

# Error Control of the Numerical Posterior with Bayes Factors in Bayesian Uncertainty Quantification

Marcos A. Capistrán\*, J. Andrés Christen<sup>\*,||</sup>, María L. Daza-Torres\*, Hugo Flores-Arguedas\*, and J. Cricelio Montesinos-López\*

**Abstract.** In this paper, we address the numerical posterior error control problem for the Bayesian approach to inverse problems or recently known as Bayesian Uncertainty Quantification (UQ). We generalize the results of Capistrán et al. (2016) to (a priori) expected Bayes factors (BF) and in a more general, infinite-dimensional setting. In this inverse problem, the unavoidable numerical approximation of the Forward Map (FM, i.e., the regressor function), arising from the numerical solution of a system of differential equations, demands error estimates of the corresponding approximate numerical posterior distribution. Our approach is to make such comparisons in the setting of Bayesian model selection and BFs. The main result of this paper is a bound on the absolute global error tolerated by the numerical solver of the FM in order to keep the BF of the numerical versus the theoretical posterior near one. For two examples, we provide a detailed analysis of the computation and implementation of the introduced bound. Furthermore, we show that the resulting numerical posterior turns out to be nearly identical from the theoretical posterior, given the control of the BF near one.

**MSC2020 subject classifications:** Primary 62F15, 62F35; secondary 62F07.

**Keywords:** inverse problems, uncertainty quantification, Bayesian inference, Bayesian model comparison, numerical analysis of ODE's and PDE's, total variation.

## 1 Introduction

Bayesian Uncertainty Quantification (UQ) is a recently coined term and has received considerable attention in the last ten years, by several researchers in many fields, as an alternative solution to some classical inverse problems (see Kaipio and Somersalo, 2005; Fox et al., 2013; Dashti and Stuart, 2016, for some reviews on the subject).

---

\*Centro de Investigación en Matemáticas (CIMAT), Jalisco S/N, Valenciana, Guanajuato, GTO, 36023, Mexico, [marcos@cimat.mx](mailto:marcos@cimat.mx)

\*Centro de Investigación en Matemáticas (CIMAT), Jalisco S/N, Valenciana, Guanajuato, GTO, 36023, Mexico, [jac@cimat.mx](mailto:jac@cimat.mx)

\*Centro de Investigación en Matemáticas (CIMAT), Jalisco S/N, Valenciana, Guanajuato, GTO, 36023, Mexico, [mdazatorres@cimat.mx](mailto:mdazatorres@cimat.mx)

\*Centro de Investigación en Matemáticas (CIMAT), Jalisco S/N, Valenciana, Guanajuato, GTO, 36023, Mexico, [hugo.flores@cimat.mx](mailto:hugo.flores@cimat.mx)

\*Centro de Investigación en Matemáticas (CIMAT), Jalisco S/N, Valenciana, Guanajuato, GTO, 36023, Mexico, [jose.montesinos@cimat.mx](mailto:jose.montesinos@cimat.mx)

<sup>||</sup>Corresponding author.

Many applications are found in image processing (Giovannelli and Idier, 2015; Zhu et al., 2011; Cai et al., 2011; Chama et al., 2012; Nissinen et al., 2011; Kozawa et al., 2012, for example), but there are great variety of applications in many fields as in geothermal prospection (Cui et al., 2011, 2019), network analysis (Hazelton, 2010; Sun et al., 2015), heat transfer (Kaipio and Fox, 2011), pollution and ecology (Keats et al., 2010; Hutchinson et al., 2017), fusion physics (Osthus et al., 2019), tumor growth (Collis et al., 2017; Kahle et al., 2019), to mention a few. As with most terms, “UQ” has a level of arbitrariness: indeed, all of statistics is partly concerned with quantifying uncertainty. However, in the inverse problems community UQ mostly refers to the theory and practice of probabilistic propagation of uncertainty and probabilistic (Bayesian) inference in the context of complex regressors involving systems of differential equations. The subject is indeed relevant to the statistical, and especially, to the Bayesian community. We address in this work what we believe is an important issue of Bayesian UQ.

In more concrete words, Bayesian UQ is Bayesian inference on a possibly infinite-dimensional parameter  $\theta$  with data  $y$ , e.g.

$$y = \mathcal{H}(\mathcal{F}(\theta)) + \epsilon; \quad \epsilon \sim N_m(0, \sigma^2 I);$$

where the regressor  $\mathcal{F}(\theta)$  or Forward Map (FM) is commonly a complex non-linear map, with input parameters  $\theta$ , defined by an initial/boundary value problem for a system of ODEs or PDEs. Consequently, to evaluate  $\mathcal{F}(\theta)$  at a single  $\theta$  value we must solve the corresponding system of ODEs or PDEs ( $\mathcal{H}$  is an observation functional or link function, formal details will be given in Section 2). However, the analytical solution of such problem is seldom available and we must work with a numerical approximation  $\mathcal{F}^{\alpha(n)}(\theta)$ , in the classic numerical analysis sense (Quarteroni and Valli, 2008; Iserles, 2009). Here  $\alpha(n)$  represents a discretization used to approximate the FM and as  $n$  increases the discretization becomes finer and the approximation error becomes smaller. This commonly involves the use of serious numerical methods and large computing power to evaluate  $\mathcal{F}^{\alpha(n)}(\theta)$ .

Bayesian UQ refers to the solution of the above problem in terms of inference on the unknown parameter  $\theta$ , by stating a prior  $\pi(\theta)$  and seeking its posterior distribution (via MCMC or any other standard or tailor-made method).

In this paper, we are concerned with the numerical error induced in the *numerical posterior* with respect to the theoretical posterior, when considering the approximate FM  $\mathcal{F}^{\alpha(n)}(\theta)$  instead of the exact, theoretical, FM  $\mathcal{F}(\theta)$ . We are mostly interested in establishing guidelines for choosing a discretization level  $\alpha(n)$  for the FM: a critical first problem to be considered when working with numerical FMs, or any other computer model.

Controlling the numerical error induced in the posterior distribution is an important issue in Bayesian UQ (Cotter et al., 2010; Yan and Guo, 2015; Lie et al., 2018; Dashti and Stuart, 2011). Capistrán et al. (2016) (CA in the following) present an approach to address the problem in the finite-dimensional case using Bayes Factors (BF) of the numerical model vs. the theoretical model. In an ODE framework, these odds are guaranteed to converge to one in the same order as the numerical solver used. For

high order solvers Capistrán et al. (2016) illustrate, by reducing the step size in the numerical solver, that there should exist a point at which the BF is nearly one, but for a fixed discretization threshold,  $\alpha(n)$  (step size) greater than zero. This threshold, in turn, depends on the numerical FM error compared to the data noise  $\sigma$ , the sample size, and other factors. This is the main point made by CA: it could be possible to calculate a threshold in the numerical error of the FM such that the numerical posterior is basically equal to the theoretical posterior so, although we are using an approximate FM, the resulting posterior is nearly numerical error-free. CA show, with some examples, that such optimal solver discretization leads to negligible differences in the numerical and the theoretical posterior since the BF is close to one, potentially saving CPU time by choosing a coarser solver.

However, the approach from CA still has some shortcomings. Firstly, it depends crucially on estimating the normalizing constants from Monte Carlo samples of the unnormalized posterior, for a range of discretizations  $\alpha(n)$ . This is a complex problem, and the subject of current research, since it is difficult to reliably estimate these normalizing constants in high dimensional problems (Moral et al., 2017; de Valpine, 2008). Secondly, CA's approach is as yet incomplete since one would need to decrease  $\alpha(n)$  systematically, calculating the normalization constant of the corresponding numerical posterior to eventually estimate the normalization constant of the theoretical posterior (see figure 2 in CA), which in turn will pinpoint a discretization at which both models are nearly identical. The main difficulty here is that only after calculating the posterior for some small step sizes, we estimate that a larger step size was sufficient, with an estimated BF close to one. That is, one has already made the CPU effort of calculating the posterior distribution for a grid of small step sizes, and therefore it renders useless the selection of the optimal step size.

To improve on CA, the idea of this paper is to consider the *expected* value of the BFs, before data is observed. We will try to bound this expected BF to find general guidelines to establish error bounds on the numerical solver, depending on the specific problem at hand and the sample design used, but not on particular data. These guidelines will be solely regarding the FM and, although perhaps conservative, they may represent useful bounds to be used in practice. Moreover, we generalize CA to an infinite-dimensional setting.

To fix ideas, we present an example. Imagine we have a metal rod subject along its length  $x$  to a heat profile described with the function  $f(x) = \sin(\pi x)$  (its forcing), and where its edges are kept at a reference temperature value (e.g. zero). The thermal conductivity of the rod  $a(x)$  along its length is unknown. The forcing is applied, and the system is left to reach equilibrium, at which point temperature measurements are taken at fixed points along the rod. This system is commonly described as the stationary heat equation in 1D

$$-\frac{d}{dx} \left( a(x) \frac{du(x)}{dx} \right) = f(x), \quad x \in (0, 1), \quad (1)$$

subject to Dirichlet boundary conditions  $u(0) = u(1) = 0$ , with forcing term  $f(x) = \sin(\pi x)$  and thermal conductivity  $a(x) > 0$ . The Forward Map in this case is: given  $a(x)$ , solve the above differential equation to find  $u$ , namely  $\mathcal{F}[a] = u$ .

Given fixed measurements locations  $x_1, x_2, \dots, x_m$ , the observation functional is  $\mathcal{H}(u) = (u(x_1), u(x_2), \dots, u(x_m))'$ . Under additive Gaussian noise, the data  $y = (y_1, y_2, \dots, y_m)'$  satisfies

$$y = \mathcal{H}(\mathcal{F}[a]) + \epsilon; \quad \epsilon \sim N_m(0, \sigma^2 I).$$

Bayesian inference is required for the unknown parameter  $a$ , the thermal conductivity. A Finite Element numerical solution (see Brenner and Scott, 2008; Reddy, 2006, for example) is required to find a numerical approximation  $\mathcal{F}^{\alpha(n)}[a]$  to the theoretical FM  $\mathcal{F}[a]$ . If prior information tells us that the rod is homogeneous then we may set  $a(x) = \theta \in \mathbb{R}^+$  and the problem is parametric inference of dimension one. In other circumstances prior information might lead us to think that the thermal conductivity is increasing linearly with length, that is  $a(x) = \theta_0 + \theta_1 x$  etc. On the other hand, little may be known about  $a(x)$  and the whole function may be needed to be inferred, in which case the parameter is infinite dimensional belonging to a functional space. A particular representation may be chosen, for example  $a(x) = \theta = \sum_{j=1}^{\infty} \beta_j \phi_j(x)$  with a finite dimensional version  $\theta_k = \sum_{j=1}^k \beta_j \phi_j(x)$  and their corresponding priors  $\pi$  and  $\pi_k$ .

There are several results published recently proving that the numerical versions of the posterior tend to the theoretical posterior (see Section 2.2 for details). That is, under some regularity conditions, as  $n$  increases, the numerical posterior becomes a better approximation to the theoretical posterior, where the limit has been adequately defined using the Hellinger or Total Variation measure norms. That is, we already know how and in what sense the approximations are consistent. However, in applications, we will still need to choose a particular (finite)  $n$ .

*Accordingly, the problem we try to address in this paper is:* How can we control the numerical error in the posterior distribution? Specifically, how can we choose a suitable discretization  $n$  to control such numerical error properly?

The basic idea is to establish the relative merit of the numerical posterior vs. the theoretical posterior using Bayesian model selection as a function of  $n$ , to keep both models with a Bayes Factor of nearly one. We prove how this is possible, for finite  $n$ , resulting in a posterior distribution with negligible numerical error w.r.t the theoretical posterior. Indeed, this depends on how significant the numerical error in the FM is in comparison to the data error  $\sigma$ , the sample size, etc. We do not discuss here the use of an approximate prior  $\pi_k$ , only a brief comment is given on this respect in Section 3.1.

After explaining our formal setting and notation in Section 2, we prove our main result in Section 3 for Banach spaces and considering any location-scale family for the distribution of the data. In Section 4 we present two examples to illustrate our approach, and in Section 5, a discussion of the paper is given.

## 2 Setting and preliminary lemmas

For completeness and to establish the theoretical setting needed for our main theorems in Section 3, we discuss here the existence and consistency of the posterior distribution in general Banach spaces. Note that in other contexts beyond the standard UQ literature,

the foundation of Bayesian theory in abstract spaces is very well known (here we outline a particular version of this general Bayesian theory see, for example, Schervish, 1997; Ghosal and van der Vaart, 2017). Here we adhere to the standard Bayesian approach in which, among many other details, the aim is to define a joint measure of both observables and unknowns, from which a posterior measure is obtained. As opposed, commonly in the UQ literature, the posterior measure is postulated as a Radon-Nikodym derivative w.r.t the prior and equal to the likelihood (e.g. Stuart, 2010; Dashti and Stuart, 2016; Bui-Thanh and Ghattas, 2014), and then proved to exist and be well defined, as opposed to standard Bayesian theory in which such Radon-Nikodym derivative w.r.t the prior is a consequence of the existence and regularity of the posterior measure (see, for example, section 1.3 and eq. (1.1) of Ghosal and van der Vaart, 2017).

In the case when  $\theta$  is a function, we need to establish a proper setting for this Bayesian inference problem in which the parameter  $\theta$  belongs to a general Banach space. The same setting also applies to the finite-dimensional case since  $\mathbb{R}^d$  is also a Banach space. A general enough and convenient setting is as follows.

Let  $Y \in \mathcal{Y} \subset \mathbb{R}^m$  be the data at hand and  $\{P_\theta : \theta \in \Theta\}$  be a family of probability models for  $Y$ . We assume that for each  $\theta \in \Theta$  the observable  $Y$  has a density  $f_\theta(y)$  w.r.t a  $\sigma$ -finite measure  $\lambda$ , namely a product of the Lebesgue and/or counting measures in  $\mathbb{R}^m$  to accommodate, possibly, continuous and/or discrete observations. That is

$$P_\theta(Y \in A) = \int_A P_\theta(dy) = \int_A f_\theta(y) \lambda(dy),$$

for all measurable  $A$ . For example,  $\mathcal{Y}$  is a product space of subsets of  $\mathbb{R}$  or  $\mathbb{Z}$ , leading to continuous and/or discrete data. This is the usual setting in parametric inference.

In any case, with the usual topological considerations, we assume  $\mathcal{Y}$  is a *Polish* space. Polish spaces include complete metric spaces that have a countable dense subset.  $\mathcal{Y}$  should be viewed as a Polish space with the standard metric in  $\mathbb{R}$  and the discrete metric in  $\mathbb{Z}$ , and  $\lambda$  then results in a Borel  $\sigma$ -finite measure on  $\mathcal{Y}$ .  $P_\theta$  is then a Radon measure for all  $\theta \in \Theta$ , since any Borel probability measure on a Polish space is Radon.

Until now, the parameter space  $\Theta$  is arbitrary. In order to define a probability measure  $\pi$  on  $\Theta$  (namely, a prior distribution) we need to define a measurable space  $(\Theta, @)$ . So far  $f_\theta(y)$  cannot be considered a conditional distribution.  $f_\theta(y)$  is already  $\lambda$  measurable and we assume also that is not only measurable in  $\theta$ , but for fixed  $y$ , continuous in  $\theta$ . More precisely, assume that  $\Theta$  is a separable Banach space,  $@$  its Borel  $\sigma$ -algebra, and  $\theta \mapsto f(y|\theta)$  is continuous for all  $y \in \mathcal{Y}$ . Therefore, since  $\mathcal{Y}$  is a *Polish* space, the posterior measure is well defined (see, for example, Christen and Pérez-Garmendia, 2020, and references therein). Namely, for any  $\pi$ -measurable  $g$  we have that the posterior measure  $\mathbb{Q}_y[\pi]$  is defined by

$$\int g(\theta) \mathbb{Q}_y[\pi](d\theta) := \frac{\int g(\theta) f(y|\theta) \pi(d\theta)}{\int f(y|\theta) \pi(d\theta)},$$

where now we use the more common notation  $f(y|\theta) = f_\theta(y)$ . Or equivalently, the Radon-Nikodim derivative of the posterior measure w.r.t the prior is proportional to

the likelihood, i.e.  $\frac{\partial \mathbb{Q}_y}{\partial \pi} \propto f(y|\theta)$ , for all  $y \in \mathcal{Y}$ ; see for example Schervish (1997), p. 16, Ghosal and van der Vaart (2017), p. 6 or Christen and Pérez-Garmendia (2020). We also let

$$Z(y) = \int f(y|\theta)\pi(d\theta),$$

the normalization constant for fixed  $y$  or the prior predictive density w.r.t  $\lambda$ , as a function of  $y$ .

To make the dependency on the prior explicit, we adopted above the notation  $\mathbb{Q}_y[\pi]$  for the posterior measure. When substituting the likelihood with a numerical approximation  $f^n(y|\theta)$ , we denote the corresponding posterior as  $\mathbb{Q}_y^n[\pi]$  with  $Z_n(y)$  its normalization constant.

## 2.1 The inverse problems setting

We follow the general Bayesian UQ setting, as explained in, for example, Scheichl et al. (2017); Dashti and Stuart (2016). Let  $\Theta$  and  $V$  be separable Banach spaces, let  $\mathcal{F} : \Theta \rightarrow V$  be the Borel measurable forward map (FM) and  $\mathcal{H} : V \rightarrow A \subseteq \mathbb{R}^{m+s}$  the Borel measurable observation operator. The composition  $\mathcal{H} \circ \mathcal{F}$  defines a Borel measurable mapping from the parameter space  $\Theta$  to the data sample space in  $\mathbb{R}^m$ , plus possibly additional parameters. Going beyond Gaussian noise assume that  $f_o(y|\eta)$  is a density for data  $y$  w.r.t.  $\lambda$  for all  $\eta \in A$ . The parametric family of sample models, as in Section 2, is defined with the family of  $\lambda$ -densities

$$f(y|\theta) = f_o(y|\mathcal{H}(\mathcal{F}(\theta))); \theta \in \Theta.$$

To fix ideas, we elaborate on the usual independent Gaussian noise case, namely

$$f_o(y|\eta) = \prod_{j=1}^m \sigma^{-1} \rho\left(\frac{y_j - \eta_j}{\sigma}\right),$$

where  $\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ , i.e.  $y_j = \mathcal{H}_j(\mathcal{F}(\theta)) + \sigma \epsilon_j$ ;  $\epsilon_j \sim N(0, 1)$  independently. If  $\sigma$  is unknown, we may take  $s = 1$  and include it as a parameter. The same holds if we had an unknown variance-covariance matrix. We do not discuss this case further in the main part of the paper. Some notes are added in Section 5 regarding the case when  $\sigma$  is unknown.

Let  $\mathcal{F}^{\alpha(n)}$  be a discretized version of the forward map  $\mathcal{F}$ , for some discretization  $\alpha$  that depends on an integer refinement  $n$ . This is the actual numerical version of the forward map defined in our computers. Let  $f^n(y|\theta) = f_o(y|\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta)))$  be the resulting discretized numerical likelihood. In the rest of the paper we take the following assumption.

**Assumption 2.1.** *Assume that, for all  $y \in \mathcal{Y}$  the observation model  $f_o(y|\eta)$  is uniformly Lipschitz continuous for each  $\eta$ , and for  $y \in \mathcal{Y}$   $\lambda$ -a.s.  $f_o(y|\eta)$  is bounded. Moreover, the FM maps  $\mathcal{H} \circ \mathcal{F}$  and  $\mathcal{H} \circ \mathcal{F}^{\alpha(n)}$  are continuous.*

If  $\mathcal{H} \circ \mathcal{F}$  and  $\mathcal{H} \circ \mathcal{F}^{\alpha(n)}$  are continuous then  $\theta \mapsto f(y|\theta)$  and  $\theta \mapsto f^n(y|\theta)$  are continuous. Therefore all requirements are met for the posterior measures  $\mathbb{Q}_y[\pi]$  and  $\mathbb{Q}_y^n[\pi]$  to be well defined (as explained in Section 2), either considering the theoretical or the numerical likelihood, respectively (with  $Z(y)$  and  $Z^n(y)$  their corresponding normalization constants). In Dashti and Stuart (2016); Scheichl et al. (2017) it is also assumed that  $\mathcal{H}(\mathcal{F}(\theta))$  is continuous.

Note that if we consider independent data with a location-scale model as

$$f_o(y|\eta) = \prod_{j=1}^m \sigma^{-1} \rho\left(\frac{y_j - \eta_j}{\sigma}\right), \quad (2)$$

where  $\rho(x)$  is uniformly Lipschitz continuous and  $\sigma$  known, the first part of assumption 2.1 is met and we only require to establish that  $\mathcal{H} \circ \mathcal{F}$  and  $\mathcal{H} \circ \mathcal{F}^{\alpha(n)}$  are continuous. Indeed the former is true if  $\rho(x)$  is Gaussian.

Assume a global error control of this numerical FM as

$$\|\mathcal{H}(\mathcal{F}(\theta)) - \mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))\|_\infty < K_0 \langle \alpha(n) \rangle^p, \quad (3)$$

for all  $\theta \in \Theta$  and some functional  $\langle \cdot \rangle$ . Note that (3) is a global bound, valid for all  $\theta \in \Theta$  and includes already the observational operator  $\mathcal{H}$ . That is, it is a global bound but it is only a statement at the locations  $\mathcal{H}_j$ s where each  $y_j$  is observed,  $j = 1, \dots, m$ . That is  $\|\mathcal{H}(\mathcal{F}(\theta)) - \mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))\|_\infty = \max_j |\mathcal{H}_j(\mathcal{F}(\theta)) - \mathcal{H}_j(\mathcal{F}^{\alpha(n)}(\theta))| < K_0 \langle \alpha(n) \rangle^p$  for all  $\theta \in \Theta$ .

Regarding the functional  $\langle \cdot \rangle$ , if we have a discretization in space and time with  $\Delta x = h_x/n$ ,  $\Delta t = h_t/n$ , then  $\alpha(n) = (\Delta x, \Delta t)$  and it could be the case that  $|\alpha(n)| = \Delta x + \Delta t$  depending on the numerical method used to solve  $\mathcal{F}(\theta)$ . There are a great variety of methods, but it is assumed that  $n$  represents the discretization level and  $|\alpha(n)| \rightarrow 0$  as  $n \rightarrow \infty$ . We assume that  $\langle \alpha(n) \rangle$  is properly defined so that  $p$  signifies the method's order.

From assumption 2.1  $f_o(y|\eta)$  is uniformly Lipschitz continuous for any given  $y$ . Then since  $|f_o(y|\eta) - f_o(y|\eta')| < L|\eta - \eta'|$  we have

$$|f^n(y|\theta) - f(y|\theta)| = |f_o(y|\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))) - f_o(y|\mathcal{H}(\mathcal{F}(\theta)))| < K_1 \langle \alpha(n) \rangle^p, \quad (4)$$

for all  $\theta \in \Theta$ , which is also a global error bound, now for the numerical likelihood, since the constant  $K_1 = LK_0$  is independent of  $\theta$ . This is the general setting needed to formally state the Bayesian UQ problem.

## 2.2 Consistency results

Once the existence of the theoretical and numerical versions of the posterior measure has been assured, the natural next question is that of approximations' consistency. That is, if and how and in what sense the numerical posterior measure  $\mathbb{Q}_y^n[\pi]$  tends to the theoretical posterior measure  $\mathbb{Q}_y[\pi]$ , as the discretization  $\alpha(n)$  becomes finer (i.e.  $n$  increases).



For example, using (4), Hosseini and Nigam (2017) prove that  $\|\mathbb{Q}_y^n[\pi] - \mathbb{Q}_y[\pi]\|_H \rightarrow 0$ , adding regularity conditions on  $-\log f(y|\theta)$ ; (see Hosseini and Nigam, 2017, and references therein), where  $\|\cdot\|_H$  is the Hellinger norm for measures. Christen and Pérez-Garmendia (2020) prove that  $\|\mathbb{Q}_y^n[\pi] - \mathbb{Q}_y[\pi]\|_{TV} \rightarrow 0$  using (4); here  $\|\cdot\|_{TV}$  is the Total Variation norm (although note that the Hellinger and Total Variation norms are equivalent, i.e. the latter implies the former and vice versa Gibbs and Su, 2002) and also that  $Z^n(y) \rightarrow Z(y)$ . That is, for a given problem, we may safely assume, after verifying the regularity conditions needed, that

$$\|\mathbb{Q}_y^n[\pi] - \mathbb{Q}_y[\pi]\| \rightarrow 0 \quad \text{and} \quad |Z^n(y) - Z(y)| \rightarrow 0 \quad (5)$$

as  $n$  increases (the discretization becomes finer), in a suitable measure norm  $\|\cdot\|$ . Moreover, Christen and Pérez-Garmendia (2020); Bui-Thanh et al. (2013) discuss the rates of convergence in some contexts, and these are proven to be the same as those for the FM numerical solver. However, besides perhaps elementary examples, authors do not provide means to compute  $\|\mathbb{Q}_y^n[\pi] - \mathbb{Q}_y[\pi]\|$  since convergence theorems depend on intractable constants.

As mentioned before, we still need to fix a discretization level to compute the posterior distribution. The previously mentioned results provide a sound theoretical basis for Bayesian UQ, but as such, they do not address the latter problem, which is the primary concern of this paper. Accordingly, next, we turn to our main result on choosing the discretization level  $n$  using Bayesian model comparison.

### 3 Expected a priori error bounds and Bayes Factors

As in Capistrán et al. (2016) in order to find reasonable guidelines for choosing a discretization level  $n$ , we compare the numerical posterior  $\mathbb{Q}_y^n[\pi]$  with the theoretical posterior  $\mathbb{Q}_y[\pi]$  using Bayesian model selection, namely Bayes Factors (BF). Assuming an equal prior probability for both models, the BF is the posterior probability odds of one model against the other, that is  $\frac{p}{1-p}$  where  $p = \frac{Z^n(y)}{Z^n(y) + Z(y)}$ , the posterior probability of the numerical model. Indeed, the BF is the ratio of the normalization constants  $\frac{Z^n(y)}{Z(y)}$  and in our case, it is the ratio of the posterior probabilities  $p$  and  $1 - p$ .

The use of Bayesian model selection and BF has not been absent of debate, leading to some stylized alternatives discussed in many specific contexts (Hoeting et al., 1999; Berger and Pericchi, 1996; Kass and Raftery, 1995; O'Hagan, 1997; Xu et al., 2019; van der Linden and Chryst, 2017). Some of these concerns relate to prior specifications and prior sensitivity. In the extreme case, improper priors cannot be used since they are established up to an arbitrary constant and  $Z(y)$  is indeterminate; additional provisions need to be taken in this context (e.g. Berger and Pericchi, 1996). Moreover, when the models are diverse (e.g. having different number of variables), it is, in general, a complicated task to specify a prior distribution that 1) conveys, in either case, the correct prior information for each model and 2) does not bias one model over other. Even trying to establish “diffuse” priors can be problematic since the meaning of “diffuse” might change from model to model. Moreover, two different models might have a BF equal to



one if they have the same predictive power, thus equal posterior probabilities for two models do not mean that the two models are the same, etc. (see Clyde et al., 2007, for a discussion and references therein).

Fortunately, the above mentioned common concerns in the use of Bayesian model comparison are not present when comparing  $\mathbb{Q}_y^n[\pi]$  with  $\mathbb{Q}_y[\pi]$ . Note first that from the convergence and error control in the FM in (3) we may conclude (with further assumed regularity conditions) that the models converge as explained in Section 2.2, namely  $\|\mathbb{Q}_y^n[\pi] - \mathbb{Q}_y[\pi]\| \rightarrow 0$ . The numerical posterior  $\mathbb{Q}_y^n[\pi]$  is not of some different nature but only a numerical approximation to  $\mathbb{Q}_y[\pi]$  and since we are assuming  $Z(y)$  is the correct sampling distribution, lack in the predictive power of  $\mathbb{Q}_y^n[\pi]$  can only be attributed to its approximate nature (we do not discuss here model fit/model adequacy, a comment is given in this respect in Section 5). Note also that the prior measure  $\pi$  is assumed proper and  $Z^n(y)/Z(y) \rightarrow 1$  as in (5). Therefore the criterion for choosing  $n$  is then sought for the numerical approximation  $\mathbb{Q}_y^n[\pi]$  to have nearly the same expected predictive power as  $\mathbb{Q}_y[\pi]$ , being all other pieces equal and knowing that  $\|\mathbb{Q}_y^n[\pi] - \mathbb{Q}_y[\pi]\| \rightarrow 0$ .

### 3.1 The expected Absolute BF

In terms of model equivalence an alternative expression conveying the same BF odds  $Z^n(y)/Z(y)$  is

$$\frac{1}{2} \left| \frac{Z^n(y)}{Z(y)} - 1 \right|, \quad (6)$$

comparing  $\mathbb{Q}_y[\pi]$  with  $\mathbb{Q}_y^n[\pi]$ . We now try to control the Bayes Factor between the discretized model and the theoretical model,  $\frac{Z^n(y)}{Z(y)}$ , through the use of what we call the Absolute BF (ABF) in (6). In order to do that, independently of the specific data at hand, we try to bound the expected ABF (the EABF),

$$\int \frac{1}{2} |Z^n(y) - Z(y)| \lambda(dy) = \int \frac{1}{2} \left| \frac{Z^n(y)}{Z(y)} - 1 \right| Z(y) \lambda(dy),$$

in terms of estimates on the error in the numerical forward map, as in (3). The idea is to keep the EABF below a small threshold (e.g.  $\frac{1}{20}$ ) so that the BF is close to one and evidence in favor of the theoretical model over the numerical model is “not worth more than a bare mention” (Kass and Raftery, 1995; Jeffreys, 1961).

**Theorem 3.1.** *With assumption 2.1, the rate of convergence in (3) and  $\phi_y(\eta) = -\log f_o(y|\eta) \in C^1$   $\lambda$ -a.s. we have*

$$\begin{aligned} \int \frac{1}{2} \left| \frac{Z^n(y)}{Z(y)} - 1 \right| Z(y) \lambda(dy) &= \int \frac{1}{2} |Z^n(y) - Z(y)| \lambda(dy) < \\ \frac{K_0 |\alpha(n)|^p}{2} \sum_{j=1}^m \int \int \left| \frac{\partial}{\partial \eta_j} \phi_y(\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))) \right| &f_o(y|\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))) \lambda(dy) \pi(d\theta). \end{aligned} \quad (7)$$

*Proof.* To bound the integral  $\int \frac{1}{2} |Z^n(y) - Z(y)| \lambda(dy)$ , note that

$$|Z^n(y) - Z(y)| = \left| \int f(y|\theta) (R_n(\theta) - 1) \pi(d\theta) \right| < \int f(y|\theta) |R_n(\theta) - 1| \pi(d\theta)$$

with  $R_n(\theta) = \frac{f^n(y|\theta)}{f(y|\theta)}$ . For  $\eta$  close enough to  $\eta_1$ , the likelihood ratio  $\frac{f_o(y|\eta)}{f_o(y|\eta_1)}$  is near 1 and

$$\left| \frac{f_o(y|\eta)}{f_o(y|\eta_1)} - 1 \right| \cong \left| \log \left( \frac{f_o(y|\eta)}{f_o(y|\eta_1)} \right) \right| = |\phi_y(\eta) - \phi_y(\eta_1)| = |\phi_y(\eta_1) - \phi_y(\eta)|.$$

With the first order Taylor approximation  $\phi_y(\eta) = \phi_y(\eta_1) + \nabla \phi_y(\eta_1) \cdot (\eta - \eta_1) + E$  we have

$$|R_n(\theta) - 1| = |\phi_y(\eta_1) - \phi_y(\eta)| = \left| \nabla \phi_y(\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))) \cdot (\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta)) - \mathcal{H}(\mathcal{F}(\theta))) + E \right|.$$

Ignoring the higher order terms in the residual  $E$  we have

$$\begin{aligned} & \int f(y|\theta) |R_n(\theta) - 1| \pi(d\theta) < \\ & \int f(y|\theta) \left| \sum_{j=1}^m \frac{\partial}{\partial \eta_j} \phi_y(\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))) (\mathcal{H}_j(\mathcal{F}^{\alpha(n)}(\theta)) - \mathcal{H}_j(\mathcal{F}(\theta))) \right| \pi(d\theta) \end{aligned}$$

and using the global error bound in (3) we have

$$\begin{aligned} & \int \frac{1}{2} |Z^n(y) - Z(y)| \lambda(dy) < \\ & \frac{K_0 |\alpha(n)|^p}{2} \int \int f_o(y|\eta) \sum_{j=1}^m \left| \frac{\partial}{\partial \eta_j} \phi_y(\mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))) \right| \pi(d\theta) \lambda(dy), \end{aligned}$$

since  $|\sum a_j(\theta) b_j(\theta)| < c \sum |a_j(\theta)|$  with  $|b_j(\theta)| < c$  for all  $\theta$ , and we obtain the result.  $\square$

Note in the last part of the above proof how the global error bound is needed for all  $j$  as in (3), nonetheless, we could restrict this bound for  $\theta \in \text{supp}\{\pi\} \subset \Theta$  since we are integrating w.r.t  $\pi(\cdot)$ .

We may now attempt to calculate the remaining double integral by changing the order of integration to calculate  $\int M(\mathcal{H}(\mathcal{F}(\theta))) \pi_k(d\theta)$  where

$$M(\eta) = \sum_{i=1}^m \int \left| \frac{\partial}{\partial \eta_i} \phi_y(\eta) \right| f_o(y|\eta) \lambda(dy). \quad (8)$$

This in general leads to no simplification, however if it happens that  $M(\eta)$  does not depend on  $\theta$  it can then be very useful. This is the case for any location-scale family and we present it next.

**Theorem 3.2.** *With the setting of Theorem 3.1, assuming independent data arising from a location-scale family, namely*

$$f_o(y|\eta) = \prod_{i=1}^m \sigma^{-1} \rho\left(\frac{y_i - \eta_i}{\sigma}\right)$$

with  $\rho$  a bounded  $C^1$  symmetric Lebesgue density in  $\mathbb{R}$  with  $\int_{-\infty}^{\infty} x^2 \rho(x) dx = 1$  then

$$\int \frac{1}{2} \left| \frac{Z^n(y)}{Z(y)} - 1 \right| Z(y) \lambda(dy) < \rho(0) \frac{K_0 |\alpha(n)|^p}{\sigma} m. \quad (9)$$

*Proof.* From (8) note that

$$\int \left| \frac{\partial}{\partial \eta_i} \phi_y(\eta) \right| f_o(y|\eta) \lambda(dy) = \int_{-\infty}^{\infty} \left| \sigma^{-1} V' \left( \frac{y_i - \eta_i}{\sigma} \right) \right| \sigma^{-1} \rho \left( \frac{y_i - \eta_i}{\sigma} \right) dy_i,$$

where  $\rho(x) = e^{V(x)}$ . The integral on the rhs is in fact equal to  $2\sigma^{-1} \int_0^{\infty} V'(x) \rho(x) dx = 2\sigma^{-1} \rho(0)$  (since  $\rho'(x) = V'(x) \rho(x)$ ), and we obtain the result since the latter does not depend on  $\eta$ .  $\square$

Since  $K_0 \langle \alpha(n) \rangle^p$  is the error in the FM (with the observation operator in (3)), measured in the same units as the  $y_j$ s, note from (9) that  $\frac{K_0 \langle \alpha(n) \rangle^p}{\sigma}$  is the relative error in the numerical FM with respect to the standard error in the observations  $\sigma$ . It makes much sense to measure  $K_0 \langle \alpha(n) \rangle^p$  with respect to  $\sigma$ , i.e. the proportion of numerical error in the FM w.r.t the error in the observations. Note that  $\frac{K_0 \langle \alpha(n) \rangle^p}{\sigma}$  is unit free. In the usual case of independent Gaussian errors, with known variance  $\sigma^2$ ,  $\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  and  $\rho(0) = \frac{1}{\sqrt{2\pi}}$ .

If we let the  $EABF \leq b$ , and for example  $b = \frac{1}{20} = 0.05$ , we expect nearly no difference in the numerical and the theoretical posterior. If we set the error in the FM  $K = K_0 \langle \alpha(n) \rangle^p$  then we require  $\rho(0) \frac{K}{\sigma} m < b$ , that is, we need the numerical error in the FM in (3) to satisfy

$$K < \frac{\sigma}{m} \frac{b}{\rho(0)}. \quad (10)$$

Our suggested procedure is to run the solver, including an after-the-fact error estimate (or *a posteriori* error estimate, we use *after-the-fact* given the conflict of terms with the Bayesian jargon, see Quarteroni and Valli, 2008; Iserles, 2009, for example), that is, after running our solver to evaluate  $\mathcal{F}^{\alpha(n)}(\theta)$ , an upper estimate  $\hat{K}$  of  $K$  is obtained, for that particular  $\theta$  value. If the error in the FM does not comply with the bound in (10), then run the solver again with a finer discretization  $\alpha(n)$ . In passing, we guarantee (3) for all  $\theta \in \Theta$ . Note that for ODEs, the RK45 method (Rungue-Kutta order 5 method of Cash and Karp, 1990, for example) produces after the fact error estimates. More recently, the discontinuous Galerkin method for PDEs may include high order solvers with after the fact error estimates (Di Pietro and Ern, 2011; Hesthaven

and Warburton, 2007). In general, error estimates for PDEs are much harder to obtain, and the usual strategy is to consider adjoint-based methods.

The numerical posterior distribution, including the approximate likelihood  $f^n(y|\theta)$  and the EABF bound is therefore

$$\pi(\theta|y) \propto \sigma^{-m} \prod_{j=1}^m \rho\left(\frac{y_j - \mathcal{H}_j(\mathcal{F}^{\alpha(n)}(\theta))}{\sigma}\right) \pi(\theta),$$

subject to  $\|\mathcal{H}(\mathcal{F}(\theta)) - \mathcal{H}(\mathcal{F}^{\alpha(n)}(\theta))\| < K = \frac{\sigma}{m} \frac{b}{\rho(0)}$  (e.g.  $b = \frac{1}{20}$ ).

A common problem in Bayesian UQ is that we also need to approximate the prior  $\pi$  with an approximate sequence of priors  $\pi_k$ . For example,  $\pi$  is a theoretical infinite dimensional prior and  $\pi_k$  are finite dimensional approximations. To choose a specific prior  $\pi_k$  Theorem 3.1 may be easily extended, resulting in the same bound for the EABF adding the term  $\|\pi_k - \pi\|_{TV}$ . However, many details need to be considered before this bound is of any theoretical or practical use, as for example, the validity of the total variation metric, the sense in which  $\pi_k$  tends to  $\pi$  etc. These problems involve many technical details and are matters of current research (Stuart, 2010; Dashti and Stuart, 2016; Bui-Thanh and Ghattas, 2014; Hairer et al., 2014; Zhou et al., 2017; Scheichl et al., 2017; Ghosal and van der Vaart, 2017).

## 4 Examples

In Sections 4.1 and 4.2 we present two examples considering Bayesian UQ problems for 1D and 2D heat equations, respectively. Using our bound in (3), we compare the resulting posterior distribution to illustrate our approach's performance.

### 4.1 A 1D heat equation inferring the thermal conductivity

Let us return to the thermal conductivity problem for the stationary heat equation in 1D, briefly mentioned in Section 1. Namely

$$-\frac{d}{dx} \left( a(x) \frac{du(x)}{dx} \right) = f(x), \quad x \in (0, 1), \quad (11)$$

subject to Dirichlet boundary conditions  $u(0) = u(1) = 0$ , with forcing term  $f(x) = \sin(\pi x)$  and thermal conductivity  $a(x)$  that varies with the space parameter  $x$ . Here, the functions  $a$  and  $f$  are assumed to be continuous on  $[0, 1]$  and  $0 < \alpha_0 \leq a(x) \leq \alpha_1 < \infty$ .

In this example, the FM is not available analytically, and a numerical FM is used. We use an error estimation in the FM to bound the EABF.

The numerical solution of (11) is computed using the Finite Element Method (FEM), which allows us to calculate a local error estimation in the  $L_2$  norm (see Babuška and Rheinboldt, 1978, for more details), given by

$$\|u_h - u\|_{L_2(I_i)} = \left( \int_{x_{i-1}}^{x_i} (u_h - u)^2 dx \right)^{1/2} \leq \frac{h^2}{\pi^2 a_{\min}^i} \|r\|_{L_2(I_i)}, \quad i = 1, \dots, n,$$

where  $n$  is the number of elements,  $u_h$  the numerical solution with step size  $h$ ,  $I_i = [x_{i-1}, x_i]$ ,  $r(x) = f(x) + \frac{d}{dx} \left( a(x) \frac{du_h(x)}{dx} \right)$  is the residual and

$$a_{\min}^i = \min_{x \in I_i} a(x) = a \left( \frac{x_i + x_{i-1}}{2} \right) (1 + O(h)) \text{ as } h \rightarrow 0.$$

Then, the error estimation  $\hat{K}_0$  is computed by

$$\hat{K}_0 = \max_{I_i} \frac{h^2}{\pi^2 a_{\min}^i} \|r\|_{L_2(I_i)}, \quad i = 1, \dots, n. \quad (12)$$

The inference problem is the estimation of the function  $a(x) = \exp(b(x))$  given observations of  $u_j = u(x_j)$  at fixed locations  $x_j$ ,  $j = 1, \dots, m$ . Certainly, the theoretical and the numerical FMs are continuous.

We simulate a synthetic data set with the true thermal conductivity is  $a(x) = k_0 - r \frac{k_0}{1 + \exp(-xa + \frac{a}{s})}$ , and error model  $y_j = u(x_j) + \sigma \varepsilon_j$ , where  $\varepsilon_j \sim N(0, 1)$ , with the following parameters  $k_0 = 5$ ,  $r = 0.9$ ,  $a = 20$ ,  $s = 2$  and  $\sigma = 0.0005$  (to maintain a 0.01 signal-to-noise ratio). The data are plotted in Figure 1(b). We consider  $m = 30$  observations at locations  $x_j$  regularly spaced between 0 and 1.

In order to define the parametric space, the function  $a$  is represented as a third-order b-spline that passes through the set of points  $\{a_i\}_{i=0}^k$ , where  $a_i = \exp(b(x_i))$ . Therefore, the parameter space is defined by  $\theta = \{b_i\}_{i=0}^k$ , where  $b_i = b(x_i)$ . In this case, since the FEM used is numerically demanding we keep the prior truncation fixed to  $k + 1$  terms; namely the control points  $\{a_i\}_{i=0}^k$ , at locations  $\{x_i\}_{i=0}^k$ , for the third-order b-spline. In this example we take  $k = 20$ .

Regarding the prior distribution for the parameters  $\{b_i\}_{i=0}^k$ , we define their prior using a Gaussian Markov random field (GMRF) with zero mean and sparse precision matrix (inverse-covariance), encoding statistical assumptions regarding the value of each element  $b_i$  based on the values of its neighbors (see details in Bardsley and Kaipio, 2013). We restrict the support of  $-\infty < \beta_1 \leq b(x) \leq \beta_2 < \infty$ , this implies that  $a_i = \exp(b_i) \in [\exp(\beta_1), \exp(\beta_2)]$ . Then the parameter space is compact and there exists a global bound for (12), complying with (3).

With the standard error and sample size used, calculating the error bound for the Forward Map (FM) as stated in (10), we require  $\hat{K}_0 < 2.1 \times 10^{-6}$ . To sample from the posterior distribution, we use the t-walk (Christen and Fox, 2010).

Regarding the numerical solver, we begin with a relatively large step size  $h = 0.02$  (considering  $n = 50$  elements in the FEM) and start the MCMC. At each iteration the FM is first computed along with its error estimation  $\hat{K}_0$ . If the solution  $u_h$  does not satisfy the estimated global bound, i.e.  $\hat{K}_0 > 2.1 \times 10^{-6}$ , we increase the number of elements by 50 ( $h = 1/(n + 50)$ ), until the bound is met. For  $h = 0.0066$ ,  $n = 150$  elements in the FEM, the bound is achieved for all iterations. For comparisons, a smaller grid is considered with  $h = 0.002$ ,  $n = 500$  elements. The results are shown in Figure 1.

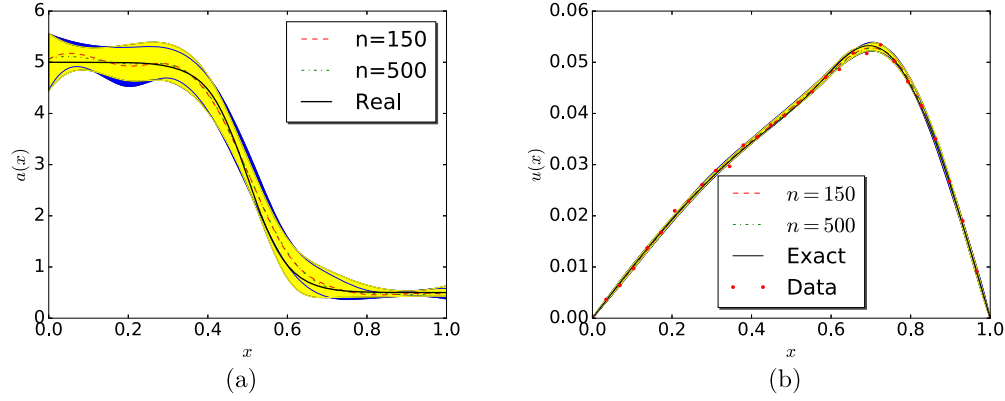


Figure 1: (a) The true conductivity  $a(x)$  (black), the posterior mean with  $n = 150$  elements (red) and  $n = 500$  elements (green) in the FEM. (b) The exact solution  $u(x)$  (black), the posterior mean with  $n = 150$  elements (red) and  $n = 500$  elements (green). Shaded areas represent the uncertainty in the model fit, as draws from the posterior distribution, using 150 elements (blue) and 500 elements (yellow). Note that, if we use a smaller step size than that required by the bound in (10), results are basically the same simply adding CPU time.

We took 50,000 iterations of the t-walk, the MCMC mixes quite well. With  $n = 150$  the sampling took 3 min and with  $n = 500$ , 16 min; in a standard 2.6 GHz processor computer. As seen in Figure 1 the conductivity is recovered and taking  $n = 500$  elements in the FEM results in basically the same posterior as for only  $n = 150$ , which already comply with the EABF bound, only resulting in unnecessary CPU effort.

## 4.2 A 2D heat equation inferring the initial condition

In this example, we consider a more complex 2D PDE inverse problem. The FM is available analytically, and a numerical FM is also used for comparisons. In this case, the numerical error is directly calculated, and we infer only two parameters.

We present a 2D heat equation problem to determine the initial conditions from observations of transient temperature measurements taken within the domain at a time  $t = t_1$ . The heat transfer PDE is given by

$$\begin{aligned} \frac{\partial u}{\partial t} &= \alpha \Delta u, \text{ in } D = (0, 1) \times (0, 1), \\ u(x, y, t) &= 0 \text{ on } \partial D, \\ u(x, y, 0) &= f(x, y). \end{aligned} \quad (13)$$

Taking the forcing term  $f(x, y) = b \sin(\pi x) \sin(\pi y) + c \sin(2\pi x) \sin(\pi y)$  as initial condition, the PDE has an analytical solution

$$u(x, y, t) = b \exp(-2\alpha\pi^2 t) \sin(\pi x) \sin(\pi y) + c \exp(-5\alpha\pi^2 t) \sin(2\pi x) \sin(\pi y).$$

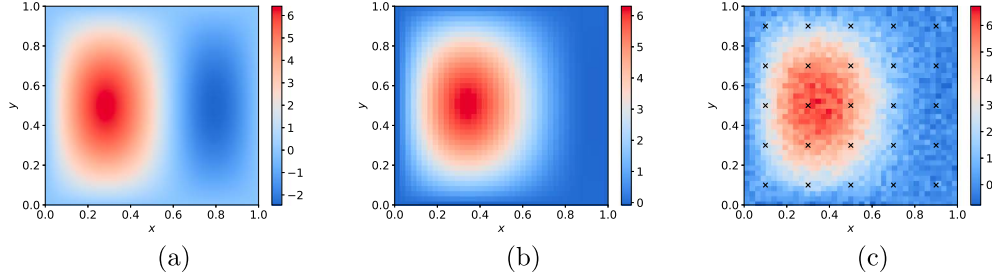


Figure 2: Heat equation in 2D, (a) exact solution at  $t = t_1$ , (b) numerical solution using finite element method with FEniCS with mesh  $40 \times 40$  with  $\Delta t = 0.067$  and (c) numerical solution with an additive Gaussian noise with variance  $\sigma = 0.3$  and data point locations.

A numerical solution of the boundary value problem represented in (13) is also computed using the Finite Element Method (FEM) within FEniCS (Martin et al., 2015), which allows us to calculate the error in the numerical solver using the exact solution.

The inferential problem is to estimate  $\theta = (b, c)$  given measurements of  $u$  at time  $t_1 = 0.3$ . A priori we took independent truncated Gamma distributions for  $b$  and  $c$  with parameters  $(2, 0.7)$  and  $(2, 0.4)$  respectively, both restricted to  $[0, 8]$ . Certainly, the theoretical and the numerical FMs are continuous, and since the support is compact we may conclude that the error bound in (3) exists for all  $\theta$ .

We simulate a synthetic data set with the error model

$$Y_i = u(x_i, y_i, t_1) + \sigma \varepsilon_i,$$

where  $\varepsilon_i \sim N(0, 1)$ ,  $i = 1, \dots, m$ ,  $\sigma = 0.3$  (using a signal to noise ratio of 5%), with  $b = 3$  and  $c = 5$ . The data are plotted in Figure 2(b). We consider  $m = 25$  observations,  $(x_i, y_i)$ ,  $i = 1, 2, \dots, m$  regularly spaced on  $D$ . Since we have an analytic solution, if we run the PDE solver we may calculate the maximum absolute error,  $K_0$ , exactly. The error bound for the FM as stated in (10) is  $\simeq 0.0015$ . To sample from the posterior distribution, we also use the t-walk (Christen and Fox, 2010).

Regarding the numerical solver we start with a large step size of  $\Delta x = \Delta y = 0.1$  and  $\Delta t = 0.268$ , and calculate  $K_0$ . If the solution does not comply with the bound, that is, if  $K_0 > 0.0015$ , a new solution is attempted by reducing the step-size in  $\Delta x$ ,  $\Delta y$  and  $\Delta t$  by half, until the global absolute errors are within the bound, i.e.,  $K_0 \leq 0.0015$ . The resulting mesh is  $\Delta x = \Delta y = 0.025$  and  $\Delta t = 0.067$ .

We compare the above FEM numerical FM with the exact FM, with 250,000 iterations of our MCMC. The results are shown in Figure 3 and in Table 1. The differences observed in both results may be attributed to the Monte Carlo sampling.



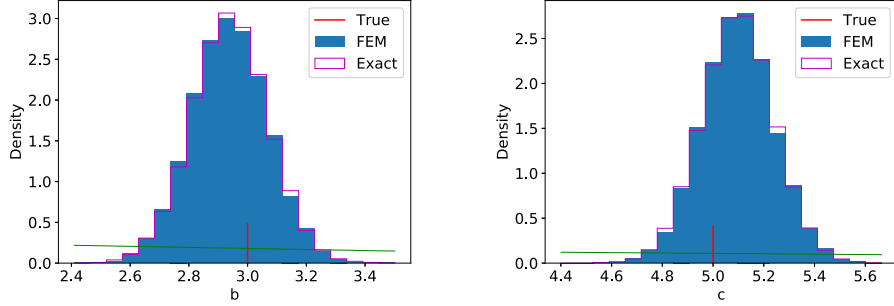


Figure 3: Comparison between numerical (blue) and theoretical (magenta) posteriors for both parameters in the initial conditions of the 2D heat equation. The prior, in this scale, appears in green.

	$b$	$c$
True	3.0	5.0
PM-Exact	2.9396	5.0966
PM-FEM	2.9377	5.0969

Table 1: Comparison of the Posterior Mean for parameters  $b$  and  $c$  using the exact FM and the FEM approximate FM.

## 5 Discussion

The generalization of results from Capistrán et al. (2016) to a priori statements and Banach parameter spaces makes our posterior error control strategy more feasible, general, and applicable. Our results lead to the error bound for the FM’s error in (10) that is quite simple to calculate, provided after-the-fact error estimates of the FM numerical approximation are available. Here we presented two examples. We have experimented with several others. In all cases, using (10) the numerical error in the posterior was controlled successfully, leading to a negligible increase in posterior numerical precision if a more precise FM was considered. This, in turn, may result in CPU time save, as cheaper/rougher solvers may be used.

We have not discussed the scenario when error parameters  $\sigma$  are not known and/or there is correlation in the data. In this case, we may consider that a priori  $\theta$  and  $\sigma$  are independent, a correlation structure for  $y$ , and equivalent results should follow; this was discussed in a previously unpublished version of this manuscript but not here (Christen et al., 2016). We only need to prove that the new likelihood follows assumption 2.1, in particular, that it is bounded  $\lambda$ -a.s.

In general, one needs to decide, sooner or later, a discretization level  $n$ . Although there are multilevel methods that work with several discretizations simultaneously (Cliffe et al., 2011; Katsiolides et al., 2018), even, in that case, one nevertheless needs, a

priori, a sense of what is a large or a small FM error. Our bound in (10) is an attempt to provide precisely that, a sense of what is a large or a small FM error in the perspective of the Bayesian inverse problem at hand, considering the noise level in the data ( $\sigma$ ) and the sample size.

Since the parameter  $\theta$  in many UQ applications is an unknown function, one may use a Bayesian nonparametric approach to estimate  $\theta$  (BNP, see, for example Müller et al., 2015, chap. 4), using a prior measure on a set of regressors with the machinery of BNP (see Giordano and Kekkonen, 2020, and references therein).

We have not discussed here the problem of model fit/model adequacy since we assumed from the onset that the actual distribution for the data  $y$  is  $Z(y)$ . The question of model adequacy is always present in applications since, indeed “all models are wrong” and certainly, it is a concern in inverse problems and computer experiments (Kennedy and O’Hagan, 2001). Generalizing our results to the scenario where the actual sampling distribution is  $Z^*(y)$ , different from  $Z(y)$ , is an interesting next step but is left for future research.

Note that decreasing FM solver precision can only be done within the stable regime of the solver used. Moreover, in real case applications, increasing the mesh size or any mesh refinements may come at a great coding effort, for example, in a large scale 3D geothermal inversion (Cui et al., 2011). Our approach only makes sense in the case where working with different discretizations for the FM is computationally feasible.

## References

- Babuška, I. and Rheinboldt, W. C. (1978). “A-posteriori error estimates for the finite element method.” *International Journal for Numerical Methods in Engineering*, 12(10): 1597–1615. [MR0488846](#). 12
- Bardsley, J. M. and Kaipio, J. (2013). “Gaussian Markov random field priors for inverse problems.” *Inverse Problems & Imaging*, 7(2): 397–416. [MR3063540](#). doi: <https://doi.org/10.3934/ipi.2013.7.397>. 13
- Berger, J. O. and Pericchi, L. R. (1996). “The Intrinsic Bayes Factor for Model Selection and Prediction.” *Journal of the American Statistical Association*, 91(433): 109–122. URL <http://www.jstor.org/stable/2291387> [MR1394065](#). doi: <https://doi.org/10.2307/2291387>. 8
- Brenner, S. and Scott, R. (2008). *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York. [MR2373954](#). doi: <https://doi.org/10.1007/978-0-387-75934-0>. 4
- Bui-Thanh, T. and 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. URL <https://doi.org/10.1137/120894877> [MR3283906](#). doi: <https://doi.org/10.1137/120894877>. 5, 12

- Bui-Thanh, T., Ghattas, O., Martin, J., and Stadler, G. (2013). “A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion.” *SIAM Journal on Scientific Computing*, 35(6): A2494–A2523. URL <https://doi.org/10.1137/12089586X> MR3126997. doi: <https://doi.org/10.1137/12089586X>. 8
- Cai, C., Mohammad-Djafari, A., Legoupil, S., and Rodet, T. (2011). “Bayesian data fusion and inversion in X-ray multi-energy computed tomography.” In *Proceedings - International Conference on Image Processing, ICIP*, 1377–1380. 2
- Capistrán, M., Christen, J., and Donnet, S. (2016). “Bayesian analysis of ODE’s: Solver optimal accuracy and Bayes factors.” *Journal of Uncertainty Quantification*, 4(1): 829–849. MR3529113. doi: <https://doi.org/10.1137/140976777>. 1, 2, 3, 8, 16
- Cash, J. R. and Karp, A. H. (1990). “A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides.” *ACM Transactions on Mathematical Software*, 16(3): 201–222. URL <http://doi.acm.org/10.1145/79505.79507> MR1070798. doi: <https://doi.org/10.1145/79505.79507>. 11
- Chama, Z., Mansouri, B., Anani, M., and Mohammad-Djafari, A. (2012). “Image recovery from Fourier domain measurements via classification using Bayesian approach and total variation regularization.” *AEU - International Journal of Electronics and Communications*, 66(11): 897–902. 2
- Christen, J., Capistrán, M., and Moreles, M. (2016). “Numerical posterior distribution error control and expected Bayes Factors in the bayesian Uncertainty Quantification of Inverse Problems.” *ArXiv e-prints*. 16
- Christen, J. and Fox, C. (2010). “A general purpose sampling algorithm for continuous distributions (the t-walk).” *Bayesian Analysis*, 5(2): 263–282. MR2719653. doi: <https://doi.org/10.1214/10-BA60>. 13, 15
- Christen, J. and Pérez-Garmendia, J. L. (2020). “Weak and TV consistency in Bayesian Uncertainty Quantification using disintegration.” *Boletín de la Sociadad Matemática Mexicana (accepted)*. 5, 6, 8
- Cliffe, K., Giles, M., Scheichl, R., and Teckentrup, A. L. (2011). “Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients.” *Computing and Visualization in Science*, 14(3): <https://doi.org/10.1007/s00791-011-0160-x>. MR2835612. doi: <https://doi.org/10.1007/s00791-011-0160-x>. 16
- Clyde, M. A., Berger, J. O., Bullard, F., Ford, E. B., Jefferys, W. H., and Luo, R. (2007). “Current Challenges in Bayesian Model Choice.” In Babu, G. J. and Feigelson, E. D. (eds.), *Statistical Challenges in Modern Astronomy IV*, volume 371 of *ASP Conference Series*, 224–240. <https://ui.adsabs.harvard.edu/abs/2007ASPC..371..224C>. 9
- Collis, J., Connor, A. J., Paczkowski, M., Kannan, P., Pitt-Francis, J., Byrne, H. M., and Hubbard, M. E. (2017). “Bayesian calibration, validation and uncertainty quantification for predictive modelling of tumour growth: A tutorial.” *Bulletin of*

- Mathematical Biology*, 79 4: 939–974. MR3628429. doi: <https://doi.org/10.1007/s11538-017-0258-5>. 2
- Cotter, S. L., Dashti, M., and Stuart, A. M. (2010). “Approximation of Bayesian inverse problems for PDEs.” *SIAM Journal on Numerical Analysis*, 48(1): 322–345. URL <https://doi.org/10.1137/090770734> MR2608372. doi: <https://doi.org/10.1137/090770734>. 2
- Cui, T., Fox, C., and O’Sullivan, M. J. (2011). “Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm.” *Water Resources Research*, 47(10). 2, 17
- Cui, T., Fox, C., and O’Sullivan, M. J. (2019). “A posteriori stochastic correction of reduced models in delayed-acceptance MCMC, with application to multiphase sub-surface inverse problems.” *International Journal for Numerical Methods in Engineering*, 118(10): 578–605. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.602%8> MR3942044. doi: <https://doi.org/10.1002/nme.6028>. 2
- Dashti, M. and Stuart, A. M. (2011). “Uncertainty Quantification and Weak Approximation of an Elliptic Inverse Problem.” *SIAM Journal on Numerical Analysis*, 49(6): 2524–2542. URL <https://doi.org/10.1137/100814664> MR2873245. doi: <https://doi.org/10.1137/100814664>. 2
- Dashti, M. and Stuart, A. M. (2016). “The Bayesian approach to inverse problems.” *Handbook of Uncertainty Quantification*, 1–118. MR3839555. 1, 5, 6, 7, 12
- de Valpine, P. (2008). “Improved estimation of normalizing constants from Markov chain Monte Carlo output.” *Journal of Computational and Graphical Statistics*, 17(2): 333–351. URL <http://www.jstor.org/stable/27594310> MR2439963. doi: <https://doi.org/10.1198/106186008X320258>. 3
- Di Pietro, D. A. and Ern, A. (2011). *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69. Springer Science & Business Media. MR2882148. doi: <https://doi.org/10.1007/978-3-642-22980-0>. 11
- Fox, C., Haario, H., and Christen, J. A. (2013). *Bayesian Theory and Applications*, Chapter 31. Inverse problems. Oxford University Press. MR3221185. 1
- Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference. Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press. MR3587782. doi: <https://doi.org/10.1017/9781139029834>. 5, 6, 12
- Gibbs, A. L. and Su, F. E. (2002). “On choosing and bounding probability metrics.” *International Statistical Review*, 70(3): 419–435. URL <http://dx.doi.org/10.1111/j.1751-5823.2002.tb00178.x> 8
- Giordano, M. and Kekkonen, H. (2020). “Bernstein–von Mises theorems and uncertainty quantification for linear inverse problems.” *SIAM/ASA Journal on Uncertainty Quantification*, 8(1): 342–373. URL <https://doi.org/10.1137/18M1226269> MR4069334. doi: <https://doi.org/10.1137/18M1226269>. 17

- Giovanelli, J.-F. and Idier, J. (2015). *Regularization and Bayesian Methods for Inverse Problems in Signal and Image Processing*. New York: John Wiley & Sons. [MR3524962](#). doi: <https://doi.org/10.1002/9781119125495>. 2
- Hairer, M., Stuart, A. M., and Vollmer, S. J. (2014). “Spectral gaps for a Metropolis–Hastings algorithm in infinite dimensions.” *Ann. Appl. Probab.*, 24(6): 2455–2490. URL <https://doi.org/10.1214/13-AAP982> [MR3262508](#). doi: <https://doi.org/10.1214/13-AAP982>. 12
- Hazelton, M. L. (2010). “Bayesian inference for network-based models with a linear inverse structure.” *Transportation Research Part B: Methodological*, 44(5): 674–685. 2
- Hesthaven, J. S. and Warburton, T. (2007). *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*, 1st edition. Springer Publishing Company, Incorporated. [MR2372235](#). doi: <https://doi.org/10.1007/978-0-387-72067-8>. 11
- Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). “Bayesian model averaging: A tutorial.” *Statistical Science*, 14(4): 382–401. [MR1765176](#). doi: <https://doi.org/10.1214/ss/1009212519>. 8
- Hosseini, B. and Nigam, N. (2017). “Well-posed Bayesian inverse problems: Priors with exponential tails.” *SIAM/ASA Journal on Uncertainty Quantification*, 5(1): 436–465. URL <https://doi.org/10.1137/16M1076824> [MR3640629](#). doi: <https://doi.org/10.1137/16M1076824>. 8
- Hutchinson, M., Oh, H., and Chen, W.-H. (2017). “A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors.” *Information Fusion*, 36: 130–148. URL <http://www.sciencedirect.com/science/article/pii/S156625351630152X> 2
- Iserles, A. (2009). *A First Course in the Numerical Analysis of Differential Equations*, volume 44. Cambridge University Press. [MR2478556](#). 2, 11
- Jeffreys, H. (1961). *Theory of Probability*, 3rd edition. Oxford, England: Oxford University Press. [MR0187257](#). 9
- Kahle, C., Lam, K., Latz, J., and Ullmann, E. (2019). “Bayesian parameter identification in Cahn–Hilliard models for biological growth.” *SIAM/ASA Journal on Uncertainty Quantification*, 7(2): 526–552. URL <https://doi.org/10.1137/18M1210034> [MR3945813](#). doi: <https://doi.org/10.1137/18M1210034>. 2
- Kaipio, J. P. and Fox, C. (2011). “The Bayesian framework for inverse problems in heat transfer.” *Heat Transfer Engineering*, 32(9): 718–753. 2
- Kaipio, J. P. and Somersalo, E. (2005). *Statistical and Computational Inverse Problems*. Applied Mathematical Sciences. Springer Science+Business Media, Incorporated. [MR2102218](#). 1

- Kass, R. E. and Raftery, A. E. (1995). “Bayes factors.” *Journal of the American Statistical Association*, 90(430): 773–795. URL <https://amstat.tandfonline.com/doi/abs/10.1080/01621459.1995.10476572> MR3363402. doi: <https://doi.org/10.1080/01621459.1995.10476572>. 8, 9
- Katsiolides, G., Müller, E. H., Scheichl, R., Shardlow, T., Giles, M. B., and Thomson, D. J. (2018). “Multilevel Monte Carlo and improved time-stepping methods in atmospheric dispersion modelling.” *Journal of Computational Physics*, 354: 320–343. URL <http://www.sciencedirect.com/science/article/pii/S0021979917307945> MR3738110. doi: <https://doi.org/10.1016/j.jcp.2017.10.035>. 16
- Keats, A., Yee, E., and Lien, F. (2010). “Information-driven receptor placement for contaminant source determination.” *Environmental Modelling and Software*, 25(9): 1000–1013. 2
- Kennedy, M. C. and O’Hagan, A. (2001). “Bayesian calibration of computer models.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3): 425–464. URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00294> MR1858398. doi: <https://doi.org/10.1111/1467-9868.00294>. 17
- Kozawa, S., Takenouchi, T., and Ikeda, K. (2012). “Subsurface imaging for anti-personal mine detection by Bayesian super-resolution with a smooth-gap prior.” *Artificial Life and Robotics*, 16(4): 478–481. 2
- Lie, H. C., Sullivan, T. J., and Teckentrup, A. L. (2018). “Random forward models and log-likelihoods in Bayesian inverse problems.” *SIAM/ASA Journal on Uncertainty Quantification*, 6(4): 1600–1629. URL <https://doi.org/10.1137/18M1166523> MR3876869. doi: <https://doi.org/10.1137/18M1166523>. 2
- Martin, S. A., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M. E., and Wells, G. N. (2015). “The FEniCS Project Version 1.5.” *Archive of Numerical Software*, 3(100): 9–23. 15
- Moral, P. D., Jasra, A., Law, K. J., and Zhou, Y. (2017). “Multilevel sequential Monte Carlo samplers for normalizing constants.” *ACM Transactions on Modeling and Computer Simulation*, 27(3): 20. MR3703050. doi: <https://doi.org/10.1145/3092841.3>
- Müller, P., Quintana, F., Jara, A., and Hanson, T. (2015). *Bayesian Nonparametric Data Analysis*. Springer International Publishing. MR3309338. doi: <https://doi.org/10.1007/978-3-319-18968-0>. 17
- Nissinen, A., Kolehmainen, V. P., and Kaipio, J. P. (2011). “Compensation of modelling errors due to unknown domain boundary in electrical impedance tomography.” *IEEE Transactions on Medical Imaging*, 30(2): 231–242. 2
- O’Hagan, A. (1997). “Properties of intrinsic and fractional Bayes factors.” *Test*, 6(1): 101–118. URL <https://doi.org/10.1007/BF02564428> MR1466435. doi: <https://doi.org/10.1007/BF02564428>. 8

- Osthus, D., Vander Wiel, S., Hoffman, N., and Wysocki, F. (2019). “Prediction uncertainties beyond the range of experience: A case study in inertial confinement fusion implosion experiments.” *SIAM/ASA Journal on Uncertainty Quantification*, 7(2): 604–633. URL <https://doi.org/10.1137/17M1158860> MR3953472. doi: <https://doi.org/10.1137/17M1158860>. 2
- Quarteroni, A. and Valli, A. (2008). *Numerical Approximation of Partial Differential Equations*, volume 23. Springer Science & Business Media. MR1299729. 2, 11
- Reddy, J. N. (2006). *Introduction to the Finite Element Method*, 3rd edition. New York: McGraw-Hill Education. URL <https://www.accessengineeringlibrary.com/content/book/9%780072466850> MR0286668. 4
- Scheichl, R., Stuart, A. M., and Teckentrup, A. L. (2017). “Quasi-Monte Carlo and multilevel Monte Carlo methods for computing posterior expectations in elliptic inverse problems.” *SIAM/ASA Journal on Uncertainty Quantification*, 5(1): 493–518. URL <https://doi.org/10.1137/16M1061692> MR3640631. doi: <https://doi.org/10.1137/16M1061692>. 6, 7, 12
- Schervish, M. J. (1997). *Theory of Statistics (Springer Series in Statistics)*, 1st edition. Springer, 1995. Corr. 2nd printing edition. URL <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0387945466> MR1354146. doi: <https://doi.org/10.1007/978-1-4612-4250-5>. 5, 6
- Stuart, A. M. (2010). “Inverse problems: A Bayesian perspective.” *Acta Numerica*, 19: 451–559. MR2652785. doi: <https://doi.org/10.1017/S0962492910000061>. 5, 12
- Sun, L., Lu, Y., Jin, J. G., Lee, D.-H., and Axhausen, K. W. (2015). “An integrated Bayesian approach for passenger flow assignment in metro networks.” *Transportation Research Part C: Emerging Technologies*, 52: 116–131. URL <http://www.sciencedirect.com/science/article/pii/S0968090X15000030> 2
- van der Linden, S. and Chryst, B. (2017). “No need for Bayes factors: A fully Bayesian evidence synthesis.” *Frontiers in Applied Mathematics and Statistics*, 3: 12. URL <https://www.frontiersin.org/article/10.3389/fams.2017.0%0012> 8
- Xu, X., Lu, P., MacEachern, S. N., and Xu, R. (2019). “Calibrated Bayes factors for model comparison.” *Journal of Statistical Computation and Simulation*, 89(4): 591–614. URL <https://doi.org/10.1080/00949655.2018.1563091> MR3904590. doi: <https://doi.org/10.1080/00949655.2018.1563091>. 8
- Yan, L. and 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. URL <https://doi.org/10.1137/140965144> MR3352614. doi: <https://doi.org/10.1137/140965144>. 2
- Zhou, Q., Hu, Z., Yao, Z., and Li, J. (2017). “A hybrid adaptive MCMC algorithm in function spaces.” *SIAM/ASA Journal on Uncertainty Quantification*, 5(1): 621–639. URL <https://doi.org/10.1137/16M1082950> MR3670264. doi: <https://doi.org/10.1137/16M1082950>. 12



Zhu, S., You, P., Wang, H., Li, X., and Mohammad-Djafari, A. (2011). “Recognition-oriented Bayesian SAR imaging.” In *2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar, APSAR 2011*, 153–156. [2](#)

### **Acknowledgments**

We thank Tan Bui-Thanh, Peter Müller, José Luis Perez Garmendia, Fernanda Méndez, the two reviewers, AE and Editor of BA for invaluable comments during the many previous drafts of the paper. This research is partially founded by CONACYT CB-2016-01-284451 and COVID19 312772, and RDECOMM and ONRG grants.