Pathwise approximations to solving the filtering problem for the stochastic chemostat

José Augusto Fontenele Magalhães[1], **Muhammad Fuady Emzir**[2] & **Francesco Corona**[1]

[1] Department of Chemical and Metallurgical Engineering, Aalto University, Espoo, FI
[2] Control and Instrumentation Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran, SA

E-mail: `augusto.magalhaes@aalto.fi`, `muhammad.emzir@kfupm.edu.sa`, `francesco.corona@aalto.fi`

5 June 2025

Abstract. This paper concerns the inverse problem of characterising the state of a bioreactor from observations. In laboratory settings, the bioreactor is represented by a device called a chemostat. We consider a differential description of the evolution of the state of the chemostat under environmental fluctuations. First, we model the state evolution as a stochastic process driven by Brownian motion. Under this model, our best knowledge about the state of the chemostat is described by its probability distribution in time, given the distribution of the initial state. The corresponding probability density function solves a deterministic partial differential equation (PDE), the Kolmogorov forward equation. While this provides a probabilistic description, incorporating an observation process allows for a more refined characterisation of the state. More formally, we are interested in obtaining the distribution of the state conditional on an observation process as the solution to a filtering problem, with the corresponding conditional probability density function solving a non-linear stochastic PDE, the Kushner–Stratonovich equation. This paper focuses on the pathwise formulation of this filtering problem in which inferences about the state are obtained conditional on a fixed stream of observations. We establish the existence and uniqueness of solutions to the governing differential equations, ensuring well-posedness before presenting numerical approximations. We then approximate the pathwise solution to the filtering problem by combining the finite difference and splitting methods for solving PDEs, and then compare the approximated solution with results from a linearisation method and a classical sequential Monte Carlo method.

Keywords: state estimation, pathwise filtering, Kolmogorov forward equation, Kushner-Stratonovich equation, stochastic chemostat

This work has been submitted to the IOP for possible publication. If accepted, copyright may be transferred without notice, after which this version may no longer be available.

Contents

1 Introduction **2**

2 Preliminaries: Stochastic filtering **5**
2.1 Stochastic differential equations (SDEs) 6
2.1.1 Existence and uniqueness of the solution. 8
2.1.2 Exit times. 9
2.1.3 Markovianity of the solution. 10
2.1.4 The Kolmogorov forward equation (KFE). 11
2.2 Stochastic filtering theory 13
2.2.1 Stochastic equation for the filter. 14
2.2.2 The Kushner-Stratonovich equation (KSE). 14

3 The chemostat **18**
3.1 Stochastic dynamics 20
3.1.1 Existence, uniqueness and non-explosivity of the solution 21
3.1.2 Markovianity and the KFE 24
3.2 Filtering problem for the stochastic chemostat 26

4 Concluding remarks **31**

1. Introduction

An inverse problem for a dynamic system refers to the construction of a set of equations given a differential description of the evolution of the system and a collection of observations related to quantities of interest. In this paper, the quantity of interest is the state variable of such a system denoted as *signal*, with inferences about the signal being obtained from indirect and noisy observations.

The study of this type of inverse problem has been the focus of considerable theoretical and numerical work under the name *stochastic filtering*. The aim is to obtain expectations with respect to the probability distribution of the signal at the current time, this distribution being conditional on noisy observations accumulated up to that time. The theoretical features of this filtering problem are well-known: such expectations require integration with respect to a stochastic process modelling the observations, and the probability density function of the conditional probability distribution of the signal solves a non-linear stochastic partial differential equation, the Kushner–Stratonovich equation (KSE). A comprehensible account of this problem is available in [1, 2].

Filtering theorists are typically concerned with analysing the differential equations governing the filtering problem. Among other topics, they study the existence, uniqueness, and growth behaviour of solutions to the stochastic differential equations that describe the evolution of the underlying process and the observations. This is crucial for numerical

approximations of the expectations of interest as most analyses of existing algorithms rely on such properties [1, 3].

Filtering practitioners have adopted the pathwise formulation of the filtering problem. This is because when observing most real-world signals, the single "path" for the observations never exactly aligns with the observation model: signals are sampled at a high frequency, with sampled points artificially connected, for instance, by straight-line segments. While this formulation enables various lines of research [4, 5], this work focuses on fixing the observation path to reduce the aforementioned KSE to a partial differential equation (PDE) with a random coefficient.

The conditional probability density function in pathwise filtering can be derived in closed-form for a specific class of problems, for instance, those with a linearly evolving observation process (the observation is a noisy, scaled version of the signal) [6, 7]. However, even in this scenario, the signal is only allowed to have a non-linear evolution under some restrictive conditions (see Beneš condition [6]), which are rarely satisfied in real-world dynamic systems.

With limited closed-form solutions, much research has been directed towards constructing algorithms to approximate the solution in more general settings. The approximations can be obtained via, but not limited to, three approaches: the extended Kalman filter (EKF), particle filters (PFs), and methods for solving PDEs.

The EKF is the classical approximation method. It linearises certain non-linearities in the model to ensure that the traditional Kalman-Bucy filter [7] can be applied. PFs, also known as sequential Monte Carlo methods, have been widely developed since their conception in the 1990s [8, 9] and are among the most popular alternatives to the EKF. In PFs, particles (also called samples) first evolve according to a desirable transition kernel and then selectively interact in accordance with their likelihood with respect to the observation path. EKF, PFs and their derivatives have practical applications across diverse domains, including target tracking in robotics [10], weather forecasting [11], and disease outbreak monitoring in epidemiology [12].

A less common approach consists of approximations via a splitting method: the PDE with a random coefficient is decomposed into two parts: a deterministic and a normalisation one. The first part is commonly solved numerically by grid-based schemes. The second part consists of finding a correcting and normalising factor based on the observation path, ensuring that the density function integrates to unity. It is generally accepted that grid-based schemes are successful in low dimensions of the state space [13, 14], being the only method which provides information over the entire state space and allows for visualisation of the state evolution without restrictive assumptions.

This study considers the practical implementation of filtering with focus on the splitting method. Our goal is to obtain the probabilities and expectations for a process performed in a bioreactor, in which one or more populations of microorganisms grow in a nutrient medium consisting of a cocktail of molecules. In particular, we consider a laboratory device representing this bioreactor: a chemostat. Prior to the genomic era, the chemostat was a major player in the use of continuous culture systems to

characterise certain kinds of microbes [15, 16]. Being a realistic and thoroughly tested continuous culture system, it has regained some interest in this post-genomic era. For instance, chemostat-grown cultures lead to reproducible and homogeneous results which are necessary to identify the roles played by genes in the biology of organisms [17].

We consider a differential description of the state of a chemostat under environmental fluctuations, specifically a chemostat subjected to Brownian motion processes with multiplicative noise, herein called stochastic chemostat. We adopt the dynamic model proposed in [18], as its differential equations have shown by means of martingale theory to admit a unique solution within a suitably chosen space. In particular, [18] ensures that their model would allow for its realisations remaining within a biologically meaningful domain (e.g. positive chemical concentration values). Finally, to describe the relationship between observations and states, the dynamic model is coupled with an observation model in which the observations are also subjected to a Brownian motion process.

The existing state-of-the-art of filtering for the stochastic chemostat comprises an approximation to the first two moments of the conditional probability distribution of the signal by EKF [19] and a more complex approximation by PFs [20]. However, to the best of our knowledge, no previous work has addressed the theoretical guarantees of existence, uniqueness, and growth behaviour of the solutions to our differential equations of interest, neither has the pathwise solution to the filtering problem been numerically realised by grid-based schemes.

The following are the contributions of this paper:

- First, Theorem 2 shows that the differential equations for the stochastic chemostat have a unique solution within a suitably chosen space. We provide an alternative proof to the results in [18] via Lyapunov functions, aligning with existing stochastic population dynamics literature (e.g. [21] on the Lotka-Volterra model). Further, our proof introduces tools that will be useful later in the analysis, particularly in establishing the well-posedness of the PDEs.
- Secondly, Proposition 2 addresses the well-posedness of the Kolmogorov forward equation for the stochastic chemostat. We also discuss the implication of this result on the pathwise form of the Kushner-Stratonovich equation. Valid conclusions from numerical simulations of these PDEs require guarantees of existence and uniqueness of their solution. To the best of our knowledge, previous studies have assumed well-posedness without formal derivation. In contrast, our work explicitly investigates if these assumptions hold.
- Lastly, we demonstrate the pathwise approximations using the splitting method, comparing them against EKF and PF results. This allows us to move beyond merely visualising the central moments or expectations over time. Instead, the shape of the entire distribution across the state space is available, providing a more comprehensive understanding of the system dynamics and the associated uncertainty.

This paper is structured as follows. Section 2 provides the theoretical foundations

for understanding the filtering problem. In particular, Section 2.1 introduces assumptions ensuring the existence, uniqueness, and growth properties of solutions to key differential equations. Section 2.2 introduces the filtering problem and its pathwise formulation, covering both the theory and approximation methods. Section 3 applies the concepts from Section 2 to the stochastic chemostat. In particular, Section 3.1 discusses the properties of solutions to the differential equations describing the evolution of the state of the stochastic chemostat. Finally, Section 3.2 presents the pathwise filtering problem for the stochastic chemostat, showing its numerical solution based on the aforementioned approximation methods. Section 4 concludes the paper with some remarks.

2. Preliminaries: Stochastic filtering

Filtering refers to a methodology in which observations and a model of a dynamic system are combined to make inferences about its state. The state variable of this dynamic system is denoted as signal. In stochastic filtering, the signal is modelled by a continuous-time stochastic process $X = (X_t, t \geq 0)$, with t denoting time. While the signal cannot be observed directly, an indirect observation can be obtained. This observation is assumed to be described by a continuous-time stochastic process $Y = (Y_t, t \geq 0)$ denoted as observation process. Moreover, Y is assumed to be a function of X corrupted by an observation noise, modelled by a stochastic process $B^Y = \left(B_t^Y, t \geq 0\right)$. For inferences on X from Y, we let $\mathcal{F}^Y = \left(\mathcal{F}_t^Y, t \geq 0\right)$ be the filtration generated by Y, where the σ-field $\mathcal{F}_t^Y = \sigma\left(Y_s, s \in [0, t]\right)$ is all the information from observations up to time t. The resulting filtering problem consists of determining the conditional distribution π_t of X_t given $\mathcal{F}_t^Y$. In a mean-square sense, the estimate $\mathbb{E}\left[X_t | \mathcal{F}_t^Y\right]$ of the value of the signal at time t, the conditional mean of X_t given $\mathcal{F}_t^Y$, is optimal. Other inferences about X from Y require evaluating quantities of the form $\mathbb{E}\left[\phi\left(X_t\right) | \mathcal{F}_t^Y\right]$ for some function ϕ over the state-space of the signal. Notably, the conditional distribution of the signal is also modelled by a stochastic process $\pi = (\pi_t, t \geq 0)$.

This section provides a compact overview of this problem and the background needed for our application. Section 2.1 overviews a number of notions and results about stochastic differential equations, such as existence, uniqueness and Markovianity of solutions. Section 2.2 concerns the filtering problem. Since the evolution of the system is described in continuous-time, we present filtering equations using a body of techniques from stochastic calculus outlined in Section 2.1. In particular, Section 2.2.1 presents these equations when we assume that 1.) the signal process $X = (X_t, t \geq 0)$ takes values in an N_X-dimensional state-space $\mathbb{S}_X$ and that X is the solution to a stochastic differential equation driven by an N_{B^X}-dimensional Brownian motion process $B^X = \left(B_t^X, t \geq 0\right)$, and 2.) the observation process $Y = (Y_t, t \geq 0)$ takes values in $\mathbb{S}_Y \subseteq \mathbb{R}^{N_Y}$ and that Y has an evolution which depends on both X and an N_{B^Y}-dimensional Brownian motion process $B^Y = \left(B_t^Y, t \geq 0\right)$. Section 2.2.2 particularises the filtering equations for a case in which the pair (X, Y) is Markovian. The main computational methods for solving these equations are also outlined.

2.1. Stochastic differential equations (SDEs)

To set the notation, we consider an identity modelling the evolution of some deterministic process $z = (z_t, t \geq 0)$,

$$dz_t = c_t\left(z, v, u, \theta\right) dt, \tag{2.1}$$

where $z_t \in \mathbb{S}_z$ is the value at time t of the process of interest, dz_t is called the differential of the process z at time t, dt is an infinitesimal time increment, $v = (v_t(z), t \geq 0)$ with $v_t \in \mathbb{S}_v$ is an auxiliary process possibly dependent on the entire process z, $u = \left(u_t\left(z_{s,s\leq t}\right), t \geq 0\right)$ with $u_t \in \mathbb{S}_u$ is a process which describes the evolution of state-feedback control actions and is possibly dependent on past values of z, and $\theta = (\theta_t = \text{const}, t \geq 0)$ with $\theta_t \in \mathbb{S}_\theta$ is a process which characterises the model parameters that are constant in time. The coefficient c_t is the functional which specifies the dynamics. At this point, it suffices to assume that $\mathbb{S}_z, \mathbb{S}_v, \mathbb{S}_u$ and $\mathbb{S}_\theta$ are complete separable metric spaces. Without loss of generality, processes u and θ are omitted hereafter. Informally, note that popular forms of ordinary differential equations (ODEs) can be recovered from (2.1). For instance, if processes z and v are smooth enough, c_t is the derivative of z with respect to time evaluated at time t. Moreover, if v is specified up to time t and takes the form $v = (v_t(z_{s,s\leq t}))$, then $c_t\left(z, v\right) = c_t\left(z, v_t\left(z_{s,s\leq t}\right)\right)$. For the particular case of first-order differential equations

$$dz_t = c_t\left(z_t, v_t\left(z_{s,s\leq t}\right)\right) dt, \tag{2.2}$$

knowing the past $(v_s, s \leq t)$ of v and any value z_s, $s \leq t$, allows to determine the conditions under which a process z satisfying ODE (2.2) exists and is unique (Picard's theorem in, for instance, [22]): typically, the initial (at $s = 0$) condition z_0 is provided.

Stochastic differential equations can be understood as deterministic differential equations into which a degree of randomness is introduced. In this case, when process v is a stochastic process $V = (V_t, t \geq 0)$ and/or the initial condition z_0 is a random variable, process z which satisfies identity (2.1) becomes a stochastic process $Z = (Z_t, t \geq 0)$. An additional source of randomness can be introduced by including another stochastic process $B = (B_t, t \geq 0)$. This procedure results in the class of differential equations known as stochastic differential equations (SDEs), whose solutions involve ordinary and stochastic integrals. We consider stochastic integration with respect to B, particularly for the case in which B is a standard Brownian motion process. We let $(\Omega, \mathcal{F}, \mathbb{P})$ be the appropriate probability space for our processes, then such B is defined as follows:

Definition 1 (Standard Brownian motion process)**.** *A stochastic process $B = (B_t, t \geq 0)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be an N_B-dimensional standard Brownian motion process if the following conditions are satisfied:*

i. $\mathbb{P}$-almost surely $B_0 = 0$; that is $\mathbb{P}\left(B_0 = 0\right) = 1$;

ii. B has independent increments; that is, $B_0, B_{t_1} - B_0, \ldots, B_{t_{N_t}} - B_{t_{N_t-1}}$ are independent, for every integer $N_t > 1$ and indices $0 < t_1 < \cdots < t_{N_t} < \infty$;

iii. *B has increments which are normally distributed with mean 0 and variance $(t-s)\, I_{N_B \times N_B}$, for all $s \in [0,t)$, where $I_{N_B \times N_B}$ is the identity matrix of size N_B; that is, $B_t - B_s \sim \mathcal{N}(0, (t-s)\, I_{N_B \times N_B})$;*

iv. *$B(\omega) = (B_t(\omega), t \in [0,\infty))$ is continuous for $\mathbb{P}$-almost every $\omega \in \Omega$, where $B(\omega)$ is the sample path, or trajectory, of B associated with ω.*

A stochastic process which only satisfies conditions (i)-(iii) is called a standard pre-Brownian motion process. It can be shown that such process admits a version, a modification, $\tilde{B}$ whose sample paths are $\mathbb{P}$-almost surely locally Hölder-continuous for every exponent $\alpha \in (0, \frac{1}{2})$; that is, for all finite times $T > 0$, there exists a constant $\beta > 0$ such that

$$\mathbb{P}\left[\omega \in \Omega : \sup_{0 \le s < t < T} \frac{\|\widetilde{B}_t(\omega) - \widetilde{B}_s(\omega)\|}{(t-s)^{\alpha}} \le \beta\right] = 1,\ \alpha \in \left(0, \frac{1}{2}\right),\ \|\cdot\| = \text{Euclidean norm}.$$

Hereafter, we choose such version to work with by setting $B = \tilde{B}$. If B has locally Hölder-continuous sample paths, it satisfies condition (iv), thereby satisfying all the conditions to classify B as the standard Brownian motion process. Proof of the existence of such B can be found in, for instance, [23, Section 2].

To provide an instrumental account of the theory of SDEs, we start by letting 1.) $C_Z = C_Z([0,\infty), \mathbb{S}_Z)$ be the space of continuous functions defined on the time index set $[0,\infty)$ and taking values in $\mathbb{S}_Z \subseteq \mathbb{R}^{N_Z}$; and 2.) $D_V = D_V([0,\infty), \mathbb{S}_V)$ be the space of right-continuous functions also on $[0,\infty)$, with left-hand limits, and taking values on a complete separable metric space $\mathbb{S}_V$ such that regular conditional probabilities exist. We use $\mathcal{B}(C_Z)$ and $\mathcal{B}(D_V)$ to denote the σ-fields of Borel sets of C_Z and D_V.

We introduce to identity (2.1) a random initial condition $z_0 = z_0(\omega)_{\omega \in \Omega}$, an $\mathbb{S}_V$-valued stochastic process $V = (V_t, t \ge 0)$, with sample paths $V(\omega) \in D_V$ for $\mathbb{P}$-almost every $\omega \in \Omega$, and a process written in terms of an N_{B^Z}-dimensional standard Brownian motion process $B^Z = \left(B_t^Z, t \ge 0\right)$, where $B_t^Z = \left(B_t^{Z,1}, \ldots, B_t^{Z,N_{B^Z}}\right)$. The resulting SDE modelling the evolution of process $Z = (Z_t, t \ge 0)$ is

$$dZ_t = c_t(Z,V)\, dt + d_t(Z,V)\, dB_t^Z, \quad Z_0 = z_0, \tag{2.3}$$

or, equivalently, in coordinate form, for each $n_Z = 1, \ldots, N_Z$,

$$dZ_t^{n_Z} = c_t^{n_Z}(Z,V)\, dt + \sum_{n_{B^Z}=1}^{N_{B^Z}} d_t^{n_Z, n_{B^Z}}(Z,V)\, dB_t^{Z_{n_{B^Z}}}, \quad Z_0^{n_Z} = z_0^{n_Z}.$$

In SDE (2.3), the functionals $c_t : C_Z \times D_V \to \mathbb{R}^{N_Z}$ and $d_t : C_Z \times D_V \to \mathbb{R}^{N_Z \times N_{B^Z}}$ are $\mathcal{B}(C_Z) \times \mathcal{B}(D_V)$-measurable coefficients; that is, we have that $(c_t^{n_Z})^{-1}(A_{Z,V}) \in \mathcal{B}(C_Z) \times \mathcal{B}(D_V)$ and $\left(d_t^{n_Z, n_{B^Z}}\right)^{-1}(A_{Z,V}) \in \mathcal{B}(C_Z) \times \mathcal{B}(D_V)$, for every $A_{Z,V} \in \mathcal{B}(\mathbb{R})$.

In the following, we limit ourselves to the class of non-anticipative functionals c_t and d_t; that is, the values of c_t and d_t depend only on the paths $V(\omega)$ and $Z(\omega)$ up to time t. For $\mathcal{B}_s(C_Z)$ - respectively, $\mathcal{B}_s(D_V)$ - denoting the smallest σ-field which makes $(Z_{s'}, s' \le s)$ - respectively, $(V_{s'}, s' \le s)$ - measurable and for $\mathcal{B}_{t^+} = \cap_{s>t} \mathcal{B}_s$ denoting the smallest σ-field containing $\mathcal{B}_s$ for all $s > t$, we define non-anticipativeness as follows:

Definition 2 (Non-anticipative functionals)**.** *A $\mathcal{B}(C_Z) \times \mathcal{B}(D_V)$-measurable functional $a_t : C_Z \times D_V \to \mathbb{R}$ is said to be non-anticipative if a_t is $\mathcal{B}_{t^+}(C_Z) \times \mathcal{B}_{t^+}(D_V)$-measurable. That is, $a_t^{-1}(A_{Z,V}) \in \mathcal{B}_{t^+}(C_Z) \times \mathcal{B}_{t^+}(D_V)$ for every $A_{Z,V} \in \mathcal{B}(\mathbb{R})$.*

If each component functional $c_t^{n_Z}$ is non-anticipative in the sense of Definition 2, then the functional $c_t = \left(c_t^1, \ldots, c_t^{n_Z}, \ldots, c_t^{N_Z}\right)$ is said to be non-anticipative. Similarly, if each entry $d_t^{n_Z, n_{B^Z}}$ is non-anticipative, then the matrix-valued functional $d_t = \left(d_t^{n_Z, n_{B^Z}}\right)$, with $n_Z = 1, \ldots, N_Z$ and $n_{B^Z} = 1, \ldots, N_{B^Z}$, is said to be non-anticipative.

2.1.1. Existence and uniqueness of the solution. The notion of a solution to SDE (2.3), as well as the conditions for its existence and uniqueness, can be stated whenever certain conditions on c_t and d_t are satisfied. These conditions are identified after noting that SDE (2.3) is a shorthand to the following stochastic integral equation:

$$Z_t = z_0 + \int_0^t c_s(Z, V) ds + \int_0^t d_s(Z, V) dB_s^Z, \tag{2.4}$$

where the first integral is a Riemann integral and the second one is an Itô integral. Itô integrals are defined for integrands which allow such integrals to be martingales. For details on the martingale theory for Itô integrals, see [24].

Based on (2.4), the notion of a solution of SDE (2.3) can be constructed after establishing the conditions needed on the collection of integrands $(c_t(Z, V), t \geq 0)$ and $(d_t(Z, V), t \geq 0)$ to allow the two integrals to be well-defined. To this end, we first endow our probability space with a notion of time: this is done by specifying which sub-σ-field of information in $\mathcal{F}$ can be gathered by time t. Secondly, we define an increasing family $(\mathcal{F}_t, t \geq 0)$ of σ-fields of $\mathcal{F}$ with $\mathcal{F}_s \subset \mathcal{F}_t$, for $s, t \in [0, \infty)$, with $s \leq t$, and such that $\mathcal{F}_0$ includes all $\mathbb{P}$-null sets in $\mathcal{F}$. Then, for the resulting filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ over which B^Z, V and z_0 are defined, we have the definition:

Definition 3 (SDE: Strong solution [25, Section 5.1])**.** *Let c_t and d_t be non-anticipative functionals. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ be a filtered probability space for B^Z, V, and z_0. A stochastic process Z is said to be a strong solution of SDE (2.3) if the following holds:*

i. Z is adapted to $(\mathcal{F}_t, t \geq 0)$; that is, each Z_t is measurable with respect to $\mathcal{F}_t$;

ii. $\mathbb{P}$-almost surely $\int_0^t \|c_s(Z, V)\| ds < \infty$ and $\int_0^t \|d_s(Z, V)\|^2 ds < \infty$, for all $t \in [0, \infty)$ and for a Euclidean norm of c_s and a Hilbert–Schmidt norm of d_s;

iii. $\mathbb{P}$-almost surely Z solves (2.4).

The non-anticipativeness of c_t and d_t, together with (i), implies that $c_t(Z, V)$ and $d_t(Z, V)$ are measurable and adapted to $\mathcal{F}_{t^+}$ for all $t \geq 0$, with $\mathcal{F}_{t^+} = \cap_{s>t} \mathcal{F}_s$ (see, [24, Section 4.2]). Imposing (ii) on the measurable $\mathcal{F}_{t^+}$-adapted process $(d_t(Z, V), t \geq 0)$ is sufficient to render the Itô integral in (2.4) a well-defined quantity. Because of (i), notably, a strong solution Z must be adapted to $(\mathcal{F}_t, t \geq 0)$: for instance, if $\mathcal{F}_t$ is the σ-field generated by B^Z, V and z_0, then the values of Z_t are determined by z_0 and by the values of B^Z and V up to t. Adaptedness to $(\mathcal{F}_t, t \geq 0)$ can be too strict and *weak*

solutions $\left(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P}, B^Z, V, z_0, Z\right)$ which satisfy SDE (2.3) can be considered, in place of seeking only process Z.

Uniqueness of solutions to an SDE can arise in different forms. One such form is pathwise uniqueness, which in the context of strong solutions is defined as:

Definition 4 (Pathwise unique strong solutions [25, Section 5.1])**.** *Pathwise uniqueness of strong solutions to SDE (2.3) holds if for any two strong solutions* $Z = (Z_t, t \geq 0)$ *and* $Z' = (Z'_t, t \geq 0)$ *in the sense of Definition 3, we have that*

$$\mathbb{P}\left(\omega \in \Omega : \; Z_t(\omega) = Z'_t(\omega) \;\; \textit{for all } t \in [0, \infty)\right) = 1.$$

For completeness, establishing existence and pathwise uniqueness of strong solutions of SDE (2.3) requires imposing the following condition on c_t, d_t, and V:

Condition A. *For all* $t \in [0, \infty)$*,* $v \in D_V$*, and pairs* $z, z' \in C_Z$*, there exists a constant* $\alpha > 0$ *independent of* t*,* v*,* z *and* z'*, such that*

$$\|c_t(z, v) - c_t(z', v)\|^2 + \|d_t(z, v) - d_t(z', v)\|^2 \leq \alpha \|z - z'\|_t^2, \tag{2.5}$$

and

$$\|c_t(z, v)\|^2 + \|d_t(z, v)\|^2 \leq \alpha \left(1 + \|z\|_t^2 + \|M_t(v)\|^2\right), \tag{2.6}$$

in which $\|a\|_t := \sup_{s \in [0,t]} \|a_s\|$*, for* $a = z - z'$ *(inequality (2.5)) and* $a = z$ *(inequality (2.6)), and* M_t *is a* $\mathcal{B}(D_V)$*-measurable functional such that* $\mathbb{E}\left[\int_0^\infty \|M_t(V)\|^2 dt\right] < \infty$*.*

Then, we have a theorem [25, Theorem 5.1.1], which we re-state here without proof.

Theorem 1 (Existence and pathwise uniqueness of strong solutions)**.** *Suppose that* $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$*,* c_t*,* d_t*,* V*,* B^Z*, and* z_0 *are given, with* c_t *and* d_t *non-anticipative. If Condition A is satisfied, then SDE (2.3) possesses a pathwise unique strong solution.*

To clarify the aforementioned notions, we provide an illustrative example, which also introduces an SDE form used in later sections.

Example 1 (SDE with additive noise)**.** *Consider a* $\mathbb{R}$*-valued process* Z*, an* 1*-dimensional standard Brownian motion process* B^Z*, and non-anticipative coefficients* c_t *and* $d_t(Z, V) = \varepsilon$ *with* $\varepsilon \in \mathbb{R}$*. Let the following SDE relate, additively,* B^Z *and* Z*:*

$$dZ_t = c_t(Z, V)\, dt + \varepsilon dB_t^Z, \quad Z_0 = z_0.$$

For $c_t \equiv 0$*, Condition A holds: the resulting SDE possesses a pathwise unique strong solution by Theorem 1, and such solution takes the common form* $Z_t = z_0 + \varepsilon B_t^Z$*.*

2.1.2. Exit times. In certain applications, it may be of interest to determine if, for $\mathbb{P}$-almost every $\omega \in \Omega$, the first time t_{e} when a sample path $Z(\omega)$ exits a certain open set $S \subset \mathbb{S}_Z$ is finite, or, equivalently, if $\mathbb{P}$-almost surely $t_{\mathrm{e}} = \infty$. With respect to our application (Section 3), this information is relevant for establishing whether the solution to an SDE only has positive sample paths; that is, $S \subseteq \mathbb{R}_+^{N_Z}$.
The example below illustrates the notion of positivity of the solution of an SDE.

Example 2 (SDE with multiplicative noise)**.** *Consider a process Z, a 1-dimensional standard Brownian motion process B^Z, and non-anticipative $c_t \equiv 0$ and $d_t(Z, V) = \varepsilon Z_t$ with $\varepsilon \in \mathbb{R}$. Let the following SDE relate, multiplicatively, B^Z and Z:*

$$dZ_t = \varepsilon Z_t dB_t^Z, \quad Z_0 = z_0. \tag{2.7}$$

It can be shown that a pathwise unique strong solution to SDE (2.7) exists, and that for any positive-valued random variable z_0, a sample path $Z(\omega) = (Z_t(\omega), t \geq 0)$ has positive real values for $\mathbb{P}$-almost every $\omega \in \Omega$.

Itô's formula [26, Theorem 6] can be used on functions from $\mathbb{S}_Z$ with continuous first and second derivatives. For the logarithm function, SDE (2.7) yields

$$d(\ln(Z_t)) = \varepsilon dB_t^Z - 1/2\varepsilon^2 t, \quad \ln(Z_0) = \ln(z_0).$$

Therefore, $Z_t = z_0 \exp\left(\varepsilon B_t^Z - 1/2\varepsilon^2 t\right)$, assuring that Z is constrained to $S \subseteq \mathbb{R}_+$.

The Khasminskii test [27] can be used to answer exit-time questions. The test is based on defining 1.) an increasing sequence of open sets $\{S_r\}$, r integer, whose closures are contained in S and such that their union equals S; and 2.) an associated sequence of stopping times $\{t_r\}$ with $t_r := \inf\{t \in [0, t_e) : Z_t \notin S_r\}$, where $\inf \emptyset = \infty$. By letting $t_\infty := \lim_{r\to\infty} t_r$, it can be shown that $\mathbb{P}$-almost surely $t_\infty \leq t_e$, whereas to verify if $Z_t \in S$ always, it suffices to show that no exit occurs at finite times ($t_\infty = \infty$).

2.1.3. Markovianity of the solution. In this section, the discussion is restricted to Markov processes, whose past and future are statistically independent given the present. If a solution to SDE (2.3) is Markovian, we can describe it in terms of certain transition probabilities. To formally define these notions, we consider the expectation of a stochastic process, conditional on the past information encoded in an appropriate collection of σ-fields. Then we have the following definition:

Definition 5 (Markov process)**.** *Let $Z = (Z_t, t \geq 0)$ be a measurable process defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$. Z is a Markov process relative to $(\mathcal{F}_t, t \geq 0)$ if for all $s \leq t$ and all bounded measurable functions ϕ, we have*

$$\mathbb{E}[\phi(Z_t) | \mathcal{F}_s] = \mathbb{E}[\phi(Z_t) | Z_s].$$

Equivalently, we can say that $\mathbb{P}(Z_t \in A_Z | \mathcal{F}_s) = \mathbb{P}(Z_t \in A_Z | Z_s)$, for any $A_Z \in \mathcal{B}(\mathbb{S}_Z)$.

In Definition 5, by letting $s = 0$, the focus is on expectations that are conditional on the $\mathcal{F}_0$-measurable N_Z-dimensional random variable Z_0, with an interest in the evolution in time of $\mathbb{E}[\phi(Z_t) | \mathcal{F}_0]$. Without loss of generality, we discuss this evolution for a specific case of SDE (2.3) with no auxiliary process V:

$$dZ_t = c_t(Z_t)\, dt + d_t(Z_t)\, dB_t^Z, \quad Z_0 = z_0, \tag{2.8}$$

whose solution $Z = (Z_t, t \geq 0)$, for a deterministic initial condition (the singleton $z_0 \in \mathbb{S}_Z$), exists, and is unique and Markov relative to $(\mathcal{F}_t, t \geq 0)$ [28, Theorem 17.2.3].

Under this setting, we construct an equivalent description of Z in terms of a Markov transition function $\mathbb{P}(\cdot, \cdot; \cdot, \cdot)$ of s, z_s, t and A_Z, defined for $0 \leq s < t$, $z_s \in \mathbb{S}_Z$,

$A_Z \in \mathcal{B}(\mathbb{S}_Z)$. For fixed s, t, and A_Z, we have that $\mathbb{P}(s, \cdot; t, A_Z)$ is a measurable function of z_s. For fixed s, t, and z_s, we have that $\mathbb{P}(s, z_s; t, \cdot)$ is a probability measure in A_Z; that is, $\mathbb{P}(s, z_s; t, A_Z) = \mathbb{P}(Z_t \in A_Z | Z_s = z_s)$, the probability of Z being in A_Z at time t, conditional on starting from z_s at time s [29, Section 6]. Consequently, using again Definition 5, for all bounded measurable functions ϕ on $\mathbb{S}_Z$, we have

$$\begin{aligned} \mathbb{E}[\phi(Z_t) | \mathcal{F}_0] &= \mathbb{E}[\phi(Z_t) | Z_0] \\ &= \int_{\mathbb{S}_Z} \phi(z) \, \mathbb{P}(0, z_0; t, dz). \end{aligned} \tag{2.9}$$

Let $p(t, \cdot)$ denote the probability density of $\mathbb{P}(0, z_0; t, \cdot)$ with respect to the Lebesgue measure. Differentiating (2.9) with respect to time yields an equation for the evolution of $\mathbb{E}[\phi(Z_t) | \mathcal{F}_0]$. The dynamics are deterministic and given in terms of p:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\phi(Z_t) | \mathcal{F}_0] &= \int_{\mathbb{S}_Z} \phi(z) \frac{\partial}{\partial t} p(t, z) \, dz \\ &= \int_{\mathbb{S}_Z} \phi(z) \, \mathcal{A}_t^* p(t, z) \, dz, \end{aligned} \tag{2.10}$$

where $\mathcal{A}_t^*$ is the operator which describes the dynamics of p [28, Section 17]. Additional characterisations [30, Section 4] lead to the Kolmogorov forward equation.

2.1.4. The Kolmogorov forward equation (KFE). In this section, we express expectations of functions of the solution to an SDE in terms of partial differential equations (PDEs). In particular, we present the forward Kolmogorov equation, describing the evolution of the probability density of a process solving SDE (2.8).

Assume the solution Z of SDE (2.8) to be a Markov process relative to $(\mathcal{F}_t, t \geq 0)$, with probability density $p(t, \cdot)$ at time t. For a given initial probability density p_0 on $\mathbb{S}_Z$, p satisfies the Kolmogorov forward, or Fokker-Planck, equation

$$\frac{\partial}{\partial t} p(t, z) = -\sum_{n_Z=1}^{N_Z} \frac{\partial}{\partial z_{n_Z}} \left[c_t^{n_Z}(z) \, p(t, z) \right] + \frac{1}{2} \sum_{n_Z=1}^{N_Z} \sum_{n'_Z=1}^{N_Z} \frac{\partial^2}{\partial z_{n_Z} \partial z_{n'_Z}} \left[D_t^{n_Z, n'_Z}(z) \, p(t, z) \right], \tag{2.11}$$

with entries $D_t^{n_Z, n'_Z}(z) = \sum_{n_{BZ}=1}^{N_{BZ}} d_t^{n_Z, n_{BZ}}(z) \, d_t^{n'_Z, n_{BZ}}(z)$ of the diffusion tensor, and initial condition $p(0, z) = p_0(z)$, provided p, c_t and D_t are sufficiently smooth. We recognise that the operator $\mathcal{A}_t^*$ in (2.10) is then adjoint to the (infinitesimal) generator of Z, whose coefficients are associated to those of SDE (2.8).

KFE: Solutions

The density p associated to a Markov process Z which solves SDE (2.8) is a *fundamental solution* to PDE (2.11). That is, with an arbitrary but fixed $T > 0$, for every $s \in [0, t)$, $t \in (0, T]$, $z_s \in \mathbb{S}_Z$, $z \in \mathbb{S}_Z$, and measurable non-negative functions p_s on $\mathbb{S}_Z$, we have:

i. $p(t, z) = \int_{\mathbb{S}_Z} p_s(z) \, \mathbb{P}(s, dz_s; t, z)$;
ii. $\lim_{t \to s^+} p(t, z) = p_s(z)$;

iii. p is continuously differentiable, once in t and twice in z;

iv. p satisfies PDE (2.11).

Existence and pathwise uniqueness of a strong solution Z of SDE (2.8), along with (2.6), ensure that the fundamental solution p, if it exists, is unique [23, Section 5.7]. We refer the reader to [31, Section 6.4] for various regularity and ellipticity conditions on c_t and D_t sufficient to guarantee existence of a fundamental solution.

The density p may also correspond to a *classical solution*. That is, p satisfies (ii)-(iv) above and there exist constants $\alpha, \beta > 0$ such that for every $z \in \mathbb{S}_Z$

$$\max_{t\in[0,T]} p\left(t,z\right) \leq \alpha \exp\left(\beta \|z\|^2\right). \tag{2.12}$$

A sufficient condition is that coefficients $c_t^{n_Z}$ and $D_t^{n_Z,n'_Z}$ are bounded and Hölder-continuous on $\mathbb{S}_Z$, and the diffusion tensor D_t is uniformly elliptic; that is, for every pair $z, z' \in \mathbb{S}_Z$, there exist a constant $\alpha > 0$ such that

$$\sum_{n_Z=1}^{N_Z} \sum_{n'_Z=1}^{N_Z} D_t^{n_Z,n'_Z}\left(z\right) z'_{n_Z} z'_{n'_Z} \geq \alpha \|z'\|. \tag{2.13}$$

In cases for which only (2.13) is satisfied, alternative sufficient conditions can ensure the existence of a unique solution with a polynomial growth instead. That is, p satisfies PDE (2.11) and there exist constants $\alpha, \beta > 0$ such that for every $z \in \mathbb{S}_Z$

$$\max_{t\in[0,T]} p\left(t,z\right) \leq \alpha \left(1 + \|z\|^{\beta}\right). \tag{2.14}$$

The conditions are that 1.) the diffusion tensor D_t is uniformly elliptic, 2.) the coefficients $c_t^{n_Z}$ and $D_t^{n_Z,n'_Z}$ satisfy (2.6) for compact subsets of $\mathbb{S}_Z$, and 3.) the functions $(\partial/\partial z_{n_Z}) c_t^{n_Z}$, $(\partial/\partial z_{n_Z}) D_t^{n_Z,n'_Z}$ and $(\partial^2/\partial z_{n_Z}\partial z_{n'_Z}) D_t^{n_Z,n'_Z}$ are bounded in $\mathbb{S}_Z$ and Hölder-continuous on compact subsets of $\mathbb{S}_Z$ [31, p. 147].

KFE: Boundary conditions

The derivation to show that the adjoint operator $\mathcal{A}_t^*$ in (2.10) has coefficients associated to those in SDE (2.8) requires Itô's formula on a continuous test function ϕ on $\mathbb{S}_Z$ with compact support [32]. This ensures that boundary terms in subsequent integration by parts vanish. Boundary terms also vanish when selecting appropriate conditions at the boundary ∂S of a set $S \subseteq \mathbb{S}_Z$ inside which z is constrained. Boundary conditions (BCs) take on a variety of forms, but note that (2.11) can always be re-written as

$$\frac{\partial}{\partial t} p\left(t,z\right) = -\sum_{n_Z=1}^{N_Z} \frac{\partial}{\partial z_{n_Z}} J^{n_Z}(t,z),$$

where $J = \left(J^1, \ldots, J^{n_Z}, \ldots, J^{N_Z}\right)$ is the probability current with components

$$J^{n_Z}\left(t,z\right) = c_t^{n_Z}\left(z\right) p\left(t,z\right) - \frac{1}{2} \sum_{n'_Z=1}^{N_Z} \frac{\partial}{\partial z_{n'_Z}} \left[D_t^{n_Z,n'_Z}\left(z\right) p\left(t,z\right)\right].$$

Once J has been defined, different BCs can be considered separately. With respect to our application (Section 3), the BCs of interest are

i. Reflecting boundary: $J(t,z)\cdot n = 0$ for all $t \in [0,\infty)$, $z \in \partial S$, and n normal to ∂S; that is, the net of probability current across ∂S is zero;
ii. Natural boundary: assume the existence of a $z \in \partial S$ for which $c_t(z) = 0$ and $d_t(z) = 0$. If process Z were to reach z, it would remain there. This scenario is called a natural boundary whenever process Z never reaches z.

2.2. Stochastic filtering theory

Thus far, we have worked with a model describing the prior probabilistic knowledge about how the process Z evolves in time, formally embedded by $\mathcal{A}_t^*$ in the KFE. The focus now shifts to the filtering problem. This section begins with the introduction of the following processes on Ω: $X = (X_t, t \geq 0)$ called the state process and $Y = (Y_t, t \geq 0)$, the observation process. Similarly to Section 2.1, we let $C_Y = C_Y([0,\infty), \mathbb{S}_Y)$ denote the space of continuous functions on $[0,\infty)$ taking values in $\mathbb{S}_Y$, $\mathbb{S}_Y \subseteq \mathbb{R}^{N_Y}$, and $D_X = D_X([0,\infty), \mathbb{S}_X)$ denote the space of $\mathbb{S}_X$-valued right-continuous functions on $[0,\infty)$ with left-hand limits, $\mathbb{S}_X$ being a complete separable metric space. For convenience, we let $\Omega = D_X \times C_Y$. Then any $\omega \in \Omega$ can be represented as a pair of sample paths $\omega = (\omega_X, \omega_Y)$, where $\omega_X \in D_X$ and $\omega_Y \in C_Y$. Lastly, we introduce the noise processes $B^X = \left(B_t^X, t \geq 0\right)$ and $B^Y = \left(B_t^Y, t \geq 0\right)$. We have not taken into account the existence of B^X and B^Y when constructing Ω, but as discussed in [33], one can extend Ω to include all ω associated to the remaining random variables and stochastic processes.

The state process describes the dynamics of a system of interest. It is modelled as a special case of SDE (2.3):

$$dX_t = f_t\left(X_{s,s\leq t}, Y_{s,s\leq t}\right) dt + g_t\left(X_{s,s\leq t}, Y_{s,s\leq t}\right) dB_t^X, \quad X_0 = x_0, \tag{2.15}$$

where X has sample paths $X(\omega) = \omega_X$, Y is a stochastic process soon to be introduced, with $Y(\omega) = \omega_Y$, process $B^X = \left(B^{X,1}, \ldots, B^{X,N_{B^X}}\right)$ is an N_{B^X}-dimensional standard Brownian motion, coefficients c_t and d_t have been respectively renamed f_t and g_t, with $f_t : D_X \times C_Y \to \mathbb{R}^{N_X}$, $g_t : D_X \times C_Y \to \mathbb{R}^{N_X \times N_{B^X}}$, and X_0 is given.

Generally, the state process cannot be directly observed (for instance, only certain components of a vector-valued process are observed). Moreover, observations are prone to error. These two aspects are represented in the following observation model:

$$dY_t = h_t\left(Y_{s,s\leq t}, X_{s,s\leq t}\right) dt + k_t\left(Y_{s,s\leq t}, X_{s,s\leq t}\right) dB_t^Y, \quad Y_0 = 0, \tag{2.16}$$

where coefficients c_t and d_t have been respectively renamed h_t and k_t, with $h_t : C_Y \times D_X \to \mathbb{R}^{N_Y}$ being the observation function and $k_t : C_Y \times D_X \to \mathbb{R}^{N_Y \times N_{B^Y}}$ denoting noise intensity, and the N_{B^Y}-dimensional standard Brownian motion process $B^Y = \left(B^{Y,1}, \ldots, B^{Y,N_{B^Y}}\right)$ is independent of B^X. We let $N_{B^Y} = N_Y$. Notably, SDE (2.16) describes the evolution of an integral form of the observations: Y is the integral process corresponding to the actually observed process $\left\{y_{t_{n_t}}\right\}_{n_t=1}^{N_t}$, where $0 < t_1 < \ldots < t_{N_t} < \infty$

is a uniform partition of a subset of $(0, \infty)$. Therefore, Y_0 being zero does not imply zero-valued observations at $t = 0$, but rather that no information is initially available.

2.2.1. Stochastic equation for the filter. In filtering, we define a stochastic process $\pi = (\pi_t, t \geq 0)$ for the conditional distribution of X_t given the observation σ-field $\mathcal{F}_t^Y = \sigma(Y_s, s \leq t)$:

$$\pi_t(A_X) := \mathbb{P}\left(X_t \in A_X | \mathcal{F}_t^Y\right) \tag{2.17}$$

$$\pi_t(\phi) \quad := \mathbb{E}_{\mathbb{P}}\left[\phi(X_t) | \mathcal{F}_t^Y\right], \tag{2.18}$$

where A_X is an arbitrary set in $\mathcal{B}(\mathbb{S}_X)$, ϕ is a measurable scalar function, and we explicitly show the measure with respect to which expectations are taken. π_t in (2.17) takes values from $\mathcal{B}(\mathbb{S}_X)$ into $[0, 1]$. π_t in (2.18) gets information about the state through ϕ, assuming that the expectation with respect to $\mathbb{P}$ exists.

The expectation process $\pi(\phi) = (\pi_t(\phi), t \geq 0)$, hereafter called filter, satisfies

$$\pi_t(\phi) = \mathbb{E}_{\mathbb{P}}[\phi(X_0)] + \mathcal{A}_t^*(\phi) + \int_0^t \Phi_s^\top d\Lambda_s, \tag{2.19}$$

where $\mathcal{A}_t^*(\phi)$ is a measurable process adapted to $(\mathcal{F}_t, t \geq 0)$, Φ_s is a measurable process adapted to $(\mathcal{F}_s, s \leq t)$, the integration is with respect to $(\Lambda_s, s \leq t)$, known as the innovation process, and $\Phi_s^\top$ denotes transpose of Φ_s. The required conditions for solvability of (2.19) for every measurable scalar ϕ are stated in [25, Theorem 8.4.1]. Since these conditions rely on additional technical definitions and cannot be easily expressed using the existing notation, they are not explicitly stated here. The remaining problem is to discover the particular form that $A_t^*(\phi)$, Φ_s and Λ_s take in (2.19) for specific choices of SDEs (2.15) and (2.16).

2.2.2. The Kushner-Stratonovich equation (KSE). We proceed by assuming the existence of a density $p_t = p(t, \cdot)$ of π_t with respect to the Lebesgue measure, for which $\pi_t(\phi)$ as defined in (2.18) is obtained via $\int_{\mathbb{S}_X} \phi(x) p_t(x) dx$. A detailed derivation of the evolution equation for this conditional density is available in [1, 33]. Since the derivation uses arguments which are standard in the theory of Markov processes, we shall state only the outline. Two common approaches for deducing this evolution equation are the change of measure and innovation process methods. We present the former method, following similar steps as in [34]. For the latter, the interested reader can refer to [25].

Changes of measure

Let us consider two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $\mathcal{F}$. We recall that $\mathbb{P}$ is called *absolutely continuous* with respect to $\mathbb{Q}$, written as $\mathbb{P} \ll \mathbb{Q}$, if every null-set of $\mathbb{P}$ is a null-set of $\mathbb{Q}$. We restrict the definition of the measure $\mathbb{P}$ to $\mathcal{F}_t$, written as $\mathbb{P}_t$; that is, $\mathbb{P}$ is restricted to all the events that can be described in terms of the behaviour of $(X_s, s \leq t)$ and $(Y_s, s \leq t)$. Let us assume the existence of a $\mathbb{Q}$ such that $\mathbb{P}_t \ll \mathbb{Q}_t$ for all

finite t. Then the Radon–Nikodym theorem [35, Section 6] guarantees the existence of a random variable L_t, unique for $\mathbb{Q}$-almost every $\omega \in \Omega$, satisfying

$$\mathbb{E}_{\mathbb{P}}\left[\phi\left(X_t\right)\right] = \mathbb{E}_{\mathbb{Q}}\left[\phi\left(X_t\right) L_t\right],$$

where $L_t := \frac{d\mathbb{P}_t}{d\mathbb{Q}_t}$ is called the Radon–Nikodym derivative of $\mathbb{P}_t$ with respect to $\mathbb{Q}_t$. In terms of $\mathbb{Q}$, the conditional expectation of ϕ, i.e. the filter $\pi_t\left(\phi\right)$, is written as

$$\begin{aligned}\pi_t(\phi) &= \frac{\mathbb{E}_{\mathbb{Q}}\left[\phi\left(X_t\right)\ L_t \mid \mathcal{F}_t^Y\right]}{\mathbb{E}_{\mathbb{Q}}\left[L_t \mid \mathcal{F}_t^Y\right]} \\ &=: \frac{\rho_t\left(\phi\right)}{\rho_t\left(\mathbb{1}\right)},\end{aligned} \tag{2.20}$$

where $\rho_t\left(\mathbb{1}\right)$ can be viewed as the normalising factor.

(2.20) is known as the Bayes' formula for stochastic processes or *Kallianpur-Striebel formula*. The hope is that L_t is simple enough to make $\mathbb{E}_{\mathbb{P}}\left[\phi(X_t)|\mathcal{F}_t^Y\right]$ more tractable to compute via the right-hand side of (2.20) than via an integration over $\mathbb{P}$. For instance, some simplification might be achieved by switching from a model with restriction $\mathbb{P}_t$ in which X and Y are coupled to a model with restriction $\mathbb{Q}_t$ in which they are independent, while preserving the distribution of X_t. For our filtering problem, we will assume the existence of a $\mathbb{Q}$ under which $(Y_s, s \leq t)$ is a standard Brownian motion process independent of $(X_s, s \leq t)$ for all finite t.

Let us assume that the pair (X, Y) solves a special case of SDEs (2.15) and (2.16):

$$dX_t = f\left(X_t\right) dt + g\left(X_t\right) dB_t^X, \quad X_0 = x_0 \tag{2.21}$$

$$dY_t = h\left(X_t\right) dt + K dB_t^Y, \quad Y_0 = 0, \tag{2.22}$$

that is, the case in which $f_t = f$, $g_t = g$ and $h_t = h$ are autonomous functions of the state X_t. Since the coefficients are autonomous, the infinitesimal operator and its adjoint become time-independent, denoted as $\mathcal{A}$ and $\mathcal{A}^*$, respectively. The observation process is driven by additive noise, i.e. $k_t(X_{s,s\leq t}, Y_{s,s\leq t}) = K$, where K is a matrix of size $N_Y \times N_Y$. Notably, the filtering problem whose observation model is driven by multiplicative noise can be converted into a problem whose observation model is driven by additive noise, via a suitably chosen stochastic flow mapping [36].

We operate under a few more technical conditions:

i. K is invertible;
ii. Condition A holds for SDE (2.21), such that this SDE has a unique strong solution $X = (X_t, t \geq 0)$ by Theorem 1;
iii. Condition A also holds for SDE (2.22) when $M_t(X) = h(X_t)$.

If we define the process $L = (L_t, 0 \leq t < \infty)$ by

$$L_t = \exp\left(\int_0^t \left(K^{-1} h\left(X_s\right)\right)^\top K^{-1} dY_s - \frac{1}{2}\int_0^t \left\|K^{-1} h\left(X_s\right)\right\|^2 ds\right), \tag{2.23}$$

then conditions (i)-(iii) are sufficient to allow L to be a standard Brownian motion process adapted to $(\mathcal{F}_t^Y, 0 \leq t < \infty)$, independent of X. Additionally, the law of the state process X under $\mathbb{Q}$ is the same as the law under $\mathbb{P}$ [1, Proposition 3.13].

(2.23), along with Itô's formula, leads to the following proposition:

Proposition 1. *Let ϕ have continuous partial derivatives of every order up to 2, all of which are bounded. Then we can write*

$$\pi_t(\phi) = \pi_0(\phi) + \int_0^t \pi_s(\mathcal{A}^*\phi)ds + \int_0^t \left(\pi_s(\phi\tilde{h}) - \pi_s(\phi)\pi_s(\tilde{h})\right)^\top \left(d\tilde{Y}_s - \pi_s(\tilde{h})\, ds\right), \quad (2.24)$$

where $\pi_0(\phi) = \mathbb{E}_{\mathbb{P}}[\phi(X_0)] = \mathbb{E}_{\mathbb{Q}}\left[\phi(X_0) L_0 | \mathcal{F}_0^Y\right]$. Hereafter, $\tilde{Y} = (\tilde{Y}_t, t \geq 0)$ is the rescaled observation process (i.e. $\tilde{Y}_t = K^{-1}Y_t$, $t \geq 0$), and $\tilde{h}$ is the rescaled observation function (i.e. $\tilde{h}(x) = K^{-1}h(x)$).

Moreover, an application of integration by parts to (2.24) results in the so-called Kushner-Stratonovich equation (KSE)

$$dp_t(x) = \mathcal{A}^* p_t(x)\, dt + p_t(x)\left(\tilde{h}(x) - \pi_t(\tilde{h})\right)^\top \left(d\tilde{Y}_t - \pi_t(\tilde{h})\, dt\right). \quad (2.25)$$

Proposition 1 concerns the evolutions of π and p in the special case of (2.19) for which $A_t^*(\phi) = \int_0^t \pi_s(\mathcal{A}^*\phi)ds$, $\Phi_s = \pi_s(\phi\tilde{h}) - \pi_s(\phi)\pi_s(\tilde{h})$, and $d\Lambda_s = d\tilde{Y}_s - \pi_s(\tilde{h})ds$. We refer the reader to [25] for the detailed derivations of (2.24) and PDE (2.25). For completeness, we outline the proof in the Supplementary Material.

KSE: Pathwise Filtering and Solutions

So far, $\pi_t(\phi)$ has been defined on an unspecified set of $\mathcal{B}(C_Y)$ of measure one. We have not yet shown $\pi_t(\phi)$ in the case of a given rescaled observation path $\tilde{Y}(\omega)$. To this end, we present the notion of pathwise filtering from [4, 37].

L_t in (2.23) is written in terms of a stochastic integral with process $\tilde{Y}$ as the integrator. However, since the integrand is independent of the observations, we reverse the role of processes X and $\tilde{Y}$: up to some finite time t, L_t is obtained by an integral whose integrator is $(X_s(\omega), s \leq t)$. By integration by parts,

$$\int_0^t \tilde{h}^\top(X_s)\, d\tilde{Y}_s = \tilde{h}^\top(X_t)\tilde{Y}_t - \int_0^t \tilde{Y}_s^\top d\tilde{h}(X_s) \quad \text{almost surely.}$$

The pathwise counterpart to (2.23) is then

$$\hat{L}_t = \exp\left[\tilde{h}^\top(X_t)\tilde{Y}_t - \int_0^t \tilde{Y}_s^\top d\tilde{h}(X_s) - \frac{1}{2}\int_0^t \|\tilde{h}(X_s)\|^2 ds\right], \quad (2.26)$$

where $\hat{L}_t$ corresponds to L_t when computed with respect to a given observation path.

The KSE in Proposition 1 is based on the Radon–Nikodym derivative L_t from (2.23). Similarly, a pathwise KSE can be derived based on $\hat{L}_t$ from (2.26). This reduces the stochastic PDE (2.25) to a PDE with a random coefficient:

$$dp_t(x) = \mathcal{A}^* p_t(x)\, dt + p_t(x)\, \hat{e}_t(\tilde{Y}_t), \quad (2.27)$$

where $\hat{e}_t : \mathbb{S}_Y \to \mathbb{R}$ is related to $\hat{L}_t$ (see [5, 33]). Once $\tilde{Y}(\omega)$ has been fixed, a solution to PDE (2.27) can be characterised as in Section 2.1.4, under additional regularity conditions on h, such as continuity and boundedness. Hereafter, $p = (p_t, t \geq 0)$ will always refer to this pathwise solution.

KSE: Pathwise Approximations to Solutions

This section contains an overview of two classes of methods to approximate p. We briefly describe each class and state related results.

Let the uniform partition of the interval $[0, \infty)$ have time step $\Delta t = t_{n_t} - t_{n_t-1}$. If a set of rescaled observations $\{\tilde{y}_{t_{n_t}}\}_{n_t=1}^{N_t}$ is available, the density $p_{t_{n_t}}$ will be approximated by $p_{n_t}^{\Delta t}$ by first defining $\tilde{Y}_{t_{n_t}} := \tilde{Y}_{t_{n_t-1}} + \tilde{y}_{t_{n_t}} \Delta t$. This corresponds to a continuous-discrete filtering problem, in which the observation process $\tilde{Y}$ from the continuous formulation is interpolated by straight-line segments between discrete observation times. Secondly, $p_{n_t-1}^{\Delta t}$ is propagated to $p_{n_t}^{\Delta t}$ according to one of the following methods:

- Particle filtering
- Finite-differences + Splitting-up approximation

Particle filtering: The density $p_{n_t}^{\Delta t}$ is constructed by density estimation techniques [38], such as histograms or kernel estimation, over a finite number of samples updated sequentially up to time t_{n_t}. The following exposition is adapted from [1]:

Let $\{X^{(n_P)}\}_{n_P=1}^{N_P}$ be a collection of N_P mutually independent stochastic processes, all independent of $\tilde{Y}$ and each being a solution to SDE (2.21). Then the pairs $\{(X^{(n_P)}, \tilde{Y})\}_{n_P=1}^{N_P}$ are identically distributed and have the same distribution as the pair $(X, \tilde{Y})$ under $\mathbb{Q}$. Additionally, for $n_P = 1, \ldots, N_P$, $0 \leq t < \infty$, define $w_t^{(n_P)}$ as

$$w_t^{(n_P)} = \exp\left(\tilde{h}^\top\left(X_t^{(n_P)}\right)\tilde{Y}_t - \int_0^t \tilde{Y}_s^\top d\tilde{h}\left(X_s^{(n_P)}\right) - \frac{1}{2}\int_0^t \|\tilde{h}\left(X_s^{(n_P)}\right)\|^2 ds\right).$$

Then the triples $\{(X^{(n_P)}, w^{(n_P)}, \tilde{Y})\}_{n_P=1}^{N_P}$ are identically distributed and have the same distribution as the triple $(X, L, \tilde{Y})$ under $\mathbb{Q}$.

For $n_P = 1, \ldots, N_P$, $w^{(n_P)}$ can be computed recursively according to

$$w_{t_{n_t}}^{(n_P)} = w_{t_{n_t-1}}^{(n_P)} \exp\left(\tilde{h}^\top\left(X_s^{(n_P)}\right)\tilde{Y}_s\Big|_{t_{n_t-1}}^{t_{n_t}} - \int_{t_{n_t-1}}^{t_{n_t}} \tilde{Y}_s^\top d\tilde{h}\left(X_s^{(n_P)}\right) - \frac{1}{2}\int_{t_{n_t-1}}^{t_{n_t}} \|\tilde{h}\left(X_s^{(n_P)}\right)\|^2 ds\right)$$

whenever a new rescaled observation $\tilde{y}_{t_{n_t}}$ becomes available.

Lastly, the N_P triples may also be used to construct the measure-valued processes $\rho^{N_P} = (\rho_t^{N_P}, t \geq 0)$ and $\pi^{N_P} = (\pi_t^{N_P}, t \geq 0)$:

$$\rho_t^{N_P} := \frac{1}{N_P}\sum_{n_P=1}^{N_P} w_t^{(n_P)}\delta_{X_t^{(n_P)}}, \quad \pi_t^{N_P} := \frac{\rho_t^{N_P}}{\rho_t^{N_P}(\mathbb{1})} = \sum_{n_P=1}^{N_P} \bar{w}_t^{(n_P)}\delta_{X_t^{(n_P)}},$$

where $\bar{w}_t^{(n_P)}$ is given by $\bar{w}_t^{(n_P)} = \left(\sum_{n_P'=1}^{N_P} w_t^{(n_P)'}\right)^{-1} w_t^{(n_P)}$, $n_P = 1, \ldots, N_P$, $0 \leq t < \infty$, and δ is the Dirac's delta generalised function; that is, $\rho_t^{N_P}$ is the empirical measure of N_P particles with positions $\{X_t^{(n_P)}\}_{n_P=1}^{N_P}$, and $\pi_t^{N_P}$ is its normalised version. Results on the convergence to ρ_t and π_t are detailed in, for instance, [1].

Finite-differences + Splitting-up approximation: a method for solving stochastic PDEs using tools inherited from solving classical PDEs. The following exposition follows closely the description given in [39]:

The transition from $p_{n_t-1}^{\Delta t}$ to $p_{n_t}^{\Delta t}$ follows a splitting-up approximation [40, 41], whose first step, called the *prediction* step, consists in solving the KFE (2.11) for a single time step Δt. In other words, we solve

$$\frac{\partial}{\partial t} p_t(x) = \mathcal{A}^* p_t(x), \quad t \in [t_{n_t-1}, t_{n_t}], \quad p_{t_{n_t-1}} = p_{n_t-1}^{\Delta t}. \tag{2.28}$$

We then define $\bar{p}_{n_t}^{\Delta t}$ as the solution evaluated at $t = t_{n_t}$, i.e. $\bar{p}_{n_t}^{\Delta t} := p_{t_{n_t}}$.

The second step, called the *correction* step, uses (2.26) and the new rescaled observation $\tilde{y}_{t_{n_t}}$ to update $\bar{p}_{n_t}^{\Delta t}$. We obtain $p_{n_t}^{\Delta t}(x)$ via $p_{n_t}^{\Delta t}(x) := \xi_{n_t} \zeta_{n_t}^{\Delta t}(x) \bar{p}_{n_t}^{\Delta t}$, where ξ_{n_t} is a normalising constant chosen such that $\int_{\mathbb{S}_X} p_{n_t}^{\Delta t}(x)\, dx = 1$, and

$$\zeta_{n_t}^{\Delta t}(x) := \exp\left(-\frac{1}{2}\Delta t \|\tilde{y}_{n_t} - \tilde{h}(x)\|^2\right). \tag{2.29}$$

Results on the convergence to $p = (p_t, t \geq 0)$ are detailed in, for instance, [42].

3. The chemostat

Let us consider a bioreactor in which one or more populations of microorganisms grow in a nutrient medium consisting of a cocktail of molecules. The chemostat (Figure 1) is a laboratory device representing this bioreactor [43]. It is widely used in experimental design due to its ability to provide controlled conditions and reproducible results [17].

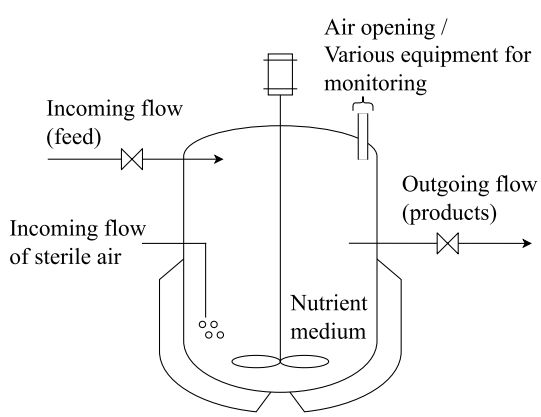

Figure 1: Schematic representation of a chemostat.

In this work, biomass refers to a single population of growing microorganisms and substrate to the cocktail of molecules. The production of biomass may result in the additional byproducts methane and carbon dioxide, collectively referred to as biogas. Under this scenario, the following reaction takes place to produce one unit of biomass and γ units of biogas from κ units of substrate:

$$\kappa\,\text{substrate} \longrightarrow \text{biomass} + \gamma\text{biogas}. \tag{3.30}$$

We begin with a differential description of the substrate and biomass concentrations in a chemostat, based on reaction (3.30).

Let $b = (b_t,\, t \geq 0)$ and $s = (s_t,\, t \geq 0)$ denote, respectively, the time evolution of concentrations of biomass and substrate (hereafter, s refers to a concentration process, rather than a time variable as in Section 2). Let $x = (x_t,\, t \geq 0)$ denote the state process, that is $x = (b, s)$. Then the differential equation for the dynamics of the chemostat is

$$d \underbrace{\begin{pmatrix} b_t \\ s_t \end{pmatrix}}_{x_t} = \underbrace{\begin{pmatrix} (\mu(\cdot) - D)\, b_t + D b_t^{in} \\ D\,(s_t^{in} - s_t) - \kappa \mu(\cdot) b_t \end{pmatrix}}_{f(x_t)} dt, \tag{3.31}$$

with initial concentrations $x_0 = (b_0, s_0)$. $b^{in} = (b_t^{in}, t \geq 0)$ and $s^{in} = (s_t^{in}, t \geq 0)$ are incoming concentrations, D is the dilution rate, and μ is the specific growth velocity, a function modelling population growth and ensuring a zero reaction rate in the absence of biomass. The derivation of ODE (3.31) is provided in the Supplementary Material.

This work assumes no incoming flow of biomass, thus $b_t^{in} = 0$, $t \geq 0$. Moreover, we proceed with a simplified version of ODE (3.31) called the minimal model [16], in which (i) μ is a function of the substrate, with $\mu(s) > 0$ for $s > 0$ and $\mu(0) = 0$, and (ii) the yield parameter κ is constant. These model assumptions are to be interpreted in the following way: (i) the velocity - or kinetics - of the reaction in (3.30) does not change as the biomass concentration changes and (ii) temperature and pressure are kept constant, thus κ is also constant.

Under these model assumptions, we now introduce two cases for which different types of organisms are being cultivated:

- Case 1 - Monotonic growth μ (Monod kinetics)
- Case 2 - Non-monotonic growth μ (Haldane kinetics)

Case 1: the Monod relation of growth [44] assumes that the growth rate is zero when there is no substrate, and tends to an upper limit when the substrate is in excess:

$$\mu(s) = \frac{\mu_{max} s}{K_s + s}. \tag{3.32}$$

Here, μ depends on the maximum growth μ_{max} and the half-saturation K_s.
Case 2: some experiments indicate that excessive substrate may inhibit the growth of a population [45]. This scenario is represented by the following growth function:

$$\mu(s) = \frac{\mu_0 s}{K_s + s + K_i s^2}. \tag{3.33}$$

Here, μ depends on the growth μ_0, half-saturation K_s and inhibitor parameter K_i.

Figure 2 illustrates these two cases with phase portraits and some highlighted phase paths for ODE (3.31) under the parameters from Table 1.

The coloured points in Figure 2 are equilibrium pairs (b_e, s_e) for which the right-hand side of ODE (3.31) is zero. Case 1 has the equilibrium point $E_1^{(1)} = (9.966, 0.345)$, known to be locally exponentially stable [16, Section 2.1.2]. Case 2 has a trivial equilibrium point $E_2^{(1)} = (0, s_0^{in})$. Additionally, it has interior equilibria $E_2^{(2)} = (0.859, 1.141)$ and $E_2^{(3)} = (1.912, 0.088)$. It is possible to verify that one of the two interior equilibria is a stable nodal point and the other is a saddle point [16, Section 2.1.3].

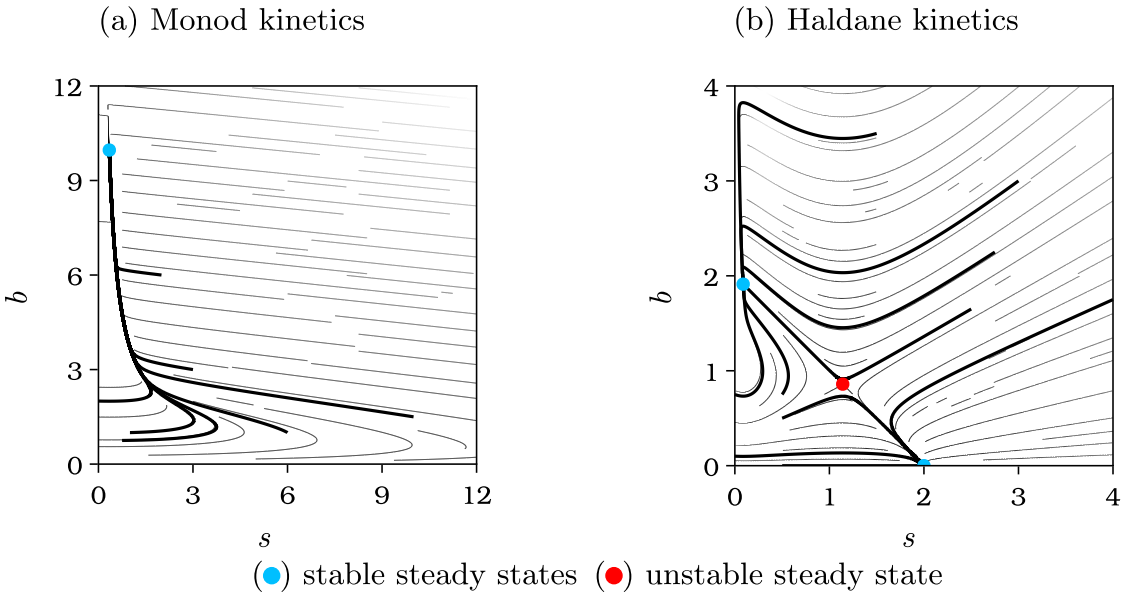

Figure 2: Phase portrait for ODE (3.31) with growth functions (a) (3.32) and (b) (3.33).

Table 1: Model parameters for ODE (3.31).

	Specific growth	κ	D (h^{-1})	$s_t^{in} = s_0^{in}$ (mgL^{-1})
Case 1	$\frac{0.3s}{10+s}$	10	0.01	100
Case 2	$\frac{5s}{0.5+s+5s^2}$	1	0.7	2

With the fundamentals of the chemostat introduced, we are now ready to apply the concepts outlined in Section 2 to this system.

Section 3.1 presents the equations governing the stochastic evolution of biomass and substrate, modeled as a process $X = (X_t, t \geq 0)$. In particular, we prove in Section 3.1.1 results on the existence, uniqueness and non-explosion of the solution X, as introduced in Sections 2.1.1 and 2.1.2, and revisit in Section 3.1.2 the Markovianity of the solution and the KFE from Sections 2.1.3 and 2.1.4.

Section 3.2 particularises the filtering equations for a case in which the pair (X, Y) is Markovian, where $Y = (Y_t, t \geq 0)$ corresponds to different observation processes possibly available for a chemostat. Additionally, we show the approximated solution to the KSE of Section 2.2.2 based on the computational methods previously outlined.

3.1. Stochastic dynamics

Let us assume that the concentration of an element changes within a time period by a random amount. As in Example 2, assume that this random amount is a function of the current concentration itself. This modelling assumption has been discussed in [18], for

which the equation governing the dynamics, in the form of SDE (2.21), is

$$d\underbrace{\begin{pmatrix} b_t \\ s_t \end{pmatrix}}_{X_t} = \underbrace{\begin{pmatrix} (\mu(s_t) - D)\, b_t \\ D(s_0^{in} - s_t) - \kappa\mu(s_t)\, b_t \end{pmatrix}}_{f(X_t)} dt + \underbrace{\begin{pmatrix} \varepsilon_b b_t & 0 \\ 0 & \varepsilon_s s_t \end{pmatrix}}_{g(X_t)} d\underbrace{\begin{pmatrix} B_t^b \\ B_t^s \end{pmatrix}}_{B_t^X}, \tag{3.34}$$

where $s_t^{in} = s_0^{in}$ was held constant, and $X_0 = (b_0,\, s_0)$.

Figure 3 presents numerical solutions of SDE (3.34) for Cases 1 and 2. The noise intensities $(\varepsilon_b, \varepsilon_s)$ are $(0.05, 0.05)$ for Case 1 and $(0.075, 0.075)$ for Case 2.

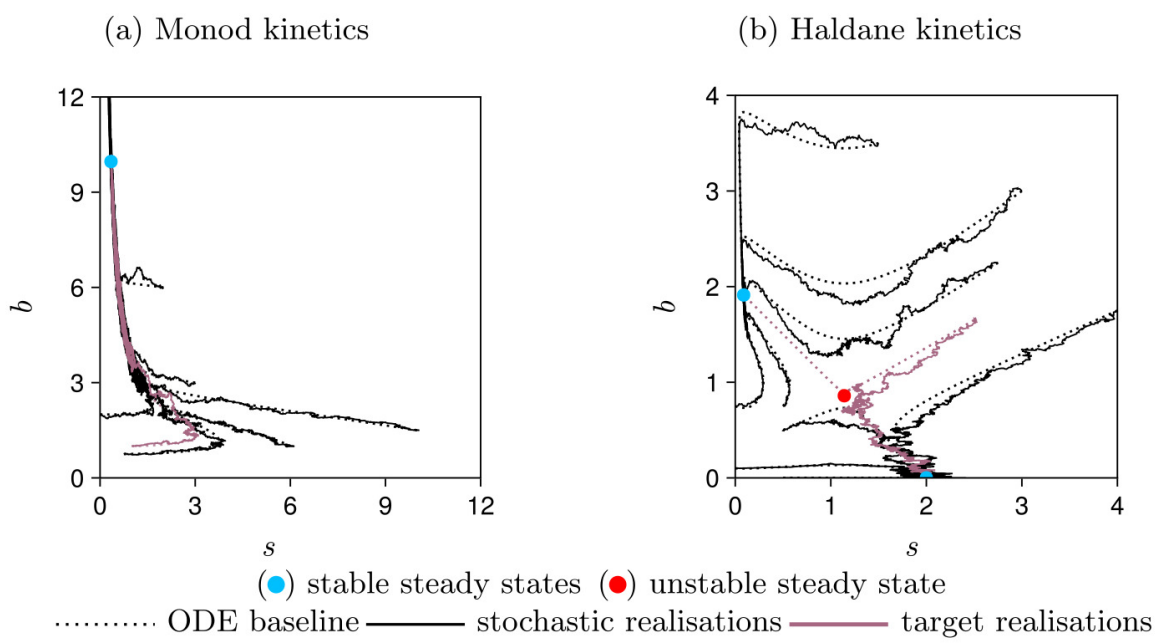

Figure 3: Solutions of SDE (3.34) with growth functions (a) (3.32) and (b) (3.33). The dashed lines in the background show the solutions of ODE (3.31), as in Figure 2. Here, a target realisation denotes a simulated trajectory which will be estimated from noisy observations later in this section.

Figure 3(a) suggests that processes X move around the trajectories of the deterministic chemostat model, with small fluctuations. Here, $D < \mu(s_0^{in})$. Under the deterministic model, this condition implies that the microbial population forming the biomass survives regardless of the initial concentrations of substrate and biomass [16]. In Figure 3(b), processes X also move around the trajectories of the deterministic chemostat model, but in order to avoid "near" extinction of the population around $E_2^{(1)} = (0, s_0^{in})$, we would have to appropriately choose X_0.

3.1.1. Existence, uniqueness and non-explosivity of the solution While numerical simulations are valuable for exploring the behaviour of the chemostat, proofs of existence and uniqueness of solutions of SDE (3.34) are required to ensure the validity of any conclusions drawn from Figure 3. Moreover, as described in Section 2.1.2, we question

whether a sample path of X almost surely exits an open set S in finite time. By letting $S := (\mathbb{R}^2)^*_+$, i.e. $S = \{(b,s) \in \mathbb{R}^2 | b > 0, s > 0\}$, we investigate whether X explodes or reaches zero/negative values in finite time. This checks for unrealistic scenarios in a chemostat, such as negative concentrations.

Since Condition A does not hold for SDE (3.34), Theorem 1 has to be modified to prove existence and uniqueness of solutions. We present the necessary modifications and show that a unique solution X exits S only at infinite time. We follow the same approach found in [21].

Theorem 2. *For any initial condition $X_0 \in S$, $S = (\mathbb{R}^2)^*_+$, there exists a unique strong solution $X = (X_t, t \geq 0)$ of SDE (3.34) with growth functions (3.32) or (3.33). Moreover, the solution only has positive sample paths.*

Proof. The coefficients f and g of SDE (3.34) satisfy the local Lipschitz continuous conditions; that is, for all $\alpha > 0$, there exist constants $\beta^f_\alpha > 0$ and $\beta^g_\alpha > 0$ such that $\|f(x) - f(x')\| \leq \beta^f_\alpha \|x - x'\|$ and $\|g(x) - g(x')\| \leq \beta^g_\alpha \|x - x'\|$, for all $\|x\|$, $\|x'\| \leq \alpha$. Therefore, there exists a unique solution $X = (X_t, t \geq 0)$ of SDE (3.34), on $t \in [0, t_e)$, where $X_0 \in S$ and t_e is the time when X_t exits S. See, for example, [27] or [46].

To show that this solution is global, we need to show that $t_e = \infty$ almost surely. Let r_0 be a sufficiently large integer such that $\frac{1}{r_0} < X_0^{n_X} < r_0$, $n_X \in \{1, 2\}$. We define an increasing sequence of open sets $\{S_r\}_{r \geq r_0}$, where $S_r := (1/r, r) \times (1/r, r)$, such that their union equals S. We also define the stopping time

$$t_r := \inf \{t \in [0, t_e) : X_t \notin S_r\},$$

where again $\inf \emptyset = \infty$. Let $t_\infty := \lim_{r \to \infty} t_r$, we then have $t_\infty \leq t_e$ almost surely. The idea is to prove that no exit occurs, i.e. $X_t \in S$ for all $t \geq 0$. We assume that the opposite result is true. Then we show that such assumption leads to a contradiction.

Assume that X will exit S almost surely. Then the random variable t_e and, consequently, the collection of random variables $\{t_r\}_{r \geq r_0}$ are bounded in probability; for any $\beta \geq 0$, there exists a constant T_β and an integer r' such that

$$\mathbb{P}(t_r \leq T_\beta) \geq 1 - \beta, \quad \text{for all } r \geq r'. \tag{3.35}$$

In accordance with the Khasminskii test for verifying non-explosion of solutions of SDEs, we construct a suitable function to prove the existence of a global solution. Define a Lyapunov function $l : \mathbb{R}^2_+ \to \mathbb{R}_+$ by

$$l(x) := \kappa (b - 1 - \ln b) + s - \iota \left(1 + \ln \frac{s}{\iota}\right), \; x = (b, s), \; \iota > 0.$$

The non-negativity of this function comes from the inequality $\alpha \geq 1 + \ln(\alpha)$ for all $\alpha > 0$. Without loss of generality, we let μ be of the Haldane type. Under the condition $X_0 = x_0$, and with $l(x_0) < \infty$, we apply Itô's formula to the process $l(X) = (l(X_t), t \geq 0)$:

$$dl(X_t) = \mathcal{A}l(X_t)\, dt + (\nabla l(X_t))^\top g(X_t)\, dB_t,$$

where $\mathcal{A}$ is the infinitesimal generator of X, whose coefficients are associated to those of SDE (3.34). Therefore, for the first term on the RHS,

$$\begin{aligned}
\mathcal{A}l\left(X_t\right) &= \begin{pmatrix} \frac{\kappa b_t-\kappa}{b_t} \\ \frac{s_t-\iota}{s_t} \end{pmatrix}^\top \begin{pmatrix} (\mu\left(s_t\right)-D)b_t \\ D(s_0^{in}-s_t)-\kappa\mu\left(s_t\right)b_t \end{pmatrix} + \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix} \varepsilon_b^2 b_t^2 & 0 \\ 0 & \varepsilon_s^2 s_t^2 \end{pmatrix}\begin{pmatrix} \frac{\kappa}{b_t^2} & 0 \\ 0 & \frac{\iota}{s_t^2} \end{pmatrix}\right) \\
&= (\kappa b_t-\kappa)\left(\frac{\mu_0 s_t}{K_s+s_t+I_s s_t^2}-D\right) \\
&\quad + \frac{s_t-\iota}{s_t}\left(Ds_0^{in}-Ds_t-\kappa\frac{\mu_0 s_t b_t}{K_s+s_t+I_s s_t^2}\right)+\frac{1}{2}(\kappa\varepsilon_b^2+\iota w_s^2) \\
&\leq -\kappa D b_t+\kappa D+Ds_0^{in}+\iota D+\iota\kappa\frac{\mu_0}{K_s}b_t+\frac{1}{2}(\kappa\varepsilon_b^2+\iota w_s^2).
\end{aligned}$$

If we let $\iota := \frac{K_s D}{\mu_0}$, the terms dependent on b_t cancel out. Consequently, $\mathcal{A}l\left(X_t\right) \leq \alpha$, where $\alpha > 0$ is a constant, and

$$dl\left(X_t\right) \leq \alpha\, dt + \varepsilon_b(b_t-\kappa)dB_t^b + \varepsilon_s(s_t-\iota)dB_t^s. \tag{3.36}$$

Let $t_r \wedge t := \min\left(t_r, t\right)$. For any $t \geq 0$ and integer $r \geq r_0$, $X_{t_r\wedge t} \in S$ by definition of t_r. Integrating both sides of (3.36) results in

$$\int_0^{t_r\wedge T_\beta} dl\left(X_t\right) \leq \int_0^{t_r\wedge T_\beta} \alpha\, dt + \int_0^{t_r\wedge T_\beta} \left(\varepsilon_b(b_t-\kappa)dB_t^b+\varepsilon_s(s_t-\iota)dB_t^s\right).$$

By taking expectations on both sides, the rightmost term disappears, resulting in

$$\begin{aligned}
\mathbb{E}_X[l\left(X_{t_r\wedge T_\beta}\right)]-l\left(x_0\right) &\leq \alpha\mathbb{E}_X[t_r\wedge T_\beta] \\
&= \alpha(t_r\wedge T_\beta).
\end{aligned} \tag{3.37}$$

Also, for every $t_r \leq T_\beta$, there is some $n_X \in \{1,2\}$ such that either $X_{t_r}^{n_X} = r$, with

$$l\left(X_{t_r}\right) \geq \left(\kappa(r-1-\ln r)\right) \wedge \left(r-\iota\left(1+\ln\frac{r}{\iota}\right)\right),$$

or $X_{t_r}^{n_X} = 1/r$, with

$$l\left(X_{t_r}\right) \geq \left(\kappa\left(\frac{1}{r}-1-\ln\frac{1}{r}\right)\right) \wedge \left(\frac{1}{r}-\iota\left(1+\ln\frac{1}{\iota r}\right)\right).$$

We combine both inequalities in the compact form $l\left(X_{t_r}\right) \geq Q\left(r\right)$, where

$$\begin{aligned}
Q\left(r\right) &:= \left[\left(\kappa(r-1-\ln r)\right) \wedge \left(r-\iota\left(1+\ln\frac{r}{\iota}\right)\right)\right] \\
&\quad \wedge \left[\left(\kappa\left(\frac{1}{r}-1+\ln r\right)\right) \wedge \left(\frac{1}{r}-\iota\left(1+\ln\frac{1}{\iota r}\right)\right)\right].
\end{aligned}$$

Under (3.35), inequality (3.37) can be further developed into

$$\begin{aligned}
l\left(x_0\right)+\alpha\left(t_r\wedge T_\beta\right) &\geq \mathbb{E}_X\left[l\left(X_{t_r\wedge T_\beta}\right)\right] \\
&\geq \mathbb{E}_X\left[\mathbb{1}\left(t_r \leq T_\beta\right) l\left(X_{t_r}\right)\right] \\
&\geq \mathbb{P}\left(t_r \leq T_\beta\right) Q\left(r\right) \\
&\geq (1-\beta)Q\left(r\right),
\end{aligned} \tag{3.38}$$

where the second and third inequalities result from the Markov's inequality. We are interested in the limiting behavior of t_r when r goes to infinity. Taking the limit on both sides of inequality (3.38) results in

$$\lim_{r\to\infty} \left(l\left(x_0\right) + \alpha(t_r \wedge T_\beta)\right) \geq (1-\beta) \lim_{r\to\infty} Q\left(r\right).$$

That is, $l\left(x_0\right) + \alpha T_\beta \geq \infty$, which is a contradiction, since $0 < l\left(x_0\right) < \infty$ and $0 \leq \alpha T_\beta < \infty$. Therefore, $t_r \to \infty$ almost surely.

We have shown that for locally Lipschitz coefficients the solution is unique up to an exit time t_e. Since $t_r \to \infty$, we have $t_e = \infty$, thus the solution X remains within S. □

Remark 1. *[47] proposes alternative driving processes to randomise the deterministic dynamics of the chemostat: for instance, a Brownian motion term $\varepsilon_b\sqrt{b_t}dB_t^b$ instead of $\varepsilon_b b_t dB_t^b$, and $\varepsilon_s\sqrt{s_t}dB_t^s$ instead of $\varepsilon_s s_t dB_t^s$. As detailed in their paper, this induces minimal change in the case of large population sizes. However, these driving processes differ drastically in the washout regime, i.e. near the boundary $b_t = 0$, as the behaviour of the two diffusion processes are different. Theorem 2 assured that X remains within S, thus the washout regime is unattainable in SDE (3.34). If X is driven by $\varepsilon_A\sqrt{a_t}dB_t^a, a \in \{b, s\}$, not only $b_t = 0$ is attainable, but also $s_t = 0$, for certain regions of the parameter space. Care needs to be taken to avoid the unrealistic scenario of a vanishing substrate concentration, despite a continuous input s^{in} process.*

3.1.2. Markovianity and the KFE Due to the Markovianity of the unique strong solution X of SDE (3.34), the evolution of X can be studied through its associated KFE, as in [14]. By solving the KFE given the initial distribution, we examine the likelihood of the state of the chemostat being within regions of interest, e.g. around $E_2^{(1)}$ or $E_2^{(3)}$.

Notably, the results from Theorem 2 alone are insufficient to ensure that the fundamental and classical solutions to the KFE discussed in Section 2.1.4 exist and are unique. For instance, the coefficients of SDE (3.34) do not satisfy Condition (2.6), and hence the sufficient condition for uniqueness of a fundamental solution according to [23, Section 5.7] is not satisfied. Moreover, the diffusion tensor associated to SDE (3.34)

$$\sum_{n_{B^X}=1}^{N_{B^X}} g^{n_X,n_{B^X}}\left(x\right) g^{n'_X,n_{B^X}}\left(x\right) = \begin{pmatrix} \varepsilon_b^2 b_t^2 & 0 \\ 0 & \varepsilon_s^2 s_t^2 \end{pmatrix},$$

herein denoted by G, does not satisfy (2.13) (uniform ellipticity). Consequently, we are unable to show the existence of a unique solution p satisfying the associated KFE with exponential (2.12) or polynomial (2.14) growths.

However, the existence of a unique solution p can be established under broader assumptions on f and G, along with certain properties of a Lyapunov function l. We present these assumptions and verify their validity for the stochastic chemostat. With the uniqueness of the solution to the KFE ensured, we present its numerical solution.

First, let us consider a local boundedness condition on coefficient f_t:

Condition B. *For all positive integer r, $f \in \mathcal{L}^{\infty}(S_r)$. That is, f is an essentially bounded measurable function on the open set S_r.*

Secondly, let us consider a relaxed version of (2.5) for coefficient G:

Condition C. *For all positive integer r, there exists a constant $\gamma_r > 0$ such that*

$$|G^{n_X,n'_X}(x) - G^{n_X,n'_X}(x')| \leq \gamma_r \|x - x'\|$$

holds for all $x, x' \in S_r$.

Lastly, let us consider a relaxed version of (2.13), again for G:

Condition D. *For all $S_r \subset S$, there exist constants $\alpha_r > 0$ and $\beta_r > 0$ such that*

$$\alpha_r \|x'\| \leq \sum_{n_X=1}^{N_X} \sum_{n'_X=1}^{N_X} G^{n_X,n'_X}(x)\, x'_{n_X} x'_{n'_X} \leq \beta_r \|x'\|$$

holds for all $x \in S$, $x' \in \mathbb{S}_X$.

These conditions yield the following result for existence and uniqueness:

Proposition 2. *There exists exactly one probability measure whose density with respect to the Lebesgue measure satisfies the KFE associated to SDE (3.34).*

Proof. Let us note that the coefficients of SDE (3.34) satisfy conditions B, C, and D (by letting $\alpha_r = r^{-2}(\varepsilon_b^2 + \varepsilon_s^2)$, $\beta_r = r^2(\varepsilon_b^2 + \varepsilon_s^2)$, and $\gamma_r = 2r \max\{\varepsilon_b^2, \varepsilon_s^2\}$).

According to [48, Theorem 3.6], such conditions are sufficient to ensure the existence of a unique solution, given that there exists a Lyapunov function l such that

$$\lim_{r\to\infty} \inf_{S_r \setminus S_{r-1}} l(x) = \infty, \tag{3.39}$$

and for all $t \in [0, \infty)$, $x \in S$, and some $\alpha > 0$, the following inequality holds:

$$\mathcal{A}l(x) \leq \alpha + \alpha l(x). \tag{3.40}$$

If we define $l(x) := (b + s + \frac{1}{b} + \frac{1}{s})$, then $\inf_{S_r \setminus S_{r-1}} l(x) = l((r-1, r-1))$ and $\mathcal{A}l(x) \leq \alpha + \alpha l(x)$, $\alpha = \max\{\mu_{\max}, D, \varepsilon_b^2, \varepsilon_s^2\}$, where $\mu_{\max}$ is a bound for the growth function μ. That is, both (3.39) and (3.40) hold, and the proof is complete. □

We present the numerical solutions of the KFE for Cases 1 and 2 in Figures 4 and 5, respectively. They are obtained by similar tools as in [14]: an upwind differencing scheme for the space discretisation, an implicit scheme for the time discretisation, reflecting boundaries along the truncated discretisation of the positive quadrant, and natural boundaries along the coordinate axes at which one of the state variables vanishes. The upwind differencing scheme is chosen to ensure the nonnegativity of probabilistic terms, as discussed in [49, Chapter 7]. Our work differs from [14] by using the Crank-Nicolson scheme instead of the implicit Euler scheme as it led to improved stability.

Each marginal curve has been normalised by its maximum value for the sake of illustration. In Figure 4, the density remains unimodal, with a relatively high spread over time due to diffusion. Notably, the distribution of the state deviates from a Gaussian, as

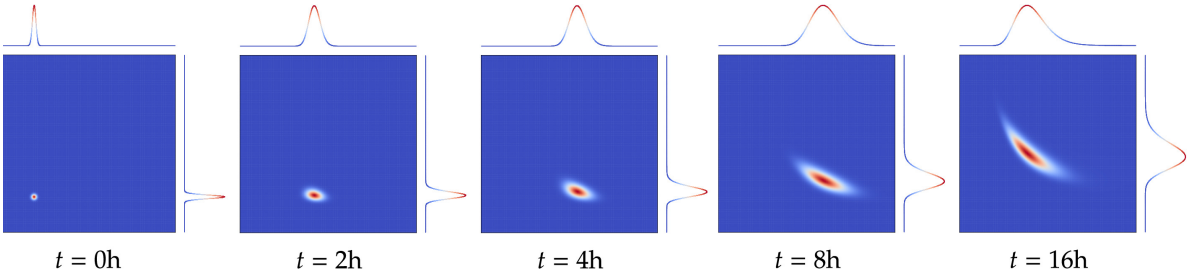

Figure 4: Case 1 - solution of the KFE for the growth function of Monod type. Initial condition is $X_0 \sim \mathcal{N}([1,\ 1]^\top, 0.05^2 I_{2\times 2})$. The discretisation in space is in the domain $(0,5] \times (0,5]$, with a count of 256^2 finite volumes. The marginal density functions are represented by marginal curves.

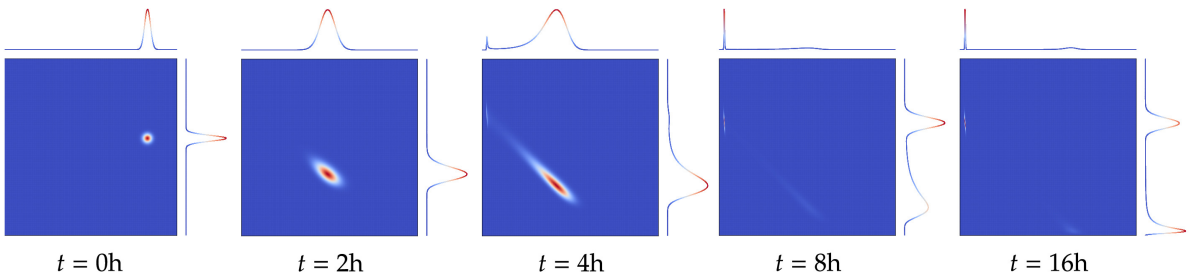

Figure 5: Case 2 - solution of the KFE for the growth function of Haldane type. Initial condition is $X_0 \sim \mathcal{N}([1.65,\ 2.5]^\top, 0.05^2 I_{2\times 2})$. The discretisation in space is in the domain $(0,3] \times (0,3]$, with a count of 512^2 finite volumes.

verifiable by computing its higher central moments. In Figure 5, the probability mass gradually splits toward the stable nodes from $t = 4$h onwards. In both cases, chemostat control may require reducing uncertainty about the probable state region.

3.2. Filtering problem for the stochastic chemostat

As discussed in Section 2.2, the state process cannot be directly observed. We introduce an observation model in which a sensor transforms the state variables in a one-dimensional signal prone to error. Specifically, we assume noisy observations of the biogas flow rate produced by the biomass after substrate consumption. This flow, proportional to the microbial activity [50, 51], is modelled by the mapping $(s, b) \mapsto \gamma\mu\,(s)\,b$, where γ is the yield parameter for biogas in (3.30). This setting is denoted by Scenario A.

Written in the form of SDE (2.22), the observation process evolves according to

$$\underbrace{d\,(\,y_t^A\,)}_{Y_t} \;=\; \underbrace{(\,\gamma\mu\,(s_t)\,b_t\,)}_{h^A(X_t)}\,dt + \underbrace{(\,\varepsilon_A\,)}_{K^A}\,d\,(\,B_t^{Y^A}\,)\,, \tag{3.41}$$

where the superscript A refers to Scenario A, and $Y_0 = 0$.

We combine Cases 1 and 2 with SDE (3.41) to design the experiments outlined in Table 2. Over a time interval of interest $[0, T]$, we simulate trajectories of the processes

X and Y using time discretisations Δt^x and Δt^y, respectively. The trajectories for Y are illustrated in Figure 6. The observations $\{y_t^A\}$ are assumed to be collected either at a high frequency ($\Delta t^y = \Delta t^x$) or at a low frequency ($\Delta t^y \gg \Delta t^x$). The initial concentrations (b_0, s_0) are $(1\,\text{mgL}^{-1}, 1\,\text{mgL}^{-1})$ for Case 1 and $(1.65\,\text{mgL}^{-1}, 2.5\,\text{mgL}^{-1})$ for Case 2.

Table 2: Additional parameters for differential equations of X and Y.

	Model parameters		Simulation parameters	
	Observation function	ε_A	Δt^x (hours)	Δt^y (hours)
1A - high-frequency	$27.5\mu(s)b$	0.01	0.01	0.01
1A - low-frequency	$27.5\mu(s)b$	$0.01\left(\frac{\Delta t^y}{\Delta t^x}\right)^{0.5}$	0.01	2
2A - high-frequency	$2.75\mu(s)b$	0.01	0.005	0.005
2A - low-frequency	$2.75\mu(s)b$	$0.01\left(\frac{\Delta t^y}{\Delta t^x}\right)^{0.5}$	0.005	2

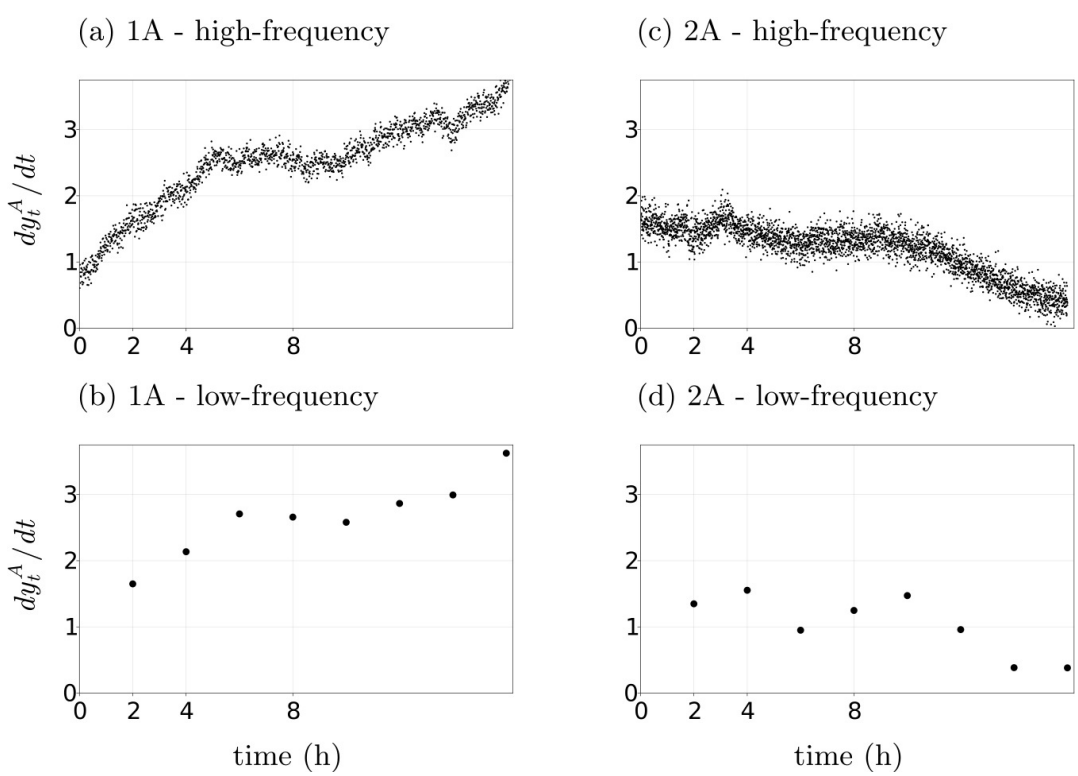

Figure 6: Observations simulated from the trajectories highlighted in Figure 3.

Additionally to Scenario A, the Supplementary Material presents the filtering problem for Cases 1 and 2 under three additional scenarios: Scenario B, with noisy observations of the substrate concentration at low frequency; Scenario C, with noisy observations of the biomass concentration at low frequency; and Scenario D, with noisy observations of both biomass and substrate concentrations at low frequency.

Without the observations, the probability density of the hidden state evolved according to KFE (2.11) as illustrated in Figures 4 and 5. In light of the observations (Figure 6), we can now approximate the solution of the pathwise version of KSE (2.25) for all the experiments in Table 2. Figures 7 and 8 show the pathwise solution $p^{\Delta t} = \left(p^{\Delta t}_{n_t}, n_t = 0, 1, \ldots\right)$ for low-frequency observations ($\Delta t = \Delta t^y \gg \Delta t^x$) via the splitting-up approximation described in Section 2.2.2.

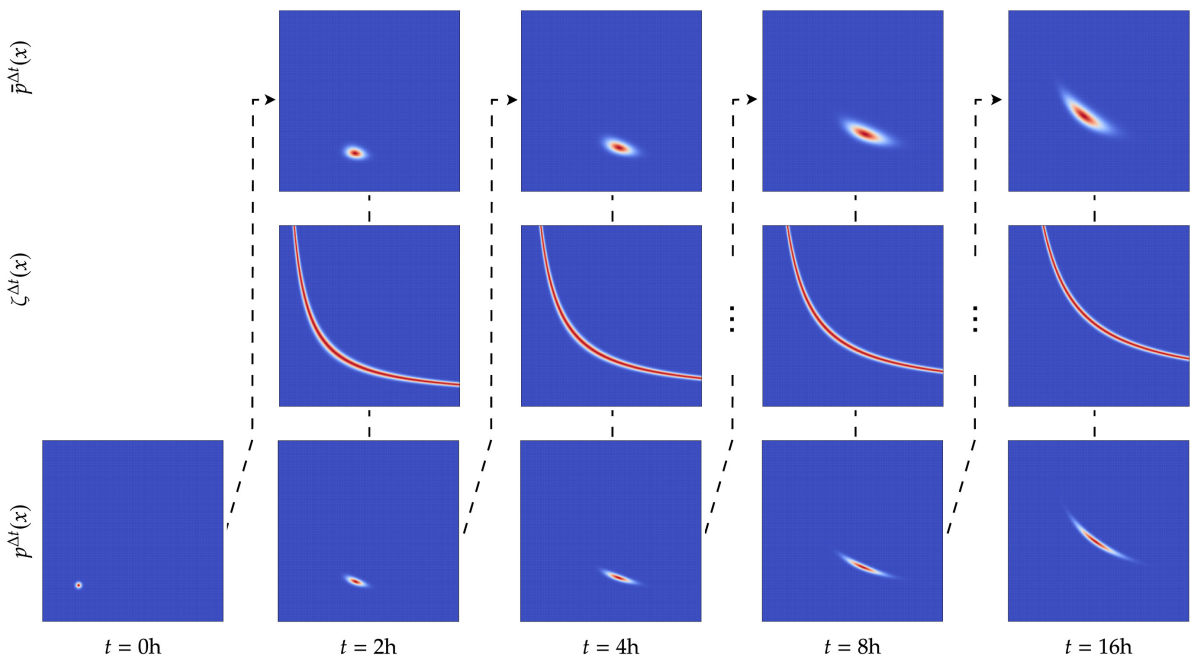

Figure 7: Splitting-up approximation for Case 1A - low-frequency observations.

Each column in Figures 7 and 8 illustrates the iterative steps of the splitting-up approximation. The top row shows the result from the prediction step based on the KFE (2.28), the middle row depicts the likelihood ratio from (2.29), and the bottom row presents the normalised product of these two, i.e. the corrected density.

We approximate these solutions via two other methods: particle filtering, discussed in Section 2.2.2, and hereafter denoted by bootstrap particle filter (BPF), and the linearisation method known as extended Kalman filter (EKF).

We analyse the similarity among the approximated solutions by computing the Hellinger distance $\mathcal{H}$ between two probability distributions, say F_1 and F_2. With respect to the Lebesgue measure, let f_1 and f_2 denote the density functions of the distributions F_1 and F_2, respectively. The square of the Hellinger distance is then defined as $\mathcal{H}^2\left(f_1, f_2\right) = \frac{1}{2}\int\left(\sqrt{f_1(x)} - \sqrt{f_2(x)}\right)^2 dx$. Note how this distance function is symmetric and is zero-valued when the compared densities are equal.

Figures 9 and 10 show the comparative results for a fixed time t. Algorithms are available in the Supplementary Material.

Figure 11 shows the square of the Hellinger distance for Case 1 between π approximated by the splitting-up method with different grid refinements and by the

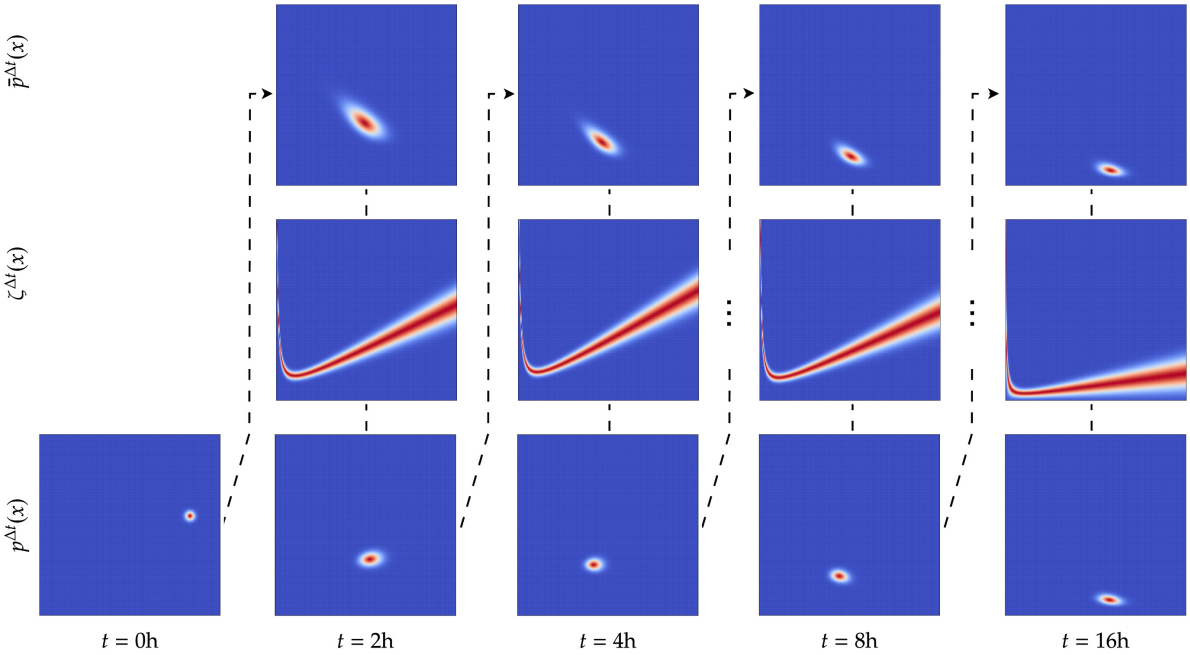

Figure 8: Splitting-up approximation for Case 2A - low-frequency observations.

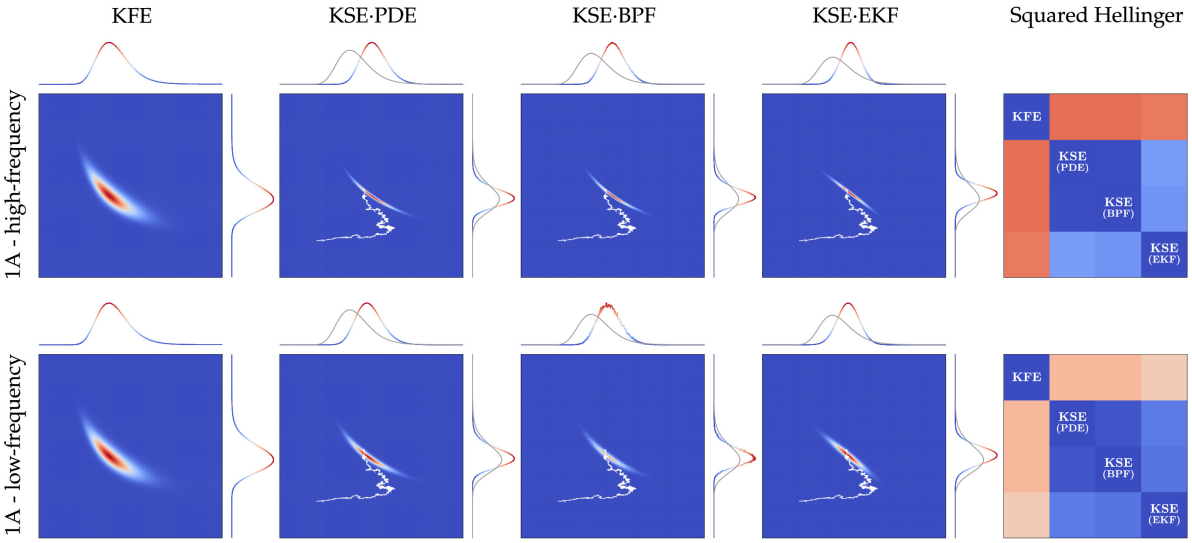

Figure 9: Case 1A - approximated solution of the KSE at $t = 16$h with the finite differences and splitting-up approximation for PDEs (KSE·PDE), the BPF (KSE·BPF) and the EKF (KSE·EKF). For the PDE method, the discretisation in space is in the domain $(0, 5.0] \times (0, 5.0]$. We consider a refinement of the grid of a total count of 256^2 finite volumes. As for the BPF, we consider $N_P = 2.5 \times 10^6$. The trajectory of X from which observations were simulated is shown in white and can also be found in Figure 3.

BPF. It is clear that the discretisation error decreases with a finer grid refinement. The distance between π approximated by the EKF and by the BPF is also shown. The EKF fails to capture higher moments of the distribution, hence its divergence from the BPF result. This has already been suggested by Figure 9, for which we can observe non-zero

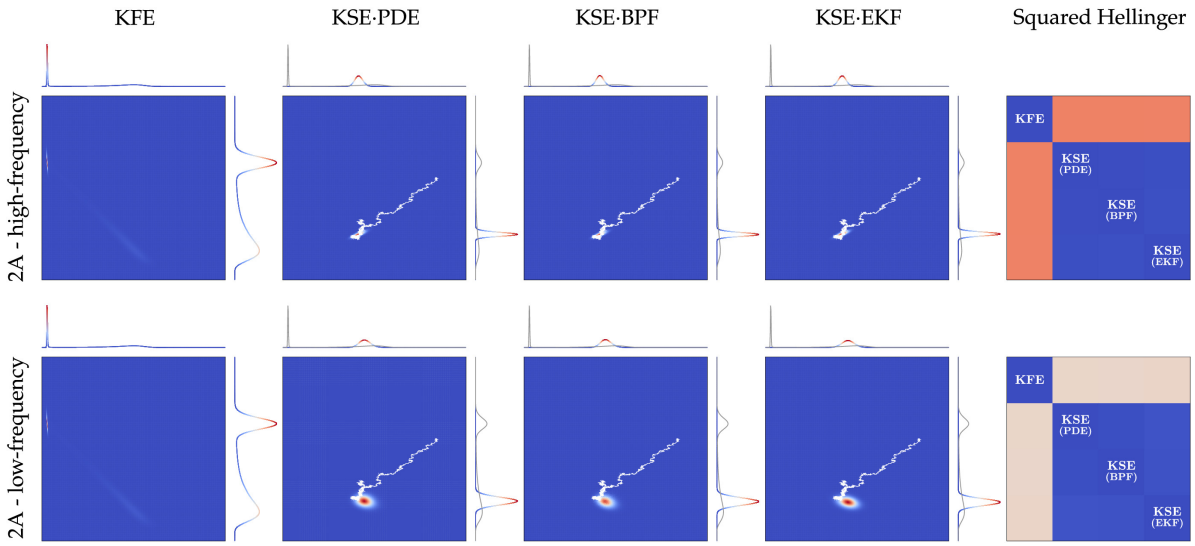

Figure 10: Case 2A - KSE·PDE, KSE·BPF, and KSE·EKF approximations at $t = 8$h. For the PDE method, the discretisation in space is in the domain $(0, 3.0] \times (0, 3.0]$. We consider a refinement of the grid of a total count of 256^2 finite volumes. As for the BPF, $N_P = 2.5 \times 10^6$. Differently from the solution of the KFE, more probability mass moves towards the washout on account of the observations.

skewness and excess kurtosis resulting from having both a non-linear dynamic process and a non-linear observation process. The shape of the conditional density for different times can be found in the Supplementary Material.

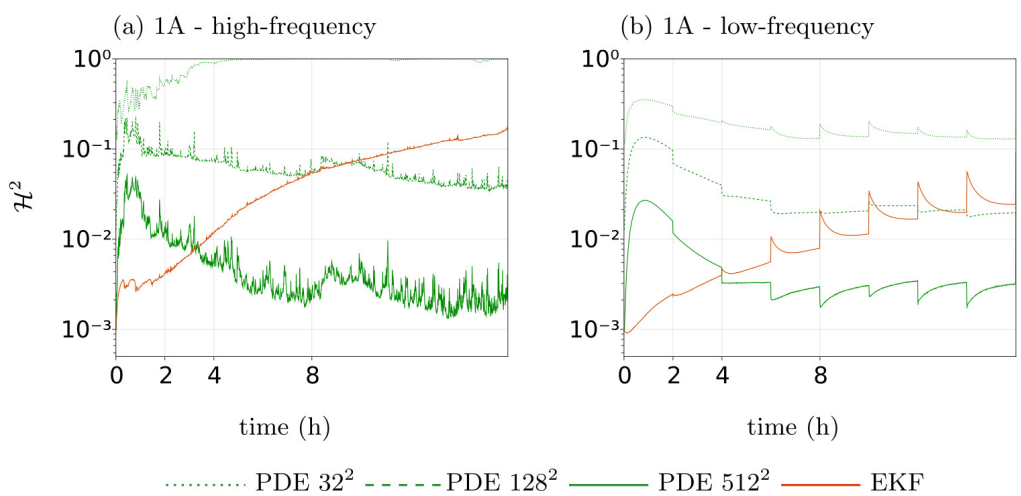

Figure 11: Case 1A - Squared Hellinger distances between π approximated by different grid refinements and the BPF (green), and between π approximated by the EKF and the BPF (orange). We consider three scenarios for the grid refinement, each with a total count of $32^2, 128^2, 512^2$ finite volumes, respectively. As for the BPF, $N_P = 2.5 \times 10^6$.

Figure 12 shows the square of the Hellinger distance between π approximated by the splitting-up method with different grid refinements and by the BPF, now for Case

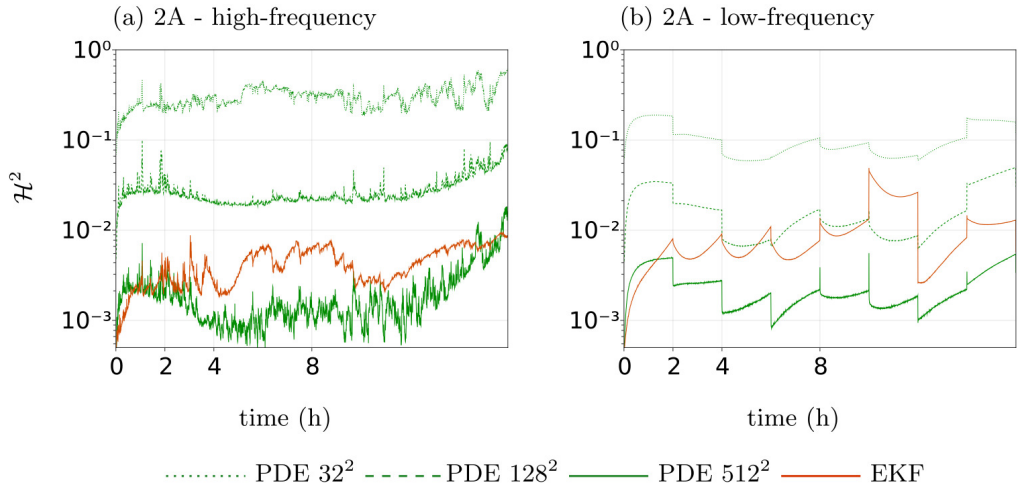

Figure 12: Case 2A - Squared Hellinger distances between π approximated by different grid refinements and the BPF (green), and between π approximated by the EKF and the BPF (orange). Same grid refinement and number of particles as in Figure 11.

2A. We again include the square of the Hellinger distance between π approximated by the EKF and by the BPF. Here, the EKF suffers in the highly non-linear region of the state-space, namely when the state is near the saddle point, $t \geq 4$h.

Examining Figure 12, we notice that the approximation with the splitting-up method combined with finite-differences on a 512^2 grid mismatches the approximation with BPF at the observation times. At $t = 8h$, for instance, the Hellinger distance between both distributions increases. The densities approximated by grid-based schemes or by the EKF assume that the solution is uniformly distributed within each finite volume. The density approximated by the BPF results from counting the number of particles falling within each finite volume. Here, we encounter the problem of degenerate finite volumes, in which no particles are present. This issue is similar to the so-called sample impoverishment, which is a reduction in the number of distinct sample values when the model noise variance for process X is very low. As a result, the particles will hardly diversify and after a few iterations they may collapse to a single point in the state space.

There are different solutions for this problem, the simplest being an increase in the number of particles. This, however, leads to an unreasonable computational load. Alternatively, a density kernel estimate can smooth the particle-based approximation and extrapolate the density values for finite volumes without any particles. Other solutions are more complicated to implement and, unlike the BPF, they might need to be tailored to the problem in hand. See [52] and [53] for more methods to prevent sample impoverishment.

4. Concluding remarks

In the first part of this work, we established mathematical guarantees for the inverse problem of state estimation in a stochastic chemostat. Specifically, we proved the

existence and uniqueness of the solution to the SDE governing the evolution of biomass and substrate concentrations, ensured trajectory positivity, and discussed the well-posedness of the KFE. Coupled with an observation model, we presented the filtering problem and established conditions for the well-posedness of the KSE. Consequently, numerical pathwise approximations inherited a theoretical guarantee of reliability.

As a complement to the first part, we presented numerical solutions to the governing SDE, the corresponding KFE, and the pathwise KSE. Approximations through Monte Carlo particles or traditional PDE-based methods imposed no restrictions on the shape of the distribution, allowing for greater accuracy compared to the linearisation method.

Data availability statement

The code used is available upon reasonable request from the corresponding author.

References

[1] Bain A and Crisan D 2008 *Fundamentals of Stochastic Filtering* (Springer New York)
[2] Crisan D and Rozovskii B 2011 *The Oxford Handbook of Nonlinear Filtering* (Oxford University Press)
[3] Lanthaler S, Mishra S and Weber F 2022 *Inverse Problems* **38** 085012
[4] Clark J 1978 The design of robust approximations to the stochastic differential equations of nonlinear filtering
[5] Pardoux E 1982 *Stochastics* **6** 193–231
[6] Beneš V E 1981 *Stochastics* **5** 65–92
[7] Kálmán R and Bucy R 1961 *Journal of Basic Engineering* **83** 95–108
[8] Gordon N J, Salmond D J and Smith A F M 1993 Novel approach to nonlinear/non-gaussian bayesian state estimation
[9] Kitagawa G 1996 *Journal of Computational and Graphical Statistics* **5** 1–25
[10] Vermaak J, Godsill S and Perez P 2005 *IEEE Transactions on Aerospace and Electronic Systems* **41** 309–332
[11] van Leeuwen P J 2009 *Monthly Weather Review* **137** 4089–4114
[12] Shaman J and Karspeck A 2012 *Proceedings of the National Academy of Sciences* **109** 20425–20430
[13] Pichler L, Masud A and Bergman L 2013 Numerical solution of the fokker–planck equation by finite difference and finite element methods—a comparative study
[14] Campillo F, Joannides M and Larramendy I 2014 *Mathematics and Computers in Simulation* **99** 37–53
[15] Bull A 2010 *Journal of Industrial Microbiology and Biotechnology* **37** 993–1021
[16] Harmand J, Lobry C, Rapaport A and Sari T 2017 *The Chemostat: Mathematical Theory of Microorganism Cultures* (John Wiley and Sons, ISTE)
[17] Hoskisson P and Hobbs G 2005 *Microbiology* **151** 3153–3159
[18] Imhof L and Walcher S 2005 *Journal of Differential Equations* **217** 26–53
[19] Hadj-Abdelkader O and Hadj-Abdelkader A 2019 *International Journal of Bioautomation* **23** 215–232
[20] Sinner P, Stiegler M, Goldbeck O, Seibold G, Herwig C and Kager J 2021 *Biotechnology and Bioengineering* **119** 575–590
[21] Mao X, Marion G and Renshaw E 2002 *Stochastic Processes and their Applications* **97** 95–110
[22] Kolmogorov A and Fomin S 1975 *Introductory Real Analysis* (Dover Publications)
[23] Karatzas I and Shreve S E 1991 *Brownian Motion and Stochastic Calculus* (Springer New York)

[24] Liptser R and Shiryaev A 2001 *Statistics of Random Processes: I. General Theory* Applications of Mathematics: Stochastic Modelling and Applied Probability Series (Springer)

[25] Kallianpur G 1980 *Stochastic Filtering Theory* Applications of mathematics (Springer)

[26] Itô K 1951 *Nagoya Mathematical Journal* **3** 55–65

[27] Khasminksii R 2012 *Stochastic Stability of Differential Equations* Stochastic Modelling and Applied Probability (Springer Berlin Heidelberg)

[28] Cohen S and Elliott R 2015 *Stochastic Calculus and Applications* Probability and Its Applications (Springer New York)

[29] Doob J 1953 *Stochastic Processes* Probability and Statistics Series (Wiley)

[30] Jazwinski A 1970 *Stochastic processes and filtering theory* 64 (Academic Press)

[31] Friedman A 1975 *Stochastic differential equations and applications. Vol. 1* (Academic Press)

[32] Pavliotis G 2014 *Stochastic Processes and Applications* (Springer New York)

[33] van Handel R 2007 *Filtering, stability, and robustness* Ph.D. thesis Princeton University

[34] Kutschireiter A, Surace S C and Pfister J 2020 *Journal of Mathematical Psychology* **94** 102307

[35] Taylor S 1973 *Introduction to Measure and Integration* (Cambridge University Press)

[36] Crisan D, Kouritzin M and Xiong J 2009 *Electronic Journal of Probability* **14** 1863–1883

[37] Davis M 1982 *Theory of Probability & Its Applications* **27** 167–175

[38] Scott D 2015 *Multivariate Density Estimation: Theory, Practice, and Visualization* Wiley Series in Probability and Statistics (Wiley)

[39] Cai Z, Le Gland F and Zhang H 1995 An adaptive local grid refinement method for nonlinear filtering

[40] Le Gland F 1989 Time discretization of nonlinear filtering equations *Proceedings of the 28th IEEE Conference on Decision and Control* vol 3 pp 2601–2606

[41] Le Gland F 1992 Splitting-up approximation for SPDE's and SDE's with application to nonlinear filtering *Stochastic Partial Differential Equations and Their Applications* (Springer Berlin Heidelberg) pp 177–187

[42] Gyöngy I and Krylov N 2003 *The Annals of Probability* **31** 564–591

[43] Smith H L and Waltman P 1995 *The Theory of the Chemostat: Dynamics of Microbial Competition* Cambridge Studies in Mathematical Biology (Cambridge University Press)

[44] Monod J 1950 La technique de culture continue: theorie et applications

[45] Andrews J 1968 *Biotechnology and bioengineering.* **10** 707–723

[46] Durrett R 1996 *Stochastic Calculus: A Practical Introduction* (Taylor and Francis)

[47] Campillo F, Joannides M and Larramendy I 2010 Stochastic models of the chemostat

[48] Manita O and Shaposhnikov S 2016 *J Dyn Diff Equat* **28** 493–518

[49] Kushner H J and Dupuis P G 1990 *Numerical Methods for Stochastic Control Problems in Continuous Time* vol 24 (Springer)

[50] Bernard O, Hadj-Sadok Z, Dochain D, Genovesi A and Steyer J P 2001 *Biotechnology and bioengineering* **75** 424–438

[51] Renard P, Dochain D, Bastin G, Naveau H and Nyns J 1988 *Biotechnology and Bioengineering* **31** 287–294

[52] Carpenter J, Cliffordy P and Fearnhead P 1999 *IEE Proc. Radar, Sonar Navig.* **146** 2–7

[53] Li T, Sun S, Sattar T and Corchado J 2014 *Expert Systems with Applications* **41** 3944–3954