STOCHASTIC ANALYSIS AND APPLICATIONS
2018, VOL. 36, NO. 2, 224–239
https://doi.org/10.1080/07362994.2017.1386571

A note on a derivation method for SDE models: Applications in biology and viability criteria

Jacky Cresson[a] and Stefanie Sonner[b]

[a]Laboratoire de Mathématiques Appliquées de Pau, Université de Pau et des Pays de l'Adour, Pau Cedex, France;
[b]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Graz, Austria

ABSTRACT
We discuss a method, which was popularized by E. J. Allen and that is frequently used in applications to construct SDE models. The derivation procedure is based on information about the elementary processes involved in the dynamics and their corresponding probabilities. We formulate criteria for the viability of the resulting models. In particular, explicit necessary and sufficient conditions are deduced for the non-negativity and/or boundedness of solutions. Moreover, we show that the class of deterministic models for which the construction leads to an admissible SDE extension is strongly limited. Several examples are presented to illustrate the implications of our results.

ARTICLE HISTORY
Received 8 September 2016
Accepted 27 September 2017

KEYWORDS
Stochastic differential equations; model validation; biological applications; Itô equations

MATHEMATICS SUBJECT CLASSIFICATION
60H10; 92B05; 60J28; 65C30

1. Introduction

Stochastic differential equations (SDEs) are frequently used in order to model systems where random effects play a significant role. Since there is no canonical way to formulate an SDE model, different construction methods have been considered. One derivation procedure that is widely used was suggested by E. Allen in [1]. It generalizes a classical method to construct ordinary differential equation (ODE) models. All possible changes δ of the considered stochastic process in a small time interval are determined, together with their corresponding probabilities p. Given this information, a system of Itô SDEs can be derived, where the coefficients in the equations are given in terms of δ and p. This method has been used to develop SDE models in various fields; a large variety of applications in mathematical biology can be found, e.g., in [1, 2]; see also [3–8].

Solutions of models in biological applications typically represent non-negative quantities, such as population densities or concentrations of chemical substrates, and hence, it is essential that they attain non-negative values. Models that do not ensure this property are not valid or break down for small values of the solution. Explicit necessary and sufficient conditions for systems of ODEs are well-known and allow to characterize the class of non-negativity preserving models. For systems of SDEs, explicit criteria have also been obtained, however, they are less known. In fact, various SDE models have been proposed and analyzed in recent years that produce undesired negative values of the solutions [9, 10].

In [9, 10], we formulated explicit necessary and sufficient conditions for the non-negativity of solutions of SDE systems and discussed several modeling applications. We now apply these

CONTACT Stefanie Sonner stefanie.sonner@uni-graz.at Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria.

results to analyze whether Allen's derivation procedure leads to SDE models that preserve non-negativity. In particular, we aim to formulate explicit necessary and sufficient conditions for the possible changes δ and probabilities p that are easy to verify and allow to classify admissible models. We further point out the limitations of the method when SDE extensions are constructed based on a given deterministic ODE model.

The interest of such a result is twofold and goes beyond mathematics.

First, our criteria can be used to distinguish viable or admissible models. This is of course fundamental in applications.

On the other hand, due to the particularity of Allen's derivation procedure, if an unrealistic model is obtained, it has direct implications on the modelling assumptions. Indeed, it is based on information about the underlying stochastic process by analyzing all possible interactions and the corresponding probabilities. This information is typically justified by arguments about the nature of the phenomenon under consideration. If the resulting SDE model is not viable, then

- either the underlying process cannot be decomposed as assumed by Allen's method and is richer,
- or some of the interaction probabilities involved have to be modified and interpreted differently.

In either case, it leads to a deeper understanding of the modeled phenomenon at hand. We give examples of these situations in Section 4 recalling classical SDE models discussed by E. J. Allen in [1] and L. J. S. Allen in [2]. Of course, mathematics cannot *in general* decide which situation applies and this is precisely where the interaction with other fields comes into play. However, one case deserves special attention. Namely, when the method is applied to construct stochastic extensions of deterministic ODE systems, since part of the coefficients in the SDE are already fixed by the deterministic model. A negative result, i.e., obtaining an SDE model that is not viable means precisely that the underlying process is not of the form assumed by the derivation method.

But our results do not only have *negative implications*. If the resulting model is not viable, they provide a pragmatic rule for modifications in order to obtain admissible models. Indeed, as our criteria yield *explicit necessary and sufficient conditions for viability*, they can be used as a guide to reformulate and modify the modeling assumptions. We provide examples for this situation in Section 4.

The paper is organized as follows: In Section 2, we recall Allen's derivation procedure and generalize it for an arbitrary number of interacting populations. In Section 3, we formulate explicit criteria that lead to viable SDE models. Moreover, we analyze the limitations of the method when SDE extensions are constructed from a given deterministic ODE model. Finally, we discuss several examples and modeling applications in Section 4. For the convenience of the reader, the general invariance criterion for SDE systems, which is the basis of our results, is given in the Appendix.

2. A derivation procedure for SDE models

In this section, we recall the modeling procedure proposed by E. Allen in [1] to derive Itô SDE models. In order to facilitate the reading of our results, we will adopt the notations in [1].

In the first step, a discrete stochastic model is developed by determining all possible changes δ of the system and the corresponding transition probabilities p in a small time interval Δt. Then, the expectation value and covariance matrix for the change of the discrete process are calculated, and based on this information the SDE model is formulated. The drift coefficient is hereby given by the expected change divided by Δt and the diffusion coefficient by the square root of the covariance matrix divided by Δt.

Allen's modeling procedure is described for models with two interacting populations in [1, 2]. We generalize it here for systems with an arbitrary number k of components. Let $X(t) = (X_1(t), \dots, X_k(t))^T$ represent the state of k interacting populations at time $t \geq 0$, where we use the superscript T to denote the transpose of a vector (or a matrix).

- **Step 1:** All possible interactions with the environment and between the populations X_i, $i = 1, \dots, k$, that lead to a change in one of the states are listed, together with the corresponding probabilities for the change up to $O((\Delta t)^2)$:
  We assume that there are m possible interactions that lead to a change in at least one of the states. The matrix $\Delta X = (\delta^{(1)}, \dots, \delta^{(m)}) \in \mathbb{R}^{m \times k}$ contains the coefficients of all changes, where
  $$\delta^{(i)} = (\delta_1^{(i)}, \dots, \delta_k^{(i)})^T \in \mathbb{R}^k,$$
  and the corresponding probabilities up to $O((\Delta t)^2)$ are given by
  $$p_i = P_i(t, X_1, \dots, X_k)\Delta t,$$
  $i = 1, \dots, m$. In addition, there is the probability of no change, $\delta^{(m+1)} = (0, \dots, 0)^T$, which is
  $$p_{m+1} = 1 - \sum_{i=1}^{m} p_i.$$
  The hypotheses are summarized in Table 1. Based on this information the coefficients in the SDE model are determined in the subsequent Steps 2 and 3.
- **Step 2:** The expected change $\mathbb{E}(\Delta X)$ of the system is computed,
  $$\mathbb{E}(\Delta X) = \sum_{i=1}^{m} p_i \delta^{(i)},$$
  and the covariance matrix up to $O((\Delta t)^2)$,
  $$\mathbb{E}(\Delta X(\Delta X)^T) = \sum_{i=1}^{m} p_i (\delta^{(i)})(\delta^{(i)})^T.$$
- **Step 3:** We define
  $$\mu = \frac{\mathbb{E}(\Delta X)}{\Delta t} = \sum_{i=1}^{m} P_i \delta^{(i)},$$
  $$V = \frac{\mathbb{E}(\Delta X(\Delta X)^T)}{\Delta t} = \sum_{i=1}^{m} P_i (\delta^{(i)})(\delta^{(i)})^T,$$
  and denote the square root of V by
  $$B(t, X_1, \dots, X_k) = \sqrt{V(t, X_1, \dots, X_k)}.$$
  Finally, the following system of Itô SDEs is considered:
  $$dX(t) = \mu(t, X_1, \dots, X_k)dt + B(t, X_1, \dots, X_k)dW(t), \tag{1}$$
  $$X(t_0) = X_0,$$

Table 1. Changes of the system and corresponding probabilities.

i	$\delta^{(i)}$	p_i
1	$\left(\delta_1^{(1)}, \dots, \delta_k^{(1)}\right)^T$	p_1
$\vdots$	$\vdots$	$\vdots$
m	$\left(\delta_1^{(m)}, \dots, \delta_k^{(m)}\right)^T$	p_m
$m+1$	$(0, \dots, 0)^T$	$1 - \sum_{i=1}^m p_i$

where $W = (W_1, \dots, W_k)^T$ with k independent standard scalar Wiener processes $W_1, \dots, W_k$, and dW denotes the corresponding Itô differential.

For details of the derivation procedure, properties of the resulting SDE models, and applications, we refer to [1, 2].

In order to simplify the presentation of our results, we introduce the following notation.

Definition 2.1. We denote by (δ, p) a couple of data in $\mathbb{R}^{m\times k} \times \mathbb{R}^k$ as in Table 1 that satisfies the properties in Step 1. This information contains the modeling assumptions and is the basis of the constructed model. We call the corresponding system of Itô equations (1) the (δ, p)-SDE model.

Remark 2.1. As indicated in [1, 2], alternative (equivalent) SDE systems can be formulated such that their drift part and covariance matrix coincide with the ones corresponding to (1).

For instance, the system

$$dX(t) = \mu(t, X_1, \dots, X_k)dt + C(t, X_1, \dots, X_k)dW^*(t), \tag{2}$$

where $W^* = (W_1^*, \dots, W_m^*)^T$ with m independent standard scalar Wiener processes W_i^*, and the matrix C is given by

$$C_{ij} = \delta_i^{(j)}\sqrt{\frac{p_j}{\Delta t}} = \delta_i^{(j)}\sqrt{P_j}$$

(see [2] for the case $k = 2$). Then, the matrix C satisfies $CC^T = V$ and the corresponding mean vector, covariance matrix, and forward Kolmogorov differential equation for the SDE systems (1) and (2) coincide.

Another diffusion matrix can be obtained using the Cholesky factorization of V. If G is the lower triangular matrix in this factorization, then G satisfies $GG^T = V$ and yields a further equivalent SDE system (see [2]).

To illustrate the method, we recall the derivation of an SDE model for two interacting populations discussed in [1].

Example 2.1. Let X_1 and X_2 denote the sizes of two populations, b_i and d_i be the corresponding birth and death rates, $i = 1, 2$, and m_{12} and m_{21} the rates at which population 1 is transformed into population 2, and vice versa. Each of these parameters can depend on time t and the population sizes X_1 and X_2.

The following table lists the possible changes δ in the population sizes along with the corresponding probabilities p.

i	$\delta^{(i)}$	p_i
1	$(-1, 0)^T$	$d_1 X_1 \Delta t$
2	$(1, 0)^T$	$b_1 X_1 \Delta t$
3	$(0, -1)^T$	$d_2 X_2 \Delta t$
4	$(0, 1)^T$	$b_2 X_2 \Delta t$
5	$(-1, 1)^T$	$m_{12} X_1 \Delta t$
6	$(1, -1)^T$	$m_{21} X_2 \Delta t$
7	$(0, 0)^T$	$1 - \sum_{i=1}^{6} p_i$

Computing the expected change and covariance matrix, the derivation procedure leads to the following (δ, p)-SDE model

$$dX = \mu(t, X_1, X_2)dt + B(t, X_1, X_2)dW,$$

where $W = (W_1, W_2)^T$ with two independent standard scalar Wiener processes W_1 and W_2. Moreover,

$$\mu = \begin{pmatrix} b_1 X_1 - d_1 X_1 - m_{12} X_1 + m_{21} X_2 \\ b_2 X_2 - d_2 X_2 - m_{21} X_2 + m_{12} X_1 \end{pmatrix}$$

and $B = \sqrt{V}$, where

$$V = \begin{pmatrix} b_1 X_1 + d_1 X_1 + m_{12} X_1 + m_{21} X_2 & -m_{12} X_1 - m_{21} X_2 \\ -m_{12} X_1 - m_{21} X_2 & b_2 X_2 + d_2 X_2 + m_{21} X_2 + m_{12} X_1 \end{pmatrix}.$$

We omitted here the dependence of the coefficients on t, X_1, X_2 in order to shorten notations.

3. Viability criteria and limitations of the method

Our aim is to analyze under which assumptions Allen's derivation procedure leads to viable SDE models. In particular, we formulate explicit necessary and sufficient conditions for the probability functions p_i and changes $\delta_j^{(i)}$ such that the resulting (δ, p)-SDE models preserve non-negativity and/or upper bounds for the solutions. We further point out the limitations of the method when SDE extensions are constructed based on a given deterministic ODE model. The results are derived from a previous invariance criterion for SDE systems formulated in [9, 10] (see Theorem A.1 in the Appendix).

3.1. Criteria for non-negativity

Here and in the sequel, we denote the positive cone in $\mathbb{R}^k$ by

$$K^+ = \left\{ y \in \mathbb{R}^k : y_i \geq 0,\ i = 1, \dots, k \right\}.$$

Definition 3.1. We say that a stochastic system

$$\begin{aligned} dX(t) &= f(t, X_1, \dots, X_k)dt + G(t, X_1, \dots, X_k)dW(t), \\ X(t_0) &= X_0, \end{aligned} \tag{3}$$

where $f : [t_0, \infty) \times \mathbb{R}^k \to \mathbb{R}^k$, $G : [t_0, \infty) \times \mathbb{R}^k \to \mathbb{R}^{k \times m}$, and $W = (W_1, \dots, W_m)^T$, preserves non-negativity if for every initial data $X_0 \in K^+$ and initial time $t_0 \geq 0$ the corresponding solution $X(t), t \geq t_0$, satisfies

$$P\left(\left\{X(t) \in K^+,\ t \in [t_0, \infty)\right\}\right) = 1,$$

i.e., it almost surely attains non-negative values.

Theorem 3.1. *Let (δ, p) be given data as in Step 1 of the model derivation. Then, the (δ, p)-SDE model (1) preserves non-negativity if and only if*

$$P_i(t, y)(\delta_j^{(i)})^2 = 0, \qquad \text{for } y \in K^+ \text{ such that } y_j = 0,\ t \geq 0, \tag{4}$$

for all $i = 1, \dots, m$.

In other words, for all $1 \leq j \leq k$, $1 \leq i \leq m$ such that $\delta_j^{(i)} \neq 0$ it follows that

$$P_i(t, y) = 0 \qquad \text{for } y \in K^+ \text{ such that } y_j = 0,\ t \geq 0.$$

The same result applies to the alternative SDE system (2).

Proof. By Theorem A.1 the SDE system (1) preserves non-negativity if and only if μ and B satisfy

$$\mu_i(t, y) \geq 0, \qquad B_{ij}(t, y) = 0, \qquad \text{for } y \in K^+ \text{ such that } y_i = 0,$$

for all $i = 1, \dots, k$, $j = 1, \dots, m$, and $t \geq 0$. Since $B = \sqrt{V}$, i.e.,

$$V_{ij}(t, y) = \sum_{l=1}^{k} B_{il}(t, y) B_{jl}(t, y), \tag{5}$$

this implies that V satisfies

$$V_{ij}(t, y) = 0, \qquad \text{for } y \in K^+ \text{ such that } y_i = 0,$$

for all $i = 1, \dots, k$, $j = 1, \dots, m$, and $t \geq 0$. We further observe that

$$V(t, y) = \sum_{i=1}^{k} P_i(t, y) \begin{pmatrix} (\delta_1^{(i)})^2 & \delta_1^{(i)}\delta_2^{(i)} & \cdots & \delta_1^{(i)}\delta_k^{(i)} \\ \delta_1^{(i)}\delta_2^{(i)} & (\delta_2^{(i)})^2 & \cdots & \delta_2^{(i)}\delta_k^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_1^{(i)}\delta_k^{(i)} & \delta_2^{(i)}\delta_k^{(i)} & \cdots & (\delta_k^{(i)})^2 \end{pmatrix},$$

which leads to condition (4).

On the other hand, if condition (4) holds, then V satisfies

$$V_{jl}(t, y) = \sum_{i=1}^{k} P_i(t, y)\delta_j^{(i)}\delta_l^{(i)} = 0,$$

for $y \in K^+$ such that either $y_l = 0$ or $y_j = 0$. Using relation (5) for $i = j$, we deduce that

$$V_{ii}(t, y) = \sum_{l=1}^{k} \left(B_{il}(t, y)\right)^2, \tag{6}$$

for any $i = 1, \dots, k$. As $V_{ii}(t, y) = 0$ for $y \in K^+$ such that $y_i = 0$, we conclude that for any $l = 1, \dots, k$, we have

$$B_{il}(t, y) = 0 \qquad \text{for } y \in K^+ \text{ such that } y_i = 0. \tag{7}$$

Moreover, condition (4) certainly implies that

$$\mu_i(t, y) = 0 \qquad \text{for } y \in K^+ \text{ such that } y_i = 0,$$

$i = 1, \dots, k$. As a consequence, the assumptions of Theorem A.1 are satisfied and system (1) preserves non-negativity.

The proof for the alternative SDE model (2) simplifies and follows by the same arguments. □

Remark 3.1. Allen's derivation procedure leads to systems of Itô SDEs, another commonly used concept is Stratonovich's interpretation of SDEs. The behavior of solutions generally depends on the choice of the interpretation, however, there is an explicit transformation formula relating the solutions of both notions (see [11] and the Appendix). It is typically difficult to decide which interpretation is more appropriate in a particular application. A long debate about this controversy has been going on, and different arguments have been given to support either Itós or Stratonovich's interpretation. For detailed discussions and possible resolutions, we refer, e.g., to [11–13].

One could ask whether Theorem 3.1 changes if the SDE models were interpreted in the sense of Stratonovich instead of Itô. In fact, the result is independent of the choice of interpretation, i.e., the necessary and sufficient conditions remain identical if Stratonovich's interpretation is used for the stochastic systems (1) or (2) (see Theorem A.1 in the Appendix).

Example 3.1. We apply Theorem 3.1 to the SDE model for two interacting populations in Example 2.1:

The system preserves non-negativity if and only if m_{12} and m_{21} satisfy $m_{12}(t, X_1, 0) = 0$ and $m_{21}(t, 0, X_2) = 0$, i.e.,

$$m_{12}(t, X_1, X_2) = X_2\tilde{m}_{12}(t, X_1, X_2), \quad m_{21}(t, X_1, X_2) = X_1\tilde{m}_{21}(t, X_1, X_2),$$

for some functions $\tilde{m}_{12}$ and $\tilde{m}_{21}$. All other probabilities certainly fulfil the required conditions.

If the transition rates m_{12} and m_{21} do not comply with these conditions, the solutions of the SDE model can attain undesired negative values.

3.2. Stochastic extensions of deterministic ODE models

Allen's method is often used to construct stochastic extensions of a given deterministic ODE model. In this case, the drift part μ is, of course, already determined. It can be deduced from Theorem 3.1 that for a large class of ODE systems that preserve non-negativity, no SDE extension can be derived by Allen's method that possesses this property.

We begin by observing a particularity of (δ, p)-SDE models.

Proposition 3.1. *Let (δ, P) be given data as in Step 1 of the derivation procedure. If the corresponding (δ, P)-SDE model preserves non-negativity, then the drift term μ satisfies*

$$\mu_j(t, y) = 0 \qquad \text{for } y \in K^+ \text{ such that } y_j = 0,\ t \geq 0,$$

for all $j = 1, \ldots, k$. The same result holds for SDE systems of the form (2).

Proof. According to Theorem 3.1, if an SDE system of the form (1) or (2) preserves non-negativity, then condition (4) holds. Moreover, the drift term is determined by

$$\mu(t, y) = \sum_{i=1}^{m} P_i(t, y)\delta^{(i)},$$

which implies the statement of the proposition. □

This proposition imposes strong constraints on the class of deterministic systems, for which a non-negativity preserving SDE extension can be constructed by Allen's method.

Corollary 3.1. *We assume that*

$$\frac{dX}{dt} = \mu(t, X)$$

is a given deterministic ODE model. If there exist $t \geq 0$, $1 \leq j \leq k$, and $y \in K^+$ with $y_j = 0$ such that

$$\mu_j(t, y) > 0,$$

then, any SDE extension derived by Allen's method does not preserve non-negativity. This applies to SDE systems of the form (1) and (2).

In particular, if one of the interaction functions μ_i contains a constant term, Allen's method leads to an SDE model that produces negative values of the solutions (see Section 4.1 for examples). In this case, the procedure has to be modified or other methods need to be applied in order to construct viable stochastic extensions.

Proof. The drift term is determined by $\mu = \sum_{i=1}^{m} P_i \delta^{(i)}$. By assumption, there exist $1 \leq i \leq m$, $1 \leq j \leq k$, $t \geq 0$, and $y \in K^+$ with $y_j = 0$ such that

$$\delta_j^{(i)} \neq 0, \qquad P_i(t, y) \neq 0. \tag{8}$$

On the other hand, by Theorem 3.1 an SDE system of the form (1) or (2) preserves non-negativity if and only if condition (4) is satisfied, and this contradicts assumption (8). □

Remark 3.2. The results of this subsection remain valid if the stochastic systems of the form (1) or (2) were interpreted in the sense of Stratonovich (see Remark 3.1).

3.3. Invariance criteria for rectangular subsets

In biological applications, the solutions often describe quantities that are not only non-negative, but also bounded by a certain maximum value, e.g., a maximum concentration or the carrying capacity of a population. Hence, the admissible ranges for the solutions are intervals of the form $[0, c]$, $c > 0$. We now formulate a more general invariance criterion that includes Theorem 3.1 as a special case. It yields explicit necessary and sufficient conditions for Allen's derivation procedure such that the resulting models preserve non-negativity and/or upper bounds for the solutions.

Definition 3.2. We call the subset $K \subset \mathbb{R}^k$ **invariant** for the stochastic system (3) if for every initial data $X_0 \in K$ and initial time $t_0 \geq 0$ the corresponding solution $X(t)$, $t \geq t_0$, satisfies

$$P\left(\{X(t) \in K,\ t \in [t_0, \infty)\}\right) = 1,$$

i.e., solutions almost surely attain values within the set K.

Theorem 3.2. *Let $I \subset \{1, \dots, k\}$ be a non-empty subset and $a_i, b_i \in \mathbb{R} \cup \{\infty\}$ be such that $b_i > a_i$, $i \in I$. Then, the set*

$$K := \{x \in \mathbb{R}^k :\ a_i \leq x_i \leq b_i,\ i \in I\}$$

is invariant for the system of SDEs (1) derived by Allen's method if and only if

$$P_i(t, y)(\delta_j^{(i)})^2 = 0 \qquad \text{for } y \in K \text{ such that } y_j = a_j \text{ or } y_j = b_j,$$

for all $t \geq 0$, $i \in I$ and $j = 1, \dots, k$.

This result equally applies to the SDE system (2) and is independent of Itô's or Stratonovich's interpretation of SDEs.

Proof. The statement can be deduced from Theorem A.1. The proof follows exactly the same lines as for Theorem 3.1 and is left to the reader. □

Remark 3.3. Similarly as in Section 3.2, we can formulate conditions such that (δ, p)-SDE models that are derived from deterministic ODE models preserve invariance. It turns out that the class of ODE systems for which the corresponding stochastic extensions preserve non-negativity and upper bounds for the solutions is seriously limited.

Example 3.2. We apply the invariance criterion to the (δ, p)-SDE model in Example 2.1 and derive conditions for the non-negativity and boundedness of the solutions.

Let κ_1 and κ_2 be given upper bounds for the populations X_1 and X_2. Then, the solution X_i attains values within the interval $[0, \kappa_i]$, $i = 1, 2$, if and only if the birth, death, and transition rates are of the following form:

$$\begin{aligned} b_i(t, X_1, X_2) &= X_i(\kappa_i - X_i)\tilde{b}_i(t, X_1, X_2), \\ d_i(t, X_1, X_2) &= X_i(\kappa_i - X_i)\tilde{d}_i(t, X_1, X_2), \\ m_{ij}(t, X_1, X_2) &= X_1 X_2(\kappa_1 - X_1)(\kappa_2 - X_2)\tilde{m}_{ij}(t, X_1, X_2), \quad i \neq j, \end{aligned}$$

for $i, j = 1, 2$, and some functions $\tilde{b}_i$, $\tilde{d}_i$, $\tilde{m}_{ij}$.

4. Applications

We discuss several of the modeling applications presented in [1, 2]. Further SDE models derived by Allen's method can be found, e.g., in [3–8].

4.1. Scalar SDE growth models

We first recall the simplest example discussed in [2], a scalar SDE model for a birth, death, and immigration process.

Example 4.1. The random variable $X(t)$ denotes the size of a population experiencing birth, death, and immigration. The following table lists the changes and corresponding probabilities, where α, β, and γ are positive constants.

i	$\delta^{(i)}$	p_i
1	1	$\alpha X \Delta t$
2	−1	$\beta X \Delta t$
3	1	$\gamma \Delta t$

The corresponding Itô SDE obtained by Allen's derivation procedure is

$$dX = ((\alpha - \beta)X + \gamma)dt + \sqrt{(\alpha + \beta)X + \gamma}\, dW,$$

or alternatively, the construction (2) leads to

$$dX = ((\alpha - \beta)X + \gamma)dt + \sqrt{\alpha X}\, dW_1 + \sqrt{\beta X}\, dW_2 + \sqrt{\gamma}\, dW_3,$$

where W, W_1, W_2, and W_3 are independent standard scalar Wiener processes.

Proposition 4.1. *The solutions corresponding to non-negative initial data of both SDEs can attain undesired negative values and these* (δ, p)*-SDE models are, in fact, not viable.*

Moreover, while the underlying deterministic ODE model

$$\frac{dX}{dt} = (\alpha - \beta)X + \gamma \tag{9}$$

certainly preserves non-negativity, no SDE extension can be derived by Allen's method that possesses this property.

Proof. The first statement is an immediate consequence of Theorem 3.1, since the probability p_3 does not fulfil the required condition. The second observation follows from Corollary 3.1. □

We can modify the modeling assumptions in order to obtain an admissible SDE model. Replacing the probability p_3, e.g., by $\gamma X \Delta t$, the procedure leads to the SDE model

$$dX = ((\alpha - \beta)X + \gamma X)dt + \sqrt{(\alpha + \beta)X + \gamma X}dW,$$

or applying the alternative construction, to

$$dX = ((\alpha - \beta)X + \gamma X)dt + \sqrt{\alpha X}dW_1 + \sqrt{\beta X}dW_2 + \sqrt{\gamma X}dW_3.$$

Both SDEs preserve non-negativity, but this change of p_3 also leads to a modification of the underlying deterministic model. The immigration rate now depends on the current size of the population, implying that individuals do not migrate to places where no other individuals of the species are present, but prefer regions with large population sizes.

Allen's construction method for SDE models is based on the assumption that the underlying stochastic process can be "decomposed" and is determined by all possible changes of the system and their corresponding probabilities. Moreover, it leads to a strong correlation of the drift and diffusion part in the SDE model. For instance, an SDE of the form

$$dX = ((\alpha - \beta)X + \gamma X)dt + \sqrt{(\alpha + \beta)X}dW \tag{10}$$

preserves non-negativity, but cannot be obtained by the derivation procedure. On the other hand, it can be deduced from the deterministic model (9) by assuming that the parameters α and β are subject to random perturbations. It would be interesting to investigate whether Allen's method can be generalized in such a way that it allows to derive stochastic models of the form (10).

The second application is the SDE extension of a logistic growth model.

Example 4.2. The classical ODE model for logistic growth is the following:

$$\frac{dX}{dt} = \alpha X\left(1 - \frac{X}{\beta}\right),$$

where $X(t)$ denotes the population size at time $t \geq 0$, α the growth rate, and β the carrying capacity of the population. Certainly, solutions emanating from initial data within the interval $[0, \beta]$ are non-negative and bounded from above by β, i.e., $[0, \beta]$ is invariant for the ODE model.

By using the modeling procedure in Section 3, two alternative SDE extensions were constructed and analyzed in [2] based on the following probability functions p_i and $\tilde{p}_i$, $i = 1, 2$, associated with the possible changes in the population size.

i	$\delta^{(i)}$	p_i	$\tilde{p}_i$
1	1	$\alpha X \Delta t$	$\alpha X \left(1 - \frac{X}{2\beta}\right) \Delta t$
2	-1	$\frac{\alpha X^2}{\beta} \Delta t$	$\frac{\alpha X^2}{2\beta} \Delta t$

These hypotheses lead to the SDE models

$$dX = \left(\alpha X \left(1 - \frac{X}{\beta}\right)\right) dt + \sqrt{\alpha X \left(1 + \frac{X}{\beta}\right)} dW,$$

and

$$dX = \left(\alpha X \left(1 - \frac{X}{\beta}\right)\right) dt + \sqrt{\alpha X} dW.$$

Proposition 4.2. *Both SDE extensions preserve non-negativity, however, unlike the deterministic growth model, solutions emanating from initial data within the interval* $[0, \beta]$ *can attain values exceeding the carrying capacity* β.

Proof. All probability functions satisfy the conditions in Theorem 3.1, but only $\tilde{p}_1$ fulfils the conditions required by Theorem 3.2 for the invariance of the interval $[0, \beta]$. □

In fact, by taking Theorem 3.2 into account we can modify the modeling assumptions and suggest alternative SDE extensions that preserve not only non-negativity but also the upper bound β for the solutions. For example, we can consider the following probability functions $\hat{p}_i$.

i	$\delta^{(i)}$	$\hat{p}_i$
1	1	$2\alpha X \left(1 - \frac{X}{\beta}\right) \Delta t$
2	-1	$\alpha X \left(1 - \frac{X}{\beta}\right) \Delta t$

Allen's derivation procedure then leads to the modified SDE model

$$dX = \left(\alpha \left(1 - \frac{X}{\beta}\right)\right) dt + \sqrt{3\alpha X \left(1 - \frac{X}{\beta}\right)} dW,$$

for which the interval $[0, \beta]$ is invariant. Different from Example 4.1, the modified probabilities $\hat{p}_i$ do not lead to a change in the drift part of the SDE (i.e., in the underlying deterministic model), but only alter the diffusion part. These modeling assumptions signify that the probabilities for growth and death vanish when the population size reaches the carrying capacity.

4.2. Models for two interacting populations

A general SDE model for two interacting populations was formulated in Example 2.1 and taken up in the subsequent sections. We now consider two concrete modeling applications that are particular cases of this model.

Example 4.3. The following ODE system describes the dynamics of an epidemic,

$$\frac{dS}{dt} = -\alpha\frac{SI}{N} + \gamma I,$$
$$\frac{dI}{dt} = \alpha\frac{SI}{N} - \gamma I,$$

where S denotes the susceptible and I the infected sub-population. The total population $N = S + I$ is preserved. It is a special case of the general model for two interacting populations, where $X_1 = S$, $X_2 = I$, $d_1 = d_2 = b_1 = b_2 = 0$ and

$$m_{12} = \alpha\frac{X_2}{N} = \alpha\frac{X_2}{X_1 + X_2}, \qquad m_{21} = \gamma.$$

Allen's derivation procedure leads to the SDE system

$$dS = \left(-\alpha\frac{SI}{N} + \gamma I\right) dt + \sqrt{\frac{1}{2}\left(\alpha\frac{SI}{N} + \gamma I\right)}(dW_1 - dW_2),$$
$$dI = \left(\alpha\frac{SI}{N} - \gamma I\right) dt + \sqrt{\frac{1}{2}\left(\alpha\frac{SI}{N} + \gamma I\right)}(-dW_1 + dW_2)$$

(see [1], sec. 5.2.2).

Theorem 3.1 and Corollary 3.1 imply the following observation.

Proposition 4.3. *The deterministic ODE model preserves non-negativity of solutions, but the stochastic extension does not possess this property and solutions can attain negative values.*

In fact, no SDE model can be constructed by Allen's derivation procedure that preserves non-negativity.

One possibility to obtain a viable stochastic model is to replace the transition rate m_{21} by γX_1, which leads to the modified SDE system

$$dS = \left(-\alpha\frac{SI}{N} + \gamma SI\right) dt + \sqrt{\frac{1}{2}\left(\alpha\frac{SI}{N} + \gamma SI\right)}(dW_1 - dW_2),$$
$$dI = \left(\alpha\frac{SI}{N} - \gamma SI\right) dt + \sqrt{\frac{1}{2}\left(\alpha\frac{SI}{N} + \gamma SI\right)}(-dW_1 + dW_2).$$

As in Example 4.1 this assumption not only alters the stochastic perturbation, but also the drift part of the SDE and hence, the underlying deterministic model. An alternative SDE model that preserves non-negativity is the system

$$dS = \left(-\alpha\frac{SI}{N} + \gamma I\right) dt + \sqrt{\left(\alpha\frac{SI}{2N}\right)}dW_1,$$
$$dI = \left(\alpha\frac{SI}{N} - \gamma I\right) dt - \sqrt{\left(\alpha\frac{SI}{2N}\right)}dW_1.$$

It can be deduced from the ODE model by assuming that only the transition rate α is subject to random perturbations, but such an SDE system cannot be obtained using Allen's derivation method.

We further discuss an SDE system modeling enzyme kinetics (see [2], secs. 7.5 and 9.5).

Example 4.4. The following ODE system describes enzyme kinetics

$$\frac{dN}{dt} = -\alpha N(\kappa - B) + \beta B,$$
$$\frac{dB}{dt} = \alpha N(\kappa - B) - (\beta + \gamma)B,$$

where N denotes the number of a nutrient molecule and B the number of molecules formed when the nutrient binds to an enzyme.

The changes and associated probabilities are listed in the following table, where α, β, γ, and κ are positive constants.

i	$\delta^{(i)}$	p_i
1	$(-1, 1)^T$	$\alpha N(\kappa - B)\Delta t$
2	$(1, -1)^T$	$\beta B \Delta t$
3	$(0, -1)^T$	$\gamma B \Delta t$

Based on the table, the drift term μ, the diffusion matrix V, and its square root can be computed. The derivation procedure leads to the Itô SDE system

$$dX = \mu(X)dt + G(X)dW,$$

where $X = (N, B)^T$, $W = (W_1, W_2, W_3)^T$ with independent standard scalar Wiener processes W_i, and

$$\mu(X) = \begin{pmatrix} -\alpha N(\kappa - B) + \beta B \\ \alpha N(\kappa - B) - (\beta + \gamma)B \end{pmatrix},$$
$$G(X) = \begin{pmatrix} -\sqrt{\alpha N(\kappa - B)} & \sqrt{\beta B} & 0 \\ \sqrt{\alpha N(\kappa - B)} & -\sqrt{\beta B} & -\sqrt{\gamma B} \end{pmatrix}.$$

It is a special case of the general model for two interacting populations, where $X_1 = N$, $X_2 = B$, $d_1 = b_1 = b_2 = 0$, $d_2 = \gamma$, and

$$m_{12} = \alpha(\kappa - B), \qquad m_{21} = \beta.$$

Similarly to the previous example, we observe that the (δ, p)-SDE model has the following properties.

Proposition 4.4. *The underlying deterministic model preserves the non-negativity of solutions. However, the probabilities p_1 and p_2 do not satisfy the conditions in Theorem 3.1 and hence, solutions of the stochastic enzyme kinetics model can attain undesired negative values.*

By Corollary 3.1 no SDE extension can be constructed by Allen's method that preserves non-negativity.

4.3. A growth model including environmental variability

An SDE extension of a simple growth model was developed in [1]. It takes stochastic fluctuations in the environment into account by considering fluctuation in the birth and death rates.

The model is based on the deterministic ODE

$$\frac{dX}{dt} = bX - dX,$$

where X denotes the population size and b and d the birth and death rates. An SDE for X is constructed by Allen's procedure based on the following table for the possible changes and corresponding probabilities.

i	$\delta^{(i)}$	p_i
1	1	$bX\Delta t$
2	−1	$dX\Delta t$

Moreover, it is assumed that the birth and death rates vary in a random manner, and SDEs are derived for b and d according to the following modeling assumptions.

i	$\delta_b^{(i)}$	$p_i^{(b)}$	$\delta_d^{(i)}$	$p_i^{(d)}$
1	α_b	$(q_b + \beta_b(b_0 - b))\Delta t$	α_d	$(q_d + \beta_d(d_0 - d))\Delta t$
2	$-\alpha_b$	$(q_b - \beta_b(b_0 - b))\Delta t$	$-\alpha_d$	$(q_d - \beta_d(d_0 - d))\Delta t$

Here, α_b, α_d, β_b, β_d, q_b, and q_d are positive constants. The derivation method leads to the SDE system

$$\begin{aligned} dX &= (bX - dX)dt + \sqrt{(b+d)X}dW_1, \\ db &= (2\alpha_b\beta_b(b_0 - b))dt + \sqrt{2\alpha_b^2 q_b}dW_2, \\ dd &= (2\alpha_d\beta_d(d_0 - d))dt + \sqrt{2\alpha_d^2 q_d}dW_3. \end{aligned}$$

Theorem 3.1 and Corollary 3.1 imply that solutions of the SDE model can attain negative values. In fact, no SDE extension can be constructed by the modeling procedure that preserves non-negativity, since the probability functions corresponding to b and d contain constant terms. We remark that the coefficients in the SDE for X actually satisfy the criterion in Theorem 3.1, however, the birth and death rates b and d can attain negative values, and this will also cause problems defining the square root in the first equation.

Acknowledgments

We would like to thank the referee for his/her valuable comments and remarks that significantly improved the writing of our article.

Funding

The first author would like to thank for the invitation to the University of Kaiserslautern and the financial support of his visit. Until May 2016 the second author was funded by the DFG grant SU807/1-1 and supported by the TU Nachwuchsring of the University of Kaiserslautern.

References

[1] Allen, E. J. 2007. *Modeling with Itô Stochastic Differential Equations (Mathematical Modeling, Theory and Applications)*, Vol. 22. Dordrecht: Springer.

[2] Allen, L. J. S. 2011. *An Introduction to Stochastic Processes with Applications to Biology*. 2nd ed. Boca Raton, FL: Chapman & Hall.
[3] Ackleh, A. S., Deng, K., and Huang, Q. 2011. Stochastic juvenile-adult models with application to a green tree frog population. *Journal of Biological Dynamics* 5(1):64–83.
[4] Allen, E. J., Allen, L. J. S., Archintega, A., and Greenwood, O. E. 2008. Construction of equivalent stochastic differential equation models. *Stochastic Analysis and Applications* 26(2):274–297.
[5] Dogan-Ciftci, E., and Allen, E. J. 2014. Derivation of several SDE systems in one- and two-locus population genetics. *Stochastic Analysis and Applications* 32(5):761–775.
[6] Tuckwell, H. C., and Le Corfec, E. 1998. A stochastic model for early HIV-1 population dynamics. *Journal of Theoretical Biology* 195(4):451–463.
[7] Vidurupola, S. W., and Allen, L. J. S. 2012. Basic stochastic models for viral infection within a host. *Mathematical Biosciences and Engineering* 9(4):915–935.
[8] Yuan, Y., and Allen, L. J. S. 2011. Stochastic models for virus and immune system dynamics. *Mathematical Biosciences* 234(2):84–94.
[9] Cresson, J., Puig, B., and Sonner, S. 2016. Stochastic models in biology and the invariance problem. *Discrete and Continuous Dynamical Systems - Series B* 21(7):2145–2168.
[10] Cresson, J., Puig, B., and Sonner, S. 2013. Validating stochastic models: Invariance criteria for systems of stochastic differential equations and the selection of a stochastic Hodgkin-Huxley type model. *International Journal of Biomathematics and Biostatistics* 2(1):111–122.
[11] Øksendal, B. 2003. *Stochastic Differential Equations: An Introduction with Applications*. 6th ed. Berlin-Heidelberg: Springer-Verlag.
[12] Braumann, C. A. 2001. Itô versus Stratonovich calculus in random population growth. *Mathematical Biosciences* 206(1):81–107.
[13] Turelli, M. 1977. Random environments and stochastic calculus. *Theoretical Population Biology* 12(2):140–178.
[14] Milian, A. 1995. Stochastic viability and comparison theorem. *Colloquium Mathematicum* LXVIII(2):297–316.

Appendix A: Invariance criteria for SDEs: A reminder

We recall a general criterion for SDE systems, which yields explicit necessary and sufficient conditions for the invariance of rectangular subsets. It slightly extends a previous result by A. Milian in [14]. For details and the proof we refer to [9, 10, 14].

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a right-continuous increasing family $F = (\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{F}$ each containing all sets of P-measure zero. We consider systems of Itô SDEs of the form

$$\begin{aligned} dX(t) &= f(t, X(t))dt + g(t, X(t))dW(t), \qquad t \in (t_0, \infty), \\ X(t_0) &= X_0, \end{aligned} \tag{A.1}$$

where $f = (f_i)_{1\leq i\leq k} : [0, \infty) \times \mathbb{R}^k \to \mathbb{R}^k$ is a Borel-measurable function, and $g = (g_{ij})_{\substack{1\leq i\leq k \\ 1\leq j\leq r}} : [0, \infty) \times \mathbb{R}^k \to \mathbb{R}^{k\times r}$ a Borel-measurable mapping into the set of all $\mathbb{R}^{k\times r}$-matrices. Furthermore, $W : [0, \infty) \times \Omega \to \mathbb{R}^r$ denotes an r-dimensional $\mathcal{F}$-adapted Wiener process and dW the corresponding Itô differential. The initial time t_0 is non-negative and $X_0 \in \mathbb{R}^k$ is the given initial data.

The initial value problem is rigorously defined through the integral equation

$$X(t) = X_0 + \int_{t_0}^{t} f(s, X(s))ds + \int_{t_0}^{t} g(s, X(s))dW(s), \qquad t \in (t_0, \infty),$$

where the last term denotes the Itô integral ([11]). An alternative notion that is frequently used in applications is Stratonovich's definition of stochastic integrals. The qualitative behavior of solutions can depend on the chosen concept, however, there is an explicit transformation formula relating the solutions of both notions. In fact, if X is a solution of system (A.1) and the SDE is interpreted in the sense of Stratonovich, then X solves the system of Itô SDEs

$$dX(t) = \Big(f(t, X(t)) + \frac{1}{2}h(t, X(t))\Big)dt + g(t, X(t))dW(t), \qquad t \in (t_0, \infty),$$

where $h = (h_i)_{1\leq i\leq k} : [0, \infty) \times \mathbb{R}^k \to \mathbb{R}^k$ is given by

$$h_i(t, x) = \sum_{j=1}^{r}\sum_{l=1}^{k} \frac{\partial g_{ij}}{\partial x_l}(t, x)g_{ij}(t, x), \qquad i = 1, \dots, k.$$

Due to the particular form of h, the invariance property for SDEs is independent of the interpretation, i.e., qualitative properties of solutions such as non-negativity and boundedness do not depend on the choice of Itô's or Stratonovich's interpretation (see [10, 11]).

In the sequel, we denote by (f, g) stochastic initial value problems of the form (A.1).

Definition A.1. We call the subset $K \subset \mathbb{R}^k$ **invariant** for the stochastic system (f, g) if for every initial data $X_0 \in K$ and initial time $t_0 \geq 0$ the corresponding solution $X(t)$, $t \geq t_0$, satisfies

$$P\left(\{X(t) \in K,\ t \in (t_0, \infty)\}\right) = 1,$$

i.e., the solution almost surely attains values within the set K.

Theorem A.1. *Let $I \subset \{1, \dots, k\}$ be a non-empty subset and $a_i, b_i \in \mathbb{R} \cup \{\infty\}$ such that $b_i > a_i$, $i \in I$. Then, the set*

$$K := \{x \in \mathbb{R}^k :\ a_i \leq x_i \leq b_i,\ i \in I\}$$

is invariant for the stochastic system (f, g) if and only if

$$\begin{aligned} f_i(t, x) &\geq 0 && \text{for } x \in K \text{ such that } x_i = a_i,\\ f_i(t, x) &\leq 0 && \text{for } x \in K \text{ such that } x_i = b_i,\\ g_{i,j}(t, x) &= 0 && \text{for } x \in K \text{ such that } x_i \in \{a_i, b_i\},\ \ j = 1, \dots, r, \end{aligned}$$

for all $t \geq 0$ and $i \in I$.

This result applies independently of Itô's or Stratonovich's interpretation of SDEs.

One particular and important case in applications is the non-negativity of solutions, i.e., the invariance of the positive cone.

Corollary A.1. *Let $I \subset \{1, \dots, k\}$ be a non-empty subset. Then, the set*

$$K^+ := \{x \in \mathbb{R}^k : x_i \geq 0,\ i \in I\}$$

is invariant for the stochastic system (f, g) if and only if

$$\begin{aligned} f_i(t, x) &\geq 0 && \text{for } x \in K^+ \text{ such that } x_i = 0,\\ g_{i,j}(t, x) &= 0 && \text{for } x \in K^+ \text{ such that } x_i = 0,\ \ j = 1, \dots, r, \end{aligned}$$

for all $t \geq 0$ and $i \in I$.

This result is valid independent of Itô's or Stratonovich's interpretation.