ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Christiansen, Marcus C.

Article

Gaussian and affine approximation of stochastic diffusion models for interest and mortality rates

Risks

Provided in Cooperation with:
MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Christiansen, Marcus C. (2013) : Gaussian and affine approximation of stochastic diffusion models for interest and mortality rates, Risks, ISSN 2227-9091, MDPI, Basel, Vol. 1, Iss. 3, pp. 81-100,
https://doi.org/10.3390/risks1030081

This Version is available at:
https://hdl.handle.net/10419/103614

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

http://creativecommons.org/licenses/by/3.0/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Risks **2013**, *1*, 81-100; doi:10.3390/risks1030081

OPEN ACCESS

risks

ISSN 2227-9091

www.mdpi.com/journal/risks

Article

Gaussian and Affine Approximation of Stochastic Diffusion Models for Interest and Mortality Rates

Marcus C. Christiansen

Institute of Insurance Science, University of Ulm, 89069 Ulm, Germany;
E-Mail: marcus.christiansen@uni-ulm.de; Tel.: +49-731-50-31186; Fax: +49-731-50-31188

Received: 8 August 2013; in revised form: 15 October 2013 / Accepted: 17 October 2013 / Published: 25 October 2013

Abstract: In the actuarial literature, it has become common practice to model future capital returns and mortality rates stochastically in order to capture market risk and forecasting risk. Although interest rates often should and mortality rates always have to be non-negative, many authors use stochastic diffusion models with an affine drift term and additive noise. As a result, the diffusion process is Gaussian and, thus, analytically tractable, but negative values occur with positive probability. The argument is that the class of Gaussian diffusions would be a good approximation of the real future development. We challenge that reasoning and study the asymptotics of diffusion processes with affine drift and a general noise term with corresponding diffusion processes with an affine drift term and an affine noise term or additive noise. Our study helps to quantify the error that is made by approximating diffusive interest and mortality rate models with Gaussian diffusions and affine diffusions. In particular, we discuss forward interest and forward mortality rates and the error that approximations cause on the valuation of life insurance claims.

Keywords: forward interest rate; forward mortality rate; life insurance; stochastic diffusion process; Gaussian approximation

---

1. Introduction

For the most part of the 20th century, actuaries used to model the lifespan of an insured stochastically, but relied on a mere deterministic prognosis of capital returns and mortality rates. The past has shown that these assumptions can vary significantly within one contract period. Especially in recent years, financial markets have experienced increased volatility, and life expectancies have risen in many

developed countries at an unforeseen rate. For two decades, actuaries have been increasingly using stochastic interest rate models, and since the last decade, stochastic mortality rate models have become more and more popular. On the one hand, the stochastic approach helps to better capture systematic risk by offering the possibility to calculate confidence intervals. On the other hand, investors and insurance regulators nowadays call for a market-consistent valuation of the insurers assets and liabilities, so that concepts from financial mathematics have entered actuarial science, where the market value of a claim is calculated as a risk-neutral expectation in a specific stochastic model.

In financial mathematics, stochastic modeling of interest rates is a common approach, and a great variety of different interest rate models can be found in the literature. The majority of authors models the interest rate with the help of diffusion processes. Typical short rate models are, for example, the Vasicek model and the Cox–Ingersoll–Ross model. The evolution of forward rates is often modeled within the Heath–Jarrow–Morton framework. Readers may refer to [1] for a comprehensive overview or to the original works, such as [2–4]. In the nineteen nineties, these models also found their way into the actuarial literature; see e.g., [5–11].

The uncertainty about future mortality probabilities always drew the attention of life insurance actuaries. However, since recently, actuaries have been increasingly using stochastic models for the mortality rate. A very popular approach is to adopt the well-known models from interest rate theory, such as the Vasicek or the Cox–Ingersoll–Ross model. A more recent development is the concept of forward mortality rates, which are frequently modeled by a Heath–Jarrow–Morton framework similarly to forward interest rates. References that use diffusion process are, e.g., [12–24], just to mention a few. Although negative mortality rates do not have a meaningful interpretation, many of those references use diffusion models, where negative values occur with a positive probability. In particular, Gaussian diffusion processes are frequently used. This is justified with the argument that Gaussian models would offer a good approximation of more appropriate models.

In the present paper, we study the asymptotics of diffusions with affine drift and a general noise term to corresponding diffusions with an affine drift term, but an affine noise term or additive noise. In the additive case, we end up with Gaussian processes, that is, the finite-dimensional projections are multivariate, normally distributed. The affine case is a generalization of the additive case and allows for better approximations, but the corresponding distributions are not always analytically tractable. We introduce a reasonable approximation concept and calculate bounds for the absolute moments of the approximation errors. Furthermore, we calculate the error that the approximating processes imply when valuating insurance claims. Our basic notion is to study asymptotics relative to the size of the noise term; the smaller the noise term, the better the approximation. Our results help to better assess the quality of Gaussian and affine approximations for interest and mortality rate models.

After introducing a basic life insurance model in Section 2, we study Gaussian approximations in Section 3 and affine approximations in Section 4 and present theoretical results for the asymptotic behavior. The proofs are put into their own section at the end; see Section 7. In Section 5, we discuss the effect of the Gaussian and affine approximations on the valuation of insurance claims. A numerical illustration is given in Section 6. In Section 8, we summarize our findings.

2. Basic Life Insurance Framework

Assume that the lifetime of a policyholder is described by a nonnegative random variable T. Time t measures the seniority of the policy (i.e., the time elapsed since policy issue). Policyholder's age at policy issue is denoted as x, so that the age at time t is $x+t$. Suppose that the distribution function $P[T \leq x+t]$ of the lifetime T has a representation of the form:

$$1 - e^{-\int_0^{x+t} m(s)\mathrm{d}s} \tag{1}$$

where m is the so-called mortality rate (or force of mortality). We consider a general life insurance contract with both survival and death benefits. We write $B(t)$ for the aggregated survival benefits minus premiums on $[0,t]$ and $c(t)$ for the benefit payments that are due in the case of a death at time t. The contract terminates at time $\omega_x = \omega - x < \infty$, at the latest, where ω is the maximal value that T can take, possibly infinite.

Let $r(t)$ be the short rate. Then, the discounted sum of future benefits minus premiums for a policyholder alive at time t is given by:

$$A(t) = \int_{(t,T)} e^{-\int_t^s r(u)\mathrm{d}u} \mathrm{d}B(s) + e^{-\int_t^T r(u)\mathrm{d}u} c(T). \tag{2}$$

In classical life insurance modeling, r and m are integrable functions, B is a right-continuous function with finite variation and c is a bounded function. Therefore, in the classical perspective, r, m, B and c are deterministic, and the prospective reserve at time t in the active state is defined as:

$$V(t; r, m) = E[A(t)|T > t] \tag{3}$$

and has the representation:

$$V(t; r, m) = \int_{(t,\omega_x]} e^{-\int_t^s r(u)+m(x+u)\mathrm{d}u} \mathrm{d}B(s) + \int_t^{\omega_x} e^{-\int_t^s r(u)+m(x+u)\mathrm{d}u} m(x+s)c(s)\mathrm{d}s. \tag{4}$$

In insurance practice, the functions r and m are not perfectly known for the future, but have to be forecasted. In order to capture the corresponding uncertainty, we construct a doubly-stochastic model by assuming that:

- $(r(t))_{t\geq 0}$ is a stochastic process with continuous paths,
- $(m(t))_{t\geq 0}$ is a stochastic process with continuous paths,
- $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration generated by the joint process, (r, m).

Let $\mathcal{F}_\infty := \sigma\big(\bigcup_{t\geq 0} \mathcal{F}_t\big) = \sigma(r(t), m(t), t \geq 0)$. We can embed the classical model, where r and m are deterministic, into the doubly-stochastic model by interpreting Formula (1) as the conditional distribution function $P[T \leq x+t|(r(t), m(t))_{t\geq 0}] := P[T \leq x+t|\mathcal{F}_\infty]$. Doing that for all $x+t$, Formula (1) uniquely defines a stochastic kernel $(A, (r, m)) \mapsto P[T \in A|(r(t), m(t))_{t\geq 0} = (r, m)]$. The prospective reserve according to Equation (4) is now interpreted as the conditional expectation:

$$V(t; r, m) = E[A(t)|T > t, (r(t), m(t))_{t\geq 0} = (r, m)]. \tag{5}$$

Moreover, B and c may here be $(\mathcal{F}_t)_{t\geq 0}$-adapted stochastic processes, as well. For a rigorous definition of the integrals in Equation (4), we assume that B is a semimartingale and c, a bounded process. In the following, we work throughout in the doubly-stochastic framework.

Suppose that we are at time s, where the information $\mathcal{F}_s$ is available. The forward interest rate, $(\rho_s(t))_{t\geq s}$, is defined via a zero-coupon bond, i.e., as the $\mathcal{F}_s$-measurable solution of:

$$E\Big[e^{-\int_s^t r(u)\mathrm{d}u}\Big|\mathcal{F}_s\Big] = e^{-\int_s^t \rho_s(u)\mathrm{d}u}, \quad t \geq s \tag{6}$$

where the underlying probability measure represents here some market measure. In the present paper, we suppose throughout that the underlying probability measure is given by the context of the application and that in the following all expectations are calculated on the basis of that probability measure. For our asymptotic approximation results, the origin of the underlying probability measure will not be relevant. In [12,13,16,19], a forward mortality rate $(\mu_s(t))_{t\geq s}$ is defined as the $\mathcal{F}_s$-measurable solution of:

$$E\Big[e^{-\int_s^t m(u)\mathrm{d}u}\Big|\mathcal{F}_s\Big] = e^{-\int_s^t \mu_s(u)\mathrm{d}u}, \quad t \geq s. \tag{7}$$

In [25], we find a spread rate concept where the forward mortality rate is either defined as the solution of:

$$E\Big[e^{-\int_s^t r(u)+m(u)\mathrm{d}u}\Big|\mathcal{F}_s\Big] = e^{-\int_s^t \rho_s(u)+\mu_s(u)\mathrm{d}u}, \quad t \geq s \tag{8}$$

or as the solution of:

$$E\Big[\int_s^t e^{-\int_s^\tau r(u)+m(u)\mathrm{d}u}\, m(\tau)\mathrm{d}\tau\Big|\mathcal{F}_s\Big] = \int_s^t e^{-\int_s^\tau \rho_s(u)+\mu_s(u)\mathrm{d}u}\, \mu_s(\tau)\mathrm{d}\tau, \quad t \geq s. \tag{9}$$

3. Gaussian Diffusion Approximation

In this section, we discuss in some generality the approximation of diffusions with affine drift and a general noise term by corresponding Gaussian diffusions with an affine drift term and additive noise.

Let X be a diffusion process that satisfies the stochastic differential equation

$$\mathrm{d}X(t) = \big(\alpha(t)\,X(t) + \beta(t)\big)\mathrm{d}t + \varsigma\,\sigma(X(t),t)\mathrm{d}W(t)\,, \quad t \geq 0, X(0) = x_0 \in \mathbb{R} \tag{10}$$

where $\varsigma \in [0,1]$ is a scaling parameter and:

(a) the column vector $W(t)$ is a d-dimensional standard Wiener process, $d \in \mathbb{N}$,
(b) the mapping, $(x,t) \mapsto \alpha(x,t)\,x + \beta(t)$, which is composed of the bounded and measurable mappings $\alpha : [0,\omega] \to \mathbb{R}$ and $\beta : [0,\omega] \to \mathbb{R}$, is jointly measurable on $\mathbb{R} \times [0,\omega]$,
(c) the mapping $\sigma : \mathbb{R} \times [0,\omega] \to \mathbb{R}^{1\times d}$ is jointly measurable on $\mathbb{R} \times [0,\omega]$,
(d) there exists a constant $K > 0$, such that

$$\|\sigma(x,t) - \sigma(y,t)\| \leq K|x-y|, \quad \|\sigma(x,t)\|^2 \leq K^2(1+|x|^2)$$

for all $x, y \in \mathbb{R}$, $t \in [0,\omega]$, where $\|\cdot\|$ is the Euclidean norm.

These assumptions are sufficient to ensure that the stochastic differential equation according to (10) has a unique solution with finite moment functions of any order; see Section 4.5 in [26]. As the dt term in

Equation (10) can be seen as an affine mapping of $X(t)$, we say that X has an affine drift term. The scaling parameter, ς, will be the reference parameter for studying asymptotic properties of X. For $\varsigma = 0$, we obtain the differential equation of the expectation function, $E[X(t)]$, of $X(t)$:

$$\mathrm{d}E[X(t)] = \big(\alpha(t)\, E[X(t)] + \beta(t)\big)\mathrm{d}t, \quad t \geq 0, E[X(0)] = x_0. \tag{11}$$

To see that this is indeed the differential equation of the expectation function, integrate Equation (10) from zero to t; take expectation, and then, differentiate again. Apart from a few specific choices of σ, the distribution of X is not analytically tractable. Calculations are much easier if $\sigma(X(t), t)$ is deterministic, since then, X satisfies a stochastic differential equation with so-called additive noise, and X is a Gaussian process, i.e., its finite dimensional projections are multivariate, normally distributed. With respect to the mean squared error, the best deterministic substitute for $\sigma(X(t), t)$ is $E[\sigma(X(t), t)]$, as the following proposition shows.

Proposition 3.1. *Let the process X^* be defined by:*

$$dX^*(t) = \big(\alpha(t)\, X^*(t) + \beta(t)\big)dt + \varsigma\, E[\sigma(X(t), t)]dW(t)\,, \quad t \geq 0, X^*(0) = x_0 \in \mathbb{R} \tag{12}$$

and for a bounded and measurable function $S : [0, \omega] \to \mathbb{R}^{1\times d}$, let Y be a process defined by:

$$dY(t) = \big(\alpha(t)\, Y(t) + \beta(t)\big)dt + \varsigma\, S(t)dW(t)\,, \quad t \geq 0, Y(0) = y_0. \tag{13}$$

Then, for all choices of y_0 and S, we have:

$$E[(X(t) - X^*(t))^2] \leq E[(X(t) - Y(t))^2]. \tag{14}$$

Proof. By applying Ito's formula:

$$\begin{aligned}(X(t) - Y(t))^2 = &(x_0 - y_0)^2 + \int_0^t 2(X(s) - Y(s))\alpha(s)\mathrm{d}s \\ &+ \frac{1}{2}\int_0^t 2\varsigma^2 \|\sigma(X(s), s) - S(s)\|^2 \mathrm{d}s + \int_0^t ...\mathrm{d}W(s)\end{aligned} \tag{15}$$

and taking expectations on both sides, we obtain:

$$\begin{aligned}E[(X(t) - Y(t))^2] = &(x_0 - y_0)^2 + \int_0^t 2E[(X(s) - Y(s))^2]\alpha(s)\mathrm{d}s \\ &+ \int_0^t \varsigma^2 E[\|\sigma(X(s), s) - S(s)]\|^2]\mathrm{d}s.\end{aligned} \tag{16}$$

The latter integral equation has the unique solution:

$$E[(X(t) - Y(t))^2] = e^{\int_0^t 2\alpha(s)\mathrm{d}s}\Big((x_0 - y_0)^2 + \int_0^t e^{-\int_0^s 2\alpha(u)\mathrm{d}u}\varsigma^2 E[\|\sigma(X(s), s) - S(s)]\|^2]\mathrm{d}s\Big). \tag{17}$$

The right-hand side is minimal if x_0 equals y_0 and $S(s)$ equals $E[\sigma(X(s), s)]$, since:

$$E[\|\sigma(X(s), s) - S(s)]\|^2] = E[\|\sigma(X(s), s) - E[\sigma(X(s), s)]\|^2] + \|E[\sigma(X(s), s)] - S(s)\|^2. \tag{18}$$

□

In the insurance applications that we have in mind, the distribution of X is difficult to get, and so, we also have difficulties in calculating $E[\sigma(X(t),t)]$. Therefore, we suggest approximating X by substituting $\sigma(X(t),t)$ with $\sigma(E[X(t)],t)$ instead of $E[\sigma(X(t),t)]$, which leads to a Gaussian diffusion $\widetilde{X}$, defined by:

$$\mathrm{d}\widetilde{X}(t) = \big(\alpha(t)\,\widetilde{X}(t) + \beta(t)\big)\mathrm{d}t + \varsigma\,\sigma(E[X(t)],t)\mathrm{d}W(t)\,, \quad t \geq 0, \widetilde{X}(0) = x_0. \tag{19}$$

Note that the expectation functions of $\widetilde{X}$ and X are equal and that they can be easily obtained from Equation (11). If the difference between $X(t)$ and $E[X(t)]$ is small, then $\widetilde{X}$ should be a good approximation for X. The fluctuation of $X(t) - E[X(t)]$ is small whenever ς is small, because a small ς means that we have only little noise. The following theorem gives an asymptotic result on the absolute moments of the approximation error $X(t) - \widetilde{X}(t)$.

Theorem 3.2. *Assume that the properties (a) to (d) hold. Then, for each $(k,l) \in \mathbb{N}_0^2$, there exists a constant $C_{kl} < \infty$ in such a way that:*

$$\mathbb{E}\big[|X(t) - \widetilde{X}(t)|^k |X(t) - E[X(t)]|^l\big] \leq C_{k,l}\,\varsigma^{2k+l}\,, \qquad \textit{for all } t \in [0,\omega], 0 \leq \varsigma \leq 1. \tag{20}$$

For the proof, see Section 7. From this theorem, we learn that approximating $X(t)$ by its expectation function, $E[X(t)]$, leads to a mean absolute error of order ς^1, and approximating $X(t)$ by the corresponding Gaussian process, $\widetilde{X}(t)$, leads to a mean absolute error of order ς^2. The same is true for the n-th roots of the n-th order moments for $n \in \mathbb{N}$. The following example shows that the order of convergence cannot be generally improved.

Example 3.3. Let $\alpha(t) = 1/2$ and $\sigma(t,x) = x$. Following the ideas of the proof of Theorem 3.2, we can show that:

$$\mathbb{E}\big[|X(t) - \widetilde{X}(t)|^2\big] = \int_0^t \mathbb{E}\big[|X(s) - \widetilde{X}(s)|^2\big]\mathrm{d}s + \varsigma^2 \int_0^t \mathbb{E}\big[|X(s) - \widehat{X}(s)|^2\big]. \tag{21}$$

This integral equation has the unique solution:

$$\mathbb{E}\big[|X(t) - \widetilde{X}(t)|^2\big] = \varsigma^2 \int_0^t e^{(t-s)}\mathbb{E}\big[|X(s) - \widehat{X}(s)|^2\big]\mathrm{d}s. \tag{22}$$

The integral on the right-hand side is positive for each t, so the left-hand side has exactly the order ς^2.

A natural question is whether approximating X by X^* instead of $\widetilde{X}$ leads to a faster convergence than ς^2. In the previous example, X^* and $\widetilde{X}$ are equivalent, so order ς^2 also applies for X^*. Hence, in general, the optimal approximation X^* does not converge faster than the approximation $\widetilde{X}$.

All in all, we can conclude that in the case that interest and mortality rates are diffusion processes of the form of Equation (10) and the noise term is 'small', the corresponding Gaussian diffusion, Equation (19), approximates the moment functions quite well. In the proof of Theorem 3.2, the estimates for the constants $C_{k,l}$ increase with an increasing time interval $[0,\omega]$, function α, Lipschitz constant K and order k of the absolute moments. The following Corollary states some bounds explicitly.

Corollary 3.4. *Under the assumptions of Theorem 3.2, for all $t \in [0, \omega]$ and $\varsigma \in [0, 1]$ we have*

$$\mathbb{E}\big[|X(t) - E[X(t)]|^2\big] \leq \varsigma^2\, 2t\, \exp\big\{2t(K^2 + \sup_{s\in[0,t]} |\alpha(s)|\,)\big\} \sup_{s\in[0,t]} \|\sigma(E[X(s)], s)\|^2, \tag{23}$$

$$\mathbb{E}\big[|X(t) - \widetilde{X}(t)|^2\big] \leq \varsigma^4\, 2K^2t^2\ \exp\big\{2t(K^2 + 2\sup_{s\in[0,t]} |\alpha(s)|\,)\big\} \sup_{s\in[0,t]} \|\sigma(E[X(s)], s)\|^2. \tag{24}$$

For the proof, see the details of the proof of Theorem 3.2.

4. Affine Diffusion Approximation

In the previous section, we simplified the diffusion process X by making the factor $\sigma(X(t), t)$ in the noise term of X deterministic. In this section, we study more general approximations by substituting $\sigma(X(t), t)$ with the first-order Taylor approximation:

$$\sigma^{lin}(X(t), t) := \sigma(E[X(t)], t) + (X(t) - E[X(t)])\, \partial_x \sigma(E[X(t)], t). \tag{25}$$

Given that the partial differential $\partial_x \sigma$ exists, we end up with the diffusion process:

$$\mathrm{d}\widehat{X}(t) = \big(\alpha(t)\, \widehat{X}(t) + \beta(t)\big)\mathrm{d}t + \varsigma\, \sigma^{lin}(\widehat{X}(t), t)\mathrm{d}W(t)\,, \quad t \geq 0, \widehat{X}(0) = x_0. \tag{26}$$

As σ^{lin} is an affine mapping in $X(t)$, we say that $\widehat{X}$ has an affine noise term, and we call $\widehat{X}$ an affine diffusion approximation of X (recall that the drift term of $\widehat{X}$ is also affine). Note that the expectation functions of $\widehat{X}$ and X are equal and that they can be easily obtained from Equation (11). The following theorem shows that $\widehat{X}$ approximates X faster than $\widetilde{X}$ approximates X.

Theorem 4.1. *Under the assumptions (a) to (d) in Section 3 and given that $\partial_{xx}\sigma(x, t)$ exists and is bounded by K, for each $(k, l) \in \mathbb{N}_0^2$, there exists a constant, $C_{kl} < \infty$, in such a way that:*

$$\mathbb{E}\big[|X(t) - \widehat{X}(t)|^k |\widehat{X}(t) - E[X(t)]|^l\big] \leq C_{k,l}\, \varsigma^{3k+l}\,, \qquad \textit{for all } t \in [0, \omega], 0 \leq \varsigma \leq 1. \tag{27}$$

For the proof, see Section 7. According to the previous section, approximating $X(t)$ by the Gaussian process $\widetilde{X}(t)$ leads to a mean absolute error of order ς^2. Theorem 4.1 shows that approximating $X(t)$ by $\widehat{X}(t)$ leads to a mean absolute error of order ς^3. The same is true for the n-th roots of the n-th order moments for $n \in \mathbb{N}$.

Corollary 4.2. *Under the assumptions of Theorem 4.1, for all $t \in [0, \omega]$ and $\varsigma \in [0, 1]$, we have:*

$$\mathbb{E}\big[|\widehat{X}(t) - E[X(t)]|^2\big] \leq \varsigma^2\, 2t\, \exp\big\{2t(K^2 + \sup_{s\in[0,t]} |\alpha(s)|\,)\big\} \sup_{s\in[0,t]} \|\sigma^{lin}(E[X(s)], s)\|^2, \tag{28}$$

$$\mathbb{E}\big[|\widehat{X}(t) - E[X(t)]|^4\big] \leq \varsigma^4\, 24t^2\, \exp\big\{2t(7K^2 + 3\sup_{s\in[0,t]} |\alpha(s)|\,)\big\} \sup_{s\in[0,t]} \|\sigma^{lin}(E[X(s)], s)\|^4, \tag{29}$$

$$\mathbb{E}\big[|X(t) - \widehat{X}(t)|^2\big] \leq \varsigma^6\, 12K^2t^2\, \exp\big\{2t(8K^2 + 4\sup_{s\in[0,t]} |\alpha(s)|\,)\big\} \sup_{s\in[0,t]} \|\sigma^{lin}(E[X(s)], s)\|^4. \tag{30}$$

While $\widetilde{X}$ has a Gaussian distribution whose mean and variance can be easily calculated, the distribution of $\widehat{X}$ must often be calculated numerically. In applications, the user has to decide decide between the high speed of convergence and analytical tractability.

5. Approximation Error When Valuating Insurance Claims

Suppose that the spot rates r and m and the forward rates ρ_s and μ_s are diffusion processes of the form of Equation (10) and that we approximate them by Gaussian or affine diffusions. Then, the approximation error for the corresponding moment functions can be described with the help of the results in the two previous sections. However, in insurance practice, it is rather important to know the approximation error that we obtain in the formulas (4) to (9). All of those valuation formulas are based on:

$$e^{-\int_s^t X(u)\mathrm{d}u} \quad \text{and} \quad e^{-\int_s^t X(u)\mathrm{d}u} X(t). \tag{31}$$

Therefore, we now study the mean approximation error in these two terms if the Gaussian and the affine diffusion approximation are used.

Theorem 5.1. *Assume that X satisfies (10) with properties (a) to (d) and that $X(t)$ is nonnegative for all $t \geq 0$. Then, there exists a constant, C, such that:*

$$E\left[\left|e^{-\int_s^t X(u)du} - e^{-\int_s^t E[X(u)]du}\right|\right] \leq C\varsigma^1, \tag{32}$$

$$E\left[\left|e^{-\int_s^t X(u)du} - e^{-\int_s^t \widetilde{X}(u)du}\right|\right] \leq C\varsigma^2 \tag{33}$$

and:

$$E\left[\left|e^{-\int_s^t X(u)du}X(t) - e^{-\int_s^t E[X(u)]du}E[X(t)]\right|\right] \leq C\varsigma^1, \tag{34}$$

$$E\left[\left|e^{-\int_s^t X(u)du}X(t) - e^{-\int_s^t \widetilde{X}(u)du}\widetilde{X}(t)\right|\right] \leq C\varsigma^2 \tag{35}$$

for all $t \in [0, \omega]$. If, additionally, $E[\exp\{|2\int_s^t \widehat{X}(u)du|\}]$ is finite, then we also have:

$$E\left[\left|e^{-\int_s^t X(u)du} - e^{-\int_s^t \widehat{X}(u)du}\right|\right] \leq C\varsigma^3, \tag{36}$$

$$E\left[\left|e^{-\int_s^t X(u)du}X(t) - e^{-\int_s^t \widehat{X}(u)du}\widehat{X}(t)\right|\right] \leq C\varsigma^3 \tag{37}$$

for all $t \in [0, \omega]$.

For the proof, see Section 7. From this theorem, we learn that the Gaussian and the affine diffusion approximation lead to valuation errors of order ς^2 and ς^3 when valuating typical insurance claims. Approximating X by its expectation function is also an option, but the speed of convergence is just ς^1.

Remark 5.2. In the present paper, we only study two-state life insurance with the states 'active' and 'dead'. However, our results can be easily generalized to multi-state life insurance whenever the corresponding transition probabilities have explicit representations with building blocks of the form of (31). This is typically the case for Markovian state processes with a hierarchical transition matrix, i.e., there are no recurrent states.

6. Numeric Illustration

We now give a numeric example that illustrates the asymptotics of the Gaussian and the affine diffusion approximation. Let X be a diffusion process that satisfies the stochastic differential equation:

$$\mathrm{d}X(t) = \big(-0.05\,X(t) + 0.2\big)\mathrm{d}t + \varsigma\,\arctan(X(t))\mathrm{d}W(t), \quad X(0) = 1. \tag{38}$$

Since X starts at $x_0 = 1$ and has a positive drift whenever $X(t) \leq 4$ and $\lim_{x\downarrow 0}\sigma(t,x) = 0$, the process X is always nonnegative. The corresponding expectation function equals:

$$E[X(t)] = 4 - 3e^{-0.05\,t}, \quad t \geq 0. \tag{39}$$

Figure 1. $\sigma(t,x)$ in green, $\sigma(t,E[X(t)])$ in red and $\sigma^{lin}(t,x)$ in blue for $t = 0$ (**left**) and $t = 10$ (**right**).

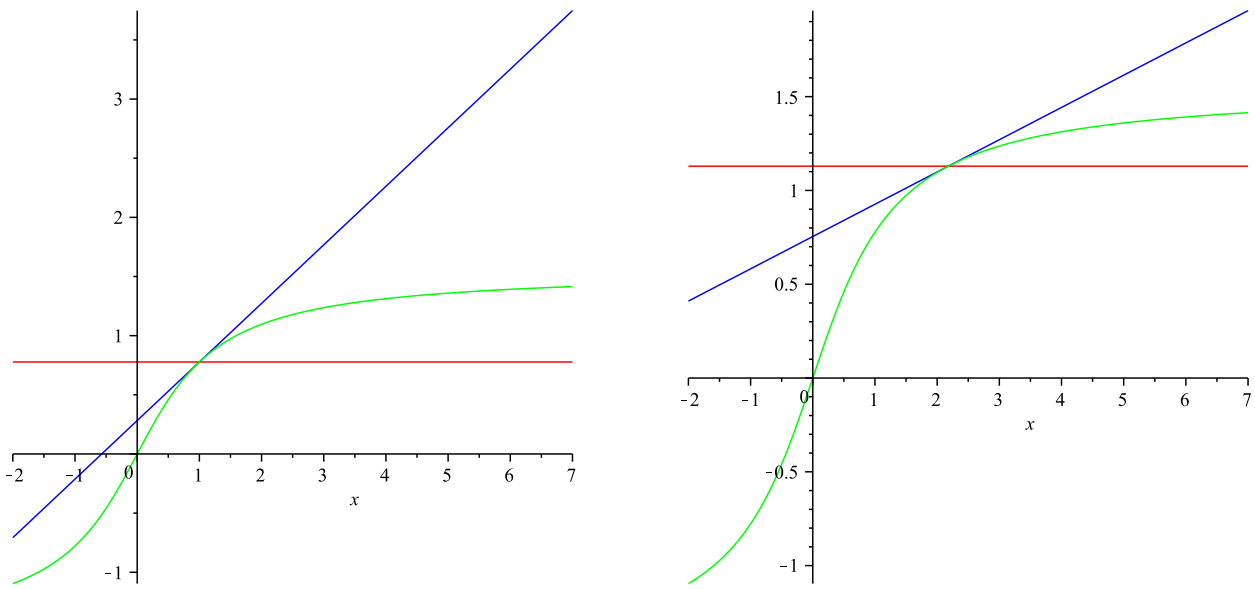

Figure 1 shows the mappings $x \mapsto \sigma(t,x)$ (green curve), $x \mapsto \sigma(t,E[X(t)])$ (red curve) and $x \mapsto \sigma^{lin}(t,x)$, for $t = 0$ and $t = 10$. We see that $\sigma(t,E[X(t)])$ and $\sigma^{lin}(t,x)$ approximate $\sigma(t,x)$ only for x close to $E[X(t)]$ quite well. Figure 2 shows a simulated path of X (green curve), the Gaussian approximation, $\widetilde{X}$ (red curve) and the affine diffusion approximation $\widehat{X}$ (blue curve). For all three processes, we used the same random numbers; that is, all paths correspond to the same path of the Brownian motion, W. The smaller ς, the closer are the processes to the expectation function, $t \mapsto E[X(t)]$. Figure 3 shows the relative differences to the reference process X. Recall that the mean absolute difference between $X(t)$ and $E[X(t)]$ is of order ς^1, that the mean absolute difference between $X(t)$ and $\widetilde{X}(t)$ is of order ς^2 and that the mean absolute difference between $X(t)$ and $\widehat{X}(t)$ is of order ς^3. Indeed, $E[X(t)]$ has the greatest approximation error here. For $\varsigma = 1$ and $\varsigma = 0.75$, the errors of the Gaussian and the affine diffusion approximations have about the same magnitude, but for $\varsigma = 0.25$, we can see a clear advantage for the affine approximation, because of its faster speed of convergence. Table 1 shows the corresponding value of:

$$e^{-0.01\int_0^{20} X(u)\mathrm{d}u} \tag{40}$$

and the corresponding values of the expectation approximation, Gaussian approximation and affine approximation:

$$e^{-0.01\int_0^{20} E[X(u)]\mathrm{d}u}, \quad e^{-0.01\int_0^{20} \widetilde{X}(u)\mathrm{d}u}, \quad e^{-0.01\int_0^{20} \widehat{X}(u)\mathrm{d}u}. \tag{41}$$

The factor 0.01 was added, because interest and mortality rates typically have magnitudes that are in the percent range. The expectation approximation clearly produces the greatest error for all sizes of ς. Interestingly, already for $\varsigma = 0.75$, the error of the Gaussian and the affine approximation is quite small. As predicted by the theorems in the previous sections, the smaller ς, the more does the faster convergence speed of $\widehat{X}$ prevail and the greater is indeed the advantage of the affine approximation.

Figure 2. Expectation function $E[X(t)]$ (black) and simulated paths of $X(t)$ (green), $\widetilde{X}(t)$ (red) and $\widehat{X}(t)$ (blue) for $\varsigma = 1$ (**top left**), $\varsigma = 0.75$ (**top right**), $\varsigma = 0.5$ (**bottom left**), $\varsigma = 0.25$ (**bottom right**).

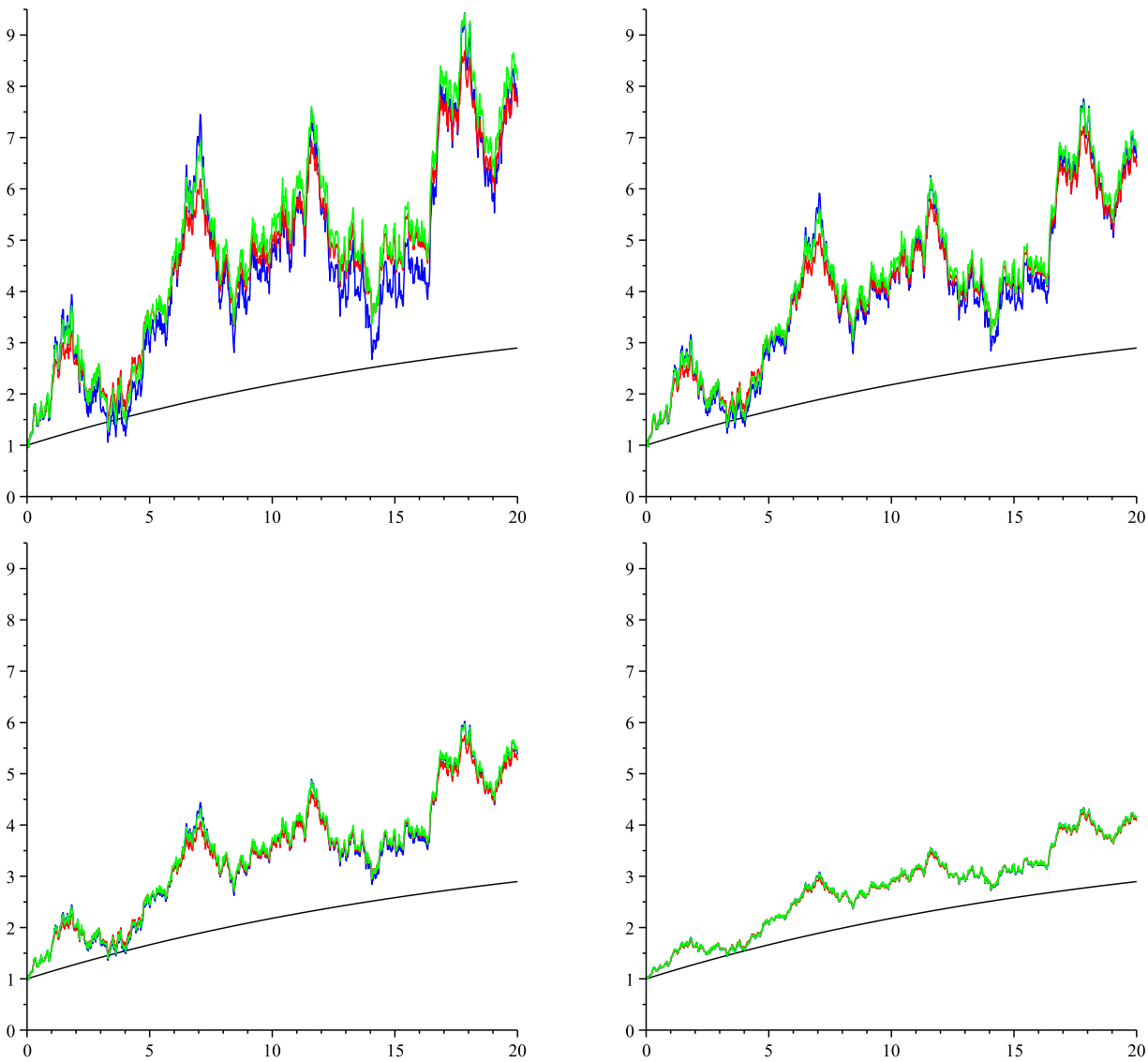

Table 2 shows Monte Carlo simulations for the expectations of (40) and (41) and for the mean absolute deviation of (41) from (40). We used 10,000 simulations. The affine approximation has the smallest

mean absolute deviations, followed by the Gaussian approximation. This conforms with our theoretical results from the previous section. Already for $\varsigma = 0.75$ the Gaussian and the affine approximation for the expectation of (40) work quite well. Surprisingly, for $\varsigma = 0.5, 0.25$ the Gaussian approximation seems to be closer to the true value than the affine approximation, but that may well be a random effect from the Monte Carlo simulation.

Figure 3. Relative difference of $E[X(t)]$ (black), $\widetilde{X}(t)$ (red) and $\widehat{X}(t)$ (blue) with respect to $X(t)$ for $\varsigma = 1$ (**top left**), $\varsigma = 0.75$ (**top right**), $\varsigma = 0.5$ (**bottom left**), $\varsigma = 0.25$ (**bottom right**).

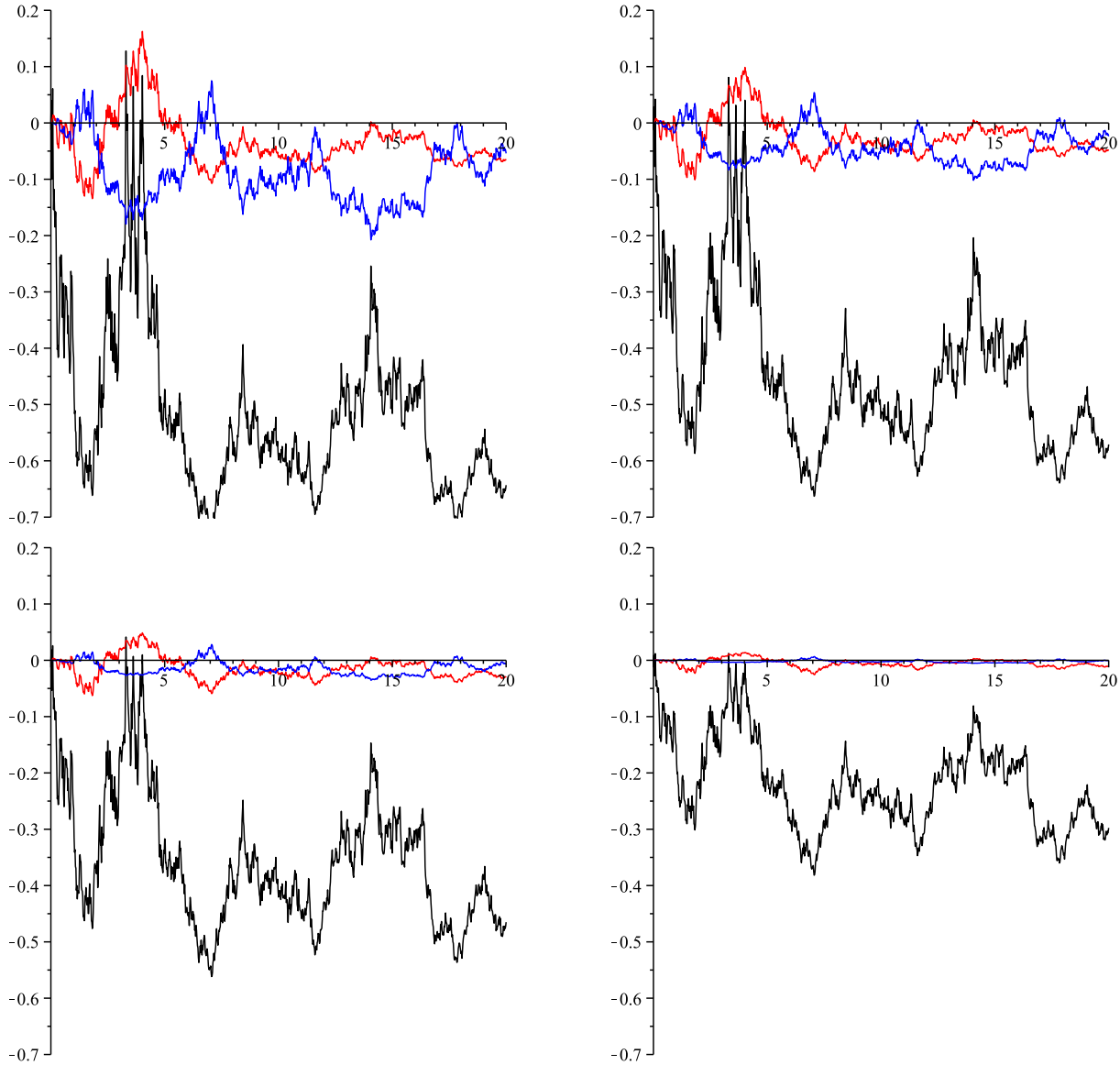

Table 1. One-path simulation for (40) and its corresponding expectation approximation, Gaussian approximation and affine approximation.

Valuation w.r.t. the Single Simulated Path	$\varsigma = 1$	$\varsigma = 0.75$	$\varsigma = 0.5$	$\varsigma = 0.25$
true value	0.37988	0.43749	0.50274	0.57575
expectation approximation	0.65644	0.65644	0.65644	0.65644
(relative error)	0.72804	0.50046	0.30573	0.14016
Gaussian approximation	0.39773	0.44968	0.50926	0.57768
(relative error)	0.04700	0.02785	0.01296	0.00336
affine approximation	0.41127	0.45185	0.50731	0.57635
(relative error)	0.08264	0.03283	0.00910	0.00105

Table 2. Monte Carlo simulation for the expectations of (40) and (41) and for the mean absolute deviation of (41) from (40).

Valuation w.r.t. 10,000 Simulations	$\varsigma = 1$	$\varsigma = 0.75$	$\varsigma = 0.5$	$\varsigma = 0.25$
true value	0.68733	0.67571	0.66712	0.65827
expectation approximation	0.65644	0.65644	0.65644	0.65644
(mean absolute deviation)	0.16151	0.13076	0.09186	0.04839
Gaussian approximation	0.70483	0.68155	0.66873	0.65835
(mean absolute deviation)	0.10065	0.05486	0.02272	0.00500
linear noise approximation	0.70280	0.68152	0.66875	0.65839
(mean absolute deviation)	0.04802	0.02486	0.00863	0.00113

7. Proofs

Proof of Theorem 3.2. For a shorter notation, define:

$$E_t^{k,l} := \mathbb{E}\big[|X(t) - \widetilde{X}(t)|^k |X(t) - E[X(t)]|^l\big] \tag{42}$$

and $\overline{X}_t := E[X(t)]$. The twice continuously differentiable function:

$$g_\varepsilon : \mathbb{R} \to (0, \infty)\,, \quad x \mapsto \begin{cases} |x| & \text{for } |x| \geq \varepsilon \\ -\frac{1}{8\varepsilon^3}x^4 + \frac{3}{4\varepsilon}x^2 + \frac{3\varepsilon}{8} & \text{for } |x| < \varepsilon \end{cases} \tag{43}$$

is for any $\varepsilon > 0$ a majorant of $x \mapsto |x|$ on $\mathbb{R}$. Thus, property (20) holds if for an $\varepsilon > 0$, there exists a constant, $C_{k,l,\varepsilon} < \infty$, with:

$$E_t^{k,l} \leq E_t^{k,l,\varepsilon} := \mathbb{E}\big[g_\varepsilon\big((X(t) - \widetilde{X}(t))^k (X(t) - \overline{X}(t))^l\big)\big] \leq C_{k,l,\varepsilon}\, \varsigma^{2k+l}\,, \qquad \forall\, t \in [0, \omega]. \tag{44}$$

In contrast to $x \mapsto |x|$, the function g_ε is differentiable at zero, which allows us to apply Ito's formula. We prove the theorem by induction. For $(k, l) = (0, 0)$, we can choose $C_{0,0} = 1$. The sequence of

induction steps is tricky here. At first, induction arguments are presented; then, we explain the sequence of induction steps. In order to avoid case differentiations, we define $< z > := \max\{z, 0\}$ and $E_t^{k,l} := 0$ for all $(k, l) \in \mathbb{Z}^2 \setminus \mathbb{N}_0^2$. Applying Ito's formula, we get:

$$
\begin{aligned}
& g_\varepsilon\big((X(t) - \widetilde{X}(t))^k (X(t) - \overline{X}(t))^l\big) - g_\varepsilon(0) \\
&= \int_0^t g_\varepsilon'\big((X(s) - \widetilde{X}(s))^k (X(s) - \overline{X}(s))^l\big)(k+l)\,(X(s) - \widetilde{X}(s))^k (X(s) - \overline{X}(s))^l\,\alpha(s)\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \varsigma^2 g_\varepsilon'\big((X(s) - \widetilde{X}(s))^k (X(s) - \overline{X}(s))^l\big) \\
&\qquad \times \Big(k(k-1)(X(s) - \widetilde{X}(s))^{<k-2>}(X(s) - \overline{X}(s))^l \left\|\sigma(X(s), s) - \sigma(\overline{X}(s), s)\right\|^2 \\
&\qquad\quad + 2kl(X(s) - \widetilde{X}(s))^{<k-1>}(X(s) - \overline{X}(s))^{<l-1>}\big(\sigma(X(s), s) - \sigma(\overline{X}(s), s)\big)\sigma(X(s), s)^T \\
&\qquad\quad + l(l-1)(X(s) - \widetilde{X}(s))^k (X(s) - \overline{X}(s))^{<l-2>}\left\|\sigma(X(s), s)\right\|^2\Big)\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \varsigma^2 g_\varepsilon''\big((X(s) - \widetilde{X}(s))^k (X(s) - \overline{X}(s))^l\big) \\
&\qquad \times \Big(k^2(X(s) - \widetilde{X}(s))^{2<k-1>}(X(s) - \overline{X}(s))^{2l}\left\|\sigma(X(s), s) - \sigma(\overline{X}(s), s)\right\|^2 \\
&\qquad\quad + 2kl(X(s) - \widetilde{X}(s))^{<2k-1>}(X(s) - \overline{X}(s))^{<2l-1>}\big(\sigma(X(s), s) - \sigma(\overline{X}(s), s)\big)\sigma(X(s), s)^T \\
&\qquad\quad + l^2(X(s) - \widetilde{X}(s))^{2k}(X(s) - \overline{X}(s))^{2<l-1>}\left\|\sigma(X(s), s)\right\|^2\Big)\mathrm{d}s \\
&\quad + \int_0^t \ldots \mathrm{d}W_s.
\end{aligned}
\tag{45}
$$

Taking expectations on both sides makes the last integral zero, since its integrand is square integrable (to see that, apply Theorem 4.5.4 in [26] and Hölders inequality). Since:

- we have $-1 \leq g_\varepsilon'(x) \leq 1$ and $0 \leq x\, g_\varepsilon''(x) \leq 3/2$ for all $x \in \mathbb{R}$,
- the continuous function, α, has a finite bound, $|\alpha(t)| \leq \alpha_{max}$, on $[0, \omega]$,
- the mapping, $\sigma(x, t)$, is Lipschitz-continuous in x with Lipschitz constant K,
- $\|\sigma(X(s), s)\| \leq \|\sigma(X(s), s) - \sigma(\overline{X}(s), s)\| + \|\sigma(\overline{X}(s), s)\| \leq K|X(s) - \overline{X}(s)| + \|\sigma(\overline{X}(s), s)\|$ for all s,
- $\|\sigma(X(s), s)\|^2 \leq 2K^2|X(s) - \overline{X}(s)|^2 + 2\|\sigma(\overline{X}(s), s)\|^2$ for all s,
- the continuous function $t \mapsto \|\sigma(\overline{X}(t), t)\|$ has a finite bound σ^{max} on $[0, \omega]$,

the expectation of the left-hand side has for $k \neq 1$ and $l \neq 1$ an upper bound of:

$$
\begin{aligned}
& \mathbb{E}\big[g_\varepsilon\big((X(t) - \widetilde{X}(t))^k (X(t) - \overline{X}(t))^l\big)\big] - \frac{3}{8}\varepsilon \\
&\leq \int_0^t (k+l)\, E_s^{k,l}\, \alpha_{max}\mathrm{d}s \\
&\quad + \frac{\varsigma^2}{2}\int_0^t \Big(k(k-1)E_s^{k-2,l+2}K^2 + 2kl\, E_s^{k-1,l+1}K^2 + 2kl\, E_s^{k-1,l}K\sigma_{max} \\
&\qquad\qquad + 2l(l-1)E_s^{k,l}K^2 + 2l(l-1)E_s^{k,l-2}\sigma_{max}^2\Big)\mathrm{d}s
\end{aligned}
$$

$$+\frac{\varsigma^2}{2}\int_0^t \frac{3}{2}\Big(k^2 E_s^{k-2,l+2}K^2 + 2kl\, E_s^{k-1,l+1}K^2 + 2kl\, E_s^{k-1,l}K\,\sigma_{max} + 2l^2 E_s^{k,l}K^2 + 2l^2 E_s^{k,l-2}\sigma_{max}^2\Big)\mathrm{d}s. \tag{46}$$

Hence, since we assumed that $\varsigma \leq 1$, we obtain:

$$E_t^{k,l} \leq E_t^{k,l,\varepsilon} \leq \frac{3}{8}\varepsilon + \int_0^t \Big(\text{const}\, E_s^{k,l} + \text{const}\,\varsigma^2\Big(E^{k-2,l+2} + E^{k-1,l+1} + E^{k-1,l} + E^{k,l-2}\Big)\Big)\mathrm{d}s. \tag{47}$$

If inequality (20) holds for $\{(k-2,l+2),(k-1,l+1),(k-1,l),(k,l-2)\}$, then we can show that:

$$E_t^{k,l} \leq \frac{3}{8}\varepsilon + \int_0^t \text{const}\, E_s^{k,l}\mathrm{d}s + \text{const}\,\varsigma^2\Big(\varsigma^{2k-4+l+2} + \varsigma^{2k-2+l+1} + \varsigma^{2k-2+l} + \varsigma^{2k+l-2}\Big). \tag{48}$$

For $\varepsilon \to 0$ and with Gronwall's inequality (*cf.* Lemma 4.5.1 in [26]), we get:

$$E_t^{k,l} \leq \text{const}\,\varsigma^2\Big(\varsigma^{2k-4+l+2} + \varsigma^{2k-2+l+1} + \varsigma^{2k-2+l} + \varsigma^{2k+l-2}\Big) \leq \text{const}\,\varsigma^{2k+l} \tag{49}$$

for all $(k,l) \in \mathbb{N}_0^2$ with $k \neq 1$ and $l \neq 1$.

Let now $k = 1$ and/or $l = 1$. Then, analogously to the above, but additionally using that $0 \leq g_\varepsilon''(x) \leq 3/(2\varepsilon)$ for $x \in \mathbb{R}$, we can show that:

$$E_t^{k,l} \leq E_t^{k,l,\varepsilon} \leq \frac{3}{8}\varepsilon + \int_0^t \text{const}\, E_s^{k,l}\mathrm{d}s + \text{const}\,\varsigma^2\Big(\varsigma^{2k-4+l+2} + \varsigma^{2k-2+l+1} + \varsigma^{2k-2+l} + \varsigma^{2k+l-2} + \frac{1}{\varepsilon}\delta_{k1}\,\varsigma^{2l+2} + \frac{1}{\varepsilon}\delta_{l1}\,\varsigma^{4k}\Big) \tag{50}$$

if inequality (20) holds for $\{(k-2,l+2),(k-1,l+1),(k-1,l),(k,l-2)\}$, where δ_{ij} is defined as one if i equals j and zero, else. Setting $\varepsilon = \varsigma^{2k+l}$ and applying Gronwall's inequality, we can conclude that:

$$E_t^{k,l} \leq \text{const}\,\varsigma^2\Big(\varsigma^{2k+l-2} + \varsigma^{2k+l-1}\Big) \leq \text{const}\,\varsigma^{2k+l} \tag{51}$$

whenever $k = 1$ and/or $l = 1$.

By using Inequalities (49) and (51), we do the induction steps along the lines: $\{(0,0),(0,1),(0,2),...\}$, $\{(1,0),(1,1),(1,2),...\}$, $\{(2,0),(2,1),(2,2),...\}$ and so forth, which ensures that being at (k,l), we already have been at $\{(k-2,l+2),(k-1,l+1),(k-1,l),(k,l-2)\}$. □

Proof of Theorem 4.1. The proof is completely analogous to the proof of Theorem 3.2, apart from the fact that the terms:

$$\|\sigma(X(s),s) - \sigma(\overline{X}(s),s)\|^2, \quad \big(\sigma(X(s),s) - \sigma(\overline{X}(s),s)\big)\sigma(X(s),s)^T, \quad \|\sigma(X(t),t)\|^2 \tag{52}$$

have to be substituted by the terms:

$$\|\sigma(X(t),t) - \sigma^{lin}(\widehat{X}(t),t)\|^2, \quad (\sigma(X(t),t) - \sigma^{lin}(\widehat{X}(t),t))\sigma^{lin}(\widehat{X}(t),t)^T, \quad \|\sigma^{lin}(\widehat{X}(t),t)\|^2 \tag{53}$$

and that the upper estimates for $\|\sigma(X(s),s) - \sigma(\overline{X}(s),s)\|$ and $\|\sigma(X(t),t)\|$ have to be replaced by the upper estimates:

$$\|\sigma(X(t),t) - \sigma^{lin}(\widehat{X}(t),t)\| \leq \|\sigma(X(t),t) - \sigma(\widehat{X}(t),t)\| + \|\sigma(\widehat{X}(t),t) - \sigma^{lin}(\widehat{X}(t),t)\|$$

$$\leq K|X(t) - \widehat{X}(t)| + \frac{1}{2}\|\partial_{xx}\sigma(\xi(t), t)\||\widehat{X}(t) - \overline{X}(t)|^2 \tag{54}$$

$$\leq K|X(t) - \widehat{X}(t)| + \frac{1}{2}K|\widehat{X}(t) - \overline{X}(t)|^2 \tag{55}$$

for $\xi(t)$ between $\overline{X}(t)$ and $\widehat{X}(t)$, and:

$$\|\sigma^{lin}(\widehat{X}(t), t)\| \leq \sup_s \|\sigma(E[X(s)], s)\| + K|\widehat{X}(t) - \overline{X}(t)| \tag{56}$$

since $|\partial_x\sigma(x, t)| \leq K$, because of the Lipschitz continuity of $x \mapsto \sigma(x, t)$. Furthermore, we have to set $\varepsilon = \varsigma^{3k+l}$ instead of $\varepsilon = \varsigma^{2k+l}$. The order of the induction steps stays the same. □

Proof of Theorem 5.1. Taylor's theorem applied on the function $x \mapsto \exp\{-x\}$ yields:

$$\exp\{x_0 + h\} = \exp\{x_0\}\Big(1 + ... + \frac{1}{k!}h^k + \frac{1}{(k+1)!}h^{k+1}\exp\{\xi\}\Big) \tag{57}$$

for $k \in \mathbb{N}$ and a proper ξ between zero and h. Setting $k = 0$, $x_0 = -\int_s^t E[X(u)]\mathrm{d}u$ and $x_0 + h = -\int_s^t X(u)\mathrm{d}u$, we get:

$$E\Big[\Big|e^{-\int_s^t X(u)\mathrm{d}u} - e^{-\int_s^t E[X(u)]\mathrm{d}u}\Big|\Big] \leq e^{-\int_s^t E[X(u)]\mathrm{d}u} E\Big[\Big|e^{\xi}\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big|\Big] \tag{58}$$

for ξ between zero and $-\int_s^t X(u) - E[X(u)]\mathrm{d}u$. Since X is nonnegative, the random variable, ξ, is bounded by $\int_s^t E[X(u)]\mathrm{d}u$. Applying Theorem 3.2, we obtain (32). Applying inequality (32) and using the fact that the continuous function $t \mapsto E[X(t)]$ is bounded on $[0, \omega]$, inequality (34) is equivalent to having an upper bound of order ς^1 for:

$$E\Big[\Big|e^{-\int_s^t X(u)\mathrm{d}u}(X(t) - E[X(t)])\Big|\Big]. \tag{59}$$

Since X is nonnegative, we can bound the previous term by:

$$E\big[|X(t) - E[X(t)]|\big] \leq \mathrm{const}\,\varsigma^1. \tag{60}$$

Thus, we showed inequality (34). Applying Taylor's formula with $k = 1$, we get:

$$e^{-\int_s^t X(u)\mathrm{d}u}e^{\int_s^t E[X(u)]\mathrm{d}u} = 1 - \int_s^t X(u) - E[X(u)]\mathrm{d}u + \frac{1}{2}\Big(\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big)^2 e^{\xi}. \tag{61}$$

Analogously, we can show that:

$$e^{-\int_s^t \widetilde{X}(u)\mathrm{d}u}e^{\int_s^t E[X(u)]\mathrm{d}u} = 1 - \int_s^t \widetilde{X}(u) - E[X(u)]\mathrm{d}u + \frac{1}{2}\Big(\int_s^t \widetilde{X}(u) - E[X(u)]\mathrm{d}u\Big)^2 e^{\widetilde{\xi}} \tag{62}$$

for $\widetilde{\xi}$ not greater than the absolute of $\int_s^t \widetilde{X}(u) - E[X(u)]\mathrm{d}u$. Taking the difference of the two equations, we obtain:

$$E\Big[\Big|e^{-\int_s^t X(u)\mathrm{d}u} - e^{-\int_s^t \widetilde{X}(u)\mathrm{d}u}\Big|\Big]e^{\int_s^t E[X(u)]\mathrm{d}u}$$
$$\leq \int_s^t E[|X(u) - \widetilde{X}(u)|]\mathrm{d}u + \frac{1}{2}E\Big[\Big(\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big)^2 e^{\xi}\Big]$$

$$+\frac{1}{2}E\Big[\Big(\int_s^t \widetilde{X}(u)-E[X(u)]\mathrm{d}u\Big)^2 e^{\widetilde{\xi}}\Big]. \tag{63}$$

According to Theorem 3.2, the first addend on the right-hand side is of order ς^2. Since $\exp\{\xi\}$ has a deterministic upper bound and:

$$E\Big[\Big(\int_s^t X(u)-E[X(u)]\mathrm{d}u\Big)^2\Big] = \int_s^t\int_s^t \mathbb{C}\mathrm{ov}(X(u),X(v))\mathrm{d}u\mathrm{d}v \tag{64}$$

$$\leq \int_s^t\int_s^t \sqrt{\mathbb{V}\mathrm{ar}(X(u))}\sqrt{\mathbb{V}\mathrm{ar}(X(v))}\mathrm{d}u\mathrm{d}v \tag{65}$$

$$\leq \mathrm{const}\,\varsigma^2. \tag{66}$$

The second addend on the right-hand side of inequality (63) is also of order ς^2. Using Hölder's inequality, the third addend on the right-hand side of inequality (63) has an upper bound of:

$$\Big(E\Big[e^{2\widetilde{\xi}}\Big]\Big)^{1/2}\Big(E\Big[\Big(\int_s^t \widetilde{X}(u)-E[X(u)]\mathrm{d}u\Big)^4\Big]\Big)^{1/2}$$

$$\leq \mathrm{const}\Big(\int_s^t\int_s^t\int_s^t\int_s^t (E[|\widetilde{X}(u_1)-E[X(u_1)]|^4])^{1/4}\ldots(E[|\widetilde{X}(u_4)-E[X(u_4)]|^4])^{1/4}\mathrm{d}u_1\ldots\mathrm{d}u_4\Big)^{1/2} \tag{67}$$

$$\leq \mathrm{const}\,\varsigma^2 \tag{68}$$

because of Theorem 4.1. Note that $E[\exp\{2\widetilde{\xi}\}]$ has a finite upper bound, since the absolute of ξ is bounded by the absolute of $\int_s^t \widetilde{X}(u)-E[X(u)]\mathrm{d}u$, where the latter term has a normal distribution, as the process $\widetilde{X}$ has a stochastic differential equation with additive noise. Hence, inequality (33) holds. Applying inequality (33) yields that inequality (35) is equivalent to having an upper bound of order ς^2 for:

$$E\Big[\Big|e^{-\int_s^t X(u)-E[X(u)]\mathrm{d}u}\big(X(t)-E[X(t)]\big)-e^{-\int_s^t \widetilde{X}(u)-E[X(u)]\mathrm{d}u}\big(\widetilde{X}(t)-E[X(t)]\big)\Big|\Big]. \tag{69}$$

With Taylor's approximation for $k=0$, the previous term has an upper bound of:

$$E\Big[\big|X(t)-E[X(t)]\big|\Big|\int_s^t X(u)-E[X(u)]\mathrm{d}u\Big|e^{\xi}\Big]$$

$$+E\Big[\big|\widetilde{X}(t)-E[X(t)]\big|\Big|\int_s^t \widetilde{X}(u)-E[X(u)]\mathrm{d}u\Big|e^{\widetilde{\xi}}\Big]. \tag{70}$$

Using the fact that $\exp\{\xi\}$ has a deterministic upper bound and by applying Hölder's inequality, the first addend has an upper bound of:

$$\big(E\big[\big|X(t)-E[X(t)]\big|^2\big]\big)^{1/2}\Big(E\Big[\Big(\int_s^t X(u)-E[X(u)]\mathrm{d}u\Big)^2\Big]\Big)^{1/2}\mathrm{const}$$

$$\leq (\mathrm{const}\,\varsigma^2)^{1/2}(\mathrm{const}\,\varsigma^2)^{1/2}\mathrm{const} \tag{71}$$

and the second addend has an upper bound of:

$$\big(E\big[\big|\widetilde{X}(t)-E[X(t)]\big|^3\big]\big)^{1/3}\Big(E\Big[\Big(\int_s^t \widetilde{X}(u)-E[X(u)]\mathrm{d}u\Big)^3\Big]\Big)^{1/3}\big(E\big[e^{3\widetilde{\xi}}\big]\big)^{1/3}$$

$$\leq \text{const}\,(\text{const}\,\varsigma^3)^{1/3}(\text{const}\,\varsigma^3)^{1/3}\,\text{const}. \tag{72}$$

Thus, we showed inequality (35). Now, we look at inequality (36), which we prove analogously to inequality (33). Applying Taylor's inequality for $k = 2$, we obtain:

$$\begin{aligned}
&E\Big[\Big|e^{-\int_s^t X(u)\mathrm{d}u} - e^{-\int_s^t \widehat{X}(u)\mathrm{d}u}\Big|\Big]e^{\int_s^t E[X(u)]\mathrm{d}u} \\
&\leq \int_s^t E[|X(u) - \widehat{X}(u)|]\mathrm{d}u + \frac{1}{2}E\Big[\Big|\Big(\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big)^2 - \Big(\int_s^t \widehat{X}(u) - E[X(u)]\mathrm{d}u\Big)^2\Big|\Big] \\
&\quad + \frac{1}{6}E\Big[\Big(\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big)^3 e^{\xi}\Big] + \frac{1}{6}E\Big[\Big(\int_s^t \widetilde{X}(u) - E[X(u)]\mathrm{d}u\Big)^3 e^{\widehat{\xi}}\Big].
\end{aligned} \tag{73}$$

According to Theorem 3.2, the first addend on the right-hand side is of order ς^3. With the help of the Binomial theorem and Hölder's inequality, the second addend has an upper bound of:

$$\begin{aligned}
&\frac{1}{2}E\Big[\Big|\int_s^t X(u) - \widehat{X}(u)\mathrm{d}u\Big|\,\Big|\int_s^t X(u) - E[X(u)] + \widehat{X}(u) - E[X(u)]\mathrm{d}u\Big|\Big] \\
&\leq \frac{1}{2}\Big(E\Big[\Big(\int_s^t X(u) - \widehat{X}(u)\mathrm{d}u\Big)^2\Big]\Big)^{1/2}\Big(E\Big[\Big(\int_s^t X(u) - E[X(u)] + \widehat{X}(u) - E[X(u)]\mathrm{d}u\Big)^2\Big]\Big)^{1/2}.
\end{aligned} \tag{74}$$

The first factor is of order ς^3, while the second factor has a finite upper bound. Similarly to the estimates for the second and third addend on the right-hand side of inequality (63), we can show that the third and fourth addend on the right-hand side of inequality (73) have upper bounds of order ς^3. Thus, we showed inequality (36). Applying inequality (36) yields that inequality (37) is equivalent to having an upper bound of order ς^3 for:

$$E\Big[\Big|e^{-\int_s^t X(u)-E[X(u)]\mathrm{d}u}\big(X(t) - E[X(t)]\big) - e^{-\int_s^t \widehat{X}(u)-E[X(u)]\mathrm{d}u}\big(\widehat{X}(t) - E[X(t)]\big)\Big|\Big]. \tag{75}$$

With Taylor's approximation for $k = 1$, the previous term has an upper bound of:

$$\begin{aligned}
&E[|X(t) - \widehat{X}(t)|] \\
&+ E\Big[\Big|\int_s^t X(u) - E[X(u)]\mathrm{d}u\big(X(t) - E[X(t)]\big) - \int_s^t \widehat{X}(u) - E[X(u)]\mathrm{d}u\big(\widehat{X}(t) - E[X(t)]\big)\Big|\Big] \\
&+ \frac{1}{2}E\Big[\Big(\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big)^2\big|X(t) - E[X(t)]\big|e^{\xi}\Big] \\
&+ \frac{1}{2}E\Big[\Big(\int_s^t \widehat{X}(u) - E[X(u)]\mathrm{d}u\Big)^2\big|\widehat{X}(t) - E[X(t)]\big|e^{\widehat{\xi}}\Big].
\end{aligned} \tag{76}$$

The first addend is of order ς^3. The second addend has an upper bound of:

$$E\Big[\Big|\int_s^t X(u) - E[X(u)]\mathrm{d}u\Big|\big|X(t) - \widehat{X}(t)\big|\Big] + E\Big[\Big|\int_s^t X(u) - \widehat{X}(u)\mathrm{d}u\Big|\big|X(t) - E[X(t)]\big|\Big] \tag{77}$$

which has an order of ς^4. To see that, apply Hölder's inequality, Theorem 3.2 and Theorem 4.1. The fourth and fifth addend of the term above have an order of ς^3, which can be verified similarly to the estimates for the second and third addend on the right-hand side of inequality (63). □

8. Conclusions

Because of nice analytical features, in the actuarial literature, stochastic models for interest and mortality rates often have the form of Gaussian diffusions with an affine drift term and additive noise. The Gaussian diffusion framework is justified with the argument that it would reasonably approximate the true development of interest and mortality rates. We studied approximation errors if appropriate models for interest and mortality rates have the form of diffusions with affine drift, but a general noise term. We calculated theoretical bounds for absolute moments and valuation formulas, showing the speed of convergence. A numerical study illustrates the approximation error, confirming the theoretical results. Our results indicate that approximation errors are reasonable if the noise term is not too large. In particular, that might be the case when modeling the mortality intensity, since demographic changes of mortality typically happen slowly, but steadily.

We generalized the Gaussian diffusion approximation with additive noise by introducing an affine diffusion approximation with an affine noise term. We found that with an affine noise term instead of additive noise, the speed of convergence can be improved from ς^2 to ς^3. However, we lose the nice analytical features of the Gaussian approach.

All in all, we identified a large class of diffusion processes that can be well approximated by Gaussian diffusions with additive noise or by affine diffusions with affine noise terms. As a consequence, we can say that modeling interest and mortality rates by Gaussian diffusions is well justified if there is evidence that the true developments of future interest and mortality rates have the form of diffusion processes with an affine drift term and an arbitrary, but small, noise term.

In the present paper, we measured approximation quality by point-wise moment errors and by mean absolute errors for some specific insurance claims. Future research should consider also other approximation measures.

Acknowledgments

The author thanks two anonymous referees whose comments greatly helped to improve the paper.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Björk, T. *Arbitrage Theory in Continuous Time*, 2nd ed.; Oxford University Press: New York, NY, USA, 2004.
2. Vasicek, O. An equilibrium characterization of the term structure. *J. Financ. Econ.* **1977**, *5*, 177–188.
3. Cox, J.C.; Ingersoll, J.E.; Ross, S.A. A theory of the term structure of interest rates. *Econometrica* **1985**, *53*, 385–407.
4. Heath, D.; Jarrow, R.; Morton, A. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica* **1992**, *60*, 77–105.

5. Beekman, J.A.; Fuelling, C.P. Interest and mortality randomness in some annuities. *Insur. Math. Econ.* **1990**, *9*, 185–196.
6. Beekman, J.A.; Fuelling, C.P. Extra randomness in certain annuity models. *Insur. Math. Econ.* **1991**, *10*, 275–287.
7. Dufresne, D. The distribution of a perpetuity, with application to risk theory and pension funding. *Scand. Actuar. J.* **1990**, *1*, 39–79.
8. De Schepper, A.; de Vijlder, F.; Goovaerts, M.J.; Kaas, R. Interest randomness in annuities certain. *Insur. Math. Econ.* **1992**, *11*, 271–281.
9. Parker, G. Two stochastic approaches for discounting actuarial functions. *ASTIN Bull.* **1994**, *24*, 167–181.
10. Norberg, R.; Møller, C.M. Thieles differential equation with stochastic interest of diffusion type. *Scand. Actuar. J.* **1996**, *1*, 37–49.
11. Persson, S.-A. Stochastic interest rate in life insurance: The principle of equivalence revisited. *Scand. Actuar. J.* **1998**, *2*, 97–112.
12. Milevsky, M.A.; Promislow, S.D. Mortality derivatives and the option to annuitise. *Insur. Math. Econ.* **2001**, *29*, 299–318.
13. Dahl, M.H. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insur. Math. Econ.* **2004**, *35*, 113–136.
14. Biffis, E. Affine processes for dynamic mortality and actuarial valuations. *Insur. Math. Econ.* **2005**, *37*, 443–468.
15. Ballotta, L.; Haberman, S. The fair valuation problem of guaranteed annuity options: The stochastic mortality environment case. *Insur. Math. Econ.* **2006**, *38*, 195–214.
16. Cairns, A.J.; Blake, D.; Dowd, K. Pricing death: Frameworks for the valuation and securitization of mortality risk. *Astin Bull.* **2006**, *36*, 79–120.
17. Dahl, M.; Møller, T. Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insur. Math. Econ.* **2006**, *39*, 193–217.
18. Schrager, D.F. Affine stochastic mortality. *Insur. Math. Econ.* **2006**, *38*, 81–97.
19. Bauer, D.; Börger, M.; Ruß, J.; Zwiesler, H.J. The volatility of mortality. *Asia-Pac. J. Risk Insur.* **2008**, *3*, 184–211.
20. Young, V.R. Pricing life insurance under stochastic mortality via the instantaneous Sharpe ratio. *Insur. Math. Econ.* **2008**, *42*, 691–703.
21. Shang, Z.; Goovaerts, M.; Dhaene, J. A recursive approach to mortality-linked derivative pricing. *Insur. Math. Econ.* **2011**, *49*, 240–248.
22. Huang, H.; Milevsky, M.A.; Salisbury, T.S. Optimal retirement consumption with a stochastic force of mortality. *Insur. Math. Econ.* **2012**, *51*, 282–291.
23. Tappe, S.; Weber, S. Stochastic mortality models: An infinite-dimensional approach. *Finance and Stochastics* (in press).
24. Christiansen, M.C.; Niemeyer, A. On the forward rate concept in multi-state life insurance. Preprint Series of the Faculty of Economics and Mathematics at the University of Ulm: Ulm, Germany, 2013.

25. Miltersen, K.R.; Persson, S.A. Is mortality dead? Stochastic forward force of mortality rate determined by no arbitrage. Working paper, University of Bergen, Norway, 2005.
26. Kloeden, P.E.; Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer: Berlin, Germany, 1992.

© 2013 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).