F. Magalhães, José Augusto; Neto, Otacilio B. L.; Corona, Francesco

Block particle filters for state estimation of stochastic reaction-diffusion systems

Published in:
IFAC-PapersOnLine

DOI:
10.1016/j.ifacol.2023.10.910

Published: 22/11/2023

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 56-2 (2023) 10270–10275

Block particle filters for state estimation of stochastic reaction-diffusion systems

José Augusto F. Magalhães * **Otacílio B. L. Neto** *
Francesco Corona *
* *School of Chemical Engineering, Aalto University, Finland.*
(e-mail: {augusto.magalhaes, otacilio.neto, francesco.corona}@aalto.fi)

Abstract: In this work, we consider a differential description of the evolution of the state of a reaction-diffusion system under environmental fluctuations. We are interested in estimating the state of the system when only partial observations are available. To describe how observations and states are related, we combine multiplicative noise-driven dynamics with an observation model. More specifically, we ensure that the observations are subjected to error in the form of additive noise. We focus on the state estimation of a Belousov-Zhabotinskii chemical reaction. We simulate a reaction conducted in a quasi-two-dimensional physical domain, such as on the surface of a Petri dish. We aim at reconstructing the emerging chemical patterns based on noisy spectral observations. For this task, we consider a finite difference representation on the spatial domain, where nodes are chosen according to observation sites. We approximate the solution to this state estimation problem with the Block particle filter, a sequential Monte Carlo method capable of addressing the associated high-dimensionality of this state-space representation.

Keywords: Estimation and filtering, Sequential Monte Carlo, Belousov-Zhabotinskii reaction

1. INTRODUCTION

Since the introduction of the "Turing pattern" concept, reaction-diffusion models have played an important role in theoretical biology (Turing, 1952). Such representations have been used to explain several naturally-occurring pattern formations: developing structures in embryos (Kondo and Miura, 2010; Green and Sharpe, 2015), skin and pigmentation patterns (Kondo, 2009; Nakamasu et al., 2009), and tumour cell growth (Ferreira et al., 2002; Hogea et al., 2008). Additionally, reaction-diffusion models have also been employed to describe wildfire spread dynamics (Asensio and Ferragut, 2002) and to synthesize textures in computer graphics (Witkin and Kass, 1991). Moreover, if the state of a reaction-diffusion system can be determined in real-time, one could use feedback theory to control the pattern-formation process for several purposes (Ardizzone et al., 2013). Adopting these models in real biochemical applications has been challenging: such models must rely on partial observations of the true system through indirect methods (e.g. light spectrography), and their typical realisations are high-dimensional stochastic systems. Consequently, the task of estimating the state is bound to incur a heavy computational burden, if at all computable.

Particle filters (PFs) are popularly used for state estimation of general non-linear stochastic dynamical systems (Chopin and Papaspiliopoulos, 2020). An ensemble of particles equipped with importance weights undergoes a sampling procedure every time a new observation is available (Kitagawa and Gersch, 1996). Due to the recursive nature of this particle representation, the ensemble is bound to present a degeneracy problem (Musso et al., 2001). Some techniques to mitigate this phenomenon include introducing diversity among particles (Gilks and Berzuini, 2001) or resampling a section of their paths over a fixed lag (Doucet et al., 2006). However, even when equipped with these techniques, classical particle filters applied to models of high-dimensional state space will suffer from degeneracy even after a few iterations (Van Leeuwen, 2003). An alternative approach consists on leveraging the decay-of-correlations property of many high-dimensional systems by performing full-state estimation based on lower-dimensional partitions of the state-space. This class of methods, denoted local particle filters, include the Multiple particle filter (MPF, Djuric et al. (2007)) and, more recently, the Block particle filter (BPF, Rebeschini and van Handel (2015)).

Herein, we aim at estimating the concentrations of chemical substances in general reactive-diffusion systems using noisy spectral observations. We focus on state-space representations arising from finite-difference approximations of the spatio-temporal dynamics. We address this high-dimensional problem by using BPFs based on disjoint partitions of the spatial domain. We illustrate the performance of the estimator for the task of reconstructing the emerging chemical patterns of a Belousov-Zhabotinskii-type reaction-diffusion system, the Oregonator (Field and Noyes, 1974). From our results, the *vanilla* BPF is unable to reconstruct the state for relatively small ensemble sizes. However, we show that the BPF is able to provide accurate estimates of the state under a specific choice of sampling procedure.

The paper is organised as follows: Section 2 overviews reaction-diffusion systems and their representation using finite difference methods. Section 3 discusses PFs for general and high-dimensional state spaces. Section 4 presents the simulation results of applying the state estimator on a benchmark reaction-diffusion model. Finally, Section 5 lists our final remarks and some future research directions.

10.1016/j.ifacol.2023.10.910

2. REACTION-DIFFUSION SYSTEMS

We consider sets of chemical species $\mathcal{S} = \{\mathcal{S}_1, \dots, \mathcal{S}_{N_\mathcal{S}}\}$ distributed over some bounded space $U \subseteq \mathbb{R}^3$ which interact according to chemical reaction networks of the form

$$\sum_{n_s=1}^{N_\mathcal{S}} \underline{S}_{n_r,n_s}\mathcal{S}_{n_s} \xrightarrow{\kappa_{n_r}} \sum_{n_s=1}^{N_\mathcal{S}} \overline{S}_{n_r,n_s}\mathcal{S}_{n_s}, \quad n_r = 1, \dots, N_R, \tag{1}$$

with rate constants $\kappa \in \mathbb{R}_{\geq 0}^{N_R}$, and matrices $\underline{S} \in \mathbb{N}^{N_R \times N_\mathcal{S}}$ and $\overline{S} \in \mathbb{N}^{N_R \times N_\mathcal{S}}$ of stoichiometric coefficients for the reactants and products, respectively. From a system-analytical perspective, this process has dynamics described in terms of concentrations $z(u,t) = ([\mathcal{S}_1](u,t), \dots, [\mathcal{S}_{N_\mathcal{S}}](u,t)) \in \mathbb{R}_{\geq 0}^{N_\mathcal{S}}$ at every space coordinates $u \in U$ and time-instant $t \in \mathbb{R}_{\geq 0}$. The evolution of concentrations can then be described as

$$\partial^t z(u,t) = \left[(\overline{S} - \underline{S})^\mathsf{T} \nu\big(z(u,t)\big) + D_z \nabla^2 z(u,t)\right] \partial t, \tag{2}$$

given reaction rates $\nu(\cdot) = (\nu_1(\cdot), \dots, \nu_{N_R}(\cdot)) \in \mathbb{R}^{N_R}$, matrix of diffusion coefficients $D_z \in \mathbb{R}^{N_\mathcal{S} \times N_\mathcal{S}}$, and Laplacian $\nabla^2 z(u,t) = \sum_{i=1}^3 \partial z^2 / \partial u_i^2$. For each n_r-th reaction in network Eq. (1), the function ν_{n_r} $(n_r = 1, \dots, N_R)$ follows the Law of Mass Action (Murray, 2002),

$$\nu_{n_r}\big(z(u,t)\big) = \kappa_{n_r} \prod_{n_s=1}^{N_\mathcal{S}} z_{n_s}(u,t)^{\underline{S}_{n_r,n_s}}. \tag{3}$$

Finally, we consider homogeneous Neumann conditions $(\partial z / \partial u)|_{u=\overline{u}} = 0$ for all boundary points $\overline{u} \in \partial U$.

The concentrations of each substance are assumed to be only observed indirectly through their mixture. Specifically, we assume a spectroscopy process emitting

$$y(u,t|\lambda) = \sum_{n_s=1}^{N_\mathcal{S}} \phi_{\mathcal{S}_{n_s}}(\lambda) z_{n_s}(u,t), \tag{4}$$

given wavelength $\lambda \in \mathbb{R}_{\geq 0}$ and functions $\phi_{\mathcal{S}_{n_s}} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ describing the characteristic response spectrum associated with each n_s-th species $(n_s = 1, \dots, N_\mathcal{S})$. This corresponds to light absorbance being linearly related to the concentration of substances as in Beer's Law (Nadler and Coifman, 2005). Hereafter, we consider the response $\Phi_{\mathcal{S}_{n_s}}(\Lambda) = (\phi_{\mathcal{S}_{n_s}}(\lambda_1), \dots, \phi_{\mathcal{S}_{n_s}}(\lambda_{N_\Lambda}))$ for specific wavelengths $\Lambda = (\lambda_1, \dots, \lambda_{N_\Lambda})$, so that the measurement process is expressed by $y(u,t) = H_z z(u,t)$ with $H_z = [\Phi_{\mathcal{S}_1} \cdots \Phi_{\mathcal{S}_{N_\mathcal{S}}}]$.

The reaction-diffusion system Eq. (1) with measurement process Eq. (4) can be represented by the state-space model

$$\partial^t z(u,t) = f_{\theta_x}^z(z(u,t), \nabla^2 z(u,t)) \partial t; \tag{5a}$$
$$y(u,t) = H_{\theta_y}^z z(u,t), \tag{5b}$$

with state $z : U \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{N_\mathcal{S}}$ and measurements $y : U \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{N_\Lambda}$. Function $f^z(\cdot)$ and matrix $H^z \in \mathbb{R}^{N_\Lambda \times N_\mathcal{S}}$ are respectively determined by fixed parameters θ_x and θ_y collected from Eqs. (1)–(4).

In general, obtaining an exact solution for state-equation Eq. (5a) is unpractical. Towards a numerical approach for its integration, we discretise the space U into a lattice graph representing evenly spaced coordinates. State-dynamics can then be approximated by a collection of ordinary differential equations by the finite-difference method (Strauss, 2007). In the following, we overview this approach.

2.1 Finite-difference approximation

We will restrict ourselves to quasi-two-dimensional spaces $U = \{u \in \mathbb{R}^2 : 0 \leq u_{i=1,2} < \overline{U}\}$ with finite sizes $\overline{U} > 0$: these correspond to chemical mixtures in squared surfaces (e.g. a petri dish). Given some length $\Delta u > 0$, we discretise U onto the square lattice $V = \{v \in \mathbb{N}^2 : 1 \leq v_{i=1,2} \leq \overline{V}\}$ of evenly-spaced grid points with size $\overline{V} = (\overline{U}/\Delta u)$. The spatially-discretised dynamics are then approximated as

$$dz^{(v)} = \left[(\overline{S} - \underline{S})^\mathsf{T} \nu\big(z^{(v)}\big) + D_z \widetilde{\nabla}^2 z^{(v)}\right] dt, \quad \forall v \in V, \tag{6}$$

with $z^{(v)}(t) = z\big((v_1, v_2)\Delta u, t\big)$, and the first-order approximation of the Laplacian $\nabla^2 z$ at $v = (v_1, v_2)$,

$$\widetilde{\nabla}^2 z^{(v)} = \sum_{v' \in \mathcal{N}(v)} \frac{1}{\Delta u^2} \left(z^{(v')} - z^{(v)}\right), \tag{7}$$

where $\mathcal{N}(v) = \{v' \in \mathbb{N}^2 : |v_1 - v_1'| + |v_2 - v_2'| = 1\}$, the 1-step neighbourhood of v. The output equation becomes $y^{(v)}(t) = H_z z^{(v)}(t)$, with $y^{(v)}(t) = y((v_1, v_2)\Delta u, t)$. This measurement process thus corresponds to a typical imaging spectrography with *pixels* assigned to grid points on V.

Finally, the reaction-diffusion model Eq. (5) is approximated by the conventional state-space representation

$$dx(t) = f_{\theta_x}(x(t))dt; \tag{8a}$$
$$y(t) = H_{\theta_y} x(t), \tag{8b}$$

with state-vector $x(t) = (z^{(v_1,v_2)}(t))_{v_1,v_2=1}^{\overline{V}} \in \mathbb{R}^{N_x}$ and output-vector $y(t) = (y^{(v_1,v_2)}(t))_{v_1,v_2=1}^{\overline{V}} \in \mathbb{R}^{N_y}$, obtained by collecting $(z^{(v)}, y^{(v)})$ at every point in the square lattice V. The output matrix is $H_{\theta_y} = [I_{\overline{V}^2} \otimes \Phi_{\mathcal{S}_1} \cdots I_{\overline{V}^2} \otimes \Phi_{\mathcal{S}_{N_\mathcal{S}}}]$, where $M_1 \otimes M_2$ denotes the Kronecker product between matrices M_1 and M_2. The model thus corresponds to $N_x = N_\mathcal{S}\overline{V}^2$ state- and $N_y = N_\Lambda \overline{V}^2$ output-variables: A high-dimensional state space when $\overline{V}$ is large (i.e. a finely discretised space with $\Delta u \ll \overline{U}$).

3. PRELIMINARIES: NON-LINEAR FILTERING

We consider the stochastic differential equation (SDE) representation of an autonomous system,

$$dX(t) = f_{\theta_x}(X(t))dt + g_{\theta_b}(X(t))dB(t), \tag{9a}$$
$$Y(t) = h_{\theta_y}(X(t)) + e(t), \tag{9b}$$

with state equation Eq. (9a) describing the time evolution of state process $X : [0, \infty) \times \Omega \to \mathbb{X}$ given its current value, stochastic driving process $B = (B(t))_{t \geq 0}$ and initial state $X(0) = \xi(\omega^*), \omega^* \in \Omega$. The output equation Eq. (9b) describes how the current state is emitted as the output-vector $y(t) \in \mathbb{Y}$, after being corrupted by a measurement noise $e(t) \in \mathbb{R}^{N_y}$. We consider Polish spaces $\mathbb{X} = \mathbb{R}^{N_x}$ and $\mathbb{Y} = \mathbb{R}^{N_y}$. The parameter vector $\theta = (\theta_x, \theta_b, \theta_y)$ determines the functions $f_{\theta_x}(\cdot)$, $g_{\theta_b}(\cdot)$, and $h_{\theta_y}(\cdot)$. For the sake of exposition, we omit the dependency on θ from here on. For all $0 \leq t' < t$, $B(t) - B(t') \sim \mathcal{N}(0, (t - t')I_{N_x})$ and $e(t) \sim \mathcal{N}(0, \Sigma_y)$ with known covariance Σ_y. Finally, we limit ourselves to linear equations $Y(t) = HX(t) + e(t)$ given matrix $H \in \mathbb{R}^{N_y \times N_x}$.

We are interested in estimating the states $(X(t))_{t \geq 0}$ based on a sequence of discrete-time measurements $(Y_1, \dots, Y_k)$, with $Y_k = Y(k\Delta t)$ given a sampling time $\Delta t > 0$. Assuming

a zero-order hold between measurements of the driving noise (that is, $g(X(t)) = g(X(t_{k-1}))$ for all $t \in [t_{k-1}, t_k)$), the discrete-time state dynamics are represented by

$$X_k = \underbrace{X_{k-1} + \int_{t_{k-1}}^{t_k} f(X(t))dt}_{F(X_{k-1})} + g(X_{k-1})\Delta B_k, \tag{10}$$

where $X_k = X(k\Delta t)$, $\Delta B_k = \mathcal{N}(0, \Delta t \cdot I_{N_x})$, and $t_k = k\Delta t$. The discrete-time output equation is $Y_k = h(X_k) + e_k$ with measurement noise $e_k \sim \mathcal{N}(0, \Sigma_y)$. The estimation problem is then formalised by the stochastic process $\pi = (\pi_k)_{k\in\mathbb{N}}$, where $\pi_k = \mathbb{P}(X_k \in \cdot | Y_1, \dots, Y_k)$ is the filtering distribution summarising the uncertainty on X_k given history $(Y_1, \dots, Y_k)$. The distributions π_k are assumed to have density p_k with respect to Lebesgue measure.

We consider the Particle filter (PF, Chopin and Papaspiliopoulos (2020)) approach of computing Monte Carlo approximations $\pi^{N_p} = (\pi_k^{N_p})_{k\in\mathbb{N}}$ by recursively sampling N_p candidates of X_k based on observation Y_k and previous $\pi_{k-1}^{N_p}$. Accounting for high-dimensional spaces, we specialise this approach to a class of localised estimators known as the BPF (Rebeschini and van Handel, 2015). In the following, we overview these approaches.

3.1 Classical particle filters

We assume that X is a Markov chain: for all $k \in \mathbb{N}$ and arbitrary Borel set $A \in \mathcal{B}(\mathbb{R}^{N_x})$,

$$\mathbb{P}(X_k \in A | X_{0:k-1}) = \mathbb{P}(X_k \in A | X_{k-1}). \tag{11}$$

We consider transition kernels $\mathrm{P}_k : \mathbb{R}^{N_x} \times \mathcal{B}(\mathbb{R}^{N_x}) \to \mathbb{R}_{\geq 0}$ such that, for all $k \in \mathbb{N}$, $x \in \mathbb{R}^{N_x}$, and any $A \in \mathcal{B}(\mathbb{R}^{N_x})$,

$$\mathrm{P}_k(x, \mathrm{A}) = \mathbb{P}(X_k \in \mathrm{A} | X_{k-1} = x). \tag{12}$$

The filter π_k can thus be computed recursively by

$$\pi_{k-1} \xrightarrow{\text{prediction}} \tilde{\pi}_k = \mathrm{P}_k \pi_{k-1} \xrightarrow{\text{correction}} \pi_k = \mathrm{C}_k \tilde{\pi}_k,$$

with the functions P_k and C_k satisfying

$$(\mathrm{P}_k \pi_{k-1})(A) \coloneqq \int \mathrm{P}_k(x, A)\pi_{k-1}(dx), \tag{13}$$

$$(\mathrm{C}_k \tilde{\pi}_k)(A) \coloneqq \frac{\int_A \mathrm{C}_k(x)\tilde{\pi}_k(dx)}{\tilde{\pi}_k(\mathrm{C}_k)}, \tag{14}$$

for non-negative functions C_k with $\tilde{\pi}_k(\mathrm{C}_k) > 0$. Here, $\tilde{\pi}_k(\mathrm{C}) \coloneqq \int \mathrm{C}(x)\tilde{\pi}_k(dx)$. In the general case, the recursion $\pi_k = \mathrm{C}_k \mathrm{P}_k \pi_{k-1}$ has no trivial solution. As such, the estimation problem can be solved by a Sequential Monte Carlo approach: The filter π_k is approximated by an empirical distribution $\mathrm{C}_k \mathrm{S}_k^{N_p} \tilde{\pi}_k$ given the sampling operator $\mathrm{S}_k^{N_p} \pi \coloneqq N_p^{-1} \sum_{n=1}^{N_p} \delta_{x^n}$ with samples $x^1, \dots, x^{N_p} \overset{i.i.d}{\sim} \pi$.

Let particles $\{X_k^{(n)}\}_{n=1}^{N_p}$ be N_p mutually independent stochastic processes, all independent of Y_k, solving Eq. (9a). The pairs $(X_k^{(n)}, Y_k)$ are identically distributed with the same distribution as (X_k, Y_k), $n = 1, \dots, N_p$. Now, let $\pi^{N_p} = (\pi_k^{N_p})_{k\in\mathbb{N}}$ be the sequence of empirical distributions

$$\pi_k^{N_p} \triangleq \frac{1}{N_p} \sum_{n=1}^{N_p} w_k^{(n)} \delta_{X_k^{(n)}}, \tag{15}$$

where to each $X_k^{(n)}$ we assign a (normalised) weight $w_k^{(n)}$. Then, we have $\pi^{N_p} \to \pi$ almost surely at a rate slightly lower than $1/\sqrt{N_p}$ (Bain and Crisan, 2008). As the filters π_k are unavailable, we sample i.i.d. $X_k^{(1)}, \dots, X_k^{(N_p)}$ from *importance distributions* ρ_k with densities of the form

$$q(x_{0:k}|y_{1:k}) = q(x_0) \prod_{k'=1}^{k} q(x_{k'}|y_{0:k'}, x_{0:k'-1}), \tag{16}$$

with x_k and y_k denoting a realisation of X and Y at time t_k, respectively. In this setup, the unnormalised weights $\widetilde{w}^{(n)}(x_{0:k}) = w^{(n)}(x_{0:k-1})\widetilde{w}_k^{(n)}$ $(n = 1, \dots, N_p)$ can be computed recursively according to

$$\widetilde{w}_k^{(n)} = \frac{p(y_k|y_{1:k-1}, x_{0:k}^{(n)})p(x_k|x_{k-1}^{(n)})}{q(x_k|y_{1:k}, x_{0:k-1}^{(n)})} \coloneqq l_k(x_k^{(n)}, y_k), \tag{17}$$

whenever new observations Y_k become available. They are normalised through $w_k = [\sum_{n'=1}^{N_p} \widetilde{w}_k^{(n')}]^{-1} \widetilde{w}_k^{(n)}$.

The choice of $q(x_{0:k}|y_{1:k})$ is arbitrary: It only needs to have a support including that of $p(x_{0:k}|y_{0:k})$. Conventionally, we sample particles from the transition distribution of the dynamics Eq. (9a), i.e. $q(x_k|y_{1:k}, x_{0:k-1}) = p(x_k|x_{k-1})$. The importance weights are then based on the likelihood,

$$\widetilde{w}_k^{(n)} \propto w_{k-1}^{(n)} p(y_k|x_k^{(n)}). \tag{18}$$

This choice of proposal density is conventional but not very informative. In practice, only a few selected particles might be assigned relevant weights. In the extreme situation, termed degeneracy, $w_k^{(n)} \approx 1$ for a single n, and $w_k^{(n')} \approx 0$ for all $n' \neq n$. The likelihood is then poorly approximated with large variance $\mathtt{Var}_\rho(\widetilde{w})$.

Optimal sampling distribution The degeneracy problem can be addressed by choosing an importance distribution that minimises the variance of the (unnormalised) weights conditioned not only on $x_{0:k}$ and $y_{0:k-1}$, but also on y_k.

Proposition 1. (Doucet et al., 2000). The (optimal importance) distribution that minimises $\mathtt{Var}_\rho(\widetilde{w})$ has density

$$q(x_k|y_{1:k}, x_{0:k-1}) = \frac{p(y_k|y_{1:k-1}, x_{0:k})p(x_k|x_{k-1})}{p(y_k|y_{1:k-1}, x_{0:k-1})}, \tag{19}$$

where the associated unnormalised weights satisfy

$$\widetilde{w}_k \propto w_{k-1} \underbrace{\int p(y_k|x_{k-1:k})p(x_k|x_{k-1})\, dx_k}_{p(y_k|x_{k-1})}. \tag{20}$$

For processes $X_k = F(X_{k-1}) + g(X_{k-1})\Delta B_k$ with measurement models $Y_k = HX_k + e_k$ (given $e_k \sim \mathcal{N}(0, \Sigma_y)$), the optimal importance distribution has the analytical form

$$X_k | Y_k, X_{k-1} \sim \mathcal{N}(M_k^{\text{opt}}, \Sigma_k^{\text{opt}}) \tag{21}$$

with mean M_k^{opt} and covariance Σ_k^{opt} computed by

$$\Sigma_B = (g(x_{k-1})g(x_{k-1})^\mathsf{T})\Delta t, \tag{22}$$

$$\Sigma_k^{\text{opt}} = (\Sigma_B^{-1} + H^\mathsf{T}\Sigma_y^{-1}H)^{-1}, \tag{23}$$

$$M_k^{\text{opt}} = \Sigma_k^{\text{opt}}(\Sigma_B^{-1}F(x_{k-1})) + H^\mathsf{T}\Sigma_y^{-1}y_k. \tag{24}$$

Moreover, the likelihood distribution satisfies

$$Y_k | X_{k-1} \sim \mathcal{N}(HF(x_{k-1}),\ \Sigma_B + H\Sigma_y H^\mathsf{T}). \tag{25}$$

For positive systems, a common modelling assumption is that $g(X_{k-1}) = \Sigma_B^{1/2} = \mathtt{diag}(X_{k-1})\Sigma_x^{1/2}$, given a $\Sigma_x^{1/2} \succ 0$: This assures a unique and positive solution X of Eq. (9a), (Yang et al., 2020). As such, variances in Eqs. (21) and (25) can be computed efficiently for high-dimensional systems when H has some special structure (such as in Section 2).

3.2 Block particle filtering

In general, regardless of the choice of importance distribution, classical particle filters face degeneracy issues when applied to high-dimensional systems (Snyder et al., 2015). A scalable solution is to design filters that are spatially-localised: Dynamics and observations at a spatial location are assumed to depend only on state-variables associated with its neighbourhood (Rebeschini and van Handel, 2015).

Consider the pair (X_k, Y_k) at each time t_k to be a random field $(X_k, Y_k)_{v\in V}$ indexed by a (finite) undirected graph $\mathcal{G} = (V, E)$. The spaces $\mathbb{X}$ and $\mathbb{Y}$ of X_k and Y_k and the filter π_k can be expressed in the product form

$$\mathbb{X} = \prod_{v\in V}\mathbb{X}^v, \quad \mathbb{Y} = \prod_{v\in V}\mathbb{Y}^v, \quad \pi_k = \bigotimes_{v\in V}\pi_k^v. \tag{26}$$

Here, π_k^v is the conditional distribution on state-space $\mathbb{X}^v$ and $\bigotimes_{v\in V}$ is the product (of measures) over a vertex set V. The transition kernel P_k in Eq. (12) and likelihood function l in Eq. (17) are then decomposed in their localised forms $\mathrm{P}_k(x, A) = \prod_{v\in V}\mathrm{P}_k^v(x, A^{(v)})$, $l_k(x, y) = \prod_{v\in V} l_k^v(x^{(v)}, y^{(v)})$. We focus on the scenario where the graph $\mathcal{G}$ represents a square lattice (as in Section 2.1) with vertex set

$$V = \{v \in \mathbb{N}^2 : 1 \le v_{i=1,2} \le \overline{V}\} \quad (\overline{V} \in \mathbb{N}). \tag{27}$$

Considering the set of non-overlapping blocks

$$\mathcal{K} = \{(v_0 + \{1, \dots, \overline{V}_b\}^2) \cap V : v_0 \in \overline{V}_b\mathbb{N}^2\}, \tag{28}$$

with size $\overline{V}_b < \overline{V}$, the vertices of $\mathcal{G}$ can be partitioned as

$$V = \bigcup_{V_b\in\mathcal{K}} V_b, \quad V_b \cap V_{b'} = \emptyset \text{ for } V_b \neq V_{b'}, \quad V_b, V_{b'} \in \mathcal{K}.$$

Under such decomposition, we define $X^{(V_b,n)} = \otimes_{v\in V_b} X^{(v,n)}$ and $Y^{(V_b)} = \otimes_{v\in V_b} Y^{(v)}$ to be the (n-th particle) state- and output-variables associated with block $V_b \in \mathcal{K}$.

Let the blocking operator be $\mathrm{B}\pi_k \coloneqq \bigotimes_{V_b\in\mathcal{K}} \mathrm{B}^{V_b}\pi_k$, where $\mathrm{B}^{V_b}\pi_k$ denotes the marginal distribution of π_k on $\prod_{v\in V_b}\mathbb{X}^v$. In this case, we compute filters π_k recursively by

$$\pi_{k-1} \xrightarrow{\text{prediction}} \tilde{\pi}_k = \mathrm{P}_k\pi_{k-1} \xrightarrow{\text{correction}} \pi_k = \mathrm{C}_k\mathrm{B}\tilde{\pi}_k.$$

The BPF algorithm is then implemented as in Algorithm 1 [†]. In this approach, the convergence properties of the particle filter become dependent on the cardinality of individual blocks, $|V_b| = \overline{V}_b^2$ ($V_b \in \mathcal{K}$), rather than on the cardinality of the original lattice, $|V| = \overline{V}^2$.

4. CASE-STUDY: THE OREGONATOR SYSTEM

In this section, we present the results obtained by the BPF (Section 3.2) when used to reconstruct the concentrations from a benchmark reaction-diffusion system. Specifically, we focus on the task of estimating the state from a bistable, oscillatory Belousov-Zhabotinskii (BZ) reaction system (Zhabotinsky, 1991). The Oregonator (Field and Noyes, 1974) consists on the simplest realistic model of the BZ reaction dynamics, with network

[†] In its standard formulation, Algorithm 1 assumes P_k^v and l_k^v to be obtained from the appropriate marginals with densities $p(\tilde{x}_k|x_{k-1}^{(n)})$ and $p(y_k|\tilde{x}_k^{(n)})$, respectively. When the optimal importance distribution is used, the marginals are taken from Eqs. (21) and (25).

Algorithm 1: Block particle filter

Initialise $\pi_0^{N_p}$ with a desirable distribution;
for $k = 1, 2, \dots, K$ **do**
  Resample $X_{k-1}^{(1)}, \dots, X_{k-1}^{(N_p)} \overset{i.i.d.}{\sim} \pi_{k-1}^{N_p}$;
  Sample $\widetilde{X}_k^{(v,1)}, \dots, \widetilde{X}_k^{(v,N_p)} \overset{i.i.d.}{\sim} \mathrm{P}_k^v(x_{k-1}^{(n)}, \cdot),\ \forall v \in V$;
  Compute $w_k^{(V_b,n)} = \dfrac{\prod_{v\in V_b} l^v\left(\tilde{x}_k^{(v,n)}, y_k^{(v)}\right)}{\sum_{n'=1}^{N_p}\prod_{v\in V_b} l^v\left(\tilde{x}_k^{(v,n')}, y_k^{(v)}\right)}$ for every $V_b \in \mathcal{K}$ and $n = 1, \dots, N_p$;
  Let $\pi_k^{N_p} = \bigotimes_{V_b\in\mathcal{K}}\sum_{n=1}^{N_p} w_k^{(V_b,n)}\delta_{\tilde{x}_k^{(V_b,n)}}$.
end

$$\begin{aligned} 2\mathcal{S}_1 &\xrightarrow{\kappa_1} \mathcal{S}_4 + \mathcal{S}_5 & \mathcal{S}_1 + \mathcal{S}_3 &\xrightarrow{\kappa_2} 2\mathcal{S}_5 \\ \mathcal{S}_1 + \mathcal{S}_4 &\xrightarrow{\kappa_3} 2\mathcal{S}_1 + 2\mathcal{S}_2 & \mathcal{S}_3 + \mathcal{S}_4 &\xrightarrow{\kappa_4} \mathcal{S}_1 + \mathcal{S}_5 \\ \mathcal{S}_2 + \mathcal{S}_6 &\xrightarrow{\kappa_5} 0.5\sigma\mathcal{S}_3, \end{aligned} \tag{29}$$

with concentrations $z = \{[\mathcal{S}_1], \dots, [\mathcal{S}_6]\}$. Typically, species $(\mathcal{S}_4, \mathcal{S}_5, \mathcal{S}_6)$ are present in high densities and thus concentrations $z_{4,5,6}(s, t)$ are assumed constant on the timescale of a few oscillations. Moreover, when z_3 is slowly varying, the dynamics of the system are summarised by the evolution of z_1 and z_2 only, with f_z represented in a scaled form [‡].

We consider the system on a quasi-two-dimensional space U with size $\overline{U} = 2$ units-of-space. We discretise U with $\Delta u = 0.02$ units-of-size and consider the finite-difference approximation of the dynamics (Section 2.1) onto lattice V with $\overline{V} = \overline{U}/\Delta u = 100$ grid-points in each dimension. The dynamics are then discretised in time with $\Delta t = 0.01$. The process is assumed to be subjected to driving noise with coefficient $g(X_{k-1}) = \sigma_x\mathtt{diag}(X_{k-1})$ given $\sigma_x = 10^{-2}$. The measurement process is defined by $H = [I_{\overline{V}^2} \otimes \Phi_{\mathcal{S}_1}\ I_{\overline{V}^2} \otimes \Phi_{\mathcal{S}_2}]$ with spectra $(\Phi_{\mathcal{S}_1}, \Phi_{\mathcal{S}_2})$ collected at 10 equally-spaced wavelengths $\lambda \in [0, 50)$ through response functions $\phi_{S_1}(\lambda) = \exp\left(-\frac{(\lambda-10)^2}{30}\right)$ and $\phi_{S_2}(\lambda) = \exp\left(-\frac{(\lambda-40)^2}{30}\right)$. The measurement noise is $e_k \sim \mathcal{N}(0, \sigma_y^2 I_{N_y})$ with $\sigma_y^2 = 10^{-5}$. Finally, we assume initial condition $X_0 \sim \delta_{x_0}$, with δ denoting the Dirac delta distribution and x_0 is the non-trivial steady-state solution the Oregonator dynamics.

For the estimation task, we design a BPF by partitioning the lattice V into blocks with size $\overline{V}_b = 5$ (with a total of $(\overline{V}/\overline{V}_b)^2 = 400$ blocks). For each block, we consider an ensemble of $N_p = 128$ particles, each assumed to evolve according to the same dynamics as the simulation system. We analyse the filtering results when particles and importance weights are obtained using the conventional ($p(x_k|x_{k-1})$) or the optimal (Eq. 19) importance density.

4.1 Estimation Results

We apply the BPF with the described configuration to estimate concentrations $z = (z_1, z_2)$ from the Oregonator model using the generated spectral data. The results are shown in Fig. 1 for different times t_k through the

[‡] From its mass-action dynamics, the Oregonator has dynamics $f_z(\cdot)$ in scaled form, defined by $[f_z]_1 = \epsilon^{-1}\left(z_1(1 - z_1) - \frac{\sigma z_2(z_1-q)}{z_1+q}\right) + D_{z_1}\nabla^2 z_1$ and $[f_z]_2 = (z_1 - z_2) + D_{z_2}\nabla^2 z_2$ (Keener and Tyson, 1986). The dimensionless constants are $(\epsilon, \sigma, q) = (0.08, 0.95, 0.0075)$. As for the diffusion coefficients, $(D_{z_1}, D_{z_2}) = (5 \times 10^{-4}, 5 \times 10^{-6})$.

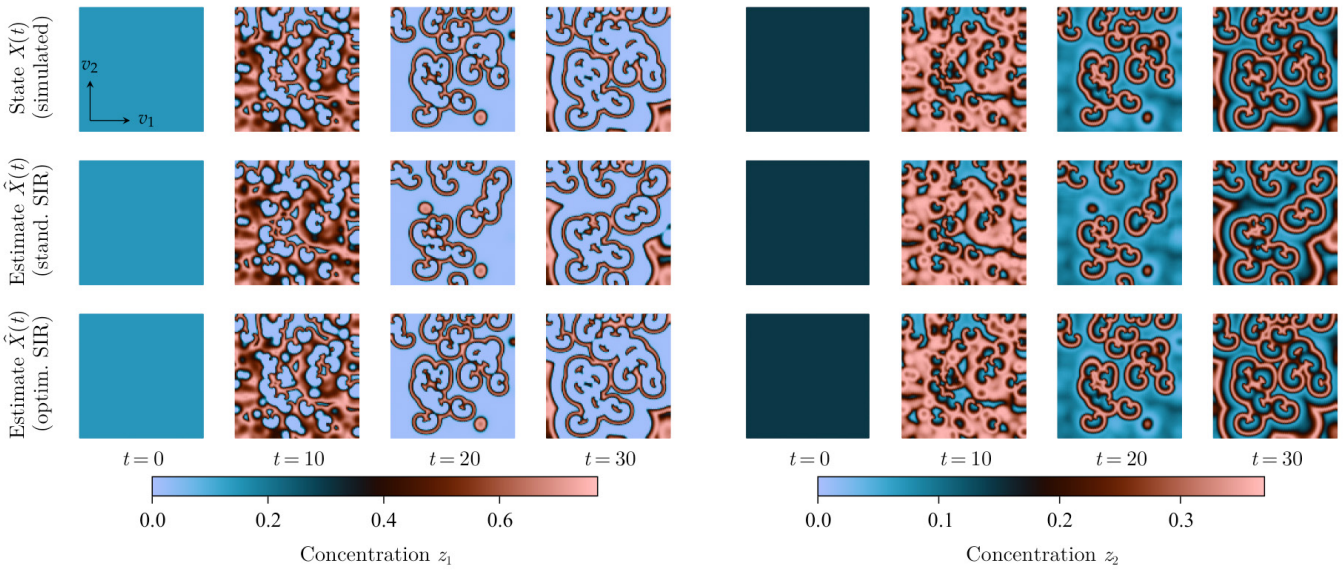

Fig. 1. Snapshots of $[X_k]_1$ (4 first columns) and $[X_k]_2$ (4 last columns) for the Oregonator model at different times t_k. First row shows a draw for X_k used to generate observation y_k. We estimated $\hat{X}_k$ by the BPF with a standard (second row) and an optimal (third row) choice of importance sampling distribution with resampling steps (SIR).

mean of the approximated conditional distribution π_k, i.e. $\hat{X}_k^{(V_b)} := \sum_{n=1}^{N_p} w_k^{(V_b,n)} X_k^{(V_b,n)}$, and the actual state X_k. From the evolution of the system, we observe the formation of oscillating "spiral-patterns" after a few iterations. The system is expected to switch back and forth between activator dominated states (species $\mathcal{S}_1$ is present in high-concentrations) and inhibitor dominated states (species $\mathcal{S}_2$ is present in high-concentrations). Initially, we observe concentrations to oscillate around the entire spatial domain. In this transient period ($t < 20$ units-of-time), the emerging patterns continuously change their shape and position. Afterwards, the patterns can be observed to emerge around fixed locations (as observed in $t \geq 20$ units-of-time). The results indicate that the BPF using the standard choice of importance distribution obtains poor estimates of the state: The particle approximations diverge during the transient region and ultimately form different patterns compared with those of the actual system. Conversely, the filter using the optimal importance distributions is able to accurately estimate the pattern-formation from the system.

In the following, we investigate the mismatch in Fig. 1 by comparative metrics between both approximations. For each block $V_b \in \mathcal{K}$, we quantify the estimation accuracy in terms of the Root Mean Square Error (RMSE),

$$\mathrm{RMSE}_k^{(V_b)} = \sqrt{\|\hat{y}_k^{(V_b)} - y_k^{(V_b)}\|_2^2}, \tag{30}$$

with predicted outputs $\hat{y}_k^{(V_b)} := \sum_{n=1}^{N_p} w_k^{(V_b,n)} (H x_k^{(V_b,n)})$, observation y_k, and $w_k^{(V_b,n)}$ and $x_k^{(V_b,n)}$ denoting, respectively, the normalised weight and the realisation of the n-th particle X_k in a block V_b. Adding the accuracy obtained at each block, we arrive at the results shown in Fig. 2.

The results (Fig. 2) show that the estimation accuracy oscillates due to the bistability of the system; crests relate to periods in which either the activator or inhibitor are dominated states, and troughs correspond to the transient periods. Moreover, this metric further highlights that accurate estimates are only obtained when the filter

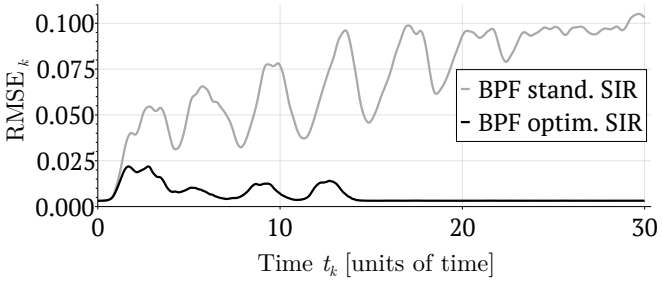

Fig. 2. RMSE_k between observation y_k and prediction $\hat{y}_k$.

considers the optimal importance distribution. In this case, the estimation error increases slightly during $t \leq 15$, before decreasing to almost zero. Conversely, we observe the estimation accuracy to deteriorate quickly when the standard importance distribution is used.

Furthermore, we compare the performance of both filters by computing the marginal likelihood function $p(y_{1:K})$ through Monte Carlo integration over the state space. Specifically, for each block $V_b \in \mathcal{K}$ we approximate the density

$$p(y_{1:K}^{(V_b)}) \approx \prod_{k=1}^{K} \left(\frac{1}{N_p} \sum_{n=1}^{N_p} w_k^{(V_b,n)} \delta_{X_k^{(V_b,n)}} \right). \tag{31}$$

Avoiding numerical issues, we approximate the marginal likelihood in the logarithmic scale. Finally, $\texttt{log}(p(y_{1:K}))$ over the full output-space is approximated by adding together the block-wise computed values, shown in Fig. 3.

Here, the marginalised likelihood is the probability of the collection of observations $y_{1:K}$ given the model in Eq. (9), without assuming a particular realisation of the state. Under such definition, approximations of $p(y_{1:k})$ should have similar results regardless of which importance sampling distribution is adopted for the filtering, provided that enough particles are used to avoid degeneracy. However, the approximation obtained from the filter with the standard choice does not match that obtained from its optimal counterpart. The relatively small probability densities of

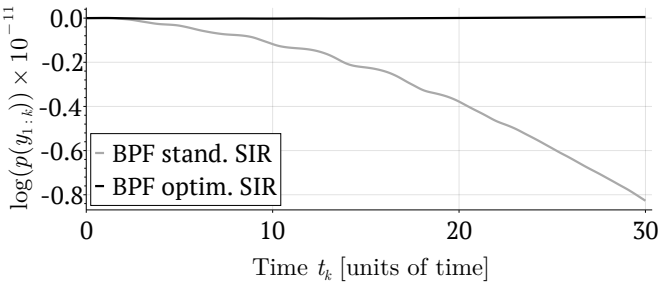

Fig. 3. Estimated log-marginal likelihood $\texttt{log}(p(y_{1:k}))$.

the collection of observations in Fig. 3 follows as a direct consequence of path degeneracy.

5. CONCLUDING REMARKS

In this work, we highlighted a sequential Monte Carlo approximation to the solution to the filtering equations (13) and (14). Under the context of a reaction-diffusion system, we examined how the BPF dealt with the high-dimensionality of the state under a standard and an optimal choice of importance sampling distribution.

BZ reaction systems are characterised by long-term unpredictability arising from an extreme sensitivity to initial conditions and process noise. Depending on these parameters, the system may or may not present oscillating patterns as in Fig. 1. Here, we presented the scenario of reconstructing the concentrations in the case of moderate values of the system noise variance. We would like to remark that the benefits from the optimal proposal will be negligible as this noise becomes small, as discussed in Snyder et al. (2015).

The BPF has some intrinsic limitations by design. For example, the bias introduced as a result of the blocking operator is not spatially homogeneous. In an extended version of this work, we intend to (i) assess any spatial inhomogeneity, (ii) investigate how some extensions to the BPF address the bias, and (iii) draw comparisons to other high-dimensional filters in a broader experimental study.

REFERENCES

Ardizzone, V., Lewandowski, P., Luk, M.H., Tse, Y.C., Kwong, N.H., Lücke, A., Abbarchi, M., Baudin, E., Galopin, E., Bloch, J., et al. (2013). Formation and control of Turing patterns in a coherent quantum fluid. *Scientific Reports*, 3(1).

Asensio, M.I. and Ferragut, L. (2002). On a wildland fire model with radiation. *International Journal for Numerical Methods in Engineering*, 54(1).

Bain, A. and Crisan, D. (2008). *Fundamentals of Stochastic Filtering*. Stochastic Modelling and Applied Probability. Springer New York.

Chopin, N. and Papaspiliopoulos, O. (2020). *An introduction to Sequential Monte Carlo*. Springer.

Djuric, P.M., Lu, T., and Bugallo, M.F. (2007). Multiple Particle Filtering. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3.

Doucet, A., Godsill, S.J., and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10.

Doucet, A., Briers, M., and Sénécal, S. (2006). Efficient Block Sampling Strategies for Sequential Monte Carlo Methods. *Journal of Computational and Graphical Statistics*, 15(3).

Ferreira, S.C., Martins, M.L., and Vilela, M.J. (2002). Reaction-diffusion model for the growth of avascular tumor. *Phys. Rev. E*, 65.

Field, R.J. and Noyes, R.M. (1974). Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction. *Journal of Chemical Physics*, 60.

Gilks, W.R. and Berzuini, C. (2001). Following a moving target: Monte Carlo inference for dynamic Bayesian models. *Journal of the Royal Statistical Society, Series B*, 63, 127–146.

Green, J.B.A. and Sharpe, J. (2015). Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development*, 142(7).

Hogea, C., Davatzikos, C., and Biros, G. (2008). An image-driven parameter estimation problem for a reaction–diffusion glioma growth model with mass effects. *Journal of Mathematical Biology*, 56(6).

Keener, J.P. and Tyson, J.J. (1986). Spiral waves in the Belousov-Zhabotinskii reaction. *Physica D: Nonlinear Phenomena*, 21(2).

Kitagawa, G. and Gersch, W. (1996). *Smoothness Prior Analysis of Time Series*. Lecture Notes in Statistics. Springer New York.

Kondo, S. (2009). How animals get their skin patterns: fish pigment pattern as a live Turing wave. *Systems Biology*.

Kondo, S. and Miura, T. (2010). Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation. *Science*, 329(5999).

Murray, J.D. (2002). *Mathematical Biology: I. An introduction*. Springer.

Musso, C., Oudjane, N., and Gland, F.L. (2001). Improving regularised particle filters. In *Sequential Monte Carlo Methods in Practice*.

Nadler, B. and Coifman, R.R. (2005). Partial least squares, Beer's law and the net analyte signal: statistical modeling and analysis. *Journal of Chemometrics*, 19(1), 45–54.

Nakamasu, A., Takahashi, G., Kanbe, A., and Kondo, S. (2009). Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. *Proceedings of the National Academy of Sciences*, 106(21).

Rebeschini, P. and van Handel, R. (2015). Can local particle filters beat the curse of dimensionality? *The Annals of Applied Probability*, 25(5).

Snyder, C., Bengtsson, T., and Morzfeld, M. (2015). Performance bounds for Particle Filters using the Optimal Proposal. *Monthly Weather Review*, 143, 4750–4761.

Strauss, W.A. (2007). *Partial differential equations: An introduction*. John Wiley & Sons, 2 edition.

Turing, A.M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B*, 237(641).

Van Leeuwen, P.J. (2003). A variance-minimizing filter for large-scale applications. *Monthly Weather Review*, 131.

Witkin, A. and Kass, M. (1991). Reaction-Diffusion Textures. *SIGGRAPH Comput. Graph.*, 25(4).

Yang, Y., Jiang, D., O'Regan, D., and Alsaedi, A. (2020). Dynamics of the Stochastic Belousov-Zhabotinskii Chemical Reaction Model. *Mathematics*, 8.

Zhabotinsky, A.M. (1991). A history of chemical oscillations and waves. *Chaos*, 1.