The geometry of multi-marginal Skorokhod Embedding

Mathias Beiglböck[1] · Alexander M. G. Cox[2] · Martin Huesmann[1]

Abstract

The Skorokhod Embedding Problem is one of the classical problems in the theory of stochastic processes, with applications in many different fields [cf. the surveys (Hobson in: Paris-Princeton lectures on mathematical finance 2010, Volume 2003 of Lecture Notes in Mathematics, Springer, Berlin, 2011; Obłój in: Probab Surv 1:321–390, 2004)]. Many of these applications have natural multi-marginal extensions leading to the *(optimal) multi-marginal Skorokhod problem*. Some of the first papers to consider this problem are Brown et al. (Probab Theory Relat Fields 119(4):558–578, 2001), Hobson (Séminaire de Probabilités, XXXII, Volume 1686 of Lecture Notes in Mathematics, Springer, Berlin, 1998), Madan and Yor (Bernoulli 8(4):509–536, 2002). However, this turns out to be difficult using existing techniques: only recently a complete solution was be obtained in Cox et al. (Probab Theory Relat Fields 173:211–259, 2018) establishing an extension of the Root construction, while other instances are only partially answered or remain wide open. In this paper, we extend the theory developed in Beiglböck et al. (Invent Math 208(2):327–400, 2017) to the multi-marginal setup which is comparable to the extension of the optimal transport problem to the multi-marginal optimal transport problem. As for the one-marginal case, this viewpoint turns out to be very powerful. In particular, we are able to show that all classical optimal embeddings have natural multi-marginal counterparts. Notably these different constructions are linked through a joint geometric structure and the classical solutions are recovered as particular cases. Moreover, our results also have consequences for the study of the martingale transport problem as well as the peacock problem.

Keywords Optimal transport · Skorokhod embedding · Multiple marginals · Martingale optimal transport · Peacocks

✉ Alexander M. G. Cox
  a.m.g.cox@bath.ac.uk

Extended author information available on the last page of the article

Mathematics Subject Classification Primary 60G42 · 60G44; Secondary 91G20

1 Introduction

The Skorokhod Embedding problem (SEP) is a classical problem in probability, dating back to the 1960s [58,59]. Simply stated, the aim is to represent a given probability as the distribution of Brownian motion at a chosen stopping time. Recently, motivated by applications in probability, mathematical finance, and numerical methods, there has been renewed, sustained interest in solutions to the SEP (cf. the two surveys [36,50]) and its multi-marginal extension, the multi-marginal SEP: Given marginal measures $\mu_0, \ldots, \mu_n$ of finite variance and a Brownian motion with $B_0 \sim \mu_0$, construct stopping times $\tau_1 \leq \ldots \leq \tau_n$ such that

$$B_{\tau_i} \sim \mu_i \text{ for all } 1 \leq i \leq n \text{ and } \mathbb{E}[\tau_n] < \infty. \tag{MSEP}$$

It is well known that a solution to (MSEP) exists iff the marginals are in convex order ($\mu_0 \preceq_c \ldots \preceq_c \mu_n$) and have finite second moment; under this condition Skorokhod's original results give the existence of solutions of the induced one period problems, which can then be pasted together to obtain a solution to (MSEP).

It appears to be significantly harder to develop genuine extensions of one period solutions: many of the classical solutions to the SEP exhibit additional desirable characteristics and optimality properties which one would like to extend to the multi-marginal case. However the original derivations of these solutions make significant use of the particular structure inherent to certain problems, often relying on explicit calculations, which make extensions difficult if not impossible. The first paper which we are aware of to attempt to extend a classical construction to the multi-marginal setting is [12], which generalised the Azéma–Yor embedding [1] to the case with two marginals. This work was further extended by Henry-Labordère et al. [30,52], who were able to extend to arbitrary (finite) marginals, under particular assumptions on the measures. Using an extension of the stochastic control approach in [25] Claisse et al. [15] constructed a two marginal extension of the Vallois embedding. Recently, Cox et al. [18] were able to characterise the solution to the general multi-marginal Root embedding through the use of an optimal stopping formulation.

Mass transport approach and general multi-marginal embedding In this paper, we develop a new approach to the multi-marginal Skorokhod problem, based on insights from the field of optimal transport.

Following the seminal paper of Gangbo and McCann [26] the mutual interplay of optimality and geometry of optimal transport plans has been a cornerstone of the field. As shown for example in [14,45,53] this in not limited to the two-marginal case but extends to the multi-marginal case where it turns out to be much harder though. Recently, similar ideas have been shown to carry over to a more probablistic context, to optimal transport problems satisfying additional linear constraints [7,28,60] and in fact to the classical Skorokhod embedding problem [2].

Building on these insights, we extend the mass transport viewpoint developed in [2] to the multi-marginal Skorokhod embedding problem. This allows us to give multi-

marginal extensions of all the classical optimal solutions to the Skorokhod problem in full generality, which we exemplify by several examples. In particular the classical solutions of Azéma–Yor, Root, Rost, Jacka, Perkins, and Vallois can be recovered as special cases. In addition, the approach allows us to derive a number of new solutions to (MSEP) which have further applications to e.g. martingale optimal transport and the peacock problem. A main contribution of this paper is that in many different cases, solutions to the multi-marginal SEP share a common *geometric structure*. In all the cases we consider, this geometric information will in fact be enough to characterise the optimiser uniquely, which highlights the flexibility of our approach.

Furthermore, our approach to the Skorokhod embedding problem is very general and does not rely on fine properties of Brownian motion. Therefore, exactly as in [2] the results of this article carry over to sufficiently regular Markov processes, e.g. geometric Brownian motion, three-dimensional Bessel process and Ornstein–Uhlenbeck processes, and Brownian motion in $\mathbb{R}^d$ for $d > 1$. As the arguments are precisely the same as in [2], we refer to [2, Section 8] for details.

Related work Interest in the multi-marginal Skorokhod problem comes from a number of directions and we describe some of these here:

- *Maximising the running maximum: the Azéma–Yor embedding*
  Suppose $(M_t)_{t \geq 0}$ is a martingale and write $\bar{M}_t := \sup_{s \leq t} M_s$. The relationship between the laws of M_1 and $\bar{M}_1$ has been studied by Blackwell and Dubins [10], Dubins and Gilat [23] and Kertz and Rösler [44], culminating in a complete classification of all possible joint laws by Rogers [55]. In particular given the law of M_1, the set of possible laws of $\bar{M}_1$ admits a maximum w.r.t. the stochastic ordering, this can be seen through the Azéma–Yor embedding. Given initial and terminal laws of the martingale, Hobson [35] gave a sharp upper bound on the law of the maximum based on an extension of the Azéma–Yor embedding to Brownian motion started according to a non-trivial initial law. These results are further extended in [12] to the case of martingales started in 0 and constrained to a specified marginal at an intermediate time point, essentially based on a further extension of the Azéma–Yor construction. The natural aim is to solve this question in the case of arbitrarily many marginals. Assuming that the marginals have ordered barycenter functions this case is included in the work of Madan and Yor [47], based on iterating the Azéma–Yor scheme. More recently, the stochastic control approach of [25] (for one marginal) is extended by Henry-Labordère et al. [30,52] to marginals in convex order satisfying an additional assumption ([52, Assumption ⊛][1]). Together with the Dambis–Dubins–Schwarz Theorem, Theorem 2.11 below provides a solution to this problem in full generality.

- *Multi-marginal Root embedding*
  In a now classical paper, Root [56] showed that for any centred distribution with finite second moment, μ, there exists a (right) *barrier* $\mathcal{R}$, i.e. a Borel subset of $\mathbb{R}_+ \times \mathbb{R}$ such that $(t, x) \in \mathcal{R}$ implies $(s, x) \in \mathcal{R}$ for all $s \geq t$, and for which $B_{\tau_{\mathcal{R}}} \sim \mu$, $\tau_{\mathcal{R}} = \inf\{t : (t, B_t) \in \mathcal{R}\}$. This work was further generalised to a large class of Markov processes by Rost [57], who also showed that this construction was optimal in that it minimised $\mathbb{E}[h(\tau)]$ for convex functions h.

---

[1] As shown by an example in [52] this condition is necessary to carry out their explicit construction.

More recent work on the Root embedding has focused on attempts to characterise the stopping region. A number of papers do this either through analytical means [16,17,27,48] or through connections with optimal stopping problems [19]. Recently the connection to optimal stopping problems has enabled Cox et al. [18] to extend these results to the multi-marginal setting. Moreover, they prove that this solution enjoys a similar optimality property to the one-marginal Root solution. The principal strategy is to first prove the result in the case of locally finitely supported measures by means of a time reversal argument. The proof is then completed in the case of general measures by a delicate limiting procedure.

As a consequence of the theoretical results in this paper, we will be able to prove similar results. In particular, the barrier structure as well as the optimality properties are recovered in Theorem 2.7. Indeed, as we will show below, the particular geometric structure of the Root embedding turns out to be archetypal for a number of multi-marginal counterparts of classical embeddings.

- *Model-independent finance*
  An important application field for the results in this paper, and one of the motivating factors behind the recent resurgence of interest in the SEP, relates to model-independent finance. In mathematical finance, one models the price process S as a martingale under a risk-neutral measure, and specifying prices of all call options at maturity T is equivalent to fixing the distribution μ of S_T. Understanding no-arbitrage price bounds for a functional γ, can often be seen to be equivalent to finding the range of $\mathbb{E}[\gamma(B)_\tau]$ among all solutions to the Skorokhod embedding problem for μ. This link between SEP and model-independent pricing and hedging was pioneered by Hobson [35] and has been an important question ever since. A comprehensive overview is given in [36].

  However, the above approach uses only market data for the maturity time T, while in practice market data for many intermediate maturities may also be available, and this corresponds to the multi-marginal SEP. While we do not pursue this direction of research in this article we emphasize that our approach yields a systematic method to address this problem. In particular, the general framework of super-replication results for model-independent finance now includes a number of important contributions, see [3,22,29,40], and most of these papers allow for information at multiple intermediate times.

- *Martingale optimal transport*
  Optimal transport problems where the transport plan must satisfy additional martingale constraints have recently been investigated, e.g. the works of Dolinsky, Ekren, Gallichon, Ghoussoub, Henry-Labordere, Hobson, Juillet, Kim, Lim, Nutz, Obłoj, Soner, Tan, Touzi in [4,7–9,13,22,24,25,38]. Besides having a natural interpretation in finance, such martingale transport problems are also of independent mathematical interest, for example – similarly to classical optimal transport – they have consequences for the investigation of martingale inequalities (see e.g. [11,30,52]). As observed in [5] one can gain insight into the martingale transport problem between two probabilities μ_1 and μ_2 by relating it to a Skorokhod embedding problem which may be considered as a continuous time version of the martingale transport problem. Notably this idea can be used to recover the known solutions of the martingale optimal transport problem in a unified fashion ([41]).

It thus seems natural that an improved understanding of an n-marginal martingale transport problem can be obtained based on the multi-marginal Skorokhod embedding problem. Indeed this is exemplified in Theorem 2.17 below, where we use a multi-marginal embedding to establish an n-period version of the martingale monotone transport plan, and recover similar results to recent work of Nutz et al. [49].

- *Construction of peacocks*

  Dating back to the work of Madan–Yor [47], and studied in depth in the book of Hirsch et al. [33], given a family of probability measures $(\mu_t)_{t\in[0,T]}$ which are increasing in convex order, a *peacock* (from the acronym PCOC "Processus Croissant pour l'Ordre Convexe") is a martingale such that $M_t \sim \mu_t$ for all $t \in [0,T]$. The existence of such a process is granted by Kellerer's celebrated theorem, and typically there is an abundance of such processes. Loosely speaking, the peacock problem is to give constructions of such martingales. Often such constructions are based on Skorokhod embedding or particular martingale transport plans, and often one is further interested in producing solutions with some additional optimality properties; see for example the recent works [31,37,42,43].

  Given the intricacies of multi-period martingale optimal transport and Skorokhod embedding, it is necessary to make additional assumptions on the underlying marginals and desired optimality properties are in general not preserved in a straight forward way during the inherent limiting/pasting procedure. We expect that an improved understanding of the multi-marginal Skorokhod embedding problem will provide a first step to tackle these range of problems in a systematic fashion.

1.1 Outline of the paper

We will proceed as follows. In Sect. 2.1, we will describe our main results. Our main technical tool is a 'monotonicity principle', Theorem 2.5. This result allows us to deduce the geometric structure of optimisers. Having stated this result, and defined the notion of 'stop-go pairs', which are important mathematical embodiment of the notion of 'swapping' stopping rules for a candidate optimiser, we will be able to deduce our main consequential results. Specifically, we will prove the multi-marginal generalisations of the Root, Rost and Azéma–Yor embeddings, using their optimality properties as a key tool in their construction. The Rost construction is entirely novel, and the solution to the Azéma–Yor embedding generalises existing results, which have only previously been given under a stronger assumption on the measures. We also give a multi-marginal generalisation of an embedding due to Hobson & Pedersen; this is, in some sense, the counterpart of the Azéma–Yor embedding; classically, this is better recognised as the embedding of Perkins [54], however for reasons we give later, this embedding has no multi-marginal extension. Moreover the proofs of these results will share a common structure, and it will be clear how to generalise these methods to provide similar results for a number of other classical solutions to the SEP.

In Sect. 2.1, we also use our methods to give a multi-marginal martingale monotone transport plan, using a construction based on a SEP-viewpoint.

The remainder of the paper is then dedicated to proving the main technical result, Theorem 2.5. In Sect. 3, we introduce our technical setup, and prove some preliminary results. As in [2], it will be important to consider the class of *randomised* multi-stopping times, and we define these in this section, and derive a number of useful properties. It is technically convenient to consider randomised multi-stopping times on a canonical probability space, where there is sufficient additional randomisation, independent of the Brownian motion, however we will prove in Lemma 3.11 that any sufficiently rich probability space will suffice. A key property of the set of randomised multi-stopping times embedding a given sequence of measures is that this set is compact in an appropriate (weak) topology, and this will be proved in Proposition 3.19; an important consequence of this is that optimisers of the multi-marginal SEP exist under relatively mild assumptions on the objective (Theorem 2.1).

In Sect. 4 we introduce the notions of color-swap pairs, and multi-colour swap pairs. These will be the fundamental constituents of the set of 'bad-pairs', or combinations of stopped and running paths that we do not expect to see in optimal solutions. In this section we define these pairs, and prove some technical properties of the sets.

In Sect. 5 we complete the proof of Theorem 2.5. In spirit this follows the proof of the corresponding result in [2], and we only provide the details here where the proof needs to adapt to account for the multi-marginal setting.

1.2 Frequently used notation

- The set of Borel (sub-)probability measures on a topological space X is denoted by $\mathcal{P}(X) / \mathcal{P}^{\leq 1}(X)$.
- $\Xi^d = \{(s_1, \ldots, s_d) : 0 \leq s_1 \leq \ldots \leq s_d\}$ denotes the ordered sequences in $[0, \infty)$ of length d.
- The d-dimensional Lebesgue measure will be denoted by $\mathcal{L}^d$.
- For a measure ξ on X we write $f(\xi)$ for the push-forward of ξ under $f : X \to Y$.
- We use $\xi(f)$ as well as $\int f \, d\xi$ to denote the integral of a function f against a measure ξ.
- $C_x(\mathbb{R}_+)$ denotes the continuous functions starting at x; $C(\mathbb{R}_+) = \bigcup_{x \in \mathbb{R}} C_x(\mathbb{R}_+)$. For $\omega \in C(\mathbb{R}_+)$ we write $\theta_s \omega$ for the path in $C_0(\mathbb{R}_+)$ defined by $(\theta_s \omega)_{t \geq 0} = (\omega_{t+s} - \omega_s)_{t \geq 0}$.
- $\mathbb{W}$ denotes Wiener measure; $\mathbb{W}_\mu$ denotes law of Brownian motion started according to a probability μ; $\mathcal{F}^0$ $(\mathcal{F}^a)$ the natural (augmented) filtration on $C_0(\mathbb{R}_+)$.
- For $d \in \mathbb{N}$ we set $\overline{C(\mathbb{R}_+)} = C(\mathbb{R}_+) \times [0, 1]^d$, $\bar{\mathbb{W}} = \mathbb{W} \otimes \mathcal{L}^d$, and $\bar{\mathcal{F}} = (\bar{\mathcal{F}}_t)_{t \geq 0}$ the usual augmentation of $(\mathcal{F}_t^0 \otimes \mathcal{B}([0, 1]^d))_{t \geq 0}$. To keep notation manageable, we suppress d from the notation since the precise number will always be clear from the context.
- X is a Polish space equipped with a Borel probability measure m. We set $\mathcal{X} := X \times C_0(\mathbb{R}_+)$, $\mathbb{P} = m \otimes \mathbb{W}$, $\mathcal{G}^0 = (\mathcal{G}_t^0)_{t \geq 0} = (\mathcal{B}(X) \otimes \mathcal{F}_t^0)_{t \geq 0}$, $\mathcal{G}^a$ the usual augmentation of $\mathcal{G}^0$.
- For $d \in \mathbb{N}$ we set $\bar{\mathcal{X}} = \mathcal{X} \times [0, 1]^d$, $\bar{\mathbb{P}} = \mathbb{P} \otimes \mathcal{L}^d$, and $\bar{\mathcal{G}} = (\bar{\mathcal{G}}_t)_{t \geq 0}$ the usual augmentation of $(\mathcal{G}_t^0 \otimes \mathcal{B}([0, 1]^d))_{t \geq 0}$. Again, we suppress d from the notation since the precise number will always be clear from the context.

- The set of stopped paths started at 0 is denoted by $S = \{(f, s) : f : [0, s] \to \mathbb{R}$ is continuous, $f(0) = 0\}$ and we define $r : C_0(\mathbb{R}_+) \times \mathbb{R}_+ \to S$ by $r(\omega, t) := (\omega_{\restriction[0,t]}, t)$. The set of stopped paths started in X is $S_\mathsf{X} = (\mathsf{X}, S) = \{(x, f, s) : f : [0, s] \to \mathbb{R}$ is continuous, $f(0) = 0$, $x \in \mathsf{X}\}$ and we define $r_\mathsf{X} : \mathsf{X} \times C_0(\mathbb{R}_+) \times \mathbb{R}_+ \to S_\mathsf{X}$ by $r_\mathsf{X}(x, \omega, t) := (x, \omega_{\restriction[0,t]}, t)$, i.e. $r_\mathsf{X} = (\mathrm{Id}, r)$.

- We use $\oplus$ for the concatenation of paths: depending on the context the arguments may be elements of S, $C_0(\mathbb{R}_+)$ or $C_0(\mathbb{R}_+) \times \mathbb{R}_+$. Specifically, $\oplus : \mathsf{Y} \times \mathsf{Z} \to \mathsf{Z}$, where Y is either S or $C_0(\mathbb{R}_+) \times \mathbb{R}_+$, and Z may be any of the three spaces. For example, if $(f, s) \in S$ and $\omega \in C_0(\mathbb{R}_+)$, then $(f, s) \oplus \omega$ is the path

$$\omega'(t) = \begin{cases} f(t) & t \leq s \\ f(s) + \omega(t - s) & t > s \end{cases}. \tag{1.1}$$

- As well as the simple concatenation of paths, we introduce a concatenation operator which keeps track of the concatenation time: if $(f, s), (g, t) \in S$, then $(f, s) \otimes (g, t) = (f \oplus g, s, s + t)$. We denote the set of elements of this form as $S^{\otimes 2}$, and inductively, $S^{\otimes i}$ in the same manner.

- Elements of $S^{\otimes i}$ will usually be denoted by $(f, s_1, \ldots, s_i)$ or $(g, t_1, \ldots, t_i)$. We define $r_i : C_0(\mathbb{R}_+) \times \Xi^i \to S^{\otimes i}$ by $r_i(\omega, s_1, \ldots, s_i)) := (\omega_{\restriction[0,s_i]}, s_1, \ldots, s_i)$. Accordingly, the set of i-times stopped paths started in X is $S_\mathsf{X}^{\otimes i} = (\mathsf{X}, S^{\otimes i})$. Elements of $S_\mathsf{X}^{\otimes i}$ are usually denoted by $(x, f, s_1, \ldots, s_i)$ or $(y, g, t_1, \ldots, t_i)$. In case of $\mathsf{X} = \mathbb{R}$ we often simply write $(f, s_1, \ldots, s_i)$ or $(g, t_1, \ldots, t_i)$ with the understanding that $f(0), g(0) \in \mathbb{R}$. In case that there is no danger of confusion we will also sometimes write $S_\mathbb{R}^{\otimes i} = S^{\otimes i}$. The operators $\oplus, \otimes$ generalise in the obvious way to allow elements of $S_\mathsf{X}^{\otimes i}$ to the left of the operator.

- For $(x, f, s_1, \ldots, s_i) \in S_\mathsf{X}^{\otimes i}$, $(h, s) \in S$ we often denote their concatenation by $(x, f, s_1, \ldots, s_i) | (h, s)$ which is the same element as $(x, f, s_1, \ldots, s_i) \otimes (h, s)$ but comes with the probabilistic interpretation of conditioning on the continuation of $(f, s_1, \ldots, s_i)$ by (h, s). In practice, this means that we will typically expect the (h, s) to be absorbed by a later $\oplus$ operation.

- The map $\mathsf{X} \times \Xi^i \ni (x, \omega, s_1, \ldots, s_i) \mapsto (x, \omega_{\llcorner[0,s_i]}, s_1, \ldots, s_i) \in S_\mathsf{X}^{\otimes i}$ will (by slight abuse of notation) also be denoted by r_i.

- We set $\tilde{r}_i : \mathsf{X} \times \Xi^i \to S_\mathsf{X}^{\otimes i} \times C(\mathbb{R}_+)$, $(x, \omega, s_1, \ldots, s_i) \mapsto ((x, \omega_{\llcorner[0,s_i]}, s_1, \ldots, s_i), \theta_{s_i}\omega)$. Then $\tilde{r}_i$ is clearly a homeomorphism with inverse map

$$\tilde{r}_i^{-1} : ((x, f, s_1, \ldots, s_i), \omega) \mapsto (x, f \oplus \omega, s_1, \ldots, s_i).$$

Hence, $\xi = \tilde{r}_i^{-1}(\tilde{r}_i(\xi))$ for any measure ξ on $\mathsf{X} \times \Xi^i$. For $1 \leq i < d$ we can extend $\tilde{r}_i$ to a map $\tilde{r}_{d,i} : \mathsf{X} \times \Xi^d \to S_\mathsf{X}^{\otimes i} \times C(\mathbb{R}_+) \times \Xi^{d-i}$ by setting

$$\tilde{r}_{d,i}(x, \omega, s_1, \ldots, s_d) = ((x, \omega_{\llcorner[0,s_i]}, s_1, \ldots, s_i), \theta_{s_i}\omega, (s_{i+1} - s_i, \ldots, s_d - s_i)).$$

- For $\Gamma_i \subseteq S^{\otimes i}$ we set $\Gamma_i^< := \{(f, s_1, \ldots, s_{i-1}, s_i) : \exists (\tilde{f}, s_1, \ldots, s_{i-1}, \tilde{s}) \in \Gamma, s_{i-1} \leq s_i < \tilde{s} \text{ and } f \equiv \tilde{f} \text{ on } [0, s_i]\}$, where we set $s_0 = 0$.

- For $(f, s_1, \ldots, s_i) \in S^{\otimes i}$ we write $\overline{f} = \sup_{r \leq s_i} f(r)$, and $\underline{f} = \inf_{r \leq s_i} f(r)$.
- For $1 \leq i < n$ and F a function on $S^{\otimes n}$ resp. $C_0(\mathbb{R}_+) \times \Xi^n$ and $(f, s_1, \ldots, s_i) \in S^{\otimes i}$ we set

$$
\begin{aligned}
F^{(f, s_1, \ldots, s_i) \otimes}(\eta, t_{i+1}, \ldots, t_n) &:= F(f \oplus \eta, s_1, \ldots, s_i, s_i + t_{i+1}, \ldots, s_i + t_n) \\
&= F((f, s_1, \ldots, s_i) \otimes (\eta, t_{i+1}, \ldots, t_n)),
\end{aligned}
$$

where $(\eta, t_{i+1}, \ldots, t_n)$ may be an element of $S^{\otimes n-i}$, or $C_0(\mathbb{R}_+) \times \Xi^{n-i}$. We similarly define

$$
\begin{aligned}
F^{(f, s_1, \ldots, s_i) \oplus}(\eta, t_{i+1}, \ldots, t_n) &:= F(f \oplus \eta, s_1, \ldots, s_{i-1}, s_i + t_{i+1}, \ldots, s_i + t_n) \\
&= F((f, s_1, \ldots, s_i) \oplus (\eta, t_{i+1}, \ldots, t_n)),
\end{aligned}
$$

where $(\eta, t_i, \ldots, t_n)$ may be an element of $S^{\otimes n-i+1}$, or $C_0(\mathbb{R}_+) \times \Xi^{n-i+1}$.
- For any j-tuple $1 \leq i_1 < \ldots < i_j \leq d$ we denote by $\text{proj}_{\mathcal{X} \times (i_1, \ldots, i_j)}$ the projection from $\mathcal{X} \times \mathbb{R}^d$ to $\mathcal{X} \times \mathbb{R}^j$ defined by

$$
(x, \omega, y_1, \ldots, y_d) \mapsto (x, \omega, y_{i_1}, \ldots, y_{i_j})
$$

and correspondingly, for $\xi \in \mathcal{P}(\mathcal{X} \times \mathbb{R}^d)$, $\xi^{(i_1, \ldots, i_j)} = \text{proj}_{\mathcal{X} \times (i_1, \ldots, i_j)}(\xi)$. When $j = 0$, we understand this as simply the projection onto $\mathcal{X}$. If $(i_1, \ldots, i_j) = (1, \ldots, j)$ we simply write $\xi^{(1, \ldots, j)} = \xi^i$.

2 Main results

2.1 Existence and monotonicity principle

In this section we present our key results and provide an interpretation in probabilistic terms. To move closer to classical probabilistic notions, in this section, we slightly deviate from the notation used in the rest of the article. We consider a Brownian motion B on some generic probability space and recall that, for each $1 \leq i \leq n$,

$$
S^{\otimes i} := \{(f, s_1, \ldots, s_i) : 0 \leq s_1 \leq \ldots \leq s_i, f \in C([0, s_i])\}.
$$

We note that $S^{\otimes i}$ carries a natural Polish topology. For a function $\gamma : S^{\otimes n} \to \mathbb{R}$ which is Borel and a sequence $(\mu_i)_{i=0}^n$ of centered probability measures on $\mathbb{R}$, increasing in convex order, we are interested in the optimization problem

$$
P_\gamma = \inf\{\mathbb{E}[\gamma((B_s)_{s \leq \tau_n}, \tau_1, \ldots, \tau_n)] : \tau_1, \ldots, \tau_n \text{ satisfy (MSEP)}\}. \quad \text{(OptMSEP)}
$$

We denote the set of all minimizers of (OptMSEP) by Opt_γ. Take another Borel measurable function $\gamma_2 : S^{\otimes n} \to \mathbb{R}$. We will be also interested in the secondary optimization problem

$$P_{\gamma_2|\gamma} = \inf\{\mathbb{E}[\gamma_2((B_s)_{s\leq\tau_n}, \tau_1, \ldots, \tau_n)] : (\tau_1, \ldots, \tau_n) \in \mathsf{Opt}_\gamma\}. \quad \text{(OptMSEP}_2)$$

Both optimization problems, (OptMSEP) and (OptMSEP$_2$),will not depend on the particular choice of the underlying probability space, provided that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq0}, \mathbb{P})$ is sufficiently rich that it supports a Brownian motion $(B_t)_{t\geq0}$ starting with law μ_0, and an independent, uniformly distributed random variable Y, which is $\mathcal{F}_0$-measurable (see Lemma 3.11). We will from now on assume that we are working in this setting. On this space, we denote the filtration generated by the Brownian motion by $\mathcal{F}^B$.

Many of the assumptions imposed on the problem can be weakened. First, the assumption that $\mathbb{E}[\tau_n] < \infty$ can be weakened, and the class of measures considered can then be extended to the class of probability measures with a finite first moment. More generally, the class of processes can be extended to include e.g. diffusions. Since all the arguments are identical to those in the single marginal setting, we do not work in this generality in this paper, but rather restrict our consideration to the case outlined above. For further details of how to extend the arguments, we refer the reader to [2, Section 7].

We will usually assume that (OptMSEP) and (OptMSEP$_2$) are *well-posed* in the sense that $\mathbb{E}\left[\gamma\left((B_s)_{s\leq\tau_n}, \tau_1, \ldots, \tau_n\right)\right]$ and $\mathbb{E}\left[\gamma_2\left((B_s)_{s\leq\tau_n}, \tau_1, \ldots, \tau_n\right)\right]$ exist with values in $(-\infty, \infty]$ for all $\tau = (\tau_1, \ldots, \tau_n)$ which solve (MSEP) and is finite for one such τ.

Theorem 2.1 *Let* $\gamma, \gamma_2 : S^{\otimes n} \to \mathbb{R}$ *be lsc and bounded from below in the sense that for some constants* $a, b, c \in \mathbb{R}_+$

$$-\left(a + bs_n + c \max_{r\leq s_n} f(r)^2\right) \leq \gamma_i(f, s_1, \ldots, s_n) \quad (2.1)$$

holds on $S^{\otimes n}$, *for* $i = 1, 2$. *Then there exists a minimizer to* (OptMSEP$_2$).

We will prove this result in Sect. 3.3.

Our main result is the monotonicity principle, Theorem 2.5, which is a *geometric characterisation* of optimizers $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_n)$ of (OptMSEP$_2$). The version we state here is weaker than the result we will prove in Sect. 5 but easier to formulate and still sufficient for our intended applications.

For two families of increasing stopping times $(\sigma_j)_{j=i}^n$ and $(\tau_j)_{j=i}^n$ with $\tau_i = 0$ we define

$$k := \inf\{j \geq i : \tau_{j+1} \geq \sigma_j\}$$

and stopping times

$$\tilde{\sigma}_j = \begin{cases} \tau_j & \text{if } j \leq k \\ \sigma_j & \text{if } j > k \end{cases}$$

and analogously

$$\tilde{\tau}_j = \begin{cases} \sigma_j & \text{if } j \leq k \\ \tau_j & \text{if } j > k \end{cases}.$$

Note that $(\tilde{\sigma}_j)_{j=i}^n$ and $(\tilde{\tau}_j)_{j=i}^n$ are again two families of increasing stopping times, since

$$\begin{aligned}
\tilde{\tau}_i &= \sigma_i \\
&\leq \tilde{\tau}_{i+1} = \mathbb{1}_{\tau_{i+1} \geq \sigma_i} \tau_{i+1} + \mathbb{1}_{\tau_{i+1} < \sigma_i} \sigma_{i+1} \\
&\leq \tilde{\tau}_{i+2} = \mathbb{1}_{\tau_{i+1} \geq \sigma_i} \tau_{i+2} + \mathbb{1}_{\tau_{i+1} < \sigma_i} (\mathbb{1}_{\tau_{i+2} \geq \sigma_{i+1}} \tau_{i+2} + \mathbb{1}_{\tau_{i+2} < \sigma_{i+1}} \sigma_{i+2}) \\
&\leq \tilde{\tau}_{i+3} = \dots,
\end{aligned} \tag{2.2}$$

and similarly for $\tilde{\sigma}_j$.

Example 2.2 To illustrate this construction, consider the following sequences of stopping times for Brownian motion started at $B_0 = 0$. Let $\sigma_j = H_{\pm(j+1)} := \inf\{t \geq 0 : |B_t| \geq j+1\}$, and $\tau_j = j$. The idea is that we want to construct a new sequence $(\tilde{\sigma}_j)$ which 'starts' with τ_0, but reverts to the original (σ_j) sequence as soon as possible. Correspondingly, we wish to construct the sequence $(\tilde{\tau}_j)$ which starts like (σ_j), but reverts to (τ_j) as soon as possible. As above, $k = \inf\{j \geq i : \tau_{j+1} \geq \sigma_j\}$ is the first time (if at all) that B leaves the interval $[-j, j]$ before time j. If this never happens, then the two sequences will just swap.

That is, if the sequences switch back, then the construction gives:

$$\begin{aligned}
(\sigma_j)_{j=0,\dots,n} &= (H_{\pm 1}, H_{\pm 2}, \dots, H_{\pm(n+1)}), \\
(\tilde{\sigma}_j)_{j=0,\dots,n} &= (0, 1, 2, \dots, k, H_{\pm(k+2)}, \dots, H_{\pm(n+1)}) \\
(\tau_j)_{j=0,\dots,n} &= (0, 1, 2, \dots, n), \\
(\tilde{\tau}_j)_{j=0,\dots,n} &= (H_{\pm 1}, H_{\pm 2}, \dots, H_{\pm(k+1)}, k+1, \dots, n).
\end{aligned}$$

Note in particular that with this swap, the $\tilde{\sigma}$ stopping times stop instantly, while the $\tilde{\tau}$ times no longer stop at time 0.

Definition 2.3 A pair $((f, s_1, \dots, s_{i-1}, s), (g, t_1, \dots, t_{i-1}, t)) \in S^{\otimes i} \times S^{\otimes i}$ constitutes an i-th stop-go pair, written $((f, s_1, \dots, s_{i-1}, s), (g, t_1, \dots, t_{i-1}, t)) \in \mathsf{SG}_i$, if $f(s) = g(t)$ and for all families of $\mathcal{F}^B$-stopping times $\sigma_i \leq \dots \leq \sigma_n$, $0 = \tau_i \leq \tau_{i+1} \leq \dots \leq \tau_n$ satisfying $0 < \mathbb{E}[\sigma_j] < \infty$ for all $i \leq j \leq n$ and $0 \leq \mathbb{E}[\tau_j] < \infty$ for all $i < j \leq n$

$$\begin{aligned}
&\mathbb{E}\left[\gamma\left(((f \oplus B)_u)_{u \leq s+\sigma_n}, s_1, \dots, s_{i-1}, s + \sigma_i, s + \sigma_{i+1}, \dots, s + \sigma_n\right)\right] \\
&\quad + \mathbb{E}\left[\gamma\left(((g \oplus B)_u)_{u \leq t+\tau_n}, t_1, \dots, t_{i-1}, t, t + \tau_{i+1}, \dots, t + \tau_n\right)\right] \\
&> \mathbb{E}\left[\gamma\left(((f \oplus B)_u)_{u \leq s+\tilde{\sigma}_n}, s_1, \dots, s_{i-1}, s, s + \tilde{\sigma}_{i+1}, \dots, s + \tilde{\sigma}_n\right)\right] \\
&\quad + \mathbb{E}\left[\gamma\left(((g \oplus B)_u)_{u \leq t+\tilde{\tau}_n}, t_1, \dots, t_{i-1}, t + \tilde{\tau}_i, t + \tilde{\tau}_{i+1}, \dots, t + \tilde{\tau}_n\right)\right],
\end{aligned} \tag{2.3}$$

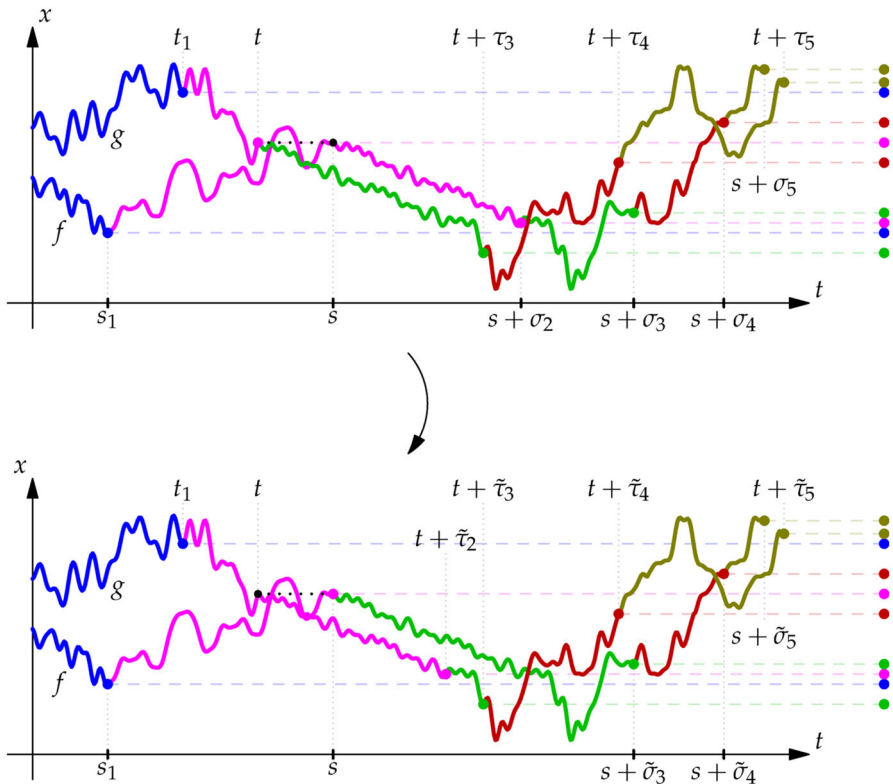**Fig. 1** We show a potential stop-go pair. In the top picture, we show the pair $((f, s_1, s), (g, t_1, t))$ with corresponding stopping times $\tau_3, \ldots, \tau_5$ and $\sigma_2, \ldots, \sigma_5$. In the bottom picture, the stopping times $\tilde{\tau}_2, \ldots \tilde{\tau}_5$ and $\tilde{\sigma}_3, \ldots, \tilde{\sigma}_5$ are shown. Note that the first time that the stopping rules can 'revert' to their original times are $\tilde{\tau}_3$ and $\tilde{\sigma}_3$

whenever both sides are well defined and the left hand side is finite. (See Fig. 1.)

A pair $((f, s_1, \ldots, s_{i-1}, s), (g, t_1, \ldots, t_{i-1}, t)) \in S^{\otimes i} \times S^{\otimes i}$ constitutes a secondary i-th stop-go pair, written $((f, s_1, \ldots, s_{i-1}, s), (g, t_1, \ldots, t_{i-1}, t)) \in \mathsf{SG}_{2,i}$, if $f(s) = g(t)$ and for all families of $\mathcal{F}^B$-stopping times $\sigma_i \leq \ldots \leq \sigma_n$, $0 = \tau_i \leq \tau_{i+1} \leq \ldots \leq \tau_n$ satisfying $0 < \mathbb{E}[\sigma_j] < \infty$ for all $i \leq j \leq n$ and $0 \leq \mathbb{E}[\tau_j] < \infty$ for all $i < j \leq n$ the inequality (2.3) holds with $\geq$ and if there is equality we have

$$\mathbb{E}\left[\gamma_2\big(((f \oplus B)_u)_{u \leq s+\sigma_n}, s_1, \ldots, s_{i-1}, s + \sigma_i, s + \sigma_{i+1}, \ldots, s + \sigma_n\big)\right]$$

$$+ \mathbb{E}\left[\gamma_2\big(((g \oplus B)_u)_{u \leq t+\tau_n}, t_1, \ldots, t_{i-1}, t \quad, t + \tau_{i+1}, \ldots, t + \tau_n\big)\right]$$

$$> \mathbb{E}\left[\gamma_2\big(((f \oplus B)_u)_{u \leq s+\tilde{\sigma}_n}, s_1, \ldots, s_{i-1}, s \quad, s + \tilde{\sigma}_{i+1}, \ldots, s + \tilde{\sigma}_n\big)\right]$$

$$+ \mathbb{E}\left[\gamma_2\big(((g \oplus B)_u)_{u \leq t+\tilde{\tau}_n}, t_1, \ldots, t_{i-1}, t + \tilde{\tau}_i, t + \tilde{\tau}_{i+1}, \ldots, t + \tilde{\tau}_n\big)\right], \quad (2.4)$$

whenever both sides are well defined and the left hand side (of (2.4)) is finite.

For $0 \leq i < j \leq n$ we define $\mathrm{proj}_{S^{\otimes i}} : S^{\otimes j} \to S^{\otimes i}$ by $(f, s_1, \ldots, s_j) \mapsto (f_{\llcorner[0,s_i]}, s_1, \ldots, s_i)$ where we take $s_0 = 0$, $S^{\otimes 0} = \mathbb{R}$, and $f_{\llcorner[0,0]} := f(0) \in \mathbb{R}$.

Definition 2.4 A set $\Gamma = (\Gamma_1, \ldots, \Gamma_n)$ with $\Gamma_i \subseteq S^{\otimes i}$ measurable for each i is called $\gamma_2|\gamma$-monotone if for each $1 \leq i \leq n$

$$\mathsf{SG}_{2,i} \cap (\Gamma_i^< \times \Gamma_i) = \emptyset,$$

where

$$\Gamma_i^< = \{(f, s_1, \ldots, s_{i-1}, u) : \text{ there exists } (g, s_1, \ldots, s_{i-1}, s)$$
$$\in \Gamma_i, s_{i-1} \leq u < s, g_{\llcorner[0,u]} = f\},$$

and $\mathrm{proj}_{S^{\otimes i-1}}(\Gamma_i) \subseteq \Gamma_{i-1}$.

Theorem 2.5 (Monotonicity principle) *Let $\gamma, \gamma_2 : S^{\otimes n} \to \mathbb{R}$ be Borel measurable, B be a Brownian motion on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with $B_0 \sim \mu_0$ and let $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_n)$ be an optimizer of* (OptMSEP$_2$). *Then there exists a $\gamma_2|\gamma$-monotone set $\Gamma = (\Gamma_1, \ldots, \Gamma_n)$ supporting $\hat{\tau}$ in the sense that $\mathbb{P}$- a.s. for all $1 \leq i \leq n$*

$$((B_s)_{s \leq \tau_i}, \tau_1, \ldots, \tau_i) \in \Gamma_i. \tag{2.5}$$

Remark 2.6 (1) We will also consider ternary or j-ary optimization problems given j Borel measurable functions $\gamma_1, \ldots, \gamma_j : S^{\otimes n} \to \mathbb{R}$ leading to ternary or j-ary i-th stop-go pairs $\mathsf{SG}_{i,3}, \ldots, \mathsf{SG}_{i,j}$ for $1 \leq i \leq n$, the notion of $\gamma_j|\ldots|\gamma_1$-monotone sets and a corresponding monotonicity principle. To save (digital) trees we leave it to the reader to write down the corresponding definitions.

(2) Intuitively, the sets Γ_i in Definition 2.4 could be simply defined to be the projections of Γ_n onto $S^{\otimes i}$, however this would not guarantee measurability of the sets $S^{\otimes i}$. Hence we need a slightly more involved statement of Theorem 2.5.

2.2 New *n*-marginal embeddings

2.2.1 The *n*-marginal Root embedding

The classical Root embedding [56] establishes the existence of a barrier (or right-barrier) $\mathcal{R} \subseteq \mathbb{R}_+ \times \mathbb{R}$ such that the first hitting time of $\mathcal{R}$ solves the Skorokhod embedding problem. A barrier $\mathcal{R}$ is a Borel set such that $(s, x) \in \mathcal{R} \Rightarrow (t, x) \in \mathcal{R}$ for all $t > s$. Moreover, the Root embedding has the property that it minimises $\mathbb{E}[h(\tau)]$ for a strictly convex function $h : \mathbb{R}_+ \to \mathbb{R}$ over all solutions to the Skorokhod embedding problem, cf. [57].

We will show that there is a unique n- marginal Root embedding in the sense that there are n barriers $(\mathcal{R}^i)_{i=1}^n$ such that for each $i \leq n$ the first hitting time of $\mathcal{R}^i$ *after* hitting $\mathcal{R}^{i-1}$ embeds μ_i.

Theorem 2.7 (*n*-marginal Root embedding, c.f. [18]) *Put* γ_i : $S^{\otimes n} \to \mathbb{R}, (f, s_1, \ldots, s_n) \mapsto h(s_i)$ *for some strictly convex function* $h : \mathbb{R}_+ \to \mathbb{R}$ *and assume that* (OptMSEP) *is well posed. Then there exist n barriers* $(\mathcal{R}^i)_{i=1}^n$ *such that defining*

$$\tau_1^{Root}(\omega) = \inf\{t \geq 0 : (t, B_t(\omega)) \in \mathcal{R}^1\}$$

and for $1 < i \leq n$

$$\tau_i^{Root}(\omega) = \inf\{t \geq \tau_{i-1}^{Root}(\omega) : (t, B_t(\omega)) \in \mathcal{R}^i\}$$

the multi-stopping time $(\tau_1^{Root}, \ldots, \tau_n^{Root})$ *minimises*

$$\mathbb{E}[h(\tilde{\tau}_i)]$$

simultaneously for all $1 \leq i \leq n$ *among all increasing families of stopping times* $(\tilde{\tau}_1, \ldots, \tilde{\tau}_n)$ *such that* $B_{\tilde{\tau}_j} \sim \mu_j$ *for all* $1 \leq j \leq n$ *and* $\mathbb{E}[\tau_n] < \infty$. *This solution is unique in the sense that for any solution* $\tilde{\tau}_1, \ldots, \tilde{\tau}_n$ *of such a barrier-type we have* $\tau_i^{Root} = \tilde{\tau}_i$ *a.s.*

Proof Fix a permutation κ of $\{1, \ldots, n\}$. We consider the functions $\tilde{\gamma}_1 = \gamma_{\kappa(1)}, \ldots, \tilde{\gamma}_n = \gamma_{\kappa(n)}$ on $S^{\otimes n}$ and the corresponding family of *n*-ary minimisation problems, (OptMSEP$_n$). Let $(\tau_1^{Root}, \ldots, \tau_n^{Root})$ be an optimiser of $P_{\tilde{\gamma}_n|\ldots|\tilde{\gamma}_1}$. By the *n*-ary version of Theorem 2.1, choose an optimizer $(\tau_1^{Root}, \ldots, \tau_n^{Root})$ of (OptMSEP$_n$) and, by the corresponding version of Theorem 2.5, a $\tilde{\gamma}_n|\ldots|\tilde{\gamma}_1$-monotone family of sets $(\Gamma_1, \ldots, \Gamma_n)$ supporting $(\tau_1^{Root}, \ldots, \tau_n^{Root})$. Hence for every $i \leq n$ we have $\mathbb{P}$-a.s.

$$((B_s)_{s \leq \tau_i}, \tau_1^{Root}, \ldots, \tau_i^{Root}) \in \Gamma_i,$$

and

$$(\Gamma_i^< \times \Gamma_i) \cap \mathsf{SG}_{i,n} = \emptyset.$$

We claim that, for all $1 \leq i \leq n$ we have

$$\mathsf{SG}_{i,n} \supseteq \{((f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i)) : f(s_i) = g(t_i), s_i > t_i\}.$$

Fix $(f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i) \in S^{\otimes i}$ satisfying $s_i > t_i$ and consider two families of stopping times $(\sigma_j)_{j=i}^n$ and $(\tau_j)_{j=i}^n$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with their modifications $(\tilde{\sigma}_j)_{j=i}^n$ and $(\tilde{\tau}_j)_{j=i}^n$ as in Sect. 2.1. Put

$$j_1 := \inf\{m \geq 1 : \kappa(m) \geq i\}$$

and inductively for $1 < a \leq n - i + 1$

$$j_a := \inf\{m \geq j_{a-1} : \kappa(m) \geq i\}.$$

Let $l = \arg\min\{a : \mathbb{P}[\sigma_{j_a} \neq \tilde{\sigma}_{j_a}] > 0\}$. By the definition of $\tilde{\sigma}_j$ and $\tilde{\tau}_j$ we have in case of $j_l = i$ the equality $\{\sigma_{j_l} \neq \tilde{\sigma}_{j_l}\} = \Omega$ and for $j_l > i$ it holds that

$$\{\sigma_{j_l} \neq \tilde{\sigma}_{j_l}\} = \bigcap_{i \leq k < j_l} \{\sigma_k > \tau_{k+1}\}.$$

As $\tau_k \leq \tau_{k+1}$, in particular, we have on $\{\sigma_{j_l} \neq \tilde{\sigma}_{j_l}\}$ the inequality $\sigma_k > \tau_k$ for every $i \leq k \leq j_l$. The strict convexity of h and $s > t$ implies

$$\mathbb{E}[h(s + \sigma_{j_l})] + \mathbb{E}[h(t + \tau_{j_l})] > \mathbb{E}[h(s + \tilde{\sigma}_{j_l})] + \mathbb{E}[h(t + \tilde{\tau}_{j_l})].$$

Hence, we get a strict inequality in (the corresponding $\kappa^{-1}(j_l)$-ary version of) (2.4) and the claim is proven.

Then we define for each $1 \leq i \leq n$

$$\mathcal{R}^i_{\mathrm{CL}} := \{(s, x) \in \mathbb{R}_+ \times \mathbb{R} : \exists (g, t_1, \ldots, t_i) \in \Gamma_i, g(t_i) = x, s \geq t_i\}$$

and

$$\mathcal{R}^i_{\mathrm{OP}} := \{(s, x) \in \mathbb{R}_+ \times \mathbb{R} : \exists (g, t_1, \ldots, t_i) \in \Gamma_i, g(t_i) = x, s > t_i\}.$$

Following the argument in the proof of Theorem 2.1 in [2], we define τ^1_{CL} and τ^1_{OP} to be the first hitting times of $\mathcal{R}^1_{\mathrm{CL}}$ and $\mathcal{R}^1_{\mathrm{OP}}$ respectively to see that actually $\tau^1_{\mathrm{CL}} \leq \tau^{Root}_1 \leq \tau^1_{\mathrm{OP}}$ and $\tau^1_{\mathrm{CL}} = \tau^1_{\mathrm{OP}}$ a.s. by the strong Markov property. Then we can inductively proceed and define

$$\tau^i_{\mathrm{CL}} := \inf\{t \geq \tau^{i-1}_{\mathrm{CL}} : (t, B_t) \in \mathcal{R}^i_{\mathrm{CL}}\}$$

and

$$\tau^i_{\mathrm{OP}} := \inf\{t \geq \tau^{i-1}_{\mathrm{CL}} : (t, B_t) \in \mathcal{R}^i_{\mathrm{OP}}\}.$$

By the very same argument we see that $\tau^i_{\mathrm{CL}} \leq \tau^{Root}_i \leq \tau^i_{\mathrm{OP}}$ and in fact $\tau^i_{\mathrm{CL}} = \tau^i_{\mathrm{OP}}$.

Finally, we need to show that the choice of the permutation κ does not matter. This follows from a straightforward adaptation of the argument of Loynes [46] (see also [2, Remark 2.3] and [18, Proof of Lemma 2.4]) to the multi-marginal set up. Indeed, the first barrier $\mathcal{R}^1$ is unique by Loynes original argument. This implies that the second barrier is unique because Loynes argument is valid for a general starting distribution of the process (t, B_t) in $\mathbb{R}_+ \times \mathbb{R}$ and we can conclude by induction. $\qquad\square$

Remark 2.8 (1) In the last theorem, the result stays the same if we take different strictly convex functions h_i for each i.

(2) Moreover, it is easy to see that the proof is simplified if one starts with the objective $\sum_{i=1}^{n} h_i(\tau_i)$, which removes the need for taking an arbitrary permutation of the indices at the start. Of course, to get the more general conclusion, one needs to consider these permutations.

Corollary 2.9 *Let $h : \mathbb{R}_+ \to \mathbb{R}$ be a strictly convex function and let $\gamma : S^{\otimes n} \to \mathbb{R}, (f, s_1, \ldots, s_n) \mapsto \sum_{i=1}^{n} h(t_i)$. Let $\tau^{Root} = (\tau_1^{Root}, \ldots, \tau_n^{Root})$ be the minimizer of Theorem 2.7. Then it also minimizes*

$$\mathbb{E}[\gamma(\tilde{\tau}_1, \ldots, \tilde{\tau}_n)]$$

among all increasing families of stopping times $\tilde{\tau}_1 \leq \ldots \leq \tilde{\tau}_n$ satisfying $B_{\tilde{\tau}_i} \sim \mu_i$ for all $1 \leq i \leq n$.

2.2.2 The *n*-marginal Rost embedding

The classical Rost embedding [57] establishes the existence of an inverse barrier (or left-barrier) $\mathcal{R} \subseteq \mathbb{R}_+ \times \mathbb{R}$ such that the first hitting time of $\mathcal{R}$ solves the Skorokhod embedding problem. An inverse barrier $\mathcal{R}$ is a Borel set such that $(t, x) \in \mathcal{R} \Rightarrow (s, x) \in \mathcal{R}$ for all $s < t$. Moreover, the Rost embedding has the property that it maximises $\mathbb{E}[h(\tau)]$ for a strictly convex function $h : \mathbb{R}_+ \to \mathbb{R}$ over all solutions to the Skorokhod embedding problem, cf. [57]. Similarly to the Root embedding it follows that

Theorem 2.10 (*n*-marginal Rost embedding) *Put $\gamma_i : S^{\otimes n} \to \mathbb{R}, (f, s_1, \ldots, s_n) \mapsto -h(s_i)$ for some strictly convex function $h : \mathbb{R}_+ \to \mathbb{R}$ and assume that (OptMSEP) is well posed. Then there exist n inverse barriers $(\mathcal{R}^i)_{i=1}^{n}$ such that defining*

$$\tau_1^{Rost}(\omega) = \inf\{t \geq 0 : (t, B_t(\omega)) \in \mathcal{R}^1\}$$

and for $1 < i \leq n$

$$\tau_i^{Rost}(\omega) = \inf\{t \geq \tau_{i-1}^{Rost}(\omega) : (t, B_t(\omega)) \in \mathcal{R}^i\}$$

the multi-stopping time $(\tau_1^{Rost}, \ldots, \tau_n^{Rost})$ maximises

$$\mathbb{E}[h(\tau_i)]$$

simultaneously for all $1 \leq i \leq n$ among all increasing families of stopping times $(\tau_1, \ldots, \tau_n)$ such that $B_{\tau_j} \sim \mu_j$ for all $1 \leq j \leq n$ and $\mathbb{E}[\tau_n] < \infty$. Moreover, it also maximises

$$\sum_{i=1}^{n} \mathbb{E}[h(\tau_i)].$$

This solution is unique in the sense that for any solution $\tilde{\tau}_1, \ldots, \tilde{\tau}_n$ of such a barrier-type we have $\tau_i^{Rost} = \tilde{\tau}_i$.

The proof of this theorem goes along the very same lines as the proof of Theorem 2.7. The only difference is that due to the maximisation we get

$$\mathsf{SG}_{i,n} \supseteq \{(f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i) : f(s_i) = g(t_i), s_i < t_i\}$$

leading to inverse barriers. We omit the details.

2.2.3 The *n*-marginal Azéma–Yor embedding

For $(f, s_1, \ldots, s_n) \in S^{\otimes n}$ we will use the notation $\bar{f}_{s_i} := \max_{0 \leq s \leq s_i} f(s)$.

Theorem 2.11 (*n*-marginal Azéma–Yor solution) *There exist n barriers* $(\mathcal{R}^i)_{i=1}^n$ *such that defining*

$$\tau_1^{AY} = \inf\{t \geq 0 : (\bar{B}_t, B_t) \in \mathcal{R}^1\}$$

and for $1 < i \leq n$

$$\tau_i^{AY} = \inf\{t \geq \tau_{i-1}^{AY} : (\bar{B}_t, B_t) \in \mathcal{R}^i\}$$

the multi-stopping time $(\tau_1^{AY}, \ldots, \tau_n^{AY})$ *maximises*

$$\mathbb{E}\left[\sum_{i=1}^n \bar{B}_{\tau_i}\right]$$

among all increasing families of stopping times $(\tau_1, \ldots, \tau_n)$ *such that* $B_{\tau_j} \sim \mu_j$ *for all* $1 \leq j \leq n$ *and* $\mathbb{E}[\tau_n] < \infty$. *This solution is unique in the sense that for any solution* $\tilde{\tau}_1, \ldots, \tilde{\tau}_n$ *of such a barrier-type we have* $\tau_i^{AY} = \tilde{\tau}_i$.

We emphasise that this result has not appeared previously in the literature in this generality; previously the most general result was due to [30,51], which proved a closely related result under an additional condition on the measures, which is not necessary here. Unlike our solution, however, the constructions of [30,51] are constructive.

Remark 2.12 In fact, similarly to the *n*-marginal Root and Rost solutions τ^{AY} simultaneously solves the optimization problems

$$\sup\{\mathbb{E}[\bar{B}_{\tilde{\tau}_i}] : \tilde{\tau}_1 \leq \ldots \leq \tilde{\tau}_n, B_{\tilde{\tau}_1} \sim \mu_1, \ldots, B_{\tilde{\tau}_n} \sim \mu_n\}$$

for each *i* which of course implies Theorem 2.11 (see also Remark 2.8.2). To keep the presentation readable, we only prove the less general version.

Proof Fix a bounded and strictly increasing continuous function $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and consider the continuous functions $\gamma(f, s_1, \ldots, s_n) = -\sum_{i=1}^n \bar{f}_{s_i}$ and $\tilde{\gamma}(f, s_1, \ldots, s_n) = \sum_{i=1}^n \varphi(\bar{f}_{s_i}) f(s_i)^2$ defined on $S^{\otimes n}$. Pick, by Theorem 2.1, a minimizer τ^{AY} of (OptMSEP$_2$) and, by Theorem 2.5, a $\tilde{\gamma}|\gamma$-monotone family of sets $(\Gamma_i)_{i=1}^n$ supporting $\tau^{AY} = (\tau_i^{AY})_{i=1}^n$ such that for all $i \leq n$

$$\mathsf{SG}_{i,2} \cap (\Gamma_i^< \times \Gamma_i) = \emptyset.$$

We claim that

$$\mathsf{SG}_{i,2} \supseteq \{((f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i)) \in S^{\otimes i} \times S^{\otimes i} : f(s_i) = g(t_i), \bar{f}_{s_i} > \bar{g}_{t_i}\}.$$
$$(2.6)$$

Indeed, pick $((f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i)) \in S^{\otimes i} \times S^{\otimes i}$ with $f(s_i) = g(t_i)$ and $\bar{f}_{s_i} > \bar{g}_{t_i}$ and take two families of stopping times $(\sigma_j)_{j=i}^n$ and $(\tau_j)_{j=i}^n$ together with their modifications $(\tilde{\sigma}_j)_{j=i}^n$ and $(\tilde{\tau}_j)_{j=i}^n$ as in Sect. 2.1. We assume that they live on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ additionally supporting a standard Brownian motion W. Observe that (as written out in the proof of Theorem 2.7) on $\{\sigma_j \neq \tilde{\sigma}_j\}$ it holds that $\sigma_j > \tau_j$. Hence, on this set we have $\bar{W}_{\sigma_j} \geq \bar{W}_{\tau_j}$. This implies that for $\omega \in \{\sigma_j \neq \tilde{\sigma}_j\}$ (and hence $\tilde{\sigma}_j = \tau_j, \tilde{\tau}_j = \sigma_j$)

$$
\begin{aligned}
&\bar{f}_{s_i} \vee (f(s_i) + \bar{W}_{\sigma_j}(\omega)) + \bar{g}_{t_i} \vee (g(t_i) + \bar{W}_{\tau_j}(\omega)) \\
&\quad \leq \bar{f}_{s_i} \vee (f(s_i) + \bar{W}_{\tilde{\sigma}_j}(\omega)) + \bar{g}_{t_i} \vee (g(t_i) + \bar{W}_{\tilde{\tau}_j}(\omega)),
\end{aligned}
\tag{2.7}
$$

with a strict inequality unless either $\bar{W}_{\sigma_j}(\omega) \leq \bar{g}_{t_i} - g(t_i)$ or $\bar{W}_{\tau_j} \geq \bar{f}_{s_i} - f(s_i)$. On the set $\{\sigma_j = \tilde{\sigma}_j\}$ we do not change the stopping rule for the j-th stopping time and hence we get a (pathwise) equality in (2.7). Thus, we always have a strict inequality in (2.4) unless a.s. either $\bar{W}_{\sigma_j}(\omega) \leq \bar{g}_{t_i} - g(t_i)$ or $\bar{W}_{\tau_j} \geq \bar{f}_{s_i} - f(s_i)$ for all j. However, in that case we have for all j such that $\mathbb{P}[\sigma_j \neq \tilde{\sigma}_j] > 0$ (there is at least one such j, namely $j = i$)

$$
\begin{aligned}
&\mathbb{E}\left[\varphi(\bar{f}_{s_i})(f(s_i) + W_{\sigma_j})^2\right] + \mathbb{E}\left[\varphi(\bar{g}_{t_i})(g(t_i) + W_{\tau_j})^2\right] \\
&\quad > \mathbb{E}\left[\varphi(\bar{f}_{s_i})(f(s_i) + W_{\tilde{\sigma}_j})^2\right] + \mathbb{E}\left[\varphi(\bar{g}_{t_i})(g(t_i) + W_{\tilde{\tau}_j})^2\right].
\end{aligned}
$$

Hence, $((f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i)) \in \mathsf{SG} \subseteq \mathsf{SG}_2$ in the first case and in the second case we have $((f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i)) \in \mathsf{SG}_2$ proving (2.6).

For each $i \leq n$ we define

$$
\mathcal{R}_{\mathrm{OP}}^i := \{(m, x) : \exists (f, s_1, \ldots, s_i) \in \Gamma_i, f(s_i) = x, \bar{f}_{s_i} < m\}
$$

and

$$
\mathcal{R}_{\mathrm{CL}}^i := \{(m, x) : \exists (f, s_1, \ldots, s_i) \in \Gamma_i, f(s_i) = x, \bar{f}_{s_i} \leq m\}
$$

with respective hitting times ($\tau^0 = 0$)

$$
\tau_{\mathrm{OP}}^i := \inf\{t \geq \tau_{\mathrm{CL}}^{i-1} : (\bar{B}_t, B_t) \in \mathcal{R}_{\mathrm{OP}}^i\}
$$

and

$$
\tau_{\mathrm{CL}}^i := \inf\{t \geq \tau_{\mathrm{CL}}^{i-1} : (\bar{B}_t, B_t) \in \mathcal{R}_{\mathrm{CL}}^i\}.
$$

We will show inductively on i that firstly $\tau_{\mathrm{CL}}^i \leq \tau_i^{AY} \leq \tau_{\mathrm{OP}}^i$ a.s. and secondly $\tau_{\mathrm{CL}}^i = \tau_{\mathrm{OP}}^i$ a.s. proving the theorem. The case $i = 1$ has been settled in [2]. So let us assume $\tau_{\mathrm{CL}}^{i-1} = \tau_{\mathrm{OP}}^{i-1}$ a.s. Then $\tau_{\mathrm{CL}}^i \leq \tau_i^{AY}$ follows from the definition of τ_{CL}^i. To show

that $\tau_i^{AY} \leq \tau_{OP}^i$ pick ω satisfying $((B_s(\omega))_{s \leq \tau_i^{AY}}, \tau_1^{AY}(\omega), \ldots, \tau_i^{AY}(\omega)) \in \Gamma_i$ and assume that $\tau_{OP}^i(\omega) < \tau_i^{AY}(\omega)$. Then there exists $s \in [\tau_{OP}^i(\omega), \tau_i^{AY}(\omega))$ such that $f := (B_r(\omega))_{r \leq s}$ satisfies $(\bar{f}, f(s)) \in \mathcal{R}_{OP}^i$. Since $\tau_{CL}^{i-1}(\omega) \leq \tau_{OP}^i(\omega) \leq s < \tau_i^{AY}(\omega)$ we have $(f, \tau_{CL}^1(\omega), \ldots, \tau_{CL}^{i-1}(\omega), s) \in \Gamma_i^<$. By definition of $\mathcal{R}_{OP}^i$, there exists $(g, t_1, \ldots, t_i) \in \Gamma_i$ such that $f(s) = g(t_i)$ and $\bar{g}_{t_i} < \bar{f}_s$, yielding a contradiction to (2.6).

Finally, we need to show that $\tau_{CL}^i = \tau_{OP}^i$ a.s. Before we proceed we give a short reminder of the case $i = 1$ from [2, Theorem 6.5]. We define

$$\tilde{\psi}_0^1(m) = \sup\{x : \exists (m, x) \in \mathcal{R}_{CL}^1\}.$$

From the definition of $\mathcal{R}_{CL}^1$, we see that $\tilde{\psi}_0^1(m)$ is increasing, and we define the right-continuous function $\psi_+^1(m) = \tilde{\psi}_0^1(m+)$, and the left-continuous function $\psi_-^1(m) = \tilde{\psi}_0^1(m-)$. It follows from the definitions of τ_{OP}^1 and τ_{CL}^1 that:

$$\tau_+ := \inf\{t \geq 0 : B_t \leq \psi_+^1(\overline{B}_t)\} \leq \tau_{CL}^1 \leq \tau_{OP}^1 \leq \inf\{t \geq 0 : B_t < \psi_-^1(\overline{B}_t)\} =: \tau_-.$$

As $\tilde{\psi}_0^1$ has at most countably many jump points (discontinuity points) it is easily checked that $\tau_- = \tau_+$ a.s. and hence $\tau_{CL}^1 = \tau_{OP}^1 = \tau_1^{AY}$. Note also that the law $\bar{\mu}^1$ of $\bar{B}_{\tau_1^{AY}}$ can have an atom only at the rightmost point of its support. Hence, with $\pi^1 := \text{Law}(\bar{B}_{\tau_1^{AY}}, B_{\tau_1^{AY}})$, the measure $\pi^1_{\llcorner\{(x,y):y<x\}}$ has a density with respect to Lebesgue measure when projected onto the first coordinate.

Defining these quantities in obvious analogy for $j \in \{2, \ldots, n\}$, we need to prove $\tau_{CL}^{i+1} = \tau_{OP}^{i+1} = \tau_{i+1}^{AY}$ assuming that π^i has continuous projection onto the horizontal axis. To do so, we decompose π^i into *free* and *trapped* particles

$$\pi_f^i := \pi^i_{\llcorner\{(m,x):x>\psi_-^i(m)\}}, \quad \pi_t^i := \pi^i_{\llcorner\{(m,x):x \leq \psi_-^i(m)\}}.$$

Here π_f^i refers to particles which are free to reach a new maximum, while π_t^i refers to particles which are trapped in the sense that they will necessarily hit $\mathcal{R}_{OP}^i$ (and thus also $\mathcal{R}_{CL}^i$) before they reach a new maximum. For particles started in π_f^i it follows precisely as above that the hitting times of $\mathcal{R}_{OP}^{i+1}$ and $\mathcal{R}_{CL}^{i+1}$ agree. For particles started in π_t^i this is a consequence of Lemma 2.13. Additionally, as above we find that $\pi^{i+1}_{\llcorner\{(x,y):y<x\}}$ has continuous projection onto the horizontal axis. □

Lemma 2.13 [41, Lemma 3.2] *Let μ be a probability measure on $\mathbb{R}^2$ such that the projection onto the horizontal axis $\text{proj}_x \mu$ is continuous (in the sense of not having atoms) and let $\psi : \mathbb{R} \to \mathbb{R}$ be a Borel function. Set*

$$R^{OP} := \{(x, y) : x > \psi(y)\}, \quad R^{CL} := \{(x, y) : x \geq \psi(y)\}.$$

Start a vertically moving Brownian motion B in μ and define

$$\tau^{\mathrm{OP}} := \inf\{t \geq 0 : (x, y + B_t) \in R^{\mathrm{OP}}\}, \quad \tau^{\mathrm{CL}} := \inf\{t \geq 0 : (x, y + B_t) \in R^{\mathrm{CL}}\}.$$

Then $\tau^{\mathrm{CL}} = \tau^{\mathrm{OP}}$ almost surely.

2.2.4 The *n*-marginal Perkins/Hobson–Pedersen embedding

For $(f, s_1, \ldots, s_n) \in S^{\otimes n}$ we will use the notation $\underline{f}_{s_i} := \min_{0 \leq s \leq s_i} f(s)$ to denote the running minimum of the path up to time s_i. (Recall also that $\bar{f}_{s_i}$ is the maximum of the path). In this section we will consider a generalisation of the embeddings of Perkins [54] and Hobson and Pedersen [39]. The construction of Perkins to solve the one-marginal problem with a trivial starting law can be shown to simultaneously minimise $\mathbb{E}[h(\bar{B}_\tau)]$ for any increasing function h, and maximise $\mathbb{E}[k(\underline{B}_\tau)]$ for any increasing function k, over all solutions of the embedding problem. Later Hobson and Pedersen [39] described a closely related construction which minimised $\mathbb{E}[h(\bar{B}_\tau)]$ over all solutions to the SEP with a general starting law. The solution of Perkins took the form:

$$\tau^P := \inf\{t \geq 0 : B_t \leq \gamma_-(\bar{B}_t) \text{ or } B_t \geq \gamma_+(\underline{B}_t)\}$$

for decreasing functions γ_+, γ_-. Hobson and Pedersen constructed, for the case of a general starting distribution, a stopping time

$$\tau^{HP} := \inf\{t \geq 0 : B_t \leq \gamma_-(\bar{B}_t) \text{ or } \overline{B}_t \geq G\}$$

where G was an appropriately chosen, $\mathcal{F}_0$-measurable random variable. (Here, recalling the discussion at the start of Sect. 2.1, we need to use the assumption that the filtration supports the Brownian motion, and an additional $\mathcal{F}_0$-measurable, independent uniform random variable; this additional information is enough then to construct a suitable G). They showed that τ^{HP} minimised $\mathbb{E}[h(\bar{B}_\tau)]$ for any increasing function, but it is clear that the second minimisation does not hold in general. In [39, Remark 2.3], the existence of a version of Perkins' construction for a general starting law is conjectured. Below we will show that the construction of Hobson and Pedersen can be generalised to the multi-marginal case, and sketch an argument that there are natural generalisations of the Perkins embedding to this situation, but argue that there is no 'canonical' generalisation of the Perkins embedding. To be more specific, for given increasing functions h, k, the embedding(s) which maximise $\mathbb{E}[k(\underline{B}_{\tau_n})]$ over all solutions to the multi-marginal embedding problem which minimise $\mathbb{E}[h(\bar{B}_{\tau_n})]$ will typically differ from the embeddings which minimise $\mathbb{E}[h(\bar{B}_{\tau_n})]$ over all maximisers of $\mathbb{E}[k(\underline{B}_{\tau_n})]$.

Theorem 2.14 (*n*-marginal 'Hobson–Pedersen' solution) *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be sufficiently rich that it supports a Brownian motion $(B_t)_{t \geq 0}$ starting with law μ_0, and an independent, uniformly distributed random variable Y, which is $\mathcal{F}_0$-measurable.*

Then there exist n left-barriers $(\mathcal{R}^i)_{i=1}^n$ and stopping times $\tau_1^* \leq \tau_2^* \leq \cdots \leq \tau_n^*$ where $\tau_i^* < \infty$ implies $B_{\tau_i^*} = \bar{B}_{\tau_i^*}$ such that

$$\tau_1^{HP} = \inf\{t \geq 0 : (\bar{B}_t, B_t) \in \mathcal{R}^1\} \wedge \tau_1^*$$

and for $1 < i \leq n$

$$\tau_i^{HP} = \inf\{t \geq \tau_{i-1}^{HP} : (\bar{B}_t, B_t) \in \mathcal{R}^i\} \wedge \tau_i^*$$

the multi-stopping time $(\tau_1^{HP}, \ldots, \tau_n^{HP})$ minimises

$$\mathbb{E}\left[\sum_{i=1}^n \bar{B}_{\tau_i}\right]$$

among all increasing families of stopping times $(\tau_1, \ldots, \tau_n)$ such that $B_{\tau_j} \sim \mu_j$ for all $1 \leq j \leq n$ and $\mathbb{E}[\tau_n] < \infty$.

Proof Fix a bounded and strictly increasing continuous function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ and consider the continuous functions $\gamma(f, s_1, \ldots, s_n) = \sum_{i=1}^n \bar{f}_{s_i}$, $\gamma_2(f, s_1, \ldots, s_n) = \sum_{i=1}^n \varphi(\bar{f}_{s_i}) f(s_i)^2$ defined on $S^{\otimes n}$. Pick, by Theorem 2.1, a minimizer τ^{HP} of (OptMSEP$_2$) and, by Theorem 2.5, a $\gamma_2|\gamma$-monotone family of sets $(\Gamma_i)_{i=1}^n$ supporting $\tau^{HP} = (\tau_i^{HP})_{i=1}^n$ such that for all $i \leq n$

$$\mathsf{SG}_{i,2} \cap (\Gamma_i^< \times \Gamma_i) = \emptyset.$$

By an essentially identical argument to that given in Theorem 2.11, we have

$$\mathsf{SG}_{i,2} \supseteq \left\{((f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i)) \in S^{\otimes i} \times S^{\otimes i} : f(s_i) = g(t_i), \bar{f}_{s_i} < \bar{g}_{t_i}\right\}.$$
$$(2.8)$$

Note that, given τ_i^{HP}, we can define stopping times $\tau_i^* := \tau_i^{HP}$ if $B_{\tau_i^{HP}} = \bar{B}_{\tau_i^{HP}}$ and to be infinite otherwise.

For each $i \leq n$ we define

$$\mathcal{R}_{OP}^i := \left\{(m, x) : \exists (f, s_1, \ldots, s_i) \in \Gamma_i, f(s_i) = x, \bar{f}_{s_i} > m, x < m\right\}$$

and

$$\mathcal{R}_{CL}^i := \left\{(m, x) : \exists (f, s_1, \ldots, s_i) \in \Gamma_i, f(s_i) = x, \bar{f}_{s_i} \geq m, x < m\right\}$$

with respective hitting times $(\tau^0 = 0)$

$$\tau_{OP}^i := \inf\{t \geq \tau_{OP}^{i-1} : (\bar{B}_t, B_t) \in \mathcal{R}_{OP}^i\}$$

and

$$\tau_{\mathrm{CL}}^i := \inf\{t \geq \tau_{\mathrm{CL}}^{i-1} : (\bar{B}_t, B_t) \in \mathcal{R}_{\mathrm{CL}}^i\}.$$

It can be shown inductively on i that firstly $\tau_{\mathrm{CL}}^i \wedge \tau_i^* \leq \tau_i^{HP} \leq \tau_{\mathrm{OP}}^i \wedge \tau_i^*$ a.s., and secondly $\tau_{\mathrm{CL}}^i \wedge \tau_i^* = \tau_{\mathrm{OP}}^i \wedge \tau_i^*$ a.s., proving the theorem. The proofs of these results are now essentially identical to the proof of Theorem 2.11. □

Of course, as before, a more general version of the statement (without the summation) can be proved, at the expense of a more complicated argument.

Remark 2.15 The result above says nothing about the uniqueness of the solution. However the following argument (also used in [2]) shows that any optimal solution (to both the primary and secondary optimisation problem in the proof of Theorem 2.14) will have the same barrier form: specifically, suppose that (τ^i) and (σ^i) are both optimal. Define a new stopping rule which, at time 0, chooses either the stopping rule (τ^i), or the stopping rule (σ^i), each with probability $1/2$. This stopping rule is also optimal (for both the primary and secondary rules), and the arguments above may be re-run to deduce the corresponding form of the optimal solution.

In fact, a more involved argument would appear to give uniqueness of the resulting barrier among the class of all such solutions; the idea is to use a Loynes-style argument as before, but applied both to the barrier and the *rate* of stopping at the maximum. The difficulty here is to argue that any stopping times of the form given above are essentially equivalent to another stopping time which simply stops at the maximum according to some rate which will be dependent only on the choice of the lower barrier (that is, in the language above, $\mathbb{P}(H_x^i < \tau_i^* = \tau_i^{HP} \leq H_{x+\varepsilon}^i)$ is independent of the choice of τ_i^{HP} for any x and $\varepsilon > 0$, where $H_x^i := \inf\{t \geq \tau_{i-1}^{HP} : B_t \geq x\}$). By identifying each of the possible optimisers with a canonical form of the optimiser, and using a Loynes-style argument which combines two stopping rules of the form above by taking the maximum of the left-barriers, and the fastest stopping rate of the rules, one can deduce that there is a unique sequence of barriers and stopping rate giving rise to an embedding of this form. We leave the details to the interested reader.

Remark 2.16 We conclude by considering informally the 'Perkins'-type construction implied by our methods. Recall that in the single marginal case, where $B_0 = 0$, the Perkins embedding simultaneously both maximises the law of the minimum, and minimises the law of the maximum. A slight variant of the methods above would suggest that one could adapt the arguments above to consider the optimiser which has the same primary objective as above, and also then aims to minimise the law of the minimum. In this case the arguments may be run to give stopping regions (for each marginal) which are barriers in the sense that it is the first hitting time of a left-barrier $\mathcal{R}$ which is left-closed in the sense that if (for a fixed x) a path with $\bar{f}_s = m, \underline{f}_s = j$ is stopped, then so too are all paths with $\bar{g}_s = m', \underline{g}_s = j'$, where $(m', -j') \prec (m, -j)$ and $\prec$ denotes the *lexicographical* ordering. With this definition, the general outline argument given above can proceed as usual, however we do not do this here since the final stage of the argument — showing that the closed and open hitting times of such

a region are equal — would appear to be much more subtle than previous examples, and so we leave this as an open problem for future work.

However, more notable is that in the multiple marginal case (and indeed, already to some extent in the case of a single marginal with a general starting law), the Perkins optimality property is no longer strictly preserved. To see why this might be the case (see also [39, Remark 2.3]) we note that, in the case of a single marginal, with trivial starting law, the embedding constructed via the double minimisation problems always stops at a time when the process sets a new minimum or a new maximum. At any given possible stopping point, the decision to stop should depend both on the current minimum, and the current maximum; however when the process is at a current maximum, both the current position and the current maximum are the same. In consequence, the decision to stop at e.g. a new maximum will only depend on the value of the minimum, and the optimisation problem relating to maximising a function of the maximum will be unaffected by the choice. In particular, it is never important which optimisation is the primary optimisation problem, and which is the secondary optimisation problem: in terms of the barrier-criteria established above, this can be seen by observing that in lexicographic ordering, $(m', -j') \prec (m, -j)$ is equivalent to $(-j', m') \prec (-j, m)$ if either $m = m'$ or $j = j'$.

On the other hand, with multiple marginals, we may have to consider possible stopping at times which do not correspond to setting a new maximum or minimum. Consider for example the case with $\mu_0 = \delta_0, \mu_1 = (\delta_1 + \delta_{-1})/2, \mu_2 = 2(\delta_2 + \delta_{-2})/5 + \delta_0/5$. In particular, the first stopping time, τ_1 must be the first hitting time of $\{-1, 1\}$, and if the process stops at 0 at the second stopping time, then to be optimal, it must stop there the first time it hits 0 after τ_1. If we consider the probability that we return to 0 after τ_1, before hitting $\{-2, 2\}$, then this is larger than $\frac{1}{5}$, and we need to choose a rule to determine which of the paths returning to 0 we should stop. It is clear that, if the primary optimisation is to minimise the law of the maximum, then this decision would only depend on the running maximum, while it will depend only on the running minimum if the primary and secondary objectives are switched. In particular, the two problems give rise to different optimal solutions. The difference here arises from the fact that we are not able to assume that all paths have either the same maximum, or the same minimum. As a consequence, we do not, in general, expect to recover a general version of the Perkins embedding, in the sense that there exists a multi-marginal embedding which minimises the law of the maximum, and maximises the law of the minimum simultaneously.

2.2.5 Further "classical" embeddings and other remarks

By combining the ideas and techniques from the previous sections and the techniques from [2, Section 6.2] we can establish the existence of n-marginal versions of the Jacka and Vallois embeddings and their siblings (replacing the local time with a suitably regular additive functional) as constructed in [2, Remark 7.13]. We leave the details to the interested reader.

We also remark that it is possible to get more detailed descriptions of the structure of the different barriers. At this point we only note that all the embeddings presented above have the nice property that their n-marginal solution restricted to the first $n - 1$

marginals is in fact the $n - 1$ marginal solution. This is a direct consequence of the extension of the Loynes argument to n-marginals as shown in the proof of Theorem 2.7. For a more detailed description of the barriers for the n-marginal Root embedding we refer to [18].

We also observe that, as in [2, Section 6.3], it is possible to deduce multi-marginal embeddings of some of the embeddings presented in the previous sections, e.g. Root and Rost, in higher dimensions. We leave the details to the interested reader.

2.2.6 A *n*-marginal version of the monotone martingale coupling

We next discuss the embedding giving rise to a multi-marginal version of the monotone martingale transport plan. Note that we need an extra assumption on the starting law μ_0, but on μ_0 only.

Theorem 2.17 (*n*-marginal martingale monotone transport plan) *Assume that* μ_0 *is continuous (in the sense that* $\mu_0(a) = 0$ *for all* $a \in \mathbb{R}$*). Let* $c : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be three times continuously differentiable with* $c_{xyy} < 0$*. Put* $\gamma_i : S^{\otimes n} \to \mathbb{R}, (f, s_1, \ldots, s_n) \mapsto c(f(0), f(s_i))$ *and assume that* (OptMSEP) *is well posed. Then there exist n barriers* $(\mathcal{R}^i)_{i=1}^n$ *such that defining*

$$\tau_1 = \inf\{t \geq 0 : (B_t - B_0, B_t) \in \mathcal{R}^1\}$$

and for $1 < i \leq n$

$$\tau_i = \inf\{t \geq \tau_{i-1} : (B_t - B_0, B_t) \in \mathcal{R}^i\}$$

the multi-stopping time $(\tau_1, \ldots, \tau_n)$ *minimises*

$$\mathbb{E}[c(B_0, B_{\tau_i})]$$

simultaneously for all $1 \leq i \leq n$ *among all increasing families of stopping times* $(\tilde{\tau}_1, \ldots, \tilde{\tau}_n)$ *such that* $B_{\tilde{\tau}_j} \sim \mu_j$ *for all* $1 \leq j \leq n$*. This solution is unique in the sense that for any solution* $\tilde{\tau}_1, \ldots, \tilde{\tau}_n$ *of such a barrier-type we have* $\tau_i = \tilde{\tau}_i$*.*

Remark 2.18 In the final stage of writing this article we learned of the work of Nutz et al. [49] on multi-period martingale optimal transport which (among various further results) provides an *n*-marginal version of the monotone martingale transport plan. Their methods are rather different from the ones employed in this article and in particular not related to the Skorokhod problem, but their solution is the same as the one presented here (see also [5]).

Proof of Theorem 2.17 The overall strategy of the proof, and in particular the first steps follow exactly the arguments encountered above. Fix a permutation κ of $\{1, \ldots, n\}$. We consider the functions $\tilde{\gamma}_1 = \gamma_{\kappa(1)}, \ldots, \tilde{\gamma}_n = \gamma_{\kappa(n)}$ on $S^{\otimes n}$ and the corresponding family of *n*-ary minimisation problems. Pick, by the n-ary version of Theorem 2.1,

an optimizer $(\tau_1, \ldots, \tau_n)$ and, by the n-ary version of Theorem 2.5, a $\tilde{\gamma}_n | \ldots | \tilde{\gamma}_1$-monotone family of sets $(\Gamma_1, \ldots, \Gamma_n)$ supporting $(\tau_1, \ldots, \tau_n)$, i.e. for every $i \leq n$ we have $\mathbb{P}$-a.s.

$$((B_s)_{s \leq \tau_i}, \tau_1, \ldots, \tau_i) \in \Gamma_i,$$

and

$$(\Gamma_i^< \times \Gamma_i) \cap \mathsf{SG}_{i,n} = \emptyset.$$

We claim that for all $1 \leq i \leq n$ we have

$$\mathsf{SG}_{i,n} \supseteq \{(f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i) : f(s_i) = g(t_i), g(0) > f(0)\}.$$

To this end, we have to consider $(f, s_1, \ldots, s_i), (g, t_1, \ldots, t_i) \in S^{\otimes i}$ satisfying $f(s_i) = g(t_i), f(s_i) - f(0) > g(t_i) - g(0)$ and consider two families of stopping times $(\sigma_j)_{j=i}^n$ and $(\tau_j)_{j=i}^n$ together with their modifications $(\tilde{\sigma}_j)_{j=i}^n$ and $(\tilde{\tau}_j)_{j=i}^n$ as in Sect. 2.1. However, since the modification of stopping times consists only of repeated swapping of the two stopping times what is effectively sufficient to prove is the following:

For $f(s) - f(0) > g(t) - g(0)$ and any stopping times ρ, σ, τ, where $\rho \leq \sigma$, we have for

$$\tilde{\sigma} := \sigma \mathbb{1}_{\rho \leq \tau} + \tau \mathbb{1}_{\rho > \tau}, \quad \tilde{\tau} := \tau \mathbb{1}_{\rho \leq \tau} + \sigma \mathbb{1}_{\rho > \tau}$$

the inequality

$$\begin{aligned}
\mathbb{E}[c(f(0), f(s) + B_\sigma)] &+ \mathbb{E}[c(g(0), g(t) + B_\tau)] \\
&> \mathbb{E}[c(f(0), f(s) + B_{\tilde{\sigma}})] + \mathbb{E}[c(g(0), g(t) + B_{\tilde{\tau}})],
\end{aligned} \tag{2.9}$$

and that this inequality is strict, provided that the set $\rho > \tau$ has positive probability.

To establish this inequality, of course only the parts were $\rho > \tau$ matters. Otherwise put, the inequality remains equally valid if we replace all of $\sigma, \tau, \tilde{\sigma}, \tilde{\tau}$ by $\tau \vee \sigma$ on the set $\rho \leq \tau$, in which case we have $\tilde{\sigma} = \tau, \tilde{\tau} = \sigma, \sigma \geq \tau$. Hence to prove (2.9) it is sufficient to show for $\alpha := \mathrm{Law}(B_\sigma), \beta := \mathrm{Law}(B_\tau)$ and $a := f(s) = g(t)$ that

$$\int c(f(0), a + x) \, d\alpha(x) + \int c(g(0), a + x) \, d\beta(x) > \int c(f(0), a + x) \, d\beta(x)$$
$$+ \int c(g(0), a + x) \, d\alpha(x).$$

To obtain this, we claim that

$$t \mapsto \int c(t, a + x) \, d\alpha(x) - \int c(t, a + x) \, d\beta(x)$$

is decreasing in t: This holds true since c_x is concave and β precedes α in the convex order (strictly if $\mathbb{P}(\rho > \tau) > 0$).

Having established the claim, we define for each $1 \leq i \leq n$

$$\mathcal{R}^i_{\mathrm{CL}} := \{(d, x) \in \mathbb{R}_+ \times \mathbb{R} : \exists (g, t_1, \ldots, t_i) \in \Gamma_i, g(t_i) = x, d \geq g(t_i) - g(0)\}$$

and

$$\mathcal{R}^i_{\mathrm{OP}} := \{(d, x) \in \mathbb{R}_+ \times \mathbb{R} : \exists (g, t_1, \ldots, t_i) \in \Gamma_i, g(t_i) = x, d > g(t_i) - g(0)\}.$$

Following the argument used above, we define τ^1_{CL} and τ^1_{OP} to be the first times the process $(B_t - B_0, B_t)_{t \geq 0}$ hits $\mathcal{R}^1_{\mathrm{CL}}$ and $\mathcal{R}^1_{\mathrm{OP}}$ respectively to see that actually $\tau^1_{\mathrm{CL}} \leq \tau_1 \leq \tau^1_{\mathrm{OP}}$.

It remains to show that $\tau^1_{\mathrm{CL}} = \tau^1_{\mathrm{OP}}$ (This has already been shown in [41, Prop. 3.1]; we present the argument for completeness). To this end, note that the hitting time of $(B_t - B_0, B_t)_{t \geq 0}$ into a barrier can equally well be interpreted as the hitting time of $(-B_0, B_t)_{t \geq 0}$ into a transformed (i.e. sheared through the transformation $(d, x) \mapsto (d - x, x)$) barrier. The purpose of this alteration is that the process $(-B_0, B_t)_{t \geq 0}$ moves only vertically and we can now apply Lemma 2.13 to establish that indeed $\tau^1_{\mathrm{CL}} = \tau^1_{\mathrm{OP}}$. Observe that at this stage the continuity assumption on μ_0 is crucial.

We then proceed by induction.

As above, uniqueness and the irrelevance of the permutation follow from Loynes' argument. □

A very natural conjecture is then that Theorem 2.17 would give rise to a solution to the *peacock* problem. The set of martingales $(S_t)_{t \in [0, T]}$ (more precisely the set of corresponding martingale measures) carries a natural topology and given $D \subseteq [0, T]$ with $T \in D$ the set of martingales with prescribed marginals $(\mu_t)_{t \in D}$ is compact (cf. [6]). By taking limits of the solutions provided above along appropriate finite discretisations $D \subseteq [0, T]$, one obtains a sequence of optimisers to the discrete problem whose limit $(S_t)_{t \in [0, T]}$ satisfies $S_t \sim \mu_t, t \in [0, T]$ and minimizes $\mathbb{E}[(S_t - S_0)^3]$ simultaneously for all $t \in [0, T]$ among all such martingales. However, since this is not the scope of the present article we leave details for future work.

We note that this also provides a continuous time extension of the martingale monotone coupling rather different from the constructions given by Henry-Labordère et al. [32] and Juillet [42].

3 Stopping times and multi-stopping times

For a Polish space X equipped with a probability measure m we define a new probability space $(\mathcal{X}, \mathcal{G}^0, (\mathcal{G}^0_t)_{t \geq 0}, \mathbb{P})$ with $\mathcal{X} := \mathsf{X} \times C_0(\mathbb{R}_+)$, $\mathcal{G}^0 := \mathcal{B}(\mathsf{X}) \otimes \mathcal{F}^0$, $\mathcal{G}^0_t := \mathcal{B}(\mathsf{X}) \otimes \mathcal{F}^0_t$, $\mathbb{P} := m \otimes \mathbb{W}$, where $\mathcal{B}(\mathsf{X})$ denotes the Borel σ-algebra on X, $\mathbb{W}$ denotes the Wiener measure, and $(\mathcal{F}^0_t)_{t \geq 0}$ the natural filtration. We denote the usual augmentation of $\mathcal{G}^0$ by $\mathcal{G}^a$. Moreover, for $* \in \{0, a\}$ we set $\mathcal{G}^*_{0-} := \mathcal{B}(\mathsf{X}) \otimes \mathcal{F}^*_0$. If we want to stress the dependence on (X, m) we write $\mathcal{G}^a(\mathsf{X}, m), \mathcal{G}^a_t(\mathsf{X}, m), \ldots$.

The natural coordinate process on $\mathcal{X}$ will be denoted by Y, i.e. for $t \geq 0$ we set

$$Y_t(x, \omega) = (x, \omega_t).$$

Note that under $\mathbb{P}$, in the case where $\mathcal{X} = \mathbb{R}$, the process Y can be interpreted as a Brownian motion with starting law m. In particular, $t \mapsto Y_t(x, \omega)$ is continuous and $\mathcal{G}_t^0 = \sigma(Y_s, s \leq t)$.

We recall

$$S := \{(f, s) \; : \; f \in C[0, s], f(0) = 0\}, \qquad S_{\mathsf{X}} := (\mathsf{X}, S)$$

and introduce the maps

$$r : C_0(\mathbb{R}_+) \times \mathbb{R}_+ \to S, \quad (\omega, t) \mapsto (\omega_{\llcorner [0,t]}, t), \tag{3.1}$$

$$r_{\mathsf{X}} : \mathcal{X} \times \mathbb{R}_+ \to S_{\mathsf{X}}, \quad (x, \omega, t) \mapsto (x, r(\omega, t)). \tag{3.2}$$

We equip $C_0(\mathbb{R}_+)$ with the topology of uniform convergence on compacts and S_{X} with the final topology inherited from $\mathcal{X} \times \mathbb{R}_+$ turning it into a Polish space. This structure is very convenient due to the following proposition which is a particular case of [20, Theorem IV. 97].

Proposition 3.1 *Optional sets / functions on $\mathcal{X} \times \mathbb{R}_+$ correspond to Borel measurable sets / functions on S_{X}. More precisely we have:*

(1) *A set $D \subseteq \mathcal{X} \times \mathbb{R}_+$ is $\mathcal{G}^0$-optional iff there is a Borel set $A \subseteq S_{\mathsf{X}}$ with $D = r_{\mathsf{X}}^{-1}(A)$.*
(2) *A process $Z = (Z_t)_{t \in \mathbb{R}_+}$ is $\mathcal{G}^0$-optional iff there is a Borel measurable $H : S_{\mathsf{X}} \to \mathbb{R}$ such that $Z = H \circ r$.*

A $\mathcal{G}^0$-optional set $A \subseteq \mathcal{X} \times \mathbb{R}_+$ is closed in $\mathcal{X} \times \mathbb{R}_+$ iff the corresponding set $r_{\mathsf{X}}(A)$ is closed in S_{X}.

Definition 3.2 A $\mathcal{G}^0$-optional process $Z = H \circ r_{\mathsf{X}}$ is called S_{X}- continuous (resp. l./u.s.c.) iff $H : S_{\mathsf{X}} \to \mathbb{R}$ is continuous (resp. l./u.s.c.).

Remark 3.3 Since the process $t \mapsto Y_t(x, \omega)$ is continuous the predictable and optional σ- algebras coincide ([20, Theorems IV.67 (c) and IV.97]). Hence, every $\mathcal{G}^0$-stopping time τ is predictable and, therefore, foretellable on the set $\{\tau > 0\}$.

Definition 3.4 Let $Z : \mathcal{X} \to \mathbb{R}$ be a measurable function which is bounded or positive. Then we define $\mathbb{E}[Z|\mathcal{G}_t^0]$ to be the unique $\mathcal{G}_t^0$-measurable function satisfying

$$\mathbb{E}[Z|\mathcal{G}_t^0](x, \omega) := Z_t^M(x, \omega) := \int Z(x, \omega_{\llcorner [0,t]} \oplus \omega') \, d\mathbb{W}(\omega').$$

Proposition 3.5 *Let $Z \in C_b(\mathcal{X})$. Then Z_t^M defines an S_{X}- continuous martingale (see Definition 3.2), $Z_\infty^M = \lim_{t \to \infty} Z_t^M$ exists and equals Z.*

Proof Up to a minor change of the probability space this is [2, Proposition 3.5]. □

3.1 Randomised stopping times

We set

$$\mathsf{M} := \{\xi \in \mathcal{P}^{\leq 1}(X \times \mathbb{R}_+) : \xi(d(x, \omega), ds) = \xi_{x,\omega}(ds)\, \mathbb{P}(d(x, \omega)), \xi_{x,\omega} \in \mathcal{P}^{\leq 1}(\mathbb{R}_+)\}$$

and equip it with the weak topology induced by the continuous and bounded functions on $X \times \mathbb{R}_+$. Each $\xi \in \mathsf{M}$ can be uniquely characterized by its cumulative distribution function $A_{x,\omega}^{\xi}(t) := \xi_{x,\omega}([0, t])$.

Definition 3.6 A measure $\xi \in \mathsf{M}$ is called randomized stopping time, written $\xi \in \mathsf{RST}$, iff the associated increasing process A^{ξ} is $\mathcal{G}^0$-optional. If we want to stress the Polish probability space $(X, \mathcal{B}(X), m)$ in the background, we write $\mathsf{RST}(X, m)$.

We remark that randomized stopping times are a subset of the so called **P**-measures introduced by Doleans [21] (for motivation and further remarks see [2, Section 3.2]).

In the sequel we will mostly be interested in a representation of randomized stopping times on an enlarged probability space. We will be interested in $(X', \mathcal{G}', (\mathcal{G}'_t)_{t \geq 0}, \mathbb{P}')$ where $X' := X \times [0, 1]$, $\mathbb{P}'(A_1 \times A_2) = \mathbb{P}(A)\mathcal{L}(A_2)$ ($\mathcal{L}$ denoting Lebesgue measure on $\mathbb{R}$), $\mathcal{G}'$ is the completion of $\mathcal{G}^0 \otimes \mathcal{B}([0, 1])$, and $(\mathcal{G}'_t)_{t \geq 0}$ the usual augmentation of $(\mathcal{G}^0_t \otimes \mathcal{B}([0, 1]))_{t \geq 0}$.

The following characterization of randomized stopping times is essentially Theorem 3.8 of [2]. The only difference is the presence of the X in the starting position, however it is easily checked that this does not affect the proof.

Theorem 3.7 *Let $\xi \in \mathsf{M}$. Then the following are equivalent:*

(1) *There is a Borel function $A : S_X \to [0, 1]$ such that the process $A \circ r_X$ is right-continuous increasing and*

$$\xi_{x,\omega}([0, s]) := A \circ r_X(x, \omega, s) \tag{3.3}$$

defines a disintegration of ξ wrt to $\mathbb{P}$.

(2) *We have $\xi \in \mathsf{RST}(X, m)$.*

(3) *For all $f \in C_b(\mathbb{R}_+)$ supported on some $[0, t]$, $t \geq 0$ and all $g \in C_b(X)$*

$$\int f(s)(g - \mathbb{E}[g|\mathcal{G}^0_t])(x, \omega)\, \xi(dx, d\omega, ds) = 0. \tag{3.4}$$

(4) *On the probability space $(X', \mathcal{G}', (\mathcal{G}'_t)_{t \geq 0}, \mathbb{P}')$, the random time*

$$\rho(x, \omega, u) := \inf\{t \geq 0 : \xi_{x,\omega}([0, t]) \geq u\} \tag{3.5}$$

defines an $\mathcal{G}'$-stopping time.

Remark 3.8 An immediate consequence of (3.4) is the closedness of RST wrt to the weak topology induced by the continuous and bounded functions on $X \times \mathbb{R}_+$ (cf. [2, Corollary 3.10] and Lemma 3.16).

3.2 Randomised multi-stopping times

In this section, we extend the results of the last section to the case of multiple stopping. Recall the notation defined in Sect. 1.2. In particular, for $d \geq 1$, recall that

$$\Xi^d := \{(s_1, \ldots, s_d) \in \mathbb{R}_+^d, s_1 \leq \ldots \leq s_d\}$$

and define M^d to consist of all $\xi \in \mathcal{P}^{\leq 1}(\mathcal{X} \times \Xi^d)$ such that

$$\xi(d(x, \omega), ds_1, \ldots, ds_d) = \xi_{x,\omega}(ds_1, \ldots, ds_d) \, \mathbb{P}(d(x, \omega)), \xi_{x,\omega} \in \mathcal{P}^{\leq 1}(\Xi^d).$$

Recall that $(\bar{\mathcal{X}}, \bar{\mathcal{G}}, (\bar{\mathcal{G}}_t)_{t \geq 0}, \bar{\mathbb{P}})$ is defined by $\bar{\mathcal{X}} = \mathcal{X} \times [0, 1]^d$, $\bar{\mathbb{P}}(A_1 \times A_2) = \mathbb{P}(A_1)\mathcal{L}^d(A_2)$, where $\mathcal{L}^d$ denotes the Lebesgue measure on $\mathbb{R}^d$ and $\bar{\mathcal{G}}_t$ is the usual augmentation of $\mathcal{G}_t^0 \otimes \mathcal{B}([0, 1]^d)$. We mostly denote $\mathcal{L}^d(du)$ by du. For $(u_1, \ldots, u_d) \in [0, 1]^d$ we often just write $(u_1, \ldots, u_d) = u$. We suppress the d- index in the notation for the extended probability space. It will either be clear from the context which d we mean or we explicitly write down the corresponding spaces.

Definition 3.9 A measure $\xi \in \mathsf{M}^d$ is called *randomised multi-stopping time*, denoted by $\xi \in \mathsf{RMST}_d$, if for all $0 \leq i \leq d - 1$

$$\tilde{r}_{i+1,i}(\xi^{(i+1)}) \in \mathsf{RST}(S_\mathsf{X}^{\otimes i}, r_i(\xi^i)). \tag{3.6}$$

We denote the subset of all randomised multi-stopping times with total mass 1 by RMST_d^1. If we want to stress the dependence on (X, m) we write $\mathsf{RMST}_d(\mathsf{X}, m)$ or $\mathsf{RMST}_d^1(\mathsf{X}, m)$.

Remark 3.10 We can understand the condition (3.6) as follows. Consider the case where $d = 2$ and $i = 1$. Then the measure ξ^2 is a sub-probability measure of the form: $\xi^2(d(x, \omega), ds_1, ds_2) = \xi_{x,\omega}(ds_1, ds_2) \, \mathbb{P}(d(x, \omega))$. Then $\tilde{r}_{2,1}(\xi^2)$ is a sub-probability measure on $S_\mathsf{X}^{\otimes i} \times \mathcal{X} \times \Xi^1$. This measure can be disintegrated against $r_1(\xi^1)$, which is a measure on $S_\mathsf{X}^{\otimes i}$, to give a measure on $\mathcal{X} \times \Xi^1$. Intuitively, this measure is the conditional law, given $((B_s)_{s \leq \tau_1}, \tau_1)$ of $((B_t - B_{\tau_1})_{t \geq 0}, \tau_2 - \tau_1)$. The condition (3.6) is then a statement that the law of this pair is then consistent with the law of a randomised stopping time.

Unlike for the randomised stopping times, there is no obvious analogue of (1), (2) or (3) of Theorem 3.7 in the multi-stopping time setting. However below we prove a representation result for randomised multi-stopping times in a similar manner to (4). The following lemma (c.f. [2, Lemma 3.11]) then enables us to conclude that, on an arbitrary probability space, all sequences of increasing stopping times can be represented as a randomised multi-stopping time on our canonical probability space.

Lemma 3.11 *Let B be a Brownian motion on some stochastic basis $(\Omega, \mathcal{H}, (\mathcal{H}_t)_{t \geq 0}, \mathbb{Q})$ with right continuous filtration. Let $\tau_1, \ldots, \tau_n$ be an increasing sequence of $\mathcal{H}$-stopping times and consider*

$$\Phi : \Omega \to C(\mathbb{R}_+) \times \Xi^d, \quad \bar{\omega} \mapsto ((B_t)_{t \geq 0}, \tau_1(\bar{\omega}), \ldots, \tau_n(\bar{\omega})).$$

Then $\xi := \Phi(\mathbb{Q})$ is a randomized multi-stopping time and for any measurable γ : $S^{\otimes n} \to \mathbb{R}$ we have

$$\int \gamma(f, s_1, \ldots, s_n) \, r_n(\xi)(d(f, s_1, \ldots, s_n)) = \mathbb{E}_{\mathbb{Q}}[\gamma((B_t)_{t \leq \tau_n}, \tau_1, \ldots, \tau_n)]. \quad (3.7)$$

If Ω is sufficiently rich that it supports a uniformly distributed random variable which is $\mathcal{H}_0$-measurable, then for $\xi \in \mathsf{RMST}$ we can find an increasing family $(\tau_i)_{1 \leq i \leq n}$ of $\mathcal{H}$-stopping times such that $\xi = \Phi(\mathbb{Q})$ and (3.7) holds.

Proof For notational convenience we show the first part for the case $n = 2$. Let $B_0 \sim m$. It then follows by [2, Lemma 3.11] that $\tilde{r}_{1,0}(\xi) \in \mathsf{RST}(\mathbb{R}, m)$. Hence, we need to show that $\tilde{r}_{2,1}(\xi) \in \mathsf{RST}(S_{\mathbb{R}}, r_1(\xi))$, i.e. we have to show that $\xi^2_{(f,s)}$ is $r_1(\xi)$–a.s. a randomized stopping time, where $(\xi^2_{(f,s)})_{(f,s)}$ denotes a disintegration of ξ^2 wrt $r_1(\xi)$. (Here and in the rest of the proof we assume $f(0) \in \mathbb{R}$ and suppress the "x" from the notation).

First we show that $\tilde{r}_1(\xi^1)(d(f, s), d\omega) = r_1(\xi^1)(d(f, s))\mathbb{W}(d\omega)$. Take a measurable and bounded $F : S_{\mathbb{R}} \times C_0(\mathbb{R}_+) \to \mathbb{R}$. Then, using the strong Markov property in the last step, we have

$$\int F((f, s), \omega) \, \tilde{r}_1(\xi^1)(d(f, s), d\omega)$$

$$= \int F((r_{\mathsf{X}}(\tilde{\omega}, s), \theta_s \tilde{\omega}) \, \xi^1(d\tilde{\omega}, ds)$$

$$= \int_{\Omega} F(r_{\mathsf{X}}(B(\omega), \tau_1(\omega)), \theta_{\tau_1(\omega)} B(\omega)) \, \mathbb{Q}(d\omega)$$

$$= \int F((f, s), \tilde{\omega}) \, r_1(\xi^1)(d(f, s))\mathbb{W}(d\tilde{\omega}) . \quad (3.8)$$

Let q be the projection from $S_{\mathbb{R}} \times C_0(\mathbb{R}_+) \times \mathbb{R}_+$ to $S_{\mathbb{R}} \times C_0(\mathbb{R}_+)$, and p be the projection from $X \times \Xi^2 \to X \times \mathbb{R}_+$, $p(\omega, s_1, s_2) = (\omega, s_1)$. Then, $q \circ \tilde{r}_{2,1} = \tilde{r}_1 \circ p$. Recalling that $\xi^1 = p(\xi^2)$ there is a disintegration of $\tilde{r}_{2,1}(\xi^{1,2})$ wrt $\tilde{r}_1(\xi^1)$ which we denote by

$$\xi^2_{(f,s_1),\omega}(ds_2) \in \mathcal{P}^{\leq 1}(\mathbb{R}_+).$$

Then, we set $\xi^2_{(f,s_1)}(d\omega, ds_2) := \xi^2_{(f,s_1),\omega}(ds_2)\mathbb{W}(d\omega)$. Since $d\tilde{r}_1(\xi^1) = dr_i(\xi^1)d\mathbb{W}$ the measures $\xi^2_{(f,s_1)}$ define a disintegration of $\tilde{r}_{2,1}(\xi^2)$ wrt $r_1(\xi^1)$. We have to show that $r_1(\xi^1)$ a.s. $\xi^2_{(f,s_1)}$ is a randomized stopping time. We will show property (2) in Theorem 3.7, where now $\mathsf{X} = S_{\mathbb{R}}$, $m = r_1(\xi)$ and accordingly $\mathcal{G}^0_t = \mathcal{B}(S_{\mathbb{R}}) \otimes \mathcal{F}^0_t$ with usual augmentation $\mathcal{G}^a_t$ (cf. Sect. 1.2).

To this end, fix $t \geq 0$ and let $g : S_{\mathbb{R}} \times C_0(\mathbb{R}_+) \to \mathbb{R}$ be bounded and measurable and set $h = \mathbb{E}_m[g|\mathcal{G}^a_t]$. Then, it holds that $\mathbb{E}_{\mathbb{Q}}[g(r_1(B, \tau_1), \theta_{\tau_1} B)|\mathcal{H}_{\tau_1+t}] = h(r_1(B, \tau_1), \theta_{\tau_1} B)$. Using rightcontinuity of the filtration $\mathcal{H}$ in the third step to conclude that $\tau_2 - \tau_1$ is an $(\mathcal{H}_{\tau_1+t})_{t \geq 0}$ stopping time, this implies

$$\int g((f,s),\omega)\,\xi^2_{(f,s),\omega}([0,t])\,r_1(\xi^1)(d(f,s))\mathbb{W}(d\omega)$$
$$= \mathbb{E}_{\mathbb{Q}}\left[g(r_1(B,\tau_1),\theta_{\tau_1}B)\mathbb{1}_{\tau_2-\tau_1\le t}\right]$$
$$= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}\left[g(r_1(B,\tau_1),\theta_{\tau_1}B)|\mathcal{H}_{\tau_1+t}\right]\mathbb{1}_{\tau_2-\tau_1\le t}\right]$$
$$= \mathbb{E}_{\mathbb{Q}}\left[h(r_1(B,\tau_1),\theta_{\tau_1}B)\mathbb{1}_{\tau_2-\tau_1\le t}\right]$$
$$= \int h((f,s),\omega)\,\xi^2_{(f,s),\omega}([0,t])\,r_1(\xi^1)(d(f,s))\mathbb{W}(d\omega).$$

This shows the first part of the lemma.

To show the second part of the lemma we start by constructing an increasing sequence of stopping times on the extended canonical probability space $(\bar{X},\bar{\mathcal{G}},(\bar{\mathcal{G}}_t)_{t\ge 0},\bar{\mathbb{P}})$. By Theorem 3.7 and the assumption that $\xi^1 \in \mathsf{RST}(X,m)$ there is a $\bar{\mathcal{G}}$ stopping time $\rho^1(x,\omega,u)=\rho^1(x,\omega,u_1)$ defining a disintegration of ξ^1 wrt $\bar{\mathbb{P}}$ via

$$\int \delta_{\rho^1}(ds_1)\,du\,.$$

By assumption, $\tilde{r}_{2,1}(\xi^2) \in \mathsf{RST}(S_X,r_1(\xi^1))$. Hence, writing $s_2' = s_2 - s_1$ we can disintegrate

$$\xi^2(d(x,\omega),ds_1,ds_2) = \int \xi^2_{r_X(x,\omega,\rho^1(x,\omega,u_1))}(\theta_{\rho^1(x,\omega,u_1)}\omega,ds_2')\delta_{\rho^1(x,\omega,u_1)}(ds_1)du_1$$

such that for $r_1(\xi^1)$ a.e. (x,f,s_1) the disintegration $\xi^2_{(x,f,s_1)}$ is a randomized stopping time. Again by Theorem 3.7 there is a stopping time $\tilde{\rho}^2_{x,f,s_1}(\tilde{\omega},u_2)$ representing $\xi^2_{(x,f,s_1)}$ as in (3.5). Then,

$$\rho^2(x,\omega,u_1,u_2) := \rho^1(x,\omega,u_1) + \tilde{\rho}^2_{r_X(x,\omega,\rho_1(x,\omega,u_1))}(\theta_{\rho_1(x,\omega,u_1)}\omega,u_2)$$

defines a $\bar{\mathcal{G}}$ stopping time such that

$$(x,\omega)\mapsto \int_{[0,1]^d}\delta_{\rho^1(x,\omega,u)}(dt_1)\delta_{\rho^2(x,\omega,u)}(dt_2)\,du$$

defines a $\mathcal{G}^a$- measurable disintegration of ξ^2 w.r.t. $\bar{\mathbb{P}}$. We proceed inductively. To finish the proof, let U be the $[0,1]^d$–valued uniform $\mathcal{H}_0$–measurable random variable. Then $\tau_i := \rho^i(B,U)$ define the required increasing family of $\mathcal{H}$ stopping times. $\qquad\square$

Remark 3.12 Lemma 3.11 shows that optimizing over an increasing family of stopping times on a rich enough probability space in (OptMSEP) is equivalent to optimizing over randomized multi-stopping times on the Wiener space.

Corollary 3.13 Let $\xi \in \mathsf{RMST}$. On the extended canonical probability space $(\bar{X},\bar{\mathcal{G}},(\bar{\mathcal{G}}_t)_{t\ge 0},\bar{\mathbb{P}})$ there exists an increasing sequence $(\rho^i)_{i=1}^d$ of $\bar{\mathcal{G}}$- stopping times such that

(1) for $u = (u_1, \ldots, u_d) \in [0, 1]^d$ and for each $1 \leq i \leq d$ we have $\rho^i(x, \omega, u) = \rho^i(x, \omega, u_1, \ldots, u_i)$;

(2)

$$(x, \omega) \mapsto \int_{[0,1]^d} \delta_{\rho^1(x,\omega,u)}(dt_1) \cdots \delta_{\rho^d(x,\omega,u)}(dt_d) \, du \tag{3.9}$$

defines a $\mathcal{G}^a$- measurable disintegration of ξ w.r.t. $\mathbb{P}$.

Next we introduce some notation to state another straightforward corollary. It is easy to see that $q^{d,i} \circ \tilde{r}_{d,i} = \tilde{r}_i \circ p^{d,i}$, where $q^{d,i}$ is the projection from $S_X^{\otimes i} \times C(\mathbb{R}_+) \times \Xi^{d-i}$ to $S_X^{\otimes i} \times C(\mathbb{R}_+)$, and $p^{d,i}$ is the projection from $X \times \Xi^d$ to $X \times \Xi^i$ defined by

$$(x, \omega, s_1, \ldots, s_d) \mapsto (x, \omega, s_1, \ldots, s_i).$$

Recalling that $\xi^i = p^{d,i}(\xi)$, it follows that there exists a disintegration of $\tilde{r}_{d,i}(\xi)$ with respect to $\tilde{r}_i(\xi^i)$, which we denote by:

$$\xi_{(x,f,s_1,\ldots,s_i),\omega}(ds_{i+1}, \ldots, ds_d) \in \mathcal{P}(\Xi^{d-i}).$$

Moreover, we set

$$\xi_{(x,f,s_1,\ldots,s_i)}(d\omega, ds_{i+1}, \ldots, ds_d)$$
$$:= \xi_{(x,f,s_1,\ldots,s_i),\omega}(ds_{i+1}, \ldots, ds_d) \, \mathbb{W}(d\omega) \in \mathcal{P}(C(\mathbb{R}_+) \times \Xi^{d-i}).$$

The map $(x, f, s_1, \ldots, s_i) \mapsto \xi_{(x,f,s_1,\ldots,s_i)}$ inherits measurability from the joint measurability of $((x, f, s_1, \ldots, s_i), \omega) \mapsto \xi_{(x,f,s_1,\ldots,s_i),\omega}$. In particular, $\xi_{(x,f,s_1,\ldots,s_i)}$ defines a disintegration of $\tilde{r}_{d,i}(\xi)$ w.r.t. $r_i(\xi^i)$, since $d\tilde{r}_i(\xi^i) = d\mathbb{W}dr_i(\xi^i)$ by the same calculation as (3.8). Following exactly the line of reasoning as in the first part of the proof of Lemma 3.11 yields

Corollary 3.14 *Let $\xi \in RMST_d(X, m)$ and $1 \leq i < d$. Then, for $r_i(\xi^i)$ a.e. $(x, f, s_1, \ldots, s_i)$ we have $\xi_{(x,f,s_1,\ldots,s_i)} \in RMST_{d-i}(\{0\}, \delta_0)$.*

Remark 3.15 We note that the last Corollary still holds for $i = 0$ by setting $S_{\mathbb{R}}^{\otimes 0} = \mathbb{R}$, $r_0(\xi^j) = m$. Then, the result says that for a disintegration $(\xi_x)_x$ of ξ w.r.t. m for m-a.e. $x \in X$ we have $\xi_x \in RMST_d$. Of course this can also trivially be seen as a consequence of $\mathbb{P} = m \otimes \mathbb{W}$.

An important property of RMST is the following Lemma.

Lemma 3.16 *RMST is closed w.r.t. the weak topology induced by the continuous and bounded functions on $X \times \Xi^d$.*

Proof We fix $0 \leq i \leq d-1$ and consider the Polish space $\tilde{X} = S_X^{\otimes i}$ with corresponding $\tilde{\mathcal{X}} = \tilde{X} \times C_0(\mathbb{R}_+)$ and $\mathbb{P} = r_i(\xi^i) \otimes \mathbb{W}$. To show the defining property (3.6) in Definition 3.9 we consider condition (2) in Theorem 3.7; the goal is to express measurability of

$Z_t(x, \omega) := \xi_{x,\omega}^{i+1}(f)$, $f \in C_b([0, t])$, $x \in S_X^{\otimes i}$, $\omega \in C_0(\mathbb{R}_+)$ in a different fashion. Note that a bounded Borel function h is $\mathcal{G}_t^0$-measurable iff for all bounded Borel functions $G : \tilde{X} \to \mathbb{R}$

$$\mathbb{E}[hG] = \mathbb{E}[h\mathbb{E}[G|\mathcal{G}_t^0]],$$

of course this does not rely on our particular setup. By a functional monotone class argument, for $\mathcal{G}_t^0$-measurability of Z_t it is sufficient to check that

$$\mathbb{E}[Z_t(G - \mathbb{E}[G|\mathcal{G}_t^0])] = 0 \qquad (3.10)$$

for all $G \in C_b(\tilde{X})$. In terms of ξ^{i+1}, (3.10) amounts to

$$0 = \mathbb{E}[Z_t(G - \mathbb{E}[G|\mathcal{G}_t^0])] = \int \mathbb{P}(dx, d\omega) \int \xi_{x,\omega}^{i+1}(ds) f(s)(G - \mathbb{E}[G|\mathcal{G}_t^0])(x, \omega)$$

$$= \int f(s)(G - \mathbb{E}[G|\mathcal{G}_t^0])(x, \omega) \, \tilde{r}_{i+1,i}(\xi^{i+1})(dx, d\omega, ds),$$

which is a closed condition by Proposition 3.5. $\qquad\qquad\square$

Given $\xi \in \mathsf{M}^d$ and $s \geq 0$ we define the random measure $\xi \wedge s$ on Ξ^d by setting for $A \subseteq \Xi^d$ and each $(x, \omega) \in X$

$$(\xi \wedge s)_{x,\omega}(A) = \int \mathbb{1}_A(s_1 \wedge s, \ldots, s_d \wedge s) \, \xi_{x,\omega}(ds_1, \ldots, ds_d).$$

Assume that $(M_s)_{s\geq 0}$ is a process on X. Then $(M_s^\xi)_{s\geq 0}$ is defined to be the *probability measure on* $\mathbb{R}^{d+1}$ such that for all bounded and measurable functions $F : \mathbb{R}^{d+1} \to \mathbb{R}$

$$\int_{\mathbb{R}^{d+1}} F(y) \, M_s^\xi(dy)$$
$$= \int_{X \times \Xi^d} F(M_0(x, \omega), M_{s_1}(x, \omega), \ldots, M_{s_d}(x, \omega)) \, (\xi \wedge s)(dx, d\omega, ds_1, \ldots, ds_d).$$

This means that M_s^ξ is the image measure of $\xi \wedge s$ under the map $M : X \times \Xi^d \to \mathbb{R}^{d+1}$ defined by

$$(x, \omega, s_1, \ldots, s_d) \mapsto (M_0(x, \omega), M_{s_1}(x, \omega), \ldots, M_{s_d}(x, \omega)).$$

We write $\lim_{s\to\infty} M_s^\xi = M_\xi$ if it exists.

3.3 The set RMST($\mu_0, \mu_1, \ldots, \mu_n$), compactness and existence of optimisers

In this subsection, we specialise our setup to $X = \mathbb{R}$, $m = \mu_0 \in \mathcal{P}(\mathbb{R})$ and $d = n$. Let $\mu_0, \mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$ be centered, in convex order and with finite second moment[2] $\int x^2 \mu_i(dx) = V_i < \infty$ for all $i \leq n$. In particular $V_i \leq V_{i+1}$. For $t \geq 0$ we set $B_t(x, \omega) = x + \omega_t$. We extend B to the extended probability space $\bar{X}$ by setting $\bar{B}(x, \omega, u) = B(x, \omega)$. By considering the martingale $\bar{B}_t^2 - t$ we immediately get (see the proof of Lemma 3.12 in [2] for more details)

Lemma 3.17 *Let* $\xi \in RMST_n$ *and assume that* $B_\xi = (\mu_0, \mu_1, \ldots, \mu_n)$. *Let* $(\rho_1, \ldots, \rho_n)$ *be any representation of* ξ *granted by Lemma* 3.11. *Then, the following are equivalent*

(1) $\bar{\mathbb{E}}[\rho^i] < \infty$ *for all* $1 \leq i \leq n$
(2) $\bar{\mathbb{E}}[\rho^i] = V_i - V_0$ *for all* $1 \leq i \leq n$
(3) $(\bar{B}_{\rho^i \wedge t})_{t \geq 0}$ *is uniformly integrable for all* $1 \leq i \leq n$.

Of course it is sufficient to test any of the above quantities for $i = n$.

Definition 3.18 We denote by RMST($\mu_0, \mu_1, \ldots, \mu_n$) the set of all randomised multi-stopping times satisfying one of the conditions in Lemma 3.17.

By pasting solutions to the one marginal Skorokhod embedding problem one can see that the set RMST($\mu_0, \mu_1, \ldots, \mu_n$) is non-empty. However, the most important property is

Proposition 3.19 *The set* RMST($\mu_0, \mu_1, \ldots, \mu_n$) *is compact wrt to the topology induced by the continuous and bounded functions on* $C(\mathbb{R}_+) \times \Xi^d$.

Proof This is a direct consequence of the compactness of RST(μ_n) established in [2, Theorem 3.14] as the set RMST($\mu_0, \mu_1, \ldots, \mu_n$) is closed. □

This result allows us to deduce one of the critical results for our optimisation problem:

Proof of Theorem 2.1 In the case where γ_1, γ_2 are bounded below, this follows from Proposition 3.19 and the Portmanteau theorem. In the case where (2.1) holds, the argument follows in an identical manner to the proof of Theorem 4.1 in [2]. □

3.4 Joinings of stopping times

We now introduce the notion of a joining; these will be used later to define new stopping times which are candidate competitors for our optimisation problem.

Definition 3.20 Let $(\mathcal{Y}, \sigma)$ be a Polish probability space. The set JOIN(m, σ) of joinings between $\mathbb{P} = m \otimes \mathbb{W}$ and σ is defined to consist of all subprobability measures $\pi \in \mathcal{P}^{\leq 1}(X \times \mathbb{R}_+ \times \mathcal{Y})$ such that

---

[2] It is possible to relax this, see [2, Section 8].

- $\text{proj}_{\mathcal{X} \times \mathbb{R}_+} (\pi_{\llcorner \mathcal{X} \times \mathbb{R}_+ \times B}) \in \mathsf{RST}(\mathsf{X}, m)$ for all $B \in \mathcal{B}(\mathcal{Y})$;
- $\text{proj}_{\mathcal{X}}(\pi) = \mathbb{P}$
- $\text{proj}_{\mathcal{Y}}(\pi) \leq \sigma$.

Example 3.21 An important example in the sequel will be the probability space $(\mathcal{X}, \mathbb{P})$ constructed from $\mathsf{X} = S_{\mathbb{R}}^{\otimes i}$ and $m = r_i(\xi^i)$ for $\xi \in \mathsf{RMST}_n^1(\mathbb{R}, \mu_0)$ and $0 \leq i < n$, where we set $S^{\otimes 0} = \mathbb{R}$, $r_0(\xi^0) = \mu_0$ leading to $\mathcal{X} = S_{\mathbb{R}}^{\otimes i} \times C(\mathbb{R}_+)$ and $\mathbb{P} = r_i(\xi^i)\mathbb{W} = \tilde{r}_i(\xi^i)$ (cf. Corollary 3.14).

4 Colour swaps, multi-colour swaps and stop-go pairs

In this section, we will define the general notion of stop-go pairs which was already introduced in a weaker form in Sect. 2.1. We will do so in two steps. First we define colour swap pairs and then we combine several colour swaps to get multi-colour swaps. Together, they build the stop-go pairs.

Our basic intuition for the different swapping rules comes from the following picture. We imagine that each of the measures $\mu_1, \ldots, \mu_n$ carries a certain colour, i.e. the measure μ_i carries colour i. The Brownian motion will be thought of being represented by a particle of a certain colour: at time zero the Brownian particle has colour 1 and when it is stopped for the i-th time it changes its colour from i to $i + 1$ (cf. Fig. 1 in Sect. 2.1).

In identifying a stop-go pair, we want to consider two sub-paths, $(f, s_1, \ldots, s_i)$ and $(g, t_1, \ldots, t_i)$, and imagine the future stopping rules, which will now be a sequence of colour changes, obtained by concatenating a path ω onto the two paths. The simplest way of creating a new stopping rule is simply to exchange the coloured tails. This will preserve the marginal law of the stopped process, while generating a new multi-stopping time. A generalisation of this rule would be to try and swap back to the original colour rule at the jth colour change, where $i < j$. In this case, one would swap the colours until the first time one of the paths would stop for the jth time, after which one attempts to revert to the previous stopping rule. Note however that this may not be possible: if the other path has not yet reached the $j - 1$st colour change, then the rules cannot be stopped, since one would have to switch from the jth colour to the $j - 1$st colour, which is not allowed. Instead, in such a case, we simply keep the swapped colourings. We call recolouring rules of this nature *colour swaps* (or $i \leftrightarrow j$ colour swaps). We will define such colour swap pairs in Sect. 4.2.

After consideration of these colour swaps, it is clear that the determination of when to revert to the original stopping rule could be determined in a more sophisticated manner. For example, instead of trying to revert only on the jth colour change, one could instead try to revert on every colour change, and revert the first time it is possible to revert. This recolouring rule gives us a second set of possible path swaps, and we call such pairs *multi-colour swaps*. We will define these recolouring rules in Sect. 4.3. Of course, a multitude of other rules can easily be created. For our purposes, colour swaps and multi-colour swaps will be sufficient, but other generalisations could easily be considered, and may be important for showing optimality in cases outside those considered in the current paper. We leave this as an avenue for future research.

An important aspect of the recolouring rules are that they provide a recipe to map from one stopping rule to another, and an important aspect that needs to be verified is that the new stopping rule does indeed define a randomised multi-stopping time.

We fix $\xi \in \mathsf{RMST}_n^1(\mathbb{R}, \mu_0)$ and $\gamma : S_\mathbb{R}^{\otimes n} \to \mathbb{R}$. As in the previous section, we denote $\xi^i = \xi^{(1,\ldots,i)} = \mathrm{proj}_{X \times (1,\ldots,i)}(\xi)$. For $(x, f, s_1, \ldots, s_i) \in S_\mathbb{R}^{\otimes i}$ we write $(f, s_1, \ldots, s_i)$ and agree on $f(0) = x \in \mathbb{R}$. For $(f, s_1, \ldots, s_i) \in S_\mathbb{R}^{\otimes i}$ and $(h, s) \in S$ we will often write $(f, s_1, \ldots, s_i)|(h, s)$ instead of $(f, s_1, \ldots, s_i) \otimes (h, s) \in S_\mathbb{R}^{\otimes i+1}$ to stress the probabilistic interpretation of conditioning the continuation of $(f, s_1, \ldots, s_i)$ on (h, s).

4.1 Coloured particles and conditional randomised multi-stopping times.

By Corollary 3.14 and Remark 3.15 (for $i = 0$), for each $0 \le i \le n$ the measure $\xi_{(f,s_1,\ldots,s_i)}$ is $r_i(\xi^i)$−a.s. a randomised multi-stopping time. For each $0 \le i \le n - 1$ we fix a disintegration $(\xi_{(f,s_1,\ldots,s_i),\omega})_{(f,s_1,\ldots,s_i),\omega}$ of $\tilde{r}_{n,i}(\xi)$ w.r.t. $\tilde{r}_i(\xi^i)$ and set $\xi_{(f,s_1,\ldots,s_i)} = \xi_{(f,s_1,\ldots,s_i),\omega} \mathbb{W}(d\omega)$. We will need to consider randomised multi-stopping times conditioned on not yet having stopped the particle of colour $i + 1$. To this end, observe that

$$\xi^{i+1}_{(f,s_1,\ldots,s_i)} := \mathrm{proj}_{C(\mathbb{R}_+) \times \Xi^1}(\xi_{(f,s_1,\ldots,s_i)})$$

defines a disintegration of $\tilde{r}_{i+1,i}(\xi^{i+1})$ wrt $r_i(\xi^i)$. By Definition 3.9, $\xi^{i+1}_{(f,s_1,\ldots,s_i)} \in \mathsf{RST}$ a.s. and we set

$$A^\xi_{(f,s_1,\ldots,s_i)}(\omega, t) := A^\xi_{(f,s_1,\ldots,s_i)}(\omega_{\llcorner[0,t]}, t) := (\xi^{i+1}_{(f,s_1,\ldots,s_i)})_\omega([0, t])$$

which is well defined for $r_i(\xi^i)$-almost every $(f, s_1, \ldots, s_i)$ by Theorem 3.7.

For $(f, s_1, \ldots, s_i) \in S_\mathbb{R}^{\otimes i}$ we define the conditional randomised multi-stopping time given $(h, s) \in S$ to be the (sub) probability measure $\xi_{(f,s_1,\ldots,s_i)|(h,s)}$ on $C(\mathbb{R}_+) \times \Xi^{n-i}$ given by

$$(\xi_{(f,s_1,\ldots,s_i)|(h,s)})_\omega([0, T_{i+1}] \times \ldots \times [0, T_n])$$
$$= \begin{cases} (\xi_{(f,s_1,\ldots,s_i)})_{h \oplus \omega} \\ \quad ((s, s + T_{i+1}] \times \ldots \times (s, s + T_n]) & \text{if } A^\xi_{(f,s_1,\ldots,s_i)}(h, s) < 1 \\ \Delta A^\xi_{(f,s_1,\ldots,s_i)}(h, s)(\xi_{(f \oplus h, s_1,\ldots,s_i, s_i+s)})_\omega \\ \quad ([s, s + T_{i+2}] \times \ldots \times [s, s + T_n]) & \text{if } A^\xi_{(f,s_1,\ldots,s_i)}(h, s) = 1, \end{cases} \qquad (4.1)$$

where $\Delta A^\xi_{(f,s_1,\ldots,s_i)}(h, s) = A^\xi_{(f,s_1,\ldots,s_i)}(h, s) - A^\xi_{(f,s_1,\ldots,s_i)}(h, s-)$. The second case in (4.1) corresponds to $(\xi^{i+1}_{(f,s_1,\ldots,s_i)})_{h \oplus \omega}$ having an atom at time s which consumes all the remaining (positive) mass, which is of course independent of ω. Note that this is non-zero only if $A^\xi_{(f,s_1,\ldots,s_i)}(h, s-) < 1$, that is, there is still mass remaining immediately before time s. This causes a δ_0 to appear in (4.2) below. Moreover, in

this case it is possible that also all particles of colour $j \in \{i+2, \ldots, n\}$ are stopped at time s by $(\xi_{(f,s_1,\ldots,s_i)}^{i+1})_{h \oplus \omega}$. This is the reason for the closed intervals in the second line on the right hand side of (4.1). Using Lemma 3.11 resp. Corollary 3.13 it is not hard to see that (4.1) indeed defines a randomized multi-stopping time (you simply have to consider the stopping times $\rho^l(\omega, u_1, \ldots, u_l)$ representing $\xi_{(f,s_1,\ldots,s_i)}$ with $u_1 > A_{(f,s_1,\ldots,s_i)}^{\xi}$ for the first case and the second case is immediate).

Accordingly, we define the normalised conditional randomised multi-stopping times, by

$$
\bar{\xi}_{(f,s_1,\ldots,s_i)|(h,s)}
$$
$$
:= \begin{cases} \frac{1}{1 - A_{(f,s_1,\ldots,s_i)}^{\xi}(h,s)} \xi_{(f,s_1,\ldots,s_i)|(h,s)} & \text{if } A_{(f,s_1,\ldots,s_i)}^{\xi}(h,s) < 1, \\ \delta_0 \xi_{(f \oplus h, s_1,\ldots,s_i, s_i+s)} & \text{if } A_{(f,s_1,\ldots,s_i)}^{\xi}(h,s-) < 1 \text{ and } A_{(f,s_1,\ldots,s_i)}^{\xi}(h,s) = 1, \\ \delta_0 \cdots \delta_0 & \text{else.} \end{cases}
$$
(4.2)

We emphasize that the construction of $\bar{\xi}_{(f,s_1,\ldots,s_i)|(h,s)}$ and $\xi_{(f,s_1,\ldots,s_i)|(h,s)}$ only relies on the fixed disintegration of $\tilde{r}_{n,i}(\xi)$ w.r.t. $\tilde{r}_i(\xi)$. In particular, the map

$$
((f, s_1, \ldots, s_i), (h, s)) \mapsto \bar{\xi}_{(f,s_1,\ldots,s_i)|(h,s)}
$$
(4.3)

is measurable.

Recall the connection of Borel sets of S_X and optional sets in $\mathcal{X} \times \mathbb{R}_+$ given by Proposition 3.1.

Definition 4.1 Let (X, m) be a Polish probability space. A set $F \subseteq S_X$ is called m-evanescent iff $r_X^{-1}(F) \subseteq \mathcal{X} \times \mathbb{R}_+$ is evanescent (wrt the probability space $(\mathcal{X}, \mathbb{P})$) iff there exists $A \subseteq \mathcal{X}$ such that $\mathbb{P}(A) = (m \otimes \mathbb{W})(A) = 1$ and $r_X(A \times \mathbb{R}_+) \cap F = \emptyset$.

By Corollary 3.14, $\xi_{(f,s_1,\ldots,s_i)} \in \mathsf{RMST}_{n-i}$ for $r_i(\xi^i)$ a.e. $(f, s_1, \ldots, s_i) \in S^{\otimes i}$. The next lemma says that for typical $(f, s_1, \ldots, s_i)|(h,s) \in S^{\otimes i+1}$ this still holds for $\bar{\xi}_{(f,s_1,\ldots,s_i)|(h,s)}$.

Lemma 4.2 Let $\xi \in \mathsf{RMST}_n^1$ and fix $0 \le i < n$.

(1) $\bar{\xi}_{(f,s_1,\ldots,s_i)|(h,s)} \in \mathsf{RMST}_{n-i}^1$ outside a $r_i(\xi^i)$-evanescent set.
(2) If $F : S^{\otimes n} \to \mathbb{R}$ satisfies $\xi(F \circ r_n) < \infty$, then the set $\{(f, s_1, \ldots, s_i)|(h,s) : \bar{\xi}_{(f,s_1,\ldots,s_i)|(h,s)}(F^{(f,s_1,\ldots,s_i)|(h,s)\oplus} \circ r_{n-i}) = \infty\}$ is $r_i(\xi^i)$-evanescent. In particular, this applies to $F(f, s_1, \ldots, s_n) = s_n$ if $\xi \in \mathsf{RMST}(\mu_0, \ldots, \mu_n)$.

Remark 4.3 Observe that a direct consequence of Corollary 3.14, assuming $\xi(F \circ r_n) < \infty$, is that $\{(f, s_1, \ldots, s_i) : \xi_{(f,s_1,\ldots,s_i)}(F^{(f,s_1,\ldots,s_i)\otimes} \circ r_{n-i}) = \infty\}$ is a $r_i(\xi^i)$ null set.

Proof of Lemma 4.2 It is apparent that $\xi_{(f,s_1\ldots,s_i)|(h,s)} \in \mathsf{RMST}$. By Corollary 3.14, (4.2), and Remark 4.3 it is sufficient to show the claims under the additional hypothesis that $A_{(f,s_1\ldots,s_i)}^{\xi}(h,s) < 1$. Hence, consider

$$U_1 = \left\{ (f, s_1, \ldots, s_i)|(h, s) : A^{\xi}_{(f, s_1 \ldots, s_i)}(h, s) < 1, \bar{\xi}_{(f, s_1, \ldots, s_i)|(h,s)} \notin \mathsf{RMST}^1_{n-i} \right\},$$

$$U_2 = \left\{ (f, s_1, \ldots, s_i)|(h, s) : A^{\xi}_{(f, s_1 \ldots, s_i)}(h, s) \right.$$
$$\left. < 1, \bar{\xi}_{(f, s_1, \ldots, s_i)|(h,s)}(F^{(f, s_1, \ldots, s_i)|(h,s)\oplus} \circ r_{n-i}) = \infty \right\}.$$

Set $A^{\xi}_{(f, s_1, \ldots, s_i)}(\omega) := \lim_{s \to \infty} A^{\xi}_{(f, s_1, \ldots, s_i)}(r(\omega, s))$. Then, $(f, s_1, \ldots, s_i)|(h, s) \in U_1$ is equivalent to $\int A^{\xi}_{(f, s_1, \ldots, s_i)}(h \oplus \omega) \, \mathbb{W}(d\omega) < 1$. Set $X = S^{\otimes i}$ and $m = r_i(\xi^i)$ and recall that the natural coordinate process on X is denoted by Y. Given a $\mathcal{G}^0$-stopping time τ on $(X, \mathcal{G}, \mathbb{P})$ we have $r_i(\xi^i)$ a.s. by the strong Markov property and the fact that ξ is almost surely a finite stopping time:

$$1 = \int A^{\xi}_{(f, s_1, \ldots, s_i)}(\omega) \, \mathbb{W}(d\omega)$$
$$= \int \left[\mathbb{1}_{\tau(\omega)=\infty} A^{\xi}_{(f, s_1, \ldots, s_i)}(\omega) + \int \mathbb{1}_{\tau(\omega)<\infty} A^{\xi}_{(f, s_1, \ldots, s_i)}(\omega_{\llcorner[0,\tau]} \oplus \tilde{\omega}) \, \mathbb{W}(d\tilde{\omega}) \right] \mathbb{W}(d\omega),$$

implying that $\mathbb{P}[((Y_s)_{s \leq \tau}, \tau) \in U_1] = 0$. Hence, the first part follows from the optional section Theorem.

Additionally, setting $\alpha(d(x, \omega), dt) = \delta_{\tau(x,\omega)}(dt) \, \mathbb{P}(d(x, \omega))$ we have

$$\int_{U_2} dr_X(\alpha)(x, h, s) \, (1 - A^{\xi}_x(h, s)) \, \bar{\xi}_{x|(h,s)}(F^{x|(h,s)\oplus}) \leq \xi(F) < \infty,$$

implying $r_X(\alpha)(U_2) = 0$. Hence, we have $\mathbb{P}[((Y_s)_{s \leq \tau}, \tau) \in U_2] = 0$ proving the claim by the optional section theorem, e.g. [20, Theorems IV 84 and IV 85] (see also Remark 5.5). □

4.2 Colour swaps

As a first step towards the definition of stop-go pairs we introduce an important building block, the colour swap pairs.

By Corollary 3.13 and Corollary 3.14, for $r_i(\xi^i)$ a.e. $(g, t_1, \ldots, t_i)$ there is an increasing sequence $(\rho^j_{(g, t_1, \ldots, t_i)})^n_{j=i+1}$ of $\bar{\mathcal{F}}^a$-stopping times such that

$$\omega \mapsto \int_{[0,1]^{n-i}} \delta_{\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{i+1}) \cdots \delta_{\rho^n_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_n) \, du$$

defines an $\mathcal{F}^a$-measurable disintegration of $\xi_{(g,t_1,\ldots,t_i)}$ w.r.t. $\mathbb{W}_{\mu_0}$. Similarly, by Lemma 4.2, outside an $r_{i-1}(\xi^{i-1})$ evanescent set, for $(f, s_1, \ldots, s_{i-1})|(h, s) \in S^{\otimes i}_{\mathbb{R}}$ such that $\bar{\xi}_{(f, s_1, \ldots, s_{i-1})|(h,s)} \neq \delta_0 \cdots \delta_0$ there is an increasing sequence $(\rho^j_{(f, s_1, \ldots, s_{i-1})|(h,s)})^n_{j=i}$ of $\bar{\mathcal{F}}^a$-stopping times such that

$$\omega \mapsto \int_{[0,1]^{n-i+1}} \delta_{\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_i) \cdots \delta_{\rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_n)\, du$$

defines an $\mathcal{F}^a$- measurable disintegration of $\bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)}$ w.r.t. $\mathbb{W}_{\mu_0}$. We make the important observation that, if $A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s) = 1$ (hence in this situation $\Delta A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s) > 0$), we have $\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)} \equiv \delta_0$.

This representation allows us to couple the two stopping rules by taking realizations of the $\rho^j_{(g,t_1,\ldots,t_i)}$ stopping times and $\rho^k_{(f,s_1,\ldots,s_i)|(h,s)}$ stopping times on the *same probability space* $\bar{\Omega}^{f\otimes h,g} := C(R_+) \times [0,1]^{n-i+1}$ where of course one of the u-coordinates is superfluous for the $\rho^j_{(g,t_1,\ldots,t_i)}$ stopping times. For $(f,s_1,\ldots,s_{i-1}),(h,s)$ and $(g,t_1,\ldots,t_i)$ as above and $n > j \geq i$ we define

$$\begin{aligned}
\Lambda_j^{f\otimes h,g} := \Big\{ &(\omega,u) \in \bar{\Omega}^{f\otimes h,g} : \rho^j_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u) \vee \rho^j_{(g,t_1,\ldots,t_i)}(\omega,u) \\
&\leq \rho^{j+1}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u) \wedge \rho^{j+1}_{(g,t_1,\ldots,t_i)}(\omega,u) \Big\}.
\end{aligned} \tag{4.4}$$

Note that this is the set where it is possible to swap the stopping rules from colour i up to colour j and *not swap* the stopping rule for colours greater than j.

The set of colour swap pairs between colour i and j, $i \leq j < n$, denoted by $\mathsf{CS}^\xi_{i\leftrightarrow j}$ is defined to consist of all $(f,s_1,\ldots,s_{i-1}) \in S_\mathbb{R}^{\otimes i-1}$, $(h,s) \in S$ and $(g,t_1,\ldots,t_i) \in S_\mathbb{R}^{\otimes i}$ such that $f \oplus h(s_{i-1}+s) = g(t_i)$, $1 - A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s) + \Delta A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s) > 0$, and

$$\begin{aligned}
&\int \gamma^{(f,s_1,\ldots,s_{i-1})|(h,s)\oplus}(\omega,s_i,\ldots,s_n)\, \bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(d\omega,ds_i,\ldots,ds_n) \\
&\quad + \int \gamma^{(g,t_1,\ldots,t_i)\otimes}(\omega,t_{i+1},\ldots,t_n)\, \xi_{(g,t_1,\ldots,t_i)}(d\omega,dt_{i+1},\ldots,dt_n) \\
&> \int \mathbb{W}(d\omega)du\, \mathbb{1}_{\Lambda_j^{f\otimes h,g}}(\omega,u) \Bigg[\int \gamma^{(f,s_1,\ldots,s_{i-1})|(h,s)\otimes}(\omega,t_{i+1},\ldots,t_j,s_{j+1},\ldots,s_n) \\
&\quad \delta_{\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{i+1}) \cdots \delta_{\rho^j_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_j) \\
&\quad \delta_{\rho^{j+1}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_{j+1}) \cdots \delta_{\rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_n) \\
&\quad + \int \gamma^{(g,t_1,\ldots,t_i)\oplus}(\omega,s_i,\ldots,s_j,t_{j+1},\ldots,t_n) \\
&\quad \delta_{\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_i) \cdots \delta_{\rho^j_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_j) \\
&\quad \delta_{\rho^{j+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{j+1}) \cdots \delta_{\rho^n_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_n) \Bigg] \\
&\quad + \int \mathbb{W}(d\omega)du \left(1 - \mathbb{1}_{\Lambda_j^{f\otimes h,g}}(\omega,u)\right) \Bigg[\int \gamma^{(f,s_1,\ldots,s_{i-1})|(h,s)\otimes}(\omega,t_{i+1},\ldots,t_n) \\
&\quad \delta_{\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{i+1}) \cdots \delta_{\rho^n_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_n) \\
&\quad + \int \gamma^{(g,t_1,\ldots,t_i)\oplus}(\omega,s_i,\ldots,s_n)\, \delta_{\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_i) \cdots \delta_{\rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_n) \Bigg].
\end{aligned} \tag{4.5}$$

Moreover, we agree that (4.5) holds in each of the following cases

(1) $\bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)} \notin \mathsf{RMST}^1_{n-i+1}$, $\xi_{(g,t_1,\ldots,t_i)} \notin \mathsf{RMST}^1_{n-i}$;
(2) the left hand side is infinite;
(3) any of the integrals appearing is not well-defined.

Then we set $\mathsf{CS}^{\xi}_i = \bigcup_{j \geq i} \mathsf{CS}^{\xi}_{i \leftrightarrow j}$.

Remark 4.4 (1) In case that $\Lambda^{f \oplus h, g}_j \neq \bar{\Omega}^{f \otimes h, g}$ it is not sufficient to only change the colours/stopping rules from colour i to j. On the complement of $\Lambda^{f \oplus h, g}$, we have to switch the whole stopping rule from colour i up to colour n in order to stay within the class of randomised multi-stopping times. This is precisely the reason for the two big integrals appearing on the right hand side of the inequality.

(2) Recall that $\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)} = \delta_0$ is possible so that it might happen that on both sides of (4.5) the stopping rule of colour i is in fact *the same* and we only change the stopping rule from colour $i+1$ onwards.

(3) In case of $\mathsf{CS}^{\xi}_{i \leftrightarrow i}$ the condition $1 - A^{\xi}_{(f,s_1,\ldots,s_{i-1})}(h,s) + \Delta A^{\xi}_{(f,s_1,\ldots,s_{i-1})}(h,s) > 0$ is not needed since there is no colour swap pair (with finite well defined integrals) not satisfying this condition.

4.3 Multi-colour swaps

Having introduced colour swap pairs we can now proceed and combine different colour swaps into multi-colour swap pairs. As described above, the aim is now to swap back *as soon as possible*. To this end, we consider for fixed $i < n$ the following partition of $\bar{\Omega}^{f \otimes h, g}$ defined in such a way that modifications of stopping rules in accordance to this partition transform randomised multi-stopping times into randomised multi-stopping times (c.f. (2.2)).

$$\bar{\Omega}^{f \otimes h, g} = \bigcup_{j=i}^{n} \left(\Lambda^{f \otimes h, g}_j \setminus \cup^{j-1}_{k=i} \Lambda^{f \otimes h, g}_k \right), \tag{4.6}$$

where

$$\Lambda^{f \otimes h, g}_n := \left\{ \rho^{i+1}_{(g,t_1,\ldots,t_i)} < \rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}, \rho^n_{(g,t_1,\ldots,t_i)} < \rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)} \right\}.$$

This is indeed a partition: The different sets are disjoint by construction. Hence, the right hand side of (4.6) is contained in the left hand side. We need to show that also the converse conclusion holds. Take $(\omega, u) \in \bar{\Omega}^{f \otimes h, g}$. If

$$\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega, u) \geq \rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u),$$

then $(\omega, u) \in \Lambda^{f \oplus h, g}_i$. Otherwise, it holds that

$$\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega, u) < \rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u) \leq \rho^{i+2}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u)$$

and either

$$\rho^{i+2}_{(g,t_1,\ldots,t_i)}(\omega, u) \geq \rho^{i+1}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u)$$

or

$$\rho^{i+2}_{(g,t_1,\ldots,t_i)}(\omega, u) < \rho^{i+1}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u) \leq \rho^{i+3}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u).$$

In the former case, we have $(\omega, u) \in \Lambda^{f\oplus h,g}_{i+1} \setminus \Lambda^{f\oplus h,g}_i$ and in the latter case we have either

$$\rho^{i+3}_{(g,t_1,\ldots,t_i)}(\omega, u) \geq \rho^{i+2}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u)$$

or

$$\rho^{i+3}_{(g,t_1,\ldots,t_i)}(\omega, u) < \rho^{i+2}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u) \leq \rho^{i+4}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega, u).$$

By induction, the claim follows. We put $\bar{\Lambda}^{f\otimes h,g}_j = \Lambda^{f\otimes h,g}_j \setminus \cup^{j-1}_{k=i} \Lambda^{f\otimes h,g}_k$. Then, the set of all *multi-colour swap pairs starting at colour i*, denoted by MCS^{ξ}_i, is defined to consist of all $(f, s_1, \ldots, s_{i-1}) \in S^{\otimes i-1}_{\mathbb{R}}$, $(h, s) \in S$, $(g, t_1, \ldots, t_i) \in S^{\otimes i}_{\mathbb{R}}$ such that $f \oplus h(s_{i-1} + s) = g(t_i)$ and

$$\int \gamma^{(f,s_1,\ldots,s_{i-1})|(h,s)\oplus}(\omega, s_i, \ldots, s_n) \, \bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(d\omega, ds_i, \ldots, ds_n)$$

$$+ \int \gamma^{(g,t_1,\ldots,t_i)\otimes}(\omega, t_{i+1}, \ldots, t_n) \, \xi_{(g,t_1,\ldots,t_i)}(d\omega, dt_{i+1}, \ldots, dt_n)$$

$$> \int \mathbb{W}(d\omega) du \sum_{j=i}^{n-1} \Bigg[\mathbb{1}_{\bar{\Lambda}^{f\otimes h,g}_j}(\omega, u)$$

$$\int \gamma^{(f,s_1,\ldots,s_{i-1})|(h,s)\otimes}(\omega, t_{i+1}, \ldots, t_j, s_{j+1}, \ldots, s_n)$$

$$\delta_{\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{i+1}) \cdots \delta_{\rho^j_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_j)$$

$$\delta_{\rho^{j+1}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_{j+1}) \cdots \delta_{\rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_n)$$

$$+ \int \gamma^{(g,t_1,\ldots,t_i)\oplus}(\omega, s_i, \ldots, s_j, t_{j+1}, \ldots, t_n)$$

$$\delta_{\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_i) \cdots \delta_{\rho^j_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_j)$$

$$\delta_{\rho^{j+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{j+1}) \cdots \delta_{\rho^n_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_n) \Bigg]$$

$$+ \int \mathbb{W}(d\omega) du \left(1 - \sum_{j=i}^{n-1} \mathbb{1}_{\bar{\Lambda}^{f\otimes h,g}_j}(\omega, u) \right) \left[\int \gamma^{(f,s_1,\ldots,s_{i-1})|(h,s)\otimes}(\omega, t_{i+1}, \ldots, t_n) \right.$$

$$\delta_{\rho^{i+1}_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_{i+1}) \cdots \delta_{\rho^n_{(g,t_1,\ldots,t_i)}(\omega,u)}(dt_n)$$

$$+ \int \gamma^{(g,t_1,\ldots,t_i)\oplus}(\omega, s_i, \ldots, s_n) \, \delta_{\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_i)$$

$$\cdots \delta_{\rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(ds_n) \Bigg]. \tag{4.7}$$

As in the case of colour swaps we agree that (4.7) holds in each of the following cases

(1) $\bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)} \notin \mathsf{RMST}^1_{n-i+1}, \xi_{(g,t_1,\ldots,t_i)} \notin \mathsf{RMST}^1_{n-i}$;
(2) the left hand side is infinite;
(3) any of the integrals appearing is not well-defined.

Remark 4.5 (1) Note that when $\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)} \equiv \delta_0$ we have $\bar{\Omega}^{f\otimes h,g} = \Lambda_i^{f\otimes h,g}$. Inserting this case into (4.7) we see that both sides agree so that there are no multi-colour swap pairs satisfying $\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)} \equiv \delta_0$.

(2) Observe that in the definition of MCS_i^ξ we do not need to impose the condition $1 - A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s) + \Delta A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s) > 0$ by Remark 4.4.

4.4 Stop-go pairs

Finally, we combine the previous two notions.

Definition 4.6 Let $\xi \in \mathsf{RMST}^1_n(\mathbb{R}, \mu_0)$. The set of stop-go pairs of colour i relative to ξ, SG_i^ξ, is defined by $\mathsf{SG}_i^\xi = \mathsf{CS}_i^\xi \cup \mathsf{MCS}_i^\xi$. We define the stop-go pairs of colour i in the wide sense by $\widehat{\mathsf{SG}}_i^\xi = \mathsf{SG}_i^\xi \cup \{(f, s_1, \ldots, s_{i-1})|(h, s) \in S_\mathbb{R}^{\otimes i} : A^\xi_{(f,s_1,\ldots,s_{i-1})}(h, s) = 1\} \times S_\mathbb{R}^{\otimes i}$.

The set of stop-go pairs relative to ξ is defined by $\mathsf{SG}^\xi := \bigcup_{1 \le i \le n} \mathsf{SG}_i^\xi$. The stop-go pairs in the wide sense are $\widehat{\mathsf{SG}}^\xi := \bigcup_{1 \le i \le n} \widehat{\mathsf{SG}}_i^\xi$.

Lemma 4.7 *Every stop-go pair is a stop-go pair in the wide sense, i.e. $\mathsf{SG}_i \subseteq \widehat{\mathsf{SG}}_i^\xi$ for any $1 \le i \le n$.*

Proof By loading notation, this follows using exactly the same argument as for the proof of [2, Lemma 5.4]. □

Remark 4.8 As in [2], we observe that the sets SG^ξ and $\widehat{\mathsf{SG}}^\xi$ are both Borel subsets of $S^{\otimes i} \times S^{\otimes i}$, since the maps given in e.g. (4.3) are measurable. In contrast, the set SG is in general just co-analytic.

5 The monotonicity principle

The aim of this section is to prove the main results, Theorem 2.5 and the closely related Theorem 5.2. The structure of this section follows closely the structure of the proof of

the corresponding results, Theorem 5.7 (resp. Theorem 5.16), in [2]. For the benefit of the reader, and to keep our presentation compact, we concentrate on those aspects of the proof where additional insight is needed to account for the multi-marginal aspects of the problem. We refer the reader to [2] for other details.

The essence of the proof is to first show that if we have a candidate optimiser ξ, and a joining rule π which identifies stop-go pairs, we can construct an infinitesimal improvement ξ^π, which will also be a candidate solution, but which will improve the objective. It will follow that the joining π will place no mass on the set of stop-go pairs. The second part of the proof shows that we can strengthen this to give a pointwise result, where we can exclude any stop-go pair from a set related to the support of the optimiser.

Important convention Throughout this section, we fix a function $\gamma : S^{\otimes n} \to \mathbb{R}$ and a measure $\xi \in \mathsf{RMST}^1(\mu_0, \mu_1, \ldots, \mu_n)$. Moreover, for each $0 \leq i \leq n - 1$ we fix a disintegration $(\xi_{(f,s_1,\ldots,s_i),\omega})_{(f,s_1,\ldots,s_i),\omega}$ of $\tilde{r}_{n,i}(\xi)$ w.r.t. $\tilde{r}_i(\xi^i)$ and set $\xi_{(f,s_1,\ldots,s_i)} = \xi_{(f,s_1,\ldots,s_i),\omega}\mathbb{W}(d\omega)$.

Recall the map $\mathrm{proj}_{S^{\otimes i}}$ from Sect. 2.1.

Definition 5.1 A family of Borel sets $\Gamma = (\Gamma_1, \ldots, \Gamma_n)$ with $\Gamma_i \subseteq S_{\mathbb{R}}^{\otimes i}$ is called (γ, ξ)-monotone iff for all $1 \leq i \leq n$

$$\widehat{\mathsf{SG}}_i^\xi \cap \left(\Gamma_i^< \times \Gamma_i\right) = \emptyset,$$

where

$$\Gamma_i^< = \{(f, s_1, \ldots, s_{i-1}, s) : \text{ there exists}$$
$$(g, s_1, \ldots, s_{i-1}, t) \in \Gamma_i, s_{i-1} \leq s < t, g_{\llcorner[0,s]} = f\},$$

and $\mathrm{proj}_{S^{\otimes i-1}}(\Gamma_i) \subseteq \Gamma_{i-1}$.

Theorem 5.2 *Assume that $\gamma : S^{\otimes n} \to \mathbb{R}$ is Borel measurable. Assume that* (OptMSEP) *is well posed and that $\xi \in \mathsf{RMST}(\mu_0, \ldots, \mu_1)$ is an optimizer. Then there exists a (γ, ξ)-monotone family of Borel sets $\Gamma = (\Gamma_1, \ldots, \Gamma_n)$ such that $r_i(\xi)(\Gamma_i) = 1$ for each $1 \leq i \leq n$.*

The proof of Theorem 5.2 is based on the following two propositions.

Proposition 5.3 *Let $\gamma : S_{\mathbb{R}}^{\otimes n} \to \mathbb{R}$ be Borel. Assume that* (OptMSEP) *is well posed and that $\xi \in \mathsf{RMST}(\mu_0, \ldots, \mu_1)$ is an optimizer. Fix $1 \leq i \leq n$ and set $X = S_{\mathbb{R}}^{\otimes i-1}$, $m = r_{i-1}(\xi^{i-1})$. Then*

$$(r_X \otimes \mathrm{Id})(\pi)(SG_i^\xi) = 0$$

for all $\pi \in \mathsf{JOIN}(r_{i-1}(\xi^{i-1}), r_i(\xi^i))$.

Proposition 5.4 *Let (X, m) and (Y, ν) be Polish probability spaces and $E \subseteq S_X \times Y$ a Borel set. Then the following are equivalent:*

(1) $(r_X \otimes \mathrm{Id})(\pi)(E) = 0$ *for all* $\pi \in \mathsf{JOIN}^1(m, \nu)$.
(2) $E \subseteq (F \times \mathcal{Y}) \cup (S_X \times N)$ *for some evanescent set* $F \subseteq S_X$ *and a ν-null set* $N \subseteq Y$.

Proof This is a straightforward modification of [2, Proposition 5.9] to the case of a general starting law (see also the proof of [2, Theorem 7.4]). □

Remark 5.5 Note that Proposition 5.4 is closely related to the classical section theorem (cf. [20, Theorems IV 84 and IV 85]) which in our setup implies the following statement:

Let $(X, \mathcal{B}, m)$ be a Polish probability space. $E \subseteq S_X$ be Borel. Then the following are equivalent:

(1) $r_X(\alpha)(E) = 0$ for all $\alpha \in \mathsf{RST}(X, m)$
(2) E is m-evanescent
(3) $\mathbb{P}(((Y_s)_{s \leq \tau}, \tau) \in E) = 0$ for every $\mathcal{G}^0$-stopping time τ.

Proof of Theorem 5.2 Fix $1 \leq i \leq n$. Set $X = S_{\mathbb{R}}^{\otimes i-1}$, $m = r_{i-1}(\xi^{i-1})$ and consider the corresponding probability space $(\mathcal{X}, \mathbb{P})$. By Proposition 5.3 $(r_X \otimes \mathrm{Id})(\pi)(\mathsf{SG}_i^\xi) = 0$ for all $\pi \in \mathsf{JOIN}^1(r_{i-1}(\xi^{i-1}), r_i(\xi^i))$. Applying Proposition 5.4 with $Y = S_{\mathbb{R}}^{\otimes i}$, $\nu = r_i(\xi^i)$ we deduce that there exists a $r_{i-1}(\xi^{i-1})$-evanescent set $\tilde{F}_i$ and a $r_i(\xi^i)$-null set N_i such that

$$\mathsf{SG}_i^\xi \subseteq (\tilde{F}_i \times S_{\mathbb{R}}^{\otimes i}) \cup (S_{\mathbb{R}}^{\otimes i} \times N_i).$$

Put $F_i := \{(g, t_1, \ldots, t_i) \in S_{\mathbb{R}}^{\otimes i} : \exists (f, t_1, \ldots, t_{i-1}, s_i) \in \tilde{F}_i, t_i \geq s_i, g \equiv f \text{ on } [0, s_i]\}$. Then, F_i is $r_{i-1}(\xi^{i-1})$-evanescent and

$$\mathsf{SG}_i^\xi \subseteq (F_i \times S_{\mathbb{R}}^{\otimes i}) \cup (S_{\mathbb{R}}^{\otimes i} \times N_i).$$

Setting $\tilde{\Gamma}_i = S_{\mathbb{R}}^{\otimes i} \setminus (F_i \cup N_i)$ we have $r_i(\xi^i)(\tilde{\Gamma}_i) = 1$ as well as $\mathsf{SG}_i^\xi \cap (\tilde{\Gamma}_i^< \times \tilde{\Gamma}_i) = \emptyset$. Define

$$\Gamma_i := \tilde{\Gamma}_i \cap \{(g, t_1, \ldots, t_i) \in S_{\mathbb{R}}^{\otimes i} : A_{(g,t_1,\ldots,t_{i-1})}^\xi (\theta_{t_{i-1}}(g)_{\llcorner[0,s]}, s)$$
$$< 1 \text{ for all } s < t_i - t_{i-1}\},$$

where $\theta_u(g)(\cdot) = g(\cdot + u) - g(u)$ as usual. Then $r_i(\xi^i)(\Gamma_i) = 1$ and $\Gamma_i^< \cap \{(g, t_1, \ldots, t_i) \in S_{\mathbb{R}}^{\otimes i} : A_{(g,t_1,\ldots,t_{i-1})}^\xi (\theta_{t_{i-1}}(g)_{\llcorner[0,t_i-t_{i-1}]}, t_i - t_{i-1}) = 1\} = \emptyset$ so that $\widehat{\mathsf{SG}}_i^\xi \cap (\Gamma_i^< \times \Gamma_i) = \emptyset$. Finally, we can take a Borel subset of Γ_i with full measure and taking suitable intersections we can assume that $\mathrm{proj}_{S_{\mathbb{R}}^{\otimes i-1}}(\Gamma_i) \subseteq \Gamma_{i-1}$. □

5.1 Proof of Proposition 5.3

For notational convenience we will only prove the statement for the colour swap pairs CS_i^ξ. As the colour swap pairs are the main building block for the multi-colour swap

pairs MCS_i^{ξ} it will be immediate how to adapt the proof for the general case. Moreover, it is clearly sufficient to show that for every $j \geq i$ we have $(r_X \otimes \mathrm{Id})(\pi)(\mathsf{CS}_{i \leftrightarrow j}^{\xi}) = 0$ for each $\pi \in \mathsf{JOIN}(r_{i-1}(\xi^{i-1}), r_i(\xi^i))$.

Working towards a contradiction, we assume that there is an index $i \leq j \leq n$ and $\pi \in \mathsf{JOIN}(r_{i-1}(\xi^{i-1}), r_i(\xi^i))$ such that $(r_X \otimes \mathrm{Id})(\pi)(\mathsf{CS}_{i \leftrightarrow j}^{\xi}) > 0$. By the definition of joinings (Definition 3.20) also $\pi_{\llcorner (r_X \otimes \mathrm{Id})^{-1}(E)} \in \mathsf{JOIN}(r_{i-1}(\xi^{i-1}), r_i(\xi^i))$ for any $E \subseteq S_{\mathbb{R}}^{\otimes i-1} \times S_{\mathbb{R}}^{\otimes i}$. Hence, by considering $(r_X \otimes \mathrm{Id})(\pi)_{\llcorner \mathsf{CS}_{i \leftrightarrow j}^{\xi}}$ we may assume that $(r_X \otimes \mathrm{Id})(\pi)$ is concentrated on $\mathsf{CS}_{i \leftrightarrow j}^{\xi}$. Recall that (by definition) there is no colour swap pair $((f, s_1, \ldots, s_{i-1}) | (h, s), (g, t_1, \ldots, t_i))$ with $A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) = 1$ and $\Delta A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) = 0$. Hence,

$$\pi \left[((f, s_1, \ldots, s_{i-1}) | (h, s), (g, t_1, \ldots, t_i)) : A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) = 1 \text{ and} \right.$$
$$\left. \Delta A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) = 0 \right] = 0. \tag{5.1}$$

We argue by contradiction and define two modifications of ξ, ξ^E and ξ^L, based on the definition of $\mathsf{CS}_{i \leftrightarrow j}^{\xi}$ such that their convex combination yields a randomised multi-stopping time embedding the same measures as ξ and leading to strictly less cost. The stopping time ξ^E will stop paths earlier than ξ, and ξ^L will stop paths later than ξ.

By Lemma 4.2 and Corollary 3.13, for $(f, s_1, \ldots, s_{i-1}) | (h, s)$ outside an $r_{i-1}(\xi^{i-1})$ - evanescent set and $(g, t_1, \ldots, t_i)$ outside an $r_i(\xi^i)$ null set there are increasing sequences of $\bar{\mathcal{G}}^a$-stopping times $(\rho_{(f, s_1, \ldots, s_{i-1}) | (h, s)}^{j})_{j=i}^n$ and $(\rho_{(g, t_1, \ldots, t_i)}^{j})_{j=i+1}^n$ defining $\bar{\mathcal{G}}^a$-measurable disintegrations of $\xi_{(f, s_1, \ldots, s_{i-1}) | (h, s)}$ and $\xi_{(g, t_1, \ldots, t_i)}$ as in (3.9).

For $B \subseteq C(\mathbb{R}_+) \times \Xi^n$ and $(g, t_1, \ldots, t_i) \in S_{\mathbb{R}}^{\otimes i}$ we set

$$B^{(g, t_1, \ldots, t_i) \oplus} := \{ (\omega, T_i, \ldots, T_n) \in C(\mathbb{R}_+) \times \Xi^{n-i+1}$$
$$: (g \oplus \omega, t_1, \ldots, t_{i-1}, t_i + T_i, \ldots, t_i + T_n) \in B \}$$

and

$$B^{(g, t_1, \ldots, t_i) \otimes} := \{ (\omega, T_{i+1}, \ldots, T_n) \in C(\mathbb{R}_+) \times \Xi^{n-i}$$
$$: (g \oplus \omega, t_1, \ldots, t_i, t_i + T_{i+1}, \ldots, t_i + T_n) \in B \}.$$

Observe that both $B^{(g, t_1, \ldots, t_i) \oplus}$ and $B^{(g, t_1, \ldots, t_i) \otimes}$ are then Borel. Note that if we define $F(g, t_1, \ldots, t_n) = \mathbb{1}_B(g, t_1, \ldots, t_n)$, then $F^{(f, s_1, \ldots, s_i) \oplus}(\eta, t_{i+1}, \ldots, t_n) = \mathbb{1}_{B^{(f, s_1, \ldots, s_i) \oplus}}(\eta, t_{i+1}, \ldots, t_n)$, and similarly for $B^{(g, t_1, \ldots, t_i) \otimes}$. Observe, that for $(f, s_1, \ldots, s_{i-1}) | (h, s)$ with $A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) = 1$ and $\Delta A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) > 0$ it follows from (4.1) that

$$\xi_{(f, s_1, \ldots, s_{i-1}) | (h, s)}(B^{(g, t_1, \ldots, t_i) \oplus}) = \Delta A_{(f, s_1, \ldots, s_{i-1})}^{\xi}(h, s) \xi_{(f \oplus h, s_1, \ldots, s_{i-1}, s_{i-1} + s)}(B^{(g, t_1, \ldots, t_i) \otimes}).$$

Then, we define the measure ξ^E by setting for $B \subseteq C(\mathbb{R}_+) \times \Xi^n$ (recall $\Lambda_j^{f \otimes h,g}$ from (4.4))

$$\xi^E(B)$$

$$:= \xi(B) - \int \xi_{(f,s_1,\ldots,s_{i-1})|(h,s)}(B^{(f \oplus h,s_1,\ldots,s_{i-1},s_{i-1}+s)\oplus})$$

$$(r_X \otimes \mathrm{Id})(\pi)(d((f,s_1,\ldots,s_{i-1})|(h,s)), d(g,t_1,\ldots,t_i))$$

$$+ \int (r_X \otimes \mathrm{Id})(\pi)(d((f,s_1,\ldots,s_{i-1})|(h,s)), d(g,t_1,\ldots,t_i))$$

$$\left(1 - A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s) + \mathbb{1}_{A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)=1} \Delta A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)\right)$$

$$\left[\int_{C(\mathbb{R}_+)} \int_{[0,1]^{n-i+1}} \mathbb{1}_{\Lambda_j^{f \otimes h,g}}(\omega,u) \mathbb{1}_{B^{(f \oplus h,s_1,\ldots,s_{i-1},s_{i-1}+s)\otimes}}\left(\omega, \rho_{(g,t_1,\ldots,t_i)}^{i+1}(\omega,u),\ldots,\right.\right.$$

$$\left.\rho_{(g,t_1,\ldots,t_i)}^{j}(\omega,u), \rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{j+1}(\omega,u), \ldots, \rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{n}(\omega,u)\right) \mathbb{W}(d\omega)\, du$$

$$+ \int_{C(\mathbb{R}_+)} \int_{[0,1]^{n-i+1}} \left(1 - \mathbb{1}_{\Lambda_j^{f \otimes h,g}}(\omega,u)\right)$$

$$\left.\mathbb{1}_{B^{(f \oplus h,s_1,\ldots,s_{i-1},s_{i-1}+s)\otimes}}\left(\omega, \rho_{(g,t_1,\ldots,t_i)}^{i+1}(\omega,u),\ldots,\rho_{(g,t_1,\ldots,t_i)}^{n}(\omega,u)\right) \mathbb{W}(d\omega)\, du\right].$$

$$(5.2)$$

Similarly we define the measure ξ^L by setting for $B \subseteq C(\mathbb{R}_+) \times \Xi^n$

$$\xi^L(B)$$

$$:= \xi(B) - \int \left(1 - A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)\right.$$

$$\left. + \mathbb{1}_{A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)=1} \Delta A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)\right) \xi_{(g,t_1,\ldots,t_i)}(B^{(g,t_1,\ldots,t_i)\otimes})$$

$$(r_X \otimes \mathrm{Id})(\pi)(d((f,s_1,\ldots,s_{i-1})|(h,s)), d(g,t_1,\ldots,t_i))$$

$$+ \int \left(1 - A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s) + \mathbb{1}_{A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)=1} \Delta A_{(f,s_1,\ldots,s_{i-1})}^{\xi}(h,s)\right)$$

$$\left[\int_{C(\mathbb{R}_+)} \int_{[0,1]^{n-i+1}} \mathbb{1}_{\Lambda_j^{f \otimes h,g}}(\omega,u) \mathbb{1}_{B^{(g,t_1,\ldots,t_i)\oplus}}\left(\omega, \rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{i}(\omega,u),\ldots,\right.\right.$$

$$\left.\rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{j}(\omega,u), \rho_{(g,t_1,\ldots,t_i)}^{j+1}(\omega,u), \ldots, \rho_{(g,t_1,\ldots,t_i)}^{n}(\omega,u)\right) \mathbb{W}(d\omega)\, du$$

$$+ \int_{C(\mathbb{R}_+)} \int_{[0,1]^{n-i+1}} \left(1 - \mathbb{1}_{\Lambda_j^{f \otimes h,g}}(\omega,u)\right)$$

$$\left.\mathbb{1}_{B^{(g,t_1,\ldots,t_i)\oplus}}\left(\omega, \rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{i}(\omega,u),\ldots,\rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{n}(\omega,u)\right) \mathbb{W}(d\omega)\, du\right]$$

$$(r_X \otimes \mathrm{Id})(\pi)(d((f,s_1,\ldots,s_{i-1})|(h,s)), d(g,t_1,\ldots,t_i)).$$

$$(5.3)$$

Then, we define a competitor of ξ by $\xi^\pi := \frac{1}{2}(\xi^E + \xi^L)$. We will show that $\xi^\pi \in \mathsf{RMST}(\mu_0, \ldots, \mu_n)$ and $\int \gamma\, d\xi > \int \gamma\, d\xi^\pi$ which contradicts optimality of ξ.

First of all note that from the definition of $\Lambda_j^{f \oplus h, g}$ in (4.4) both $\xi^E, \xi^L \in \mathsf{RMST}$ (also compare (2.2)). Hence, also $\xi^\pi \in \mathsf{RMST}$. Next we show that $\xi^\pi \in \mathsf{RMST}(\mu_0, \ldots, \mu_n)$.

For bounded and measurable $F : C(\mathbb{R}) \times \Xi^n \to \mathbb{R}$ (5.2) and (5.3) imply by using (4.1) and (4.2)

$$
\begin{aligned}
2 \int & F\, d(\xi - \xi^\pi) \\
&= \int (r_X \otimes r_i)(\pi)(d((f, s_1, \ldots, s_{i-1})|(h, s)), d(g, t_1, \ldots, t_i)) \\
&\quad \left(1 - A^\xi_{(f, s_1, \ldots, s_{i-1})}(h, s) + \mathbb{1}_{A^\xi_{(f, s_1, \ldots, s_{i-1})}(h,s)=1} \Delta A^\xi_{(f, s_1, \ldots, s_{i-1})}(h, s)\right) \\
&\quad \times \left(\int F^{(f, s_1, \ldots, s_{i-1})|(h,s)\oplus}(\omega, S_i, \ldots, S_n)\, \bar{\xi}_{(f, s_1, \ldots, s_{i-1})|(h,s)}(d\omega, dS_i, \ldots, dS_n)\right. \\
&\quad + \int F^{(g, t_1, \ldots, t_i)\otimes}(\omega, T_{i+1}, \ldots, T_n)\, \xi_{(g, t_1, \ldots, t_i)}(d\omega, dT_{i+1}, \ldots, dT_n) \\
&\quad - \int_{C(\mathbb{R}_+)} \int_{[0,1]^{n-i+1}} \mathbb{W}(d\omega) du\, \mathbb{1}_{\Lambda_j^{f\otimes h, g}}(\omega, u) \\
&\quad \left[\int F^{(f, s_1, \ldots, s_{i-1})|(h,s)\otimes}(\omega, T_{i+1}, \ldots, T_j, S_{j+1}, \ldots, S_n)\right. \\
&\quad \delta_{\rho^{i+1}_{(g, t_1, \ldots, t_i)}(\omega, u)}(dT_{i+1}) \cdots \delta_{\rho^j_{(g, t_1, \ldots, t_i)}(\omega, u)}(dT_j) \\
&\quad \delta_{\rho^{j+1}_{(f, s_1, \ldots, s_{i-1})|(h,s)}(\omega, u)}(dS_{j+1}) \cdots \delta_{\rho^n_{(f, s_1, \ldots, s_{i-1})|(h,s)}(\omega, u)}(dS_n) \\
&\quad - \int F^{(g, t_1, \ldots, t_i)\oplus}(\omega, S_i, \ldots, S_j, T_{j+1}, \ldots, T_n) \\
&\quad \delta_{\rho^i_{(f, s_1, \ldots, s_{i-1})|(h,s)}(\omega, u)}(dS_i) \cdots \delta_{\rho^j_{(f, s_1, \ldots, s_{i-1})|(h,s)}(\omega, u)}(dS_j) \\
&\quad \left. \delta_{\rho^{j+1}_{(g, t_1, \ldots, t_i)}(\omega, u)}(dT_{j+1}) \cdots \delta_{\rho^n_{(g, t_1, \ldots, t_i)}(\omega, u)}(dT_n)\right] \\
&\quad - \int \mathbb{W}(d\omega) du \left(1 - \mathbb{1}_{\Lambda_j^{f\otimes h, g}}(\omega, u)\right)\left[\int F^{(f, s_1 \ldots, s_{i-1})|(h,s)\otimes}(\omega, T_{i+1}, \ldots, T_n)\right. \\
&\quad \delta_{\rho^{i+1}_{(g, t_1, \ldots, t_i)}(\omega, u)}(dT_{i+1}) \cdots \delta_{\rho^n_{(g, t_1, \ldots, t_i)}(\omega, u)}(dT_n) \\
&\quad - \int F^{(g, t_1, \ldots, t_i)\oplus}(\omega, S_i, \ldots, S_n) \\
&\quad \left.\left. \delta_{\rho^i_{(f, s_1, \ldots, s_{i-1})|(h,s)}(\omega, u)}(dS_i) \cdots \delta_{\rho^n_{(f, s_1, \ldots, s_{i-1})|(h,s)}(\omega, u)}(dS_n)\right]\right) .
\end{aligned}
\tag{5.4}
$$

Next we show that (5.4) extends to nonnegative F satisfying $\xi(F) < \infty$ in the sense that it is well defined with a value in $[-\infty, \infty)$. To this end, we will show that

$$\int (r_X \otimes r_i)(\pi)(d((f, s_1, \ldots, s_{i-1})|(h, s)), d(g, t_1, \ldots, t_i))$$

$$\left(1 - A^\xi_{(f,s_1,\ldots,s_{i-1})}(h, s) + \mathbb{1}_{A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s)=1} \Delta A^\xi_{(f,s_1,\ldots,s_{i-1})}(h, s) \right)$$

$$\times \left[\int F^{(f,s_1,\ldots,s_{i-1})|(h,s)\oplus}(\omega, S_i, \ldots, S_n) \, \bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(d\omega, dS_i, \ldots, dS_n) \right.$$

$$\left. + \int F^{(g,t_1,\ldots,t_i)\otimes}(\omega, T_{i+1}, \ldots, T_n) \, \xi_{(g,t_1,\ldots,t_i)}(d\omega, dT_{i+1}, \ldots, dT_n) \right] < \infty$$

$$(5.5)$$

Since $\pi \in \mathsf{JOIN}(r_{i-1}(\xi^{i-1}), r_i(\xi^i))$ the integral of the second term in the square brackets in (5.5) is bounded by $\xi(F)$, and hence is finite. To see that the first term is also finite, write $\mathrm{proj}_{X \times \mathbb{R}_+}(\pi) =: \pi_1$ and note that $\pi_1 \in \mathsf{RST}(S_{\mathbb{R}}^{\otimes i-1}, r_{i-1}(\xi^{i-1}))$. Hence, the disintegration $(\pi_1)_{(f,s_1,\ldots,s_{i-1})}$ of π_1 wrt $r_{i-1}(\xi^{i-1})$ is a.s. in RST. Fix $(f, s_1, \ldots, s_{i-1}) \in S_{\mathbb{R}}^{\otimes i-1}$ and assume $\alpha := (\pi_1)_{(f,s_1,\ldots,s_{i-1})} \in \mathsf{RST}$ (which holds on a set of measure one). In case that $\alpha_\omega(\mathbb{R}_+) < 1$ we extend it to a probability on $[0, \infty]$ by adding an atom at ∞. We denote the resulting randomized stopping time still by α. Then, we can calculate using the strong Markov property of the Wiener measure for the first equality and $F \geq 0$ for the first inequality

$$\int F^{(f,s_1,\ldots,s_{i-1})\otimes}(\omega, s_i, \ldots, s_n) \, \xi_{(f,s_1,\ldots,s_{i-1})}(d\omega, ds_i, \ldots, ds_n)$$

$$= \iint F^{(f,s_1,\ldots,s_{i-1})\otimes}(\omega_{\llcorner[0,t]} \oplus \theta_t \omega, s_i, \ldots, s_n)$$

$$(\xi_{(f,s_1,\ldots,s_{i-1})})_{\omega_{\llcorner[0,t]}\oplus\theta_t\omega}(ds_i, \ldots, ds_n)\alpha_\omega(dt)\mathbb{W}(d\omega)$$

$$= \iint F^{(f,s_1,\ldots,s_{i-1})\otimes}(\omega_{\llcorner[0,t]} \oplus \tilde{\omega}, s_i, \ldots, s_n)$$

$$(\xi_{(f,s_1,\ldots,s_{i-1})})_{\omega_{\llcorner[0,t]}\oplus\tilde{\omega}}(ds_i, \ldots, ds_n)\alpha_\omega(dt)\mathbb{W}(d\omega)\mathbb{W}(d\tilde{\omega})$$

$$\geq \iint \mathbb{1}_{\{(\omega,t):t\leq s_i<\infty\}} F^{(f,s_1,\ldots,s_{i-1})\otimes}(\omega_{\llcorner[0,t]} \oplus \tilde{\omega}, s_i, \ldots, s_n)$$

$$(\xi_{(f,s_1,\ldots,s_{i-1})})_{\omega_{\llcorner[0,t]}\oplus\tilde{\omega}}(ds_i, \ldots, ds_n)\alpha_\omega(dt)\mathbb{W}(d\omega)\mathbb{W}(d\tilde{\omega})$$

$$= \iint \mathbb{1}_{\{(\omega,t):t<\infty\}} F^{(f,s_1,\ldots,s_{i-1})|r(\omega,t)\oplus}(\tilde{\omega}, s_i', \ldots, s_n')$$

$$\xi_{(f,s_1,\ldots,s_{i-1})|r(\omega,t)}(d\tilde{\omega}, ds_i', \ldots, ds_n')\alpha(d\omega, dt)$$

Hence,

$$\int (r_X \otimes r_i)(\pi)(d((f, s_1, \ldots, s_{i-1})|(h, s)), d(g, t_1, \ldots, t_i))$$

$$\left(1 - A^\xi_{(f,s_1,\ldots,s_{i-1})}(h, s) + \mathbb{1}_{A^\xi_{(f,s_1,\ldots,s_{i-1})}(h,s)=1} \Delta A^\xi_{(f,s_1,\ldots,s_{i-1})}(h, s) \right)$$

$$\times \int F^{(f,s_1,\ldots,s_{i-1})|(h,s)\oplus}(\omega, S_i, \ldots, S_n)\, \bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(d\omega, dS_i, \ldots, dS_n)$$

$$= \int (r_X \otimes r_i)(\pi)(d((f, s_1, \ldots, s_{i-1})|(h, s)), d(g, t_1, \ldots, t_i)) \times$$

$$\int F^{(f,s_1,\ldots,s_{i-1})|(h,s)\oplus}(\omega, S_i, \ldots, S_n)\, \xi_{(f,s_1,\ldots,s_{i-1})|(h,s)}(d\omega, dS_i, \ldots, dS_n)$$

$$\leq \iint F^{(f,s_1,\ldots,s_{i-1})\otimes}(\omega, s_i, \ldots, s_n)\, \xi_{(f,s_1,\ldots,s_{i-1})}(d\omega, ds_i, \ldots, ds_n)$$

$$r_{i-1}(\xi^{i-1})(d((f, s_1, \ldots, s_{i-1}))$$

$$= \xi(F) < \infty.$$

Applying (5.4) to $\mathsf{t}_n(\omega, t_1, \ldots, t_n) = t_n$, and observing that all the terms on the right-hand side cancel implies that $\xi^\pi(\mathsf{t}_n) = \xi(\mathsf{t}_n) < \infty$. Taking $F(\omega, s_1, \ldots, s_n) = G(\omega(s_j))$ for $0 \leq j \leq n$ with $s_0 := 0$ for bounded and measurable $G : \mathbb{R} \to \mathbb{R}$ the right hand side vanishes. For $j < i$ this follows since $\xi^{i-1} = (\xi^\pi)^{i-1}$ as we have not changed any stopping rule for colours prior to i. For $j \geq i$ this follows from the fact that π is concentrated on pairs $((f, s_1, \ldots, s_{i-1})|(h, s), (g, t_1, \ldots, t_i))$ satisfying $f \oplus h(s_i + s) = g(t_i)$. Hence, we have shown that $\xi^\pi \in \mathsf{RMST}(\mu_0, \ldots, \mu_n)$.

Using that $\xi^\pi(\gamma^-), \xi(\gamma^-) < \infty$, by well posedness of (OptMSEP), we can apply (5.4) to $F = \gamma$ to obtain by the Definition of $\mathsf{CS}_{i \leftrightarrow j}^\xi$ that

$$\int (\gamma \circ r_n)^{(f,s_1,\ldots,s_{i-1})|(h,s)\oplus}(\omega, S_i, \ldots, S_n)\, \bar{\xi}_{(f,s_1,\ldots,s_{i-1})|(h,s)}(d\omega, dS_i, \ldots, dS_n)$$

$$+ \int (\gamma \circ r_n)^{(g,t_1,\ldots,t_i)\otimes}(\omega, T_{i+1}, \ldots, T_n)\, \xi_{(g,t_1,\ldots,t_i)}(d\omega, dT_{i+1}, \ldots, dT_n)$$

$$- \int_{C(\mathbb{R}_+)} \int_{[0,1]^{n-i+1}} \mathbb{1}_{\Lambda_j^{f \otimes h, g}}(\omega, u)$$

$$\Bigg[\int (\gamma \circ r_n)^{(f,s_1,\ldots,s_{i-1})|(h,s)\otimes}(\omega, T_{i+1}, \ldots, T_j, S_{j+1}, \ldots, S_n)$$

$$\delta_{\rho_{(g,t_1,\ldots,t_i)}^{i+1}(\omega,u)}(dT_{i+1}) \cdots \delta_{\rho_{(g,t_1,\ldots,t_i)}^{j}(\omega,u)}(dT_j)\, \delta_{\rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{j+1}(\omega,u)}(dS_{j+1})$$

$$\cdots \delta_{\rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{n}(\omega,u)}(dS_n)$$

$$- \int (\gamma \circ r_n)^{(g,t_1,\ldots,t_i)\oplus}(\omega, S_i, \ldots, S_j, T_{j+1}, T_n)$$

$$\delta_{\rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{i}(\omega,u)}(dS_i) \cdots \delta_{\rho_{(f,s_1,\ldots,s_{i-1})|(h,s)}^{j}(\omega,u)}(dS_j)\, \delta_{\rho_{(g,t_1,\ldots,t_i)}^{j+1}(\omega,u)}(dT_{j+1})$$

$$\cdots \delta_{\rho_{(g,t_1,\ldots,t_i)}^{n}(\omega,u)}(dT_n) \Bigg] \mathbb{W}(d\omega)\, du$$

$$- \int \left(1 - \mathbb{1}_{\Lambda_j^{f \otimes h, g}}(\omega, u)\right) \Bigg[\int (\gamma \circ r_n)^{(f,s_1,\ldots,s_{i-1})|(h,s)\otimes}(\omega, T_{i+1}, \ldots, T_n)$$

$$\delta_{\rho_{(g,t_1,\ldots,t_i)}^{i+1}(\omega,u)}(dT_{i+1}) \cdots \delta_{\rho_{(g,t_1,\ldots,t_i)}^{n}(\omega,u)}(dT_n)$$

$$- \int (\gamma \circ r_n)^{(g,t_1,\ldots,t_i)\oplus}(\omega, S_i, \ldots, S_n)\, \delta_{\rho^i_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(dS_i)$$

$$\cdots \delta_{\rho^n_{(f,s_1,\ldots,s_{i-1})|(h,s)}(\omega,u)}(dS_n)\bigg] \mathbb{W}(d\omega)du$$

is $(r_X \otimes \mathrm{Id})(\pi)$ a.s. strictly positive applying Lemma 4.2. Hence, we arrive at the contradiction $\int \gamma \, d\xi^\pi < \int \gamma d\xi$. □

5.2 Secondary optimization (and beyond)

We set $\bar{\mu} = (\mu_0, \ldots, \mu_n)$ and denote by $\mathsf{Opt}(\gamma, \bar{\mu})$ the set of optimizers of (OptMSEP). If $\pi \mapsto \int \gamma d\pi$ is lower semicontinuous then $\mathsf{Opt}(\gamma, \bar{\mu})$ is a closed subset of $\mathsf{RMST}(\mu_0, \mu_1, \ldots, \mu_n)$ and therefore also compact.

Definition 5.6 (Secondary stop-go pairs) Let $\gamma, \gamma' : S_{\mathbb{R}}^{\otimes n} \to \mathbb{R}$ be Borel measurable. The set of secondary stop-go pairs of colour i relative to ξ, short $\mathsf{SG}_{2,i}^\xi$, consists of all $(f, s_1, \ldots, s_{i-1})|(h,s) \in S_{\mathbb{R}}^{\otimes i}, (g, t_1, \ldots, t_i) \in S_{\mathbb{R}}^{\otimes i}$ such that $f \oplus h(s_{i-1} + s) = g(t_i)$ and either $((f, s_1, \ldots, s_{i-1})|(h,s), (g, t_1, \ldots, t_i)) \in \mathsf{SG}_i^\xi$ or equality holds in (4.5) for γ, and strict inequality holds in (4.5) for γ', or equality holds in (4.7) for γ, and strict inequality holds in (4.7) for γ'. As before we agree that $((f, s_1, \ldots, s_{i-1})|(h,s), (g, t_1, \ldots, t_i)) \in \mathsf{SG}_i^\xi$ if either of the integrals in (4.5) or (4.7) is infinite or not well defined.

We also define the secondary stop-go pairs of colour i relative to ξ in the wide sense, $\widehat{\mathsf{SG}}_{2,i}^\xi$, by $\widehat{\mathsf{SG}}_{2,i}^\xi = \mathsf{SG}_{2,i}^\xi \cup \{(f, s_1, \ldots, s_{i-1})|(h,s) \in S_{\mathbb{R}}^{\otimes i} : A_{(f,s_1,\ldots,s_{i-1})}^\xi(h,s) = 1\} \times S_{\mathbb{R}}^{\otimes i}$.

The set of secondary stop-go pairs relative to ξ is defined by $\mathsf{SG}^\xi := \bigcup_{1 \leq i \leq n} \mathsf{SG}_i^\xi$. The secondary stop-go pairs in the wide sense are $\widehat{\mathsf{SG}}^\xi := \bigcup_{1 \leq i \leq n} \widehat{\mathsf{SG}}_i^\xi$.

Theorem 5.7 (Secondary minimisation) *Let $\gamma, \gamma' : S_{\mathbb{R}}^{\otimes n} \to \mathbb{R}$ be Borel measurable. Assume that $\mathsf{Opt}(\gamma, \bar{\mu}) \neq \emptyset$ and that $\xi \in \mathsf{Opt}(\gamma, \bar{\mu})$ is an optimiser for*

$$P_{\gamma'|\gamma}(\bar{\mu}) = \inf_{\pi \in \mathsf{Opt}(\gamma, \bar{\mu})} \int \gamma' \, d\pi. \tag{5.6}$$

Then, for any $i \leq n$ there exists a Borel set $\Gamma_i \subseteq S^{\otimes i}$ such that $r_i(\xi^i)(\Gamma_i) = 1$ and

$$\widehat{\mathsf{SG}}_{2,i}^\xi \cap (\Gamma_i^< \times \Gamma_i) = \emptyset. \tag{5.7}$$

Theorem 5.7 follows from a straightforward modification of Proposition 5.3 by the same proof as for Theorem 5.2 using Proposition 5.4. We omit further details.

Remark 5.8 Of course the previous theorem can be applied repeatedly to a sequence of functions $\gamma, \gamma', \gamma'', \ldots$ and sets $\mathsf{Opt}(\gamma, \bar{\mu}), \mathsf{Opt}(\gamma'|\gamma, \bar{\mu}), \mathsf{Opt}(\gamma''|\gamma'|\gamma, \bar{\mu}), \ldots$ leading to $\mathsf{SG}_3^\xi, \mathsf{SG}_4^\xi, \ldots$. We omit further details.

5.3 Proof of main result

We are now able to conclude, by observing that our main result is now a simple consequence of previous results.

Proof of Theorem 2.5 Since any $\xi \in \mathsf{RMST}(\mu_0, \ldots, \mu_n)$ induces via Lemma 4.2 and Corollary 3.13 a sequence of stopping times as used for the definition of stop-go pairs in Sect. 2.1 the result follows from Theorem 5.7. □

References

1. Azéma, J., Yor, M.: Une solution simple au probleme de Skorokhod. In: Séminaire de probabilités XIII, pp. 90–115. Springer (1979)
2. Beiglböck, M., Cox, A.M.G., Huesmann, M.: Optimal transport and Skorokhod embedding. Invent. Math. **208**(2), 327–400 (2017)
3. Beiglböck, M., Cox, A.M.G., Huesmann, M., Perkowski, N., Prömel, D.J.: Pathwise superreplication via Vovk's outer measure. Finance Stoch. **21**(4), 1141–1166 (2017)
4. Beiglböck, M., Henry-Labordère, P., Penkner, F.: Model-independent bounds for option prices: a mass transport approach. Finance Stoch. **17**(3), 477–501 (2013)
5. Beiglböck, M., Henry-Labordère, P., Touzi, N.: Monotone martingale transport plans and Skorokhod embedding. Stoch. Process. Appl. **127**, 3005–3013 (2017)
6. Beiglböck, M., Huesmann, M., Stebegg, F.: Root to Kellerer. In: Donati-Martin, C., Lejay, A., Rouault, A. (eds.) Séminaire de Probabilités XLVIII, pp. 1–12. Springer, Berlin (2016)
7. Beiglböck, M., Juillet, N.: On a problem of optimal transport under marginal martingale constraints. Ann. Probab. **44**(1), 42–106 (2016)
8. Beiglböck, M., Nutz, M., Touzi, N.: Complete duality for martingale optimal transport on the line. Ann. Probab. **45**(5), 3038–3074 (2017)
9. Beiglboeck, M., Lim, T., Obłój, J.: Dual attainment for the martingale transport problem. Bernoulli **25**, 1640–1658 (2018)
10. Blackwell, D., Dubins, L.: A converse to the dominated convergence theorem. Ill. J. Math. **7**, 508–514 (1963)
11. Bouchard, B., Nutz, M.: Arbitrage and duality in nondominated discrete-time models. Ann. Appl. Probab. **25**(2), 823–859 (2015)
12. Brown, H., Hobson, D., Rogers, L.C.G.: The maximum maximum of a martingale constrained by an intermediate law. Probab. Theory Relat. Fields **119**(4), 558–578 (2001)
13. Campi, L., Laachir, I., Martini, C.: Change of numeraire in the two-marginals martingale transport problem. Finance Stoch. **21**(2), 471–486 (2017)
14. Carlier, G., Ekeland, I.: Matching for teams. Econ. Theory **42**(2), 397–418 (2010)
15. Claisse, J., Guo, G., Henry-Labordere, P.: Robust hedging of options on local time. (2015). arXiv:1511.07230
16. Cox, A.M.G., Wang, J.: Optimal robust bounds for variance options. arXiv:1308.4363
17. Cox, A.M.G., Wang, J.: Root's barrier: construction, optimality and applications to variance options. Ann. Appl. Probab. **23**(3), 859–894 (2013)
18. Cox, A.M.G., Obłój, J., Touzi, N.: The Root solution to the multi-marginal embedding problem: an optimal stopping and time-reversal approach. Probab. Theory Relat. Fields **173**, 211–259 (2018)
19. De Angelis, T.: From optimal stopping boundaries to Rost's reversed barriers and the Skorokhod embedding. Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **54**(2), 1098–1133 (2018)

20. Dellacherie, C., Meyer, P.-A.: Probabilities and potential, volume 29 of North-Holland Mathematics Studies. North-Holland Publishing Co., Amsterdam (1978)
21. Doléans, C.: Existence du processus croissant naturel associé à un potentiel de la classe (d). Probab. Theory Relat. Fields **9**(4), 309–314 (1968)
22. Dolinsky, Y., Soner, H.M.: Martingale optimal transport and robust hedging in continuous time. Probab. Theory Relat. Fields **160**(1–2), 391–427 (2014)
23. Dubins, L.E., Gilat, D.: On the distribution of maxima of martingales. Proc. Am. Math. Soc. **68**(3), 337–338 (1978)
24. Ekren, I., Soner, H.M.: Constrained optimal transport. Arch. Ration. Mech. Anal. **227**(3), 929–965 (2018)
25. Galichon, A., Henry-Labordère, P., Touzi, N.: A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. Ann. Appl. Probab. **24**(1), 312–336 (2014)
26. Gangbo, W., McCann, R.: The geometry of optimal transportation. Acta Math. **177**(2), 113–161 (1996)
27. Gassiat, P., Oberhauser, H., dos Reis, G.: Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs. Stoch. Process. Appl. **125**(12), 4601–4631 (2015)
28. Ghoussoub, N., Kim, Y.-H., Lim, T.: Structure of optimal martingale transport plans in general dimensions. Ann. Probab. **47**(1), 109–164 (2019)
29. Guo, G., Tan, X., Touzi, N.: Tightness and duality of martingale transport on the Skorokhod space. Stoch. Process. Appl. **127**(3), 927–956 (2017)
30. Henry-Labordère, P., Obłój, J., Spoida, P., Touzi, N.: The maximum maximum of a martingale with given *n* marginals. Ann. Appl. Probab. **26**(1), 1–44 (2016)
31. Henry-Labordère, P., Tan, X., Touzi, N.: An explicit martingale version of the one-dimensional Brenier's Theorem with full marginals constraint. Stoch. Process. Appl. **126**(9), 2800–2834 (2016)
32. Henry-Labordère, P., Touzi, N.: An explicit martingale version of the one-dimensional brenier theorem. Finance Stoch. **20**(3), 635–668 (2016)
33. Hirsch, F., Profeta, C., Roynette, B., Yor, M.: Peacocks and Associated Martingales, with Explicit Constructions, Volume 3 of Bocconi & Springer Series. Springer, Milan; Bocconi University Press, Milan (2011)
34. Hobson, D.: The maximum maximum of a martingale. In: Séminaire de Probabilités, XXXII, Volume 1686 of Lecture Notes in Mathematics, pp. 250–263. Springer, Berlin (1998)
35. Hobson, D.: Robust hedging of the lookback option. Finance Stoch. **2**, 329–347 (1998)
36. Hobson, D.: The Skorokhod embedding problem and model-independent bounds for option prices. In: Paris–Princeton Lectures on Mathematical Finance 2010, Volume 2003 of Lecture Notes in Mathematics, pp. 267–318. Springer, Berlin (2011)
37. Hobson, D.: Mimicking martingales. Ann. Appl. Probab. **26**(4), 2273–2303 (2016)
38. Hobson, D., Neuberger, A.: Robust bounds for forward start options. Math. Finance **22**(1), 31–56 (2012)
39. Hobson, D., Pedersen, J.L.: The minimum maximum of a continuous martingale with given initial and terminal laws. Ann. Probab. **30**(2), 978–999 (2002)
40. Hou, Z., Obłój, J.: Robust pricing-hedging dualities in continuous time. Finance Stoch. **22**(3), 511–567 (2018)
41. Huesmann, M., Stebegg, F.: Monotonicity preserving transformations of MOT and SEP. Stoch. Processes Appl. **128**, 1114–1134 (2017)
42. Juillet, N.: Martingales associated to peacocks using the curtain coupling. Electron. J. Prob. **23** (2018)
43. Källblad, S., Tan, X., Touzi, N.: Optimal Skorokhod embedding given full marginals and Azéma–Yor peacocks. Ann. Appl. Probab. **27**(2), 686–719 (2017)
44. Kertz, R., Rösler, U.: Martingales with given maxima and terminal distributions. Israel J. Math. **69**(2), 173–192 (1990)
45. Kim, Y.-H., Pass, B.: Wasserstein barycenters over Riemannian manifolds. Adv. Math. **307**, 640–683 (2017)
46. Loynes, R.M.: Stopping times on Brownian motion: some properties of Root's construction. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **16**, 211–218 (1970)
47. Madan, D.B., Yor, M.: Making Markov martingales meet marginals: with explicit constructions. Bernoulli **8**(4), 509–536 (2002)
48. McConnell, T.R.: The two-sided Stefan problem with a spatially dependent latent heat. Trans. Am. Math. Soc. **326**(2), 669–699 (1991)
49. Nutz, M., Stebegg, F., Tan, X.: Multiperiod martingale transport (2017). arXiv:1703.10588

50. Obłój, J.: The Skorokhod embedding problem and its offspring. Probab. Surv. **1**, 321–390 (2004)
51. Obłój, J., Spoida, P.: An iterated Azéma–Yor type embedding for finitely many marginals. Ann. Probab. **45**(4), 2210–2247 (2017)
52. Obłój, J., Spoida, P., Touzi, N.: Martingale inequalities for the maximum via pathwise arguments. In: In memoriam Marc Yor—Séminaire de Probabilités XLVII, Volume 2137 of Lecture Notes in Math., pp. 227–247. Springer (2015)
53. Pass, B.: Multi-marginal optimal transport: theory and applications. ESAIM Math. Model. Numer. Anal. **49**(6), 1771–1790 (2015)
54. Perkins, E.: The Cereteli–Davis solution to the H^1-embedding problem and an optimal embedding in Brownian motion. In: Seminar on stochastic processes, 1985, pp. 172–223. Springer (1986)
55. Rogers, L.C.G.: The joint law of the maximum and terminal value of a martingale. Probab. Theory Relat. Fields **95**(4), 451–466 (1993)
56. Root, D.H.: The existence of certain stopping times on Brownian motion. Ann. Math. Statist. **40**, 715–718 (1969)
57. Rost, H.: Skorokhod stopping times of minimal variance. In: Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975), Lecture Notes in Mathematics, Vol. 511, pp. 194–208. Springer, Berlin (1976)
58. Skorohod, A.V.: Issledovaniya po teorii sluchainykh protsessov (Stokhasticheskie differentsialnye uravneniya i predelnye teoremy dlya protsessov Markova). Izdat. Kiev. Univ., Kiev (1961)
59. Skorokhod, A.V.: Studies in the Theory of Random Processes. Translated from the Russian by Scripta Technica, Inc., Addison-Wesley Publishing Co. Inc, Reading, MA (1965)
60. Zaev, D.A.: On the Monge–Kantorovich problem with additional linear constraints. Mat. Zametki **98**(5), 664–683 (2015)

Affiliations

Mathias Beiglböck[1] · Alexander M. G. Cox[2] · **Martin Huesmann[1]**

Mathias Beiglböck
mathias.beiglboeck@univie.ac.at

Martin Huesmann
martin.huesmann@univie.ac.at

[1]   Universität Wien, Vienna, Austria

[2]   University of Bath, Bath, UK