Stochastic state estimation via incremental iterative sparse polynomial chaos based Bayesian-Gauss-Newton-Markov-Kalman filter

Bojana Rosić
Applied Mechanics and Data Analysis
University of Twente
Netherlands

September 17, 2019

In this paper is proposed a novel incremental iterative Gauss-Newton-Markov-Kalman filter method for state estimation of dynamic models given noisy measurements. The filter is constructed by projecting the random variable representing the unknown state onto the subspace generated by data. The approximation of projection, i.e. the conditional expectation of the state given data, is evaluated by minimising the expected Bregman's loss. The mathematical formulation of the proposed filter is based on the construction of an optimal nonlinear map between the observable and parameter (state) spaces via a convergent sequence of linear maps obtained by successive linearisation of the observation operator in a Gauss-Newton-like form. To allow automatic linearisation of the dynamical system in a sparse form, the smoother is designed in a hierarchical setting such that the forward map and its linearised counterpart are estimated in a Bayesian manner given a forecasted data set. For this purpose the relevance vector machine approach is used. To improve the algorithm convergence, the smoother is further reformulated in its incremental form in which the current and intermediate states are assimilated before the initial one, and the corresponding posterior estimates are taken as pseudo-measurements. As the latter ones are random variables, and not deterministic any more, the novel stochastic iterative filter is designed to take this into account. To correct the bias in the posterior outcome, the procedure is built in a predictor-corrector form in which the predictor phase is used to assimilate noisy measurement data, whereas the corrector phase is constructed to correct the mean bias. The resulting filter is further discretised via time-adapting sparse polynomial chaos expansions obtained either via modified Gram-Schmidt orthogonalisation or by a carefully chosen nonlinear mapping, both of which are estimated in a Bayesian manner by promoting the sparsity of the outcomes. The time adaptive basis with non-Gaussian arguments is further mapped to the polynomial chaos one by a suitably chosen isoprobabilistic transformation. Finally, the proposed method is tested on a chaotic nonlinear Lorenz 1984 system.

1 Introduction

Probabilistic inverse estimation is gaining momentum in computational practice today. Bayes's rule as given in its classical form often cannot be used in practice because the evaluation of the posterior distribution requires the use of slowly convergent random walk strategies such as Markov chain Monte Carlo-like algorithms [10, 25, 24]. On the other hand, its linear approximation in the form of a Kalman filter [14] became a very important industrial tool for the prediction/forecast of the system state describing various types of dynamical systems. However, Kalman filters are not good at coping with highly nonlinear system responses, and many attempts have been made to resolve this issue. The vast majority of studies on this subject can be broadly classified into two groups: stochastic strategies based on the sequential Monte Carlo algorithm also known as particle/ensemble filters (e.g. [20, 7]), and deterministic methods based on the linearisation of the measurement operator such as extended [9, 13] and unscented [27, 18] Kalman filters. The former theories are based on the approximation of the posterior distribution via a convex combination of the Dirac delta measure such that the corresponding filter requires only few simulation calls. But, it is well known that the ensemble in the particle form may collapse, which is especially evident for small ensembles. On the other hand, the deterministic filters based on the first order Taylor expansion of the measurement operator may become inaccurate when used in a highly nonlinear setting.

It is well known that the Bayesian update is theoretically based on the notion of conditional expectation [3]. Here the conditional expectation is not only used as a theoretical basis, but also as a basic computational tool for the identification of the initial state of the dynamical system. Being a unique optimal projector for all Bregman's loss functions, the conditional expectation allows the estimation of the posterior moments by finding an optimal map between the measurement and the parameter/state space that minimises the expected Bregman's loss. Therefore, being able to numerically approximate conditional expectations, one can build various filtering techniques for the state assimilation. To accommodate the nonlinearities present in the estimation problem, in this paper an iterative version of the filter in the Gauss-Newton form is suggested for the backpropagation of information on the state in the current time moment to the initial one. Several previous studies have investigated the linearisation idea by building the filter either as an iterative version of ensemble Kalman filters as presented in [22, 2], or procedures coming from the randomised likelihood (e.g. [6]) and maximum a posteriori error estimate (e.g. [28]). In this paper the iterative filtering technique is based on the approximation of the conditional expectation of the state given observation, as well as its inverse map, via a sequence of linearised maps obtained by minimising the corresponding expected quadratic Bregman's loss functions, or by using Bayesian estimation. In this manner the Gauss-Newton filtering procedure obtains its hierarchical structure and does not require special differentiation techniques as the estimation of the Jacobian comes as the by-product. To improve the local convergence, the Gauss-Newton estimation is here improved by substituting the direct state estimation with the incremental one based on the pseudo-time discretisations. The idea is to build the optimal map between the observation and the initial state as a composition of linearised maps displaying the intermediate state posteriors chacareterised by pseudo-time discretisations. In contrast to the direct estimation this approach takes the estimated intermediate states as pseudo-measurements for the preceding ones. Hence, the dynamic of the filter's incremental form is driven by pseudo-time stepping in which the global optimal linear map of one update step is substituted by few optimal local maps obtained by splitting the update step into smaller increments (pseudo-update steps). As the pseudo-measurements are random variables and not

deterministic ones, here is suggested a novel stochastic Gauss-Newton filter for the state estimation in a predictor-corrector form.

In contrast to most sampling approaches to Bayesian updating that typically start from the classical formulation involving conditional measures and densities, the conditional expectation as the computationally prime object allows a direct estimation of the posterior random variable in a functional approximation form. As a stochastic Gauss-Newton filter operates on random variables, not densities, its numerical implementation is achieved by discretising the random variables of consideration via time dependent polynomial chaos expansion (PCEs). The time adaptive nature of discretisation is used to prevent an overestimation of the measurement prediction after long-time integration, which is known to be a side-effect of the classical polynomial chaos representations. Therefore, the observation random variables are first discretised in a non-Gaussian basis, which is further transformed to the Gaussian one by a nonlinear isoprobabilistic transformation. The non-Gaussian basis is chosen either as an orthogonal one by employing the stochastic modified Gram-Schmidt orthogonalisation as already discussed by [11] for purely uncertainty quantification purposes, or as a non-orthogonal one taking the form of a nonlinear polynomial map between two consecutive states. To promote for sparsity, the functional representations are estimated in a data-driven Bayesian way by using the relevance vector machine approach [26]. By using the sparse time dependent PCE approximations, the filter is finally designed in its minimal form that is estimated by using a minimal number of model evaluations.

The paper is organised as follows: Section 2 gives a concise introduction to the Bayesian state estimation of the abstract dynamical system. Section 3 considers the approximate Bayesian estimation from a conditional expectation point of view. Numerical approximations of conditional expectation are shortly studied in Section 4, and hence the Gauss-Newton filtering procedure is introduced. The Bayesian point of view on the Gauss-Newton filter is further studied in Section 5, whereas its incremental version in predictor-corrector form is discussed in Section 6. The filter discretisation and its computational form are given in Section 7. Here the filter is studied from the perspective of time adaptive sparse random variable discretisations. The paper is concluded with Section 8.

2 Model problem

Let the state of the dynamical system $x \in \mathbb{R}^d$ satisfying the nonlinear initial value problem

$$\dot{x} = f(x,t), \quad x_0 = x(0) \tag{1}$$

be observed in time moments $0 \leq t_k \leq T$, $t_k = k\Delta t, k \in \mathbb{N}_0$ given time increment Δt via

$$y_k = Y(x_n) \tag{2}$$

in which Y is a nonlinear observation operator, whereas x_n, $n \in \mathbb{N}_0$ either denotes the current state when $n = k$, or an unknown previous state when $x_n = x_{k-r}$ for $r \in \mathbb{N}$, respectively. Assuming that y_k is possibly not measured in its full component form, i.e. $y_k \in \mathbb{R}^m$, $m \leq d$, the goal is to estimate the state x_n given noisy measurements

$$y^{mes} = Y(x^{tru}) + \hat{\varepsilon} \tag{3}$$

in which x^{tru} denotes the so-called truth, whereas $\hat{\varepsilon}$ stands for the corresponding realisation of the measurement noise.

Formally, in a Bayesian setting the unknown state x_n in Eq. (3) is modelled as a random variable (a priori knowledge or forecast)

$$x_{nf}(\omega) : \Omega \to \mathbb{R}^d \tag{4}$$

on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with the set of elementary events Ω, a σ-algebra of measurable events $\mathcal{F}$, and a probability measure $\mathbb{P}$. The common

choice is to assume that $x_{nf} \in \mathcal{X} := L_2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$, the space of real valued random variables with finite variance. As x_n is a random variable, so is the observation in Eq. (2), here obtaining the form of

$$\mathcal{Y} \ni y_{kf}(\omega) = Y(x_{nf}(\omega)) + \varepsilon_k(\omega) \tag{5}$$

in which $\varepsilon_k(\omega) \sim \mathcal{N}(0, C_{\varepsilon_k})$ forecasts the measurement error usually taking the form of zero-mean Gaussian noise with covariance C_{ε_k}.

Assuming that x_n and y_k have a joint probability density function $\pi(x_n, y_k)$, one may use Bayes's theorem in its density form

$$\pi_{x|y}(x_n|y_k) = \frac{\pi(x_n, y_k)}{P(y_k)} = \frac{\pi_{y|x}(y_k|x_n)\pi_x(x_n)}{P(y_k)} \tag{6}$$

to incorporate (assimilate) new information y^{mes} into the probabilistic description given in Eqs. (4)-(5). Here, $\pi_x(x_n)$ denotes the prior density function, $\pi_{y|x}(y_k|x_n)$ is the likelihood, the form of which depends on the measurement error, and $P(y_k) = \int_\Omega \pi(x_n, y_k) dx_n$ is the normalisation factor or evidence. If both the prior and the likelihood are conjugate, i.e. belong to the exponential family of distributions with predefined statistics, the posterior $\pi_{x|y}(x_n|y_k)$ in Eq. (6) can be analytically evaluated. Otherwise, the estimation boils down to computationally intense random walk algorithms of the Markov chain Monte Carlo type. However, both computations essentially lead to the extraction of neccessary information from the posterior by evaluating some form of expectation w.r.t. the posterior, an example of which is the conditional mean

$$\mathbb{E}(x_n|y_k) = \int_\Omega x_n \pi_{x|y}(x_n|y_k) \mathrm{d}x_n. \tag{7}$$

Having done so, one may avoid expensive evaluation of the full posterior by targeting a direct calculation of desired estimates such as the one given in Eq. (7). To achieve this, one may design filtering procedures based on conditional expectation as further described.

3 Conditional expectation

The conditional expectation is defined as the unique optimal projector for all Bregman's loss functions (BLFs) [5]

$$x^* := \mathbb{E}(x|\mathfrak{B}) = \underset{\hat{x} \in L_2(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{R}^d)}{\arg\min} \mathbb{E}(\mathcal{D}_\phi(x, \hat{x})) \tag{8}$$

over all $\mathcal{B}$-measurable random variables $\hat{x}$ in which $\mathcal{B} := \sigma(y)$ is the sub-σ-algebra generated by measurement y. The Bregman's loss function is defined as

Definition 3.1. *Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a strictly convex, differentiable function. Then the Bregman loss function $\mathcal{D}_\phi : \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}_+ := [0, +\infty)$ is defined as*

$$\mathcal{D}_\phi(x, y) = \mathcal{H}(x) - \mathcal{H}(y) = \phi(x) - \phi(y) - \langle x - y, \nabla\phi(y) \rangle \tag{9}$$

in which $\mathcal{H}(x) = \phi(y) + \langle x - y, \nabla\phi(y) \rangle$ is hyperplane tangent to ϕ at point y.

The optimality in Eq. (8) then follows from [1]

Theorem 3.2. *Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a strictly convex, differentiable function and let $\mathcal{D}_\phi$ be the corresponding BLF. Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be an arbitrary probability space and let $\mathfrak{B}$ be a sub-σ-algebra of $\mathfrak{F}$. Let x be any $\mathfrak{F}$-measurable random variable taking values in $\mathbb{R}^d$ for which both $\mathbb{E}(x)$ and $\mathbb{E}(\phi(x))$ are finite. Then, among all $\mathfrak{B}$-measurable random variables, the conditional expectation is the unique minimiser (up to a.s. equivalence) of the expected Bregman loss, i.e.*

$$x^* := \mathbb{E}(x|\mathfrak{B}) = \underset{\hat{x} \in L_2(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{R}^d)}{\arg\min} \mathbb{E}(\mathcal{D}_\phi(x, \hat{x})). \tag{10}$$

The proof of the theorem can be shortly sketched as follows:

Proof. Let $\hat{x}$ be any $\mathfrak{B}$-measurable random variable, and $x^* = \mathbb{E}(x|\mathfrak{B})$, then one has

$$\begin{aligned} \mathbb{E}(\mathcal{D}_\phi(x, \hat{x})) - \mathbb{E}(\mathcal{D}_\phi(x, x^*)) = \mathbb{E}(\phi(x^*) - \phi(\hat{x}) \\ -\langle x - \hat{x}, \nabla\phi(\hat{x}) \rangle + \langle x - x^*, \nabla\phi(x^*) \rangle. \end{aligned} \tag{11}$$

Using the law of total expectation, e.g. $\mathbb{E}(x) = \mathbb{E}(\mathbb{E}(x|\mathfrak{B}))$, one may further state

$$\begin{aligned}\mathbb{E}(\langle x-\hat{x}, \nabla\phi(\hat{x})\rangle) &= \mathbb{E}\left(\mathbb{E}(\langle x-\hat{x}, \nabla\phi(\hat{x})\rangle|\mathfrak{B})\right)\\ &= \mathbb{E}(\langle \mathbb{E}(x|\mathfrak{B})-\hat{x}, \nabla\phi(\hat{x})\rangle)\\ &= \mathbb{E}(x^*-\hat{x}, \nabla\phi(\hat{x})) \qquad (12)\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbb{E}(\langle x-x^*, \nabla\phi(\hat{x})\rangle) &= \mathbb{E}(\mathbb{E}(\langle x-x^*, \nabla\phi(\hat{x})\rangle|\mathfrak{B}))\\ &= \mathbb{E}(x^*-x^*, \nabla\phi(\hat{x}))\\ &\equiv 0. \qquad (13)\end{aligned}$$

Following this, the relation in Eq. (11) reduces to

$$\begin{aligned}&\mathbb{E}(\mathcal{D}_\phi(x,\hat{x})) - \mathbb{E}(\mathcal{D}_\phi(x,x^*)) =\\ &\mathbb{E}(\phi(x^*)-\phi(\hat{x})-\langle x^*-\hat{x}, \nabla\phi(\hat{x})\rangle)\\ &= \mathbb{E}(\mathcal{D}_\phi(x^*,\hat{x})). \qquad (14)\end{aligned}$$

□

The last relation in Eq. (14) defines the Bregman Pythagorean inequality

$$\mathbb{E}(\mathcal{D}_\phi(x,\hat{x})) \geq \mathbb{E}(\mathcal{D}_\phi(x,x^*)) + \mathbb{E}(\mathcal{D}_\phi(x^*,\hat{x})) \qquad (15)$$

such that one may state

Theorem 3.3. *Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a strictly convex, differentiable function and let $\mathcal{D}_\phi$ be the corresponding BLF. Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be an arbitrary probability space and let $\mathfrak{B}$ be a sub-σ-algebra of $\mathfrak{F}$. Let $\hat{x}$ and x be any $\mathfrak{F}$-measurable random variable taking values in $\mathbb{R}^d$ for which both pairs $(\mathbb{E}(\hat{x}), \mathbb{E}(x))$ and $(\mathbb{E}(\phi(\hat{x})), \mathbb{E}(\phi(x)))$ are finite. Then, we have*

$$\mathbb{E}(\mathcal{D}_\phi(x,\hat{x})) \geq \mathbb{E}(\mathcal{D}_\phi(x,x^*)) + \mathbb{E}(\mathcal{D}_\phi(x^*,\hat{x})) \qquad (16)$$

in which the unique point x^* is called the Bayesian projection of x onto $\mathfrak{B}$ and is defined as following

$$x^* := \mathbb{E}(x|\mathfrak{B}) = P_{\mathfrak{B}}x = \underset{\hat{x}\in L_2(\Omega,\mathcal{B},\mathbb{P};\mathbb{R}^d)}{\arg\min}\ \mathbb{E}(\mathcal{D}_\phi(x,\hat{x})) \qquad (17)$$

Note that if we took $x^* = \mathbb{E}(x)$ then the term $\mathbb{E}(\mathcal{D}_\phi(x,x^*))$ is known as the Bregman's variance

$$\mathrm{var}_\phi(x) = \mathbb{E}(\mathcal{D}_\phi(x||\mathbb{E}(x))) = \mathbb{E}(\phi(x)) - \phi(\mathbb{E}(x)) \geq 0 \qquad (18)$$

for which holds (see [1])

Theorem 3.4. *Let x be a random variable with mean $\mathbb{E}(x)$ and variance $var(x)$. The Bregman variance $var_\phi(x) \neq var(x)$ is then defined as follows*

$$\begin{aligned}var_\phi(x) &= \mathbb{E}(\mathcal{D}_\phi(x||\mathbb{E}(x)))\\ &= \mathbb{E}(\phi(x)) - \phi(\mathbb{E}(x)) \geq 0. \qquad (19)\end{aligned}$$

From inequality Eq. (14) one may further state

$$\begin{aligned}var_\phi(x) &= \mathbb{E}(\mathcal{D}_\phi(x||\mathbb{E}(x)))\\ &= \mathbb{E}(\mathcal{D}_\phi(x,\hat{x})) - \mathbb{E}(\mathcal{D}_\phi(\mathbb{E}(x),\hat{x}))\\ &\geq 0 \qquad (20)\end{aligned}$$

for any random variable $\hat{x}$. This then leads to

$$\mathbb{E}(x) = \underset{\hat{x}\in L_2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)}{\arg\min}\ \mathbb{E}(\mathcal{D}_\phi(x,\hat{x})) \qquad (21)$$

which is the same minimum point for any expected Bregman's divergence.

The key result of the previous theorems justifies using a mean as a representative of a random variable, particularly in a Bayesian estimation.

In a special case when ϕ takes the quadratic form, i.e. $\phi(x) = \frac{1}{2}\|x\|^2_{L_2}$, the Bregman's divergence in Eq. (9) modifies to the squared-Euclidean distance

$$\mathcal{D}_\phi(x||y) = \|x-y\|^2. \qquad (22)$$

In such a case the Bregman Pythagorean theorem Eq. (15) reduces to the classical Pythagorean theorem as already discussed by the author and co-workers in [17].

Following the authors previous works, the conditional expectation $E(x|y)$ of a random variable x given the measurement y in terms of Bregman's

quadratic loss functions is an orthogonal projection $P_{\mathcal{B}}(x)$ of x onto the subspace $L_2(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{R}^d)$ of all random variables consistent with the data y, i.e. generated by the sub-sigma algebra $\mathcal{B} := \sigma(y)$. This further means that x can be orthogonally decomposed into two components x_p and x_o:

$$x = x_p + x_o \tag{23}$$

in which the projected part reads $x_p := P_{\mathcal{B}}(x)$, whereas the orthogonal component x_o equals $(I - P_{\mathcal{B}})x$.

As an observation y^{mes} arrives, the first term in Eq. (23), x_p, is altered by the data y^{mes}, whereas the latter one, x_o, embodies the remaining (residuals) of the prior information x_f. This idea leads to the analogy of x_p with $\mathbb{E}(x_f|y^{mes})$ and of x_o with $x_f(\omega) - \mathbb{E}(x_f|y_f)$ in which y_f takes the form given in Eq. (5) such that

$$x_a = \mathbb{E}(x_f|y^{mes}) + (x_f - \mathbb{E}(x_f|y_f)) \tag{24}$$

holds. This is the filtering form of the decomposition given in Eq. (23), in which the indices a and f are used to denote the assimilated (posterior) state and forecast (prior) state, respectively. Following the Doob-Dynkin lemma, the previous equation can be rewritten as

$$x_a = \varphi(y^{mes}) + (x_f - \varphi(y_f)), \tag{25}$$

in which the conditional expectation $\mathbb{E}(x_f|y^{mes})$ is represented by a measurable map $\varphi(y^{mes})$, and similarly $\mathbb{E}(x_f|y_f)$ is expressed as $\varphi(y_f)$. By rearranging the terms in Eq. (25) one obtains

$$x_a = x_f + \varphi(y^{mes}) - \varphi(y_f), \tag{26}$$

the general form that is further used to construct the nonlinear filtering procedure. The advantage of Eq. (26) compared to Eq. (6) is that all quantities of consideration are given in terms of random variables, and not probability measures. Hence, it is easier to functionally approximate and computationally manipulate Eq. (26) than Eq. (6), as further discussed.

4 Optimal map

To obtain the maximal information gain in Eq. (26), the task is to find the optimal map φ among all measurable maps $\mathcal{Y} \to \mathcal{X}$. However, this step is not computationally tractable, and thus additional approximations are required. The simplest possible choice is to consider a linear approximation

$$\mathbb{E}(x_f|y_f) \approx Ky_f + b \tag{27}$$

in which the map coefficients (K, b) are obtained by minimising the orthogonal component in Eq. (24), i.e.

$$\begin{aligned} &\arg\min_{K,b} \mathbb{E}(\|x_f - \mathbb{E}(x_f|y_f)\|_2^2) \\ &= \arg\min_{K,b} \mathbb{E}(\|x_f - (Ky_f + b)\|_2^2). \end{aligned} \tag{28}$$

From the optimality condition

$$\forall \chi: \quad \mathbb{E}(\langle x_f - (Ky_f + b), \chi \rangle) = 0. \tag{29}$$

one obtains

$$\begin{aligned} &\mathbb{E}(\langle x_f - Ky_f - b, y_f \rangle) = 0 \\ &\mathbb{E}\left(x_f - Ky_f - b\right) = 0 \end{aligned} \tag{30}$$

which further results in a linear Gauss-Markov-Kalman filter equation

$$x_a(\omega) = x_f(\omega) + K(y^{mes} - y_f(\omega)), \tag{31}$$

specified by the well-known Kalman gain

$$K = C_{x_f, y_f}(C_{y_f})^{\dagger}. \tag{32}$$

Here, $\dagger$ denotes the pseudo-inverse, C_{x_f,y_f} is the covariance between the prior x_f and the observation forecast y_f, and $C_{y_f} = C_{Y(x_f)} + C_{\varepsilon}$ is the auto-covariance of y_f consisting of forecast covariance $C_{Y(x_f)}$ and the measurement covariance C_{ε}.

Even though computationally cheap, the previous formula uses only pieces of provided information in

y^{mes} and may lead to over– or under– estimation in highly nonlinear systems. Namely, the term y_f in Eq. (31) is essentially nonlinear and does not comply with the linear approximation of the map φ. To resolve nonlinearity, let the measurement operator Y be the Fréchet differentiable with Lipschitz continuous derivative $H := \partial Y/\partial x$ such that

$$Y(x) \approx Y(\check{x}) + H(x - \check{x}) =: Y_\ell(x) \tag{33}$$

holds. Following this assumption, one may further state

$$y \approx Y_\ell(x) + \varepsilon =: y_\ell(x) \tag{34}$$

in which $y_\ell(x)$ represents the linearised measurement around the point $\check{x}$. As Y_ℓ is linear, the new Gauss-Markov-Kalman formula obtains a similar form to the one given in Eq. (31) and reads

$$\begin{aligned} x_a &= x_f + K_\ell(y^{mes} - y_\ell(x_f)) \\ &= x_f + K_\ell(y^{mes} - Y(\check{x}) - H(x_f - \check{x}) - \varepsilon). \end{aligned} \tag{35}$$

Here, x_f is the forecast parameter, $y_\ell(x_f)$ is the forecast value of linearised measurement around the point $\check{x}$ given the prior x_f, ε is the model of the measurement error, and K_ℓ is the corresponding Kalman gain calculated via

$$\begin{aligned} K_\ell &= C_{x_f y_\ell}(C_{y_\ell})^\dagger \\ &= C_{x_f} H^T (H C_{x_f} H^T + C_\varepsilon)^\dagger. \end{aligned} \tag{36}$$

Note that in a special case when all distributions of consideration are known to be Gaussian, the last formula obtains a similar form to the extended Kalman filter [8, 23].

The map in Eq. (35) is not optimal as it highly depends on the choice of the point $\check{x}$. Obviously, $\check{x}$ taken as $\mathbb{E}(x_f)$ is not always the best choice. To find an optimal linearisation point, one may introduce the sequence of the first order approximants

$$Y_\ell^{(i)}(x) := Y(\check{x}^{(i)}) + H^{(i)}(x - \check{x}^{(i)}) \tag{37}$$

with

$$H^{(i)} := \frac{\partial Y}{\partial x}\Big|_{\check{x}^{(i)}}, \tag{38}$$

and

$$y_\ell^{(i)}(x) = Y_\ell^{(i)}(x) + \varepsilon. \tag{39}$$

As a result, the optimal map Y is iteratively found via the sequence of Kalman gains

$$\begin{aligned} K_\ell^{(i)} &= C_{x_f y_\ell^{(i)}} C^\dagger_{y_\ell^{(i)}} \\ &= C_{x_f}(H^{(i)})^T (H^{(i)} C_{x_f} (H^{(i)})^T + C_\varepsilon)^\dagger, \end{aligned} \tag{40}$$

and subsequently the posterior state is estimated via an iterative procedure

$$x_a^{(i+1)} = x_f + K_\ell^{(i)}(y^{mes} - y_\ell^{(i)}(x_f)), \tag{41}$$

$$\check{x}^{(i)} = \mathbb{E}(x_a^{(i+1)}), \tag{42}$$

here called the Gauss-Newton-Markov-Kalman filter. Under Gaussianity assumptions one may show that the previous equation represents the Gauss-Newton procedure for the maximum aposteriori estimate (MAP) as shown in [2]. Note that no such assumption is made here.

The convergence properties of the algorithm can be studied via fixed point theorem [12], according to which the algorithm has local convergence characterised by a spectral radius of $\rho(K_\ell^{(i)} H^{(i)})$.

5 Bayesian estimation of optimal map

In the form given in Eq. (41) the Gauss-Newton-Kalman filter has two drawbacks: first the filter requires the time consuming evaluation of the Jacobian $H^{(i)}$, and second the filter is biased as it assumes that

$$\mathbb{E}\left[(Y(x_f))^k\right] = \mathbb{E}\left[(Y(\check{x}) + H(\check{x})(x_f - \check{x}))^k\right] \tag{43}$$

holds for $k = 1, .., n$. Therefore, the straightforward linearisation is not the best possible choice. Instead,

one may search for the optimal linear map in a similar setting as given in Section 4.

In numerical practice the measurement operator $Y(x_f)$ is encoded in the corresponding computer software/simulator of the physical model, and hence is not explicitly known. But, using the classical uncertainty quantification procedures (e.g. the pseudo-spectral method or similar) one may obtain $z_f := Y(x_f)$ given x_f in a non-intrusive way. In such a case both z_f and x_f are known, and hence the estimation of the measurement operator $Y(x_f)$ in a linearised form becomes simple. It only requires an estimation of the map $\varphi_y : x_f \mapsto z_f$, i.e. the conditional expectation $\mathbb{E}(z_f|x_f)$. By taking the Bregman's squared loss function, as already discussed, the parameterised map $\varphi_y(\beta)$ can be estimated by minimising

$$\beta^* = \arg\min_{\beta} \mathbb{E}(\|z_f - \varphi_y(x_f, \beta)\|_2^2). \tag{44}$$

In a special affine case

$$\mathbb{E}(z_f|x_f) \approx \check{H}(x_f - \check{x}) + h =: \varphi_y(x_f, \beta), \tag{45}$$

with $\beta := (\check{H}, h)$ the previous optimisation problem reduces to

$$\arg\min_{\check{H},h} \mathbb{E}(\|z_f - (\check{H}(x_f - \check{x}) + h)\|_2^2), \tag{46}$$

the solution of which

$$\check{H} = C_{Y(x_f),x_f} C_{x_f}^{\dagger} \tag{47}$$

represents the approximation of the Jacobian, and

$$h := \mathbb{E}(Y(x_f)) - \mathbb{E}(x_f - \check{x}) \tag{48}$$

is the linear constant. Note that if $Y(x_f)$ is originally linear described by the true Jacobian H, then using the formula in Eq. (47) one has that

$$\check{H} = C_{Y(x_f),x_f} C_{x_f}^{\dagger} = H C_{x_f} C_{x_f}^{\dagger} \equiv H. \tag{49}$$

Similarly, for the inverse map $z_f \mapsto x_f$ holds

$$C_{x_f,z_f} C_{z_f}^{\dagger} \equiv H^{\dagger}. \tag{50}$$

Employing the previous two relations one may conclude that the Jacobian of the forward map is equal to the inverse Kalman gain when the observation $y_f = z_f + \varepsilon$ does not contain the measurement/modelling/approximation error ε.

However, note that Eq. (46) holds only if linearisation is done once as in the extended Kalman filter procedure. Otherwise, given $z_a^{(i)} := Y(x_a^{(i)})$ one solves the following problem

$$\arg\min_{\check{H}^{(i)},h^{(i)}} \mathbb{E}(\|z_a^{(i)} - (\check{H}^{(i)}(x_a^{(i)} - \check{x}^{(i)}) + h^{(i)})\|_2^2), \tag{51}$$

such that the filter in Eq. (41) obtains its unbiased form

$$x_a^{(i+1)} \;\; = \;\; x_f + K_{\ell}^{(i)}(y^{mes} - y_h^{(i)}(x_f)) \tag{52}$$

in which

$$y_h^{(i)}(x_f) : \;\; = \;\; \check{H}^{(i)}(x_f - \check{x}^{(i)}) + h^{(i)} + \varepsilon. \tag{53}$$

Note that previously we have assumed that we know random variables z_f and x_f resp. $z_a^{(i)}, x_a^{(i)}$, which is often not the case. Instead, in numerical simulations we may only know their samples. Let us denote the set of samples of the variable $x_a^{(i)}$ by $x^{sim} := (x_a^{(i)}(\omega_i)_{i=1}^N)$. Similarly, let us denote the set of forecasted samples by $z^{sim} := (z_a^{(i)}(\omega_i)_{i=1}^N)$ such that

$$z_a^{(i)}(\omega_i) = \varphi_y(x_a^{(i)}(\omega_i)) + \varepsilon_y(\omega_i) \tag{54}$$

holds. In such a case, the approximation of the Jacobian can be estimated from

$$z_a^{(i)}(\omega_i) = \check{H}^{(i)}(x_a^{(i)}\omega_i) - \check{x}^{(i)}) + h^{(i)} + \varepsilon_y(\omega_i) \tag{55}$$

in a Bayesian framework given measurement data $d^{sim} := (x^{sim}, z^{sim})$ by assuming that the pair $h, \check{H}$ and the approximation error ε_y are unknown, and

hence modelled as uncertain. In a Bayesian setting the map parameters $\beta := (\check{H}, h, \varepsilon_y)$ can be estimated as:

$$\pi_{\beta|d^{sim}}(\beta|d^{sim}) \propto \pi_{d^{sim}|\beta}(d^{sim}|\beta)\pi_\beta(\beta) \tag{56}$$

in which $\pi_\beta(\beta)$ is a joint prior distribution on β here factorised according to $\pi_\beta = \pi_{\check{H}}(\check{H})\pi_h(h)\pi_{\varepsilon_y}(\varepsilon_y)$. The prior information can be imposed further such that each element of the prior is of Gaussian type. As Eq. (55) is of linear type, the Bayesian estimation in such a case reduces to the Kalman filter estimate. For this purpose one may assume that the prior mean for the Jacobian is close to the inverse of the previously estimated Kalman gain, see Eq. (49) and Eq. (50). To include more information into the prior such as sparsity of the matrix, the prior has to be carefully designed, as discussed in Section 7.2.

Note that same type of approach can be also used for the estimation of the Kalman gain in Eq. (32). Following Eq. (54) one may pose the following problem: given samples $(x_f(\omega_i), y_h^{(i)}(x_f(\omega_i))$ estimate $\mathbb{E}(x_f|y_f) = \varphi(y_h^{(i)})$ such that

$$x_f = \mathbb{E}(x_f|y_f) + \varepsilon_x = \varphi(y_h^{(i)}) + \varepsilon_x^{(i)} \tag{57}$$

holds. Assuming linear map

$$\varphi_y(y_h^{(i)}) = K^{(i)} y_h^{(i)} + b \tag{58}$$

and given the data set $d^{sim} := (x_f(\omega_i), y_h^{(i)}(\omega_i))$ one may use Bayes's rule to estimate $\beta := (K, b, \varepsilon_x)$ in a similar manner as in Eq. (56). The numerical advantage of Bayes's rule compared to Eq. (28) lies in the prior knowledge which can be imposed on the Kalman gain, e.g. the sparsity information on the mapping coefficients as discussed in Section 7.2. This further allow us to use the previously described filter in a "hierarchical sense" for both solving the inverse problem, as well as for estimating the optimal linear map. In particular, the hierarchical approach is interesting when one would like to estimate the approximation/modelling/linearisation error ϵ as further discussed in Section 7.2. However, note that by using Bayes's rule to obtain a Kalman gain we do not satisfy the orthogonality condition, and hence we do not have a Kalman filter estimate as understood in the classical sense.

6 Predictor-corrector Bayesian-Gauss-Newton-Markov-Kalman filter for backpropagation

To estimate the initial condition of the dynamical system given in Eq. (1), one may use the previously designed filter in the following form:

$$x_{0,a}^{(i+1)} = x_{0,f} + K_\ell^{(i)}(y^{mes} - y_\ell^{(i)}(x_{0,f})), \tag{59}$$

in which $x_{0,f}$ is the a priori random variable describing the initial condition at t_0, y^{mes} is the measurement at the time T and $y_\ell := \check{H}^{(i)}(x_f - \check{x}^{(i)}) + \check{h}^{(i)} + \varepsilon$ is the forecasted linearised measurement at T and in iteration (i). In a similar manner one may also estimate any state between t_0 and T. Considering the identification of all states equidistantly separated by the update time step $\Delta\tau$, the Gauss-Newton-Markov-Kalman filter is schematically described in Alg. (1)-Alg. (2), and depicted in Fig. (1). After initialisation of the prior variable, one approximates the forward map $x_f \mapsto y_f$ by the linearised operator estimated either in a classical way, see Eq. (51), or in a Bayesian manner, see Eq. (56). Once the linearised measurement is found, one may estimate the inverse map linearly again in two different manners: by projection or by Bayes's rule. Once both maps are estimated one may assimilate the state using the measurement data, and hence update the linearisation point. This method of estimating the state will be called direct smoothing (DS) further on. A numerical example is shown in Fig. (2). Here, the smoothing algorithm with a window size of two days is used to estimate the second component of the Lorenz 1984 system (for model details see the Appendix) given noisy full state measurement data, see Alg. (2). Clearly, the linear

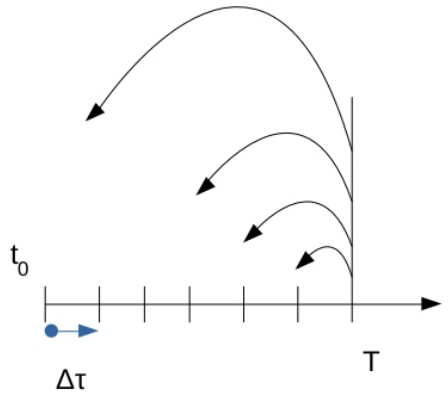


Figure 1: Schematic representation of direct backward propagation

update observed in the upper plot fails to properly estimate the state in any other time moment than the time of the measurement itself. On the other hand, the nonlinear filtering counterpart taking the iterative form as described previously produces satisfying results, see the lower plot in Fig. (2). This also holds for all three Lorenz components as depicted in Fig. (3).

In general the Gauss-Newton procedure is known to be convergent when the residuals are assumed to be small. However, if the state x_n is estimated given y_k, in which k is many times larger than n (e.g. estimation of the initial state after long time integration), and/or the system is highly nonlinear, the direct estimation can be a problem. Fig. (4) depicts an example of filter divergence when estimating the initial condition of the Lorenz 1984 system given the state measured after 96 hours. To overcome this, the large "update step", i.e. the time interval $[t_n, t_k]$, is split into smaller update steps defined by pseudo-time moments $t_n \leq \tau_\ell \leq t_k, \tau_\ell = t_n + \ell\Delta\tau$ via $\Delta\tau = c\Delta t$ stepping in which Δt is the time discretisation step, and $1 \leq c \in \mathbb{N}$. In this way one divergent Gauss-Newton iteration is substituted by several convergent ones, and the direct estimation is substituted by an incremental one.

The initial value estimation via a pseudo-time stepping Gauss-Newton procedure can be done in different ways. Here, two variants are considered: the mean-based and the random variable-based smoothing. Both start with filtering of the current state x_k given the measurement data y_k^{mes} at t_k via

$$x_{k,a}^{(i+1)} = x_{k,f} + K_k^{(i)}(y_k^{mes} - y_{kh}^{(i)}(x_{k,f})) \tag{60}$$

in which $x_{k,f}$ is the prior knowledge on the current state, and $y_{kh}^{(i)}(x_{k,f})$ is the measurement prediction. As $y_{kh}^{(i)}(x_{k,f})$ is linear in the state $x_{k,f}$, the iterative filter in Eq. (60) consists of only one iteration. Fig. (5) shows the posterior probability density function of x_a of the current state x after six days of integration given the perturbed full measurement data $x_m = x_t + \hat{\varepsilon}$ and the measurement noise with $C_\varepsilon = (0.1x_t)^2 I$.

Once converged, the a posteriori state $x_{k,a}$ is adopted as a pseudo-measurement for the preceding state $x_{k-\Delta\tau}$ at the time $t_k - \Delta\tau$. However, this could be done in at least two different ways: i) by assuming that the posterior mean is a pseudo-measurement and the posterior covariance is the measurement/modelling error describing our confidence in the "measured" value, or ii) by assuming that $x_{k,a}$ is an uncertain "perfect" measurement, see Fig. (6).

6.1 Gaussian based pseudo-measurement

Instead of evaluating the initial condition in Eq. (59) directly one may use the "smoothing" procedure in which the intermediate states are estimated before the desired one, see Fig. (6). In other words, the first unknown state x_k at the measurement time t_k is estimated via Eq. (60), whereas the preceding state $x_{k-\Delta\tau}$ at the time $t_k - \Delta\tau$ is further evaluated given the Gaussian approximation $x_{k,a}^g \sim \mathcal{N}(\bar{x}_{k,a}, C_{x_{k,a}})$ of

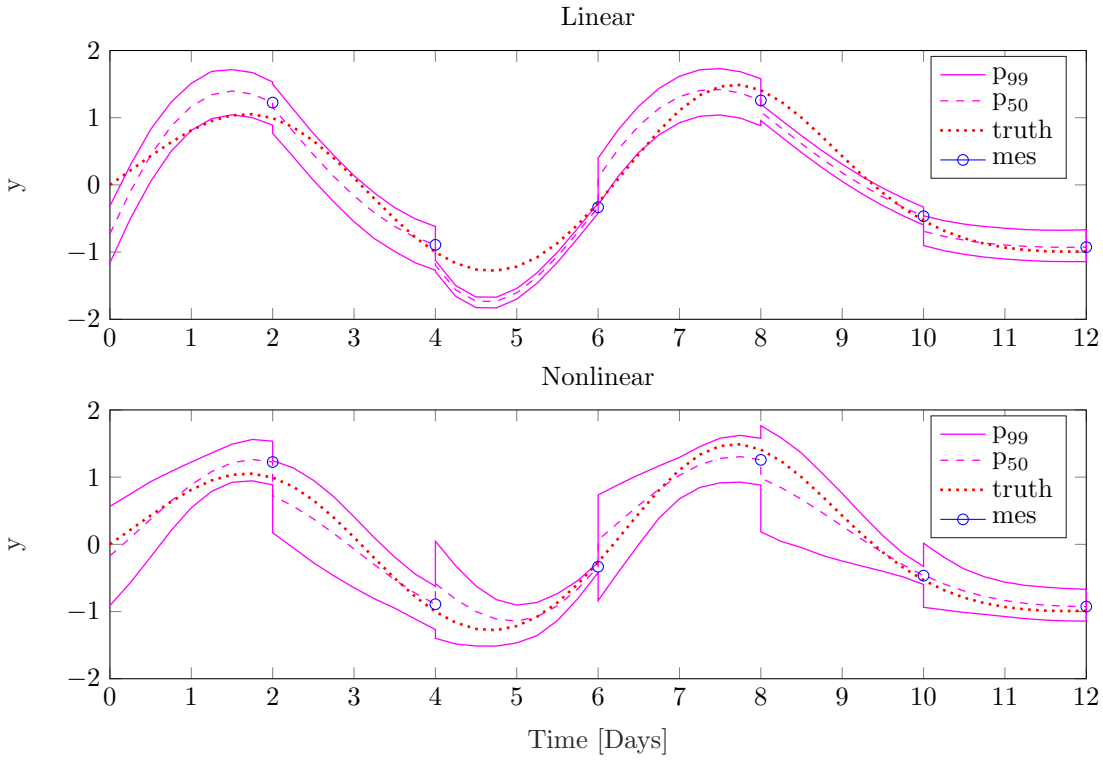


Figure 2: Linear and nonlinear smoothing of the second component of the Lorenz 1984 system

Algorithm 1 Direct smoothing (DS): Bayesian-Gauss-Newton-Markov-Kalman filter, backpropagation

function BGNMK($x_{0,f}, y^{mes}, @integ, \Delta t, \Delta\tau, t_0, T, \varepsilon$)▷ Where $x_{0,f}$ - prior on initial value, t_0 beginning of time interval, T end of time interval, y^{mes} - measurements, $integ$ - forward function (model) handle, Δt - time integration step, $\Delta\tau$ - update step, ε - measurement error

for $tt = t_0 : \Delta\tau : T$ **do** ▷ Update the assimilation time

Set prior

$x_f = integ(x_{0,f}, \Delta t, t_0, tt)$ ▷ integrate ODE system from t_0 to tt by time step Δt

Update

$x_a = \mathrm{GNMK}(x_f, y^{mes}, @integ, \Delta t, tt, T, \varepsilon)$

end for

end function

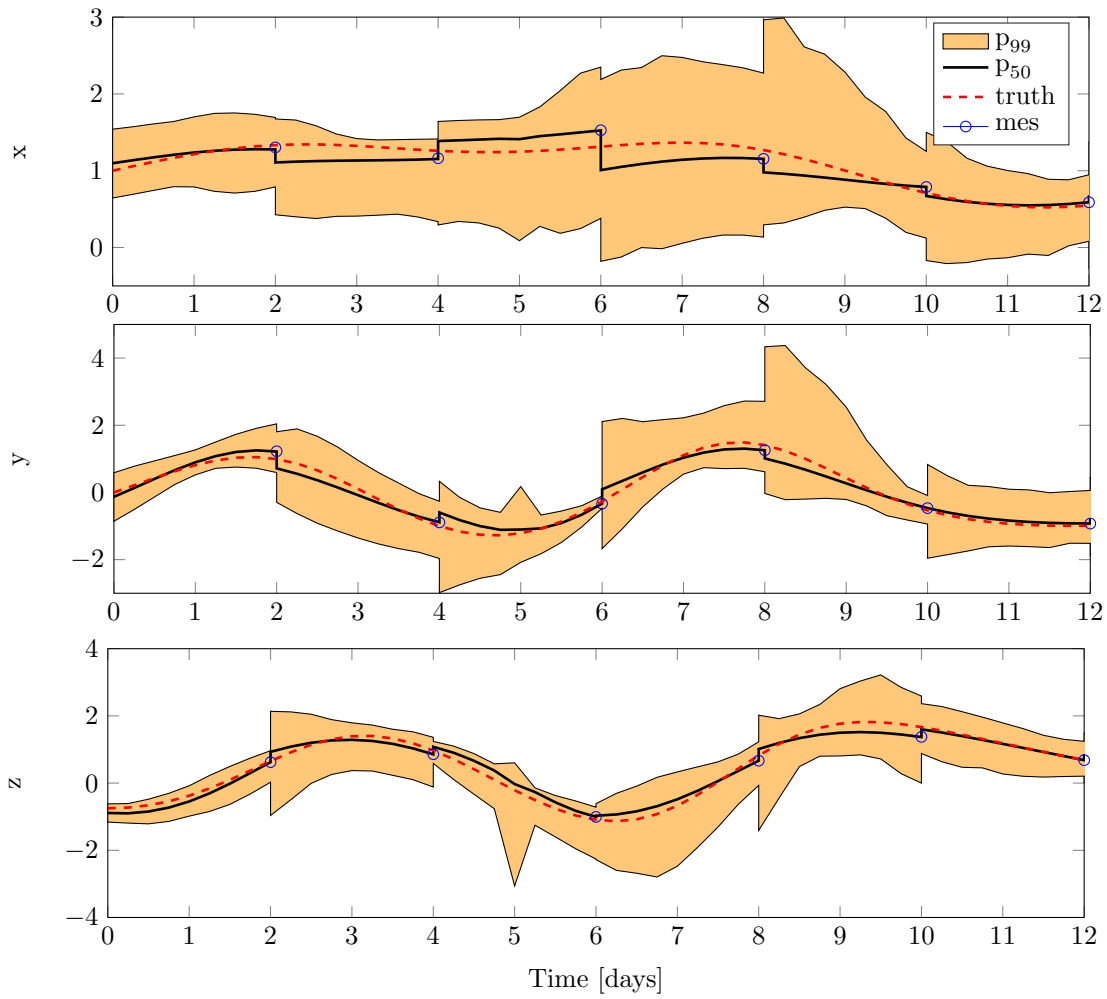


Figure 3: The backward estimation of the Lorenz 1984 state with an updating window size of two days

Algorithm 2 Direct smoothing (DS): Bayesian-Gauss-Newton-Markov-Kalman filter, backpropagation

function GNMK($x_f, y^{mes}, @integ, \Delta t, \Delta\tau, tt, T, \varepsilon$) ▷ Where $x_{0,f}$ - prior on initial value, t_0 beginning of time interval, T end of time interval, y^{mes} - measurements, *integ* - forward function (model) handle, Δt - time integration step, $\Delta\tau$ - update step, ε - measurement error

Set linearisation point

$\mathring{x}^{(0)} = \mathbb{E}(x_f), \quad x_a^{(0)} = x_f$

Set $i = 0$, err $= 2 \cdot$ tol, maxiter $= 100$

while $i \leq$ maxiter & err $\leq$ tol **do**

Predict measurement

$z_a^{(0)} = integ(x_a^{(i)}, \Delta t, tt, T)$ ▷ integrate ODE system from tt to T

Approximate forward map $x_a^{(i)} \mapsto z_a^{(i)} := Y(x_a^{(i)})$ **by**

$\varphi_y(x_a^{(i)}) = \mathring{H}^{(i)}(x_a^{(i)} - \mathring{x}^{(i)}) + \mathring{h}^{(i)}$

Estimate forward map coefficients $\beta := (\mathring{H}^{(i)}, \mathring{h}^{(i)})$ **by**

- projection:

$\mathring{H}^{(i)} = C_{z_a^{(i)}, x_a^{(i)}} C^{\dagger}_{x_a^{(i)}}, \quad \mathring{h}^{(i)} = \mathbb{E}(z_a^{(i)}) - \mathring{H}^{(i)}(x_a^{(i)} - \mathring{x}^{(i)}),$

- or by Bayes's rule

given data $d^{sim} = (x_a(\omega_j)^{(i)}, z_a(\omega_j)^{(i)}), j = 1, ..., N$ (see Section 7.2)

update $\pi_{\beta|d^{sim}}(\beta|d^{sim}) \propto \pi_{d^{sim}|\beta}(d^{sim}|\beta)\pi_\beta(\beta)$

Linearise predicted measurement

$y_\ell^{(i)}(x_f) = \mathring{H}^{(i)}(x_f - \mathring{x}^{(i)}) + \mathring{h}^{(i)} + \varepsilon$

Approximate inverse map $y_\ell^{(i)} \mapsto x_f$ **by**

$\varphi(y_\ell^{(i)}) = K^{(i)} y_\ell^{(i)} + b^{(i)}$

Estimate inverse map coefficients $w := (K^{(i)}, b^{(i)})$ **by**

- projection:

$K^{(i)} = C_{x_f, y_\ell^{(i)}} C^{\dagger}_{y_\ell^{(i)}}, \quad b = \mathbb{E}(x_f) - K^{(i)}\mathbb{E}(y_\ell^{(i)})$

- or by Bayes's rule

given data $d^{sim} = (x_f(\omega_j), y_\ell^{(i)}(\omega_{(j)}), j = 1, ..., N$ (see Section 7.2)

update $\pi_{w|d^{sim}}(w|d^{sim}) \propto \pi_{d^{sim}|w}(d^{sim}|w)\pi_w(w)$

Update state

$x_a^{(i+1)} = x_f + K^{(i)}(y^{mes} - y_\ell^{(i)}),$

Update linearisation point

$i = i + 1;$

$\mathring{x}^{(i)} = \mathbb{E}(x_a^{(i)})$

Convergence criterion ▷ e.g. mean based

err $= \|\mathbb{E}(x_a^{(i)}) - \mathbb{E}(x_a^{(i-1)})\| \cdot \|\mathbb{E}(x_a^{(i-1)})\|^{-1}$

end while

end function

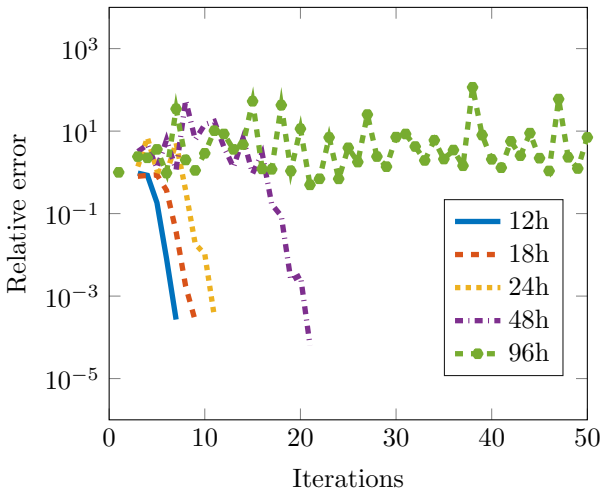


Figure 4: Convergence of posterior estimate of the initial condition $\boldsymbol{x}_0$ w.r.t. time at which the measurement data arrive

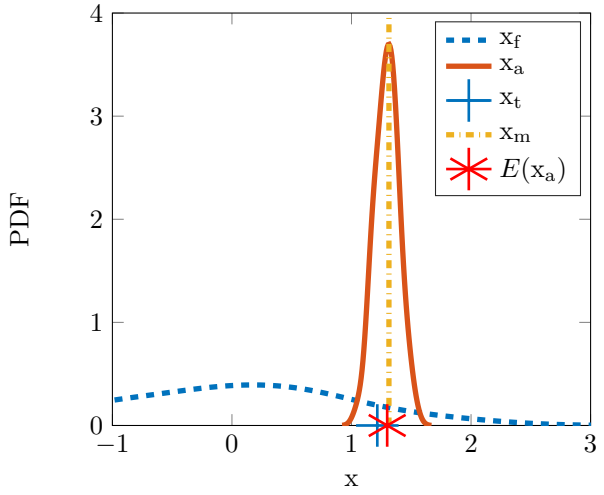


Figure 5: Update of the current state x at $t = 6$ days

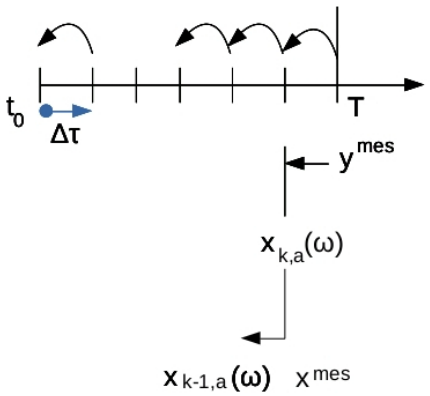


Figure 6: The schematic representation of pseudo-backward propagation

the convergent $x_{k,a}$ such that

$$x^{(i+1)}_{k-\Delta\tau,a} = x_{k-\Delta\tau,f} + K^{(i)}_{k-\Delta\tau}(x^g_{k,a} - y^{(i)}_{kh}(x_{k-\Delta\tau,f})) \tag{61}$$

holds. Here, $x_{k-\Delta\tau,f}$ is the apriori assumption on the state at the time $t_{k-\Delta\tau}$, $y^{(i)}_{kh}(x_{k-\Delta\tau,f})$ is the linearised measurement operator (i.e. the linearised forward map $x_{k-\Delta\tau,f} \mapsto x_{k,f}$) around the point $\mathring{x}^{(i)}$ in iteration (i):

$$y^{(i)}_{kh}(x_{k-\Delta\tau,f}) = \mathring{H}^{(i)}(x_{k-\Delta\tau,f} - \mathring{x}^{(i)}) + \mathring{h}^{(i)}. \tag{62}$$

The map coefficients $\mathring{H}^{(i)}, \mathring{h}^{(i)}$ are estimated either by the projection algorithm or by Bayesian update similar to those depicted in Alg. (1)-Alg. (2), whereas the linearisation point is chosen as

$$\mathring{x}^{(i+1)} = \mathbb{E}(x^{(i+1)}_{k-\Delta\tau,a}). \tag{63}$$

Decoupling $x^g_{k,a}$ into the mean $\bar{x}_{k,a}$ and perturbation $\varepsilon_{k,f} \sim \mathcal{N}(0, C_{x_{k,a}})$ parts, one may rewrite Eq. (61) to

$$x^{(i+1)}_{k-\Delta\tau,a} = x_{k-\Delta\tau,f} + K^{(i)}_{k-\Delta\tau}(\bar{x}_{k,a} - (y^{(i)}_{kh} + \varepsilon_{k,f})), \tag{64}$$

thanks to the symmetry of the Gaussian distribution representing $\varepsilon_{k,f}$. In this manner Eq. (64) can be understood as the state estimation given deterministic measurement $\bar{x}_{k,a}$ at the time t_k. Hence, the algorithm of pseudo-time stepping is only a slight extension of the one presented in Alg. (1). The new procedure requires estimation of the current state, after which the original filter is called, see Alg. (3).

Rewriting Eqs. (60)-(65) for all preceding states, one obtains the general form of a smoothing iterative filter:

$$x^{(i+1)}_{\ell-1,a} = x_{\ell-1,f} + K^{(i)}_{\ell}(\bar{x}_{\ell,a} - (y^{(i)}_{\ell,h}(x_{\ell,f}) + \varepsilon_{\ell})), \quad (65)$$

for all $\ell = k, k-\Delta\tau, .., k-n\Delta\tau$. The last formula further can be generalised by taking into account all estimated states from the time moment t_k to the current time t_n as measurements, similarly to the classical smoothing algorithm.

Unfortunately, the estimate in Eq. (65) is biased due to nonlinearity of the time-dependent problem. If not corrected, the bias becomes propagated through the model with each new update as shown in Fig. (7) on the example of the first Lorenz 1984 component. The mean value deteriorates from the measured one with each update such that the deviation becomes larger with the reduction of the update step size $\Delta\tau$ in contrast to expectations.

The posterior $x_{k,a}$ in Eq. (65) has the mean $\bar{x}_{k,a}$ that differs from the true posterior mean $\bar{x}^{true}_{k,a}$ according to the error

$$\epsilon_k = \bar{x}^{true}_{k,a} - \bar{x}_{k,a}, \quad (66)$$

which further becomes propagated in time with the state integration/assimilation. Hence, Eq. (64) (and similarly Eq. (65)) have to be corrected for the amount given in Eq. (66).

The correction scheme is schematically depicted in Fig. (8) and is of the predictor-corrector type. The predictor phase starts with

- the prior assumption on the state $x_{k-\Delta\tau,f}$ at the time $t_{k-\Delta\tau}$ with $\Delta\tau$ being the backpropagation increment.
- The state $x_{k-\Delta\tau,f}$ is integrated forward ($\mathcal{I}_{\Delta\tau}$ in Fig. (8) denotes the integration operator over time interval $\Delta\tau$ from $t_{k-\Delta\tau}$ to t_k) to obtain the current prior state $x_{k,f}$ at the time t_k.
- The current state $x_{k,f}$ is further assimilated with the measurement data x^{mes}_k in a linear direct GMK manner (in Fig. (8) denoted by $\mathcal{U}_L$) to obtain the posterior $x_{k,a}$. x^{mes}_k may represent the real data only for the state that is being measured, otherwise these are pseudo-measurement data. For example, if we update in the time interval $[t_0, T]$ given measurement data at the time T, then the measurement x^{mes}_k at $t_k = T$ is the real measurement y^{mes}. Otherwise, if $t_k < T$ our measurement at t_k is the posterior estimate obtained by incremental backpropagation of the posterior at $t_k + \Delta\tau$.
- The assimilated current state $x_{k,a}$ is then used as a pseudo-measurement for the assimilation of $x_{k-\Delta t,f}$ state via iterative GMK (see Eq. (64)), in Fig. (8) denoted by backward update operator $\mathcal{U}_{-\Delta\tau}$.

With this the corrector phase starts by

- integrating forward the estimate $x_{k-\Delta\tau,a}$ via $\mathcal{I}_{\Delta\tau}$ to obtain the prior on the current state $x^a_{k,f}$ at t_k given posterior $x_{k-\Delta\tau,a}$ at $t_{k-\Delta\tau}$.
- Furthermore, the newly obtained estimate $x^f_{k,a}$ is used as a prior for a second turn of updating the current state at t_k given measurement x^{mes}_k. The update is performed using linear direct GMK rule to obtain $x^a_{k,a}$.
- The difference between the prior $x^f_{k,a}$ and posterior $x^a_{k,a}$ estimates then defines the correction

Algorithm 3 Pseudo-smoothing I (PS): incremental BGNMK (iBGNMK) with Gaussian approximation

function IGNMK($x_{0,f}, y^{mes}, @integ, \Delta t, \Delta\tau, t_0, T, \varepsilon$) ▷ Where $x_{0,f}$ - prior on initial value, t_0 beginning of time interval, T end of updating interval, y^{mes} - measurement at T, $integ$ - forward function handle, Δt - time integration step, $\Delta\tau$ - update step, ε - measurement error

  Predict current state at T

    $x_f = integ(x_{0,f}, \Delta t, t_0, T)$ ▷ integrate ODE system from t_0 to T by time step Δt

  Predict measurement

    $y_f = I_x(x_f) + \varepsilon$ ▷ I_x is the indicator operator in case $\dim(x_f) > \dim(y^{mes})$

  Update current state at T **given** y^{mes}

    $x_a = x_f + C_{x_f,y_f} C^{-1}_{y_f,y_f}(y^{mes} - y_f)$

  for $tt = T - \Delta\tau : -\Delta\tau : t_0$ **do**

    Set pseudo-measurement

      $x_a^g = \text{Gaussian}(x_a)$

    Decompose pseudo-measurement:

      to the mean value $\bar{x}_a = \mathbb{E}(x_a^g)$

      and the fluctuation term $\varepsilon_f := x_a^g - \bar{x}_a$

    Set preceding prior

      $x_f = integ(x_{0,f}, \Delta t, t_0, tt)$ ▷ integrate ODE system from t_0 to tt by time step Δt

    Update preceding state

      $x_a = GNMK(x_f, \bar{x}_a, @integ, \Delta t, \Delta\tau, tt, tt + \Delta\tau, \mathring{x}, \varepsilon_f)$

  end for

end function

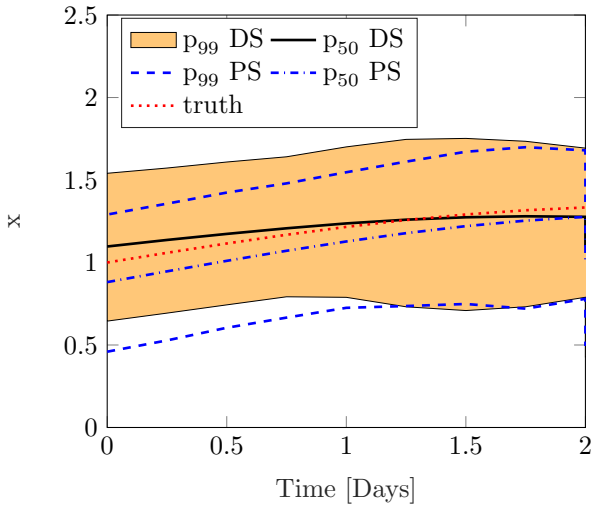


Figure 7: Bias propagation over time for the first Lorenz 1984 component. DS is the direct simulation estimate given in Eq. (59) and PS is the pseudo-estimate given in Eq. (64) with $\Delta\tau = 6h$. p_n denotes $n\%$ quantile.

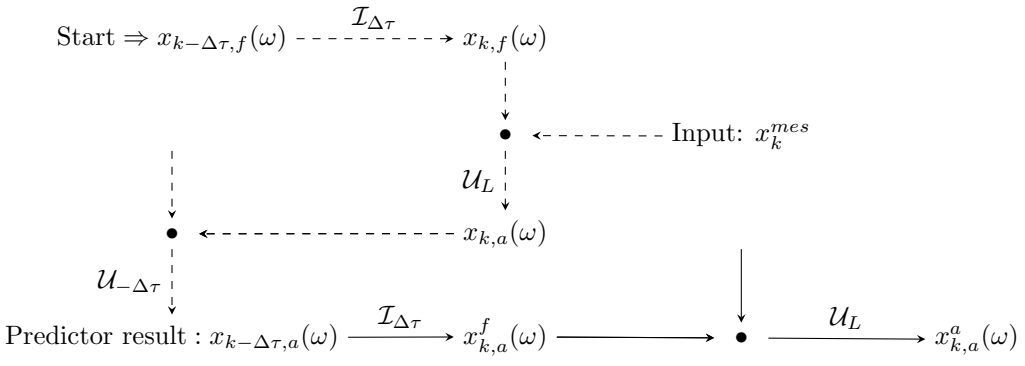


Figure 8: The scheme of bias correction

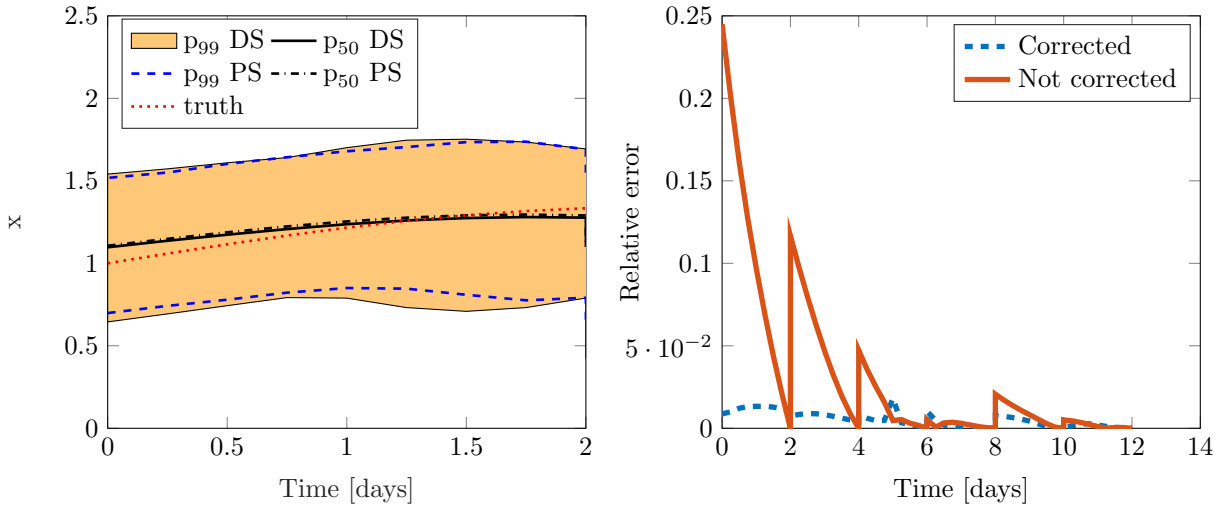


Figure 9: The bias propagation in time (left) and the bias correction in time (right)

error. This is the corrector phase. The process is further repeated for $x_{k-2\Delta\tau,f}$ given the measurement $x^{mes}_{k-\Delta\tau}$ adopted as the corrected version of $x_{k-\Delta\tau,a}$.

To estimate the correction error, let the converged posterior estimate in Eq. (64) be denoted by $x_{k-\Delta\tau,a}$ (beginning of the corrector phase in Fig. (8)) such that

$$x_{k-\Delta\tau,a} = x^{true}_{k-\Delta\tau,a} + e_{k-\Delta\tau} \tag{67}$$

holds, in which $e_{k-\Delta\tau}$ denotes the bias error at the time $t_{k-\Delta\tau}$. Propagating the a posteriori estimate $x_{k-\Delta\tau,a}$ by time step $\Delta\tau$ forward [1], one obtains the forecast estimate $x^f_{k,a}$ at t_k such that

$$\begin{aligned} x^f_{k,a} &= \mathring{H}_k(x_{k-\Delta\tau,a} - \mathring{x}_k) + \mathring{h}_k \qquad (68)\\ &= \mathring{H}_k(x^{true}_{k-\Delta\tau,a} + e_{k-\Delta\tau} - \mathring{x}_k) + \mathring{h}_k \\ &= \mathring{H}_k(x^{true}_{k-\Delta\tau,a} - \mathring{x}_k) + \mathring{h}_k + \mathring{H}_k e_{k-\Delta\tau} \\ &= x^{f,true}_{k,a} + \mathring{H}_k e_{k-\Delta\tau} \qquad (69) \end{aligned}$$

[1] this may include several time discretisation steps Δt

holds. Here, $\mathring{H}_k$ and $\mathring{h}_k$ are converged parameters of the forward map, and $x^{f,true}_{k,a}$ denotes the forecast of the exact a posteriori estimate. The analysis step at time moment t_k is then given by

$$\begin{aligned} x^a_{k,a} &= x^f_{k,a} + K(x^{mes}_k - x^f_{k,a} - \varepsilon_{k,f}) \qquad (70)\\ &= x^{f,true}_{k,a} + \mathring{H}_k e_{k-\Delta\tau} + \\ &\quad K(x^{mes}_k - x^{f,true}_{k,a} - \mathring{H}_k e_{k-\Delta\tau} - \varepsilon_{k,f}) \\ &= x^{f,true}_{k,a} + K(x^{mes}_k - x^{f,true}_{k,a}) \\ &\quad + \mathring{H}_k(I-K)e_{k-\Delta\tau} - K\varepsilon_{k,f} \\ &= x^{a,true}_{k,a} + \mathring{H}_k(I-K)e_{k-\Delta\tau} - K\varepsilon_{k,f}. \end{aligned}$$

Here, $x^{a,true}_{k,a}$ is the assimilated value of $x^{f,true}_{k,a}$. By subtracting the previous two equations

$$\begin{aligned} x^f_{k,a} - x^a_{k,a} &= x^{f,true}_{k,a} - x^{a,true}_{k,a} + \qquad (71)\\ &\quad \mathring{H}_k e_{k-\Delta\tau} - \mathring{H}_k(I-K)e_{k-\Delta\tau} \\ &\quad + K\varepsilon_{k,f} \end{aligned}$$

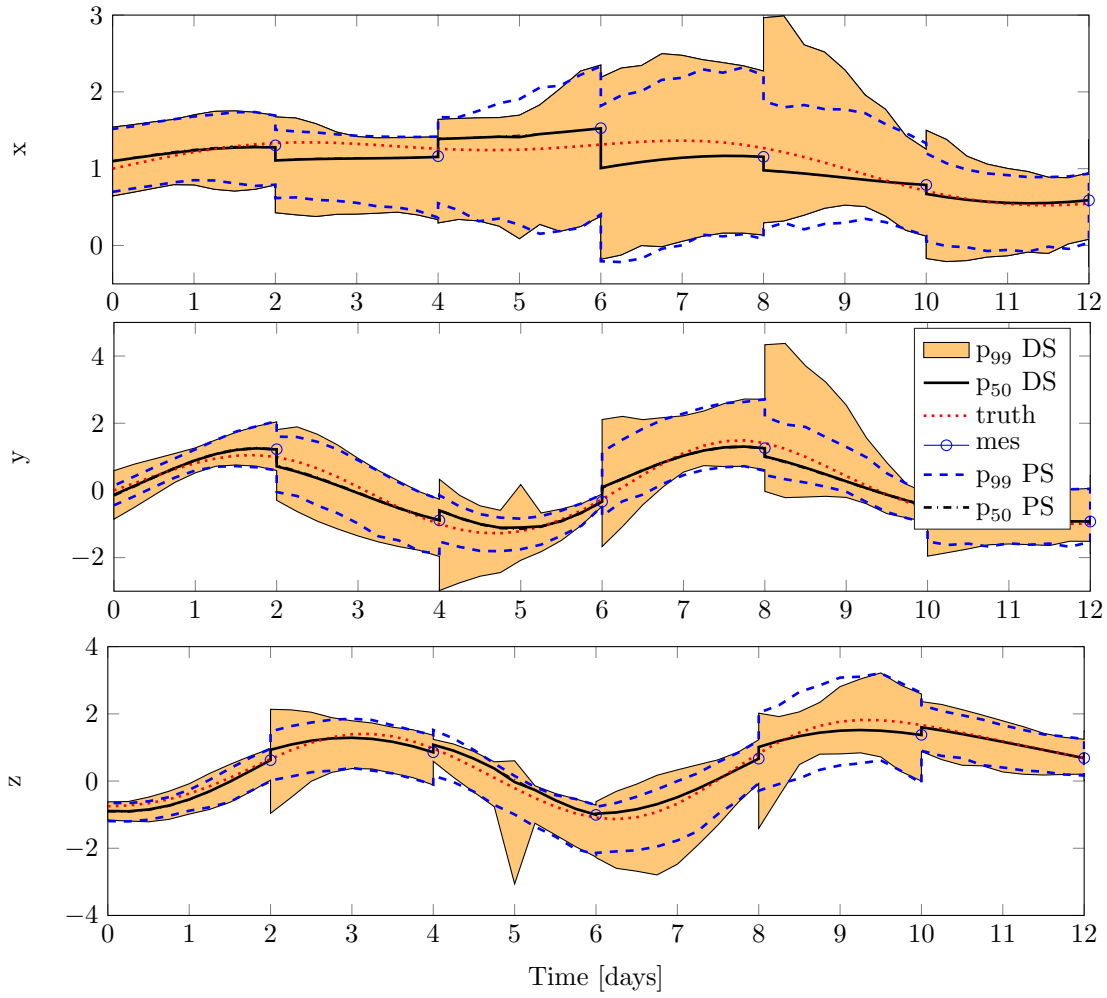


Figure 10: The mean based corrected estimate of the Lorenz 1984 state every two days backwards. The prior is obtained starting from the initial condition and the measurement has the coefficient of the variance equal to 10%.

and taking the mathematical expectation one obtains

$$\mathbb{E}(x^f_{k,a} - x^a_{k,a}) = \mathbb{E}(x^{f,true}_{k,a} - x^{a,true}_{k,a}) + \mathring{H}_k \bar{e}_{k-\Delta\tau} - \mathring{H}_k(I-K)\bar{e}_{k-\Delta\tau}. \quad (72)$$

Furthermore,

$$\mathbb{E}(x^f_{k,a} - x^a_{k,a}) = \mathbb{E}(x^{f,true}_{k,a} - x^{true}_{k,a}) + \mathbb{E}(x^{true}_{k,a} - x^{a,true}_{k,a}) + \mathring{H}_k K \bar{e}_{k-\Delta\tau} \quad (73)$$

in which

$$\mathbb{E}(x^{f,true}_{k,a} - x^{true}_{k,a}) \stackrel{!}{=} 0 \quad (74)$$
$$\mathbb{E}(x^{a,true}_{k,a} - x^{true}_{k,a}) \stackrel{!}{=} 0$$

due to unbiased requirement. This further gives

$$\mathbb{E}(x^f_{k,a} - x^a_{k,a}) = \mathring{H}_k K \bar{e}_{k-\Delta\tau}.$$

Hence, the mean bias error for the assimilated state $x_{k-\Delta\tau,a}$ at the end of the predictor phase reads:

$$\bar{e}_{k-\Delta\tau} = (\mathring{H}_k K)^{-1}\mathbb{E}(x^f_{k,a} - x^a_{k,a}). \quad (75)$$

In a similar manner one may correct the variance of the posterior by considering the second moment in Eq. (72).

Introducing the estimated error in Eq. (75) to the update in Eq. (64) one obtains the unbiased solution as shown in Fig. (9a) for the update of the first Lorenz 1984 component. The total correction over a period of 12 days is shown in Fig. (9b), in which are depicted the relative errors of the biased and unbiased pseudo-estimated states compared to the direct estimated state following Eq. (59). As one may notice the error is decreasing for several orders of magnitudes when the correction is introduced.

6.2 Random-variable based pseudo-measurement

The previous estimation did not take into consideration the full uncertainty in the pseudo-measurement. Hence, the estimate does not have correct variance as only the Gaussian approximation of the measurement is considered. This can be seen in Fig. (10) in which the corrected pseudo-estimate is compared to the direct one.

However, by taking the current aposteriori estimate $x_{k,a}$ at the time t_k— obtained by assimilating the measurement data y^{mes} at t_k via linear GMK filter— as uncertain non-Gaussian pseudo-measurement, the estimation of the preceding state $x_{k-\Delta\tau}$ in a back-propagation manner ($t_k \to t_{k-\Delta\tau}$) becomes stochastic as the measurement is a random variable. Following this, one may further state

$$x^{(i+1)}_{k-\Delta\tau,a} = x_{k-\Delta\tau,f} + K^{(i)}_{k-\Delta\tau}(x_{k,a} - y^{(i)}_{kh}(x_{k-\Delta\tau,f})), \quad (76)$$

similarly to Eq. (64). However, in contrast to Eq. (64) the pseudo-measurement $x_{k,a}$ is taken in its full form, and not only as a Gaussian approximation. This further means that $y^{(i)}_{kh}(x_{k-\Delta\tau,f})$ is a "perfect" linearised version of the time-discretised model in Eq. (1) around point $\mathring{x}^{(i)}_k$

$$y^{(i)}_{kh}(x_{k-\Delta\tau,f}) = \mathring{H}^{(i)}_k(x_{k-\Delta\tau,f} - \mathring{x}^{(i)}_k) + \mathring{h}^{(i)}_k + \epsilon_k, \quad (77)$$

and similarly $K^{(i)}_{k-\Delta\tau}$ is the "perfect" Kalman gain given as

$$K^{(i)}_{k-\Delta\tau} = C_{x_{k-\Delta\tau,f}, y^{(i)}_{kh}} C^{\dagger}_{y^{(i)}_{kh}}. \quad (78)$$

Notice that ϵ_k represents the modelling/ discretisation error, the estimate of which is further described in Section 7.2.

The posterior estimate in Eq. (76) has different second order statistics than those specified by the "classical" Kalman filter in the previous section. To simplify the notation let $x_f := x_{k-\Delta\tau,f}$, $x_a := x_{k-\Delta\tau,a}$, $y_f := y^{(i)}_{kh}(x_{k-\Delta\tau,f})$ and $y_m := x_{k,a}$, then the mean

value of the converged posterior reads

$$\bar{x}_a = \bar{x}_f + K_{k-\Delta\tau}(\bar{y}_m - \bar{y}_f), \tag{79}$$

whereas the covariance follows from

$$\begin{aligned} C_{x_a} = & C_{x_f} + K_{k-\Delta\tau} C_{y_m} K_{k-\Delta\tau}^T \\ & + K_{k-\Delta\tau} C_{y_f} K_{k-\Delta\tau}^T \\ & - 2K_{k-\Delta\tau} C_{y_f,y_m} K_{k-\Delta\tau}^T \\ & - C_{x_f,y_f} K_{k-\Delta\tau}^T \\ & - K_{k-\Delta\tau} C_{x_f,y_f}^T \\ & + K_{k-\Delta\tau} C_{x_f,y_m} \\ & + C_{x_f,y_m} K_{k-\Delta\tau}^T. \end{aligned} \tag{80}$$

In the previous equations the index $^{(i)}$ is avoided, as the last two equations are written for $i = i_{conv}$ in which i_{conv} is the number of iterations of the converged estimate. In Eq. (80) note that

$$\begin{aligned} & -2K_{k-\Delta\tau} C_{y_f,y_m} K_{k-\Delta\tau}^T \\ & = -2K_{k-\Delta\tau} \mathring{H}_k^{(i)} C_{x_f,y_m} K_{k-\Delta\tau}^T \\ & = -2C_{x_f,y_m} K_{k-\Delta\tau}^T \end{aligned} \tag{81}$$

as the Kalman gain is optimal [2], i.e. $K_{k-\Delta\tau}\mathring{H}_k^{(i)} = I$. Having that $K_{k-\Delta\tau} = C_{x_f,y_f} C_{y_f}^\dagger$ and after substituting the last equation in Eq. (80) one obtains

$$\begin{aligned} C_{x_{k-\Delta\tau,a}} = & C_{x_f} + C_{x_f,y_f} C_{y_f}^\dagger C_{y_m} (C_{y_f}^\dagger)^T C_{x_f,y_f}^T \\ & + C_{x_f,y_f} C_{y_f}^\dagger C_{y_f} (C_{y_f}^\dagger)^T C_{x_f,y_f}^T \\ & - C_{x_f,y_f} (C_{y_f}^\dagger)^T C_{x_f,y_f}^T \\ & - C_{x_f,y_f} C_{y_f}^\dagger C_{x_f,y_f}^T \\ = & C_{x_f} + C_{x_f,y_f} C_{y_f}^\dagger \\ & (C_{y_m} - C_{y_f}) (C_{y_f}^\dagger)^T C_{x_f,y_f}^T \end{aligned} \tag{82}$$

which in the original notation reads

$$\begin{aligned} C_{x_{k-\Delta\tau,a}} = & C_{x_{k-\Delta\tau,f}} + C_{x_{k-\Delta\tau,f}} \mathring{H}^T C_{y_{kh}^{(i)}}^\dagger \\ & (C_{x_{k,a}} - C_{y_{kh}^{(i)}}) C_{y_{kh}^{(i)}}^\dagger \mathring{H}^T C_{x_{k-\Delta\tau,f}}{}^T \end{aligned} \tag{83}$$

Using the estimation in Eq. (76) one obtains the correct estimate of the posterior variance as obtained by the direct simulation, see Fig. (11) for the comparison of the update obtained by direct iteration (DS) and the pseudo (PS) one. Note that the pseudo-updating is here performed every 6 hours. The same estimate is also depicted earlier in Fig. (2), in which the iterative pseudo-estimation is compared to the linear pseudo-estimation every 6 hours. The pseudo-nonlinear posterior estimate converges faster than the direct one, see Fig. (12) for comparison of the number of iterations neccessary to achieve the relative error in the posterior mean of magnitude 1e-3. Usually the posterior converges very fastly after two or three iterations up to tolerance on the first decimal. However, this number raises up to ten iterations if the accuracy is up to 1e-3 in all three components. On the other hand, a direct iteration of the initial condition requires up to 50 iterations for the same accuracy. This behaviour also depends on the discretisation of the previously described filters which will be discussed later.

The random variable updating does not introduce bias into the estimation, and hence the bias correction introduced earlier does not change much the posterior estimate, see Fig. (13). A small difference between the corrected and original estimates exists due to numerical integration of discretisation errors.

7 Iterative polynomial chaos filter

The advantage of the filtering approach as presented in Eq. (52), Eq. (64) and Eq. (76) compared to the other Bayesian numerical procedures lies in the simplicity of the posterior variable estimation. Once the random variables appearing in Eq. (52) are approximated using the standard Galerkin functional approximation tools in their minimal form, the filtering procedure reduces to the purely algebraic method for

[2] In numerical computations $K_{k-\Delta\tau}\mathring{H}_k^{(i)} \approx I$

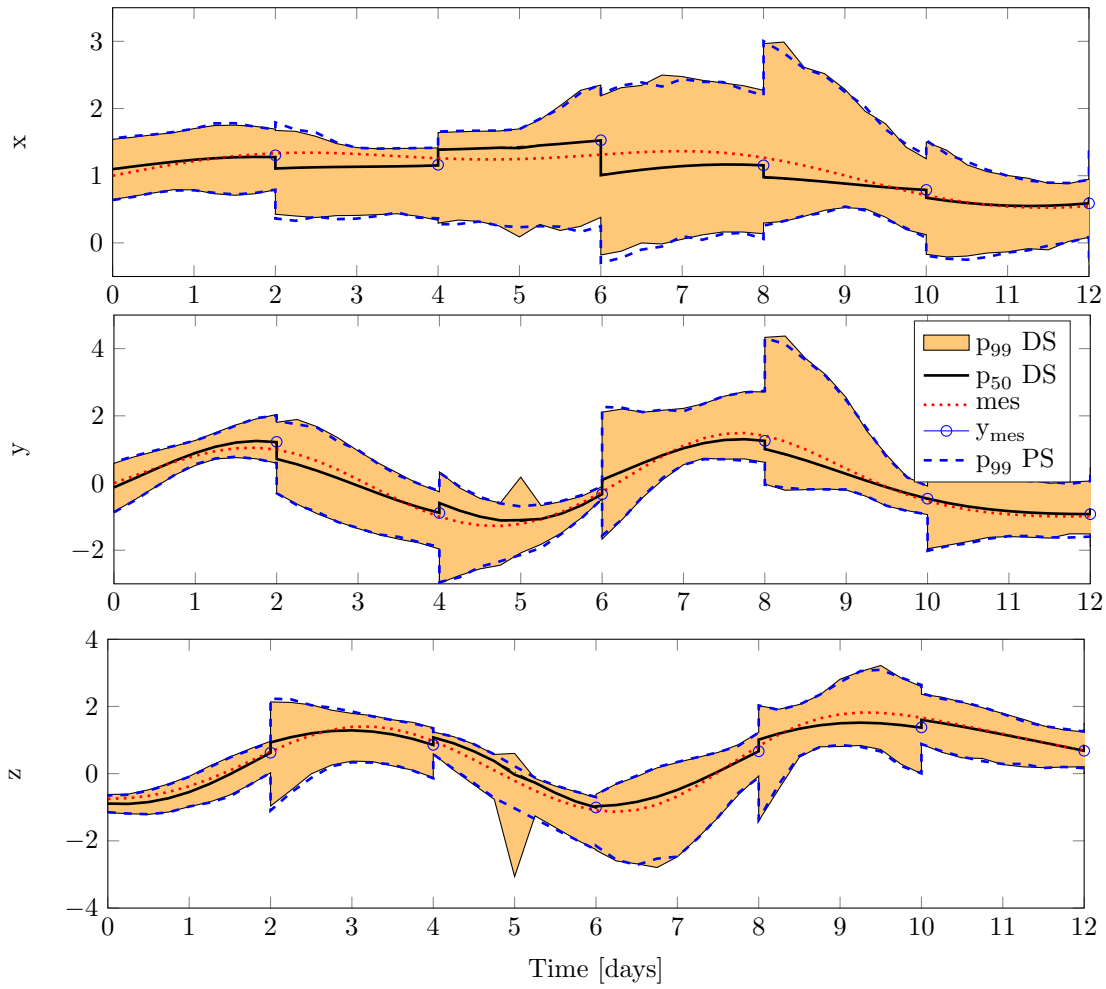


Figure 11: The pseudo-estimation of the Lorenz 1984 state every two days backwards using the random variable algorithm. The pseudo-update is made every 6 hours.

Algorithm 4 Pseudo-smoothing II (PS): incremental RV-based GNMK (irvGNMK)

function IRVGNMK($x_{0,f}, y^{mes}, @integ, \Delta t, \Delta\tau, t_0, T, \mathring{x}, \varepsilon$) ▷ Where $x_{0,f}$ - prior on initial value, t_0 beginning of time interval, T end of updating interval, y^{mes} - measurement at T, *integ* - forward function handle, Δt - time integration step, $\Delta\tau$ - update step, ε - measurement error
  Predict current state at T
    $x_f = integ(x_{0,f}, \Delta t, t_0, T)$ ▷ integrate ODE system from t_0 to T by time step Δt
  Predict measurement
    $y_f = I_x(x_f) + \varepsilon$ ▷ I_x is the indicator operator in case $\dim(x_f) > \dim(y^{mes})$
  Update current state at T **given** y^{mes}
    $x_a = x_f + C_{x_f,y_f} C^{-1}_{y_f,y_f}(y^{mes} - y_f)$
  for $tt = T - \Delta\tau : -\Delta\tau : t_0$ **do**
    Set pseudo-measurement
      $x_a^m = x_a$
    Set preceding prior
      $x_f = integ(x_{0,f}, \Delta t, t_0, T)$ ▷ integrate ODE system from t_0 to tt by time step Δt
    Update preceding state
      $x_a = orvGNMK(x_f, x_a^m, @integ, \Delta t, \Delta\tau, tt, T, \varepsilon)$
  end for
end function

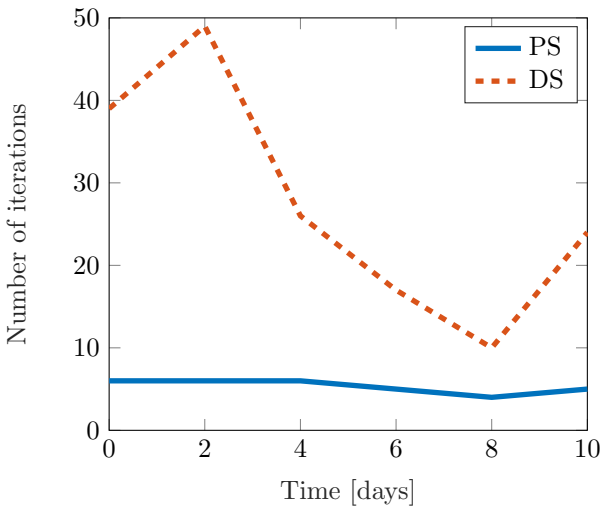


Figure 12: Number of iterations neccessary to achieve posterior accuracy of 1e-3 in the mean

estimating the posterior variable. However, in high-dimensional problems, or when using commercial softwares, sometimes it is not possible to use spectral, but pseudo-spectral approximations. Therefore, here the focus is put on the discretisation of random variables in a data-driven sparse functional approximation form. In this light the optimal approximations of the state variable, their numerical evaluations using the minimal number of sample points, as well as an efficient estimation of forward and inverse maps, i.e. the Jacobian of linearised forward maps, as well as Kalman gain are discussed here.

7.1 Random variable discretisations

For the purpose of discretisation, the random variables appearing in Eq. (52) can be expressed in terms of some known simpler kind of random variables, as previously studied by the author and colleagues in a purely linear setting, see [21]. This can be achieved

Algorithm 5 Pseudo-smoothing (PS) II: incremental BGNMK filter, backpropagation

function ORVGNMK($x_f, x_a^m, @integ, \Delta t, \Delta\tau, tt, T, \varepsilon$) ▷ Where $x_{0,f}$ - prior on initial value, t_0 beginning of time interval, T end of time interval, y^{mes} - measurements, *integ* - forward function (model) handle, Δt - time integration step, $\Delta\tau$ - update step, ε - measurement error

Set linearisation point

$\mathring{x}^{(0)} = \mathbb{E}(x_f), \quad x_a^{(0)} = x_f$

Set $i = 0$, err $= 2 \cdot$ tol, maxiter $= 100$

while $i \leq$ maxiter & err $\leq$ tol **do**

Predict measurement

$z_a^{(0)} = integ(x_a^{(i)}, \Delta t, tt, T)$ ▷ integrate ODE system from tt to T

Approximate forward map $x_a^{(i)} \mapsto z_a^{(i)} := Y(x_a^{(i)})$ **by**

$\varphi_y(x_a^{(i)}) = \mathring{H}^{(i)}(x_a^{(i)} - \mathring{x}^{(i)}) + \mathring{h}^{(i)}$

Estimate forward map coefficients $\beta := (\mathring{H}^{(i)}, \mathring{h}^{(i)})$ **by**

- projection:

$\mathring{H}^{(i)} = C_{z_a^{(i)}, x_a^{(i)}} C^{\dagger}_{x_a^{(i)}}, \quad \mathring{h}^{(i)} = \mathbb{E}(z_a^{(i)}) - \mathring{H}^{(i)}(x_a^{(i)} - \mathring{x}^{(i)}),$

- or by Bayes's rule

given data $d^{sim} = (x_a(\omega_j)^{(i)}, z_a(\omega_j)^{(i)}), j = 1, ..., N$ (see Section 7.2)

update $\pi_{\beta|d^{sim}}(\beta|d^{sim}) \propto \pi_{d^{sim}|\beta}(d^{sim}|\beta)\pi_\beta(\beta)$

Linearise predicted measurement

$y_\ell^{(i)}(x_f) = \mathring{H}^{(i)}(x_f - \mathring{x}^{(i)}) + \mathring{h}^{(i)} + \varepsilon$

Approximate inverse map $y_\ell^{(i)} \mapsto x_f$ **by**

$\varphi(y_\ell^{(i)}) = K^{(i)} y_\ell^{(i)} + b^{(i)}$

Estimate inverse map coefficients $w := (K^{(i)}, b^{(i)})$ **by**

- projection:

$K^{(i)} = C_{x_f, y_\ell^{(i)}} C^{\dagger}_{y_\ell^{(i)}}, \quad b = \mathbb{E}(x_f) - K^{(i)}\mathbb{E}(y_\ell^{(i)})$

- or by Bayes's rule

given data $d^{sim} = (x_f(\omega_j), y_\ell^{(i)}(\omega_{(j)}), j = 1, ..., N$ (see Section 7.2)

update $\pi_{w|d^{sim}}(w|d^{sim}) \propto \pi_{d^{sim}|w}(d^{sim}|w)\pi_w(w)$

Update state

$x_a^{(i+1)} = x_f + K^{(i)}(x_a^m - y_\ell^{(i)}),$

Update linearisation point

$i = i + 1;$

$\mathring{x}^{(i)} = \mathbb{E}(x_a^{(i)})$

Convergence criterion ▷ e.g. mean based

err $= \|\mathbb{E}(x_a^{(i)}) - \mathbb{E}(x_a^{(i-1)})\| \cdot \|\mathbb{E}(x_a^{(i-1)})\|^{-1}$

end while

end function

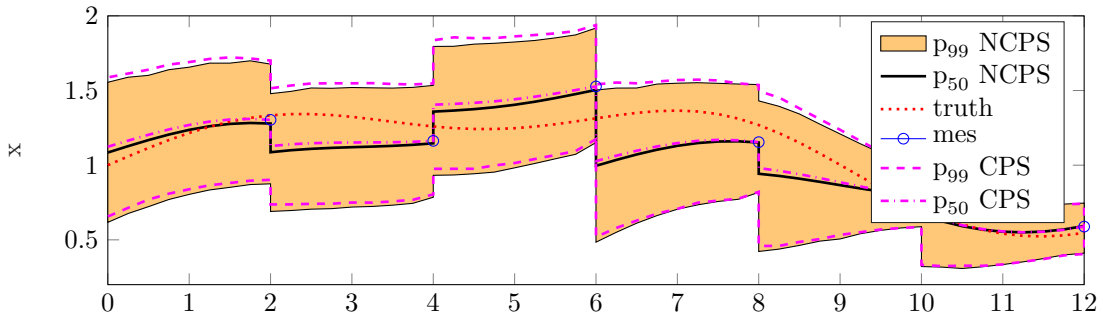


Figure 13: The random variable pseudo-update of the Lorenz 1984 first component with (CPS) and without (NSP) correction.

by introducing a truncated polynomial chaos approximation of the state variable

$$x(\omega) \approx \hat{\boldsymbol{x}}(\omega) = \sum_{\alpha \in \mathcal{J}_x} \boldsymbol{x}^{(\alpha)} \Psi_\alpha(\boldsymbol{\vartheta}(\omega)), \tag{84}$$

in which Ψ_α are multi-variate polynomials in random variables $\boldsymbol{\vartheta}$ as arguments. The random variables $\boldsymbol{\vartheta}$ represent the parameterisation of the prior uncertainties in the initial conditions or even model parameters. They are usually taken as independent, uncorrelated random variables of some simpler kind such as for example normal or uniform random variables corresponding to the Askey scheme as discussed in [29]. In a similar manner, one may approximate the predicted error

$$\varepsilon(\omega) \approx \hat{\boldsymbol{\varepsilon}}(\omega) = \sum_{\alpha \in \mathcal{J}_\varepsilon} \boldsymbol{\varepsilon}^{(\alpha)} \Psi_\alpha(\boldsymbol{\eta}(\omega)) \tag{85}$$

in which $\boldsymbol{\eta}(\omega)$ and $\boldsymbol{\vartheta}(\omega)$ are assumed to be independent and uncorrelated. Collecting all random variables of consideration, the global discretisation of the state reads

$$x(\omega) \approx \hat{\boldsymbol{x}}(\omega) = \sum_{\alpha \in \mathcal{J}_\Psi} \boldsymbol{x}^{(\alpha)} \Psi_\alpha(\boldsymbol{\xi}(\omega)), \tag{86}$$

in which $\boldsymbol{\xi}(\omega) := (\boldsymbol{\vartheta}, \boldsymbol{\eta})$

When dealing with time-dependent systems, the approximation as given previously is not optimal when the time integration of the nonlinear system before the update is too long. In such a case the state becomes highly non-Gaussian and requires high-order polynomial chaos approximations. Fig. (14) shows the decrease of the state PCE accuracy in time, and its improvement with the increase of the polynomial order. Similarly, the non-Gaussianity increases the number of sampling points neccessary for the estimation of PCE coefficients as the sparsity of the solution decreases, see Fig. (15).

To resolve this problem, the idea is to change the basis in Eq. (86) to

$$\hat{\boldsymbol{x}}_k(\omega) = \sum_{\alpha \in \mathcal{J}_\Phi} \boldsymbol{x}_k^{(\alpha)} \Phi_\alpha(\boldsymbol{\zeta}(\omega)) = \boldsymbol{\Phi}_k \boldsymbol{v}_k, \tag{87}$$

in which the random variable $\boldsymbol{\zeta}(\omega)$ follows the distribution of the last known state $\boldsymbol{x}_{k-n}(\omega)$ for which the lower order approximation in Eq. (86) is still suitable, and $\Phi_\alpha(\boldsymbol{\zeta}(\omega))$ are the basis functions chosen either as orthogonal via a modified Gram-Schmidt process, or non-orthogonal ones as polynomial maps of the last known state.

The basis transformation starts with the definition of new random variables $\boldsymbol{\zeta}(\omega)$ driven by the evolution law in Eq. (1) such that

$$\boldsymbol{\zeta}(\omega) = g(\boldsymbol{\xi}(\omega)) \tag{88}$$

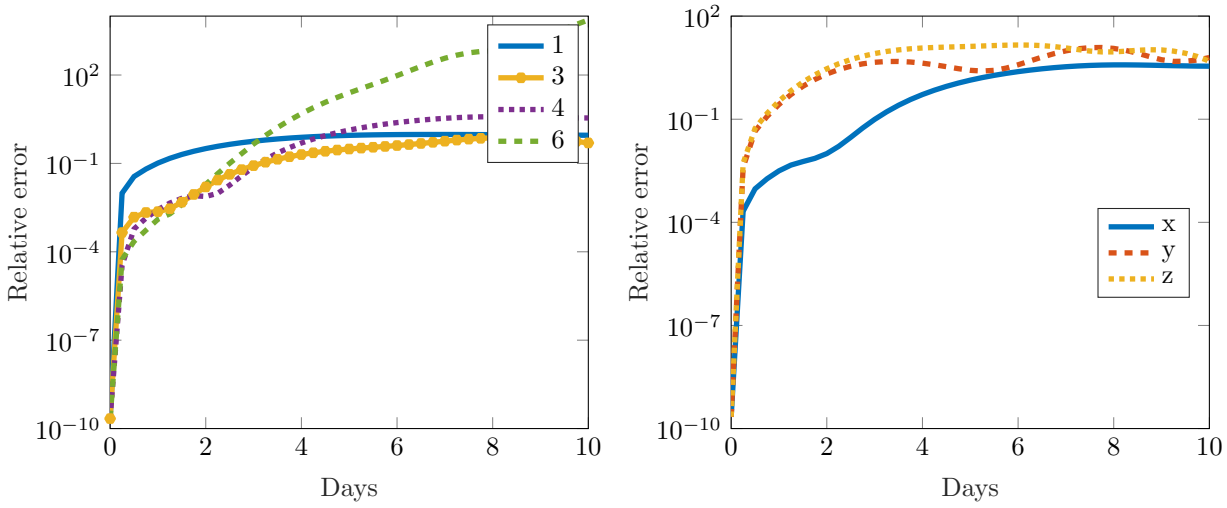


Figure 14: Relative error of a) the first state PCE w.r.t. to the polynomial order for 100 randomly chosen samples b) the state PCE for $p = 4$ and 100 randomly chosen points

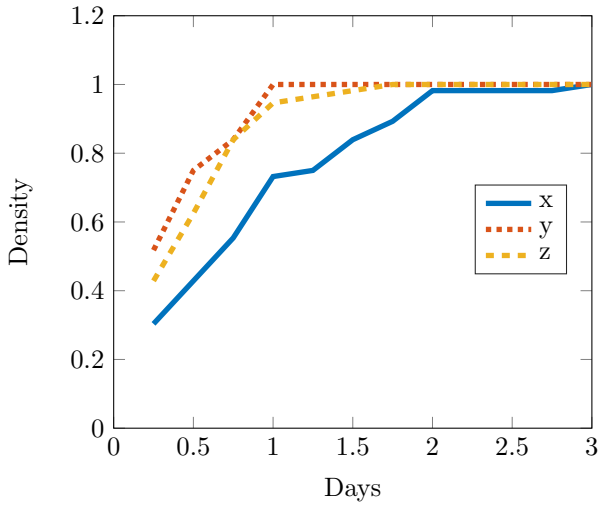


Figure 15: State sparsity in time for fixed polynomial order $p = 4$

holds, in which $g(\boldsymbol{\xi}(\omega))$ describes a nonlinear transformation of the initial random variables $\boldsymbol{\xi}(\omega)$ over some predefined period of time. Let t_{k-n} be the last time moment in which the classical PCE basis can be used to approximate the state $\boldsymbol{x}_{k-n}$. Then, given a small number N of model trajectories $(\boldsymbol{x}_{k-n}(\boldsymbol{\xi}(\omega_i))_{i=1}^N$ for $(\boldsymbol{\xi}(\omega_i))_{i=1}^N$ one may estimate the state coefficients $\boldsymbol{x}_{k-n}^{(\alpha)}$ in the original basis $\Psi_\alpha(\boldsymbol{\xi}(\omega))$. Since $\boldsymbol{x}_{k-n}(\omega)$ is fully defined, one may take $\boldsymbol{\zeta}(\omega) := \boldsymbol{x}_{k-n}(\omega)$. By arranging $\boldsymbol{\zeta}(\omega)$ into multivariate polynomial form, we may define the new basis $\Phi_\alpha(\boldsymbol{\zeta}(\omega))$ using the modified Gram-Schmidt (MGS) orthogonalisation process, for more details please see [11]. In such a case the new state $\boldsymbol{x}_k(\omega)$ at time t_k can be estimated given a small number of trajectories $(\boldsymbol{x}_k(\boldsymbol{\xi}(\omega_i))_{i=1}^N$ and their corresponding basis functions $\boldsymbol{\Phi}(\boldsymbol{\zeta}(\omega_i))$. Having

$$\begin{aligned} \hat{\boldsymbol{x}}_k(\omega_i) &= \sum_{\alpha \in \mathcal{J}_\Phi} \boldsymbol{x}_k^{(\alpha)} \Phi_\alpha(\boldsymbol{\zeta}(\omega_i)) \\ &= \sum_{\alpha \in \mathcal{J}_\Phi} \boldsymbol{x}_k^{(\alpha)} \Phi_\alpha(g(\boldsymbol{\xi}(\omega_i))) \end{aligned} \tag{89}$$

one may estimate the coefficients $\boldsymbol{x}_k^{(\alpha)}$ via Bayesian regression as described in Section 7.2. Here, $\mathcal{J}_\Phi$ is a new multi-index set defined by a polynomial order that is lower than the corresponding Hermite one. This procedure further allows the evaluation of a large number of samples of $\boldsymbol{x}_k$ as the large number of samples of $\boldsymbol{\zeta}$ resp. $\boldsymbol{\xi}$ is known, and hence one may repeat the process to estimate the next unknown state in time t_{k+1}.

Fig. (16) shows the accuracy of the MGS for the polynomial order $p = 3$ and 100 randomly chosen samples w.r.t. the solution obtained from 10^6 Monte Carlo runs. In comparison to the Bayesian regression on classical PCE depicted in Fig. (14) one may note that the accuracy of the MGS solution improves by an order of magnitude for the same number of samples. The dependence of the MGS solution on the number of samples and the polynomial order can be seen in Fig. (17) and Fig. (18), respectively. As expected, the accuracy improves with the sample number. Similar holds for the polynomial order. Finally, the sparsity of the newly obtained approximation is shown in Fig. (19), where it is observed that the first state is much sparser than the other two.

The basis estimation via Gram-Schmidt orthogonalisation can be computationally demanding. Thus, a much more efficient solution is to consider the non-orthogonal basis. The simplest choice is to observe the current state $\boldsymbol{x}_k(\omega)$ as a nonlinear map of the last known one $\boldsymbol{x}_{k-n}(\omega)$, i.e.

$$\boldsymbol{x}_k(\omega) = \sum_{\alpha \in \mathcal{J}_\Upsilon} \boldsymbol{x}_k^{(\alpha)} \Upsilon_\alpha(\boldsymbol{x}_{k-n}(\omega)) \tag{90}$$

in which the coefficients $\boldsymbol{x}_k^{(\alpha)}$ are obtained via regression described in Section 7.2. Here, $\Upsilon_\alpha(\boldsymbol{x}_{k-n}(\omega))$ are taken to be the non-orthogonal multivariate polynomials defined as:

$$\begin{aligned} \Upsilon_\alpha(\boldsymbol{x}_{k-n}(\omega)) &= \boldsymbol{x}_{k-n}^{(\alpha)} \\ &= x_{k-n}^{\alpha_1} y_{k-n}^{\alpha_2} z_{k-n}^{\alpha_3} \end{aligned} \tag{91}$$

with (α) being the multi-index set similarly defined to the one that describes the classical PCE.

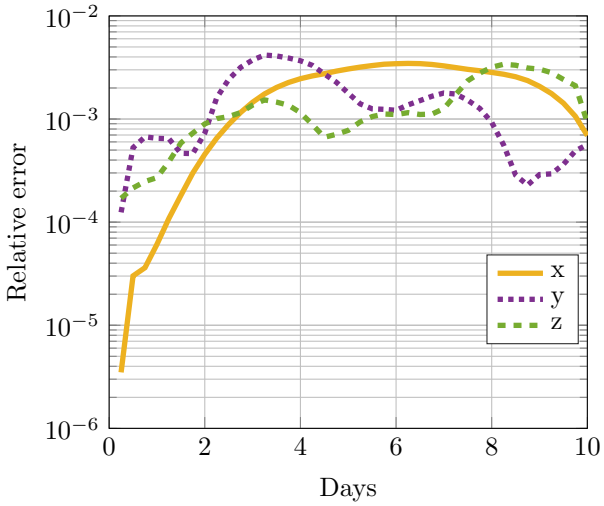


Figure 16: Accuracy of the MGS basis in time for all three Lorenz states

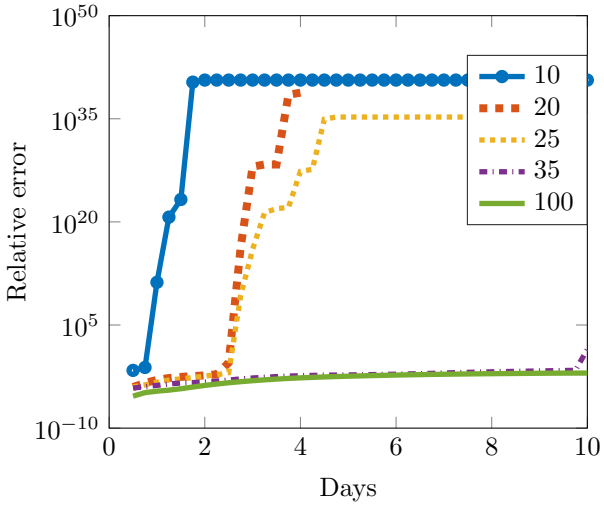


Figure 17: Accuracy of the MGS solution w.r.t. the number of randomly chosen samples for the first Lorenz state for $p = 3$

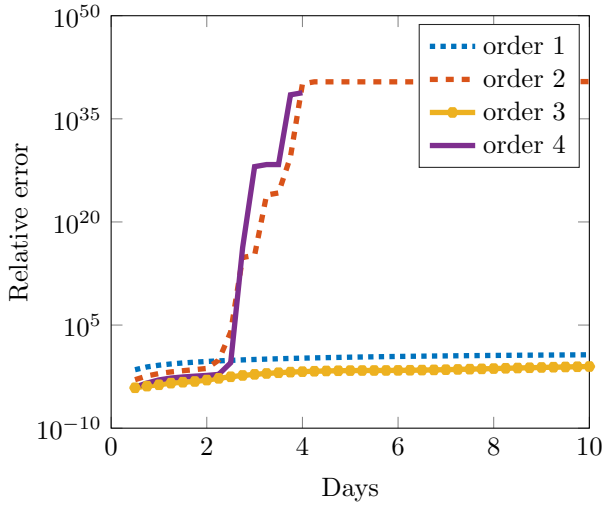


Figure 18: Accuracy of the MGS solution w.r.t. the polynomial order

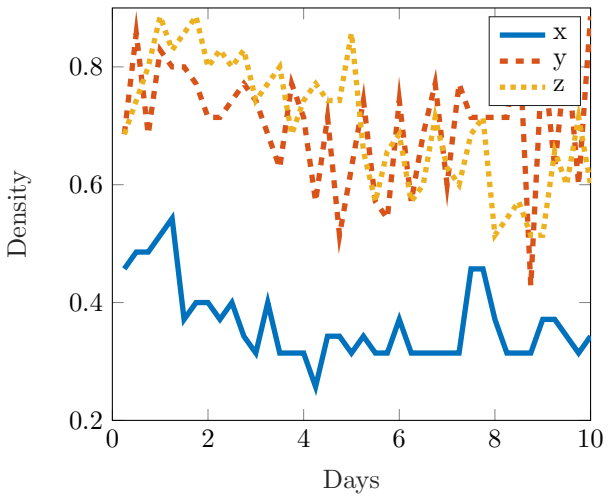


Figure 19: Sparsity of the MGS Lorenz state in time

Fig. (20)a) shows the accuracy of the third and fourth order nonlinear map (NMAP) approximations (of order 3 (NMAP3) resp. order 4 NMAP4) compared to the solution obtained by regressing on the fixed Hermite polynomial basis of fourth order (PCE), and the MGS solution of fourth order. While the PCE solution is not accurate enough, both the MGS and the nonlinear map solutions give similar results for the same order of approximation. In the beginning lower order nonlinear map solution (NMAP3) matches the solution obtained by fixed regression. In contrast to the PCE solution, the error stabilises over time and does not over-estimate the Lorenz state. Furthermore, the accuracy of the NMAP4 solution is tested on different data set sizes in Fig. (20)b). The experiment shows that even a low number of samples (56 samples) can be used to achieve the desired accuracy, see Fig. (20)c).

By using approximations in Eq. (89) and in Eq. (90) one may use a small number of solution trajectories of $\boldsymbol{x}_{k-n}(\omega)$ to estimate a large number of samples $\boldsymbol{x}_k(\omega)$. The approximations are made adaptive such that the last known basis is used in a current time, and the Kullback-Leibler divergence is used to estimate the error compared to the validation set. If the error is bigger than tolerance then the basis is adaptively modified.

Even though both of the previous approximations are significantly better than the original basis, they are not suitable to be used in the filtering process due to correlated arguments, and in the latter case also due to the non-orthogonality. Therefore, to compute the time dependent polynomial chaos approximations, the previous approximations at the update time are transformed such that the non-Gaussian correlated random variables $\zeta(\omega)$ are mapped to uncorrelated Gaussian ones via nonlinear transformation. The main idea of the transformation process is to map the state variable $\boldsymbol{x}_{k-n}(\omega_x)$ by an isoprobabilistic map

to a Gaussian random variable $\theta(\omega_\theta), \omega_\theta \in \Omega_\theta$, i.e.

$$T: \quad \boldsymbol{x}_{k-n}(\omega_x) \mapsto \theta(\omega_\theta) \tag{92}$$

such that the approximations rewrite to the PCE with multivariate Hermite orthogonal basis $\Psi_\alpha(\boldsymbol{\theta}(\omega_\theta))$:

$$\breve{\boldsymbol{x}}_k(\omega_\theta) = \sum_{\alpha \in \mathcal{K}} x_k^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}(\omega_\theta)) \tag{93}$$

characterised by much lower cardinality than the one in Eq. (89) or Eq. (90).

Due to simplicity reasons, the transformation in Eq. (92) is assumed to be of the Nataf-type, which shows good performance for this kind of problem. The other more general type of transformations are the current state of the research and will be discussed in another paper.

The Nataf transform is a composition of maps $T = T_1 \circ T_2$ in which the first one T_1 maps the vector of non-Gaussian random variables $\boldsymbol{\zeta}$ with marginal cumulative distributions $F_{\boldsymbol{\zeta}}$ to the vector of correlated standard Gaussian variables κ via inverse cumulative distribution of the standard normal $\Phi_{\mathcal{N}}^{-1}$:

$$T_1: \quad \boldsymbol{\zeta}(\omega_\zeta) \to \boldsymbol{\kappa}(\omega_\theta) := \Phi^{-1}(F_{\boldsymbol{\zeta}}(\boldsymbol{\zeta})), \tag{94}$$

whereas the second one T_2 maps correlated random variables into uncorrelated ones

$$T_2: \quad \boldsymbol{\kappa}(\omega_\theta) \to \boldsymbol{\theta}(\omega_\theta) = \boldsymbol{C}_\kappa^{-1/2} \boldsymbol{\kappa}(\omega_\theta). \tag{95}$$

Here, the factor $\boldsymbol{C}_\kappa^{-1/2}$ is evaluated using the Cholesky decomposition. To perform the step in Eq. (94), one requires knowledge on the cumulative distribution function (cdf) $F_{\boldsymbol{\zeta}}$. As this information is not accessible, but only instances of the random variable $(\boldsymbol{\zeta}(\omega_j))_{j=1}^M$ are known, one may use the kernel density estimator as the one presented in [4] to obtain $F_{\boldsymbol{\zeta}}$. In addition, $F_{\boldsymbol{\zeta}}$ is interpolated in a Bayesian manner (see Section 8) using the polynomial of order 3.

Hence, for the further process of assimilation one may rewrite Eq. (90) to the orthogonal polynomial chaos expansion expressed in terms of newly estimated standard random variables:

$$\hat{\boldsymbol{x}}_{k-\Delta\tau,a}^{(i+1)} = \hat{\boldsymbol{x}}_{k-\Delta\tau,f} + \hat{\boldsymbol{K}}_{k-\Delta\tau}^{(i)} (\hat{\boldsymbol{x}}_{k,a} - \hat{\boldsymbol{y}}_{kh}^{(i)}) \tag{96}$$

in which

$$y_{kh}(\omega) \approx \hat{\boldsymbol{y}}_{kh}(\omega) = \sum_{\alpha \in \mathcal{K}} \boldsymbol{y}_{k\ell}^{(\alpha)} \Psi_\alpha(\boldsymbol{\theta}(\omega)) + \hat{\varepsilon}(\omega) \tag{97}$$

i.e.

$$\hat{\boldsymbol{y}}_{kh}(\omega) = \hat{\mathring{H}}(\hat{x}_{k-\Delta\tau,f} - \mathring{x}) + \hat{\mathring{h}} + \hat{\varepsilon}. \tag{98}$$

The accuracy of the transformed solution in a Gaussian basis (tMGS- transformed modified Gram-Schmidt process) compared to non-Gaussian ones (denoted by MGS in plot) w.r.t. to the polynomial order is shown in Fig. (21)a). As expected, the Gaussian basis requires higher polynomial order to achieve the same accuracy as the non-Gaussian one.

The comparison of the transformed approach to the classical MGS one is depicted in Fig. (21b). Here, four different types of solutions are considered. The solutions denoted by MGS and tMGS (transformed MGS) are obtained by integrating original samples of the state in time, whereas solutions MGSresamp and tMGSresamp are obtained by sampling the polynomial chaos approximations that are further integrated in time. In the latter case the approximation error gets propagated in time, and hence the solution is less accurate than the corresponding sampled solution. The reason to investigate the second case lies in the updating procedure. After the update of the state is made one does not have the original state samples coming from sampling the initial condition. Instead, one samples the newly obtained polynomial chaos approximation.

The discretised posterior in Eq. (96) is described by both the state random variables as well as the variables describing the measurement noise. The number of the latter ones increases with the number of measurements, and hence the cardinality of the posterior PCE grows. However, the dimension increase can be

avoided by same transformation process as described before. In Fig. (22)a) the accuracy of the transformed state for NMAP estimate with respect to the polynomial order is depicted. The sparsity of the newly obtained state is depicted in Fig. (22)b) and is slightly higher than the one described by the MGS procedure.

Finally, the assimilation results can be significantly different than those obtained using the classical PCE. In Fig. (23) one may see a comparison of the assimilated state using the direct iteration with non-adaptive classical polynomial chaos expansion of order 4 (DS) and the pseudo-state (PS) update (frequency of update is 6h) using the transformed nonlinear map estimate of same order. Clearly, the DS estimation leads to the overestimation of the posterior variance already after one day of estimation, as expected. This is due to the inaccuracy of the state approximations. On the other hand, the transformed nonlinear map estimate and the one based on the transformed modified Gram-Schmidt estimation are giving very close results. Fig. (24) depicts the mean and variance relative errors between these two solutions.

The modified basis approach results in a stable posterior variance with respect to the pseudo-time step size, see Fig. (25) in which are depicted posterior bounds for two different updating step sizes. This, however, does not hold for the classical PCE. In addition, the modified updating procedure is robust with respect to the measurement noise as presented in Fig. (26). Here, the posterior 99% regions are shown for three different values of the measurement noise with c_ε being the coefficient of the variation of noise.

7.2 Sparse polynomial chaos approximations

The Gauss-Newton-Markov-Kalman filter requires repeated evaluations of the forward problem. To reduce the overall computational burden, the propagation of the uncertainty through the forward problem can be achieved in a data-driven non-intrusive spectral setting. Given the approximation of the state in a polynomial chaos setting

$$\hat{\boldsymbol{x}}_f(\omega) = \sum_{\alpha\in\mathcal{J}_\Psi} \boldsymbol{x}_f^{(\alpha)} \Psi_\alpha(\boldsymbol{\xi}(\omega)) = \boldsymbol{\Psi}\boldsymbol{v}, \tag{99}$$

the goal is to estimate the unknown coefficients $\boldsymbol{v}$ given N samples $(\boldsymbol{x}_f(\omega_i))_{i=1}^N$, i.e.

$$\boldsymbol{x}_f(\omega_i) = \sum_{\alpha\in\mathcal{J}_\Psi} \boldsymbol{x}_f^{(\alpha)} \Psi_\alpha(\boldsymbol{\xi}(\omega_i)) \tag{100}$$

for $i = 1, ..., N$. In a vector form the previous relation reads

$$\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{v}. \tag{101}$$

In a case when $N \leq P := \text{card}\, \mathcal{J}_\Psi$, the system in Eq. (101) is undetermined, and requires additional information. As a priori knowledge on the current state exists (e.g. for small time step sizes the subsequent state is close to the current one), one may model the unknown coefficients $\boldsymbol{v}$ a priori as random variables in $L_2(\Omega_v, \mathcal{F}, \mathbb{P}; \mathbb{R}^P)$, i.e.

$$\boldsymbol{v}(\omega_v) := [v^{(\alpha)}(\omega_v)] : \Omega_v \to \mathbb{R}^P,$$

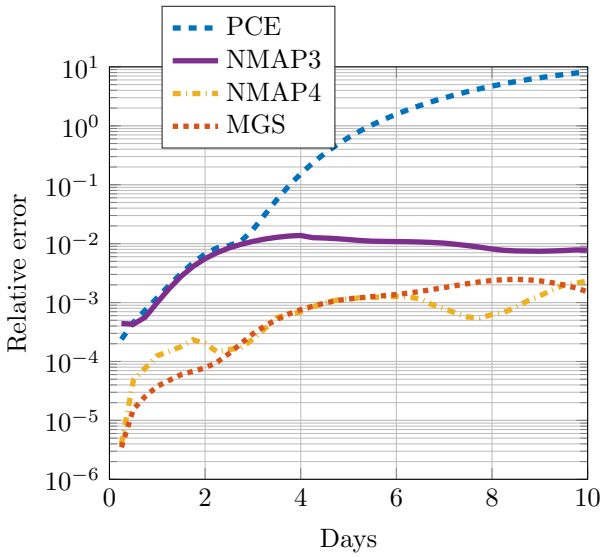
and further use the linear Gauss-Markov-Kalman filtering procedure as previously described to determine their conditional mean. As the coefficients can be both positive and negative, one may assume that the prior $\boldsymbol{v}(\omega_v)$ is normally distributed

$$\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}) \sim e^{-\frac{\|\boldsymbol{v}\|^2}{2}}$$

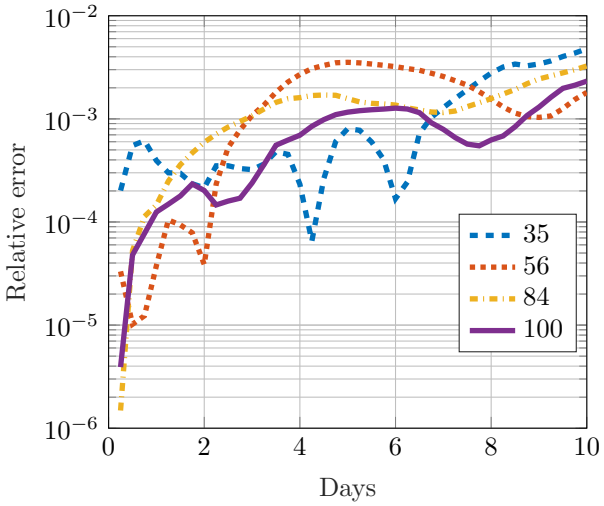
resulting in

$$\boldsymbol{v}_a(\omega_v) = \boldsymbol{v}_f(\omega_v) + \boldsymbol{K}(\boldsymbol{u} - \boldsymbol{\Psi}\boldsymbol{v}_f(\omega_v)). \tag{102}$$

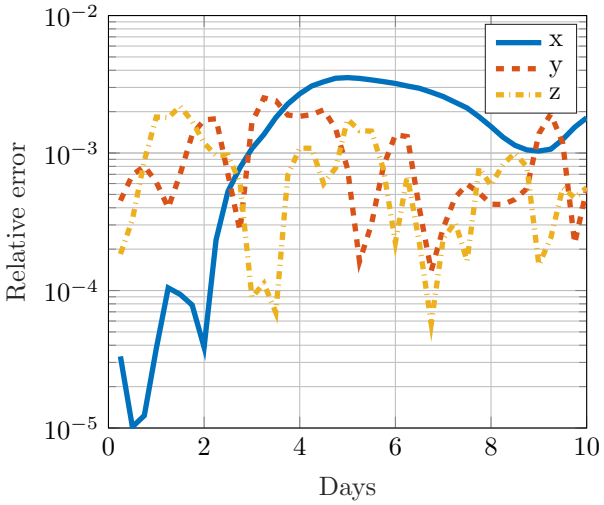
Having the Kalman filter on both the updating and forecasting levels, the last equation together with Eq. (76) forms the hierarchical structure of the iterative Gauss-Newton-Kalman filter. However, such estimation still requires a large number of samples as all coefficients in the polynomial chaos expansion are taken into consideration even those close to zero. To promote for sparsity, see Fig. (15), the prior distribution has to be concentrated more around the zero value. This can be achieved by taking a Laplace prior



a) Comparison of accuracy



b) Accuracy w.r.t. to the number of training points used in regression



c) State accuracy

Figure 20: Non-orthogonal approximation of solution

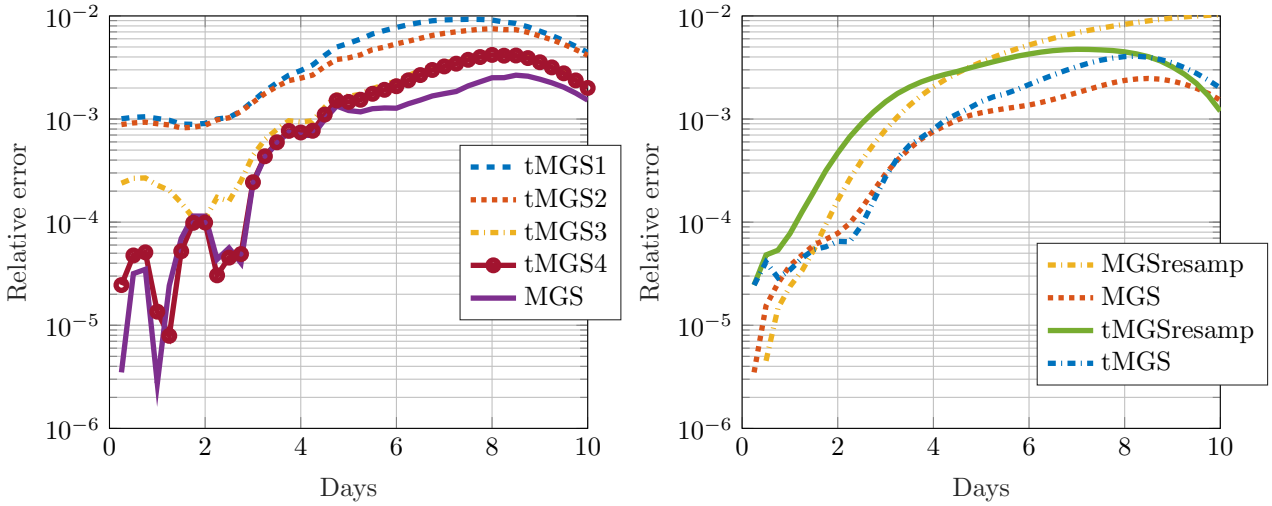


Figure 21: a) Accuracy of transformed MGS solution in time w.r.t. to polynomial order b) Comparison of accuracies of different MGS approaches

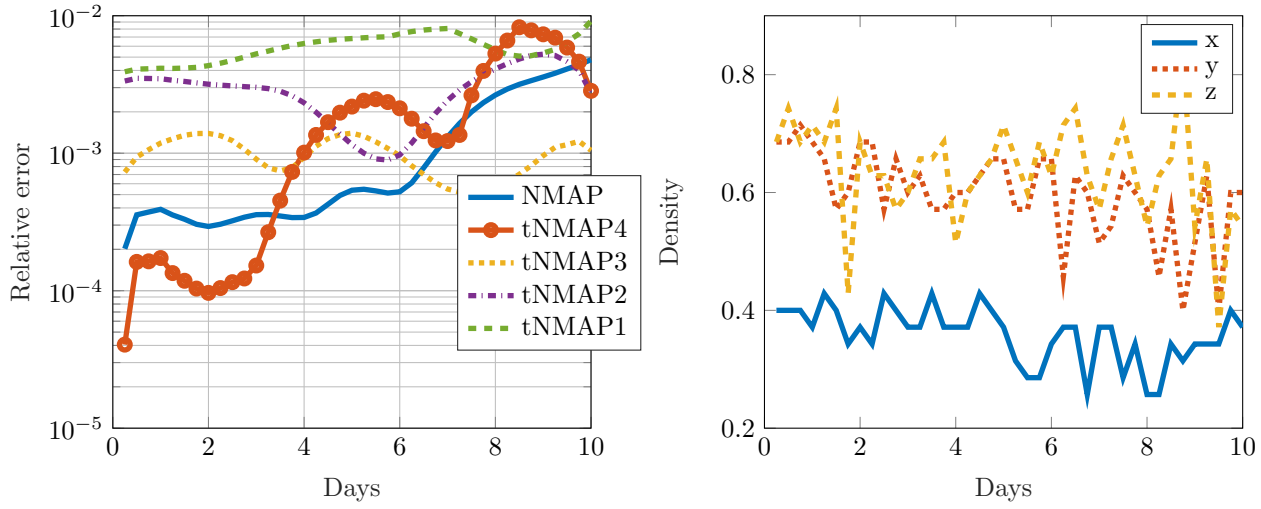


Figure 22: State sparsity of the non-orthogonal approximation of solution

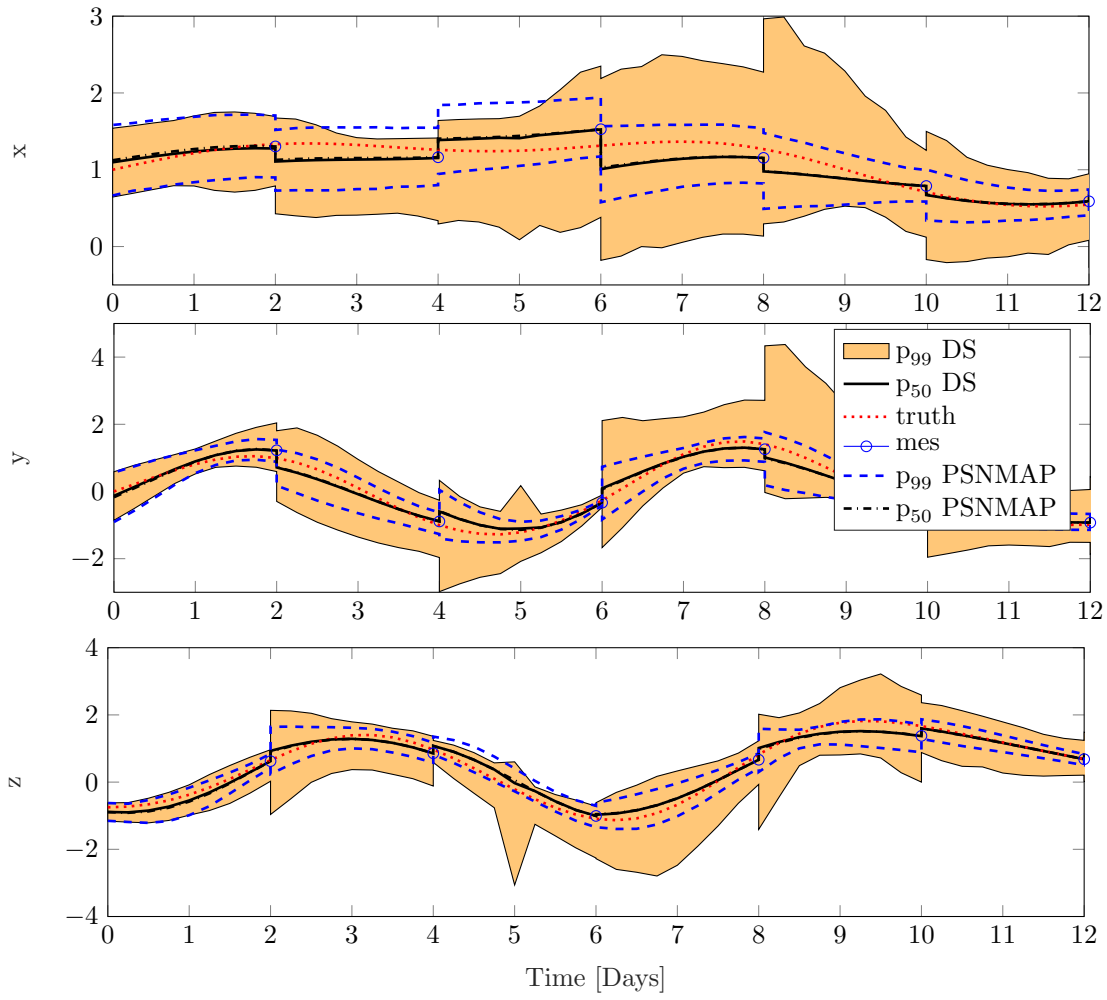


Figure 23: Update of the Lorenz 1984 state backwards every two days. The forecast is estimated using the nonlinear map.

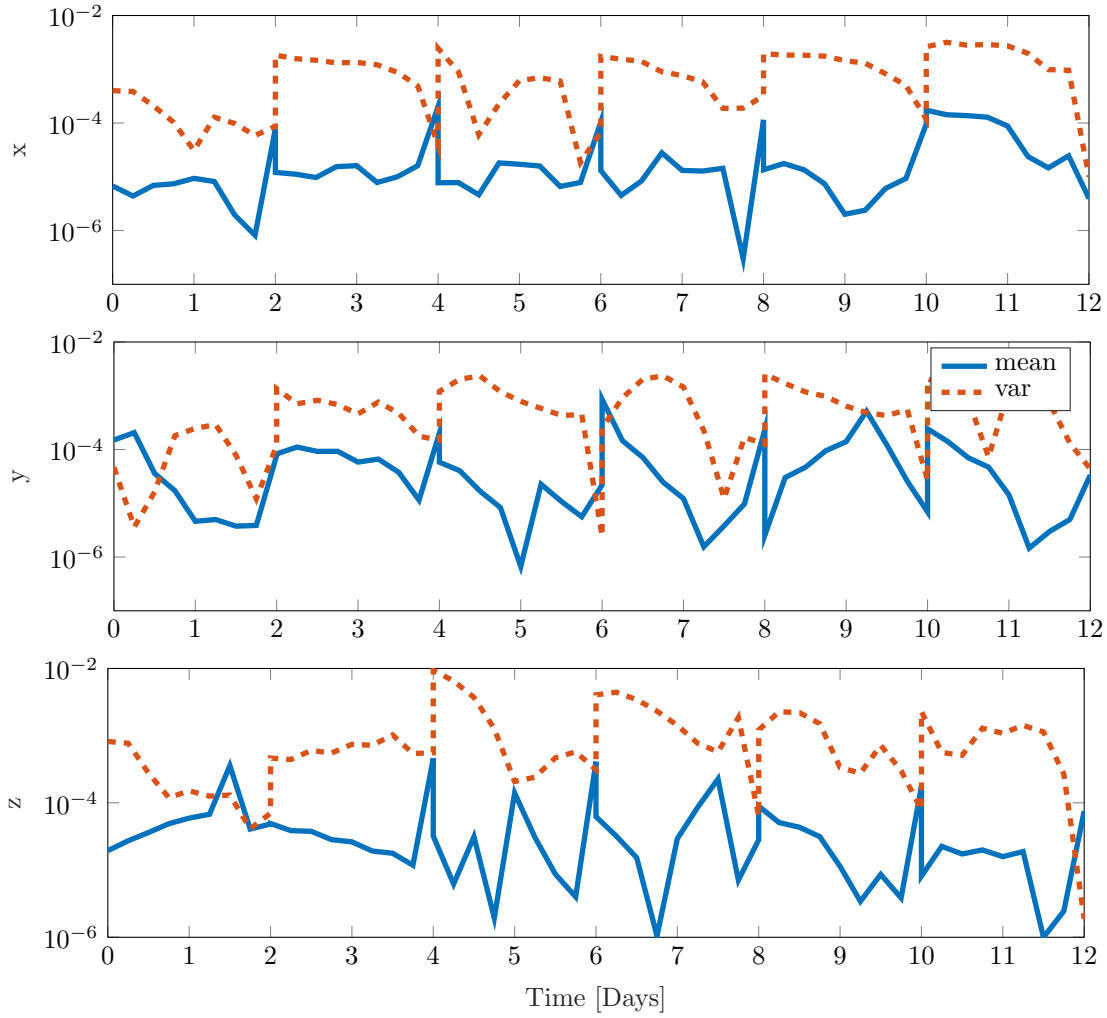


Figure 24: Relative error between the posterior mean and variance of the transformed MGS based solution with respect to the transformed nonlinear map one. Both are obtained by using the square root algorithm

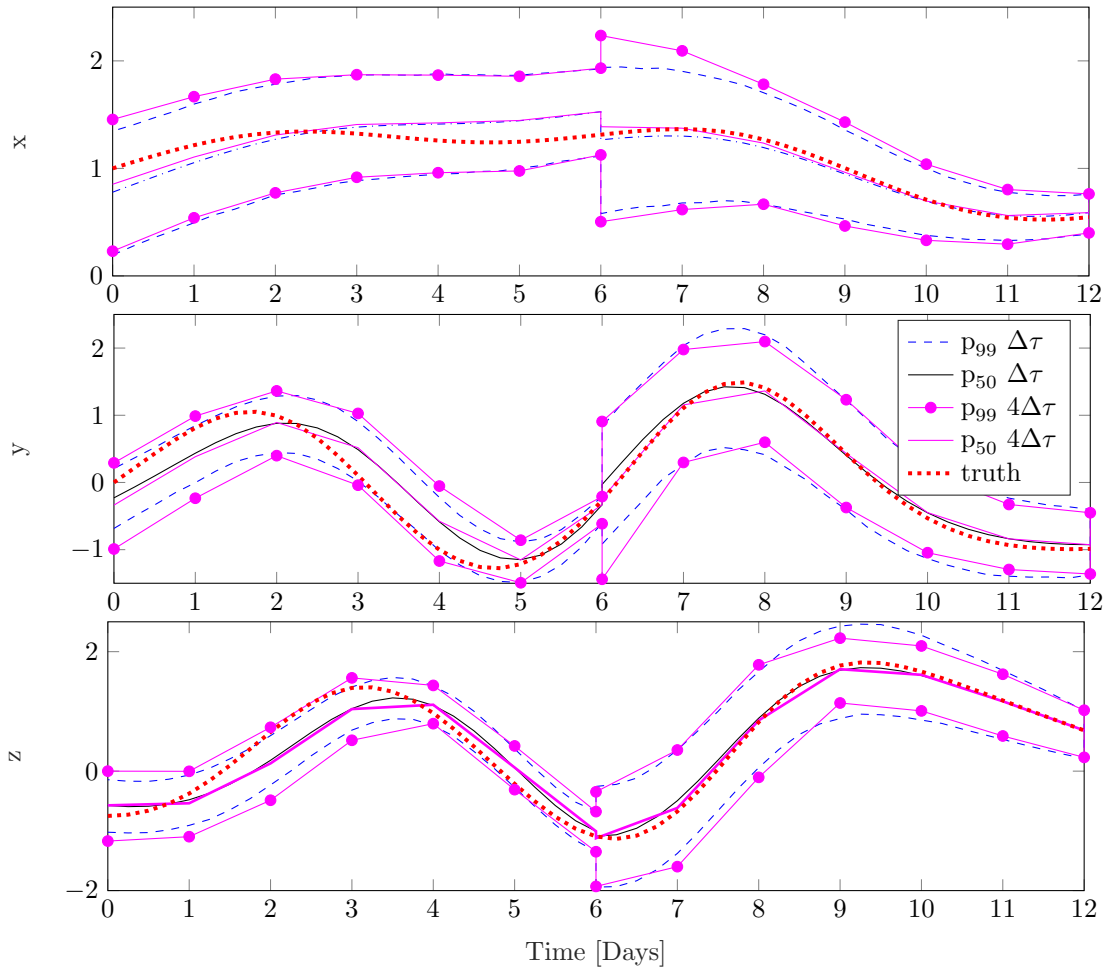


Figure 25: The robustness of the square-root update of the full state with respect to the time step size

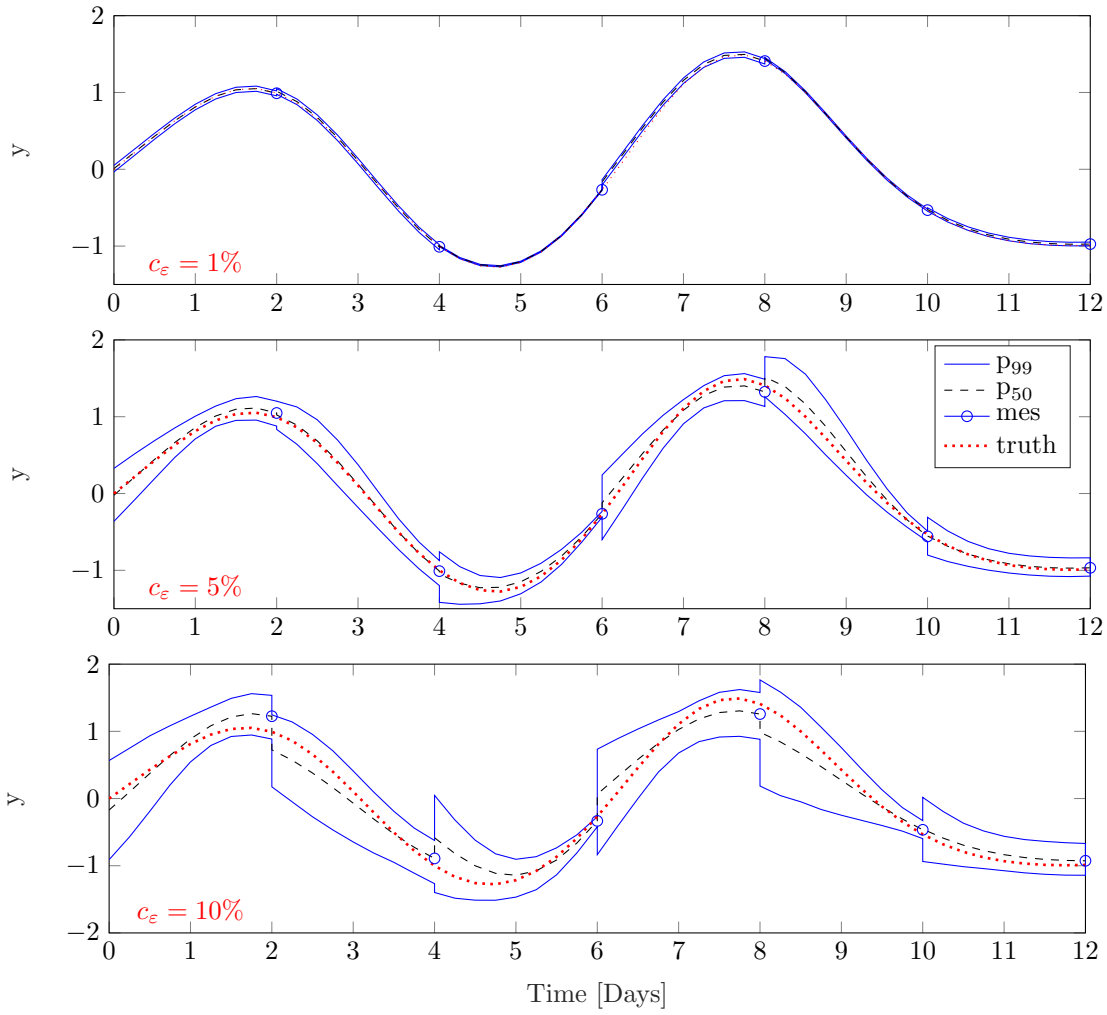


Figure 26: Estimation of the Lorenz state with respect to different coefficients of the variation of noise c_ε

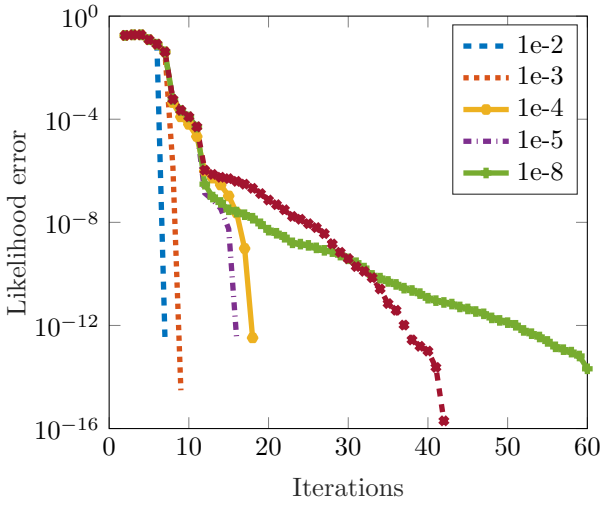


Figure 27: Convergence of marginal likelihood error with respect to the priorly assumed regression error $\mathcal{N}(0, \sigma^2)$ for the Lorenz 1984 example

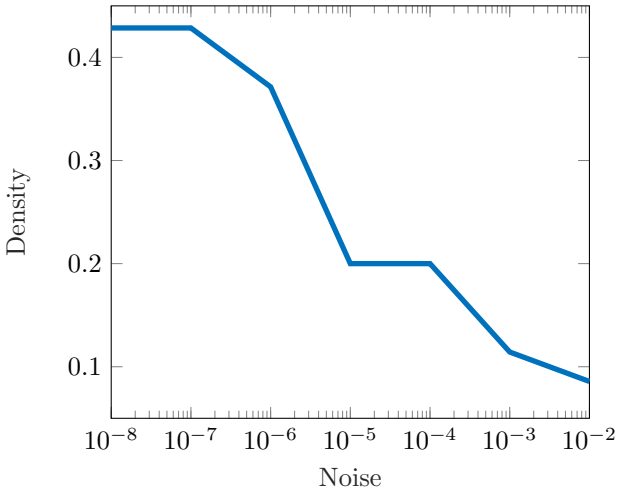


Figure 28: Lorenz 1984 state sparsity with respect to the data noise level

$$\boldsymbol{v} \sim e^{-\|\boldsymbol{v}\|_1}$$

to model the unknown coefficients. As the work with a Laplace prior is computationally difficult, in this paper we use the corresponding hyperprior instead as advocated in relevance vector machine approach, see [26]. The hyperprior is modelled as

$$p(\boldsymbol{v}|\varpi) = \prod_{\alpha \in \mathcal{J}} p(\boldsymbol{v}^{(\alpha)}|\varpi_\alpha), \quad \boldsymbol{v}^{(\alpha)} \sim \mathcal{N}(0, \varpi_\alpha^{-1})$$

with ϖ_α being the precision of each PCE coefficient modelled by a Gamma prior $p(\varpi_\alpha)$. By marginalising over ϖ one obtains the overall prior

$$p(\boldsymbol{v}) = \prod_{\alpha \in \mathcal{J}} \int p(\boldsymbol{v}^{(\alpha)}|\varpi_\alpha) p(\varpi_\alpha) \mathrm{d}\varpi_\alpha$$

which is further simplified by taking the most probable values for $\boldsymbol{\varpi}$, i.e. $\boldsymbol{\varpi}_{MP}$.

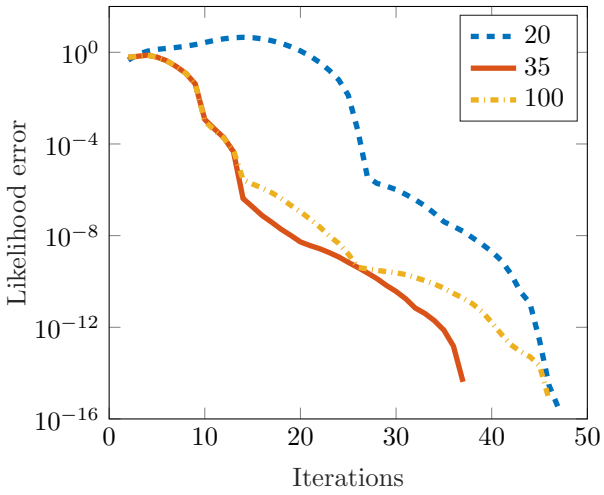


Figure 29: Convergence of the marginal likelihood error with respect to the number of data points for a fixed regression error prior. The state is the second component of the Lorenz 1984 system.

To estimate the coefficients in Eq. (101) we further use Bayes's rule in a form

$$p(\boldsymbol{v}, \boldsymbol{\varpi}, \boldsymbol{\sigma}^2|\boldsymbol{u}) = \frac{p(\boldsymbol{u}|\boldsymbol{v}, \boldsymbol{\varpi}, \boldsymbol{\sigma}^2)}{p(\boldsymbol{u})} p(\boldsymbol{v}, \boldsymbol{\varpi}, \boldsymbol{\sigma}^2) \quad (103)$$

in which the coefficients $\boldsymbol{v}$, the precision $\boldsymbol{\varpi}$ and the regression error $\boldsymbol{\sigma}^2$ are assumed to be uncertain. For computational reasons the posterior is further factorised into

$$p(\boldsymbol{v}, \boldsymbol{\varpi}, \boldsymbol{\sigma}^2|\boldsymbol{u}) = p(\boldsymbol{v}|\boldsymbol{u}, \boldsymbol{\varpi}, \boldsymbol{\sigma}^2) p(\boldsymbol{\varpi}, \boldsymbol{\sigma}^2|\boldsymbol{u})$$

in which the first factoring term is the convolution of normals $p(\boldsymbol{v}|\boldsymbol{u}, \boldsymbol{\varpi}, \boldsymbol{\sigma}^2) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, whereas the second factoring term $p(\boldsymbol{\varpi}, \boldsymbol{\sigma}^2|\boldsymbol{u})$ cannot be computed analytically, and thus is approximated by delta function $p(\boldsymbol{\varpi}, \boldsymbol{\sigma}^2|\boldsymbol{u}) \approx \delta(\boldsymbol{\varpi}_{MP}, \boldsymbol{\sigma}_{MP})$. The estimate $(\boldsymbol{\varpi}_{MP}, \boldsymbol{\sigma}_{MP})$ is obtained from

$$p(\boldsymbol{\varpi}, \boldsymbol{\sigma}^2|\boldsymbol{u}) \propto p(\boldsymbol{u}|\boldsymbol{\varpi}, \boldsymbol{\sigma}^2) p(\boldsymbol{\varpi}) p(\boldsymbol{\sigma}^2)$$

by maximising the evidence (marginal likelihood)

$$p(\boldsymbol{u}|\boldsymbol{\varpi}, \boldsymbol{\sigma}^2) = \int p(\boldsymbol{u}|\boldsymbol{v}, \boldsymbol{\sigma}^2) p(\boldsymbol{v}|\boldsymbol{\varpi}) \mathrm{d}\boldsymbol{v}$$

taking the form

$$\begin{aligned} p(\boldsymbol{u}|\boldsymbol{\varpi}, \boldsymbol{\sigma}^2) &= (2\pi)^{-P/2} (\boldsymbol{R})^{-1/2} \\ &\quad \exp\left(-\frac{1}{2} \boldsymbol{u}^T (\boldsymbol{R})^{-1} \boldsymbol{u}\right) \end{aligned} \quad (104)$$

in which $\boldsymbol{R} = B^{-1} + \boldsymbol{\Psi} A^{-1} \boldsymbol{\Psi}^T$ with $B = \boldsymbol{\sigma}^{-2} I_P$, and $A = \mathrm{diag}(\varpi_\alpha)_{\alpha \in \mathcal{J}}$. By optimality criteria

$$\frac{\partial p(\boldsymbol{u}|\boldsymbol{\varpi}, \boldsymbol{\sigma}^2)}{\partial \boldsymbol{\varpi}} = \boldsymbol{0}$$

and

$$\frac{\partial p(\boldsymbol{u}|\boldsymbol{\varpi}, \boldsymbol{\sigma}^2)}{\partial \boldsymbol{\sigma}^2} = \boldsymbol{0}$$

one may iteratively obtain the values for $\boldsymbol{\varpi}$ and $\boldsymbol{\sigma}^2$. The number of iterations neccessary to achieve the desired accuracy depends on the value of the measurement noise if not marginalised, see Fig. (27). If the signal is clean, the convergence is faster and vice versa. Likewise, the sparsity increases with the increase of the noise magnitude, see Fig. (28). For a higher noise magnitude more polynomial chaos terms can be considered as zero, and vice versa. In addition, the convergence also depends on the size of the data set, see Fig. (29) on the example of a randomly chosen (i.e. Monte Carlo) data set. The convergence is hence faster when more samples are available.

7.3 Sparse optimal maps

Once the functional approximation of the forecasted state is computed, the discretisation of the coefficients of the forward (e.g. Jacobian $\hat{\hat{\boldsymbol{H}}}$) and inverse maps (i.e. Kalman gain $\hat{\boldsymbol{K}}^{(i)}_{k-\Delta\tau}$) in Eq. (96) is the only remaining operation before having full discretisation of the posterior variable. This can be simply achieved by using the direct projection method in which Jacobian and Kalman gains are computed directly by employing Eq. (47) and Eq. (40), respectively, and the formula for the evaluation of the respective covariance matrices:

$$\begin{aligned} \boldsymbol{C}_{q,w} &= \mathbb{E}((\hat{\boldsymbol{q}} - \bar{\boldsymbol{q}}) \otimes (\hat{\boldsymbol{w}} - \bar{\boldsymbol{w}})) \\ &= \sum_{\alpha,\beta \in \mathcal{K}} \mathbb{E}(\Psi_\alpha \Psi_\beta) \boldsymbol{q}^{(\alpha)} \otimes \boldsymbol{w}^{(\beta)} - \bar{\boldsymbol{q}} \otimes \bar{\boldsymbol{w}}. \end{aligned} \quad (105)$$

The last relation can be further rewritten in a matrix form as

$$\boldsymbol{C}_{q,w} = \tilde{\boldsymbol{Q}}_f \boldsymbol{\Delta} \tilde{\boldsymbol{W}}_f^T \quad (106)$$

in which $(\boldsymbol{\Delta})_{\alpha\beta} = \mathbb{E}(\Psi_\alpha \Psi_\beta) = \mathrm{diag}(\alpha!)$ and $\tilde{\boldsymbol{Q}}$ is equal to $\boldsymbol{Q} := (..., \boldsymbol{x}^{(\alpha)}, ...)^T$ without the mean part. Similar holds for $\boldsymbol{W}$.

However, in case of high dimensional problems this approach can be expensive. Therefore, the estimation of a linearised map in Eq. (55) can be rephrased in a similar setting as described in the previous section. Given samples of the a posteriori estimate of

the state $\boldsymbol{x}_{na}^{(i)}(\omega_j)$ in i-th iteration, one may evaluate the samples of the measurement forecast $\boldsymbol{u}(\omega) = [Y(\boldsymbol{x}_{na}^{(i)}(\omega_j))]_{j=1}^N$ such that

$$\boldsymbol{u}(\omega) = \sum_{\alpha\in\mathcal{K}} \boldsymbol{u}^{(\alpha)}\Psi_\alpha(\boldsymbol{\xi}(\omega)) = \boldsymbol{\Psi}\boldsymbol{v} \tag{107}$$

holds. Hence, Bayesian regression as introduced earlier can be used for the estimation of unknown sparse coefficients $\boldsymbol{v}$. In this regard

$$Y_k(\boldsymbol{x}_{na}^{(i)}) = \mathring{\boldsymbol{H}}^{(i)}(\boldsymbol{x}_{na}^{(i)} - \check{\boldsymbol{x}}^{(i)}) + \mathring{\boldsymbol{h}}_k + \boldsymbol{\epsilon}_k \tag{108}$$

holds in which $\mathring{\boldsymbol{H}}^{(i)}, \mathring{\boldsymbol{h}}_k$ and ϵ are unknown, and are to be estimated from underdetermined data. However, in contrast to the problem in the previous section, here one aims at estimating the matrix parameter type. To reduce the estimation to the same form as in Eq. (107), one may vectorise the previous equation to

$$\boldsymbol{u}_x = \boldsymbol{X}\boldsymbol{q}_x + \boldsymbol{\epsilon}_x \tag{109}$$

in which

$$\begin{aligned} \boldsymbol{X} &= [\boldsymbol{1} \quad (\boldsymbol{x}_{na}^{(i)}(\omega_j) - \check{\boldsymbol{x}}^{(i)})^T]_{j=1}^N \in \mathbb{R}^{N\times(d+1)} \\ \boldsymbol{q}_x &= [\boldsymbol{h}_x; (\check{H}^{(i)}(1,:))^T] \end{aligned} \tag{110}$$

and $\boldsymbol{\epsilon}_x = \epsilon_x \otimes \boldsymbol{e}$. Here, $\boldsymbol{x}_{na}(\omega_j) \in \mathbb{R}^d$ is the state sample, $\boldsymbol{e} = [1, 0, ..., 0]^T$ and ϵ_x is the approximation error of the first state. Similarly, one may write

$$\begin{aligned} \boldsymbol{u}_y &= \boldsymbol{X}\boldsymbol{q}_y + \boldsymbol{\epsilon}_y \\ \boldsymbol{u}_z &= \boldsymbol{X}\boldsymbol{q}_z + \boldsymbol{\epsilon}_z. \end{aligned}$$

In these forms Eqs. (110)- (111) can be also solved in a sparse Bayesian setting.

The Jacobian estimated in this manner is slightly better than the estimate obtained using Eq. (47) as can be seen in Fig. (30). Here, the relative errors of regression (J_{reg}) and covariance (J_{cov}) type of Jacobians compared to the analytical value of Jacobian are depicted. Both Jacobians converge very fast, already after 3 iterations, whereas their accuracy deteriorates with the increase of the length of pseudo-update step as expected.

Besides promoting sparsity in the polynomial chaos approximations and the Jacobian, one may also use the Bayesian method to estimate the Kalman gain by solving the linear system

$$\boldsymbol{x}_{k-\Delta\tau,f}(\omega_j) = \boldsymbol{K}_{k-\Delta\tau}^{(i)}\boldsymbol{y}_{kf}^{(i)}(\omega_j) + \boldsymbol{b}^{(i)} + \boldsymbol{\epsilon}_K \tag{111}$$

given the set of sample points $(\boldsymbol{x}_{nf}^{(i)}(\omega_j), \boldsymbol{y}_{kf}^{(i)}(\omega_j))$. Collecting samples of each of the states into vectors $\boldsymbol{u}_1, \boldsymbol{u}_2$ and $\boldsymbol{u}_3$ respectively for x, y and z one may rewrite the previous equation as

$$\boldsymbol{u}_m = \boldsymbol{W}\boldsymbol{k}_m + \boldsymbol{\epsilon}_m, \quad m = 1, ..3 \tag{112}$$

in which $\boldsymbol{W} = [\boldsymbol{1} \quad (\boldsymbol{y}_{kf}^{(i)}(\omega_j))^T] \in \mathbb{R}^{N\times(d+1)}$ and $\boldsymbol{k}_m = [\boldsymbol{b}^{(i)}; (\boldsymbol{K}_\ell^{(i)}(m,:))^T] \in \mathbb{R}^d$. The unknown coefficients $\boldsymbol{k}_m$ can be then evaluated by using the Bayes's rule.

8 Conclusion

We have developed the iterative incremental predictor-corrector Gauss-Newton-Markov-Kalman smoothing algorithm for the non-Gaussian state estimation given noisy measurements. The method is based on the nonlinear local approximation of the conditional expectation, and is mathematically generalised to take into account possible measurement uncertainty. The resulting update equation is discretised by using the time-adaptive polynomial chaos expansion in terms of the standard normal random variables, the number of which matches the state dimension. These are obtained by isoprobabilistic transformation of the non-Gaussian posterior random variable expressed in terms of generalised polynomials of the last known state. The adjustment of the basis functions is achieved via modified Gram-Schmidt as well as nonlinear mapping algorithm such that the desired updating accuracy does not

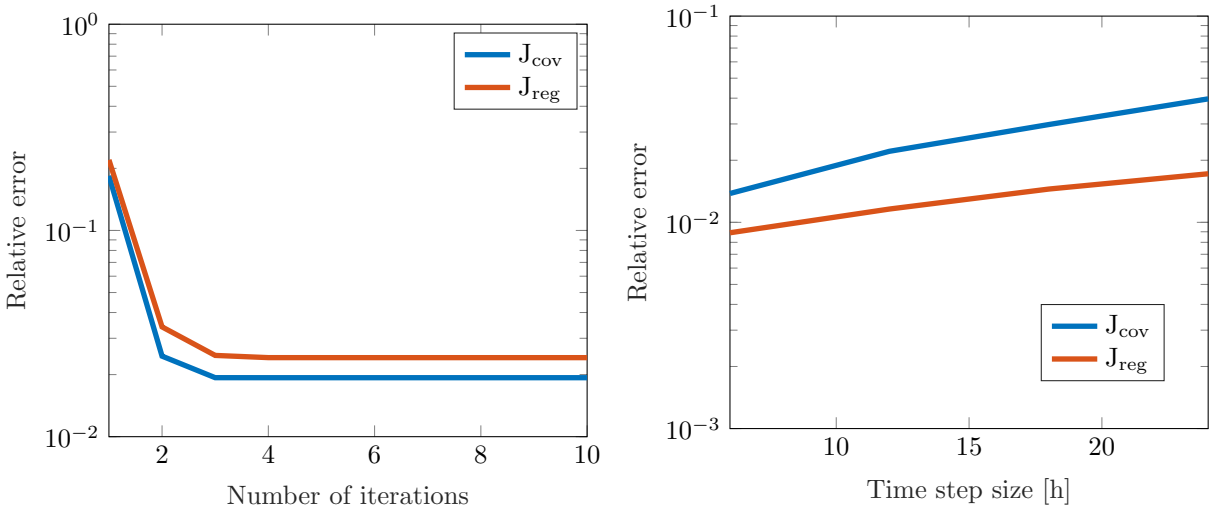


Figure 30: Accuracy of the approximated Jacobian compared to the exact Jacobian

change when the measurement frequency is too low. The resulting Kalman-type update formula for the PCE coefficients can be efficiently computed solely within the PCE. As it does not rely on sampling, the method is robust, fast and exact.

As compared to Monte Carlo, the method is not directly affected by sampling error. However, the method accuracy involves regression error, the truncation error of polynomial approximations (PCE, approximation of optimal map and approximation of inverse map) and errors characterising the transformation of the non-Gaussian random variables. The polynomial approximations here are all evaluated in a data learning setting via Bayes's rule given randomly chosen samples. This may lead to potential over-estimation of some of the polynomial coefficients. However, note that the PCE approximations can be easily exchanged with a fully deterministic Galerkin algorithm for the state estimation obtained given the variational form of the stochastic ordinary differential equations as previously studied in [19].

The updating procedure has been applied to a low-dimensional state estimation problem of the chaotic Lorenz-84 system. It is shown that the method is robust and able to estimate the initial state of the Lorenz-84 system even when the updating step is large and the measurement noise is high. The extension of the presented method to more realistic applications is currently ongoing research. As the numerical complexity of the method increases with the state dimension, the future plan is to consider low-rank techniques as well as to implement more efficient adaptive sampling strategies. This would then allow the use of quadratic approximations in the iterative form. Finally, the proposed method is based on the approximation of the conditional expectation of the state and not its higher moments. The further step is to also include the conditional expectation of the second moment into the updating process as well.

Acknowledgment The author greatly appreciates partial financial funding by the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG) as part of priority programs GRK 2075, SPP 1886 and SPP 1748.

9 Appendix: The Lorenz 1984 system

For the numerical evaluation of the estimation method described in the previous, here we consider the well-known "Lorenz-84" model [15, 16]. It is described by a set of three state variables $\boldsymbol{x} = (x, y, z)^T$. Here x represents a symmetric, globally averaged westerly wind current, whereas y and z represent the cosine and sine phases of a chain of superposed large-scale eddies transporting heat polewards. The state evolution is described by the following set of ordinary differential equations (ODEs):

$$\begin{aligned} \frac{dx}{dt} &= -ax - y^2 - z^2 + aF_1 \\ \frac{dy}{dt} &= -y + xy - bxz + F_2 \\ \frac{dz}{dt} &= -z + xz + bxy, \end{aligned} \tag{113}$$

in which F_1 and F_2 represent known thermal forcings, and a and b are fixed constants.

In the numerical experiment considered in the paper the initial condition of the "unknown truth" is $(1.0, 0.0, -0.75)$, the thermal forcings are set to $F_1 = 8$ and $F_2 = 1$, whereas the parameters are set to $a = 0.25$ and $b = 4$. Given the initial values, the previous system is integrated forward in time using an adaptive embedded Runge-Kutta (RK) scheme of orders 4 and 5.

As the Lorenz-84 model shows chaotic behaviour and is very sensitive to the initial conditions, we model them as independent Gaussian random variables:

$$\begin{aligned} x_0(\omega) &\sim \mathcal{N}(x_0, \sigma_1) \\ y_0(\omega) &\sim \mathcal{N}(y_0, \sigma_2) \\ z_0(\omega) &\sim \mathcal{N}(z_0, \sigma_3) \end{aligned} \tag{114}$$

with $x_0 = y_0 = z_0 = 0$ and standard deviations $\sigma_1 = \sigma_2 = \sigma_3 = 1$.

References

[1] A. Banerjee, X. Guo, and H. Wang. On the optimality of conditional expectation as a bregman predictor. *IEEE Trans. Information Theory*, 51(7):2664–2669, 2005.

[2] B. M. Bell. The iterated Kalman Smoother as a Gauss-Newton Method. *SIAM Journal on Optimization*, 4(3):626–636, 1994.

[3] A. Bobrowski. *Functional analysis for probability and stochastic processes: an introduction.* Cambridge University Press, Cambridge, Cambridge, UK, 2005.

[4] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *Ann. Statist.*, 38(5):2916–2957, 10 2010.

[5] L.M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200 – 217, 1967.

[6] Y. Chen and D. S. Oliver. Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Mathematical Geosciences*, 44(1):1–26, 2012.

[7] N. Chustagulprom, S. Reich, and M. Reinhardt. A hybrid ensemble transform particle filter for nonlinear and spatially extended dynamical systems. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):592–608, 2016.

[8] G. A. Einicke. *Smoothing, filtering and prediction: estimating the past, present and future.* InTech, 2012.

[9] G. A. Einicke and B. Langford. Robust extended Kalman filtering. *IEEE Transactions on Signal Processing*, 47(9):2596–2599, 1999.

[10] D. Gamerman. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman & Hall, Boca Raton, USA, 2 edition, May 1997.

[11] Marc Gerritsma, Jan-Bart van der Steen, Peter Vos, and George Karniadakis. Time-dependent generalized polynomial chaos. *J. Comput. Phys.*, 229(22):8333–8363, November 2010.

[12] S. Gratton, A. S. Lawless, and N. K. Nichols. Approximate Gauss-Newton methods for nonlinear least squares problems. *SIAM Journal on Optimization*, 18(1):106–132, 2007.

[13] X. Kai, C. Wei, and L. Liu. Robust extended Kalman filtering for nonlinear systems with stochastic uncertainties. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(2):399–405, 2010.

[14] R.E. Kalman. A new approach to linear filtering and prediction problems. *ASME. J. Basic Eng.*, 82(1):35–45, 1960.

[15] E. N. Lorenz. Irregularity: a fundamental property of the atmosphere. *Tellus A*, 36(2):98–110, 1984.

[16] Edward N. Lorenz. A look at some details of the growth of initial uncertainties. *Tellus A*, 57(1):1–11, 2005.

[17] H. G. Matthies, E. Zander, B. Rosić, and A. Litvinenko. Parameter estimation via conditional expectation: a Bayesian inversion. *Advanced Modeling and Simulation in Engineering Sciences*, 3(1):1–21, 2016.

[18] R. Van Der Merwe and E. A. Wan. The square-root unscented Kalman filter for state and parameter-estimation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01)*, volume 6, pages 3461–3464. IEEE, 2001.

[19] O. Pajonk, B. Rosić, A. Litvinenko, and H. G. Matthies. A deterministic filter for non-Gaussian Bayesian estimation. *Physica D: Nonlinear Phenomena*, 241(7):775–788, 2012.

[20] B. Ristić, S. Aurlampalam, and N. Gordon. *Beyond the Kalman filter: particle filters for tracking applications*. Artech House Publishers, Boston, 2004.

[21] B. Rosić, O. Pajonk, A. Litvinenko, and H. G. Matthies. Sampling-free linear Bayesian update of polynomial chaos represenations. *Journal of Computational Physics*, 231(17):5761–5787, 2012.

[22] P. Sakov, D. Oliver, and L. Bertino. An iterative EnKF for strongly nonlinear systems. *Monthly Weather Review*, 140(6):1988–2004, 2012.

[23] D. Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[24] A. F. M. Smith and G. O. Roberts. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(1):3–23, 1993.

[25] H. A. Tchelepi, H. Bazargan, and M. A. Christie. Efficient Markov chain Monte Carlo sampling using polynomial chaos expansion. In *Proceedings of the SPE Reservoir Simulation Symposium*, The Woodlands, Texas, United States, 2013. online.

[26] M. E. Tipping. Sparse bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.

[27] E. A. Wan and R. Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. IEEE, 2000.

[28] K. Wang, T. Bui-Thanh, and O. Ghattas. A randomized maximum a posterior method for posterior sampling of high dimensional nonlinear Bayesian inverse problems. *arXiv preprint arXiv:1602.03658*, 2016.

[29] Dongbin Xiu. *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton University Press, Princeton, NJ, USA, 2010.