

# Bayesian updating of bacterial microfilms under hybrid uncertainties with a novel surrogate model

Lukas Fritsch<sup>1,2</sup>, Hendrik Geisler<sup>3\*</sup>, Jan Grashorn<sup>4</sup>,  
Felix Klemp<sup>3</sup>, Meisam Soleimani<sup>3</sup>, Matteo Broggi<sup>1</sup>,  
Philipp Junker<sup>2,3</sup>, Michael Beer<sup>1,2,5,6</sup>

<sup>1</sup>Institute for Risk and Reliability, Leibniz University Hannover,  
Callinstr. 34, Hannover, 30167, Germany.

<sup>2</sup>International Research Training Group (IRTG) 2657, Leibniz  
University Hannover, Appelstr. 11/11a, Hannover, 30167, Germany.

<sup>3</sup>Institute for Continuum Mechanics, Leibniz University Hannover, An  
der Universität 1, Garbsen, 30823, Germany.

<sup>4</sup>Chair of Engineering Materials and Building Preservation,  
Helmut-Schmidt-University, Holstenhofweg 85, Hamburg, 22043,  
Germany.

<sup>5</sup>Department of Civil and Environmental Engineering, University of  
Liverpool, Liverpool, L69 3GH, UK.

<sup>6</sup>International Joint Research Center for Resilient Infrastructure &  
International Joint Research Center for Engineering Reliability and  
Stochastic Mechanics, Tongji University, 1239 Siping road, Shanghai,  
200092, People's Republic of China.

\*Corresponding author(s). E-mail(s): [geisler@ikm.uni-hannover.de](mailto:geisler@ikm.uni-hannover.de);  
Contributing authors: [fritsch@irz.uni-hannover.de](mailto:fritsch@irz.uni-hannover.de); [grashorn@hsu-hh.de](mailto:grashorn@hsu-hh.de);  
[klemp@ikm.uni-hannover.de](mailto:klemp@ikm.uni-hannover.de); [soleimani@ikm.uni-hannover.de](mailto:soleimani@ikm.uni-hannover.de);  
[broggi@irz.uni-hannover.de](mailto:broggi@irz.uni-hannover.de); [junker@ikm.uni-hannover.de](mailto:junker@ikm.uni-hannover.de);  
[beer@irz.uni-hannover.de](mailto:beer@irz.uni-hannover.de);

## Abstract

Accurate modeling of bacterial biofilm growth is essential for understanding their complex dynamics in biomedical, environmental, and industrial settings. These dynamics are shaped by a variety of environmental influences, including the presence of antibiotics, nutrient availability, and inter-species interactions, all of

which affect species-specific growth rates. However, capturing this behavior in computational models is challenging due to the presence of hybrid uncertainties, a combination of epistemic uncertainty (stemming from incomplete knowledge about model parameters) and aleatory uncertainty (reflecting inherent biological variability and stochastic environmental conditions). In this work, we present a Bayesian model updating (BMU) framework to calibrate a recently introduced multi-species biofilm growth model. To enable efficient inference in the presence of hybrid uncertainties, we construct a reduced-order model (ROM) derived using the Time-Separated Stochastic Mechanics (TSM) approach. TSM allows for an efficient propagation of aleatory uncertainty, which enables single-loop Bayesian inference, thereby avoiding the computationally expensive nested (double-loop) schemes typically required in hybrid uncertainty quantification. The BMU framework employs a likelihood function constructed from the mean and variance of stochastic model outputs, enabling robust parameter calibration even under sparse and noisy data. We validate our approach through two case studies: a two-species and a four-species biofilm model. Both demonstrate that our method not only accurately recovers the underlying model parameters but also provides predictive responses consistent with the synthetic data.

**Keywords:** Bayesian updating, hybrid uncertainty, bacterial biofilms, time-separated stochastic mechanics, model calibration

## 1 Introduction

Bacterial biofilms are structured microbial communities whose growth dynamics are influenced by environmental conditions, nutrient availability, antibiotics, and inter-species interactions [1]. A key feature of biofilms is their remarkable resilience: they can exhibit up to 1000-fold greater tolerance to antibiotics and environmental stressors compared to planktonic (free-floating) bacteria [2]. This inherent robustness contributes to the widespread presence of biofilms across diverse settings, such as natural ecosystems, industrial systems, and clinical environments. In industrial and environmental contexts, biofilms can play beneficial roles, such as in wastewater treatment processes [3]. However, they are also associated with numerous challenges, including persistent infections [4, 5], medical device contamination [6], and infrastructure bio-fouling [7]. One particularly important area is oral biofilm formation, which can lead to infections around dental implants [8–11].

In many environments, biofilms are composed of multiple microbial species that compete for resources and respond collectively to external cues [12–14]. Understanding the dynamics of such multispecies biofilms and modeling these systems is critical for any application. Some fundamental interaction principles are outlined by James et al. [15] and they serve as a theoretical foundation of how the species interact, and also how they do not interact in some scenarios. Depending on the specific area of application, different aspects are modeled and different parameters are used to describe the physical and chemical processes of the biofilm growth. Ouidir et al. [16] review different approaches for the modeling of such systems and classify them by application

to wastewater treatment, soil, and biomedical applications. Specifically, they focus on biofilms in the oral cavity which is important for dental applications, also highlighted in [14, 17].

A recently proposed continuum model by Klemp et al. [18], derived from the extended Hamilton principle, captures multi-species biofilm growth by introducing abstract material parameters. However, these parameters are typically unknown and subject to uncertainty. Combined with the inherent stochasticity of biological processes, this presents a significant challenge for constructing predictive and physically meaningful models. Accurate parameter inference is thus essential to identify underlying biofilm properties from data and enable robust modeling.

Various strategies for calibrating biological models have been proposed. Frequentist approaches, such as those reviewed by Read et al. [19], rely on statistical tests (e.g., Kolmogorov-Smirnov) to compare model output distributions with data. Other works follow similar strategies [20–23]. In contrast, Bayesian inference provides a probabilistic framework for parameter estimation that naturally incorporates uncertainty [24, 25]. While Bayesian methods offer interpretability and flexibility, their computational cost remains a key limitation. Recent developments, including modern Markov Chain Monte Carlo (MCMC), sequential Monte Carlo, and Approximate Bayesian Computation (ABC), have helped mitigate this burden [24].

Despite their advantages, parameter calibration is often neglected in biofilm modeling. For example, Shewa et al. [23] and Rittmann et al. [26] note that default literature values are frequently used instead of performing parameter calibration. Recent studies have demonstrated the value of Bayesian inference for biofilm models, including parameter estimation with quorum sensing [27] and the inference of rheological properties from experimental data [28]. In the field of computational mechanics, Willmann et al. [29] propose a Bayesian framework for the calibration of a model of multi-physics biofilm model introduced in Ref. [30]. The authors present an approach that can also handle unavoidable uncertainty and reduces the computational cost through the use of a Gaussian process surrogate for the log-likelihood. Further, different discrepancy metrics are introduced and compared in their study.

The problem to estimate parameters under uncertainty with limited experimental data is also present in many other fields of applications. One such application is the estimation of parameters of constitutive models in computational mechanics. Wollner et al. [31] present a Bayesian inference framework which they apply to the parameter estimation of the hyperelastic Ogden model. A summary of approaches common in applications in mechanical and civil engineering are given in Ref. [32]. There, again, we come back to Bayesian inference which allows for the handling of the unknown parameters through a posterior probability distribution of said parameters conditioned on the observed data. In engineering applications this is commonly referred to as Bayesian model updating (BMU) and related to the contribution of Beck and Katafygiotis [33].

Building on recent advancements in BMU for engineering applications, we extend these methods to the biofilm growth model introduced by Klemp et al. [18]. In such systems, it is essential to account for inherent variability caused by biological randomness and stochastic environmental influences. This intrinsic variability introduces

aleatory uncertainty, which significantly increases the complexity of stochastic model updating. For robust and reliable inference, it is crucial to represent both epistemic (due to limited knowledge) and aleatory (inherent randomness) uncertainties, collectively referred to as hybrid uncertainty [34, 35]. However, hybrid uncertainty poses major computational challenges in Bayesian inference, where propagating stochastic variability through complex models typically requires expensive double-loop procedures [32, 36, 37]. To address this, we employ a reduced-order model (ROM) based on the Time-Separated Stochastic Mechanics (TSM) methodology [38, 39]. TSM approximates the stochastic dynamics by expanding the model response with respect to the uncertain parameters around their expected values and solving a sequence of deterministic evolution equations for the expansion coefficients. This separation of temporal and stochastic components enables efficient forward simulation under aleatory uncertainty, eliminating the need for repeated sampling in time.

The aim of this work is to calibrate a continuum model of biofilm growth, influenced by nutrients and antibiotics, as introduced in [18]. We employ a Bayesian updating approach to infer key model parameters under hybrid uncertainties. We model our unknown parameters as parametric probability-boxes (p-boxes) with unknown mean values and fixed coefficients of variations, in order to model natural variability as aleatory uncertainty. Then, a TSM-ROM is derived for an efficient propagation of the aleatory uncertainties as to not use a double-loop of full-order model calls. We construct a likelihood function based on summary statistics, i.e., mean and variance, of model responses.

In this paper, we begin with a review of the theoretical background on Bayesian model updating, the treatment of hybrid uncertainties and TSM in section 2. Subsequently, in section 3, we present the application of TSM-ROM to the biofilm model. Finally, we demonstrate the efficacy of our methodology through two case studies, illustrating its accuracy and computational efficiency in section 4. Our results highlight the potential for broader applications in uncertainty-aware modeling of biological systems.

## 2 Background

### 2.1 Bayesian model updating

Stochastic model updating aims to refine a forward model  $\mathcal{M}(\boldsymbol{\theta})$  by inferring unknown parameters  $\boldsymbol{\theta}$  based on observed data [40]. Here,  $\boldsymbol{\theta}$  is a realization of the  $M$ -dimensional random variable  $\Theta$ , defined over a parameter space  $\mathcal{D}_\Theta \subset \mathbb{R}^M$ . The forward model thus defines a mapping from the parameter space to the output space  $\mathcal{M} : \mathcal{D}_\Theta \rightarrow \mathbb{R}^{N_{\text{out}}}$ .

To relate the model predictions  $\mathbf{y} = \mathcal{M}(\boldsymbol{\theta})$  to the observed data  $\mathbf{D}$ , a discrepancy term  $\boldsymbol{\varepsilon}$  is introduced:

$$\mathbf{D} = \mathcal{M}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}. \quad (1)$$

This term accounts for measurement noise and model inaccuracies, acknowledging that computational models are idealized approximations of reality [40]. A common assumption is that  $\boldsymbol{\varepsilon}$  follows a zero-mean Gaussian distribution with covariance  $\boldsymbol{\Sigma}_\varepsilon$ , i.e.,  $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$ .

The Bayesian model updating (BMU) approach [33] addresses challenges such as incomplete data, observation noise, and model-form uncertainty by treating the parameters  $\boldsymbol{\theta}$  as random variables. Prior knowledge about these parameters is encoded in a prior distribution  $p(\boldsymbol{\theta})$ .

A central component of Bayesian inference is the likelihood function  $\mathcal{L}(\boldsymbol{\theta}) = p(\mathbf{D}|\boldsymbol{\theta})$ , which quantifies the probability of observing the data  $\mathbf{D}$  for a given parameter realization  $\boldsymbol{\theta}$ . Essentially, the likelihood serves as a stochastic measure of fit between model predictions and observations. Its specific form depends on modeling assumptions for the discrepancy term. Under the Gaussian noise assumption, the likelihood is itself a Gaussian distribution centered at the model output and evaluated at the data [33, 40]. Its value increases as the data and model assumptions align more closely.

Bayes' theorem then combines the prior and likelihood to yield the posterior distribution  $p(\boldsymbol{\theta}|\mathbf{D})$  over the parameters:

$$p(\boldsymbol{\theta}|\mathbf{D}) = \frac{\mathcal{L}(\boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{D})}, \quad (2)$$

where  $p(\mathbf{D}) = \int \mathcal{L}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$  is the model evidence. Since this normalizing constant is independent of a fixed set of observations  $\mathbf{D}$  and often intractable, the unnormalized posterior is typically used in practice:

$$p(\boldsymbol{\theta}|\mathbf{D}) \propto \mathcal{L}(\boldsymbol{\theta}) p(\boldsymbol{\theta}). \quad (3)$$

Note that in eq. (2) and eq. (3) the densities are implicitly also conditioned on the model assumptions.

### 2.1.1 Bayesian model updating in the presence of hybrid uncertainties

The previously introduced Bayesian model updating approach treats unknown parameters as random variables with prior distributions that are refined into posterior distributions using observational data. This process only reduces epistemic uncertainty, assuming that all variability can be captured probabilistically. However, many real-world systems exhibit not only epistemic but also aleatory uncertainty: inherent randomness that cannot be reduced through further data collection. While epistemic uncertainty refers to fixed-but-unknown quantities, aleatory uncertainty reflects stochastic variation between experimental or simulation outcomes, even under identical conditions.

A comprehensive review of model updating under different uncertainty types is provided in [32], where parameters are categorized into four types based on the presence and combination of aleatory and epistemic uncertainty. An overview of these categories is illustrated in fig. 1.

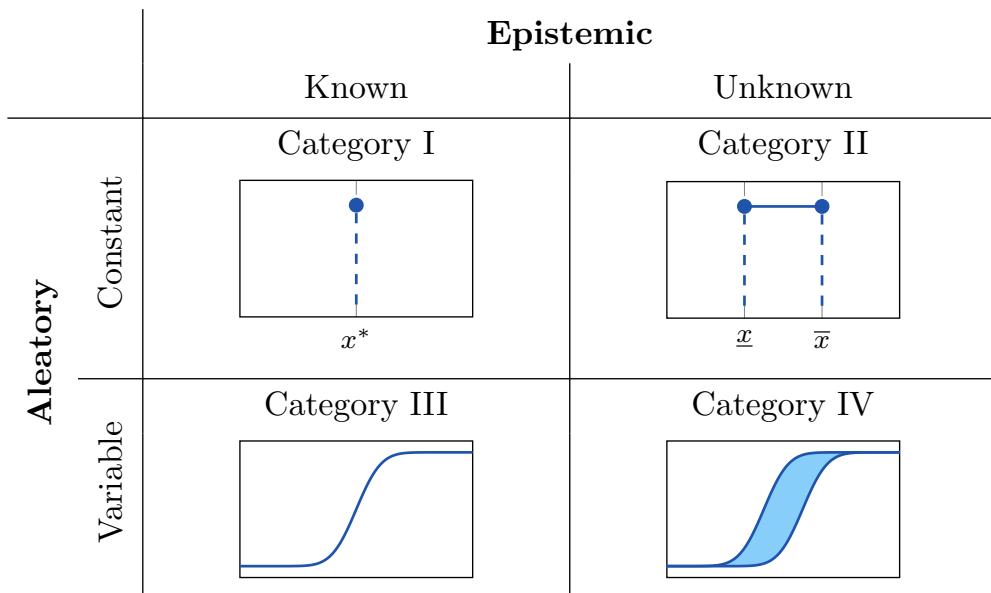
In this work, we focus on the specific challenges posed by parameters of category IV, which involve both types of uncertainty and are thus described by imprecise probabilities [35]. A common representation of such parameters is the probability box

(p-box), in which aleatory uncertainty is modeled as a random variable, while epistemic uncertainty is expressed as interval bounds on the distribution parameters.

In these cases, model outputs are inherently stochastic. A deterministic mapping like in eq. (1) is no longer sufficient, and commonly a stochastic forward model based on Monte Carlo simulation  $\hat{\mathcal{M}} : \mathcal{D}_{\Theta} \rightarrow \mathbb{R}^{N_{\text{out}} \times N_{\text{samples}}}$  is defined, where  $N_{\text{samples}}$  denotes the number of Monte Carlo samples used to propagate aleatory uncertainty for each parameter configuration  $\theta$ . Each input returns a sequence of outputs:

$$\mathcal{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{N_{\text{samples}}}\} = \hat{\mathcal{M}}(\theta). \quad (4)$$

This setup typically results in a nested (double-loop) approach: an outer loop explores different realizations of  $\theta$ , while the inner loop samples the stochastic model output using the deterministic model  $\mathcal{M}$ . Examples for this are the updating approaches in Refs. [32, 36, 37]. Although this allows the construction of p-boxes for the model response, it comes at a high computational cost, particularly in Bayesian inverse problems, which require many model evaluations across the parameter space.



**Fig. 1:** Parameters with different combinations of aleatory and epistemic uncertainties.

In the presence of hybrid uncertainties, constructing the likelihood function becomes especially challenging, since model outputs are no longer deterministic but represent probability distributions. In Approximate Bayesian Computation (ABC), for example, model predictions are compared with observations using distance metrics to approximate a likelihood function [32, 41, 42]. Alternatively, summary statistics such as the mean and variance can be used to define the likelihood function.

### 2.1.2 Formulation of the likelihood function

Given the deterministic model in eq. (1) and a corresponding observation  $\mathbf{D}$ , and assuming Gaussian-distributed measurement noise, the likelihood function can be written as:

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^{N_{\text{out}}} \det \boldsymbol{\Sigma}_{\epsilon}}} \exp \left[ -\frac{1}{2} (\mathbf{D} - \mathbf{y})^\top \boldsymbol{\Sigma}_{\epsilon}^{-1} (\mathbf{D} - \mathbf{y}) \right]. \quad (5)$$

Note that eq. (5) is relating a single multi-variate observation  $\mathbf{D}$  to a single deterministic model output  $\mathbf{y}$ , where  $\boldsymbol{\Sigma}_{\epsilon}$  describes the expected spread due to the additive errors. This can be extended to the stochastic model  $\hat{\mathcal{M}}$  by using for example the predicted mean  $\boldsymbol{\mu}_{\mathbf{y}}$  and covariance  $\boldsymbol{\Sigma}_{\mathbf{y}}$  of the stochastic predictions in eq. (4):

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^{N_{\text{out}}} \det \boldsymbol{\Sigma}_{\mathbf{y}}}} \exp \left[ -\frac{1}{2} (\mathbf{D} - \boldsymbol{\mu}_{\mathbf{y}})^\top \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} (\mathbf{D} - \boldsymbol{\mu}_{\mathbf{y}}) \right]. \quad (6)$$

Since the likelihood typically has limited support within the parameter space  $\mathcal{D}_{\boldsymbol{\Theta}}$ , it is convenient to consider its logarithmic form to improve numerical stability:

$$\log \mathcal{L}(\boldsymbol{\theta}) = -\log \left( \sqrt{(2\pi)^{N_{\text{out}}} \det \boldsymbol{\Sigma}_{\mathbf{y}}} \right) - \frac{1}{2} (\mathbf{D} - \boldsymbol{\mu}_{\mathbf{y}})^\top \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} (\mathbf{D} - \boldsymbol{\mu}_{\mathbf{y}}). \quad (7)$$

This formulation extends naturally to multiple uncorrelated observations  $\mathbf{D}^k$ ,  $k \in \{1, \dots, m\}$ , by summing the log-likelihoods:

$$\log \mathcal{L}(\boldsymbol{\theta}) = \sum_k -\log \left( \sqrt{(2\pi)^{N_{\text{out}}} \det \boldsymbol{\Sigma}_{\mathbf{y}}^k} \right) - \frac{1}{2} (\mathbf{D}^k - \boldsymbol{\mu}_{\mathbf{y}}^k)^\top \boldsymbol{\Sigma}_{\mathbf{y}, k}^{-1} (\mathbf{D}^k - \boldsymbol{\mu}_{\mathbf{y}}^k). \quad (8)$$

Here, the measurements  $\mathbf{D}^k$  can correspond to either repeated observations of the same quantity or, as in this paper, values taken at different time steps. The quantities  $\boldsymbol{\mu}_{\mathbf{y}}^k$  and  $\boldsymbol{\Sigma}_{\mathbf{y}}^k$  thus denote the model-predicted mean and covariance at each measurement time step  $t_k$ . Note that in stochastic forward models, where Monte Carlo sampling is used, eq. (8) would need the mentioned double-loop approach, as estimating the means and covariances requires multiple samples. For this reason, we will implement a TSM-ROM to facilitate faster parameter inference. Details on the exact setup of the likelihood will be given in section 4.

### 2.1.3 Transitional Markov Chain Monte Carlo

In practice, the posterior distribution is analytically intractable, primarily due to its implicit dependence on the forward model through the likelihood function [40]. Moreover, when multiple measurements are incorporated, the posterior typically does not conform to any standard probability distribution, which becomes clear when considering how the log-likelihood function is constructed in eq. (8). To address this, Markov Chain Monte Carlo (MCMC) methods are widely used, as they enable sampling from

distributions of arbitrary shape. A key advantage of MCMC techniques is their ability to draw samples directly from the unnormalized posterior distribution, as given in eq. (3).

The most common MCMC algorithm is the Metropolis-Hastings algorithm [43, 44] which generates samples from the unnormalized posterior by starting a random walk algorithm with Markov chains from some initial samples and then accepting or rejecting new samples based on a proposal distribution.

In this paper, we apply the Transitional Markov Chain Monte Carlo (TMCMC) algorithm [45] to draw samples from eq. (3). TMCMC is an advanced MCMC method designed to sample from multimodal distributions by sampling from a sequence of intermediate distributions rather than from the posterior directly, a process which is also known as annealing [46]. These transitional densities are defined as:

$$p^j(\boldsymbol{\theta}) \propto \mathcal{L}(\boldsymbol{\theta})^{\beta_j} p(\boldsymbol{\theta}) \quad (9)$$

where  $j \in \{1, 2, \dots, m\}$  represents the transition steps, with the corresponding tempering parameter  $\beta_j$  progressing from  $\beta_0 = 0$  to  $\beta_m = 1$  through intermediate values  $\beta_0 = 0 < \beta_1 < \beta_2 < \dots < \beta_m = 1$ . This gradual transition allows the prior density  $p(\boldsymbol{\theta})$  to smoothly evolve into the posterior density  $\mathcal{L}(\boldsymbol{\theta})p(\boldsymbol{\theta})$ , as noted in Ref. [40]. To make use of the formulation of the log-likelihood in eq. (8), the transitional densities in eq. (9) are also formulated in a logarithmic expression. We chose TMCMC for its robustness against posteriors with small support due to the annealing procedure.

## 2.2 Time-separated Stochastic Mechanics

In order to reduce the high computational cost of the double-loop approach, a surrogate model or ROM can be used to replace the inner loop and thus decrease the computational cost, as also mentioned in [32]. Faes et al. [47] discuss different surrogate modeling techniques for the propagation of hybrid uncertainties modeled with p-boxes. Further, Reiser et al. [48] present a two-step Bayesian framework for surrogate-based inference that propagates both epistemic and aleatoric uncertainties from surrogate model training to parameter inference.

Here, we use the Time-separated Stochastic Mechanics (TSM) [38, 39] to efficiently handle the intrinsic aleatory uncertainty and reduce the updating from a double to a single-loop algorithm. The main idea of the TSM is to replace a forward model  $\mathcal{M}(\boldsymbol{\theta})$  with the parameters  $\boldsymbol{\theta}$  by a surrogate  $\mathcal{M}_S$ . The surrogate is defined such, that the first  $p$  partial derivatives of the surrogate coincide with the forward model at the expectation of the parameters. For a scalar parameter  $\theta$  with expectation  $\langle \theta \rangle$  this reads as

$$\frac{\partial^i}{\partial \theta^i} \mathcal{M} \Big|_{\langle \theta \rangle} = \frac{\partial^i}{\partial \theta^i} \mathcal{M}_S \Big|_{\langle \theta \rangle} \quad \forall i \in \{0, \dots, p\}. \quad (10)$$

For an algebraic model, this coincides with a Taylor series. Forward models involving derivatives in time need a special treatment. This is presented in more detail in Section 3.2. The approach is advantageous for the application to Bayesian Model Updating as the approximation is best near to the expectation of the parameters.

Often, much of the probability mass is indeed collected around the expectation. In comparison to other surrogate models, as the Polynomial Chaos Expansion and Stochastic Collocation Method, only a very limited number of function evaluations needed. In fact, for many problems a linear or quadratic approximation suffices.

### 3 Application to biofilm growth

In this paper, we apply Bayesian model updating to identify parameters in the evolution of biofilm growth of a model introduced by Klempt et al. [18]. This model incorporates the growth of multiple species under the influence of nutrients and antibiotics, as well as their interactions. The growth is represented by the concentration of the biofilms over time, denoted by  $\phi$ . Additionally,  $\psi$  is defined as the percentage of living bacteria. The volume occupied by living bacteria from species  $l$  is expressed as  $\bar{\phi}_l = \phi_l \psi_l$ . The model derives from the extended Hamilton principle [49], which leads for the special case of a local, quasi-static, isothermal model with no external forces to

$$\frac{\partial \Psi}{\partial \xi} + \frac{\partial \Delta^s}{\partial \dot{\xi}} + \frac{\partial c}{\partial \xi} = 0, \quad (11)$$

with the vector of internal variables

$$\xi = \begin{pmatrix} \phi \\ \psi \end{pmatrix}, \quad (12)$$

the energy density function  $\Psi$ , the dissipation function  $\Delta^s$  and the constraint function  $c$ .

The temporal evolution of biofilm concentration is determined by the energy density function and the dissipation function. The energy density function is defined as

$$\Psi = -\frac{1}{2} c^* \bar{\phi} \cdot \mathbf{A} \cdot \bar{\phi} + \frac{1}{2} \alpha^* \psi \cdot \mathbf{B} \cdot \psi \quad (13)$$

and consists of two terms. The first term, where  $c^*$  represents nutrients, which promote an increase in living bacteria, while the second term, where  $\alpha^*$  signifies antibiotics, results in a decrease.

Coefficient matrices  $\mathbf{A}$  and  $\mathbf{B}$  are crucial in characterizing the material behavior of the species. The matrix  $\mathbf{A}$  is a symmetric matrix designed to capture the interactions between different biofilm species and the effects of nutrients on their growth. Its off-diagonal elements represent inter-species interactions, while the diagonal elements account for the effects of nutrients on individual species growth. Generally,  $\mathbf{A}$  can be expressed for multiple species as

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}. \quad (14)$$

The matrix  $\mathbf{B}$  is a diagonal matrix that characterizes the impact of antibiotics on the viability of the biofilm species. Its diagonal components represent the susceptibility of each species to antibiotics. The general form of  $\mathbf{B}$  for multiple species is assumed to be

$$\mathbf{B} = \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix}. \quad (15)$$

The dissipation function is modeled as a function of  $\dot{\bar{\phi}}$  and  $\dot{\phi}$ :

$$\Delta^s = \Delta^s(\dot{\bar{\phi}}, \dot{\phi}) = \frac{1}{2} \dot{\bar{\phi}} \cdot \boldsymbol{\eta} \cdot \dot{\bar{\phi}} + \frac{1}{2} \dot{\phi} \cdot \boldsymbol{\eta} \cdot \dot{\phi} \quad (16)$$

with the diagonal viscosity matrix

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 & 0 & \cdots & 0 \\ 0 & \eta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_n \end{pmatrix}. \quad (17)$$

The choice of modeling the dissipation as a function of the rate of living bacteria  $\dot{\bar{\phi}}$  and not directly of the rate of the state variables  $\dot{\phi}$  and  $\dot{\psi}$  leads to a deeply linked system of equation, resulting in complex model behavior.

To limited the growth to a finite amount, the total volume of all species is limited by the constraint function

$$c = \gamma \left( \sum_{l=0}^n \phi_l - 1 \right) = 0 \quad (18)$$

with the Lagrange multiplicator  $\gamma$ . This leads to the governing evolution equations in their strong forms:

$$0 = -c^* \psi_i \left( a_{ii} \bar{\phi}_i + \sum_j^{n-1} a_{ij} \bar{\phi}_j \right) + \eta_i (\dot{\phi}_i \psi_i^2 + \bar{\phi}_i \dot{\psi}_i + \dot{\phi}_i) + \gamma \quad (19)$$

$$0 = -c^* \phi_i \left( a_{ii} \bar{\phi}_i + \sum_j^{n-1} a_{ij} \bar{\phi}_j \right) + \alpha^* \psi_i b_i + \eta_i (\dot{\psi}_i \phi_i^2 + \bar{\phi}_i \dot{\phi}_i) + \gamma \quad (20)$$

$$0 = \sum_{l=0}^n \phi_l - 1 \quad (21)$$

In the following, the parameters in  $\mathbf{A}$  and  $\mathbf{B}$ , which describe the behavior of the species are referred to as the unknown, stochastic parameters  $\boldsymbol{\theta}$ . The observations  $\mathbf{D}$  are the volume percentage for each species  $\phi_l$  and the percentage of living cells for each species  $\psi_l$  for each time step.

To ensure, that the internal variables remained in the interval  $\xi_i \in [0, 1]$ , the penalty method was applied. The penalty terms  $K_p(\frac{1}{\xi_1^2(1-\xi_1)^2})$  were added for all internal variables to the free energy density in [Equation 13](#).

### 3.1 Uncertainty model for biofilm growth

For the uncertainty model involved in the parameter modeling of the biofilm growth, we treat these parameters  $\boldsymbol{\theta}$  as category IV, specifically using p-boxes. This complex setup with mixed uncertainties is designed to replicate the uncertainties in real-world biofilm growth due to both inherent randomness (aleatory uncertainty) and data scarcity (epistemic uncertainty). The hybrid nature of these uncertainties justifies the application of a ROM using TSM to more efficiently handle both types of uncertainties in inverse problems.

Based on this reasoning, the biofilm growth model is implemented in a stochastic manner, as shown in eq. (4). The TSM-ROM replaces the Monte Carlo simulation  $\hat{M}$  and returns a sequence of outputs for a given input  $\boldsymbol{\theta}$ . The sequence of outputs at a given time instance is stochastic due to the aleatoric uncertainty inherent to biofilm growth. This is true even if the epistemic uncertainty is reduced to zero. The epistemic uncertainty is connected to the mean value of  $\boldsymbol{\theta}$ , whose exact distribution is unknown a priori and is inferred via Bayesian updating.

### 3.2 TSM-ROM for biofilm growth

Equations (19), (20) and (21) constitute a nonlinear system of equations for the evolution of the variables  $\psi$  and  $\phi$ . The input parameter  $\boldsymbol{\theta}$  is split into an part with epistemic uncertainty  $\boldsymbol{\theta}^{(0)}$  and a part with aleatoric uncertainties  $\tilde{\boldsymbol{\theta}}$  with the expectation  $\langle \tilde{\boldsymbol{\theta}} \rangle = \mathbf{0}$ . In the following, the input parameters are referred to as  $\boldsymbol{\theta}^{\text{TSM}} = \boldsymbol{\theta}^{(0)} + \tilde{\boldsymbol{\theta}}$ . The variables  $\psi$  and  $\phi$  are expressed by a linear Taylor series in the aleatoric uncertainties

$$\psi^{\text{TSM}} = \psi^{(0)} + \tilde{\boldsymbol{\theta}} \cdot \psi^{(1)} \quad (22)$$

$$\phi^{\text{TSM}} = \phi^{(0)} + \tilde{\boldsymbol{\theta}} \cdot \phi^{(1)}. \quad (23)$$

The Taylor series for the variables  $\psi$  and  $\phi$  constitute the surrogate model  $\mathcal{M}_S$ . No solution of the systems of equations is needed for a new parameter value  $\boldsymbol{\theta}$ . While a higher order model is simple to implement, a linear series already often suffices [39]. For simplicity, let us refer to the system of equations (19) to (21) as  $\mathcal{G}(\dot{\phi}, \dot{\psi}, \phi, \psi, \boldsymbol{\theta}^{\text{TSM}})$ . The system of equations naturally depends on both, the variables themselves and their time derivatives. The time derivative of the Taylor series results trivially as

$$\dot{\psi}^{\text{TSM}} = \dot{\psi}^{(0)} + \tilde{\boldsymbol{\theta}} \cdot \dot{\psi}^{(1)} \quad (24)$$

$$\dot{\phi}^{\text{TSM}} = \dot{\phi}^{(0)} + \tilde{\boldsymbol{\theta}} \cdot \dot{\phi}^{(1)}. \quad (25)$$

because the random term  $\tilde{\boldsymbol{\theta}}$  is time-independent. All Taylor series are set into the system of equations. The zeroth order terms can be calculated as

$$\mathcal{G}(\dot{\phi}^{\text{TSM}}, \dot{\psi}^{\text{TSM}}, \phi^{\text{TSM}}, \psi^{\text{TSM}}, \boldsymbol{\theta}^{\text{TSM}})|_{\tilde{\boldsymbol{\theta}}=\mathbf{0}}. \quad (26)$$

Unsurprisingly, the original system of equations evaluated at  $\boldsymbol{\theta}^{\text{TSM}} = \mathbf{0}$  results. For the derivation of the first order term,

$$\frac{d}{d\tilde{\boldsymbol{\theta}}} \mathcal{G}(\dot{\boldsymbol{\phi}}^{\text{TSM}}, \dot{\boldsymbol{\psi}}^{\text{TSM}}, \boldsymbol{\phi}^{\text{TSM}}, \boldsymbol{\psi}^{\text{TSM}}, \boldsymbol{\theta}^{\text{TSM}}) \Big|_{\tilde{\boldsymbol{\theta}}=\mathbf{0}}. \quad (27)$$

is calculated. This results in a nonlinear system of equations that allow to compute  $\dot{\boldsymbol{\phi}}^{(1)}$  and  $\dot{\boldsymbol{\psi}}^{(1)}$ . As these equations depend on  $\dot{\boldsymbol{\phi}}^{(0)}$  and  $\dot{\boldsymbol{\psi}}^{(0)}$ , a staggered solution scheme naturally arises.

The model and its updating procedure are implemented in the open-source programming language *Julia*. Performance tests for a single deterministic model prediction and for the TSM-ROM implementation indicate that one model run with two species requires approximately 3 ms, whereas solving the complete TSM-ROM takes about 40 ms on the same hardware. It should be noted that, to fully represent aleatory uncertainties, MC simulations require multiple, often hundreds of model runs. In contrast, the TSM-ROM computes stochastic moments directly, making it considerably more efficient for mixed-uncertainty models and well-suited for multi-query applications.

## 4 Numerical Experiments

To illustrate the application of BMU in determining the *material properties* governing biofilm growth, we present two case studies. The primary objective in both cases is to infer the matrices  $\mathbf{A}$  and  $\mathbf{B}$ , which characterize biofilm evolution through the energy density function outlined in eq. (13).

The first case study involves a model with two species and five unknown material parameters. The second case study extends this framework to four species with 14 parameters. In both scenarios, the BMU approach utilizes TMCMC, implemented via the Julia package `UncertaintyQuantification.jl` [50]. To compare model predictions with experimental data, we employ the likelihood function given in eq. (8), assessing discrepancies in the volume fractions of living bacteria, denoted by  $\bar{\Phi}$ , at predetermined discrete time steps.

To build the likelihood function, the response of the TSM biofilm model is compared to experimental data. Specifically, we consider the mean response feature of the  $l$ -th species at the  $k$ -th time step, denoted as  $\mu_{\mathbf{y},l}^k$ . Similarly, its variance is denoted as  $\sigma_{\mathbf{y},l}^k$ . In this context, the output features are defined as the volume fractions of  $n$  living bacterial species at various discrete time steps  $k$ . For a single sample  $\boldsymbol{\theta}_j$ , i.e. one full simulation of the TSM-ROM model, these features are structured in a matrix form:

$$\mathbf{y}_j = \begin{pmatrix} \bar{\phi}_1^1 & \bar{\phi}_1^2 & \cdots & \bar{\phi}_1^m \\ \bar{\phi}_2^1 & \bar{\phi}_2^2 & \cdots & \bar{\phi}_2^m \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\phi}_n^1 & \bar{\phi}_n^2 & \cdots & \bar{\phi}_n^m \end{pmatrix}, \quad (28)$$

where  $\bar{\phi}_l^k$  denotes the volume fraction of species  $l$  at time step  $k$ . Here,  $n$  represents the total number of species, while  $m$  is the number of selected discrete time steps used

for comparison with observational data. For better readability, we omitted the index  $j$  in  $\bar{\phi}_l^k$ .

It is important to note that  $m \ll N$ , where  $N$  denotes the total number of time steps in the complete analysis. Typically, a large number of steps such as  $N = 1000$  is used when performing the time integration of the TSM model. This is chosen to ensure numerical stability and accuracy, as a smaller time step size  $\Delta t$  helps to control numerical errors during the calculation.

While  $N$  allows for detailed modeling of the biofilm dynamics over time,  $m$  is strategically chosen to facilitate the construction of the likelihood function by providing a manageable subset of time steps. For instance, one might select  $m = 20$  time steps for comparing simulation responses with experimental data. This selection ensures efficient and meaningful statistical inference without overwhelming computational resources. Thus, a dataset  $\mathbf{D}$  is constructed from  $N_{\text{data}} = m$  unique experiments, each representing a realization that terminates at different time steps. Our aim is to replicate the setup of *in vitro* experiments in our *in silico* experiments, i.e., simulating multiple experiments started from the same initial conditions, but stopped at different time instances. This approach accounts for both aleatory uncertainty, arising from inherent randomness in biofilm growth, and epistemic uncertainty due to limited observational data.

Finally, the likelihood function from eq. (8) is computed for each row within the feature matrix detailed in eq. (28), as described above. In preliminary testing we found that using the full covariance matrix is unnecessary, as correlations between the model outputs are implicitly handled by the model. Therefore, the likelihood is reduced to only the diagonal elements of the covariance matrix, and the expression in eq. (7) can be simplified. The full likelihood, constructed from the  $l \in \{1, \dots, n\}$  species and  $k \in \{1, \dots, m\}$  time steps, thus reads

$$\log \mathcal{L}(\boldsymbol{\theta}) = \sum_l \sum_k -\log \left[ 2\pi \sigma_{\mathbf{y},l}^k \right] - \frac{1}{2} \sigma_{\mathbf{y},l}^{k,-2} \left( \bar{\phi}_l^k - \mu_{\mathbf{y},l}^k \right)^2. \quad (29)$$

where  $\sigma_{\mathbf{y},l}^k$  is the  $i$ -th diagonal element of  $\Sigma_{\mathbf{y}}^k$  at time step  $t_k$ . Further comment will be given below.

#### 4.1 Case I: Two-Species Biofilm Model

In this first case study, we focus on biofilm growth comprising  $n = 2$  interacting species. Specifically, this case aligns with case 5 in Ref. [18], augmented by the addition of an interaction parameter  $a_{12}$  to adeptly capture inter-species interactions.

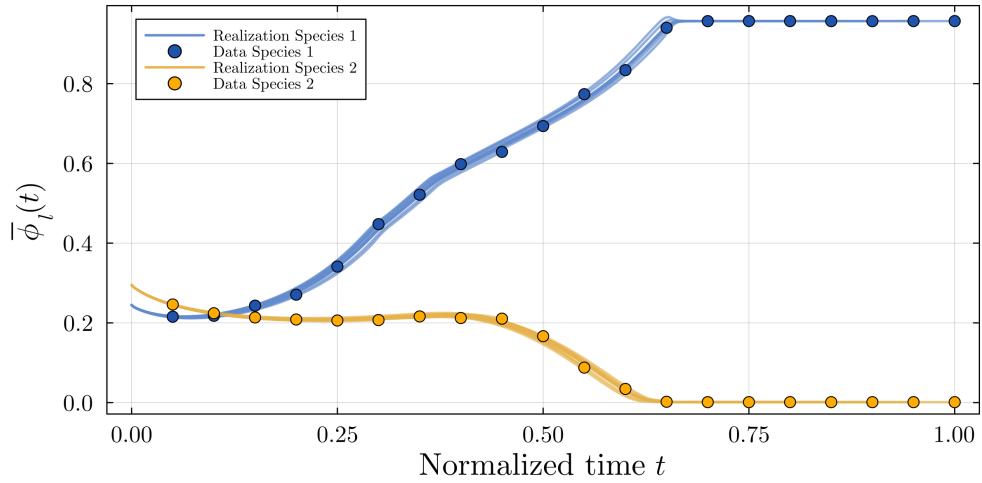
For the two-species scenario, the material behavior of the biofilm is characterized by the two matrices in eqs. (14) and (15) with  $n = 2$ . The symmetry inherent in matrix  $\mathbf{A}$  reduces the dimensionality of the problem to five independent parameters  $\boldsymbol{\theta} = [a_{11}, a_{12}, a_{22}, b_1, b_2]$ . Each of the five parameters is modeled as a parametric p-box with a Normal distribution with an unknown mean and a coefficient of variation of 0.5%.

The dataset  $\mathbf{D}$  comprises  $N_{\text{data}} = 20$  individual simulations, each concluding at different time steps, as shown in fig. 2. These data points correspond to model outputs generated from 20 parameter samples  $\boldsymbol{\theta}$ , randomly selected from a larger set of 1000

samples drawn from the Normal distributions of the respective parameters with *true* means.

The solid lines represent the time evolution of the 20 samples of volume fractions of living bacteria for both species,  $\bar{\phi}_1$  and  $\bar{\phi}_2$ . The scatters present the 20 randomly selected realizations at  $m = 20$  evenly spaced time steps within the interval  $t \in [0, N = 1000]$ . Since the initial conditions at  $t = 0$  are known, the first time step is chosen at  $t = 50$ , as depicted in fig. 2. It is important to highlight that the dataset  $\mathcal{D}$  only consist of the 20 discrete data points. Further, we note that the time steps were chosen as a balance between accuracy by using a large number of time instances to compare the data and realism by using as little data as possible.

The different constant simulation parameters used in case I, e.g., the viscosity  $\eta$ , different initial conditions, nutrients and antibiotics, are summarized in table 1. Further, for all the following cases, we use a penalty term  $K_p = 10^{-4}$  and initial  $\psi_i = 0.999$  for all species. Similarly, table 2 shows the chosen prior distributions of the mean values of the five parameters that are inferred with the proposed BMU approach. Here, uninformative priors modeled by uniform distributions are chosen.



**Fig. 2:** Dataset with  $N_{\text{data}} = 20$  volume fractions of living bacteria for two species,  $\bar{\phi}_1(t)$  and  $\bar{\phi}_2(t)$ . Individual realizations are shown as yellow and blue lines, each ending at different steps, indicated by the dots. The data is generated from 1000 samples of the underlying true distribution of parameters with mean values  $\boldsymbol{\theta}^* = [1, 0.1, 1, 1, 2]$  along with a CoV of 0.5%.

The Bayesian model updating with the described procedure is performed to calibrate the model parameters. The resulting samples of the posterior are visualized in fig. 3. We observe that the posterior bounds are much tighter compared to the priors, indicating a significant reduction in uncertainty due to the incorporation of data. The posterior distributions generally peak around the “true” parameter values used

**Table 1:** Values of simulation parameters for case I.

	Variable	Unit	Value
viscosity	$\eta_1$	[ $\frac{\text{kg}}{\text{ms}}$ ]	1
viscosity	$\eta_2$	[ $\frac{\text{kg}}{\text{ms}}$ ]	2
initial	$\phi_1$	[ $\cdot$ ]	0.25
initial	$\phi_2$	[ $\cdot$ ]	0.30
nutrients	$c^*$	[ $\frac{\text{m}^2}{\text{s}^2}$ ]	100
antibiotics	$\alpha^*$	[ $\frac{\text{m}^2}{\text{s}^2}$ ]	10
number of time steps	$N$	[ $\cdot$ ]	1000
time step size	$\Delta t$	[s]	$10^{-4}$
number of data points	$N_{\text{data}}$	[ $\cdot$ ]	20
number of aleatory samples	$N_{\text{samples}}$	[ $\cdot$ ]	500
number of posterior samples	$N_{\text{posterior}}$	[ $\cdot$ ]	5000
coefficient of variation	CoV	[ $\%$ ]	0.5

**Table 2:** Uniform prior ranges  $\mathcal{U}(a, b)$  for the mean values of the parameters in case I.

Parameter	$\theta_1 = a_{11}$	$\theta_2 = a_{12}$	$\theta_3 = a_{22}$	$\theta_4 = b_1$	$\theta_5 = b_2$
Range	(0, 3)	(0, 0.5)	(0, 3)	(0, 3)	(0, 3)

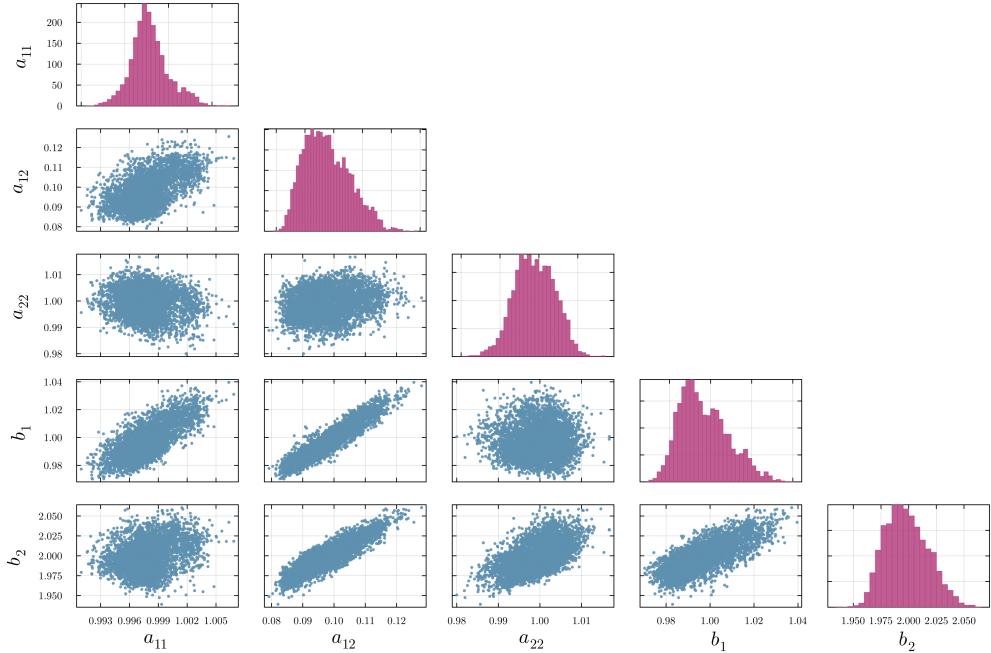
to construct the dataset, while still capturing some spread that reflects the inherent variability in the data. Specifically, the dataset was generated using mean values  $\boldsymbol{\theta}^* = [1, 0.1, 1, 1, 2]$  along with a CoV of 0.5%, and the results demonstrate that our Bayesian updating framework is capable of accurately resolving these values. Moreover, dependencies between parameters become apparent, both in the scatter plots and in the Pearson correlation coefficients  $\rho$  calculated from the posterior samples. For instance, we observe strong linear correlations between  $a_{12}$  and  $b_1$  ( $\rho = 0.934$ ), as well as between  $a_{12}$  and  $b_2$  ( $\rho = 0.898$ ). Moderate correlations are also evident between  $a_{11}$  and  $b_1$  ( $\rho = 0.725$ ), and between  $b_1$  and  $b_2$  ( $\rho = 0.759$ ). When considering the governing equations of the biofilm growth given by eqs. (19) to (21), these correlations seem reasonable as the respective parameters jointly influence the growth and degradation behavior of the biofilm.

Figure 4 visualizes the model responses corresponding to the posterior samples, alongside the data  $\mathbf{D}$  used for the updating. There, we show the responses for the two individual quantities  $\boldsymbol{\phi}$  and  $\boldsymbol{\psi}$  as well as the combined measure  $\bar{\boldsymbol{\phi}} = \boldsymbol{\phi}\boldsymbol{\psi}$ . It shall be noted that the updating was only performed using the output  $\bar{\boldsymbol{\phi}}$  and the corresponding data. We visualize  $\boldsymbol{\phi}$  and  $\boldsymbol{\psi}$  along with the respective data to validate the calibrated model. In general, we observe a good agreement between the model outputs and the data: the posterior-informed simulations can reproduce similar behavior as the observed dataset while appropriately handling the variability introduced by random realizations and differing termination times across trajectories. This holds true

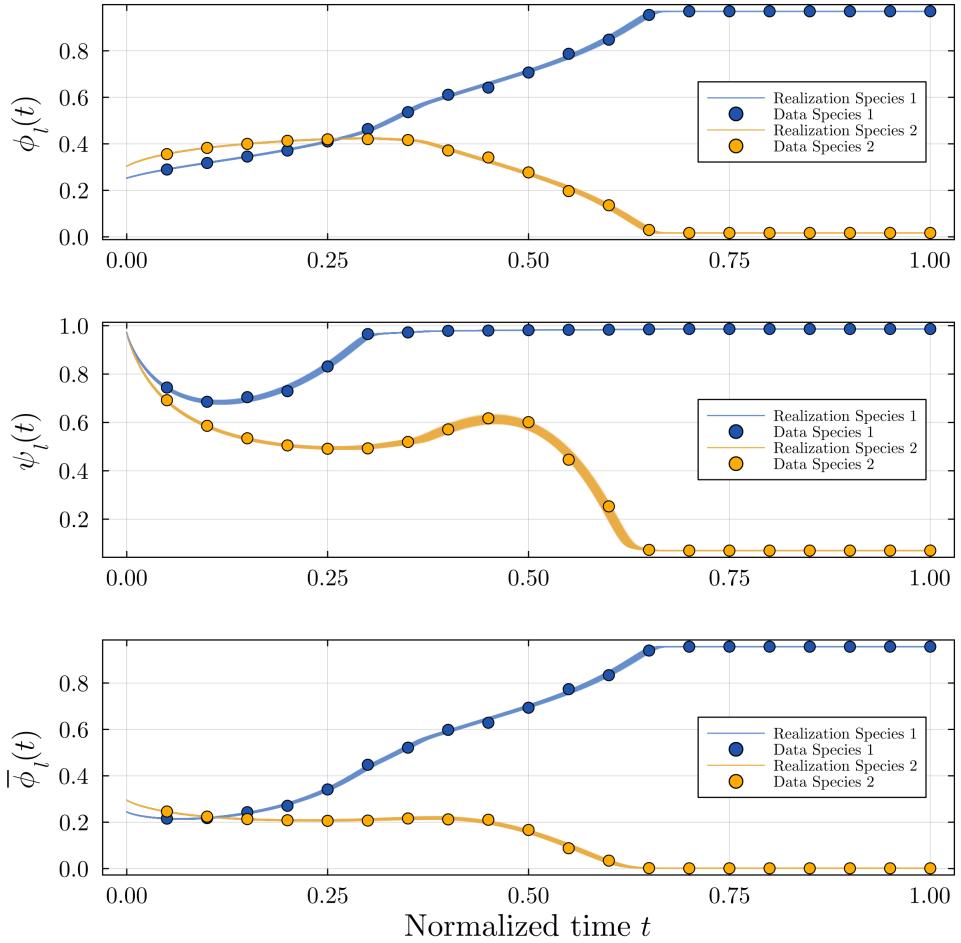
for all measures, also for the two quantities used for the validations which highlights the robustness of the parameter estimation.

The tight output prediction interval observed in fig. 4 is a contrast to fig. 5. There, the model response  $\bar{\phi}$  when sampling the parameters  $\theta$  from the prior distributions given in table 2 is visualized. It can be seen that the response for all possible parameters from the prior covers a wide range of solutions.

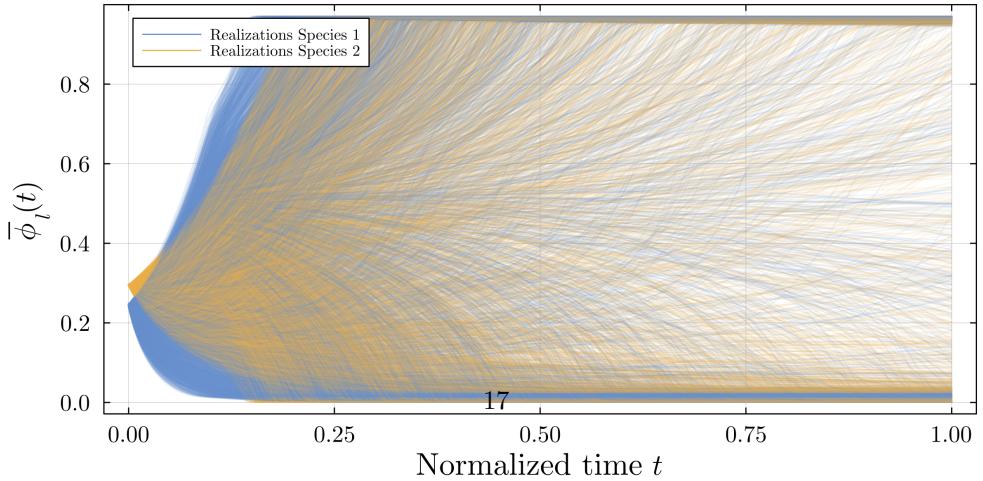
Lastly, to show that using only the diagonal elements of the estimated covariance matrix does not decrease the accuracy of the updating, fig. 6 shows the resulting p-boxes. Figure 6a shows results with only the diagonal elements, while fig. 6b shows results obtained from using the full covariance. Both results do not diverge from each other, and the resulting intervals are very similar. Thus, in the following, we will concentrate on only using the diagonal elements.



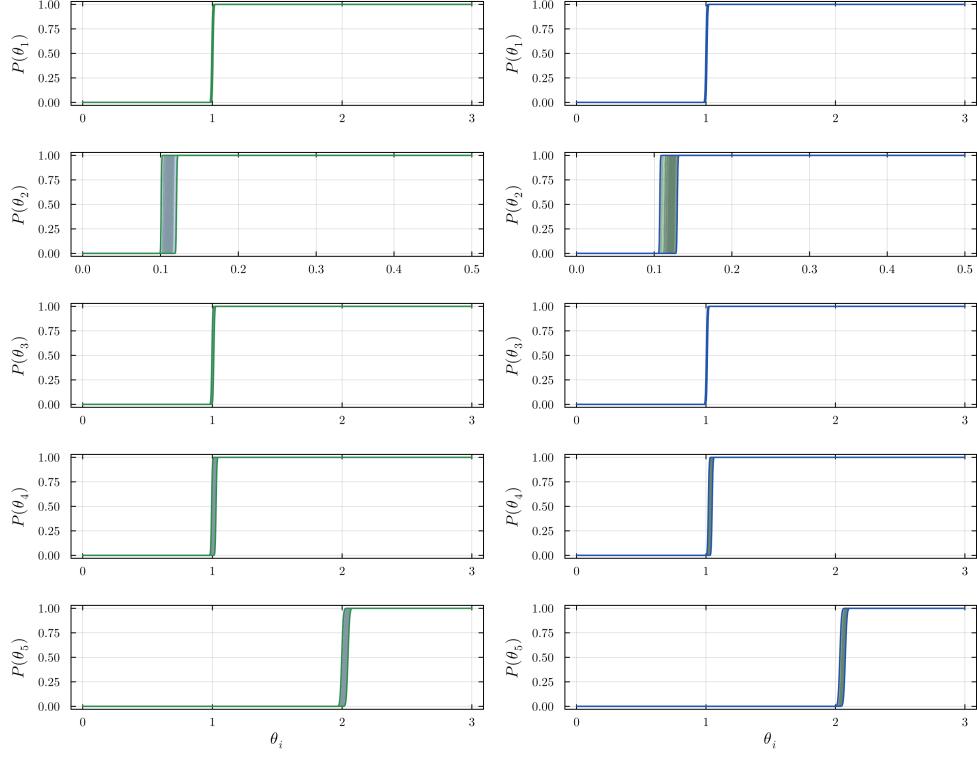
**Fig. 3:** Posterior samples of the mean values of the five material parameters  $\theta = [a_{11}, a_{12}, a_{22}, b_1, b_2]$  of case I.



**Fig. 4:** Model outputs  $\phi$ ,  $\psi$  and  $\bar{\phi}$  corresponding to the input given by the posterior samples of case I in fig. 3. Note: Only  $\bar{\phi}$  along with its associated data are used for the model calibration, the outputs  $\phi$  and  $\psi$  serve as a validation.



**Fig. 5:** Model realizations corresponding to the input given by the prior samples of case I from table 2.



(a) Diagonal covariance

(b) Full covariance

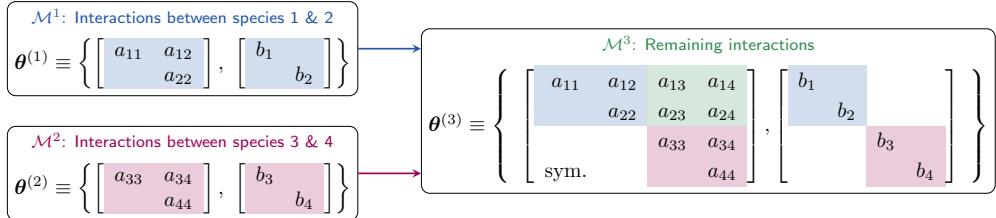
**Fig. 6:** Differences in the updated p-boxes between using the diagonal covariance (fig. 6a) and the full covariance (fig. 6b) in the likelihood

## 4.2 Case II: Four-Species Biofilm Model

As a second case study, we consider a biofilm model with  $n = 4$  interacting species and a total of 14 unknown parameters. To render Bayesian inference tractable, we employ a hierarchical (multilevel) updating strategy in three stages. First, we decompose the full four-species system into two simpler two-species submodels:  $\mathcal{M}^1$  captures the interactions of species 1 and 2 (five parameters), and  $\mathcal{M}^2$  captures the interactions of species 3 and 4 (five parameters). These submodels are calibrated *in parallel* using uninformative priors, yielding posterior distributions on the reduced parameter spaces  $\mathcal{D}_\Theta^1 \subset \mathbb{R}^5$  and  $\mathcal{D}_\Theta^2 \subset \mathbb{R}^5$ . A schematic of this multilevel updating is shown in fig. 7.

Next, in model  $\mathcal{M}^3$ , we assemble the full four-species interaction matrices  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{4 \times 4}$ , where the ten parameters already inferred in the submodels are highlighted (blue for  $\mathcal{M}^1$ , red for  $\mathcal{M}^2$ ). We now fix the corresponding ten parameters at their respective maximum-a-posteriori (MAP) estimates obtained from Steps 1 and 2. As a result, only the four remaining cross-block interaction parameters  $a_{13}, a_{14}, a_{23}, a_{24}$  are treated as uncertain in  $\mathcal{M}^3$  and assigned uninformative uniform priors, yielding

a reduced inference subspace  $\mathcal{D}_\Theta^3 \subset \mathbb{R}^4$ . A final Bayesian update is then performed solely over this four-dimensional subspace.



**Fig. 7:** Visualization of the hierarchical updating procedure on two different levels: Updating of two model  $\mathcal{M}^1$  and  $\mathcal{M}^2$  is performed using simpler two-species models, subsequently  $\mathcal{M}^3$  is used to update the remaining interactions.

In practice, Steps 1 and 2 follow the same two-species Bayesian updating procedure described in Case I, starting from uninformative priors. In Step 3, instead of reusing full posterior distributions, we carry forward only the MAP estimates from  $\mathcal{M}^1$  and  $\mathcal{M}^2$ , thereby focusing inference in  $\mathcal{M}^3$  entirely on the remaining four interaction parameters. This hierarchical structure reduces the dimensionality of each inference step—from 14 parameters in total to two problems of dimension 5, followed by one of dimension 4. Such a staged reduction in dimensionality is expected to improve computational efficiency and convergence, as also observed in [51], where the authors highlight the impact of parameter dimension on the performance of TMCMC algorithms.

Afterwards, we look at a modified setup of  $\mathcal{M}^3$  to validate the calibrated parameters with new data in a different setup. For this, we apply the antibiotics only after  $t = 0.5$  to check if our calibrated are model parameters are robust to this change in the setup. We denote this as the model  $\mathcal{M}_{\text{val}}^3$ . The selected simulation parameters of the submodels are summarized in table 3. All prior distributions are chosen as  $\mathcal{U}(0, 3)$ .

Values in table 3 vary between  $\mathcal{M}^1$ ,  $\mathcal{M}^2$  and  $\mathcal{M}^3$  since the different species have different sensitivities to nutrients and antibiotics. In  $\mathcal{M}^2$  the antibiotics were reduced to not have a zero-concentration of the bacteria, since this would lead to non-informative data. Moreover,  $\mathcal{M}^2$  has a longer experimental duration (5000 instead of 2500 time steps) due to slower growth of the microfilms. We simulated  $\mathcal{M}^1$  for a shorter duration because the concentrations became almost stationary and did not result in any further information gain.

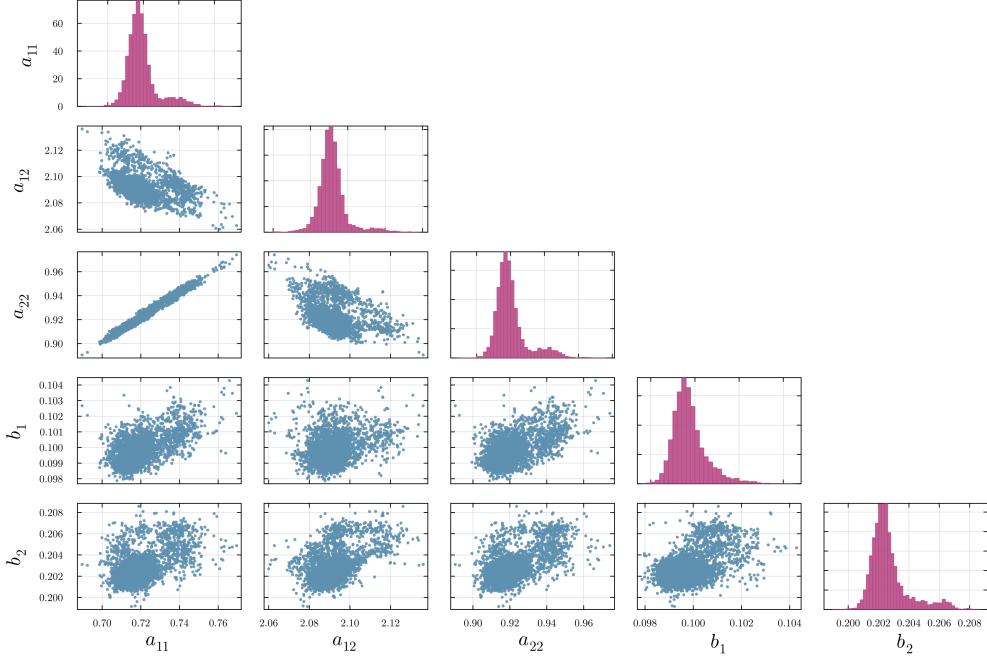
#### 4.2.1 Interaction of Species 1 and 2

First, the parameter set  $\theta^{(1)}$  is inferred using the two-species model  $\mathcal{M}^1$ . For this first submodel, the antibiotic parameter is set to  $\alpha = 100 \frac{\text{m}^2}{\text{s}^2}$ . We employ the same likelihood and same approach as in case I for the updating. The resulting posterior samples are shown in fig. 8. Here, again, for every parameter, a single peak along with some spread around that can be observed. Notably, we observe a strong correlation between the mean values of the parameters  $a_{11}$  and  $a_{22}$ .

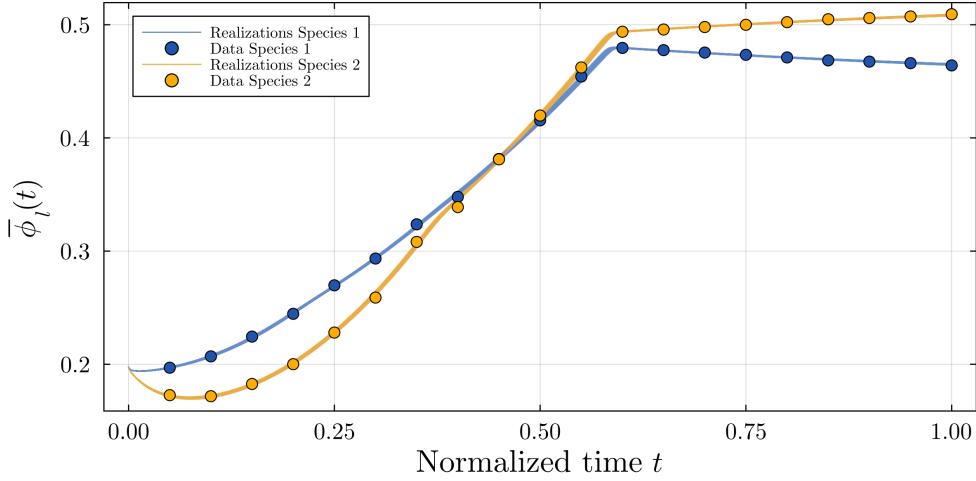
**Table 3:** Values of simulation parameters for the submodels of case II.

	Variable	Unit	$\mathcal{M}^1$	$\mathcal{M}^2$	$\mathcal{M}^3$	$\mathcal{M}_{\text{val}}^3$
viscosity	$\eta_1$	[ $\frac{\text{kg}}{\text{ms}}$ ]	1.0	1.0	1.0	1.0
viscosity	$\eta_2$	[ $\frac{\text{kg}}{\text{ms}}$ ]	1.0	1.0	1.0	1.0
viscosity	$\eta_3$	[ $\frac{\text{kg}}{\text{ms}}$ ]	-	-	1.0	1.0
viscosity	$\eta_4$	[ $\frac{\text{kg}}{\text{ms}}$ ]	-	-	1.0	1.0
initial	$\phi_1$	[ $\cdot$ ]	0.2	0.2	0.02	0.02
initial	$\phi_2$	[ $\cdot$ ]	0.2	0.2	0.02	0.02
initial	$\phi_3$	[ $\cdot$ ]	-	-	0.02	0.02
initial	$\phi_4$	[ $\cdot$ ]	-	-	0.02	0.02
nutrients	$c^*$	[ $\frac{\text{m}^2}{\text{s}^2}$ ]	100	100	25	25
antibiotics	$\alpha^*$	[ $\frac{\text{m}^2}{\text{s}^2}$ ]	100	10	0	50 $\mathbb{I}[t > 500]$
number of time steps	$N$	[ $\cdot$ ]	2500	5000	750	1500
time step size	$\Delta t$	[ $\text{s}$ ]	$10^{-5}$	$10^{-5}$	$10^{-4}$	$10^{-4}$
number of data points	$N_{\text{data}}$	[ $\cdot$ ]	20	20	20	20
number of aleatory samples	$N_{\text{samples}}$	[ $\cdot$ ]	500	500	500	500
number of posterior samples	$N_{\text{posterior}}$	[ $\cdot$ ]	5000	5000	5000	-
coefficient of variation	CoV	[ $\%$ ]	0.5	0.5	0.5	0.5

Figure 9 visualizes the model output corresponding to the posterior samples along with the data points used for the updating. Good agreement between the model response and the data can be observed.



**Fig. 8:** Posterior samples of the mean values of the five parameter in the set  $\theta^{(1)}$  updated with the two-species model  $\mathcal{M}^1$ .



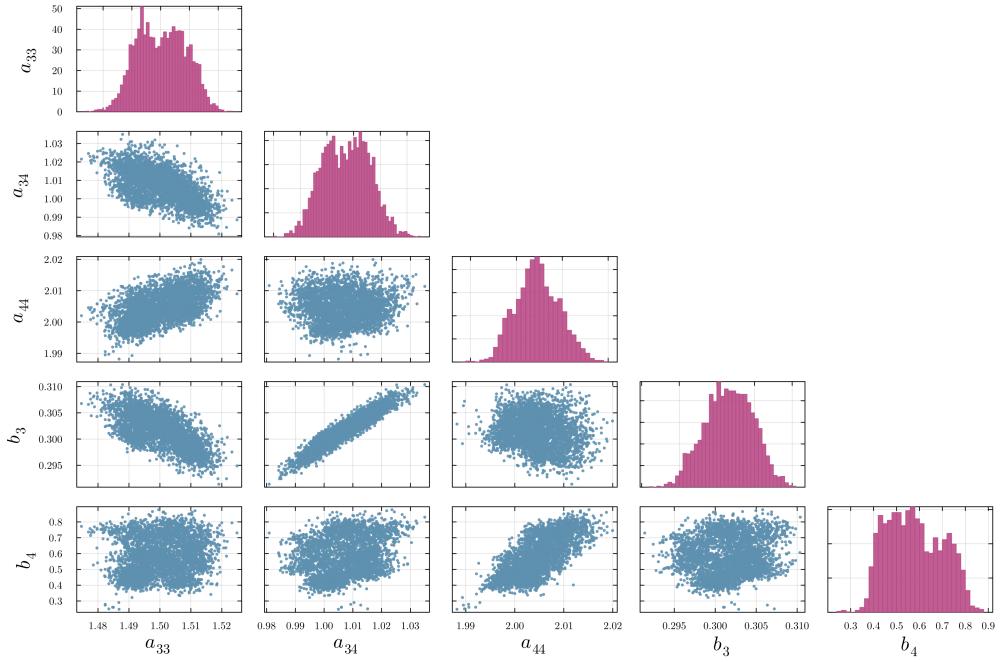
**Fig. 9:** Comparison of the model output of model  $\mathcal{M}^1$  corresponding to calibrated posterior samples (shaded) and the data (scatter).

#### 4.2.2 Interaction of Species 3 and 4

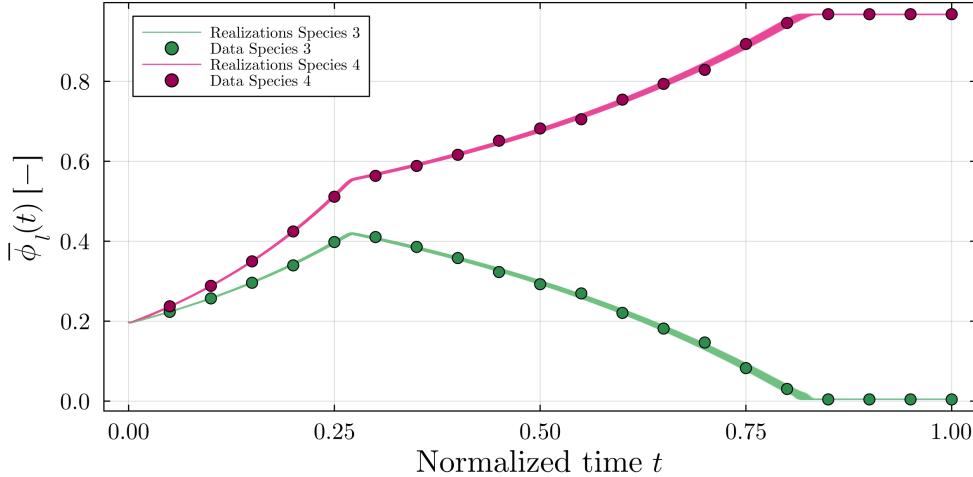
The same approach is applied to infer the mean values of the parameter set  $\theta^{(2)}$  with the second two-species model,  $\mathcal{M}^2$ . The results of the model calibration are shown in figs. 10 and 11, which again show the posterior samples and model outputs, respectively. Here, an antibiotic concentration of  $\alpha = 10 \frac{\text{m}^2}{\text{s}^2}$  is applied to build the data set and perform the updating.

It can be observed that the peak are not as sharp as in the case of the first parameter set. However, for all parameters but  $b_4$ , the spread around the peak is still small where comparing the ranges of the posterior samples to the prior range, i.e.,  $[0, 3]$ . Only for the antibiotic sensitivity of the forth species,  $b_4$ , the range of posterior samples is rather large.

The model output in fig. 11, however, shows very good agreement with the data points that are used for the Bayesian updating of the second parameter set.



**Fig. 10:** Posterior samples of the mean values of the five parameter in the set  $\theta^{(2)}$  updated with the two-species model  $\mathcal{M}^2$ .

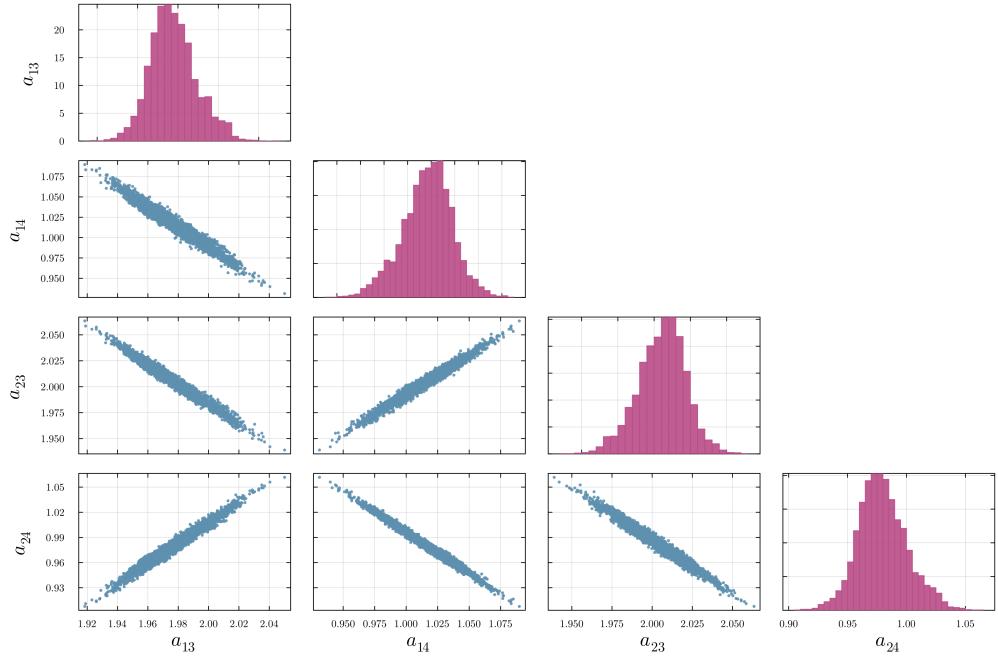


**Fig. 11:** Comparison of the model output of model  $\mathcal{M}^2$  corresponding to calibrated posterior samples (shaded) and the data (scatter).

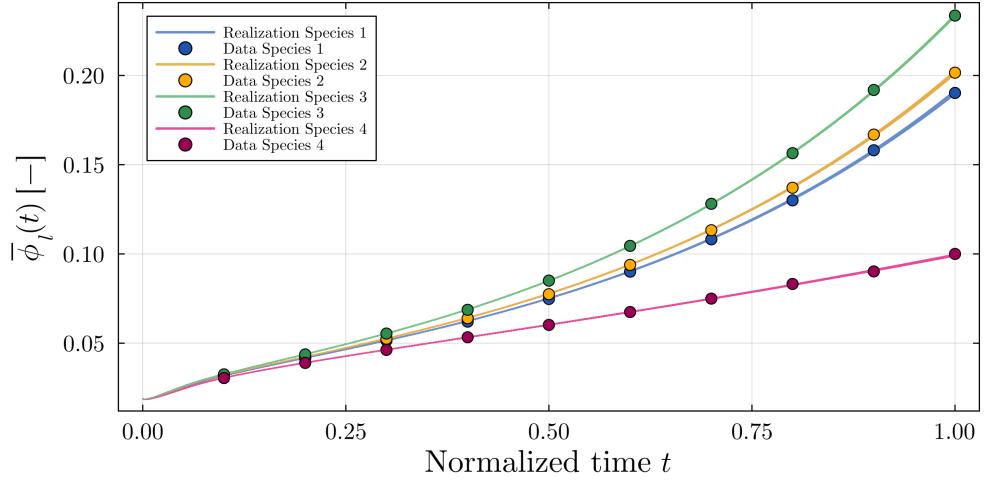
#### 4.2.3 Remaining Interactions

After the first two parameter sets are inferred using the two submodels  $\mathcal{M}^1$  and  $\mathcal{M}^2$ , the remaining interaction parameters can be determined with the final four-species model  $\mathcal{M}^3$ . In this setup, we fix the parameters in  $\boldsymbol{\theta}^{(1)}$  and  $\boldsymbol{\theta}^{(2)}$  to their respective MAP estimates. Thus, only the remaining four interaction parameters are inferred, denoted as  $\boldsymbol{\theta}^{(3)}$ . Since the remaining interaction parameters are only in the matrix  $\mathbf{A}$ , the term that is not dependent on the antibiotic concentration, we set the latter to  $\alpha = 0$ .

The results of this third and last updating are shown in figs. 12 and 13. Here, sharp and distinct peaks along with a strong linear correlation can be observed for all four parameters. Again, a good agreement between posterior model response and data is observed after the updating.



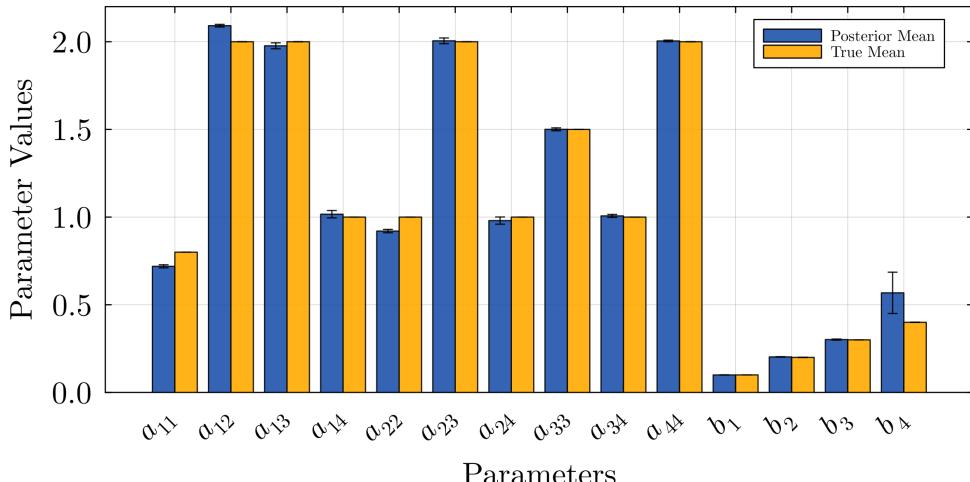
**Fig. 12:** Posterior samples of the mean values of the parameter set  $\theta^{(3)}$  updated with the four-species model  $\mathcal{M}^3$ .



**Fig. 13:** Comparison of the model output of model  $\mathcal{M}^3$  corresponding to calibrated posterior samples (shaded) and the data (scatter).

#### 4.2.4 Comparison of the identified posterior mean with true mean

In fig. 14, we compare the identified material parameters with the true parameter values used for the data generation. In addition, an error bar is added to the posterior means highlighting the standard deviation of the posterior. For almost all parameters, the values are very similar. The largest difference is obtained for the parameter  $b_4$  connected to the sensitivity to antibiotics of the fourth biofilm. In addition, a rather high standard deviation of the posterior results. This gives an important hint that more data is needed for a better identification.



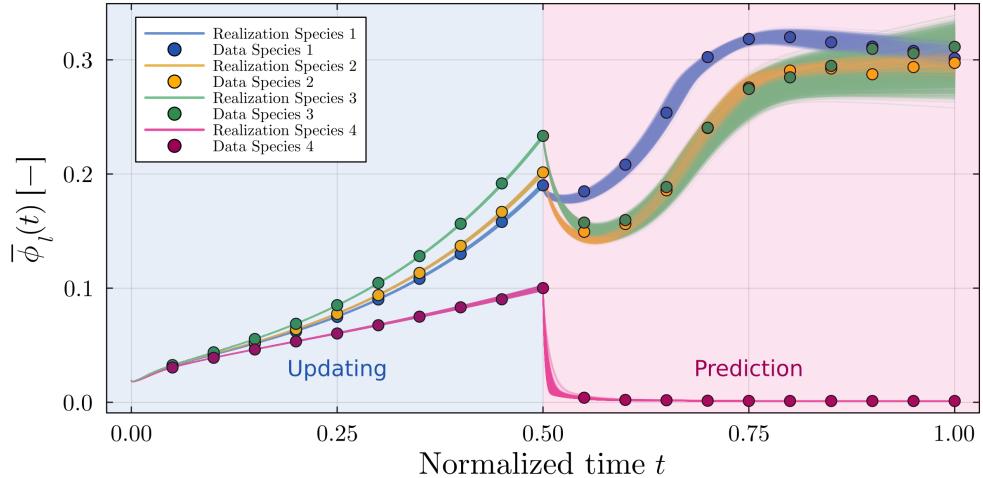
**Fig. 14:** Comparison of the mean values of the identified parameters and true values used to generate the data used in case II. The error bars reflect the standard deviation of the posterior samples of the respective parameters.

#### 4.2.5 Validation with time-dependent antibiotics

As a final part, a validation setup is considered in order to see how robust the calibrated model parameters are to a changed setup. Specifically, a time-dependent application of antibiotics is considered in the validation case. As indicated for  $\mathcal{M}_{\text{val}}^3$  in table 3, antibiotics are applied  $t = 0.5$ . For  $t < 0.5$ , the setup is identical with model  $\mathcal{M}^3$  used to calibrate the final interaction parameters.

The result of the validation is shown in fig. 15. There, the model responses up until the antibiotics are applied are the same in for  $\mathcal{M}^3$ . When the antibiotics are applied, a rapid change of the antibiotic concentration can be observed. Again, we see a good agreement between the data and the predicted model response. Here, the data is only used to validate the predicted response by comparison; no additional model calibration is performed. We note that the application of the antibiotics leads to an increased variability in the model response for the same set of inputs. This validation also shows that calibrating a physical model is useful in settings where the model is used in a

predictive setting. Since the underlying model parameters were updated, a change in environmental conditions can be modeled, even if no data is available for this change.



**Fig. 15:** Comparison of the model output of model  $\mathcal{M}_{\text{val}}^3$  corresponding to calibrated posterior samples (shaded) and the data (scatter). The output for  $t < 0.5$  corresponds to  $\mathcal{M}^3$  from fig. 13.

## 5 Conclusion

In this paper, we presented a Bayesian updating approach for biofilm growth models that accounts for hybrid uncertainties, incorporating both epistemic (unknown parameters) and aleatory (biological variability) uncertainty via a probabilistic formulation. Traditional double-loop approaches to uncertainty quantification are often inefficient in this context. To address this, we employed a reduced-order model based on Time-separated Stochastic Mechanics (TSM), enabling the propagation of aleatory uncertainty with only a single model evaluation, thus eliminating the need for nested simulation loops. By leveraging a Taylor-decomposition-based representation of the stochastic process, our approach allows direct computation of the outputs mean and variance, which are then used in a Gaussian likelihood function for inference. This significantly reduces computational costs while preserving accuracy in capturing uncertainty effects compared to a Monte Carlo approach.

The proposed methodology was validated through two representative case studies. The first involves a two-species biofilm model with five parameters, where monolithic updating was used to infer parameters governing growth, interaction, and antibiotic sensitivity. The second expands to a four-species system with fourteen parameters, employing a hierarchical inference strategy to decompose the high-dimensional problem into tractable sub-tasks. In both cases, the model successfully recovered the

true parameters and revealed meaningful interdependencies among them, highlighting the ability of the method to capture complex inter-species dynamics. Additionally, a validation study using time-dependent antibiotic application confirmed that the inferred parameters retain predictive power under varying experimental conditions. This supports the robustness and generalizability of the calibrated model.

Furthermore, our results show that employing the TSM-ROM approach for Bayesian updating is robust to nonlinearities and suitable to deal with a large number of uncertain parameters. Moreover, since the TSM-ROM directly captures the output's uncertainties, the necessity for expensive multi-query simulations is removed and the updating is less demanding in terms of computation time. The proposed approach therefore has the ability to reduce the time it takes to infer model parameters. Thus, the TSM-ROM approach can be used to accelerate the verification of the presented biofilm model against *in vitro* biofilm data, as for example presented by [52].

Overall, the TSM-ROM approach offers a computationally efficient and robust framework for Bayesian inference in complex, uncertainty-laden biological systems. In future research, the computational demand can further be addressed by using more efficient updating schemes like variational inference [53] or more efficient sampling strategies [54].

## Acknowledgements

The work has been funded by the German Research Foundation (DFG) within the framework of the International Research Training Group IRTG 2657 “Computational Mechanics Techniques in High Dimensions” under grant number 433082294. The work has been funded by the European Union (ERC, Gen-TSM, project number 101124463). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The work has been funded by dtec.bw - Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union - NextGenerationEU. The work has been funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) through the project grant SFB/TRR-298-SIIRI (Project-ID 426335750).

## References

- [1] Donlan, R.: Biofilms: Microbial Life on Surfaces. *Emerging Infectious Disease journal* **8**(9), 881 (2002) <https://doi.org/10.3201/eid0809.020063>
- [2] Kang, X., Yang, X., He, Y., Guo, C., Li, Y., Ji, H., Qin, Y., Wu, L.: Strategies and materials for the prevention and treatment of biofilms. *Materials Today Bio* **23**, 100827 (2023) <https://doi.org/10.1016/j.mtbio.2023.100827>
- [3] Chattopadhyay, I., J, R.B., Usman, T.M.M., Varjani, S.: Exploring the role of microbial biofilm for industrial effluents treatment. *Bioengineered* **13**(3), 6420–6440 (2022) <https://doi.org/10.1080/21655979.2022.2044250>
- [4] Klapper, I., Dockery, J.: Mathematical Description of Microbial Biofilms. *SIAM Review* **52**(2), 221–265 (2010) <https://doi.org/10.1137/080739720>
- [5] Shree, P., Singh, C.K., Sodhi, K.K., Surya, J.N., Singh, D.K.: Biofilms: Understanding the structure and contribution towards bacterial resistance in antibiotics. *Medicine in Microecology* **16**, 100084 (2023) <https://doi.org/10.1016/j.medmic.2023.100084>
- [6] Khatoon, Z., McTiernan, C.D., Suuronen, E.J., Mah, T.-F., Alarcon, E.I.: Bacterial biofilm formation on implantable devices and approaches to its treatment and prevention. *Heliyon* **4**(12), e01067 (2018) <https://doi.org/10.1016/j.heliyon.2018.e01067>
- [7] Melo, L.F., Bott, T.R.: Biofouling in water systems. *Experimental Thermal and Fluid Science* **14**(4), 375–381 (1997) [https://doi.org/10.1016/S0894-1777\(96\)00139-2](https://doi.org/10.1016/S0894-1777(96)00139-2)
- [8] Paquette, D.W., Brodala, N., Williams, R.C.: Risk Factors for Endosseous Dental Implant Failure. *Dental Clinics of North America* **50**(3), 361–374 (2006) <https://doi.org/10.1016/j.cden.2006.05.002>
- [9] Kommerein, N., Stumpp, S.N., Müsken, M., Ehlert, N., Winkel, A., Häussler, S., Behrens, P., Buettner, F.F.R., Stiesch, M.: An oral multispecies biofilm model for high content screening applications. *PLOS ONE* **12**(3), e0173973 (2017) <https://doi.org/10.1371/journal.pone.0173973>
- [10] Rath, H., Feng, D., Neuweiler, I., Stumpp, N.S., Nackenhorst, U., Stiesch, M.: Biofilm formation by the oral pioneer colonizer *Streptococcus gordonii*: An experimental and numerical study. *FEMS Microbiology Ecology* **93**(3), 010 (2017) <https://doi.org/10.1093/femsec/fix010>
- [11] Feng, D., Neuweiler, I., Nogueira, R., Nackenhorst, U.: Modeling of Symbiotic Bacterial Biofilm Growth with an Example of the *Streptococcus–Veillonella* sp. System. *Bulletin of Mathematical Biology* **83**(5), 48 (2021) <https://doi.org/10.1007/s12307-021-01070-w>

1007/s11538-021-00888-2

- [12] Moons, P., Michiels, C.W., Aertsen, A.: Bacterial interactions in biofilms. *Critical Reviews in Microbiology* **35**(3), 157–168 (2009) <https://doi.org/10.1080/10408410902809431>
- [13] Nadell, C.D., Xavier, J.B., Foster, K.R.: The sociobiology of biofilms. *FEMS Microbiology Reviews* **33**(1), 206–224 (2009) <https://doi.org/10.1111/j.1574-6976.2008.00150.x>
- [14] Yang, L., Liu, Y., Wu, H., Høiby, N., Molin, S., Song, Z.-j.: Current understanding of multi-species biofilms. *International Journal of Oral Science* **3**(2), 74–81 (2011) <https://doi.org/10.4248/IJOS11027>
- [15] James, G.A., Beaudette, L., Costerton, J.W.: Interspecies bacterial interactions in biofilms. *Journal of Industrial Microbiology* **15**(4), 257–262 (1995) <https://doi.org/10.1007/bf01569978>
- [16] Ouidir, T., Gabriel, B., Nait Chabane, Y.: Overview of multi-species biofilms in different ecosystems: Wastewater treatment, soil and oral cavity. *Journal of Biotechnology* **350**, 67–74 (2022) <https://doi.org/10.1016/j.jbiotec.2022.03.014>
- [17] Marsh, P.D.: Dental plaque: Biological significance of a biofilm and community life-style. *Journal of Clinical Periodontology* **32**(s6), 7–15 (2005) <https://doi.org/10.1111/j.1600-051X.2005.00790.x>
- [18] Klemp, F., Geisler, H., Soleimani, M., Junker, P.: A Continuum Multi-Species Biofilm Model with a Novel Interaction Scheme. *Biofilm paper* (2025)
- [19] Read, M.N., Alden, K., Timmis, J., Andrews, P.S.: Strategies for calibrating models of biology. *Briefings in Bioinformatics* (2020) <https://doi.org/10.1093/bib/bby092>
- [20] Gábor, A., Banga, J.R.: Robust and efficient parameter estimation in dynamic models of biological systems. *BMC Systems Biology* **9**(1), 74 (2015) <https://doi.org/10.1186/s12918-015-0219-2>
- [21] Mitra, E.D., Hlavacek, W.S.: Parameter estimation and uncertainty quantification for systems biology models. *Current Opinion in Systems Biology* **18**, 9–18 (2019) <https://doi.org/10.1016/j.coisb.2019.10.006>
- [22] Mary-Huard, T., Robin, S.: In: Stumpf, M.P.H., Balding, D., Girolami, M. (eds.) *Introduction to Statistical Methods for Complex Systems*, 1st edn., pp. 15–38. Wiley, Chichester, West Sussex (2011). <https://doi.org/10.1002/9781119970606.ch2>
- [23] Shewa, W.A., Sun, L., Bossy, K., Dagnew, M.: Biofilm characterization and

- dynamic simulation of advanced rope media reactor for the treatment of primary effluent. Water Environment Research **96**(11), 11150 (2024) <https://doi.org/10.1002/wer.11150>
- [24] Robert, C.P., Marin, J.-M., Rousseau, J.: In: Stumpf, M.P.H., Balding, D., Girolami, M. (eds.) Bayesian Inference and Computation, 1st edn., pp. 39–65. Wiley, Chichester, West Sussex (2011). <https://doi.org/10.1002/9781119970606.ch3>
- [25] Wilkinson, D.J.: Bayesian methods in bioinformatics and computational systems biology. Brief Bioinform **8**(2), 109–116 (2007) <https://doi.org/10.1093/bib/bbm007>
- [26] Rittmann, B.E., Boltz, J.P., Brockmann, D., Daigger, G.T., Morgenroth, E., Sørensen, K.H., Takács, I., van Loosdrecht, M., Vanrolleghem, P.A.: A framework for good biofilm reactor modeling practice (GBRMP). Water Science and Technology **77**(5), 1149–1164 (2018) <https://doi.org/10.2166/wst.2018.021>
- [27] Taghizadeh, L., Karimi, A., Presterl, E., Heitzinger, C.: Bayesian inversion for a biofilm model including quorum sensing. Computers in Biology and Medicine **117**, 103582 (2020) <https://doi.org/10.1016/j.combiomed.2019.103582>
- [28] Nooranidoost, M., Cogan, N.G., Stoodley, P., Gloag, E.S., Hussaini, M.Y.: Bayesian estimation of *Pseudomonas aeruginosa* viscoelastic properties based on creep responses of wild type, rugose, and mucoid variant biofilms. Biofilm **5**, 100133 (2023) <https://doi.org/10.1016/j.bioflm.2023.100133>
- [29] Willmann, H., Nitzler, J., Brandstätter, S., Wall, W.A.: Bayesian calibration of coupled computational mechanics models under uncertainty based on interface deformation. Advanced Modeling and Simulation in Engineering Sciences **9**(1), 24 (2022) <https://doi.org/10.1186/s40323-022-00237-5>
- [30] Willmann, H., Wall, W.A.: Inverse analysis of material parameters in coupled multi-physics biofilm models. Advanced Modeling and Simulation in Engineering Sciences **9**(1), 7 (2022) <https://doi.org/10.1186/s40323-022-00220-0>
- [31] Wollner, M.P., Rolf-Pissarczyk, M., Holzapfel, G.A.: A reparameterization-invariant Bayesian framework for uncertainty estimation and calibration of simple materials. Computational Mechanics (2025) <https://doi.org/10.1007/s00466-024-02573-2>
- [32] Bi, S., Beer, M., Cogan, S., Mottershead, J.: Stochastic Model Updating with Uncertainty Quantification: An Overview and Tutorial. Mechanical Systems and Signal Processing **204**, 110784 (2023) <https://doi.org/10.1016/j.ymssp.2023.110784>
- [33] Beck, J.L., Katafygiotis, L.S.: Updating Models and Their Uncertainties. I: Bayesian Statistical Framework. J. Eng. Mech. **124**(4), 455–461 (1998) [https://doi.org/10.1061/\(ASCE\)1084-0606\(1998\)124:4\(455\)](https://doi.org/10.1061/(ASCE)1084-0606(1998)124:4(455))

//doi.org/10.1061/(ASCE)0733-9399(1998)124:4(455)

- [34] Kiureghian, A.D., Ditlevsen, O.: Aleatory or epistemic? Does it matter? Structural Safety **31**(2), 105–112 (2009) <https://doi.org/10.1016/j.strusafe.2008.06.020>
- [35] Beer, M., Ferson, S., Kreinovich, V.: Imprecise probabilities in engineering analyses. Mechanical Systems and Signal Processing **37**(1-2), 4–29 (2013) <https://doi.org/10.1016/j.ymssp.2013.01.024>
- [36] Bi, S., Broggi, M., Beer, M.: The role of the Bhattacharyya distance in stochastic model updating. Mechanical Systems and Signal Processing **117**, 437–452 (2019) <https://doi.org/10.1016/j.ymssp.2018.08.017>
- [37] Kitahara, M., Bi, S., Broggi, M., Beer, M.: Nonparametric Bayesian stochastic model updating with hybrid uncertainties. Mechanical Systems and Signal Processing **163**, 108195 (2022) <https://doi.org/10.1016/j.ymssp.2021.108195>
- [38] Geisler, H., Junker, P.: Time-separated stochastic mechanics for the simulation of viscoelastic structures with local random material fluctuations. Computer Methods in Applied Mechanics and Engineering **407**, 115916 (2023) <https://doi.org/10.1016/j.cma.2023.115916>
- [39] Geisler, H., Erdogan, C., Nagel, J., Junker, P.: A new paradigm for the efficient inclusion of stochasticity in engineering simulations: Time-separated stochastic mechanics. Comput Mech **75**(1), 211–235 (2025) <https://doi.org/10.1007/s00466-024-02500-5>
- [40] Lye, A., Cicirello, A., Patelli, E.: Sampling methods for solving Bayesian model updating problems: A tutorial. Mechanical Systems and Signal Processing **159**, 107760 (2021) <https://doi.org/10.1016/j.ymssp.2021.107760>
- [41] Turner, B.M., Van Zandt, T.: A tutorial on approximate Bayesian computation. Journal of Mathematical Psychology **56**(2), 69–85 (2012) <https://doi.org/10.1016/j.jmp.2012.02.005>
- [42] Lye, A., Ferson, S., Xiao, S.: Comparison between Distance Functions for Approximate Bayesian Computation to Perform Stochastic Model Updating and Model Validation under Limited Data. ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng. **10**(2), 03124001 (2024) <https://doi.org/10.1061/AJRUA6.RUENG-1223>
- [43] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of State Calculations by Fast Computing Machines. The Journal of Chemical Physics **21**(6), 1087–1092 (1953) <https://doi.org/10.1063/1.1699114>
- [44] Hastings, W.K.: Monte Carlo sampling methods using Markov chains and their

- applications. *Biometrika* **57**(1), 97–109 (1970) <https://doi.org/10.1093/biomet/57.1.97>
- [45] Ching, J., Chen, Y.-C.: Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging. *J. Eng. Mech.* **133**(7), 816–832 (2007) [https://doi.org/10.1061/\(ASCE\)0733-9399\(2007\)133:7\(816\).1](https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816).1)
  - [46] Kirkpatrick, S., Gelatt Jr, C.D., Vecchi, M.P.: Optimization by simulated annealing. *science* **220**(4598), 671–680 (1983)
  - [47] Faes, M., Daub, M., Marelli, S., Patelli, E., Beer, M.: Engineering analysis with probability boxes: A review on computational methods **93**, 102092 <https://doi.org/10.1016/j.strusafe.2021.102092>
  - [48] Reiser, P., Aguilar, J.E., Guthke, A., Bürkner, P.-C.: Uncertainty quantification and propagation in surrogate-based Bayesian inference. *Stat Comput* **35**(3), 66 (2025) <https://doi.org/10.1007/s11222-025-10597-8>
  - [49] Junker, P., Balzani, D.: An extended hamilton principle as unifying theory for coupled problems and dissipative microstructure evolution. *Continuum Mechanics and Thermodynamics* **33**(4), 1931–1956 (2021)
  - [50] Behrensdorf, J., Gray, A., Perin, A., Grashorn, J., Luttmann, M., Broggi, M., Agarwal, G., Fritsch, L., Mett, F., Knipper, L.: FriesischScott/UncertaintyQuantification.Jl: V0.12.0. Zenodo (2025). <https://doi.org/10.5281/zenodo.14901342>
  - [51] Betz, W., Papaioannou, I., Straub, D.: Transitional Markov Chain Monte Carlo: Observations and Improvements. *J. Eng. Mech.* **142**(5), 04016016 (2016) [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0001066](https://doi.org/10.1061/(ASCE)EM.1943-7889.0001066)
  - [52] Heine, N., Bittroff, K., Szafrański, S.P., Duitscher, M., Behrens, W., Vollmer, C., Mikolai, C., Kommerein, N., Debener, N., Frings, K., et al.: Influence of species composition and cultivation condition on peri-implant biofilm dysbiosis in vitro (2025)
  - [53] Rubio, P.-B., Chamoin, L., Louf, F.: Real-time Bayesian data assimilation with data selection, correction of model bias, and on-the-fly uncertainty propagation. *Comptes Rendus Mécanique* **347**(11), 762–779 (2019) <https://doi.org/10.1016/j.crme.2019.11.004>
  - [54] Igea, F., Cicirello, A.: Cyclical Variational Bayes Monte Carlo for efficient multi-modal posterior distributions evaluation. *Mechanical Systems and Signal Processing* **186**, 109868 (2023) <https://doi.org/10.1016/j.ymssp.2022.109868>