

From Bachelier to Dupire via Optimal Transport

Mathias Beiglböck, Gudmund Pammer, Walter Schachermayer

May 17, 2021

Abstract

Famously mathematical finance was started by Bachelier in his 1900 PhD thesis where – among many other achievements – he also provides a formal derivation of the Kolmogorov forward equation. This forms also the basis for Dupire’s (again formal) solution to the problem of finding an arbitrage free model calibrated to the volatility surface. The later result has rigorous counterparts in the theorems of Kellerer and Lowther. In this survey article we revisit these hallmarks of stochastic finance, highlighting the role played by some optimal transport results in this context.

Keywords: Bachelier, Dupire’s formula, Kellerer’s theorem, optimal transport, martingales, peacocks.

1 Bachelier’s work relating Brownian motion to mass transport and the heat equation

In this section, which is mainly dedicated to the historic point of view, we follow [54] and point out that Bachelier already had some thoughts on “*horizontal transport of probability measures*” in his dissertation “*Théorie de la spéculation*” [4], which he defended in Paris in 1900.

In this thesis he was the first to consider a *mathematical model* of Brownian motion. Bachelier argued using infinitesimals by visualizing Brownian motion $(W(t))_{t \geq 0}$ as an infinitesimal version of a random walk. His 19th century style argument runs as follows. Suppose that the grid in space is given by

$$\dots, x_{n-2}, x_{n-1}, x_n, x_{n+1}, x_{n+2}, \dots \quad (1.1)$$

having the same (infinitesimal) distance $\Delta x = x_n - x_{n-1}$, for all n , and such that at time t these points have (infinitesimal) probabilities

$$\dots, p_{n-2}^t, p_{n-1}^t, p_n^t, p_{n+1}^t, p_{n+2}^t, \dots \quad (1.2)$$

under the random walk under consideration. What are the probabilities

$$\dots, p_{n-2}^{t+\Delta t}, p_{n-1}^{t+\Delta t}, p_n^{t+\Delta t}, p_{n+1}^{t+\Delta t}, p_{n+2}^{t+\Delta t}, \dots \quad (1.3)$$

of these points at time $t + \Delta t$?

The random walk moves half of the mass p_n^t , sitting on x_n at time t , to the point x_{n+1} . En revanche, it moves half of the mass p_{n+1}^t , sitting on x_{n+1} at time t , to the point x_n . We thus may calculate the net difference between $p_n^t/2$ and $p_{n+1}^t/2$, which Bachelier identifies with

$$-\frac{1}{2} \frac{\partial p^t}{\partial x}(x), \quad (1.4)$$

where we let $x = x_n = x_{n+1}$ which is legitimate for Bachelier as x_n and x_{n+1} only differ by an infinitesimal.

This amount of mass is transported from the interval $(-\infty, x_n]$ to $[x_{n+1}, \infty)$ during the time interval $(t, t + \Delta t)$. In Bachelier's own words, this is very nicely captured by the following quote from his thesis:

“Each price x during an element of time radiates towards its neighboring price an amount of probability proportional to the difference of their probabilities. I say proportional because it is necessary to account for the relation of Δx to Δt . The above law can, by analogy with certain physical theories, be called the law of radiation or diffusion of probability.”

Passing formally to the continuous limit and – using today's terminology – denoting by

$$P_t(x) = \int_{-\infty}^x p_t(z) dz \quad (1.5)$$

the distribution function associated to the Gaussian density function $p_t(x)$, Bachelier thus deduces in this intuitively convincing way the relation

$$\frac{\partial P}{\partial t} = \frac{1}{2} \frac{\partial p}{\partial x}, \quad (1.6)$$

where we have normalized the relation between Δx and Δt to obtain the constant $1/2$. By differentiating (1.6) with respect to x one obtains the usual heat equation

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} \quad (1.7)$$

for the density function $p(t, x)$. Of course, the heat equation was known to Bachelier, and he notes regarding (1.7) *“C'est une équation de Fourier.”*

Bachelier has thus derived, on a formal level, the Kolmogorov forward equation, also known as Fokker-Planck-equation, for the propagation of a probability density p under Brownian motion. The forward equation will also play an important role subsequently and we take the opportunity to note that Bachelier's argument can equally well be applied to the more general process with increments $dX_t = \sigma(t, X_t) dW_t$ to arrive at the PDE

$$\frac{\partial}{\partial t} P = \frac{1}{2} \frac{\partial}{\partial x} (\sigma^2 p), \quad \frac{\partial}{\partial t} p = \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2 p). \quad (1.8)$$

But let us still remain with the form (1.6) of the heat equation and analyze its message in terms of *“horizontal transport of probability measures”*. One may ask: what is the “velocity field”, acting on the set of probabilities on \mathbb{R} , which moves the probability density $p_t(\cdot)$ to the probability density $p_{t+dt}(\cdot)$? Following Bachelier's intuition and keeping in mind that the mass sitting at time t on x equals $p_t(x)$, the velocity of this move at the point x has to be equal to

$$-\frac{1}{2} \frac{\frac{\partial p_t}{\partial x}(x)}{p_t(x)} \quad (1.9)$$

which has the natural interpretation as the “speed” of the horizontal transport induced by $p_t(x)$. We thus encounter *in nuce* the “score function” $\frac{p'_t(x)}{p_t(x)} = \frac{\nabla p_t(x)}{p_t(x)}$ where the nabla notation ∇ indicates that this is a vector field which makes perfect sense in the n -dimensional case too.

At this stage we can relate Bachelier’s work with the more recent notion of the *Wasserstein metric* $\mathcal{W}_2(\cdot, \cdot)$, at least intuitively and at an infinitesimal level. One may ask: what is the necessary kinetic energy needed to transport $p_t(\cdot)$ to $p_{t+dt}(\cdot)$? Knowing the speed (1.9) and the usual formula for the kinetic energy, we obtain for the Wasserstein distance between the two infinitesimally close probabilities p_t and p_{t+dt}

$$\frac{\mathcal{W}_2(p_t, p_{t+dt})}{dt} := \frac{1}{2} \left(\int_{\mathbb{R}} \left(\frac{p'_t(x)}{p_t(x)} \right)^2 dx \right)^{\frac{1}{2}} \quad (1.10)$$

For a formal definition of the Wasserstein distance $\mathcal{W}_2(\cdot, \cdot)$ we refer, e.g., to [57]. While for the finite version of the Wasserstein distance between two probability measures one has to find an optimal transport plan, the situation is simpler – and very pleasant – in the case of the infinitesimal transport induced by the vector field (1.9). This infinitesimal transport is automatically optimal. Intuitively this corresponds to the geometric insight in the one-dimensional case that the transport lines of infinitesimal length cannot cross each other. For a thorough treatment of the geometry of absolutely continuous curves of probabilities such as $(p_t(\cdot))_{t \geq 0}$ above we refer to the lecture notes [3].

We finish the section by returning to Bachelier’s thesis. The *rapporteur* of Bachelier’s dissertation was no lesser a figure than Henri Poincaré. Apparently he was aware of the enormous potential of the section “*Rayonnement de la probabilité*” in Bachelier’s thesis, when he added to his very positive report the handwritten phrase: “*On peut regretter que M. Bachelier n’ait pas développé davantage cette partie de sa thèse.*” That is: One might regret that Mr. Bachelier did not develop further this part of his thesis. Truly prophetic words!

2 Dupire’s Formula

We now turn to a well-known and more recent topic in Mathematical Finance continuing the early achievements of Bachelier.

Suppose that in a financial market we know the prices of “many” European options on a given (highly liquid) stock S . What can we deduce from this data about the prices of exotic, i.e., path-dependent options?

This question leads to the following mathematical idealization: suppose we know the prices of *all* European call options, i.e., the price $C(t, x)$ of every call option with strike price x and maturity t , for every $0 \leq t \leq T$ and $x \in \mathbb{R}_+$. Our task is to analyze the set of all possible (local) martingale measures for the stock price processes which are compatible with this data. Once we have a hand on the relevant set of martingale measures, we can price arbitrary exotic options by taking expectations.

To make the question meaningful, it is a good idea to restrict the class of processes under consideration, e.g., to continuous, Markovian martingales.

We make the economically meaningful assumptions that the function $(t, x) \mapsto C(t, x)$ is sufficiently smooth, as well as strictly convex in the variable x and strictly increasing in the variable t , to allow for the subsequent formal manipulations.

The first observation is that the knowledge of $C(t, x)$, for $0 \leq t \leq T$ and $x > 0$ is tantamount to the knowledge of the marginal probabilities $(\mu_t)_{0 \leq t \leq T}$ of the underlying stock price process

under a martingale measure which determines the prices, via the formula

$$C(t, x) = \mathbb{E}_{\mu_t}[(S_t - x)_+]. \quad (2.1)$$

This observation goes back to the Breeden and Litzenberger [13].

If the measures μ_t are absolutely continuous with respect to Lebesgue measure with a continuous density function $p_t(x)$ then (2.1) amounts to the relation

$$p_t(x) = C_{xx}(t, x), \quad x > 0, \quad (2.2)$$

as one verifies via integration by parts.

In a very influential and highly cited paper from 1994 [18] (compare also the work of Derman and Kani [17]) Dupire considered diffusion processes of the form

$$\frac{dS_t}{S_t} = \sigma(t, S_t) dW_t, \quad 0 \leq t \leq T, \quad (2.3)$$

where the “local volatility” $\sigma(\cdot, \cdot)$ is modeled as a *deterministic* function of t and x , and W_t is a Brownian motion adapted to its natural filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$. It turns out that there is the beautiful and strikingly simple “Dupire’s formula” which relates $\sigma(\cdot, \cdot)$ to the given option prices $C(t, x)$, namely

$$\frac{\sigma^2(t, x)}{2} = \frac{C_t(t, x)}{x^2 C_{xx}(t, x)}. \quad (2.4)$$

Indeed, the Fokker-Planck equation implies – at least on a formal level (cf. (1.8)) – that $p_t(x)$ satisfies the PDE

$$\frac{\partial}{\partial t} p_t(x) = \frac{\partial^2}{\partial x^2} \left(\frac{\sigma^2(t, x)}{2} p_t(x) \right) \quad (2.5)$$

Integrating in x , using (2.2), and changing the order of derivatives quickly yields (2.4).

We note that this beautiful argument is very much in line with Bachelier’s reasoning in (1.6) and (1.7) above pertaining to the case of constant volatility σ . We note in passing that Bachelier used instead of the wording “volatility” the more colorful term “nervousness of the market”.

Of course, Dupire’s formal arguments need proper regularity assumptions in order to be justified. There are two aspects: existence and uniqueness of the martingales fitting the given option prices $C(t, x)$. As regards the former, the question of existence amounts to a remarkable theorem by Kellerer [38, 39]: given a family $(\mu_t)_{0 \leq t \leq T}$ of probability distributions on \mathbb{R} which is increasing in convex order, there is a Markov martingale having these probabilities as marginals. By increasing in convex order we mean that each μ_t has finite first moment and that $\mu_t(f)$ is nondecreasing in t , for every convex function f on \mathbb{R} . Kellerer’s theorem extends earlier work of Strassen [55] who establish the discrete time version of the result. We also note that the convex order condition on the marginal distributions is necessary as easily follows from Jensen’s theorem.

Kellerer’s theorem goes far beyond the simple formula (2.4) and has been further refined, notably by Lowther in an impressive series of papers [47, 46, 48]. We shall review these results in the subsequent sections.

However, from an application point of view, the existence question is not of primordial relevance. After all, the function $C(t, x)$ is an idealization of reality which has to be estimated

from a finite set of given European option prices. In this context, it does not harm to make strong regularity assumptions on the smoothness and convexity (in the variable x) of the function $C(t, x)$ which justify the above argument. Under such assumptions Dupire's solution (2.4) does make sense and the issue of existence is settled.

A different issue is the question of uniqueness. As we shall see below, this question is challenging and relevant – at least from a mathematical point of view – even in very regular settings, such as the Bachelier or the Black-Scholes model.

In order to formulate existence and uniqueness results for a process with given marginals, one has to specify the class of processes with respect to which we want to establish existence and uniqueness. Under proper regularity assumptions the unique solution should, of course, equal Dupire's solution. Dupire's process is a *martingale with continuous paths, enjoying the Markov property*. Is Dupire's solution unique within this class? In a veritable tour de force Lowther [47, 46] has shown that the answer is yes, provided that we replace the word *Markov* by the word *strong Markov* and one restricts to continuous processes.

We also refer to Theorem 6.1 in [30] where a slightly different version of this theorem, credited to Pierre, is proved. These theorems settle the question of uniqueness in a very satisfactory way. We shall discuss Lowther's theorem in more detail in Section 6.

To the best of our knowledge the following question remained open: is it really necessary to add the adjective *strong* to the word *Markov* in Lowther's uniqueness theorem? At least, if one is willing to accept strong regularity assumptions on the function $C(\cdot, \cdot)$ and the resulting process S as defined in (2.3) one may ask whether the Markov property alone is sufficient. We will focus on this question in the next section.

3 An eye-opening example

The subsequent example is known since the work by Dynkin and Jushkevich in the fifties (see [19]).

Example 3.1. *There is an \mathbb{R}_+ -valued, continuous, Markov martingale which fails to be strongly Markovian.*

Proof. We define the process $S = (S_t)_{0 \leq t \leq 1}$ by starting at $S_0 = 1$ and subsequently proceeding in two steps. For $t \in [0, \frac{1}{2}]$ the process S is a stopped geometric Brownian motion, i.e.,

$$S_t = \exp\left(B_{t \wedge \tau} - \frac{t \wedge \tau}{2}\right) \quad 0 \leq t \leq \frac{1}{2}, \quad (3.1)$$

where B is a standard Brownian motion and τ is the first moment when S hits the level 2.

For $\frac{1}{2} \leq t \leq 1$, we distinguish two cases. If S has been stopped, i.e., if $S_{\frac{1}{2}} = 2$, the process S simply remains constant at the level 2. If this is not the case, the process continues to follow a geometric Brownian motion, i.e.,

$$S_t = \exp\left(B_t - \frac{t}{2}\right), \quad \frac{1}{2} \leq t \leq 1 \quad \text{on the set} \quad \left\{\tau > \frac{1}{2}\right\}, \quad (3.2)$$

Obviously S is a continuous martingale. The crucial feature is its Markovian nature: the Markov property follows from the fact that, for every fixed (deterministic) time $\frac{1}{2} \leq t \leq 1$, the probability for the geometric Brownian motion S_t to be equal to 2 is zero on the set $\{\tau > \frac{1}{2}\}$.

Hence, for every fixed $\frac{1}{2} \leq t \leq 1$, the conditional law of $(S_u)_{t \leq u \leq 1}$ is *almost surely* determined by the present value S_t of the process.

Why does S fail to be strongly Markovian? On the set $\{\tau > \frac{1}{2}\}$ define the stopping time ϑ as the first instance $u > \frac{1}{2}$ when S_u equals the value 2, which happens with positive probability during the interval $(\frac{1}{2}, 1)$. The process S therefore, at time $\vartheta \wedge 1$, takes the value 2 on a non-negligible part of the set $\{\tau > \frac{1}{2}\}$. Of course, the random variable S_τ equals 2 on the set $\{\tau \leq \frac{1}{2}\}$ too. Hence there is no strong Markovian prescription for the process S what to do after time ϑ : without further information on the past the process S cannot decide whether it should remain constant or continue to move on as a geometric Brownian motion. \square

Let us apply this example to the pricing of options of the form $(S_T - x)_+$. Fix $x \geq 2$. It is straightforward to calculate its price $P^x(t, z)$ at time t , conditionally on $S_t = z$, which is defined via

$$P^x(t, z) = \mathbb{E}[(S_T - x)_+ | S_t = z], \quad \leq t \leq T, z \in \mathbb{R}. \quad (3.3)$$

Letting $z = 2$, we find

$$P^x(t, z) = 0, \quad 0 < t \leq 1. \quad (3.4)$$

On the other hand, for $z \neq 2$ and $\frac{1}{2} \leq t < 1$, the prices $P^x(t, z)$ are given by the usual Black-Scholes formula and therefore strictly positive. Hence, for $\frac{1}{2} \leq t < 1$, the option prices $z \rightarrow P^x(t, z)$ are discontinuous at $z = 2$. They also fail to be increasing and convex in the variable x which a reasonable option pricing regime should certainly satisfy. On the other hand we note that these option prices – strange as they might be – do not violate the no arbitrage principle as they were legitimately derived from a martingale.

The marginal probabilities of the process S have an atom at the point 2 which is rather unpleasant. One may ask whether it is possible to construct variants of the above example which have more regular marginals.

Here is a straightforward modification: fix an uncountable compact K in \mathbb{R}_+ with zero Lebesgue measure. For example one may take the classical Cantor set $K = \{1 + \sum_{n=1}^{\infty} \frac{\varepsilon_n}{3^n} : \varepsilon_n \in \{0, 2\}\}$. We can modify the construction of Example 3.1 in three steps. For $0 \leq t \leq \frac{1}{3}$, let $(S_t)_{0 \leq t \leq \frac{1}{3}}$ be geometric Brownian motion starting at $S_0 = 1$. For $\frac{1}{3} \leq t \leq \frac{2}{3}$, let $(S_t)_{\frac{1}{3} \leq t \leq \frac{2}{3}}$ continue to be geometric Brownian motion, but stopped at the first hitting time τ of the compact set K . For $\frac{2}{3} \leq t \leq 1$ we again distinguish between the cases $\{\tau \leq \frac{2}{3}\}$ and $\{\tau > \frac{2}{3}\}$. On the former set the process S remains constant, i.e., $S_t = S_{\frac{2}{3}}$. On its complement $\{\tau > \frac{2}{3}\}$ the process S continues to follow a geometric Brownian motion. The process S thus enjoys all the features of Example 3.1 and, in addition, has continuous marginals. Note, however, that these marginals are not given by densities as they are not absolutely continuous with respect to Lebesgue measure.

Turning back to the context of Example 3.1 there is another continuous Markovian martingale with the same marginals as S , inducing reasonable option prices. In fact, there is a continuous *strong* Markovian martingale with this property and which is unique in this latter class (Theorem 4.1 below).

We only give an informal, verbal description of this strong Markov process. For $0 \leq t \leq \frac{1}{2}$ let S be defined as in Example 3.1. For $\frac{1}{2} \leq t \leq 1$ we again distinguish two cases. On the set $\{S_t \neq 2\}$ define S to be geometric Brownian motion, but now we stop this motion when S_t hits the value 2. On the other hand, on the set $\{S_t = 2\}$ the process S starts an excursion from

$S_t = 2$ into the geometric Brownian motion (3.2) with a certain intensity rate. We are free to choose this rate in such a way that the mass remaining at the atom $\{S_t = 2\}$ equals precisely the constant mass which is prescribed by the given marginals of the process S .

We thus have indicated the construction of another continuous martingale having the same marginals as the process S in Example 3.1. One may check that the latter construction is strongly Markovian – as opposed to the above construction in Example 3.1 – and that the option prices are increasing and strictly convex in the variable z as they should be. It will follow from Theorem 4.1 below that the latter martingale is the unique strong Markov solution for the given marginals.

Note that this answers the question raised at the end of the last section: In Lowther’s uniqueness theorem, it is not sufficient to consider Markovian (but not necessarily *strongly* Markovian) martingales: as we have just seen, there exist two distinct continuous Markov martingales with the same 1-dimensional distributions.

In view of Dupire’s formula, this leads to the next question. It seems natural to conjecture that, provided the call prices are sufficiently regular in t and x , there should be only one continuous Markov martingale matching these prices. Correspondingly one would ask: can one obtain a similar example as above with absolutely continuous (or even more regular) marginals?

To the surprise of the present authors it turned out that the answer is “yes”, even when we pass to the “most regular” situation when S is a Brownian motion, i.e., in the Bachelier model (or the Black-Scholes model). The construction is more involved but resting on the above developed intuition, see the accompanying paper [9].

4 Uniqueness of Dupire’s diffusion

There is a huge literature on one-dimensional processes inducing a given family of marginal distributions (see [38, 49, 29, 30, 7, 46, 26, 20, 34, 50, 2, 6, 36] among others). In particular, the late Marc Yor and his co-authors Hirsch, Profeta, and Roynette wrote the beautiful book [28] on “peacocks”. This is a pun on the French acronym PCOC, for “Processus Croissant pour l’Ordre Convexe” and a *peacock* is a stochastic process $(X_t)_{t \geq 0}$ for which the family of laws $\text{law}(X_t), t \geq 0$ is increasing in the convex order. We take here the liberty to use the word peacock also for a family of probabilities $(\mu_t)_{t \geq 0}$ that increases in convex order.¹

To comply with this literature we find it more natural to pass from the multiplicative setting (2.3) to the additive setting of a martingale diffusion

$$dX_t = \sigma(t, X_t) dW_t, \quad 0 \leq t \leq T. \quad (4.1)$$

Hence we consider now processes taking possibly values in all of \mathbb{R} and switch to the notation X instead of the “stock price” S . We note, however, that this change is only for notational reasons, and everything below could also be done in the multiplicative setting of the previous sections.

Given a peacock $(\mu_t)_{t \geq 0}$ we may define option prices via

$$C(t, x) = \mathbb{E}_{\mu_t}[(X_t - x)_+], \quad (4.2)$$

where $\mu_t, t \geq 0$ denote the marginal probability measures and $x \in \mathbb{R}$. The ‘multiplicative’ formula (2.4) becomes in the additive setting

$$\frac{\sigma^2(t, x)}{2} = \frac{C_t(t, x)}{C_{xx}(t, x)}. \quad (4.3)$$

¹Fortunately, “Probabilités Croissant pour l’Ordre Convexe” still yields the same acronym.

We can now cite Lowther’s complete solution of the uniqueness problem within the class of continuous, strong Markov martingales. We stress (and admire) that this theorem does not require any additional regularity assumptions [47, Theorem 1.2].

Theorem 4.1 (Lowther). *Let $X = (X_t)_{0 \leq t \leq 1}$ and $Y = (Y_t)_{0 \leq t \leq 1}$ be \mathbb{R} -valued, continuous, strong Markov martingales. If X and Y have the same marginal distributions they also have the same joint distributions.*

The proof of this theorem is highly technical and its presentation goes far beyond the scope of the present paper. Instead, we formulate a “toy” version of the theorem under strong regularity assumptions. We then analyze why the notion of *strong* Markovianity is key in the above theorem and finally give some hints on the strategy of the proof of Theorem 4.1.

Here is the “toy” version of this theorem imposing strong regularity assumptions which make life easier:

Assumption 4.2. *We suppose that the process X is given by $X_0 = x_0$ and (4.1), where $\sigma(t, x)$ is sufficiently smooth to guarantee that there is a unique strong solution X . We also suppose that X_T has finite second moment and that the function $C(t, x)$ is strictly convex in the variable x , strictly increasing in the variable t , and satisfies standard Itô smoothness assumptions, i.e., twice continuously differentiable in x and once continuously differentiable in t . We also assume that, for every $x \in \mathbb{R}$, the pricing function $(t, z) \rightarrow P^{X,x}(t, z)$ defined via*

$$P^{X,x}(t, z) = \mathbb{E}[(X_T - x)_+ | X_t = z], \quad 0 \leq t \leq T, z \in \mathbb{R}. \quad (4.4)$$

also satisfies these standard Itô assumptions.

These assumptions are strong enough to guarantee that the function $C(t, x)$ indeed satisfies Equation (4.3).

Theorem 4.3. *Let $X = (X_t)_{0 \leq t \leq T}$ satisfy Assumption 4.2. Let $Y = (Y_t)_{0 \leq t \leq T}$ be another continuous Markov martingale such that X_t and Y_t have the same distribution, for every $0 \leq t \leq T$. For fixed strike price $x \in \mathbb{R}$, let $P^{Y,x}(t, z)$ be the corresponding option prices defined via*

$$P^{Y,x}(t, z) = \mathbb{E}[(Y_T - x)_+ | Y_t = z], \quad 0 \leq t \leq T, z \in \mathbb{R}, \quad (4.5)$$

and assume that, for every x , the functions $(t, z) \rightarrow P^{Y,x}(t, z)$ also satisfy the above standard Itô smoothness assumptions. Then $P^{X,x}(t, z) = P^{Y,x}(t, z)$, for all t, x, z and the processes X and Y have the same joint distributions.

Proof. As the function $(t, z) \rightarrow P^{Y,x}(t, z)$ is **assumed** to satisfy the standard Itô conditions we may apply Itô’s formula to obtain

$$dP^{Y,x}(t, Y_t) = P_t^{Y,x}(t, Y_t) dt + P_z^{Y,x}(t, Y_t) dY_t + \frac{1}{2} P_{zz}^{Y,x}(t, Y_t) d\langle Y \rangle_t, \quad (4.6)$$

where $\langle Y \rangle_t$ denotes the quadratic variation process of the continuous, square integrable martingale Y (see [42, Theorem 7.6.4]). By (4.5) the process $(P^{Y,x}(t, Y_t))_{0 \leq t \leq 1}$ is a martingale. The martingale condition implies that the drift term vanishes so that the equality $\frac{1}{2} P_{zz}^{Y,x}(t, Y_t) d\langle Y \rangle_t = -P_t^{Y,x}(t, Y_t) dt$ holds true in the following sense:

$$\frac{1}{2} \int_A P_{zz}^{Y,x}(t, Y_t) d\langle Y \rangle_t = - \int_A P_t^{Y,x}(t, Y_t) dt, \quad (4.7)$$

whenever we integrate over a predictable set $A \subseteq [0, T] \times \mathcal{C}[0, T]$. In particular, the finite measure $d\langle Y \rangle_t$ on the predictable sigma-algebra of subsets of $[0, T] \times \mathcal{C}[0, T]$ has to be absolutely continuous with respect to the measure $d\langle X \rangle_t = \sigma^2(t, \omega_t) dt$. Here we have used the assumption that $t \rightarrow C(t, z)$ is strictly increasing in t so that $\frac{\sigma^2(t, z)}{2} = \frac{C_t(t, z)}{C_{zz}(t, z)}$ is strictly positive. We thus may define the function

$$\frac{\rho^2(t, z)}{2} := \frac{P_t^{Y, x}(t, z)}{P_{zz}^{Y, x}(t, z)}$$

so that $d\langle Y \rangle_t = \rho^2(t, Y_t) dt$. We therefore must have that Y may be represented as in (4.1), with σ replaced by ρ . This implies the relation

$$C_t(t, x) = \frac{\rho^2(t, x)}{2} C_{xx}(t, x). \quad (4.8)$$

Comparing with (4.3) we obtain $\rho^2 = \sigma^2$ which shows the identity of X and Y in distribution. \square

The above theorem provides a sufficient set of regularity assumptions to substantiate the statement in Dupire's paper [18]: "...we can recover, up to technical regularity assumptions, a unique diffusion process".

Of course, one could do some massaging of the above argument to somewhat weaken the very strong Assumptions 4.2 which we have imposed. But there is a long and thorny road, going far beyond simple cosmetic changes, to arrive at Lowther's Theorem 4.1.

In Theorem 4.3 the strong regularity assumptions implied in particular the strong Markov property of the process X . We stress once more that, in the setting of Lowther's theorem 4.1, the *strong* Markov property is the key assumption.

Passing to Lowther's notation and looking at (4.4), a crucial step in the above argument is to start from a convex, increasing, and Lipschitz-one function $g(z)$, such as $g(z) = (z - x)_+$, to its conditional expectations

$$f(t, z) = \mathbb{E}[g(Y_T) | Y_t = z], \quad 0 \leq t \leq T, z \in \mathbb{R}. \quad (4.9)$$

In order to start a chain of arguments one has to verify that $f(t, z)$ is a "nice" function. When looking at Example 3.1 and its variants we have seen that in this case, for $g(z) = (z - x)_+$, this is not at all the case. Its conditional expectation $f(t, z)$ lacked each of the following desired properties: continuity, monotonicity and convexity in x .

Contrary to this lamentable breakdown of regularity, we shall verify in Corollary 5.3 that the strong Markov property guarantees that the following three properties are inherited from $g(\cdot)$ to each $f(t, \cdot)$: convexity, monotonicity, and 1-Lipschitz continuity (which serves as a more quantitative version of continuity). This preservation of regularity is a decisive feature of Lowther's proof.

5 Coupling strong Markov processes

What is the salient property which distinguishes the strong Markov property from the Markov property in our context? While the former condition allows for Lowther's uniqueness theorem to hold true, in Example 3.1 we have seen that there may be different continuous Markov martingales inducing the same marginals. The following well-known concept is the key to understanding the difference.

Definition 5.1. For probability measures π^1 and π^2 on \mathbb{R} , we say that π^2 dominates π^1 of first order if, for every $a \in \mathbb{R}$, we have $\pi^2[a, \infty[\geq \pi^1[a, \infty[$.

We shall show in the next proposition that the strong Markov property of a continuous martingale implies that the transition probabilities $(\pi_x^{s,t})_{x \in \mathbb{R}}$

$$\pi_x^{s,t}[A] = \mathbb{P}[X_t \in A | X_s = x], \quad (5.1)$$

where $s < t$ and A is a Borel set in \mathbb{R} , are increasing of first order in the variable x , for every $s < t$.

We follow Hobson who applied a well-known technique, namely the ‘‘joys of coupling’’ (to quote his paper [33]) in the present context.

Proposition 5.2. Let $X = (X_t)_{0 \leq t \leq T}$ be a continuous strong Markov process with transition probabilities $\pi_x^{s,t}[\cdot]$. Then, for $0 \leq s \leq t \leq T$, and for $x < y$ the probability $\pi_y^{s,t}$ dominates $\pi_x^{s,t}$ of first order.

Proof. Fix s, t and $x < y$ as above and let $(X_u^x)_{s \leq u \leq t}$ and $(X_u^y)_{s \leq u \leq t}$ be independent copies of the process X , starting at $X_s^x = x$ and $X_s^y = y$, both defined on the same filtered probability base. Define the stopping time τ as the first moment u when the process X_u^x equals the process X_u^y , if this happens for some $u \in [s, t]$; otherwise we let $\tau = \infty$.

Define the process \tilde{X}^x by

$$\tilde{X}_u^x = \begin{cases} X_u^x & \text{for } s \leq u \leq \tau, \\ X_u^y & \text{for } \tau < u \leq t. \end{cases}$$

We clearly have

$$\tilde{X}_t^x \leq X_t^y, \quad \text{for all } 0 \leq t \leq T. \quad (5.2)$$

Indeed, if $\tau = \infty$, the paths of $(X_u^x)_{s \leq u \leq t} = (\tilde{X}_u^x)_{s \leq u \leq t}$ and $(X_u^y)_{s \leq u \leq t}$ never touch, so that we even have a strict inequality by continuity of the processes. If $\tau < \infty$, then \tilde{X}^x and X^y have ‘‘joined’’ at time τ , and follow the same trajectory. Hence $\tilde{X}_u^x = X_u^y$, for $\tau \leq u \leq t$.

Inequality 5.2 implies that the law of X_t^y dominates the law of \tilde{X}_t^x in first order. \square

Corollary 5.3. Let $X = (X_t)_{0 \leq t \leq T}$ be a continuous strong Markov process with marginal laws μ_t and transition probabilities $\pi_x^{s,t}[\cdot]$. Let $0 \leq s \leq t \leq T$ and $z \mapsto g(z)$ be a measurable μ_t -integrable function, and define the conditional expectation similarly as in (4.9)

$$f(s, x) = \mathbb{E}[g(X_t) | X_s = x], \quad x \in \mathbb{R}. \quad (5.3)$$

Then the following assertions hold true.

- (i) If $z \mapsto g(z)$ is increasing, then so is $x \mapsto f(s, x)$, for every $0 \leq s \leq t \leq T$.
- If, in addition, we assume that X is a martingale, we also have the following two assertions.
 - (ii) If $z \mapsto g(z)$ is 1-Lipschitz, then so is $x \mapsto f(s, x)$, for every $0 \leq s \leq t \leq T$.
 - (iii) If $z \mapsto g(z)$ is convex, then so is $x \mapsto f(s, x)$, for every $0 \leq s \leq t \leq T$.

Proof. Assertion (i): this is just a reformulation of Proposition 5.2.

Assertion (ii): If g is 1-Lipschitz, then $g(x) + x$ is increasing. As X is a martingale we have $\mathbb{E}[X_t | X_s = x] = x$. By (i) $f(s, x) + x$ is increasing. By the same token $f(s, x) - x$ is decreasing which readily shows that $x \mapsto f(s, x)$ is Lipschitz.

Assertion (iii): We follow the proof of [33, Theorem 3.1]. For convex g and $x < y < z$ we have to show that

$$(z - x)f(s, y) \leq (z - y)f(s, x) + (y - x)f(s, z). \quad (5.4)$$

Fix $x < y < z$ and choose three independent copies X^x, X^y, X^z of the process X , starting at time s from the initial values x, y , and z . To simplify notation we denote the resulting triple of processes (X^x, X^y, X^z) by (X, Y, Z) . We define coupling times similarly as above. Let τ^x be the first moment $u > s$ when $X(u) = Y(u)$; similarly τ^z is defined as the first moment when Y and Z meet. Finally, let $\tau = \tau^x \wedge \tau^z \wedge t$. This time we leave the processes unchanged; rather we shall argue on the three disjoint (up to null sets) sets $\{\tau = \tau^x\}$, $\{\tau = \tau^z\}$, and $\{\tau = t\}$.

We start with the latter on which we have $X_t < Y_t < Z_t$. By the convexity of g we have

$$(Z_t - X_t)g(Y_t) \leq (Z_t - Y_t)g(X_t) + (Y_t - X_t)g(Z_t). \quad (5.5)$$

so that

$$\mathbb{E}[(Z_t - X_t)g(Y_t) - ((Z_t - Y_t)g(X_t) + (Y_t - X_t)g(Z_t)); \tau = t] \leq 0. \quad (5.6)$$

On $\{\tau = \tau^x\}$ we have $X_t = Y_t$ so that the last term in inequality (5.5) vanishes. Also the first and the middle term are equal so that (5.5) holds true (with equality) on the set $\{\tau = \tau^x\}$. In particular,

$$\mathbb{E}[(Z_t - X_t)g(Y_t) - ((Z_t - Y_t)g(X_t) + (Y_t - X_t)g(Z_t)); \tau = \tau^x] \leq 0. \quad (5.7)$$

The same reasoning applies to $\{\tau = \tau^z\}$.

$$\mathbb{E}[(Z_t - X_t)g(Y_t) - ((Z_t - Y_t)g(X_t) + (Y_t - X_t)g(Z_t)); \tau = \tau^z] \leq 0. \quad (5.8)$$

Summing over these three sets we obtain

$$\mathbb{E}[(Z_t - X_t)g(Y_t) - ((Z_t - Y_t)g(X_t) + (Y_t - X_t)g(Z_t))] \leq 0. \quad (5.9)$$

Finally we use independence and the martingality of X, Y , and Z to obtain

$$(z - x)\mathbb{E}[g(Y_t)|Y_s = y] \leq (z - y)\mathbb{E}[g(X_T)|X_s = x] + (y - x)\mathbb{E}[g(Z_t)|Z_s = z], \quad (5.10)$$

which is tantamount to (5.4). \square

We can reformulate the message of Corollary 5.3 (ii) in the spirit of Bachelier by considering the Wasserstein cost $\mathcal{W}_1(\pi_x^{s,t}[\cdot], \pi_y^{s,t}[\cdot])$ of the horizontal transport of the conditional probability measures $\pi_x^{s,t}[\cdot]$ to $\pi_y^{s,t}[\cdot]$.

Recall that, for probabilities μ, ν on the real line the Wasserstein-1 distance is given by

$$\mathcal{W}_1(\mu, \nu) := \inf_{\pi \in \text{cpl}(\mu, \nu)} \int |x - y| d\pi(x, y),$$

where $\text{cpl}(\mu, \nu)$ denotes the set of all probabilities on \mathbb{R}^2 having μ, ν as marginal measures, see e.g. [57] for an extensive overview of the field of optimal transport.

Definition 5.4. *Let π be a probability on \mathbb{R}^2 and write μ for its projection onto the first coordinate and $(\pi_x)_x$ for the respective disintegration so that $\pi = \int_{\mathbb{R}} \pi_x d\mu(x)$. Then π is called a Lipschitz-kernel if for all x, y in a set X with $\mu(X) = 1$*

$$\mathcal{W}_1(\pi_x, \pi_y) \leq |x - y|. \quad (5.11)$$

We call π a *martingale coupling* if $\int y d\pi_x(y) = x$, μ -a.s. It is then straight-forward to see that for a martingale coupling π the following are equivalent:

- (i) π is a Lipschitz-kernel.
- (ii) for all x, y in a set X with $\mu(X) = 1$ we have $\mathcal{W}_1(\pi_x, \pi_y) = |x - y|$.
- (iii) for all x, y in a set X with $\mu(X) = 1$, $x \leq y$ the measure π_x is dominated in first order by π_y .

Definition 5.5. *Let X be an \mathbb{R} -valued Markovian process. Then X has the Markov Lipschitz property if for all $s \leq t$, the law of (X_s, X_t) is a Lipschitz kernel.*

To give yet another characterisation of Lipschitz-Markov processes, recall that a process X is Markov if and only if for all $s \leq t$ and every bounded measurable function f there is a measurable function g such that

$$\mathbb{E}[f(X_s)|\mathcal{F}_s] = g(X_t).$$

A process X is Lipschitz-Markov if and only if for all $s \leq t$ and every 1-Lipschitz function f there is a 1-Lipschitz function g such that

$$\mathbb{E}[f(X_s)|\mathcal{F}_s] = g(X_t).$$

This is a straightforward consequence of the Kantorovich-Rubinstein theorem that provides a dual characterization of the Wasserstein-1 distance through 1-Lipschitz functions.

We now can resume the crucial role of the strong Markov property

Corollary 5.6. *Let M be a continuous Markov martingale. Then M is Lipschitz-Markov if and only if it is strong Markov.*

Proof. By Proposition 5.2 / Corollary 5.3 every strong Markov-martingale is Lipschitz Markov. That a Lipschitz Markov martingale is strongly Markov is proved in the same way as one establishes the strong Markov property for Feller processes. See e.g. [45, Theorem 1.68]. \square

To the best of our knowledge, Lipschitz kernels play a crucial role in all known proofs of Kellerer's theorem. The decisive property is the following:

Proposition 5.7. *Consider the space of $\mathcal{P}(\mathcal{D}[0, 1])$ of probability measures on the Skorokhod space equipped with the convergence of finite dimensional distributions. Then the set of Lipschitz-Markov martingales is closed.*

In contrast, the set of Markov-martingales is not closed. See e.g. [7] for the (simple) proof of Proposition 5.7.

6 Continuity of the martingale solution

An important question in the present context is the following: under which conditions on a peacock $(\mu_t)_{0 \leq t \leq 1}$, as defined in Section 4 above, is there a strong Markov martingale with *continuous trajectories* having the given marginals. We only focus on the one-dimensional case and mention that the corresponding question for higher dimensions remains wide open. The one-dimensional case, however, is fully understood by now, again by the definitive work of Lowther [46, Theorem 1.3]:

Theorem 6.1 (Lowther). *Let $(\mu_t)_{t \geq 0}$ be a peacock and assume that $t \mapsto \mu_t$ is weakly continuous and that each μ_t has convex support. Then there exists a unique strongly continuous Markov martingale X such that $X_t \sim \mu_t, t \geq 0$.*

We do not show Lowther's theorem in full generality, but again we want to isolate a sufficient set of assumptions that allows us to present a (comparably simple) self-contained proof of the continuity theorem.

Remark 6.2. *A key ingredient of the proof is that for probabilities μ, ν in convex order there exists a continuous martingale $(X_t)_{0 \leq t \leq 1}$, $X_0 \sim \mu$, $X_1 \sim \nu$ which is strongly Markovian (and hence Lipschitz Markov). For instance we can take X to be a stretched Brownian motion, that is, a solution to a continuous time martingale transport problem, see [5]. Another possibility would be to apply an appropriate deterministic time change to Root's solution [53] (see [16] for the case of a non-trivial starting distribution) of the Skorokhod embedding problem. We note that the martingale transport approach is also applicable to measures μ, ν defined on $\mathbb{R}^d, d > 1$. This could be interesting in view of a possible multi-dimensional extension of Lowther's theorem but this is not within the scope of the present article.*

Assumption 6.3. *Let $(\mu_t)_{0 \leq t \leq 1}$ be a one-dimensional peacock centered at zero, with densities $p_t(x)$ and finite second moments $m_2^2(\mu_t) = \int_{-\infty}^{\infty} x^2 d\mu_t(x) = \int_{-\infty}^{\infty} x^2 p_t(x) dx$ such that the function $t \rightarrow m_2^2(\mu_t)$ is continuous.*

We assume that there is a – bounded or unbounded – open interval $I \subseteq \mathbb{R}$ supporting each μ_t such that, for each compact subset $K \subseteq I$, the Lebesgue densities $p_t(x)$ of μ_t are bounded away from zero, uniformly in $x \in K$ and $t \in [0, 1]$.

It will be convenient to suppose (w.l.g. via a deterministic time change) that $t \rightarrow m_2^2(\mu_t)$ is affine. More precisely we may assume that $m_2^2(\mu_{t+h}) - m_2^2(\mu_t) = h$ so that for all martingales M with $\text{law}(M_t) = \mu_t$ and $\text{law}(M_{t+h}) = \mu_{t+h}$

$$\mathbb{E} \left[(M_{t+h} - M_t)^2 \right] = h. \tag{6.1}$$

Theorem 6.4. *Under the above assumptions there is a continuous strong Markov martingale $M = (M_t)_{0 \leq t \leq 1}$ with given marginals $(\mu_t)_{0 \leq t \leq 1}$.*

There is an obvious and well-known strategy for the proof. We want to obtain the desired martingale M as a limit of approximations which fit the peacock $(\mu_t)_{0 \leq t \leq 1}$ on finitely many points of time. As in [30] it is convenient to do so along the ordered set \mathcal{S} of finite partitions $S = \{s_0, \dots, s_n\}$ of $[0, 1]$, where $0 = s_0 < \dots < s_n = 1$.

For each $S \in \mathcal{S}$ we choose a continuous strong Markov martingale M^S having the given marginals at each time $s_i \in S$, the existence of M^S is a direct consequence of Remark 6.2.

Identifying the martingales M^S with their induced measures on the path space $\mathcal{C}[0, 1]$ this family of measures is tight, if considered on Skorohod space $\mathcal{D}[0, 1]$, equipped with the topology of convergence of finite dimensional distributions. Hence we can find a cluster point M in the set $\mathcal{P}(\mathcal{D}[0, 1])$ of probability measures on $\mathcal{D}[0, 1]$, see e.g. [7] for the straightforward argument. By refining the filter \mathcal{S} we may suppose that M is a limit point.

We fix such a limiting process M which by Proposition 5.7 is a Lipschitz Markov martingale. These arguments again are standard by now and, e.g., well presented in the papers [30, 7]. A

priori the martingale M has càdlàg trajectories. Our present task is to show the *continuity* of the trajectories of the limiting process M under the above assumptions.

We first give a general criterion for the continuity of a limiting martingale M which is somewhat reminiscent of the classical Kolmogorov continuity criterion [51, Theorem 1.8].

Proposition 6.5. *Let $(M^i)_{i \in I}$ be a net of \mathbb{R} -valued continuous strong Markov martingales $M^i = (M^i)_{0 \leq t \leq 1}$ and M its limit in the set of probabilities on Skorohod space $\mathcal{P}(\mathcal{D}[0, 1])$ with respect to convergence in finite dimensional distribution. Suppose that there are constants $C_1 > 0$ and $\beta > 0$ such that, for every $0 \leq t_0 < t_0 + h \leq 1$ and every $i \in I$*

$$\|(M_t^i)_{t_0 \leq t \leq t_0+h}\|_{BMO_1} := \sup_{\tau} \{ \|\mathbb{E}[|M_{t_0+h}^i - M_{\tau}^i| \mid \mathcal{F}_{\tau}]\|_{\infty} \} \leq C_1 h^{\beta}. \quad (6.2)$$

Then the martingale M has continuous trajectories.

In (6.2) τ runs through the $[t_0, t_0 + h]$ -valued stopping times with respect to the natural filtration of $(M_t^i)_{t_0 \leq t \leq t_0+h}$. As M^i is strong Markov, condition (6.2) is tantamount to the requirement that the first moments $m_1(\pi_x^{i, \tau, t_0+h}) = \int_{-\infty}^{\infty} |y - x| d\pi_x^{i, \tau, t_0+h}(y)$ of the transition probabilities π_x^{i, τ, t_0+h} from $M_{\tau}^i = x$ to $M_{t_0+h}^i$ satisfy

$$m_1(\pi_x^{i, \tau, t_0+h}) = \int_{-\infty}^{\infty} |y - x| d\pi_x^{i, \tau, t_0+h}(y) \leq C_1 h^{\beta}, \quad (6.3)$$

for μ_{τ}^i -almost all $x \in \mathbb{R}$, where μ_{τ}^i denotes the law of M_{τ}^i .

An important feature of the BMO norms for *continuous* martingales is that, by the John-Nirenberg inequality, all BMO_q norms are equivalent, for $1 \leq q < \infty$ (see, e.g. [37, Corollary 2.1]). Applying this fact to the present context, inequality (6.2) is equivalent to the existence of a constant $C_q > 0$ (for some or, equivalently, for all $1 \leq q < \infty$) such that

$$\|(M_t^i)_{t_0 \leq t \leq t_0+h}\|_{BMO_q} := \sup_{\tau} \left\{ \left\| \mathbb{E}[|M_{t_0+h}^i - M_{\tau}^i|^q \mid \mathcal{F}_{\tau}]^{\frac{1}{q}} \right\|_{\infty} \right\} \leq C_q h^{\beta}. \quad (6.4)$$

Proof of Proposition 6.5. Suppose that M fails to be continuous and let us work towards a contradiction to (6.4) when $q > \frac{1}{\beta}$. Assume M has jumps of size bigger than $3a > 0$ with probability bigger than $\kappa > 0$, i.e.

$$\mathbb{P}(\{\exists t \in [0, 1]: |M_t - M_{t-}| \geq 3a\}) > \kappa.$$

As M has càdlàg paths, there is $h_0 > 0$ such that for all $0 < h \leq h_0$

$$\mathbb{P}(\{\exists k \in \mathbb{N}: |M_{(kh) \wedge 1} - M_{((k-1)h) \wedge 1}| \geq 2a\}) > \kappa.$$

By the pigeon hole principle, we can find for each $0 < h \leq h_0$ a time $t_0 \in [0, 1]$ with

$$\mathbb{P}(\{|M_{(t_0+h) \wedge 1} - M_{t_0}\}| \geq 2a\}) > h\kappa. \quad (6.5)$$

In view of the convergence in finite dimensional distributions of $(M^i)_{i \in I}$ to M , we find for each $0 < h \leq h_0$ a time $t_0 \in [0, 1]$ (w.l.g. $t_0 + h \leq 1$) and an index $i \in I$ with

$$\mathbb{P}(\{|M_{t_0+h}^i - M_{t_0}^i| \geq a\}) \geq h\kappa. \quad (6.6)$$

Fixing such an index $i \in I$, it follows that there is a set $A \subseteq \mathbb{R}$ of positive measure with respect to the law of $M_{t_0}^i$ such that, for all $x \in A$,

$$\mathbb{P}(\{|M_{t_0+h}^i - M_{t_0}^i| \geq a \mid M_{t_0}^i = x\}) \geq h\kappa. \quad (6.7)$$

For $x \in A$ we therefore have

$$m_q(\pi_x^{i,t_0,t_0+h}) = \left(\int_{-\infty}^{\infty} |y-x|^q d\pi_x^{i,t_0,t_0+h}(y) \right)^{\frac{1}{q}} \geq a(h\kappa)^{\frac{1}{q}}. \quad (6.8)$$

As $q > \frac{1}{\beta}$ we can choose $0 < h \leq h_0$ sufficiently small, with $a(h\kappa)^{\frac{1}{q}} > C_q h^\beta$, and arrive at the desired contradiction to (6.4)

$$C_q h^\beta \geq \|(M_t^i)_{t_0 \leq t \leq t_0+h}\|_{BMO_q} \geq m_q(\pi_x^{i,t_0,t_0+h}) \geq a(h\kappa)^{\frac{1}{q}} > C_q h^\beta.$$

□

Turning back to the above defined family $(M^S)_{S \in \mathcal{S}}$ of martingales we shall establish an inequality of the type (6.3) for the transition probabilities π_x^{S,t_0,t_0+h} , using the fact that the functions $x \mapsto \pi_x^{S,t_0,t_0+h}$ are Lipschitz-Markov.

Lemma 6.6. *Let $(\mu_t)_{0 \leq t \leq 1}$ be a peacock satisfying Assumption 6.3. Fix a compact set $K \subseteq I$. For all $h > 0$ sufficiently small and $x \in K$, there is a constant $D > 0$ such that for all $S \in \mathcal{S}$ and $0 \leq t_0 \leq t_0 + h \leq 1$, with $t_0, t_0 + h \in \mathcal{S}$, the first moments of the transition measures π_x^{S,t_0,t_0+h} can be estimated by*

$$m_1\left(\pi_x^{S,t_0,t_0+h}\right) := \mathbb{E}_{\mathbb{P}}\left[|M_{t_0+h}^S - M_{t_0}^S| \mid M_{t_0}^S = x\right] = \int_{-\infty}^{\infty} |y-x| d\pi_x^{S,t_0,t_0+h}(y) \leq Dh^{\frac{1}{4}}, \quad (6.9)$$

where \mathbb{P} denotes the law of the martingale M^S .

Proof. We first suppose that $I = \mathbb{R}$. By (6.1) and Jensen's inequality we have

$$\mathbb{E}_{\mathbb{P}}[|M_{t_0+h}^S - M_{t_0}^S|] \leq h^{\frac{1}{2}}.$$

We may rewrite this inequality in the form

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[|M_{t_0+h}^S - M_{t_0}^S|] &= \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[|M_{t_0+h}^S - M_{t_0}^S| \mid M_{t_0}^S]] = \int_{-\infty}^{\infty} m_1\left(\pi_x^{S,t_0,t_0+h}\right) d\mu_{t_0}(x) \\ &= \int_{-\infty}^{\infty} F(x) d\mu_{t_0}(x) \leq h^{\frac{1}{2}}, \end{aligned} \quad (6.10)$$

where, alleviating the notation from π_x^{S,t_0,t_0+h} to π_x , and

$$F(x) = \int_{-\infty}^{\infty} |x-y| d\pi_x(y).$$

Claim: The function $x \mapsto F(x)$ satisfies the estimate

$$|F(x) - F(x+k)| \leq 2k, \quad x \in \mathbb{R}, \quad k > 0. \quad (6.11)$$

Indeed,

$$\begin{aligned} |F(x+k) - F(x)| &= \left| \int_{-\infty}^{\infty} |y - (x+k)| d\pi_{x+k}(y) - \int_{-\infty}^{\infty} |y - x| \pi_x(y) \right| \\ &\leq k + \left| \int_{-\infty}^{\infty} |y - (x+k)| d\pi_x(y) - \int_{-\infty}^{\infty} |y - x| d\pi_x(y) \right| \leq 2k \end{aligned}$$

proving the claim. In the first inequality we have used the fact that π is a Lipschitz Markov kernel.

Choose a compact interval $[a, b]$ such that $K \subseteq [a+1, b-1]$ and denote by $l > 0$ a lower bound for the density function p_{t_0} of μ_{t_0} on $[a, b]$.

We want to estimate $F(x_0)$, for $x_0 \in K$, and start by showing the rough estimate $F(x_0) \leq 2$. Using (6.11) we otherwise have $\int_{x_0}^{x_0+1} F(x) dx \geq 1$ and we arrive at the following contradiction to (6.10) for small enough h

$$1 \leq \int_{x_0}^{x_0+1} F(x) dx \leq \frac{1}{l} \int_{x_0}^{x_0+1} F(x) dp_{t_0}(x) dx \leq \frac{1}{l} h^{\frac{1}{2}}.$$

If $F(x_0) \leq 2$ we may argue similarly, using again (6.11) and elementary geometry, to obtain, for $x_0 \in K$,

$$\frac{1}{8}(F(x_0))^2 \leq \int_{x_0}^{x_0+1} F(x) dx \leq \frac{1}{l} \int_{x_0}^{x_0+1} F(x) d\mu_{t_0}(x) \leq \frac{1}{l} h^{\frac{1}{2}},$$

yielding the desired estimate for h sufficiently small

$$\sup_{x \in K} m_1(\pi_x) = \sup_{x \in K} F(x) \leq Dh^{\frac{1}{4}},$$

where the constant $D > 0$ only depends on the compact set K , but not on h .

Finally we have to come back to our assumption $I = \mathbb{R}$ which allowed us to imbed the compact set K into the interval $[a, b] \subseteq I$ such that $K \subseteq [a+1, b-1]$. If I is only one- or two-sided bounded we have to reason slightly more carefully, as we can imbed the compact set K only into an interval $[a, b] \subseteq I$ such that $[a+\varepsilon, b-\varepsilon]$ contains K . But no difficulties arise from replacing 1 by ε and it is straightforward to adapt the above argument also to this situation. \square

Under the assumptions of Lemma 6.6, Tschebyscheff's inequality and (6.9) allow us to control the difference between medians and means, since for fixed $0 < \delta < \frac{1}{4}$

$$\pi_x^{S, t_0, t_0+h} \left(y: |y - x| \geq h^\delta \right) \leq Dh^{\frac{1}{4}-\delta}$$

for $h > 0$ sufficiently small, $x \in K$ and feasible $S \in \mathcal{S}$. Hence, we have in the setting of Lemma 6.6 that

$$\left| \text{median}(\pi_x^{S, t_0, t_0+h}) - \text{mean}(\pi_x^{S, t_0, t_0+h}) \right| \leq h^\delta. \quad (6.12)$$

Lemma 6.7. *Let $0 < \delta < \frac{1}{4}$. Under the assumptions of the Lemma 6.6 the same conclusion as in (6.9) holds true for every $[t_0, t_0+h]$ -valued stopping time τ (by possibly changing the constant D to a different C) for $x \in K$*

$$m_1 \left(\pi_x^{S, \tau, t_0+h} \right) := \mathbb{E}_{\mathbb{P}} \left[\left| M_{t_0+h}^S - M_\tau^S \right| \mid M_\tau^S = x \right] = \int_{-\infty}^{\infty} |y - x| d\pi_x^{S, \tau, t_0+h}(y) \leq Ch^\delta. \quad (6.13)$$

Proof. By Corollary 5.6 we have that $x \mapsto \pi^{S,t_0,t}$ is Lipschitz Markov for all $s \in [t_0, t_0 + h]$. From this, we can deduce the continuity of the map

$$(x, t) \mapsto \int_{-\infty}^{\infty} |z - y| d\pi^{S,t,t_0+h}(z).$$

Therefore, it suffices to show (6.13) for deterministic $\tau \equiv t \in [t_0, t_0 + h]$ due to the strong Markov property. To this end, let \tilde{K} be a compact interval in I , containing the compact set K in its interior and fix the constant D from Lemma 6.6, applied to \tilde{K} .

To do so, find $x \in I$ such that y equals the median of the measure $\pi_x^{S,t_0,t}$, that is

$$\pi_x^{S,t_0,t}(-\infty, y] \leq \frac{1}{2} \leq \pi_x^{S,t_0,t}(-\infty, y).$$

Since $x \mapsto \pi_x^{S,t_0,t}$ is Lipschitz Markov by Corollary 5.6, such an x exists. We may use x to obtain the following estimate

$$\begin{aligned} & \mathbb{E}_{\mathbb{P}} \left[\left| M_{t_0+h}^S - y \right| \middle| M_{\tau}^S = y \right] \\ & \leq \mathbb{E}_{\mathbb{P}} \left[(M_{t_0+h}^S - y)_+ \middle| M_{\tau}^S \geq y, M_{t_0}^S = x \right] + \mathbb{E}_{\mathbb{P}} \left[(M_{t_0+h}^S - y)_- \middle| M_{\tau}^S \leq y, M_{t_0}^S = x \right] \\ & \leq 2\mathbb{E}_{\mathbb{P}} \left[\left| M^{S,t_0,t_0+h} - y \right| \middle| M_{t_0}^S = x \right] \\ & \leq 2\mathbb{E}_{\mathbb{P}} \left[\left| M^{S,t_0,t_0+h} - x \right| \middle| M_{t_0}^S = x \right] + 2|x - y|. \end{aligned} \quad (6.14)$$

Here, we used the first item of Corollary 5.3 for the first inequality and y being the median of $\pi_x^{S,t_0,\tau}$ for the second. Note that for h sufficiently small we have by (6.12) that $x \in \tilde{K}$. Applying the estimates (6.9) and (6.12) to (6.14) we find

$$\mathbb{E}_{\mathbb{P}} \left[\left| M_{t_0+h}^S - y \right| \middle| M_{\tau}^S = y \right] \leq 2Dh^{\frac{1}{4}} + 2h^{\delta} \leq (2D + 2)h^{\delta},$$

which concludes the proof. \square

Proof of Theorem 6.4. We have to combine Proposition 6.5 and Lemma 6.7 with a stopping argument. To this end, let $(K_n := [a_n, b_n])_{n \in \mathbb{N}}$ be an increasing sequence of compact intervals exhausting the interval $I =]a, b[$, where $a, b \in [-\infty, \infty]$. We let $K_1 = \emptyset$. Define the stopping times $\tau^{S,n}$ as the first moment when the continuous martingale M^S leaves K_n , write $M^{S,n}$ for the stopped process, and denote the differences $M^{S,n+1} - M^{S,n}$ by $\Delta M^{S,n}$.

For the processes $\Delta M^{S,n}$ the assumptions of Proposition 6.5 still hold true as a consequence of Lemma 6.7 and the strong Markov property of M^S . Consider the process

$$\tilde{M}_t^S := (\Delta M_t^{S,n})_{n \in \mathbb{N}}, \quad t \in [0, 1],$$

taking values in $\mathbb{R}^{\mathbb{N}}$. Moreover, let

$$m^S := \inf \left\{ k \in \mathbb{N} : \left| \Delta M^{S,k} \right|_{\infty} = 0 \right\}$$

be the smallest integer such that the entire trajectory of $(M_t^S(\omega))_{0 \leq t \leq 1}$ is contained in K_{m^S} . We know already that for every $n \in \mathbb{N}$, $M^{S,n}$ (and therefore $\Delta M^{S,n}$) allows for f.d.d. convergent subnets along \mathcal{S} where the limits have by Proposition 6.5 continuous versions. We want to argue that similarly the pair $(\tilde{M}^S, m^S)_{S \in \mathcal{S}}$ taking values in $\mathcal{C}[0, 1]^{\mathbb{N}} \times \mathbb{N}$ admits a convergent subnet, too. For this reason, it is sufficient to show the following.

Claim: For every $\varepsilon > 0$ there is $n \in \mathbb{N}$ such that, for every $S \in \mathcal{S}$,

$$\sup_{S \in \mathcal{S}} \mathbb{P}(m^S > n) < \varepsilon. \quad (6.15)$$

Indeed, for each $n \in \mathbb{N}$ (sufficiently large), there are maximal positive numbers α_n, β_n with $\alpha_n + \beta_n \leq 1$ such that the probability measure

$$\alpha_n \delta_{a_n} + \beta_n \delta_{b_n} + (1 - \alpha_n - \beta_n) \delta_{\frac{\text{mean}(\mu_1) - \alpha_n a_n - \beta_n b_n}{1 - \alpha_n - \beta_n}}$$

is dominated by μ_1 in convex order. Since μ_1 puts no mass to the boundary of I , there is for each $\varepsilon > 0$ an index $N \in \mathbb{N}$ such that $\alpha_n + \beta_n < \varepsilon$ for all $n \geq N$. The law of $M^{S,n}$ is dominated in the convex order by μ_1 . By maximality of α_n and β_n we find uniformly for all $S \in \mathcal{S}$

$$\mathbb{P}(m^S > n) = \mathbb{P}(\tau^{S,n} < \infty) = \mathbb{P}(M^{S,n} \in \{a_n, b_n\}) \leq \alpha_n + \beta_n < \varepsilon,$$

which yields the claim (6.15).

By passing to a subnet, still denoted by \mathcal{S} , we thus obtain that $(\tilde{M}^S, m^S)_{S \in \mathcal{S}}$ admits a limit (\tilde{M}, m) w.r.t. convergence of finite dimensional distributions, where

$$\tilde{M}_t = (\Delta M_t^n)_{n \in \mathbb{N}}, \quad t \in [0, 1],$$

and $(\Delta M_t^{S,n})_{S \in \mathcal{S}}$ has ΔM_t^n as its f.d.d. limit. Due to Proposition 6.5 we may choose a version of \tilde{M} taking values in $\mathcal{C}[0, 1]^{\mathbb{N}}$. Consider the following process

$$\hat{M}_t := \sum_{n=1}^m \Delta M_t^n, \quad t \in [0, 1],$$

which has continuous trajectories as m is integer-valued finite. Note that f.d.d. convergence of \tilde{M}^S to \tilde{M} yields f.d.d. convergence of

$$M^S = \sum_{n=1}^{m^S} \Delta M^{S,n} \rightarrow \sum_{n=1}^m \Delta M^n = \hat{M}.$$

We conclude that \hat{M} and M coincide in law, and thus the Lipschitz Markov martingale M has a version with continuous paths. \square

7 Overview of related results in the literature

In the last sections we have focused on specific aspects of the theory of mimicking processes. In this final section we provide an overview of related results in the literature and give some context for theorems discussed above.

An early influential result is the work of Strassen [55] who established that there exists a submartingale with marginals μ_1, \dots, μ_N if and only if these measures are increasing in the increasing convex order induced by increasing convex functions. He also proved that there exists a martingale with values in \mathbb{R}^d and marginals μ_1, \dots, μ_N if and only if these measures are increasing in convex order. In [38, 39], Kellerer managed to extend Strassen's result on the existence of submartingales to an arbitrary family of marginals. As an important particular case, this yields the existence of a mimicking martingale if the marginals increase in convex order. Over time a number of authors have given new approaches to Kellerer's theorem see [29, 30, 47, 46, 48, 7, 11]. As discussed extensively above the work Lowther adds substantial new developments. He characterizes when the Markov martingale can be chosen to be continuous, as well as adding a clear-cut uniqueness part to Kellerer's original result, complementing the formal uniqueness result of Dupire [18]. We also recall from above, that the question whether the natural extension of Kellerer's result to higher dimension holds true remains completely open.

Given a continuum of marginals which increase in convex order (and maybe satisfies additional technical conditions), different authors have provided specific constructions of (not necessarily Markovian) martingales that match these marginals. A main motivation stems from the calibration problem in mathematical finance. An additional goal has often been to give constructions that optimize particular functionals given the martingale and marginal constraints since this yields robust bounds on option prices. Madan and Yor [49] and K'allblad, Tan, and Touzi [36] establish a time continuous version of the Azema-Yor embedding. Hobson [31] established a continuous time version of the martingale coupling constructed in [32]. Henry-Labordere, Tan, and Touzi [27] as well as Br'uckerhoff, Huesmann, and Juillet [14] provide continuous time versions of the shadow coupling (originally introduced in [8]). Richard, Tan, and Touzi [52] give a continuous time version of the Root solution to the Skorokhod embedding problem. In a slightly different but related direction, Boubel and Juillet [12] consider a continuum of marginals on the real line that do not satisfy an order condition and construct a canonical Markov-process matching these marginals. We also refer to the book [28] of Hirsch, Profeta, Roynette, and Yor that collects a variety of related constructions.

The problem of finding martingales with given marginals has received specific attention in the case where these marginals equal the ones of Brownian motion. Hamza-Klebaner [26] posed the challenge of constructing martingales with Brownian marginals that differ from Brownian motion, so called *fake Brownian motions*. Non-continuous solutions can be found in Madan-Yor [49], Hamza-Klebaner [26], Hobson [34], and Fan-Hamza-Klebaner [20] whereas continuous (but non-Markovian) fake Brownian motions were constructed by Oleszkiewicz [50], Albin [2], Baker-Donati-Yor [6] and Hobson [31]. As already noted, the accompanying article [10] establishes that there exists a Markovian martingale with continuous paths that has Brownian marginals.

A somewhat different direction arises if one starts with marginals that do not merely satisfy a structural condition (specifically, monotonicity in convex order) but rather assumes that a set of marginals is generated from an Ito-diffusion

$$dX_t = \sigma_t dB_t + \mu_t dt \tag{7.1}$$

and one seeks a Markovian diffusion

$$d\hat{X}_t = \hat{\sigma}_t(X_t) dB_t + \mu_t(X_t) dt$$

that mimicks the evolution of X in the sense that $X_t \sim \hat{X}_t$ for each $t \geq 0$. The process \hat{X} is then considered to be the *Markovian projection* of X . This line of research goes back essentially to the work of Krylov [41] and Gy'ongy [25]. Of course, also the work of Dupire [18] can be seen as a formal contribution to this line of research. A rigorous justification of Dupire's formula under rather general assumption is obtained by Klebaner [40]. A very general theorem on mimicking aspects of Ito processes is given by Brunick and Shreve [15]. Recently, Lacker, Shkolnikov, and Zhang [44] show that the results of [25, 15] can be established directly from the superposition principle of Trevisan [56] (or Figalli [21] in the case where (7.1) has bounded coefficients). (Notably, the main focus of the work [44] is a mimicking result that shows that conditional time marginals of an Ito-process can be matched by a solution of a conditional McKean-Vlasov SDE with Markovian coefficients.)

In the mathematical finance community, Markovian (local vol) models are often considered to exhibit dynamics that are not particularly realistic. There has been significant interest to combine the convenience that the local vol model offers in terms of calibration with more realistic dynamics that are exhibited by other classes of financial models. That is, given a Markovian model $d\hat{X}_t = \hat{\sigma}_t(\hat{X}_t) dS_t$ that represents marked data, one would like to “reconstruct” a more realistic model $dX_t = \sigma_t dB_t$ and thus to “invert” the Markovian projection. A concrete way to perform this inversion is the *stochastic local volatility model* see the work of Guyon and Henry-Labordere [22, 23, 24]. However, it is remarkably delicate to establish existence and uniqueness results for the resulting SDEs. Partial solutions were given by Jourdain and Zhou [35] and by Lacker Shkolnikov and Zhang in [43]. The problem is also discussed by Acciaio and Guyon [1] who consider it an important open problem to establish existence of the stochastic local volatility model under fairly general assumptions.

References

- [1] B. Acciaio and J. Guyon. Short communication: inversion of convex ordering: local volatility does not maximize the price of VIX futures. *SIAM J. Financial Math.*, 11(1):1–13, 2020.
- [2] J. Albin. A continuous non-Brownian motion martingale with Brownian motion marginal distributions. *Statist. Probab. Lett.*, 78(6):682–686, 2008.
- [3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [4] L. Bachelier. Théorie de la spéculation. *Ann. Sci. École Norm. Sup. (3)*, 17:21–86, 1900.
- [5] J. Backhoff-Veraguas, M. Beiglböck, M. Huesmann, and S. Källblad. Martingale Benamou-Brenier: a probabilistic perspective. *Ann. Probab.*, 48(5):2258–2289, 2020.
- [6] D. Baker, C. Donati-Martin, and M. Yor. A sequence of albin type continuous martingales with brownian marginals and scaling. In *Séminaire de probabilités XLIII*, pages 441–449. Springer, 2011.
- [7] M. Beiglböck, M. Huesmann, and F. Stebegg. Root to Kellerer. *Séminaire de Probabilités XLVIII*, pages 1–12, 2016.

- [8] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016.
- [9] M. Beiglböck, G. Lowther, G. Pammer, and W. Schachermayer. Faking brownian motion through continuous Markov martingales. *ArXiv e-prints*, 2021.
- [10] M. Beiglböck, G. Pammer, and W. Schachermayer. From bachelier to modern finance. *ArXiv e-prints*, 2021.
- [11] M. Beiglboeck and N. Juillet. Shadow couplings. *Trans. Amer. Math. Soc.*, to appear, 2021.
- [12] C. Boubel and N. Juillet. The markov-quantile process attached to a family of marginals. *ArXiv e-prints*, 2018.
- [13] D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 51(4):621–51, 1978.
- [14] M. Brückerhoff, M. Huesmann, and N. Juillet. Shadow martingales – a stochastic mass transport approach to the peacock problem. *ArXiv e-prints*, 2020.
- [15] G. Brunick and S. Shreve. Mimicking an Itô process by a solution of a stochastic differential equation. *Ann. Appl. Probab.*, 23(4):1584–1628, 2013.
- [16] A. M. G. Cox and J. Wang. Root’s Barrier: Construction, Optimality and Applications to Variance Options. *Ann. Appl. Probab.*, 23(3):859–894, 2013.
- [17] E. Derman and I. Kani. The volatility smile and its implied tree. *Goldman-Sachs, Quantitative strategies research notes*, pages 1 – 21, 1994.
- [18] B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.
- [19] E. Dynkin and A. Jushkevich. Strong Markov processes. *Teor. Veroyatnost. i Primenen.*, 1:149–155, 1956.
- [20] J. Y. Fan, K. Hamza, and F. Klebaner. Mimicking self-similar processes. *Bernoulli*, 21(3):1341–1360, 2015.
- [21] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.
- [22] J. Guyon and P. Henry-Labordère. The smile calibration problem solved. *SSRN*: <https://ssrn.com/abstract=1885032>, 2011.
- [23] J. Guyon and P. Henry-Labordère. Being particular about calibration. *Risk*, 25, 2012.
- [24] J. Guyon and P. Henry-Labordère. Nonlinear option pricing. *CRC Press*, 2013.
- [25] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields*, 71(4):501–516, 1986.
- [26] K. Hamza and F. Klebaner. A family of non-Gaussian martingales with Gaussian marginals. *J. Appl. Math. Stoch. Anal.*, pages Art. ID 92723, 19, 2007.

- [27] P. Henry-Labordère, X. Tan, and N. Touzi. An explicit martingale version of the one-dimensional Brenier's theorem with full marginals constraint. *Stochastic Process. Appl.*, 126(9):2800–2834, 2016.
- [28] F. Hirsch, C. Profeta, B. Roynette, and M. Yor. *Peacocks and associated martingales, with explicit constructions*, volume 3 of *Bocconi & Springer Series*. Springer, Milan; Bocconi University Press, 2011.
- [29] F. Hirsch and B. Roynette. A new proof of Kellerer's theorem. *ESAIM Probab. Stat.*, 16:48–60, 2012.
- [30] F. Hirsch, B. Roynette, and M. Yor. Kellerer's theorem revisited. In Springer, editor, *Asymptotic Laws and Methods in Stochastics. Volume in Honour of Miklos Csorgo*, Fields Institute Communications Series, 2014.
- [31] D. Hobson. Mimicking martingales. *The Annals of Applied Probability*, 26(4):2273–2303, 2016.
- [32] D. Hobson and M. Klimmek. Robust price bounds for the forward starting straddle. *Finance Stoch.*, 9(1):189–214, Apr. 2015.
- [33] D. G. Hobson. Volatility misspecification, option pricing and superreplication via coupling. *Ann. Appl. Probab.*, 8(1):193–205, 1998.
- [34] D. G. Hobson. Fake exponential Brownian motion. *Statist. Probab. Lett.*, 83(10):2386–2390, 2013.
- [35] B. Jourdain and A. Zhou. Existence of a calibrated regime switching local volatility model. *Math. Finance*, 30(2):501–546, 2020.
- [36] S. Källblad, X. Tan, and N. Touzi. Optimal Skorokhod embedding given full marginals and Azéma-Yor peacocks. *Ann. Appl. Probab.*, 27(2):686–719, 2017.
- [37] N. Kazamaki. *Continuous exponential martingales and BMO*. Springer, 2006.
- [38] H. Kellerer. Markov-Komposition und eine Anwendung auf Martingale. *Math. Ann.*, 198:99–122, 1972.
- [39] 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)*, pages 341–374. Academia, Prague, 1973.
- [40] F. Klebaner. Option price when the stock is a semimartingale. *Electron. Comm. Probab.*, 7:79–83, 2002.
- [41] N. Krylov. Once more about the connection between elliptic operators and Itô's stochastic equations. In *Statistics and control of stochastic processes (Moscow, 1984)*, Transl. Ser. Math. Engrg., pages 214–229. Optimization Software, New York, 1985.
- [42] H.-H. Kuo. *Introduction to stochastic integration*. Universitext. Springer, New York, 2006.

- [43] D. Lacker, M. Shkolnikov, and J. Zhang. Inverting the Markovian projection, with an application to local stochastic volatility models. *Ann. Probab.*, 48(5):2189–2211, 2020.
- [44] D. Lacker, M. Shkolnikov, and J. Zhang. Superposition and mimicking theorems for conditional McKean-Vlasov equations. *ArXiv e-prints*, 2020.
- [45] T. M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.
- [46] G. Lowther. Fitting martingales to given marginals. *ArXiv e-prints*, Aug. 2008.
- [47] G. Lowther. A generalized backward equation for one dimensional processes. *ArXiv e-prints*, Mar. 2008.
- [48] G. Lowther. Properties of expectations of functions of martingale diffusions. *ArXiv e-prints*, Jan. 2008.
- [49] D. Madan and M. Yor. Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536, 2002.
- [50] K. Oleszkiewicz. On fake Brownian motions. *Statist. Probab. Lett.*, 78(11):1251–1254, 2008.
- [51] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [52] A. Richard, X. Tan, and N. Touzi. On the Root solution to the Skorokhod embedding problem given full marginals. *SIAM J. Control Optim.*, 58(4):1874–1892, 2020.
- [53] D. H. Root. The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.*, 40:715–718, 1969.
- [54] W. Schachermayer. Introduction to the mathematics of financial markets. In *Lectures on probability theory and statistics (Saint-Flour, 2000)*, volume 1816 of *Lecture Notes in Math.*, pages 107–179. Springer, Berlin, 2003.
- [55] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.
- [56] D. Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.*, 21:Paper No. 22, 41, 2016.
- [57] C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.