



Peacocks nearby: Approximating sequences of measures

Stefan Gerhold*, I. Cetin Gülüm

TU Wien, Austria

Received 8 January 2016; received in revised form 2 October 2017; accepted 4 July 2018

Available online 19 July 2018

Abstract

A peacock is a family of probability measures with finite mean that increases in convex order. It is a classical result, in the discrete time case due to Strassen, that any peacock is the family of one-dimensional marginals of a martingale. We study the problem whether a given sequence of probability measures can be approximated by a peacock. In our main results, the approximation quality is measured by the infinity Wasserstein distance. Existence of a peacock within a prescribed distance is reduced to a countable collection of rather explicit conditions. This result has a financial application (developed in a separate paper), as it allows to check European call option quotes for consistency. The distance bound on the peacock then takes the role of a bound on the bid–ask spread of the underlying. We also solve the approximation problem for the stop-loss distance, the Lévy distance, and the Prokhorov distance.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

A celebrated result, first proved by Strassen in 1965,¹ states that, for a given sequence of probability measures $(\mu_n)_{n \in \mathbb{N}}$, there exists a martingale $M = (M_n)_{n \in \mathbb{N}}$ such that the law of M_n is μ_n for all n , if and only if all μ_n have finite mean and $(\mu_n)_{n \in \mathbb{N}}$ is increasing in convex order (see [Definition 2.1](#)). Such sequences, and their continuous time counterparts, are nowadays

* Corresponding author.

E-mail addresses: sgerhold@fam.tuwien.ac.at (S. Gerhold), ismail.cetin.gueluem@gmx.at (I.C. Gülüm).

¹ See Theorem 8 in [36]. (Another result from that paper, relative to the usual stochastic order instead of the convex order, is also sometimes referred to as Strassen's theorem; see [23].)

referred to as peacocks, a pun on the French acronym PCOC, for “Processus Croissant pour l’Ordre Convexe” [15]. For further references on Strassen’s theorem and its predecessors, see the appendix of [6], p. 380 of Dellacherie and Meyer [8], and [1]. A constructive proof, and references to earlier constructive proofs, are given in Müller and Rüschendorf [27].

The theorem gave rise to plenty of generalizations, one of the most famous being Kellerer’s theorem [19,20]. It states that, for a peacock $(\mu_t)_{t \geq 0}$ with index set \mathbb{R}^+ , there is a Markov martingale $M = (M_t)_{t \geq 0}$ such that $M_t \sim \mu_t$ for all $t \geq 0$. Several proofs and ramifications of Kellerer’s theorem can be found in the literature. Hirsch and Roynette [16] construct martingales as solutions of stochastic differential equations and use an approximation argument. Lowther [25,26] shows that under some regularity assumptions there exists an ACD martingale with marginals $(\mu_t)_{t \geq 0}$. Here, ACD stands for “almost-continuous diffusion”, a condition implying the strong Markov property and stochastic continuity. Beiglböck, Huesmann and Stebegg [2] use a certain solution of the Skorokhod problem, which is Lipschitz–Markov, to construct a martingale which is Markov. The recent book by Hirsch, Profeta, Roynette, and Yor [15] contains a wealth of constructions of peacocks and associated martingales.

The main question that we consider in this paper is the following: given $\epsilon > 0$, a metric d on \mathcal{M} – the set of all probability measures on \mathbb{R} with finite mean – and a sequence of measures $(\mu_t)_{t \in T}$ in \mathcal{M} , when does a sequence $(\nu_t)_{t \in T}$ in \mathcal{M} exist, such that $d(\mu_t, \nu_t) \leq \epsilon$ and such that the sequence $(\nu_t)_{t \in T}$ is a peacock? Here T is either \mathbb{N} or the interval $[0, 1]$. Once we have constructed a peacock, we know, from the results mentioned above, that there is a martingale (with certain properties) with these marginals. We thus want to find out when there is a martingale M such that the law of M_t is close to μ_t for all t . We will state necessary and sufficient conditions when d is the infinity Wasserstein distance, the stop-loss distance, the Prokhorov distance, and the Lévy distance.

The infinity Wasserstein distance is a natural analogue of the p -Wasserstein distance. Besides the dedicated probability metrics literature (e.g., [32,33]), it has been studied in an optimal transport setting in [5]. It also has applications in graph theory, where it is referred to as the bottleneck distance (see p. 216 of [10]). A well-known alternative representation of the infinity Wasserstein distance shows some similarity to the Lévy distance. The stop-loss distance was introduced by Gerber in [11] and has been studied in actuarial science (see for instance [7,18]).

For both of these metrics, we translate existence of a peacock within ϵ -distance into a more tractable condition: There has to exist a real number (with the interpretation of the desired peacock’s mean) that satisfies a countable collection of finite-dimensional conditions, each explicitly expressed in terms of the call functions $x \mapsto \int (y-x)^+ \mu_t(dy)$ of the given sequence of measures. For the infinity Wasserstein distance, the existence proof is not constructive, as it uses Zorn’s lemma. For the stop-loss distance, the problem is much simpler, and our proof is short and constructive. Note, though, that the result about the infinity Wasserstein distance admits a financial application, which was the initial motivation for this work. The problem is similar to the one considered by Davis and Hobson [6]: given a set of European call option prices with different maturities on the same underlying, we want to know when there is a model which is consistent with these prices. In contrast to Davis and Hobson we allow a bid–ask spread, bounded by some constant, on the underlying. This application will be developed in the companion paper [12].

Our proof approach is similar for both metrics: we will construct minimal and maximal elements (with respect to the convex order) in closed balls, and then use these elements to derive our conditions. In the case of the infinity Wasserstein distance, we will make use of the lattice structure of certain subsets of closed balls.

The Lévy distance was first introduced by Lévy in 1925 (see [22]). Its importance is partially due to the fact that d^L metrizes weak convergence of measures on \mathbb{R} . The Prokhorov distance,

first introduced in [31], is a metric on measures on an arbitrary separable metric space, and is often referred to as a generalization of the Lévy metric, since d^p metrizes weak convergence on any separable metric space. For these two metrics, peacocks within ϵ -distance *always* exist, and can be explicitly constructed.

The definition of the infinity Wasserstein distance yields a coupling representation, and it is a natural question whether – assuming the existence of a peacock nearby – there is a filtered probability space on which the coupling can be realized with a martingale. For finite sequences of measures with finite support, we show in [Theorem 9.2](#) that this is true, thus extending (a special case of) Strassen’s theorem in a novel direction.

The structure of the paper is as follows. Section 2 specifies our notation and introduces the most important definitions. Section 3 contains our main result on approximation by peacocks using the infinity Wasserstein distance. Its proof is given in Section 4, and a continuous time version can be found in Section 5. In Section 6 we will treat the approximation problem for the stop-loss distance. After collecting some well-known facts on the Lévy and Prokhorov distances in Section 7, we will prove a criterion for approximation by peacocks under these metrics in Section 8. Section 9 presents our novel extension of (a special case of) Strassen’s theorem, and some related open problems that we propose to tackle in future work.

2. Notation and preliminaries

Let \mathcal{M} denote the set of all probability measures on \mathbb{R} with finite mean. We start with the definition of convex order.

Definition 2.1. Let μ, ν be two measures in \mathcal{M} . Then we say that μ is smaller in convex order than ν , in symbols $\mu \leq_c \nu$, if for every convex function $\phi : \mathbb{R} \rightarrow \mathbb{R}$ we have $\int \phi d\mu \leq \int \phi d\nu$, whenever both integrals are finite.² A family of measures $(\mu_t)_{t \in T}$ in \mathcal{M} , where $T \subseteq [0, \infty)$, is called peacock, if $\mu_s \leq_c \mu_t$ for all $s \leq t$ in T (see Definition 1.3 in [15]).

Intuitively, $\mu \leq_c \nu$ means that ν is more dispersed than μ , as convex integrands tend to emphasize the tails. By choosing $\phi(x) = x$ resp. $\phi(x) = -x$, we see that $\mu \leq_c \nu$ implies that μ and ν have the same mean. As mentioned in the introduction, Strassen’s theorem asserts the following:

Theorem 2.2 (Strassen [36]). *For any peacock with $T = \mathbb{N}$, there is a martingale whose family of one-dimensional marginal laws coincides with it.*

The converse implication is of course true as well, as a trivial consequence of Jensen’s inequality. As mentioned in the introduction, the equivalence also holds for time index set $T = \mathbb{R}^+$ [16,19,20]. For $\mu \in \mathcal{M}$ and $x \in \mathbb{R}$ we define

$$R_\mu(x) = \int_{\mathbb{R}} (y - x)^+ \mu(dy) \quad \text{and} \quad F_\mu(x) = \mu((-\infty, x]).$$

² The apparently stronger requirement that the inequality $\int \phi d\mu \leq \int \phi d\nu$ holds for convex ϕ whenever it makes sense, i.e., as long as both sides exist in $(-\infty, \infty]$, leads to an equivalent definition. This can be seen by the following argument, similar to Remark 1.1 in [15]: Assume that the inequality holds if both sides are finite, and let ϕ (convex) be such that $\int \phi d\mu = \infty$. We have to show that then $\int \phi d\nu = \infty$. Since ϕ is the envelope of the affine functions it dominates, we can find convex ϕ_n with $\phi_n \uparrow \phi$ pointwise, and such that each ϕ_n is C^2 and ϕ_n'' has compact support. By monotone convergence, we then have $\int \phi d\nu = \lim \int \phi_n d\nu \geq \lim \int \phi_n d\mu = \int \phi d\mu = \infty$. Note that the convexity of ϕ guarantees that $\int \phi d\nu > -\infty$.

We call R_μ the call function of μ , as in financial terms it is the (undiscounted) price of a call option with strike x , written on an underlying with risk-neutral law μ at maturity. (It is also known as integrated survival function [27].) The mean of a measure μ will be denoted by $\mathbb{E}\mu = \int y \mu(dy)$. The following proposition summarizes important properties of call functions. All of them are well known. In particular, the equivalence in part (iv) has been used a lot to investigate the convex order; see, e.g., [29].

Proposition 2.3. *Let μ, ν be two measures in \mathcal{M} . Then:*

- (i) R_μ is convex, decreasing and strictly decreasing on $\{R_\mu > 0\}$. Hence the right derivative of R_μ always exists and is denoted by R'_μ .
- (ii) $\lim_{x \rightarrow \infty} R_\mu(x) = 0$ and $\lim_{x \rightarrow -\infty} (R_\mu(x) + x) = \mathbb{E}\mu$. In particular, if $\mu([a, \infty)) = 1$ for $a > -\infty$, then $\mathbb{E}\mu = R_\mu(a) + a$.
- (iii) $R'_\mu(x) = -1 + F_\mu(x)$ and $R_\mu(x) = \int_x^\infty (1 - F_\mu(y)) dy$, for all $x \in \mathbb{R}$.
- (iv) $\mu \leq_c \nu$ holds if and only if $R_\mu(x) \leq R_\nu(x)$ for all $x \in \mathbb{R}$ and $\mathbb{E}\mu = \mathbb{E}\nu$.
- (v) For $x_1 \leq x_2 \in \mathbb{R}$, we have $R_\mu(x_2) - R_\mu(x_1) = \int_{x_1}^{x_2} R'_\mu(y) dy$.

Conversely, if a function $R : \mathbb{R} \rightarrow \mathbb{R}$ satisfies (i) and (ii), then there exists a probability measure $\mu \in \mathcal{M}$ with finite mean such that $R_\mu = R$.

As for (v), note that R'_μ is increasing, thus integrable, and that the fundamental theorem of calculus holds for right derivatives. See [4] for a short proof. The other assertions of Proposition 2.3 are proved in [16], Proposition 2.1, and [15], Exercise 1.7. For a metric d on \mathcal{M} , denote by $B^d(\mu, \epsilon)$ the closed ball with respect to d , with center μ and radius ϵ . Then our main question is:

Problem 2.4. Given $\epsilon > 0$, a metric d on \mathcal{M} , and a sequence $(\mu_n)_{n \in \mathbb{N}}$ in \mathcal{M} , when does there exist a peacock $(\nu_n)_{n \in \mathbb{N}}$ with $\nu_n \in B^d(\mu_n, \epsilon)$ for all n ?

Note that this can also be phrased as

$$d_\infty((\mu_n)_{n \in \mathbb{N}}, (\nu_n)_{n \in \mathbb{N}}) \leq \epsilon,$$

where

$$d_\infty((\mu_n)_{n \in \mathbb{N}}, (\nu_n)_{n \in \mathbb{N}}) = \sup_{n \in \mathbb{N}} d(\mu_n, \nu_n)$$

defines a metric on $\mathcal{M}^{\mathbb{N}}$ (with possible value infinity). For some results on this kind of infinite product metric, we refer to [3]. Clearly, a solution to Problem 2.4 settles the case of finite sequences $(\mu_n)_{n=1, \dots, n_0}$, too, by extending the sequence with $\mu_n := \mu_{n_0}$ for $n > n_0$.

To fix ideas, consider the case where the given sequence $(\mu_n)_{n=1,2}$ has only two elements. We want to find measures $\nu_n \in B^d(\mu_n, \epsilon)$, $n = 1, 2$, such that $\nu_1 \leq_c \nu_2$. Intuitively, we want ν_1 to be as small as possible and ν_2 to be as large as possible, in the convex order. Recall that a peacock has constant mean, which is fixed as soon as ν_1 is chosen. We will denote the set of probability measures on \mathbb{R} with mean $m \in \mathbb{R}$ by \mathcal{M}_m . These considerations lead us to the following problem.

Problem 2.5. Suppose that a metric d on \mathcal{M} , a measure $\mu \in \mathcal{M}$ and two numbers $\epsilon > 0$ and $m \in \mathbb{R}$ are given. When are there two measures $\mu^{\min}, \mu^{\max} \in B^d(\mu, \epsilon) \cap \mathcal{M}_m$ such that

$$\mu^{\min} \leq_c \nu \leq_c \mu^{\max}, \quad \text{for all } \nu \in B^d(\mu, \epsilon) \cap \mathcal{M}_m ?$$

We now recall the definition of the infinity Wasserstein distance³ W^∞ , and its connection to call functions.

Definition 2.6. The infinity Wasserstein distance is the mapping $W^\infty : \mathcal{M} \times \mathcal{M} \rightarrow [0, \infty]$ defined by

$$W^\infty(\mu, \nu) = \inf \|X - Y\|_\infty,$$

where the infimum is taken over all probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and random pairs (X, Y) with marginals given by μ and ν .

For various other probability metrics and their relations, see [13,32]. We will use the words “metric” and “distance” for mappings $\mathcal{M} \times \mathcal{M} \rightarrow [0, \infty]$ in a loose sense. Since all our results concern *concrete* metrics, there is no need to give a general definition (as, e.g., Definition 1 in Zolotarev [37]). The metric W^∞ has the following representation in terms of call functions (see, e.g., [24], p. 127):

$$W^\infty(\mu, \nu) = \inf \left\{ h > 0 : R'_\mu(x - h) \leq R'_\nu(x) \leq R'_\mu(x + h), \forall x \in \mathbb{R} \right\}. \tag{2.1}$$

By (2.1) and Proposition 2.3(iii), W^∞ can also be written as

$$W^\infty(\mu, \nu) = \inf \left\{ h > 0 : F_\mu(x - h) \leq F_\nu(x) \leq F_\mu(x + h), \forall x \in \mathbb{R} \right\}.$$

We will see below (Proposition 3.2) that, when d is the infinity Wasserstein distance, Problem 2.5 has a solution (μ^{\min}, μ^{\max}) if and only if $|m - \mathbb{E}\mu| \leq \epsilon$. As an easy consequence, given $(\mu_n)_{n=1,2}$, the desired “close” peacock $(\nu_n)_{n=1,2}$ exists if and only if there is an m with $|m - \mathbb{E}\mu_1| \leq \epsilon$, $|m - \mathbb{E}\mu_2| \leq \epsilon$ such that the corresponding measures $\mu_1^{\min}, \mu_2^{\max}$ satisfy $\mu_1^{\min} \leq_c \mu_2^{\max}$. Then, $(\nu_1, \nu_2) = (\mu_1^{\min}, \mu_2^{\max})$ is a possible choice.

Besides the infinity Wasserstein distance, we will solve Problems 2.4 and 2.5 also for the stop-loss distance (Proposition 6.1), for index sets \mathbb{N} and $[0, 1]$ (see Theorems 3.5, 5.1, 6.3, and 6.5). For the Lévy distance and the Prokhorov distance we will use different techniques and solve Problem 2.4 for index set \mathbb{N} (see Corollary 8.4 and Theorem 8.5).

3. Approximation by peacocks: infinity Wasserstein distance (discrete time)

We now start to investigate the interplay between the infinity Wasserstein distance and the convex order. Recall that \mathcal{M}_m denotes the set of probability measures on \mathbb{R} with mean m . It is a well known fact that the ordered set (\mathcal{M}_m, \leq_c) is a lattice for all $m \in \mathbb{R}$, with least element δ_m (Dirac delta). See for instance [21,28]. The lattice property means that, given any two measures $\mu, \nu \in \mathcal{M}_m$, there is a unique supremum, denoted by $\mu \vee \nu$, and a unique infimum, denoted by $\mu \wedge \nu$, with respect to convex order. It is easy to prove that the corresponding call functions are $R_{\mu \vee \nu} = R_\mu \vee R_\nu$ and $R_{\mu \wedge \nu} = \text{conv}(R_\mu, R_\nu)$. Here and in the following $\text{conv}(R_\mu, R_\nu)$ denotes the convex hull of R_μ and R_ν , i.e., the largest convex function that is majorized by $R_\mu \wedge R_\nu$.

In the following we will denote balls with respect to W^∞ by B^∞ . The next lemma shows that $(B^\infty(\mu, \epsilon) \cap \mathcal{M}_m, \leq_c)$ is a sublattice of (\mathcal{M}_m, \leq_c) , which will be important afterwards. Recall that two measures can be comparable w.r.t. convex order only if their means agree. This accounts for the relevance of sublattices of the form $(B^\infty(\mu, \epsilon) \cap \mathcal{M}_m, \leq_c)$ for our problem: If a peacock $(\nu_n)_{n \in \mathbb{N}}$ satisfying $\nu_n \in B^\infty(\mu_n, \epsilon)$ for all $n \in \mathbb{N}$ exists, then we necessarily have $\nu_n \in B^\infty(\mu_n, \epsilon) \cap \mathcal{M}_m, n \in \mathbb{N}$, with $\mathbb{E}\nu_1 = \mathbb{E}\nu_2 = \dots = m$.

³ The name “infinite Wasserstein distance” is also in use, but “infinity Wasserstein distance” seems to make more sense (cf. “infinity norm”).

Lemma 3.1. *Let $\epsilon > 0$ and $\mu, \nu_1, \nu_2 \in \mathcal{M}$ such that $\mathbb{E}\nu_1 = \mathbb{E}\nu_2 = m$. Then if $\nu_1, \nu_2 \in B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$ we have $\nu_1 \vee \nu_2 \in B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$ and $\nu_1 \wedge \nu_2 \in B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$.*

Proof. Denote the call functions of ν_1 and ν_2 with R_1 and R_2 . We start with $\nu_1 \vee \nu_2$. It is easy to check that $R : x \mapsto R_1(x) \vee R_2(x)$ is a call function satisfying $R'(x) \in \{R'_1(x), R'_2(x)\}$ for all $x \in \mathbb{R}$. By Proposition 2.3(ii), it is also clear that $\nu_1 \vee \nu_2 \in \mathcal{M}_m$. This proves the assertion.

As for the infimum, we will first assume that there exists $x_0 \in \mathbb{R}$ such that $R_1(x) \leq R_2(x)$ for $x \leq x_0$ and $R_2(x) \leq R_1(x)$ for $x \geq x_0$. Then there exist $x_1 \leq x_0$ and $x_2 \geq x_0$ such that the convex hull of R_1 and R_2 can be written as (see [30])

$$\text{conv}(R_1, R_2)(x) = \begin{cases} R_1(x), & x \leq x_1, \\ R_1(x_1) + \frac{R_2(x_2) - R_1(x_1)}{x_2 - x_1}(x - x_1), & x \in [x_1, x_2], \\ R_2(x), & x \geq x_2. \end{cases}$$

Now observe that for all $x \in [x_1, x_2]$

$$\begin{aligned} R'_\mu(x - \epsilon) &\leq R'_2(x) \leq R'_2(x_2 - \epsilon) \\ &\leq \frac{R_2(x_2) - R_1(x_1)}{x_2 - x_1} \\ &\leq R'_1(x_1) \leq R'_1(x) \leq R'_\mu(x + \epsilon), \end{aligned}$$

and hence $\text{conv}(R_1, R_2)'(x) \in [R'_\mu(x - \epsilon), R'_\mu(x + \epsilon)]$. Therefore $\nu_1 \wedge \nu_2 \in B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$.

For the general case, note that for all $x \in \mathbb{R}$ we have by [30] that either $\text{conv}(R_1, R_2)(x) = R_\mu(x) \wedge R_\nu(x)$, or that x lies in an interval I such that $\text{conv}(R_1, R_2)$ is affine on I . If the latter condition is the case, then we can derive bounds for the right-derivative $\text{conv}(R_1, R_2)'(x)$, $x \in I$, exactly as before. The situation is clear if either $\text{conv}(R_1, R_2)(x) = R_1(x)$ or $\text{conv}(R_1, R_2)(x) = R_2(x)$. \square

We now show that the sublattice $(B^\infty(\mu, \epsilon) \cap \mathcal{M}_m, \leq_c)$ contains a least and a greatest element with respect to the convex order. This is the subject of the following proposition, which solves Problem 2.5 for the infinity Wasserstein distance. As for the assumption $m \in [\mathbb{E}\mu - \epsilon, \mathbb{E}\mu + \epsilon]$ in Proposition 3.2, it is necessary to ensure that $B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$ is not empty. Indeed, if $W^\infty(\mu_1, \mu_2) \leq \epsilon$ for some $\mu_1, \mu_2 \in \mathcal{M}$, then by (2.1), Proposition 2.3(ii), (v), and the continuity of call functions, we obtain

$$R_{\mu_1}(x + \epsilon) \leq R_{\mu_2}(x) \leq R_{\mu_1}(x - \epsilon), \quad x \in \mathbb{R}. \tag{3.1}$$

By part (ii) of Proposition 2.3, it follows that $|\mathbb{E}\mu_1 - \mathbb{E}\mu_2| \leq \epsilon$.

Proposition 3.2. *Given $\epsilon > 0$, a measure $\mu \in \mathcal{M}$ and $m \in [\mathbb{E}\mu - \epsilon, \mathbb{E}\mu + \epsilon]$, there exist unique measures $S(\mu), T(\mu) \in B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$ such that*

$$S(\mu) \leq_c \nu \leq_c T(\mu) \quad \text{for all } \nu \in B^\infty(\mu, \epsilon) \cap \mathcal{M}_m.$$

The call functions of $S(\mu)$ and $T(\mu)$ are explicitly given by

$$R_\mu^{\min}(x) = R_{S(\mu)}(x) = \left(m + R_\mu(x - \epsilon) - (\mathbb{E}\mu + \epsilon)\right) \vee R_\mu(x + \epsilon), \tag{3.2}$$

$$R_\mu^{\max}(x) = R_{T(\mu)}(x) = \text{conv}\left(m + R_\mu(\cdot + \epsilon) - (\mathbb{E}\mu - \epsilon), R_\mu(\cdot - \epsilon)\right)(x). \tag{3.3}$$

To highlight the dependence on ϵ and m we will sometimes write $S(\mu; m, \epsilon)$ and $R_\mu^{\min}(\cdot; m, \epsilon)$, respectively $T(\mu; m, \epsilon)$ and $R_\mu^{\max}(\cdot; m, \epsilon)$.

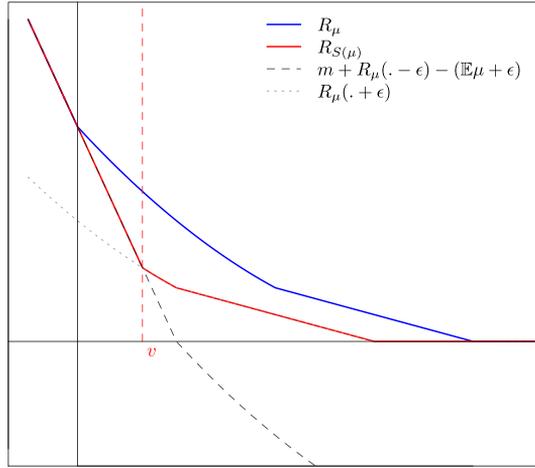


Fig. 1. Illustration of the function $R_\mu^{\min} = R_{S(\mu)}$ (lower solid curve): to the left of v , $R_{S(\mu)}$ is as steep as possible and to the right of v , $R_{S(\mu)}$ is as small as possible.

Proof. We define R_μ^{\min} and R_μ^{\max} by the right hand sides of (3.2) resp. (3.3), and argue that the associated measures $S(\mu)$ and $T(\mu)$ have the stated property. Clearly R_μ^{\min} is a call function, and we have

$$\begin{aligned} \mathbb{E}R_\mu^{\min} &= \lim_{x \rightarrow -\infty} \left(m + R_\mu(x - \epsilon) - (\mathbb{E}\mu + \epsilon) + x \right) \vee \left(R_\mu(x + \epsilon) + x \right) \\ &= m \vee (\mathbb{E}\mu - \epsilon) = m. \end{aligned}$$

From the convexity of R_μ we can deduce the existence of $v \in \mathbb{R} \cup \{\pm\infty\}$ such that

$$R_\mu^{\min}(x) = \begin{cases} m + R_\mu(x - \epsilon) - (\mathbb{E}R_\mu + \epsilon), & x \leq v, \\ R_\mu(x + \epsilon) & x \geq v. \end{cases}$$

Hence we get that $(R_\mu^{\min})'(x) \in [R'_\mu(x - \epsilon), R'_\mu(x + \epsilon)]$ for all x . By (2.1), the measure associated with R_μ^{\min} lies in $B^\infty(\mu, \epsilon) \cap \mathcal{M}_m$. To the left of v , R_μ^{\min} is as steep as possible (where steepness refers to the absolute value of the right derivative), and to the right of v it is as flat as possible (see Fig. 1). From this and convexity, it is easy to see that $S(\mu)$ is a least element.

Similarly we can show that $\mathbb{E}R_\mu^{\max} = m$, and thus it suffices to show that

$$(R_\mu^{\max})'(x) \in [R'_\mu(x - \epsilon), R'_\mu(x + \epsilon)].$$

But this can be done exactly as in Lemma 3.1. \square

Remark 3.3. It is not hard to show that

$$R_\mu^{\max}(x) = \begin{cases} m + R_\mu(x + \epsilon) - (\mathbb{E}\mu - \epsilon), & x \leq x_1, \\ R_\mu(x_1 + \epsilon) + \frac{(\mathbb{E}\mu - \epsilon) - m}{2\epsilon}(x - x_1 - 2\epsilon), & x \in [x_1, x_1 + 2\epsilon], \\ R_\mu(x - \epsilon), & x \geq x_1 + 2\epsilon, \end{cases}$$

where

$$x_1 = \inf \left\{ x \in \mathbb{R} : R'_\mu(x + \epsilon) \geq -\frac{m - (\mathbb{E}\mu - \epsilon)}{2\epsilon} \right\}.$$

Before formulating our first main theorem, we recall that a peacock is a sequence of probability measures with finite mean and increasing w.r.t. convex order (Definition 2.1). We now give a simple reformulation of this property. For a given sequence of call functions $(R_n)_{n \in \mathbb{N}}$, define, for $N \in \mathbb{N}$ and $x_1, \dots, x_N \in \mathbb{R}$,

$$\Phi_N(x_1, \dots, x_N) = R_1(x_1) + \sum_{n=2}^N \left(R_n(x_n) - R_n(x_{n-1}) \right) - R_{N+1}(x_N). \tag{3.4}$$

Proposition 3.4. *A sequence of call functions $(R_n)_{n \in \mathbb{N}}$ with constant mean defines a peacock if and only if $\Phi_N(x_1, \dots, x_N) \leq 0$ for all $N \in \mathbb{N}$ and $x_1, \dots, x_N \in \mathbb{R}$.*

Proof. According to Proposition 2.3(iv), we need to check whether the sequence of call functions increases. Let $n \in \mathbb{N}$ be arbitrary. If we set the n th component of (x_1, \dots, x_{n+1}) to an arbitrary $x \in \mathbb{R}$ and let all others tend to ∞ , we get

$$\Phi_{n+1}(\infty, \dots, \infty, x, \infty) = R_n(x) - R_{n+1}(x).$$

The sequence of call functions thus increases, if Φ is always non-positive. Conversely, assume that $(R_n)_{n \in \mathbb{N}}$ increases. Then, for $N \in \mathbb{N}$ and $x_1, \dots, x_N \in \mathbb{R}$,

$$\begin{aligned} \Phi_N(x_1, \dots, x_N) &\leq R_1(x_1) + \sum_{n=2}^N R_{n+1}(x_n) - \sum_{n=2}^N R_n(x_{n-1}) - R_{N+1}(x_N) \\ &= R_1(x_1) + \sum_{n=3}^{N+1} R_n(x_{n-1}) - \sum_{n=2}^N R_n(x_{n-1}) - R_{N+1}(x_N) \\ &= R_1(x_1) - R_2(x_1) \leq 0. \quad \square \end{aligned}$$

We now extend the definition of Φ_N for $x_1, \dots, x_N \in \mathbb{R}$, $m \in \mathbb{R}$, and $\epsilon > 0$ as follows, using the notation from Proposition 3.2:

$$\begin{aligned} \Phi_N(x_1, \dots, x_N; m, \epsilon) &= R_1^{\min}(x_1; m, \epsilon) \\ &\quad + \sum_{n=2}^N \left(R_n(x_n + \epsilon \sigma_n) - R_n(x_{n-1} + \epsilon \sigma_n) \right) \\ &\quad - R_{N+1}^{\max}(x_N; m, \epsilon). \end{aligned} \tag{3.5}$$

Here, R_1^{\min} is the call function of $S(\mu_1; m, \epsilon)$, R_{N+1}^{\max} is the call function of $T(\mu_{N+1}; m, \epsilon)$, and

$$\sigma_n = \text{sgn}(x_{n-1} - x_n) = \begin{cases} 1, & \text{if } x_{n-1} > x_n, \\ 0, & \text{if } x_{n-1} = x_n, \\ -1, & \text{if } x_{n-1} < x_n, \end{cases} \tag{3.6}$$

depends on x_{n-1} and x_n . Clearly, for $\epsilon = 0$ and $\mathbb{E}\mu_1 = \mathbb{E}\mu_2 = \dots = m$, we recover (3.4):

$$\Phi_N(x_1, \dots, x_N; m, 0) = \Phi_N(x_1, \dots, x_N), \quad N \in \mathbb{N}, \quad x_1, \dots, x_N \in \mathbb{R}. \tag{3.7}$$

The following theorem gives an equivalent condition for the existence of a peacock within W^∞ -distance ϵ of a given sequence of measures, thus solving Problem 2.4 for the infinity Wasserstein distance, and is our first main result. Note that the functions Φ_N defined in (3.5) have explicit expressions in terms of the given call functions, as R^{\min} and R^{\max} are explicitly given by (3.2) and (3.3). The existence criterion we obtain is thus rather explicit; the existence proof is not constructive, though, as mentioned in the introduction.

Theorem 3.5. Let $\epsilon > 0$ and $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{M} such that

$$I := \bigcap_{n \in \mathbb{N}} [\mathbb{E}\mu_n - \epsilon, \mathbb{E}\mu_n + \epsilon]$$

is not empty. Denote by $(R_n)_{n \in \mathbb{N}}$ the corresponding call functions, and define Φ_N by (3.5). Then there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ such that

$$W^\infty(\mu_n, \nu_n) \leq \epsilon, \quad \text{for all } n \in \mathbb{N}, \tag{3.8}$$

if and only if for some $m \in I$ and for all $N \in \mathbb{N}$, $x_1, \dots, x_N \in \mathbb{R}$, we have

$$\Phi_N(x_1, \dots, x_N; m, \epsilon) \leq 0. \tag{3.9}$$

In this case it is possible to choose $\mathbb{E}\nu_1 = \mathbb{E}\nu_2 = \dots = m$.

The proof of Theorem 3.5 is given in Section 4, building on Theorem 4.1 and Corollary 4.2 below. In view of our intended application (see [12]), we now give an alternative formulation of Theorem 3.5, which avoids the existential quantification “for some $m \in I$ ”. Note that the expressions inside the suprema in (3.10)–(3.12) are similar to Φ_N , defined in (3.5). Corollary 3.6 is proved towards the end of Section 4.

Corollary 3.6. Let $\epsilon > 0$ and $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{M} such that

$$I := \bigcap_{n \in \mathbb{N}} [\mathbb{E}\mu_n - \epsilon, \mathbb{E}\mu_n + \epsilon]$$

is not empty. Denote by $(R_n)_{n \in \mathbb{N}}$ the corresponding call functions. Then there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ such that (3.8) holds if and only if

$$\sup_{\substack{N_1 \in \mathbb{N} \\ x_1, \dots, x_{N_1} \in \mathbb{R}}} \left\{ R_1(x_1 + \epsilon) + \sum_{n=2}^{N_1} R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n) - R_{N_1+1}(x_{N_1} - \epsilon) \right\} \leq 0, \tag{3.10}$$

$$\sup_{\substack{N_1 \in \mathbb{N} \\ x_1, \dots, x_{N_1} \in \mathbb{R}}} \left\{ R_1(x_1 - \epsilon) + \sum_{n=2}^{N_1} R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n) - R_{N_1+1}(x_{N_1} + \epsilon) + \mathbb{E}\mu_{N_1+1} - \mathbb{E}\mu_1 \right\} \leq 2\epsilon, \tag{3.11}$$

$$\begin{aligned} &\sup_{\substack{N_1 \in \mathbb{N} \\ x_1, \dots, x_{N_1} \in \mathbb{R}}} \left\{ R_1(x_1 + \epsilon) + \sum_{n=2}^{N_1} R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n) - R_{N_1+1}(x_{N_1} + \epsilon) + \mathbb{E}\mu_{N_1+1} \right\} + \\ &\sup_{\substack{N_2 \in \mathbb{N} \\ y_1, \dots, y_{N_2} \in \mathbb{R}}} \left\{ R_1(y_1 - \epsilon) + \sum_{n=2}^{N_2} R_n(y_n + \epsilon\sigma_n) - R_n(y_{n-1} + \epsilon\sigma_n) - R_{N_2+1}(y_{N_2} - \epsilon) - \mathbb{E}\mu_1 \right\} \leq 2\epsilon. \end{aligned} \tag{3.12}$$

For $\epsilon = 0$, condition (3.9) is equivalent to the sequence of call functions (R_n) being increasing, see Proposition 3.4. For $\epsilon > 0$, analogously to the proof of Proposition 3.4, we see that

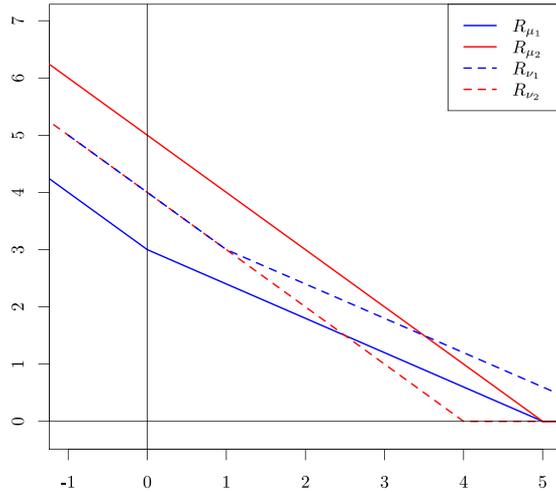


Fig. 2. The call functions of μ_1 (lower solid curve) and μ_2 (upper solid curve) from Example 3.7, for $m = 4$ and $\epsilon = 1$. The call function of v_1 is the call function of μ_1 shifted to the right by one. Similarly, shifting the call function of μ_2 by one to the left yields the call function of v_2 .

(3.9) implies

$$R_n(x + \epsilon) \leq R_{n+1}(x - \epsilon), \quad x \in \mathbb{R}, n \in \mathbb{N}. \tag{3.13}$$

It is clear that (3.13) is necessary for the existence of the peacock $(v_n)_{n \in \mathbb{N}}$, since, by (3.1) and Proposition 2.3(iv),

$$R_n(x + \epsilon) \leq R_{v_n}(x) \leq R_{v_{n+1}}(x) \leq R_{n+1}(x - \epsilon), \quad x \in \mathbb{R}, n \in \mathbb{N}.$$

On the other hand, it is easy to show that (3.13) is not sufficient for (3.9):

Example 3.7. Fix $m > 1$ and $\epsilon = 1$ and define two measures

$$\mu_1 = \frac{2}{m+1} \delta_0 + \frac{m-1}{m+1} \delta_{m+1}, \quad \mu_2 = \delta_{m+1},$$

where δ denotes the Dirac delta. It is simple to check that (3.13) is satisfied, i.e.

$$R_{\mu_1}(x + \epsilon) \leq R_{\mu_2}(x - \epsilon), \quad x \in \mathbb{R}.$$

Now assume that we want to construct a peacock $(v_n)_{n=1,2}$ such that $W^\infty(\mu_n, v_n) \leq 1$. Then the only possible mean for this peacock is m , since $\mathbb{E}\mu_1 = m - 1$ and $\mathbb{E}\mu_2 = m + 1$ (see the remark before Proposition 3.2). Therefore the peacock has to satisfy $v_n \in B^\infty(\mu_n, 1) \cap \mathcal{M}_m, n = 1, 2$, and the only possible choice is

$$v_1 = \frac{2}{m+1} \delta_1 + \frac{m-1}{m+1} \delta_{m+2}, \quad v_2 = \delta_m.$$

But since $R_{v_1}(x) > R_{v_2}(x)$ for $x \in (1, m + 2)$, $(v_n)_{n=1,2}$ is not a peacock; see Fig. 2.

If the sequence $(\mu_n)_{n=1,2}$ has just two elements, then it suffices to require (3.9) only for $N = 1$. It then simply states that there is an $m \in I$ such that $R_1^{\min}(x; m, \epsilon) \leq R_2^{\max}(x; m, \epsilon)$ for all x , which is clearly necessary and sufficient for the existence of $(v_n)_{n=1,2}$.

Example 3.8. Unsurprisingly, the peacock from [Theorem 3.5](#) is in general not unique: Let $\epsilon > 0$ and consider the constant sequences $R_n(x) = (-x)^+$, $n \in \mathbb{N}$, and

$$P_n(x, c) = \begin{cases} -x, & x \leq -\epsilon, \\ \epsilon - \frac{\epsilon(x + \epsilon)}{c + \epsilon}, & -\epsilon \leq x \leq c, \\ 0, & x \geq c. \end{cases}$$

Then, for any $c \in [0, \epsilon]$, it is easy to verify that the sequence of call functions $P_n(\cdot, c)$ defines a peacock satisfying [\(3.8\)](#).

4. Proof and ramifications of [Theorem 3.5](#)

The following theorem furnishes the main step for the induction proof of [Theorem 3.5](#), given at the end of the present section. In each induction step, the next element of the desired peacock should be contained in a certain ball, it should be larger in convex order than the previous element (ν in [Theorem 4.1](#)), and it should be as small as possible in order not to hamper the existence of the subsequent elements. This leads us to search for a least element of the set A_μ^ν defined in [\(4.1\)](#). The conditions defining this least element translate into inequalities on the corresponding call function. Part [\(ii\)](#) of [Theorem 4.1](#) states that, at each point of the real line, at least one of the latter conditions becomes an equality.

Theorem 4.1. *Let μ, ν be two measures in \mathcal{M} such that the set*

$$A_\mu^\nu := \left\{ \theta \in B^\infty(\mu; \epsilon) : \nu \leq_c \theta \right\} \tag{4.1}$$

is not empty.

(i) *The set A_μ^ν contains a least element $S_\nu(\mu)$ with respect to \leq_c , i.e. for every $\theta \in A_\mu^\nu$ we have*

$$\nu \leq_c S_\nu(\mu) \leq_c \theta.$$

Equivalently, if

$$R_\nu(x) \leq R_{T(\mu; \mathbb{E}\nu, \epsilon)}(x), \quad x \in \mathbb{R},$$

where $T(\mu)$ was defined in [\(3.3\)](#), there exists a pointwise smallest call function R^ which is greater than R_ν and satisfies*

$$(R^*)'(x) \in [R'_\mu(x - \epsilon), R'_\mu(x + \epsilon)] \text{ for all } x \in \mathbb{R}.$$

(ii) *The call function R^* is a solution of the following variational type inequality:*

$$\min \left\{ R^*(x) - R_\nu(x), (R^*)'(x) - R'_\mu(x - \epsilon), R'_\mu(x + \epsilon) - (R^*)'(x) \right\} = 0, \quad x \in \mathbb{R}. \tag{4.2}$$

Proof. The equivalence in [\(i\)](#) follows from [Proposition 2.3\(iv\)](#); note that the existence of $T(\mu; \mathbb{E}\nu, \epsilon)$ follows from $A_\mu^\nu \neq \emptyset$. We now argue that $S_\nu(\mu)$ exists. An easy application of Zorn’s lemma shows that there exist minimal elements in A_μ^ν . If θ_1 and θ_2 are two minimal elements of A_μ^ν then, according to [Lemma 3.1](#), the measure $\theta_1 \wedge \theta_2$ lies in $B^\infty(\nu, \epsilon) \cap \mathcal{M}_{\mathbb{E}\nu}$. Moreover, the convex function R_ν nowhere exceeds R_{θ_1} and R_{θ_2} , and hence we have $R_\nu \leq \text{conv}(R_{\theta_1} \wedge R_{\theta_2}) = R_{\theta_1 \wedge \theta_2}$. Therefore $\theta_1 \wedge \theta_2$ lies in A_μ^ν . Now clearly $\theta_1 \wedge \theta_2 \leq_c \theta_1$ and $\theta_1 \wedge \theta_2 \leq_c \theta_2$, and from the minimality we can conclude that $\theta_1 \wedge \theta_2 = \theta_1 = \theta_2$.

Now let θ^* be the unique minimal element and let $\theta \in A_\mu^v$ be arbitrary. Exactly as before we can show that $\theta^* \wedge \theta$ lies in A_μ^v . Moreover $\theta^* = \theta^* \wedge \theta \leq_c \theta$ and therefore θ^* is the least element of A_μ^v .

It remains to show (ii). We set

$$R^*(x) = \inf\{R_\theta(x) : \theta \in A_\mu^v\}. \tag{4.3}$$

Clearly R^* is a decreasing function with $\lim_{x \rightarrow \infty} R^*(x) = 0$ and $\lim_{x \rightarrow -\infty} R^*(x) + x = \mathbb{E}v$. We will show that R^* is convex, which is equivalent to the convexity of the epigraph \mathcal{E} of R^* . Pick two points $(x_1, y_1), (x_2, y_2) \in \mathcal{E}$. Then there exist measures $\theta_1, \theta_2 \in A_\mu^v$ such that $R_{\theta_1}(x_1) \leq y_1$ and $R_{\theta_2}(x_2) \leq y_2$. Using Lemma 3.1 once more, we get that $\theta := \theta_1 \wedge \theta_2 \in A_\mu^v$ and $R_\theta(x_i) \leq y_i, i = 1, 2$. Therefore, the whole segment with endpoints (x_1, y_1) and (x_2, y_2) lies in the epigraph of R_θ and hence in \mathcal{E} . This implies that R^* is a call function, and as we already know that A_μ^v has a least element $S_v(\mu)$, the measure associated to R^* has to be $S_v(\mu)$. Also, we can therefore conclude that the infimum in (4.3) is attained for all x .

Now assume that (4.2) is wrong. Since all functions appearing in (4.2) are right-continuous, there must then exist an open interval (a, b) where (4.2) does not hold, i.e. $R^*(x) > R_v(x)$ and $(R^*)'(x) \in (R'_\mu(x - \epsilon), R'_\mu(x + \epsilon))$ for all $x \in (a, b)$.

Case 1: There exists an open interval $I \subseteq (a, b)$ where R^* is strictly convex. Then we can pick $x_1 \in I$ and $h_1 > 0$ such that $x_1 + h_1 \in I$ and such that the tangent

$$P_1(x) := R^*(x_1) + (R^*)'(x_1)(x - x_1), \quad x \in [x_1, x_1 + h_1]$$

satisfies $R_v(x) < P_1(x) < R^*(x)$ for $x \in (x_1, x_1 + h_1]$. Also, since $(R^*)'(x_1) > R'_\mu(x_1 - \epsilon)$ and since R'_μ is right-continuous, we can choose h_1 small enough to guarantee $(R^*)'(x_1) \geq R'_\mu(x_1 + h_1 - \epsilon)$. Next pick $x_2 \in (x_1, x_1 + h_1)$, such that $R'_\mu(\cdot + \epsilon)$ is continuous at x_2 and set

$$P_2(x) := R^*(x_2) + (R^*)'(x_2)(x - x_2), \quad x \in [x_2 - h_2, x_2].$$

We can choose h_2 small enough to ensure that $R_v(x) < P_2(x) < R^*(x)$ and $(R^*)'(x_2) \leq R'_\mu(x_2 - h_2 + \epsilon)$. Also, if x_1 and x_2 are close enough together, then there is an intersection of P_1 and P_2 in (x_1, x_2) . Now the function

$$\tilde{R}(x) := \begin{cases} P_1(x) \vee P_2(x), & x \in [x_1, x_2], \\ R^*(x), & \text{otherwise,} \end{cases}$$

is a call function which is strictly smaller than R^* and satisfies $\tilde{R}'(x) \in [R'_\mu(x - \epsilon), R'_\mu(x + \epsilon)]$ for all $x \in \mathbb{R}$. This is a contradiction to (4.3). See Fig. 3 for an illustration.

Case 2: If there is no open interval in (a, b) where R^* is strictly convex, then R^* has to be affine on some closed interval $I \subseteq (a, b)$ (see p. 7 in [34]). Therefore, there exist k, d in \mathbb{R} such that

$$R^*(x) = kx + d, \quad x \in I.$$

By Proposition 2.3(ii), the slope k has to lie in the open interval $(-1, 0)$, since R^* is greater than R_v on I . We set

$$a_1 := \sup\{x \in \mathbb{R} : (R^*)'(x) < k\} > -\infty, \\ b_1 := \inf\{x \in \mathbb{R} : (R^*)'(x) > k\} < \infty;$$

the finiteness of these quantities follows from Proposition 2.3(ii). From the convexity of R_v and the fact that $R_v \leq R^*$, we get that $R^*(x) > R_v(x)$ for all $x \in (a_1, b_1)$, as well as $(R^*)'(x) > R'_\mu(x - \epsilon)$ for all $x \in (a_1, b)$ and $(R^*)'(x) < R'_\mu(x + \epsilon)$ for all $x \in (a, b_1)$. We

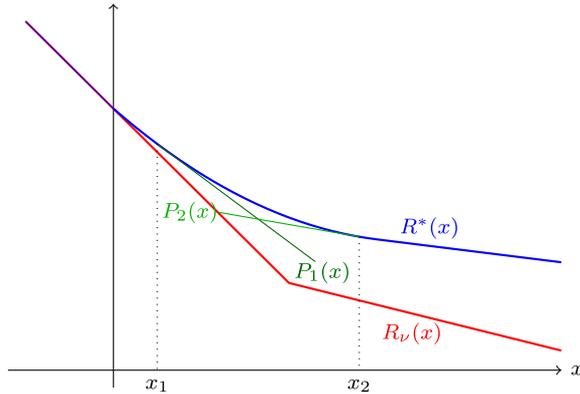


Fig. 3. Case 1 of the proof of [Theorem 4.1](#). If R^* is strictly convex, then we can deform it using two appropriate tangents, contradicting minimality of the associated measure.

now define lines P_1 and P_2 , with analogous roles as in Case 1. Their definitions depend on the behavior of $(R^*)'$ at a_1 and b_1 .

If $(R^*)'(a_1-) < k$, then we set $x_1 = a_1$ and $P_1(x) = R^*(x_1) + k_1(x - x_1)$ for $x \geq x_1$, with an arbitrary $k_1 \in ((R^*)'(x_1-), k)$; see [Fig. 4](#).

If, on the other hand, $(R^*)'(a_1-) = k$, then we can find $x_1 < a_1$ such that $R^*(x_1) > R_\nu(x_1)$ and $(R^*)'(x_1) > R'_\mu(x_1 - \epsilon)$. In this case we define

$$P_1(x) := R^*(x_1) + (R^*)'(x_1)(x - x_1), \quad x \geq x_1.$$

Similarly, if $(R^*)'(b_1) > k$, then we define $x_2 = b_1$ and $P_2(x) = R^*(x_2) + k_2(x - x_2)$ for $x \leq x_2$ and for $k_2 \in (k, (R^*)'(b_1))$, and otherwise we can find $x_2 > b_1$ such that $R^*(x_2) > R_\nu(x_2)$ and $(R^*)'(x_2) < R'_\mu(x_2 + \epsilon)$. We then set

$$P_2(x) := R^*(x_2) + (R^*)'(x_2)(x - x_2), \quad x \leq x_2.$$

We can choose $h_1, h_2 > 0, \tilde{d} < d$ and k_1, k_2 such that the function

$$\tilde{R}(x) := \begin{cases} P_1(x), & x \in [x_1, x_1 + h_1], \\ kx + \tilde{d}, & x \in [x_1 + h_1, x_2 - h_2] \\ P_2(x), & x \in [x_2 - h_2, x_2], \\ R^*(x), & \text{otherwise,} \end{cases}$$

is a call function which is strictly smaller than R^* but not smaller than R_ν . Also, if h_1 and h_2 are small enough we have $\tilde{R}'(x) \in [R'_\mu(x - \epsilon), R'_\mu(x + \epsilon)]$ for all $x \in \mathbb{R}$, which is a contradiction to [\(4.3\)](#). \square

In part (i) of [Theorem 4.1](#), we showed that A_μ^ν has a least element. The weaker statement that it has an infimum follows from [\[21\]](#), p. 162; there it is shown that any subset of the lattice (\mathcal{M}_m, \leq_c) has an infimum. (The stated requirement that the set be bounded from below is always satisfied, as the Dirac delta δ_m is the least element of (\mathcal{M}_m, \leq_c) .) This infimum is, of course, given by the least element $S_\nu(\mu)$ that we found.

If $\nu = \delta_m$, then $S_\nu(\mu) = S(\mu)$, the least element from [Proposition 3.2](#). In this case we have

$$(R^*)'(x) = \begin{cases} R'_\mu(x - \epsilon), & x < x^*, \\ R'_\mu(x + \epsilon), & x \geq x^*, \end{cases}$$

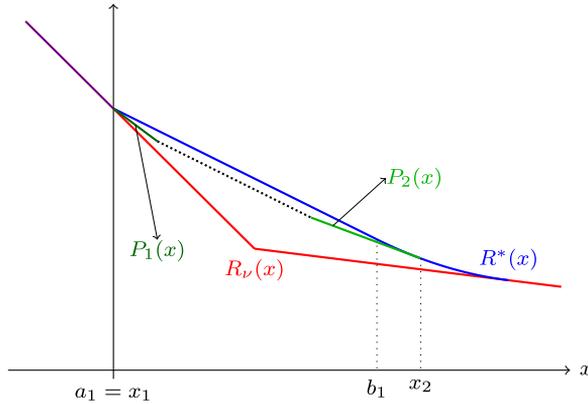


Fig. 4. Case 2 of the proof of Theorem 4.1, with $(R^*)'(a_1-) < k$ and $(R^*)'(b_1) = k$.

where x^* is the unique solution of

$$m + R_\mu(x - \epsilon) - (\mathbb{E}\mu + \epsilon) = R_\mu(x + \epsilon).$$

The following corollary establishes an alternative representation of the inequality (4.2), which we will use to prove Theorem 3.5. Note that, in general, (4.2) has more than one solution, not all of which are call functions. However, R^* is always a solution.

Corollary 4.2. Assume that the conditions from Theorem 4.1 hold and denote the call function of $S_\nu(\mu)$ by R^* . Then for all $x \in \mathbb{R}$ there exists $y \in \mathbb{R} \cup \{\pm\infty\}$ such that

$$R^*(x) = R_\nu(y) - R_\mu(y + \epsilon\sigma) + R_\mu(x + \epsilon\sigma),$$

where $\sigma = \text{sgn}(y - x)$. Here and in the following we set $R(\infty) = 0$ for all call functions R and

$$R_1(-\infty \pm \epsilon) - R_2(-\infty \pm \epsilon) := \lim_{x \rightarrow -\infty} (R_1(x \pm \epsilon) - R_2(x \pm \epsilon)),$$

for call functions R_1 and R_2 .

Proof. By Theorem 4.1 we know that R^* is a solution of (4.2). Let x be an arbitrary real number. If $R^*(x) = R_\nu(x)$, then the above relation clearly holds for $y = x$. Otherwise, we have $R^*(x) > R_\nu(x)$, and one of the other two expressions on the left hand side of (4.2) must vanish at x . First we assume that $(R^*)'(x) = R'_\mu(x + \epsilon)$. Define

$$y := \inf\{z \geq x : (R^*)'(z) < R'_\mu(z + \epsilon)\}.$$

If $y < \infty$, then by definition $(R^*)'(y) < R'_\mu(y + \epsilon)$. By (4.2), we have $R^*(y) = R_\nu(y)$. It follows that

$$\begin{aligned} R^*(z) &= R^*(y) + \int_y^z (R^*)'(x) dx = R_\nu(y) + \int_y^z R'_\mu(x + \epsilon) dx = \\ &= R_\nu(y) - R_\mu(y + \epsilon) + R_\mu(z + \epsilon), \quad \text{for all } z \in [x, y]. \end{aligned}$$

If $y = \infty$, then this equation, i.e. $R^*(z) = R_\mu(z + \epsilon)$, $z \geq x$, also holds.

If, on the other hand, $(R^*)'(x) = R'_\mu(x - \epsilon)$, then we similarly define $y := \sup\{z \leq x : (R^*)'(z) > R'_\mu(z - \epsilon)\}$. If $y > -\infty$ then $(R^*)'(y-) > R'_\mu((y - \epsilon)-)$ and hence $R^*(y) = R_\nu(y)$

by (4.2). Therefore we can write

$$R^*(z) = R_\nu(y) - R_\mu(y - \epsilon) + R_\mu(z - \epsilon), \quad \text{for all } z \in [y, x].$$

If $y = -\infty$ then $(R^*)'(z) = R'_\mu(z - \epsilon)$ for all $z \leq x$. The above equation holds if we take the limit $y \rightarrow \infty$ on the right hand side. \square

Corollary 4.3. Using Proposition 3.2 and Theorem 4.1, for a given sequence of measures $(\mu_n)_{n \in \mathbb{N}}$ in \mathcal{M} , we inductively define the measures

$$\theta_1 = S(\mu_1; m, \epsilon), \quad \theta_k = S_{\theta_{k-1}}(\mu_k), \quad k \geq 2,$$

if the sets

$$\{v \in B^\infty(\mu_k, \epsilon) : \theta_{k-1} \leq_c v\}$$

are not empty. Then the following relation holds:

$$R_{\theta_n}(x) = R_{\theta_{n-1}}(y) - R_{\mu_n}(y + \epsilon\sigma) + R_{\mu_n}(x + \epsilon\sigma),$$

where $n \geq 2$, $y \in \mathbb{R} \cup \{\pm\infty\}$ depends on x and $\sigma = \text{sgn}(y - x)$.

Proof. The result follows by simply applying Theorem 4.1 and Corollary 4.2 with $\nu = \theta_{n-1}$ and $\mu = \mu_n$. \square

The next corollary will be useful later on in Theorem 9.2 and is an easy consequence of (4.2).

Corollary 4.4. Let μ, ν be as in Theorem 4.1 and additionally assume that both measures have finite support. Then $S_\nu(\mu)$ has finite support too.

Proof. By (iii) of Proposition 2.3, the finiteness of the support of a measure θ is equivalent to R'_θ having a finite range. Therefore, we can partition the real line into a finite number of intervals I_1, \dots, I_N such that for all $n \in \{1, \dots, N\}$ the functions $R'_\nu, R'_\mu(\cdot - \epsilon)$ and $R'_\mu(\cdot + \epsilon)$ are constant on I_n . Since $R_{S_\nu(\mu)}$ solves (4.2), we can conclude that $R'_{S_\nu(\mu)}$ takes at most three distinct values on each I_n . Hence, $R'_{S_\nu(\mu)}$ is piecewise constant and $S_\nu(\mu)$ has finite support. \square

We can now prove Theorem 3.5, our main result on approximation by peacocks. We first prove the “if” direction, which, unsurprisingly, is the more difficult one.

Proof of Theorem 3.5. Suppose that (3.9) holds for some $m \in I$ and all $N \in \mathbb{N}, x_1, \dots, x_N \in \mathbb{R}$. We will inductively construct a sequence $(P_n)_{n \in \mathbb{N}}$ of call functions, which will correspond to the measures $(\nu_n)_{n \in \mathbb{N}}$. Define $P_1 = R_1^{\min}(\cdot; m, \epsilon)$. For $N = 1$, (3.9) guarantees that $R_1^{\min}(x) \leq R_2^{\max}(x)$. Note that the continuity of the R_n guarantee that (3.9) also holds for $x_n \in \{\pm\infty\}$, if we set $\text{sgn}(\infty - \infty) = \text{sgn}(-\infty + \infty) = 0$. We can now use Theorem 4.1 together with Corollary 4.2, with $R_\nu = R_1^{\min}$ and $R_\mu = R_2$, to construct a call function P_2 , which satisfies

$$P_2(x) = R_1^{\min}(x_1) + R_2(x + \epsilon\sigma) - R_2(x_1 + \epsilon\sigma), \quad x \in \mathbb{R},$$

where $\sigma = \text{sgn}(x_1 - x)$, and x_1 depends on x . If we use (3.9) we get that

$$R_1^{\min}(x_1) + R_2(x + \epsilon\sigma) - R_2(x_1 + \epsilon\sigma) \leq R_n^{\max}(x; m, \epsilon), \quad n \geq 3, \quad x_1, x \in \mathbb{R}.$$

Hence $P_2(x) \leq R_n^{\max}(x)$ for all $x \in \mathbb{R}$ and for all $n \geq 3$. Now suppose that we have already constructed a finite sequence (P_1, \dots, P_N) such that $P_n \leq P_{n+1}, 1 \leq n < N$, and such that

$P_N \leq R_n^{\max}$ for all $x \in \mathbb{R}$ and for all $n \geq N + 1$. Then by induction we know that for all $x \in \mathbb{R}$ there exists (x_1, \dots, x_{N-1}) such that

$$P_N(x) = R_1^{\min}(x_1) + \sum_{n=2}^{N-1} (R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n)) + R_N(x + \epsilon\sigma_N) - R_N(x_{N-1} + \epsilon\sigma_N),$$

with $\sigma_N = \text{sgn}(x_{N-1} - x)$. In particular, we have $P_N \leq R_{N+1}^{\max}$. We can therefore again use [Corollary 4.2](#), with $R_\mu = R_{N+1}$ and $R_\nu = P_N$, to construct a call function P_{N+1} , such that

$$P_{N+1}(x) = R_1^{\min}(x_1) + \sum_{n=2}^N (R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n)) + R_{N+1}(x + \epsilon\sigma_{N+1}) - R_{N+1}(x_N + \epsilon\sigma_{N+1}),$$

where $\sigma_{N+1} = \text{sgn}(x_N - x)$ and (x_1, \dots, x_N) depend on x . Assumption [\(3.9\)](#) guarantees that $P_{N+1} \leq R_n^{\max}$ for all $n \geq N + 1$.

We have now constructed a sequence of call functions, such that $P_n \leq P_{n+1}$. Their associated measures, which we will denote by ν_n , satisfy $W^\infty(\mu_n, \nu_n) \leq \epsilon$ and $\nu_n \leq_c \nu_{n+1}$. Thus we have constructed a peacock with mean m .

We proceed to the proof of the (easier) “only if” direction of [Theorem 3.5](#). Thus, assume that $(\nu_n)_{n \in \mathbb{N}}$ is a peacock such that $W^\infty(\mu_n, \nu_n) \leq \epsilon$ and set $m = \mathbb{E}\nu_1$. Denote the call function of ν_n by P_n . We will show by induction that [\(3.9\)](#) holds. For $N = 1$ we have

$$R_1^{\min}(x; m, \epsilon) \leq P_1(x) \leq P_2(x) \leq R_2^{\max}(x; m, \epsilon), \quad x \in \mathbb{R},$$

by [Proposition 3.2](#).

For $N = 2$ and $x_1 \leq x_2$ we have

$$\begin{aligned} R_1^{\min}(x_1; m, \epsilon) + R_2(x_2 - \epsilon) - R_2(x_1 - \epsilon) &\leq P_2(x_1) + \int_{x_1}^{x_2} R_2'(z - \epsilon) dz \\ &\leq P_2(x_1) + \int_{x_1}^{x_2} P_2'(z) dz \\ &= P_2(x_2) \leq P_3(x_2) \leq R_3^{\max}(x_2; m, \epsilon). \end{aligned}$$

Similarly, if $x_2 \leq x_1$,

$$\begin{aligned} R_1^{\min}(x_1; m, \epsilon) + R_2(x_2 + \epsilon) - R_2(x_1 + \epsilon) &\leq P_2(x_1) - \int_{x_2}^{x_1} R_2'(z + \epsilon) dz \\ &\leq P_2(x_1) - \int_{x_2}^{x_1} P_2'(z) dz \\ &= P_2(x_2) \leq P_3(x_2) \leq R_3^{\max}(x_2; m, \epsilon). \end{aligned}$$

If [\(3.9\)](#) holds for $N - 1$ and $x_{N-1} \leq x_N$, then

$$\begin{aligned} R_1^{\min}(x_1; m, \epsilon) + \sum_{n=2}^N (R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n)) \\ \leq P_{N-1}(x_{N-1}) + R_N(x_N - \epsilon) - R_N(x_{N-1} - \epsilon) \\ \leq P_N(x_{N-1}) + \int_{x_{N-1}}^{x_N} P_N'(z) dz \end{aligned}$$

$$\leq P_{N+1}(x_N) \leq R_{N+1}^{\max}(x_N; m, \epsilon).$$

The case where $x_{N-1} \geq x_N$ can be dealt with similarly. \square

Proof of Corollary 3.6. First, by going through the proof of [Theorem 3.5](#) a second time, we see that $R_{N+1}^{\max}(x_N; m, \epsilon)$, in the definition of Φ_N can be replaced by

$$\widetilde{R}_{N+1}^{\max}(x_N; m, \epsilon) := \left(m + R_{N+1}(x_N + \epsilon) - (\mathbb{E}\mu - \epsilon) \right) \wedge R_{N+1}(x_N - \epsilon),$$

which is $R_{N+1}^{\max}(x_N; m, \epsilon)$ without the convex envelope.

Next, we can split up [\(3.9\)](#) into four inequalities according to the different components of $\widetilde{R}_{N+1}^{\max}$ and R_1^{\min} . In two of these inequalities m does not appear, and these are exactly equations [\(3.10\)](#) and [\(3.11\)](#). The remaining two inequalities are given by

$$\begin{aligned} & R_1(x_1 - \epsilon) + m - (\mathbb{E}\mu_1 + \epsilon) + \sum_{n=2}^N R_n(x_n + \epsilon\sigma_n) \\ & \quad - R_n(x_{n-1} + \epsilon\sigma_n) - R_{N+1}(x_N + \epsilon) \leq 0, \\ & R_1(x_1 + \epsilon) + \sum_{n=2}^N R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n) \\ & \quad - \left(R_{N+1}(x_N + \epsilon) + m - (\mathbb{E}\mu_{N+1} - \epsilon) \right) \leq 0. \end{aligned}$$

In particular, m can only exist if

$$\begin{aligned} & \sup_{\substack{N_1 \in \mathbb{N} \\ x_1, \dots, x_{N_1} \in \mathbb{R}}} \left\{ R_1(x_1 + \epsilon) + \sum_{n=2}^{N_1} R_n(x_n + \epsilon\sigma_n) - R_n(x_{n-1} + \epsilon\sigma_n) \right. \\ & \quad \left. - R_{N_1+1}(x_{N_1} + \epsilon) + \mathbb{E}\mu_{N_1+1} \right\} - \epsilon \leq \\ & \inf_{\substack{N_2 \in \mathbb{N} \\ y_1, \dots, y_{N_2} \in \mathbb{R}}} \left\{ -R_1(y_1 - \epsilon) - \sum_{n=2}^{N_2} R_n(y_n + \epsilon\sigma_n) - R_n(y_{n-1} + \epsilon\sigma_n) \right. \\ & \quad \left. + R_{N_2+1}(y_{N_2} - \epsilon) + \mathbb{E}\mu_1 \right\} + \epsilon, \end{aligned} \tag{4.4}$$

in which case m can be chosen arbitrarily from the closed interval with bounds given by the left hand side resp. right hand side of [\(4.4\)](#). A simple modification of [\(4.4\)](#) yields [\(3.12\)](#). \square

Remark 4.5. In [Theorem 3.5](#), it is actually not necessary that the balls centered at the measures μ_n are all of the same size. The theorem easily generalizes to the following result: For $m \in \mathbb{R}$, a sequence of non-negative numbers $(\epsilon_n)_{n \in \mathbb{N}}$, and a sequence of measures $(\mu_n)_{n \in \mathbb{N}}$ in \mathcal{M} , define

$$\begin{aligned} \Phi_N(x_1, \dots, x_N; m, \epsilon_1, \dots, \epsilon_{N+1}) &= R_1^{\min}(x_1; m, \epsilon_1) \\ & \quad + \sum_{n=2}^N \left(R_n(x_n + \epsilon_n\sigma_n) - R_n(x_{n-1} + \epsilon_n\sigma_n) \right) \\ & \quad - R_{N+1}^{\max}(x_N; m, \epsilon_{N+1}), \\ & N \in \mathbb{N}, x_1, \dots, x_N \in \mathbb{R}, \end{aligned} \tag{4.5}$$

with σ_n defined in (3.6), and assume that

$$I := \bigcap_{n \in \mathbb{N}} [\mathbb{E}\mu_n - \epsilon_n, \mathbb{E}\mu_n + \epsilon_n]$$

is not empty. Then there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ such that

$$W^\infty(\mu_n, \nu_n) \leq \epsilon_n, \quad \text{for all } n \in \mathbb{N},$$

if and only if for some $m \in I$ and for all $N \in \mathbb{N}, x_1, \dots, x_N \in \mathbb{R}$, we have

$$\Phi_N(x_1, \dots, x_N; m, \epsilon_1, \dots, \epsilon_{N+1}) \leq 0.$$

To prove this result, simply replace ϵ by ϵ_n in the proof of Theorem 3.5.

Remark 4.6. If a probability metric is comparable with the infinity Wasserstein distance, then our Theorem 3.5 implies a corresponding result about that metric (but, of course, not an “if and only if” condition). Denote by W^p the p -Wasserstein distance ($p \geq 1$), defined by

$$W^p(\mu, \nu) = \inf \left(\mathbb{E}[|X - Y|^p] \right)^{1/p}, \quad \mu, \nu \in \mathcal{M}.$$

The infimum is taken over all probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and random pairs (X, Y) with marginals given by μ and ν . Clearly, we have that for all $\mu, \nu \in \mathcal{M}$ and $p \geq 1$

$$W^\infty(\mu, \nu) \geq W^p(\mu, \nu).$$

Hence, given a sequence $(\mu_n)_{n \in \mathbb{N}}$, (3.9) is a sufficient condition for the existence of a peacock $(\nu_n)_{n \in \mathbb{N}}$, such that $W^p(\mu_n, \nu_n) \leq \epsilon$ for all $n \in \mathbb{N}$. But since the balls with respect to W^p are in general strictly larger than the balls with respect to W^∞ , we cannot expect (3.9) to be necessary.

5. Approximation by peacocks: infinity Wasserstein distance (continuous time)

In this section we will formulate a version of Theorem 3.5 for continuous index sets. We generalize the definition of Φ_N from (3.5) as follows. For finite sets $\mathcal{T} = \{t_1, \dots, t_{N+1}\} \subseteq [0, 1]$ with $t_1 < t_2 < \dots < t_{N+1}$, we set

$$\begin{aligned} \Phi_{\mathcal{T}}(x_1, \dots, x_N; m, \epsilon) &= R_{t_1}^{\min}(x_1; m, \epsilon) \\ &\quad + \sum_{n=2}^N \left(R_{t_n}(x_n + \epsilon \sigma_n) - R_{t_{n-1}}(x_n + \epsilon \sigma_n) \right) \\ &\quad - R_{t_{N+1}}^{\max}(x_N; m, \epsilon). \end{aligned} \tag{5.1}$$

Here, $R_{t_1}^{\min}$ is the call function of $S(\mu_{t_1}; m, \epsilon)$, $R_{t_{N+1}}^{\max}$ is the call function of $T(\mu_{t_{N+1}}; m, \epsilon)$, and $\sigma_n = \text{sgn}(x_{n-1} - x_n)$ depends on x_{n-1} and x_n . Using $\Phi_{\mathcal{T}}$, we can now formulate a necessary and sufficient condition for the existence of a peacock within ϵ -distance. The continuity assumption (5.2) occurs in the proof in a natural way; we do not know to which extent it can be relaxed.

Theorem 5.1. Assume that $(\mu_t)_{t \in [0,1]}$ is a family of measures in \mathcal{M} such that

$$I := \bigcap_{t \in [0,1]} [\mathbb{E}\mu_t - \epsilon, \mathbb{E}\mu_t + \epsilon]$$

is not empty and such that

$$\lim_{s \uparrow t} \mu_s = \mu_t \text{ weakly, } t \in [0, 1]. \tag{5.2}$$

Then there exists a peacock $(\nu_t)_{t \in [0,1]}$ with

$$W^\infty(\mu_t, \nu_t) \leq \epsilon, \text{ for all } t \in [0, 1],$$

if and only if there exists $m \in I$ such that for all finite sets $\mathcal{T} = \{t_1, \dots, t_{N+1}\} \subset \mathbb{Q} \cap [0, 1]$ with $t_1 < t_2 < \dots < t_{N+1}$, and for all $x_1, \dots, x_N \in \mathbb{R}$ we have that

$$\Phi_{\mathcal{T}}(x_1, \dots, x_N; m, \epsilon) \leq 0. \tag{5.3}$$

In this case it is possible to choose $\mathbb{E}\nu_t = m$ for all $t \in [0, 1]$.

Proof. By Theorem 3.5, condition (5.3) is clearly necessary for the existence of such a peacock. In order to show that it is sufficient, fix $m \in I$ such that (5.3) holds. We will first construct ν_q for $q \in \mathcal{D}$, where

$$\mathcal{D} = \{a2^{-b} \in [0, 1] : a, b \in \mathbb{N}_0\}.$$

For $n \in \mathbb{N}$, define measures (recall the notation from Theorem 4.1)

$$\theta_0^{(n)} = \mu_0 \text{ and } \theta_k^{(n)} = S_{\theta_{k-1}^{(n)}}(\mu_{k2^{1-n}}), \quad 1 \leq k \leq 2^{n-1}.$$

Condition (5.3) guarantees that these measures exist. Obviously,

$$\theta_k^{(n)} \leq_c \theta_{k+1}^{(n)}, \quad n \in \mathbb{N}, \quad 0 \leq k < 2^{n-1}. \tag{5.4}$$

We show by induction on k that

$$\theta_k^{(n)} \leq_c \theta_{2k}^{(n+1)}, \quad n \in \mathbb{N}, \quad 0 \leq k \leq 2^{n-1}. \tag{5.5}$$

For $k = 1$, we have

$$\begin{aligned} \theta_1^{(n)} &= S_{\mu_0}(\mu_{2^{1-n}}) \\ &\leq_c S_{\theta_1^{(n+1)}}(\mu_{2^{1-n}}) = \theta_2^{(n+1)}, \quad n \in \mathbb{N}. \end{aligned}$$

For $k \geq 1$, we obtain

$$\begin{aligned} \theta_{k+1}^{(n)} &= S_{\theta_k^{(n)}}(\mu_{(k+1)2^{1-n}}) \\ &\leq_c S_{\theta_{2k}^{(n+1)}}(\mu_{(k+1)2^{1-n}}) \\ &\leq_c S_{\theta_{2k+1}^{(n+1)}}(\mu_{(k+1)2^{1-n}}) = \theta_{2k+2}^{(n+1)}, \quad n \in \mathbb{N}, \end{aligned}$$

where the first “ \leq_c ” follows from the induction hypothesis and the definition of $S(\cdot)$. Thus, (5.5) is true.

For $q = a2^{-b} \in \mathcal{D}$, define

$$\nu_q^{(n)} = \theta_{a2^{n-b-1}}^{(n)} \in B^\infty(\mu_q, \epsilon), \quad n > b.$$

By (5.5), we have $\nu_q^{(n)} \leq_c \nu_q^{(n+1)}$, $n > b$. Let R_n be the call function associated to $\nu_q^{(n)}$. Then we have

$$R_{S(\mu_q; m, \epsilon)} \leq R_n \leq R_{n+1} \leq R_{T(\mu_q; m, \epsilon)}, \quad n > b, \tag{5.6}$$

and thus the bounded and increasing sequence (R_n) converges pointwise to a function R . As a limit of decreasing convex functions, R is also decreasing and convex and together with (5.6) we see that R is a call function with $\lim_{x \rightarrow -\infty} R(x) + x = m$. Therefore R can be associated to a measure $\nu_q \in \mathcal{M}_m$.

Next, we will show that $\nu_q \in B^\infty(\mu_q, \epsilon)$. From the convexity of the R_n we get that

$$\begin{aligned} R'(x) &= \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} \frac{R_n(x+h) - R_n(x)}{h} \\ &\geq \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} R'_n(x+h) \\ &\geq \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} R'_{\mu_q}(x+h-\epsilon) = R'_{\mu_q}(x-\epsilon), \end{aligned}$$

and similarly

$$\begin{aligned} R'(x) &= \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} \frac{R_n(x+h) - R_n(x)}{h} \\ &\leq \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} R'_n(x) \\ &\leq \lim_{n \rightarrow \infty} R'_{\mu_q}(x+\epsilon) = R'_{\mu_q}(x+\epsilon), \end{aligned}$$

thus $W^\infty(\nu_q, \mu_q) \leq \epsilon$.

For two elements $q < q'$ of \mathcal{D} , it is an immediate consequence of (5.4) that $\nu_q^{(n)} \leq_c \nu_{q'}^{(n)}$ for large n , and therefore $\nu_q \leq_c \nu_{q'}$. It follows that $(\nu_t)_{t \in \mathcal{D}}$ is a peacock. Now pick $t \in [0, 1] \setminus \mathcal{D}$ and a sequence $\mathcal{D} \ni q_n \nearrow t$. The sequence of call functions corresponding to ν_{q_n} increases and converges to a call function, which is clearly independent of the choice of q_n . Denote the associated measure by ν_t ; it satisfies $\mathbb{E}\nu_t = m$. Fix $x \in \mathbb{R}$ and define

$$\mathcal{H} = \{h \in \mathbb{R} : F_{\mu_t} \text{ continuous at } x+h-\epsilon\}.$$

Note that $(0, \infty) \setminus \mathcal{H}$ is countable. We obtain

$$\begin{aligned} R'_{\nu_t}(x) &= \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} \frac{R_{\nu_{q_n}}(x+h) - R_{\nu_{q_n}}(x)}{h} \\ &\geq \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} R'_{\nu_{q_n}}(x+h) \\ &\geq \lim_{h \downarrow 0} \lim_{n \rightarrow \infty} R'_{\mu_{q_n}}(x+h-\epsilon) \\ &= \lim_{h \downarrow 0, h \in \mathcal{H}} \lim_{n \rightarrow \infty} R'_{\mu_{q_n}}(x+h-\epsilon) \\ &= \lim_{h \downarrow 0} R'_{\mu_t}(x+h-\epsilon) = R'_{\mu_t}(x-\epsilon), \end{aligned}$$

where the last but one equality follows from (5.2). Similarly we see that $R'_{\nu_t}(x) \leq R'_{\mu_t}(x+\epsilon)$. We have shown that $\nu_t \in B^\infty(\mu_t, \epsilon)$ for all $t \in [0, 1]$. From the definition of ν_t we have $\nu_q \leq_c \nu_t$ for $q < t, q \in \mathcal{D}$ and $\nu_t \leq_c \nu_p$ for $p > t, p \in \mathcal{D}$. This implies $\nu_s \leq_c \nu_t$ for all $0 \leq s \leq t \leq 1$, and thus $(\nu_t)_{t \in [0,1]}$ is a peacock with mean m . \square

6. Approximation by peacocks: stop-loss distance

The stop-loss distance [7, 11, 18] is defined as

$$d^{\text{SL}}(\mu, \nu) = \sup_{x \in \mathbb{R}} |R_\mu(x) - R_\nu(x)|, \quad \mu, \nu \in \mathcal{M}.$$

We will denote closed balls with respect to d^{SL} by B^{SL} . In the following proposition, we use the same notation for least elements as in the case of the infinity Wasserstein distance; no confusion should arise.

Proposition 6.1. *Given $\epsilon > 0$, a measure $\mu \in \mathcal{M}$ and $m \in [\mathbb{E}\mu - \epsilon, \mathbb{E}\mu + \epsilon]$, there exists a unique measure $S(\mu) \in B^{\text{SL}}(\mu, \epsilon) \cap \mathcal{M}_m$, such that*

$$S(\mu) \leq_c \nu, \quad \text{for all } \nu \in B^{\text{SL}}(\mu, \epsilon) \cap \mathcal{M}_m.$$

The call function of $S(\mu)$ is given by

$$R_{\mu}^{\min}(x) = R_{S(\mu)}(x) = (m - x)^+ \vee (R_{\mu}(x) - \epsilon). \tag{6.1}$$

To highlight the dependence on ϵ and m we will sometimes write $S(\mu; m, \epsilon)$ or $R_{\mu}^{\min}(\cdot; m, \epsilon)$.

Proof. It is easy to check that $R_{S(\mu)}$ defines a call function, and by (ii) of Proposition 2.3 we have

$$\begin{aligned} \mathbb{E}R_{S(\mu)} &= \lim_{x \rightarrow -\infty} R_{S(\mu)}(x) + x \\ &= \lim_{x \rightarrow -\infty} (m \vee (R_{\mu}(x) + x - \epsilon)) \\ &= m \vee (\mathbb{E}\mu - \epsilon) = m. \end{aligned}$$

The rest is clear. \square

Remark 6.2. The set $B^{\text{SL}}(\mu, \epsilon) \cap \mathcal{M}_m$ does not contain a greatest element. To see this, take an arbitrary $\nu \in B^{\text{SL}}(\mu, \epsilon) \cap \mathcal{M}_m$ and define $x_0 \in \mathbb{R}$ as the unique solution of $R_{\nu}(x) = \frac{1}{2}\epsilon$. Then for $n \in \mathbb{N}$ define new call functions

$$R_n(x) = \begin{cases} (x - x_0) \frac{R_{\nu}(x_0 + n) - R_{\nu}(x_0)}{n} + R_{\nu}(x_0), & x \in [x_0, x_0 + n], \\ R_{\nu}(x), & \text{otherwise.} \end{cases}$$

It is easy to check that R_n is indeed a call function and the associated measures θ_n lie in $B^{\text{SL}}(\mu, \epsilon) \cap \mathcal{M}_m$. Furthermore, from the convexity of R_{ν} we can deduce that $R_{\nu} \leq R_n \leq R_{n+1}$, and hence $\nu \leq_c \theta_n \leq_c \theta_{n+1}$. The call functions R_n converge to a function R which is not a call function since $R(x) = R_{\nu}(x_0) = \frac{\epsilon}{2}$ for all $x \geq x_0$. Therefore no greatest element can exist. However, it is true that a measure ν is in $B^{\text{SL}}(\mu, \epsilon)$ if and only if $R_{\mu}^{\min}(\cdot; \mathbb{E}\nu, \epsilon) \leq R_{\nu} \leq R_{\mu} + \epsilon$.

Theorem 6.3. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{M} such that*

$$I := \bigcap_{n \in \mathbb{N}} [\mathbb{E}\mu_n - \epsilon, \mathbb{E}\mu_n + \epsilon],$$

is not empty. Denote by $(R_n)_{n \in \mathbb{N}}$ the corresponding call functions. Then there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ such that

$$d^{\text{SL}}(\mu_n, \nu_n) \leq \epsilon, \quad n \in \mathbb{N}, \tag{6.2}$$

if and only if

$$R_k(x) \leq R_n(x) + 2\epsilon, \quad \text{for all } k \leq n \text{ and } x \in \mathbb{R}. \tag{6.3}$$

Proof. We first argue that (6.3) is equivalent to the assertion

$$\text{There is } m \in I \text{ such that } R_k^{\min}(x; m, \epsilon) \leq R_n(x) + \epsilon, \quad \text{for all } k \leq n \text{ and } x \in \mathbb{R}, \quad (6.4)$$

where R_k^{\min} denotes the call function of $S(\mu_k; m, \epsilon)$. Indeed, by (6.1), (6.4) clearly implies (6.3), and the converse implication follows from the obvious estimate $(m - x)^+ \leq R_n(x) + \epsilon$, valid for arbitrary $m \in I$.

Now suppose that (6.4) holds. We will define the measures ν_n via their call functions P_n . Define $P_1(x) = R_1^{\min}(x; m, \epsilon)$ and

$$P_n(x) = \max\{P_{n-1}(x), R_n^{\min}(x; m, \epsilon)\}, \quad n \geq 2. \quad (6.5)$$

It is easily verified that P_n is a call function and satisfies

$$R_n^{\min}(x) \leq P_n(x) \leq R_n(x) + \epsilon, \quad x \in \mathbb{R}, \quad (6.6)$$

and therefore ν_n , the measure associated to P_n , satisfies $\nu_n \in B^{\text{SL}}(\mu_n, \epsilon)$. Furthermore $P_n \leq P_{n+1}$, and thus $(\nu_n)_{n \in \mathbb{N}}$ is a peacock with mean m .

Now assume that $(\nu_n)_{n \in \mathbb{N}}$ is a peacock such that $d^{\text{SL}}(\mu_n, \nu_n) \leq \epsilon$. We will denote the call function of ν_n by P_n and set $m = \mathbb{E}\nu_1 \in I$. Then for $k \leq n$ and $x \in \mathbb{R}$ we get with Proposition 6.1

$$R_k^{\min}(x; m, \epsilon) \leq P_k(x) \leq P_n(x) \leq R_n(x) + \epsilon. \quad \square$$

Note that (6.4) trivially holds for $k = n$. Moreover, unwinding the recursive definition (6.5) and using (6.1), we see that P_n has the explicit expression

$$P_n(x) = \max\{(m - x)^+, R_1(x) - \epsilon, \dots, R_n(x) - \epsilon\}, \quad x \in \mathbb{R}, n \in \mathbb{N}.$$

The following proposition shows that the peacock from Theorem 6.3 is never unique.

Proposition 6.4. *In the setting of Theorem 6.3, suppose that (6.4) holds. Then there are infinitely many peacocks satisfying (6.2).*

Proof. Define P_n as in the proof of Theorem 6.3, and fix $x_0 \in \mathbb{R}$ with $P_1(x_0) < \epsilon$. For arbitrary $c \in (0, 1)$, we define

$$G(x) = \begin{cases} P_1(x_0), & x \leq x_0, \\ P_1(x_0) + cP_1'(x_0)(x - x_0), & x \geq x_0. \end{cases}$$

Thus, in a right neighborhood of x_0 , the graph of G is a line that lies above P_1 . We then put $\tilde{P}_n = P_n \vee G$, for $n \in \mathbb{N}$. It is easy to see that (\tilde{P}_n) is an increasing sequence of call functions with mean m , and thus defines a peacock. Moreover, we have

$$\tilde{P}_n \leq (R_n + \epsilon) \vee G \leq R_n + \epsilon,$$

by (6.6) and the fact that $G \leq \epsilon$. The lower estimate $\tilde{P}_n \geq P_n \geq R_n - \epsilon$ is also obvious. \square

Theorem 6.3 easily extends to continuous index sets.

Theorem 6.5. *Assume that $(\mu_t)_{t \in [0,1]}$ is a family of measures in \mathcal{M} such that*

$$I := \bigcap_{t \in [0,1]} [\mathbb{E}\mu_t - \epsilon, \mathbb{E}\mu_t + \epsilon]$$

is not empty. Denote the call function of μ_t by R_t . Then there exists a peacock $(\nu_t)_{t \in [0,1]}$ with

$$d^{\text{SL}}(\mu_t, \nu_t) \leq \epsilon, \quad \text{for all } t \in [0, 1],$$

if and only if

$$R_s(x) \leq R_t(x) + 2\epsilon, \quad \text{for all } 0 \leq s < t \leq 1 \text{ and } x \in \mathbb{R}. \tag{6.7}$$

Proof. As in the discrete case (Theorem 6.3), (6.7) is equivalent to the statement

$$\begin{aligned} &\text{There is } m \in I \text{ such that } R_s^{\min}(x; m, \epsilon) \leq R_t(x) + \epsilon, \\ &\text{for all } 0 \leq s < t \leq 1 \text{ and } x \in \mathbb{R}. \end{aligned} \tag{6.8}$$

If (6.8) holds, then we set

$$P_t(x) = \sup_{s \leq t} R_s^{\min}(x; m, \epsilon), \quad t \in [0, 1].$$

Then P_t is a call function which satisfies $R_t^{\min}(x; m, \epsilon) \leq P_t(x) \leq R_t(x) + \epsilon$ for $x \in \mathbb{R}$. The rest can be done as in the proof of Theorem 6.3. \square

7. Lévy distance and Prokhorov distance: preliminaries

The Lévy distance is a metric on the set of all measures on \mathbb{R} , defined as

$$d^L(\mu, \nu) = \inf \left\{ h > 0 : F_\mu(x - h) - h \leq F_\nu(x) \leq F_\mu(x + h) + h, \forall x \in \mathbb{R} \right\}.$$

Its importance is partially due to the fact that d^L metrizes weak convergence of measures on \mathbb{R} . The Prokhorov distance is a metric on measures on an arbitrary separable metric space (S, ρ) . For measures μ, ν on S it can be written as

$$d^P(\mu, \nu) = \inf \left\{ h > 0 : \nu(A) \leq \mu(A^h) + h, \text{ for all closed sets } A \subseteq S \right\},$$

where $A^h = \{x \in S : \inf_{a \in A} \rho(x, a) \leq h\}$. The Prokhorov distance is often referred to as a generalization of the Lévy metric, since d^P metrizes weak convergence on any separable metric space. Note, though, that d^L and d^P do not coincide when $(S, \rho) = (\mathbb{R}, |\cdot|)$. It is easy to see ([17], p. 36) that the Prokhorov distance of two measures on \mathbb{R} is an upper bound for their Lévy distance:

Lemma 7.1. *Let μ and ν be two probability measures on \mathbb{R} . Then $d^L(\mu, \nu) \leq d^P(\mu, \nu)$.*

For further information concerning these metrics, their properties and their relations to other metrics, we refer the reader to [17] (p.27 ff). Now we define slightly different distances d_p^L and d_p^P on the set of probability measures on \mathbb{R} , which in general are not metrics in the classical sense (recall the remark after Definition 2.6). These distances are useful for two reasons: First, it will turn out that balls with respect to d^L and d^P can always be written as balls w.r.t. d_p^L and d_p^P , see Lemma 7.2. Second, the function d_p^P has a direct link to minimal distance couplings which are especially useful for applications, see Proposition 7.4. For $p \in [0, 1]$ we define

$$d_p^L(\mu, \nu) := \inf \left\{ h > 0 : F_\mu(x - h) - p \leq F_\nu(x) \leq F_\mu(x + h) + p, \forall x \in \mathbb{R} \right\} \tag{7.1}$$

and

$$d_p^P(\mu, \nu) := \inf \left\{ h > 0 : \nu(A) \leq \mu(A^h) + p, \text{ for all closed sets } A \subseteq S \right\}. \tag{7.2}$$

It is easy to show (using complements) that $d_p^P(\mu, \nu) = d_p^P(\nu, \mu)$ (see e.g. Proposition 1 in [9]). Note that $d_p^P(\mu, \nu) = 0$ does not imply that $\mu = \nu$. We will refer to d_p^L as the modified Lévy

distance, and to d_p^P as the modified Prokhorov distance.⁴ The following Lemma explains the connection between the Lévy distance d^L and the modified Lévy distance d_p^L , resp. the Prokhorov distance d^P and the modified Prokhorov distance d_p^P .

Lemma 7.2. *Let $\mu \in \mathcal{M}$. Then for every $\epsilon \in [0, 1]$ we have*

$$B^L(\mu, \epsilon) = B_\epsilon^L(\mu, \epsilon) \quad \text{and} \quad B^P(\mu, \epsilon) = B_\epsilon^P(\mu, \epsilon).$$

Proof. For $\nu \in \mathcal{M}$, the assertion $\nu \in B^P(\mu, \epsilon)$ is equivalent to

$$\mu(A) \leq \nu(A^{\epsilon+\delta}) + \epsilon + \delta, \quad \delta > 0, \quad A \subseteq \mathbb{R} \text{ closed}, \tag{7.3}$$

whereas $\nu \in B_\epsilon^P(\mu, \epsilon)$ means that

$$\mu(A) \leq \nu(A^{\epsilon+\delta}) + \epsilon, \quad \delta > 0, \quad A \subseteq \mathbb{R} \text{ closed}. \tag{7.4}$$

Obviously, (7.4) implies (7.3). Now suppose that (7.3) holds, and let $\delta \downarrow 0$. Notice that $A^{\epsilon+\delta_1} \subseteq A^{\epsilon+\delta_2}$ for $\delta_1 \leq \delta_2$. The continuity of ν then gives

$$\mu(A) \leq \nu(A^\epsilon) + \epsilon \leq \nu(A^{\epsilon+\delta}) + \epsilon \quad \delta > 0, \quad A \subseteq \mathbb{R} \text{ closed},$$

and thus $B^P(\mu, \epsilon) = B_\epsilon^P(\mu, \epsilon)$. Replacing A by intervals $(-\infty, x]$ for $x \in \mathbb{R}$ in (7.3) and (7.4) proves that $B^L(\mu, \epsilon) = B_\epsilon^L(\mu, \epsilon)$. \square

Similarly to Lemma 7.1 we can show that the modified Lévy distance of two measures never exceeds the modified Prokhorov distance.

Lemma 7.3. *Let μ and ν be two probability measures on \mathbb{R} and let $p \in [0, 1]$. Then*

$$d_p^L(\mu, \nu) \leq d_p^P(\mu, \nu).$$

Proof. We set $\epsilon = d_p^P(\mu, \nu)$. Then for any $x \in \mathbb{R}$ and all $n \in \mathbb{N}$ we have

$$\begin{aligned} F_\nu(x) &= \nu((-\infty, x]) \leq \mu\left(\left(-\infty, x + \epsilon + \frac{1}{n}\right]\right) + p \\ &= F_\mu\left(x + \epsilon + \frac{1}{n}\right) + p, \end{aligned}$$

and by the symmetry of d^P the above relation also holds with μ and ν interchanged. This implies that $d_p^L(\mu, \nu) \leq \epsilon$. \square

The following coupling representation of d_p^P was first proved by Strassen and was then extended by Dudley [9,36].

Proposition 7.4. *Given measures μ, ν on \mathbb{R} , $p \in [0, 1]$, and $\epsilon > 0$ there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with random variables $X \sim \mu$ and $Y \sim \nu$ such that*

$$\mathbb{P}(|X - Y| > \epsilon) \leq p, \tag{7.5}$$

if and only if

$$d_p^P(\mu, \nu) \leq \epsilon. \tag{7.6}$$

⁴ Note that our definition of the modified Prokhorov distance does *not* agree with the Prokhorov-type metric π_λ from [32] and [33].

8. Approximation by peacocks: Prokhorov distance and Lévy distance

In this section we will prove peacock approximation results, first for the modified Prokhorov distance and later on for the modified Lévy distance, the Prokhorov distance, and the Lévy distance. It turns out that [Problem 2.4](#) always has a solution for these distances, regardless of the size of ϵ . In the following we denote the quantile function of a measure $\mu \in \mathcal{M}$ by G_μ , i.e.

$$G_\mu(p) = \inf \{x \in \mathbb{R} : F_\mu(x) \geq p\}, \quad p \in [0, 1].$$

Proposition 8.1. *Let $\mu \in \mathcal{M}$, $p \in (0, 1]$, and $m \in \mathbb{R}$. Then the set*

$$B_p^P(\mu, 0) \cap \mathcal{M}_m$$

is not empty. Moreover, this set contains at least one measure with bounded support.

Proof. The statement is clear for $p = 1$, and so we focus on $p \in (0, 1)$. Given a measure μ we set $I = [G_\mu(\frac{p}{4}), G_\mu(1 - \frac{p}{4})]$. We will first define a measure η with bounded support which lies in $B_p^P(\mu, 0)$, and then we will modify it to obtain a measure θ with mean m . We set

$$F_\eta(x) := \begin{cases} 0, & x < G_\mu(\frac{p}{4}), \\ F_\mu(x), & x \in I, \\ 1, & x \geq G_\mu(1 - \frac{p}{4}), \end{cases}$$

which is clearly a distribution function of a measure η . Note that η has bounded support, so in particular η has finite mean. Next we define

$$\theta = \left(1 - \frac{p}{2}\right)\eta + \frac{p}{2}\delta_w,$$

where w is chosen such that $\mathbb{E}\theta = m$. Since η has bounded support, we can deduce that θ also has bounded support. Now for every closed set $A \subseteq \mathbb{R}$ we have

$$\begin{aligned} \theta(A) &\leq \left(1 - \frac{p}{2}\right)\eta(A) + \frac{p}{2} \\ &\leq \left(1 - \frac{p}{2}\right)\eta(A \cap \text{int}(I)) + p \\ &\leq \mu(A) + p, \end{aligned}$$

where $\text{int}(I)$ denotes the interior of I . For the last inequality, note that μ and η are equal on $\text{int}(I)$. The last equation implies that $\theta \in B_p^P(\mu, 0) \cap \mathcal{M}_m$. \square

Note that in [Proposition 8.1](#) it is not important that μ has finite mean. The statement is true for all measures on \mathbb{R} . The same is true for all subsequent results of this section.

Proposition 8.2. *Let $v \in \mathcal{M}$ be a measure with bounded support and $p \in (0, 1)$. Then for all measures $\mu \in \mathcal{M}$ there exists a measure $\theta \in B_p^P(\mu, 0)$ with bounded support such that $v \leq_c \theta$.*

Proof. Fix $\mu, v \in \mathcal{M}$ and $p \in (0, 1)$, and set $m = \mathbb{E}v$. Then, by [Proposition 8.1](#), there is a measure $\theta_0 \in B_{p/2}^P(\mu, 0) \cap \mathcal{M}_m$ which has bounded support. For $n \in \mathbb{N}$ we define

$$\theta_n = \left(1 - \frac{p}{2}\right)\theta_0 + \frac{p}{4}\delta_{m-n} + \frac{p}{4}\delta_{m+n}.$$

These measures have bounded support and mean m . Furthermore, for $A \subseteq \mathbb{R}$ closed, we have

$$\begin{aligned} \theta_n(A) &\leq \left(1 - \frac{p}{2}\right)\theta_0(A) + \frac{p}{2} \\ &\leq \theta_0(A) + \frac{p}{2} \leq \mu(A) + p, \quad n \in \mathbb{N}, \end{aligned}$$

and hence $\nu_n \in B_p^P(\mu, 0)$ for all $n \in \mathbb{N}$. Now observe that for all $n \in \mathbb{N}$ and $x \in [m - n, m + n]$ we have

$$R_{\theta_n}(x) = \left(1 - \frac{p}{2}\right)R_{\theta_0}(x) + \frac{p}{4}(m + n - x), \tag{8.1}$$

which tends to infinity as n tends to infinity. Outside of the support of θ_n (i.e. outside the interval $[m - n, m + n]$) the call function of θ_n equals the call function of the Dirac measure δ_m with mass at m . Therefore there has to exist $n_0 \in \mathbb{N}$ such that $\nu \leq_c \theta_{n_0}$. \square

In Proposition 8.2 it is important that $p > 0$. For $p = 0$ the limit in (8.1) is finite.

Theorem 8.3. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{M} , $\epsilon > 0$, and $p \in (0, 1]$. Then, for all $m \in \mathbb{R}$ there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ with mean m such that*

$$d_p^P(\mu_n, \nu_n) \leq \epsilon.$$

Proof. If $p = 1$ then $B_p^P(\mu, 0)$ contains all probability measures on \mathbb{R} , which is easily seen from the definition of d_p^P , and the result is trivial. So we consider the case $p < 1$. Since $B_p^P(\mu, 0) \subseteq B_p^P(\mu, \epsilon)$, it suffices to prove the statement for $\epsilon = 0$. By Proposition 8.1, there exists a measure $\nu_1 \in B_p^P(\mu_1, 0) \cap \mathcal{M}_m$ with bounded support. By Proposition 8.2 there exists a measure $\nu_2 \in B_p^P(\mu_2, 0)$ such that $\nu_1 \leq_c \nu_2$. Since ν_2 has again finite support, we can proceed inductively to finish the proof. \square

Setting $\epsilon = p \in (0, 1]$ in the previous result, we obtain the following corollary.

Corollary 8.4. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{M} and $\epsilon > 0$. Then, for all $m \in \mathbb{R}$ there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ with mean m such that*

$$d^P(\mu_n, \nu_n) \leq \epsilon.$$

Proof. By Lemma 7.2 we have $B^P(\mu, \epsilon) = B_\epsilon^P(\mu, \epsilon)$ for all $\mu \in \mathcal{M}$ and $\epsilon \in [0, 1]$. The result now easily follows from Theorem 8.3. \square

Since balls with respect to the modified Prokhorov metric are smaller than balls with respect to the Lévy metric, we get the following corollary.

Theorem 8.5. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{M} , $\epsilon > 0$, and $p \in (0, 1]$. Then, for all $m \in \mathbb{R}$ there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ with mean m such that*

$$d_p^L(\mu_n, \nu_n) \leq \epsilon.$$

In particular, there exists a peacock $(\nu_n)_{n \in \mathbb{N}}$ with mean m such that

$$d^L(\mu_n, \nu_n) \leq \epsilon.$$

Proof. Fix $\epsilon > 0$ and $p \in (0, 1]$, and let $(\nu_n)_{n \in \mathbb{N}}$ be the peacock from Theorem 8.3 resp. Corollary 8.4. Then by Lemma 7.3 resp. Lemma 7.1, we have $\nu_n \in B_p^L(\mu_n, \epsilon)$ resp. $\nu_n \in B^L(\mu_n, \epsilon)$ for all $n \in \mathbb{N}$. \square

For $\mu \in \mathcal{M}$, $\epsilon \geq 0$, $p \in (0, 1)$, and $m \in \mathbb{R}$, the set $B_p^L(\mu, \epsilon) \cap \mathcal{M}_m$ always contains a least element with respect to \leq_c , with an explicit call function. See Section 2.4.3 in [14].

9. A variant of Strassen’s theorem

So far, we discussed the problem of approximating a given sequence of measures (μ_n) by a peacock (ν_n) . If the distance is measured by W^∞ , then the existence of such a peacock has two consequences: First, there is a probability space with a martingale M^* with marginals (ν_n) (by Strassen’s theorem). Second, the definition of W^∞ implies that for each $\epsilon' > \epsilon$ there is a probability space supporting processes \hat{M} and \hat{X} satisfying $\hat{\mathbb{P}}[|\hat{M}_n - \hat{X}_n| > \epsilon'] = 0$ for all n . It is now a natural question whether a martingale M with marginals (ν_n) can be found such that there is an adapted process X satisfying $\mathbb{P}[|M_n - X_n| > \epsilon'] = 0$. We answer this question affirmatively for finite sequences of measures with finite support. This restriction suffices for the financial application that motivated our study (see [12]), and it allows to replace “for all $\epsilon' > \epsilon \dots \mathbb{P}[|M_n - X_n| > \epsilon'] = 0$ ” simply by $\mathbb{P}[|M_n - X_n| > \epsilon] = 0$. The result (Theorem 9.2) is a consequence of Theorem 3.5 and the following lemma.

Lemma 9.1. *Let $\epsilon > 0$. Let $(\nu_n)_{n=1, \dots, n_0}$ be a peacock, and $(\mu_n)_{n=1, \dots, n_0}$ be a sequence of measures in \mathcal{M} . Assume that there is a finite filtered probability space $(\Omega^*, \mathcal{F}^*, (\mathcal{F}_n^*)_{1 \leq n \leq n_0}, \mathbb{P}^*)$ with a martingale M^* satisfying $M_n^* \sim \nu_n$ for $1 \leq n \leq n_0$.*

Assume further that there is a finite probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ supporting two processes \hat{M} and \hat{X} satisfying $\hat{M}_n \sim \nu_n$, $\hat{X}_n \sim \mu_n$ for $1 \leq n \leq n_0$ and

$$\hat{\mathbb{P}}[|\hat{M}_n - \hat{X}_n| > \epsilon] = 0, \quad n = 1, \dots, n_0. \tag{9.1}$$

Then there is a finite filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{1 \leq n \leq n_0}, \mathbb{P})$ with processes M and X combining all properties mentioned, i.e.:

- M is a martingale
- X is adapted
- $M_n \sim \nu_n, X_n \sim \mu_n, \quad n = 1, \dots, n_0,$
- $\mathbb{P}[|M_n - X_n| > \epsilon] = 0, \quad n = 1, \dots, n_0.$

Proof. Let $n' \in \{1, \dots, n_0\}$ and assume, inductively, that we have already constructed a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$ that satisfies the requirements, where the conditions concerning X hold for $n < n'$, i.e. there are processes $M = (M_n)_{1 \leq n \leq n_0}$ and $X = (X_n)_{1 \leq n < n'}$ such that

- M is a martingale
- X is adapted
- $M_n \sim \nu_n, \quad n = 1, \dots, n_0,$
- $X_n \sim \mu_n, \quad 1 \leq n < n',$
- $\mathbb{P}[|M_n - X_n| > \epsilon] = 0, \quad 1 \leq n < n'.$

Note that in case $n' = 1$ (induction base) we may simply take $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P}) = (\Omega^*, \mathcal{F}^*, (\mathcal{F}_n^*), \mathbb{P}^*)$. Let $z \in \mathbb{R}$ be an arbitrary member of the image of $M_{n'}$, and define

$$U := A_1 \cup \dots \cup A_m := (M_{n'})^{-1}(z),$$

where A_1, \dots, A_m are (distinct) atoms of $\mathcal{F}_{n'}$. We denote the preimage of z in $\hat{\Omega}$ by

$$\{\hat{\omega}_1, \dots, \hat{\omega}_l\} := \hat{M}_{n'}^{-1}(z).$$

As $M_{n'} \sim \mu_{n'} \sim \hat{M}_{n'}$, we have

$$\mathbb{P}[U] = \mathbb{P}[A_1 \cup \dots \cup A_m] = \hat{\mathbb{P}}[\{\hat{\omega}_1, \dots, \hat{\omega}_l\}]. \tag{9.2}$$

To make room for an appropriate $X_{n'}$ on a new filtered probability space, whose constituents will be denoted by Ω', \mathcal{F}' etc., we divide each “old” atom

$$A_r =: \{\omega_{r1}, \dots, \omega_{rk_r}\}, \quad 1 \leq r \leq m,$$

into l “new” atoms

$$A_r^{(i)} := \{\omega_{r1}^{(i)}, \dots, \omega_{rk_r}^{(i)}\}, \quad 1 \leq r \leq m, \quad 1 \leq i \leq l.$$

Then, define

$$\Omega' := (\Omega \setminus U) \cup \bigcup_{\substack{1 \leq r \leq m \\ 1 \leq i \leq l}} A_r^{(i)}$$

and $\mathcal{F}' := 2^{\Omega'}$. We let $\mathbb{P}' := \mathbb{P}$ on $\Omega \setminus U$ and

$$\mathbb{P}'[\omega_{rj}^{(i)}] := \frac{\mathbb{P}[\omega_{rj}] \hat{\mathbb{P}}[\hat{\omega}_i]}{\sum_{i'=1}^l \hat{\mathbb{P}}[\hat{\omega}_{i'}]}, \quad 1 \leq r \leq m, \quad 1 \leq i \leq l, \quad 1 \leq j \leq k_r.$$

The sigma-algebra $\mathcal{F}'_{n'}$ is generated by the atoms of $\mathcal{F}'_{n'}$, but with each atom A_r replaced by the atoms $A_r^{(1)}, \dots, A_r^{(l)}$. Similarly, we define \mathcal{F}'_n for $n < n'$ and $n > n'$. E.g., if A_1 decomposes into atoms $A_1 = B \cup \tilde{B}$ in $\mathcal{F}'_{n'+1}$, then we replace B and \tilde{B} by $B \cap A_1^{(1)}, \dots, B \cap A_1^{(l)}$ and $\tilde{B} \cap A_1^{(1)}, \dots, \tilde{B} \cap A_1^{(l)}$, respectively, and so on. Clearly, this defines a filtered probability space $(\Omega', \mathcal{F}', (\mathcal{F}'_n), \mathbb{P}')$. On this space, we define M' like M , forgetting that the atoms A_1, \dots, A_m were split: $M'_n := M_n$ for all n on $\Omega \setminus U$ and

$$M'_n(\omega_{rj}^{(i)}) := M_n(\omega_{rj}), \quad 1 \leq r \leq m, \quad 1 \leq i \leq l, \quad 1 \leq j \leq k_r, \quad 1 \leq n \leq n_0.$$

Thus, the adapted process M' has the same marginal laws as M . Now we verify that M' is a martingale. Let $n_1 > n'$. (The cases of time points n_1, n_2 in other positions relative to n' work very similarly, but need additional cumbersome notation.) First, let A' be any atom of $\mathcal{F}'_{n'}$, distinct from $A_r^{(1)}, \dots, A_r^{(l)}, 1 \leq r \leq m$. Then we compute

$$\begin{aligned} \mathbb{E}'[M'_{n_1} | A'] &= \frac{\sum_{\omega \in A'} M'_{n_1}(\omega) \mathbb{P}'[\omega]}{\mathbb{P}'[A']} \\ &= \frac{\sum_{\omega \in A'} M_{n_1}(\omega) \mathbb{P}[\omega]}{\mathbb{P}[A']} \\ &= \mathbb{E}[M_{n_1} | A'] = M_{n'}(A') = M'_{n'}(A'). \end{aligned}$$

For $r \in \{1, \dots, m\}$ and $i \in \{1, \dots, l\}$, we have

$$\begin{aligned} \mathbb{E}'[M'_{n_1} | A_r^{(i)}] &= \frac{\sum_{j=1}^{k_r} M'_{n_1}(\omega_{rj}^{(i)}) \mathbb{P}'[\omega_{rj}^{(i)}]}{\sum_{j=1}^{k_r} \mathbb{P}'[\omega_{rj}^{(i)}]} \\ &= \frac{\sum_{j=1}^{k_r} M_{n_1}(\omega_{rj}) \mathbb{P}[\omega_{rj}]}{\sum_{j=1}^{k_r} \mathbb{P}[\omega_{rj}]} \\ &= \mathbb{E}[M_{n_1} | A_r] = M_{n'}(A_r) = M'_{n'}(A_r^{(i)}). \end{aligned}$$

Therefore, M' is a martingale. Now we define the process $(X'_n)_{1 \leq n < n'}$ as X on $\Omega \setminus U$, and

$$X'_n(\omega_{rj}^{(i)}) := X_n(\omega_{rj}), \quad 1 \leq r \leq m, \quad 1 \leq i \leq l, \quad 1 \leq j \leq k_r, \quad 1 \leq n < n'.$$

As for n' , we put

$$X'_{n'}(\omega_{rj}^{(i)}) := \hat{X}_{n'}(\hat{\omega}_i), \quad 1 \leq r \leq m, \quad 1 \leq i \leq l, \quad 1 \leq j \leq k_r. \tag{9.3}$$

To make the definition complete, let $X'_{n'} := M_{n'}$ on $\Omega \setminus U$, although this is of no relevance, because this definition will be overwritten when we continue the construction for the next element of the image of $M_{n'}$. As the right hand side of (9.3) is independent of j , the process $(X'_{n'})_{1 \leq n \leq n'}$ is adapted to $(\mathcal{F}'_n)_{1 \leq n \leq n'}$. We now show that the random variables

$$X'_{n'}|_U \quad \text{and} \quad \hat{X}_{n'}|_{\{\hat{\omega}_1, \dots, \hat{\omega}_l\}}$$

have the same law. Indeed, for $1 \leq i \leq l$ we have

$$\begin{aligned} \mathbb{P}'[X'_{n'}|_U = \hat{X}_{n'}(\hat{\omega}_i)] &= \sum_{r=1}^m \sum_{j=1}^{k_r} \sum_{i': X'_{n'}(\omega_{rj}^{(i')}) = \hat{X}_{n'}(\hat{\omega}_i)} \mathbb{P}'[\omega_{rj}^{(i')}] \\ &= \sum_{r=1}^m \sum_{j=1}^{k_r} \sum_{i': \hat{X}_{n'}(\hat{\omega}_{i'}) = \hat{X}_{n'}(\hat{\omega}_i)} \frac{\hat{\mathbb{P}}[\hat{\omega}_{i'}] \mathbb{P}[\omega_{rj}]}{\sum_{i''=1}^l \hat{\mathbb{P}}[\hat{\omega}_{i''}]} \\ &= \frac{1}{\sum_{i''=1}^l \hat{\mathbb{P}}[\hat{\omega}_{i''}]} \sum_{r=1}^m \sum_{j=1}^{k_r} \mathbb{P}[\omega_{rj}] \sum_{i': \hat{X}_{n'}(\hat{\omega}_{i'}) = \hat{X}_{n'}(\hat{\omega}_i)} \hat{\mathbb{P}}[\hat{\omega}_{i'}] \\ &= \hat{\mathbb{P}}[\hat{X}_{n'}|_{\{\hat{\omega}_1, \dots, \hat{\omega}_l\}} = \hat{X}_{n'}(\hat{\omega}_i)], \end{aligned}$$

where we used (9.2) in the last inequality. It remains to verify

$$\mathbb{P}'[|M'_n - X'_{n'}| > \epsilon] = 0, \quad 1 \leq n \leq n'.$$

From the definition of M' and X' , this is clear for $n < n'$, and for $n = n'$ it is obvious that $|M'_n - X'_{n'}| \leq \epsilon$ on $\Omega \setminus U$. For an arbitrary element $\omega_{rj}^{(i)}$, we have

$$\begin{aligned} |M'_{n'}(\omega_{rj}^{(i)}) - X'_{n'}(\omega_{rj}^{(i)})| &= |M_{n'}(\omega_{rj}) - \hat{X}_{n'}(\hat{\omega}_i)| \\ &= |\hat{M}_{n'}(\hat{\omega}_i) - \hat{X}_{n'}(\hat{\omega}_i)| \leq \epsilon. \end{aligned}$$

The last inequality follows from (9.1), as we may assume w.l.o.g. that $\hat{\mathbb{P}}$ puts mass on all elements of $\hat{\Omega}$.

Recall that U was defined as the preimage of z . Repeating the procedure we just described for all values in the range of $M_{n'}$ completes the induction step. \square

For the formulation of the main result of this section, recall the definition of Φ_N in (3.5). Theorem 9.2 holds for $\epsilon = 0$, too; then it is just a special case of Strassen’s theorem (recall Proposition 3.4 and (3.7)).

Theorem 9.2 (A Variant of Strassen’s Theorem). *Let $\epsilon > 0$ and $(\mu_n)_{n=1, \dots, n_0}$ be a sequence of measures in \mathcal{M} with finite support such that*

$$I := \bigcap_{1 \leq n \leq n_0} [\mathbb{E}\mu_n - \epsilon, \mathbb{E}\mu_n + \epsilon] \neq \emptyset.$$

Then the following conditions are equivalent:

- (i) For some $m \in I$ and for all $1 \leq N < n_0$ and $x_1, \dots, x_N \in \mathbb{R}$, we have

$$\Phi_N(x_1, \dots, x_N; m, \epsilon) \leq 0.$$

(ii) There is a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{1 \leq n \leq n_0}, \mathbb{P})$ supporting two processes M and X such that

- M is a martingale w.r.t. $(\mathcal{F}_n)_{1 \leq n \leq n_0}$
- X is adapted to $(\mathcal{F}_n)_{1 \leq n \leq n_0}$
- $M_n \sim \nu_n, \quad n = 1, \dots, n_0,$
- $X_n \sim \mu_n, \quad n = 1, \dots, n_0,$
- $\mathbb{P}[|M_n - X_n| > \epsilon] = 0, \quad n = 1, \dots, n_0.$

Proof. Suppose that (ii) holds. Since $\mathbb{P}[|M_n - X_n| > \epsilon] = 0$, we have $W^\infty(\mu_n, \nu_n) \leq \epsilon$. As M is a martingale, (ν_n) is a peacock, and so (i) follows from (the easy implication of) [Theorem 3.5](#).

Now assume that (i) holds. Then [Theorem 3.5](#) yields a peacock $(\nu_n)_{1 \leq n \leq n_0}$ satisfying $W^\infty(\mu_n, \nu_n) \leq \epsilon$ for $1 \leq n \leq n_0$. Using [Corollary 4.4](#), we see that the finiteness of the support of the μ_n implies that we can choose $(\nu_n)_{1 \leq n \leq n_0}$ with finite support, too. From Strassen’s theorem we get a filtered probability space $(\Omega^*, \mathcal{F}^*, (\mathcal{F}_n^*), \mathbb{P}^*)$ with a martingale M^* satisfying $M_n^* \sim \nu_n$ for $1 \leq n \leq n_0$. Moreover, as $W^\infty(\mu_n, \nu_n) \leq \epsilon$, there is a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ with two processes \hat{M} and \hat{X} satisfying $\hat{M}_n \sim \nu_n, \hat{X}_n \sim \mu_n$ for all n and

$$\hat{\mathbb{P}}[|\hat{M}_n - \hat{X}_n| > \epsilon] = 0, \quad n = 1, \dots, n_0.$$

(This is an easy consequence of [Proposition 7.4](#) and the finiteness of the supports of μ_n and ν_n .) We may assume that both Ω^* and $\hat{\Omega}$ are finite. Indeed, we may clearly replace them by the finite sets

$$\text{all intersections of sets from } \{(M_n^*)^{-1}(z) : z \in \text{supp}(\nu_n), 1 \leq n \leq n_0\}$$

respectively

$$\begin{aligned} &\text{all intersections of sets from } \{\hat{M}_n^{-1}(z) : z \in \text{supp}(\nu_n), 1 \leq n \leq n_0\} \\ &\text{and } \{\hat{X}_n^{-1}(z) : z \in \text{supp}(\mu_n), 1 \leq n \leq n_0\} \end{aligned}$$

and update the sigma-algebras and the filtration of Ω^* accordingly. The assertion then follows from [Lemma 9.1](#). \square

In future work, we intend to prove an appropriate version of [Theorem 9.2](#) (possibly featuring d_0^p or d_p^p instead of W^∞) for infinite sequences of general probability measures. Also, a natural problem is to extend our peacock approximation results to other distances, such as the p -Wasserstein distance W^p ($p \geq 1$). Note that a related problem (involving the *sum* of the W^2 -distances of all sequence elements) has been solved in [\[35\]](#).

References

- [1] B. Armbruster, A short proof of Strassen’s theorem using convex analysis. Preprint, available at <http://users.iems.northwestern.edu/~armbruster/>, 2013.
- [2] M. Beiglböck, M. Huesmann, F. Stebegg, Root to Kellerer, in: Séminaire de Probabilités XLVIII, in: Lecture Notes in Math., vol. 2168, Springer, Cham, 2016, pp. 1–12.
- [3] C.R. Borges, The sup metric on infinite products, Bull. Aust. Math. Soc. 44 (3) (1991) 461–466.
- [4] M.W. Botsko, R.A. Gosser, Stronger versions of the fundamental theorem of calculus, Amer. Math. Monthly 93 (4) (1986) 294–296.
- [5] T. Champion, L. De Pascale, P. Juutinen, The ∞ -Wasserstein distance: Local solutions and existence of optimal transport maps, SIAM J. Math. Anal. 40 (1) (2008) 1–20.
- [6] M.H.A. Davis, D.G. Hobson, The range of traded option prices, Math. Finance 17 (1) (2007) 1–14.

- [7] N. De Pril, J. Dhaene, Error bounds for compound Poisson approximations of the individual risk model, *Astin Bull.* 22 (02) (1992) 135–148.
- [8] C. Dellacherie, P.A. Meyer, *Probabilities and Potential. C*, in: North-Holland Mathematics Studies, vol. 151, North-Holland Publishing Co., Amsterdam, 1988, p. xiv+416 Potential theory for discrete and continuous semigroups, Translated from the French by J. Norris.
- [9] R.M. Dudley, Distances of probability measures and random variables, *Ann. Math. Statist.* 39 (1968) 1563–1572.
- [10] H. Edelsbrunner, J.L. Harer, *Computational Topology*, American Mathematical Society, Providence, RI, 2010, p. xii+241. An introduction.
- [11] H.U. Gerber, *An Introduction to Mathematical Risk Theory*, Vol. 8, SS Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania Philadelphia, 1979.
- [12] S. Gerhold, I.C. Gülüm, Consistency of option prices under bid-ask spreads, arXiv preprint arXiv:1608.05585v1, 2016.
- [13] A.L. Gibbs, F.E. Su, On choosing and bounding probability metrics, *Internat. Statist. Rev.* 70 (3) (2002) 419–435.
- [14] I.C. Gülüm, Consistency of Option Prices under Bid-Ask Spreads and Implied Volatility Slope Asymptotics (Ph.D. thesis), TU Wien, 2016.
- [15] F. Hirsch, C. Profeta, B. Roynette, M. Yor, Peacocks and Associated Martingales, with Explicit Constructions, in: *Bocconi & Springer Series*, vol. 3, Springer, Milan; Bocconi University Press, Milan, 2011, p. xxxii+384.
- [16] F. Hirsch, B. Roynette, A new proof of Kellerer's theorem, *ESAIM Probab. Stat.* 16 (2012) 48–60.
- [17] P.J. Huber, E.M. Ronchetti, *Robust Statistics*, second ed., in: *Wiley Series in Probability and Statistics*, John Wiley & Sons, Inc., Hoboken, NJ, 2009.
- [18] R. Kaas, A. Van Heerwaarden, M. Goovaerts, On stop-loss premiums for the individual model, *Astin Bull.* 18 (01) (1988) 91–97.
- [19] H.G. Kellerer, Markov-Komposition und eine Anwendung auf Martingale, *Math. Ann.* 198 (1972) 99–122.
- [20] H.G. Kellerer, Integraldarstellung von Dilationen, in: *Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* (Tech. Univ., Prague, 1971; Dedicated to the Memory of Antonín Špaček), Academia, Prague, 1973, pp. 341–374.
- [21] R.P. Kertz, U. Rösler, Complete lattices of probability measures with applications to martingale theory, 35 (2000), 153–177.
- [22] P. Lévy, *Calcul des probabilités*, Vol. 9, Gauthier-Villars Paris, 1925.
- [23] T. Lindvall, On Strassen's theorem on stochastic domination, *Electron. Commun. Probab.* 4 (1999) 51–59 (electronic). Review, pointing out and correcting an error, available at <http://www.ams.org/mathscinet-getitem?mr=1711599>.
- [24] T. Lindvall, *Lectures on the Coupling Method*, Dover Publications, Inc, 2002, p. xiv+257. Corrected reprint of the 1992 original.
- [25] G. Lowther, Fitting martingales to given marginals, arXiv preprint arXiv:0808.2319, 2008.
- [26] G. Lowther, Limits of one-dimensional diffusions, *Ann. Probab.* 37 (1) (2009) 78–106.
- [27] A. Müller, L. Rüschendorf, On the optimal stopping values induced by general dependence structures, *J. Appl. Probab.* 38 (3) (2001) 672–684.
- [28] A. Müller, M. Scarsini, Stochastic order relations and lattices of probability measures, *SIAM J. Optim.* 16 (4) (2006) 1024–1043 (electronic).
- [29] A. Müller, D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, in: *Wiley Series in Probability and Statistics*, John Wiley & Sons, Ltd., Chichester, 2002, p. xii+330.
- [30] A.M. Oberman, The convex envelope is the solution of a nonlinear obstacle problem, *Proc. Amer. Math. Soc.* 135 (6) (2007) 1689–1694 (electronic).
- [31] Y.V. Prokhorov, Convergence of random processes and limit theorems in probability theory, *Theory Probab. Appl.* 1 (2) (1956) 157–214.
- [32] S.T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, in: *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley & Sons, Ltd., Chichester, 1991, p. xiv+494.
- [33] S.T. Rachev, L. Rüschendorf, A. Schief, Uniformities for the convergence in law and in probability, *J. Theoret. Probab.* 5 (1) (1992) 33–44.
- [34] A.W. Roberts, D.E. Varberg, *Convex Functions*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1973, p. xx+300. *Pure and Applied Mathematics*, Vol. 57.
- [35] L. Rüschendorf, The Wasserstein distance and approximation theorems, *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* 70 (1) (1985) 117–129.
- [36] V. Strassen, The existence of probability measures with given marginals, *Ann. Math. Stat.* 36 (1965) 423–439.
- [37] V.M. Zolotarev, Metric distances in spaces of random variables and of their distributions, *Mat. Sb. (NS)* 101(143) (3) (1976) 416–454, 456.