technische universität
dortmund

Weak approximation of SDEs by discrete-time processes

Henryk Zähle

Weak approximation of SDEs by discrete-time processes

Henryk Zähle*

Abstract

We consider the martingale problem related to the solution of an SDE on the line. It is shown that the solution of this martingale problem can be approximated by solutions of the corresponding time-discrete martingale problems under some conditions. This criterion is especially expedient for establishing the convergence of population processes to SDEs. We also show that the criterion yields a weak Euler scheme approximation of SDEs under fairly weak assumptions on the driving force of the approximating processes.

Key words and phrases. Stochastic differential equation, martingale problem, Doob-Meyer decomposition, discrete-time process, weak convergence, Galton-Watson process, Euler scheme

MSC 2000 subject classification. 60H10, 60J80, 60F17

---

*Postal address: Dortmund University of Technology, Faculty of Mathematics, Vogelpothsweg 87, D-44227 Dortmund, Germany; Email: henryk.zaehle@math.uni-dortmund.de

1 Introduction

It is well-known that a re-scaled version of the classical Galton-Watson process (GWP) with offspring variance σ^2 weakly converges to the unique solution of the following one-dimensional stochastic differential equation (SDE)

$$dX_t = \sigma \sqrt{|X_t|} \, dW_t \tag{1}$$

where W is a one-dimensional Brownian motion (cf. [14]). One might ask whether it is possible to approximate more general SDEs, driven by a Brownian motion, by generalized GWPs. In [21] it will be shown that this is actually possible. In fact, in [21] the solution of the SDE

$$dX_t = \delta(t, X_t)dt + \sigma(t, X_t)\sqrt{|X_t|} \, dW_t \tag{2}$$

is weakly approximated by two different types of population-size-dependent GWPs (in the sense of [5], [8], [10], [11]) with immigration, where δ and σ are suitable nonnegative continuous functions on $\mathbb{R}_+ \times \mathbb{R}$. Here the methods of [14] do not apply anymore (cf. Section 3). In the present article, we establish a general criterion for the weak approximation of SDEs by discrete-time processes, which is the crux of the analysis of [21].

To be exact, we focus on the following one-dimensional SDE

$$dX_t = b(t, X_t)dt + a(t, X_t)dW_t, \qquad X_0 = x_0 \tag{3}$$

where $x_0 \in \mathbb{R}$ and W is a one-dimensional Brownian motion. The coefficients a and b are continuous functions on $\mathbb{R}_+ \times \mathbb{R}$ satisfying

$$|a(t,x)| + |b(t,x)| \leq K(1 + |x|) \qquad \forall \, t \in \mathbb{R}_+, \; x \in \mathbb{R} \tag{4}$$

for some finite constant $K > 0$. We assume that SDE (3) has a weak solution. It means that there exists a triplet $\{X; W; (\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})\}$ where $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is a filtered probability space with $(\mathcal{F}_t)$ satisfying the usual conditions, $W = (W_t : t \geq 0)$ is an $(\mathcal{F}_t)$-Brownian motion and $X = (X_t : t \geq 0)$ is a real-valued continuous $(\mathcal{F}_t)$-adapted process such that $\mathbb{P}$-almost surely,

$$X_t = x_0 + \int_0^t b(r, X_r)dr + \int_0^t a(r, X_r)dW_r \qquad \forall \, t \geq 0.$$

Here the latter is an Itô-integral. Moreover we require the solution to be weakly unique, which means that any two solutions coincide in law. For instance, the existence of a unique weak solution is implied by Lipschitz continuity of b in x (uniformly in t) and

$$|a(t,x) - a(t,x')| \leq h(|x - x'|) \qquad \forall \, t \in \mathbb{R}_+, \; x, x' \in \mathbb{R} \tag{5}$$

for some strictly increasing $h : \mathbb{R}_+ \to \mathbb{R}_+$ with $\int_0^{0+} h^{-2}(u)du = \infty$. Note that (5) and Lipschitz continuity of b even imply the existence of a strongly unique strong solution (Yamada-Watanabe criterion [20]). But the notion of strong solutions and strong uniqueness is beyond our interest.

Our starting point is the fact that any weak solution of (3) is a solution of the following martingale problem and vice versa (cf. Section 5.4.B of [9], or Theorem 1.27 of [1]).

Definition 1.1 *A tuple* $\{X; (\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})\}$ *is said to be a solution of the (a,b,x_0)-martingale problem if* $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ *is a filtered probability space with* $(\mathcal{F}_t)$ *satisfying the usual conditions and* $X = (X_t : t \geq 0)$ *is a real-valued continuous* $(\mathcal{F}_t)$*-adapted process such that*

$$M_t = X_t - x_0 - \int_0^t b(r, X_r) dr \tag{6}$$

provides a (continuous, mean-zero) square-integrable $(\mathcal{F}_t)$*-martingale with compensator*

$$\langle M \rangle_t = \int_0^t a^2(r, X_r) dr. \tag{7}$$

The solution is said to be unique if any two solutions coincide in law.

In view of the weak equivalence of the SDE to the martingale problem, discrete-time processes solving the discrete analogue (Definition 2.1) of the (a,b,x_0)-martingale problem should approximate weakly the unique solution of SDE (3). Theorem 2.2 below shows that this is true under an additional assumption on the moments of the increments (condition (10)).

Note that the characterization of discrete or continuous population processes as solutions of martingale problems of the form (6)-(7) respectively (8)-(9) is fairly useful and also common (see e.g. [16], [17], [19]). Especially for real-valued discrete-time processes these characterizations are often easy to see, so that, according to the criterion, the only thing to check is condition (10). Also note that the conditions of the famous criterion of Stroock and Varadhan for the weak convergence of Markov chains to SDEs ([18] Theorem 11.2.3) are different. In particular, in our framework we do not insist on the Markov property of the approximating processes (cf. the discussion at the end of Section 4). Another alternative approach to the discrete-time approximation of SDEs can be found in the seminal paper [13], see also references therein. In [13] general conditions are given under which the convergence in distribution $(Y_\alpha, Z_\alpha) \to (Y, Z)$ in the cádlàg space implies convergence in distribution $\int Y_\alpha dZ_\alpha \to \int Y dZ$ of the corresponding stochastic integrals in the cádlàg space.

In Section 3 we will demonstrate that the criterion of Theorem 2.2 yields an easy proof of the convergence result discussed at the beginning of the Introduction. Moreover, in Section 4 we will apply our criterion to obtain a weak Euler scheme approximation of SDEs under fairly weak assumptions on the driving force of the approximating processes.

2 Main result

We will regard discrete-time processes as continuous-time cádlàg processes. For this reason we denote by $D(\mathbb{R})$ the space of cádlàg functions from $\mathbb{R}_+$ to $\mathbb{R}$. We equip $D(\mathbb{R})$ with the topology generated by the Skohorod convergence on compacts and consider it as a measurable space with respect to its Borel σ-algebra. Moreover we set $t_n^\epsilon = n\epsilon$ for every $n \in \mathbb{N}_0$ and $\epsilon > 0$.

For every $\alpha \in \mathbb{N}$ we fix some $\epsilon_\alpha > 0$ such that $\epsilon_\alpha \to 0$. For the sake of clarity we also set $t_n^\alpha = t_n^{\epsilon_\alpha} (= n\epsilon_\alpha)$ for all $n \in \mathbb{N}_0$. Now suppose a_α and b_α are measurable functions on

$\mathbb{R}_+ \times \mathbb{R}$ such that $\|a - a_\alpha\|_\infty$ and $\|b - b_\alpha\|_\infty$ converge to 0 as $\alpha \to \infty$, where $\|.\|_\infty$ is the usual supremum norm. Let $(x_\alpha) \subset \mathbb{R}$ satisfy $x_\alpha \to x_0$, and suppose that X^α is a solution of the following $(\epsilon_\alpha, a_\alpha, b_\alpha, x_\alpha)$-martingale problem for every $\alpha \geq 1$. Here we write $n_\alpha(t)$ for the largest $n \in \mathbb{N}_0$ with $t_n^\alpha \leq t$.

Definition 2.1 *Suppose* $X^\alpha = (X_t^\alpha : t \geq 0)$ *is a real-valued process on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *whose trajectories are constant on the intervals* $[t_n^\alpha, t_{n+1}^\alpha)$, $n \in \mathbb{N}_0$. *Then* X^α *is called solution of the* $(\epsilon_\alpha, a_\alpha, b_\alpha, x_\alpha)$-*martingale problem if*

$$M_t^\alpha = X_t^\alpha - x_\alpha - \sum_{i=0}^{n_\alpha(t)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^\alpha) \, \epsilon_\alpha \tag{8}$$

provides a (zero-mean) square-integrable martingale (w.r.t. the natural filtration) with compensator

$$\langle M^\alpha \rangle_t = \sum_{i=0}^{n_\alpha(t)-1} a_\alpha^2(t_i^\alpha, X_{t_i^\alpha}^\alpha) \, \epsilon_\alpha. \tag{9}$$

The X^α could be defined on different probability spaces $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbb{P}_\alpha)$. However we assume without loss of generality $\Omega_\alpha = D(\mathbb{R})$, $\mathcal{F}_\alpha = \mathcal{B}(D(\mathbb{R}))$ and that X^α is the coordinate process of $\mathbb{P}_\alpha$ (each cádlàg process induces a corresponding law on $D(\mathbb{R})$). We further assume that there are some $q > 2$ and $\delta > 1$ such that

$$\mathbb{E}_\alpha \left[|X_{t_n^\alpha}^\alpha - X_{t_{n-1}^\alpha}^\alpha|^q \right] \leq C_T \left(1 + \mathbb{E}_\alpha \left[|X_{t_{n-1}^\alpha}^\alpha|^q \right] \right) \epsilon_\alpha^\delta \tag{10}$$

for every $\alpha \geq 1$ and $n \in \mathbb{N}$ with $t_n^\alpha \leq T$, where $C_T > 0$ is some finite constant that may depend on T. (By an induction on n, (10) implies immediately that $\mathbb{E}_\alpha[|X_{t_n^\alpha}^\alpha|^q] < \infty$ for all α and n. Lemma 5.1 will provide an even stronger statement.) The following theorem shows that X^α converges in distribution to the unique solution of (3).

Theorem 2.2 *Suppose SDE (3) subject to (4) has a unique weak solution, and denote by* $\mathbb{P}$ *the corresponding law on* $D(\mathbb{R})$. *Moreover, let* $\mathbb{P}_\alpha$ *be the law (on* $D(\mathbb{R})$*) of* X^α *subject to (8)-(10). Then* $\mathbb{P}_\alpha \Rightarrow \mathbb{P}$ *as* $\alpha \to \infty$.

Here $\Rightarrow$ symbolizes weak convergence. The proof of Theorem 2.2 will be carried out in Section 5. The finiteness of the q-th moments for some $q > 2$ is not always necessary, it is true. From time to time the finiteness of the second moments is sufficient. However for a general statement involving convenient moment conditions as (10), a weakening of $q > 2$ to $q = 2$ is hardly possible. The assumption $q > 2$ is common in the theory of functional, time-discrete approximations of SDEs, SDDEs and SPDEs (see e.g. [15], [19]).

3 Example 1: Convergence of re-scaled GWP to (1)

As a first application of Theorem 2.2, we show that a re-scaled GWP weakly converges to Feller's branching diffusion ([4]), i.e., to the solution of SDE (1). Lindvall [14] showed this approximation via the convergence of the finite-dimensional distributions, for which the shape

of the Laplace transforms of the transition probabilities is essential. Here we shall exploit the martingale property of the Galton-Watson process (with offspring variance σ^2). The latter is an $\mathbb{N}_0$-valued Markov process $Z = (Z_n : n \in \mathbb{N}_0)$ that can be defined recursively as follows. Choose an initial state $Z_0 \in \mathbb{N}$ and set $Z_n = \sum_{i=1}^{Z_{n-1}} N_{n-1,i}$ for all $n \geq 1$, where $\{N_{n,i} : n \geq 0, \ i \geq 1\}$ is a family of i.i.d. $\mathbb{N}_0$-valued random variables with mean 1 and variance σ^2. In addition we require that the 4-th moment of $N_{1,1}$ is finite. Thereby Z_n has a finite 4-th moment for every $n \in \mathbb{N}_0$. Actually, in [14] the finiteness of the 4-th moments was not required. On the other hand, the methods used there break down when considering a population-size-dependent branching intensity or an additional general immigration into the system. In contrast, the procedure below still works in that cases (cf. [21]).

Setting $Z_{t_n^\epsilon}^\epsilon = \epsilon Z_n$ we obtain a re-scaled version, Z^ϵ, of Z. Recall $t_n^\epsilon = n\epsilon$, hence Z^ϵ is a process having $\epsilon\mathbb{N}_0 = \{0, \epsilon, 2\epsilon, \ldots\}$ as both its index set and its state space. Now pick $(\epsilon_\alpha) \subset \mathbb{R}_+$ such that $\epsilon_\alpha \to 0$, and recall our convention $t_n^\alpha = t_n^{\epsilon_\alpha}$ and that $\lfloor t \rfloor_\epsilon$ denotes the largest element s of $\epsilon\mathbb{N}_0$ with $s \leq t$. Regard the process Z^{ϵ_α} as continuous-time process, X^α, by setting $X_t^\alpha = Z_{\lfloor t \rfloor_{\epsilon_\alpha}}^{\epsilon_\alpha}$, and suppose $X_0^\alpha = \lfloor x_0 \rfloor_{\epsilon_\alpha}$. The latter requires that Z_0 actually depends on α. The domain of X^α is denoted by $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbb{P}_\alpha)$. It is easy to see that M^α defined in (8) provides a (zero-mean) square-integrable martingale. Moreover the compensator of M^α is given by $\langle M^\alpha \rangle_t = \sigma^2 \sum_{i=0}^{n_\alpha(t)-1} X_{t_i^\alpha}^\alpha \, \epsilon_\alpha$ since in this case,

$$\mathbb{E}_\alpha\left[\left((M_{t_n^\alpha}^\alpha)^2 - \langle M^\alpha \rangle_{t_n^\alpha}\right) - \left((M_{t_{n-1}^\alpha}^\alpha)^2 - \langle M^\alpha \rangle_{t_{n-1}^\alpha}\right)\Big| \mathcal{F}_{t_{n-1}^\alpha}^{X^\alpha}\right] = 0$$

can be checked easily with help of

$$\mathbb{E}_\alpha\left[X_{t_n^\alpha}^\alpha | X_{t_{n-1}^\alpha}^\alpha\right] = X_{t_{n-1}^\alpha}^\alpha \qquad \text{and} \qquad \mathbb{V}\mathrm{ar}_\alpha\left[X_{t_n^\alpha}^\alpha | X_{t_{n-1}^\alpha}^\alpha\right] = \sigma^2 X_{t_{n-1}^\alpha}^\alpha \, \epsilon_\alpha. \tag{11}$$

The formulae in (11) are immediate consequences of the well-known moment formulae for Z (see [7] p.6) and $(\mathcal{F}_{t_n^\alpha}^{X^\alpha})$ denotes the natural filtration induced by X^α. Hence X^α solves the $(\epsilon_\alpha, a, b, x_\alpha)$-martingale problem of Definition 2.1 with $a(t, x) = \sqrt{|x|}$, $b \equiv 0$ and $x_\alpha = \lfloor x_0 \rfloor_{\epsilon_\alpha}$. It remains to show (10). To this end we state the following lemma.

Lemma 3.1 *Assume $\xi_1, \xi_2, \ldots$ are independent random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[\xi_i] = 0$ and $\sup_{i \in \mathbb{N}} \mathbb{E}[\xi_i^4] < \infty$. Let ν be a further random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ being independent of (ξ_i), taking values in $\mathbb{N}$ and satisfying $\mathbb{E}[\nu^4] < \infty$. Then there is some finite constant $C > 0$, depending only on the second and the fourth moments of the ξ_i, such that $\mathbb{E}[(\sum_{i=1}^\nu \xi_i)^4] \leq C\mathbb{E}[\nu^2]$.*

Proof By the finiteness of the fourth moments the law of total expectation yields

$$\mathbb{E}\left[\left(\sum_{i=1}^\nu \xi_i\right)^4\right] = \sum_{n \in \mathbb{N}} \sum_{i_1=1}^n \sum_{i_2=1}^n \sum_{i_3=1}^n \sum_{i_4=1}^n \mathbb{E}\left[\xi_{i_1}\xi_{i_2}\xi_{i_3}\xi_{i_4}\right] \mathbb{P}[\nu = n].$$

Since the ξ_i are independent and centered, the summand on the right-hand side might differ from 0 only if either $i_1 = i_2 = i_3 = i_4$, or $i_1 = i_2$ and $i_3 = i_4 \neq i_1$, or $i_1 = i_3$ and $i_2 = i_4 \neq i_1$,

or $i_1 = i_4$ and $i_2 = i_3 \neq i_1$. Hence,

$$\mathbb{E}\left[\left(\sum_{i=1}^{\nu} \xi_i\right)^4\right] \leq \sum_{n \in \mathbb{N}} \left\{ (n + 3n(n-1)) \sup_{i,j \in \mathbb{N}} \mathbb{E}\left[\xi_i^2 \xi_j^2\right] \right\} \mathbb{P}[\nu = n] \leq 4 \sup_{i,j \in \mathbb{N}} \mathbb{E}\left[\xi_i^2 \xi_j^2\right] \mathbb{E}\left[\nu^2\right].$$

This yields the claim of the lemma with $C = 4 \sup_{i,j \in \mathbb{N}} \mathbb{E}[\xi_i^2 \xi_j^2]$. $\qquad\square$

With help of Lemma 3.1 we obtain

$$\mathbb{E}_\alpha\left[|X_{t_n^\alpha}^\alpha - X_{t_{n-1}^\alpha}^\alpha|^4\right] = \mathbb{E}_\alpha\left[\Big|\sum_{i=1}^{\epsilon_\alpha^{-1} X_{t_{n-1}^\alpha}^\alpha} (\epsilon_\alpha N_{n-1,i} - \epsilon_\alpha)\Big|^4\right] = \mathbb{E}_\alpha\left[\Big|\sum_{i=1}^{\epsilon_\alpha^{-1} X_{t_{n-1}^\alpha}^\alpha} (N_{n-1,i} - 1)\Big|^4\right] \epsilon_\alpha^4$$

$$\leq C \, \mathbb{E}_\alpha\left[(\epsilon_\alpha^{-1} X_{t_{n-1}^\alpha}^\alpha)^2\right] \epsilon_\alpha^4 \leq C\left(1 + \mathbb{E}_\alpha\left[(X_{t_{n-1}^\alpha}^\alpha)^4\right]\right) \epsilon_\alpha^2$$

for some suitable constant $C > 0$. This shows that (10) holds, too. Hence the assumptions of Theorem 2.2 are fulfilled, and the theorem implies that X^α converges in distribution to the unique solution of (1).

4 Example 2: Weak Euler scheme approximation of (3)

As a second application of Theorem 2.2, we establish a weak Euler scheme approximation of SDE (3). Our assumptions are partially weaker then the assumptions of classical results on weak functional Euler scheme approximations. A standard reference for Euler schemes is the monograph [12]; see also references therein. As before we suppose that a and b are continuous functions on $\mathbb{R}_+ \times \mathbb{R}$ satisfying (4), and that SDE (3) possesses a unique weak solution. Now let $\epsilon > 0$, recall the notation introduced in Section 2 and consider the following stochastic difference equation (weak Euler scheme)

$$X_{t_n^\epsilon}^\epsilon - X_{t_{n-1}^\epsilon}^\epsilon = b(t_{n-1}^\epsilon, X_{t_{n-1}^\epsilon}^\epsilon)\,\epsilon + a(t_{n-1}^\epsilon, X_{t_{n-1}^\epsilon}^\epsilon)\,V_{t_n^\epsilon}^\epsilon, \qquad X_{t_0^\epsilon}^\epsilon = x_\epsilon. \tag{12}$$

Here (x_ϵ) is a sequence in $\mathbb{R}$ satisfying $x_\epsilon \to x_0$ as $\epsilon \to 0$, and $V^\epsilon = \{V_{t_n^\epsilon}^\epsilon : n \in \mathbb{N}\}$ is a family of independent centered random variables with variance ϵ and $\mathbb{E}_\epsilon[|V_{t_n^\epsilon}^\epsilon|^q] \leq C\epsilon^{q/2}$ for all $n \in \mathbb{N}$, $\epsilon \in (0,1]$, some $q > 2$ and some finite constant $C > 0$, where $(\Omega_\epsilon, \mathcal{F}_\epsilon, \mathbb{P}_\epsilon)$ denotes the domain of V^ϵ. For instance, one may set $V_{t_n^\epsilon}^\epsilon = \sqrt{\epsilon}\,\xi_n$ where $\{\xi_n : n \in \mathbb{N}\}$ is a family of independent centered random variables with variance 1 and q-th moment being bounded uniformly in n. Note that we do not require that the random variables $\{V_{t_n^\epsilon}^\epsilon : n \in \mathbb{N}\}$ are identically distributed. Below we will see that the independence is necessary neither.

By virtue of (4), $X_{t_n^\epsilon}^\epsilon$ has a finite q-th moment if $X_{t_{n-1}^\epsilon}^\epsilon$ has. It follows by induction that the solution $X^\epsilon = (X_{t_n^\epsilon}^\epsilon : n \in \mathbb{N}_0)$ of (12) is q-integrable, and hence square-integrable. Equation (12) is obviously equivalent to the stochastic sum equation

$$X_{t_n^\epsilon}^\epsilon = x_\epsilon + \sum_{i=0}^{n-1} b(t_i^\epsilon, X_{t_i^\epsilon}^\epsilon)\,\epsilon + \sum_{i=0}^{n-1} a(t_i^\epsilon, X_{t_i^\epsilon}^\epsilon)\,V_{t_{i+1}^\epsilon}^\epsilon. \tag{13}$$

Suppose (ϵ_α) is an arbitrary sequence with $\epsilon_\alpha \in (0,1]$ and $\epsilon_\alpha \to 0$, set $x_\alpha = x_{\epsilon_\alpha}$ and recall our convention $\mathbb{E}_\alpha = \mathbb{E}_{\epsilon_\alpha}$, $X^\alpha \equiv X^{\epsilon_\alpha}$, $t_n^\alpha = t_n^{\epsilon_\alpha}$. Then it is easy to see that M^α defined in (8) provides a (mean-zero) square-integrable $(\mathcal{F}_t^{X^\alpha})$-martingale. Moreover $M_{t_n^\alpha}^\alpha$ coincides with the second sum on the right-hand side of (13). Therefore we also obtain

$$
\begin{aligned}
\langle M^\alpha \rangle_{t_n^\alpha} &= \sum_{i=1}^{n} \mathbb{E}_\alpha \left[\left(a(t_{i-1}^\alpha, X_{t_{i-1}^\alpha}^\alpha) \, V_{t_i^\alpha}^{\epsilon_\alpha} \right)^2 \Big| \mathcal{F}_{t_{n-1}^\alpha}^{X^\alpha} \right] \\
&= \sum_{i=0}^{n-1} a^2(t_i^\alpha, X_{t_i^\alpha}^\alpha) \, \mathbb{E}_\alpha \left[(V_{t_{i+1}^\alpha}^{\epsilon_\alpha})^2 \right] = \sum_{i=0}^{n-1} a^2(t_i^\alpha, X_{t_i^\alpha}^\alpha) \, \epsilon_\alpha
\end{aligned}
\tag{14}
$$

which shows that X^α solves the $(\epsilon_\alpha, a, b, x_\alpha)$-martingale problem of Definition 2.1. For an application of Theorem 2.2 it thus remains to show (10). But (10) follows from

$$
\begin{aligned}
&\mathbb{E}_\alpha \left[|X_{t_n^\alpha}^\alpha - X_{t_{n-1}^\alpha}^\alpha|^q \right] \\
&\leq 2^{q-1} \left\{ \mathbb{E}_\alpha \left[|b(t_{n-1}^\alpha, X_{t_{n-1}^\alpha}^\alpha)\epsilon_\alpha|^q \right] + \mathbb{E}_\alpha \left[|a(t_{n-1}^\alpha, X_{t_{n-1}^\alpha}^\alpha)|^q \right] \mathbb{E}_\alpha \left[|V_{t_n^\alpha}^{\epsilon_\alpha}|^q \right] \right\} \\
&\leq 2^{q-1} \left\{ K 2^{q-1} \left(1 + \mathbb{E}_\alpha[|X_{t_{n-1}^\alpha}^\alpha|^q] \right) \epsilon_\alpha^q + K 2^{q-1} \left(1 + \mathbb{E}_\alpha[|X_{t_{n-1}^\alpha}^\alpha|^q] \right) C \epsilon_\alpha^{q/2} \right\}
\end{aligned}
\tag{15}
$$

for which we used (12), the independence of $X_{t_{n-1}^\alpha}^\alpha$ of V^{ϵ_α}, (4) and $\mathbb{E}_\alpha[|V_{t_n^\alpha}^{\epsilon_\alpha}|^q] \leq C\epsilon_\alpha^{q/2}$. Hence Theorem 2.2 ensures that X^α converges in distribution to the unique solution of SDE (3).

As mentioned above, the independence of the random variables $\{V_{t_n^\epsilon}^\epsilon : n \in \mathbb{N}\}$ is not necessary. The independence was used for (14), (15) and the martingale property of M^α. But these relations may be valid even if the $V_{t_n^\epsilon}^\epsilon$ are not independent. For instance, let $\{\xi_n(i) : n, i \in \mathbb{N}\}$ be an array of independent centered random variables with variance 1 and q-th moments being bounded above by some $C > 0$ uniformly in n, i, for some $q > 2$. Then the martingale property of M^α and the main statements of (14) and (15) remain true for $V_{t_1^\epsilon}^\epsilon = \sqrt{\epsilon} \, \xi_1(1)$ and $V_{t_n^\epsilon}^\epsilon = \sqrt{\epsilon} \, \xi_n(f_n(V_{t_1^\epsilon}^\epsilon, \ldots, V_{t_{n-1}^\epsilon}^\epsilon))$, $n \geq 2$, where f_n is any measurable mapping from $\mathbb{R}^{n-1}$ to $\mathbb{N}$. This follows from the following relations which can be shown easily with help of the functional representation theorem for conditional expectations respectively by conditioning:

$$
\mathbb{E}_\alpha \left[V_{t_n^\alpha}^{\epsilon_\alpha} | \mathcal{F}_{t_{n-1}^\alpha}^{X^\alpha} \right] = 0, \qquad \mathbb{E}_\alpha \left[(V_{t_{i+1}^\alpha}^{\epsilon_\alpha})^2 | \mathcal{F}_{t_{n-1}^\alpha}^{X^\alpha} \right] = \epsilon_\alpha \qquad (1 \leq i \leq n-1)
$$

and

$$
\mathbb{E}_\alpha \left[|a(t_{n-1}^\alpha, X_{t_{n-1}^\alpha}^\alpha) V_{t_n^\alpha}^{\epsilon_\alpha}|^q \right] \leq C\epsilon_\alpha^{q/2}.
$$

If the $\xi_n(i)$ are not identically distributed, then the $V_{t_n^\epsilon}^\epsilon$ are typically not independent. In particular, the approximating process X^ϵ may be non-Markovian.

5 Proof of Theorem 2.2

Theorem 2.2 is an immediate consequence of Propositions 5.2, 5.5 and the weak equivalence of the martingale problem to the SDE. For the proofs of the two propositions we note that there exist $K' > 0$ and $\alpha_0 \geq 1$ such that for all $\alpha \geq \alpha_0$, $t \geq 0$ and $x \in \mathbb{R}$,

$$
|a_\alpha(t, x)| + |b_\alpha(t, x)| \leq K'(1 + |x|).
\tag{16}
$$

This is true since we assumed (4) and uniform convergence of a_α and b_α to the coefficients a and b, respectively. Throughout this section we will frequently use the well-known inequality $|\sum_{i=1}^{m} y_i|^p \leq m^{p-1} \sum_{i=1}^{m} |y_i|^p$ for all $m \in \mathbb{N}$, $p \geq 1$ and $y_1, \ldots, y_m \in \mathbb{R}$. As a first consequence of (16) we obtain Lemma 5.1. For every $x \in \mathbb{R}_+$ we write $\lfloor x \rfloor_\epsilon$ for the largest element of $\epsilon \mathbb{N}_0 = \{0, \epsilon, 2\epsilon, \ldots\}$ which is smaller than or equal to x. Moreover we assume without loss of generality $\epsilon_\alpha \leq 1$.

Lemma 5.1 *For $q > 2$ and $\delta > 1$ satisfying (10) and every $T > 0$,*

$$\sup_{\alpha \geq \alpha_0} \mathbb{E}_\alpha \left[\sup_{t \leq T} |X_t^\alpha|^q \right] < \infty. \tag{17}$$

Proof First of all note that for the proof it actually suffices to require $q \geq 2$ and $\delta \geq 1$. Set $S = \sup_{\alpha \geq \alpha_0} |x_\alpha|^q$ and $S_t^\alpha = \mathbb{E}_\alpha[\max_{1 \leq i \leq n_\alpha(t)} |M_{t_i}^\alpha - M_{t_{i-1}}^\alpha|^q]$. Using Proposition A.1 and (16) we obtain for all $t > 0$ and $\alpha \geq \alpha_0$,

$$\mathbb{E}_\alpha \left[\sup_{i \leq n_\alpha(t)} |X_{t_i}^\alpha|^q \right] \tag{18}$$

$$\leq 3^{q-1} \left\{ \mathbb{E}_\alpha \left[\sup_{i \leq n_\alpha(t)} |M_{t_i}^\alpha|^q \right] + S + \mathbb{E}_\alpha \left[\left(\sum_{i=0}^{n_\alpha(t)-1} |b_\alpha(t_i^\alpha, X_{t_i}^\alpha)|\epsilon_\alpha \right)^q \right] \right\}$$

$$\leq 3^{q-1} C_q \left\{ \mathbb{E}_\alpha \left[\left| \sum_{i=0}^{n_\alpha(t)-1} a_\alpha^2(t_i^\alpha, X_{t_i}^\alpha)\epsilon_\alpha \right|^{\frac{q}{2}} \right] + S_t^\alpha + S + \mathbb{E}_\alpha \left[\left[\sum_{i=0}^{n_\alpha(t)-1} |b_\alpha(t_i^\alpha, X_{t_i}^\alpha)|\epsilon_\alpha \right]^q \right] \right\}$$

$$\leq k_q \left\{ \mathbb{E}_\alpha \left[\left[\sum_{i=0}^{n_\alpha(t)-1} (K'(1 + |X_{t_i}^\alpha|))^2 \epsilon_\alpha \right]^{\frac{q}{2}} \right] + S_t^\alpha + S + \mathbb{E}_\alpha \left[\left[\sum_{i=0}^{n_\alpha(t)-1} K'(1 + |X_{t_i}^\alpha|)\epsilon_\alpha \right]^q \right] \right\}$$

where C_q is independent of t and α, and $k_q = 3^{q-1} C_q$. By Hölder's inequality we get

$$\mathbb{E}_\alpha \left[\left[\sum_{i=0}^{n_\alpha(t)-1} (K'(1 + |X_{t_i}^\alpha|))^2 \epsilon_\alpha \right]^{q/2} \right]$$

$$\leq \mathbb{E}_\alpha \left[\left(\sum_{i=0}^{n_\alpha(t)-1} (2K'^2(1 + |X_{t_i}^\alpha|^2))^{q/2} \right) \left(\sum_{i=0}^{n_\alpha(t)-1} \epsilon_\alpha^{(q/2)/(q/2-1)} \right)^{q/2-1} \right]$$

$$\leq \mathbb{E}_\alpha \left[\left(\sum_{i=0}^{n_\alpha(t)-1} 2^{q/2-1} (2K'^2)^{q/2} (1 + |X_{t_i}^\alpha|^q) \right) n_\alpha(t)^{q/2-1} \epsilon_\alpha^{q/2} \right]$$

$$\leq c_q t^{q/2} + c_q t^{q/2-1} \sum_{i=0}^{n_\alpha(t)-1} \mathbb{E}_\alpha \left[\sup_{j \leq i} |X_{t_j}^\alpha|^q \right] \epsilon_\alpha$$

where $c_q = 2^{q/2-1} (2K'^2)^{q/2}$. Analogously, with $\bar{c}_q = 2^{q-1} K'^q$,

$$\mathbb{E}_\alpha \left[\left[\sum_{i=0}^{n_\alpha(t)-1} K'(1 + |X_{t_i}^\alpha|)\epsilon_\alpha \right]^q \right] \leq \bar{c}_q t^q + c_q t^{q-1} \sum_{i=0}^{n_\alpha(t)-1} \mathbb{E}_\alpha \left[\sup_{j \leq i} |X_{t_j}^\alpha|^q \right] \epsilon_\alpha.$$

8

Moreover, by (10) and (16) we obtain for all $t \leq T$ and $\alpha \geq \alpha_0$,

$$
\begin{aligned}
S_t^\alpha &\leq \sum_{i=1}^{n_\alpha(t)} \mathbb{E}_\alpha \Big[|M_{t_i^\alpha}^\alpha - M_{t_{i-1}^\alpha}^\alpha|^q \Big] \\
&\leq 2^{q-1} \sum_{i=1}^{n_\alpha(t)} \mathbb{E}_\alpha \Big[|X_{t_i^\alpha}^\alpha - X_{t_{i-1}^\alpha}^\alpha|^q + |b_\alpha(t_{i-1}^\alpha, X_{t_{i-1}^\alpha}^\alpha)|^q \epsilon_\alpha^q \Big] \\
&\leq 2^{q-1} \sum_{i=0}^{n_\alpha(t)-1} \left\{ C_T \Big(1 + \mathbb{E}_\alpha \Big[|X_{t_i^\alpha}^\alpha|^q \Big] \Big) \epsilon_\alpha^\delta + \mathbb{E}_\alpha \Big[K'(1 + |X_{t_i^\alpha}^\alpha|^q) \Big] \epsilon_\alpha^q \right\} \\
&\leq c_{q,T} \, t \, + \, c_{q,T} \sum_{i=0}^{n_\alpha(t)-1} \mathbb{E}_\alpha \Big[\sup_{j \leq i} |X_{t_j^\alpha}^\alpha|^q \Big] \epsilon_\alpha
\end{aligned}
$$

where $c_{q,T} = 2^{q-1}(C_T + K')$. By all account we have for all $t \leq T$ and $\alpha \geq \alpha_0$,

$$
\begin{aligned}
\mathbb{E}_\alpha \Big[\sup_{i \leq n_\alpha(t)} |X_{t_i^\alpha}^\alpha|^q \Big] &\leq k_q \left\{ S + (c_q + \bar{c}_q + c_{q,T})(t^{q-1} \vee 1) \Big(1 + \epsilon_\alpha \sum_{i=0}^{n_\alpha(t)-1} \mathbb{E}_\alpha \Big[\sup_{j \leq i} |X_{t_j^\alpha}^\alpha|^q \Big] \Big) \right\} \\
&\leq (k_q S + C_{q,T}) + C_{q,T} \, \epsilon_\alpha \sum_{i=0}^{n_\alpha(t)-1} \mathbb{E}_\alpha \Big[\sup_{j \leq i} |X_{t_j^\alpha}^\alpha|^q \Big]
\end{aligned}
$$

where $C_{q,T} = k_q(c_q + \bar{c}_q + c_{q,T})(T^{q-1} \vee 1)$. An application of Lemma A.2 yields

$$
\mathbb{E}_\alpha \Big[\sup_{s \leq t} |X_s^\alpha|^q \Big] = \mathbb{E}_\alpha \Big[\sup_{i \leq n_\alpha(t)} |X_{t_i}^\alpha|^q \Big] \leq (k_q S + C_{q,T})(1 + C_{q,T} \, \epsilon_\alpha)^{n_\alpha(t)} + (C_{q,T} \, \epsilon_\alpha)^{n_\alpha(t)} S, \quad (19)
$$

where we emphasize that the constants k_q, S and $C_{q,T}$ are independent of $t \leq T$ and $\alpha \geq \alpha_0$. This proves Lemma 5.1 since $\limsup_{\alpha \to \infty}(1 + C_{q,T}\epsilon_\alpha)^{n_\alpha(t)}$ is bounded by $\exp(tC_{q,T})$ (note that $n_\alpha(t) = \lfloor t/\epsilon_\alpha \rfloor_1 \leq t/\epsilon_\alpha$). $\qquad \square$

Proposition 5.2 *If $(\mathbb{P}_\alpha)$ is tight then the coordinate process of any weak limit point, that has no mass outside of $C(\mathbb{R})$, is a solution of the (a, b, x_0)-martingale problem of Definition 1.1.*

Proof We consider a weakly convergent subsequence whose limit, $\mathbb{P}$, has no mass outside of $C(\mathbb{R})$. By an abuse of notation, we denote this subsequence by $(\mathbb{P}_\alpha)$ either. We further write X for the coordinate process of $\mathbb{P}$. Since X is $\mathbb{P}$-almost surely continuous, we know ([3] Theorem 3.7.8) that

$$
\mathbb{P}_\alpha \circ \pi_{t_1,\ldots,t_k}^{-1} \Rightarrow \mathbb{P} \circ \pi_{t_1,\ldots,t_k}^{-1} \quad (20)
$$

for all $t_1, \ldots, t_k \in \mathbb{R}_+$, where $\pi_{t_1,\ldots,t_k} : D(\mathbb{R}) \to \mathbb{R}^k$ is the usual coordinate projection. In the remainder of the proof we will show in three steps that M defined in (6) is square-integrable, provides an $(\bar{\mathcal{F}}_t^X)$-martingale and has $\langle M \rangle$ defined in (7) as compensator. Here $(\bar{\mathcal{F}}_t^X)$ denotes the natural augmentation of the filtration $(\mathcal{F}_t^X)$ induced by X.

Step 1. With help of Fatou's lemma as well as (20) and (17) we obtain for every $T > 0$,

$$\sup_{t \leq T} \mathbb{E}\left[|X_t|^q\right] \leq \sup_{t \leq T} \liminf_{N \to \infty} \lim_{\alpha \to \infty} \mathbb{E}_\alpha[|X_t^\alpha|^q \wedge N] \leq \sup_{t \leq T} \sup_{\alpha \geq \alpha_0} \mathbb{E}_\alpha[|X_t^\alpha|^q] < \infty. \tag{21}$$

Taking (4) into account we conclude that M defined in (6) is square-integrable.

Step 2. We next show that M is an $(\bar{\mathcal{F}}_t^X)$-martingale. It suffices to show that M is an $(\mathcal{F}_t^X)$-martingale, see [2] p. 75. The latter is true if and only if

$$\mathbb{E}\left[\left(X_{t+s} - X_t - \int_t^{t+s} b(r, X_r)dr\right) \prod_{i=1}^l h_i(X_{t_i})\right] = 0 \tag{22}$$

holds for all $0 \leq t_1 < \cdots \leq t_l \leq t$, $s \geq 0$, $l \geq 1$ and bounded $h_1, \ldots, h_l \in C(\mathbb{R})$ (do not confuse t_i and t_i^α). Since X^α solves the $(\epsilon_\alpha, a_\alpha, b_\alpha, x_\alpha)$-martingale problem, we have

$$\mathbb{E}_\alpha\left[\left(X_{t+s}^\alpha - X_t^\alpha - \sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^\alpha)\epsilon_\alpha\right) \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right] = 0. \tag{23}$$

We are going to verify (22) by showing that the left-hand side of (23) converges to the left-hand side of (22) as $\alpha \to \infty$. We begin with proving

$$\lim_{\alpha \to \infty} \mathbb{E}_\alpha\left[X_u^\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right] = \mathbb{E}\left[X_u \prod_{i=1}^l h_i(X_{t_i})\right] \tag{24}$$

for every $u \geq 0$, which together with (28) below implies the required convergence. To this end we set $x^{(N)} = (-N \vee x) \wedge N$ for all $x \in \mathbb{R}$ and $N > 0$. The right-hand side of

$$\left|\mathbb{E}_\alpha\left[X_u^{\alpha,(N)} \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right] - \mathbb{E}_\alpha\left[X_u^\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right]\right| \leq \mathbb{E}_\alpha\left[\left|X_u^{\alpha,(N)} - X_u^\alpha\right| \prod_{i=1}^l \|h_i\|_\infty\right]$$

can be estimated, for every $T \geq u$, by

$$\sup_{r \leq T} \sup_{\alpha' \geq \alpha_0} \mathbb{E}_{\alpha'}\left[|X_r^{\alpha'}| \mathbb{1}_{|X_r^{\alpha'}| > N}\right] \prod_{i=1}^l \|h_i\|_\infty$$

which tends to 0 as $N \to \infty$ since $\{X_r^{\alpha'} : r \leq T, \ \alpha' \geq 1\}$ is uniformly integrable by (17). Therefore we have

$$\lim_{N \to \infty} \mathbb{E}_\alpha\left[X_u^{\alpha,(N)} \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right] = \mathbb{E}_\alpha\left[X_u^\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right] \quad \text{uniformly in } \alpha \geq \alpha_0 \tag{25}$$

(and uniformly in $u \leq T$, for every $T > 0$). By (20) we further obtain for every $N > 0$,

$$\lim_{\alpha \to \infty} \mathbb{E}_\alpha\left[X_u^{\alpha,(N)} \prod_{i=1}^l h_i(X_{t_i}^\alpha)\right] = \mathbb{E}\left[X_u^{(N)} \prod_{i=1}^l h_i(X_{t_i})\right] \tag{26}$$

10

since the mapping $(x_1, \ldots, x_{l+1}) \mapsto x_{l+1}^{(N)} \prod_{i=1}^l h_i(x_i)$ from $\mathbb{R}^{l+1}$ to $\mathbb{R}$ is bounded and continuous. This is the reason why we introduced the truncation $x^{(N)}$. By virtue of (21), an application of the dominated convergence theorem gives

$$\lim_{N \to \infty} \mathbb{E}\left[X_u^{(N)} \prod_{i=1}^l h_i(X_{t_i}) \right] = \mathbb{E}\left[X_u \prod_{i=1}^l h_i(X_{t_i}) \right] \tag{27}$$

which along with (25) and (26) implies (24). It remains to show

$$\lim_{\alpha \to \infty} \mathbb{E}_\alpha\left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^\alpha)\epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] = \mathbb{E}\left[\int_t^{t+s} b(r, X_r)dr \prod_{i=1}^l h_i(X_{t_i}) \right]. \tag{28}$$

Taking (16) and $(n_\alpha(t + s) - n_\alpha(t))\epsilon_\alpha \leq s + \epsilon_\alpha$ into account we obtain analogously to (25) and (27),

$$\lim_{N \to \infty} \mathbb{E}_\alpha\left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^{\alpha,(N)})\epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] \tag{29}$$

$$= \mathbb{E}_\alpha\left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^\alpha)\epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] \quad \text{uniformly in } \alpha \geq \alpha_0$$

respectively

$$\lim_{N \to \infty} \mathbb{E}\left[\int_t^{t+s} b(r, X_r^{(N)})dr \prod_{i=1}^l h_i(X_{t_i}) \right] = \mathbb{E}\left[\int_t^{t+s} b(r, X_r)dr \prod_{i=1}^l h_i(X_{t_i}) \right]. \tag{30}$$

By the uniform convergence of b_α to b and $(n_\alpha(t + s) - n_\alpha(t))\epsilon_\alpha \leq s + \epsilon_\alpha$, we also have

$$\mathbb{E}_\alpha\left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^{\alpha,(N)})\epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] = \mathbb{E}_\alpha\left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b(t_i^\alpha, X_{t_i^\alpha}^{\alpha,(N)})\epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] + o_\alpha(1). \tag{31}$$

Moreover we have

$$\mathbb{E}_\alpha\left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b(t_i^\alpha, X_{t_i^\alpha}^{\alpha,(N)})\epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] = \mathbb{E}_\alpha\left[\int_t^{t+s} b(r, X_r^{\alpha,(N)})dr \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] + o_\alpha(1) \tag{32}$$

which is a consequence of the dominated convergence theorem and

$$\left| \sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b(t_i^\alpha, X_{t_i^\alpha}^{\alpha,(N)})\epsilon_\alpha - \int_t^{t+s} b(r, X_r^{\alpha,(N)})dr \right|$$

$$\leq \int_{\lfloor t \rfloor_{\epsilon_\alpha}}^{\lfloor t+s \rfloor_{\epsilon_\alpha} - \epsilon_\alpha} \left| b(\lfloor r \rfloor_{\epsilon_\alpha}, X_r^{\alpha,(N)}) - b(r, X_r^{\alpha,(N)}) \right| dr + o_\alpha(1)$$

together with the fact that b is bounded and uniformly continuous on $[0, t+s] \times [-N, N]$. Finally we get by (20) and the dominated convergence theorem and (17),

$$
\lim_{\alpha \to \infty} \mathbb{E}_\alpha \left[\int_t^{t+s} b(r, X_r^{\alpha,(N)}) dr \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] = \int_t^{t+s} \lim_{\alpha \to \infty} \mathbb{E}_\alpha \left[b(r, X_r^{\alpha,(N)}) \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] dr
$$

$$
= \int_t^{t+s} \mathbb{E} \left[b(r, X_r^{(N)}) \prod_{i=1}^l h_i(X_{t_i}) \right] dr = \mathbb{E} \left[\int_t^{t+s} b(r, X_r^{(N)}) dr \prod_{i=1}^l h_i(X_{t_i}) \right] \tag{33}
$$

which along with (31) and (32) implies

$$
\lim_{\alpha \to \infty} \mathbb{E}_\alpha \left[\sum_{i=n_\alpha(t)}^{n_\alpha(t+s)-1} b_\alpha(t_i^\alpha, X_{t_i^\alpha}^{\alpha,(N)}) \epsilon_\alpha \prod_{i=1}^l h_i(X_{t_i}^\alpha) \right] = \mathbb{E} \left[\int_t^{t+s} b(r, X_r^{(N)}) dr \prod_{i=1}^l h_i(X_{t_i}) \right].
$$

This, (29) and (30) ensure (28).

`Step 3`. It remains to show (7). By the uniqueness of the Doob-Meyer decomposition, M has the required compensator if and only if

$$
\mathbb{E} \left[\left(M_{t+s}^2 - M_t^2 - \int_t^{t+s} a^2(r, X_r) dr \right) \prod_{i=1}^l h_i(X_{t_i}) \right] = 0 \tag{34}
$$

holds for all $0 \le t_1 < \cdots \le t_l \le t$, $s \ge 0$, $l \ge 1$ and bounded $h_1, \ldots, h_l \in C(\mathbb{R})$. Now, the discrete analogue of equation (34) for $\mathbb{E}_\alpha$, a_α and X^α holds. Proceeding similarly to the proof of (22) one can show that the left-hand side of this equation converges to the left-hand side of (34) as $\alpha \to \infty$. Therefore we obtain (34). For the sake of brevity we omit the details. It should be mentioned, however, that we now need uniform integrability of $\{(X_r^\alpha)^2 : r \le t+s, \ \alpha \ge 1\}$. This is why we established (17) for q being strictly larger than 2. □

The assumptions of Proposition 5.2 can be checked with help of the following two lemmas, where $\mathbb{Q}_\alpha$ and $\mathbb{Q}$ refer to any laws on $D(\mathbb{R})$, and Y^α and Y are the respective coordinate processes. By an abuse of notation, we denote the corresponding expectations by $\mathbb{Q}_\alpha$ and $\mathbb{Q}$ either. The first lemma follows from [3] Theorem 3.8.8 and [3] Theorem 3.8.6 (b)⇒(a) along with Prohorov's theorem. Lemma 5.4 is more or less standard and can be proved with help of the continuity criterion 3.10.3 in [3]; we omit the details.

Lemma 5.3 *Assume (Y_t^α) is tight in $\mathbb{R}$ for every rational $t \ge 0$. Let $m > 0$, $\gamma > 1$ and assume for every $T > 0$ there is some finite constant $C_T > 0$ such that for all $\alpha \ge 1$ and $t, h \ge 0$ with $0 \le t - h$ and $t + h \le T$,*

$$
\mathbb{Q}_\alpha \left[|Y_{t-h}^\alpha - Y_t^\alpha|^{m/2} |Y_t^\alpha - Y_{t+h}^\alpha|^{m/2} \right] \le C_T \, h^\gamma. \tag{35}
$$

Then $(\mathbb{Q}_\alpha)$ is tight.

Lemma 5.4 *Let $m > 0$, $\gamma > 1$ and assume for every $T > 0$ there is some finite constant $C_T > 0$ such that for all $\alpha \geq 1$ and $0 \leq t' \leq t'' \leq T$,*

$$\limsup_{\alpha \to \infty} \mathbb{Q}_\alpha \left[|Y_{t'}^\alpha - Y_{t''}^\alpha|^m \right] \leq C_T (t'' - t')^\gamma. \tag{36}$$

Then if $\mathbb{Q}_\alpha \Rightarrow \mathbb{Q}$, the limit $\mathbb{Q}$ has no mass outside of $C(\mathbb{R})$.

Proposition 5.5 *$(\mathbb{P}_\alpha)$ is tight and each limit point has no mass outside of $C(\mathbb{R})$.*

Proof Let $q > 2$ and $\delta > 1$ satisfy (10). Using techniques as in the proof of Lemma 5.1 we can find a finite constant $C_q > 0$ such that for every $0 \leq t' \leq t''$ and $\alpha \geq \alpha_0$,

$$\mathbb{E}_\alpha[|X_{t'}^\alpha - X_{t''}^\alpha|^q]$$

$$\leq C_q \Bigg\{ \epsilon_\alpha^{q/2} \, \mathbb{E}_\alpha \Bigg[\Big(\sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} (1 + |X_{t_i^\alpha}^\alpha|)^2 \Big)^{q/2} \Bigg] + \epsilon_\alpha^q \, \mathbb{E}_\alpha \Bigg[\Big(\sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} (1 + |X_{t_i^\alpha}^\alpha|) \Big)^q \Bigg]$$

$$+ \sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} \Big(1 + \mathbb{E}_\alpha \left[|X_{t_i^\alpha}^\alpha|^q \right] \Big) \epsilon_\alpha^{\delta \wedge q} \Bigg\}.$$

Applying Hölder's inequality to each of the first two summands on the right-hand side, using (17) and setting $\gamma = (q/2) \wedge \delta$ we may continue with

$$\leq C_q \Bigg\{ \epsilon_\alpha^{q/2} \Big(\sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} \mathbb{E}_\alpha \left[(1 + |X_{t_i^\alpha}^\alpha|)^{2q/2} \right] \Big) \Big(\sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} 1^{q/(q-2)} \Big)^{q/2-1} +$$

$$\epsilon_\alpha^q \Big(\sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} \mathbb{E}_\alpha \left[(1 + |X_{t_i^\alpha}^\alpha|)^q \right] \Big) \Big(\sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} 1^{q/(q-1)} \Big)^{q-1} + \sum_{i=n_\alpha(t')}^{n_\alpha(t'')-1} \Big(1 + \mathbb{E}_\alpha \left[|X_{t_i^\alpha}^\alpha|^q \right] \Big) \epsilon_\alpha^{\delta \wedge q} \Bigg\}$$

$$\leq C_{q,T} \Big\{ \epsilon_\alpha^{q/2} \left(n_\alpha(t'') - n_\alpha(t') \right)^{q/2} + \epsilon_\alpha^q \left(n_\alpha(t'') - n_\alpha(t') \right)^q + \epsilon_\alpha^{\delta \wedge q} \left(n_\alpha(t'') - n_\alpha(t') \right) \Big\}$$

$$\leq \bar{C}_{q,T} \left(\epsilon_\alpha \left(n_\alpha(t'') - n_\alpha(t') \right) \right)^\gamma$$

$$\leq \bar{C}_{q,T} \left((t'' - t') + \epsilon_\alpha \right)^\gamma \tag{37}$$

where $C_{q,T}, \bar{C}_{q,T} > 0$ are some finite constants being independent of $t', t'' \leq T$ and $\alpha \geq \alpha_0$. Then Lemma 5.4 ensures that any weak limit point of $(\mathbb{P}_\alpha)$ has no mass outside of $C(\mathbb{R})$. At this point it is essential that we required q and δ to be strictly larger than 2 respectively 1.

Toward the verification of tightness of $(\mathbb{P}_\alpha)$ we use Hölder's inequality to get

$$\mathbb{E}_\alpha \left[|X_{t-h}^\alpha - X_t^\alpha|^{q/2} |X_t^\alpha - X_{t+h}^\alpha|^{q/2} \right] \leq \mathbb{E}_\alpha \left[|X_{t-h}^\alpha - X_t^\alpha|^q \right]^{1/2} \mathbb{E}_\alpha \left[|X_t^\alpha - X_{t+h}^\alpha|^q \right]^{1/2}. \tag{38}$$

If $h \geq \epsilon_\alpha/2$ then (37) implies that both factors on the right-hand side of (38) are bounded by $\bar{C}_{q,T}(3h)^{\gamma/2}$. If $h < \epsilon_\alpha/2$ then at least one of these factors vanishes since X^α is constant on intervals of length ϵ_α. Hence,

$$\mathbb{E}_\alpha \left[|X_{t-h}^\alpha - X_t^\alpha|^{q/2} |X_t^\alpha - X_{t+h}^\alpha|^{q/2} \right] \leq (\bar{C}_{q,T}^2 \, 3^\gamma) \, h^\gamma$$

for all $\alpha \geq \alpha_0$ and $t, h \geq 0$ with $t + h \leq T$. That is, (35) holds with $m = q$. Therefore Lemma 5.3 ensures tightness of $(\mathbb{P}_\alpha)$. $\square$

13

A Appendix

Here we give two auxiliary results. We first recall a square function inequality for martingales. Let $M = (M_n : n \in \mathbb{N}_0)$ be an $(\mathcal{F}_n)$-martingale on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The corresponding compensator is given by $\langle M \rangle_n = \sum_{i=1}^n \mathbb{E}[(M_i - M_{i-1})^2 | \mathcal{F}_{i-1}]$.

Proposition A.1 *([6] Theorem 2.11) For every $q > 0$ there is some finite constant $C_q > 0$ depending only on q such that*

$$\mathbb{E}\left[\max_{1 \leq i \leq n} |M_i|^q\right] \leq C_q \left(\mathbb{E}\left[\langle M \rangle_n^{q/2}\right] + \mathbb{E}\left[\max_{1 \leq i \leq n} |M_i - M_{i-1}|^q\right]\right).$$

The second result is a Gronwall lemma for functions with discrete domain. It can be proven by means of iterating (39) n-times. We omit the proof since it is more or less well-known.

Lemma A.2 *Suppose g is a mapping from $\mathbb{N}_0$ to $\bar{\mathbb{R}}_+ = [0, \infty]$ with $g(0) < \infty$. If there are finite constants $c_0, c_1 \geq 0$ such that*

$$g(n) \leq c_0 + c_1 \sum_{i=0}^{n-1} g(i) \qquad \forall \, n \leq N, \tag{39}$$

then

$$g(n) \leq c_0(1 + c_1)^n + c_1^n g(0) < \infty \qquad \forall \, n \leq N.$$

References

[1] Cherny, A.S. and Engelbert, H.-J. (2005) Singular Stochastic Differential Equations, *Lecture Notes in Mathematics*, vol. 1858, Springer.

[2] Dellacherie, C. and Meyer, P.-A. (1980) *Probabilités et potentiel. Chapitres V à VIII.* Hermann, Paris. Théorie des martingales.

[3] Ethier, S.N. and Kurtz, T.G. (1986) *Markov processes: Characterization and convergence.* Wiley, New York.

[4] Feller, W. (1951) Diffusion processes in genetics. In *Proc. Second Berkely Symp. Math. Statist. Probab.* Univ. of California Press, Berkely.

[5] Fujimagari, T. (1976) Controlled Galton-Watson process and its asymptotic behavior. *Koda Math. Sem. Rep.*, 27:11-18.

[6] Hall, P. and Heyde, C.C. (1980) *Martingale Limit Theory and Its Application.* Academic Press, New York.

[7] Harris, T.E. (1963) *The theory of branching processes.* Springer-Verlag, Berlin.

[8] Höpfner, R. (1985) On some class of population-size-dependent Galton-Watson process. *J. Appl. Prob.*, 22:25–36.

[9] Karatzas, I. and Shreve, S.E. (1991) *Brownian motion and stochastic calculus*. Springer-Verlag, New York.

[10] Klebaner, F.C. (1984) On population-size-dependent branching processes. *Adv. Appl. Prob.*, 16:30–55.

[11] Klebaner, F.C. (1984) Geometric rate of growth in population-size-dependent branching processes. *J. Appl. Prob.*, 21:40–49.

[12] Kloeden, P.E. and Platen, E. (1992) *Numerical solutions of stochastic differential equations*. Springer-Verlag, Berlin.

[13] Kurtz, T.G. and Protter, P. (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Prob.*, 19(3):1035–1070.

[14] Lindvall, T. (1972) Convergence of critical Galton-Watson branching processes. *J. Appl. Prob.*,9:445–450.

[15] Lorenz, R. (2006) *Weak approximation of stochastic delay differential equations with bounded memory by discrete time series*. PhD thesis, Humbold University of Berlin.

[16] Méléard, S. and Roelly, S. (1993) Interacting measure branching processes. Some bounds for the support. *Stochastics and Stochastics Reports*, 44:103–121.

[17] Roelly-Coppoletta, S. (1986) A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17:43–65.

[18] Stroock, D.W. and Varadhan S.R.S. (1979) *Multidimensional diffusion processes*. Springer-Verlag, Berlin.

[19] Sturm, A. (2003) On convergence of population processes in random environments to the stochastic heat equation with colored noise. *Electronic Journal of Probability*, 8(6):1–39.

[20] Yamada, T. and Watanabe, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167.

[21] Zähle, H. (2008) Approximation of SDEs by population-size-dependent Galton-Watson processes. *Preprint*.