On strong causal binomial approximation for stochastic processes

Nikolai Dokuchaev

Department of Mathematics & Statistics, Curtin University,

GPO Box U1987, Perth, 6845 Western Australia

Submitted: November 4, 2013. Revised: April 20, 2014

Abstract

This paper considers binomial approximation of continuous time stochastic processes. It is shown that, under some mild integrability conditions, a process can be approximated in mean square sense and in other strong metrics by binomial processes, i.e., by processes with fixed size binary increments at sampling points. Moreover, this approximation can be causal, i.e., at every time it requires only past historical values of the underlying process. In addition, possibility of approximation of solutions of stochastic differential equations by solutions of ordinary equations with binary noise is established. Some consequences for the financial modelling and options pricing models are discussed.

Key words: stochastic processes, Donsker Theorem, binomial approximation, discretisation of Ito equations, incomplete market, complete market.

MSC 2010 classification: Primary: 60F17, 94A14 39A50: Secondary : 91G20, 91B70,

---

This is a pre-copy-editing, author-produced PDF of an article "On strong causal binomial approximation for stochastic processes" published in *Discrete and Continuous Dynamical Systems – Series B (DCDS-B)* **20** (2014), No.6, pp.1549-1562 following peer review (DOI: 10.3934/dcdsb.2014.19.1549). The definitive publisher-authenticated version is available online at: http://aimsciences.org/journals/contentsListnew.jsp?pubID=693.

1 Introduction

This paper considers approximation of continuous time stochastic processes by binomial piecewise affine processes. Usually, these problems are studied in the framework of the Functional Limit Theorems and weak convergence, i.e., convergence in distributions. The classical result is the Donsker Theorem that establishes weak convergence for some particular processes with given distributions (Donsker [10], 1952). Currently, there are many results on the weak convergence for many types of underlying continuous time processes and approximating discrete time processes; see, e.g., [4, 5, 6, 13, 17, 20, 21], and the bibliography there. As far as we know, the convergence of discrete time processes to continuous time processes in the strong sense was not considered in the literature, including approximation in the mean square sense, in L_q-norm, or in probability.

The possibility to approximate a continuous time process by discrete time processes appears to be important for applications in financial modelling. This possibility allows to replace the options pricing for continuous time market models by the options pricing for discrete time market models. The models based on binomial processes, i.e., with fixed size increments, are especially important for this purpose. The reason for this is that the corresponding discrete time market models are usually complete and allow uniquely defined prices for the derivatives. These prices can be conveniently calculated by the so-called binomial trees method which represents a special case of the finite difference methods for PDEs. There are many works on this topic; see, e.g., [2, 3, 11, 16], and the bibliography there. Again, the approximation for the financial models was considered in the weak sense (i.e., in distributions).

With respect to the options pricing problem, the particular distributions of the approximating binomial processes is not really important, since the pricing formula is based on an artificial martingale (or risk neutral) measure rather than on the measure generated by the observed prices. The only part of the Donsker Theorem used in this framework was the existence of the binomial approximating processes. In the present paper, we address only this aspect of the Donsker Theorem: the existence of binomial approximations, without specifications of their distributions.

We consider binomial approximation of continuous time stochastic processes in L_q-norm, where $q \in [1, +\infty)$; this is a strong convergence that includes convergence in mean square and implies convergence in probability. In Section 2, we show that a general L_q-integrable stochastic process can be approximated in L_q-norm by pathwise continuous processes with fixed size binary increments for the sampling points; in particular, continuous Itô processes and processes with jumps are covered (Theorem 2.1). Moreover, we show that this approximation can be *causal*, i.e., the value of an approximating process at each time is calculated using only the past historical values of the underlying process; in other words, the approximating process is adapted to the filtration generated by the underlying process. This can be interpreted as a strong version of Donsker Theorem.

It can be noted that we consider approximation in a different setting than in the cited papers on weak convergence of discrete time processes, where some particular distributions were assumed for the underlying continuous time process and for the approximating processes. We do not assume a particular distribution or certain dynamic properties such as independence or correlation of the increments. We rather suggest an algorithm that allows to construct the approximating processes of the prescribed binomial type from the current observations of the underlying process. Therefore, our approximation result is not in the framework of the Functional Limit Theorem; it does not establish convergence of particular distributions.

In a more general setting, we consider approximation of a stochastic process by solutions of ordinary differential equations with a given drift coefficient and with binary noise (Theorems 2.2 and 3.2). In particular, we found that the solution of a stochastic Itô equation can be approximated by solutions of ordinary differential equations with the same drift coefficient and with binary noise replacing the driving Wiener process (Corollary 2.1). Currently, there are many methods of discretization of stochastic differential equations such as Euler-Maruyama discretization; see, e.g., [1, 12, 15]. Theorems 2.2 and 3.2 could be a useful addition to the existing methods of discretization of stochastic differential equations.

It appears that these approximation results have some implications for financial modelling and for the general pricing theory. To illustrate this, we show that existing of binomial approximations with fixed rate of changes implies that, for a incomplete continuous time market model, there exists a complete model such that these two models are statistically indistinguishable, given the presence of any non-zero errors in the measurements. This feature is non-trivial. It is well known that the market completeness is not a robust property: there are arbitrarily small random contaminations of the coefficients that can convert a complete market model into a incomplete one. We are presenting an "inverse" property: the market incompleteness is also non-robust, meaning that there exist arbitrarily small contaminations that can convert an incomplete model into a complete one.

The paper is organized as follows. In Section 2, we consider approximation by continuous binomial processes and by the solutions of related ODEs with binary noise inputs. In Section 3, we consider some useful modifications of the main result, including approximation by piecewise constant approximating processes and by the solutions of related ODEs with binary jumps. In addition, we discuss in Section 3 approximation of positively valued processes by the processes similar to the prices in Cox-Ross-Rubinstein model, and a possibility of dynamic replication of the changes of the diffusion coefficient. In Section 4, we discuss some implications for financial modelling and pricing theory.

2 The problem setting and the main result

Consider a standard probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where Ω is a set of elementary events, $\mathbf{P}$ is a probability measure, and $\mathcal{F}$ is a $\mathbf{P}$-complete σ-algebra of events.

Let $T > 0$ be given, $q \in [1, +\infty)$. Let X be the set of real valued stochastic processes such that $\|x\|_X \triangleq \left(\mathbf{E}\int_0^T |x(t)|^q dt\right)^{1/q} < +\infty$ for $x \in X$.

Let X_c be the set of all processes $x \in X$ such that for any $x \in X_c$ there exists $\theta = \theta(x(\cdot)) \in [0, T)$ such that the mapping $x : [\theta, T] \to L_q(\Omega, \mathcal{F}, \mathbf{P})$ is continuous.

For $x \in X_c$, we denote $\|x\|_{X_c} \triangleq \|x\|_X + (\mathbf{E}|x(T)|^q)^{1/q}$. Clearly, this value is uniquely defined for any $x \in X_c$.

Let $x \in X$ be given, and let $\mathcal{F}_t$ be the filtration generated by $x(t)$.

Let $\mathcal{Y}_n = \mathcal{Y}_n(x(\cdot))$ be the set of pathwise continuous $\mathcal{F}_t$-adapted real valued processes $y(t)$ such that $y(0) = x(0)$ and that there exists $d > 0$ such that either $y(t) = y(t_k) + d(t - t_k)$ for $t \in [t_k, t_{k+1})$ or $y(t) = y(t_k) - d(t - t_k)$ for $t \in [t_k, t_{k+1})$, where $t_k = kT/n$, $k = 0, ..., n$, $n = 1, 2, ...$ The sequence $\{y(t_k)\}$ represents a path of a so-called binomial tree.

Let $\mathcal{Y} = \cup_{n\geq 1}\mathcal{Y}_n$.

Our main result can be formulated as the following.

Theorem 2.1 *(i) For any $x \in X$ and $\varepsilon > 0$, there exists $y \in \mathcal{Y}$ such that*

$$\|x - y\|_X \le \varepsilon.$$

(ii) For any $x \in X_c$ and $\varepsilon > 0$, there exists $y \in \mathcal{Y}$ such that

$$\|x - y\|_{X_c} \le \varepsilon.$$

Proof. Without a loss of generality, we assume that $x(t)$ is defined for $t < 0$ and that $x(t) = x(0)$ for $t < 0$.

Let

$$\bar{x}_m(t) \triangleq \min(\max(x(t), -m), m), \quad m = 1, 2, ...$$

Clearly, $|\bar{x}_m| \le m$ and $\bar{x}_m(t) = x(t)$ if and only if $|x(t)| \le m$.

Let

$$\mathrm{x}_{m,p}(t) \triangleq \frac{1}{\varepsilon_p} \int_{t-\varepsilon_n}^{t} \bar{x}_m(s)ds, \quad \varepsilon_p = 1/p, \quad p = 1, 2, ...$$

Clearly, this process is pathwise absolutely continuous and such that

$$\operatorname{ess\,sup}_t |d\mathrm{x}_{m,p}(t)/dt| \le 2\varepsilon_k^{-1} \sup_{t\in[0,T]} |\bar{x}_m(t)| \le 2mp.$$

Let any $K \ge 0$ be selected. Let $M_{m,p} \triangleq 2mp + K$. (For the proof of Theorem 2.1, it suffices to use $K = 0$; we need $K > 0$ for the proof of the next theorem).

Let us consider $n = 1, 2, ...$ Let $t_k = kT/n$, $k = 0, ..., n$. Let the process $y(t) = y_{n,m,p}(t)$ be defined such that $y(0) = \mathrm{x}_{m,p}(0)$, $y(t) = y(t_k) + M_{m,p}(t - t_k)$ for $t \in [t_k, t_{k+1})$ if $y(t_k) \le \mathrm{x}_{m,p}(t_k)$, and $y(t) = y(t_k) - M_{m,p}(t - t_k)$ for $t \in [t_k, t_{k+1})$, if $y(t_k) \ge \mathrm{x}_{m,p}(t_k)$. Clearly, $y \in \mathcal{Y}_p$.

Let $\delta = \delta(n) = t_{k+1} - t_k = T/n$. Let us show that

$$|y(t) - \mathrm{x}_{m,p}(t)| \le 2M_{m,p}\delta, \quad t \in [0, T]. \tag{2.1}$$

Clearly, (2.1) holds for $t = t_0$. It suffices to show that if $|y(t_k) - \mathrm{x}_{m,p}(t_k)| \le 2M_{m,p}\delta$ then

$$|y(t) - \mathrm{x}_{m,p}(t)| \le 2M_{m,p}\delta, \quad t \in [t_k, t_{k+1}], \quad k = 0, ..., n. \tag{2.2}$$

Let $M_1 \triangleq 2M_{m,p}\delta$. For $t \in [t_k, t_{k+1}]$, let $M = M_{m,p}(t - t_k)$. We have to consider several possible scenarios.

(i) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in [0, M_1]$ and $\mathrm{x}_{m,p}(t) \geq \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) + M - \varepsilon$, where $\varepsilon \in [0, M]$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) + M - \varepsilon - y(t_k) - M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) - \varepsilon \in [-\varepsilon, M_1 - \varepsilon]. \end{aligned}$$

(ii) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in (M_1, 2M_1]$ and $\mathrm{x}_{m,p}(t) \geq \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) + M - \varepsilon$ again, where $\varepsilon \in [0, M]$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) + M - \varepsilon - y(t_k) - M \\ &= = \mathrm{x}_{m,p}(t_k) - y(t_k) - \varepsilon \in (M_1 - \varepsilon, 2M_1 - \varepsilon]. \end{aligned}$$

(iii) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in [0, M_1]$ and $\mathrm{x}_{m,p}(t) < \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) - M + \varepsilon$, where $\varepsilon \in (0, M]$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) - M + \varepsilon - y(t_k) - M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) - 2M + \varepsilon \in [-2M + \varepsilon, M_1 - 2M + \varepsilon]. \end{aligned}$$

(iv) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in (M_1, 2M_1]$ and $\mathrm{x}_{m,p}(t) < \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) - M + \varepsilon$ again, where $\varepsilon \in (0, M]$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) - M + \varepsilon - y(t_k) - M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) - 2M + \varepsilon \in (M_1 - 2M + \varepsilon, 2M_1 - 2M + \varepsilon]. \end{aligned}$$

(v) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in [-M_1, 0)$ and $\mathrm{x}_{m,p}(t) \leq \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) - M + \varepsilon$, where $\varepsilon \in [0, M]$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) - M + \varepsilon - y(t_k) + M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) + \varepsilon \in [-M_1 + \varepsilon, \varepsilon). \end{aligned}$$

(vi) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in [-2M_1, -M_1)$ and $\mathrm{x}_{m,p}(t) \leq \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) - M + \varepsilon$, where $\varepsilon \in [0, M]$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) - M + \varepsilon - y(t_k) + M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) + \varepsilon \in [-2M_1 + \varepsilon, -M_1 + \varepsilon). \end{aligned}$$

(vii) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in [-M_1, 0)$ and $\mathrm{x}_{m,p}(t) > \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) + M - \varepsilon$, where $\varepsilon \in [0, M)$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) + M - \varepsilon - y(t_k) + M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) + 2M - \varepsilon \in [2M - M_1 - \varepsilon, 2M - \varepsilon). \end{aligned}$$

(viii) Assume that $\mathrm{x}_{m,p}(t_k) - y(t_k) \in [-2M_1, -M_1)$ and $\mathrm{x}_{m,p}(t) > \mathrm{x}_{m,p}(t_k)$. In this case, $\mathrm{x}_{m,p}(t) = \mathrm{x}_{m,p}(t_k) + M - \varepsilon$, where $\varepsilon \in [0, M)$. Hence

$$\begin{aligned} \mathrm{x}_{m,p}(t) - y(t) &= \mathrm{x}_{m,p}(t_k) + M - \varepsilon - y(t_k) + M \\ &= \mathrm{x}_{m,p}(t_k) - y(t_k) + 2M - \varepsilon \in [2M - 2M_1 - \varepsilon, 2M - M_1 + \varepsilon). \end{aligned}$$

By the assumptions, $M_1 \le M$. It follows from (i)-(viii) that $\mathrm{x}_{m,p}(t) - y(t) \in (-2M_1, 2M_1)$ for all possible scenarios. Hence (2.2) holds and therefore (2.1) holds.

We are now in the position to complete the proof.

Let $\|\cdot\|_{\mathcal{X}} \triangleq \|\cdot\|_X$ for the proof of statement (i), and let $\|\cdot\|_{\mathcal{X}} \triangleq \|\cdot\|_{X_c}$ for the proof of statement (ii).

By the Lebesgue's Dominated Convergence Theorem, we have that

$$\|x - \bar{x}_m\|_{\mathcal{X}} \to 0 \quad \text{as} \quad m \to +\infty.$$

By the Lebesgue's Dominated Convergence Theorem again, for any m,

$$\|\bar{x}_m - \mathrm{x}_{m,p}\|_{\mathcal{X}} \to 0 \quad \text{as} \quad p \to +\infty.$$

In addition, it follows from (2.1) that, for any m and p,

$$\|\mathrm{x}_{m,p} - y_{n,m,p}\|_{\mathcal{X}} \to 0 \quad \text{as} \quad n \to +\infty.$$

Let $\varepsilon > 0$ be given. It suffices to show that there exists n, m, p such that

$$\|x - y\|_{\mathcal{X}} \le \varepsilon, \tag{2.3}$$

for $y = y_{n,m,p}$ constructed as described above.

Let m be such that

$$\|x - \bar{x}_m\|_{\mathcal{X}} \le \frac{\varepsilon}{3}. \tag{2.4}$$

Further, let $p = p(m)$ be such that

$$\|\bar{x}_m - \mathrm{x}_{m,p}\|_{\mathcal{X}} \le \frac{\varepsilon}{3}. \tag{2.5}$$

Finally, let $n = n(m,p)$ be such that $2T^{1/q} M_{m,p}\delta \le \varepsilon/3$, where $\delta = \delta(n) = T/n$. In this case, it follows from (2.1) that

$$\|\mathrm{x}_{m,p} - y_{n,m,p}\|_{\mathcal{X}} \le \frac{\varepsilon}{3}. \tag{2.6}$$

Estimates (2.4)-(2.6) imply (2.3). This completes the proof. □

Remark 2.1 *Theorem 2.1 does not require any information on the evolution and the distribution of $x(\cdot)$. Respectively, this theorem does not suggest how to select the set (n, m, p) for a given ε. If (n, m, p) is selected, then the values of the process $y(t)$ at any time t are computed using the historical observations of $x(s)|_{s \le t}$, according to the algorithm described in the proof of Theorem 2.1; in other words, the approximating process is $\mathcal{F}_t$-adapted, and its choice is causal.*

Remark 2.2 *If $x(t)$ is a bounded process, then one can use $x(t)$ directly instead of $\bar{x}_m(t)$. If $\bar{x}_m(t)$ is absolutely continuous and $\bar{c} = \operatorname{ess\,sup}_t |d\bar{x}_m(t)/dt| < +\infty$, then one can construct $y(t)$ using $\bar{x}_m(t)$ instead of $\mathrm{x}_{m,p}(t)$ and $\bar{c}$ instead of $M_{m,p}$.*

We remind that the paths of the processes $y \in \mathcal{Y}_n$ are piecewise affine and continuous; therefore, these paths are absolutely continuous, left-differentiable and right-differentiable; they are differentiable at all $t \neq t_k$, $k = 0, ..., n$, $t_k = kT/n$.

We denote by $D_t^{\pm}$ the left-hand side derivative or right-hand side derivative respectively.

Example 2.1 Theorem 2.1 is oriented on non-differentiable stochastic processes. However, it will be useful to expose some properties of the processes $y(t)$ using the following toy examples.

(i) Assume that $x(t) \equiv 0$. In this case, the approximating processes $y(t)$ defined in Theorem 2.1 can be considered with fixed $m = 1$; they are periodic functions oscillating about $x(t)$ and such that $|D_t^{\pm} y(t)| \equiv c$, where $c > 0$. This c can be selected arbitrarily.

(ii) Assume that $x(t) = 0$ for $t < T/2$ and $x(t) = 1$ for $t \ge T/2$. In this case, it suffices to use fixed $m = 1$ again. The approximating processes $y(t)$ defined in Theorem 2.1

oscillate about zero for $t < T/2$, and $|D_t^{\pm} y(t)| \to +\infty$ as $\|y - x\|_X \to 0$, i.e., as $p \to +\infty$ and $n \to +\infty$. This is because $\operatorname{ess\,sup}_{t\in[0,T]} |d\mathrm{x}_{m,p}(t)/dt| \to +\infty$ as $p \to +\infty$. Since $|D_t^{\pm} y(t)|$ is constant in $t \in [0, T]$, this shows that the approximation suggested does not track the rate of change for the underlying process.

Approximation via solutions of ODEs with binary noise

Let $f(x, t) : \mathbf{R} \times [0, T] \to \mathbf{R}$ be a continuous function such that $|f(x, t)| + |\partial f(x, t)/\partial x| \le c_f$ for some $c_f > 0$.

Again, we assume that $x \in X$ is given, and that $\mathcal{F}_t$ is the filtration generated by x.

Let $\mathcal{U}_n$ be the set of real valued processes $u(t)$ such that $u(0) = x(0)$ and that

$$u(t) = u(0) + \int_0^t f(u(s), s)ds + y(t), \tag{2.7}$$

where $y \in \mathcal{Y}_n$.

Clearly, u is uniquely defined pathwise continuous and $\mathcal{F}_t$-adapted process that satisfy the ordinary differential equation (ODE)

$$\frac{du}{dt}(t) = f(u(t), t) + \eta(t)$$

with binary noise $\eta(t) = D_t^{\pm} y(t)$. The choice of the right hand derivative D_t^+ or left-hand derivative D_t^- here does not affect the solution of the ODE, since these derivatives coincides everywhere except the points $t_k = \delta k$, $\delta = T/n$.

The process η is adapted to the filtration generated by x, and its properties are defined by the properties of x.

Let $\mathcal{U} = \cup_{n\ge 1} \mathcal{U}_n$.

Theorem 2.2 *(i) For any $x \in X$ and $\varepsilon > 0$, there exists $u \in \mathcal{U}$ such that*

$$\|x - u\|_X \le \varepsilon.$$

(ii) For any $x \in X_c$ and $\varepsilon > 0$, there exists $u \in \mathcal{U}$ such that

$$\|x - u\|_{X_c} \le \varepsilon.$$

Clearly, Theorem 2.2 applied with $f \equiv 0$ gives Theorem 2.1; therefore, Theorem 2.2 represents a generalization of Theorem 2.1.

Proof of Theorem 2.2. We use the processes $\bar{x}_m$ and $\mathrm{x}_{m,p}$ from the proof of Theorem 2.1. We modify the construction of $y(t) = y_{n,m,p}(t)$ as the following. We select $K = \sup_{x,t} |f(x,t)|$, i.e., $M_{m,p} = 2mp + \sup_{x,t} |f(x,t)|$. For $n \in \{1, 2, ...\}$, we set $t_k = kT/n$, $k = 0, ..., n$, and define step functions $\theta(s)$ such that $\theta(s) = t_k$ if $s \in [t_k, t_{k+1})$.

Let us construct the processes $r(t) = r_{n,m,p}(t)$, $y(t) = y_{n,m,p}(t)$, and $u(t) = u_{n,m,p}(t)$ as the following. We assume that $r(0) = y(0) = 0$, $u(0) = x(0)$, and

$$r(t) \triangleq \mathrm{x}_{m,p}(t) - x(0) - \int_0^t f(u(\theta(s)), s)ds,$$
$$u(t) \triangleq u(0) + \int_0^t f(u(s), s)ds + y(t).$$

Here the process $y(t)$ is defined as the following: $y(t) = y(t_k) + M_{m,p}(t - t_k)$ for $t \in [t_k, t_{k+1})$ if $y(t_k) \le r(t_k)$, and $y(t) = y(t_k) - M_{m,p}(t - t_k)$ for $t \in [t_k, t_{k+1})$, if $y(t_k) > r(t_k)$.

Clearly, $y \in \mathcal{Y}_n$, and the processes r, u, and y, can be constructed consequently on the intervals $[t_k, t_{k+1}]$, $k = 0, 1, 2, ...$

Let $\delta = t_{k+1} - t_k = T/n$. Similarly to (2.1), we obtain that

$$|y(t) - r(t)| \le 2M_{m,p}\delta, \quad t \in [0, T].$$

Let

$$\widetilde{r}(t) = \widetilde{r}_{n,m,p}(t) \triangleq x(t) - x(0) - \int_0^t f(u(s), s)ds.$$

Clearly,

$$\begin{aligned}
|\widetilde{r}(t) - r(t)| &\le \int_0^T |f(u(\theta(s)), s) - f(u(s), s)|ds \\
&\le T \sup_s |f(u(\theta(s)), s) - f(u(s), s)| \\
&\le c_f T \sup_s |u(\theta(s)) - u(s)| \\
&\le c_f T \sup_s \left| \int_{\theta(s)}^s f(u(r), r)dr + y(s) - y(\theta(s)) \right| \\
&\le c_f(c_f\delta + M_{m,p}\delta) = c_f(c_f + M_{m,p})\frac{T}{n}.
\end{aligned}$$

It follows that, for all t,

$$|\mathrm{x}_{m,p}(t) - u(t)| = |\widetilde{r}(t) - y(t)| \le |\widetilde{r}(t) - r(t)| + |r(t) - y(t)| \to 0 \quad \text{as} \quad n \to +\infty.$$

The remaining part of the proof follows the proof of Theorem 2.1. □

Corollary 2.1 *Let $f(x,t) : \mathbf{R} \times [0,T] \to \mathbf{R}$ and $b(x,t) : \mathbf{R} \times [0,T] \to \mathbf{R}$ be continuous bounded functions such that the derivatives $\partial f(x,t)/\partial x$ and $\partial b(x,t)/\partial x$ are also bounded. Let $w(t)$ be a standard Wiener process, and let the evolution of x be described by the Itô equation*

$$dx(t) = f(x(t),t)dt + b(x(t),t)dw(t). \tag{2.8}$$

By Theorem 2.2, the process x can be approximated in L_q-norm by the solutions of ordinary differential equations (2.7), where $y \in \mathcal{Y}$.

The approximation of solutions of Itô stochastic differential equations by the solutions of ordinary differential equations with binary noise could be a useful addition to the existing methods of discretization such as Euler-Maruyama discretization; see, e.g., [1, 12, 15].

Remark 2.3 *Theorem 2.2 and Corollary 2.1 imply that the solutions of Itô equations (2.8) are statistically indistinguishable from the solutions of ordinary equations (2.7) with small enough $\delta = t_{k+1} - t_k$, given the presence of an arbitrarily small errors in the measurements of the processes. A related feature is discussed in detail in Section 4 below.*

3 Some modifications

3.1 Approximation using piecewise constant processes

We assume again that $x \in X$ is given, and that $\mathcal{F}_t$ is the filtration generated by x.

Let $\mathcal{Z}_n$ be the set of pathwise right-continuous piecewise-constant $\mathcal{F}_t$-adapted real valued processes $y(t)$ such that $y(0) = x(0)$ and that there exists $d > 0$ such that either $y(t) = y(t_k) + d$ for $t \in [t_k, t_{k+1})$ or $y(t) = y(t_k) - d$ for $t \in [t_k, t_{k+1})$, where $t_k = kT/n$, $k = 0, ..., n$, $n = 1, 2, ...$ The sequence $\{y(t_k)\}$ represents a path of a binomial tree again.

Let $\mathcal{Z} = \cup_{n \ge 1} \mathcal{Z}_n$.

Theorem 3.1 *The statement of Theorem 2.1 holds with $\mathcal{Y}$ replaced by $\mathcal{Z}$.*

Proof of Theorem 3.1 repeats the proof of Theorem 2.1 with the following changes. For $n = 1, 2, ...,$ let $t_k = kT/n$, $k = 0, ..., n$, and $\delta = t_{k+1} - t_k$, we define processes $y(t) = y_{n,m,p}(t)$ such that $y(t) = x(0) = \mathrm{x}_{m,p}(0)$ for $t \in [t_0, t_1]$, $y(t) = y(t_k) + M_{m,p}\delta$ for $t \in [t_k, t_{k+1})$ if $y(t_k) \le \mathrm{x}_{m,p}(t_k)$, and $y(t) = y(t_k) - M_{m,p}\delta$ for $t \in [t_k, t_{k+1})$, if $y(t_k) \ge \mathrm{x}_{m,p}(t_k)$, $k = 1, 2, ...$ Here $M_{m,p} = 2mp$. Similarly to the proof of Theorem 2.1, we obtain that

$$|y(t_k) - \mathrm{x}_n(t_k)| \le 2M_{m,p}\delta, \quad k = 0, 1, ..., n.$$

It follows that

$$|y(t) - \mathrm{x}_n(t)| \le 4M_{m,p}\delta, \quad t \in [0, T].$$

The remaining proof repeats the proof of Theorem 2.1. □

Further, let $f(x, t) : \mathbf{R} \times [0, T] \to \mathbf{R}$ be a continuous function that is bounded together with the derivative $\partial f(x, t)/\partial x$. Let $\mathcal{V}_n$ be the set of real valued processes $v(t)$ such that $u(0) = x(0)$ and that

$$v(t) = v(0) + \int_0^t f(v(s), s)ds + z(t), \tag{3.1}$$

where $z \in \mathcal{Z}_n$.

Clearly, v is uniquely defined; it is a $\mathcal{F}_t$-adapted process with jumps at the times t_k.

Let $\mathcal{V} = \cup_{n \ge 1} \mathcal{V}_n$.

The following theorem represents a modification of Theorem 2.2.

Theorem 3.2 *The statements of Theorem 2.2 and Corollary 2.1 hold with $\mathcal{U}$ replaced by $\mathcal{V}$.*

Again, Theorem 3.2 applied with $f \equiv 0$ gives Theorem 3.1; therefore, Theorem 3.2 represents a generalization of Theorem 3.1.

Proof of Theorem 3.2 repeats the proof of Theorem 2.2 with the changes similar to the changes that were done in the proof of Theorem 3.1. □

3.2 Binomial approximation of $\log x(t)$

In financial modelling, it is common to approximate positively valued stochastic processes by binomial processes with the rate of change decreasing near zero such that their logarithm have the constant rate of change. We need to modify our approach to cover these problems.

Let $x \in X$ be given such that $x(t) > 0$ for all t and that the process $\log x(t)$ belongs to X. Let $\mathcal{F}_t$ be the filtration generated by $x(t)$.

Let $\mathcal{Y}_n^+ = \mathcal{Y}_n^+(x(\cdot))$ be the set of pathwise continuous piecewise affine $\mathcal{F}_t$-adapted real valued processes $y(t)$ such that $y(0) = x(0)$ and that there exists $d_1 \in (0, \delta)$ and $d_2 > 0$ such that either $y(t_{k+1}) = y(t_k)(1 - d_1\delta)$ or $y(t_{k+1}) = y(t_k)(1 + d_2\delta)$. Here $t_k = kT/n$, $k = 0, ..., n$, $n = 1, 2, ...$, and $\delta = t_{k+1} - t_k = T/n$.

Let $\mathcal{Y}^+ = \cup_{n\geq 1}\mathcal{Y}_n^+$, and let $\mathcal{Y}_{log}^+$ be the set of all processes $\eta(t)$ such that $\eta(t) = \log y(t)$, where $y \in \mathcal{Y}^+$.

In financial modelling, binomial processes from $\mathcal{Y}^+$ are used for positively valued stochastic Itô processes with lognormal distributions describing the evolution of the stock prices in the Black-Scholes market model. We will be using these processes in the next section addressing the financial applications.

Theorem 3.3 *The statement of Theorems 3.1 holds with $x(t)$ replaced by $\log x(t)$ and with $\mathcal{Y}$ replaced by $\mathcal{Y}_{log}^+$.*

Proof. The proof requires a small modification of the proof of Theorems 3.1. We select d_1 and d_2 such that $\log(1 - d_1\delta) = -M_{m,p}\delta$ and $\log(1 + d_2\delta) = M_{m,p}\delta$, where $\delta = t_{k+1} - t_k$ and $M_{m,p} = 2mp$ are selected similarly to the proof of Theorem 2.1. We define $\eta \in \mathcal{Y}_{log}^+$ by selecting

$$\eta(t_k) = \log x(t_0) + \sum_{i=1}^{k} \xi_i\delta,$$

where ξ_i take values $\pm M_{m,p}$ selected similarly to the proof of Theorems 2.1. Consider representation $\xi_i\delta = \log(1 + \zeta_i\delta)$, where ζ_i take values $-d_1$ or d_2. We have that

$$\eta(t_k) = \log x(t_0) + \sum_{i=1}^{k} \log(1 + \zeta_i\delta).$$

Then $\eta(t) = \log y(t)$, where $y \in \mathcal{Y}^+$ is such that $y(t_0) = x(t_0)$ and $y(t_k) = x(t_0)\prod_{i=0}^k(1+\zeta_i\delta)$ for $k > 0$. This completes the proof. □

A similar result can be obtained for piecewise constant approximations.

3.3 Approximation with dynamically adjusted sizes of the binary increments

It could be interesting to consider approximating sequences of processes with dynamically adjusted sizes of the increments that can replicate the changes in the evolution law for the underlying process, such as the dynamics of the volatility for the stock prices. So far, Example 2.1(ii) shows that this feature is not feasible for the approximating processes from $\mathcal{Y}$.

It appears that, for the case of underlying processes from some more narrow classes, the algorithm described in Theorem 2.1 can be extended on the approximating processes with dynamically adjusted sizes of the increments.

Let $H_{\theta,q}$ be the set of all processes $x \in X$ such that there exists $\varepsilon_0 > 0$, $q \in (0, 1]$, $\theta > 0$, $C > 0$, and a $\mathcal{F}_t$-adapted stochastic process $\sigma(t)$ such that $0 \le \sigma(t, \omega) \le C$ for all $t \in [0, T]$ and $\omega \in \Omega$, and that

$$\sup_{(t,\omega)\in[0,T]\times\Omega} \frac{|x(t,\omega) - x(t-\varepsilon,\omega)|}{\varepsilon^q} \le \sigma(t-\theta,\omega) \tag{3.2}$$

for all $\varepsilon \in (0, \varepsilon_0)$.

For $x \in H_{\theta,q}$, the approximating processes with dynamically adjusted sizes of binary increments can be constructed as follows. For $n = 1, 2, ...$, we select $y(t) = y(t_k) \pm \delta^{q-1}\sigma(t_k, \omega)(t - t_k)$, $t \in [t_k, t_{k+1})$, where $\delta = t_k - t_{k+1} = T/n$. We can skip construction of the processes $\bar{x}_m$. We construct the process $\mathrm{x}_{m,p}(t)$ using x instead of $\bar{x}_m$, and observe that $\operatorname{ess\,sup}_t |d\mathrm{x}_{m,p}(t)/dt| \le \varepsilon^{q-1}\sigma(t-\theta)$ for $\varepsilon = 1/p$. The approximating properties can be established as before, with sufficiently small $\delta \le \min(\varepsilon_0, \theta)$.

The Hölder type condition (3.2) for $H_{\theta,q}$ is close to the Hölder regularity property for the trajectories of the continuous Itô processes; see, e.g., [14]. Unfortunately, the presence of the supremum over ω still makes condition (3.2) too restrictive; this condition is not satisfied even for Itô processes with constant diffusion coefficients, including a Wiener process.

Condition (3.2) could be reasonable for models using a causal estimator for σ based on historical observations of x, under a hypothesis that the currently calculated $\sigma(t, \omega)$ satisfies

(3.2) on the time interval $[t, t+\theta]$. Then condition (3.2) ensures that

$$\sup_{(t,\omega)\in[0,T]\times\Omega} \frac{|x(t+\tau,\omega) - x(t+\tau-\varepsilon,\omega)|}{\varepsilon^q} \le \sigma(t+\tau-\theta,\omega)$$

for all $\varepsilon \in (0, \varepsilon_0)$ and $\tau \in [0, \theta]$. Since $\sigma(t+\tau-\theta)$ is a $\mathcal{F}_t$-measurable random variable, we have that the value of the upper boundary is known at time t.

If we interpret $\sigma(t-\theta)$ as an analog of the diffusion coefficient at time t, then condition (3.2) can be interpreted as the requirement that the diffusion coefficient is predictable on the time horizon θ. In other word, this is a requirement that there is some stability in the evolution law for x. This could be a reasonable requirement for many models.

4 Applications for the financial modelling

In quantitative finance, the classical discrete time Cox-Ross-Rubinstein model of a single-stock financial market includes a bond or money market account with the price B_k and a single risky asset with the price S_k. In this model, the process B_k is assumed to be non-random or risk-free and is used as a numéraire, and S_k is assumed to be a binomial stochastic process, $k = 0, 1, 2, ...$ For simplicity, we assume that $B_k = \rho^k$ for some $\rho \ge 1$. Let $\widetilde{S}_k = B_k^{-1} S_k$ be the discounted price process. For the Cox-Ross-Rubinstein model, $\widetilde{S}_{k+1} = \widetilde{S}_k(1 + \zeta_{k+1})$, where ζ_k takes only two values, $-d_1$ and d_2, such that $d_1 \in (0, 1)$ and $d_2 > 0$; see, e.g., [7], Chapter 3, and [8]. This model is a so-called complete market where any claim can be replicated and where there is a unique martingale (risk-neutral) measure equivalent to the historical measure. For complete market models, the price of a derivative is defined via the expectation of the payoff by this unique martingale measure.

We consider also a continuous time model that includes a bond or money market account with the price $\mathrm{B}(t) = e^{rt}$ and a single risky asset with the price $\mathsf{S}(t)$, $t \in [0, T]$. Here $r \ge 0$ is given and known, and $\mathsf{S}(t)$ is assumed to be a stochastic Ito process such that

$$d\mathsf{S}(t) = \mathsf{S}(t)[a(t)dt + \sigma(t)dw(t)],$$

where $a(t)$ is some appreciation rate process, $\sigma(t) = \sigma(t, \omega)$ is a random volatility process, and $w(t)$ is a Wiener process. We assume that $\sigma(t)$ is independent on the increments $w(\theta) - w(\tau)$, $\theta > \tau \ge t$. We assume for simplicity that $a \in \mathbf{R}$ is a constant.

Let $\widetilde{\mathsf{S}}(t) = \mathrm{B}(t)^{-1}\mathsf{S}(t)$ be the discounted price process.

The classical Black and Scholes continuous time model represents a special case of this model with a constant volatility $\sigma > 0$. In this case, the market model is complete; see, e.g., [7], Chapter 5. The Donsker theorem allows to approximate the process $\widetilde{\mathsf{S}}(t)$ by binomial processes $y \in \mathcal{Y}^+$ in distributions. This allows to replace pricing of derivatives in the continuous time setting by the pricing in the discrete time setting via binomial trees (i.e., finite differences).

The pricing of derivatives is usually more difficult for the so-called incomplete market models where a martingale measure is not unique. Some important examples of market incompleteness arise when the volatility $\sigma = \sigma(t, \omega)$ is time varying, random, and not adapted to the filtration generated by $w(t)$. In this case, a straightforward discrete time approximation leads to discrete time market models such as binomial models with dynamically adjusted sizes (i.e., random sizes) of the binary increments; see, e.g., [2] and discussion in Section 3.3. These binomial models are incomplete.

On the other hand, Theorem 3.3 implies that, for any continuous time market model, including models with random volatility $\sigma = \sigma(t, \omega)$, there exists a process $y \in \mathcal{Y}^+$ such that y and $\widetilde{\mathsf{S}}$ are statistically indistinguishable, given that there is a non-zero measurements error, for instance, a rounding error, or any arbitrarily small error.

This y can be used for construction of a complete discrete time market that approximates the original incomplete market as follows.

For $n = 1, 2, ...$, consider the sets $\{t_k\}_{k=0}^n$ such that $t_k = k\delta$, $\delta = T/n$.

Let $B_k = e^{rt_k} = \mathrm{B}(t_k)$, $\widetilde{S}_k = y(t_k)$, and $S_k = B_k\widetilde{\mathsf{S}}(t_k)$. Consider a discrete time market model with the stock prices S_k, with the discounted prices $\widetilde{S}_k$, and with the bond prices B_k. By the definitions, this is a Cox-Ross-Rubinstein model, i.e., a complete market model.

This leads to a counterintuitive conclusion that the incomplete markets are indistinguishable from the complete markets by econometric methods, i.e., in the terms of the market statistics.

Let us elaborate this conclusion. It is known that the market completeness is not a robust property: small random deviations of the coefficients convert a complete market model into a incomplete one. Thanks to Theorem 3.3 and approximation scheme described above, we can claim now that market incompleteness is also non-robust: small deviations can convert an incomplete model into a complete one. More precisely, it implies that, for any incomplete market from a wide class of models, there exists a complete market model with arbitrarily close discrete sets of the observed processes.

It can be further illustrated as the following. Assume that we collect the marked data (the prices) for $t \in [0, T]$, with the purpose to test the following hypotheses $\mathbf{H}_0$ and $\mathbf{H}_A$ about the stock price evolution:

$\mathbf{H}_0$: For any sampling interval δ, the discrete time market with the stock prices $S_k = \mathsf{S}(t_k)$ is incomplete; and

$\mathbf{H}_A$: There exists a sampling interval δ such that the discrete time market with the stock prices $S_k = \mathsf{S}(t_k)$ is complete.

Here $\delta = T/n$, $n = 1, 2, ...$, $t_k = k\delta$, $k = 0, 1, ..., n$.

According to Theorem 3.3, it is impossible to reject $\mathbf{H}_A$ hypothesis based solely on the market data collected, for a random volatility process $\sigma(t, \omega)$ generating an incomplete continuous time market and incomplete discrete time markets based on the sampled prices.

It can be noted that we can replace the hypothesis $\mathbf{H}_0$ by a hypothesis assuming a particular stochastic price model, such as a Markov chain model for the volatility, Heston model, etc.

Due to rounding errors, the statistical indistinguishability leading to this conclusion cannot be fixed via the sample increasing since the statistics for the incomplete market models can be arbitrarily close to the statistics of the alternative complete models.

It must be clarified that this conclusion has rather a purely theoretical value. Unfortunately, the causal binomial approximation described above is not particularly useful for practical options pricing since the process $\{S_k\}$ does not represent the price of a tradable assets in the market model with the price $\{\mathsf{S}(t_k)\}$; the values S_k represent the prices of a tradable assets for the new discrete time market model only. In addition, the hedging strategies for the discrete time market with the prices S_k do not give the same output when applied to the prices samples of the original prices $\mathsf{S}(t_k)$, since small errors for single transactions could be accumulated into a significant error even if S_k is close to $\mathsf{S}(t_k)$ a.e.. This is because a close approximation requires small sampling intervals, large number of periods, and a large number of transactions. Each transaction will generate a small but non-zero error since we allow that $S_k \neq \mathsf{S}(t_k)$ even if these values are close. Furtehrmore, using of the new discrete time complete market model with the price process $\{S_k\}$ for the options pricing would lead to overpricing if $\sigma(t,\omega)$ is random and time variable, since the process S_k is constructed such that its rate of change is constant; see Example 2.1. It can be noted that approximation in the class $H_{\theta,q}$ described in Section 3.3 allows to replicate stochastic and time varying volatility; however, this approximation does not lead to approximating complete markets.

The results of this section on statistical indistinguishability of the complete binomial markets and incomplete markets was presented on The Quantitative Methods in Finance conference in Sydney in December 2013. A related result was obtained in [9], where the approximation was considered in a diffusion setting that allowed to approximate the dynamics of the original volatility as well, via approximation by diffusion processes.

5 Discussion and future development

The approach suggested in this paper allows many more modifications. We outline below some possible straightforward modifications as well as more challenging problems and possible applications that we leave for the future research.

(i) Instead of binomial processes, other processes could be used for approximation, for instance, trinomial processes.

(ii) It could be interesting to extend the construction from Section 3.3 on a case where condition (3.2) is replaced by a weaker condition that covers Itô processes.

(iii) Ordinary differential equations (2.7) and equations with jumps (3.1) can be investigated with random noise with preselected distributions rather than with the processes $y(t)$ and $z(t)$ defined by the approximation procedure. For example, equation (2.7) could be considered for a binary white noise $\eta = dy/dt$ that is a piecewise constant stochastic process. Equation (3.1) could be considered for a random input represented by binary white noise $\eta = dz/dt$ such that $\eta(t) = \pm c\,\delta(t - t_k)$ for $t \in (t_{k-1}, t_{k+1})$, where $\delta(t)$ is a delta-function, $c = c_\eta \in \mathbf{R}$.

(iv) For equations (2.7) and (3.1) with random noise, optimal stochastic control problems as well as stability and instability could be investigated, as was done for Euler-Maruyama discretization; see, e.g., [18] and the bibliography there.

(v) Approximation of classical stochastic differential plant equations by equations (2.7) or (3.1) with random noise could be applied for solution of optimal stochastic control problems, in particular, for optimal portfolio selection problems. Some related results for Euler-Maruyama discretization were obtained in [19].

Acknowledgment

This work was supported by Australian Research Council grant DP120100928 to the author.

References

[1] V. Abramov, F. Klebaner, R. Liptser, The Euler-Maruyama approximations for the CEV model, *Discrete and Continuous Dynamical Systems. Series B*, **16** (1) (2011), 1–14.

[2] E. Akyildirim, Y. Dolinsky, H.M. Soner, Approximating stochastic volatility by recombinant trees, preprint, arXiv:1205.3555, 2012.

[3] K.I. Amin, On the computation of continuous time option prices using discrete approximations, *Journal of Financial and Quantitative Analysis*, **26** (04) (1991), 477–495.

[4] P. Billingsley, Convergence of probability measures, Wiley, New York, 1968.

[5] S. Borovkova, R. Burton, and H. Dehling, Limit theorems for functionals of mixing processes with application to U-statistics and dimension estimation, *Trans. Amer. Math. Soc.*, **353** (2001), 4261–4318.

[6] J. Dedecker, C. Prieur, New dependence coefficients. Examples and applications to statistics. *Probab. Theory and Relat. Fields*, **132** (2005), 203–236.

[7] N.G. Dokuchaev, *Mathematical finance: core theory, problems, and statistical algorithms*, Routledge, London and New York, 2007.

[8] N. Dokuchaev, Discrete time market with serial correlations and optimal myopic strategies, *European Journal of Operational Research*, **177**(2)(2007), 1090–1104.

[9] N. Dokuchaev, On statistical indistinguishability of the complete and incomplete markets, preprint, arXiv:1209.4695, 2012.

[10] M.D. Donsker, Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems, *Annals of Mathematical Statistics*, **23** (1952), 277–281.

[11] D. Heath, R. Jarrow, A. Morton, (Bond pricing and the term structure of interest rates: a discrete time approximation. *Journal of Financial and Quantitative Analysis*, **25**(4) (1990), 419–440.

[12] D. J. Higham, X. Mao, A. M. Stuart, Strong convergence of numerical methods for nonlinear stochastic differential equations, *SIAM J. Num. Anal.*, 40(3) (2002), 1041–1063.

[13] I. A. Ibragimov, Some limit theorems for stationary processes, *Theory of probability and its applications* , **7** (1962), 349–382.

[14] I.A. Ibragimov, Properties of sample functions of stochastic processes and embedding theorems. *Theory of probability and its applications*, **18** (3) (1973), 442–453.

[15] P. E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.

[16] D. B. Nelson, K. Ramaswamy, (1990). Simple binomial processes as diffusion approximations in financial models. *Review of Financial Studies*, **3** (3) (1990), 393–430.

[17] R. Nickl, M. Reiß, J. Söhl, M. Trabs, High-frequency Donsker theorems for Lévy measures, preprint, arXiv:1310.252, 2013.

[18] A. Rodkina, N. Dokuchaev, Instability and stability of solutions of systems of nonlinear stochastic difference equations with diagonal noise. *Journal of Difference Equations and Applications, in press*, 2013; available at http://dx.doi.org/10.1080/10236198.2013.815748.

[19] A. Rodkina, N. Dokuchaev, On asymptotic optimality of Merton's myopic portfolio strategies for discrete time market, preprint, arXiv:1403.4329, 2014.

[20] C. Tudor, S. Torres, Donsker theorem for the Rosenblatt process and a binary market model, preprint, arXiv:math/0703085, 2007.

[21] A. van der Vaart, H. van Zanten, Donsker theorems for diffusions: Necessary and sufficient conditions, *Annals of Probability*, **33** (4) (2005), 1422–1451.