# Approximation of SDEs by population-size-dependent Galton-Watson processes 

Henryk Zähle

Fakultät für Mathematik

Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund

# Approximation of SDEs by population-size-dependent Galton-Watson processes 

Henryk Zähle*


#### Abstract

A certain class of stochastic differential equations, containing the Cox-Ingersoll-Ross model and the geometric Brownian motion, is considered. The corresponding solutions are approximated weakly by discrete-time population-size-dependent Galton-Watson processes with immigration. The long-time behavior of the limiting processes is also investigated.


Key words and phrases. Stochastic differential equation, Galton-Watson process, population-size-dependent branching, weak convergence, martingale problem, Doob-Meyer decomposition, Cox-Ingersoll-Ross model

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

[^0]
## 1 Introduction

A critical Galton-Watson process (GWP) with finite offspring variance $\sigma^{2}>0$ and initial state $Z_{0}$ is an $\mathbb{N}_{0}$-valued Markov process $Z=\left(Z_{n}: n \in \mathbb{N}_{0}\right)$ that can be defined recursively as follows:

$$
\begin{equation*}
Z_{n}=\sum_{i=1}^{Z_{n-1}} N_{n-1, i} \quad(n \in \mathbb{N}) \tag{1}
\end{equation*}
$$

where $\left\{N_{n, i}: n \geq 0, i \geq 1\right\}$ is a family of i.i.d. $\mathbb{N}_{0}$-valued random variables with mean 1 and variance $\sigma^{2}$. The state $Z_{n}$ can be seen as the number of individuals of a certain population in the $n$-th generation; $N_{n, i}$ is the random number of direct descendants of individual $i$ of the $n$-th generation. Background on GWPs is given in $[1,2,3,11,12,23]$ and others.

It is known that $Z$, re-scaled by the factor $\epsilon$ in both time and space, converges weakly ( $\epsilon \downarrow$ 0 ) to a time-continuous homogeneous Markov process with continuous samples and Jiřina's branching property [15], see [9, 22, 24]. The limiting process is known as Feller's branching diffusion and it is the unique solution of the following stochastic differential equation (SDE)

$$
\begin{equation*}
d X_{t}=\sigma \sqrt{\left|X_{t}\right|} d W_{t} . \tag{2}
\end{equation*}
$$

Here, $W$ is a one-dimensional Brownian motion. The solution of (2) can be regarded as a total mass approximation of a "large" system of individuals having "small" individual mass and branching in "quick" succession. To get an intuition why the square-root is the right coefficient in (2), note that the variance of $X_{t}$ equals $\sigma^{2} X_{0} t$. This expression resembles the variance of $Z_{n}$ which is given by $\sigma^{2} Z_{0} n$.

In this article, we focus on the approximation of more general SDEs by (generalized) timediscrete GWPs. We consider the following equation

$$
\begin{equation*}
d X_{t}=\left(\delta(t)+\mu(t) X_{t}\right) d t+\sigma\left(t, X_{t}\right) \sqrt{\left|X_{t}\right|} d W_{t}, \quad X_{0}=y \geq 0 \tag{3}
\end{equation*}
$$

where $\delta: \mathbb{R}_{+} \rightarrow \mathbb{R}_{+}, \mu: \mathbb{R}_{+} \rightarrow \mathbb{R}$ and $\sigma: \mathbb{R}_{+} \times \mathbb{R} \rightarrow \mathbb{R}_{+}$are continuous functions. We assume that there is some constant $K>0$ satisfying

$$
\begin{gather*}
\left|\sigma^{2}(t, x)\right| \leq K(1+|x|) \quad \forall t \in \mathbb{R}_{+}, x \in \mathbb{R},  \tag{4}\\
\delta(t)+|\mu(t)|<K \quad \forall t \in \mathbb{R}_{+} . \tag{5}
\end{gather*}
$$

Moreover we assume that there exists a unique weak solution of (3), which is the case, e.g., if (14) holds for $a(t, x)=\sigma(t, x) \sqrt{|x|}$ (cf. Section 4). For instance, one can choose $\sigma(t, x)=\nu|x|^{\gamma}$ with $\nu \in \mathbb{R}_{+}$and $0 \leq \gamma \leq 1 / 2$. In particular, Feller's branching diffusion with drift, i.e., the Cox-Ingersoll-Ross model [5], and the geometric Brownian motion match the requirements. In Section 2 we present a weak approximation of SDE (3) by a re-scaled version of the GWP defined in (1) with additional immigration, where the branching mechanism is no longer critical and where the offspring variance may depend on the time and the current state of the process. To some extent, the approximating process belongs to the class of population-size-dependent Galton-Watson processes in the sense of $[10,13,17,18]$. Note that there are
several results on the approximation of SDEs by different types of GWPs, see [4, 7, 8, 24] and others. However, up to the authors best knowledge, the approximation in Section 2 has not been made rigorous so far. The proof of our approximation result avails a martingale method (given in [28]) similar to those of [25, 26]. The martingale method will be described in Section 4 , the proof itself will be carried out in Section 5. The method of [28] allows in particular a weakening of the usual assumption of a common offspring law. In Section 3 we add to the approximation result a discussion of the long-time behavior of the limiting process (3).

We emphasize that the approximation of (3) in Section 2 is time-discrete, i.e., the lifetime of an individual is fix (and not governed by an exponential clock as for instance in Section 9.4 of $[8]$ ). In the scope of numerical simulations of SDEs, practitioners typically prefer to work in discrete time so that the result is interesting for applications. In contrast to the classical Euler scheme (see, e.g., [19]), the presented approximation of SDE (3) has the advantage that also the approximating process is nonnegative for all time. In many situations this is an desirable property, for instance, in the context of the Cox-Ingersoll-Ross short term interest rate model [5]. Note that we have to modify the "conventional" re-scaling of GWPs [24] since we allow for an immigration $\delta \neq 0$. If $\delta=0$ then one could also consider the "conventional" re-scaling, cf. Remark 2.4.

We conclude the Introduction with some notation. We denote by $D(\mathbb{R})$ and $D\left(\mathbb{R}_{+}\right)$the spaces of cádlàg functions from $\mathbb{R}_{+}$to $\mathbb{R}$ respectively to $\mathbb{R}_{+}$. We equip these spaces with the topology generated by the Skohorod convergence on compacts and we consider them as measurable spaces with respect to their Borel $\sigma$-algebras. For every $\epsilon>0$ we set $\epsilon \mathbb{N}_{0}=\left\{\epsilon n: n \in \mathbb{N}_{0}\right\}$. For every $x \in \mathbb{R}_{+}$we let $\lfloor x\rfloor_{\epsilon}$ be the largest element of $\epsilon \mathbb{N}_{0}$ which is smaller than or equal to $x$. Finally we set $t_{n}^{\epsilon}=n \epsilon$ and $x_{k}^{\epsilon}=k \epsilon^{2}$ for every $n, k \in \mathbb{N}_{0}$.

## 2 Main result

Let $\delta, \mu$, and $\sigma$, be functions as discussed subsequent to (3). Moreover let $\bar{X}^{\epsilon}$ be an $\epsilon^{2} \mathbb{N}_{0}-$ valued Markov process with index set $\epsilon \mathbb{N}_{0}$, which is defined via the transition probabilities as follows. First, set $\bar{X}_{t_{0}^{6}}^{\epsilon}=\lfloor y\rfloor_{\epsilon^{2}}$. Then, for every $n \in \mathbb{N}$, the one-step transition probability of $\bar{X}^{\epsilon}$,

$$
p_{\epsilon}\left(t_{n-1}^{\epsilon}, t_{n}^{\epsilon} ; x_{k}^{\epsilon}, .\right)=\operatorname{prob}\left[\bar{X}_{t_{n}^{\epsilon}}^{\epsilon}=. \mid \bar{X}_{t_{n-1}^{\epsilon}}^{\epsilon}=x_{k}^{\epsilon}\right],
$$

is defined to be the law of the random variable

$$
\begin{equation*}
\epsilon\left[\delta\left(t_{n-1}^{\epsilon}\right)\right\rfloor_{\epsilon}+\sum_{i=1}^{\epsilon^{-2} x_{k}^{\epsilon}} \epsilon^{2} N_{n-1, i}^{\epsilon}\left(x_{k}^{\epsilon}\right) . \tag{6}
\end{equation*}
$$

Here, $\left\{N_{n, i}^{\epsilon}\left(x_{k}^{\epsilon}\right): i \in \mathbb{N}, n, k \in \mathbb{N}_{0}\right\}$ is any family of independent, 4-integrable, $\mathbb{N}_{0}$-valued random variables such that, for every $i, n \in \mathbb{N}$ and $k \in \mathbb{N}_{0}$, the random variable $N_{n-1, i}^{\epsilon}\left(x_{k}^{\epsilon}\right)$ has
(i) mean $1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right)$,
(ii) variance $\sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right) / \epsilon$,
(iii) fourth moment being bounded above by $C_{4}\left(1+\left(x_{k}^{\epsilon}\right)^{3}\right) / \epsilon^{\eta}$,
where $\sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right)=\sigma^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right)+\epsilon^{2} K$, and $\eta<5$ and $C_{4}>0$ are some constants being independent of $\epsilon, i, n$, and $k$. The constant $K$ was introduced subsequent to (3). We assume $\epsilon \leq 1 /(2 K)$, so that

$$
\begin{equation*}
1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right) \geq 1 / 2>0 \quad\left(n \in \mathbb{N}_{0}\right) \tag{7}
\end{equation*}
$$

(note that $\mu$ may take negative values). Condition (7) and the use of $\sigma_{\epsilon}^{2}$ (instead of $\sigma^{2}$ ) ensure the existence of a distribution on $\mathbb{N}_{0}$ satisfying (i)-(iii). This will be discussed in detail in the Appendix A. The existence of a Markov chain corresponding to $p_{\epsilon}$ is given by the classical theory of Markov chains.

As in (1), the sum in (6) can be seen as a reproduction of the individuals of generation $n-1$, but now the mass of each single individual is only $\epsilon^{2}$. We emphasize that $\epsilon^{-2} x_{k}^{\epsilon}$ is an integer and represents the number of individuals of generation $n-1$. The variance of the number of offspring now depends on the time and the current total mass. Note however that we do not assume that the numbers of offspring of the individuals of a given generation are identically distributed, so that the approximating process is a bit more general than the classical population-size dependent GWP ( $[10,13,17,18]$, etc.). The first summand in (6) can be regarded as a time-dependent immigration into the system.

We may and do regard the Markov chain $\left(\bar{X}_{t_{n}^{\epsilon}}^{\epsilon}: n \in \mathbb{N}_{0}\right)$ as a time-continuous Markov process $\left(X_{t}^{\epsilon}: t \geq 0\right)$ with trajectories in the cádlàg space $D(\mathbb{R})$ by setting $X_{t}^{\epsilon}=\bar{X}_{\lfloor t]_{\epsilon}}^{\epsilon}$. The unique continuous solution $X$ of (3) can be seen as a cádlàg process too. Hence both $X^{\epsilon}$ and $X$ induce laws on $D(\mathbb{R})$. We denote them by $\mathbb{P}_{\epsilon}$ and $\mathbb{P}$, respectively. Now we can state the main result which holds for every sequence $\left(\epsilon_{\alpha}\right)$ with $\epsilon_{\alpha}>0$ and $\epsilon_{\alpha} \rightarrow 0$.

Theorem 2.1 $\mathbb{P}_{\epsilon_{\alpha}} \Rightarrow \mathbb{P}$ as $\alpha \rightarrow \infty$.
Here, $\Rightarrow$ symbolizes weak convergence. The proof is relegated to Section 5 .
Remark 2.2 Since the laws $\mathbb{P}_{\epsilon_{\alpha}}$ have support in $D\left(\mathbb{R}_{+}\right)$, Theorem 2.1 shows in particular that the unique solution of $\operatorname{SDE}(3)$ is nonnegative.

Remark 2.3 It might be possible to relax the 4-integrability of the offspring numbers $N_{n, i}^{\epsilon}\left(x_{k}^{\epsilon}\right)$. However, this integrability condition typically does not raise any problem. The basic problem is to find any distribution on $\mathbb{N}_{0}$ satisfying (i)-(ii). Once having found such a distribution, the 4-integrability as well as the validity of (iii) often follow automatically, see the Appendix A. For this reason (and for the sake of clarity) we desisted from weakening the 4 -integrability, and we rather decided to work with the convenient criterion of Theorem 4.2.

Remark 2.4 In the scope of the approximation of (2) by a re-scaled version of (1), both the index set and the state space of the re-scaled GWP can be chosen to be $\epsilon \mathbb{N}_{0}$ (cf. [24]). If however $\delta \neq 0$ in the context of $(3)$ then $\epsilon \mathbb{N}_{0}$ is excluded from being the state space. Indeed: The first summand in (6) has to be an element of the state space, and, moreover, this summand divided by $\epsilon$ has to converge to $\delta\left(t_{n-1}^{\epsilon}\right)$, as $\epsilon \rightarrow 0$, to make sure an application of Theorem 4.2 (or of any other related criterion). In general these two requirements contradict each other. If $\delta=0$ in the context of (3) then, in turn, the state space of $\bar{X}^{\epsilon}$ can be chosen to be $\epsilon \mathbb{N}_{0}$ (instead of $\epsilon^{2} \mathbb{N}_{0}$ ). Of course, in this case (6) and (i)-(iii) have to be modified accordingly:

$$
\sum_{i=1}^{x_{k}^{\epsilon} / \epsilon} \epsilon N_{n-1, i}^{\epsilon}\left(x_{k}^{\epsilon}\right)
$$

with $\left\{N_{n, i}^{\epsilon}\left(x_{k}^{\epsilon}\right): i \in \mathbb{N}, n, k \in \mathbb{N}_{0}\right\}$ any family of independent, 4-integrable, $\mathbb{N}_{0}$-valued random variables such that, for every $i, n \in \mathbb{N}$ and $k \in \mathbb{N}_{0}$, the random variable $N_{n-1, i}^{\epsilon}\left(x_{k}^{\epsilon}\right)$ has
(i) mean $1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right)$,
(ii) variance $\sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right)$,
(iii) fourth moment being bounded above by $C_{4}\left(1+\left(x_{k}^{\epsilon}\right)^{3}\right) / \epsilon^{\eta}$, where $\eta<2$ and $x_{k}^{\epsilon}=k \epsilon$, and $\sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right)$ and $C_{4}$ are as before.

## 3 Long-time behavior of the limiting process

If $\mathrm{SDE}(3)$ is regraded as the dynamics of the evolution of a population system, it is natural to ask whether the population may become (temporally) extinct. In this section, we take up this question. In fact, we are going to study the long-time behavior of the solution of (3) for a particular set of coefficients. The question of (temporal) extinction then corresponds to the question whether the stopping time $T$, defined below, is finite. Let $\delta \geq 0, \mu \in \mathbb{R}, \nu>0$, and $0 \leq \gamma \leq 1 / 2$. Further, let $X$ be the unique weak solution of

$$
\begin{equation*}
d X_{t}=\delta+\mu X_{t}+\nu\left|X_{t}\right|^{\gamma+1 / 2} d W_{t}, \quad X_{0}=y>0 \tag{8}
\end{equation*}
$$

We emphasize that (8) is a special case of (3), and that $X$ is nonnegative by Theorem 2.1. Let $T=\inf \left\{t \geq 0: X_{t}=0\right\}$ be the first hitting time of the level 0 , where we use the convention $\inf \emptyset=\infty$. With the help of general results on the long-time behavior of SDEs $([6,16])$ we will show that, depending on the choice of $\delta, \mu, \nu$, and $\gamma$, one of the following statements holds:
(a) $\mathbb{P}[T=\infty]=\mathbb{P}\left[\inf _{t \geq 0} X_{t}=0\right]=\mathbb{P}\left[\sup _{t \geq 0} X_{t}=\infty\right]=1$
(b) $\mathbb{P}[T=\infty]=\mathbb{P}\left[\inf _{t \geq 0} X_{t}>0\right]=\mathbb{P}\left[\lim _{t \rightarrow \infty} X_{t}=\infty\right]=1$
(c) $\mathbb{P}[T=\infty]=\mathbb{P}\left[\lim _{t \rightarrow \infty} X_{t}=0\right]=1$
(d) $\mathbb{P}[T<\infty]=\mathbb{P}\left[X_{t}=0 \quad \forall t \geq T\right]=1$
(e) $\mathbb{P}[T<\infty]=\mathbb{P}\left[\sup _{t \geq 0} X_{t}=\infty\right]=1$
(f) $\mathbb{P}[T<\infty]=1-\mathbb{P}\left[\lim _{t \uparrow T} X_{t}=\infty\right]=\frac{\pi(y)}{\pi(0)}$ and $\mathbb{P}\left[X_{t}=0 \forall t \geq T \mid T<\infty\right]=1$
(g) $\mathbb{P}[T<\infty]=1-\mathbb{P}\left[\lim _{t \uparrow T} X_{t}=\infty\right]=\frac{\pi(y)}{\pi(0)}$ and $\mathbb{P}\left[\sup _{t \geq 0} X_{t}=\infty\right]=1$
where

$$
\pi(x)=\int_{x}^{\infty} \exp \left(-\frac{2}{\nu^{2}} \int_{1}^{s} \frac{\delta+\mu z}{z^{2 \gamma+1}} d z\right) d s
$$

Note that the statements (a)-(g) contradict each other.
Theorem 3.1 The following table shows, in dependence on the choice of $\delta, \mu, \nu$, and $\gamma$, which of the statements (a)-(g) holds.

|  | $\gamma=0$ |  | $\gamma \in(0,1 / 2)$ | $\gamma=1 / 2$ |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| (a) | $\delta \geq \nu^{2} / 2$ | $\mu<0$ | $\delta>0$ | $\mu \leq 0$ | $\delta \geq 0$ | $\mu=\nu^{2} / 2$ |
|  | $\delta=\nu^{2} / 2$ | $\mu=0$ |  |  | $\delta>0$ | $\mu<\nu^{2} / 2$ |
| (b) | $\delta \geq \nu^{2} / 2$ | $\mu>0$ | $\delta>0$ | $\mu>0$ | $\delta \geq 0$ | $\mu>\nu^{2} / 2$ |
|  | $\delta>\nu^{2} / 2$ | $\mu=0$ |  |  |  |  |
| (c) |  |  |  |  | $\delta=0$ | $\mu<\nu^{2} / 2$ |
| (d) | $\delta=0$ | $\mu \leq 0$ | $\delta=0$ | $\mu \leq 0$ |  |  |
| (e) | $\delta \in\left(0, \nu^{2} / 2\right)$ | $\mu \leq 0$ |  |  |  |  |
| (f) | $\delta=0$ | $\mu>0$ | $\delta=0$ | $\mu>0$ |  |  |
| (g) | $\delta \in\left(0, \nu^{2} / 2\right)$ | $\mu>0$ |  |  |  |  |

Proof (of Theorem 3.1) We adopt the notation of [16], and we choose $I=(0, \infty)$ and $c=1$. Since $X$ is finite for all times $t \geq 0$ ("no explosion in finite time"), the stopping time $T$ defined above coincides with the first exit time of the interval $I$, i.e. with $S=\inf \left\{t \geq 0: X_{t} \notin I\right\}$ (which is studied in [16, Section 5.5.C]). The scale function and the speed measure are given by

$$
p(x)=\int_{c}^{x} \exp \left(-\frac{2}{\nu^{2}} \int_{c}^{s} \frac{\delta+\mu z}{z^{2 \gamma+1}} d z\right) d s, \quad m(d x)=\frac{2 / \nu^{2}}{p^{\prime}(x) x^{2 \gamma+1}} d x \quad(x \in I) .
$$

We further set $v(\xi)=\int_{c}^{\xi}(p(\xi)-p(x)) m(d x)$ for $\xi \in I$. We concentrate on the case $\gamma=0$ (for $\gamma \in(0,1 / 2)$ and $\gamma=1 / 2$ one can proceed in the same line). If $\delta \geq \nu^{2} / 2$ and $\mu<0$, or $\delta=\nu^{2} / 2$ and $\mu=0$, then $p(0+)=-\infty$ and $p(\infty-)=\infty$, so that [16, Proposition 5.5.22(a)] implies (a). If $\delta \geq \nu^{2} / 2$ and $\mu>0$, or $\delta>\nu^{2} / 2$ and $\mu=0$, then $p(0+)=-\infty$ and $p(\infty-)<\infty$, so that [16, Proposition 5.5.22(c)] implies (b). If $\delta \in\left[0, \nu^{2} / 2\right)$ and $\mu \leq 0$, then $v(0+)<\infty$ and $p(\infty-)=\infty$. Thus [16, Proposition 5.5.32(iii)] implies $\mathbb{P}[T<\infty]=1$. If in addition $\delta=0$, then it is not hard to show that the assumptions of Theorem 2.13 of [6] are fulfilled, so that this theorem and the nonnegativity of $X$ imply that the second probability in (d) equals one
too. If, on the other hand, in addition $\delta$ is strictly inbetween 0 and $\nu^{2} / 2$, then it is not hard to show that the assumptions of Theorem 2.12 of [6] are fulfilled (for every level $a>0$ ), so that this theorem and the finiteness of $X$ imply that the second probability in (e) equals one too.

If $\delta=0$ and $\mu>0$, then $p(0+)>-\infty$ and $p(\infty-)<\infty$. Therefore [16, Proposition $5.5 .22(\mathrm{~d})]$ implies

$$
\mathbb{P}\left[\lim _{t \uparrow T} X_{t}=0\right]=1-\mathbb{P}\left[\lim _{t \uparrow T} X_{t}=\infty\right]=\frac{\pi(y)}{\pi(0)}
$$

For the first statement of (f) it remains to show $\mathbb{P}\left[\lim _{t \uparrow T} X_{t}=0\right]=\mathbb{P}[T<\infty]$. By Bayes' theorem, it suffices to show

$$
\begin{equation*}
\mathbb{P}\left[\lim _{t \uparrow T} X_{t}=0 \mid T<\infty\right]=1 \tag{9}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbb{P}\left[T<\infty \mid \lim _{t \uparrow T} X_{t}=0\right]=1 \tag{10}
\end{equation*}
$$

Assertion (9) is obvious. To verify (10) we will use Theorem 2.13 of [6]. One can check that the assumptions of this theorem are satisfied for every level $a>0$, so that this theorem implies $\mathbb{P}\left[T_{(0, a)}<\infty\right]=1$ for every $a>0$, where $T_{(0, a)}=\inf \left\{t \geq 0: X_{t} \notin(0, a)\right\}$. In particular, $\mathbb{P}\left[T_{(0, a)}<\infty \forall a \in \mathbb{N}\right]=1$, and therefore $\mathbb{P}\left[T_{(0, a)}<\infty \forall a \in \mathbb{N} \mid \lim _{t \uparrow T} X_{t}=0\right]=1$. Also, we clearly have $\mathbb{P}\left[\sup _{0 \leq t \leq T} X_{t}<\infty \mid \lim _{t \uparrow T} X_{t}=0\right]=1$. Hence,

$$
\begin{equation*}
\mathbb{P}\left[T_{(0, a)}<\infty \forall a \in \mathbb{N}, \text { and } \sup _{0 \leq t \leq T} X_{t}<\infty \mid \lim _{t \uparrow T} X_{t}=0\right]=1 \tag{11}
\end{equation*}
$$

Since $\left\{T_{(0, a)}<\infty \forall a \in \mathbb{N}\right.$, and $\left.\sup _{0 \leq t \leq T} X_{t}<\infty\right\} \subset\{T<\infty\}$, equation (11) implies (10). The second statement of (f) can be shown with the help of [6, Theorem 2.13] as in the proof of (d).

If $\delta \in\left(0, \nu^{2} / 2\right)$ and $\mu>0$, then the first statement of $(\mathrm{g})$ can be obtained in the same way as the first statement of (f) with the following change. This time equation (10) follows from

$$
\begin{aligned}
\mathbb{P}\left[T<\infty \mid \lim _{t \uparrow T} X_{t}=0\right] & =1-\mathbb{P}\left[T=\infty \mid \lim _{t \uparrow T} X_{t}=0\right] \\
& =1-\mathbb{P}\left[T=\infty, \lim _{t \uparrow T} X_{t}=0\right] \mathbb{P}\left[\lim _{t \uparrow T} X_{t}=0\right]^{-1} \\
& =1
\end{aligned}
$$

The last step is due to $\mathbb{P}\left[T=\infty, \lim _{t \uparrow T} X_{t}=0\right]=0$, which in turn follows from the second statement of $(\mathrm{g})$. The latter can be shown with the help of [6, Theorem 2.12] as in the proof of (e).

## 4 A general criterion for the approximation of SDEs

In this Section, we recall from [28] a general criterion for the weak convergence of step processes with fix equidistant jump times to the solution of $\operatorname{SDE}$ (12). We emphasize that there are several other related criteria that could in part be used for our purposes too (see, for instance, $[8,14,20,21,25,26]$ and references cited therein). We consider the solution of the following one-dimensional SDE:

$$
\begin{equation*}
d X_{t}=b\left(t, X_{t}\right) d t+a\left(t, X_{t}\right) d W_{t}, \quad X_{0}=y \tag{12}
\end{equation*}
$$

where $y \in \mathbb{R}$, and $a$ and $b$ are continuous functions on $\mathbb{R}_{+} \times \mathbb{R}$ satisfying

$$
\begin{equation*}
|a(t, x)|+|b(t, x)| \leq K(1+|x|) \quad \forall t \in \mathbb{R}_{+}, x \in \mathbb{R} \tag{13}
\end{equation*}
$$

for some constant $K>0$. We presuppose the existence of a weak solution of (12). That means, there exists a triplet $\left\{X ; W ;\left(\Omega, \mathcal{F},\left(\mathcal{F}_{t}\right), \mathbb{P}\right)\right\}$ where $\left(\Omega, \mathcal{F},\left(\mathcal{F}_{t}\right), \mathbb{P}\right)$ is a filtered probability space with $\left(\mathcal{F}_{t}\right)$ satisfying the usual conditions, $W=\left(W_{t}: t \geq 0\right)$ is an $\left(\mathcal{F}_{t}\right)$-Brownian motion, and $X=\left(X_{t}: t \geq 0\right)$ is a real-valued continuous $\left(\mathcal{F}_{t}\right)$-adapted process such that, for every $t \geq 0$,

$$
\int_{0}^{t}\left(\left|b\left(s, X_{r}\right)\right|+a^{2}\left(r, X_{r}\right)\right) d r<\infty \quad \mathbb{P} \text {-a.s. }
$$

and

$$
X_{t}=y+\int_{0}^{t} b\left(r, X_{r}\right) d r+\int_{0}^{t} a\left(s, X_{r}\right) d W_{r} \quad \mathbb{P} \text {-a.s. }
$$

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

$$
\begin{equation*}
\left|a(t, x)-a\left(t, x^{\prime}\right)\right| \leq h\left(\left|x-x^{\prime}\right|\right) \quad \forall t \in \mathbb{R}_{+}, x, x^{\prime} \in \mathbb{R} \tag{14}
\end{equation*}
$$

for some strictly increasing $h: \mathbb{R}_{+} \rightarrow \mathbb{R}_{+}$with $\int_{0}^{0+} h^{-2}(u) d u=\infty$. Note that (14) and Lipschitz continuity of $b$ even imply the existence of a strongly unique strong solution (YamadaWatanabe criterion [27]). But the notion of strong solutions and strong uniqueness is beyond our interest.

Now, for every $\alpha \in \mathbb{N}$ we fix some $\epsilon_{\alpha}>0$ such that $\epsilon_{\alpha} \rightarrow 0$, and we set $t_{n}^{\alpha}=t_{n}^{\epsilon_{\alpha}}\left(=n \epsilon_{\alpha}\right)$ for all $n \in \mathbb{N}_{0}$. We further let $a_{\alpha}$ and $b_{\alpha}$ be measurable functions on $\mathbb{R}_{+} \times \mathbb{R}$ such that $\left\|a-a_{\alpha}\right\|_{\infty}$ and $\left\|b-b_{\alpha}\right\|_{\infty}$ converge to 0 as $\alpha \rightarrow \infty$, where $\|\cdot\|_{\infty}$ is the usual supremum norm. We consider a sequence $\left(y_{\alpha}\right) \subset \mathbb{R}$ satisfying $y_{\alpha} \rightarrow y$, and we suppose that $X^{\alpha}$ is a solution of the following $\left(\epsilon_{\alpha}, a_{\alpha}, b_{\alpha}, y_{\alpha}\right)$-martingale problem for each $\alpha \geq 1$. Here, $n_{\alpha}(t)$ denotes the largest $n \in \mathbb{N}_{0}$ with $t_{n}^{\alpha} \leq t$.

Definition 4.1 Suppose $X^{\alpha}=\left(X_{t}^{\alpha}: t \geq 0\right)$ is a real-valued process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ whose trajectories are constant on the intervals $\left[t_{n}^{\alpha}, t_{n+1}^{\alpha}\right), n \in \mathbb{N}_{0}$. Then $X^{\alpha}$ is called solution of the $\left(\epsilon_{\alpha}, a_{\alpha}, b_{\alpha}, y_{\alpha}\right)$-martingale problem if

$$
\begin{equation*}
M_{t}^{\alpha}=X_{t}^{\alpha}-y_{\alpha}-\sum_{i=0}^{n_{\alpha}(t)-1} b_{\alpha}\left(t_{i}^{\alpha}, X_{t_{i}^{\alpha}}^{\alpha}\right) \epsilon_{\alpha} \tag{15}
\end{equation*}
$$

provides a (zero-mean) square-integrable martingale (with respect to the natural filtration) with compensator

$$
\begin{equation*}
\left\langle M^{\alpha}\right\rangle_{t}=\sum_{i=0}^{n_{\alpha}(t)-1} a_{\alpha}^{2}\left(t_{i}^{\alpha}, X_{t_{i}^{\alpha}}^{\alpha}\right) \epsilon_{\alpha} . \tag{16}
\end{equation*}
$$

The $X^{\alpha}$ could be defined on different probability spaces $\left(\Omega_{\alpha}, \mathcal{F}_{\alpha}, \mathbb{P}_{\alpha}\right)$. However we assume without loss of generality $\Omega_{\alpha}=D(\mathbb{R}), \mathcal{F}_{\alpha}=\mathcal{B}(D(\mathbb{R}))$, and that $X^{\alpha}$ 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 $\gamma>1$ such that

$$
\begin{equation*}
\mathbb{E}_{\alpha}\left[\left|X_{t_{n}^{\alpha}}^{\alpha}-X_{t_{n-1}^{\alpha}}^{\alpha}\right|^{q}\right] \leq C_{T}\left(1+\mathbb{E}_{\alpha}\left[\left|X_{t_{n-1}^{\alpha}}^{\alpha}\right|^{q}\right]\right) \epsilon_{\alpha}^{\gamma} \tag{17}
\end{equation*}
$$

for every $\alpha, n \in \mathbb{N}$ with $t_{n}^{\alpha} \leq T$, where $C_{T}>0$ is any constant that may depend on $T$. By an induction on $n$, (17) implies in particular that $\mathbb{E}_{\alpha}\left[\left|X_{t_{n}^{\alpha}}^{\alpha}\right|^{q}\right]<\infty$ for all $\alpha$ and $n$. That is, the processes $X^{\alpha}$ have to be $q$-integrable. The following theorem shows that $X^{\alpha}$ converges in distribution to the unique solution of (12).

Theorem 4.2 [28, Theorem 2.2] Suppose SDE (12) subject to (13) 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 $X^{\alpha}$ subject to (15)-(17). Then, $\mathbb{P}_{\alpha} \Rightarrow \mathbb{P}$ as $\alpha \rightarrow \infty$.

## 5 Proof of Theorem 2.1

We adopt the notation introduced in Sections 1 and 2. However, for the sake of clarity, we write $X^{\alpha}, \mathbb{P}_{\alpha}, \sigma_{\alpha}, t_{i}^{\alpha}, x_{i}^{\alpha}$ and $y_{\alpha}$ in place of $X^{\epsilon_{\alpha}}, \mathbb{P}_{\epsilon_{\alpha}}, \sigma_{\epsilon_{\alpha}}, t_{i}^{\epsilon_{\alpha}}, x_{i}^{\epsilon_{\alpha}}$ and $\lfloor y\rfloor_{\epsilon_{\alpha}^{2}}$, respectively. We intend to show that $X^{\alpha}$ converges in distribution to the solution of (3), i.e., that $\mathbb{P}_{\alpha} \Rightarrow \mathbb{P}$. To this end we set

$$
a_{\alpha}(t, x)=\sigma_{\alpha}(t, x) \sqrt{x} \quad \text { and } \quad b_{\alpha}(t, x)=\lfloor\delta(t)\rfloor_{\epsilon_{\alpha}}+\mu(t) x,
$$

and we note that $a_{\alpha}(t, x)$ and $b_{\alpha}(t, x)$ converge uniformly in $(t, x)$ to $a(t, x)=\sigma(t, x) \sqrt{x}$ and $b(t, x)=\delta(t)+\mu(t) x$, respectively, as $\alpha \rightarrow \infty$. Moreover we have $y_{\alpha} \rightarrow y$. By Theorem 4.2 it thus suffices to show that $X^{\alpha}$ solves the $\left(\epsilon_{\alpha}, a_{\alpha}, b_{\alpha}, y_{\alpha}\right)$-martingale problem (Definition 4.1), and that $X^{\alpha}$ satisfies (17). The key tools will be the following two lemmas. The proof of the first lemma is relegated to the end of this Section.

Lemma 5.1 There is some constant $C>0$ such that, for all $\alpha, n \in \mathbb{N}$,

$$
\begin{equation*}
\mathbb{E}_{\alpha}\left[\left(X_{t_{n}^{\alpha}}^{\alpha}-X_{t_{n-1}^{\alpha}}^{\alpha}\right)^{4}\right] \leq C\left(1+\mathbb{E}_{\alpha}\left[\left(X_{t_{n-1}^{\alpha}}^{\alpha}\right)^{4}\right]\right) \epsilon_{\alpha}^{\gamma} \tag{18}
\end{equation*}
$$

with $\gamma=\min \{2 ; 6-\eta\}>1$. In particular, the process $X^{\alpha}$ is 4-integrable.
Lemma 5.2 For all $\alpha, n \in \mathbb{N}$,

$$
\begin{align*}
\mathbb{E}_{\alpha}\left[X_{t_{n}^{\alpha}}^{\alpha} \mid X_{t_{n-1}^{\alpha}}^{\alpha}\right] & =\epsilon_{\alpha}\left\lfloor\delta\left(t_{n-1}^{\alpha}\right)\right\rfloor_{\epsilon_{\alpha}}+\left(1+\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right)\right) X_{t_{n-1}^{\alpha}}^{\alpha}  \tag{19}\\
\operatorname{Var}_{\alpha}\left[X_{t_{n}^{\alpha}}^{\alpha} \mid X_{t_{n-1}^{\alpha}}^{\alpha}\right] & =\epsilon_{\alpha} \sigma_{\alpha}^{2}\left(t_{n-1}^{\alpha}, X_{t_{n-1}^{\alpha}}^{\alpha}\right) X_{t_{n-1}^{\alpha}}^{\alpha}  \tag{20}\\
\mathbb{E}_{\alpha}\left[\left(X_{t_{n}^{\alpha}}^{\alpha}\right)^{2} \mid X_{t_{n-1}^{\alpha}}^{\alpha}\right] & \left.=\epsilon_{\alpha} \sigma_{\alpha}^{2}\left(t_{n-1}^{\alpha}, X_{t_{n-1}^{\alpha}}^{\alpha}\right) X_{t_{n-1}^{\alpha}}^{\alpha}+\left(\epsilon_{\alpha} \mid \delta\left(t_{n-1}^{\alpha}\right)\right\rfloor_{\epsilon_{\alpha}}+\left(1+\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right)\right) X_{t_{n-1}^{\alpha}}^{\alpha}\right)^{2}(, \tag{21}
\end{align*}
$$

Proof Since the sum in (6) consists of $\left(\epsilon^{-2} x_{k}^{\epsilon}\right)$-many independent summands, each of which has mean $\epsilon^{2}+\epsilon^{3} \mu\left(t_{n-1}^{\alpha}\right)$ and variance $\epsilon^{3} \sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right)$, the formulae (19) and (20) are obvious. Further, (21) follows immediately from (19) and (20).

Now, let $M^{\alpha}$ and $\left\langle M^{\alpha}\right\rangle$ be the processes defined in (15) and (16), respectively. $M^{\alpha}$ is squareintegrable since $X^{\alpha}$ is, and it is an $\left(\mathcal{F}_{t}^{X^{\alpha}}\right)$-martingale since

$$
\begin{aligned}
& \mathbb{E}_{\alpha}\left[M_{t_{n}^{\alpha}}^{\alpha}-M_{t_{n-1}^{\alpha}}^{\alpha} \mid \mathcal{F}_{t_{n-1}^{\alpha}}^{X^{\alpha}}\right] \\
& \quad=\mathbb{E}_{\alpha}\left[X_{t_{n}^{\alpha}}^{\alpha}-X_{t_{n-1}^{\alpha}}^{\alpha}-b_{\alpha}\left(t_{n-1}^{\alpha}, X_{t_{n-1}^{\alpha}}^{\alpha}\right) \epsilon_{\alpha} \mid X_{t_{n-1}^{\alpha}}^{\alpha}\right] \\
& \quad=\mathbb{E}_{\alpha}\left[X_{t_{n}^{\alpha}}^{\alpha} \mid X_{t_{n-1}^{\alpha}}^{\alpha}\right]-\epsilon_{\alpha}\left\lfloor\delta\left(t_{n-1}^{\alpha}\right)\right\rfloor_{\epsilon_{\alpha}}-\left(1+\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right)\right) X_{t_{n-1}^{\alpha}}^{\alpha}
\end{aligned}
$$

vanishes by (19). With the help of (19) and (21) we further obtain by straightforward calculations

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

This shows that $\left\langle M^{\alpha}\right\rangle$ is the compensator of $M^{\alpha}$ since the Doob-Meyer decomposition is unique. Hence $X^{\alpha}$ solves the $\left(\epsilon_{\alpha}, a_{\alpha}, b_{\alpha}, y_{\alpha}\right)$-martingale problem. Moreover, by (18) we have (17) with $q=4$. This completes the proof of Theorem 2.1.

For the proof of Lemma 5.1 we need the following lemma. Its proof is elementary and will be omitted.

Lemma 5.3 Assume $\xi_{1}, \xi_{2}, \ldots$ are independent random variables on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ with $\tilde{\mathbb{E}}\left[\xi_{i}\right]=0$ and $\sup _{i \in \mathbb{N}} \tilde{\mathbb{E}}\left[\xi_{i}^{4}\right]<\infty$. Then, for every $\nu \in \mathbb{N}$,

$$
\tilde{\mathbb{E}}\left[\left(\sum_{i=1}^{\nu} \xi_{i}\right)^{4}\right] \leq \nu \sup _{i \in \mathbb{N}} \tilde{\mathbb{E}}\left[\xi_{i}^{4}\right]+3 \nu^{2} \sup _{i, j \neq i} \tilde{\mathbb{E}}\left[\xi_{i}^{2}\right] \tilde{\mathbb{E}}\left[\xi_{j}^{2}\right]
$$

Proof (of Lemma 5.1) Plainly,

$$
\begin{align*}
& \mathbb{E}_{\alpha}\left[\left(X_{t_{n}^{\alpha}}^{\alpha}-X_{t_{n-1}^{\alpha}}^{\alpha}\right)^{4}\right] \\
& \quad=\sum_{(k, m) \in \mathbb{N}_{0}^{2}}\left(x_{m}^{\alpha}-x_{k}^{\alpha}\right)^{4} \mathbb{P}_{\alpha}\left[X_{t_{n}^{\alpha}}^{\alpha}=x_{m}^{\alpha} \mid X_{t_{n-1}^{\alpha}}^{\alpha}=x_{k}^{\alpha}\right] \mathbb{P}_{\alpha}\left[X_{t_{n-1}^{\alpha}}^{\alpha}=x_{k}^{\alpha}\right] \tag{22}
\end{align*}
$$

If $(\bar{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ denotes the domain of the random variables $N_{n-1, i}^{\epsilon}\left(x_{k}^{\epsilon}\right)$ in (6), we also obtain with the help of (5),

$$
\begin{align*}
& \sum_{m \in \mathbb{N}_{0}}\left(x_{m}^{\alpha}-x_{k}^{\alpha}\right)^{4} \mathbb{P}_{\alpha}\left[X_{t_{n}^{\alpha}}^{\alpha}=x_{m}^{\alpha} \mid X_{t_{n-1}^{\alpha}}^{\alpha}=x_{k}^{\alpha}\right] \\
& \quad=\overline{\mathbb{E}}\left[\left(\epsilon_{\alpha}\left\lfloor\delta\left(t_{n-1}^{\alpha}, x_{k}^{\alpha}\right)\right\rfloor_{\epsilon_{\alpha}}+\sum_{i=1}^{\epsilon_{\alpha}^{-2} x_{k}^{\alpha}} \epsilon_{\alpha}^{2} N_{n-1, i}\left(x_{k}^{\alpha}\right)-x_{k}^{\alpha}\right)^{4}\right]  \tag{23}\\
& \quad \leq 3^{3}\left\{\left(\epsilon_{\alpha} K\right)^{4}+\overline{\mathbb{E}}\left[\left(\sum_{i=1}^{\epsilon_{\alpha}^{-2} x_{k}^{\alpha}} \epsilon_{\alpha}^{2}\left[N_{n-1, i}\left(x_{k}^{\alpha}\right)-\left(1+\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right)\right]\right)^{4}\right]+\left(\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right) x_{k}^{\alpha}\right)^{4}\right\}\right.
\end{align*}
$$

We assume without loss of generality $0<\epsilon_{\alpha} \leq 1$, and we set

$$
S_{l}^{\alpha}=\sup _{i \in \mathbb{N}} \overline{\mathbb{E}}\left[\left[N_{n-1, i}\left(x_{k}^{\alpha}\right)-\left(1+\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right)\right]^{l}\right] \quad(l=2,4) .\right.
$$

By (4) and (ii)-(iii) of Section 2, we have

$$
S_{2}^{\alpha} \leq K\left(1+x_{k}^{\alpha}\right) / \epsilon_{\alpha} \quad \text { and } \quad S_{4}^{\alpha} \leq C_{4}\left(1+\left(x_{k}^{\alpha}\right)^{3}\right) \epsilon^{-\eta}
$$

Therefore we obtain with the help of Lemma 5.3,

$$
\begin{align*}
& \overline{\mathbb{E}}\left[\left(\sum_{i=1}^{\epsilon_{\alpha}^{-2} x_{k}^{\alpha}} \epsilon_{\alpha}^{2}\left[N_{n-1, i}\left(x_{k}^{\alpha}\right)-\left(1+\epsilon_{\alpha} \mu\left(t_{n-1}^{\alpha}\right)\right]\right)^{4}\right]\right. \\
& \quad \leq \epsilon_{\alpha}^{8}\left(\left(\epsilon_{\alpha}^{-2} x_{k}^{\alpha}\right) S_{4}^{\alpha}+3\left(\epsilon_{\alpha}^{-2} x_{k}^{\alpha}\right)^{2}\left(S_{2}^{\alpha}\right)^{2}\right) \\
& \quad \leq \epsilon_{\alpha}^{8}\left(\left(\epsilon_{\alpha}^{-2} x_{k}^{\alpha}\right) C_{4}\left(1+\left(x_{k}^{\alpha}\right)^{3}\right) \epsilon_{\alpha}^{-\eta}+3\left(\epsilon_{\alpha}^{-2} x_{k}^{\alpha}\right)^{2}\left(K\left(1+x_{k}^{\alpha}\right) / \epsilon_{\alpha}\right)^{2}\right) \\
& \quad \leq(C / 2)\left(1+\left(x_{k}^{\alpha}\right)^{4}\right)\left(\epsilon_{\alpha}^{6-\eta}+\epsilon_{\alpha}^{2}\right) \tag{24}
\end{align*}
$$

for some suitable constant $C>0$. Now (22), (23), (5), (24), and Fubini's theorem, imply (18).

## A Appendix

The construction of the approximating process $\bar{X}^{\epsilon}$ in Section 2 relies on the existence of a distribution on $\mathbb{N}_{0}$ satisfying (i)-(iii) of Section 2. The existence of such a distribution is not completely obvious. Basically the problem is to find any distribution on $\mathbb{N}_{0}$ with mean $m$ and variance $v$ for given $m, v>0$. Provided

$$
\begin{equation*}
m<m^{2}+v \tag{25}
\end{equation*}
$$

the counting density of the requested distribution can be chosen as

$$
\begin{equation*}
p_{0}=1-\sum_{j=1}^{\infty} p_{j}, \quad p_{k}=c q^{k-1} \quad(k \geq 1) \tag{26}
\end{equation*}
$$

with

$$
c=\frac{4 m^{3}}{\left(v+m^{2}+m\right)^{2}} \quad \text { and } \quad q=\frac{v+m^{2}-m}{v+m^{2}+m}
$$

(note that condition (25) is more or less necessary for the existence of the requested distribution since $m^{2}+v=\sum_{k=1}^{\infty} k^{2} p_{k} \geq \sum_{k=1}^{\infty} k p_{k}=m$ ). Condition (25) implies $0<c, q<1$ and $q \leq 1-c$, which in turn ensures that $\left(p_{k}\right)$ provides a counting density of a probability measure. Note that the corresponding probability generating function is given by

$$
\begin{equation*}
g(t)=1-\frac{c}{1-q}+\frac{c t}{1-q t} \quad(0<t<1) . \tag{27}
\end{equation*}
$$

Let $N$ be any random variable whose law has the counting density (26). Using (27) and

$$
\begin{equation*}
\mathbb{E}[N(N-1) \cdots(N-k+1)]=g^{(k)}(1) \quad(k \in \mathbb{N}) \tag{28}
\end{equation*}
$$

we easily obtain $\mathbb{E}[N]=m$ and $\operatorname{Var}[N]=v$. With the help of (27) and (28) we further obtain

$$
\begin{align*}
\mathbb{E}\left[N^{4}\right] & =g^{(4)}(1)+6 g^{(3)}(1)+7 g^{\prime \prime}(1)+g^{\prime}(1) \\
& =\frac{24 c q^{3}}{(1-q)^{5}}+\frac{36 c q^{2}}{(1-q)^{4}}+\frac{14 c q}{(1-q)^{3}}+\frac{c}{(1-q)^{2}} \\
& \leq 36 \sum_{k=1}^{4} \frac{4 m^{3}\left(v+m^{2}-m\right)^{k-1}}{(2 m)^{k+1}} \tag{29}
\end{align*}
$$

In the context of (i)-(iii) of Section 2, we now set

$$
\begin{equation*}
m=1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right) \quad \text { and } \quad v=\sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right) / \epsilon \tag{30}
\end{equation*}
$$

By (5) we have $\epsilon \mu()+.\epsilon K>0$, and therefore we obtain

$$
\begin{aligned}
m & =1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right) \\
& <1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right)+\epsilon \mu\left(t_{n-1}^{\epsilon}\right)+\epsilon^{2} \mu^{2}\left(t_{n-1}^{\epsilon}\right)+\epsilon K+\sigma^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right) / \epsilon \\
& =\left(1+\epsilon \mu\left(t_{n-1}^{\epsilon}\right)\right)^{2}+\sigma_{\epsilon}^{2}\left(t_{n-1}^{\epsilon}, x_{k}^{\epsilon}\right) / \epsilon \\
& =m^{2}+v .
\end{aligned}
$$

That is, condition (25) holds. Therefore the counting density in (26) generates a distribution with mean $m$ and variance $v$ as in (30). Moreover, using (30), (29), (4), (5), and (7), one can easily show that the fourth moment is bounded above by $C_{4}\left(1+\left|x_{k}^{\epsilon}\right|^{3}\right) \epsilon^{-3}$ for all $n \in \mathbb{N}$, $k \in \mathbb{N}_{0}$, and some suitable universal constant $C_{4}>0$. By all account, we have established a distribution on $\mathbb{N}_{0}$ satisfying (i)-(iii) of Section 2.

## References

[1] Asmussen, S. and Hering, H. (1983) Branching processes. Birkhäuser, Boston.
[2] Athreya, K.B. and Jagers, P. (1997) Classical and Modern Branching Processes. SpringerVerlag, New York.
[3] Athreya, K.B. and Ney, P.E. (1972) Branching processes. Springer-Verlag, Berlin.
[4] Buckholtz, P.G. and Wasan, M.T. (1982) Diffusion approximation of the two-type GaltonWatson process with mean matrix close to the identity. J. Multivariate Anal. 12:493-507.
[5] Cox, J., Ingersoll. J. and Ross, S. (1985) A theory of the term structure of interest rates. Econometrica, 53:385-408.
[6] Cherny, A.S. and Engelbert, H.-J. (2005) Singular Stochastic Differential Equations, Lecture Notes in Mathematics, vol. 1858, Springer.
[7] Dyakonova, E.E. (1999) Diffusion approximation of branching migration processes. J. Math. Sci., 93(4):511-514.
[8] Ethier, S.N. and Kurtz, T.G. (1986) Markov processes: Characterization and convergence. Wiley, New York.
[9] Feller, W. (1951) Diffusion processes in genetics. In Proc. Second Berkely Symp. Math. Statist. Probab. Univ. of California Press, Berkely.
[10] Fujimagari, T. (1976) Controlled Galton-Watson process and its asymptotic behavior. Koda Math. Sem. Rep., 27:11-18.
[11] Haccou, P., Jagers, P. and Vatutin, V.A. (2005) Branching Processes: Variation, Growth and Extinction of Populations. Cambridge University Press, Cambridge.
[12] Harris, T.E. (1963) The theory of branching processes. Springer-Verlag, Berlin.
[13] Höpfner, R. (1985) On some class of population-size-dependent Galton-Watson processes. J. Appl. Prob., 22:25-36.
[14] Jacod, J. and Shiryaev, A.N. (2002) Limit theorems for stochastic processes. SpringerVerlag, Berlin.
[15] Jiřina, M. (1958) Stochastic branching processes with continuous state space. Czech. J. Math., 8:292-313.
[16] Karatzas, I. and Shreve, S.E. (1991) Brownian motion and stochastic calculus. SpringerVerlag, New York.
[17] Klebaner, F.C. (1984) On population-size-dependent branching processes. Adv. Appl. Prob., 16:30-55.
[18] Klebaner, F.C. (1984) Geometric rate of growth in population-size-dependent branching processes. J. Appl. Prob., 21:40-49.
[19] Kloeden, P.E. and Platen, E. (1992) Numerical solutions of stochastic differential equations. Springer-Verlag, Berlin.
[20] Kurtz, T.G. and Protter, P. (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. Ann. Prob., 19(3):1035-1070.
[21] Kushner, H.J. (1984) Approximation and weak convergence methods for random processes. The MIT Press Series in Signal Processing, Optimization, and Control, Cambridge.
[22] Lamperti, J. (1966) The limit of a sequence of branching processes. Z. Wahrscheinlichkeitstheorie verw. Geb., 7:271-288.
[23] Le Gall, J.-F. (1999) Spatial branching processes, random snakes and partial differential equations. In Lecture Notes Math. ETH Zürich, Birkhäuser.
[24] Lindvall, T. (1972) Convergence of critical Galton-Watson branching processes. J. Appl. Prob.,9:445-450.
[25] Rebolledo, R. (1979) La méthod des martingales appliquée à l'étude de la convergence en loi de processus. Bull. Soc. Math. France Mem. 62, 125 pages.
[26] Stroock, D.W. and Varadhan, S.R.S. (1979) Multidimensional diffusion processes. Springer-Verlag, Berlin.
[27] Yamada, T. and Watanabe, S. (1971). On the uniqueness of solutions of stochastic differential equations. J. Math. Kyoto Univ., 11:155-167.
[28] Zähle, H. (2008) Weak approximation of SDEs by discrete-time processes. J. Appl. Math. Stoch. Analysis, Vol. 2008, 14 pages.

## Preprints ab 2008

2009-01 Henryk Zähle
Approximation of SDEs by population-size-dependent Galton-Watson processes

Winfried Hazod
MEHLER SEMIGROUPS, ORNSTEIN-UHLENBECK PROCESSES AND BACKGROUND DRIVING LÉVY PROCESSES ON LOCALLY COMPACT GROUPS AND ON HYPERGROUPS

Karl Friedrich Siburg, Pavel A. Stoimenov
Symmetry of functions and exchangeability of random variables
Ina Kirsten Voigt
Voronoi Cells of Discrete Point Sets
Michael Lenzinger and Ben Schweizer
Effective reaction rates of a thin catalyst layer
Michael Voit
Bessel convolutions on matrix cones: Algebraic properties and random walks

Margit Rösler and Michael Voit
Limit theorems for radial random walks on $p \times q$-matrices as $p$ tends to infinity

Michael Voit
Central Limit Theorems for Radial Random Walks on
$p \times q$ Matrices for $p \rightarrow \infty$
Michael Voit
Limit theorems for radial random walks on homogeneous spaces with growing dimensions

2008-17 Ansgar Steland and Henryk Zähle
Sampling inspection by variables: nonparametric setting
Guy Bouchitté and Ben Schweizer
Homogenization of Maxwell's equations with split rings
Wilfried Hazod
Multiple selfdecomposable laws on vector spaces and on groups:
The existence of background driving processes
Wilfried Hazod
Mixing of generating functionals and applications to (semi-)stability
of probabilities on groups

| 2008-13 | Wilfried Hazod Probability on Matrix-Cone Hypergroups: Limit Theorems and Structural Properties |
| :---: | :---: |
| 2008-12 | Michael Lenzinger and Ben Schweizer <br> Two-phase flow equations with outflow boundary conditions in the hydrophobic-hydrophilic case |
| 2008-11 | Karl Friedrich Siburg <br> Geometric proofs of the two-dimensional Borsuk-Ulam theorem |
| 2008-10 | Peter Becker-Kern, Wilfried Hazod <br> Mehler hemigroups and embedding of discrete skew convolution |
| 2008-09 | Karl Friedrich Siburg, Pavel A. Stoimenov Gluing copulas |
| 2008-08 | Karl Friedrich Siburg, Pavel A. Stoimenov A measure of mutual complete dependence |
| 2008-07 | Karl Friedrich Siburg, Pavel A. Stoimenov A scalar product for copulas |
| 2008-06 | Flavius Guiaş Generalized Becker-Döring Equations Modeling the Time Evolution of a Process of Preferential Attachment with Fitness |
| 2008-05 | Benjamin Fine, Alexei Myasnikov, Gerhard Rosenberger Generic Subgroups of Group Amalgams |
| 2008-04 | Ben Schweizer <br> Homogenization of the Prager model in one-dimensional plasticity |
| 2008-03 | Benjamin Fine, Miriam Hahn, Alexander Hulpke, Volkmar große Rebel, Gerhard Rosenberger, Martin Scheer <br> All Finite Generalized Tetrahedron Groups |
| 2008-02 | Benjamin Fine, Gerhard Rosenberger <br> An Epic Drama: The Development of the Prime Number Theorem |
| 2008-01 | Henryk Zähle <br> Weak approximation of SDEs by discrete-time processes |


[^0]:    *Postal address: Technische Universität Dortmund, Fakultät für Mathematik, Vogelpothsweg 87, D-44227 Dortmund, Germany; Email: henryk.zaehle@math.tu-dortmund.de

