THE UNIVERSITY OF WARWICK

Queueing Syst (2013) 73:447–460
DOI 10.1007/s11134-013-9351-0

# The strong weak convergence of the quasi-EA

**Stefano Peluchetti · Gareth O. Roberts · Bruno Casella**

Received: 10 March 2011 / Revised: 14 August 2012 / Published online: 17 March 2013


**Abstract** In this paper, we investigate the convergence of a novel simulation scheme to the target diffusion process. This scheme, the Quasi-EA, is closely related to the Exact Algorithm (EA) for diffusion processes, as it is obtained by neglecting the rejection step in EA. We prove the existence of a myopic coupling between the Quasi-EA and the diffusion. Moreover, an upper bound for the coupling probability is given. Consequently we establish the convergence of the Quasi-EA to the diffusion with respect to the total variation distance.

**Keywords** Simulation of SDEs · Biased Brownian Motion · Exact simulation

**Mathematics Subject Classification** 65C05 · 65C35

## 1 Introduction

In this paper, we shall present two convergence results about a novel simulation scheme to the target diffusion process. This scheme is closely related to the Exact Algorithm (EA) for the simulation of diffusion process which was introduced in [1]. In particular, the simulation scheme we consider is essentially EA without the acceptance–rejection correction.

This scheme (which we call the Quasi-EA) is studied for two reasons. Firstly, we are interested in the properties of Quasi-EA as a simulation scheme in its own right. Secondly, a thorough understanding of Quasi-EA contributes to a fuller understanding of EA scheme itself.

S. Peluchetti · G. O. Roberts (✉) · B. Casella
Department of Statistics, University of Warwick, Coventry CV4 7AL, UK
e-mail: gareth.o.roberts@warwick.ac.uk

The main appeal of EA is that it allows for the exact simulation (i.e., free from any time discretisation error) of any skeleton of the diffusion sample path. Moreover, it is possible to simulate exactly from some classes of path-dependent functionals of the diffusion process. Because EA plays a central rule in our work, we briefly introduce the main ideas behind EA. For a more exhaustive exposition of EA, we refer to [2] and [3].

We consider the diffusion process $Y$ a unique strong solution of the Stochastic Differential Equation (SDE)

$$\begin{aligned} \mathrm{d}Y_t &= b\left(Y_t\right)\mathrm{d}t + \sigma\left(Y_t\right)\mathrm{d}B_t \quad 0 \le t \le T \\ Y_0 &= y, \end{aligned} \tag{1}$$

where $B$ is scalar Brownian Motion (BM) on the bounded time interval.

We make the mild assumptions that $\sigma$ is continuously differentiable and strictly positive, that $b$ is continuously differentiable, and that the diffusion is non-explosive. These assumptions are more than enough to guarantee the existence of a unique strong solution to the SDE (see for example Sect. V.10 of [10]). These assumptions also guarantee the existence and uniqueness of a bijective function $\eta$ such that the transformed diffusion process $X_t := \eta\left(Y_t\right)$ satisfies the SDE

$$\begin{aligned} \mathrm{d}X_t &= \alpha\left(X_t\right)\mathrm{d}t + \mathrm{d}B_t \quad 0 \le t \le T \\ X_0 &= x := \eta\left(y\right). \end{aligned} \tag{2}$$

The drift coefficient $\alpha$ takes the form:

$$\alpha(x) = \frac{b(\eta^{-1}(x))}{\sigma(\eta^{-1}(x))} - \sigma'(\eta^{-1}(x)),$$

and inherits continuous differentiability from $b$ and $\sigma$.

From now on the SDE (2) will be our starting point.

Let $\mathbb{Q}_T^x$ and $\mathbb{W}_T^x$ denote the law of the diffusion $X$ and the law of a BM, respectively, on $[0, T]$ both started at $x$. From now on the following hypotheses are assumed to hold:

- $\forall x \in \mathbb{X}\ \mathbb{Q}_T^x \ll \mathbb{W}_T^x$ and the Radon–Nikodym derivative is given by Girsanov's formula:

$$\frac{\mathrm{d}\mathbb{Q}_T^x}{\mathrm{d}\mathbb{W}_T^x}(\omega) = \exp\left\{\int_0^T \alpha\left(\omega_s\right)\mathrm{d}\omega_s - \frac{1}{2}\int_0^T \alpha^2\left(\omega_s\right)\mathrm{d}s\right\} \tag{3}$$

- $\alpha$ is continuously differentiable on $\mathbb{X}$;
- $\alpha^2 + \alpha'$ is bounded below on $\mathbb{X}$.

We introduce Biased Brownian Motion (BBM) $Z$ and its law $\mathbb{Z}_T^x$. This process is defined as a BM on $[0, T]$ started at $x$ and conditioned on having its terminal value $Z_T$ distributed according to the density

$$h_{x,T}(u) := \eta_{x,T} \times \exp\left\{A(u) - \frac{(u-x)^2}{2T}\right\}. \tag{4}$$

Here $A(u) := \int_c^u \alpha(r)\,\mathrm{d}r$ for some $c \in \mathbb{X}$ and the normalising constant $\eta_{x,T}$ is assumed to be finite. Hence conditionally on the value of $Z_T$ the process $Z$ is distributed as a Brownian Bridge (BB). Given the hypothesis, it is possible to prove that

$$\frac{\mathrm{d}\mathbb{Q}_T^x}{\mathrm{d}\mathbb{Z}_T^x}(\omega) = \eta_{x,T} \exp\{-A(x)\} \exp\left\{-\int_0^T \frac{\alpha^2+\alpha'}{2}(\omega_s)\,\mathrm{d}s\right\} \tag{5}$$

$$\propto \exp\left\{-\int_0^T \phi(\omega_s)\,\mathrm{d}s\right\} \leq 1, \tag{6}$$

where $\phi(u) := \left(\alpha^2(u) + \alpha'(u)\right)/2 - \inf_{r\in\mathbb{X}}\left(\alpha^2(r)+\alpha'(r)\right)/2$. Equation (6) suggests the use of a rejection sampling algorithm to generate realisations from $\mathbb{Q}_T^x$. However, it is not possible to generate a sample from $Z$, $Z$ being an infinite dimensional variate, and moreover it is not possible to compute analytically the value of the integral in (6). For ease of exposition, we only consider the case of EA1, where $\alpha^2+\alpha'$ is assumed to be bounded. It should be noted that this hypothesis can be weakened or even removed; see [2,3].

We denote by $m$ the finite supremum of $\phi$. Let $\Phi = (\mathcal{X}, \Psi)$ be a unit rate Poisson Point Process (PPP) on $[0,T]\times[0,m]$, and let $N$ be the number of points of $\Phi$ below $\phi(\omega_s)$ conditionally on a path $\omega$. If the event $\Gamma$ is defined as $\Gamma(\omega,\Phi) := \{N=0\}$ it follows that

$$\mathcal{P}r[\Gamma|\omega] = \exp\left\{-\int_0^T \phi(\omega_s)\,\mathrm{d}s\right\}. \tag{7}$$

This equivalence is an immediate consequence of the properties of PPPs. The lhs of (7) is the probability that no points of $\Phi$ fall in the region of the plane bounded by 0 and $\phi(\omega_s)$, a region whose area is given by the integral on the rhs of (7). Thus there is no need to calculate the integral analytically. We have implicitly assumed that it was possible to generate the infinite dimensional variate $\omega \sim \mathbb{Z}_T^x$. However to compute the acceptance event $\Gamma$, we only need to know the value of $Z$ on a finite set of random times only, this set corresponding to the time coordinates of the realisations of $\Phi$. It is then possible to exchange the order in which $Z$ and $\Phi$ are generated, obtaining Algorithm 1.

**Algorithm 1** Exact Algorithm 1 (EA1)

1. SAMPLE $Z_T \sim h_{x,T}$
2. SAMPLE $\Phi = (\mathcal{X}, \Psi)$
3. SAMPLE $\left\{Z_{\chi_j}; 1 \leq j \leq |\mathcal{X}|\right\} \sim \mathbb{Z}_T^x \mid Z_T$
4. IF $\Gamma$ RETURN $\mathcal{S} := \left\{Z_{\chi_j}; 1 \leq j \leq |\mathcal{X}|\right\}, Z_T$
   ELSE GOTO 1

As already observed, step 3 of Algorithm 1 results in the generation of independent BBs. EA1 returns a skeleton $\mathcal{S}$ distributed according to the true law of the diffusion $X$. Furthermore, from [3] it is known that the conditional law $\mathbb{Z}_T^x \mid \mathcal{S}$ can be expressed as the product measure of independent BBs. Using this result it is possible to complete $\mathcal{S}$ on any finite arbitrary number of times if required and even to generate (exactly) certain functionals of the path of $X$.

The remaining part of this paper is organised as follows. In Sect. 2, the Quasi-EA is motivated and introduced and the connection with EA is examined. In Sect. 3, the two main theorems are proved. We start by establishing a local convergence result that is further extended by means of the maximal coupling inequality. We show that the Quasi-EA is an accurate continuous time approximation of the law of the diffusion process. In fact, we prove the existence of a myopic (sequential) coupling between the diffusion and the simulation scheme. The existence of such a coupling implies the convergence with respect to the total variation distance, a strong form of weak convergence according to [5]. Section 4 concludes the paper with some remarks about possible future research on the topic and practical considerations about our scheme.

The Euler scheme does not generally converge with respect to the total variation distance; see [9]. However, under mild technical conditions, the Euler scheme does converge with respect to total variation distance if the diffusion process has a constant diffusion term (see [6]). Genon-Catalot [4] extended this result to prove that the rate of convergence is of order $\Delta^{1/2}$, where $\Delta$ is the length of the (equally spaced) discretisation interval.

One of the main reasons for the interest in total variation distance is the well-known connection it shares with coupling. Essentially the total variation distance between two stochastic process laws is small if and only if there exists a probability space in which the probability that the two processes are identical is close to unity. However in the context of Markov processes, there is particular interest in jointly Markov (or *sequential* or *myopic*) couplings, which explicitly utilise Markov structure. A sequential coupling is a coupling in which, for each time increment in turn, we try to maximise the probability that the two processes stay together. We therefore focus on the existence of a sequential (or myopic) coupling between the diffusion and the simulation scheme.

Although total variation convergence is equivalent to the existence of a coupling, there is no guarantee that such a coupling is sequential. The existence of a sequential coupling clearly implies total variation convergence, however the converse is false in general. To illustrate this in the context of diffusion sample path approximation, consider

$$\mathrm{d}Y_t = \alpha \mathrm{d}t + \mathrm{d}B_t \tag{8}$$

$$\mathrm{d}X_t = \mathrm{d}B_t, \tag{9}$$

where $\alpha$ is a constant, $B_t$ is a scalar BM on $[0, T]$, and both $X_t$ and $Y_t$ start at the same value. Let $[0, T]$ be partitioned in intervals of length $\Delta$. Then elementary calculations show that the total variation distance between $Y_\Delta$ and $X_\Delta$ is of order $\mathcal{O}\left(\Delta^{1/2}\right)$, hence the probability of the coupling succeeds is $\left(1 - \Delta^{1/2}\right)^{1/\Delta}$ and this last quantity tends to 0 as $\Delta \downarrow 0$.

The major finding of this paper in Theorem 1 is to show that, unlike the Euler scheme, the Quasi-EA algorithm can be shown to exibit a sequential coupling with the true diffusion which is successful with probability converging to 1 as $\Delta \to 0$, on the whole interval $[0, T]$ with rate $\Delta^{1/2}$.

## 2 The quasi-EA

The main idea behind the construction of Quasi-EA is very simple. In step 2 of Algorithm 1, we sample a PPP with unit rate on $[0, T] \times [0, m]$. The acceptance rate in EA increases as $T \downarrow 0$, as the likelihood that no points from $\Phi$ are sampled increases too. Clearly in this eventuality, the proposed path is accepted.

As the acceptance rate can be interpreted as a measure of the quality of the proposal measure, we develop a scheme that always accept the proposal variate distributed as $\mathbb{Z}_T^x$. Given the previous considerations, this approximation is accurate only if $T$ is quite small. Hence the time interval $[0, T]$ is partitioned into smaller intervals on which the scheme is applied sequentially.

We now describe the scheme more precisely. The time interval $[0, T]$ is divided into $n$ smaller intervals having the same length $\Delta = T/n$. The continuous time scheme $Y$ is defined by the following equations

$$Y_0 = x \tag{10}$$

$$Y_{i\Delta} \sim h_{Y_{i\Delta},\Delta} \quad (i = 1, \ldots, n) \tag{11}$$

$$Y_s \sim \mathbb{BB}\left(Y_{i\Delta}, Y_{(i+1)\Delta}, \Delta\right) \quad (i\Delta < s < (i+1)\Delta) \tag{12}$$

where $\mathbb{BB}(x, \mathrm{d}y, t)$ is the measure of a BB starting at $(0, x)$ and ending at $(t, \mathrm{d}y)$. It can be easily seen the process $Y$ thus defined consists of $n$ sequential BBMs. The simulation of the Quasi-EA involves the sampling of the sequence of random variables in (11) only. However we are going to prove a stronger result than the convergence of the discretized process $\{Y_{i\Delta}; i = 0, \ldots, n\}$ to the diffusion process. We shall prove that the law of the continuous time process $Y$ defined by Eqs. (10)–(12) converges with respect to the total variation distance to the law of the diffusion process $X$ as $n \uparrow \infty$. This result suggests that the BBs are good process to *fill-in the gaps* between the simulated values of the discretized process.

We denote by $\mathbb{Y}_{0,T}^{x,n}$ the probability measure induced by the Quasi-EA scheme $Y$ started at $x$ and consisting of $n$ steps on $[0, T]$. Consistently with the previous notation $\mathbb{Y}_{0,\Delta}^{y}$ denotes the probability measure induced by this scheme on the single step $[0, \Delta]$ when $Y_0 = y$. To sum up

$$\{Y_s; 0 \le s \le T \mid Y_0 = x\} \sim \mathbb{Y}_{0,T}^{x,n} \tag{13}$$

$$\{Y_s; 0 \le s \le \Delta \mid Y_0 = y\} \sim \mathbb{Y}_{0,\Delta}^{y} = \mathbb{Z}_{\Delta}^{y}. \tag{14}$$

## 3 Two convergence results

The two convergence theorems require the two following lemmas.

**Lemma 1** *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function, such that for all sufficiently small $\Delta > 0$*

$$\frac{1}{\sqrt{2\pi\Delta}} \int_{\mathbb{R}} |f(y)| \, \mathrm{e}^{-\frac{(y-x)^2}{2\Delta}} \, \mathrm{d}y < \infty \tag{15}$$

*then (for any fixed $x \in \mathbb{R}$)*

$$\lim_{\Delta \downarrow 0} \frac{1}{\sqrt{2\pi\Delta}} \int_{\mathbb{R}} f(y) \mathrm{e}^{-\frac{(y-x)^2}{2\Delta}} \, \mathrm{d}y = f(x). \tag{16}$$

This is proved by elementary dominated convergence arguments. We omit the proof.

For ease of exposition, we assume that the state space of $X$ is $\mathbb{R}$, although this is easily generalised.

The following conditions, to be used selectively in the following results, are now introduced

- $(E1)\, \exists k \in (0,1), \exists c \in \mathbb{R}_+ : \alpha(u) \le c + \frac{k}{T} u \; (u \ge 0)$ *and* $\alpha(u) \ge c + \frac{k}{T} u \; (u < 0)$
- $(E2)\, \exists k \in (0,1), \exists c \in \mathbb{R}_+ : |\alpha(u)| \le c + \frac{k}{T} |u|, u \in \mathbb{R}$
- $(S1)$ $\alpha$ is twice continuously differentiable on $\mathbb{R}$.

**Lemma 2** *If condition $(E1)$ holds, then $\forall r > 0$*

$$\sup_{0 \le s \le T} \mathbb{E}_{X \sim h_{x,s}} [\mathrm{e}^{r|X|}] < \infty. \tag{17}$$

*Proof* From the definition of $A$, we get

$$A(u) \le r|u| + \frac{k}{2T} u^2 \; (u \in \mathbb{R})$$

and so

$$h_{x,s}(u) \le \exp\left\{ -\frac{(u-x)^2}{2s} + r|u| + \frac{k}{2T} u^2 \right\} \; (u \in \mathbb{R}, 0 \le s \le T).$$

If $u \ge 0$

$$h_{x,s}(u) \le \exp\left\{ -\frac{(u-\mu_+)^2}{2\sigma_+^2} \right\} r_+,$$

where $\mu_+ = \left(\frac{x+2sr}{1-\frac{sk}{T}}\right)$, $\sigma_+ = \frac{s}{\left(1-\frac{sk}{T}\right)}$, $r_+ = \exp\left\{-\frac{x^2+x+2sr}{1-\frac{sk}{T}}\right\}$. Similarly if $u < 0$

$$h_{x,s}(u) \leq \exp\left\{-\frac{(u-\mu_-)^2}{2\sigma_-^2}\right\} r_-,$$

where $\mu_- = \left(\frac{x-2sr}{1-\frac{sk}{T}}\right)$, $\sigma_+ = \sigma_-$, $r_- = \exp\left\{-\frac{x^2+x-2sr}{1-\frac{sk}{T}}\right\}$.

As $k \in (0,1)$ it follows that $r_+, r_-, \mu_+, \mu_-, \sigma_+, \sigma_-$ are bounded for $s \in (0, T]$ and the result follows. □

For any two probability measures $\mathbb{M}$, $\mathbb{N}$ on a measurable space $(E, \mathcal{E})$, let $||\mathbb{M}-\mathbb{N}||$ be their total variation metric, that is

$$||\mathbb{M}-\mathbb{N}|| := \sup_{A\in\mathcal{E}} |\mathbb{M}(A) - \mathbb{N}(A)|. \tag{18}$$

Finally we require the following proposition concerning total variation distance. Again no proof will be provided for these elementary facts.

**Proposition 1** *(i) Suppose that $\mathbb{P}_1$ is a probability measure and $B$ a $\mathbb{P}_1$-measurable set. Define $\mathbb{P}_2(A) = \mathbb{P}_1(A \cap B)/\mathbb{P}_1(B)$. Then*

$$\|P_1 - \mathbb{P}_2\| = \mathbb{P}_1(B^c).$$

*(ii) Let $\mathbb{P}_3$ be a probability measure with*

$$\frac{\mathrm{d}\mathbb{P}_3}{\mathrm{d}\mathbb{P}_1} \geq \mathrm{e}^{-\delta}$$

*for some non-negative constant $\delta$. Then*

$$\|\mathbb{P}_1 - \mathbb{P}_3\| \leq 1 - \mathrm{e}^{-\delta} \leq \delta.$$

We are now ready to state the following localised result:

**Theorem 1** *If condition $(S1)$ holds and one of conditions $(E1)$ and $(E2)$ holds, then for any fixed $x$ the law of the BBM $\mathbb{Z}_\Delta^x$ converges towards the law of the diffusion process $\mathbb{Q}_\Delta^x$ with respect to the total variation metric as $\Delta \downarrow 0$:*

$$\lim_{\Delta\downarrow 0} \left|\left|\mathbb{Z}_\Delta^x - \mathbb{Q}_\Delta^x\right|\right| = 0. \tag{19}$$

*Moreover, for all $\Delta$ sufficiently small, $\varepsilon > 0$*

$$\left|\left|\mathbb{Z}_\Delta^x - \mathbb{Q}_\Delta^x\right|\right| \leq k_x \Delta^{3/2-\varepsilon}, \tag{20}$$

*where the leading order constant $k_x$ is a continuous function of $x$ (i.e., the rate of convergence is at least $\mathcal{O}_x\left(\Delta^{3/2-\varepsilon}\right)$ for any $\varepsilon > 0$).*

*Proof* To ease the notation let $\mathbb{W} = \mathbb{W}^x_\Delta$ and $\mathbb{Q} = \mathbb{Q}^x_\Delta$ in the scope of this proof, and in fact generally suppress any unneccessary dependence on the initial state $x$. We set $B$ to be the event

$$B = \left\{ \sup_{0 \le s \le \Delta} |\omega_s - \omega_0| \le \Delta^{1/2-\varepsilon} \right\}.$$

We introduce a new probability measure $\mathbb{Q}^B$ via the following Radon–Nikodym derivative

$$\frac{\mathrm{d}\mathbb{Q}^B}{\mathrm{d}\mathbb{Q}} = 1_B / q_{x,\Delta}$$
$$q_{x,\Delta} = \mathbb{Q}[B] \ .$$

Similarly define $\mathbb{Z}^B$ to be $\mathbb{Z}$ conditioned on the event $B$. The triangle inequality gives

$$\begin{aligned} \|\mathbb{Z} - \mathbb{Q}\| &\le \left\|\mathbb{Z}^B - \mathbb{Q}^B\right\| + \left\|\mathbb{Q}^B - \mathbb{Q}\right\| + \left\|\mathbb{Z}^B - \mathbb{Z}\right\| \\ &= T_1 + T_2 + T_3, \end{aligned}$$

and we shall deal separately with the three terms.

First note that from (S1), there exists a constant $k_1$ such that

$$\left| \frac{\alpha^2 + \alpha'}{2}(z) - \frac{\alpha^2 + \alpha'}{2}(x) \right| \le k_1 |x - z| \le k_1 \Delta^{1/2-\epsilon}$$

at least for $z$ such that $|z - x| \le \Delta^{1/2-\epsilon}$. For $T_1$ note that if $\frac{\mathrm{d}\mathbb{Q}^B(\omega)}{\mathrm{d}\mathbb{W}^B} = f(\omega)$ then

$$\inf_\omega f(\omega) \ge \mathrm{e}^\delta \sup_\omega f(\omega),$$

where $\delta = 2k_1 \Delta^{3/2-\epsilon}$. Therefore by Proposition 1(ii),

$$T_1 \le k_1 \Delta^{3/2-\epsilon}.$$

By Proposition 1(i), the proof will thus be complete if we can demonstrate that $\mathbb{Q}(B^C)$ and $\mathbb{Z}(B^C)$ can be similarly bounded. From the definition of $\mathbb{Z}$ and an application of the Cauchy–Schwarz inequality, we get

$$\mathbb{Z}(B^C) = \frac{\mathbb{E}_\mathbb{W}[1_{\{B^C\}} \mathrm{e}^{A(w_\Delta)}]}{\mathbb{E}_\mathbb{W}[\mathrm{e}^{A(w_\Delta)}]} \le \frac{\mathbb{W}(B^C)^{\frac{1}{2}} \mathbb{E}_\mathbb{W}[\mathrm{e}^{2A(w_\Delta)}]^{\frac{1}{2}}}{\mathbb{E}_\mathbb{W}[\mathrm{e}^{A(w_\Delta)}]} \ .$$

Now recall that by assumption $(E1)$ (or $(E2)$), $A$ is bounded by a quadratic function so that for all sufficiently small $\Delta$, $\mathbb{E}_\mathbb{W}[\mathrm{e}^{2A(w_\Delta)}]$ is finite and bounded. Thus by Lemma 1 $\exists k_2 > 0$ s.t.

$$\mathbb{Z}(B^C) \le k_2 \mathbb{W}(B^C)^{\frac{1}{2}}.$$

Similar considerations can be applied to $\mathbb{Q}(B^C)$ yielding:

$$\mathbb{Q}(B^C) \le \frac{\mathbb{W}(B^C)^{\frac{1}{2}} \mathbb{E}_{\mathbb{W}} \left[ \mathrm{e}^{2A(w_\Delta) - 2\int_0^\Delta \phi(\omega_s)\mathrm{d}s} \right]^{\frac{1}{2}}}{\mathbb{E}_{\mathbb{W}} \left[ \mathrm{e}^{A(w_\Delta) - \int_0^\Delta \phi(\omega_s)\mathrm{d}s} \right]}$$

and as $\phi$ is a positive function by Lemma 1 $\exists k_3 > 0$ s.t.

$$\mathbb{Q}(B^C) \le k_3 \mathbb{W}(B^C)^{\frac{1}{2}}.$$

On the other hand, we can use the standard fluctuation result for BM:

$$\begin{aligned} \mathbb{W}(B^C) &\le 2\mathbb{W}(\sup_{0\le s\le \Delta} w_s \ge x + \Delta^{1/2-\epsilon}) \\ &= 4\mathbb{W}(w_\Delta \ge x + \Delta^{1/2-\epsilon}) = \mathcal{O}(\exp\{-\Delta^{-2\epsilon}/2\}). \end{aligned}$$

thus completing the proof. □

The extension of this localised result to the global case relies on the maximal coupling inequality. The coupling method (see [11]) is already prevalent in the field of SDEs, mainly in the multi-dimensional case. Our approach is very similar to that of [7]. Given the relevance of the coupling method, it is sensible to briefly introduce its basic elements. We recall:

**Definition 1** Let $(E, \mathcal{E})$ be a Polish measurable space, and $\mathbb{M}, \mathbb{N}$ be two probability measures on $(E, \mathcal{E})$. We state that a probability measure $\hat{\mathbb{P}}$ on $\left(E^2, \mathcal{E}^2\right)$ is a coupling of $(\mathbb{M}, \mathbb{N})$ if its marginals are $\mathbb{M}$ and $\mathbb{N}$. We also say that a random object $\left(\Omega', \mathcal{F}', \mathbb{P}', \left(X', Y'\right)\right)$, where $\left(\Omega', \mathcal{F}', \mathbb{P}'\right)$ is a probability space and $\left(X', Y'\right)$ is a $\mathcal{F}'/\mathcal{E}^2$-measurable function, is a coupling of $(\mathbb{M}, \mathbb{N})$ if the image measure $\mathbb{P}'\left(X', Y'\right)^{-1}$ is a coupling of $(\mathbb{M}, \mathbb{N})$.

The power of the coupling argument comes from the following Lemma

**Lemma 3** *Let $||\mathbb{M} - \mathbb{N}||$ be the total variation metric, that is*

$$||\mathbb{M} - \mathbb{N}|| := \sup_{A\in\mathcal{E}} |\mathbb{M}(A) - \mathbb{N}(A)| \tag{21}$$

(Coupling Inequality) *For any coupling* $\left(\Omega', \mathcal{F}', \mathbb{P}', \left(X', Y'\right)\right)$ *of* $(\mathbb{M}, \mathbb{N})$

$$||\mathbb{M} - \mathbb{N}|| \le \mathbb{P}'\left[X' \ne Y'\right] \tag{22}$$

(Maximal Coupling Equality) *There is a coupling* $\left(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \left(\hat{X}, \hat{Y}\right)\right)$ *of* $(\mathbb{M}, \mathbb{N})$ *s.t.*

$$||\mathbb{M} - \mathbb{N}|| = \hat{\mathbb{P}}\left[\hat{X} \ne \hat{Y}\right] \tag{23}$$

*This coupling is called the maximal coupling of* $(\mathbb{M}, \mathbb{N})$.

**Theorem 2** (Main Convergence Theorem) *If in addition to condition* $(S1)$ *conditions* $(E2)$ *and* $(S2)$ *also hold, where*

- $(S2)$ $\alpha'$ *is sub-quadratic, that is*

$$\left|\alpha'(u)\right| \leq c\left(1+u^2\right) \quad (u \in \mathbb{R}) \tag{24}$$

*or condition* $(E1)$ *and* $(S3)$ *also hold, where*

- $(S3)$ $\alpha$ *and* $\alpha'$ *are sub-exponential, that is that:*

$$|\alpha(u)|, \left|\alpha'(u)\right| \leq c\left(1+\mathrm{e}^{c|u|}\right) \quad (u \in \mathbb{R}) \tag{25}$$

*then there exist a (myopic) coupling* $\left(\widehat{\mathbb{P}}, \hat{X}, \hat{Y}\right)$ *of* $\left(\mathbb{Q}^x_{0,T}, \mathbb{Y}^{x,n}_{0,T}\right)$ *such that*

$$\lim_{n\to\infty} \hat{\mathbb{P}}\left[\hat{X} \neq \hat{Y}\right] = 0 \tag{26}$$

*and the rate of convergence is at least* $\mathcal{O}\left(\Delta^{1/2-\varepsilon}\right)$ *for any* $\varepsilon > 0$. *As a consequence of the coupling inequality,* $\mathbb{Y}^{x,n}_{0,T}$ *converges towards* $\mathbb{Q}^x_{0,T}$ *with respect to the total variation metric with the same rate of convergence.*

*Remark 1* If $(E2)$ holds then condition $(S2)$ is a weak assumption. As $\alpha^2 + \alpha'$ is bounded below, this additional condition means that the drift coefficient cannot oscillate too quickly as $|u| \to \infty$. Moreover in most diffusion models condition $(E1)$ is satisfied, as otherwise the diffusion would exhibit explosive behaviour.

*Proof* We build a probability space $\left(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}\right)$ and two measurable functions

$$\begin{aligned}\left(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}\right) &\to \tilde{X}(s, x, y) \quad s \in (0, \Delta] \\ &\to \tilde{Y}(s, x, y) \quad s \in (0, \Delta]\end{aligned}$$

that define the maximal coupling of $\left(\mathbb{Q}^x_{0,\Delta}, \mathbb{Y}^y_{0,\Delta}\right)$. A coupling of $\left(\mathbb{Q}^x_{0,T}, \mathbb{Y}^{x,n}_{0,T}\right)$ that starts from this maximal coupling is constructed on the single time interval $(0, \Delta]$. The initial step is defined by setting $\hat{X}_0 = \hat{Y}_0 = x$ and $\hat{X}_s = \tilde{X}(s, x, x)$, $\hat{Y}_s = \tilde{Y}(s, x, x)$ for $s \in (0, \Delta]$. Now suppose that $\left(\hat{X}, \hat{Y}\right)$ is a coupling of $\left(\mathbb{Q}^x_{0,i\Delta}, \mathbb{Y}^{x,n}_{0,i\Delta}\right)$. Let

$$\begin{aligned}\hat{X}_{i\Delta+s} &:= \tilde{X}\left(s, \hat{X}_{i\Delta}, \hat{Y}_{i\Delta}\right) \quad s \in (0, \Delta] \\ \hat{Y}_{i\Delta+s} &:= \tilde{Y}\left(s, \hat{X}_{i\Delta}, \hat{Y}_{i\Delta}\right) \quad s \in (0, \Delta]\end{aligned}$$

independent of $\left(\hat{X}, \hat{Y}\right)$ on $[0, i\Delta]$. From the time homogeneity and Markov property of the processes $X, Y$ (limited to the set of times $\{i\Delta; i = 1, \ldots, n-1\}$) it follows that $\left(\hat{X}, \hat{Y}\right)_{s\in(i\Delta,(i+1)\Delta]}$ is a coupling of $\left(\mathbb{Q}^{\hat{X}_{i\Delta}}_{i\Delta,(i+1)\Delta}, \mathbb{Y}^{\hat{Y}_{i\Delta}}_{i\Delta,(i+1)\Delta}\right)$ and that the extended process $\left(\hat{X}, \hat{Y}\right)_{s\in[0,(i+1)\Delta]}$ is a coupling of $\left(\mathbb{Q}^{x}_{0,(i+1)\Delta}, \mathbb{Y}^{x,n}_{0,(i+1)\Delta}\right)$. The induction step is thus satisfied. No measurability problem arise in the definition of $\tilde{X}, \tilde{Y}$ and they can be chosen to be jointly measurable in $(x, y)$, see [7] for the technical details. The coupling inequality yields

$$\begin{aligned}
&\left|\left|\mathbb{Y}^{x_0,n}_{0,T} - \mathbb{Q}^{x_0}_{0,T}\right|\right| \\
&\quad \le \hat{\mathbb{P}}\left[\exists s \in [0, T] : \hat{X}_s \ne \hat{Y}_s\right] \\
&\quad \le \hat{\mathbb{P}}\left[\hat{X}_s \ne \hat{Y}_s \text{ on } [0, \Delta]\right] \\
&\sum_{i=0}^{n-2} \hat{\mathbb{P}}\left[\hat{X} \ne \hat{Y} \text{ on } ((i+1)\,\Delta, i\Delta] \mid \hat{X}_{i\Delta} = \hat{Y}_{i\Delta}\right]
\end{aligned}$$

from the structure of the coupling $\left(\hat{X}, \hat{Y}\right)$, and the generic term of this last quantity is exactly the maximal coupling of $\left(\mathbb{Q}^{\hat{X}_{i\Delta}}_{i\Delta,(i+1)\Delta}, \mathbb{Y}^{\hat{Y}_{i\Delta}}_{i\Delta,(i+1)\Delta}\right)$. Hence, it is possible to use the maximal coupling equality. By defining a family of sub-probability measures on $\mathbb{R}$ by $\mathbb{S}_i(\mathcal{A}) := \hat{\mathbb{P}}\left[\hat{X}_{i\Delta} = \hat{Y}_{i\Delta} \in \mathcal{A}\right]$ and by using this last consideration

$$\begin{aligned}
&\left|\left|\mathbb{Y}^{x_0,n}_{0,T} - \mathbb{Q}^{x_0}_{0,T}\right|\right| \\
&\quad \le \left|\left|\mathbb{Y}^{x_0}_{0,\Delta} - \mathbb{Q}^{x_0}_{0,\Delta}\right|\right| + \sum_{i=0}^{n-2} \int \left|\left|\mathbb{Y}^{s}_{0,\Delta} - \mathbb{Q}^{s}_{0,\Delta}\right|\right| \mathrm{d}\mathbb{S}_i(s) \\
&\quad \le k_{x_0}\Delta^{3/2-\varepsilon} + \Delta^{3/2-\varepsilon} \sum_{i=0}^{n-2} \int k_s \mathrm{d}\mathbb{S}_i(s)
\end{aligned}$$

Similarly to the proof of Theorem 1

$$k_x := \sup_{|y|\le|x|+1} \frac{\left|\frac{\alpha^2+\alpha'}{2}(y) - \frac{\alpha^2+\alpha'}{2}(x)\right|}{|y-x|}.$$

We see that the behaviour in the tails as $|x| \to \infty$ is determined by

$$\overline{\lim}_{|x|\to\infty} \tilde{k}_x := \overline{\lim}_{|x|\to\infty} \frac{\left|\frac{\alpha^2+\alpha'}{2}(x)\right|}{|x|}$$

If condition $(S2)$ holds the positive function $\tilde{k}_x$ can diverge at most linearly. So $k_x \le k(1+|x|)$ and

$$\int k_s \mathrm{d}\mathbb{S}_i(s) \leq k + k \int |s| \, \mathrm{d}\mathbb{S}_i(s) \leq k + k \int |s| \, \mathrm{d}\mathbb{Q}^{x_0}_{i\Delta}(s),$$

where the last integral is the absolute momentum of $X_{i\Delta} : \mathbb{E}[|X_{i\Delta}| \mid X_0 = x_0]$. It is a known result that $(E2)$ implies

$$\mathbb{E}[|X_{i\Delta}| \mid X_0 = x_0] \leq \mathbb{E}\left[\sup_{0\leq s\leq T} |X_s|^2 \mid X_0 = x_0\right] < \infty$$

$\forall i \in \mathbb{N}$. Therefore

$$\begin{aligned}
&\left|\left|\mathbb{Y}^{x_0,n}_{0,T} - \mathbb{Q}^{x_0}_{0,T}\right|\right| \\
&\quad \leq k_{x_0}\Delta^{3/2-\varepsilon} + \Delta^{3/2-\varepsilon}\sum_{i=0}^{n-2} k\left(1 + \mathbb{E}\left[\sup_{0\leq s\leq T}|X_s|^2 \mid X_0 = x_0\right]\right) \\
&\quad \leq \mathrm{d}\Delta^{3/2-\varepsilon} n = \frac{d}{T}\Delta^{1/2-\varepsilon}
\end{aligned}$$

for $\Delta$ sufficiently small.

If instead condition $(S3)$ holds, by looking at $\overline{\lim}_{|x|\to\infty}\tilde{k}_x$, we obtain the sub-exponential growth condition on $k_x$, that is

$$k_x \leq k\left(1 + \mathrm{e}^{k|x|}\right) \; (x \in \mathbb{R}).$$

As a consequence

$$\int k_s d\mathbb{S}_i(s) \leq k + k \int \mathrm{e}^{k|s|} \mathrm{d}\mathbb{S}_i(s) \leq k + k \int \mathrm{e}^{k|s|} d\mathbb{Z}^{x_0}_{i\Delta}(s).$$

Again, the last integral is $\mathbb{E}\left[\mathrm{e}^{k|Y_{i\Delta}|} \mid Y_0 = x_0\right]$. As condition $(E1)$ holds, Lemma 2 implies that

$$\sup_{i=1,\ldots,n-1} \mathbb{E}\left[\mathrm{e}^{k|Y_{i\Delta}|} \mid Y_0 = x_0\right] < \infty.$$

Finally

$$\begin{aligned}
&\left|\left|\mathbb{Y}^{x_0,n}_{0,T} - \mathbb{Q}^{x_0}_{0,T}\right|\right| \\
&\quad \leq k_{x_0}\Delta^{3/2-\varepsilon} + \Delta^{3/2-\varepsilon}\sum_{i=0}^{n-2} k\left(1 + \sup_{i=1,\ldots,n-1}\mathbb{E}\left[\mathrm{e}^{k|Y_{i\Delta}|} \mid Y_0 = x_0\right]\right) \\
&\quad \leq \mathrm{d}\Delta^{3/2-\varepsilon} n = \frac{d}{T}\Delta^{1/2-\varepsilon}.
\end{aligned}$$

□

## 4 Conclusion

In this paper, we proved two convergence results about the Quasi-EA simulation scheme.

We shown the convergence of the law of the BBM $Z$ to the law of the diffusion process $X$ when both are started at the same value and the time interval $[0, \Delta]$ shrinks to zero. The convergence is obtained with respect to the total variation distance and an upper bound for the rate of the convergence is shown to be $\mathcal{O}_x\left(\Delta^{3/2-\varepsilon}\right)$ $\forall\varepsilon > 0$. The notation underlines that this speed of convergence is not necessarily uniform in $x$.

We also extend this convergence to the global case of a fixed time interval $[0, T]$. In this case, $[0, T]$ is uniformly partitioned in $n$ intervals and the convergence is obtained as $n \uparrow \infty$. The main difficulty that the starting point for $X$ and $Y$ on each single interval is not the same anymore is overcome using the coupling method. We are thus able to construct a successful myopic coupling of the Quasi-EA and the diffusion. Consequently, we obtain the convergence with respect to the total variation distance and an upper bound for the rate of convergence $\mathcal{O}\left(\Delta^{1/2-\varepsilon}\right)$ $\forall\varepsilon > 0$.

Very efficient algorithms to sample from the parametric family of densities $\{h_{x,T}\}_{x\in\mathbb{X}}$ are introduced in [8]. However, while the quasi-EA is more efficient than the Euler scheme, a brief simulation study suggests that Predictor–Corrector schemes result in a more accurate simulation method.

As already noted, the hypothesis used in the derivation of these results include the conditions that permit the simulation of $X$ using EA3. It is possible to weaken the condition used in our work and prove the convergence of the Quasi-EA even in models where EA3 can not be applied, and this will be the focus of future research. This could be worked out in future research. The main contribute in this paper is to obtain an insight into the role of the BBM $Z$ in the context of EA.

**Acknowledgments** All three authors would like to thank EPSRC for support through grant EP/D505607/2. Roberts would additionally like to thank EPSRC for support through grants EP/20620/01 and EP/S61577/01.



## References

1. Beskos, A., Roberts, G.O.: Exact simulation of diffusions. Ann. Appl. Probab. **15**, 2422–2444 (2005)
2. Beskos, A., Papaspiliopoulos, O., Roberts, G.O.: A new factorisation of diffusion measure and finite sample path construction. Methodol. Comput. Appl. Probab. **10**(1), 85–104 (2008)
3. Beskos, A., Papaspiliopoulos, O., Roberts, G.O.: Retrospective exact simulation of diffusion sample paths with applications. Bernoulli **12**, 1077–1098 (2006)
4. Genon-Catalot, V.: Personal comunication (2007)
5. Jacka, S.D., Roberts, G.O.: On strong forms of weak convergence. Stoch. Proc. Appl. **67**, 41–53 (1997)
6. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer, New York (1987)
7. Odasso, C.: Méthodes de couplage pour des équations stochastiques de type Navier-Stokes et Schrödinger. PhD thesis, Institut de Recherche Mathématique de Rennes & Ecole Normale Supérieure de Cachan, antenne de Bretagne (2005)

8. Peluchetti, S.: An analysis of the efficiency of the exact algorithm. PhD thesis, IMQ-Universitá Commerciale Luigi Bocconi (2007)
9. Perez, J.A., Glynn, P., Goodman, J.: Numerical schemes for simulating SDEs under the total variation norm (2006)
10. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes and Martingales. Wiley, New York (1987)
11. Thorisson, H.: Coupling, Stationarity, and Regeneration. Springer, New York (2000)