Probabilistic Numerical Methods for Non-Linear Partial Differential Equations: Strong Form Solutions

Jon Cockayne[1], Chris J. Oates[2], Tim Sullivan[3], Mark Girolami[4]

1: University of Warwick, 2: University of Technology Sydney, 3: Free University of Berlin, 4: Alan Turing Institute for Data Science

Abstract

Recent work by (1) develops methods for capturing and propagating discretisation error in the numerical solution of Ordinary and Partial Differential Equations.

We are developing similar models for **strong form solution of linear and nonlinear Partial Differential Equations**. This allows us to propagate the error through other numerical procedures.

In particular, we study solution of inverse problems. Using a probabilistic model for the solution, we are able to capture the numerical error in solving the forward problem in the inferences drawn from the inverse problem.

Classical Solvers

We wish to solve a system of operator equations for $u(\boldsymbol{x})$ defined on Ω (with boundary $\partial\Omega$):

$$\begin{aligned} \mathcal{A}u(\mathbf{x}) &= g(\mathbf{x}) & \mathbf{x} &\in \Omega \\ \mathcal{B}u(\mathbf{x}) &= 0 & \mathbf{x} &\in \partial\Omega \end{aligned}$$

(NB we can easily adapt this for nonzero BCs). The most widely used classical system for solving such equations is the Finite Element method. We construct a *fine mesh* over the problem domain and construct an approximation like:

$$\hat{u}(\mathbf{x}) = \sum_i w_i \phi_i(\mathbf{x})$$

Under certain conditions $|\hat{u} - u| \to 0$ as the mesh width $h \to 0$. However, making the mesh arbitrarily fine is computationally infeasible in complex, high-dimensional problems. Capturing the effect of the imperfect mesh is therefore critical.

Probabilistic Solvers

Instead, we try to construct a **probability model** for u which captures that uncertainty directly, as a probability measure over solutions to the PDE. This allows us to:

- **Propagate inaccuracy through subsequent computations.**
- **Understand the effect on inferences we make.**
- **Develop new methods utilizing the measure of inaccuracy** (for example: minimize posterior uncertainty with a limited computational budget).

Linear Forward Problem

We place a Gaussian Process prior over u with kernel k, and condition this on observations of the right-hand-side of the system. For linear problems the posterior is available in closed form.

$$\begin{aligned} u|\mathbf{g} &\sim N(\mu, \boldsymbol{\Sigma}) \\ \mu &= \bar{\mathcal{A}}K(\mathbf{x}, X_0)\left[\mathcal{A}\bar{\mathcal{A}}K(X_0)\right]^{-1}\mathbf{g} \\ \boldsymbol{\Sigma} &= K(X) - \bar{\mathcal{A}}K(X, X_0)\left[\mathcal{A}\bar{\mathcal{A}}K(X_0)\right]^{-1}\mathcal{A}K(X_0, X) \end{aligned}$$

Thanks to linearity the posterior is a **Gaussian Process** for the solution to the PDE in the Reproducing Kernel Hilbert Space associated with k, which includes uncertainty due to the number of evaluations of the RHS.

An illustration of this process, taken from a 1-D slice through through the solution to Example 1, is given in Fig. 1. With a 10x10 mesh we see that the mean of the posterior distribution is within 1SD of the true solution.

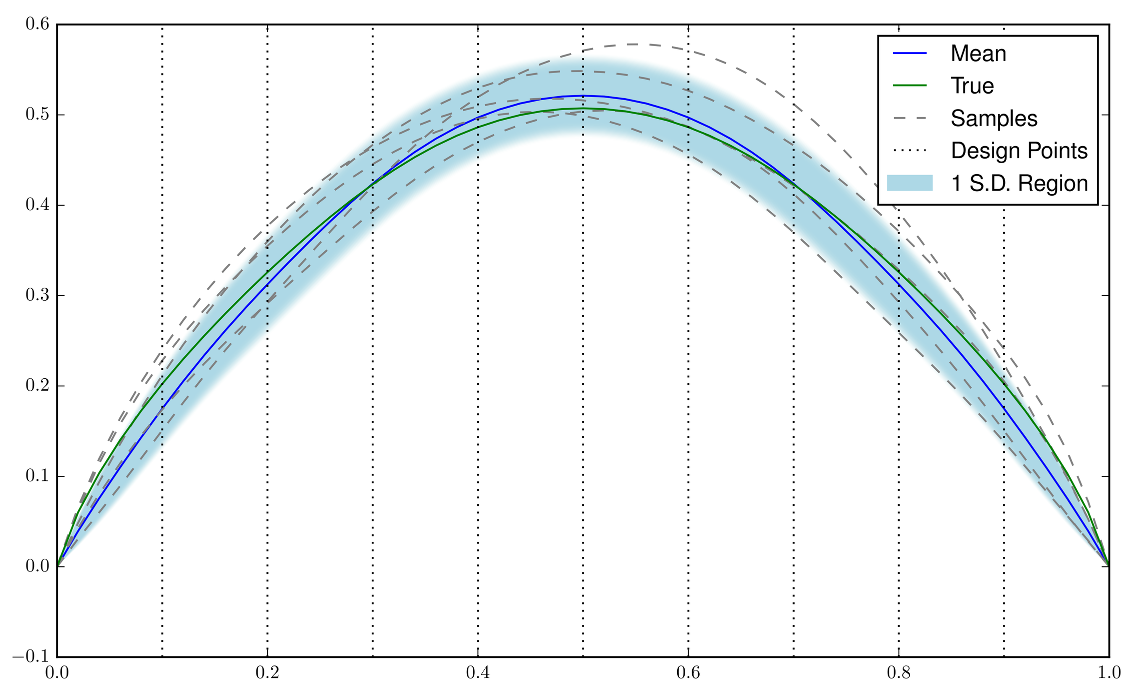


Fig 1: posterior measure for the forward solution of Example 1, along the line $x_1 = x_2$

Linear Inverse Problem

We are interested in performing inference on parameters θ of a PDE model, based on (possibly noisy) observations u_i of the solution at locations x_i.

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

The likelihood $p(y|\theta)$ can be found as:

$$p(y|\theta) := \prod_i p(y_i|u(x_i;\theta))$$

which is dependent on an exact value for u at each x_i. Because we have a measure of uncertainty in the forward problem, instead we use an **approximate likelihood incorporating the solver error**, similar to (1):

$$p_{\mathrm{PN}}(\theta|y) = \int \prod_i p(y_i|u)p(u|\theta)\mathrm{d}u$$

For **linear** operators this can be evaluated explicitly, and we can draw samples from the posterior $p(\theta|y)$ with MCMC.

Example 1

$$\begin{aligned} \nabla \cdot \kappa(\mathbf{x};\theta)\nabla u(\mathbf{x}) &= 0 & &\text{in } [0,1]^2 \\ u(\mathbf{x}) &= x_1 & &\text{at } x_2 = 0 \\ u(\mathbf{x}) &= 1 - x_1 & &\text{at } x_2 = 1 \\ \frac{\partial u}{\partial x_1} &= 0 & &\text{at } x_1 = 0, 1 \end{aligned}$$

We approximate κ with a KL expansion, truncated at six terms:

$$\kappa(\mathbf{x};\theta) = \exp\left(\sum_{k=1}^{6} \frac{\theta_k}{k^2}\cos(2\pi(s_k x_1 + r_k x_2))\right)$$

True data is generated using $\theta_k = 1$ and corrupted with Gaussian noise. We then perform MCMC to infer the values of θ_k. Results are shown in Fig. 2, below.

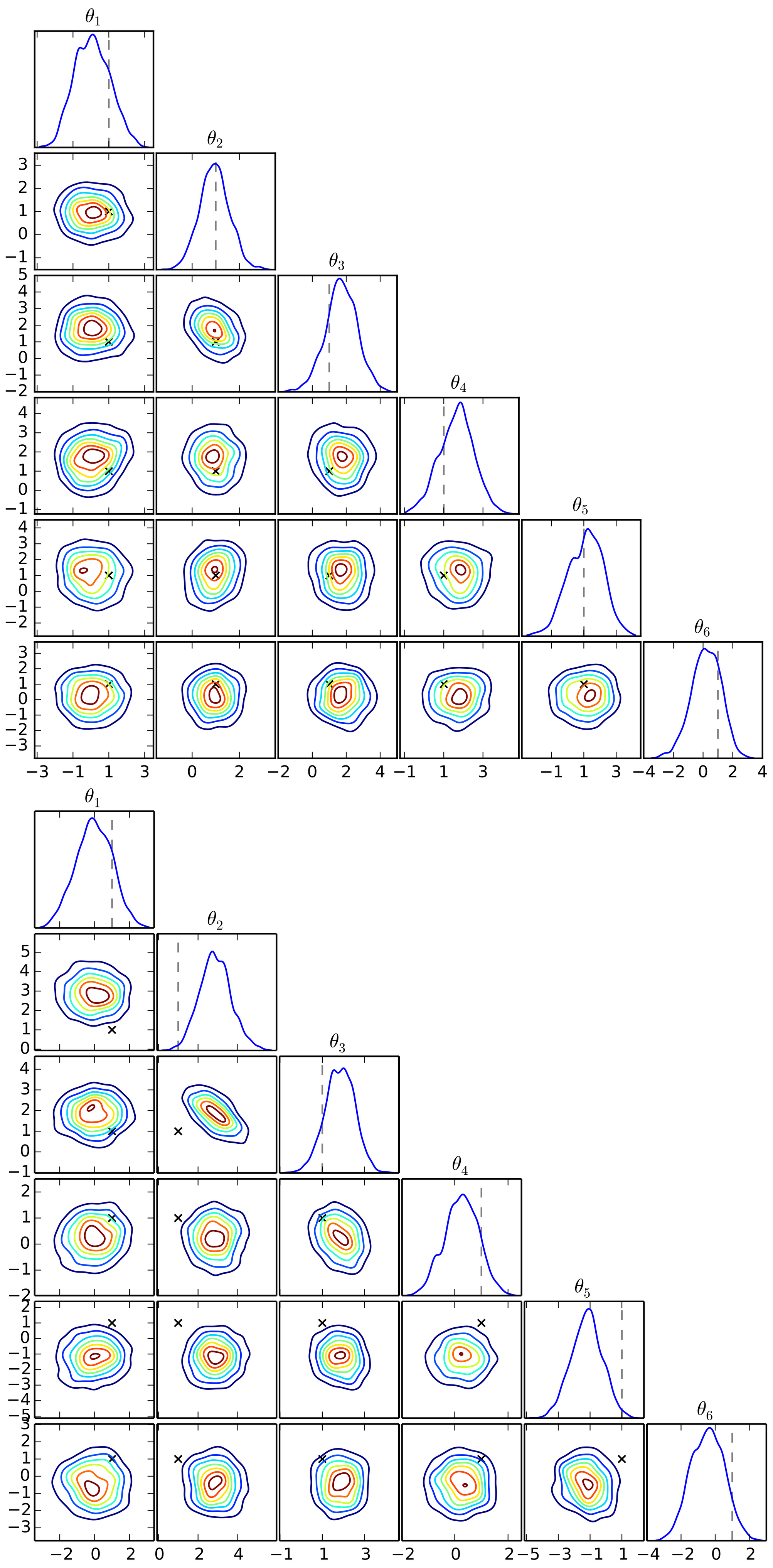


Fig 2: Comparison of posterior measures in inverse problem, n each case on an 10x10 mesh. Dashed lines / black crosses denote the true values.

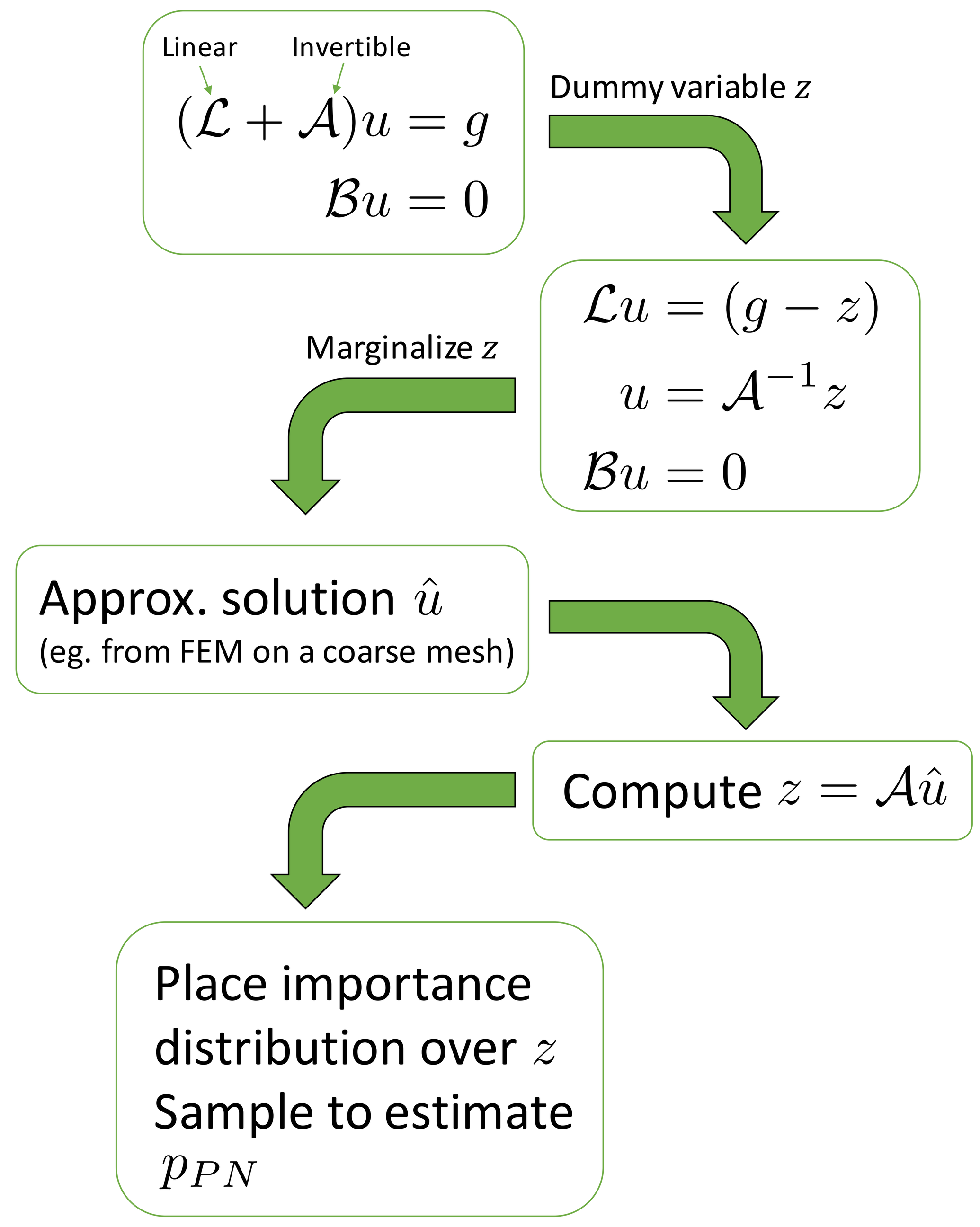
In the PN case (top) we see that **the posterior is wider, and encompasses the true value**, reflecting the relative courseness of the solver.

Using FEM (bottom) gives posteriors which are **concentrated on incorrect values** and in some cases not covering the true value.

Nonlinear Problems

When the system is **nonlinear**, the posterior over u is no longer a GP, and so we must resort to a sampling methodology.

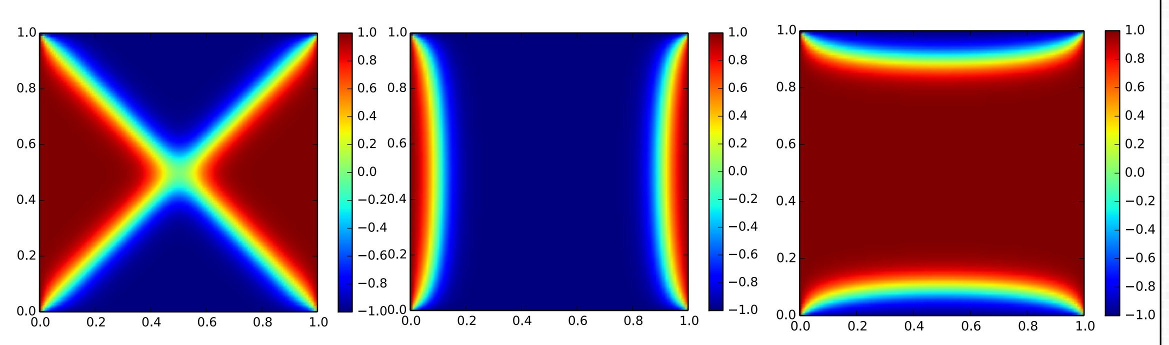
We focus on a class of PDEs which can be decomposed into a sum of linear and nonlinear operators. We further assume that the nonlinear operators are monotonic and invertible.



Example 2 (Steady State Allen-Cahn)

$$\begin{aligned} -\delta\nabla^2 u + \delta^{-1}(u^3 - u) &= 0 & (x, y) &\in (0,1)^2 \\ u &= 1 & x = 0, x &= 1, y \in (0,1) \\ u &= -1 & y = 0, y &= 1 \end{aligned}$$

The forward solution is **non-unique**; 3 solutions exist, as shown below (found using techniques in (2)):



This is easily handled in the inverse problem by placing a **mixture distribution** over possible values of z at each δ visited. Classical inverse problem samplers would require either a-priori knowledge of the 'correct' solution, or an ad-hoc method for determining it for each δ.

Posterior for δ is shown in Fig. 3

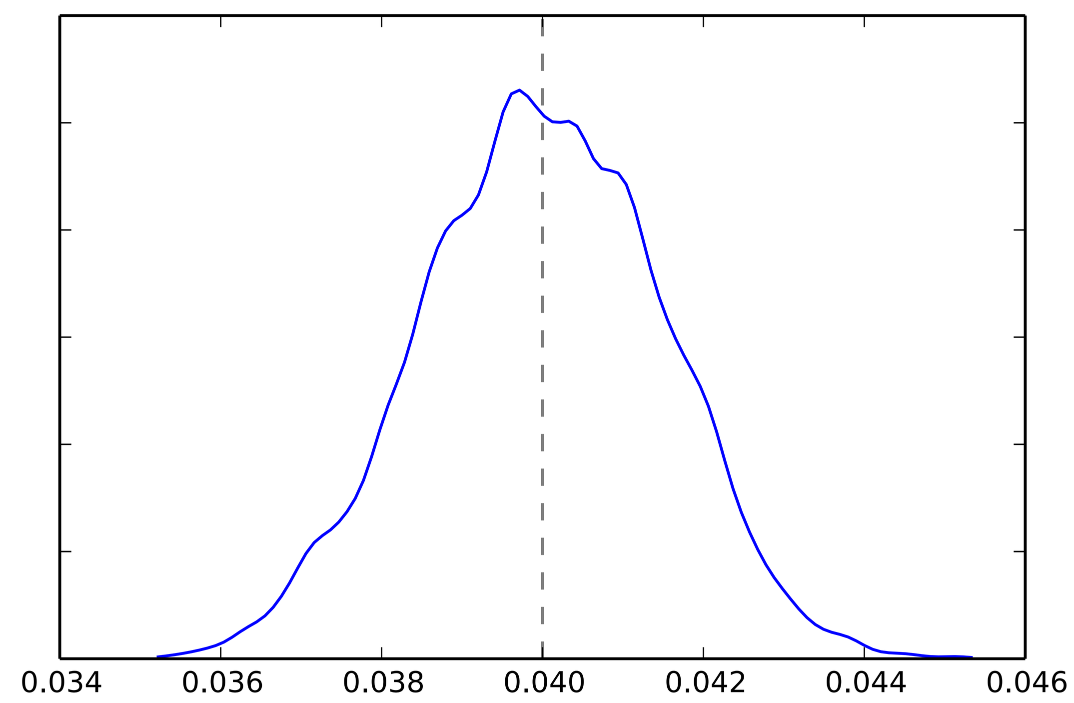


Fig 3: Inference of δ in Example 2. The posterior is concentrated around the true value, without needing to supply the 'true' solution.

References

(1) Conrad, P.R., Girolami, M., Särkkä, S., Stuart, A., and Zygalakis, K. (2015). Probability measures for numerical solutions of differential equations. arXiv preprint arXiv:1506.04592.

(2) Funke, S. W., & Farrell, P. E. (2013). Deflation techniques for finding distinct solutions of nonlinear partial differential equations´, 1–22.