n). Observe that, under (2.20), if $r_{\mu_i} < \infty$ then $\xi_n(y) = y$ and $K_n(y) = 0$ for $y \geq r_{\mu_n}$ and otherwise ξ_n is unbounded and $K_n(y) \to 0$ as $y \to \infty$. For $n = 1$, these are known properties of the barycentre function in (2.19). For the general case, let us reason by contradiction. Suppose $n > 1$ is the first index where the above property is violated. If $r_{\mu_n} < \infty$ then, noting that (2.20) implies $r_{\mu_i} \leq r_{\mu_{i+1}}$, we have, for $y \geq r_{\mu_n}$, $\zeta_i^{n-1}(y) = y$ and

$$\bar{K}_n(\zeta_1^{n-1}(y) \wedge \zeta, \dots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y) = \frac{c_n(\zeta)}{y - \zeta},$$

which attains its minimum equal to zero at any $\zeta \in [r_{\mu_n}, y]$ giving, by (2.6), $\xi_n(y) = y$. Similarly, if $r_{\mu_n} = +\infty$ but ξ_n is bounded, say $\xi_n < \lambda$, then for y large enough $\zeta_i^{n-1}(y) > \lambda$, and hence for $\zeta < \lambda$

$$\bar{K}_n(\zeta_1^{n-1}(y) \wedge \zeta, \dots, \zeta_{n-1}^{n-1}(y) \wedge \zeta, \zeta, y) = \frac{c_n(\zeta)}{y - \zeta},$$

which would be minimized by taking $\zeta = \lambda$ giving a contradiction.

REMARK 2.10 [Assumption $\circledast$(ii)]. By Remark 2.8 above, Assumption $\circledast$(i) implies that the function in (2.9) admits a minimizer $\zeta^\star$ on $(l_{\mu_i}, y]$ which satisfies $\zeta^\star < y$. Assumption $\circledast$(ii) then states that this minimizer is unique. We note however that in general, only assuming (2.20), the minimizer might not be unique and/or might satisfy $\zeta^\star = y$. The latter fact has been overlooked in [4] where $\zeta^\star < y$ is required for the arguments to hold; see Section 4.1 for details. Our assumption ensures in particular that we may rely on results in [4].

2.5. *Examples*. We turn now to examples. The first two, Examples 2.11 and 2.12, respectively, show that we recover the stopping boundaries obtained by Madan and Yor [18] and by Brown, Hobson and Rogers [4]. In particular, the case $n = 1$ corresponds to the solution of Azéma and Yor [1]. Example 2.14 serves to construct rich family of examples of marginals which satisfy Assumption $\circledast$.

EXAMPLE 2.11 (Madan and Yor [18]). Recall the definition of the barycentre function b_i from (2.18). Madan and Yor [18] consider the "increasing mean residual value" case, that is,

$$b_1 \leq b_2 \leq \cdots \leq b_n. \tag{2.24}$$

We will now show that our main result reproduces their result if Assumption $\circledast$ is in place. In fact, as can be seen below, our definitions of ξ_i and K_i [cf. (2.6) and

(2.7)], respectively, reproduce the correct stopping boundaries in the general case, showing that Assumption ⊛ is not necessary; cf. also Section 4. More precisely, we have

$$\xi_i = b_i^{-1}, \qquad K_i(y) = \frac{c_i(b_i^{-1}(y))}{y - b_i^{-1}(y)} =: \mu_i^{\mathrm{HL}}\big([y,\infty)\big), \qquad i = 1,\dots,n, \tag{2.25}$$

where b_i^{-1} denotes the right-continuous inverse of b_i and μ_i^{HL} is the Hardy–Littlewood transform of μ_i; cf. Carraro, El Karoui and Obłój [6].

We argue the claim by induction on n. For $n = 1$, it holds by definition. Now assume the claim holds for all $i \le n-1$. By (2.24), $\zeta_i^{n-1} = \xi_{n-1}$ for all $i \le n-1$. The optimization problem for ξ_n in (2.6) then becomes

$$\begin{aligned}
\xi_n(y) &\in \operatorname*{arg\,min}_{\zeta \le y} \left\{ \frac{c_n(\zeta)}{y-\zeta} - \mathbb{1}_{\{\zeta > b_{n-1}^{-1}(y)\}} \left[\frac{c_{n-1}(\zeta)}{y-\zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} \right] \right\} \\
&\in \operatorname*{arg\,min}_{\zeta \le y} \left\{ \min_{\zeta \le b_{n-1}^{-1}(y)} \frac{c_n(\zeta)}{y-\zeta}, \right. \\
&\qquad \left. \min_{\zeta \ge b_{n-1}^{-1}(y)} \left(\frac{c_n(\zeta)}{y-\zeta} - \left[\frac{c_{n-1}(\zeta)}{y-\zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} \right] \right) \right\}.
\end{aligned}$$

It is clear that the first minimum is $A_1 = \frac{c_n(b_n^{-1}(y))}{y - b_n^{-1}(y)}$ since $b_n^{-1}(y) \le b_{n-1}^{-1}(y)$.

As for the second minimum, we set

$$F(\zeta) := \frac{c_n(\zeta)}{y-\zeta} - \left[\frac{c_{n-1}(\zeta)}{y-\zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} \right]$$

and we see by direct calculation that for almost all $\zeta \in \mathbb{R}$

$$\begin{aligned}
&(y-\zeta)^2 F'(\zeta) \\
&\quad = \big(b_n(\zeta) - y\big)\mu_n\big([\zeta,\infty)\big) - \big(b_{n-1}(\zeta) - y\big)\mu_{n-1}\big([\zeta,\infty)\big) \\
&\quad = c_n(\zeta)\frac{b_n(\zeta) - y}{b_n(\zeta) - \zeta} - c_{n-1}(\zeta)\frac{b_{n-1}(\zeta) - y}{b_{n-1}(\zeta) - \zeta}.
\end{aligned}$$

By (2.24), we conclude therefore

$$(y-\zeta)^2 F'(\zeta) \ge \big(c_n(\zeta) - c_{n-1}(\zeta)\big)\frac{b_{n-1}(\zeta) - y}{b_{n-1}(\zeta) - \zeta} \ge 0,$$

where the last inequality follows from the nondecrease of the μ_i's in the convex order. Hence, F is nondecreasing, and it follows that it attains its minimum at the left boundary, that is, $A_2 = \frac{c_n(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} - [\frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)}] = \frac{c_n(b_{n-1}^{-1})(y)}{y - b_{n-1}^{-1}(y)}$. Consequently, by (2.19), $\min\{A_1, A_2\} = A_1$ and (2.25) follows.

EXAMPLE 2.12 (Brown, Hobson and Rogers [4]). In the case of $n = 2$, our definition of ξ_1 and ξ_2 clearly recovers the stopping boundaries in the main result of Brown, Hobson and Rogers [4]. However, our embedding is not as general as their embedding because we enforce Assumption ⊛; see also the discussion in Section 4.

EXAMPLE 2.13 (Locally no constraints). In general, we have

$$K_n(y) \le \mu_n^{\mathrm{HL}}([y, \infty)), \tag{2.26}$$

which holds by the fact that the distribution of the maximum in the n-marginal problem cannot be larger (in stochastic order) than in the 1-marginal problem where it is bounded by μ_n^{HL}. However, if for some $y \ge 0$

$$\xi_n(y) = b_n^{-1}(y) \quad \text{and} \quad \iota_n(\xi_n(y), y) = 0 \tag{2.27}$$

then it follows from Theorem 2.3 that

$$K_n(y) = \frac{c_n(b_n^{-1}(y))}{y - b_n^{-1}(y)} = \mu_n^{\mathrm{HL}}([y, \infty)), \tag{2.28}$$

that is, locally at level of maximum y the intermediate laws have no impact on the distribution of the terminal maximum as compared with the (one marginal) Azéma–Yor embedding.

EXAMPLE 2.14 (Generic measures). We provide now a method to produce generic families of marginals satisfying Assumption ⊛ based on a reversed procedure: first draw suitable boundaries, then run embedding and read off the marginals.

Let $n \in \mathbb{N}$ and consider a family of stopping boundaries ξ_i, $i \le n$, which are continuous, strictly increasing on $[0, \infty)$ with $\xi_i(y) = y$ for all $y \ge \inf\{y : \xi_i(y) = y\}$. Further, we assume that for some fixed $y_0 > 0$

$$\xi_n(y) < \cdots < \xi_1(y) \qquad \forall y \in (0, y_0). \tag{2.29}$$

It follows that ζ_i^n, as defined in (2.5), are given by ξ_n on $[0, y_0]$. Let τ_i be the stopping times defined in (2.8). We assume τ_n is such that $(B_{t\wedge\tau_n} : t \ge 0)$ is a uniformly integrable martingale and denote μ_i the distribution of B_{τ_i}. A simple sufficient condition for the uniform integrability is that $\xi_n(0) > -\infty$ and $\inf\{y : \xi_i(y) = y\} < \infty$, $i \le n$, since then, by construction, $(B_{t\wedge\tau_n} : t \ge 0)$ is a bounded martingale and all μ_i are compactly supported, more precisely $l_{\mu_i} = \xi_i(0)$ and $r_{\mu_i} = \inf\{y : \zeta_1^n(y) = y\}$.[4] Assumption ⊛(i), except the strict inequality between

[4]More generally, an excursion theoretical computation shows that it is sufficient to impose conditions on the asymptotic behavior of ζ_1^n, for example, that, for some $\alpha \in (0, 1)$, $\zeta_1^n(y) \ge -y^{-1/(1-\alpha)}$ for y small enough and $\zeta_1^n(y) \ge \alpha y$ for y large enough.

the call prices, follows from uniform integrability of $(B_{t\wedge\tau_n} : t \geq 0)$. The strict inequality $c_i > c_{i-1}$ on (l_{μ_i}, r_{μ_i}) follows from properties of ξ_i in (2.29) giving, for any $2 \leq i \leq n$,

$$\mathbb{P}(B_{\tau_{i-1}} \leq x, B_{\tau_i} > x) \geq \mathbb{P}\big(B_{\tau_{i-1}} \leq \min\{x, y_0\}, B_{\tau_i} > x\big) > 0, \qquad x \in (l_{\mu_i}, r_{\mu_i}).$$

It remains to argue that Assumption $\circledast$(ii) holds and that the stopping boundaries obtained in (2.6), say $\tilde{\xi}_i$, are in fact the original ones: $\tilde{\xi}_i = \xi_i$, $1 \leq i \leq n$. Our proof relies on the optimality properties obtained in Henry-Labordère et al. [13] and some further results in Obłój, Spoida and Touzi [22].

We can argue by induction on n since adding subsequent boundaries does not change the previous ones nor the embedded marginals. For $n = 1$, this problem is (2.19) and is known to have a unique solution $b_1^{-1}(y) = \tilde{\xi}_1(y)$ which then equals $\xi_1(y)$. Suppose, by induction, that $\tilde{\xi}_i = \xi_i$ for all $i < n$. Let $(Z_t : t \leq t_n)$ be a continuous time change of $(B_{t\wedge\tau_n} : t \geq 0)$ with $Z_{t_i} = B_{\tau_i}$, for example,

$$Z_t := B_{\tau_i \wedge (\tau_{i-1} \vee \frac{t-t_{i-1}}{t_i - t})} \qquad \text{for } t_{i-1} < t \leq t_i, i = 1, \dots, n.$$

It follows from Lemma 4.1 and Theorem 3.3 in Henry-Labordère et al. [13] that $\zeta_1^n \leq \cdots \leq \zeta_n^n$ solve the global minimization problem (2.11) for $0 < y < r_{\mu_n}$. The optimization problem (2.19) is a subproblem of (2.11) with the first $n-1$ coordinates taken to be $\zeta_i^{n-1} \wedge \zeta_n$. By the global optimization property above, $\zeta_n^n = \xi_n$ is necessarily a solution. It remains to show that it is the only minimizer. Suppose that for some $y \in (0, r_{\mu_n})$ there is another one, say $\tilde{\xi}_n(y) \neq \xi_n(y)$. Let $\tilde{\zeta}_i^n$ be the associated functions in (2.5), in particular $\tilde{\zeta}_n^n(y) \neq \zeta_n^n(y)$, and consider the pathwise inequality in [13], Proposition 3.1, associated with $(\tilde{\zeta}_i^n)_{i\leq n}$. Evaluating this inequality on Z, inspecting Proposition 3.2 in [22] and its proof, we see that there is a positive probability of the inequality being strict. This means $\tilde{\zeta}_i^n$ do not solve the global optimization problem (2.11), which is a contradiction since ζ_i^n do but both attain the same value.

2.6. *Properties of ξ_n and K_n.* Under Assumption $\circledast$, we establish the continuity of ξ_n for $n \geq 2$ (cf. Lemma 2.15), and prove monotonicity of ξ_n for $n \geq 1$; cf. Lemma 2.16. In Lemma 2.18, we derive an ODE for K_n which will be later used to identify the distribution of the maximum of the embedding from Definition 2.2. Recall from Section 2.1 that the main quantities are defined in an iterative manner, that is, the embedding of the first n_1 marginals in the n_2-marginals embedding problem, $n_2 > n_1$, coincides with the n_1-marginals embedding problem. Hence, it is natural to prove our results by induction over the number of marginals n.

LEMMA 2.15 (Continuity of c^n and ξ_n). *Let $n \geq 2$ and let Assumption $\circledast$ hold. Then the mappings*

$$c^n : \{(x, y) \in \mathbb{R} \times \mathbb{R}_+ : x < y\} \to \mathbb{R},$$
$$\xi_n : \mathbb{R}_+ \to \mathbb{R}$$

are continuous.

PROOF. We prove the claim by induction over n. Let us start with the induction basis $n = 1, 2$. Continuity of c^1 is the same as continuity of c_1 and continuity of c^2 is proven by Brown, Hobson and Rogers [4]; cf. Lemma 3.5 therein. As for continuity of ξ_2, we note that our Assumption $\circledast$(ii) precisely rules out discontinuities of ξ_2 as shown by Brown, Hobson and Rogers [4], Section 3.5. By induction hypothesis, we assume continuity of $c^1, \dots, c^{n-1}$ and $\xi_2, \dots, \xi_{n-1}$.

Note that the only way ξ_1 enters into the definition of c^n is through ζ_1^{n-1} and that $\frac{c_1(\xi_1(y)\wedge x)}{y-\xi_1(y)\wedge x}$ is a continuous function. Continuity of c^n then follows from its definition and the assumed continuity of $\xi_2, \dots, \xi_{n-1}$.

It remains to argue continuity of ξ_n, which we prove by contradiction. Suppose ξ_n is discontinuous at some $y > 0$ and let $y_k \to y$ with $\xi_n(y_k) \to \tilde{\xi} \neq \xi_n(y)$. Recalling (2.23), we note that $\tilde{\xi}$ is finite. By Assumption $\circledast$(ii), we necessarily have $\xi_n(y) < y$ and

$$\frac{c^n(\tilde{\xi}, y)}{y - \tilde{\xi}} > \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)}. \tag{2.30}$$

If $\tilde{\xi} < y$ then using continuity of c^n, we obtain a contradiction with optimality of $\xi_n(y_k)$ for k large enough. If $\tilde{\xi} = y$ then, by Assumption $\circledast$(i) and continuity of c_n and c_{n-1}, we have $c_n > c_{n-1} + \varepsilon$ on a small neighborhood of y and some $\varepsilon > 0$. It follows that we can make the term $\frac{c_n(\zeta)-c_{n-1}(\zeta)}{y-\zeta}$ in the sum in (2.3) uniformly large in ζ. Then the convergence $\xi_n(y_k) \to \tilde{\xi} = y$ again contradicts optimality of $\xi_n(y_k)$ for k large enough. $\square$

LEMMA 2.16 (Monotonicity of ξ_n). *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold. Then*

$$y \mapsto \xi_n(y) \qquad \textit{is nondecreasing on } [0, \infty). \tag{2.31}$$

PROOF. The claim for $n = 1, 2$ follows from Brown, Hobson and Rogers [4]. Assume by induction hypothesis that we have proven monotonicity of $\xi_1, \dots, \xi_{n-1}$.

We prove monotonicity locally and note that we may exclude the set of y's which are discontinuity points of ξ_1 and that it is enough to consider $y < r_{\mu_n}$. Let us then fix a $y < r_{\mu_n}$ where $\xi_1, \dots, \xi_n$ are continuous and recall that $\xi_n(y) < y$. The reasoning is similar to the arguments of Brown, Hobson and Rogers [4], Lemma 3.2.

It suffices then to argue that $\xi_n(y + \delta) \geq \xi_n(y)$ for δ small enough. We will first consider the case when $\xi_n(y) \neq \xi_j(y)$ for all $j < n$. By continuity of ξ_n it follows that there is an $\varepsilon > 0$ such that

$$\xi_n(\tilde{y}) \neq \xi_j(\tilde{y}) \quad \text{and} \quad \jmath_n(y) = \jmath_n(\tilde{y}) \qquad \forall \tilde{y} \in [y - \varepsilon, y + \varepsilon] \text{ and } j < n. \tag{2.32}$$

Consider the image of $I := [y - \varepsilon, y + \varepsilon]$ via ξ_n, that is, let $\delta_\pm \geq 0$ be such that $[\xi_n(y) - \delta_-, \xi_n(y) + \delta_+] = \{\xi_n(\tilde{y}) : \tilde{y} \in I\}$. If $\delta_- = 0$, then the statement is automatically true so assume $\delta_- > 0$. Consider $\tilde{y} \in I$ and $\zeta_\pm < \tilde{y}$ with $\xi_n(y) - \delta_- \leq$

$\zeta_- < \xi_n(\tilde{y}) < \zeta_+ \le \xi_n(y) + \delta_+$ and $\lambda\zeta_- + (1-\lambda)\zeta_+ = \xi_n(\tilde{y})$. Then, using (2.15) and Assumption ⊛(ii), we have

$$\begin{aligned}
&\lambda c^n(\zeta_-, \tilde{y}) + (1-\lambda)c^n(\zeta_+, \tilde{y}) \\
&\quad > \lambda(\tilde{y} - \zeta_-)\frac{c^n(\xi_n(\tilde{y}), \tilde{y})}{\tilde{y} - \xi_n(\tilde{y})} + (1-\lambda)(\tilde{y} - \zeta_+)\frac{c^n(\xi_n(\tilde{y}), \tilde{y})}{\tilde{y} - \xi_n(\tilde{y})} \\
&\quad = c^n\big(\xi_n(\tilde{y}), \tilde{y}\big).
\end{aligned}$$

Further, from (2.14) and (2.32), for $\zeta \in [\xi_n(y) - \delta_-, \xi_n(y) + \delta_+]$, the difference

$$c^n(\zeta, \tilde{y}) - c^n(\zeta, y) = \zeta\big(K_{J_n(y)}(y) - K_{J_n(y)}(\tilde{y})\big) + \tilde{y}K_{J_n(y)}(\tilde{y}) - yK_{J_n(y)}(y)$$

is a linear function in ζ. It follows that $c^n(\cdot, y)$, and hence also any $c^n(\cdot, \tilde{y})$, is a (strictly) convex function for $\zeta \in (\xi_n(y) - \delta_-, \xi_n(y) + \delta_+)$. By definition, the supporting tangent at $\xi_n(y)$ to $c^n(\cdot, y)$ intersects the x-axis in y and is given by

$$l_1(\zeta) = c^n\big(\xi_n(y), y\big) - \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)}\big(\zeta - \xi_n(y)\big).$$

Consequently, the supporting tangent to $c^n(\cdot, \tilde{y})$ at $\xi_n(y)$ is given by

$$\begin{aligned}
l_2(\zeta) = {}& c^n\big(\xi_n(y), y\big) - \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)}\big(\zeta - \xi_n(y)\big) \\
& + \tilde{y}K_{J_n(y)}(\tilde{y}) - yK_{J_n(y)}(y) + \zeta\big(K_{J_n(y)}(y) - K_{J_n(y)}(\tilde{y})\big).
\end{aligned}$$

Taking $\tilde{y} = y + \delta$, $\delta < \varepsilon$, evaluating at $\zeta = \tilde{y}$ and simplifying we obtain

$$l_2(\tilde{y}) = -\frac{c^n(\xi_n(y), y)}{y - \xi_n(y)}\delta + \delta K_{J_n(y)}(y) = -\delta\frac{c_n(\xi_n(y)) - c_{J_n(y)}(\xi_n(y))}{y - \xi_n(y)} \le 0.$$

By local (strict) convexity of $c^n(\cdot, \tilde{y})$, the supporting tangent in any $\zeta \in [\xi_n(y) - \delta_-, \xi_n(y))$ is strictly negative when evaluated at $\tilde{y}$. By definition, $\xi_n(\tilde{y})$ is such that the supporting tangent at that point is zero in $\tilde{y}$. It follows, since $\xi_n(\tilde{y}) \in [\xi_n(y) - \delta_-, \xi_n(y) + \delta_+]$, that $\xi_n(\tilde{y}) \ge \xi_n(y)$ as required.

Now we relax the assumption (2.32). Assume that there exists a $\delta > 0$ such that $\xi_n(y) > \xi_n(y + \delta)$. We derive a contradiction to the special case as follows. Set $y_0 := y$ and $y_n := y + \delta$. Recall that ξ_n is continuous. Now we can choose $y_0 < y_1 < \cdots < y_{n-1} < y_n$ such that $\xi_n(y_0) > \xi_n(y_1) > \cdots > \xi_n(y_{n-1}) > \xi_n(y_n)$. Set $x_i := \xi_n(y_i)$, $i = 0, \dots, n$. Observe that by monotonicity of ξ_k, $k < n$ the graph of ξ_k intersects with at most one rectangle $(x_i, x_{i-1}) \times (y_{i-1}, y_i)$, $i = 1, \dots, n$. Consequently, there must exist at least one integer j such that the rectangle $R := (x_j, x_{j-1}) \times (y_{j-1}, y_j)$ is disjoint with the graph of every ξ_k, $k < n$. By construction and continuity of $y \mapsto \xi_n(y)$ R is not disjoint with the graph of ξ_n. Inside this rectangle R, the conditions of the special case (2.32) are satisfied. Recalling that $\xi_n(y_j) = x_j < x_{j-1} = \xi_n(y_{j-1})$ and by continuity of $y \mapsto \xi_n(y)$, we can find two points $s_1 < s_2$ such that $z_1 = \xi_n(s_1) > \xi_n(s_2) = z_2$ and $(z_1, s_1) \in R$, $(z_2, s_2) \in R$, which gives the desired contradiction. □

REMARK 2.17 (Properties of $\imath_n$). Observe that, from the definition $\imath_n$ in (2.12) and the above monotonicity result, both $\imath_n(\cdot, y)$ and $\imath_n(x, \cdot)$ are piecewise constant with at most $n-1$ jumps and further for all $y \geq 0$

$$\imath_n(\cdot, y) \quad \text{is left-continuous and nondecreasing} \tag{2.33}$$

and for all $x \in \mathbb{R}$

$$\imath_n(x, \cdot) \quad \text{is right-continuous and nonincreasing.} \tag{2.34}$$

LEMMA 2.18 (ODE for K_n). *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold. Then*

$$y \mapsto K_n(y) \qquad \textit{is locally Lipschitz continuous and nonincreasing}$$

on $(0, r_{\mu_n})$ and in particular is a.e. differentiable. We have

$$K_n'(y) + \frac{K_n(y)}{y - \xi_n(y)} = K_j'(y) + \frac{K_{\jmath_n(y)}(y)}{y - \xi_n(y)} \qquad \textit{for } j = \jmath_n(y), \tag{2.35}$$

for almost all $y \geq 0$. Further,

$$K_n(y) + c_n'\big(\xi_n(y)\big) - c_j'\big(\xi_n(y)\big) - K_j(y) = 0 \qquad \textit{for } j = \jmath_n(y), \tag{2.36}$$

for all y such that $\xi_n(y)$ is not an atom of $\mu_1, \ldots, \mu_n$.

PROOF. The proof is reported in the Appendix. $\square$

3. Proof of the main result. In this section, we prove the main result, Theorem 2.3. The key step is the identification of the distribution of the maximum; cf. Proposition 3.4.

Let $n \in \mathbb{N}$. For convenience, we set

$$M_0 := 0, \qquad M_i := B_{\tau_i}, \qquad \bar{M}_i := \bar{B}_{\tau_i}, \qquad i = 1, \ldots, n, \tag{3.1}$$

where τ_i, given in Definition 2.2, can be represented according to two cases:

$$\tau_i = \begin{cases} \inf\{t \geq \tau_{i-1} : B_t \leq \xi_i(\bar{B}_t)\}, & \text{if } B_{\tau_{i-1}} > \xi_i(\bar{B}_{\tau_{i-1}}), \quad \text{(3.2a)} \\ \tau_{i-1}, & \text{else.} \quad \text{(3.2b)} \end{cases}$$

In the results which follow, we say that $\xi_n(y)$ is strictly increasing at y if either $y = 0$ or else for any $y' < y$ we have $\xi_n(y') < \xi_n(y)$. Put differently, these are the points y such that the left continuous inverse ξ_n^{-1} satisfies $\xi_n^{-1}(\xi_n(y)) = y$. We note that if a property holds for all such y then it holds $d\xi_n(y)$-a.e.

3.1. *Basic properties of the embedding*. Our first result shows that there is a "strong relation" between M and $\bar{M}$.

LEMMA 3.1 (Relations between M and $\bar{M}$). *Let $n \in \mathbb{N}$ and suppose Assumption ⊛ holds. Then the following implications hold*:

$$M_n > \xi_n(y) \implies \bar{M}_n \geq y, \tag{3.3}$$

$$M_n \geq \xi_n(y) \implies \bar{M}_n \geq y \qquad \textit{if } \xi_n \textit{ is strictly increasing at } y. \tag{3.4}$$

For $y \geq 0$ such that $\jmath_n(y) \neq 0$, we have

$$M_{\jmath_n(y)} \geq \xi_n(y) > \xi_{\jmath_n(y)}(y) \implies M_n \geq \xi_n(y), \tag{3.5}$$

$$\bar{M}_{\jmath_n(y)} < y, \bar{M}_n \geq y \implies M_n \geq \xi_n(y), \tag{3.6}$$

$$\bar{M}_{\jmath_n(y)} \geq y, M_{\jmath_n(y)} < \xi_n(y) \implies M_n < \xi_n(y). \tag{3.7}$$

If ξ_n is strictly increasing at $y \geq 0$ and $\jmath_n(y) = 0$, then the following holds:

$$M_n \geq \xi_n(y) \iff \bar{M}_n \geq y. \tag{3.8}$$

PROOF. The results are easily verified for $n = 1$, so we consider $n \geq 2$. Write $\jmath = \jmath_n$. We have

$$\xi_{\jmath(y)}(y) < \xi_n(y) \leq \xi_i(y), \qquad i = \jmath(y) + 1, \ldots, n.$$

In the following, we are using monotonicity of ξ_1 and continuity and monotonicity of $\xi_2, \ldots, \xi_n$; cf. Lemmas 2.15 and 2.16.

Case $\jmath(y) \neq 0$. As for implication (3.3) assume that $M_n > \xi_n(y)$ and $\bar{M}_n < y$ holds. In this case, M_n cannot be at the boundary ξ_n. There has to be a $j < n$ such that $M_n = M_j$, $\bar{M}_n = \bar{M}_j$ and $M_j = \xi_j(\bar{M}_j) = \xi_j(y')$ for some $y' < y$. However, this cannot be true because $\xi_n(y') \leq \xi_n(y) < \xi_j(y') = M_n$, and hence case (3.2a) of the definition of $\tau_1, \ldots, \tau_n$ would have been triggered.

Implication (3.4) follows by the same arguments as for implication (3.3).

Implication (3.5) now follows from implication (3.3) applied for $\jmath(y)$ and the fact that either $M_n = M_{\jmath(y)}$ [case (3.2b)] or M moves to a point at the boundary $\xi_i(y') \geq \xi_n(y)$ for some $i = \jmath(y) + 1, \ldots, n$, $y' \geq y$ [case (3.2a)].

Implication (3.6) holds because if M increases its maximum at time $\jmath(y)$, which is $< y$, to some $y' \geq y$ at time n, it will hit a boundary point $\xi_i(y') \geq \xi_n(y)$ for some $i = \jmath(y) + 1, \ldots, n$.

Implication (3.7) holds because from $\bar{M}_{\jmath(y)} \geq y$ and $M_{\jmath(y)} < \xi_n(y)$ it follows that $M_{\jmath(y)} = \xi_i(y') < \xi_n(y) \leq \xi_j(y')$ for some $i \leq \jmath(y)$, $y' \geq y$, $j > \jmath(y)$. From this, it follows that M will stay where it is until time n; cf. case (3.2b).

Case $\jmath(y) = 0$. Assume that ξ_n is strictly increasing at y and that $\bar{M}_n \geq y$ holds. In this case, M_n must be at a boundary point $\xi_i(y') \geq \xi_n(y)$ for some $i = 1, \ldots, n$, $y' \geq y$. The converse direction is just (3.4), together giving (3.8). □

As an application of Lemma 3.1, we obtain the following result.

LEMMA 3.2 (Contributions to the maximum). *Let $n \in \mathbb{N}$ and suppose Assumption $\circledast$ holds. Fix $y \geq 0$ and assume ξ_n is strictly increasing at y. Then, if $\jmath_n(y) \neq 0$,*

$$\mathbb{P}[\bar{M}_n \geq y] = \mathbb{P}[M_n \geq \xi_n(y)] - \mathbb{P}[M_{\jmath_n(y)} \geq \xi_n(y)] \tag{3.9}$$

$$+ \mathbb{P}[\bar{M}_{\jmath_n(y)} \geq y] \tag{3.10}$$

and if $\jmath_n(y) = 0$,

$$\mathbb{P}[\bar{M}_n \geq y] = \mathbb{P}[M_n \geq \xi_n(y)]. \tag{3.11}$$

PROOF. Write $\jmath = \jmath_n$.

Case $\jmath(y) \neq 0$. First, let us compute

$$\begin{aligned}
&\mathbb{P}[\bar{M}_n \geq y] - \mathbb{P}[M_n \geq \xi_n(y)] \\
&\quad \overset{(3.4)}{=} \mathbb{P}[\bar{M}_n \geq y] - \mathbb{P}[M_n \geq \xi_n(y), \bar{M}_n \geq y] \\
&\quad = \mathbb{P}[\bar{M}_n \geq y, M_n < \xi_n(y)] \\
&\quad = \mathbb{P}[\bar{M}_n \geq y, M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geq y] \\
&\qquad + \mathbb{P}[\bar{M}_n \geq y, M_n < \xi_n(y), \bar{M}_{\jmath(y)} < y] \\
&\quad \underset{(3.6)}{\overset{(3.5)}{=}} \mathbb{P}[M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)].
\end{aligned}$$

Second, let us compute

$$\begin{aligned}
&\mathbb{P}[\bar{M}_{\jmath(y)} \geq y] - \mathbb{P}[M_{\jmath(y)} \geq \xi_n(y)] \\
&\quad = \mathbb{P}[\bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} \geq \xi_n(y)] \\
&\qquad + \mathbb{P}[\bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)] \\
&\qquad - \mathbb{P}[M_{\jmath(y)} \geq \xi_n(y)] \\
&\quad \overset{(3.3)}{=} \mathbb{P}[\bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)] \\
&\quad \overset{(3.7)}{=} \mathbb{P}[M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)].
\end{aligned}$$

Comparing these two equations yields the claim.

Case $\jmath(y) = 0$. The claim follows directly from (3.8). □

3.2. *Law of the maximum.* Our next goal is to identify the distribution of M_n. We will achieve this by deriving an ODE for $\mathbb{P}[\bar{M}_n \geq \cdot]$ using excursion theoretical results (cf. Lemma 3.3), and link it to the ODE satisfied by K_n; cf. Lemma 2.18.

LEMMA 3.3 (ODE for the maximum). *Let $n \in \mathbb{N}$ and suppose Assumption $\circledast$ holds. Then the mapping*

$$y \mapsto \mathbb{P}[\bar{M}_n \geq y]$$

is locally Lipschitz continuous on $(0, r_{\mu_n})$ and a.e. differentiable with

$$\begin{aligned} &\frac{\partial \mathbb{P}[\bar{M}_n \geq y]}{\partial y} + \frac{\mathbb{P}[\bar{M}_n \geq y]}{y - \xi_n(y)} \\ &\quad = \frac{\mathbb{P}[\bar{M}_{\jmath_n(y)} \geq y]}{y - \xi_n(y)} + \frac{\partial \mathbb{P}[\bar{M}_j \geq y]}{\partial y}\Big|_{j=\jmath_n(y)}. \end{aligned} \tag{3.12}$$

PROOF. Write $\jmath = \jmath_n$. The cases $n = 1, 2$ are true by Brown, Hobson and Rogers [4]. Assume by induction hypothesis that we have proven the claim for $i = 1, \dots, n-1$.

Case $\jmath(y) \neq 0$. We have

$$\begin{aligned} &\mathbb{P}[\bar{M}_n \geq y + \delta, \bar{M}_{\jmath(y)} < y + \delta] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y] \\ &\quad = \mathbb{P}[\bar{M}_n \geq y + \delta, \bar{M}_{\jmath(y)} < y] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y] \\ &\qquad + \underbrace{\mathbb{P}[\bar{M}_n \geq y + \delta, y \leq \bar{M}_{\jmath(y)} < y + \delta]}_{=0 \text{ for } \delta > 0 \text{ small enough by definition of } \jmath(y)} \end{aligned} \tag{3.13}$$

since, by right-continuity of ξ_1 and continuity of ξ_i, for $j = \jmath(y)$, we have $\xi_j < \xi_n$ on some open neighborhood of y.

For $r > 0$, we define

$$\bar{\xi}_j(r) := \max_{k: j < k \leq n} \{\xi_k(r) : \xi_k(y) = \xi_n(y)\},$$

$$\underline{\xi}_j(r) := \min_{k: j < k \leq n} \{\xi_k(r) : \xi_k(y) = \xi_n(y)\}$$

and note that

$$\bar{\xi}_{\jmath(y)}(r) \to \xi_n(y), \qquad \underline{\xi}_{\jmath(y)}(r) \to \xi_n(y) \qquad \text{as } r \to y \tag{3.14}$$

by continuity of ξ_i at y for $i = 2, \dots, n$.

Let $\delta > 0$. We have by excursion theoretical result (cf., e.g., Rogers [23]),

$$\begin{aligned} &\mathbb{P}[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y] \exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_{\jmath(y)}(r)}\right) \\ &\quad \leq \mathbb{P}[\bar{M}_n \geq y + \delta, \bar{M}_{\jmath(y)} < y] \\ &\quad \leq \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y] \exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_{\jmath(y)}(r)}\right). \end{aligned} \tag{3.15}$$

Combining the above gives

$$
\begin{aligned}
&\frac{\mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_{J(y)} < y+\delta] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{J(y)} < y]}{\delta} \\
&\overset{(3.13),(3.15)}{\leq} \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{J(y)} < y] \frac{\exp(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r-\underline{\xi}_{J(y)}(r)}) - 1}{\delta} \qquad (3.16)\\
&\xrightarrow[\text{as } \delta\downarrow 0]{\text{by (3.14)}} -\frac{\mathbb{P}[\bar{M}_n \geq y, \bar{M}_{J(y)} < y]}{y-\xi_n(y)}
\end{aligned}
$$

and analogously

$$
\begin{aligned}
&\frac{\mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_{J(y)} < y+\delta] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{J(y)} < y]}{\delta} \\
&\overset{(3.13),(3.15)}{\geq} \mathbb{P}[\bar{M}_n \geq y, \bar{M}_{J(y)} < y] \frac{\exp(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r-\bar{\xi}_{J(y)}(r)}) - 1}{\delta} \qquad (3.17)\\
&\xrightarrow[\text{as } \delta\downarrow 0]{\text{by (3.14)}} -\frac{\mathbb{P}[\bar{M}_n \geq y, \bar{M}_{J(y)} < y]}{y-\xi_n(y)}.
\end{aligned}
$$

Hence, from (3.16) and (3.17) it follows that the right-derivative of

$$
y \mapsto \mathbb{P}[\bar{M}_n \geq y, \bar{M}_j < y]|_{j=J(y)} \tag{3.18}
$$

exists. Similar arguments for $\delta < 0$ show that the left-derivative exists and is the same as the right-derivative. In particular, (3.18) is locally Lipschitz continuous.

Observe the obvious equality

$$
\mathbb{P}[\bar{M}_n \geq y] = \mathbb{P}[\bar{M}_j \geq y] + \mathbb{P}[\bar{M}_n \geq y, \bar{M}_j < y]. \tag{3.19}
$$

Taking $j = J(y)$ in (3.19) and fixing it, we conclude by induction hypothesis that $y \mapsto \mathbb{P}[\bar{M}_n \geq y]$ is locally Lipschitz continuous and a.e. differentiable with

$$
\frac{\partial \mathbb{P}[\bar{M}_n \geq y]}{\partial y} = \frac{\partial \mathbb{P}[\bar{M}_j \geq y]}{\partial y}\Big|_{j=J_n(y)} + \frac{\mathbb{P}[\bar{M}_{J_n(y)} \geq y] - \mathbb{P}[\bar{M}_n \geq y]}{y-\xi_n(y)}.
$$

Case $J(y) = 0$. For $\delta > 0$, we have by excursion theoretical results

$$
\begin{aligned}
&\mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_1 < y+\delta] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_1 < y] \\
&\quad = \mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_1 < y+\delta] - \mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_1 < y] \\
&\qquad + \mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_1 < y] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_1 < y] \qquad (3.20)\\
&\quad \leq \int_y^{y+\delta} \mathbb{P}[\bar{M}_1 \in \mathrm{d}s] \frac{(\xi_1(s) - \underline{\xi}_1(s))^+}{s-\underline{\xi}_1(s)} \exp\biggl(-\int_s^{y+\delta} \frac{\mathrm{d}r}{r-\underline{\xi}_1(r)}\biggr) \\
&\qquad + \mathbb{P}[\bar{M}_n \geq y, \bar{M}_1 < y]\biggl[\exp\biggl(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r-\underline{\xi}_1(r)}\biggr) - 1\biggr].
\end{aligned}
$$

Similarly, we have

$$\begin{aligned}
&\mathbb{P}[\bar{M}_n \geq y+\delta, \bar{M}_1 < y+\delta] - \mathbb{P}[\bar{M}_n \geq y, \bar{M}_1 < y] \\
&\quad \geq \int_y^{y+\delta} \mathbb{P}[\bar{M}_1 \in \mathrm{d}s] \frac{(\xi_1(s) - \bar{\xi}_1(s))^+}{s - \bar{\xi}_1(s)} \exp\left(-\int_s^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_1(r)}\right) \\
&\qquad + \mathbb{P}[\bar{M}_n \geq y, \bar{M}_1 < y]\left[\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_1(r)}\right) - 1\right].
\end{aligned} \tag{3.21}$$

From (3.20) and (3.21), it follows that the right-derivative of

$$y \mapsto \mathbb{P}[\bar{M}_n \geq y, \bar{M}_1 < y] \tag{3.22}$$

exists. Similar arguments for $\delta < 0$ show that the left-derivative exists and is the same as the right-derivative except possibly when $\xi_1(y-) \neq \xi_1(y)$. Local Lipschitz continuity of (3.22) then follows from (3.20) and (3.21). Now we can conclude from (3.19)–(3.21) applied with $j = 1$ that $y \mapsto \mathbb{P}[\bar{M}_n \geq y]$ is locally Lipschitz continuous and a.e. its derivative reads

$$\begin{aligned}
&\frac{\partial \mathbb{P}[\bar{M}_n \geq y]}{\partial y} \\
&\quad \overset{(3.14)}{=} \frac{\partial \mathbb{P}[\bar{M}_1 \geq y]}{\partial y} - \frac{\partial \mathbb{P}[\bar{M}_1 \geq y]}{\partial y} \frac{(\xi_1(y) - \xi_n(y))^+}{y - \xi_n(y)} \\
&\qquad - \frac{\mathbb{P}[\bar{M}_n \geq y] - \mathbb{P}[\bar{M}_1 \geq y]}{y - \xi_n(y)},
\end{aligned}$$

which implies by induction hypothesis

$$\frac{\mathbb{P}[\bar{M}_n \geq y]}{y - \xi_n(y)} + \frac{\partial \mathbb{P}[\bar{M}_n \geq y]}{\partial y} = 0.$$

This completes the proof. □

Finally, we argue that $\mathbb{P}[\bar{M}_n \geq y] = K_n(y)$ holds for all $y \geq 0$.

PROPOSITION 3.4 (Law of the maximum). *Let $n \in \mathbb{N}$ and let Assumption ⊛ hold. Then, for all $y \geq 0$, we have*

$$\mathbb{P}[\bar{M}_n \geq y] = K_n(y). \tag{3.23}$$

PROOF. The case $n = 1$ holds by the Azéma–Yor embedding. Assume by induction hypothesis that

$$K_i(y) = \mathbb{P}[\bar{M}_i \geq y], \qquad i = 1, \dots, n-1; y \geq 0.$$

In Lemmas 2.18 and 3.3, we derived an ODE for K_n and $\mathbb{P}[\bar{M}_n \geq \cdot]$, respectively, in terms of $K_1, \dots, K_{n-1}$ and $\mathbb{P}[\bar{M}_1 \geq \cdot], \dots, \mathbb{P}[\bar{M}_{n-1} \geq \cdot]$, respectively. Taking the difference between the two ODEs and using the induction hypothesis and the equality $K_n(0) = \mathbb{P}[\bar{M}_n \geq 0] = 1$ (cf. Remark 2.8), we see that

$$\begin{aligned}&\big(\mathbb{P}[\bar{M}_n \geq y] - K_n(y)\big)'\\&\quad= -\frac{\mathbb{P}[\bar{M}_n \geq y] - K_n(y)}{y - \xi_n(y)}, \qquad y \in (0, r_{\mu_n}),\end{aligned}$$

subject to $\mathbb{P}[\bar{M}_n \geq 0] - K_n(0) = 0$. Continuity of ξ_n and Gronwall's lemma imply that the above equation has a unique absolutely continuous solution given by 0. This means the desired equality holds for $y < r_{\mu_i}$. This concludes the proof when $r_{\mu_n} = +\infty$ and otherwise we conclude using Remark 2.9. $\square$

3.3. *Embedding property.* In this section, we prove that the stopping times $\tau_1, \dots, \tau_n$ from Definition 2.2 embed the laws $\mu_1, \dots, \mu_n$ if Assumption $\circledast$ is in place. This allows us to complete the proof of our main theorem.

PROOF OF THEOREM 2.3. The case $n = 1$ is just the Azéma–Yor embedding. For the induction hypothesis, assume that the claim holds for all $i \leq n - 1$.

Recall from Remark 2.8 that $\xi_n(0) = l_{\mu_n}$ and $\xi_n(r_{\mu_n}) = r_{\mu_n}$. By continuity, it follows that ξ_n maps $[0, r_{\mu_n}]$ onto $[l_{\mu_n}, r_{\mu_n}]$. Further, by continuity and monotonicity, for a given ζ, the set of y such that $\xi_n(y) = \zeta$ is a closed interval and is $d\xi_n$ negligible. It follows that the set of y such that $\xi_n(y)$ is an atom of $\mu_1, \dots, \mu_n$ is $d\xi_n$ negligible. We then have

$$\begin{aligned}&\mathbb{P}\big[M_n \geq \xi_n(y)\big] - \mathbb{P}\big[M_{J_n(y)} \geq \xi_n(y)\big] + \mathbb{P}[\bar{M}_{J_n(y)} \geq y]\\&\quad\overset{\text{Lemma 3.2}}{=} \mathbb{P}[\bar{M}_n \geq y]\\&\quad\overset{\text{Proposition 3.4}}{=} K_n(y)\\&\quad\overset{(2.36)}{=} -c_n'\big(\xi_n(y)\big) + c_{J_n(y)}'\big(\xi_n(y)\big) + K_{J_n(y)}(y), \qquad d\xi_n\text{-a.e.}\end{aligned}$$

This implies by the induction hypothesis that

$$\mathbb{P}\big[M_n \geq \xi_n(y)\big] = -c_n'\big(\xi_n(y)\big) = \mu_n\big([\xi_n(y), \infty)\big), \qquad d\xi_n\text{-a.e.}$$

The embedding property follows.

Given Proposition 3.4 above, it remains to show the required uniform integrability property. We apply a result from Azéma, Gundy and Yor [2] which states that if

$$\lim_{x\to\infty} x\mathbb{P}\big[|\bar{B}|_{\tau_n} \geq x\big] = 0 \qquad \text{where } |\bar{B}|_t = \sup_{s \leq t} |B_s|, \tag{3.24}$$

then $(B_{\tau_n \wedge t})_{t \geq 0}$ is uniformly integrable.

Let us verify (3.24). Set $H_x = \inf\{t > 0 : B_t = x\}$ and let ξ_i^{-1} denote the left-continuous inverse of ξ_i with $\xi_i^{-1}(x) = 0$ for $x \le l_{\mu_i}$. Then

$$\begin{aligned}\mathbb{P}\big[|\bar{B}|_{\tau_n} \ge x\big] &\le \mathbb{P}[H_{-x} < H_{\max_{i\le n}\xi_i^{-1}(-x)}] + \mathbb{P}[\bar{B}_{\tau_n} \ge x] \\ &= \frac{\max_{i\le n}\xi_i^{-1}(-x)}{x + \max_{i\le n}\xi_i^{-1}(-x)} + K_n(x).\end{aligned}$$

From the definition of ξ_n (cf. Remark 2.8), we have

$$0 \le \max_{i\le n}\xi_i^{-1}(-x) \underset{x\to\infty}{\longrightarrow} 0$$

and hence, recalling the definition of μ_n^{HL} in (2.25),

$$\begin{aligned}\lim_{x\to\infty} x\mathbb{P}\big[|\bar{B}|_{\tau_n} \ge x\big] \le \lim_{x\to\infty} xK_n(x) &\le \lim_{x\to\infty} x\frac{c_n(b_n^{-1}(x))}{x - b_n^{-1}(x)} \\ &= \lim_{x\to\infty} x\mu_n^{\mathrm{HL}}\big([x,\infty)\big) = 0.\end{aligned}$$

This completes the proof. □

4. Discussion of Assumption ⊛ and extensions. In this section, we focus on our main technical assumption so far: the condition (ii) in Assumption ⊛. We construct a simple example of probability measures μ_1, μ_2, μ_3 which violate the condition and where the stopping boundaries ξ_1, ξ_2, ξ_3, obtained via (2.6), fail to embed (μ_1, μ_2, μ_3). It follows that the assumption is not merely technical but does rule out certain type of interdependence between the marginals. If it is not satisfied, then it may not be enough to perturb the measures slightly to satisfy it. The example also provides a counterexample to Brown, Hobson and Rogers [4] showing a silent assumption, which is not automatically satisfied, that $\xi_2(y) < y$ for $y < r_{\mu_2}$ was used in the proofs therein.

We then present an extension of our embedding, in the case $n = 3$, which works in greater generality. More precisely, we show how to modify the optimization problem from which ξ_3 is determined in order to obtain the embedding property. The general embedding, as compared to the embedding in the presence of Assumption ⊛(ii), gains an important degree of freedom and becomes less explicit. In consequence, it is also much harder to implement in practice, to the point that we do not believe this is worth pursuing for $n > 3$. This is also why, as well as for the sake of brevity, we keep the discussion in the section rather formal.

4.1. *Counterexamples for Assumption ⊛(ii).* In Figure 2, we define measures via their potentials

$$U\mu : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto U\mu(x) := -\int_{\mathbb{R}} |u - x|\mu(\mathrm{d}u). \tag{4.1}$$

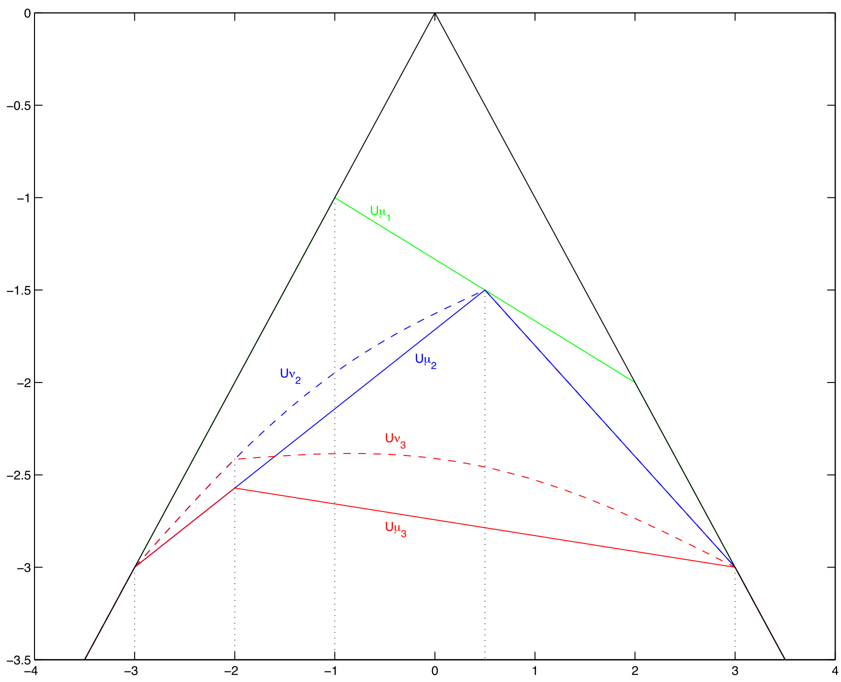

FIG. 2. *Potentials of* $\mu_1, \mu_2, \mu_3, \nu_2$ *and* ν_3.

We refer to Obłój [20], Proposition 2.3, for useful properties of $U\mu$.

The measures with potentials illustrated in Figure 2 are given as

$$\mu_1(\{-1\}) = \frac{2}{3}, \qquad \mu_1(\{2\}) = \frac{1}{3}, \tag{4.2a}$$

$$\mu_2(\{-3\}) = \frac{2}{7}, \qquad \mu_2\left(\left\{\frac{1}{2}\right\}\right) = \frac{18}{35}, \qquad \mu_2(\{3\}) = \frac{1}{5}, \tag{4.2b}$$

$$\mu_3(\{-3\}) = \frac{2}{7}, \qquad \mu_3(\{-2\}) = \frac{9}{35}, \qquad \mu_3(\{3\}) = \frac{16}{35}. \tag{4.2c}$$

Observe that the embedding for (μ_1, μ_2, μ_3) is unique: We write $H_{a,b}$ for the exit time of $[a, b]$ and denote $H_{a,b} \circ \theta_\tau := \inf\{t > \tau : B_t \notin (a, b)\}$. Then the embedding $(\tau_1, \tau_2', \tau_3)$ can be written as

$$\begin{aligned} &\tau_1 = H_{-1,2}, \qquad \tau_2' = H_{-3,\frac{1}{2}} \circ \theta_{\tau_1} \mathbb{1}_{\{B_{\tau_1}=-1\}} + H_{\frac{1}{2},3} \circ \theta_{\tau_1} \mathbb{1}_{\{B_{\tau_1}=2\}}, \\ &\tau_3 = H_{-2,3} \circ \theta_{\tau_2'}. \end{aligned} \tag{4.3}$$

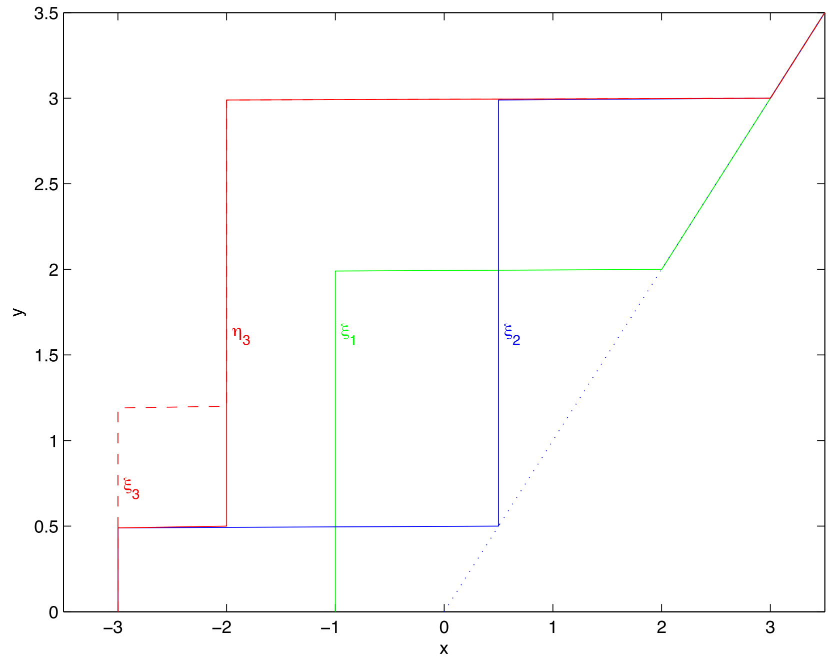

FIG. 3. *We illustrate the* (*unique*) *boundaries* ξ_1, ξ_2, η_3 *required for the embedding of* (μ_1, μ_2, μ_3) *from* (4.2a)–(4.2c) *and the stopping boundary* ξ_3 *obtained from* (2.6). *In order to ensure the embedding for* μ_2, *the mass stopped at* τ_2 *in* -1 *on the event* $\{\bar{B}_{\tau_2} \in (1/2, 2)\}$ *is diffused to* -3 *or to* $1/2$ *at* τ_2', *without affecting the maximum*: $\bar{B}_{\tau_2} = \bar{B}_{\tau_2'}$. *Note that the case* $\xi_2(y) = y$, *here for* $y = 1/2$, *is possible and required to define the embedding. After* τ_2', *we need to define* τ_3 *which embeds* μ_3 *which here is implied directly by* (4.3). *In Section* 4.2, *we develop arguments which generalize this.*

As mentioned earlier, our construction yields the same first two stopping boundaries as the method of Brown, Hobson and Rogers [4]. In this case (cf. Figure 3),

$$\xi_1(y) := \begin{cases} -1, & \text{if } y \in [0, 2), \\ y, & \text{else,} \end{cases} \qquad \xi_2(y) := \begin{cases} -3, & \text{if } y \in \left[0, \dfrac{1}{2}\right), \\ \dfrac{1}{2}, & \text{if } y \in \left[\dfrac{1}{2}, 3\right), \\ y, & \text{else.} \end{cases}$$

This already shows that our embedding fails to embed μ_2. This is easily seen comparing the definition of τ_2 in (2.8) which uses ξ_2 above with (4.3). In Section 4.2, we will recall from Brown, Hobson and Rogers [4] how the stopping time τ_2 has to be modified into τ_2', giving the stopping time in (4.3).

In consequence, the optimization problem (2.6), or equivalently (2.15), is set up wrongly and does not return the third (unique) stopping boundary which is required for the embedding of (μ_1, μ_2, μ_3). Indeed, a direct computation shows that $\frac{c^3(\zeta, y)}{y-\zeta}$ is minimized for $\zeta < \xi_1(y) \wedge \xi_2(y)$ which turns the problem into the

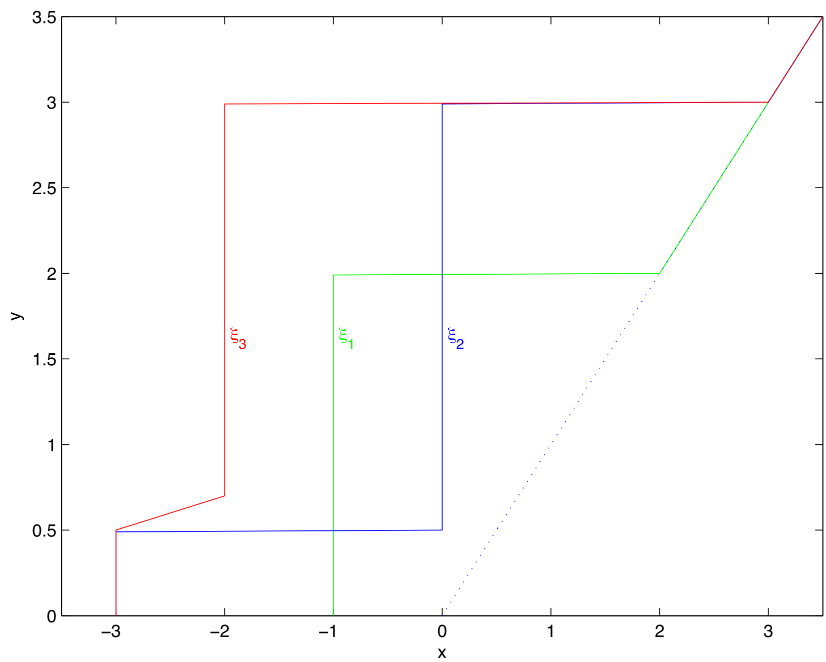

FIG. 4. *Example of boundaries* ξ_1, ξ_2, ξ_3 *for which the associated stopping times* $(\tau_1, \tau_2', \tau_3)$ *give measures* (μ_1, μ_2, μ_3) *which satisfy Assumption* ⊛(i) *but not Assumption* ⊛(ii).

minimization of $\zeta \mapsto \frac{c_3(\zeta)}{y-\zeta}$ which is attained by $\xi_3(y) = -3 < -2$, for $y \in [0, 1.2)$ due an atom of μ_3 at -3. Consequently, there will be a positive probability to hit -3 after τ_2'. This contradicts (4.2c). This, together with the correct third boundary η_3, is illustrated in Figure 3. The error is coming from the fact that our optimization problem fails to account for the additional diffusion of mass between τ_2 and τ_2' which was necessary to embed μ_2 correctly.

The above example is singular with the measures violating both Assumption ⊛(i), since the call prices are not strictly ordered, as well as ⊛(ii), and further the embedding being unique. Small modifications lead to more "regular" examples and also show that a "small perturbation" to (μ_1, μ_2, μ_3) may not be enough to remove the problem. Indeed, first consider measures (μ_1, ν_2, ν_3) defined by their potentials in Figure 2. The embedding is no longer unique but a similar reasoning to the one above holds and our optimization procedure still returns ξ_3 which embeds an incorrect marginal. Second, consider boundaries (ξ_1, ξ_2, ξ_3) given in Figure 4 and let τ_1, τ_2 be defined by (2.8),

$$\tau_2' = \inf\{t \geq \tau_2 : B_t \in \{-3, 0\}\} \qquad \text{on } \{\bar{B}_{\tau_1} = \bar{B}_{\tau_2} \in (0.5, 2)\}$$

and $\tau_2' = \tau_2$ elsewhere. We put $\tau_3 = \inf\{t \geq \tau_2' : B_t \leq \xi_3(\bar{B}_t)\}$ and let μ_1, μ_2, μ_3 be the distributions of B_{τ_1}, $B_{\tau_2'}$ and B_{τ_3}, respectively. By construction we see that the measures satisfy Assumption ⊛(i). However, using Definition 2.1 to derive stopping boundaries will reproduce ξ_1 and ξ_2 but will still yield a wrong ξ_3 as the optimization problem will not account for mass embedded in -3 by τ_2'. In Section 4.2, we develop arguments which generalize (2.6) accordingly.

Finally, we note that in the first example with (μ_1, μ_2, μ_3) in (4.2a)–(4.2c), where the embedding is unique, we have $\xi_2(1/2) = 1/2$ and one can verify that $\mathbb{P}(\bar{M}_2 = 1/2) = 4/21$, where we use the notation in (3.1). It particular, the function $y \to \mathbb{P}(\bar{M}_2 \geq y)$ is not locally Lipschitz continuous at $y = 1/2$, and hence Lemmas 3.4 and 3.6 in Brown, Hobson and Rogers [4] cannot hold true.[5]

4.2. *Sketch for general embedding in the case $n = 3$.* In the example of the measures (μ_1, μ_2, μ_3) from (4.2a)–(4.2c), the (unique) embedding could still be seen as a type of "iterated Azéma–Yor type embedding" although it does not satisfy the relations from Lemma 3.1. Consequently, one might conjecture that a modification of the optimization problem (2.6) and a relaxation of Lemma 3.1 might lead to a generally applicable embedding. We now explain in which sense this is true. Our aim is to outline new ideas and arguments which are needed. The technical details quickly become very involved and lengthy. In the sake of brevity, but also to better illustrate the main points, we restrict ourselves to a formal discussion and the case $n = 3$.

Consider now a case of general measures (μ_1, μ_2, μ_3) where Assumption $\circledast$(ii) is possibly violated. Recall from Brown, Hobson and Rogers [4] that in general, if $M_1 \leq \xi_2(\bar{M}_1)$, so that $\tau_2 = \tau_1$, but $\xi_2(y-) < M_1 < \xi_2(y)$ for some $y \leq \bar{M}_1$ then this mass is further diffused to a stopping time $\tau_2' \geq \tau_2$ with $\bar{B}_{\tau_2} = \bar{B}_{\tau_2'}$. In the case of measures in (4.2a)–(4.2c), we have $\tau_2' > \tau_2$ on $\bar{B}_{\tau_1} \in (\frac{1}{2}, 2]$. The existence of τ_2' is established by Brown, Hobson and Rogers [4] by showing that the relative parts of the mass which are further diffused have the same mass, mean and are in convex order. In general, there will be infinitely many such stopping times τ_2'. Although this is not true for measures in (4.2a)–(4.2c) because their embedding was unique, it is true for measures (μ_1, ν_2, ν_3) which are defined via their potentials in Figure 2.

Let ξ_1 and ξ_2 be defined as in (2.6) and let $\tau_2^{\mathrm{BHR}} = \tau_2'$ be the general second stopping time in [4]. We then redefine (3.1) putting $M_2 = B_{\tau_2^{\mathrm{BHR}}}$ and $\bar{M}_2 = \bar{B}_{\tau_2^{\mathrm{BHR}}}$. Now our goal is to define an embedding $\tilde{\tau}_3$ for the third marginal based on some stopping boundary $\tilde{\xi}_3$ as a first exit time,

$$\tilde{\tau}_3 := \begin{cases} \inf\{t \geq \tau_2^{\mathrm{BHR}} : B_t \leq \tilde{\xi}_3(\bar{B}_t)\}, & \text{if } B_{\tau_2^{\mathrm{BHR}}} > \tilde{\xi}_3(\bar{B}_{\tau_2^{\mathrm{BHR}}}), \\ \tau_2^{\mathrm{BHR}}, & \text{else,} \end{cases} \tag{4.4}$$

[5] Upon inspection, one sees that the proof of Lemma 3.6 therein implicitly uses the fact that $\xi_2(y) < y$ which does not hold in all generality. It seems that to complement the analysis in [4] one would need to argue that $y \to \mathbb{P}(\bar{M}_2 \geq y)$ and K_y have the same jumps which becomes involved due to different scenarios of ξ_2 either jumping to the diagonal or "creeping" continuously to the diagonal.

and prove that this is a valid embedding of μ_3. We observe that now the choice of τ_2' in the definition of τ_2^{BHR} may matter for the subsequent embedding.[6] Similarly, as in the embedding of Brown, Hobson and Rogers [4], we expect that this will be only possible if the procedure which produces $\tilde{\xi}_3$ yields a continuous $\tilde{\xi}_3$. Otherwise, an additional step, producing a stopping time $\tau_3' \geq \tilde{\tau}_3$ would be required and further complicate the presentation.

With this, a more canonical approach in the context of Lemma 3.1 is to write

$$\mathbb{P}[\bar{M}_3 \geq y] = \mathbb{P}[M_3 \geq \tilde{\xi}_3(y)] + \text{“error-term,”} \tag{4.5}$$

which we formalize in (4.23). As it will turn out, this “error-term” provides a suitable “book-keeping procedure” to keep track of the masses in the embedding. We proceed along the lines of the proof of our main result. For simplicity, we further assume that ξ_2 has only one discontinuity, that is, $\underline{z} := \xi_2(\underline{y}-) < \xi_2(\underline{y}+) := \bar{z}$ for some $\underline{y} \geq 0$ and we let $\bar{y} := \xi_1^{-1}(\bar{z})$. As explained below, this is not restrictive since our procedure is localized. If $\bar{y} \leq \underline{y}$, then μ_1 can be “ignored” and the results of Brown, Hobson and Rogers [4] apply. Hence, we assume $\bar{y} > \underline{y}$.

4.2.1. *Redefining ξ_3 and K_3.* Define the following auxiliary terms:

$$F(\zeta, y; \tau_2') := \mathbb{1}_{\{\bar{M}_1 \geq y\}}(\zeta - M_2)^+, \tag{4.6}$$

$$f^{\mathrm{iAY}}(\zeta, y; \tau_2') := \mathbb{E}[F(\zeta, y; \tau_2')]. \tag{4.7}$$

As the notation underlines, these quantities may depend on the additional choice of stopping time τ_2' between τ_2 and τ_3. Note that for $\zeta \in [\underline{z}, \bar{z}]$ and $y \in [\underline{y}, \bar{y}]$,

$$\frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\zeta, y; \tau_2') = \mathbb{P}[\bar{M}_1 \geq y, M_2 < \zeta] \tag{4.8}$$

and

$$\begin{aligned} \frac{\partial f^{\mathrm{iAY}}}{\partial y}(\zeta, y; \tau_2') &= -\mathbb{E}\left[\frac{\mathbb{1}_{\{\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\}}}{\mathrm{d}y}\right]\zeta + \mathbb{E}\left[\frac{\mathbb{1}_{\{\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\}}}{\mathrm{d}y} M_2\right] \\ &= -(\zeta - \alpha(\zeta, y; \tau_2'))\frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta]}{\mathrm{d}y}, \end{aligned} \tag{4.9}$$

where

$$\alpha(\zeta, y; \tau_2') := \mathbb{E}[M_2 | \bar{M}_1 = y, M_2 < \zeta], \tag{4.10a}$$

$$\beta(\zeta, y; \tau_2') := \mathbb{E}[M_2 | \bar{M}_1 = y, M_2 \geq \zeta]. \tag{4.10b}$$

---

[6]Appropriate choice of τ_2' will also be crucial to maintain optimality of the embedding along the lines on Henry-Labordère et al. [13]. A natural conjecture will be to use first hitting times, with boundaries depending on the level of $\bar{M}_{\tau_2}$ and exhibiting some monotonicity. A recent work of Beiglböck, Cox and Huesmann [3], and also Guo, Tan and Touzi [12], provides new insights which might allow to answer this question in all generality.

With these definitions, we have by the properties of τ_2'

$$\begin{aligned}
&\alpha(\zeta, y; \tau_2')\frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta]}{\mathrm{d}y} + \beta(\zeta, y; \tau_2')\frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y, M_2 \geq \zeta]}{\mathrm{d}y} \\
&\quad = \xi_1(y)\frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y]}{\mathrm{d}y}.
\end{aligned} \tag{4.11}$$

We now redefine ξ_3 and K_3 from (2.15) and (2.7), respectively, and denote the new definition by $\tilde{\xi}_3$ and $\tilde{K}_3$. To this end, introduce the function

$$\tilde{c}^3(\zeta, y) := \begin{cases} c_3(\zeta) - f^{\mathrm{iAY}}(\zeta, y; \tau_2'), & \text{if } \underline{z} \leq \zeta \leq \bar{z}, \underline{y} \leq y \leq \bar{y}, \quad (4.12a) \\ c^3(\zeta, y), & \text{else.} \quad (4.12b) \end{cases}$$

We have that $\tilde{c}^3$ is continuous and $\tilde{c}^3 \leq c^3$. The inequality follows from the properties of τ_2' since for $\zeta \in [\underline{z}, \bar{z}]$ and $y \in [\underline{y}, \bar{y}]$ we have

$$\begin{aligned}
f^{\mathrm{iAY}}(\zeta, y; \tau_2') &= \mathbb{E}\big[(\zeta - M_2)^+ \mathbb{1}_{\{\bar{M}_1 \geq y\}}\big] = \mathbb{E}\big[\big\{(M_2 - \zeta)^+ - (y - \zeta)\big\}\mathbb{1}_{\{\bar{M}_1 \geq y\}}\big] \\
&= \mathbb{E}\big[(M_2 - \zeta)^+ \mathbb{1}_{\{\bar{M}_1 \geq y\}}\big] - (y - \zeta)K_1(y) \\
&\geq \begin{cases} c_1(\zeta) - (y - \zeta)K_1(y), & \text{if } \zeta > \xi_1(y), \\ 0, & \text{else.} \end{cases}
\end{aligned} \tag{4.13}$$

Continuity inside the region is clear and on the boundaries we check it as follows. First, we have $f^{\mathrm{iAY}}(\underline{z}, y; \tau_2') = 0$ and $f^{\mathrm{iAY}}(\zeta, \bar{y}; \tau_2') = 0$, as required. For $\zeta = \bar{z}$, we have

$$\mathbb{E}\big[(M_2 - \zeta)^+ \mathbb{1}_{\{\bar{M}_1 \geq y\}}\big] = \mathbb{E}\big[(M_1 - \zeta)^+\big]$$

so by the above $\tilde{c}^3(\bar{z}, y) = c^3(\bar{z}, y)$. Finally, for $y = \underline{y}$, we compute

$$\mathbb{E}\big[(\zeta - M_2)^+ \mathbb{1}_{\{\bar{M}_2 \geq \underline{y}\}}\big] = \mathbb{E}\big[(\zeta - M_2)^+ \mathbb{1}_{\{\bar{M}_1 \geq \underline{y}\}}\big] \qquad \text{which then implies that}$$

$$\mathbb{E}\big[(M_2 - \zeta)^+ \mathbb{1}_{\{\bar{M}_1 \geq \underline{y}\}}\big] = c_2(\zeta) - (\underline{y} - \zeta)\big(K_2(\underline{y}) - K_1(\underline{y})\big),$$

which combined with (4.13) and continuity of c^3 establishes the claim.

As before, let

$$\tilde{\xi}_3(y) := \arg\min_{\zeta \leq y} \frac{\tilde{c}^3(\zeta, y)}{y - \zeta} \tag{4.14}$$

and

$$\tilde{K}_3(y) := \frac{\tilde{c}^3(\tilde{\xi}_3(y), y)}{y - \tilde{\xi}_3(y)}. \tag{4.15}$$

It is clear that a discontinuity of ξ_2 results in a local perturbation of c^3 into $\tilde{c}^3$ and in consequence of ξ_3 into $\tilde{\xi}_3$. If ξ_2 has multiple discontinuities the construction above applies to each of them giving a global definition of $\tilde{c}^3$. Then $\tilde{K}_3$ and $\tilde{\xi}_3$ are defined as above.

4.2.2. *Law of the maximum.* In the following, we assume that $y \in [\underline{y}, \bar{y}]$ and $\tilde{\xi}_3(y), \zeta \in [\underline{z}, \bar{z}]$. Otherwise, $\tilde{c}^3 = c^3$ and the arguments from Sections 2 and 3 apply. We have $\tilde{\xi}_3(y) < \xi_2(y)$ and $\bar{M}_1 = \bar{M}_2$ on $\{\bar{M}_1 \in [\underline{y}, \bar{y}]\}$.

Note the obvious decomposition

$$\mathbb{P}[\bar{M}_3 \geq y] = \mathbb{P}[\bar{M}_1 < y, \bar{M}_3 \geq y] + \mathbb{P}[\bar{M}_1 \geq y].$$

We compute by similar excursion theoretical arguments as in the proof of Lemma 3.3,

$$\begin{aligned} &\frac{\partial \mathbb{P}[\bar{M}_1 < y, \bar{M}_3 \geq m]}{\partial y}\Big|_{m=y} \\ &\quad =: p\big(\tilde{\xi}_3(y), y; \tau_2'\big) \qquad (4.16)\\ &\quad = \frac{\mathbb{P}[M_2 \geq \tilde{\xi}_3(y), \bar{M}_1 \in \mathrm{d}y]}{\mathrm{d}y} \frac{\beta(\tilde{\xi}_3(y), y; \tau_2') - \tilde{\xi}_3(y)}{y - \tilde{\xi}_3(y)}. \end{aligned}$$

In analogy to (3.13), and because $\tilde{\xi}_3(y) < \xi_2(y)$,

$$\frac{\partial \mathbb{P}[\bar{M}_1 < m, \bar{M}_3 \geq y]}{\partial y}\Big|_{m=y} = -\frac{\mathbb{P}[\bar{M}_3 \geq y] - \mathbb{P}[\bar{M}_1 \geq y]}{y - \tilde{\xi}_3(y)}.$$

Hence, combining the above

$$\begin{aligned} &\frac{\partial}{\partial y}\mathbb{P}[\bar{M}_3 \geq y] \\ &\quad = p\big(\tilde{\xi}_3(y), y; \tau_2'\big) - \frac{\mathbb{P}[\bar{M}_3 \geq y] - \mathbb{P}[\bar{M}_1 \geq y]}{y - \tilde{\xi}_3(y)} + \frac{\partial \mathbb{P}[\bar{M}_1 \geq y]}{\partial y} \qquad (4.17)\\ &\quad \overset{(3.12)}{=} -\frac{\mathbb{P}[\bar{M}_3 \geq y]}{y - \tilde{\xi}_3(y)} - \frac{\tilde{\xi}_3(y) - \xi_1(y)}{y - \tilde{\xi}_3(y)} \frac{\partial \mathbb{P}[\bar{M}_1 \geq y]}{\partial y} + p\big(\tilde{\xi}_3(y), y; \tau_2'\big). \end{aligned}$$

In the redefined domain the first-order condition for optimality of $\tilde{\xi}_3(y)$ reads

$$\tilde{K}_3(y) + c_3'\big(\tilde{\xi}_3(y)\big) - \frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}\big(\tilde{\xi}_3(y), y; \tau_2'\big) = 0. \qquad (4.18)$$

By similar calculations as in (A.9) below, we have

$$\begin{aligned} \tilde{K}_3'(y) &\overset{(4.18)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} - \frac{\frac{\partial f^{\mathrm{iAY}}}{\partial y}(\tilde{\xi}_3(y), y; \tau_2')}{y - \tilde{\xi}_3(y)} \\ &\overset{(4.9)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} + \frac{\tilde{\xi}_3(y) - \alpha(\tilde{\xi}_3(y), y; \tau_2')}{y - \tilde{\xi}_3(y)} \frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y, M_2 < \tilde{\xi}_3(y)]}{\mathrm{d}y} \\ &\overset{(4.11)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} + \frac{\tilde{\xi}_3(y) - \xi_1(y)}{y - \tilde{\xi}_3(y)} \frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y]}{\mathrm{d}y} \qquad (4.19) \end{aligned}$$

$$+\frac{\beta(\tilde{\xi}_3(y), y; \tau_2') - \tilde{\xi}_3(y)}{y - \tilde{\xi}_3(y)} \frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y, M_2 \geq \tilde{\xi}_3(y)]}{\mathrm{d}y}$$

$$\overset{(4.16)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} - \frac{\tilde{\xi}_3(y) - \xi_1(y)}{y - \tilde{\xi}_3(y)} \frac{\partial \mathbb{P}[\bar{M}_1 \geq y]}{\partial y} + p(\tilde{\xi}_3(y), y; \tau_2').$$

Consequently, by comparing (4.17) and (4.19), and in conjunction with Proposition 3.4, we obtain

$$\tilde{K}_3(y) = \mathbb{P}[\bar{M}_3 \geq y], \qquad \forall y \geq 0. \tag{4.20}$$

4.2.3. *Embedding property.* Having found the distribution of the maximum, the final step is to prove the embedding property. To achieve this, we will need that $\tilde{\xi}_3$ is nondecreasing.

Recall the first-order condition of optimality of $\tilde{\xi}_3$ in (4.18) and then the second-order condition for optimality of $\tilde{\xi}_3(y)$ reads

$$c_3''(\tilde{\xi}_3(y)) - \frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta^2}(\tilde{\xi}_3(y), y; \tau_2') \geq 0. \tag{4.21}$$

Now, differentiating (4.18) in y yields

$$\tilde{K}_3'(y) + c_3''(\tilde{\xi}_3(y))\tilde{\xi}_3'(y) - \frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta^2}(\tilde{\xi}_3(y), y; \tau_2')\tilde{\xi}_3'(y)$$
$$- \frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta\, \partial y}(\tilde{\xi}_3(y), y; \tau_2') = 0$$

or equivalently,

$$\tilde{\xi}_3'(y) \underbrace{\left(c_3''(\tilde{\xi}_3(y)) - \frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta^2}(\tilde{\xi}_3(y), y; \tau_2')\right)}_{\geq 0 \text{ by } (4.21)}$$
$$= -\tilde{K}_3'(y) + \frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta\, \partial y}(\tilde{\xi}_3(y), y; \tau_2').$$

In order to formally infer

$$\tilde{\xi}_3'(y) \geq 0$$

we require

$$-\tilde{K}_3'(y) + \frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta\, \partial y}(\tilde{\xi}_3(y), y; \tau_2') \geq 0. \tag{4.22}$$

Direct computation shows that

$$\frac{\partial^2 f^{\mathrm{iAY}}}{\partial \zeta\, \partial y}(\zeta, y; \tau_2') = -\frac{\mathbb{P}[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta]}{\mathrm{d}y}$$

and by (4.20),

$$-\tilde{K}_3'(y) = \frac{\mathbb{P}[\bar{M}_3 \in \mathrm{d}y]}{\mathrm{d}y},$$

which implies (4.22), and hence that $\tilde{\xi}_3$ is nondecreasing.

By definition of the embedding in (4.4), and since $\tilde{\xi}_3$ is nondecreasing, we have

$$\begin{aligned}
\mathbb{P}[\bar{M}_3 \geq y] &= \mathbb{P}\big[M_3 \geq \tilde{\xi}_3(y)\big] + \mathbb{P}\big[\bar{M}_3 \geq y, M_3 < \tilde{\xi}_3(y)\big] \\
&= \mathbb{P}\big[M_3 \geq \tilde{\xi}_3(y)\big] + \mathbb{P}\big[\bar{M}_1 \geq y, M_2 < \tilde{\xi}_3(y)\big] \\
&\overset{(4.8)}{=} \mathbb{P}\big[M_3 \geq \tilde{\xi}_3(y)\big] + \frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}\big(\tilde{\xi}_3(y), y; \tau_2'\big)
\end{aligned} \tag{4.23}$$

and then, by (4.20), (4.18) and (4.23),

$$-c_3'\big(\tilde{\xi}_3(y)\big) = \mathbb{P}\big[M_3 \geq \tilde{\xi}_3(y)\big],$$

which is the desired embedding property.

The above construction hinged on the appropriate choice of the auxiliary term F in (4.6) whose expectation allows for the error book keeping, as suggested in (4.5). We identified the correct F by analysing the "error terms" which cause strict inequality for $(B_u : u \leq \tau_2')$ in the pathwise inequality (4.1) of Henry-Labordère et al. [13]. This is natural since this inequality is used to prove optimality of our embedding. It gives an upper bound but fails to be sharp if condition (ii) in Assumption $\circledast$ does not hold. In order to recover a sharp bound one has to look at the error terms causing strict inequality when Assumption $\circledast$ fails. The same principle applies for $n > 3$. However, then interactions between discontinuities of boundaries $\xi_2, \tilde{\xi}_3$, etc. come into play and the relevant terms become very involved. The construction would become increasingly technical and implicit and we decided to stop at this point.

APPENDIX: PROOF OF LEMMA 2.18

In order to prove Lemma 2.18, we require to prove, inductively, several auxiliary results along the way. We now state and prove a Lemma which contains the statement of Lemma 2.18.

LEMMA A.1. *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold.*

For $\zeta \in \mathbb{R}$, the mapping

$$y \mapsto c^n(\zeta, y) \qquad \textit{is locally Lipschitz continuous and nondecreasing}$$

on $(\zeta \vee 0, \infty)$. Further, for almost all $y \geq 0$,

$$\frac{\partial c^n}{\partial y}(\zeta, y)\Big|_{\zeta = \xi_n(y)} = K_{J_n(y)}(y) + \big(y - \xi_n(y)\big)K'_{J_n(y)}(y), \tag{A.1}$$

where K_j' denotes the derivative of K_j.

The mapping $c^n(\cdot, y)$ is locally Lipschitz continuous on $(-\infty, y)$ and

$$K_n(y) + c_n'\big(\xi_n(y)\big) - c_j'\big(\xi_n(y)\big) - K_j(y) = 0 \qquad \text{for } j = \jmath_n(y) \tag{A.2}$$

and all y such that $\xi_n(y)$ is not an atom of $\mu_1, \dots, \mu_n$.

The mapping

$$y \mapsto K_n(y) \qquad \text{is locally Lipschitz continuous and nonincreasing}$$

on $(0, r_{\mu_n})$, $K_n(y) = 0$ for $y > r_{\mu_n}$, and for almost all $y \geq 0$ we have

$$K_n'(y) + \frac{K_n(y)}{y - \xi_n(y)} = K_{\jmath_n(y)}'(y) + \frac{K_{\jmath_n(y)}(y)}{y - \xi_n(y)}. \tag{A.3}$$

PROOF. We prove the claim by induction over n. The induction basis $n = 1$ holds by definition, representation of ξ_1 in (2.18)–(2.19) and Lemma 2.6 of Brown, Hobson and Rogers [4]. Now assume that the claim holds for all $i = 1, \dots, n-1$.

Recall from Remark 2.17 that $\imath_n(\zeta, \cdot)$ is piecewise constant with at most $n-1$ jumps. Consider y from an interval of constancy of $\imath_n(\zeta, \cdot)$ where $i = \imath_n(\zeta, y)$. We have, using (2.14),

$$c^n(\zeta, y) = (y - \zeta)K_i(y) + c_n(\zeta) - c_i(\zeta). \tag{A.4}$$

It follows by induction hypothesis that $c^n(\zeta, \cdot)$ is locally Lipschitz continuous. Further, $y \to c^n(\zeta, y)/(y - \zeta)$ is nonincreasing. These holds on each interval of constancy of $\imath_n(\zeta, \cdot)$, and hence, by continuity of $c^n(\cdot, \cdot)$ in Lemma 2.15, for all $y > \zeta$. Analogous arguments show that $c^n(\cdot, y)$ is locally Lipschitz continuous on $(-\infty, y)$.

Analogously to the above, to show that $c^n(\zeta, \cdot)$ is nondecreasing it suffices to show $c^n(\zeta, y) \leq c^n(\zeta, \tilde{y})$ for $y \leq \tilde{y}$ in an interval of constancy of $\imath_n(\zeta, \cdot)$. We then have $\imath_n(\zeta, y) = \imath_n(\zeta, \tilde{y}) = i$ and in particular $\xi_i(\tilde{y}) < \zeta$. It follows that

$$\begin{aligned} c^n(\zeta, y) - c^n(\zeta, \tilde{y}) &= (y - \zeta)K_i(y) - (\tilde{y} - \zeta)K_i(\tilde{y}) \\ &= \big(y - \xi_i(\tilde{y})\big)K_i(y) - \big(\tilde{y} - \xi_i(\tilde{y})\big)K_i(\tilde{y}) \\ &\quad + \big(\zeta - \xi_i(\tilde{y})\big)\big(K_i(\tilde{y}) - K_i(y)\big) \\ &\overset{(2.15)}{\leq} c^i\big(\xi_i(\tilde{y}), y\big) - c^i\big(\xi_i(\tilde{y}), \tilde{y}\big) \\ &\quad + \big(\zeta - \xi_i(\tilde{y})\big)\big(K_i(\tilde{y}) - K_i(y)\big) \leq 0, \end{aligned}$$

by the induction hypothesis, using monotonicity of $c^i(\xi_i(\tilde{y}), \cdot)$ and the fact that K_i is nonincreasing. This establishes monotonicity of $c^n(\zeta, \cdot)$.

To prove (A.1), first note that by monotonicity, for any $\zeta \in \mathbb{R}$, the right and left derivatives of $c^n(\zeta, \cdot)$ exist everywhere on $(\zeta \vee 0, \infty)$ and agree almost everywhere. Recalling that $\imath_n(\zeta, \cdot)$ is right-continuous [see (2.34)], the right-derivative is computed directly from the above expression for $c^n(\zeta, y)$ and is given by

$$K_{\jmath_n(y)}(y) + (y - \zeta)K_{\jmath_n(y)}'(y+), \tag{A.5}$$

where the last derivate, by induction hypothesis, is well defined everywhere. Noting that $K'_{\jmath_n(y)}(y+) = K'_{\jmath_n(y)}(y)$ a.e., we see that (A.1) specifies the right derivative dy-a.e. Similarly, for almost all y such that $\tilde{y} \to \imath_n(\xi_n(y), \tilde{y})$ does not jump in $\tilde{y} = y$ the left derivative is also given by the same expression. Otherwise, we compute the left derivative directly, writing $k = \lim_{\tilde{y} \nearrow y} \imath_n(\xi_n(y), \tilde{y}) > \imath_n(\xi_n(y), y) = \jmath_n(y) = \jmath_k(y)$,

$$\begin{aligned}
&\lim_{\delta \uparrow 0} \frac{1}{\delta}\big(-c_k\big(\xi_n(y)\big) + \big(y + \delta - \xi_n(y)\big)K_k(y+\delta) \\
&\qquad + c_{\jmath_n(y)}\big(\xi_n(y)\big) - \big(y - \xi_n(y)\big)K_{\jmath_n(y)}(y)\big) \\
&\quad \overset{\xi_n(y)=\xi_k(y)}{=} \lim_{\delta \uparrow 0} \frac{1}{\delta}\big(\big(y + \delta - \xi_n(y)\big)K_k(y+\delta) - \big(y - \xi_n(y)\big)K_k(y)\big) \qquad \text{(A.6)} \\
&\quad = K_k(y) + \big(y - \xi_n(y)\big)K'_k(y-) \\
&\quad \overset{\text{(A.3)}}{=} K_{\jmath_n(y)}(y) + \big(y - \xi_n(y)\big)K'_{\jmath_n(y)}(y),
\end{aligned}$$

dy-a.e., by induction hypothesis since $n > k > \jmath_n(y) = \jmath_k(y)$ and $\xi_n(y) = \xi_k(y)$. We conclude that for almost all $y > 0$, $\tilde{y} \to c^n(\xi_n(y), \tilde{y})$ is differentiable at $\tilde{y} = y$ with derivative given by (A.1), as required.

We move to properties of K_n. As observed above, the mapping $y \mapsto \frac{c^n(\zeta, y)}{y - \zeta}$ is nonincreasing, and hence for $\delta > 0$

$$\begin{aligned}
K_n(y+\delta) &= \inf_{\zeta \le y+\delta} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} \le \inf_{\zeta \le y} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} \\
&\le \inf_{\zeta \le y} \frac{c^n(\zeta, y)}{y-\zeta} = K_n(y)
\end{aligned}$$

proving that K_n is nonincreasing.

Using that ξ_n is continuous, we may take δ small enough so $\xi_n(y+\delta) < y$ and then, using monotonicity of $c^n(\zeta, \cdot)$ we obtain

$$\begin{aligned}
K_n(y) &\le \frac{c^n(\xi_n(y+\delta), y)}{y - \xi_n(y+\delta)} \le \frac{c^n(\xi_n(y+\delta), y+\delta)}{y - \xi_n(y+\delta)} \\
&= K_n(y+\delta)\left(1 + \frac{\delta}{y - \xi_n(y+\delta)}\right),
\end{aligned}$$

from which the local Lipschitz continuity of K_n follows.

In order to prove (A.2), we consider $y \ge 0$ such that $\xi_n(y)$ is not an atom of any $\mu_1, \dots, \mu_n$. In particular, we have that c_i are differentiable at $\xi_n(y)$. If y is such that $\jmath_n(y) = \imath_n(\xi_n(y), y) = \imath_n(\xi_n(y)+, y)$, that is, $\xi_n(y)$ is in the interior of an interval of constancy of $\imath_n(\cdot, y)$, then the representation (A.4) holds for ζ in a neighborhood of $\xi_n(y)$. By its definition, $\xi_n(y)$ is (the unique) minimizer in (2.15). Using the first-order condition and (2.17) yields directly (A.2).

Suppose now $j := \jmath_n(y) = \imath_n(\xi_n(y), y) < \imath_n(\xi_n(y)+, y) =: k$, which implies in particular that $\xi_j(y) < \xi_k(y) = \xi_n(y)$. Then the first-order condition for $\xi_n(y)$ tells us that

$$\frac{\partial}{\partial\zeta}\frac{c^n(\zeta, y)}{y-\zeta}\Big|_{\zeta=\xi_n(y)-} \le 0 \le \frac{\partial}{\partial\zeta}\frac{c^n(\zeta, y)}{y-\zeta}\Big|_{\zeta=\xi_n(y)+}.$$

These can be written explicitly as

$$\begin{aligned}
\frac{\partial}{\partial\zeta}\frac{c^n(\zeta, y)}{y-\zeta}\Big|_{\zeta=\xi_n(y)-} &= \frac{c_n'(\xi_n(y)) - c_j'(\xi_n(y))}{y-\xi_n(y)} + \frac{c_n(\xi_n(y)) - c_j(\xi_n(y))}{(y-\xi_n(y))^2} \\
&\overset{(2.17)}{=} \frac{c_n'(\xi_n(y)) - c_j'(\xi_n(y))}{y-\xi_n(y)} + \frac{K_n(y) - K_j(y)}{y-\xi_n(y)} \le 0, \\
\frac{\partial}{\partial\zeta}\frac{c^n(\zeta, y)}{y-\zeta}\Big|_{\zeta=\xi_n(y)+} &= \frac{c_n'(\xi_n(y)) - c_k'(\xi_n(y))}{y-\xi_n(y)} + \frac{c_n(\xi_n(y)) - c_k(\xi_n(y))}{(y-\xi_n(y))^2} \\
&\overset{(2.17)}{=} \frac{c_n'(\xi_n(y)) - c_k'(\xi_n(y))}{y-\xi_n(y)} + \frac{K_n(y) - K_k(y)}{y-\xi_n(y)} \ge 0.
\end{aligned}$$

Subtracting the two inequalities, we obtain

$$c_k'(\xi_n(y)) - c_j'(\xi_n(y)) + K_k(y) - K_j(y) \le 0.$$

However, this holds with equality by induction hypothesis since $\xi_n(y) = \xi_k(y)$ and $\jmath_k(y) = j$. Consequently, we see that equality also holds in the two inequalities above. This establishes (A.2).

Finally, we prove the claimed ODE for K_n. By absolute continuity, K_n is differentiable for almost all $y \ge 0$ and we have

$$\begin{aligned}
K_n'(y) &= \lim_{\delta\searrow 0}\frac{1}{\delta}\left[\frac{c^n(\xi_n(y+\delta), y+\delta)}{y+\delta-\xi_n(y+\delta)} - \frac{c^n(\xi_n(y), y)}{y-\xi_n(y)}\right] \\
&= \lim_{\delta\searrow 0}\frac{1}{\delta}\left[\left(\frac{1}{y+\delta-\xi_n(y+\delta)} - \frac{1}{y-\xi_n(y)}\right)c^n(\xi_n(y+\delta), y+\delta)\right. \\
&\qquad \left. + \frac{c^n(\xi_n(y+\delta), y+\delta) - c^n(\xi_n(y), y)}{y-\xi_n(y)}\right] \\
&= \frac{\xi_n'(y+) - 1}{y-\xi_n(y)}K_n(y) \\
&\qquad + \frac{1}{y-\xi_n(y)}\left(\lim_{\delta\searrow 0}\frac{c^n(\xi_n(y+\delta), y+\delta) - c^n(\xi_n(y), y)}{\delta}\right).
\end{aligned}$$

The main technical difficulty comes from the possibility that $\xi_n(y) = \xi_k(y)$ for some $k < n$. We present the arguments for this case and leave the other (much easier) case to the reader.

By assumption the last limit exists, and hence we can compute it using some "convenient" sequence $\delta_m \downarrow 0$ where δ_m is such that $\jmath_n(y+\delta_m) = l$ for all $m \in \mathbb{N}$. Note that by continuity of $\xi_1, \dots, \xi_n$ at y we have that either $l = \jmath_n(y)$ or l is such that $\xi_l(y) = \xi_n(y)$ and then $\jmath_n(y) = \jmath_l(y)$. In the latter case, by (2.17) and continuity of c^n (see Lemma 2.15), we obtain

$$\frac{c_n(\xi_n(y)) - c_{\jmath_n(y)}(\xi_n(y))}{y - \xi_n(y)} + K_{\jmath_n(y)}(y) = \frac{c_n(\xi_n(y)) - c_l(\xi_n(y))}{y - \xi_n(y)} + K_l(y). \tag{A.7}$$

Recall (2.34). For δ_m small enough such that $\imath_n(\xi_n(y), y+\delta_m) = \jmath_n(y)$, we obtain

$$\begin{aligned}
&c^n\big(\xi_n(y+\delta_m), y+\delta_m\big) - c^n\big(\xi_n(y), y+\delta_m\big) \\
&\quad = c_n\big(\xi_n(y+\delta_m)\big) - c_l\big(\xi_n(y+\delta_m)\big) + \big(y+\delta_m - \xi_n(y+\delta_m)\big)K_l(y+\delta_m) \\
&\qquad - c_n\big(\xi_n(y)\big) + c_{\jmath_n(y)}\big(\xi_n(y)\big) - \big(y+\delta_m - \xi_n(y)\big)K_{\jmath_n(y)}(y+\delta_m) \\
&\quad \overset{(A.7)}{=} c_n\big(\xi_n(y+\delta_m)\big) - c_l\big(\xi_n(y+\delta_m)\big) + \big(y+\delta_m - \xi_n(y+\delta_m)\big)K_l(y+\delta_m) \\
&\qquad - c_n\big(\xi_n(y)\big) + c_l\big(\xi_n(y)\big) - \big(y - \xi_n(y)\big)\big(K_l(y) - K_{\jmath_n(y)}(y)\big) \\
&\qquad - \big(y+\delta_m - \xi_n(y)\big)K_{\jmath_n(y)}(y+\delta_m).
\end{aligned}$$

From this, for almost all $y \ge 0$, we obtain using the induction hypothesis

$$\begin{aligned}
&\lim_{m\to\infty} \frac{c^n(\xi_n(y+\delta_m), y+\delta_m) - c^n(\xi_n(y), y+\delta_m)}{\delta_m} \\
&\quad = \xi_n'(y+)\big[c_n'\big(\xi_n(y)+\big) - c_l'\big(\xi_n(y)+\big) - K_l(y)\big] \\
&\qquad + K_l(y) + \big(y-\xi_n(y)\big)K_l'(y) - K_{\jmath_n(y)}(y) - \big(y-\xi_n(y)\big)K_{\jmath_n(y)}'(y) \qquad \text{(A.8)} \\
&\quad \underset{(A.3)}{\overset{(A.2)}{=}} -\xi_n'(y+)K_n(y).
\end{aligned}$$

The use of (A.2) above is a priori only justified for $y \ge 0$ such that $\xi_n(y)$ is not an atom of $\mu_1, \dots, \mu_n$. However, for a fixed atom ζ, the set $\{y : \xi_n(y) = \zeta\}$ is a closed interval on whose interior $\xi_n' \equiv 0$ and it follows that (A.8) extends to almost all $y \ge 0$. Together with (A.1) this yields, for almost all $y \ge 0$,

$$\begin{aligned}
K_n'(y) &= \frac{\xi_n'(y) - 1}{y - \xi_n(y)} K_n(y) + \frac{1}{y - \xi_n(y)}\left(-K_n(y)\xi_n'(y) + \frac{\partial c^n}{\partial y}\big(\xi_n(y), y\big)\right) \\
&= -\frac{K_n(y)}{y - \xi_n(y)} + \frac{1}{y - \xi_n(y)}\big(K_{\jmath_n(y)}(y) + \big(y - \xi_n(y)\big)K_{\jmath_n(y)}'(y)\big),
\end{aligned} \tag{A.9}$$

which completes the proof. □

REFERENCES

[1] AZÉMA, J. and YOR, M. (1979). Une solution simple au problème de Skorokhod. In *Séminaire de Probabilités*, *XIII* (*Univ. Strasbourg*, *Strasbourg*, 1977/78) (C. Dellacherie, P. A. Meyer and M. Weil, eds.). *Lecture Notes in Math.* **721** 90–115. Springer, Berlin. MR0544782

[2] AZÉMA, J., GUNDY, R. F. and YOR, M. (1980). Sur l'intégrabilité uniforme des martingales continues. In *Seminar on Probability*, *XIV* (*Paris*, 1978/1979) (*French*). *Lecture Notes in Math*. **784** 53–61. Springer, Berlin. MR0580108

[3] BEIGLBÖCK, M., COX, A. M. G. and HUESMANN, M. (2015). Optimal transport and Skorokhod embedding. Available at arXiv:1307.3656v3.

[4] BROWN, H., HOBSON, D. and ROGERS, L. C. G. (2001). The maximum maximum of a martingale constrained by an intermediate law. *Probab. Theory Related Fields* **119** 558–578. MR1826407

[5] BROWN, H., HOBSON, D. and ROGERS, L. C. G. (2001). Robust hedging of barrier options. *Math*. *Finance* **11** 285–314. MR1839367

[6] CARRARO, L., EL KAROUI, N. and OBŁÓJ, J. (2012). On Azéma–Yor processes, their optimal properties and the Bachelier-drawdown equation. *Ann*. *Probab*. **40** 372–400. MR2917776

[7] COX, A. M. G. and OBŁÓJ, J. (2011). Robust hedging of double touch barrier options. *SIAM J. Financial Math*. **2** 141–182. MR2772387

[8] COX, A. M. G. and OBŁÓJ, J. (2011). Robust pricing and hedging of double no-touch options. *Finance Stoch*. **15** 573–605. MR2833100

[9] COX, A. M. G., OBŁÓJ, J. and TOUZI, N. (2015). The Root solution to the multi-marginal embedding problem: An optimal stopping and time-reveral approach. Available at arXiv:1505.03169.

[10] COX, A. M. G. and WANG, J. (2013). Root's barrier: Construction, optimality and applications to variance options. *Ann*. *Appl*. *Probab*. **23** 859–894. MR3076672

[11] GALICHON, A., HENRY-LABORDÈRE, P. and TOUZI, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann*. *Appl*. *Probab*. **24** 312–336. MR3161649

[12] GUO, G., TAN, X. and TOUZI, N. (2017). On the monotonicity principle of optimal Skorokhod embedding problem. *SIAM J. Control Optim*. **54** 2478–2489.

[13] HENRY-LABORDÈRE, P., OBŁÓJ, J., SPOIDA, P. and TOUZI, N. (2016). The maximum maximum of a martingale with given *n* marginals. *Ann*. *Appl*. *Probab*. **26** 1–44. MR3449312

[14] HOBSON, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris–Princeton Lectures on Mathematical Finance* 2010 (R. A. Carmona, I. Çinlar, E. Ekeland, E. Jouini, J. A. Scheinkman and N. Touzi, eds.). *Lecture Notes in Math*. **2003** 267–318. Springer, Berlin. MR2762363

[15] HOBSON, D. and NEUBERGER, A. (2012). Robust bounds for forward start options. *Math*. *Finance* **22** 31–56. MR2881879

[16] HOBSON, D. G. (1998). Robust hedging of the lookback option. *Finance Stoch*. **2** 329–347.

[17] LOYNES, R. M. (1970). Stopping times on Brownian motion: Some properties of Root's construction. *Z. Wahrsch*. *Verw*. *Gebiete* **16** 211–218. MR0292170

[18] MADAN, D. B. and YOR, M. (2002). Making Markov martingales meet marginals: With explicit constructions. *Bernoulli* **8** 509–536. MR1914701

[19] MONROE, I. (1972). On embedding right continuous martingales in Brownian motion. *Ann*. *Math*. *Stat*. **43** 1293–1311. MR0343354

[20] OBŁÓJ, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surv*. **1** 321–390. MR2068476

[21] OBŁÓJ, J. (2010). Skorokhod Embedding. In *Encyclopedia of Quantitative Finance* (R. Cont, ed.) 1653–1657. Wiley, Chichester.

[22] OBŁÓJ, J., SPOIDA, P. and TOUZI, N. (2015). Martingale inequalities for the maximum via pathwise arguments. In *In Memoriam Marc Yor—Séminaire n Probabilités XLVII* (C. Donati-Martin, A. Lejay and A. Rouault, eds.). *Lecture Notes in Math*. **2137** 227–247. Springer, Cham. MR3444301

[23] ROGERS, L. C. G. (1989). A guided tour through excursions. *Bull*. *Lond*. *Math*. *Soc*. **21** 305–341. MR0998631

[24] SKOROKHOD, A. V. (1965). *Studies in the Theory of Random Processes*. Addison-Wesley, Reading, MA. MR0185620

[25] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat*. **36** 423–439. MR0177430

MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
AWB, ROQ, WOODSTOCK ROAD
OXFORD OX2 6GG
UNITED KINGDOM
E-MAIL: Jan.Obloj@maths.ox.ac.uk
Peter.Spoida@maths.ox.ac.uk