Journal of Theoretical Biology 501 (2020) 110319

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Statistical approach for parameter identification by Turing patterns

Alexey Kazarnikov [a,c,*], Heikki Haario [a,b]

[a] Department of Mathematics and Physics, LUT University, Yliopistonkatu 34, 53850 Lappeenranta, Finland
[b] Finnish Meteorological Institute, FI-00101, P.O. Box 503, Helsinki, Finland
[c] Southern Mathematical Institute of the Vladikavkaz Scientific Centre of the Russian Academy of Sciences, 362027 Vladikavkaz, Russia

ARTICLE INFO

Article history:
Received 29 January 2020
Revised 16 April 2020
Accepted 4 May 2020
Available online 13 May 2020

Keywords:
Pattern formation
Reaction-diffusion systems
Parameter identification
Model identification
Generalized correlation integral
MCMC

ABSTRACT

Prevailing theories in biological pattern formation, such as in morphogenesis or multicellular structures development, have been based on purely chemical processes, with the Turing models as the prime example. Recent studies have challenged the approach, by underlining the role of mechanical forces. A quantitative discrimination of competing theories is difficult, however, due to the elusive character of the processes: different mechanisms may result in similar patterns, while patterns obtained with a fixed model and fixed parameter values, but with small random perturbations of initial values, will significantly differ in shape, while being of the "same" type. In this sense each model parameter value corresponds to a family of patterns, rather than a fixed solution. For this situation we create a likelihood that allows a statistically sound way to distinguish the model parameters that correspond to given patterns. The method allows us to identify model parameters of reaction-diffusion systems by using Turing patterns only, i.e., the steady-state solutions of the respective equations without the use of transient data or initial values. The method is tested with three classical models of pattern formation: the FitzHugh-Nagumo model, Gierer-Meinhardt system and Brusselator reaction-diffusion system. We quantify the accuracy achieved by different amounts of training data by Bayesian sampling methods. We demonstrate how a large enough ensemble of patterns leads to detection of very small but systematic structural changes, practically impossible to distinguish with the naked eye.

1. Introduction

Reaction-diffusion systems form an important class of mathematical models capable of reproducing pattern formation (Koch and Meinhardt, 1994; Maini et al., 1997; Murray, 1993; Kondo and Miura, 2010), first introduced by Alan Turing in 1952 (Turing, 1952) as a qualitative model of biological morphogenesis. An important feature of these models is an existence of diffusion-driven instability (or Turing instability), which can generate stable patterns of morphogen concentrations from small initial perturbations of the homogeneous steady states. After a first experimental observation of Turing patterns by Castets et al. (1990) in chlorite-iodine-malonic acid (CIMA) reaction, Turing mechanism has found a wide range of applications in modeling of chemical systems (Facchini et al., 2009; Szalai and De Kepper, 2008), explaining growth and development of biological populations (Xu et al., 2015; Tang et al., 2016; Owen and Lewis, 2001; Upadhyay et al., 2014), studying the behavior of microorganism colonies (Lee et al., 1993; Vilas et al., 2012), explaining animal skin pattern formation (Kondo and Asai, 1995; Madzvamuse et al., 2002), designing neural networks (Zhao and Huang, 2014; Dong et al., 2017) and image processing (Nomura et al., 2011; Ebihara et al., 2003).

Mathematical models allow to establish a connection between a pattern observed on a macroscopic scale and a hypothetical underlying mechanism. But different mechanisms may result in similar patterns. As a topical example, in developmental biology the *de novo* formation of periodic structures similar to purely chemical Turing patterns can also be obtained due to mechano-chemical model with only one diffusing morphogen (Mercker et al., 2015; Brinkmann et al., 2018; Mercker et al., 2013). Another example concerns comparison of the classical Turing patterns (close-to-equilibrium patterns) with far-from-equilibrium patterns obtained due to hysteresis in the structure of model nonlinearities (Härting et al., 2017). These mechanisms are often not amenable to direct experimental verification. A computational approach allowing model calibration to a certain pattern will enable not only model identification based on experimental pattern data but also comparison of different models (and mechanisms) by checking how well a specific model can be fitted to the data produced by models based on another mechanism.

* Corresponding author at: Department of Mathematics and Physics, LUT University, Yliopistonkatu 34, 53850 Lappeenranta, Finland.
E-mail address: kazarnikov@gmail.com (A. Kazarnikov).

https://doi.org/10.1016/j.jtbi.2020.110319
0022-5193/

However, model parameter identification by pattern data only is challenging. Patterns obtained with fixed model parameter values but small random perturbations of the initial values will significantly differ in location and shape (Zhang and Tian, 2014; Murray, 1993; Lefèvre and Mangin, 2010), while being of the "same" type. In this sense, for unknown or randomized initial values, each model parameter corresponds to a family of patterns rather than a fixed solution. This rules out the use of standard estimation methods such as least squares. Changes of model parameters do affect pattern formation, but it is difficult to quantify exactly how much the parameters should change in order to cause a statistically significant deviation from that caused by perturbed initial values only. Also, it is non-trivial to create a cost function that would allow one to reliably quantify the model parameters that correspond to a given pattern data set. Most typically, one has to resort to tedious and subjective hand-tuning.

The aim of the current paper is to present a solution for such problems. The situation is analogous to the identification of chaotic systems: in both cases slightly different initial values lead to different solutions which, however, can be considered to belong to the same family of solutions. So we modify the recently developed statistical approach for parameter studies of chaotic ODE systems (Haario, 2015) to the non-chaotic reaction-diffusion systems studied here. The method is tested with the FitzHugh-Nagumo, Gierer-Meinhardt and Brusselator models, that exhibit the formation of Turing patterns. We demonstrate how the approach provides a cost function that enables a statistically sound identification of the model parameters by steady-state pattern data only. A remarkable feature of the presented approach is a strong sensitivity with respect to even small but systematic changes in model behaviour, practically impossible to detect with the "naked eye".

Quantitative issues, related to comparing theoretical and experimental data for the case of reaction-diffusion systems and reconstructing system dynamics from observational data are intensively studied in literature at the present time. One may recall, for example, analysis of correlation between pattern formation and theoretical models (Miura et al., 2000), numerical and experimental approaches to design and control of pattern formation (Vilas et al., 2012; Escala et al., 2015; Ghorai and Poria, 2016; Vanag and Epstein, 2008; Horváth et al., 1999; Chakravarti et al., 1995), quantitative analysis of noise impact on patterns (Barrass et al., 2006; Li, 2011) and parameter estimation for reaction-diffusion systems (Kramer and Bollt, 2013; Garvie et al., 2010; Campillo-Funollet et al., 2019). However, these approaches assume either known initial values or transient data. To the best of our knowledge, the approach discussed here is unique in that model parameters are identified by steady-state pattern data only, with unknown initial values. This is the situation often faced in experimental work.

Generalized recurrence plots are also successfully employed for identifying the dynamics of reaction-diffusion systems. For example, in Facchini et al. (2009) generalized recurrence plot concept is used to detect different regimes in the formation of spot patterns in Belousov-Zhabotinsky reaction. In Mocenni et al. (2010) the same approach is used to detect structural changes of Turing patterns in Schnakenberg model and study stability properties of spiral waves in complex Ginzburg-Landau equation. However, it should be pointed out that our focus is different: while the mentioned works aim at the detection of domains with characteristically different behaviour, our aim is to distinguish local changes within a given domain of behaviour. More exactly, we identify the distribution of parameters that produce solutions undistinguishable from those of a given data set.

The rest of the paper is organized as follows. In Results we demonstrate the approach using classical reaction-diffusion models. We first show how the model parameters can be robustly estimated by the steady-state pattern data. Then the accuracy of the approach is studied by determining the model parameter distributions by Bayesian sampling algorithms. Finally we discuss approaches to minimize the amount of data needed for successful model identification. The Discussion summarizes the potential of the approach and points out possible next steps. In Material and Methods we discuss the finite difference approximation of the equations under study, provide a detailed description of statistical approach for parameter identification and discuss technical details about numerical simulations.

2. Results

Intuitively, our motive is to develop a tool to quantify the "sameness" of the qualitatively similar patterns created by a given pattern formation process. In a more formal sense, the task is to create a statistical likelihood that quantifies the variability of a given data set of patterns. With a likelihood available, we can identify the parameters of models that try to describe the pattern formation process. We quantify the posterior distribution of the parameters, i.e., find "all" the parameter values that make the model agree with the likelihood function of the data. As always, the likelihood is conditional on data: a large data set should provide more accurate information than a limited one. While this paper demonstrates the approach by using several models with synthetic data, the likelihood function can be equally well created by real data. This provides a way to quantitatively compare the performance of different models suggested to explain various pattern formation processes.

Next, we present three reaction-diffusion models used for the examples, give a schematic overview on how the likelilhood function is constructed by pattern data, demonstrate the use of it for parameter estimation, and determine the accuracy of the estimates by MCMC sampling methods both for large and limited data sets. Further technical details are given in the section Materials and methods.

2.1. Equations under study

A general two-component reaction-diffusion system may be written in the form

$$\begin{aligned} v_t &= \nu_1 \Delta v + f(v,w), \\ w_t &= \nu_2 \Delta w + g(v,w), \end{aligned} \tag{1}$$

where $v = v(\boldsymbol{x},t)$ and $w = w(\boldsymbol{x},t)$ are unknown chemical concentrations, Δ is the Laplace operator, ν_1, $\nu_2 > 0$ are fixed diffusion coefficients, the nonlinear functions $f(v,w)$ and $g(v,w)$ represent local chemical reactions, $\boldsymbol{x} = (x_1, x_2)$ and $t > 0$.

In the present paper we consider three well-known reaction-diffusion systems: the FitzHugh-Nagumo model (FitzHugh, 1961; Nagumo et al., 1962), the Gierer-Meinhardt activator-inhibitor system (Gierer and Meinhardt, 1972) and the Brusselator reaction-diffusion system (Prigogine and Lefever, 1968). The FitzHugh-Nagumo model was developed as a two-component reduction of the Hodgkin-Huxley nerve impulse propagation model in squid giant axon and is given by the following reaction terms

$$\begin{aligned} f(v,w) &= \varepsilon(w - \alpha v), \\ g(v,w) &= -v + \mu w - w^3, \end{aligned} \tag{2}$$

where $v(\boldsymbol{x},t)$ is the recovery variable, $w(\boldsymbol{x},t)$ is the membrane potential, $\mu \in \mathbb{R}, \alpha \geqslant 0$ and $\varepsilon > 0$ are control parameters. The Gierer-Meinhardt activator-inhibitor system describes tentacle formation in Hydra freshwater polyp and is given by equations

$$\begin{aligned} f(v,w) &= -\mu_v v + \tfrac{v^2}{w}, \\ g(v,w) &= -\mu_w w + v^2, \end{aligned} \tag{3}$$

where $v(\boldsymbol{x},t)$ and $w(\boldsymbol{x},t)$ denote the activator and inhibitor concentrations and $\mu_v, \mu_w > 0$ are activator and inhibitor decay rates. Finally, the Brusselator reaction-diffusion system is a simple model describing the periodic chemical reaction between two reagents being constantly supplied into the reactor. It is given by the following kinetics

$$\begin{aligned} f(v,w) &= A - (B+1)v + v^2 w, \\ g(v,w) &= Bv - v^2 w, \end{aligned} \tag{4}$$

where $v(\boldsymbol{x},t)$ and $w(\boldsymbol{x},t)$ are reaction intermediates concentrations and A, B are constant concentrations of reagents, being supplied to the reactor.

We consider the equations under study in squared unit domain Ω supplied with homogeneous Neumann (zero-flux) boundary conditions. As initial conditions we take homogeneous steady state (v_0, w_0) of the system (1), given by following condition

$$f(v_0, w_0) = g(v_0, w_0) = 0,$$

perturbed with a small uniform random noise $U(0,\delta)$ for all $\boldsymbol{x} \in \Omega$

$$v(\boldsymbol{x},0) = v_0 + U(0,\delta),\ w(\boldsymbol{x},0) = w_0 + U(0,\delta),\ \delta < 1,$$

where $(v_0, w_0) = (0,0)$ for model (2), $(v_0, w_0) = (\frac{\mu_w}{\mu_v}, \frac{\mu_w}{\mu_v^2})$ for system (3) and $(v_0, w_0) = (A, \frac{B}{A})$ for Eqs. (4) respectively.

In the special case when $\alpha = 0$ and $\varepsilon = 1$, system (2) transforms into a model system, which could be named as the Rayleigh reaction-diffusion system. The bifurcational behavior of the Rayleigh system was studied in Kazarnikov and Revina (2016); Kazarnikov et al., 2015; Kazarnikov and Revina, 2016. The system was considered in an arbitrary bounded domain Ω, supplied with Dirichlet or mixed boundary conditions. The Liapunov-Schmidt reduction was used to obtain explicit expressions in the form of power series for the spatially-inhomogeneous auto-oscillation and stationary regimes, which branch from homogeneous zero steady state due to a Hopf bifurcation or monotonous instability.

The results of bifurcational analysis, obtained for the special case of the Rayleigh reaction-diffusion system, may be naturally generalized for the case of Turing instability. However, in the present paper we study the formation of Turing patterns far away from the instability threshold. Therefore we can not apply analytical approaches but focus on numerical techniques.

2.2. Parameter identification

We assume to have a training set of data that consists of N_{set} patterns, all qualitatively similar, as produced by the same underlying process. We quantify the variability of the patterns by subsampling. We divide the data to n_{ens} subsets, each consisting of N samples of patterns. For any given subset pair we compute the N^2 distances between all their samples, and construct the empirical cumulative distribution function (eCDF, see, e.g., van der Vaart, 1998) of the distances. It turns out that these eCDF vectors are normally distributed (see the discussion on this and the computational details in Materials and Methods). To empirically estimate the distribution, the eCDF vectors are computed between all the $n_{ens}(n_{ens}-1)/2$ pairs of the different subsets. The mean and covariance of the ensuing set of eCDF vectors can then be calculated. So we arrive at a Gaussian likelihood for the "feature vector", the eCDF of distances between patterns from two data subsets of length N (Fig. 1). We call it the correlation integral likelihood (CIL).

Let us demonstrate how the correlation integral likelihood defines an effective cost function to estimate the parameters of reaction-diffusion models producing Turing patterns. We employ the classical systems: FitzHugh-Nagumo model (2), Gierer-Meinhardt system (3) and Brusselator reaction-diffusion system (4). For simplicity, we consider only two parameters to be identified for each equation system, grouped in a vector $\boldsymbol{\theta}$. All other model parameters are fixed, see Table 1 for the values. When the parameter values of a reaction-diffusion system are fixed, different realizations of random initial conditions result in different Turing patterns in numerical simulations (Fig. 2), which, however, belong to one type of patterns. At the same time, changes in model parameters naturally affect final shape, appearance and type of patterns being observed in the system (Fig. 3).

We begin with creating synthetic data. The training set of patterns is obtained by simulating the respective reaction-diffusion systems with given control parameter vector values $\boldsymbol{\theta}_0$. In this example, we generate an ensemble of $n_{ens} = 46$ subsets, each consisting of $N = 500$ patterns. For the numerical simulations we discretize the spatial terms in the equations on a square mesh of 64×64 points and arrive at a set of 8192 equations in the finite-dimensional ODE system. See Materials and Methods for technical details about the simulations performed.

Next we omit the true value $\boldsymbol{\theta}_0$ of each case, and only use the simulated pattern data to create the respective likelihoods. The training set gives us $N_{pair} = \frac{46(46-1)}{2} = 1035$ different subset pairs. For each pair we compute the N^2 distances between the pattern samples to create the respective eCDF vector. This gives us in total 1035 sets of distances, each with 250000 values (please see Fig. 1 for general idea and Materials and Methods for a detailed discussion). Next, the bin values are found to construct the cumulative distribution functions of the distances. As always with histograms or empirical CDF functions, some hand-tuning is needed to find a suitable M, the number of bins. We denote by R_0 the maximum of all computed distances. The bin values are given by the logarithmic scale $R_k = b^{-k}R_0, k = 1, \ldots, M$. With R_0 and M given, the value of b can be found by the requirement that the smallest radius $R_M = b^{-M}R_0$ is somewhat larger than the minimum of all the distances computed by the training set. With the bin values thus fixed, we get all the 1035 eCDF vectors, and can compute the mean and covariance of them. We note that a final check for the R_k values is to verify that the resulting covariance matrix is not singular, which might happen if some R_k values are too large or too small. All parameter values for the current experiment are listed in Table 1.

With the feature vectors, and their mean and covariance thus calculated, we have the likelihood $N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. We define the cost function as the negative log-likelihood function

$$f(\boldsymbol{\theta}) = (\boldsymbol{\mu}_0 - \boldsymbol{y}(\boldsymbol{\theta}))\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{y}(\boldsymbol{\theta})), \tag{5}$$

Next we demonstrate how minimization of this cost function with respect to $\boldsymbol{\theta}$, even starting far away from $\boldsymbol{\theta}_0$, converges close to the true parameter values. Note that our cost function is stochastic, due to random initial values used in each integration. For the minimization task we employ the method of Differential Evolution (DE), see Storn and Price (1997) and Materials and Method. We run the algorithm for 100 iterations, which is enough to achieve convergence of the population close to the "true" value $\boldsymbol{\theta}_0$. Fig. 4 demonstrates the parameter values for a few given generations. We see how the populations of 60 elements each converge from a crude initial guess into the vicinity of the correct parameter values in each case. The final parameter estimate can be taken as the mean of the elements of the last population.

Similar results were observed using different sets of pattern data, but we omit further details here. We can conclude that the approach provides a robust algorithm to identify the model parameters that correspond to a given data set of patterns, without the need of transient data or knowledge of initial values of the respective differential equations.

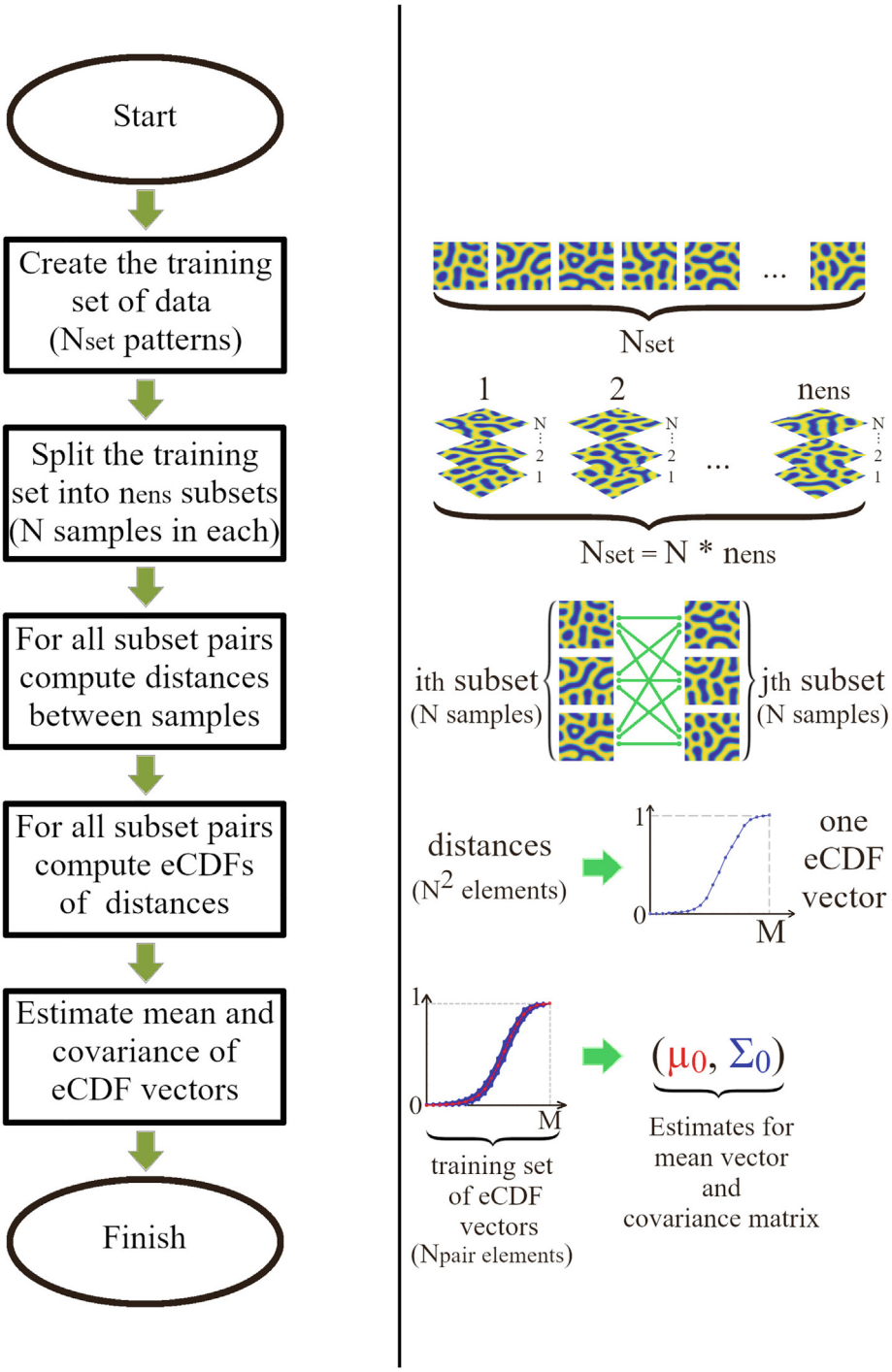

Fig. 1. Creation of the correlation integral likelihood. We first divide the training set of N_{set} patterns to n_{ens} subsets, each consisting of N samples. Next, for any given subset pair we compute N^2 distances between all their samples and construct the empirical cumulative distribution function (eCDF). This gives us $Npair = n_{ens}(n_{ens}-1)/2$ samples of eCDF vectors, from which the mean and covariance of the eCDF vector are estimated.

2.3. Parameter posteriors

After the optimization we can continue by Bayesian sampling methods to find "all" the parameter values that agree with the given data patterns. That is, we find the posterior distribution of the parameters $\boldsymbol{\theta}$ that produce feature vectors $\boldsymbol{y}(\boldsymbol{\theta})$ belonging to the distribution $N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. For the definition of $\boldsymbol{y}(\boldsymbol{\theta})$ see the Eq. (10) in Materials and Methods and related discussion there. In each case, the likelihood and the sampled parameters are the same as those used in the above optimization step, see Table 1.

As pointed out in Materials and Methods, the statistical distribution of the feature vectors (10) can theoretically shown to be Gaussian. To numerically verify the hypothesis, we apply the chi-square criterion, which states that if the Gaussian hypothesis is true then

$$(\boldsymbol{\mu}_0 - \boldsymbol{y}(\boldsymbol{\theta}_0))\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_0 - \boldsymbol{y}(\boldsymbol{\theta}_0)) \sim \chi^2_M, \tag{6}$$

where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are mean vector and covariance matrix of the training set, χ^2_M is a chi-squared distribution with M degrees of freedom. Examples of the density function of χ^2_M and the corresponding

Table 1
Model parameters (estimated and fixed), constants for creating the CIL likelihood, constants for DE optimization and resulting estimated parameter values.

Parameter	FHN	GM	BRS
$\boldsymbol{\theta}$	(μ, ε)	(μ_v, μ_w)	(A, B)
$\boldsymbol{\theta}_0$	$(1, 10)$	$(0.5, 1)$	$(4.5, 6.96)$
	$\nu_1 = 0.05$, $\nu_2 = 0.00028$, $\alpha = 1$	$\nu_1 = 0.00025$, $\nu_2 = 0.01$	$\nu_1 = 0.0016$ $\nu_2 = 0.0132$
R_0	1.23	6.3	1.9
M	18	18	18
b	1.031	1.050	1.040
N	500	500	500
n_{ens}	46	46	46
$\bar{\boldsymbol{\theta}}_0$	$(1.0003, 10.051)$	$(0.503, 1.008)$	$(4.514, 6.984)$

Here FHN, GM and BRS denote the FitzHugh-Nagumo model, Gierer-Meinhardt system and Brusselator reaction-diffusion system respectively.

empirical histogram of (6) are shown in Fig. 5. Individual components y_m of the vectors $\boldsymbol{y}(\boldsymbol{\theta})$ are also normally distributed, which can be verified by any scalar Normality test.

The parameter posterior distribution can now be constructed by Markov Chain Monte Carlo (MCMC) sampling methods. We generate a "chain" of parameter values $\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \ldots, \boldsymbol{\theta}^n$ whose empirical distribution approaches the posterior distribution. A new point $\boldsymbol{\theta}^{n+1}$ in chain is generated by the following rule. First, a new candidate $\boldsymbol{\theta}^*$ is generated by a Gaussian proposal distribution with a prescribed covariance matrix. The reaction-diffusion system is simulated N times for the proposed parameter vector $\boldsymbol{\theta}^*$, and a CIL vector $\boldsymbol{y}^{k,\theta^*}$ is calculated using distances between the simulated patterns and those from a randomly chosen subset $\boldsymbol{s}^k$ from the training set. The proposed parameter $\boldsymbol{\theta}^*$ is finally accepted or rejected, depending on the value of the likelihood for $\boldsymbol{y}^{k,\theta^*}$. For more details on MCMC methods see Robert and Casella (2004). We use the adaptive Metropolis-Hastings algorithm which updates the covariance matrix in order to improve the proposal during the sampling (Haario et al., 2001; Haario et al., 2006). The first member $\boldsymbol{\theta}^1$ of the chain is given as the result of the optimization step. Moreover, the parameter values of the populations on final iteration steps of DE can be used to create an initial proposal covariance.

We sample a parameter chain of length 4000. The results shown in Fig. 6 demonstrate that the sampling performs well, producing a strong correlation but clearly limited variability of the parameter values. Fig. 6 also shows examples of Turing patterns, obtained for the verification values from inside the parameter posterior and slightly outside it. We can observe that while the algorithm is able to detect structural changes in the patterns due to slightly changed parameter values, it would be practically impossible to distinguish any systematic differences of those pictures with the naked eye.

2.4. Limited data

We have to note that our previous training set contains a high number of Turing pattern examples, of the order 23.000. Such an amount of data allows to keep the number of patterns for one feature vector, as well as the number of pairs of the vectors, high enough to produce smooth eCDF curves and accurate estimates of $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$. It is, however, most likely an "overkill" for many practical purposes. Let us next consider situations where we possess a limited amount of information. First we assume that we have a limited set $S = \{\boldsymbol{s}_i\}_{i=1}^{N_{set}}$ of patterns available. In this situation it is possible to use a bootstrapping method that allows us to keep the number of trajectory pairs in training set high enough by resampling. However, posterior distributions of model parameters naturally grow larger as the amount of data N decreases. Next we consider the situation when data patterns are scaled by min-max normalization. This may be considered as switching to the 2D "observations" of patterns,

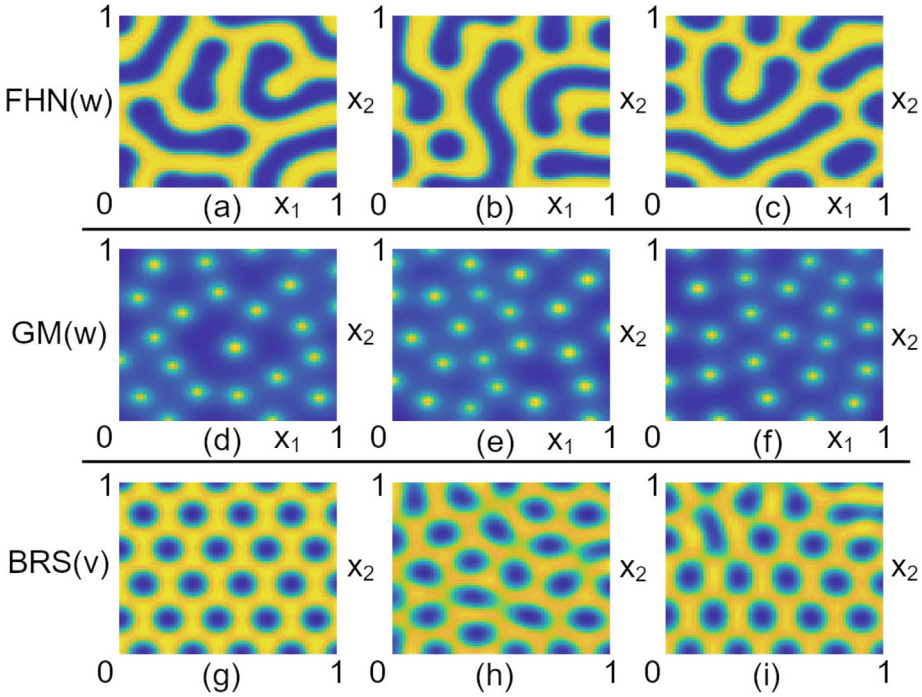

Fig. 2. Turing patterns obtained from direct simulations with fixed parameter values and different initial conditions, taken as small random perturbations of the homogeneous steady state. (a)-(c) labyrinthine-type patterns (variable w) in the FitzHugh-Nagumo model ($\nu_1 = 0.05, \nu_2 = 0.00028, \alpha = 1, \varepsilon = 10, \mu = 1$), (d)-(f) isolated spot peaks (variable w) in the Gierer-Meinhardt system ($\nu_1 = 0.00025, \nu_2 = 0.01, \mu_v = 0.5, \mu_w = 1$) and (g)-(i) hexagons (variable v) in the Brusselator reaction diffusion-system ($\nu_1 = 0.0016, \nu_2 = 0.0131, A = 4.5, B = 6.96$). Here FHN, GM and BRS denote the FitzHugh-Nagumo model, Gierer-Meinhardt system and Brusselator reaction-diffusion system respectively.

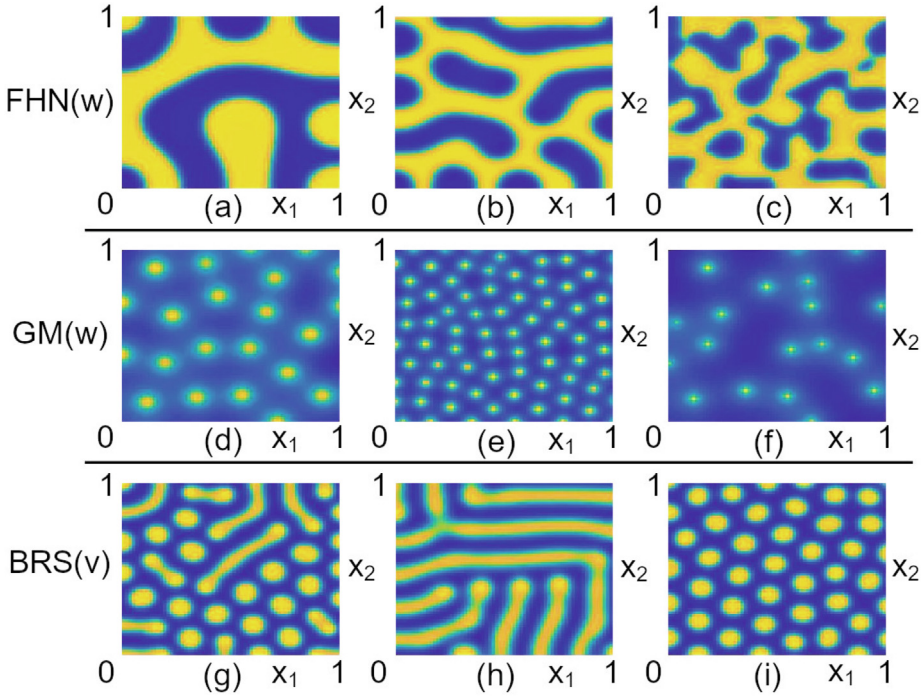

Fig. 3. Turing patterns obtained from direct simulations with different parameter values and fixed initial conditions. Diffusion coefficients and plotted system variables are the same as in Fig. 2. The FitzHugh-Nagumo model: (a) $\alpha = 1, \varepsilon = 1, \mu = 1$, (b) $\alpha = 1, \varepsilon = 5.5, \mu = 1$, (c) $\alpha = 1, \varepsilon = 6.45, \mu = 1.5$. The Gierer-Meinhardt system: (d) $\mu_v = 0.35, \mu_w = 1.0$, (e) $\mu_v = 0.75, \mu_w = 4.0$, (f) $\mu_v = 1.0, \mu_w = 0.75$. The Brusselator reaction-diffusion system: (g) $A = 5, B = 13$, (h) $A = 4.5, B = 8.72$, (i) $A = 4.5, B = 13.29$. Here FHN, GM and BRS denote the FitzHugh-Nagumo model, Gierer-Meinhardt system and Brusselator reaction-diffusion system respectively.

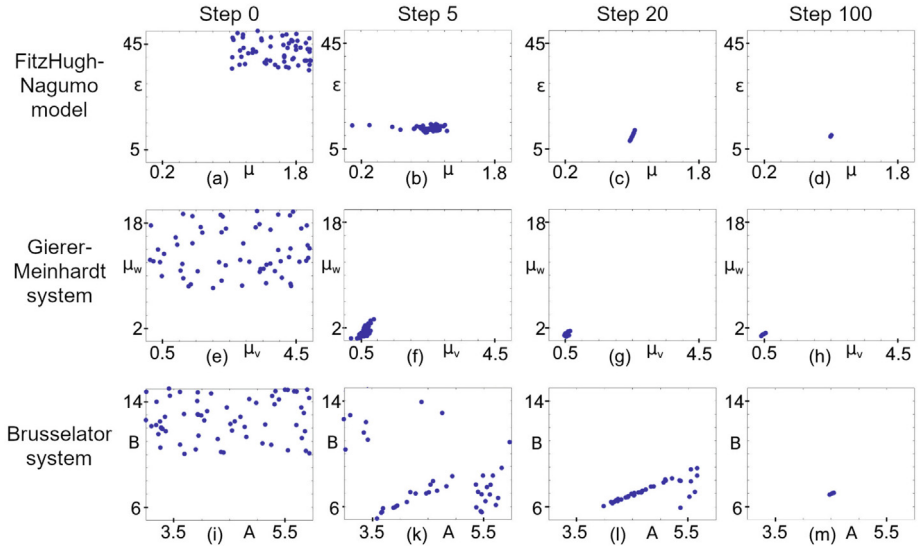

Fig. 4. Parameter identification of the reaction-diffusion systems. The optimization of stochastic cost function $f(\boldsymbol{\theta})$ by the Differential Evolution algorithm. The FitzHugh-Nagumo model: (a) step 0 (initial population), (b) step 5, (c) step 20, (d) step 100 (final population). The Gierer-Meinhardt system: (e) step 0 (initial population), (f) step 5, (g) step 20, (h) step 100 (final population). The Brusselator reaction-diffusion system: (i) step 0 (initial population), (k) step 5, (l) step 20, (m) step 100 (final population). True parameter values $\boldsymbol{\theta}_0$ are given in Table 1.

which could be treated as grayscale images of the data. In this situation the approach performs well, however the posterior regions become larger as some amount of information has been lost. Finally we test the method in the situation of a sparse mesh when the spatial resolution of the picture decreases.

We split the set S into two equal parts S_1 and S_2, each containing $N = \frac{N_{set}}{2}$ elements. Then, we create the training set for statistics of $\boldsymbol{y}(\boldsymbol{\theta}_0)$ by the following algorithm.

1. Create a set $\tilde{S_1}$ by sampling N patterns with replacement from data set S_1.
2. Create a set $\tilde{S_2}$ by sampling N patterns with replacement from data set S_2.
3. Use the sets $\tilde{S_1}$ and $\tilde{S_2}$ to compute the correlation integral vector $\boldsymbol{y}(\tilde{S_1}, \tilde{S_2}, \boldsymbol{\theta}_0)$.
4. Randomizing $\tilde{S_1}$ and $\tilde{S_2}$, repeat steps 1–3 until N_{pair} vectors $\boldsymbol{y}(\boldsymbol{\theta}_0)$ for the training set have been created.

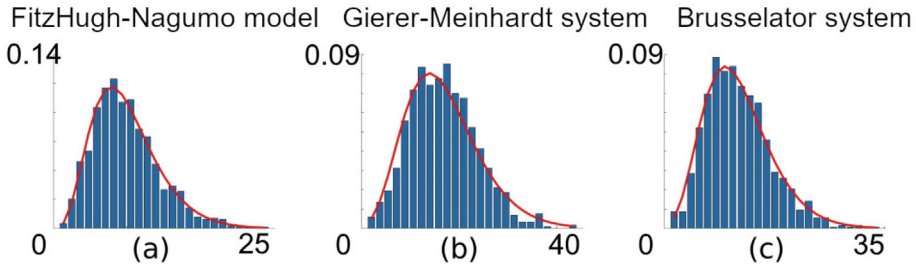

Fig. 5. Normality test for the set of CIL vectors. Density function χ^2_M with the corresponding empirical histograms: (a) the FitzHugh-Nagumo model, (b) the Gierer-Meinhardt system and (c) the Brusselator reaction-diffusion system.

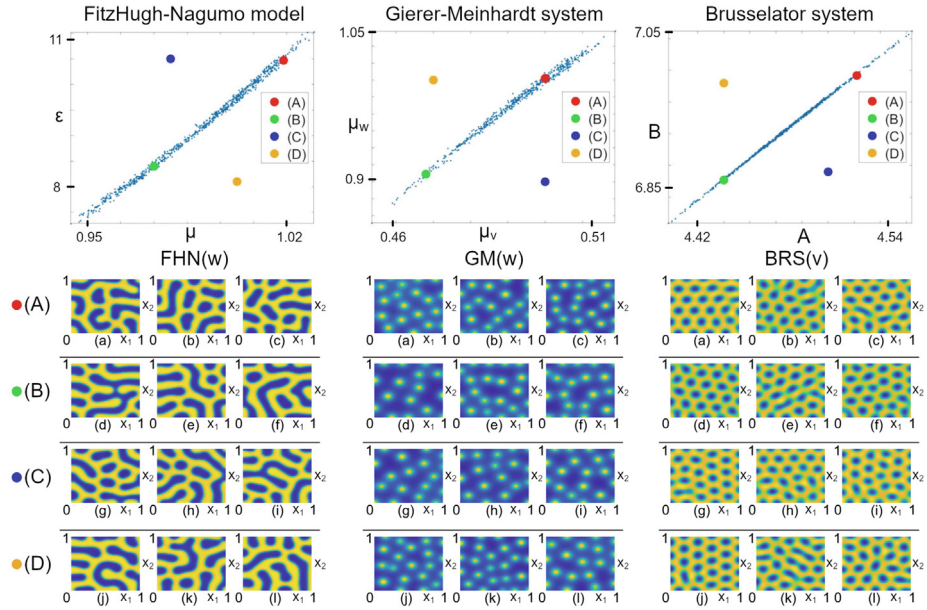

Fig. 6. Posterior distributions of model parameters and Turing patterns obtained for the verification values. Posterior distributions of model parameters with verification values inside and outside the region: (left) the FitzHugh-Nagumo model (verification values are: (A) $\mu = 1.01, \varepsilon = 10.47$, (B) $\mu = 0.97, \varepsilon = 8.4$, (C) $\mu = 0.98, \varepsilon = 10.5$ and (D) $\mu = 1.0, \varepsilon = 8.1$), (center) the Gierer-Meinhardt system (verification values are: (A) $\mu_v = 0.5, \mu_w = 1.002$, (B) $\mu_v = 0.47, \mu_w = 0.89$, (C) $\mu_v = 0.5, \mu_w = 0.88$ and (D) $\mu_v = 0.47, \mu_w = 1.0$) and (right) the Brusselator reaction-diffusion system (verification values are: (A) $A = 4.522, B = 6.992$, (B) $A = 4.42, B = 6.827$, (C) $A = 4.5, B = 6.84$ and (D) $A = 4.44, B = 7.0$). Turing patterns, obtained from direct simulations of the equations under study for verification values of control parameters, lying inside the posterior distribution region (points (A) and (B)) and outside it (points (C) and (D)). Initial conditions are taken as small random perturbations of the homogeneous steady state. The values of fixed model parameters are listed in Table 1. Left column: the FitzHugh-Nagumo model (variable w). Central column: the Gierer-Meinhardt system (variable w). Right column: the Brusselator reaction-diffusion system (variable v).

The vectors $\boldsymbol{y}(\boldsymbol{\theta}_0)$ constructed in the above way again follow a multivariate Gaussian distribution, as can be once more verified by the chi-square test. The bootstrapping is able produce estimates for mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$, which are needed for the construction of parameter posterior distribution. We tested the impact of decreasing data using training sets containing 1000, 500 and 50 patterns. The results of the MCMC sampling for the model parameters are shown in Fig. 7. We see how the decrease of original data samples results in larger parameter posterior distributions. But this is only natural: decreasing amount of data always results in increasing uncertainty. However, we note that the value $N_{set} = 50$ is at the lower limit where the bootstrapping approach still works in a numerically stable manner, because a smaller set of distances gives too crude approximations for the eCDF vectors (see Fig. 8).

Finally, we can again visually verify the patterns created by the model using parameter values at the tails or outside the sampled distribution. As the verification points we select a few examples as shown in Fig. 9, using the least accurate case obtained with $N_{set} = 50$ data patterns and show the examples of Turing patterns obtained for these verification values. We might observe that the patterns corresponding to the parameters at the tails of the posterior are yet difficult to distinguish from the reference patterns with the naked eye, while the patterns created using the parameter values outside the distribution now are clearly different.

Recall that the Turing patterns used as data so far are the steady state solutions of the reaction-diffusion equations. But while the solutions are 3D concentration fields in the domain Ω, the visual pattern information actually is only 2D, in the sense that the concentrations are replaced by grayscale images, see (Nomura et al., 2011; Ebihara et al., 2003). We can take this into account by removing the absolute concentration values from the data by normalization. We scale all samples s_i by min-max values, i.e. by the transformation

$$s_i = (v_i(x), w_i(x)) \to \tilde{s}_i = (\tilde{v}_i(x), \tilde{w}_i(x)),$$
$$\tilde{v}_i(x) = \frac{v_i(x) - v_i^{min}}{v_i^{max} - v_i^{min}}, \; \tilde{w}_i(x) = \frac{w_i(x) - w_i^{min}}{w_i^{max} - w_i^{min}},$$

where v_i^{min}, w_i^{min} and v_i^{max}, w_i^{max} are minimum and maximum values of corresponding functions on a squared domain Ω.

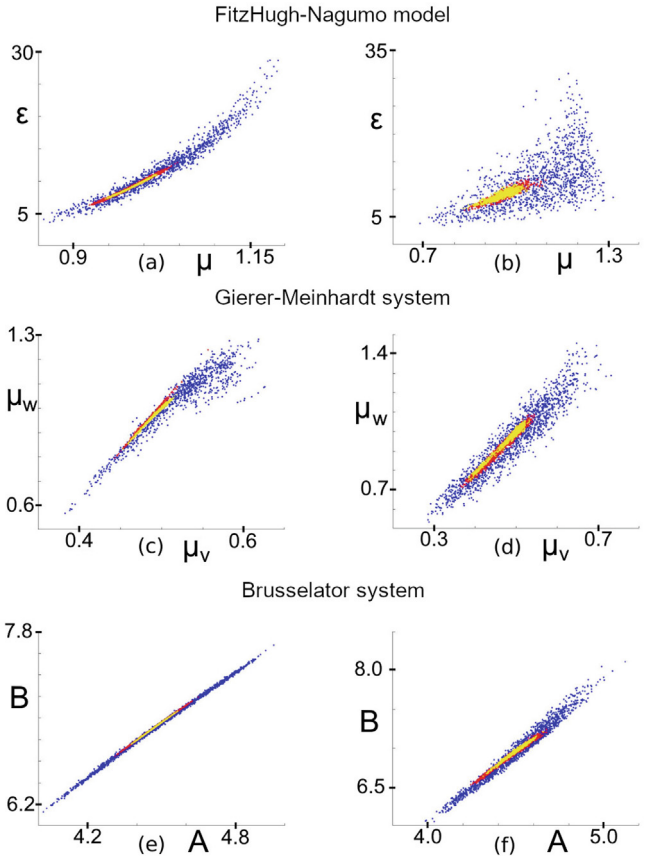

Fig. 7. Posterior distribution of model parameters for the different number of samples in the training set ($N_{set} = 1000$ (yellow), $N_{set} = 500$ (red) and $N_{set} = 50$ (blue)). Left column: non-normalized case. Right column: min-max normalized case. Values of all model parameters are given in Table 1 for non-normalized data and in Table 2 for min-max normalized data. The FitzHugh-Nagumo model: (a) and (b), the Gierer-Meinhardt system: (c) and (d) and the Brusselator reaction-diffusion system: (e) and (f).

We can create the likelihoods for the normalized data in the very same way as earlier, and successfully test the normality of the correlation integral vectors, again by the chi-square criterion. All parameter values used for the construction of the likelihood for the normalized case are listed in Table 2.

The sampled parameter distributions for both normalized and non-normalized cases are shown in Fig. 7. It can be seen that the posterior distribution region is larger in the former case, naturally since the data there is less informative. However, if the number of data samples is large enough, the variability of the model parameters remains very low and the overall conclusion remains: with a large enough training set the algorithm finds small systematic changes of pattern formation, far below what one intuitively can see by naked eye, even if only 2D information of the patterns is employed.

We also tested the robustness of proposed approach in the situation with larger spatial step h. We repeated all the experiments from this and previous sections in a square mesh of 32×32 points (and spatial step size $h = 1/31 = 0.0322$). Considering less dense meshes may be inappropriate due to the size of spatial step with respect to the domain size. However, we have found that in this case the statistical approach worked in numerically stable manner, reproducing similar results and conclusions for the spatial mesh of 64×64 points. Posterior distribution of model parameters for the different values of N_{set} and Brusselator reaction-diffusion system are shown in Fig. 10.

It is naturally possible to sample other model parameters as well, such as the diffusion coefficients v_1, v_2. The approach works as expected, however a higher number of sampled parameters results in longer MCMC chains. This, however, is left for future consideration.

Finally we consider the least accurate case of $N_{set} = 50$ Turing patterns together with min-max normalization and apply the CIL approach to estimate the values of control parameters from the data. We use the same likelihood construction as described in Parameter identification section, define the negative log-likelihood as a stochastic cost function and minimize it with respect to $\boldsymbol{\theta}$, starting with initial parameter values far away from $\boldsymbol{\theta}_0$. All parameters for the current experiment are given in Table 3. We can conclude that the parameter identification performs well again.

As noted above, the number of $N_{set} = 50$ patterns was observed to be a lower limit for reliable results. While the number n_{ens} of disjoint subsets, of size $N = N_{set}/2$ that can be drawn from the training set, is not a limiting factor, a too small value of N^2 leads to too noisy values for the calculated eCDF vectors. A way around the limitation is to employ a different strategy, the so-called synthetic likelihood (Price et al., 2018; Wood, 2010). In this approach, the likelihood is computationally re-created for every new parameter value, and the data tested against the likelihood. In our situation, even the data of one pattern could be used. Naturally, the results remain less accurate with fewer data. Moreover, the re-creation of the likelihood at every estimation step considerably increases the computational cost. If enough parallel processing power is available, the simulations can be performed, but nevertheless they need more resources than the off-line likelihood construction employed here.

3. Discussion

In this paper we study the problem of detecting local changes of stationary Turing patterns obtained as solutions of reaction-diffusion equations with randomized initial conditions. With a

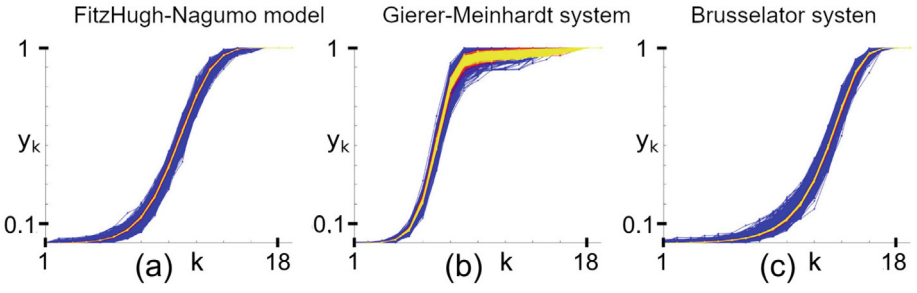

Fig. 8. The correlation integral vector components y_k for different number of samples in the training set ($N_{set} = 1000$ (yellow), $N_{set} = 500$ (red) and $N_{set} = 50$ (blue)). The FitzHugh-Nagumo model (a), the Gierer-Meinhardt system (b) and the Brusselator reaction-diffusion system (c).

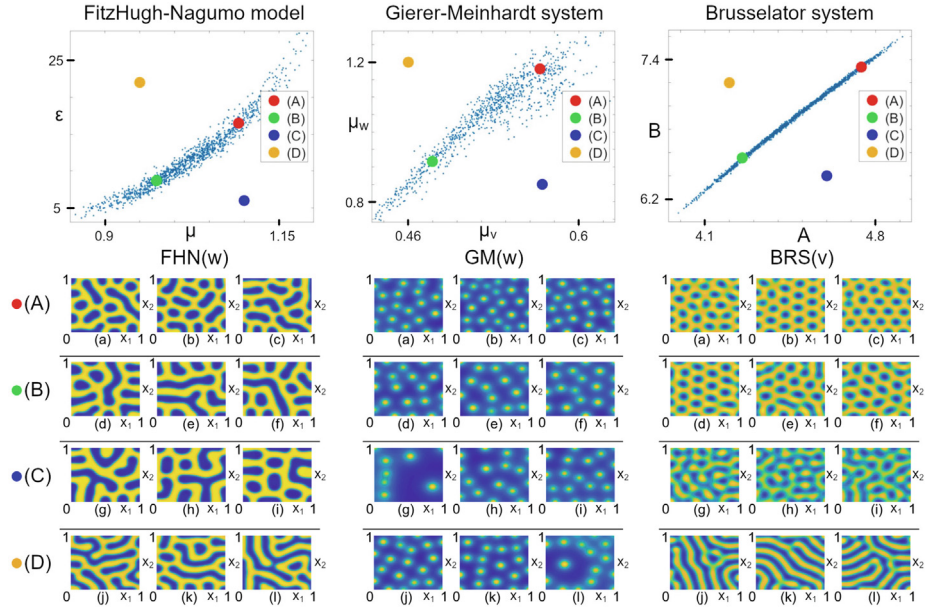

Fig. 9. Posterior distributions of model parameters and Turing patterns, obtained from direct simulations of the equations under study for the verification values. Posterior distributions of model parameters with verification values inside and outside the region for the case $N_{set} = 50$: (left) the FitzHugh-Nagumo model (verification values are: (A) $\mu = 1.092, \varepsilon = 16.5$, (B) $\mu = 0.97, \varepsilon = 8.7$, (C) $\mu = 1.1, \varepsilon = 6.0$ and (D) $\mu = 0.95, \varepsilon = 22.0$), (center) the Gierer-Meinhardt system (verification values are: (A) $\mu_v = 0.57, \mu_w = 1.18$, (B) $\mu_v = 0.48, \mu_w = 0.91$, (C) $\mu_v = 0.76, \mu_w = 0.85$ and (D) $\mu_v = 0.46, \mu_w = 1.2$) and (right) the Brusselator reaction-diffusion system (verification values are: (A) $A = 4.742, B = 7.335$, (B) $A = 4.253, B = 6.554$, (C) $A = 4.6, B = 6.4$ and (D) $A = 4.2, B = 7.2$). Turing patterns, obtained from direct simulations of the equations under study for verification values of control parameters, lying inside the posterior distribution region for the case $N_{set} = 50$ (points (A) and (B)) and outside it (points (C) and (D)). Non-normalized case. Initial conditions are taken as small random perturbations of the homogeneous steady state. The values of fixed model parameters are listed in Table 1. Left column: the FitzHugh-Nagumo model (variable w). Central column: the Gierer-Meinhardt system (variable w). Right column: the Brusselator reaction-diffusion system (variable v).

Table 2
Parameter values used for creating the training set for statistics of generalized correlation integral vector $\boldsymbol{y}(\theta_0)$ in the case of min-max normalized data.

Parameter	FHN	GM	BRS
R_0	0.81	0.47	0.7
M	32	32	32
b	1.016	1.05	1.025
N	500	500	500
n_{ens}	46	46	46

Here FHN, GM and BRS denote the FitzHugh-Nagumo model, Gierer-Meinhardt system and Brusselator reaction-diffusion system respectively.

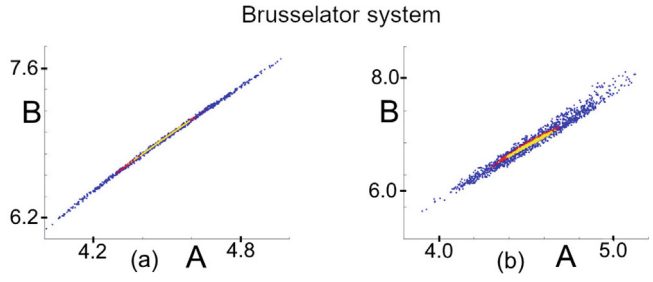

Fig. 10. Posterior distributions of model parameters for the case of sparse mesh. Posterior distributions of model parameters for the case of sparse mesh of 32×32 points and different number of samples in the training set ($N_{set} = 1000$ (yellow), $N_{set} = 500$ (red) and $N_{set} = 50$ (blue)). The Brusselator reaction-diffusion system: non-normalized case (a) and min-max normalized case (b). The values of model control parameter vector θ_0 and fixed model parameters are given in Table 1.

large enough set of pattern data we can create a feature vector that follows a multinomial Gaussian distribution. This allows a statistically sound way to determine if a given set of patterns belongs to the same family of patterns as the training data. Consequently, we

Table 3
"True" and estimated model parameters and constants used for creating the CIL likelihood in the case of $N_{set} = 50$ with min-max data normalization.

Parameter	FHN	GM	BRS
θ_0	$(1, 10)$	$(0.5, 1)$	$(4.5, 6.96)$
N_{set}	50	50	50
R_0	0.81	0.47	0.7
M	32	32	32
b	1.016	1.05	1.025
N	25	25	25
$\bar{\theta}_0$	$(1.067, 10.27)$	$(0.46, 0.9)$	$(4.47, 6.86)$

Fixed model parameters and constants for DE optimization are the same as in Table 1 as well as abbreviations FHN, GM and BRS.

can use the Gaussian likelihood as a cost function for parameter identification, which can be done by methods of statistical optimization. Next we can apply Markov Chain Monte Carlo (MCMC) sampling methods to determine the distribution of those parameters of a reaction-diffusion system that produce patterns similar to data.

The method was tested using the well-known FitzHugh-Nagumo, Gierer-Meinhardt and Brusselator reaction-diffusion systems. The amount of training data was varied from 23000 to 50 patterns. The method was used as a tool for estimating parameters of the reaction-diffusion systems, only using a collection of steady-state solutions as data, without the knowledge of initial state values. The performance of the method was satisfactory in all cases: a large amount of data leads to an extremely accurate detection of changes in model parameters, practically impossible to detect visually, while a modest amount of training patterns leads to the same level of detection as, roughly speaking, might be observed with the naked eye. Implementing an efficient numerical solver is crucial because numerical simulation of the model is repeated

many times to construct parameter posterior distribution by MCMC methods or to run Differential Evolution (DE) stochastic optimization.

Possible future applications of the approach include model identification in developmental biology based on patterns from real experimental data. While we here have used steady-state data only, the technique allows the inclusion of dynamic spatio-temporal patterns, with some necessary modifications. Moreover, it is possible to further minimize the amount of data needed, all the way to one pattern only, by combining the correlation integral likelihood (CIL) feature vectors with the idea of synthetic likelihoods. This, however, will come at the cost of considerably increased computational demands.

4. Materials and methods

4.1. Correlation integral likelihood

Let us recall the distance concept introduced in Haario (2015) and Springer et al. (2019) for chaotic dynamical systems, and adapt the use of it to quantify pattern formation in reaction-diffusion systems. Consider a chaotic ODE system, defined by the following equation

$$\frac{d\boldsymbol{s}}{dt} = \boldsymbol{F}(\boldsymbol{s}, \boldsymbol{\theta}),\ \boldsymbol{s}(t)|_{t=0} = \boldsymbol{x} \tag{7}$$

where $\boldsymbol{s} \in \mathbb{R}^n$ is a system state, $\boldsymbol{x} \in \mathbb{R}^n$ is a vector of initial conditions and $\boldsymbol{\theta} \in \mathbb{R}^d$ is a parameter vector. Such systems (for example, the classical Lorenz system) are highly sensitive to initial conditions and model parameters, small changes in these values lead to widely diverging trajectories, which makes long-term predictions impossible in general. The work in Haario (2015) deals with the following question: how to define a practically computable distance concept for chaotic trajectories, i.e. how to define the distribution of model parameters for which the variability due to the differences in parameter values is indistinguishable from the chaotic variability of the system, as quantified by a given amount of training data.

The distance concept is based on the notion of correlation integral vector. In the following, we use the notation $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{x}, t)$ for a trajectory $\boldsymbol{s}$ of chaotic system (7) that depends on a model parameter vector $\boldsymbol{\theta}$, input values $\boldsymbol{x}$ and time t. Generally, $\boldsymbol{x}$ may include any variables whose small changes lead to chaotic unpredictability, e.g., the tolerances of the numerical solver. In the current paper, however, we denote the initial values of an ODE system by $\boldsymbol{x}$ only.

Denote by $\boldsymbol{s}_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, N$ the points of the trajectory $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{x}, t)$, evaluated for consecutive time points $t = t_i$. For a fixed $R > 0$ set

$$C(R, N) = \frac{1}{N^2} \sum_{1 \leqslant i,j \leqslant N} \#(||\boldsymbol{s}_i - \boldsymbol{s}_j|| < R).$$

where $\boldsymbol{s}_i = \boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{x}, t_i)$ and $\boldsymbol{s}_j = \boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{x}, t_j)$. Above $||.||$ denotes the Euclidean distance. So $C(R, N)$ gives the fraction of pairs of points with a distance less than R. As the radius R tends to zero, $C(R, N) \sim R^{\gamma}$, where γ is the classical *correlation dimension*, given by the formula (see Cencini et al., 2009)

$$\gamma = \lim_{R \to 0} \frac{\log C(R, N)}{\log R}. \tag{8}$$

We are, however, not interested in the intrinsic fractal dimension of a given trajectory as given by the limit $R \to 0$, but want to characterize the distance between different trajectories, using all relevant scales R.

Suppose we have an ensemble of different trajectories $\boldsymbol{s}^k = \boldsymbol{s}^k(\boldsymbol{\theta}, \boldsymbol{x}^k, t), k = 1, \ldots, n_{ens}$, each evaluated at time points $t_i, i = 1, \ldots, N$. For fixed $R > 0$ and different trajectories $\boldsymbol{s}^k, \boldsymbol{s}^l$ set

$$C(R, N, \boldsymbol{s}^k, \boldsymbol{s}^l) = \frac{1}{N^2} \sum_{1 \leqslant i,j \leqslant N} \#(||\boldsymbol{s}_i^k - \boldsymbol{s}_j^l|| < R), \tag{9}$$

where $\boldsymbol{s}_i^k = \boldsymbol{s}^k(\boldsymbol{\theta}, \boldsymbol{x}^k, t_i)$ and $\boldsymbol{s}_j^l = \boldsymbol{s}^l(\boldsymbol{\theta}, \boldsymbol{x}^l, t_j)$. We assume furthermore that the system state $\boldsymbol{s}$ remains bounded, and all the trajectory points $\boldsymbol{s}_i^k, \boldsymbol{s}_j^l$ are samples from the same underlying fixed attractor (i.e., they are measurements given by a stationary time series, or obtained by integrating a fixed chaotic system with perturbed initial values $\boldsymbol{x}^k, \boldsymbol{x}^l$). Choose $R_0 > 0$ such that $||\boldsymbol{s}_i^k - \boldsymbol{s}_j^l|| < R_0$ for all $k, l = 1, \ldots, n_{ens}, i, j = 1, \ldots, N, k \neq l$. For a given integer M, the *generalized correlation integral vector* $\boldsymbol{y}^{k,l}(\boldsymbol{\theta}) \in \mathbb{R}^M$ of the pair $\boldsymbol{s}^k, \boldsymbol{s}^l$ is given by the components

$$y_m^{k,l} = C(R_m, N, \boldsymbol{s}^k, \boldsymbol{s}^l),\ m = 1, 2, \ldots, M, \tag{10}$$

where $R_m = b^{-m} R_0$ with $b > 1$ (for the numerical evaluations we follow the usual settings found in literature (Cencini et al., 2009), see also the examples below).

The correlation integral vector $\boldsymbol{y}^{k,l}(\boldsymbol{\theta}) = (y_1^{k,l}, y_2^{k,l}, \ldots, y_M^{k,l})$ is stochastic, as the evaluations are done by randomized initial conditions $\boldsymbol{x}^k$ and $\boldsymbol{x}^l$. For simplicity, we omit indexes k and l and initial data $\boldsymbol{x}^k, \boldsymbol{x}^l$ in the following discussion. It turns out that the distribution of the correlation integral vector is Gaussian. Intuitively, as the expression (9) consists of averages, the Central Limit Theorem might be expected to hold. More exactly, the expressions (9) and (10) define the empirical cumulative distribution function of the respective set of distances, evaluated at bin values R_m. The basic form of the Donsker's theorem tells that empirical cumulative distribution functions of i.i.d scalar random numbers asymptotically tend to a Brownian bridge (Donsker, 1951; Donsker, 1952). In more general settings, that cover our situation, the Gaussianity is established by theorems of the so-called U-statistics, see (Borovkova et al., 2001; Neumeyer, 2004). In Borovkova et al. (2001), especially, the Gaussianity of the numerical construction used to estimate the correlation dimension (8) of the classical Lorenz attractor is discussed.

The *Correlation Integral Likelihood (CIL)* is defined as the Gaussian distribution $N(\boldsymbol{\mu}_0, \Sigma_0)$, where $\boldsymbol{\mu}_0 \in \mathbb{R}^M$ and $\Sigma_0 \in \mathbb{R}^{M \times M}$ are the mean and covariance of the vector $\boldsymbol{y}(\boldsymbol{\theta})$. Numerically, $\boldsymbol{\mu}_0$ and Σ_0 are estimated as the mean and covariance of the set of M dimensional vectors (10) given by all the pairs $k, l = 1, 2, \ldots, n_{ens}$.

In this paper we use the correlation integral likelihood in order to create a statistical distribution that quantifies the variability of Turing patterns within a given reaction-diffusion system (1). Certain modifications to the approach in Haario (2015) are necessary. First, by applying the Method of Lines we reduce an infinite-dimensional reaction-diffusion system (1) to a finite ODE system (see the next subsection), which makes it possible applying the CIL concept.

Next, we modify the concept of trajectory $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\theta}, \boldsymbol{x}, t)$. In Haario (2015), the samples $\boldsymbol{s}_i$ are vectors from a numerically computed trajectory $\boldsymbol{s}$, evaluated at time points t_i. So they actually are samples from the chaotic attractor in the phase space. In the case of reaction-diffusion system (1), when model parameters belong to a pure Turing domain, numerical simulations converge to stable stationary states (Turing patterns). These stationary states belong to a global attracting set of the system considered. Therefore, we choose inhomogeneous steady states of the reaction-diffusion systems to be the samples $\boldsymbol{s}_i$. Each subset of N patterns is obtained by solving the Eq. (1) N times with a fixed model parameter vector $\boldsymbol{\theta}$ but randomized initial values $\boldsymbol{x}$.

Finally, we choose appropriately the distance. For $\boldsymbol{s}_i, \boldsymbol{s}_j$ we define the distance by L_2-norm formula

$$||\boldsymbol{s}_i - \boldsymbol{s}_j||^2 = \int_\Omega [(v_i - v_j)^2 + (w_i - w_j)^2] dx_1 dx_2,$$

where $\boldsymbol{s}_i = (v_i(\boldsymbol{x}), w_i(\boldsymbol{x}))$ and $\boldsymbol{s}_j = (v_j(\boldsymbol{x}), w_j(\boldsymbol{x}))$. In numerical experiments this integral is approximated by the trapezoidal rule.

The MATLAB code, which was used for all numerical simulations, discussed in this paper is publicly available on GitHub. Please see the link: https://github.com/AlexeyKazarnikov/CILNumericalCode.

4.2. Numerical solution of the equations

To project an infinite-dimensional system onto a finite-dimensional system we apply the Method of Lines (MOL). We create the decomposition of domain Ω by equidistant grid with fixed step size $h = 1/(M_{dim} - 1), M_{dim} \in \mathbb{N}$ and arrive at a finite set of points

$$\{(x_1^i, x_2^j) : x_1^i = (i-1)h, x_2^j = (j-1)h, i,j = 1, \dots, M_{dim}\},$$

Next, the Laplace operator is discretized by the five-point stencil (Grossmann et al., 2007; Hupkes and Van Vleck, 2016)

$$\Delta u \approx \frac{u_{i+1\,j} + u_{i-1\,j} + u_{i\,j+1} + u_{i\,j-1} - 4u_{i\,j}}{h^2} = \nabla_5^2 u_{i\,j},$$

and an infinite-dimensional reaction-diffusion system (1) is reduced to a finite set of $2M_{dim}^2$ ordinary differential equations (ODE), which is expressed by the system

$$\begin{aligned} \dot{v}_{i\,j}(t) &= \nu_1 \nabla_5^2 v_{i\,j}(t) + f(v_{i\,j}(t), w_{i\,j}(t)), \\ \dot{w}_{i\,j}(t) &= \nu_2 \nabla_5^2 w_{i\,j}(t) + g(v_{i\,j}(t), w_{i\,j}(t)), \end{aligned} \tag{11}$$

where $i,j = 1, 2, \dots, M_{dim}$ and Neumann boundary conditions are taken into account by applying central difference scheme (Grossmann et al., 2007), which leads to assumptions

$$\begin{aligned} &v_{0\,j}(t) \equiv v_{1\,j}(t),\ v_{M_{dim}+1\,j}(t) \equiv v_{M_{dim}\,j}(t), \\ &w_{0\,j}(t) \equiv w_{1\,j}(t),\ w_{M_{dim}+1\,j}(t) \equiv w_{M_{dim}\,j}(t), \\ &v_{i\,0}(t) \equiv v_{i\,1}(t),\ v_{i\,M_{dim}+1}(t) \equiv v_{i\,M_{dim}}(t), \\ &w_{i\,0}(t) \equiv w_{i\,1}(t),\ w_{i\,M_{dim}+1}(t) \equiv w_{i\,M_{dim}}(t). \end{aligned}$$

In all numerical experiments, being discussed in the current paper, we use $M_{dim} = 64$ (with the exception of a sparse grid case with $M_{dim} = 32$, which is discussed in Limited Data).

When the control parameters of a reaction-diffusion model are fixed and belong to Turing domain, patterns are obtained from numerical simulations as steady states of the MOL ODE system. Perturbations of initial conditions result in convergence of numerical solution to different steady states.

Computing steady states for the CIL vectors requires running model simulations with the same parameter values but different initial conditions. This task is very suitable for modern GPUs, which allow performing computation of a large number of solutions in parallel. This enables a significant improvement of performance, making the proposed approach more suitable for numerical applications. For our experiments we implemented an efficient parallel algorithm for computing steady state solutions of reaction-diffusion systems on GPU. The code was run on a laptop with Nvidia GeForce GTX 980 SLI, where computing of all steady states (typically 500 simulations) for one correlation integral vector $\boldsymbol{y}(\theta)$ required approximately 12.4 s.

In all examples considered in this paper the integration of MOL system was done by explicit Runge-Kutta method (RK4) with fixed time step $h = 0.0065$. Numerical solution was computed in the time interval $[0, 100]$. Following the idea, proposed in Campillo-Funollet et al. (2019), we evaluated at final time point $T = 100$ the time derivative of the solution by finite difference scheme. We checked that for the chosen time interval the L_2-norm of the time derivative was less than $\delta = 10^{-3}$ for all the systems considered, therefore we assumed that the steady state was reached. For several randomly chosen values of control parameters, used in numerical experiments, we also integrated the systems on the interval $[0, 2T]$ and checked that the difference between the solutions at both times was less than δ.

Where possible, we verified that the patterns obtained are similar to those obtained by other authors and available in literature (see Painter et al., 2012) for the case of the Gierer-Meinhardt system (3 and Peña and Pérez-García (2001) for the case of the Brusselator reaction-diffusion system (4)).

Table 4
Constants for DE optimization.

Parameter	FHN	GM	BRS
D_0	$[1,2] \times [35,50]$	$[0.1,5] \times [8,20]$	$[3,6] \times [10,15]$
N_p	60	60	60
η	0.8	0.8	0.8
ξ	0.5	0.5	0.5

Here FHN, GM and BRS denote the FitzHugh-Nagumo model, Gierer-Meinhardt system and Brusselator reaction-diffusion system respectively.

4.3. Differential evolution

For minimizing the stochastic cost function $f(\theta)$ (defined in Eq. (5)) we employ the method of Differential Evolution (DE). It is a stochastic, population-based optimization algorithm (Storn and Price, 1997). We start from a population of parameter vectors $\{\boldsymbol{\theta}_{1,0}, \boldsymbol{\theta}_{2,0}, \dots, \boldsymbol{\theta}_{N_p,0}\}$, where N_p is the size of a population (we use $N_p = 60$). Initial population is randomly distributed on the interval D_0 (which is selected such that it does not contain the original point $\boldsymbol{\theta}_0$). On the initial step, the best candidate $\boldsymbol{\theta}_0^*$ is defined as element with lowest value of the cost function $f(\boldsymbol{\theta})$.

The algorithm is run iteratively. On the each iteration or "generation" G the population goes through three phases: mutation, recombination and selection. During the mutation phase the search space is expanded by generating the donor elements $\boldsymbol{g}_{i,G+1} = \boldsymbol{\theta}_G^* + \eta(\boldsymbol{\theta}_{i_1,G} - \boldsymbol{\theta}_{i_2,G}), i = 1, \dots, N_p$. Here $\boldsymbol{\theta}_G^*$ is the best candidate in the G-th iteration, η is a constant and indexes i_1 and i_2 are selected randomly for each donor element $\boldsymbol{g}_{i,G+1}$. Next, during the recombination phase trial elements $\boldsymbol{t}_{i,G+1}$ are created from previous elements by the following rule: a trial element $\boldsymbol{t}_{i,G+1}$ is taken to be $\boldsymbol{g}_{i,G+1}$ or $\boldsymbol{\theta}_{i,G}$ with probability ξ. Finally, on the selection phase the elements $\boldsymbol{\theta}_{i,G}$ are compared with the trial elements and those with lowest function values are admitted to the next generation. All parameter values used in the DE optimization are presented in the Table 4.

We run the algorithm for 100 iterations, which is enough to achieve convergence of the population close to the "true" value $\boldsymbol{\theta}_0$. The final parameter estimate can be taken as the mean of the elements of the last population.

CRediT authorship contribution statement

Alexey Kazarnikov: Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Heikki Haario:** Conceptualization, Formal analysis, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research was funded by the Centre of Excellence of Inverse Modelling and Imaging, Academy of Finland, decision number 312122. We would like to thank Anna Marciniak-Czochra, Antti Hannukainen and Teemu Häkkinen for usefull discussions.

References

Barrass, I., Crampin, E.J., Maini, P.K., 2006. Mode transitions in a model reaction–diffusion system driven by domain growth and noise. Bull. Math. Biol. 68 (5), 981–995. https://doi.org/10.1007/s11538-006-9106-8.

Borovkova, S., Burton, R., Dehling, H., 2001. Limit theorems for functionals of mixing processes with applications to u-statistics and dimension estimation. Trans. Am. Math. Soc. 353 (11), 4261–4318. URL: http://www.jstor.org/stable/2693737.

Brinkmann, F., Mercker, M., Richter, T., Marciniak-Czochra, A., 2018. Post-turing tissue pattern formation: Advent of mechanochemistry. PLOS Comput. Biol. 14 (7), 1–21. https://doi.org/10.1371/journal.pcbi.1006259.

Campillo-Funollet, E., Venkataraman, C., Madzvamuse, A., 2019. Bayesian parameter identification for turing systems on stationary and evolving domains. Bull. Math. Biol. 81 (1), 81–104. https://doi.org/10.1007/s11538-018-0518-z.

Castets, V., Dulos, E., Boissonade, J., De Kepper, P., 1990. Experimental evidence of a sustained standing turing-type nonequilibrium chemical pattern. Phys. Rev. Lett. 64, 2953–2956. https://doi.org/10.1103/PhysRevLett.64.2953.

Cencini, M., Cecconi, F., Vulpiani, A., 2009. Chaos: From Simple Models To Complex Systems. World Scientific Publishing Co.

Chakravarti, S., Marek, M., Ray, W.H., 1995. Reaction-diffusion system with brusselator kinetics: Control of a quasiperiodic route to chaos. Phys. Rev. E 52, 2407–2423. https://doi.org/10.1103/PhysRevE.52.2407.

Dong, T., Xu, W., Liao, X., 2017. Hopf bifurcation analysis of reaction–diffusion neural oscillator system with excitatory-to-inhibitory connection and time delay. Nonlinear Dyn. 89 (4), 2329–2345. https://doi.org/10.1007/s11071-017-3589-8.

Donsker, M., 1951. An Invariance Principle for Certain Probability Limit Theorems. American Mathematical Society, Memoirs.

Donsker, M.D., 1952. Justification and extension of doob's heuristic approach to the Kolmogorov-Smirnov theorems. Ann. Math. Statist. 23 (2), 277–281. https://doi.org/10.1214/aoms/1177729445.

Ebihara, M., Mahara, H., Sakurai, T., Nomura, A., Miike, H., 2003. Image processing by a discrete reaction-diffusion system. Proc. Visualization, Imaging, Image Process. 396.

Escala, D.M., Guiu-Souto, J., Munuzuri, A.P., 2015. Externally controlled anisotropy in pattern-forming reaction-diffusion systems. Chaos 25, (6). https://doi.org/10.1063/1.4922303 064309.

Facchini, A., Rossi, F., Mocenni, C., 2009. Spatial recurrence strategies reveal different routes to Turing pattern formation in chemical systems. Phys. Lett., Section A: General, At. Solid State Phys. 373 (46), 4266–4272. https://doi.org/10.1016/j.physleta.2009.09.049.

FitzHugh, R., 1961. Impulses and physiological states in theoretical models of nerve membrane. Biophys. J . 1 (6), 445–466. https://doi.org/10.1016/S0006-3495(61)86902-6.

Garvie, M.R., Maini, P.K., Trenchea, C., 2010. An efficient and robust numerical algorithm for estimating parameters in turing systems. J. Comput. Phys. 229 (19), 7058–7071. https://doi.org/10.1016/j.jcp.2010.05.040. URL: http://www.sciencedirect.com/science/article/pii/S0021999110003128.

Ghorai, S., Poria, S., 2016. Pattern formation and control of spatiotemporal chaos in a reaction diffusion prey-predator system supplying additional food. Chaos, Solitons Fractals 85, 57–67. https://doi.org/10.1016/j.chaos.2016.01.013.

Gierer, A., Meinhardt, H., 1972. A theory of biological pattern formation. Kybernetik 12 (1), 30–39. https://doi.org/10.1007/BF00289234.

Grossmann, C., Gorg, H.-R., Stynes, M., 2007. Numerical Treatment of Partial Differential Equations. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-71584-9.

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli 7 (2), 223–242. URL: https://projecteuclid.org:443/euclid.bj/1080222083.

Haario, H., Laine, M., Mira, A., Saksman, E., 2006. DRAM: Efficient adaptive MCMC. Stat. Comput. 16 (4), 339–354. https://doi.org/10.1007/s11222-006-9438-0.

Haario, H., Kalachev, L., Hakkarainen, J., 2015. Generalized correlation integral vectors: A distance concept for chaotic dynamical systems. Chaos 25 (6). https://doi.org/10.1063/1.4921939.

Härting, S., Marciniak-Czochra, A., Takagi, I., 2017. Stable patterns with jump discontinuity in systems with turing instability and hysteresis. Discrete Continuous Dyn. Syst. — A 37 (2), 757–800. https://doi.org/10.3934/dcds.2017032.

Horváth, A.K., Dolnik, M., Muñuzuri, A.P., Zhabotinsky, A.M., Epstein, I.R., 1999. Control of turing structures by periodic illumination. Phys. Rev. Lett. 83, 2950–2952. https://doi.org/10.1103/PhysRevLett.83.2950.

Hupkes, H.J., Van Vleck, E.S., 2016. Travelling waves for complete discretizations of reaction diffusion systems. J. Dyn. Diff. Eqs. 28 (3), 955–1006. https://doi.org/10.1007/s10884-014-9423-9.

Kazarnikov, A.V., Revina, S.V., 2016. Vozniknovenie avtokolebanii v sisteme Releya s diffusiei [the onset of auto-oscillations in Rayleigh system with diffusion]. Bull. South Ural State University. Ser. Math. Model., Programm. Comput. Softw. 9 (2), 16–28. 10/14529/mmp160202.

Kazarnikov, A.V., Revina, S.V., 2016. Asimptotika statsionarnykh reshenii sistemy Releya s diffusiei [asymptotics of stationary solutions of Rayleigh reaction-diffusion system], Izvestiya VUZov. Severo-Kavkazskii region 191, 13–19. 10/18522/0321-3005-2016-3-13-19.

Kazarnikov, A.V., Revina, S.V., Haario, H., 2015. Numerical and asymptotical analysis of Rayleigh reaction-diffusion system. In: Bai, Z., Krukier, L.A., Muratova, G.V. (Eds.), Numerical algebra with applications. Proceedings of Fourth China-Russia Conference. Southern Federal University Publishing, pp. 114–119.

Koch, A.J., Meinhardt, H., 1994. Biological pattern formation: from basic mechanisms to complex structures. Rev. Mod. Phys. 66, 1481–1507. https://doi.org/10.1103/RevModPhys.66.1481.

Kondo, S., Asai, R., 1995. A reaction-diffusion wave on the skin of the marine angelfish pomacanthus. Nature 376, 765–768.

Kondo, S., Miura, T., 2010. Reaction-diffusion model as a framework for understanding biological pattern formation. Science 329 (5999), 1616–1620. https://doi.org/10.1126/science.1179047. URL: https://science.sciencemag.org/content/329/5999/1616.

Kramer, S., Bollt, E.M., 2013. Spatially dependent parameter estimation and nonlinear data assimilation by autosynchronization of a system of partial differential equations. Chaos 23 (3). https://doi.org/10.1063/1.4812722.

Lee, K.J., McCormick, W.D., Ouyang, Q., Swinney, H.L., 1993. Pattern formation by interacting chemical fronts. Science 261 (5118), 192–194. https://doi.org/10.1126/science.261.5118.192.

Lefèvre, J., Mangin, J.-F., 2010. A reaction-diffusion model of human brain development. PLOS Comput. Biol. 6 (4), 1–10. https://doi.org/10.1371/journal.pcbi.1000749.

Li, A.-W., 2011. Impact of noise on pattern formation in a predator–prey model. Nonlinear Dyn. 66 (4), 689–694. https://doi.org/10.1007/s11071-010-9941-x.

Madzvamuse, A., Thomas, R.D.K., Maini, P.K., Wathen, A.J., 2002. A numerical approach to the study of spatial pattern formation in the ligaments of arcoid bivalves. Bull. Math. Biol. 64 (3), 501–530. https://doi.org/10.1006/bulm.2002.0283.

Maini, P.K., Painter, K.J., Nguyen Phong Chau, H., 1997. Spatial pattern formation in chemical and biological systems. J. Chem. Soc., Faraday Trans. 93, 3601–3610. https://doi.org/10.1039/A702602A.

Mercker, M., Hartmann, D., Marciniak-Czochra, A., 2013. A mechanochemical model for embryonic pattern formation: coupling tissue mechanics and morphogen expression. PLOS ONE 8 (12), 1–6. https://doi.org/10.1371/journal.pone.0082617.

Mercker, M., Köthe, A., Marciniak-Czochra, A., 2015. Mechanochemical symmetry breaking in hydra aggregates. Biophys. J . 108 (9), 2396–2407. https://doi.org/10.1016/j.bpj.2015.03.033.

Miura, T., Komori, M., Shiota, K., 2000. A novel method for analysis of the periodicity of chondrogenic patterns in limb bud cell culture: correlation of in vitro pattern formation with theoretical models. Anat. Embryol. 201 (5), 419–428. https://doi.org/10.1007/s004290050329.

Mocenni, C., Facchini, A., Vicino, A., 2010. Identifying the dynamics of complex spatio-temporal systems by spatial recurrence properties. Proc. Nat. Acad. Sci. USA 107 (18), 8097–8102. https://doi.org/10.1073/pnas.0910414107.

Murray, J.D., 1993. Mathematical Biology II: Spatial Models and Biomedical Applications. Springer-Verlag. https://doi.org/10.1007/0-387-22438.

Nagumo, J., Arimoto, S., Yoshizawa, S., 1962. An active pulse transmission line simulating nerve axon. Proc. IRE 50(10), 2061–2070. https://doi.org/10.1109/JRPROC.1962.288235.

Neumeyer, N., 2004. A central limit theorem for two-sample u-processes. Stat. Prob. Lett. 67 (1), 73–85. https://doi.org/10.1016/j.spl.2002.12.001. URL: www.sciencedirect.com/science/article/pii/S0167715204000112.

Nomura, A., Ichikawa, M., Okada, K., Miike, H., Sakurai, T., 2011. Edge detection algorithm inspired by pattern formation processes of reaction-diffusion systems. Int. J. Circuits, Syst. Signal Process. 5, 105–115.

Owen, M.R., Lewis, M.A., 2001. How predation can slow, stop or reverse a prey invasion. Bull. Math. Biol. 63 (4), 655. https://doi.org/10.1006/bulm.2001.0239.

Painter, K.J., Hunt, G.S., Wells, K.L., Johansson, J.A., Headon, D.J., 2012. Towards an integrated experimental-theoretical approach for assessing the mechanistic basis of hair and feather morphogenesis. Interface Focus 2 (4), 433–450. https://doi.org/10.1098/rsfs.2011.0122.

Peña, B., Pérez-García, C., 2001. Stability of turing patterns in the brusselator model. Phys. Rev. E 64,. https://doi.org/10.1103/PhysRevE.64.056213 056213.

Price, L.F., Drovandi, C.C., Lee, A., Nott, D.J., 2018. Bayesian synthetic likelihood. J. Comput. Graphical Stat. 27 (1), 1–11. https://doi.org/10.1080/10618600.2017.1302882.

Prigogine, I., Lefever, R., 1968. Symmetry breaking instabilities in dissipative systems. ii. J. Chem. Phys. 48 (4), 1695–1700. https://doi.org/10.1063/1.1668896.

Robert, C.P., Casella, G., 2004. Monte Carlo Statistical Methods. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4757-4145-2_1.

Springer, S., Haario, H., Shemyakin, V., Kalachev, L., Shchepakin, D., 2019. Robust parameter estimation of chaotic systems. Inverse Problems Imag. 13 (6), 1189. https://doi.org/10.3934/ipi.2019053.

Storn, R., Price, K., 1997. Differential Evolution – a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11 (4), 341–359. https://doi.org/10.1023/A:1008202821328.

Szalai, I., De Kepper, P., 2008. Pattern formation in the ferrocyanide-iodate-sulfite reaction: the control of space scale separation. Chaos 18 (2). https://doi.org/10.1063/1.2912719. 0–9.

Tang, X., Song, Y., Zhang, T., 2016. Turing-Hopf bifurcation analysis of a predator–prey model with herd behavior and cross-diffusion. Nonlinear Dyn. 86 (1), 73–89. https://doi.org/10.1007/s11071-016-2873-3.

Turing, A.M., 1952. The chemical basis of morphogenesis. Philos. Trans. R. Soc. London. Ser. B, Biolog. Sci. 237 (641), 37–72.

Upadhyay, R.K., Roy, P., Datta, J., 2014. Complex dynamics of ecological systems under nonlinear harvesting: Hopf bifurcation and Turing instability. Nonlinear Dyn. 79, 2251–2270. https://doi.org/10.1007/s11071-014-1808-0.

Vanag, V.K., Epstein, I.R., 2008. Design and control of patterns in reaction-diffusion systems. Chaos 18 (2). https://doi.org/10.1063/1.2900555.

van der Vaart, A.W., 1998. Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press. https://doi.org/10.1017/CBO9780511802256.

Vilas, C., Balsa-Canto, E., Garcia, M.G., 2012. Dynamic optimization of distributed biological systems using robust and efficient numerical techniques. BMC Syst. Biol. 6 (79).

Wood, S.N., 2010. Statistical inference for noisy nonlinear ecological dynamic systems. Nature 466 (7310), 1102–1104. https://doi.org/10.1038/nature09319.

Xu, J., Yang, G., Xi, H., Su, J., 2015. Pattern dynamics of a predator-prey reaction-diffusion model with spatiotemporal delay. Nonlinear Dyn. 81 (4), 2155–2163. https://doi.org/10.1007/s11071-015-2132-z.

Zhang, Q., Tian, C., 2014. Pattern dynamics in a diffusive Rössler model. Nonlinear Dyn. 78 (2), 1489–1501. https://doi.org/10.1007/s11071-014-1530-y.

Zhao, H., Huang, X., 2014. Turing instability and pattern formation of neural networks with reaction-diffusion terms. Nonlinear Dyn. 76 (1), 115–124. https://doi.org/10.1007/s11071-013-1114-2.