Ensemble-marginalized Kalman filter for linear time-dependent PDEs with noisy boundary conditions: application to heat transfer in building walls

Marco Iglesias\textsuperscript{1,5}, Zaid Sawlan\textsuperscript{2,5}, Marco Scavino\textsuperscript{2,4}, Raúl Tempone\textsuperscript{2} and Christopher Wood\textsuperscript{3}

1 School of Mathematical Sciences, University of Nottingham, Nottingham, United Kingdom
2 King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Science and Engineering Division (CEMSE), Thuwal 23955-6900, Saudi Arabia
3 Department of Architecture and Built Environment, University of Nottingham, Nottingham, United Kingdom
4 Instituto de Estadística (IESTA), Universidad de la República, Montevideo, Uruguay

E-mail: marco.iglesias@nottingham.ac.uk, zaid.sawlan@kaust.edu.sa, marco.scavino@kaust.edu.sa, raul.tempone@kaust.edu.sa and christopher.wood@nottingham.ac.uk

Received 13 December 2017, revised 7 April 2018
Accepted for publication 3 May 2018
Published 22 May 2018

Abstract
In this work, we present the ensemble-marginalized Kalman filter (EnMKF), a sequential algorithm analogous to our previously proposed approach (Ruggeri et al 2017 Bayesian Anal. 12 407–33, Iglesias et al 2018 Int. J. Heat Mass Transfer 116 417–31), for estimating the state and parameters of linear parabolic partial differential equations in initial-boundary value problems when the boundary data are noisy. We apply EnMKF to infer the thermal properties of building walls and to estimate the corresponding heat flux from real and synthetic data. Compared with a modified ensemble Kalman filter (EnKF) that is not marginalized, EnMKF reduces the bias error, avoids the collapse of the ensemble without needing to add inflation, and converges to the mean field posterior using 50\% or less of the ensemble size required by EnKF. According to our results, the marginalization technique in EnMKF is key to performance improvement with smaller ensembles at any fixed time.

Keywords: ensemble Kalman filter, linear PDEs, heat equation, nuisance boundary parameters marginalization, heat flux measurements, thermal resistance, heat capacity

(Some figures may appear in colour only in the online journal)
List of notation

\[ \theta \]
Parameter vector

\[ A_{\theta} \]
Forward model matrix operator

\[ T_{k} \]
State vector at time \( t_k \)

\[ B_{\theta} \]
Control matrix operator

\[ u_{k} \]
Control vector at time \( t_k \)

\[ F \]
Control evolution matrix operator

\[ H \]
Observation matrix operator

\[ y_{k} \]
Observation vector at time \( t_k \)

\[ y_{1:k} = (y_{1}, \ldots, y_{k}) \]
Observation vectors from time \( t_1 \) to \( t_k \)

\[ z_{k} \]
Control vector at time \( t_k \)

\[ z_{1:k} = (z_{1}, \ldots, z_{k}) \]
Control vectors from time \( t_1 \) to \( t_k \)

\[ X_{k} \]
Augmented state-parameter vector at time \( t_k \)

\[ w_{k} \]
Centered Gaussian noise of the state vector \( T_{k} \) with covariance matrix \( W_{k} \)

\[ v_{k} \]
Centered Gaussian noise of the observation vector \( y_{k} \) with covariance matrix \( V_{k} \)

\[ q_{k} \]
Centered Gaussian noise of the control vector \( u_{k} \) with covariance matrix \( Q_{k} \)

\[ c_{k} \]
Centered Gaussian noise of the observed control vector \( z_{k} \) with covariance matrix \( C_{k} \)

\[ T_{\theta k} \]
\( E[T_{k} | \theta, z_{1:k}, y_{1:k}] \)

\[ T_{\theta k-1} \]
\( E[T_{k} | \theta, z_{1:k}, y_{1:k-1}] \)

\[ P_{\theta k} \]
Cov\( (T_{k} | \theta, z_{1:k}, y_{1:k}) \)

\[ P_{\theta k-1} \]
Cov\( (T_{k} | \theta, z_{1:k}, y_{1:k-1}) \)

\[ u_{\theta k} \]
\( E[u_{k} | z_{1:k}] \)

\[ P_{u_{\theta k}} \]
Cov\( (u_{k} | z_{1:k}) \)

\[ P_{k | k} \]
Cov\( (X_{k} | z_{1:k}, y_{1:k-1}) \)

1. Introduction

Uncertainty in the thermophysical properties of building envelopes has been recognized as one of the main sources of uncertainty when simulating the thermal performance of buildings [3–7]. Simulations of this kind are essential to the development of cost-effective retrofitting measures aimed at improving the energy efficiency of existing buildings. Moreover, accurate uncertainty analyses of the energy performance of the housing stock are vital for drafting policies to reduce global carbon emissions [3, 8]. Most existing approaches use simplified heat diffusion models to quantify uncertainty in the thermophysical properties of walls [8, 9]. However, recent works have suggested that a PDE-constrained data-assimilation approach is required to accurately estimate the thermophysical properties of walls [2, 10]. Most existing statistical approaches that infer thermophysical properties of walls [8–10] assume that the boundary temperatures, \( T_{\text{int}}(t) \) and \( T_{\text{ext}}(t) \), can be approximated by their corresponding measurements recorded on the internal and external surfaces of a wall by temperature
sensors. Although these approaches ignore the uncertainty in those measurements, they enable the construction of a well-defined parameter-to-measurements map that can be inverted via existing frameworks [11]. However, we recently showed that ignoring the uncertainty in boundary temperatures can lead to biased estimates of the thermophysical properties and, thus, reduce the accuracy of uncertainty analyses in energy-performance evaluations of buildings [2]. In order to account for the uncertainty of boundary measurements within a PDE-constrained Bayesian framework, we developed a marginalization technique that assimilates the available data all at once. This approach was both (i) accurate in identifying thermophysical properties of walls, and (ii) useful for developing off-line experimental design strategies to reduce the measurement campaigns.

Following these recent advances, practitioners would substantially benefit from a data assimilation framework capable of updating the probabilistic knowledge of thermophysical properties as new measurements become available. Such a framework could be used to assess in real time whether a specific level of accuracy or uncertainty has been achieved, in order to reduce the long measurement campaigns that often involve several weeks [5]. The objective of this work is to develop a sequential Bayesian approach to infer the thermal properties of walls and estimate the corresponding heat flux within a computationally tractable framework that accounts for the uncertainty in the measurements of the boundary temperatures.

Bayesian filtering [12, 13] can be used to compute the desired distribution for a joint state-parameter assimilation problem. However, sampling is required to approximate the desired distribution when the Bayesian posterior cannot be expressed in a closed form. For example, particle filters use importance sampling to represent the posterior distribution. Unfortunately, particle filters may become computationally prohibitive [14] and, thus, unsuitable for practical implementation in real-time data assimilation. As an alternative to fully Bayesian samplers, the ensemble Kalman filter (EnKF) is the preferred method for data assimilation in partially observed systems where the underlying dynamics are nonlinear and the state space is high-dimensional [15]. Initially proposed for state estimation [16], modified versions of EnKF have been proposed to address joint state-parameter estimation [17]. The EnKF methodology uses a Gaussian approximation for the predictive step in the Bayesian filtering framework; this enables the characterization of the Bayesian posterior within the analysis step as a Gaussian distribution with mean and covariance computed via standard Kalman-like formulas. Different approaches to characterize this Gaussian distribution have led to several EnKF algorithms including the stochastic EnKF and the ensemble square-root Kalman filter that are reviewed elsewhere [15, 18].

The Gaussian approximations within the EnKF framework do not lead to algorithms with asymptotic convergence to the desired posterior when the underlying dynamics of the state-parameter system are nonlinear [19]. Nevertheless, EnKF has been successfully used in numerous applications, including oceanography, meteorology, subsurface flow and engineering [20, 21]. In most of these studies, the success of EnKF was reliant on ad-hoc fixes such as covariance inflation and localization [20, 22, 23]. More recently, regularized versions of EnKF were proposed that merged ideas from particle filters with iterative regularization techniques [10, 24, 25]. For example, a regularizing EnKF was applied for parameter estimation of thermal properties of walls [10]. This approach, however, focused entirely on the Bayesian parameter estimation of the thermophysical properties and did not consider joint state-parameter estimation. Consequently, the resulting algorithm required that the heat transfer model was run from the initial start time until the current assimilation time to update the state of the system (i.e. the wall-temperature profile). A joint state-parameter estimation approach for the real-time assimilation of in situ measurements that enabled sequential updating of both the parameters and the state would be much more efficient.

In this paper, we study the joint state-parameter estimation problem within a setting of linear PDEs with noisy boundary conditions motivated by the aforementioned application which involves
the mathematical modeling of heat transfer through walls. More specifically, we are interested in
the real-time uncertainty quantification of thermophysical properties of walls via the sequential
assimilation of in situ measurements of the wall’s surface temperature and heat flux. We propose an
ensemble-marginalized Kalman filter (EnMKF) to approximate the joint state-parameter Bayesian
estimation problem for a partially observed linear system that is obtained from the discretization
of linear time-dependent PDEs. In this general setting, we obtain a linear forward operator that
depends nonlinearly on a vector of parameters while the observation operator is linear and inde-
dependent of all parameters. In our specific application, we also adapt EnMKF and EnKF to handle
an observation operator that depends linearly on the thermal resistance parameter.

The remainder of the paper is organized as follows. In section 2, we introduce the state-
space formulation that can be used to analyze a class of linear PDE problems. We present
the filtering approach for estimating the unknown state-parameter vectors in section 3. In
particular, section 3.3 provides a Bayesian formulation for the proposed marginalized filtering
algorithms. Section 3.4 contains our novel ensemble-marginalized Kalman filter (EnMKF). In
section 3.5, we consider a slightly modified EnKF where the boundary conditions are sampled
from their filtered distribution, and we describe the primary differences between EnMKF and
the modified EnKF. The performances of the EnMKF and the modified EnKF algorithms are
compared using real and synthetic data in section 4. Conclusions are summarized in section 5.

2. General formulation

Consider the linear time-dependent parabolic initial-boundary value problem on the domain
\(D \times [0, t_N] \subset \mathbb{R}^{d+1}\):

\[
\begin{align*}
\frac{\partial T}{\partial t} + L_\theta T &= 0, \ x \in D, \ t \in [0, t_N] \\
T(x, t) &= u(x, t), \ x \in \partial D, \ t \in [0, t_N] \\
T(x, 0) &= T_0(x), \ x \in D,
\end{align*}
\]

(1)

where \(L_\theta T = -\sum_{ij=1}^{d} a_{ij}(x) \frac{\partial^2 T}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial T}{\partial x_i} + c(x)T\) and \(\theta = (a_{11}, \ldots, a_{dd}, b_1, \ldots, b_d, c)\). Our goal is to estimate \(\theta\) and \(T\) sequentially in time given partial observations of \(T(x, t)\) and measurements of \(u(x, t)\). As a special case, we consider in section 4 the heat equation to model the heat transfer process through building walls and, given real measurements, we estimate the thermal properties of a brick wall.

By discretizing (1) using finite differences or finite elements [1], we can derive the discrete
state-space model

\[
\begin{align*}
\mathbf{T}_k &= A_\theta \mathbf{T}_{k-1} + B_\theta \mathbf{u}_k + \mathbf{w}_k, \\
\mathbf{y}_k &= H \mathbf{T}_k + \mathbf{v}_k,
\end{align*}
\]

(2)

where \(\mathbf{T}_k\) is the state vector at time \(t_k\). \(\mathbf{y}_k\) corresponds to measurements of \(T_k\). \(A_\theta\) is the
forward model operator that depends on the parameter vector \(\theta\), \(H\) is a linear observational
map, \(B_\theta\) is the control matrix operator, and we assume that the control vector \(\mathbf{u}_k\) follows
another state observation system:

\[
\begin{align*}
\mathbf{u}_k &= F \mathbf{u}_{k-1} + \mathbf{q}_k, \\
\mathbf{z}_k &= \mathbf{u}_k + \mathbf{e}_k,
\end{align*}
\]

(3)
where \( F \) is a user-defined linear evolution operator (see section 3.1) and \( z_k \) is the observed control vector at time \( t_k \). We further assume that \( w_k, v_k, q_k, c_k \) are independent centered Gaussian random vectors with covariances \( W_k, V_k, Q_k, C_k \), respectively.

The joint Bayesian estimation of \( \theta \) and \( T_k \), in a sequential filtering approach, as the observations \( \{ z_{1:k}, y_{1:k} \} \) become available can be now posed in terms of the computation of the conditional density \( p(T_k, \theta | z_{1:k}, y_{1:k}) \). Following the marginalization technique \([1, 2]\) and ideas from marginalized particle filters \([26–28]\), we exploit the linear structure of (2) and propose an ensemble-marginalized Kalman filter (EnMKF) that provides an approximation of \( p(T_k, \theta | z_{1:k}, y_{1:k}) \). From (2), it is easy to approximate the conditional distribution \( p(T_k, \theta | u_k, y_{1:k}) \). In order to incorporate the error of the boundary conditions, we use the tractability of the distribution \( p(u_k | z_{1:k}) \). In other words, by noticing from (3) that \( u_k \) follows a linear dynamic system independent of (2), and under further Gaussian assumptions on \( u_0 \), it follows that the distribution \( p(u_k | z_{1:k}) \) can be expressed in closed form by using the Kalman filter \([29]\). The desired joint distribution \( p(T_k, \theta, u_k | z_{1:k}, y_{1:k}) \) is then approximated by integrating out \( u_k \) from the fully joint distribution \( p(T_k, \theta, u_k | z_{1:k}, y_{1:k}) \) as follows:  

\[
\begin{align*}
K_k &= P_{u|k}^{-1} \left( \Sigma_{u|k}^{-1} + C_k \right)^{-1}, \\
u_{k|k} &= u_{k|k-1} + K_k \left( z_k - u_{k|k-1} \right), \\
\Sigma_{u|k} &= (I - K_k) P_{u|k}^{-1}.
\end{align*}
\]

3. Methodology

In this section, we introduce the ensemble-marginalized Kalman filter (EnMKF) as a state-parameter estimation algorithm for the partially observed linear system (2). The boundary conditions \( u_k \) are estimated first by applying the classical Kalman filter to (3). Next, we derive a marginalized Kalman filter for the state \( T_k \), assuming that \( \theta \) is known. Then, this intermediate result is generalized to develop a joint filter for the state \( T_k \) and parameter \( \theta \) based on the EnKF framework. EnMKF differs from EnKF in the prediction step; in EnMKF, the prediction covariance of \( T_k \) is inflated by the expectation of the analysis covariance matrix of \( u_k \) with respect to \( \theta \).

3.1. Kalman filtering for \( u_k \)

The state-space model for \( u_k \) in (3) is linear and independent from \( T_k \). Therefore, we can implement Kalman filter given the observation \( z_k \) as follows:

1. Prediction step:
\[
\begin{align*}
\hat{u}_{k|k-1} &= Fu_{k-1|k-1}, \\
\Sigma_{u|k}^{-1} &= FP_{u|k}^{-1} F^T + Q_k.
\end{align*}
\]

2. Analysis step:
\[
\begin{align*}
K_k^u &= P_{u|k}^{-1} \left( P_{u|k}^{-1} + C_k \right)^{-1}, \\
u_{k|k} &= \hat{u}_{k|k-1} + K_k \left( z_k - \hat{u}_{k|k-1} \right), \\
P_{u|k} &= (I - K_k) P_{u|k}^{-1}.
\end{align*}
\]

where \( K_k^u \) is the Kalman gain matrix of the state-space model for the control vector \( u_k \).

Assuming \( u_0 \) is Gaussian, the Kalman filter provides the exact filtering distribution as \( N(u_{0|k}, P_{0|k}) \).

In many applications, the forward operator \( F \) is unknown and autoregressive models must be considered \([30]\). Some proposals for the autoregressive forward models are:
• Random walk model AR(1)
  \[ u_k = u_{k-1} + q_k. \]

• Random increment model AR(2)
  \[ u_k = 2u_{k-1} - u_{k-2} + q_k, \]
  with \( q_k \sim N(0, Q_k). \)

In section 4.2, we impose AR(1) and AR(2) on the boundary temperatures, \( T_{\text{int}}(t) \) and \( T_{\text{ext}}(t) \), and compare the filtered results with the real data.

**Remark.** Although the state-space model for \( u_k \) is independent from \( T_k \), we apply the Kalman filter within the time propagation of either EnMfK or EnKF.

### 3.2. Marginalized Kalman filter for \( T_k|\theta \)

Before deriving the filtering algorithm for the joint state-parameter vector, we derive the marginalized Kalman filter for the conditional state \( T_k|\theta \) given the observations \( (z_{1:k}, y_{1:k}) \).

Assuming \( T_0 \) is Gaussian, the Kalman filter is sufficient to estimate the filtering distribution of \( T_k \) given \( \theta \). We derive the conditional mean and the conditional covariance for \( T_k|\theta \) given the observations \( (z_{1:k}, y_{1:k}) \):

\[
T_{k|k-1} = E[T_k|\theta, z_{1:k}, y_{1:k-1}] = E[A_0 T_{k-1} + B_0 u_k + w_k|\theta, z_{1:k}, y_{1:k-1}]
= A_0 E[T_{k-1}|\theta, z_{1:k-1}, y_{1:k-1}] + B_0 E[u_k|z_{1:k}]
= A_0 T_{k-1|k-1} + B_0 u_{k|k}.
\]

\[
P_{k|k-1} = \text{Cov}[T_k|\theta, z_{1:k}, y_{1:k-1}]
= \text{Cov}[A_0 T_{k-1} + B_0 u_k + w_k|\theta, z_{1:k}, y_{1:k-1}]
(\text{using the mutual independence of } T_{k-1}, u_k \text{ and } w_k)
= \text{Cov}[A_0 T_{k-1}|\theta, z_{1:k-1}, y_{1:k-1}] + \text{Cov}[B_0 u_k|\theta, z_{1:k}] + \text{Cov}[w_k]
= A_0 P_{k-1|k-1} A_0^T + B_0 P_{u|k} B_0^T + W_k.
\]

From (6) and (7), the marginalized Kalman filter for \( T_k|\theta \) is summarized by

1. **Prediction step:**
   \[
   T_{k|k-1} = A_0 T_{k-1|k-1} + B_0 u_{k|k},
   \]
   \[
   P_{k|k-1} = A_0 P_{k-1|k-1} A_0^T + B_0 P_{u|k} B_0^T + W_k,
   \]
   where \( T_{k|k-1} \) is the state prediction at time \( k \), given \( \theta \) and the observations \( (z_{1:k}, y_{1:k-1}) \), and \( P_{k|k-1} \) is the corresponding covariance matrix.

2. **Analysis step:**
   \[
   K_k = P_{k|k-1} H^T (HP_{k|k-1} H^T + V_k)^{-1},
   \]
   \[
   T_{k|k} = T_{k|k-1} + K_k (y_k - HT_{k|k-1}),
   \]
   \[
   P_{k|k} = (I - K_k H) P_{k|k-1},
   \]
   where \( T_{k|k} \) is the state estimation at time \( k \), given \( \theta \) and the observations \( (z_{1:k}, y_{1:k}) \), and \( K_k \) is the Kalman gain.
Remark.  The prediction step can be generalized under the Bayesian approach and corresponds to the following equation:

\[
p(T_k|\theta, z_{1:k}, y_{1:k-1}) = \int p(T_k|\theta, z_{1:k}, T_{k-1}) p(T_{k-1}|\theta, z_{1:k-1}, y_{1:k-1}) \, dT_{k-1}
\]

\[
= \int \left( \int p(T_k|\theta, u_k, T_{k-1}) p(u_k|z_{1:k}) \, du_k \right) p(T_{k-1}|\theta, z_{1:k-1}, y_{1:k-1}) \, dT_{k-1},
\]

where the marginalization of \(u_k\) appears explicitly. The analysis step is generalized under the Bayesian approach as

\[
p(T_k|\theta, z_{1:k}, y_{1:k}) = \frac{p(y_k|T_k) p(T_k|\theta, z_{1:k}, y_{1:k-1})}{\int p(y_k|T_k) p(T_k|\theta, z_{1:k}, y_{1:k-1}) \, dT_k}
\]

\[
\propto p(y_k|T_k) p(T_k|\theta, z_{1:k}, y_{1:k-1}).
\]

3.3. Marginalized Bayesian filtering for \(T_k\) and \(\theta\)

We now derive filtering algorithms to estimate the joint state-parameter vector, \(X_k = [\theta \ T_k]'\), given the observations \((z_{1:k}, y_{1:k})\) up to time \(k\). In this case, the evolution of the joint vector, \(X_k\), is nonlinear and approximate filtering algorithms are needed. We derive first the underlying Bayesian formulation for the filtering algorithms. The static parameters, \(\theta\), have a trivial evolution in the prediction step and therefore we neglect their time subscript.

1. Initialization:

\[
\theta \sim \rho_\theta(\theta),
\]

\[
T_0 \sim \rho_{T_0}(T_0),
\]

where \(\rho_\theta(\cdot)\) and \(\rho_{T_0}(\cdot)\) are the prior distributions of \(\theta\) and \(T_0\), respectively.

2. Prediction step:

\[
p(X_k|z_{1:k}, y_{1:k-1}) = p(T_k, \theta|z_{1:k}, y_{1:k-1})
\]

\[
= p(T_k|\theta, z_{1:k}, y_{1:k-1}) p(\theta|z_{1:k}, y_{1:k-1})
\]

\[
= \left( \int p(T_k|\theta, z_{1:k}, T_{k-1}) p(T_{k-1}|\theta, z_{1:k-1}, y_{1:k-1}) \, dT_{k-1} \right) p(\theta|z_{1:k-1}, y_{1:k-1})
\]

\[
= \int p(T_k|\theta, z_{1:k}, T_{k-1}) p(T_{k-1}|\theta, z_{1:k-1}, y_{1:k-1}) \, dT_{k-1}
\]

\[
= \int \left( \int p(T_k|\theta, u_k, T_{k-1}) p(u_k|z_{1:k}) \, du_k \right) p(X_{k-1}|z_{1:k-1}, y_{1:k-1}) \, dT_{k-1}.
\]

3. Analysis step:

\[
p(X_k|z_{1:k}, y_{1:k}) \propto p(y_k|X_k) p(X_k|z_{1:k}, y_{1:k-1})
\]

\[
= p(y_k|T_k) p(X_k|z_{1:k}, y_{1:k-1}).
\]
3.4. EnMKF for $\theta$ and $T_k$

Here, the main idea is to consider an ensemble version of the marginalized Kalman filter introduced in 3.2 to approximate the Bayesian filter 3.3. We initialize the algorithm by drawing independently $M$ samples from the prior distributions of the parameters and the initial state, $\{\theta_i\}^M_{i=1}$ and $\{T_0\}^M_{i=1}$. Then, in the prediction step, we evolve each state vector $T_i$ in time given the parameters $\theta_i$ to obtain the state vectors at time $t_i$. The measurements at time $t_i$ are then utilized in the analysis step to update both states and parameters. We iterate the prediction and analysis steps until we use all measurements and reach the final time point $t_N$.

The evolution of the joint vector, $X_k$, is nonlinear because $T_k$ depends on $\theta$ non-linearly. Therefore, we can not use Kalman filter directly. The ensemble Kalman framework is introduced to handle the nonlinear dependence on $\theta$, and it uses a Gaussian approximation to update the ensemble members in the analysis step. In our EnMKF algorithm, we exploit the conditional linearity of the joint vector to derive its prediction mean and covariance.

Following the stochastic EnKF scheme [17, 20], we assume a trivial evolution of the static parameters and denote the samples from $p(\theta|y_{1:k-1})$ by $\theta_i^{k-1}$. Given the ensemble size $M$, we compute the predicted state $T_i^{k|k-1} = E[T_k^i|\theta_i^{k-1}, z_{1:k}, y_{1:k-1}]$ for each $i = 1, \ldots, M$ as

$$
T_i^{k|k-1} = A\theta_i^{k-1} T_{i|k-1} + B\theta_i^{k-1} u_k^i.
$$

The prediction covariance matrix, $P_{i|k-1}$, is usually approximated by the sample covariance matrix of $X_{i|k-1}$:

$$
P_{i|k-1} = \frac{1}{M-1} \sum_{i=1}^M (X_{i|k-1} - \bar{X}_{i|k-1})(X_{i|k-1} - \bar{X}_{i|k-1})',
$$

where $X_{i|k-1} = \bar{X}_{i|k-1}$. This approximation is needed due to the nonlinearity. Instead, we derive the covariance matrix of $X_{i|k-1}$ using the law of total covariance by conditioning on $\theta_{i|k-1}$:

$$
P_{i|k-1} = Cov(X_i|z_{1:k}, y_{1:k-1})
= Cov\left(E[X_i|\theta_{i|k-1}, z_{1:k}, y_{1:k-1}]\right) + E\left[Cov(X_i|\theta_{i|k-1}, z_{1:k}, y_{1:k-1})\right]
= Cov\left(E\left[T_k^i|\theta_{i|k-1}, z_{1:k}, y_{1:k-1}\right]\right) + E\left[Cov\left(T_k^i|\theta_{i|k-1}, z_{1:k}, y_{1:k-1}\right)\right]
= Cov\left(E\left[T_k^i|\theta_{i|k-1}, z_{1:k}, y_{1:k-1}\right]\right) + E\left[Cov\left(T_k^i|\theta_{i|k-1}, z_{1:k}, y_{1:k-1}\right)\right]
= \frac{1}{M-1} \sum_{i=1}^M \left(\theta_{i|k-1} - \bar{\theta}_{k-1}\right)\left(\theta_{i|k-1} - \bar{\theta}_{k-1}\right)' + \frac{1}{M} \sum_{i=1}^M \left(\theta_{i|k-1} - \bar{\theta}_{k-1}\right)\left(T_{i|k-1} - T_{k-1}\right)\left(\theta_{i|k-1} - \bar{\theta}_{k-1}\right)'
$$

$$
+ \frac{1}{M} \sum_{i=1}^M \left(\theta_{i|k-1} - \bar{\theta}_{k-1}\right)A_{i|k-1} B_{i|k-1}' W_k + B_{i|k-1} P_{i|k-1} B_{i|k-1}' + W_k
$$

where $\bar{\theta}_{k-1} = \frac{1}{M} \sum_{i=1}^M \theta_{i|k-1}$ and $P_{i|k-1}$ must be computed using the Kalman update equations:

$$
P_{i|k} = (I - P_{i|k-1}H') (HP_{i|k-1}H + V_k)^{-1} H P_{i|k-1},
$$

$$
P_{i|k-1} = A_{i|k-1} P_{i|k-1} A_{i|k-1}' + B_{i|k-1} P_{i|k} B_{i|k-1}' + W_k
$$

Computing and storing these covariance matrices $P_{i|k}$ and $P_{i|k-1}$ for each $i = 1, \ldots, M$ may be burdensome, especially when $\dim(T_k)$ is large. Therefore, we consider the following approximation of the prediction covariance matrix:
$P_{k|k-1} \approx \frac{1}{M} \sum_{i=1}^{M} (X_{i}^j - \bar{X}_{i|k-1})(X_{i}^j - \bar{X}_{i|k-1})^T + \frac{1}{M} \sum_{i=1}^{M} \begin{bmatrix} 0 & 0 & \tilde{B}_{\theta_i|k-1} P_{k|k-1} \tilde{B}_{\theta_i|k-1}^T + W_k \end{bmatrix}$,

which requires the computation of $B_{\theta_i|k-1} P_{k|k-1} \tilde{B}_{\theta_i|k-1}^T$ for each $i = 1, \ldots, M$ in addition to the standard prediction covariance of EnKF $\rho_0$. The additional term in (11) can be interpreted as inflation with respect to $u_k$ and $w_k$. The full prediction covariance (10) provides an extra term in inflation with respect to $T_{k-1}$. However, it is computationally intractable except in very simple cases. The standard practice in EnKF is to consider the sample covariance (9) only. Our implementation (11) inflates the covariance in a tractable way. In the next section, we introduce another algorithm where we sample $u_k$; therefore, the inflation is not needed.

After applying the prediction step in (8) and (11), the analysis step for $X_k$ is computed by

$$K_k = P_{k|k-1} \mathcal{H}' \left( \mathcal{H} P_{k|k-1} \mathcal{H}' + V_k \right)^{-1}.$$  

$$X_{i|k} = X_{i|k-1} + K_k \left( y_k + v_k^i - \mathcal{H} X_{i|k-1} \right),$$

where $\mathcal{H} = \begin{bmatrix} 0 & H_{m \times (p+n)} \end{bmatrix}$ is the observation operator that maps the augmented vector $X_k$ to the corresponding observation space and $v_k^i \sim N(0, V_k)$ is used to perturb the observations.

A clear advantage of EnMKF is that it runs sequentially on time without the need to restart from the initial state. At a given time $t_k$, the static parameters are updated in the analysis step using measurements at time $t_k$. We summarize the complete algorithm of EnMKF in algorithm 1.

---

**Algorithm 1.** Ensemble-marginalized Kalman filter (EnMKF).

1: **draw** an initial ensemble $\theta_i', i = 1, \ldots, M$ from the prior distribution $\rho_0(\theta)$
2: **draw** independently an initial ensemble $T_{0i}', i = 1, \ldots, M$ from a normal prior distribution $N(T_0, P_0)$
3: for $k = 1$ to number of time observation $N$ do
4: run the Kalman filter for $u_k$ (equations (4) and (5)) to obtain $u_{ik}$ and $P_{ik}^u$
5: run the prediction step for $i = 1, \ldots, M$ 
   
   $T_{i|k-1}^c = A_{\theta_i|k-1} T_{i|k-1} + B_{\theta_i|k-1} u_{ik}$
6: compute the prediction covariance matrix $P_{ik|k-1}$ using (11)
7: run the analysis step for each $X_i, i = 1, \ldots, M$
   
   $K_k = P_{ik|k-1} \mathcal{H}' \left( \mathcal{H} P_{ik|k-1} \mathcal{H}' + V_k \right)^{-1}$
   
   $X_{i|k} = X_{i|k-1} + K_k \left( y_k + v_k^i - \mathcal{H} X_{i|k-1} \right), v_k^i \sim N(0, V_k)$
8: end for

---

In general, ensemble Kalman filters suffer from bias errors, especially with nonlinear systems. A common bias error is due to the Kalman formula that assumes Gaussianity. Nevertheless, the incorporation of the boundary temperature uncertainties will reduce the total bias error in the estimated parameters [2]. We show in section 4.7 that EnMKF admits smaller bias errors than the following EnKF algorithm.

### 3.5. Ensemble Kalman filter for $\theta$ and $T_k$

An alternative approach to solve our problem (2) and (3), is to apply an adequately modified EnKF to the joint vector $X_k = [\theta \ T_k]'$. In this modified EnKF, the control vector $u_k$ is
sampled from its distribution, which was determined during the prediction step. Thus, we sample \( u'_k \) from \( N(u_{k|k}, P_{k|k}) \) for each \( i = 1, \ldots, M \). Then, the prediction step differs from (8) and is given by

\[
T_{k|k-1}^i = A_{\theta_{k-1}} \bar{T}_{k-1|k-1}^i + B_{\theta_{k-1}} u'_k.
\]

In this case, the prediction covariance matrix is approximated by

\[
\frac{1}{M-1} \sum_{i=1}^{M} (X_{k|k-1}^i - X_{k|k-1})(X_{k|k-1}^i - X_{k|k-1})' + \begin{bmatrix} 0 & 0 \\ 0 & W_k \end{bmatrix}.
\] (14)

The analysis step is the same as the one that we developed for EnMKF, which is summarized by equations (12) and (13).

Both EnMKF and EnKF are Monte Carlo implementations of Bayesian filtering. The main difference appears in the prediction step, especially in the integration with respect to \( u_k \). In EnMKF, we use the fact that \( u_k | z_{1:k} \) and \( T_k | \theta, u_k, T_{k-1} \) have Gaussian distributions and, therefore, the resulting distribution is Gaussian with mean (6) and covariance (7). In EnKF 3.5, the same integral is approximated with the Monte Carlo method by sampling from \( p(u_k | z_{1:k}) \).

The modified EnKF approach avoids the computation of the matrices \( B_{\theta_{k-1}} P_{k|k} B_{\theta_{k-1}}' \) that are needed in formula (11) of EnMKF. However, the additional term in EnMKF can be considered inflation induced from the law of total covariance by conditioning on the unknown parameters \( \theta \). Similarly, the joint distribution of the state and parameters can be split into the distribution of the conditional state and the marginal distribution of the parameters [31]. In our case, the linear evolution of the state and control vectors simplifies the conditional distribution to the conditional mean (6) and the conditional covariance (7).

4. Real-world application: heat transfer across a solid brick wall

In this section, we show how to formulate and implement the EnMKF method to deal with the state and parameter estimation problem in a real-world heat-transfer application, which is summarized in section 4.1. Next, we compare the results obtained with the EnMKF algorithm 1 to those obtained with the modified EnKF method from section 3.5 using experimental and synthetic data. We also propose stopping criteria that indicate when the ensemble parameters are stationary with adequate non-zero variance.

We consider the problem of describing the thermal performance of a solid wall given in situ measurements. This can be posed as a state and parameter estimation problem constrained by the heat equation. For simplicity, we consider the case of a single-layer wall with homogeneous density, thermal conductivity, and specific heat capacitance. Therefore, we characterize the thermophysical properties of the wall in terms of its heat capacity per unit area (J m\(^{-2}\) K) and thermal resistance (m\(^2\) K W\(^{-1}\)). These are denoted by \( \rho C \) and \( R \), respectively. Under the standard assumption of uni-directional heat flux across the wall’s thickness, the forward model is described by the internal temperature profile of the wall, denoted by \( T(x, t) \). This profile is the solution to the 1D heat equation [32]:

\[
\begin{align*}
\frac{\rho C}{L} \frac{\partial T}{\partial t} &= \frac{\partial^2 T}{\partial x^2}, \quad x \in (0, L), \quad t \in [0, t_N] \\
T(0, t) &= T_{in}(t), \quad t \in [0, t_N] \\
T(L, t) &= T_{ext}(t), \quad t \in [0, t_N] \\
T(x, 0) &= T_0(x), \quad x \in (0, L),
\end{align*}
\] (15)
where $L$ is the wall thickness, $T_{\text{int}}$ and $T_{\text{ext}}$ are the internal and external wall surface temperatures, respectively, and $T_0(x)$, is the initial temperature of the wall. The goal is to estimate the thermophysical properties $\theta = (R, \rho C)$ given boundary temperature measurements $T_{\text{int}}$ and $T_{\text{ext}}$ as well as measurements of the internal and external boundary heat flux defined via:

$$
\begin{align}
F_{\text{int}}(t) &= -\frac{L}{R} \frac{\partial T}{\partial x} \big|_{x=0}, \\
F_{\text{ext}}(t) &= -\frac{L}{R} \frac{\partial T}{\partial x} \big|_{x=L}.
\end{align}
$$

(16)

An experimental setting consistent with (15) and (16) can be designed by placing temperature sensors and heat flux meters on the internal and external surfaces of the wall [2, 10]. These quantities are then monitored for a period of time which often consist of at least two weeks. In order to reduce the effect of solar radiation (not considered in (15)), measurement campaigns are usually conducted on north-facing walls during the winter [33]. For simplicity, we also assume that the initial condition, $T_0(x)$, is well-approximated by the piecewise linear function

$$
\begin{align}
T_{\text{int}}(0) + 2 \frac{x}{L} (\tau_0 - T_{\text{int}}(0)) & \quad \text{if } 0 < x < \frac{L}{2} \\
\tau_0 + 2 \frac{T_{\text{int}}(0) - \tau_0}{\frac{L}{2}} (x - \frac{L}{2}) & \quad \text{if } \frac{L}{2} < x < L,
\end{align}
$$

(17)

where we assume $\tau_0 = 16.1 \, ^\circ \text{C}$, according to the estimate we obtained in our previous study [2].

To formulate the discrete state observation system, we consider the following uniform space-time discretization:

$x_0 = 0, x_1 = \Delta x, \ldots, x_i = i\Delta x, \ldots, x_n = n\Delta x = L,$

$t_0 = 0, t_1 = \Delta t, \ldots, t_i = j\Delta t, \ldots, t_N = N\Delta t,$

and denote by $T_k = (T(x, \Delta x), \ldots, T((n-1)\Delta x, k\Delta t)).$ $T_{\text{int}, k} = T_{\text{int}}(k\Delta t)$ and $T_{\text{ext}, k} = T_{\text{ext}}(k\Delta t)$ the inner, internal and external surface temperatures at time $t_k$, respectively. The inner temperature vector $T_k$ can be written as a linear function of $T_{\text{int}, k}$ and $T_{\text{ext}, k}$ as follows [2]:

$$
\begin{pmatrix}
\tilde{T}_k \\
\end{pmatrix}
= \begin{pmatrix}
\tilde{A}_\theta \\
\tilde{B}_\theta,\text{int} \\
\tilde{B}_\theta,\text{ext}
\end{pmatrix}
\begin{pmatrix}
T_{\text{int}, k} \\
T_{\text{ext}, k}
\end{pmatrix},
$$

where the matrices $\tilde{A}_\theta, \tilde{B}_\theta,\text{int}$, and $\tilde{B}_\theta,\text{ext}$ are specified exactly (see [2]).

We let $T_k = [T_{\text{int}, k} \quad T_{\text{ext}, k}]^t$ be the temperature vector on the closed interval $[0, L]$ and

$$
\begin{pmatrix}
0 \\
\tilde{A}_\theta \\
0
\end{pmatrix}, \quad 
\begin{pmatrix}
1 \\
\tilde{B}_\theta,\text{int} \\
0
\end{pmatrix}, \quad 
\begin{pmatrix}
0 \\
\tilde{B}_\theta,\text{ext} \\
1
\end{pmatrix}.
$$

Then, the time evolution equation of the state vector $T_k$ is given by

$$
T_k = A\theta T_{k-1} + B\theta,\text{int} T_{\text{int}, k} + B\theta,\text{ext} T_{\text{ext}, k}.
$$

(18)

The heat-flux vector $F_{k} = [F_{\text{int}, k} \quad F_{\text{ext}, k}]^t$ is approximated by the first difference

$$
\begin{align}
F_{\text{int}, k} &= \frac{1}{R} \frac{1}{2\Delta t} (3T_{\text{int}, k} - 4T_{1, k} + T_{2, k}) \\
F_{\text{ext}, k} &= \frac{1}{R} \frac{1}{2\Delta t} (3T_{\text{ext}, k} - 4T_{n-1, k} + T_{n-2, k}).
\end{align}
$$

Therefore, the observation matrix operator takes the form...
\[ H = \frac{1}{2\Delta x} \begin{bmatrix} 3 & -4 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & -4 & 3 \end{bmatrix} \]

and the observation equation is

\[ y_k = \frac{1}{R} HT_k + v_k, \quad (19) \]

where \( v_k \sim N(0, V_k) \) is the measurements noise and \( H = [0 \quad H] \) is the observation operator that maps the augmented vector \( \bar{X}_k \) to the observation space.

We define the observation matrix operator without \( R \) to avoid dependence of the Kalman gain on this unknown parameter. Instead, we keep the Kalman gain fixed in each iteration of each layer. Each of these properties can then be included in the unknown vector of parameters described by suitable modifications to (15) that incorporate the thermophysical properties on thermophysical properties of a much wider class of walls. Multilayer walls, for example, can be described by suitable modifications to (15) that incorporate the thermophysical properties on each layer. Each of these properties can then be included in the unknown vector of parameters \( \theta \) that is inferred within the EnMKF.

**Algorithm 2.** EnMKF algorithm to estimate \( R, \rho C, F_{\text{int}} \) and \( F_{\text{ext}} \).

1: draw an initial ensemble \( \Theta^i = (R^i, \rho C^i), i = 1, \ldots, M \) from a prior distribution \( \rho_0(\Theta) \)
2: approximate the initial condition \( T_0 \) using (17)
3: draw an initial ensemble \( T_{0i}, i = 1, \ldots, M \) from the normal prior distribution \( N(T_0, P_0) \)
4: define the augmented ensemble \( X^i = [\log(R^i) \quad \log(\rho C^i) \quad T^i]' \), \( i = 1, \ldots, M \)
5: for \( k = 1 \) to number of time observation \( N \) do
6: run the Kalman filter for \( T_{\text{int}, k} \) and \( T_{\text{ext}, k} \) to obtain \( T_{\text{int}, k|k}, T_{\text{ext}, k|k}, P_{\text{int}, k|k} \) and \( P_{\text{ext}, k|k} \)
7: run the prediction step for \( i = 1, \ldots, M \)
   \[ T_{i,k-1}^j \approx X_{i,k-1}^j + B_{\theta_{i-1}, \text{int}} T_{\text{int}, k|k} + B_{\theta_{i-1}, \text{ext}} T_{\text{ext}, k|k} \]
8: compute the prediction covariance matrix
   \[ P_{i,k-1} \approx \frac{1}{M} \sum_{i=1}^{M} (X_{i,k-1}^j - X_{k-1}^j)(X_{i,k-1}^j - X_{k-1}^j)' \]
9: run the analysis step for each \( X^i, i = 1, \ldots, M \)
   \[ K_k = P_{i,k-1}^{-1} (\bar{H} P_{i,k-1}^{-1} \bar{H}' + V_k)^{-1} \]
   \[ X_{i,k} = X_{i,k-1}^j + K_k (R^j y_k + v_k) - \bar{H} X_{k-1}^j \]
   \[ v_{i,k} \sim N(0, V_k) \]
10: estimate the heat flux and its sample mean and sample covariance
    \[ F_k = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (\bar{F}_i - \bar{F}_k)'(\bar{F}_i - \bar{F}_k) \]
    \[ \bar{F}_k = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} F_i \]
11: end for
4.1. Experimental data set

We use real data from an experiment conducted inside an environmental chamber in the Energy Technologies Building at Nottingham University Innovation Park. The chamber was divided into two rooms by a 215 mm thick partition wall. The aim of the experiment was to estimate the thermal properties of a brick section of the partition wall. The temperature in both rooms was controlled to resemble internal conditions in Room 1 and external conditions in Room 2. The wall boundary temperature and heat-flux measurements were recorded every minute by sensors placed on the surface of the brick wall [2].

Figure 1 shows the time series of the temperature and heat flux measurements correspond­ing to the two sides of the brick wall, identified by internal and external measurements. The measurements contain noise that we can analyze, in a non-sequential framework, by using a smoothing spline method to fit a curve to each time series. Then, we approximate the noise by the difference between the measurements and the smooth curve values. The noise variance is estimated to be around 0.01 for the temperature measurements. The noise variances for the internal and external heat flux measurements are estimated to be 20 and 5, respectively. Therefore, the covariance matrix $V_k$ is approximated by

$$V_k = \begin{bmatrix} 20 & 0 \\ 0 & 5 \end{bmatrix}.$$ 

Alternatively, we can approximate $V_k$ by using prior knowledge on the accuracy of the measuring devices in different situations or by assuming a prior model with unknown hyperparameters and estimate them sequentially along with the quantities of interest. Further statistical analyses and details about these experimental data are provided in [2].

4.2. Kalman filtering for $u_k$

The first step in implementing EnMKF is to run a Kalman filter for the control vector $u_k$ in (3), which in this case, consists of the boundary conditions $T_{\text{int},k}$ and $T_{\text{ext},k}$. The forward-evolution operator for the boundary conditions is unknown but noisy measurements are available. We consider the two autoregressive models $AR(1)$ and $AR(2)$ from section 3.1 and use them with the measurements to apply the Kalman filter. Figures 2 and 3 show the results from the Kalman filter with each model and the real temperature measurements. In the experiments with EnMKF and EnKF presented below, we use only the results from the Kalman filter with $AR(1)$.

We also estimate the variance for the boundary conditions $T_{\text{int},k}$ and $T_{\text{ext},k}$. For example, figure 4 shows the estimated mean and confidence bands for the internal temperature measurements using Kalman filter with AR(1). The estimated variances capture the variability of the data and therefore we can marginalize the uncertainty of the boundary conditions accurately.

4.3. EnMKF results

We run EnMKF algorithm 2 with $M = 100$ and using the complete experimental data set, $N = 6900$. We initially sample $R$ and $\rho C$ from the uniform priors, $U(0.17, 0.36)$ and $U(234000, 431000)$, and sample $T_0$ from the normal prior distribution, $N(T_0, 0.01I_n)$.

Figure 5 shows the convergence of the ensemble parameters $R$ and $\rho C$ with respect to time as more observations are incorporated. The thermal-resistance mean converges to $0.31 \text{ m}^2 \text{ K/W}$ and the mean of the heat capacity per unit area converges to $3.11 \times 10^5 \text{ J m}^{-2} \text{ K}$. As we increase the ensemble size, we obtain results that are more accurate and consistent.
with the MCMC results in [2]. We further analyze the convergence of EnMKF with respect to the ensemble size $M$ and compare it with the convergence of the modified EnKF in section 4.5.

We also compute the heat flux from the estimated temperature and parameters by EnMKF. Figures 6 and 7 show that the estimated heat fluxes match the real measurements suitably. The corresponding variances for the estimated heat fluxes converge to be smaller than 1. We recall that the heat flux measurements are very noisy and it would not be optimal to recover

\[\text{Figure 1.} \text{ Raw temperature and heat-flux measurements. Temperature in room 2 imitates outdoor weather conditions.}\]

\[\text{Figure 2.} \text{ Internal temperature measurements (red dots) and Kalman filter mean with AR(1) and AR(2).}\]
the measurements variance through the variance of the estimates. Figures 6 and 7 indicate that our estimates are unbiased and therefore smaller variances are preferable.

4.4. EnKF results

We apply the modified EnKF approach introduced in section 3.5, with ensemble size $M = 100$, to infer the parameters $R$ and $\rho C$ as in the previous example. Figure 8 shows that the ensemble of $R$ and $\rho C$ collapse at time $t = 1734$ minutes. After this time, the parameters remain fixed at
the biased estimates \( R = 0.3 \text{ m}^2 \text{K/W} \) and \( \rho C = 3.42 \times 10^5 \text{ J/m}^2 \text{K} \). Thus, the estimated heat fluxes do not match the measurements. In figures 9 and 10, we show only the sample mean of the heat-flux ensemble because the variances of the internal and external heat fluxes are approximately 390 and 130, respectively. These values are much larger than the estimated variances produced by the EnMKF approach.
We can gradually increase the ensemble size $M$ to mitigate the ensemble collapse and reduce the bias error in the estimated parameters. Figures 11–13 show the estimated parameters and heat-flux mean when $M = 300$. Although the bias error of the estimated parameters is reduced, the corresponding variances still vanish as time increases, which is not a realistic feature. While the internal and external heat fluxes variances are reduced to 375 and 124, respectively, they are still very large values when compared with the estimated variances produced by EnMKF.

**Figure 7.** Estimated heat flux ensemble (solid lines) using EnMKF with ensemble size $M = 100$ compared with real raw measurements (red dots) in Room 2.

**Figure 8.** Propogation of the thermal resistance $R$ (left) and heat capacity $\rho C$ (right) ensemble with respect to time using EnKF with ensemble size $M = 100$. 

We can gradually increase the ensemble size $M$ to mitigate the ensemble collapse and reduce the bias error in the estimated parameters. Figures 11–13 show the estimated parameters and heat-flux mean when $M = 300$. Although the bias error of the estimated parameters is reduced, the corresponding variances still vanish as time increases, which is not a realistic feature. While the internal and external heat fluxes variances are reduced to 375 and 124, respectively, they are still very large values when compared with the estimated variances produced by EnMKF.
Finally, we compare the EnMKF method and the modified EnKF method by analyzing the convergence of the estimated parameters at a fixed time with respect to the ensemble size $M$.

Figures 14–16 show the convergence of the parameters-ensemble mean and standard deviation with ensemble size $M$ at time $t' = 3000$ minutes. The ensemble mean converges linearly for both methods but EnMKF has a better error constant and, therefore, is more reliable for small ensemble sizes. Moreover, due to the collapse of the ensemble, the convergence of the ensemble standard deviation for EnKF occurs only after $M > 1000$.

4.5. Comparing EnMKF and EnKF

Finally, we compare the EnMKF method and the modified EnKF method by analyzing the convergence of the estimated parameters at a fixed time with respect to the ensemble size $M$. Figures 14–16 show the convergence of the parameters-ensemble mean and standard deviation with ensemble size $M$ at time $t' = 3000$ minutes. The ensemble mean converges linearly for both methods but EnMKF has a better error constant and, therefore, is more reliable for small ensemble sizes. Moreover, due to the collapse of the ensemble, the convergence of the ensemble standard deviation for EnKF occurs only after $M > 1000$. 
4.6. Stopping criteria

EnMKF solves the collapse problem that usually arises with EnKF algorithms for state-parameter estimation. Ideally, we want the ensemble of each parameter to converge with time to become samples from the posterior distribution. Assuming this posterior distribution is unimodal, the convergence of our algorithm can be confirmed by computing the ensemble mean.

Figure 11. Propagation of the thermal resistance $R$ (left) and heat capacity $\rho C$ (right) ensemble with respect to time using EnKF with ensemble size $M = 300$.

Figure 12. Estimated heat-flux mean using EnKF with ensemble size $M = 300$ compared with real raw data measurements in room 1.
and standard deviation for each parameter. The time convergence of these estimated values can be used to suggest stopping criteria for ongoing measurement campaigns. For example, when the difference between two consecutive means and two consecutive standard deviations are relatively small, then the measurement campaign should be stopped.

In figures 17 and 18, we see that the ensemble mean and standard deviation for $R$ and $\rho C$ stabilize after 5000 minutes. This indicates that the measurements collected up to that time are sufficient and the experiment can be stopped.
We treat the boundary conditions as random variables that are marginalized in EnMKF and sampled in EnKF. In both EnKF and EnMKF, the bias error of the estimated parameters and the variability of the estimated state are reduced by increasing the ensemble size. However, in EnKF, the bias error induced by the boundary conditions depends on the ensemble size, $M$; with EnMKF, that bias error is eliminated independently from $M$. Therefore, EnMKF reduces the total bias error in comparison with the EnKF. To strengthen this conclusion, we use a synthetic data set similar to the experimental data set from Nottingham University (figure 1). The synthetic data are generated by assuming $R = 0.3106$ m$^2$ K/W, $\rho C = 3.2 \times 10^5$ J m$^{-2}$ K.

**Figure 15.** Convergence of the ensemble mean error of the estimated parameters $R$ (left) and $\rho C$ (right) at time $t' = 3000$ with respect to the ensemble size $M$.

**Figure 16.** Convergence of the ensemble standard deviation of the estimated parameters $R$ (left) and $\rho C$ (right) at time $t' = 3000$ with respect to the ensemble size $M$.

### 4.7. Bias error analysis

We treat the boundary conditions as random variables that are marginalized in EnMKF and sampled in EnKF. In both EnKF and EnMKF, the bias error of the estimated parameters and the variability of the estimated state are reduced by increasing the ensemble size. However, in EnKF, the bias error induced by the boundary conditions depends on the ensemble size, $M$; with EnMKF, that bias error is eliminated independently from $M$. Therefore, EnMKF reduces the total bias error in comparison with the EnKF. To strengthen this conclusion, we use a synthetic data set similar to the experimental data set from Nottingham University (figure 1). The synthetic data are generated by assuming $R = 0.3106$ m$^2$ K/W, $\rho C = 3.2 \times 10^5$ J m$^{-2}$ K,
smooth boundary conditions, and the initial condition given in (17). We solve the heat equation and compute the heat flux at the boundaries for each minute. The temperature and heat-flux time series are then perturbed by Gaussian white noise. The resulting synthetic data are presented in figure 19.

We apply the Kalman filter to the perturbed temperature series and obtain the filtered mean temperatures and their corresponding variances. Then, we apply EnMKF and EnKF to estimate the parameters $R$ and $\rho C$ using the uniform priors:

$$R \sim U(0.28, 0.36), \quad \rho C \sim U(301\,000, 376\,000).$$

Figure 17. Time convergence for the thermal resistance mean (left) and heat capacity mean (right) using EnMKF with ensemble size $M = 100$.

Figure 18. Time convergence for the thermal resistance standard deviation (left) and heat capacity standard deviation (right) using EnMKF with ensemble size $M = 100$. 

Figure 20 shows the convergence of the sample mean of the estimated parameters using EnMKF and EnKF at time $t' = 2000$ with respect to the ensemble size. The total bias error obtained by EnMKF is smaller and more stable than the error produced by the modified EnKF. For large ensemble sizes, both methods provide similar results.

In this case, the synthetic data have Gaussian noise. In the previous sections, the real noise is not Gaussian. Nevertheless, the results show that the influence of Gaussianity assumption is not very significant when the sequential framework is applied to this heat-transfer problem.
5. Summary and conclusions

We studied the filtering problem for a state observation system with random control vector and unknown static parameters. The control vectors may represent the boundary conditions or the source function of the underlying PDE model. By creating an artificial linear dynamic model for the control vector, we applied Kalman filter to obtain a Gaussian distribution of the random control vectors at each time step. The resulting distributions were then incorporated to generate a marginalized Kalman filter for the conditional state. The ensemble-marginalized Kalman filter (EnMKF) was derived using conditional probability, in particular, the law of total covariance. The main feature of EnMKF appeared in the prediction covariance which is approximated by the sample covariance of the augmented vector inflated by $E_\Theta [\text{Cov}(B_\theta u_k)]$.

Another sample-based EnKF algorithm was also proposed where the prediction covariance is only approximated by the sample covariance.

To deal with high-dimensional joint state-parameter estimation problems, fully Bayesian techniques such as particle filters are not computationally affordable. The two introduced algorithms are not fully Bayesian; thus, they may handle nonlinear state-parameter systems and high-dimensional problems. EnKF usually requires additional adjustments in order to provide better approximations to the desired posterior distribution. EnMKF, on the other hand, completely avoids the ensemble-collapse phenomenon.

We applied both algorithms to estimate the thermal properties of a solid brick wall from experimental data. We showed that EnMKF provides better results than the modified EnKF. Due to the marginalization technique in EnMKF, we prevented the collapse of the parameters ensemble without artificial inflation. EnMKF had other advantages as well, such as a reduced heat-flux variance and more accurate fit of the measurements. We also considered a synthetic data example and compared the bias error from the parameters estimated using EnMKF and EnKF. The results confirmed that EnMKF provides a smaller bias error when the ensemble size is not very large. Finally, we proposed stopping criteria to reduce the time and cost required for real experiments, based on the estimated parameters reaching a stationary period.

While the numerical results are focused on the heat-transfer application, the EnMKF can be used for a wider PDE-constrained data assimilation problems in which the boundary conditions are subject to uncertainty; these include, for example, flood predictions via hydrological models [34].

Acknowledgments

Z Sawlan, M Scavino and R Tempone are members of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering. R Tempone received support from the KAUST CRG3 Award Ref: 2281 and the KAUST CRG4 Award Ref: 2584.

ORCID iDs

Marco Iglesias https://orcid.org/0000-0002-8952-717X
Zaid Sawlan https://orcid.org/0000-0002-4361-2491

References

[1] Ruggeri F, Sawlan Z, Scavino M and Tempone R 2017 A hierarchical Bayesian setting for an inverse problem in linear parabolic PDEs with noisy boundary conditions Bayesian Anal. 12 407–33
Iglesias M, Sawlan Z, Scavino M, Tempone R and Wood C 2018 Bayesian inferences of the thermal properties of a wall using temperature and heat flux measurements Int. J. Heat Mass Transfer 116 417–31
[3] Li F G et al 2015 Solid-wall U-values; heat flux measurements compared with standard assumptions Build. Res. Inf. 43 238–52
[4] Cesaratto P G and De Carli M 2013 A measuring campaign of thermal conductance in situ and possible impacts on net energy demand in buildings Energy Build. 59 29–36
[5] Doran S Detr framework project report: field investigations of the thermal performance of construction elements as built, Building Research Establishment Client Report (78132)
[6] Asdrubali F, D’Alessandro F, Baldinelli G and Bianchi F 2014 Evaluating in situ thermal transmittance of green buildings masonries—a case study Case Stud. Constr. Mater. 1 53–9
[7] Hong S H, Oreszczyn T and Ridley I 2006 The impact of energy efficient refurbishment on the space heating fuel consumption in English dwellings Energy Build. 38 1171–81
[8] Biddulph P, Gori V, Elwell C A, Scott C, Rye C, Lowe R and Oreszczyn T 2014 Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements Energy Build. 78 10–6
[9] Gori V, Marincioni V, Biddulph P and Elwell C A 2017 Inferring the thermal resistance and effective thermal mass distribution of a wall from in situ measurements to characterise heat transfer at both the interior and exterior surfaces Energy Build. 135 398–409
[10] De Simon L, Iglesias M A, Jones B and Wood C 2017 Quantifying uncertainty in thermal properties of walls by means of Bayesian inversion (submitted) (arXiv:1710.02976)
[11] Kaipio J and Somersalo E 2005 Statistical and Computational Inverse Problems (Berlin: Springer)
[12] Doucet A, Godsill S and Andrieu C 2000 On sequential Monte Carlo sampling methods for Bayesian filtering Stat. Comput. 10 197–208
[13] Asch M, Bocquet M and Nodet M 2016 Data Assimilation: Methods, Algorithms, and Applications (Philadelphia, PA: SIAM)
[14] Bickel P, Li B and Bengtsson T 2008 Sharp failure rates for the bootstrap particle filter in high dimensions Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K Ghosh (Bethesda, MD: Institute of Mathematical Statistics Collections) (https://doi.org/10.1214/074921708000000228)
[15] Evensen G 2003 The ensemble Kalman filter: theoretical formulation and practical implementation Ocean Dyn. 53 343–67
[16] Burgers G, Jan van Leeuwen P and Evensen G 1998 Analysis scheme in the ensemble Kalman filter Mon. Weather Rev. 126 1719–24
[17] Evensen G 2009 The ensemble Kalman filter for combined state and parameter estimation IEEE Control Syst. 29 83–104
[18] Law K, Stuart A and Zygalakis K 2015 Data Assimilation: a Mathematical Introduction vol 62 (Berlin: Springer)
[19] Kwiatkowski E and Mandel J 2015 Convergence of the square root ensemble Kalman filter in the large ensemble limit SIAM/ASA J. Uncertain. Quantif. 3 1–17
[20] Katzfuss M, Stroud J R and Wikle C K 2016 Understanding the ensemble Kalman filter Am. Stat. 70 350–7
[21] Iglesias M A, Law K J and Stuart A M 2013 Ensemble Kalman methods for inverse problems Inverse Problems 29 045001
[22] Li H, Kalnay E and Miyoshi T 2009 Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter Q. J. R. Meteorol. Soc. 135 523–33
[23] Anderson J L 2007 An adaptive covariance inflation error correction algorithm for ensemble filters Tellus A 59 210–24.
[24] Iglesias M, Park M and Tretyakov M 2017 Bayesian inversion in resin transfer molding (submitted) (arXiv:1707.03575)
[25] Iglesias M A 2016 A regularizing iterative ensemble Kalman method for PDE-constrained inverse problems Inverse Problems 32 025002
[26] Schon T, Gustafsson F and Nordlund P-J 2005 Marginalized particle filters for mixed linear/nonlinear state-space models IEEE Trans. Signal Process. 53 2279–89
[27] Doucet A, De Freitas N, Murphy K and Russell S 2000 Rao-Blackwellised particle filtering for dynamic Bayesian networks Proc. of the 16th Conf. on Uncertainty in artificial intelligence, (Morgan Kaufmann Publishers Inc.) pp 176–83

[28] Šmidl V and Hofman R 2011 Marginalized particle filtering framework for tuning of ensemble filters Mon. Weather Rev. 139 3589–99

[29] Kalman R E 1960 A new approach to linear filtering and prediction problems, J. Basic Eng. 82 35–45

[30] Pagano M 1974 Estimation of models of autoregressive signal plus white noise Ann. Stat. 2 99–108

[31] Stroud J R, Katzfuss M and Wikle C K 2018 A Bayesian adaptive ensemble Kalman filter for sequential state and parameter estimation Mon. Wea. Rev. 146 373–86

[32] Yunus C and Afshin J 2011 Heat and Mass Transfer: Fundamentals and Applications (New Delhi: Tata McGraw-Hill)

[33] ISO 9869-1:2014 2014 Thermal insulation—building elements—in situ thermal resistance and thermal transmittance Technical Report The British Standards Institution

[34] Pappenberger F, Matgen P, Beven K J, Henry J-B, Pfister L, Fraiport P 2006 Influence of uncertain boundary conditions and model structure on flood inundation predictions Adv. Water Resour. 29 1430–49