Diffusivity Estimation for Activator–Inhibitor Models: Theory and Application to Intracellular Dynamics of the Actin Cytoskeleton

Gregor Pasemann[1] · Sven Flemming[2] · Sergio Alonso[3] · Carsten Beta[2] · Wilhelm Stannat[1]

Abstract

A theory for diffusivity estimation for spatially extended activator–inhibitor dynamics modeling the evolution of intracellular signaling networks is developed in the mathematical framework of stochastic reaction–diffusion systems. In order to account for model uncertainties, we extend the results for parameter estimation for semilinear stochastic partial differential equations, as developed in Pasemann and Stannat (Electron J Stat 14(1):547–579, 2020), to the problem of joint estimation of diffusivity and parametrized reaction terms. Our theoretical findings are applied to the estimation of effective diffusivity of signaling components contributing to intracellular dynamics of the actin cytoskeleton in the model organism *Dictyostelium discoideum*.

Keywords Parametric drift estimation · Stochastic reaction–diffusion systems · Maximum likelihood estimation · Actin cytoskeleton dynamics

---

Communicated by Philipp M Altrock.

✉ Gregor Pasemann
   pasemann@math.tu-berlin.de

   Sven Flemming
   svenflem@uni-potsdam.de

   Sergio Alonso
   s.alonso@upc.edu

   Carsten Beta
   beta@uni-potsdam.de

   Wilhelm Stannat
   stannat@math.tu-berlin.de

1 Institute of Mathematics, Technische Universität Berlin, 10623 Berlin, Germany

2 Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany

3 Department of Physics, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

Springer

Mathematics Subject Classification 60H15 · 62F10

1 Introduction

The purpose of this paper is to develop the mathematical theory for statistical inference methods for the parameter estimation of stochastic reaction–diffusion systems modeling spatially extended signaling networks in cellular systems. Such signaling networks are one of the central topics in cell biology and biophysics as they provide the basis for essential processes including cell division, cell differentiation, and cell motility (Peter 2017). Nonlinearities in these network may cause rich spatiotemporal behavior including the emergence of oscillations and waves (Beta and Kruse 2017). Furthermore, alterations and deficiencies in the network topology can explain many pathologies and play a key role in diseases such as cancer (Condeelis et al. 2005). Here we present a method to estimate both diffusivity and reaction terms of a stochastic reaction–diffusion system, given space–time structured data of local concentrations of signaling components. We mainly focus on the estimation of diffusivity, whose precision can be increased by simultaneous calibration of the reaction terms.

To test this approach, we use fluorescence microscopy recordings of the actin dynamics in the cortex of cells of the social amoeba *Dictyostelium discoideum*, a well-established model organism for the study of a wide range of actin-dependent processes (Annesley and Fisher 2009). A recently introduced stochastic reaction–diffusion model could reproduce many features of the dynamical patterns observed in the cortex of these cells including excitable and bistable states (Alonso et al. 2018; Flemming et al. 2020; Moreno et al. 2020). In combination with the experimental data, this model will serve as a specific test case to exemplify our mathematical approach. Since in real-world applications the available data will not allow for calibrating and validating detailed mathematical models, in this paper we will be primarily interested in minimal models that are still capable of generating all observed dynamical features at correct physical magnitudes. The developed estimation techniques should in practice be as robust as possible w.r.t. uncertainty and even misspecification of the unknown real dynamics.

The impact of diffusion and reaction in a given model will be of fundamentally different structure and it is one of the main mathematical challenges to separate these impacts in the data in order to come to valid parameter estimations. On the more mathematical side, diffusion corresponds to a second-order partial differential operator—resulting in a strong spatial coupling in the given data, whereas the reaction corresponds to a lower order, in fact 0 order, in general resulting in highly nonlinear local interactions in the data. For introductory purposes, let us assume that our data is given in terms of a space- and time-continuous field $X(t, x)$ on $[0, T] \times \mathcal{D}$, where T is the terminal time of our observations and $\mathcal{D} \subset \mathbb{R}^2$ a rectangular domain that corresponds to a chosen data segment in a given experiment. Although in practice the given data will be discrete w.r.t. both space and time, we will be interested in applications where the resolution is high enough in order to approximate the data by such a continuous field. Our standing assumption is that $X(t, x)$ is generated by a dynamical

system of the form

$$\partial_t X(t, x) = \theta_0 \Delta X(t, x) + \mathcal{F}_X(t, x), \tag{1}$$

where Δ is the Laplacian, given by $\Delta X(t, x) = \partial_{x_1}^2 X(t, x) + \partial_{x_2}^2 X(t, x)$, $x = (x_1, x_2)$, which captures the diffusive spreading in the dynamics of $X(t, x)$. The intensity of the diffusion is given by the diffusivity θ_0. Finally, $\mathcal{F}$ is a generic term, depending on the solution field $X(t, x)$, which describes all non-diffusive effects present in $X(t, x)$, whether they are known or unknown. A natural approach to extract θ_0 from the data is to use a "cutting-out estimator" of the form

$$\hat{\theta}_0 = \frac{\int_0^T \int_{\mathcal{D}} Y(t, x) \partial_t X(t, x) \mathrm{d}x \mathrm{d}t}{\int_0^T \int_{\mathcal{D}} Y(t, x) \Delta X(t, x) \mathrm{d}x \mathrm{d}t} = \frac{\int_0^T \langle Y, \partial_t X \rangle \mathrm{d}t}{\int_0^T \langle Y, \Delta X \rangle \mathrm{d}t}, \tag{2}$$

where $Y(t, x)$ is a suitable test function. In the second fraction of (2), we use the functional form for readability. In particular, we write $X = X(t, x)$ for the solution field. We will also write $X_t = X(t, \cdot)$ for the (spatially varying) solution field at a fixed time t. In order to ease notation, we will use this functional form from now on throughout the paper. It is possible to derive (2) from a least squares approach by minimizing $\theta \mapsto \|\partial_t X - \theta \Delta X\|^2$ with a suitably chosen norm. If the non-diffusive effects described by $\mathcal{F}$ are negligible, we see by plugging (1) into (2) that $\hat{\theta}_0$ is close to θ_0. If a sound approximation $\overline{\mathcal{F}}$ to $\mathcal{F}$ is known, the estimator can be made more precise by substituting $\partial_t X$ by $\partial_t X - \overline{\mathcal{F}}_X$ in (2). A usual choice for Y is a reweighted spectral cutoff of X, which leads to the spectral approach described below.

Under additional model assumptions, e.g., if (1) is in fact a stochastic partial differential equation (SPDE) driven by Gaussian white noise, a rather developed parameter estimation theory for θ_0 has been established in Pasemann and Stannat (2020) on the basis of maximum likelihood estimation (MLE).

In this paper, we are interested in further taking into account also those parts of $\mathcal{F}_X$ corresponding to local nonlinear reactions. As a particular example, we will focus on a recently introduced stochastic reaction–diffusion system of FitzHugh–Nagumo type that captures many aspects of the dynamical wave patterns observed in the cortex of motile amoeboid cells (Flemming et al. 2020),

$$\partial_t U = D_U \Delta U + k_1 U(u_0 - U)(U - u_0 a) - k_2 V + \xi, \tag{3}$$
$$\partial_t V = D_V \Delta V + \epsilon(bU - V). \tag{4}$$

Here, we identify $\theta_0 = D_U$ and the only observed data is the activator variable, i.e., $X = U$.

Therefore, in this example the non-diffusive part of the dynamics will be further decomposed as $\mathcal{F} = F + \xi$, where ξ is Gaussian white noise and $F = F(U)$ encodes the non-Markovian reaction dynamics of the activator. The inhibitor component V in the above reaction–diffusion system is then incorporated for minimal modeling purposes to allow the formation of traveling waves in the activator variable U that are

indeed observed in the time evolution of the actin concentration. This model and its dynamical features is explained in detail in Sect. 2.1.1.

As noted before, it is desirable to include this additional knowledge into the estimation procedure (2) by subtracting a suitable approximation $\overline{\mathcal{F}}$ of the—in practice—unknown $\mathcal{F}$. Although (3), (4) suggest an explicit parametric form for $\overline{\mathcal{F}}$, it is a priori not clear how to quantify the nuisance parameters appearing in the system. Thus, an (approximate) model for the data is known *qualitatively*, based on the observed dynamics, but not *quantitatively*. In order to resolve this issue, we extend (2) and adopt a joint maximum likelihood estimation of θ_0 and various nuisance parameters.

The field of statistical inference for SPDEs is rapidly growing, see Cialenco (2018) for a recent survey. The spectral approach to drift estimation was pioneered by Hübner et al. (1993), Huebner and Rozovskii (1995) and subsequently extended by various works, see, e.g., Huebner et al. (1997), Lototsky and Rosovskii (1999), Lototsky and Rozovskii (2000) for the case of non-diagonalizable linear evolution equations. In Cialenco and Glatt-Holtz (2011), the stochastic Navier–Stokes equations have been analyzed as a first example of a nonlinear evolution equation. This has been generalized by Pasemann and Stannat (2020) to semilinear SPDEs. Joint parameter estimation for linear evolution equations is treated in Huebner (1993), Lototsky (2003), see also Piterbarg and Rozovskii (1996) for a discussion. Besides the spectral approach, other measurement schemes have been studied. See, e.g., Pospíšil and Tribe (2007), Bibinger and Trabs (2020), Bibinger and Trabs (2019), Chong (2019a), Chong (2019b), Khalil and Tudor (2019), Cialenco and Huang (2019), Cialenco et al. (2020), Cialenco and Kim (2020), Kaino and Uchida (2019) for the case of discrete observations in space and time. Recently, the local approach has been worked out in Altmeyer and Reiß (2020) for linear equations, was subsequently generalized in Altmeyer et al. (2020b) to the semilinear case and applied to a stochastic cell repolarization model in Altmeyer et al. (2020a).

The paper is structured as follows: In Sect. 2, we give a theory for joint diffusivity and reaction parameter estimation for a class of semilinear SPDEs and study the spatial high-frequency asymptotics. Special emphasis is put on the FitzHugh–Nagumo system. In Sect. 3, the biophysical context for these models is discussed. The performance of our method on simulated and real data is evaluated in Sect. 4.

2 Maximum Likelihood Estimation for Activator–Inhibitor Models

In this section, we develop a theory for parameter estimation for a class of semilinear SPDE using a maximum likelihood ansatz. The application we are aiming at is an activator–inhibitor model as in Flemming et al. (2020). More precisely, we show under mild conditions that the diffusivity of such a system can be identified in finite time given high spatial resolution and observing only the activator component.

2.1 The Model and Basic Properties

Let us first introduce the abstract mathematical setting in which we are going to derive our main theoretical results. We work in spatial dimension $d \geq 1$. Given a bounded domain $\mathcal{D} = [0, L_1] \times \cdots \times [0, L_d] \subset \mathbb{R}^d$, $L_1, \ldots, L_d > 0$, we consider the following parameter estimation problem for the semilinear SPDE

$$\mathrm{d}X_t = \big(\theta_0 \Delta X_t + F_{\theta_1,\ldots,\theta_K}(X)\big) \mathrm{d}t + B\mathrm{d}W_t \tag{5}$$

with periodic boundary conditions for Δ on the Hilbert space $H = \bar{L}^2(\mathcal{D}) = \{u \in L^2(\mathcal{D})| \int_{\mathcal{D}} u\mathrm{d}x = 0\}$, together with initial condition $X_0 \in H$. We allow the non-linear term F to depend on additional (nuisance) parameters $\theta_1, \ldots, \theta_K$ and write $\theta = (\theta_0, \ldots, \theta_K)^T$, $\theta_{1:K} = (\theta_1, \ldots, \theta_K)$ for short. Without further mentioning it, we assume that $\theta \in \Theta$ for a fixed parameter space Θ, e.g., $\Theta = \mathbb{R}_+^K$. Next, W is a cylindrical Wiener process modeling Gaussian space–time white noise, that is, $\mathbb{E}[\dot{W}(t, x)] = 0$ and $\mathbb{E}[\dot{W}(t, x)\dot{W}(s, y)] = \delta(t - s)\delta(x - y)$. In order to introduce spatial correlation, we use a dispersion operator of the form $B = \sigma(-\Delta)^{-\gamma}$ with $\sigma > 0$ and $\gamma > d/4$, describing spectral decay of the noise intensity. Here, σ is the overall noise intensity, and γ quantifies the decay of the noise for large frequencies in Fourier space. In addition, γ determines the spatial smoothness of X, see Sect. 2.1.2. The condition $\gamma > d/4$ ensures that the covariance operator BB^T is of trace class, which is a standard assumption for well-posedness of (5), cf. Liu and Röckner (2015).

Denote by $(\lambda_k)_{k\geq 0}$ the eigenvalues of $-\Delta$, ordered increasingly, with corresponding eigenfunctions $(\Phi_k)_{k\geq 0}$. It is well known (Weyl 1911; Shubin 2001) that $\lambda_k \asymp \Lambda k^{2/d}$ for a constant $\Lambda > 0$, i.e., $\lim_{k\to\infty} \lambda_k/(\Lambda k^{2/d}) = 1$. The proportionality constant Λ is known explicitly [see, e.g., Shubin (2001, Proposition 13.1)] and depends on the domain $\mathcal{D}$. Let $P_N : H \to H$ be the projection onto the span of the first N eigenfunctions, and set $X^N := P_N X$. For later use, we denote by I the identity operator acting on H. For $s \in \mathbb{R}$, we write $H^s := D((-\Delta)^{s/2})$ for the domain of $(-\Delta)^{s/2}$, which is given by

$$(-\Delta)^{s/2}x = \sum_{k=1}^{\infty} \lambda_k^{s/2}\langle\Phi_k, x\rangle\Phi_k,$$

and abbreviate $|\cdot|_s := |\cdot|_{H^s}$ for the norm on that space whenever convenient. We assume that the initial condition X_0 is regular enough, i.e., it satisfies $\mathbb{E}[|X_0|_s^p] < \infty$ for any $s \geq 0$, $p \geq 1$, without further mentioning it in the forthcoming statements. We will use the following general class of conditions with $s \geq 0$ in order to describe the regularity of X:

(A_s) For any $p \geq 1$, it holds

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t|_s^p\right] < \infty. \tag{6}$$

Our standing assumption is that X is well posed in the sense that there exists a proba-
bilistically and analytically weak solution $X \in C(0, T; H)$ to (5), unique in the sense
of probability law, such that (A_0) holds. This is a consequence, for example, of the
assumptions from Liu and Röckner (2015, Theorem 5.1.3).

2.1.1 An Activator–Inhibitor Model

An important example for our analysis is given by the following FitzHugh–Nagumo
type system of equations in $d \leq 2$ [cf. Flemming et al. (2020)]:

$$dU_t = (D_U \Delta U_t + k_1 f(|U_t|_{L^2}, U_t) - k_2 V_t)dt + B dW_t, \tag{7}$$
$$dV_t = (D_V \Delta V_t + \epsilon(b U_t - V_t))dt, \tag{8}$$

together with sufficiently smooth initial conditions. Here, f is a bistable third-order
polynomial $f(x, u) = u(u_0 - u)(u - a(x)u_0)$, and $a \in C_b^1(\mathbb{R}, \mathbb{R})$ is a bounded
and continuously differentiable function with bounded derivative. The boundedness
condition for a is not essential to the dynamics of U and can be realized in practice
by a suitable cutoff function.

The FitzHugh–Nagumo system Fitzhugh (1961), Nagumo et al. (1962) originated as
a minimal model capable of generating excitable pulses mimicking action potentials
in neuroscience. Its two components U and V are called *activator* and *inhibitor*,
respectively.

The spatial extension of the Fitzhugh–Nagumo system, obtained via diffusive cou-
pling, is used to model propagation of excitable pulses and two-phase dynamics. In
the case of the two phases, low and high concentration of the activator U are realized
as the stable fixed points of the third-order polynomial f at 0 and u_0. The unstable
fixed point $au_0, a \in (0, 1)$, separates the domains of attraction of the two stable fixed
points. The interplay between spatial diffusion, causing a smoothing out of concen-
tration gradients with rate D_U, and the local reaction forcing f, causing convergence
of the activator to one of the stable phases with rate k_1, leads to the formation of tran-
sition phases between regions with low or high concentration of U. The parameters
determine shape and velocity of the transition phases, e.g., low values of a enhance the
growth of regions with high activator concentration. This corresponds to the excitable
regime, as explained in Flemming et al. (2020).

Conversely, a high concentration of the inhibitor V leads to a decay in the activator
U, with rate k_2. In the excitable regime, this mechanism leads to moving activator
wave fronts. The inhibitor is generated with rate ϵb in the presence of U and decays at
rate ϵ. Its spatial evolution is determined by diffusion with rate D_V. Finally, choosing
a as a functional depending on the total activator concentration introduces a feedback
control that allows to stabilize the dynamics.

A detailed discussion of the relevance for cell biology is given in Sect. 3. More
information on the FitzHugh–Nagumo model and related models can be found in
Ermentrout and Terman (2010).

For this model, we can find a representation of the above type (5) as follows: Using
the variation of constants formula, the solution V to (8) with initial condition $V_0 = 0$

can be written as $V_t = \epsilon b \int_0^t e^{(t-r)(D_V \Delta - \epsilon I)} U_r dr$. Inserting this representation into (7) yields the following reformulation

$$
\begin{aligned}
dU_t &= \Big(D_U \Delta U_t + k_1 U_t (u_0 - U_t)(U_t - au_0) \\
&\quad - k_2 \epsilon b \int_0^t e^{(t-r)(D_V \Delta - \epsilon I)} U_r dr \Big) dt + B dW_t \\
&= (\theta_0 \Delta U_t + \theta_1 F_1(U_t) + \theta_2 F_2(U_t) + \theta_3 F_3(U)(t)) \, dt + B dW_t
\end{aligned} \tag{9}
$$

of the activator–inhibitor model (7), (8) by setting $\theta_0 = D_U$, $\theta_1 = k_1 u_0 \bar{a}$, $\theta_2 = k_1$, $\theta_3 = k_2 \epsilon b$, $\overline{F} = 0$ for some $\bar{a} > 0$ and

$$
F_1(U) = -\frac{a(|U|_{L^2})}{\bar{a}} U(u_0 - U), \tag{10}
$$

$$
F_2(U) = U^2(u_0 - U), \tag{11}
$$

$$
F_3(U)(t) = -\int_0^t e^{(t-r)(D_V \Delta - \epsilon I)} U_r dr. \tag{12}
$$

Here $e^{D_V \Delta - \epsilon I}$ is the semigroup generated by $D_V \Delta - \epsilon I$. Note that F_3 now depends on the whole trajectory of U, so that the resulting stochastic evolution Eq. (9) is no longer Markovian.

For the activator–inhibitor system (7), (8), we can verify well-posedness directly. For completeness, we state the optimal regularity results for both U and V, but our main focus lies on the observed variable $X = U$.

Proposition 1 *Let $\gamma > d/4 + 1/2$. Then there is a unique solution (U, V) to (7), (8). Furthermore, U satisfies (A_s) for any $s < 2\gamma - d/2 + 1$, and V satisfies (A_s) for any $s < 2\gamma - d/2 + 3$.*

The proof is deferred to "Appendix A.1."

2.1.2 Basic Regularity Results

In the semilinear SPDE model (5), the nonlinear term F is assumed to satisfy [cf. Altmeyer et al. (2020b)]:

$(F_{s,\eta})$ There is $b > 0$ and $\epsilon > 0$ such that

$$
|(-\Delta)^{\frac{s-2+\eta+\epsilon}{2}} F_{\theta_{1:K}}(Y)|_{C(0,T;H)} \le c(\theta_{1:K})(1 + |(-\Delta)^{\frac{s}{2}} Y|_{C(0,T;H)})^b
$$

for $Y \in C(0, T; H^s)$, where c depends continuously on $\theta_{1:K}$.

In particular, if $F(Y)(t) = F(Y_t)$, this simplifies to

$$
|F_{\theta_{1:K}}(Y)|_{s-2+\eta+\epsilon} \le c(\theta_{1:K})(1 + |Y|_s)^b \tag{13}
$$

for $Y \in H^s$. In order to control the regularity of X, we apply a splitting argument (see also Cialenco and Glatt-Holtz 2011; Pasemann and Stannat 2020; Altmeyer et al. 2020b) and write $X = \overline{X} + \widetilde{X}$, where $\overline{X}$ is the solution to the linear SPDE

$$\mathrm{d}\overline{X}_t = \theta_0 \Delta \overline{X}_t \mathrm{d}t + B \mathrm{d}W_t, \quad \overline{X}_0 = 0, \tag{14}$$

where W is the same cylindrical Wiener process as in (5), and $\widetilde{X}$ solves a random PDE of the form

$$\mathrm{d}\widetilde{X} = (\theta_0 \Delta \widetilde{X}_t + F_{\theta_{1:K}}(\overline{X} + \widetilde{X})(t))\mathrm{d}t, \quad \overline{X}_0 = X_0. \tag{15}$$

Lemma 2 *The process $\overline{X}$ is Gaussian, and for any $p \geq 1$, $s < 2\gamma - d/2 + 1$:*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |\overline{X}_t|_s^p\right] < \infty. \tag{16}$$

Proof This is classical, see, e.g., Da Prato and Zabczyk (2014), Liu and Röckner (2015). □

Proposition 3

(1) Let (A_s) and $(F_{s,\eta})$ hold. Then for any $p \geq 1$:

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |\widetilde{X}_t|_{s+\eta}^p\right] < \infty. \tag{17}$$

In particular, if $s + \eta < 2\gamma - d/2 + 1$, then $(A_{s+\eta})$ is true.

(2) Let $G : C(0, T; H) \supset D(G) \to C(0, T; H)$ be any function such that $(F_{s,\eta})$ holds for G. Then for $s < 2\gamma - d/2 + 1$ and $p \geq 1$:

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |G(X)(t)|_{s+\eta-2}^p\right] < \infty. \tag{18}$$

In particular,

$$\mathbb{E}\int_0^T |G(X)(t)|_{s+\eta-2}^2 \mathrm{d}t < \infty. \tag{19}$$

Proof

(1) For $t \in [0, T]$ and $\epsilon > 0$,

$$|\widetilde{X}_t|_{s+\eta} \leq |S(t)X_0|_{s+\eta} + \int_0^t |S(t - r)F_{\theta_{1:K}}(X)(t)|_{s+\eta}\mathrm{d}r$$

$$\leq |X_0|_{s+\eta} + \int_0^t (t - r)^{-1+\epsilon/2}|F_{\theta_{1:K}}(X)(t)|_{s-2+\eta+\epsilon}\mathrm{d}r$$

$$\leq |X_0|_{s+\eta} + \frac{2}{\epsilon} T^{\frac{\epsilon}{2}} \sup_{0 \leq t \leq T} |F_{\theta_{1:K}}(X)(t)|_{s-2+\eta+\epsilon}$$

$$\leq |X_0|_{s+\eta} + \frac{2}{\epsilon} T^{\frac{\epsilon}{2}} c(\theta_{1:K})(1 + |X|_{C(0,T;H^s)})^b,$$

where $\theta_1, \ldots, \theta_K$ are the true parameters. This implies (17). If $s + \eta < 2\gamma - d/2 + 1$, then a bound as in (17) holds for $\overline{X}$ by Lemma 2, and the claim follows.

(2) This follows from

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |G(X)(t)|^p_{s+\eta-2} \right] \leq c\mathbb{E}\left[\left(1 + \sup_{0 \leq t \leq T} |X_t|_s \right)^{bp} \right] < \infty. \qquad (20)$$

$\square$

These regularity results form the basis for the asymptotic analysis of diffusivity estimation, as explained in the next section.

2.2 Statistical Inference: The General Model

The projected process $P_N X$ induces a measure $\mathbb{P}_\theta$ on $C(0, T; \mathbb{R}^N)$. Heuristically [see Liptser and Shiryayev (1977, Section 7.6.4)], we have the following representation for the density with respect to $\mathbb{P}_{\overline{\theta}}^N$ for an arbitrary reference parameter $\overline{\theta} \in \Theta$:

$$\frac{d\mathbb{P}_\theta^N}{d\mathbb{P}_{\overline{\theta}}^N}(X^N) = \exp\left(-\frac{1}{\sigma^2} \int_0^T \left\langle (\theta_0 - \overline{\theta}_0) \Delta X_t^N, (-\Delta)^{2\gamma} dX_t^N \right\rangle \right.$$

$$- \frac{1}{\sigma^2} \int_0^T \left\langle P_N (F_{\theta_{1:K}} - F_{\overline{\theta}_{1:K}})(X), (-\Delta)^{2\gamma} dX_t^N \right\rangle$$

$$+ \frac{1}{2\sigma^2} \int_0^T \left\langle (\theta_0 - \overline{\theta}_0) \Delta X_t^N + P_N (F_{\theta_{1:K}} - F_{\overline{\theta}_{1:K}})(X), \right.$$

$$\left. (-\Delta)^{2\gamma} \left[(\theta_0 + \overline{\theta}_0) \Delta X_t^N + P_N (F_{\theta_{1:K}} + F_{\overline{\theta}_{1:K}})(X) \right] \right\rangle dt \Bigg).$$

By setting the score (i.e., the gradient with respect to θ of the log likelihood) to zero, and by formally substituting the (fixed) parameter γ by a (free) parameter α, we get the following maximum likelihood equations:

$$\hat{\theta}_0^N \int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 dt = \int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_{\hat{\theta}_{1:K}^N}(X) \rangle dt$$

$$- \int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, dX_t^N \rangle,$$

$$-\hat{\theta}_0^N \int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, \partial_{\theta_i} P_N F_{\hat{\theta}_{1:K}^N}(X) \rangle dt$$

$$
= -\int_0^T \langle (-\Delta)^{2\alpha} P_N F_{\hat{\theta}_{1:K}^N}(X), \partial_{\theta_i} P_N F_{\hat{\theta}_{1:K}^N}(X) \rangle \mathrm{d}t
$$
$$
+ \int_0^T \langle (-\Delta)^{2\alpha} \partial_{\theta_i} P_N F_{\hat{\theta}_{1:K}^N}(X), \mathrm{d}X_t^N \rangle.
$$

Any solution $(\hat{\theta}_0^N, \ldots, \hat{\theta}_K^N)$ to these equations is a (joint) maximum likelihood estimator (MLE) for $(\theta_0, \ldots, \theta_K)$. W.l.o.g. we assume that the MLE is unique, otherwise fix any solution. We are interested in the asymptotic behavior of this estimator as $N \to \infty$, i.e., as more and more spatial information (for fixed $T > 0$) is available. While identifiability of $\theta_1, \ldots, \theta_K$ in finite time depends in general on additional structural assumptions on F, the diffusivity θ_0 is expected to be identifiable in finite time under mild assumptions. Indeed, the argument is similar to Cialenco and Glatt-Holtz (2011), Pasemann and Stannat (2020), but we have to take into account the dependence of $\hat{\theta}_0^N$ on the other estimators $\hat{\theta}_1^N, \ldots, \hat{\theta}_K^N$. Note that the likelihood equations give the following useful representation for $\hat{\theta}_0^N$:

$$
\hat{\theta}_0^N = \frac{-\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, \mathrm{d}X_t^N \rangle + \int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_{\hat{\theta}_{1:K}^N}(X) \rangle \mathrm{d}t}{\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t}. \tag{21}
$$

By plugging in the dynamics of X according to (5), we obtain the following decomposition:

$$
\hat{\theta}_0^N - \theta_0 = \frac{\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_{\hat{\theta}_{1:K}^N}(X) \rangle \mathrm{d}t}{\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t}
$$
$$
- \frac{\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_{\theta_{1:K}}(X) \rangle \mathrm{d}t}{\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t} - \frac{\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, B\mathrm{d}W_t^N \rangle}{\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t}. \tag{22}
$$

The right-hand side vanishes whenever for large N the denominator grows faster than the numerator in each of the three fractions. In principle, strong oscillation of the reaction parameter estimates $\hat{\theta}_{1:K}^N$ may influence the convergence rate for the first term, so in order to exclude undesirable behavior, we assume that $\hat{\theta}_{1:K}^N$ is bounded in probability[1]. This is a mild assumption which is in particular satisfied if the estimators for the reaction parameters are consistent. In Sect. 2.3, we verify this condition for the case that F depends linearly on $\theta_{1:K}$. Regarding the third term, we exploit the martingale structure of the noise in order to capture the growth in N. Different noise models may be used in (5) without changing the result, as long as the numerator grows slower than the denominator. For example, the present argument directly generalizes

---

[1] A sequence of estimators $(\hat{\theta}^N)_{N\in\mathbb{N}}$ is called bounded in probability (or tight) if $\sup_{N\in\mathbb{N}} \mathbb{P}(|\hat{\theta}^N| > M) \to 0$ as $M \to \infty$.

to noise of martingale type[2]. Now, the growth of the denominator can be quantified as follows:

Lemma 4 *Let $\alpha > \gamma - d/4 - 1/2$, let further $\eta, s_0 > 0$ such that (A_s) and $(F_{s,\eta})$ are true for $s_0 \leq s < 2\gamma + 1 - d/2$. Then*

$$\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 dt \asymp \mathbb{E} \int_0^T |(-\Delta)^{1+\alpha} \overline{X}_t^N|_H^2 dt \tag{23}$$

$$\asymp C_\alpha N^{\frac{2}{d}(2\alpha - 2\gamma + 1) + 1} \tag{24}$$

in probability, with

$$C_\alpha = \frac{T \Lambda^{2\alpha - 2\gamma + 1} d}{2\theta(4\alpha - 4\gamma + 2 + d)}. \tag{25}$$

Proof Using Proposition 3 (i), the proof is exactly as in Pasemann and Stannat (2020, Proposition 4.6). □

Theorem 5 *Assume that the likelihood equations are solvable for $N \geq N_0$, assume that $(\hat{\theta}_i^N)_{N \geq N_0}$ is bounded in probability for $i = 1, \ldots, K$. Let $\alpha > \gamma - d/4 - 1/2$ and $\eta, s_0 > 0$ such that (A_s) and $(F_{s,\eta})$ hold for any $s_0 \leq s < 2\gamma + 1 - d/2$. Then the following is true:*

(1) $\hat{\theta}_0^N$ is a consistent estimator for θ_0, i.e., $\hat{\theta}_0^N \xrightarrow{\mathbb{P}} \theta_0$.

(2) If $\eta \leq 1 + d/2$, then $N^r(\hat{\theta}_0^N - \theta_0) \xrightarrow{\mathbb{P}} 0$ for any $r < \eta/d$.

(3) If $\eta > 1 + d/2$, then

$$N^{\frac{1}{2} + \frac{1}{d}}(\hat{\theta}_0^N - \theta_0) \xrightarrow{d} \mathcal{N}(0, V), \tag{26}$$

with $V = 2\theta_0(4\alpha - 4\gamma + d + 2)^2 / (T d \Lambda^{2\alpha - 2\gamma + 1}(8\alpha - 8\gamma + d + 2))$.

Proof By means of the decomposition (22), we proceed as in Pasemann and Stannat (2020). Denote by $\hat{\theta}_0^{\text{full}, N}$ the estimator which is given by (21) if the $\hat{\theta}_1^N, \ldots, \hat{\theta}_K^N$ are substituted by the true values $\theta_1, \ldots, \theta_K$. In this case, the estimation error simplifies to

$$\hat{\theta}_0^{\text{full}, N} - \theta_0 = -c_N \frac{\int_0^T \langle (-\Delta)^{1+2\alpha-\gamma} X_t^N, dW_t^N \rangle}{\sqrt{\int_0^T |(-\Delta)^{1+2\alpha-\gamma} X_t^N|_H^2 dt}} \tag{27}$$

with

$$c_N = \frac{\sqrt{\int_0^T |(-\Delta)^{1+2\alpha-\gamma} X_t^N|_H^2 dt}}{\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 dt}.$$

---

[2] The generalization of our results to the case of multiplicative noise depends crucially on the noise model, see, e.g., Cialenco and Lototsky (2009), Cialenco (2010). This is beyond the scope of the present work.

By Lemma 4, the rescaled prefactor $c_N N^{1/2+1/d}$ converges in probability to $\sqrt{C_{2\alpha-\gamma}/C_\alpha}$. The second factor converges in distribution to a standard normal distribution $\mathcal{N}(0,1)$ by the central limit theorem for local martingales (see Liptser and Shiryayev 1989, Theorem 5.5.4 (I); Jacod and Shiryaev 2003, Theorem VIII.4.17). This proves (26) for $\hat\theta_0^{\mathrm{full},N}$. To conclude, we bound the bias term depending on $\hat\theta_1^N, \ldots, \hat\theta_K^N$ as follows, using $|P_N Y|_{s_2} \le \lambda_N^{(s_2-s_1)/2} |P_N Y|_{s_1}$ for $s_1 < s_2$: Let $\delta > 0$. Then

$$\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_{\hat\theta_{1:K}^N}(X)\rangle \mathrm{d}t$$

$$\le \left(\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t\right)^{\frac12} \left(\int_0^T |(-\Delta)^\alpha P_N F_{\hat\theta_{1:K}^N}(X)|_H^2 \mathrm{d}t\right)^{\frac12}$$

$$\lesssim N^{\frac1d(2\alpha-2\gamma+1)+\frac12} \left(\int_0^T |(-\Delta)^\alpha P_N F_{\hat\theta_{1:K}^N}(X)|_H^2 \mathrm{d}t\right)^{\frac12}$$

$$\lesssim N^{\frac2d(2\alpha-2\gamma+1)+1-\frac{\eta-\delta}{d}} \left(\int_0^T |(-\Delta)^{\gamma+\frac12-\frac d4-1+\frac{\eta-\delta}{2}} P_N F_{\hat\theta_{1:K}^N}(X)|_H^2 \mathrm{d}t\right)^{\frac12},$$

so using $(F_{2\gamma+1-d/2-\delta,\eta})$,

$$\frac{\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_{\hat\theta_{1:K}^N}(X)\rangle \mathrm{d}t}{\int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t} \lesssim c(\hat\theta_{1:K}^N) N^{-(\eta-\delta)/d}.$$

As $c(\hat\theta_{1:K}^N)$ is bounded in probability and $\delta > 0$ is arbitrarily small, the claim follows. The remaining term involving the true parameters $\theta_1, \ldots, \theta_K$ is similar. This concludes the proof. $\qquad\square$

It is clear that a Lipschitz condition on F with respect to $\theta_{1:K}$ allows to bound $\hat\theta_0^N - \hat\theta_0^{\mathrm{full},N}$ in terms of $|\hat\theta_{1:K}^N - \theta_{1:K}| N^{-(\eta-\delta)/d}$ for $\delta > 0$, using the notation from the previous proof. In this case, consistency of $\hat\theta_i^N$, $i = 1, \ldots, K$, may improve the rate of convergence of $\hat\theta_0^N$. However, as noted before, in general we cannot expect $\hat\theta_i^N$, $i = 1, \ldots, K$, to be consistent as $N \to \infty$.

2.3 Statistical Inference: The Linear Model

We put particular emphasis on the case that the nonlinearity F depends linearly on its parameters:

$$\mathrm{d}X_t = \left(\theta_0 \Delta X_t + \sum_{i=1}^K \theta_i F_i(X) + \overline{F}(X)\right)\mathrm{d}t + B\mathrm{d}W_t. \tag{28}$$

This model includes the FitzHugh–Nagumo system in the form (9). We state an additional verifiable condition, depending on the contrast parameter $\alpha \in \mathbb{R}$, which guarantees that the likelihood equations are well posed, among others.

(L_α) The terms $F_1(Y), \ldots, F_K(Y)$ are well defined as well as linearly independent in $L^2(0, T; H^{2\alpha})$ for every non-constant $Y \in C(0, T; C(\mathcal{D}))$.

In particular, condition (L_α) implies for $i = 1, \ldots, K$ that

$$\int_0^T |(-\Delta)^\alpha F_i(X)|_H^2 \mathrm{d}t > 0. \tag{29}$$

For linear SPDEs, similar considerations have been made first in Huebner (1993), Chapter 3. The maximum likelihood equations for the linear model (28) simplify to

$$A_N(X)\hat{\theta}^N(X) = b_N(X), \tag{30}$$

where

$$A_N(X)_{0,0} = \int_0^T |(-\Delta)^{1+\alpha} X_t^N|_H^2 \mathrm{d}t,$$

$$A_N(X)_{0,i} = A_N(X)_{i,0} = -\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N F_i(X) \rangle \mathrm{d}t,$$

$$A_N(X)_{i,j} = \int_0^T \langle (-\Delta)^{2\alpha} P_N F_i(X), P_N F_j(X) \rangle \mathrm{d}t$$

for $i, j = 1, \ldots, K$, and

$$b_N(X)_0 = -\int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, \mathrm{d}X_t^N \rangle$$
$$+ \int_0^T \langle (-\Delta)^{1+2\alpha} X_t^N, P_N \overline{F}(X) \rangle \mathrm{d}t,$$

$$b_N(X)_i = \int_0^T \langle (-\Delta)^{2\alpha} P_N F_i(X), \mathrm{d}X_t^N \rangle$$
$$- \int_0^T \langle (-\Delta)^{2\alpha} P_N F_i(X), P_N \overline{F}(X) \rangle \mathrm{d}t$$

for $i = 1, \ldots, K$.

In order to apply Theorem 5, we need that the estimators $\hat{\theta}_1^N, \ldots, \hat{\theta}_K^N$ are bounded in probability.

Proposition 6 *In the setting of this section, let $\eta, s_0 > 0$ such that (A_s) and $(F_{s,\eta})$ are true for $s_0 \leq s < 2\gamma + 1 - d/2$. For $\gamma - d/4 - 1/2 < \alpha \leq \gamma \wedge (\gamma - d/4 - 1/2 + \eta/2) \wedge (\gamma - d/8 - 1/4 + \eta/4)$, let (L_α) be true. Then the $\hat{\theta}_i^N$, $i = 0, \ldots, K$, are bounded in probability.*

The proof of Proposition 6 is given in "Appendix A.2." We note that the upper bound on α can be relaxed in general, depending on the exact asymptotic behavior of $A_N(X)_{i,i}, i = 1, \ldots, K$. Proposition 6 together with Theorem 5 gives conditions for $\hat{\theta}_0^N$ to be consistent and asymptotically normal in the linear model (28). In particular, we immediately get for the activator–inhibitor model (7), (8), as the linear independency condition (L_α) is trivially satisfied and η can be chosen arbitrarily close to 2:

Theorem 7 *Let $\gamma > d/4$. Then $\hat{\theta}_0^N$ has the following properties in the activator–inhibitor model* (7), (8)*:*

(1) In $d = 1$, let $\gamma - 3/4 < \alpha \leq \gamma$. Then $\hat{\theta}_0^N$ is a consistent estimator for θ_0, which is asymptotically normal as in (26)*.*

(2) In $d = 2$, let $\gamma - 1 < \alpha < \gamma$. Then $\hat{\theta}_0^N$ is a consistent estimator for θ_0 with optimal convergence rate, i.e., $N^r(\hat{\theta}_0^N - \theta_0) \xrightarrow{\mathbb{P}} 0$ for any $r < 1$.

So far, we have presented a theory of parameter estimation for stochastic reaction–diffusion models, with special emphasis on activator–inhibitor systems. In the next chapter, the context of this class of models for intracellular actin dynamics is discussed.

3 Application to Activator–Inhibitor Models of Actin Dynamics

The actin cytoskeleton is a dense polymer meshwork at the inner face of the plasma membrane that determines the shape and mechanical stability of a cell. Due to the continuous polymerization and depolymerization of the actin filaments, it displays a dynamic network structure that generates complex spatiotemporal patterns. These patterns are the basis of many essential cellular functions, such as endocytic processes, cell shape changes, and cell motility (Blanchoin et al. 2014). The dynamics of the actin cytoskeleton is controlled and guided by upstream signaling pathways, which are known to display typical features of non-equilibrium systems, such as oscillatory instabilities and the emergence of traveling wave patterns (Peter 2017; Beta and Kruse 2017). Here we use giant cells of the social amoeba *D. discoideum* that allow us to observe these cytoskeletal patterns over larger spatial domains (Gerhardt et al. 2014). Depending on the genetic background and the developmental state of the cells, different types of patterns emerge in the cell cortex. In particular, pronounced actin wave formation is observed as the consequence of a mutation in the upstream signaling pathway—a knockout of the RasG-inactivating RasGAP NF1—which is present for instance in the commonly used laboratory strain AX2 (Veltman et al. 2016). Giant cells of NF1-deficient strains thus provide a well-suited setting to study the dynamics of actin waves and their impact on cell shape and division (Flemming et al. 2020).

In Fig. 1a, b, we show a normal-sized and a giant *D. discoideum* cell in the wave-forming regime for comparison. Images were recorded by confocal laser scanning microscopy and display the distribution of mRFP-LimEΔ, a fluorescent marker for filamentous actin, in the cortex at the substrate-attached bottom membrane of the cell. As individual actin filaments are not resolved by this method, the intensity of the fluorescence signal reflects the local cortical density of filamentous actin. Rectangular
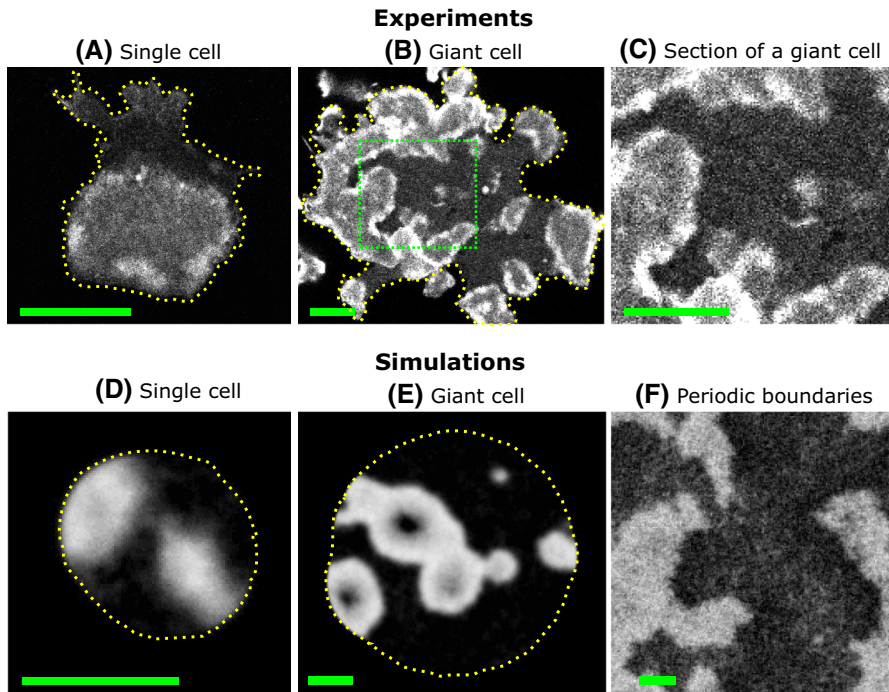
Experiments

(A) Single cell   **(B)** Giant cell   **(C)** Section of a giant cell

Simulations

(D) Single cell   **(E)** Giant cell   **(F)** Periodic boundaries

Fig. 1 Actin waves in experiments (top) and model simulations (bottom). **a** Normal-sized cell with a circular actin wave. **b** Giant cell with several fragmented actin waves. **c** Subsection of the cortical area of the giant cell shown in (**b**), indicated as dotted green rectangle. Experimental images are confocal microscopy recordings of mRFP-LimEΔ expressing *D. discoideum* AX2 cells, see Gerhardt et al. (2014). (Bottom) Simulations of the stochastic reaction–diffusion model (3), (4) in a (**d**) small and **e** large domain, defined by a dynamically evolving phase field and **f** with periodic boundary conditions. For details on the phase field simulations, see Flemming et al. (2020). (Scale bars, 10 μm) Details on the numerical implementation can be found in "Appendix B"

subsections of the inner part of the cortex of giant cells as displayed in panel (C) were used for data analysis in Sect. 4.

Many aspects of subcellular dynamical patterns have been addressed by reaction–diffusion models. While some models rely on detailed modular approaches (Beta et al. 2008; Peter 2017), others have focused on specific parts of the upstream signaling pathways, such as the phosphatidylinositol lipid signaling system (Arai et al. 2010) or Ras signaling (Fukushima et al. 2019). To describe wave patterns in the actin cortex of giant *D. discoideum* cells, the noisy FitzHugh–Nagumo type reaction–diffusion system (3), (4), combined with a dynamic phase field, has been recently proposed (Flemming et al. 2020).

In contrast to the more detailed biochemical models, the structure of this model is rather simple. Waves are generated by noisy bistable/excitable kinetics with an additional control of the total amount of activator U. This constraint dynamically regulates the amount of U around a constant level in agreement with the corresponding biological restrictions. Elevated levels of the activator represent typical cell front markers,

such as active Ras, PIP3, Arp2/3, and freshly polymerized actin that are also concentrated in the inner part of actin waves. On the other hand, markers of the cell back, such as PIP2, myosin II, and cortexillin, correspond to low values of U and are found outside the wave area (Schroth-Diez et al. 2009). Tuning of the parameter b allows to continuously shift from bistable to excitable dynamics, both of which are observed in experiments with *D. discoideum* cells. In Fig. 1d–f, the results of numerical simulations of this model displaying excitable dynamics are shown. Examples for bounded domains that correspond to normal-sized and giant cells are shown, as well as results with periodic boundary conditions that were used in the subsequent analysis.

Model parameters, such as the diffusivities, are typically chosen in an ad hoc fashion to match the speed of intracellular waves with the experimental observations. The approach introduced in Sect. 2 now allows us to estimate diffusivities from data in a more rigorous manner. On the one hand, we may test the validity of our method on in silico data of model simulations, where all parameters are predefined. On the other hand, we can apply our method to experimental data, such as the recordings of cortical actin waves displayed in Fig. 1c. This will yield an estimate of the diffusivity of the activator U, as dense areas of filamentous actin reflect increased concentrations of activatory signaling components. Note, however, that the estimated value of D_U should not be confused with the molecular diffusivity of a specific signaling molecule. It rather reflects an effective value that includes the diffusivities of many activatory species of the signaling network and is furthermore affected by the specific two-dimensional setting of the model that neither includes the kinetics of membrane attachment/detachment nor the three-dimensional cytosolic volume.

4 Diffusivity Estimation on Simulated and Real Data

In this section, we apply the methods from Sect. 2 to synthetic data obtained from a numerical simulation and to cell data stemming from experiments as described in Sect. 3. We follow the formalism from Theorem 5 and perform a Fourier decomposition on each data set. Set $\phi_k(x) = \cos(2\pi kx)$ for $k \leq 0$ and $\phi_k(x) = \sin(2\pi kx)$ for $k > 0$, then $\Phi_{k,l}(x, y) = \phi_k(x/L_1)\phi_l(y/L_2)$, $k, l \in \mathbb{Z}$, form an eigenbasis for $-\Delta$ on the rectangular domain $\mathcal{D} = [0, L_1] \times [0, L_2]$. The corresponding eigenvalues are given by $\lambda_{k,l} = 4\pi^2((k/L_1)^2 + (l/L_2)^2)$. As before, we choose an ordering $((k_N, l_N))_{N \in \mathbb{N}}$ of the eigenvalues (excluding $\lambda_{0,0} = 0$) such that $\lambda_N = \lambda_{k_N, l_N}$ is increasing.

In the sequel, we will use different versions of $\hat{\theta}_0^N$ which correspond to different model assumptions on the reaction term F, concerning both the effects included in the model and a priori knowledge on the parametrization. While all of these estimators enjoy the same asymptotic properties as $N \to \infty$, it is reasonable to expect that they exhibit huge qualitative differences for fixed $N \in \mathbb{N}$, depending on how much knowledge on the generating dynamics is incorporated. In order to describe the model nonlinearities that we presume, we use the notation F_1, F_2, F_3 as in (10), (11), (12). As a first simplification, we substitute F_1 by $\widetilde{F}_1$ given by

$$\widetilde{F}_1(U) = -U(u_0 - U) \tag{31}$$

in all estimators below. This corresponds to an approximation of the function a by an effective average value $\bar{a} > 0$. While this clearly does not match the full model, we will see that it does not pose a severe restriction as $a(U)$ tends to stabilize in the simulation. Recall the explicit representation (21) of $\hat{\theta}_0^N$. As before, K is the number of nuisance parameters appearing in the nonlinear term F. We construct the following estimators which capture qualitatively different model assumptions:

(1) The *linear estimator* $\hat{\theta}_0^{\mathrm{lin},N}$ results from presuming $K = 0$ and $F = 0$.
(2) The *polynomial* or *Schlögl estimator* $\hat{\theta}_0^{\mathrm{pol},N}$, where $K = 0$ and

$$
\begin{aligned}
F(u) &= k_1 u(u_0 - u)(u - \bar{a}u_0) \\
&= -k_1 \bar{a} u_0 u(u_0 - u) + k_1 u^2(u_0 - u) \\
&= \theta_1 \widetilde{F}_1(u) + \theta_2 F_2(u)
\end{aligned}
$$

for known constants $k_1, u_0, \bar{a} > 0, \theta_1 = k_1 u_0 \bar{a}, \theta_2 = k_1$. The corresponding SPDE (5) is called stochastic Nagumo equation or stochastic Schlögl equation and arises as the limiting case $\epsilon \to 0$ of the stochastic FitzHugh–Nagumo system.
(3) The *full* or *FitzHugh–Nagumo estimator* $\hat{\theta}_0^{\mathrm{full},N}$, where $K = 0$ and

$$
\begin{aligned}
F(u) &= k_1 u(u_0 - u)(u - \bar{a}u_0) - k_2 v \\
&= \theta_1 \widetilde{F}_1(u) + \theta_2 F_2(u) + \theta_2 F_3(u)
\end{aligned}
\tag{32}
$$

with $\theta_1 = k_1 u_0 \bar{a}$, $\theta_2 = k_1$ and $\theta_3 = k_2 \epsilon b$, where v is given by $v_t = \int_0^t e^{(t-r)(D_V \Delta - \epsilon I)} u_r \mathrm{d}r$. As before, $k_1, k_2, u_0, \bar{a}, D_v, \epsilon > 0$ are known.

Furthermore, we modify $\hat{\theta}_0^{\mathrm{full},N}$ in order to estimate different subsets of model parameters at the same time. We use the notation $\hat{\theta}_0^{i,N}$, where i is the number of simultaneously estimated parameters. More precisely, we set $\hat{\theta}_0^{1,N} = \hat{\theta}_0^{\mathrm{full},N}$, and additionally:

(1) The estimator $\hat{\theta}_0^{2,N}$ results from $K = 1$ and F_{θ_1} given by (32) for known $\theta_2, \theta_3 > 0$. This corresponds to an unknown $\bar{a}$.
(2) The estimator $\hat{\theta}_0^{3,N}$ results from $K = 2$ and F_{θ_1,θ_2} given by (32). Only θ_3 is known.
(3) The estimator $\hat{\theta}_0^{4,N}$ results from $K = 3$ and $F_{\theta_1,\theta_2,\theta_3}$ given by (32). All three parameters $\theta_1, \theta_2, \theta_3$ are unknown.

In all estimators in this section, we set the regularity adjustment $\alpha = 0$. This is a reasonable choice if the driving noise in (7), (8) is close to white noise.

It is worthwhile to note that $\hat{\theta}_0^{\mathrm{lin},N}$ is invariant under rescaling the intensity of the data, i.e., substituting X by cX, $c > 0$. This has the advantage that we do not need to know the physical units of the data. In fact, the intensity of fluorescence microscopy data may vary due to different expression levels of reporter proteins within a cell population, or fluctuations in the illumination. While invariance under intensity rescaling is a desirable property, the fact that nonlinear reaction terms are not taken into account may outweigh this advantage, especially if the SPDE model is close to the true generating process of the data. This is the case for synthetic data. The
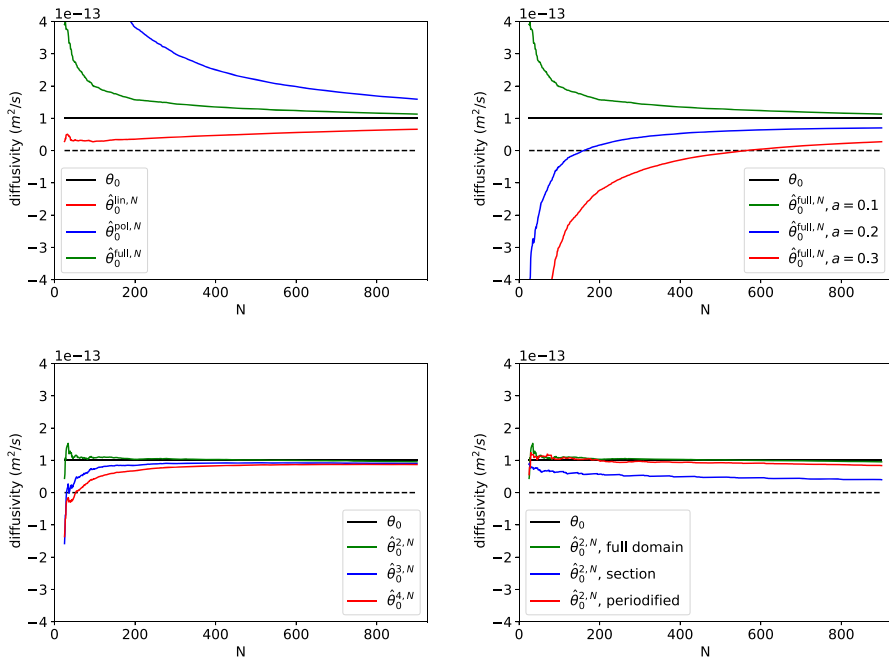
Fig. 2 Performance of diffusivity estimators on simulated data under different model assumptions in the spatial high-frequency regime. Solid black line is plotted at the true parameter $\theta_0 = 1 \times 10^{-13}$, dashed black line is plotted at zero. In all displays, we restrict to $N \geq 25$ in order to avoid artifacts stemming from low resolution

discussion in Sect. 4.1 shows that even if the model specific correction terms in (21) vanish asymptotically, their effect on the estimator may be huge in the non-asymptotic regime, especially at low resolution level. However, real data may behave differently, and a detailed nonlinear model may not reveal additional information on the underlying diffusivity, see Sect. 4.4.

4.1 Performance on Synthetic Data

First, we study the performance of the mentioned estimators on simulations. The numerical scheme is specified in "Appendix B." While we have perfect knowledge on the dynamical system which generates the data in this setting, it is revelatory to compare the different versions of $\hat{\bar{\theta}}_0^N$ which correspond to varying levels of model mis-specification. The simulation shows that $a(|U_t|_{L^2})$ fluctuates around a value slightly larger than 0.15. We demonstrate the effect of qualitatively different model assumptions on our method in Fig. 2 (top left) by comparing the performance of $\hat{\theta}_0^{\text{lin},N}$, $\hat{\theta}_0^{\text{pol},N}$, and $\hat{\theta}_0^{\text{full},N}$. The result can be interpreted as follows: As $\hat{\theta}_0^{\text{lin},N}$ does not see any information on the wave fronts, the steep gradient at the transition phase leads to a low diffusivity estimate. On the other hand, $\hat{\theta}_0^{\text{pol},N}$ incorporates knowledge on the wave fronts as they appear in the Schlögl model, but the decay in concentration due to the
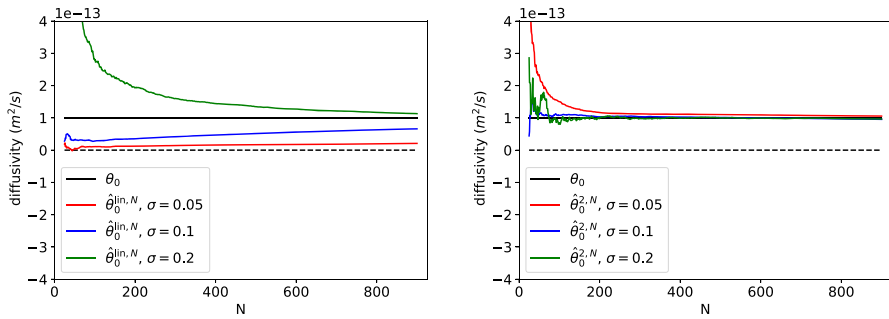
Fig. 3 Sensitivity of (left) $\hat{\theta}_0^{\mathrm{lin},N}$ and (right) $\hat{\theta}_0^{2,N}$ to different noise levels. Solid black line is plotted at $\theta_0 = 1 \times 10^{-13}$, and dashed black line is plotted at zero. As before, we restrict to $N \geq 25$ in the plots

presence of the inhibitor is mistaken as additional diffusion. Finally, $\hat{\theta}_0^{\mathrm{full},N}$ contains sufficient information on the dynamics to give a precise estimate. In Fig. 2 (top right), we show the effect of wrong a priori assumptions on $\bar{a}$ in $\hat{\theta}_0^{\mathrm{full},N}$. Even for $N = 800$, the precision of $\hat{\theta}_0^{\mathrm{full},N}$ clearly depends on the choice of $\bar{a}$. Remember that there is no true $\bar{a}$ in the underlying model, rather, $\bar{a}$ serves as an approximation for $a(|U_t|_{L^2})$. Better results can be achieved with $\hat{\theta}_0^{2,N}$, $\hat{\theta}_0^{3,N}$, and $\hat{\theta}_0^{4,N}$, see Fig. 2 (bottom left): $\hat{\theta}_0^{2,N}$ has no knowledge on $\bar{a}$ and recovers the diffusivity precisely, and even $\hat{\theta}_0^{4,N}$ performs better than the misspecified $\hat{\theta}_0^{\mathrm{full},N}$ from the top right panel of Fig. 2.

4.2 Discussion of the Periodic Boundary

In Fig. 2 (bottom right), we sketch how the assumption of periodic boundary conditions influences the estimate. While $\hat{\theta}_0^{2,N}$ works very well on the full domain of 200×200 pixels with periodic boundary conditions, it decays rapidly if we just use a square section of 75×75 pixels. In fact, the boundary conditions are not satisfied on that square section. This leads to the presence of discontinuities at the boundary. These discontinuities, if interpreted as steep gradients, lower the observed diffusivity. Hence, a first guess to improve the quality is to mirror the square section along each axis and glue the results together. In this manner, we construct a domain with 150×150 pixels, on which $\hat{\theta}_0^{2,N}$ performs well. We emphasize that, while this periodification procedure is a natural approach, its performance will depend on the specific situation, because the dynamics at the transition edges will still not obey the true underlying dynamics. Furthermore, by modifying the data set as explained, we change its resolution, and consequently, a different amount of spectral information may be included into $\hat{\theta}_0^{2,N}$ for interpretable results.

4.3 Effect of the Noise Intensity

In Fig. 3, we study the effect of varying the noise level in the simulation. We compare $\hat{\theta}_0^{\mathrm{lin},N}$, which is agnostic to the reaction model, to $\hat{\theta}_0^{2,N}$, which incorporates a detailed
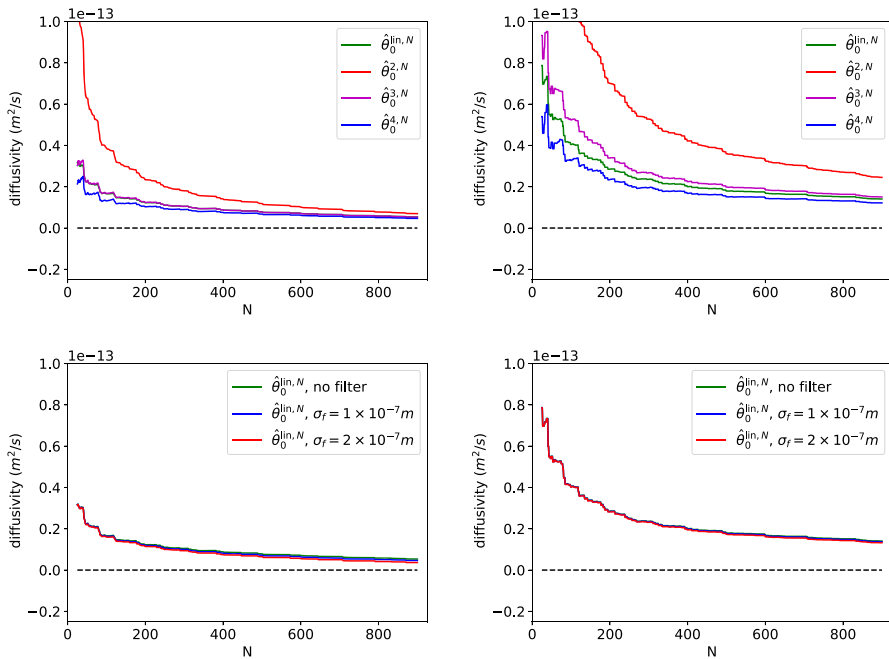
Fig. 4 In all displays, we restrict to $N \geq 25$ in order to avoid artifacts stemming from low resolution. Dashed black line is plotted at zero (top). Performance of different diffusivity estimators on (top left) cell data and (top right) periodified cell data (bottom). The effects of applying a kernel with bandwidth $\bar{\sigma}$ is shown for (bottom left) not periodified and (bottom right) periodified data

reaction model. While $\hat{\theta}_0^{2,N}$ performs well regardless of the noise level, the quality of $\hat{\theta}_0^{\text{lin},N}$ tends to improve for larger σ. In this sense, a large noise amplitude hides the effect of the nonlinearity. This is in line with the observations made in Pasemann and Stannat (2020, Section 3). We note that the dynamical features of the process change for $\sigma = 0.2$: In this case, due to the strong fluctuations stemming from the noise, the model is no longer capable of generating traveling waves.

4.4 Performance on Real Data

A description of the experimental setup can be found in "Appendix B." The concentration in the data is represented by grey values ranging from 0 to 255 at every pixel. This range is standardized to the unit interval [0, 1], in order to match the stable fixed points of the bistable polynomial f in the reference case $u_0 = 1$. Note that this is necessary for all estimators except $\hat{\theta}_0^{\text{lin},N}$. We compare $\hat{\theta}_0^{\text{lin},N}$ with $\hat{\theta}_0^{2,N}$, $\hat{\theta}_0^{3,N}$, and $\hat{\theta}_0^{4,N}$, which are more flexible than $\hat{\theta}_0^{\text{pol},N}$ and $\hat{\theta}_0^{\text{full},N}$. In Fig. 4 (top left) the behavior of these four estimators on a sample cell is shown. Interestingly, the model-free linear estimator $\hat{\theta}_0^{\text{lin},N}$ is close to $\hat{\theta}_0^{3,N}$ and $\hat{\theta}_0^{4,N}$, which impose very specific model assumptions. This pattern can be observed across different cell data sets. In particular, this is notably different from the performance of these three estimators on synthetic

data. This discrepancy seems to indicate that the lower-order reaction terms in the activator–inhibitor model are not fully consistent with the information contained in the experimental data. This can have several reasons; for example, it is possible that a more detailed model reduction of the known signaling pathway inside the cell is needed. On the contrary, $\hat{\theta}_0^{2,N}$ seems to be comparatively rigid due to its a priori choices for θ_2 and θ_3, but it eventually approaches the other estimators. Variations in the value of u_0 have an impact on the results for small N but not on the asymptotic behavior. In Fig. 4 (top right), the cell from Fig. 4 (top left) is periodified before evaluating the estimators. As expected from the discussion in Sect. 4.2, the estimates rise, but the order of magnitude does not change drastically.

4.5 Invariance under Convolution

Given a function $k \in L^1(\mathcal{D})$, define $T_k : H^s(\mathcal{D}) \to H^s(\mathcal{D})$, $s \in \mathbb{R}$, via $u \mapsto k * u = \int_{\mathcal{D}} k(\cdot - x)u(x)\mathrm{d}x$, where k and u are identified with their periodic continuation. It is well known that T_k commutes with Δ, i.e., $T_k \circ \Delta = \Delta \circ T_k$. Thus, if X is a solution to a semilinear stochastic PDE with diffusivity θ, the same is true for $T_k X$: While the nonlinearity and the dispersion operator may be changed by T_k, the diffusive part is left invariant, in particular the diffusivity of X and $T_k X$ is the same. Based on this observation, a comparison between the effective diffusivity of X and $T_k X$ for different choices of k may serve as an indicator if the assumption that a data set X is generated by a semilinear SPDE (5) is reasonable in the first place, and if the diffusion indeed can be considered to be homogeneous and isotropic. We use a family of periodic kernels $k = k_{\bar{\sigma}}$, $\bar{\sigma} > 0$, which are normed in $L^1(\mathcal{D})$ and coincide on the reference rectangle $[-L_1/2, L_1/2] \times [-L_2/2, L_2/2]$ with a Gaussian density with standard deviation $\bar{\sigma}$. In Fig. 4 (bottom), the effects of applying $T_{k_{\bar{\sigma}}}$ for different bandwidths $\bar{\sigma}$ are shown for one data set and its periodification. While the diffusivity of the data without periodification on the left-hand panel is slightly affected by the kernel, more precisely, its tendency to fall is enlarged, the graphs for the effective diffusivity of the periodified data are virtually indistinguishable. Periodification seems to be compatible with the expected invariance under convolution, even if the periodified data are not generated by a semilinear SPDE but instead by joining smaller patches of that form. In total, these observations are in accordance with the previous sections and suggest that the statistical analysis of the data based on a semilinear SPDE model is reasonable.

4.6 The Effective Diffusivity of a Cell Population

We compare the estimated diffusivity for a cell population consisting of 36 cells. The boundaries in space and time of all samples are selected in order to capture only the interior dynamics within a cell, and consequently, the data sets differ in their size. On the one hand, the estimated diffusivity tends to stabilize in time, i.e., the number of frames in a sample, corresponding to the final time T, does not affect the result much. On the other hand, the size of each frame, measured in pixels, determines the number of eigenfrequencies that carry meaningful information on the process. Thus, we expect that N can be chosen larger for samples with high spatial resolution. We formalize this
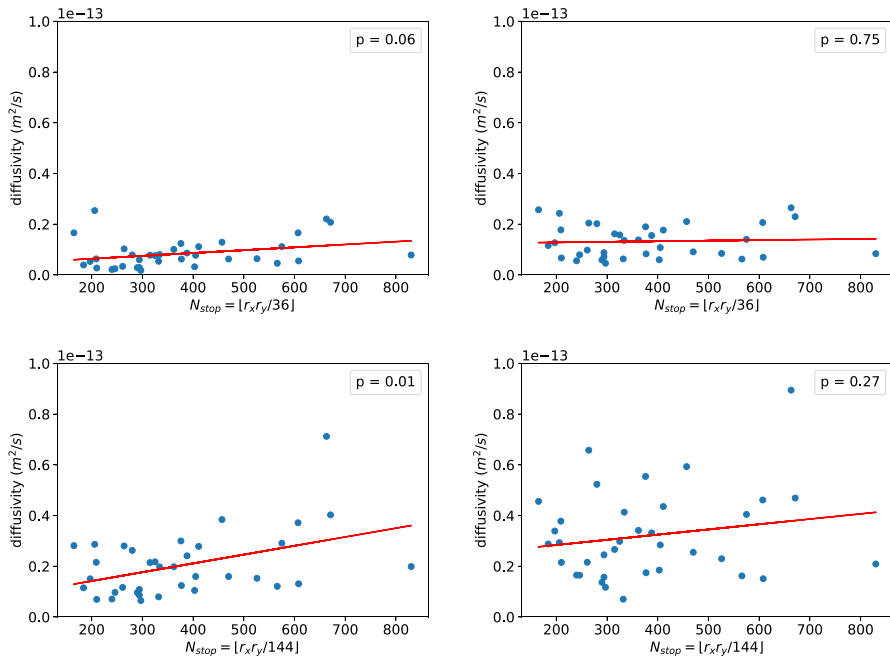
Fig. 5 The samples are evaluated (top) without or (bottom) with periodification. $\hat{\theta}_0^{3,N}$ with (left) $N = N_{\text{const}}$ or (right) $N = N_{\text{stop}}$ is plotted against N_{stop}. The least squares fit is shown in red. The p value in each plot corresponds to a t-test whose null hypothesis states that the slope of the regression line is zero. Clearly, the slope is more notable in the case $N = N_{\text{const}}$

intuition with the following heuristic: Let r_x and r_y denote the number of pixels in each row and column, resp., of every frame in a sample. Let $N_{\text{stop}} = \lfloor 4 r_x r_y / M^2 \rfloor$, where $M \in \mathbb{N}$ is a parameter representing the number of pixels needed for a sine or cosine to extract meaningful information. That is, a square of dimensions $r_x \times r_y = M \times M$ leads to $N_{\text{stop}} = 4$, so in this case only the eigenfunctions $\Phi_{k,l}, k, l \in \{-1, 1\}$ are taken into account, whose period length is M in both dimensions. In our evaluations, we choose $M = 12$ for the data without periodification and $M = 24$ for the periodified data sets. We mention that the cells are also heterogeneous with respect to their characteristic length $\overline{dx}$ and time $\overline{dt}$, given in meters per pixel and seconds per frame for each data set.[3] However, a detailed quantitative analysis of the resulting discretization error is beyond the scope of our work. In Fig. 5, we compare the results for $\hat{\theta}_0^{3,N_{\text{stop}}}$ with $\hat{\theta}_0^{3,N_{\text{const}}}$ within the cell population, where $N_{\text{const}} = 899$ is independent of the sample resolution. Evaluating the estimator at N_{stop} decorrelates the estimated diffusivity from the spatial extension of the sample. The results for different cells have the same order

---

[3] Further tests indicate that the estimated diffusivity correlates with the characteristic diffusivity $\overline{dx}^2 / \overline{dt}$. A more detailed examination of the discretization effects and other dependencies which are not directly related to the underlying process is left for future research. Also, it would be interesting to understand the extent to which the effective diffusivity in the data is scale dependent.
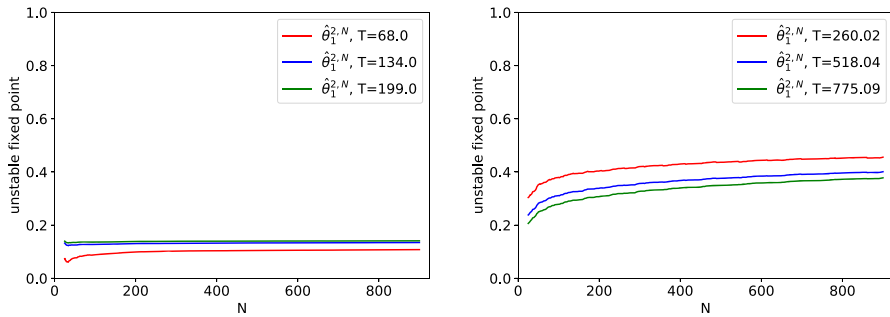
Fig. 6 Estimation of the reaction parameter θ_1 on (left) simulated data and (right) cell data. All times are given in seconds, relative to the first frame. As before, we restrict to $N \geq 25$

of magnitude, which indicates that the effective diffusivity can be used for statistical inference for cells within a population or between populations in future research.

4.7 Estimating θ_1

When solving the linear MLE equations in order to obtain $\hat{\theta}_0^{2,N}$, we simultaneously get an estimate $\hat{\theta}_1^{2,N}$ for $\theta_1 = k_1 u_0 \bar{a}$. Note that $u_0 = 1$ by convention, and $\theta_2 = k_1 = 1$ is treated as known quantity in this case. Thus, θ_1 can be identified with $\bar{a}$. In Fig. 6 we show the results for $\hat{\theta}_1^{1,N}$ on simulated data (left) and on a cell data sample (right). Note that in general, we cannot expect to observe increased precision for $\hat{\theta}_1^{2,N}$ as N grows, because the reaction term is of order zero.[4] However, it is informative to consider also the large time regime $T \to \infty$, i.e., to include more and more frames to the estimation procedure. In the case of simulated data, a oscillates around an average value close to 0.15, which should be considered to be the ground truth for $\bar{a}$. This effective value $\bar{a}$ is recovered well, even for small T, with increasing precision as T grows. Clearly, this depends heavily on the model assumptions. In the case of cell data, the results are rather stable. This indicates that it may be reasonable to use the concept of an "effective unstable fixed point $\bar{a}$ of the reaction dynamics, conditioned on the model assumptions included in $\hat{\theta}_0^{2,N}$," when evaluating cell data statistically.

4.8 The Case of Pure Noise Outside the Cell

If the data set does not contain parts of the cell but rather mere noise, the estimation procedure still returns a value. This "observed diffusivity" (see Fig. 7) originates completely from white measurement noise. More precisely, the appearance and vanishing of singular pixels is interpreted as instantaneous (i.e., within the time between two frames) diffusion to the steady state. Thus, the observed diffusivity in this case can be expected to be even larger than the diffusivity inside the cell. In this section, we give a

---

[4] According to Huebner and Rozovskii (1995), the maximum likelihood estimator for the coefficient of a linear-order zero perturbation to a heat equation with known diffusivity converges only with logarithmic rate in $d = 2$.
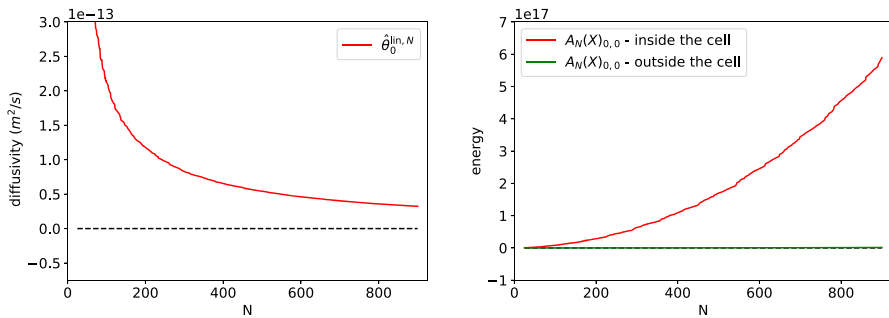
Fig. 7 As before, we restrict to $N \geq 25$. (left) Effective diffusivity outside the cell, plot for one data set. Dashed line is plotted at zero. (right) Comparison of the energy inside and outside the cell. Both data sets have the same spatial and temporal extensions

heuristical explanation for the order of magnitude of the effective diffusivity outside the cell.

We work in dimension $d = 2$. Assume that a pixel has width $x > 0$. This value is determined by the spatial resolution of the data. For simplicity, we approximate it by a Gaussian density $\phi_0(y)$ with standard deviation $\sigma_0 = \frac{x}{2}$. This way, the inflection points of the (one-dimensional marginal) density match the sharp edges of the pixel. Now, using ϕ_0 as an initial condition for the heat equation on the whole space $\mathbb{R}^2$, the density ϕ_t after time t is also a Gaussian density, with standard deviation $\sigma_t = \sqrt{\sigma_0^2 + 2\theta t}$, obtained by convolution with the heat kernel. The maximal value $f_{\max}^t$ of ϕ_t is attained at $y = 0$ with $f_{\max}^t = (2\pi\sigma_t^2)^{-1} = (2\pi(\sigma_0^2 + 2\theta t))^{-1}$. Now, if we observe after time $t > 0$ at the given pixel an intensity decay by a factor $b > 0$, i.e.,

$$bf_{\max}^t \leq f_{\max}^0, \tag{33}$$

this leads to an estimate for the diffusivity of the form

$$\theta \geq (b - 1)\frac{\sigma_0^2}{2t}. \tag{34}$$

For example, set $t = 0.97s$ and $x = 2.08 \times 10^{-7}$ m, as in the data set from Fig. 7 (left). The intensity decay factor varies between different pixels in the data set, reasonable values are given for $b \leq 30$. If $b = 30$, we get $\theta \geq 1.6 \times 10^{-13}$ m^2/s, for $b = 20$, we get $\theta \geq 1 \times 10^{-13}$ m^2/s, and for $b = 15$, we have $\theta \geq 7.8 \times 10^{-14}$ m^2/s. This matches the observed diffusivity outside the cell from Fig. 7, which is indeed of order 1×10^{-13} m^2/s: For example, with $N = N_{\text{stop}} = \lfloor 4r_x r_y/M^2 \rfloor$ and $M = 12$, as in Sect. 4.6, we get $N_{\text{stop}} = 165$ and $\hat{\theta}_0^{\text{lin}, N_{\text{stop}}} = 1.36 \times 10^{-13}$ m^2/s for this data set consisting of pure noise. In total, this gives a heuristical explanation for the larger effective diffusivity outside the cell compared to the estimated values inside the cell.

It is important to note that even if the effective diffusivity outside the cell is larger, this has almost no effect on the estimation procedure inside the cell. This is because

the total energy $A_N(X)_{0,0}$ of the noise outside the cell is several orders of magnitude smaller than the total energy of the signal inside the cell, see Fig. 7 (right).

5 Discussion and Further Research

In this paper, we have extended the mathematical theory of parameter estimation of stochastic reaction–diffusion system to the joint estimation problem of diffusivity and parametrized reaction terms within the variational theory of stochastic partial differential equations. We have in particular applied our theory to the estimation of effective diffusivity of intracellular actin cytoskeleton dynamics.

Traditionally, biochemical signaling pathways were studied in a purely temporal manner, focusing on the reaction kinetics of the individual components and the sequential order of the pathway, possibly including feedback loops. Relying on well-established biochemical methods, many of these temporal interaction networks could be characterized. However, with the recent progress in the *in vivo* expression of fluorescent probes and the development of advanced live cell imaging techniques, the research focus has increasingly shifted to studying the full spatiotemporal dynamics of signaling processes at the subcellular scale. To complement these experiments with modeling studies, stochastic reaction–diffusion systems are the natural candidate class of models that incorporate the relevant degrees of freedom of intracellular signaling processes. Many variants of this reaction–diffusion framework have been proposed in an empirical manner to account for the rich plethora of spatiotemporal signaling patterns that are observed in cells. However, the model parameters in such studies are oftentimes chosen in an ad hoc fashion and tuned based on visual inspection, so that the patterns produced in model simulations agree with the experimental observations. A rigorous framework that allows to estimate the parameters of stochastic reaction–diffusion systems from experimental data will provide an indispensable basis to refine existing models, to test how well they perform, and to eventually establish a new generation of more quantitative mathematical models of intracellular signaling patterns.

The question of robustness of the parameter estimation problem with respect to specific modeling assumptions of the underlying stochastic evolution equation is an important problem in applications that needs to be further investigated in future research. In particular, this applies to the dependence of diffusivity estimation on the domain and its boundary. In this work, we based our analysis on a Fourier decomposition on a rectangular domain with periodic boundary conditions. A natural, boundary-free approach is using local estimation techniques as they have been developed and used in Altmeyer and Reiß (2020), Altmeyer et al. (2020b), Altmeyer et al. (2020a). An additional approach aiming in the same direction is the application of a wavelet transform.

It is a crucial task to gather further information on the reaction term from the data. Principally, this cannot be achieved in a satisfactory way on a finite time horizon, so the long-time behavior of maximum likelihood-based estimators needs to be studied in the context of stochastic reaction–diffusion systems. We will address this issue in detail in future work.

To conclude, statistical inference for stochastic partial differential equations is an emerging field, which increasingly attracts the attention of mathematical research. When applied to experimental data coming from microscopy observation, it will provide a beneficial tool for the quantitative analysis of subcellular pattern formation.

Appendix A. Additional Proofs

A.1 Proof of Proposition 1

We prove Proposition 1 by a series of lemmas. First, we show that (7), (8) is well posed in $\mathcal{H} = L^2 \times L^2$. First note that for any $u, v \in L^2(\mathcal{D})$ and $x \in \mathbb{R}$:

$$\langle \partial_u f(x, u)v, v \rangle_{L^2} \lesssim |v|_{L^2}^2 \tag{35}$$

as $\partial f(x, u)$ is bounded from above uniformly in x and u.

Lemma 8 *There is a unique $\mathcal{H}$-valued solution (U, V) to (7), (8), and*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |(U_t, V_t)|_{\mathcal{H}}^p \right] < \infty \tag{36}$$

for any $p \geq 1$.

In particular,

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |U_t|_{L^2}^p \right] < \infty, \quad \mathbb{E}\left[\sup_{0 \leq t \leq T} |V_t|_{L^2}^p \right] < \infty \tag{37}$$

for any $p \geq 1$.

Proof of Lemma 8 This follows from Liu and Röckner (2015, Theorem 5.1.3). In order to apply this result, we have to test the conditions $(H1)$, $(H2')$, $(H3)$, $(H4')$ (i.e.,

hemicontinuity, local monotonicity, coercivity, and growth) therein. Set $\mathcal{V} = H^1 \times H^1$. Define $A(U, V) = (A_1(U, V), A_2(U, V))$ with

$$A_1(U, V) = D_U \Delta U + k_1 f(|U|_{L^2}, U) - k_2 V, \tag{38}$$

$$A_2(U, V) = D_V \Delta V + \epsilon(bU - V). \tag{39}$$

As $B = (-\Delta)^{-\gamma}$ is constant and of Hilbert–Schmidt type, we can neglect it in the estimates. In order to check the conditions , we have to look separately at A_1 and A_2. The statements for A_2 is trivial by linearity, so we test the parts of the conditions corresponding to A_1. $(H1)$ is clear as $f(x, u)$ is a polynomial in u and continuous in x. $(H2)$ follows from (35) via

$$
\begin{aligned}
&_{H^{-1}}\langle A_1(U_1, V_1) - A_1(U_2, V_2), U_1 - U_2 \rangle_{H^1} \\
&\quad \lesssim k_{1\,H^{-1}}\langle f(|U_1|_{L^2}, U_1) - f(|U_2|_{L^2}, U_2), U_1 - U_2 \rangle_{H^1} \\
&\qquad + k_2 |V_1 - V_2|_{L^2} |U_1 - U_2|_{L^2} \\
&\quad \lesssim k_{1\,H^{-1}}\langle \partial_u f(|U_1|_{L^2} \widetilde{V})(U_1 - U_2), U_1 - U_2 \rangle_{H^1} \\
&\qquad + k_{1\,H^{-1}}\langle \partial_x f(\tilde{x}, U_2)(|U_1|_{L^2} - |U_2|_{L^2}), U_1 - U_2 \rangle_{H^1} \\
&\qquad + C\|(U_1, V_1) - (U_2, V_2)\|_{\mathcal{H}}^2 \\
&\quad \lesssim k_1 |\partial_x f(\tilde{x}, U_2)|_{L^2} \big| \|U_1\|_{L^2} - \|U_2\|_{L^2} \big| |U_1 - U_2|_{L^2} \\
&\qquad + C\|(U_1, V_1) - (U_2, V_2)\|_{\mathcal{H}}^2 \\
&\quad \lesssim (1 + |\partial_x f(\tilde{x}, U_2)|_{L^2})\|(U_1, V_1) - (U_2, V_2)\|_{\mathcal{H}}^2
\end{aligned}
$$

for some $\tilde{x} \in \mathbb{R}$ and $\widetilde{V} : \mathcal{D} \to \mathbb{R}$, and $|\partial_x f(\tilde{x}, U_2)|_{L^2} = |u_0 a'(\tilde{x})U_2(u_0 - U_2)|_{L^2} \lesssim |U_2|_{L^2} + |U_2^2|_{L^2}$. With $|U_2^2|_{L^2} = |U_2|_{L^4}^2 \lesssim |U_2|_{H^1}^2$, we see that local monotonicity as in $(H2)$ is satisfied by taking $\rho((u, v)) = c|u|_{H^1}^2 \leq c|(u, v)|_{\mathcal{V}}^2$. Now, for $(H3)$,

$$
\begin{aligned}
&_{H^{-1}}\langle A_1(U, V), U \rangle_{H^1} \leq -D_U |U|_{H^1}^2 + k_{1\,H^{-1}}\langle f(|U|_{L^2}, U), U \rangle_{H^1} \\
&\qquad + k_2 |U|_{L^2} |V|_{L^2} \\
&\quad \lesssim -D_U |U|_{H^1}^2 + k_1 \langle \partial_u f(|U|_{L^2}, \widetilde{V})U, U \rangle_{L^2} \\
&\qquad + C\|(U, V)\|_{\mathcal{H}}^2 \\
&\quad \lesssim -D_U |U|_{H^1}^2 + C|U|_{L^2}^2 + C\|(U, V)\|_{\mathcal{H}}^2
\end{aligned}
$$

for some $\widetilde{V} : \mathcal{D} \to \mathbb{R}$, again using (35). Finally,

$$|A_1(U)|_{H^{-1}}^2 \lesssim D_U |U|_{H^1}^2 + k_1 |f(|U|_{L^2}, U)|_{H^{-1}}^2 + k_2 |V|_{L^2}^2,$$

so it remains to control $|f(|U|_{L^2}, U)|_{H^{-1}}^2 \lesssim |f(|U|_{L^2}, U)|_{L^1}^2 \lesssim |U|_{L^1}^2 + |U^2|_{L^1}^2 + |U^3|_{L^1}^2$, and we have $|U|_{L^1}^2 \lesssim |U|_{L^2}^2$ as well as $|U^2|_{L^1}^2 = |U|_{L^2}^4$ and $|U^3|_{L^1}^2 = |U|_{L^3}^6 \lesssim |U|_{H^{1/3}}^6 \lesssim |U|_{H^1}^2 |U|_{L^2}^4$. Thus, $(H4)$ is true. Putting things together, we get that Liu and Röckner (2015) is applicable for sufficiently large $p \geq 1$, and the claim follows. $\square$

In order to improve (37) to the H^1-norm, it suffices to prove coercivity, i.e., $(H3)$ from Liu and Röckner (2015), for the Gelfand triple $H^2 \subset H^1 \subset L^2$.

Lemma 9 *The solution* (U, V) *to* (7), (8) *satisfies* (A_1), *i.e., for any* $p \geq 1$:

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |U_t|^p_{H^1}\right] < \infty, \quad \mathbb{E}\left[\sup_{0 \leq t \leq T} |V_t|^p_{H^1}\right] < \infty. \tag{40}$$

Proof This is done by

$$
\begin{aligned}
{}_{L^2}\langle A_1(U, V), U\rangle_{H^2} &\leq -D_U |U|^2_{H^2} + k_1 \langle \partial_u f(|U|_{L^2}, U)\nabla U, \nabla U\rangle_{L^2} \\
&\quad + k_2 |U|_{H^1}|V|_{H^1} \\
&\lesssim -D_U |U|^2_{H^2} + C|U|^2_{H^1} + C\|(U, V)\|^2_{H^1 \times H^1},
\end{aligned}
$$

where (35) has been used componentwise. By Liu and Röckner (2015, Lemma 5.1.5) we immediately obtain (40). $\qquad\square$

Remember that by integrating (8), we can write (7) as

$$\mathrm{d}U_t = (D_U \Delta U_t + \theta_1 F_1(U_t) + \theta_2 F_2(U_t) + \theta_3 F_3(U)(t))\mathrm{d}t + B\mathrm{d}W_t, \tag{41}$$

where we adopted the notation from (10), (11), (12).

Lemma 10 *The process* (U, V) *satisfies* (A_s) *for some* $s > 1$.

Proof Let $U = \overline{U} + \widetilde{U}$ be the decomposition of U into its linear and nonlinear part as in (14),(15). We will prove

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |\widetilde{U}_t|^p_{W^{s,q}}\right] < \infty \tag{42}$$

for any $p, q \geq 1$ and $s < 2$. As $\gamma > d/4$ in $d \leq 2$, there is $s > 1$ such that (42) is true for $\overline{U}$ with $q = 2$, and the claim follows. Let us now prove (42). First, note that by Lemma 9 and the Sobolev embedding theorem in dimension $d \leq 2$, (42) is true for U_t with any $p, q \geq 1$ and $s = 0$. For $k \in \mathbb{N}$, by $|U^k_t|^p_{L^q} = |U_t|^{kp}_{L^{kq}}$, we see that polynomials in U satisfy (42) with $s = 0$, too. In particular, using that a is bounded,

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |F_i(U_t)|^p_{L^q}\right] < \infty \tag{43}$$

for $i = 1, 2$ and $p, q \geq 1$. A simple calculation shows that the same is true for F_3. As in the proof of Proposition 3, we see that

$$\sup_{0 \leq t \leq T} |\widetilde{U}_t|_{\eta, q} \leq |U_0|_{\eta, q} + \frac{2}{\epsilon} T^{\frac{\epsilon}{2}} \sup_{0 \leq t \leq T} |\theta_1 F_1(U) + \theta_2 F_2(U) + \theta_3 F_3(U)|_{L^q}$$

with $\eta < 2$, $q \geq 1$ and $\epsilon = 2 - \eta$, and the proof is complete. $\qquad\square$

Proof of Proposition 1 In $d \leq 2$, H^s is an algebra for $s > 1$, i.e., $uv \in H^s$ for $u, v \in H^s$ Adams and Fournier (2003). Together with the assumption that a is bounded, it follows immediately that $(F_{s,\eta})$ holds for F_1, F_2 separately (with $K = 0$) for any $s > 1$ and $\eta < 2$. For F_3, we have for any $\delta > 0$:

$$\sup_{0 \leq t \leq T} |F_3(U_t)|_{s+2-\delta} \leq \sup_{0 \leq t \leq T} \int_0^t |e^{(t-r)(D_V \Delta - \epsilon I)} U_r|_{s+2-\delta} \mathrm{d}r \tag{44}$$

$$\leq \sup_{0 \leq t \leq T} \int_0^t (t-r)^{-1+\delta/2} |U_r|_s \mathrm{d}r \tag{45}$$

$$\leq \frac{2}{\delta} T^{\frac{\delta}{2}} \sup_{0 \leq t \leq T} |U_t|_s, \tag{46}$$

so F_3 satisfies $(F_{s,\eta})$ even with $s \in \mathbb{R}$, $\eta < 4$. It is now clear that the $(F_{s,\eta})$ holds for $F = \theta_1 F_1 + \theta_2 F_2 + \theta_3 F_3$ for any $s > 1$ and $\eta < 2$ (with $K = 3$). The claim follows from Proposition 3 (i) for U and (ii) for V, noting that $V = \epsilon b F_3(U)$. □

A.2 Proof of Proposition 6

This section is devoted to proving Proposition 6. A similar argument can be found in Huebner (1993), Chapter 3 for linear SPDEs. We start with an auxiliary statement, which characterizes the rate of $\det(A_N(X))$. For simplicity of notation, we abbreviate $a_i^N := A_N(X)_{i,i}$ for $i = 0, \ldots, K$. While there is a trival upper bound $\det(A_N(X)) \lesssim a_0^N \ldots a_K^N$ obtained by the Cauchy–Schwarz inequality, the corresponding lower bound requires more work. Our argument is geometric in nature: A Gramian matrix measures the (squared) volume of a parallelepiped spanned by the vectors defining the matrix. The argument takes place in the separable Hilbert space $L^2(0, T; H^{2\alpha})$, and we find a (uniform in N) lower bound on the $(K+1)$-dimensional volume of the parallelepiped spanned by the vectors $P_N(-\Delta)X$, $P_N F_1(X), \ldots, P_N F_K(X)$. While the latter K components (and thus their K-dimensional volume) converge, the first component gets eventually orthogonal to the others as N grows. This is formalized as follows:

Lemma 11 *Let $\eta, s_0 > 0$ such that (A_s) and $(F_{s,\eta})$ are true for $s_0 \leq s < 2\gamma + 1 - d/2$. Let $\gamma - d/4 - 1/2 < \alpha \leq \gamma - d/4 - 1/2 + \eta/2$. Under assumption (L_α), we have*

$$\det(A_N(X)) \gtrsim a_0^N \ldots a_K^N. \tag{47}$$

In particular, $A_N(X)$ is invertible if N is large enough.

Proof The argument is pathwise, so we fix a realization of X. We abbreviate $\langle \cdot, \cdot \rangle_\alpha := \langle \cdot, \cdot \rangle_{L^2(0,T;H^{2\alpha})}$ and $\|\cdot\|_\alpha := |\cdot|_{L^2(0,T;H^{2\alpha})}$; further, we write $F_0(X) = \Delta X$ in order to unify notation. By a simple normalization procedure, it suffices to show that

$$\liminf_{N \to \infty} \det \left(\left(\left\langle \frac{P_N F_i(X)}{\|P_N F_i(X)\|_\alpha}, \frac{P_N F_j(X)}{\|P_N F_j(X)\|_\alpha} \right\rangle_\alpha \right)_{i,j=0,\ldots K} \right) > 0. \tag{48}$$

Let $\epsilon > 0$ and choose $N_0, N_1 \in \mathbb{N}$ with $N_0 \leq N_1$ such that:

- $\|P_N F_i(X) - F_i(X)\|_\alpha < \epsilon \|F_i(X)\|_\alpha$ and $\|F_i(X)\|_\alpha / \|P_N F_i(X)\|_\alpha < 2$ for $i = 1, \ldots, K$ and $N \geq N_0$. This is possible due to $\alpha \leq \gamma - d/4 - 1/2 + \eta/2$.
- $\|P_{N_0} F_0(X) / \|P_N F_0(X)\|_\alpha\|_\alpha < \epsilon$ for $N \geq N_1$. This is possible as $\|P_N F_0(X)\|_\alpha$ diverges due to $\alpha > \gamma - d/4 - 1/2$.

Now one performs a Laplace expansion of (48) in the first column. Each entry of the matrix in (48) can be bounded by 1 trivially, but we can say more:

$$
\left| \left\langle \frac{P_N F_0(X)}{\|P_N F_0(X)\|_\alpha}, \frac{P_N F_i(X)}{\|P_N F_i(X)\|_\alpha} \right\rangle_\alpha \right| \leq \left| \left\langle \frac{P_{N_0} F_0(X)}{\|P_N F_0(X)\|_\alpha}, \frac{P_{N_0} F_i(X)}{\|P_N F_i(X)\|_\alpha} \right\rangle_\alpha \right|
$$
$$
+ \left| \left\langle \frac{(P_N - P_{N_0}) F_0(X)}{\|P_N F_0(X)\|_\alpha}, \frac{(P_N - P_{N_0}) F_i(X)}{\|P_N F_i(X)\|_\alpha} \right\rangle_\alpha \right|
$$
$$
\leq \left\| \frac{P_{N_0} F_0(X)}{\|P_N F_0(X)\|_\alpha} \right\|_\alpha + \left\| \frac{(I - P_{N_0}) F_i(X)}{\|P_N F_i(X)\|_\alpha} \right\|_\alpha
$$
$$
\leq \epsilon + 2\epsilon
$$

for $N \geq N_1$ with $i = 1, \ldots, K$, where we used the Cauchy–Schwarz inequality and the choice of N_0, N_1. Consequently, in order for (48) to be true, it suffices to show for the $(0, 0)$-minor:

$$
\liminf_{N \to \infty} \det \left(\left(\left\langle \frac{P_N F_i(X)}{\|P_N F_i(X)\|_\alpha}, \frac{P_N F_j(X)}{\|P_N F_j(X)\|_\alpha} \right\rangle \right)_{i,j=1,\ldots K} \right) > 0. \tag{49}
$$

By (L_α), we have

$$
\det \left(\left(\left\langle \frac{F_i(X)}{\|F_i(X)\|_\alpha}, \frac{F_j(X)}{\|F_j(X)\|_\alpha} \right\rangle_\alpha \right)_{i,j=1,\ldots K} \right) > 0, \tag{50}
$$

so (49) holds true by continuity for $i = 1, \ldots, K$. $\qquad\square$

Proof of Proposition 6 As before, we formally write $F_0(X) = \Delta X$ in order to unify notation. By plugging in the dynamics of X, we see that for $i = 0, \ldots, K$,

$$
b_N(X)_i = \int_0^T \langle (-\Delta)^{2\alpha - \gamma} P_N F_i(X), \mathrm{d}W_t^N \rangle + \sum_{k=0}^K \theta_k A_N(X)_{i,k}. \tag{51}
$$

Thus, with $\bar{b}_N(X)_i = \int_0^T \langle (-\Delta)^{2\alpha - \gamma} P_N F_i(X), \mathrm{d}W_t^N \rangle$ and $\theta = (\theta_0, \ldots, \theta_K)^T$, this read as

$$
A_N(X)\hat{\theta}^N(X) = \bar{b}_N(X) + A_N(X)\theta, \tag{52}
$$

i.e.,

$$
\hat{\theta}^N - \theta = A_N^{-1}(X)\bar{b}_N(X). \tag{53}
$$

Using the explicit representation for the inverse matrix $A_N(X)^{-1}$, we have

$$\hat{\theta}_j^N - \theta_j = \frac{1}{\det(A_N(X))} \sum_{i=0}^{K} (-1)^{i+j} \det(\overline{A}_N^{(i,j)}(X)) \bar{b}_N(X)_i, \tag{54}$$

where $\overline{A}_N^{(i,j)}(X)$ results from $A_N(X)$ by erasing the ith row and the jth column. Remember that by Lemma 4, $a_0^N \asymp C_\alpha N^{\frac{2}{d}(2\alpha-2\gamma+1)+1}$. Further, due to $\alpha \leq \gamma - d/4 - 1/2 + \eta/2$, $a_i^N/a_0^N \to 0$ in probability for all $i = 1, \ldots, K$. Now, by the Cauchy–Schwarz inequality, each summand in the numerator can be bounded by terms of the form

$$\prod_{k=0}^{K} a_k^N \frac{|\bar{b}_N(X)_i|}{\sqrt{a_i^N a_j^N}} \tag{55}$$

for $i, j = 0, \ldots, K$. Thus, by Lemma 11, in order to prove the claim it remains to find a uniform bound in probability for the terms of the form $|\bar{b}_N(X)_i|/\sqrt{a_i^N a_j^N}$. Now, for $i = 1, \ldots, K$,

$$\mathbb{E}\left[\bar{b}_N(X)_i^2\right] = \mathbb{E}\left[\left(\int_0^T \langle(-\Delta)^{2\alpha-\gamma} P_N F_i(X), dW_t^N\rangle\right)^2\right] \tag{56}$$

$$= \mathbb{E}\left[\int_0^T |(-\Delta)^{2\alpha-\gamma} P_N F_i(X)|_H^2 dt\right] < \infty \tag{57}$$

uniformly in N as $\alpha < \gamma - d/8 - 1/4 + \eta/4$. Further, $(a_0^N)^{-1/2}$ converges to zero in probability, while $(a_i^N)^{-1/2}$ converges to a positive constant for $i = 1, \ldots, K$. For $i = 0$, the reasoning is similar: By Proposition 3 and Lemma 4, we have

$$\mathbb{E}\left[\bar{b}_N(X)_0^2\right] \asymp \mathbb{E}\left[\int_0^T |(-\Delta)^{1+2\alpha-\gamma} \overline{X}_t^N|_H^2 dt\right] \asymp C_{2\alpha-\gamma} N^{\frac{2}{d}(4\alpha-4\gamma+1)+1}. \tag{58}$$

Together with $a_0^N \sim N^{\frac{2}{d}(2\alpha-2\gamma+1)+1}$ and $\alpha \leq \gamma$, this yields that $\bar{b}_N(X)_0^2/a_0^N$ (and consequently $|\bar{b}_N(X)_0|/\sqrt{a_0^N a_j^N}$ for $j = 0, \ldots, K$) is bounded in probability. The claim follows. $\qquad\square$

Appendix B. Methods

Numerical Simulation

For the evaluation in Sect. 4 and Fig. 1f, we simulated (7), (8) on a square with side $L = 75$, with spatial increment $dx = 0.375$, for $t \in [T_0, T_1)$, $T_0 = 500$, $T_1 = 700$,

with temporal increment $dt = 0.01$, using an explicit finite difference scheme, of which we recorded every 100th frame. We choose $T_0 = 500$ in order to avoid artifacts stemming from zero initial conditions at $T = 0$. We set $b = 0.2$, $\gamma = 0$, $\sigma = 0.1$ and $a(x) = 0.5 - b + 0.5(x/(0.33u_0 L^2) - 1)$. In Sect. 4.3, we use additional simulations with $\sigma = 0.05$ and $\sigma = 0.2$.

For single and giant cell simulations in Fig. 1d, e, we solve (7), (8) on a square with side $L = 75$, with spatial increment $dx = 0.15$, for $t \in [0, T]$, $T = 1000$, with temporal increment $dt = 0.002$, using an explicit finite difference scheme together with a phase field model Flemming et al. (2020) to account for the interior of the single and giant cells, corresponding, respectively, to an area of $A_0 = 113\,\mu m^2$ and $A_0 = 2290\,\mu m^2$. We used $b = 0.4$ and $a(x) = 0.5 - b + 0.5(x/(0.25u_0 A_0) - 1)$. The noise terms are chosen as in Flemming et al. (2020), with $\sigma = 0.1$ and $k_\eta = 0.1$ in the notation therein.

For both simulations, the remaining parameters are $D_U = 0.1$, $D_V = 0.02$, $k_1 = k_2 = 1$, $u_0 = 1$, $\epsilon = 0.02$. The unit length is $1\,\mu m$, and the unit time is $1s$.

Experimental Settings

For experiments, *D. discoideum* AX2 cells expressing mRFP-LimEΔ as a marker for F-actin were used. Cell culture and electric pulse-induced fusion to create giant cells were performed as described in Gerhardt et al. (2014), Flemming et al. (2020). For live cell imaging, an LSM 780 (Zeiss, Jena) with a 63x objective lens (alpha Plan-Apochromat, NA 1.46, Oil Korr M27, Zeiss Germany) was used. The spatial and temporal resolution was adjusted for each individual experiment to acquire image series with the best possible resolution, while protecting the cells from photodamage. Images series were saved as 8-bit tiff files. For image series where the full range of 256 pixel values was not utilized, the image histograms were optimized in such a way that the brightest pixels corresponded to a pixel value of 255.

References

Altmeyer, R., Bretschneider, T., Janák, J., Reiß, M.: Parameter estimation in an SPDE model for cell repolarization. Preprint: arXiv:2010.06340 [math.ST] (2020)

Altmeyer, R., Cialenco, I., Pasemann, G.: Parameter estimation for semilinear SPDEs from local measurements. Preprint: arXiv:2004.14728 [math.ST] (2020)

Adams, R.A., Fournier, J.J.F.: Sobolev Spaces, second ed., Pure and Applied Mathematics, vol. 140. Elsevier/Academic Press, Cambridge (2003)

Annesley, Sarah J., Fisher, Paul R.: Dictyostelium discoideum—a model for many reasons. Mol. Cell. Biochem. **329**(1–2), 73–91 (2009)

Altmeyer, R., Reiß, M.: Nonparametric estimation for linear SPDEs from local measurements. Ann. Appl. Probab. (2020) (**to appear**)

Alonso, Sergio, Stange, Maike, Beta, Carsten: Modeling random crawling, membrane deformation and intracellular polarity of motile amoeboid cells. PLoS ONE **13**(8), e0201977 (2018)

Arai, Yoshiyuki, Shibata, Tatsuo, Matsuoka, Satomi, Sato, Masayuki J., Yanagida, Toshio, Ueda, Masahiro: Self-organization of the phosphatidylinositol lipids signaling system for random cell migration. Proc. Natl. Acad. Sci. **107**(27), 12399–12404 (2010)

Beta, Carsten, Amselem, Gabriel, Bodenschatz, Eberhard: A bistable mechanism for directional sensing. New J. Phys. **10**(8), 083015 (2008)

Blanchoin, Laurent, Boujemaa-Paterski, Rajaa, Sykes, Cécile, Plastino, Julie, Dynamics, Actin: Architecture, and mechanics in cell motility. Physiol. Rev. **94**(1), 235–263 (2014)

Beta, Carsten, Kruse, Karsten: Intracellular oscillations and waves. Ann. Rev. Condens. Matter Phys. **8**(1), 239–264 (2017)

Bibinger, M., Trabs, M.: On central limit theorems for power variations of the solution to the stochastic heat equation. In: Steland, A., Rafajłowicz, E., Okhrin, O. (eds.) Stochastic Models, Statistics and Their Applications. Springer Proceedings in Mathematics and Statistics, vol. 294, pp. 69–84. Springer, Berlin (2019)

Bibinger, M., Trabs, M.: Volatility estimation for stochastic PDEs using high-frequency observations. Stochast. Process. Their Appl. **130**(5), 3005–3052 (2020)

Cialenco, I., Delgado-Vences, F., Kim, H.-J.: Drift estimation for discretely sampled SPDEs, stochastics and partial differential equations: analysis and computations (2020) (**to appear**)

Cialenco, I., Glatt-Holtz, N.: Parameter estimation for the stochastically perturbed Navier–Stokes equations. Stochast. Process. Appl. **121**(4), 701–724 (2011)

Cialenco, I., Huang, Y.: A Note on parameter estimation for discretely sampled SPDEs. Stochast. Dyn. (2019). https://doi.org/10.1142/S0219493720500161 (**to appear**)

Chong, C.: High-frequency analysis of parabolic stochastic PDEs. Ann. Stat. (2019) (**to appear**)

Chong, C.: High-frequency analysis of parabolic stochastic PDEs with multiplicative noise: part I. Preprint: arXiv:1908.04145 [math.PR] (2019)

Cialenco, Igor: Parameter estimation for SPDEs with multiplicative fractional noise. Stoch. Dyn. **10**(4), 561–576 (2010)

Cialenco, I.: Statistical inference for SPDEs: an overview. Stat. Inference Stoch. Process. **21**(2), 309–329 (2018)

Cialenco, I., Kim, H.-J.: Parameter estimation for discretely sampled stochastic heat equation driven by space-only noise revised. Preprint: arXiv:2003.08920 [math.PR] (2020)

Cialenco, Igor, Lototsky, Sergey V.: Parameter estimation in diagonalizable bilinear stochastic parabolic equations. Stat. Inference Stoch. Process. **12**(3), 203–219 (2009)

Condeelis, John, Singer, Robert H., Segall, Jeffrey E.: THE GREAT ESCAPE: when cancer cells hijack the genes for chemotaxis and motility. Annu. Rev. Cell Dev. Biol. **21**(1), 695–718 (2005)

Devreotes, P.N., Bhattacharya, S., Edwards, M., Iglesias, P.A., Lampert, T., Miao, Y.: excitable signal transduction networks in directed cell migration. Ann. Rev. Cell Dev. Biol. **33**(1), 103–125 (2017)

Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics and its Applications, vol. 152, 2nd edn. Cambridge University Press, Cambridge (2014)

Ermentrout, G.B., Terman, D.H.: Mathematical Foundations of Neuroscience, Interdisciplinary Applied Mathematics, vol. 35. Springer, New York (2010)

Flemming, Sven, Font, Francesc, Alonso, Sergio, Beta, Carsten: How cortical waves drive fission of motile cells. Proc. Natl. Acad. Sci. **117**(12), 6330–6338 (2020)

Fitzhugh, R.: Impulses and physiological states in theoretical models of nerve membrane. Biophys. J. **1**, 445–466 (1961)

Fukushima, S., Matsuoka, S., Ueda, M.: Excitable dynamics of Ras triggers spontaneous symmetry breaking of PIP3 signaling in motile cells. J. Cell Sci. **132**, 1–12 (2019)

Gerhardt, Matthias, Ecke, Mary, Walz, Michael, Stengl, Andreas, Beta, Carsten, Gerisch, Günther: Actin and PIP3 waves in giant cells reveal the inherent length scale of an excited state. J. Cell Sci. **127**(20), 4507–4517 (2014)

Hübner, M., Khasminskii, R., Rozovskii, B.L.: In: Cambanis, S., Ghosh, J.K., Karandikar, R.L., Sen, P.K. (eds) Two Examples of Parameter Estimation for Stochastic Partial Differential Equations, Stochastic Processes, pp. 149–160. Springer, New York (1993)

Huebner, M., Lototsky, S., Rozovskii, B.L.: Asymptotic Properties of an Approximate Maximum Likelihood Estimator for Stochastic PDEs, Statistics and Control of Stochastic Processes (Y. M. Kabanov, B. L. Rozovskii, and A. N. Shiryaev, eds.), World Sci. Publ., (1997), pp. 139–155

Huebner, M., Rozovskii, B.L.: On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's. Probab. Theory Related Fields **103**(2), 143–163 (1995)

Huebner, M.: Parameter Estimation for Stochastic Differential Equations, ProQuest LLC, Ann Arbor (1993). Thesis (Ph.D.)–University of Southern California

Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes, Grundlehren der Mathematischen Wissenschaften, vol. 288. Springer, Berlin (2003)

Khalil, Z.M., Tudor, C.: Estimation of the drift parameter for the fractional stochastic heat equation via power variation. Mod. Stoch. Theory Appl. **6**(4), 397–417 (2019)

Kaino, Y., Uchida, M.: Parametric estimation for a parabolic linear SPDE model based on sampled data. Preprint: arXiv:1909.13557 [math.ST] (2019)

Lototsky, S.: Parameter Estimation for Stochastic Parabolic Equations: Asymptotic Properties of a Two-Dimensional Projection-Based Estimator. Stat. Inference Stoch. Process. **6**(1), 65–87 (2003)

Lototsky, S.V., Rosovskii, B.L.: Spectral asymptotics of some functionals arising in statistical inference for SPDEs. Stoch. Process. Appl. **79**(1), 69–94 (1999)

Lototsky, S., Rozovskii, B.L.: In: Korolyuk, V., Portenko, N., Syta, H. (eds) Parameter Estimation for Stochastic Evolution Equations with Non-Commuting Operators, Skorokhod's Ideas in Probability Theory, Institute of Mathematics of the National Academy of Science of Ukraine, Kiev, pp. 271–280 (2000)

Liu, W., Röckner, M.: Stochastic Partial Differential Equations: An Introduction. Universitext. Springer, New York (2015)

Liptser, R.S., Shiryayev, A.N.: Statistics of Random Processes I (General Theory), Applications of Mathematics, vol. 5. Springer, New York (1977)

Sh, R., Liptser, A., Shiryayev, N.: Theory of Martingales, Mathematics and its Applications (Soviet Series), vol. 49. Kluwer Academic Publishers Group, Dordrecht (1989)

Moreno, E., Flemming, S., Font, F., Holschneider, M., Beta, C., Alonso, S.: Modeling cell crawling strategies with a bistable model: From amoeboid to fan-shaped cell motion. Physica D Nonlinear Phenomena **412**, (2020)

Nagumo, A., Arimoto, S., Yoshizawa, S.: An active pulse transmission line simulating nerve axon. Proc. IRE **50**(10), 2061–2070 (1962)

Piterbarg, L., Rozovskii, B.: Maximum likelihood estimators in the equations of physical oceanography. In: Adler, R.J., Mĺuller, P., Rozovskii, B.L. (eds.) Stochastic Modelling in Physical Oceanography, Progress in Probability vol. 39, pp. 397–421. Birkhäuser Boston, Boston (1996)

Pasemann, G., Stannat, W.: Drift estimation for stochastic reaction–diffusion systems. Electron. J. Stat. **14**(1), 547–579 (2020)

Pospíšil, Jan, Tribe, Roger: Parameter estimates and exact variations for stochastic heat equations driven by space–time white noise. Stoch. Anal. Appl. **25**(3), 593–611 (2007)

Schroth-Diez, Britta, Gerwig, Silke, Ecke, Mary, Hegerl, Reiner, Diez, Stefan, Gerisch, Günther: Propagating waves separate two states of actin organization in living cells. HFSP J. **3**(6), 412–427 (2009)

Shubin, M.A.: Pseudodifferential Operators and Spectral Theory, 2nd edn. Springer, Berlin (2001)

Veltman, Douwe M., Williams, Thomas D., Bloomfield, Gareth, Chen, Bi-Chang, Betzig, Eric, Insall, Robert H., Kay, Robert R.: A plasma membrane template for macropinocytic cups. Elife **5**, e20085 (2016)

Weyl, H.: Über die asymptotische Verteilung der Eigenwerte. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse 110–117 (1911)