



UNIVERSITY OF LEEDS

This is a repository copy of *Bayesian Probabilistic Numerical Methods in Time-Dependent State Estimation for Industrial Hydrocyclone Equipment*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/140561/>

Version: Accepted Version

Article:

Oates, CJ, Cockayne, J, Aykroyd, RG orcid.org/0000-0003-3700-0816 et al. (1 more author) (2019) Bayesian Probabilistic Numerical Methods in Time-Dependent State Estimation for Industrial Hydrocyclone Equipment. *Journal of the American Statistical Association*, 114 (528). pp. 1518-1531. ISSN 0162-1459

<https://doi.org/10.1080/01621459.2019.1574583>

© 2019 American Statistical Association. This is an author produced version of a paper published in *Journal of the American Statistical Association*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Bayesian Probabilistic Numerical Methods in Time-Dependent State Estimation for Industrial Hydrocyclone Equipment

Chris. J. Oates*

School of Mathematics, Statistics and Physics, Newcastle University
Alan Turing Institute

and

Jon Cockayne

Department of Statistics, University of Warwick

and

Robert G. Aykroyd

School of Mathematics, University of Leeds

and

Mark Girolami

Department of Mathematics, Imperial College London

Alan Turing Institute

*Correspondence should be addressed to the first author. Chris. J. Oates is Chair in Statistics, School of Mathematics, Statistics and Physics, Newcastle University, NE1 7RU, UK, as well as Turing Fellow and Group Leader, Alan Turing Institute, London NW1 2DB, UK. (E-mail: chris.oates@ncl.ac.uk) Jon Cockayne is a PhD student, Department of Statistics, University of Warwick, CV4 7AL, UK. (E-mail: j.cockayne@warwick.ac.uk) Robert G. Aykroyd is Senior Lecturer, School of Mathematics, University of Leeds, LS2 9JT, UK. (E-mail: r.g.aykroyd@leeds.ac.uk) Mark Girolami is Chair in Statistics, Department of Mathematics, Imperial College London, SW7 2AZ, UK. (E-mail: m.girolami@imperial.ac.uk) This research was supported by the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers and by the Key Technology Partnership program at the University of Technology Sydney. CJO and MG were supported by the Lloyd's Register Foundation programme on data-centric engineering at

Abstract

The use of high-power industrial equipment, such as large-scale mixing equipment or a hydrocyclone for separation of particles in liquid suspension, demands careful monitoring to ensure correct operation. The fundamental task of state-estimation for the liquid suspension can be posed as a time-evolving inverse problem and solved with Bayesian statistical methods. In this paper, we extend Bayesian methods to incorporate statistical models for the error that is incurred in the numerical solution of the physical governing equations. This enables full uncertainty quantification within a principled computation-precision trade-off, in contrast to the over-confident inferences that are obtained when all sources of numerical error are ignored. The method is cast within a sequential Monte Carlo framework and an optimised implementation is provided in `Python`.

Keywords: Inverse Problems, Electrical Tomography, Partial Differential Equations, Probabilistic Meshless Methods, Sequential Monte Carlo

the Alan Turing Institute, UK. MG was supported by the EPSRC grants [EP/K034154/1, EP/R018413/1, EP/P020720/1, EP/L014165/1], an EPSRC Established Career Fellowship [EP/J016934/1] and a Royal Academy of Engineering Research Chair in Data Centric Engineering. The collection of tomographic data was supported by an EPSRC grant [GR/R22148/01]. This material was based upon work partially supported by the National Science Foundation under Grant DMS-1127914 to the Statistical and Applied Mathematical Sciences Institute. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

Hydrocyclones provide a simple and inexpensive method for removing solids from liquids, as well as separating two liquids according to their relative densities (assuming equal fluid resistances) (Gutierrez et al. 2000). They have widespread applications, including in areas such as environmental engineering and the petrochemical industry (Sripriya et al. 2007). In particular, they have few moving parts, can handle large volumes and are relatively inexpensive to maintain. This makes them ideal as part of a continuous process in hazardous industrial settings and contrasts with alternatives, such as filters and centrifuges, which are more susceptible to breakdown and/or have higher running costs. The physical principles governing the hydrocyclone are simple; a mixed input is forced into a cone-shaped tank at high pressure, to create a circular rotation. This rotation forces less-dense material to the centre and denser material to the periphery of the tank. The less-dense material in the core can then be extracted from the top (overflow) and the denser material removed from the bottom of the tank (underflow). This mechanism is illustrated in Fig. 1.

Continual monitoring of the hydrocyclone is essential in most industrial applications, since the input flow rate is an important control parameter that can be adjusted to maximise the separation efficiency of the equipment. Our focus in this work is on state estimation for the internal fluid. Indeed, the high pressures that are often involved necessitate careful observation of the internal fluid dynamics to ensure safety in operation (Bradley 2013).

1.1 Statistical Challenges

Direct observation of the internal flow of the fluids is difficult or impossible due to, for example, the reinforced walls of the hydrocyclone and the opacity of the mixed component. Under correct operation, the output (overflow and underflow) can be measured and tested

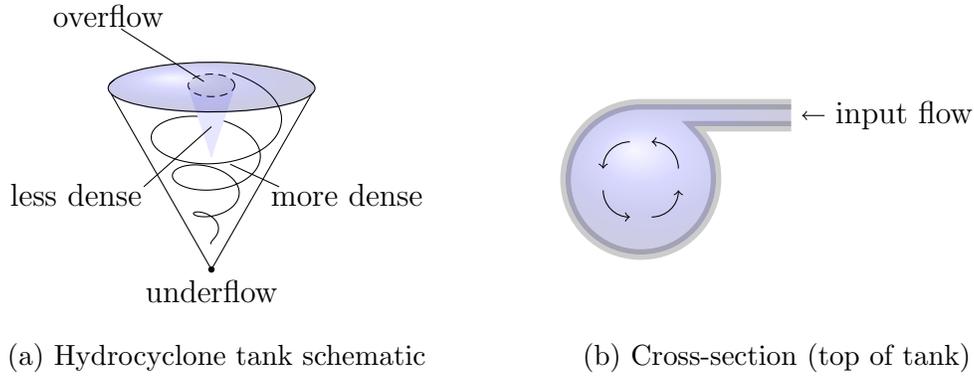


Figure 1: A simplified schematic description of typical hydrocyclone equipment. (a) The tank is cone-shaped with overflow and underflow pipes positioned to extract the separated contents. (b) Fluid, a mixture to be separated, is injected at high pressure at the top of the tank to create a vortex. Under correct operation, less-dense materials are directed toward the centre of the tank and denser materials are forced to the peripheries of the tank.

for purity, but advanced indication of a potential loss of efficiency is desirable, if not essential in most industrial contexts. Such a warning allows for adjustment and hence avoidance of impending catastrophic failure. One possible technique for monitoring the internal flow is electrical impedance tomography (EIT). The target of an EIT analysis is the electrical conductivity field a^\dagger of the physical object; the conductivities of different fluid components will in general differ and this provides a means to measure the fluid constituents. This technique has many applications in medicine, as well as industry, as it provides a non-invasive method to estimate internal structure from external measurements (the *inverse problem*) (Gutierrez et al. 2000). Further, it is ideal for industrial processes as it is possible to collect data at rates of several hundred frames per second, hence allowing real-time monitoring and control of sensitive systems. However, the rapid acquisition of data requires equally rapid analysis and the nature of EIT requires that low-accuracy approximations to the physical

governing equations are needed to keep pace with incoming data in the monitoring context (Hamilton & Hauptmann 2018). This is due to the computational demands that are posed by the repeated solution of physical governing equations (the *forward problem*) in evaluation of the statistical likelihood. However, in standard approaches, the error introduced by a crude discretisation of the physical governing equations is not accounted for and may lead to an over-optimistic view of the precision of results. This could lead to misleading interpretations of the results and hence potentially dangerous mis-control; it is therefore important to account for the presence of an unknown and non-negligible discretisation error in interpretation of the statistical output.

1.2 Probabilistic Numerical Methods

Probabilistic numerics (Hennig et al. 2015) is an emergent research field that aims to model the uncertainty in the solution space of the physical equations that arises when the forward problem is only approximately solved. In contrast to conventional emulation methods (Kennedy & O’Hagan 2001), which are *extrusive* in the sense that the physical equations are treated as a black box, probabilistic numerical methods are *intrusive* and seek to model the error introduced in the numerical solution due to discretisation of the original continuous physical equations. Thus a probabilistic numerical method provides *uncertainty quantification for the forward problem* that is meaningful, reflective of the specific discretisation scheme employed, and enables a principled computation-precision trade-off, where the presence of an unknown discretisation error is explicitly accounted for by marginalisation over the unknown solution to the forward problem (Briol et al. 2018, Cockayne et al. 2017). This paper contributes a rigorous assessment of probabilistic numerical methods for the Bayesian solution of an important inverse problem in industrial process monitoring, detailed next.

1.3 Our Contributions

The scientific problem that we consider is Bayesian state estimation for the time-evolving conductivity field of internal fluid using data obtained via EIT. The Bayesian approach to inverse problems is well-studied (Stuart 2010, Nouy & Soize 2014) and in particular the application of statistical methods to EIT is now well-understood (Kaipio et al. 1999, 2000, Watznig & Fox 2009, Dunlop & Stuart 2016, Yan & Guo 2015, Aykroyd 2015, Stuart & Teckentrup 2016) with sophisticated computational methods proposed (Kaipio et al. 2000, Vauhkonen et al. 2001, Polydorides & Lionheart 2002, Schwab & Stuart 2012, Schillings & Schwab 2013, Beskos et al. 2015, Chen & Schwab 2015, Hyvonen & Leinonen 2015, Chen & Schwab 2016*a,b,c*). In this paper, probabilistic numerical methods are proposed and investigated as a natural approach to uncertainty quantification with a computation-precision trade-off, wherein numerical error in the approximate solution of the forward problem is explicitly modelled and accounted for in a full Bayesian solution to the inverse problem of interest. At present, the literature on probabilistic numerical methods for partial differential equations consists of Owhadi (2015), Chkrebtii et al. (2016), Owhadi (2017), Owhadi & Zhang (2017), Cockayne et al. (2016*a,b*), Conrad et al. (2017), Raissi et al. (2017). This paper goes further than the most relevant work in Cockayne et al. (2016*a,b*), which tackled Bayesian inverse problems based on EIT with probabilistic numerical methods, in several aspects:

- The inversion problem herein is more challenging than the (static) problems considered in previous work, in that we aim to recover the temporal evolution of the true unknown conductivity field a^\dagger based on indirect and noise-corrupted observations at a finite set of measurement times. To address this challenge, a (descriptive, rather than mechanistic) Markovian prior model a for the field is developed, which is shown

to admit a filtering formulation (Todescato et al. 2017). This permits a sequential Monte Carlo method (particle filter) to be exploited for efficient data assimilation (Law et al. 2015).

- The filtering formulation introduces additional challenges due to the fact that numerical (discretisation) error in solution of the forward problem will be propagated through computations performed at earlier time points to later time points, as well as the possibility that numerical errors can accumulate within the computations. A detailed empirical investigation of the computation-precision trade-off is undertaken based on the use of probabilistic numerical methods for solution of the EIT governing equations.
- Real experimental data are analysed, generated by one of the present authors in a controlled laboratory experiment. These data consist of 2,401 individual voltage measurements taken at discrete spatial and temporal intervals over the boundary of the vessel, and are used to demonstrate the efficacy of the approach under realistic experimental conditions.

In particular, this paper constitutes one of the first serious applications of probabilistic numerical methods to a challenging real-world problem, where proper quantification of uncertainty is crucial.

1.4 Overview of the Paper

The structure of the paper is as follows: Sec. 2 contains the mathematical, statistical and computational methodological development. Sec. 3 reports our experimental results and Sec. 4 discusses their implications for further research and for future application to industrial processes.

2 Methods

In Sec. 2.1 we introduce the physical model and make the inversion problem formal. Then in Sec. 2.2 we recall the Bayesian approach to inversion, with an extension to a time-evolving unknown parameter. Sec. 2.3 introduces probabilistic models for numerical error incurred in discretisation of the physical governing equations. The final section, 2.5 develops a sequential Monte Carlo method for efficient computation.

2.1 Abstraction of the Inverse Problem

The physical equations that model the measurement process are presented below, following the recent comprehensive treatment in Dunlop & Stuart (2016).

2.1.1 Set-Up

Consider a bounded, open domain $D \subset \mathbb{R}^d$ with smooth boundary denoted ∂D . Let $\bar{D} = D \cup \partial D$. The domain represents a physical object and our parameter of interest is the conductivity field $a : \bar{D} \rightarrow \mathbb{R}$ of that object. Here $a(x)$ denotes the conductivity at spatial location $x \in \bar{D}$. Consider m electrodes fixed to ∂D , the region of contact of electrode $i \in \{1, \dots, m\}$ being denoted $E_i \subset \partial D$. A current stimulation pattern $\mathbf{I} = (I_i)_{i=1}^m \in \mathbb{R}^m$ is passed, via the electrodes, through the object. Note that from physical conservation of current we have $\sum_{i=1}^m I_i = 0$.

The electrical potential $u : \bar{D} \rightarrow \mathbb{R}$ over the domain, induced by the current stimulation

pattern, can be described by the following partial differential equation (PDE):

$$\begin{aligned}
& \nabla \cdot (a \nabla u) = 0 \quad \text{in } D \\
& \int_{E_i} a \nabla u \cdot \mathbf{n} d\sigma = I_i \\
& u = U_i \quad \text{on } E_i \\
& a \nabla u \cdot \mathbf{n} = 0 \quad \text{on } \partial D \setminus \cup_{i=1}^m E_i.
\end{aligned} \tag{1}$$

Here \mathbf{n} is the *outward* unit normal, which corresponds to the convention that $I_i > 0$ refers to current flow *out* of the domain, and $d\sigma$ represents an infinitesimal boundary element. The quantities U_i on the electrodes E_i will constitute the measurements. Known as the *complete electrode model* (CEM), this PDE¹ was first studied in Cheng et al. (1989). For a suitable fixed field a , existence of a solution u is guaranteed and, under the additional condition that $\sum_{i=1}^m U_i = 0$, uniqueness of the solution u can also be established (Somersalo et al. 1992). Thus the forward problem is well-defined.

The true conductivity field a^\dagger is considered to be unknown and is the object of interest. In contrast to most work on EIT, in our context a^\dagger is time-dependent and we extend the notation as $a^\dagger(x, t)$ for, with no loss in generality, a time index $t \in [0, 1]$. In order to estimate a^\dagger , measurements $y_{j,k}$ are obtained under distinct stimulation patterns $\mathbf{I}_j \in \mathbb{R}^m$, $j = 1, \dots, J$, modelled as

$$y_{j,k} = \mathcal{P}_{j,k} u^\dagger + \epsilon_{j,k} \in \mathbb{R}^m \tag{2}$$

¹The mathematical formulation in Eqn. 1 assumes, as we do in this work, that contact impedance at the electrodes can be neglected. For the case of imperfect electrodes, the reader is referred to Aykroyd (2018).

where the projections

$$\mathcal{P}_{j,k}u^\dagger := \begin{bmatrix} u(x_1^E; \mathbf{I}_j, a^\dagger(\cdot, t_k)) \\ \vdots \\ u(x_m^E; \mathbf{I}_j, a^\dagger(\cdot, t_k)) \end{bmatrix}$$

are defined for each stimulation pattern j and each discrete time point t_k , $k = 1, \dots, n$, and the $\epsilon_{j,k}$ represent error in the measurement. Here $u(\cdot; \mathbf{I}, a)$ denotes the solution of the PDE with conductivity field a and stimulation pattern \mathbf{I} , while x_i^E is a point central to the electrode E_i . Thus $u^\dagger = u(\cdot; \mathbf{I}, a^\dagger)$ is the solution of the PDE defined by the true field a^\dagger and, for fixed j, k , the vector $\mathcal{P}_{j,k}u^\dagger$ contains the quantities U_i in the CEM with conductivity field $a^\dagger(\cdot, t_k)$ and stimulation pattern $\mathbf{I} = \mathbf{I}_j$.

In the absence of further conditions on a^\dagger , the inverse problem is ill-posed. Indeed, the infinite-dimensional field a^\dagger cannot be uniquely recovered from a finite dataset. (Recall the seminal work of Hadamard (1902), who defined an inverse problem to be *well-posed* if (i) a solution exists, (ii) the solution is unique, and (iii) the solution varies continuously as the data are varied.) To proceed, the inverse problem must therefore be regularised (Tikhonov & Arsenin 1977).

2.2 The Bayesian Approach to Inversion

In this section we exploit Bayesian methods to regularise the inverse problem (Stuart 2010). Sec. 2.2.1 introduces the prior model, Sec. 2.2.2 casts posterior computation as a filtering problem and Sec. 2.2.3 reviews mathematical analysis for numerical approximation of the posterior that accounts for numerical error in the PDE solution method.

2.2.1 Prior Model for the Conductivity Field

In this paper we interpret Eqn. 1 in the strong form, which in particular requires the existence of ∇a^\dagger on D . This information will be encoded into a prior distribution: Let $\{\phi_i\}_{i=1}^\infty$ be an orthonormal basis for a separable Hilbert space H with norm $\|\cdot\|_H$. It is assumed that $H \subset C^1(\bar{D})$, where $C^m(S)$ is used to denote the set of m -times continuously differentiable functions from S to \mathbb{R} . Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and for measurable $v : \Omega \rightarrow \mathbb{R}$ denote $\mathbb{E}v = \int v d\mathbb{P}$.

Model Assumption 1. *Let $\alpha > 1/2$ and $\omega \in \Omega$. Our prior model is expressed as a separable Karhounen-Loéve expansion:*

$$\log a(x, t; \omega) = \sum_{i=1}^{\infty} i^{-\alpha} \psi_i(t; \omega) \phi_i(x)$$

where the $\omega \mapsto \psi_i(\cdot; \omega)$ are modelled as independent Gaussian processes with mean functions $m_{\psi,i}$ and covariance functions $k_{\psi,i}$ such that

$$m_{\psi}^{\max} := \sup_{i \in \mathbb{N}} \sup_{t \in [0,1]} |m_{\psi,i}(t)| < \infty, \quad k_{\psi}^{\max} := \sup_{i \in \mathbb{N}} \sup_{t \in [0,1]} k_{\psi,i}(t, t) < \infty.$$

The logarithm is used to ensure positivity of the conductivity field, as is considered standard in Bayesian approaches to (static) EIT (Dunlop & Stuart 2016). Henceforth the probability argument $\omega \in \Omega$ will be left implicit.

This prior construction ensures that ∇a exists in D . To see this, we have the following result:

Proposition 1. *For fixed $t \in [0, 1]$, almost surely $a(\cdot, t)$ exists in $C^1(\bar{D})$.*

Note that, in particular, this result justifies point evaluation of ∇a in the algorithms that we present; since $x \mapsto \nabla a(x, t)$ is almost surely continuous, such point evaluations are almost surely well-defined. All proofs are reserved for Appendix A.

This paper imparts weak prior assumptions, in the sense of smoothness, on the time-evolution of the random field:

Model Assumption 2. *The ψ_i are modelled as Brownian with mean functions $m_{\psi,i}(t) = 0$ and covariance functions $k_{\psi,i}(t, t') = \lambda \min(t + \tau, t' + \tau)$, for all $t, t' \in [0, 1]$, for some fixed $\lambda > 0$ and $\tau \geq 0$.*

This prior model allows for flexible and data-driven estimation of the temporal evolution of the unknown conductivity field. At the same time, this choice allows estimation to be cast as a filtering problem (see Sec. 2.2.2) due to the following important fact:

Proposition 2 (Due to Wiener (1949)). *The increments $\psi_i(t + s) - \psi_i(t)$ are independent with distribution $N(0, \lambda(s + \tau))$, for all $0 \leq t \leq t + s \leq 1$.*

An immediate consequence is that the field $a(\cdot, t)$ itself is a Markov process. To see this, let

$$k_\phi(x, x') = \sum_{i=1}^{\infty} i^{-2\alpha} \phi_i(x) \phi_i(x').$$

For convenience, we let $\theta = \log a$ in the sequel. Then:

Corollary 1. *The increments $\theta_\Delta(\cdot) := \theta(\cdot, t + s) - \theta(\cdot, t)$ are independent Gaussian random fields with mean function $m_\Delta(x) = 0$ and covariance function $k_\Delta(x, x') = \lambda(s + \tau)k_\phi(x, x')$, for all $x, x' \in D$ and all $0 \leq t \leq t + s \leq 1$.*

Let Γ_s denote the distribution of the increment θ_Δ over the time interval $[t, t + s]$. The infinite-dimensional nature of the random variable θ_Δ precludes the use of standard density notation, due to the non-existence of a Lebesgue measure in the infinite-dimensional

context (p.143 Yamasaki 1985). Instead, the distribution Γ_s is formalised through its Radon-Nikodym derivative

$$\frac{d\Gamma_s}{d\gamma}(\theta_\Delta) \propto \exp\left(-\frac{1}{2}\|\theta_\Delta\|_{k_\Delta}^2\right) = \exp\left(-\frac{1}{2\lambda(s+\tau)}\|\theta_\Delta\|_{k_\phi}^2\right)$$

with respect to abstract Wiener measure γ (Gross 1967), where $\|\cdot\|_k$ denotes the Cameron-Martin norm based on the covariance function k . The reader unfamiliar with Radon-Nikodym notation is referred to the accessible introduction in Halmos & Savage (1949).

2.2.2 Formulation as a Filtering Problem

Denote $\theta_k = \theta(\cdot, t_k)$. Then the directed acyclic graph representation of the conditional independence structure of the statistical model (Lauritzen 1996) is as follows:

$$\begin{array}{ccccccccc} \theta_1 & \rightarrow & \theta_2 & \rightarrow & \dots & \rightarrow & \theta_{n-1} & \rightarrow & \theta_n \\ \downarrow & & \downarrow & & & & \downarrow & & \downarrow \\ Y_{\cdot,1} & & Y_{\cdot,2} & & \dots & & Y_{\cdot,n-1} & & Y_{\cdot,n} \end{array}$$

Let Π_n represent the posterior distribution over the conductivity field based on the data $y_{\cdot,k}$ for $k \leq n$. Then statistical inference is naturally formulated as a filtering problem at linear cost (Särkkä 2013, Todescato et al. 2017):

$$\frac{d\Pi_n}{d\Pi_{n-1}}(\theta) \propto p(y_{\cdot,n}|\mathbb{I}, \theta_n).$$

Here the Radon-Nikodym notation has been used on the LHS, while p has the conventional interpretation as a p.d.f. with respect to Lebesgue measure, here representing the likelihood model specified by the distributional model for $\epsilon_{j,k}$ in Eqn. 2. The solution to the filtering problem is the n -step posterior distribution:

$$\frac{d\Pi_n}{d\Pi_0}(\theta) = \prod_{k=1}^n \frac{d\Pi_k}{d\Pi_{k-1}}(\theta)$$

where the reference measure Π_0 is the prior distribution for the conductivity field given in Assumption 1. From Prop. 2, the prior marginal on $(\theta_1, \dots, \theta_n)$, denoted $\Pi_{0,1:n}$, can be decomposed as follows:

$$\frac{d\Pi_{0,1:n}}{d(\gamma \times \dots \times \gamma)}(\theta_1, \dots, \theta_n) \propto \frac{d\Pi_{0,1}}{d\gamma}(\theta_1) \prod_{k=2}^n \frac{d\Gamma_{t_k - t_{k-1}}}{d\gamma}(\theta_k - \theta_{k-1}),$$

where $\gamma \times \dots \times \gamma$ denotes the product of n abstract Wiener measures and the initial distribution $\Pi_{0,1}$ is computed as

$$\frac{d\Pi_{0,1}}{d\gamma}(\theta_1) \propto \exp\left(-\frac{1}{2\lambda(t_1 + \tau)} \|\theta_1\|_{k_\phi}^2\right)$$

as a direct consequence of Assumption 2. Later we use $\Pi_{n,n+1}$ to denote the marginal of Π_n over the components $\theta(\cdot, t_{n+1})$; the so-called *posterior predictive* distribution. This is simply a convolution of Π_n with the centred Gaussian field described in Cor. 1 and is of industrial relevance since it allows anticipation of the future dynamics and thus for intelligent hazard control.

2.2.3 Numerical Error and its Analysis

The likelihood model in Eqn. 2 depends on the projections $\mathcal{P}_{j,k}u$ which in turn depend on the exact solution $u(x_i^E; I_j, a(\cdot, t_k))$ of the PDE for given inputs I_j and $a(\cdot, t_k)$. In general the exact solution of the PDE is unavailable in closed-form and numerical methods are used to obtain discrete approximations, for instance based on a finite element or collocation basis (Quarteroni & Valli 2008). The assessment of the error introduced through discretisation is well-studied, with sophisticated theories for worst-case and average-case errors and beyond (Novak & Wozniakowski 2008, 2010).

Several papers have leveraged these analyses to consider the impact of discretisation error in the forward problem on the inferences that are made for the inverse problem (Schwab

& Stuart 2012, Schillings & Schwab 2013, 2014, Nouy & Soize 2014, Bui-Thanh & Ghattas 2014, Chen & Schwab 2015, 2016*c,a*, Nagel & Sudret 2016). These analyses all focus on static inverse problems (i.e. for a single time point). However, the generalisation of these theoretical results to the temporal context introduces considerable technical difficulties. Indeed, the filtering formulation is such that error in an approximation of Π_1 will be propagated and lead to an error in the approximation of Π_n whenever $n \geq 2$. Numerical approximation of Π_n thus involves $n - 1$ sources of discretisation error and analysis in the time-evolving setting must account for propagation and accumulation of these discretisation errors. However, standard worst-case error analyses (such as those listed above) are inappropriate for temporal problems, since in general the worst-case scenario will not be realised simultaneously by all numerical methods involved in the computational work-flow.

Presented with such an inherently challenging problem, our novel approach - described in the next section - to *model* discretisation error as an unknown random variable and propagate *uncertainty due to discretisation* through computation has appeal on philosophical, technical and practical levels.

2.3 Probabilistic Numerical Methods

Recall that the exact solution $u(\cdot; a, I)$ to the PDE is unavailable in closed-form. In this section we view numerical solution of Eqn. 1 not as a forward problem, but as an inverse problem in its own right (called a *sub-inverse* problem in this work) and provide full quantification of solution uncertainty that arises from the discretisation of this PDE via a collocation-type method. Sec. 2.3.1 introduces a prior model for u while Sec. 2.3.2 completes the specification of this sub-inverse problem associated with solution of the PDE. Then, Sec. 2.3.3 demonstrates how solution uncertainty can be propagated through the original inverse problem by marginalisation over the unknown exact solution u of the

PDE. Sec. 2.4 establishes theoretical properties of the proposed method.

2.3.1 Prior Model for the Potential Field

In this section we again adopt Bayesian methods to make the sub-inverse problem well-posed. The chief task is to construct a prior for u , the potential field. In principle, the physical governing equations, together with the prior for the conductivity field a , induce a unique prior for the potential field. The relationship between these probabilities has been explored in the context of stochastic PDEs; see Lord et al. (2014) for a book-length treatment. However, the task of characterising (or even approximating) the implied distribution on u is highly non-trivial². For this reason, we follow Cockayne et al. (2016a,b) and treat the two unknown fields as independent under the prior model. In particular, we encode independence across time points into the prior model for u , a choice that is algorithmically convenient. This allows us, in the following, to leave the time index implicit. This has a natural statistical interpretation of encoding only partial information into the prior - and can be both statistically and pragmatically justified (Potter & Anderson 1983).

To reduce notation in this and the following section, we consider a fixed conductivity field $a \in C^1(\bar{D})$ and a fixed current stimulation pattern $I \in \mathbb{R}^m$; these will each be left implicit.

Model Assumption 3. *The unknown solution u to Eqn. 1 is modelled as a Gaussian process with mean function $m_u(x) = 0$ and covariance function*

$$k_u(x, x') = \int k_u^0(x, z)k_u^0(z, x')dz \quad (3)$$

such that $k_u^0 \in C^{2 \times 2}(\bar{D} \times \bar{D})$ is a positive-definite kernel.

²In principle this is characterised by the Green's function of the PDE, but if the Green's function was known we would not have needed to discretise the PDE.

This minimal assumption ensures that, *under the prior*, the differential $\nabla \cdot (a\nabla u)$ is well-defined over D . Indeed, in general:

Proposition 3. *If $k_u^0 \in C^{\beta \times \beta}(\bar{D} \times \bar{D})$ with $\beta \in \mathbb{N}$, then almost surely $u \in C^\beta(\bar{D})$.*

2.3.2 Probabilistic Meshless Method

Next we obtain a posterior distribution over the solution u to the PDE in Eqn. 1. In particular this requires us to be explicit about the nature of our “data” for this sub-inverse problem. The mathematical justification for our approach below is provided in the information-based complexity literature on linear elliptic PDEs of the form $Au = f$ on D , $Bu = g$ on ∂D (Werschulz 1996, Novak & Wozniakowski 2008, Cialenco et al. 2012). In this framework, limited data $f_i = f(x_i^A)$, $g_i = g(x_i^B)$ are provided on the forcing term f and the boundary term g ; the mathematical problem is then optimal recovery of the solution u from these data, under a loss function that must be specified. This is a particular example of a *linear information* problem, since the f_i and g_i are linear projections of the unknown solution u of interest; see Novak & Wozniakowski (2008) for a book length treatment.

The data with which we work, in the above sense, are linear projections obtained at *collocation* points $\{x_i^A\}_{i=1}^{n_A} \subset D$ and $\{x_i^B\}_{i=1}^{n_B} \subset \partial D$:

$$\begin{aligned} \mathcal{L}_i u &:= \nabla \cdot a(x_i^A) \nabla u(x_i^A) &= 0 & \quad i = 1, \dots, n_A \\ \mathcal{L}_{n_A+i} u &:= a(x_i^B) \nabla u(x_i^B) \cdot \mathbf{n}(x_i^B) &= 0 & \quad i = 1, \dots, n_B \\ \mathcal{L}_{n_A+n_B+i} u &:= \int_{E_i} a \nabla u \cdot \mathbf{n} d\sigma &= I_i & \quad i = 1, \dots, m. \end{aligned}$$

Here $\mathcal{L} = [\mathcal{L}_1, \dots, \mathcal{L}_{n_A+n_B+m}]$ is a linear operator from $C^2(\bar{D})$ to $\mathbb{R}^{n_A+n_B+m}$. For a function $h(\cdot, \cdot) \in C^{2 \times 2}(\bar{D} \times \bar{D})$, in a slight abuse of notation, $\mathcal{L}h$ will be used to denote action of \mathcal{L} on the first argument, while the notation $\bar{\mathcal{L}}h$ denotes action on the second argument.

The composition $\mathcal{L}\bar{\mathcal{L}}h$ is understood as a matrix with (i, j) th element $\mathcal{L}_i\bar{\mathcal{L}}_jh \in \mathbb{R}$. In this notation, the data can be expressed as $\mathcal{L}u = [0^\top, \mathbf{I}^\top]^\top$ where $\mathbf{I} = [I_1, \dots, I_m]^\top$. The posterior over u is obtained by conditioning the prior measure on these data. Recall that $\mathcal{P}u = [u(x_1^E), \dots, u(x_m^E)]^\top$. For our purposes, it is sufficient to obtain the posterior over the finite dimensional vector $\mathcal{P}u$:

$$\begin{aligned} \mathcal{P}u \left| \mathcal{L}u = \begin{bmatrix} 0 \\ \mathbf{I} \end{bmatrix} \right. &\sim \text{N}(\mu, \Sigma) \\ \mu &= [\mathcal{P}\bar{\mathcal{L}}k_u][\mathcal{L}\bar{\mathcal{L}}k_u]^{-1} \begin{bmatrix} 0 \\ \mathbf{I} \end{bmatrix} \\ \Sigma &= [\mathcal{P}\bar{\mathcal{P}}k_u] - [\mathcal{P}\bar{\mathcal{L}}k_u][\mathcal{L}\bar{\mathcal{L}}k_u]^{-1}[\mathcal{L}\bar{\mathcal{P}}k_u] \end{aligned} \quad (4)$$

This distribution represents uncertainty due to the finite amount of computation that is afforded to numerical solution of the PDE in Eqn. 1. Eqn. 4 was termed a *probabilistic meshless method* in Cockayne et al. (2016a,b). Note that the *maximum a posteriori* estimate μ is identical to the point estimate provided by symmetric collocation (Fasshauer 1996) and this point estimator (only) was considered in the context of Bayesian PDE-constrained inverse problems in Marzouk & Xiu (2009), Yan & Guo (2015). The point estimator μ requires that the (n_A+n_B+m) -dimensional square matrix $\mathcal{L}\bar{\mathcal{L}}k_u$ is inverted; since this is also the computational bottleneck in computation of Σ , it follows that the probabilistic meshless method has essentially the same computational cost as its non-probabilistic counterpart. Considerable theoretical advances in the numerical analysis of these probabilistic numerical methods (for static problems) have since been made in Owhadi (2017). For non-degenerate kernels k_u , the matrix $\mathcal{L}\bar{\mathcal{L}}k_u$ is of full rank provided that no two collocation points are coincidental.

The selection of collocation points can be formulated as a problem of statistical experi-

mental design. Indeed, adaptive refinement strategies, that target an appropriate functional of the posterior covariance Σ until a pre-specified tolerance is met, can be considered (see Cockayne et al. 2016a). For brevity in this paper we simply considered the collocation points to be fixed.

2.3.3 Marginal Likelihood

The natural approach to define a data distribution is through marginalisation over the unknown solution u to the PDE. This marginalisation can be performed in closed form under a Gaussian measurement error model:

Model Assumption 4. *The measurement errors $\epsilon_{j,k}$ are independent $N(0, \sigma^2 I)$.*

Consider a stimulation pattern I_j applied at time $t_k \in [0, 1]$. Define $P_{j,k} = \mathcal{P}u(\cdot; I_j, a(\cdot, t_k))$ and denote by μ_j, Σ_j the output of the probabilistic meshless method (Eqn. 4) for the input stimulation pattern I_j . Then the marginal distribution of the data $y_{j,k}$, given the measurement error standard deviation σ , admits a density as follows:

$$\begin{aligned} p^*(y_{j,k} | I_j, a(\cdot, t_k), \sigma) &= \int N(y_{j,k} | P_{j,k}, \sigma^2 I) N(P_{j,k} | \mu_j, \Sigma_j) dP_{j,k} \\ &= N(y_{j,k} | \mu_j, \sigma^2 I + \Sigma_j) \end{aligned} \tag{5}$$

where we have used the shorthand of $N(\cdot | \mu_j, \Sigma_j)$ for the p.d.f. of $N(\mu_j, \Sigma_j)$. Eqn. 5 has the clear interpretation of inflating the measurement error covariance $\sigma^2 I$ by an additional amount Σ_j to reflect additional uncertainty due to discretisation error in the numerical solution of the PDE in Eqn. 1. This distinguishes the probabilistic approach from other applications of collocation methods in the solution of Bayesian PDE-constrained inverse problems, where uncertainty due to discretisation is ignored (Marzouk & Xiu 2009, Yan & Guo 2015). Eqn. 5 also appears in the emulation literature for static problems (e.g.

Calvetti et al. 2017). However, emulation methods treat the PDE as a perfect black-box and, as a result, the matrices Σ_j obtained from emulation do not reflect the fact that the PDE must be discretised³.

This paper proposes to base statistical inferences on the posterior distribution Π_n^* defined recursively via

$$\frac{d\Pi_n^*}{d\Pi_{n-1}^*}(\theta) \propto p^*(y_{\cdot,n} | \mathbb{I}, \exp(\theta_n), \sigma), \quad \Pi_0^* = \Pi_0.$$

In particular we will be most interested in the posterior predictive distribution $\Pi_{n,n+1}^*$ obtained with these probabilistic numerical methods, where discretisation uncertainty is explicitly modelled. Unlike $\Pi_{n,n+1}$, the posterior predictive distribution $\Pi_{n,n+1}^*$ can be exactly computed, since it does not require the exact solution of the PDE.

2.4 Theoretical Properties

The theoretical analysis of Cockayne et al. (2016a) can be exploited to assess the consistency of the probabilistic meshless method in Eqn. 4, in the case where the field a is fixed. Define the *fill distance* $h := \min\{h_A, h_B\}$ where $h_A = \sup_{x \in D} \min_i \|x - x_i^A\|_2$ and $h_B = \sup_{x \in \partial D} \min_i \|x - x_i^B\|_2$. Then we outline the following result:

Proposition 4. *Let B_ϵ denote a Euclidean ball of radius $\epsilon > 0$ centred on $\mathcal{P}u$ in \mathbb{R}^m , where u is the true solution of the PDE and \mathcal{P} was as previously defined. Then, under the assumptions of Cockayne et al. (2016a), which include that $H(k_u^0)$ is norm-equivalent to the Sobolev space $H^\beta(D)$, then the mass afforded to $\mathbb{R}^m \setminus B_\epsilon$ in the posterior $\mathcal{N}(\mu_j, \Sigma_j)$ is $O(\epsilon^{-2} h^{2\beta-4-d})$ for $h > 0$ sufficiently small.*

³The typical usage of emulators is to reduce the total number of forward problems that must be solved. This consideration is orthogonal to the present work and the two approaches could be combined.

This result ensures asymptotic agreement between the probabilistic numerical approach to the inverse problem and the (unavailable) exact approach based on the exact solution of the PDE in Eqn. 1 in the limit $h \rightarrow 0$ of infinite computation. Empirical evidence for the appropriateness of the uncertainty quantification for static EIT experiments and finite computation was presented in Cockayne et al. (2016a).

2.5 Computation via Sequential Monte Carlo

The log-normal prior on the conductivity field precludes a closed-form posterior. However, the filtering formulation of Sec. 2.2.2 suggests a natural approach to computation based on particle filters, otherwise known as sequential Monte Carlo (SMC) methods (Del Moral 2004).

Let $\Pi_{0,1} \ll \Pi'_{0,1}$ where $\Pi'_{0,1}$ is a user-chosen importance distribution on $C^1(\bar{D})$ (and could be Π_0). The method begins with N independent draws $\theta_0^{(1)}, \dots, \theta_0^{(N)}$ from $\Pi'_{0,1}$; each draw $\theta_0^{(i)}$ is associated with an importance weight

$$w_0^{(i)} \propto \frac{d\Pi_{0,1}}{d\Pi'_{0,1}}(\theta_0^{(i)})$$

such that $\sum_{i=1}^N w_0^{(i)} = 1$. This provides an empirical approximation $\sum_{i=1}^N w_0^{(i)} \delta(\theta_0^{(i)})$ to the prior marginal distribution $\Pi_{0,1}$ that becomes exact as N is increased. Let $t_0 := t_1$. Then, at each iteration $n = 1, 2, \dots$ of the SMC algorithm, the following steps are performed:

1. **Re-sample:** Particles $\tilde{\theta}_n^{(1)}, \dots, \tilde{\theta}_n^{(N)}$ are generated as a random sample (with replacement) of size N from the empirical distribution $\sum_{i=1}^N w_{n-1}^{(i)} \delta(\theta_{n-1}^{(i)})$.
2. **Move:** Each particle $\tilde{\theta}_n^{(i)}$ is updated to $\theta_n^{(i)}$ according to a Markov transition M_{n-1} that leaves Π_{n-1}^* invariant. (Details are provided in Appendix B.)

3. **Re-weight:** The next set of weights are defined as

$$w_n^{(i)} \propto p^*(y_{\cdot,n} | \mathbf{I}, \exp(\theta_n^{(i)}), \sigma) \frac{d\Gamma_{t_n - t_{n-1}}(\theta_n^{(i)} - \tilde{\theta}_n^{(i)})}{d\gamma}$$

and such that $\sum_{i=1}^N w_n^{(i)} = 1$.

The output after n iterations is an empirical approximation $\sum_{i=1}^N w_n^{(i)} \delta(\theta_n^{(i)})$ to the posterior distribution Π_n^* . The posterior predictive distribution $\Pi_{n,n+1}^*$ can be obtained from similar methods, as

$$\frac{d\Pi_{n,n+1}^*}{d\Pi_n^*}(\theta) \propto \frac{d\Gamma_{t_{n+1} - t_n}(\theta_{n+1} - \theta_n)}{d\gamma}. \quad (6)$$

The re-sample step in the above procedure does not in general need to occur at each iteration, only when the effective sample size is small; see Del Moral (2004). Theoretical analysis of SMC methods in the context of infinite-dimensional state spaces is provided in Beskos et al. (2015). For this work we considered a fairly standard SMC method, but several extensions are possible and include, in particular, stratified or quasi Monte Carlo re-sampling methods (Gerber & Chopin 2015). One extension which we explored was to introduce fictitious intermediate distributions between Π_{n-1}^* and Π_n^* following Chopin (2002), which we found to improve the performance of SMC in this context. For the experiments reported in the paper, 100 intermediate distributions were used, defined by tempering on a linear temperature ladder, c.f. Kantas et al. (2014), Beskos et al. (2015).

This completes our methodological development. Optimised Python scripts are available to reproduce these results at: https://github.com/jcockayne/hydrocyclone_code. Next, we report empirical results based on data from a controlled EIT experiment.

3 Results

This section considers data from a laboratory experiment designed to investigate the temporal mixing of two liquids. The experiment was conducted by one of the present authors and carefully controlled, to enable assessment of statistical methods and to mimic the salient features of industrial hydrocyclone equipments.

3.1 Experimental Protocol

In the experiment, a cylindrical perspex tank of diameter 15cm and height 30cm was used with a single ring of $m = 8$ electrodes, each measuring approximately 1cm wide by 3cm high. The electrodes start at the bottom of the tank, with the initial liquid level exactly at the top of the electrodes. Hence there is translation invariance in the vertical direction and the contents are effectively a single 2D region, meaning that electrical conductivity can be modelled as a 2D field. The experimental set-up is depicted in Fig. 2.

At the start of the experiment, a mixing impeller was used to create a rotational flow. This was then removed and, after a few seconds, concentrated potassium chloride solution was carefully injected into the tap water initially filling the tank. Data was then collected at regular time intervals until it was assumed that the liquid had fully mixed. Further details of the experiment can be found in West et al. (2005). These data were previously analysed (with non-probabilistic numerical methods) in Aykroyd & Cattle (2007).

This experiment mimics the situation when a hydrocyclone moves from an in-control regime to an out-of-control regime, in that initially there is a well defined core which gradually disappears as the liquids merge together. Performing the experiment in the laboratory allowed careful control of experimental conditions and, in particular, a lack of electrical interference from other equipment. A similar experimental set-up for data-

generation was recently employed in Hyvonen & Leinonen (2015).

There are several widely accepted data collection ‘protocols’ for EIT (Isaacson 1986). A protocol specifies the sequence of electrodes that are used to create the electric field, as well as the sequence of electrodes used to measure the resulting electric potential. In this experiment the ‘reference protocol’ was used, where a drive current is passed between a reference electrode and each of the other electrodes in turn allowing a maximum of $J = 7$ linearly independent current patterns. For each current pattern, the U_1, \dots, U_m were measured up to a common additive constant⁴, so that without loss of generality E_1 is the ‘reference’ electrode and $U_1 \equiv 0$. This permits a total of $7 \times 7 = 49$ measurements $y_{\cdot,k}$, obtained at each time point t_k in the experiment.

3.2 Experimental Results

The proposed statistical approach, based on probabilistic numerical methods, was used to make inferences on the unknown conductivity field a^\dagger based on this realistic experimental dataset.

The assumption that k_u^0 has two continuous derivatives is sufficient for the prior to be well-defined (Prop. 3). However, the theoretical result in Prop. 4 requires a more regular kernel with at least $\beta > 2 + d/2$ (weak) derivatives to ensure contraction of the (static) posterior. In reality, molecular diffusion implies that clear boundaries are not expected to be present in the true conductivity field. Thus it is reasonable to assume that both the conductivity field a and the electrical potential u will be fairly smooth in the interior D . For these reasons, the kernels employed for experiments below were of squared-exponential form, since this trivially meets all smoothness requirements, including smoothness of the solution u in D .

⁴This reflects the fact that it is voltage that is actually measured, which is the difference of two potentials.

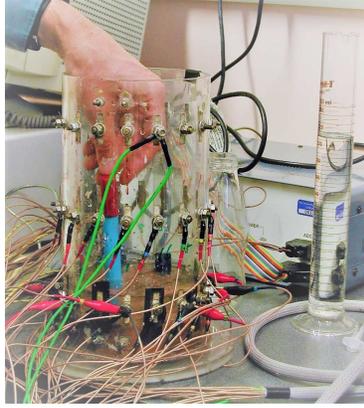


Figure 2: Experimental set-up: A cylindrical perspex tank containing tap water was stirred before an amount of potassium chloride was injected. Electrodes positioned around the tank measured voltages, which can be related through a partial differential equation to the internal conductivity field. The inverse problem consists of estimating the internal conductivity field from the voltages that were measured. (Only the bottom ring of electrodes were used for the data analysed in this paper.) Photo reproduced from West et al. (2005).

3.2.1 Static Recovery Problem

First, we calibrated our probabilistic numerical methods by analysing the static recovery problem. This prior for θ was taken to be Gaussian, with a mean of zero and a squared-exponential covariance function

$$k_a(\mathbf{x}, \mathbf{x}') := \varphi_a \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|_2^2}{2\ell_a^2}\right)$$

where φ_a controls the magnitude of fields drawn from the prior, while the length-scale ℓ_a controls how rapidly those functions vary. Since the main aim here is to assess the probabilistic meshless method, rather than the performance at state estimation, we simply fixed $\varphi_a = 1$ and $\ell_a = 0.3$. Note that while it is common in EIT problems to use priors

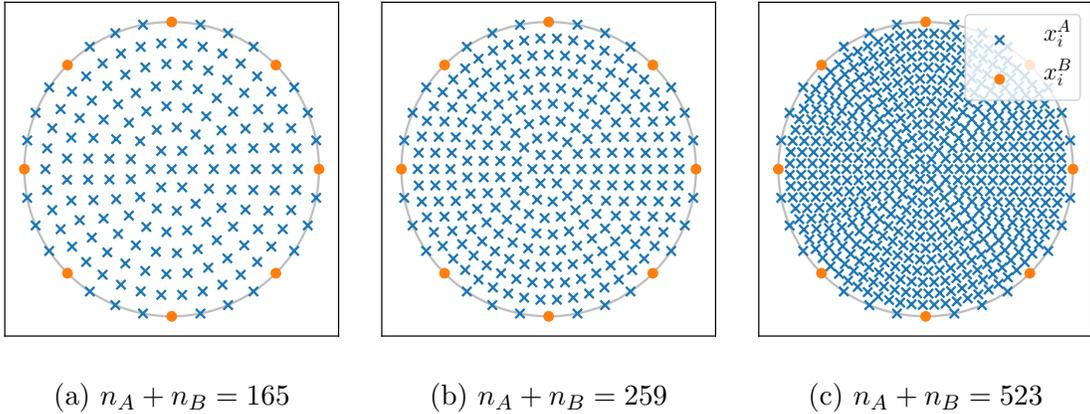


Figure 3: Typical sets of collocation points x_i^A and x_i^B that were used to discretise the PDE.

which promote hard edges in drawn samples, owing to applications in medicine, here a smooth prior is appropriate. For all experiments in this paper the parameter σ , that describes technical measurement error, was set to $\sigma = 1.0$ based on analysis of a technical replicate dataset. For the probabilistic meshless method, the prior model was centered and a squared-exponential covariance function was used, with $\varphi_u = 100$ to match the scale of measurements in the dataset, and $\ell_u = 0.211$, a value chosen by empirical Bayes based upon a high-quality reference sample. The collocation points were chosen on concentric circles, as shown in Fig. 3 for increasing values of n_A and n_B .

For illustration, we first considered simulated data and a coarse collocation method which did *not* model discretisation error. This was compared to a reference posterior, obtained using a brute-force symmetric collocation forward solver with a large number of collocation points. (Of course, it is impractical to use a large number of collocation points in the applied context due to the associated computational cost.) The result, shown in Fig. 4, was a posterior that did not contain the true data-generating field a^\dagger in its region

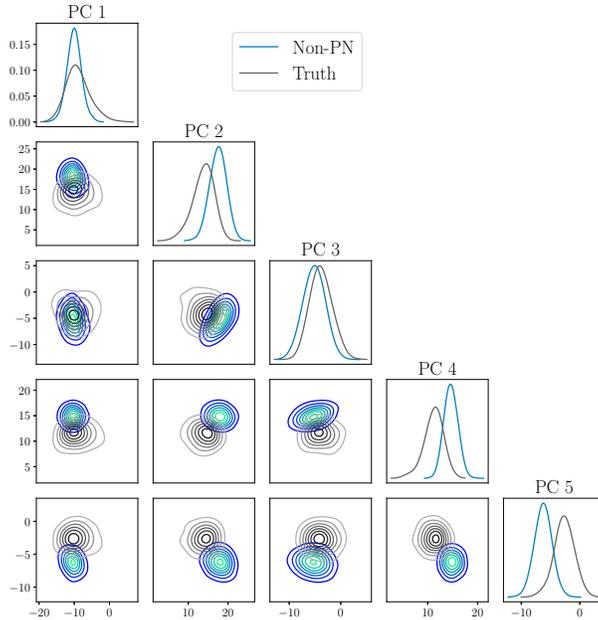


Figure 4: Failure case: Here a small number $n_A + n_B = 165$ of collocation points was used to discretise the PDE, but the uncertainty due to discretisation was not modelled. The reference posterior distribution (grey) was compared to the approximation to the posterior obtained when the PDE is discretised (blue) and the discretisation error is not modelled (“Non-PN”). Projections onto principal components (PC) of the reference posterior are displayed. It is observed that the approximated posterior is highly biased.

of support. This result was observed to be typical for $n_A + n_B < 250$ and motivates the formal uncertainty quantification for discretisation error that is provided by probabilistic numerical methods in this paper.

The experimental data used for this assessment were obtained as a single frame (time point 14) from the larger temporal dataset. In Fig. 5(a) we show the posterior mean estimate, together with its posterior variance, for a reference conductivity field generated

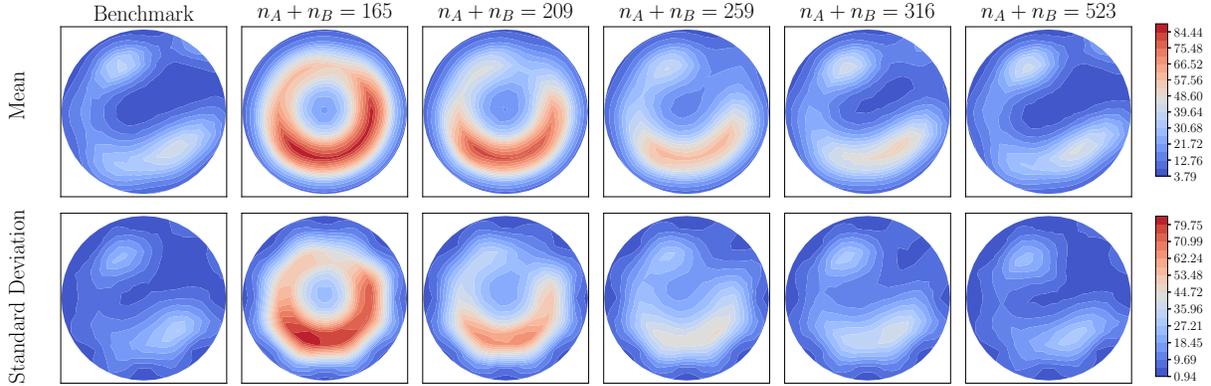


Figure 5: Posterior means and standard-deviations for the recovered conductivity field. The first column represents the reference solution, obtained using a symmetric collocation forward solver with a large number of collocation points. The remaining columns represent the recovered field when probabilistic numerical methods are used based on $n_A + n_B$ collocation points as illustrated in Fig. 3.

using a high-quality symmetric collocation forward solver with $n_A + n_B = 207$ collocation points. Adjacent, in Fig. 5 we show the posterior mean and variance for the conductivity field obtained with probabilistic numerical methods for increasing values of $n_A + n_B$. It is seen that both the posterior mean and posterior standard deviation produced with the probabilistic numerical method converge to the reference posterior as the number of collocation points is increased. However, at coarse resolution, the posterior variance is inflated to reflect the contribution of a discretisation uncertainty to each numerical solution of the forward problem. This provides automatic protection against the erroneous results seen in Fig. 4.

In Fig. 6 we plot the number $n_A + n_B$ of collocation basis points versus the integrated posterior standard-deviation for the unknown conductivity field. These results demon-

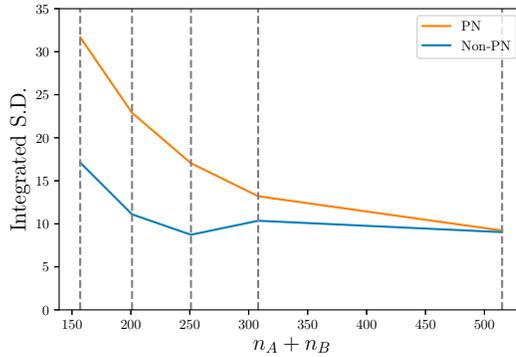


Figure 6: Posterior standard-deviation for the conductivity field, integrated over the domain D , as a function of the number $n_A + n_B$ of collocation points. The blue curve represents the standard case where error due to discretisation of the PDE is not quantified (“Non-PN”) whilst the red curve represents the case where a probabilistic numerical method is used to provide uncertainty quantification for the PDE solution itself (“PN”).

strate the computation-precision trade-off that is made possible with probabilistic numerical methods, and are consistent with the preliminary investigation in Cockayne et al. (2016*a,b*). Next, we turn to the temporal problem that motivates this research.

3.2.2 Temporal Recovery Problem

For this experiment, data were obtained at 49 regular time intervals. Times 1-10 were obtained before injection of the potassium chloride solution, while the injection occurred rapidly, between frames 10 and 11. The remaining time points 12-49 capture the diffusion and rotation of the liquids, which is the behaviour that we hope to recover.

The parameter λ controls the temporal smoothness of the conductivity field in the prior model. Three fixed values, $\lambda \in \{10, 100, 1000\}$ were considered in turn, representing

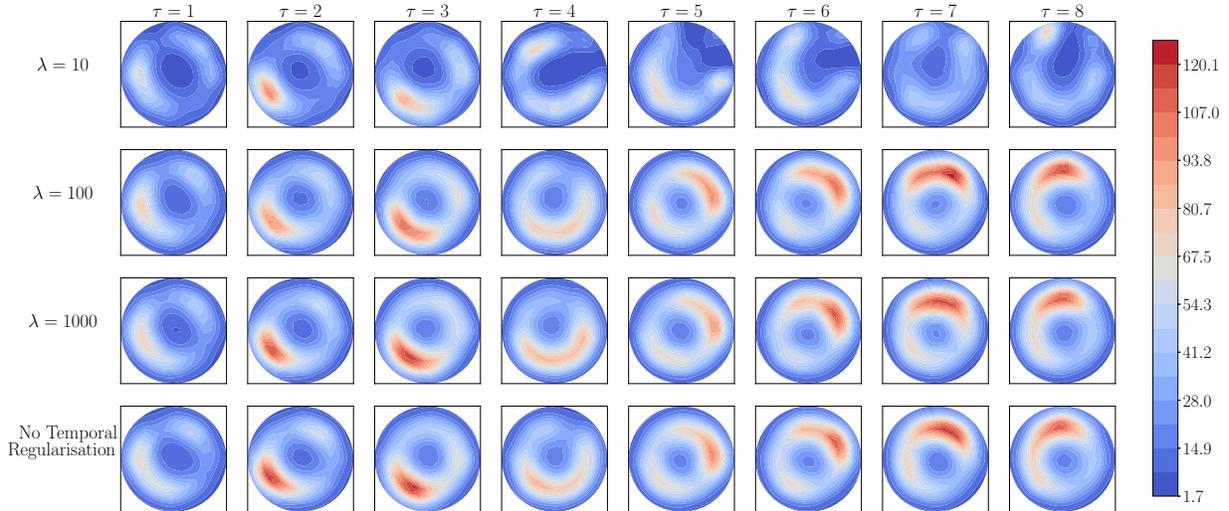


Figure 7: Posterior mean for the conductivity field $a(\cdot, t)$, shown as a function of the time index. Here we consider the dependence of the recovered field on the choice of the temporal covariance parameter λ . The value $n_A + n_B = 209$ was used.

decreasing levels of smoothness. The case of no temporal regularisation was also displayed. Our method was applied to estimate the time-evolution of the field. Results are shown in Fig. 7. The counter-clockwise rotation of the fluid was first clearly seen for $\lambda = 100$, whilst the value $\lambda = 10$ represented too much temporal regularisation, which caused this information to be lost. On the other hand, the predictive posterior in Eqn. 6 is trivial in the limit of large λ , so that in our context smaller values of λ are preferred. It is expected that an analogous calibration can be performed in the real-world context.

To assess whether the problems of bias and over-confidence due to discretisation can be mitigated in the temporal context, where discretisation errors are propagated and accumulate over time, we fixed $\lambda = 100$ and inspected the posterior over the coefficients ψ_i at the final time point t_n . Results in Fig. 8 confirmed that the posterior Π_n^* (red) was inflated

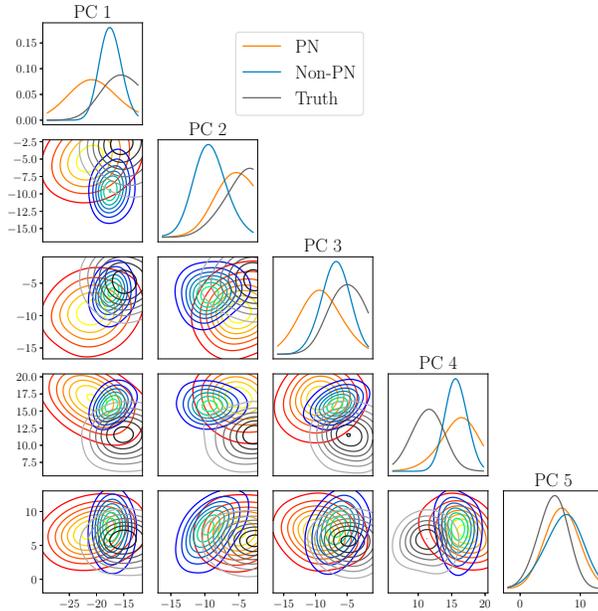


Figure 8: Posterior distribution over the coefficients ψ_i , at the final time point t_n . Here a small number $n_A + n_B = 165$ of collocation points was used to discretise the PDE. The reference posterior distribution (grey) was compared to the approximation to the posterior obtained when discretisation of the PDE is not modelled (“Non-PN”) and modelled (“PN”). Projections onto principal components (PC) of the reference posterior are displayed.

relative to the standard approximate posterior (blue) and tended to cover more of the true posterior Π_n (grey) in its effective support. This provides empirical evidence to support the use of the proposed posterior Π_n^* .

In Fig. 9 we again plot the number $n_A + n_B$ of collocation basis points versus the integrated posterior standard-deviation for the unknown conductivity field, again at the final time point. These results demonstrate a computation-precision trade-off similar to that which was observed for the static recovery problem. Compared to the static recovery

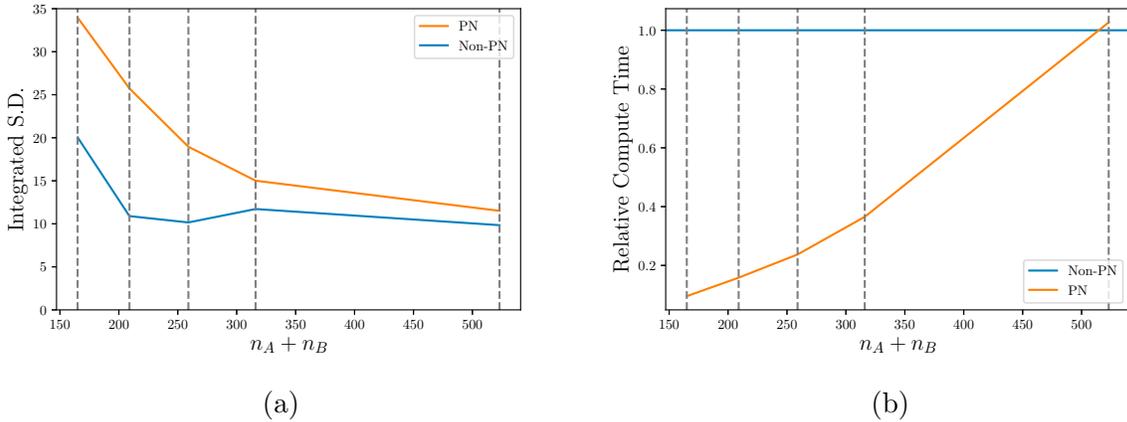


Figure 9: Temporal recovery problem: (a) Posterior standard-deviation for the conductivity field $a(\cdot, t_n)$ at the final time point t_n , integrated over the domain D . (b) Computational time required by the proposed method, relative to the (non-probabilistic) symmetric collocation method applied on a resolved grid.

problem in Fig. 6, however, we observed greater inflation of the posterior standard deviation when probabilistic numerical methods were used. This reflects the fact that we have constructed a full probability model for the effect of discretisation error, which is able to capture how these errors propagate and accumulate within the computational output.

4 Discussion

The motivation for this research was industrial process monitoring, but the associated methodological development was general. In particular, we addressed the important topic of how to perform uncertainty quantification for numerical error due to discretisation of the physical governing equations specified through a PDE. Typically this source of error is

ignored, or its contribution bounded through detailed numerical analysis, such as Schwab & Stuart (2012). However, in the temporal setting, theoretical bounds are difficult to obtain due to propagation and accumulation of errors, so that it is unclear how to proceed.

In this work we proposed a statistical solution, wherein a probabilistic numerical method was used to provide uncertainty quantification for the discretisation error associated with a collocation-type numerical method. Aided by sequential Monte Carlo sampling methods, it was shown how this model for *discretisation uncertainty* can be employed in the temporal context. The result was a more comprehensive quantification of uncertainty, that accounts for both statistical uncertainty and for propagation and accumulation of discretisation uncertainty in the final output. For our motivating industrial application, this work is expected to facilitate more reliable anticipation and pro-active control of the hydrocyclone, to ensure safety in operation (Bradley 2013). Beyond that, it is anticipated that the mitigation of bias and over-confidence observed in our experimental results is a feature of probabilistic numerical methods in general.

The methods that we pursued differ in a fundamental sense to techniques that seek to *emulate* the forward model. Emulation, as well as dimension reduction methods, have been widely used in static recovery problems to reduce the computational cost of repeatedly solving the governing PDE (Marzouk et al. 2007, Marzouk & Najm 2009, Cotter et al. 2010, Schwab & Stuart 2012, Cui et al. 2016, Chen & Schwab 2016*c,b*). Notably Stuart & Teckentrup (2016), Calvetti et al. (2017) considered integrating the emulator uncertainty into inference for model parameters. However, to train an emulator it is usually required to have access to a training set of parameters a for which the exact solution u of the PDE is provided. Thus the focus of emulation is related to generalisation in the a domain, as opposed to quantification of discretisation uncertainty in the u domain. An interesting extension of this work would be to combine these two complementary techniques; this

would be expected to reduced the computational cost of the proposed method.

The principal limitation of our approach was that a Markov temporal evolution of the conductivity field $a(\cdot, t)$ was assumed. Physical consideration suggest that the Markov assumption is incorrect, since time-derivatives of all orders of this field will vary continuously and thus encode information that is useful. However, it is not clear how this information can be encoded into a prior for the temporal evolution of the conductivity field whilst preserving the computationally convenient filtering framework. On the other hand, the Markov prior can be statistically justified in that it represents an encoding of partial prior information (Potter & Anderson 1983). It remains a problem for future work to investigate the potential loss of estimation and predictive efficiency as a result of encoding only partial information into the prior model.

The second limitation we highlight is that the prior model for the potential field $u(\cdot, t)$ did not include a temporal component. This choice was algorithmically convenient, as it de-coupled each of the forward problems of solving the PDE, such that each time a probabilistic numerical method was called, it could be implemented “out of the box”. Nevertheless, a temporal covariance structure in the parameter a implies that there also exists such structure in u and the effect of not encoding this aspect of prior information should be further investigated.

Overall, we are excited by the prospect of new and more powerful methods for uncertainty quantification that can deal with both statistical and discretisation error in a unified analytical framework.

Acknowledgements: The authors are grateful for detailed suggestions from the Associate Editor and two anonymous Reviewers. For the numerical results reported in Section 3 we thank T. J. Sullivan for the use of computing facilities at the Freie Universität Berlin,

funded by the Excellence Initiative of the German Research Foundation.

A Proof of Results in the Main Text

Proof of Proposition 1. Let $L_H^2 = \{v : \bar{D} \times \Omega \rightarrow \mathbb{R} \text{ s.t. } \mathbb{E}\|v\|_H^2 < \infty\}$, which is a Banach space with norm $(\mathbb{E}\|\cdot\|_H^2)^{1/2}$; see section 2.4 of Dashti & Stuart (2016). Following Thm. 2.10 in Dashti & Stuart (2016), consider the partial sums

$$\log a^N(\cdot, t) = \sum_{i=1}^N i^{-\alpha} \psi_i(t) \phi_i(\cdot).$$

For $N > M$ we have

$$\begin{aligned} \mathbb{E}\|\log a^N(\cdot, t) - \log a^M(\cdot, t)\|_H^2 &= \mathbb{E} \sum_{i=M+1}^N i^{-2\alpha} |\psi_i(t)|^2 \\ &\leq \sum_{i=M+1}^N i^{-2\alpha} [(m_\psi^{\max})^2 + k_\psi^{\max}] \\ &\leq [(m_\psi^{\max})^2 + k_\psi^{\max}] \sum_{i=M+1}^{\infty} i^{-2\alpha}. \end{aligned}$$

Since $\alpha > 1/2$, the RHS vanishes as $M \rightarrow \infty$. Thus, as L_H^2 is a Banach space, $\log a(\cdot, t)$ exists as an L_H^2 limit. It follows that $\log a(\cdot, t)$, and hence $a(\cdot, t)$, takes values almost surely in $C^1(\bar{D})$. \square

Proof of Corollary 1. From direct algebra:

$$\begin{aligned} \theta_\Delta(x) = \theta(x, t+s) - \theta(x, t) &= \sum_{i=1}^{\infty} i^{-\alpha} [\psi_i(t+s) - \psi_i(t)] \phi_i(x) \\ &= \sqrt{\lambda(s+\tau)} \sum_{i=1}^{\infty} i^{-\alpha} \xi_i \phi_i(x) \end{aligned}$$

where the ξ_i are independent $N(0, 1)$. From the Karhounen-Loève theorem (Loève 1977), this is recognised as a Gaussian random field with mean function $m_\Delta(x) = 0$ and covariance function

$$\begin{aligned} k_\Delta(x, x') &= \lambda(s + \tau) \sum_{i=1}^{\infty} i^{-2\alpha} \phi_i(x) \phi_i(x') \\ &= \lambda(s + \tau) k_\phi(x, x'), \end{aligned}$$

as claimed. □

Proof of Proposition 3. Let $H(k)$ denote the reproducing kernel Hilbert space associated with a kernel k . In Cialenco et al. (2012), Lemma 2.2, it was established that a generic integral-type kernel k_u , as in Eqn. 3, corresponds to the covariance function for a Gaussian process that takes values almost surely in $H(k_u^0)$. To complete the proof, Corr. 4.36 (p131) in Steinwart & Christmann (2008) establishes that if $k_u^0 \in C^{\beta \times \beta}(\bar{D} \times \bar{D})$ then $H(k_u^0) \subset C^\beta(\bar{D})$. □

Proof of Proposition 4. Let $\mu(x)$ and $\sigma(x)$ denote, respectively, the posterior mean and standard deviation of $u(x)$ under the probabilistic meshless method. Prop. 4.1 of Cockayne et al. (2016a) established that the posterior mean $\mu(x)$ satisfies $|\mu(x) - u(x)| \leq \sigma(x) \|u\|_{H(k_u)}$ and Prop. 4.2 of Cockayne et al. (2016a) established that the posterior standard deviation $\sigma(x)$ satisfies $\sigma(x) \leq Ch^{\beta-2-d/2}$ for some constant C independent of $x \in D$. In particular we have $\|\mu_j - \mathcal{P}u\|_2 = O(h^{\beta-2-d/2})$. Lastly, Thm. 4.3 of Cockayne et al. (2016a) established a generic rate of contraction for the mass of a Gaussian distribution of $O(\epsilon^{-2}h^{2\beta-4-d})$, as required. (Note that these results are specific consequences of more general results found in Lem. 3.4 in Cialenco et al. (2012) and Secs. 11.3 and 16.3 of Wendland (2005).) □

B Details of the Markov Kernel Used

This appendix contains a description of the Markov transition kernel that was used. Indeed, for the Markov transition kernel M_{n-1} used in the **Move** step in Section 2.5, we employed the pre-conditioned Crank–Nicholson method. This will now be described.

Let Π be a probability distribution on a measurable space (Θ, \mathcal{B}) , such that the Radon–Nikodym derivative $d\Pi/d\Pi_0$ is well-defined for a fixed reference distribution Π_0 . Recall that a Markov transition kernel M which leaves Π invariant is a function $M : \Theta \times \mathcal{B} \rightarrow [0, 1]$ such that

1. the map $\theta \mapsto M(\theta, B)$ is \mathcal{B} -measurable for all $B \in \mathcal{B}$
2. the map $B \mapsto M(\theta, B)$ is a probability measure on (Θ, \mathcal{B}) for all $\theta \in \Theta$
3. invariance; $\Pi(B) = \int M(\theta, B)d\Pi(\theta)$ for all $B \in \mathcal{B}$.

The pre-conditioned Crank–Nicholson method (with step size $\beta \in (0, 1)$)

$$\begin{aligned} \theta^* &:= \sqrt{(1 - \beta^2)}\theta + \beta\xi, & \xi &\sim \Pi_0 \\ \theta_{\text{new}} &= \begin{cases} \theta^* & \text{with probability } \alpha(\theta, \theta^*) = \min \left\{ 1, \frac{d\Pi}{d\Pi_0}(\theta^*) / \frac{d\Pi}{d\Pi_0}(\theta) \right\} \\ \theta & \text{otherwise} \end{cases} \end{aligned}$$

for generating the next state θ_{new} of the Markov chain, given the current state is θ , corresponds to a Markov transition kernel

$$M(\theta, B) = \int 1[\theta^* \in B]\alpha(\theta, \theta^*) + 1[\theta \in B](1 - \alpha(\theta, \theta^*))d\Pi_0(\xi)$$

that leaves Π invariant. The associated Markov chain has been shown to have non-vanishing acceptance probability when Θ is a Hilbert space and Π_0 is a Gaussian distribution (Thm. 6.4 of Cotter et al. 2013). This was the Markov transition kernel that we employed, with

β tuned to achieve an acceptance rate of between 10%–25%, Θ being the state space of θ_n and Π_0 being the prior $\Pi_{0,1}$, defined in the main text. Nevertheless, it is not the unique Markov transition kernel that could be used; see Cotter et al. (2013) for several examples of Markov transition kernels that are well-defined in the Hilbert space context.

References

- Aykroyd, R. (2018), ‘A statistical approach to the inclusion of electrode contact impedance uncertainty in electrical tomography reconstruction’, *International Journal of Tomography and Simulation* **31**(1), 56–67.
- Aykroyd, R. G. (2015), *Industrial Tomography: Systems and Applications*, Woodhead Publishing, chapter Statistical Image Reconstruction, pp. 401–428.
- Aykroyd, R. G. & Cattle, B. A. (2007), ‘A boundary-element approach for the complete-electrode model of EIT illustrated using simulated and real data’, *Inverse Problems in Science and Engineering* **15**(5), 441–461.
- Beskos, A., Jasra, A., Muzaffer, E. A. & Stuart, A. M. (2015), ‘Sequential Monte Carlo methods for Bayesian elliptic inverse problems’, *Statistics and Computing* **25**(4), 727–737.
- Bradley, D. (2013), *The Hydrocyclone: International Series of Monographs in Chemical Engineering*, Vol. 4, Elsevier.
- Briol, F.-X., Oates, C. J., Girolami, M., Osborne, M. A. & Sejdinovic, D. (2018), ‘Probabilistic integration: A role in statistical computation? (with discussion)’, *Statistical Science*. To appear.

- Bui-Thanh, T. & Ghattas, O. (2014), ‘An analysis of infinite dimensional Bayesian inverse shape acoustic scattering and its numerical approximation’, *SIAM/ASA Journal on Uncertainty Quantification* **2**(1), 203–222.
- Calvetti, D., Dunlop, M. M., Somersalo, E. & Stuart, A. M. (2017), ‘Iterative updating of model error for Bayesian inversion’, *arXiv:1707.04246* .
- Chen, P. & Schwab, C. (2015), ‘Sparse-grid, reduced-basis Bayesian inversion’, *Computer Methods in Applied Mechanics and Engineering* **297**, 84–115.
- Chen, P. & Schwab, C. (2016a), Adaptive sparse grid model order reduction for fast Bayesian estimation and inversion, in ‘Sparse Grids and Applications (Stuttgart 2014)’, Springer, pp. 1–27.
- Chen, P. & Schwab, C. (2016b), ‘Model order reduction methods in computational uncertainty quantification’, *Handbook of Uncertainty Quantification*, Springer .
- Chen, P. & Schwab, C. (2016c), ‘Sparse-grid, reduced-basis Bayesian inversion: Nonaffine-parametric nonlinear equations’, *Journal of Computational Physics* **316**, 470–503.
- Cheng, K.-S., Isaacson, D., Newell, J. C. & Gisser, D. G. (1989), ‘Electrode models for electric current computed tomography’, *IEEE Transactions on Biomedical Engineering* **36**(9), 918–924.
- Chkrebtii, O. A., Campbell, D. A., Calderhead, B. & Girolami, M. A. (2016), ‘Bayesian solution uncertainty quantification for differential equations’, *Bayesian Analysis* **11**(4), 1239–1267.
- Chopin, N. (2002), ‘A sequential particle filter method for static models’, *Biometrika* **89**(3), 539–552.

- Cialenco, I., Fasshauer, G. E. & Ye, Q. (2012), ‘Approximation of stochastic partial differential equations by a kernel-based collocation method’, *International Journal of Computer Mathematics* **89**(18), 2543–2561.
- Cockayne, J., Oates, C., Sullivan, T. & Girolami, M. (2016a), ‘Probabilistic meshless methods for Bayesian inverse problems’, *arXiv:1605.07811* .
- Cockayne, J., Oates, C., Sullivan, T. & Girolami, M. (2016b), Probabilistic numerical methods for PDE-constrained Bayesian inverse problems, *in* G. Verdoolaege, ed., ‘Proceedings of the 36th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering’, Vol. 1853 of *AIP Conference Proceedings*.
- Cockayne, J., Oates, C., Sullivan, T. & Girolami, M. (2017), ‘Bayesian probabilistic numerical methods’, *arXiv:1702.03673* .
- Conrad, P. R., Girolami, M., Särkkä, S., Stuart, A. & Zygalakis, K. (2017), ‘Statistical analysis of differential equations: Introducing probability measures on numerical solutions’, *Statistics and Computing* **27**(4), 1065–1082.
- Cotter, S. L., Dashti, M. & Stuart, A. M. (2010), ‘Approximation of Bayesian inverse problems for PDEs’, *SIAM Journal on Numerical Analysis* **48**(1), 322–345.
- Cotter, S. L., Roberts, G. O., Stuart, A. M., White, D. et al. (2013), ‘MCMC methods for functions: Modifying old algorithms to make them faster’, *Statistical Science* **28**(3), 424–446.
- Cui, T., Marzouk, Y. & Willcox, K. (2016), ‘Scalable posterior approximations for large-scale Bayesian inverse problems via likelihood-informed parameter and state reduction’, *Journal of Computational Physics* **315**, 363–387.

- Dashti, M. & Stuart, A. M. (2016), *Handbook of Uncertainty Quantification*, Springer, chapter The Bayesian approach to inverse problems, pp. 311–428.
- Del Moral, P. (2004), *Feynman-Kac Formulae*, Springer.
- Dunlop, M. M. & Stuart, A. M. (2016), ‘The Bayesian formulation of EIT: Analysis and algorithms’, *Inverse Problems and Imaging* **10**(4), 1007–1036.
- Fasshauer, G. E. (1996), Solving partial differential equations by collocation with radial basis functions, in A. L. Méhauté, C. Rabut & L. L. Schumaker, eds, ‘Surface Fitting and Multiresolution Methods. Vol. 2 of the Proceedings of the 3rd International Conference on Curves and Surfaces held in Chamonix-Mont-Blanc, June 27-July 3’, Vanderbilt University Press, Nashville, TN, pp. 131–178.
- Gerber, M. & Chopin, N. (2015), ‘Sequential quasi Monte Carlo’, *Journal of the Royal Statistical Society: Series B* **77**(3), 509–579.
- Gross, L. (1967), Abstract Wiener spaces, in ‘Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory, Part 1’, The Regents of the University of California.
- Gutierrez, J., Dyakowski, T., Beck, M. & Williams, R. (2000), ‘Using electrical impedance tomography for controlling hydrocyclone underflow discharge’, *Powder Technology* **108**(2), 180–184.
- Hadamard, J. (1902), ‘Sur les problèmes aux dérivées partielles et leur signification physique’, *Princeton University Bulletin* pp. 49–52.
- Halmos, P. & Savage, L. (1949), ‘Application of the Radon-Nikodym theorem to the theory of sufficient statistics’, *The Annals of Mathematical Statistics* **20**(2), 225–241.

- Hamilton, S. J. & Hauptmann, A. (2018), ‘Deep D-bar: Real time electrical impedance tomography imaging with deep neural networks’, *IEEE Transactions on Medical Imaging*, to appear .
- Hennig, P., Osborne, M. A. & Girolami, M. (2015), Probabilistic numerics and uncertainty in computations, in ‘Proceedings of the Royal Society A’, Vol. 471, The Royal Society.
- Hyvonen, N. & Leinonen, M. (2015), ‘Stochastic Galerkin finite element method with local conductivity basis for electrical impedance tomography’, *SIAM/ASA Journal on Uncertainty Quantification* **3**(1), 998–1019.
- Isaacson, D. (1986), ‘Distinguishability of conductivity by electric current computed tomography’, *IEEE Trans. Med. Imaging* **MI-5**, No. 2, 91–95.
- Kaipio, J. P., Kolehmainen, V., Somersalo, E. & Vauhkonen, M. (2000), ‘Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography’, *Inverse Problems* **16**(5), 1487.
- Kaipio, J. P., Kolehmainen, V., Vauhkonen, M. & Somersalo, E. (1999), ‘Inverse problems with structural prior information’, *Inverse Problems* **15**(3), 713.
- Kantas, N., Beskos, A. & Jasra, A. (2014), ‘Sequential Monte Carlo methods for high-dimensional inverse problems: A case study for the Navier–Stokes equations’, *SIAM/ASA Journal on Uncertainty Quantification* **2**(1), 464–489.
- Kennedy, M. C. & O’Hagan, A. (2001), ‘Bayesian calibration of computer models’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **63**(3), 425–464.
- Lauritzen, S. L. (1996), *Graphical Models*, Clarendon Press.

- Law, K., Stuart, A. & Zygalakis, K. (2015), *Data Assimilation: A Mathematical Introduction*, Vol. 62, Springer.
- Loève, M. (1977), ‘Probability theory’.
- Lord, G., Powell, C. & Shardlow, T. (2014), *An introduction to computational stochastic PDEs*, number 50, Cambridge University Press.
- Marzouk, Y. M. & Najm, H. N. (2009), ‘Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems’, *Journal of Computational Physics* **228**(6), 1862–1902.
- Marzouk, Y. M., Najm, H. N. & Rahn, L. A. (2007), ‘Stochastic spectral methods for efficient Bayesian solution of inverse problems’, *Journal of Computational Physics* **224**(2), 560–586.
- Marzouk, Y. & Xiu, D. (2009), ‘A stochastic collocation approach to Bayesian inference in inverse problems’, *Communications in Computational Physics* **6**(4), 826–847.
- Nagel, J. B. & Sudret, B. (2016), ‘Spectral likelihood expansions for Bayesian inference’, *Journal of Computational Physics* **309**, 267–294.
- Nouy, A. & Soize, C. (2014), ‘Random field representations for stochastic elliptic boundary value problems and statistical inverse problems’, *European Journal of Applied Mathematics* **25**(03), 339–373.
- Novak, E. & Wozniakowski, H. (2008), *Tractability of Multivariate Problems. Vol. I: Linear Information*, Vol. 6 of *EMS Tracts in Mathematics*, European Mathematical Society (EMS), Zürich.

- Novak, E. & Wozniakowski, H. (2010), *Tractability of Multivariate Problems. Vol II: Standard Information for Functionals*, Vol. 12 of *EMS Tracts in Mathematics*, European Mathematical Society (EMS), Zürich.
- Owhadi, H. (2015), ‘Bayesian numerical homogenization’, *Multiscale Modeling & Simulation* **13**(3), 812–828.
- Owhadi, H. (2017), ‘Multi-grid with rough coefficients and multiresolution operator decomposition from hierarchical information games’, *SIAM Review* **59**(1), 99–149. To appear.
- Owhadi, H. & Zhang, L. (2017), ‘Gamblers for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients’, *Journal of Computational Physics* **347**, 99–128.
- Polydorides, N. & Lionheart, W. R. (2002), ‘A MATLAB toolkit for three-dimensional electrical impedance tomography: A contribution to the Electrical Impedance and Diffuse Optical Reconstruction Software project’, *Measurement Science and Technology* **13**(12), 1871.
- Potter, J. & Anderson, B. (1983), ‘Statistical inference with partial prior information’, *IEEE Transactions on Information Theory* **29**(5), 688–695.
- Quarteroni, A. & Valli, A. (2008), *Numerical Approximation of Partial Differential Equations*, Springer Science & Business Media.
- Raissi, M., Perdikaris, P. & Karniadakis, G. (2017), ‘Inferring solutions of differential equations using noisy multi-fidelity data’, *Journal of Computational Physics* **335**, 736–746.
- Särkkä, S. (2013), *Bayesian Filtering and Smoothing*, Vol. 3, Cambridge University Press.

- Schillings, C. & Schwab, C. (2013), ‘Sparse, adaptive Smolyak quadratures for Bayesian inverse problems’, *Inverse Problems* **29**(6), 065011.
- Schillings, C. & Schwab, C. (2014), ‘Sparsity in Bayesian inversion of parametric operator equations’, *Inverse Problems* **30**(6), 065007.
- Schwab, C. & Stuart, A. M. (2012), ‘Sparse deterministic approximation of Bayesian inverse problems’, *Inverse Problems* **28**(4), 045003.
- Somersalo, E., Cheney, M. & Isaacson, D. (1992), ‘Existence and uniqueness for electrode models for electric current computed tomography’, *SIAM Journal on Applied Mathematics* **52**(4), 1023–1040.
- Sripriya, R., Kaulaskar, M., Chakraborty, S. & Meikap, B. (2007), ‘Studies on the performance of a hydrocyclone and modeling for flow characterization in presence and absence of air core’, *Chemical Engineering Science* **62**(22), 6391–6402.
- Steinwart, I. & Christmann, A. (2008), *Support Vector Machines*, Springer Science & Business Media.
- Stuart, A. M. (2010), ‘Inverse problems: A Bayesian perspective’, *Acta Numerica* **19**, 451–559.
- Stuart, A. M. & Teckentrup, A. L. (2016), ‘Posterior consistency for Gaussian process approximations of Bayesian posterior distributions’, *arXiv:1603.02004* .
- Tikhonov, A. & Arsenin, V. (1977), *Solutions of Ill-Posed Problems*, New York: Winston.
- Todescato, M., Carron, A., Carli, R., Pillonetto, G. & Schenato, L. (2017), ‘Efficient spatio-temporal Gaussian regression via Kalman filtering’, *arXiv:1705.01485* .

- Vauhkonen, M., Lionheart, W. R., Heikkinen, L. M., Vauhkonen, P. J. & Kaipio, J. P. (2001), ‘A MATLAB package for the EIDORS project to reconstruct two-dimensional EIT images’, *Physiological Measurement* **22**(107).
- Watzenig, D. & Fox, C. (2009), ‘A review of statistical modelling and inference for electrical capacitance tomography’, *Measurement Science and Technology* **20**(5), 052002.
- Wendland, H. (2005), *Scattered Data Approximation*, Vol. 17 of *Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press.
- Werschulz, A. G. (1996), ‘The complexity of definite elliptic problems with noisy data’, *Journal of Complexity* **12**(4), 440–473.
- West, R. M., Meng, S., Aykroyd, R. G. & Williams, R. A. (2005), ‘Spatial-temporal modeling for electrical impedance imaging of a mixing process’, *Review of Scientific Instruments* **76**(7), 073703.
- Wiener, N. (1949), *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, MIT Press Cambridge.
- Yamasaki, Y. (1985), *Measures on Infinite Dimensional Spaces*, World Scientific, Singapore.
- Yan, L. & Guo, L. (2015), ‘Stochastic collocation algorithms using ℓ_1 -minimization for Bayesian solution of inverse problems’, *SIAM Journal on Scientific Computing* **37**(3), A1410–A1435.