POSTERIOR CONTRACTION RATES FOR NON-PARAMETRIC STATE AND DRIFT ESTIMATION

SEBASTIAN REICH* AND PAUL J. ROZDEBA

Institute of Mathematics, University of Potsdam
Karl-Liebknecht-Str. 24/25
D-14476 Potsdam, Germany

(Communicated by Colin John Cotter)

ABSTRACT. We consider a combined state and drift estimation problem for the linear stochastic heat equation. The infinite-dimensional Bayesian inference problem is formulated in terms of the Kalman–Bucy filter over an extended state space, and its long-time asymptotic properties are studied. Asymptotic posterior contraction rates in the unknown drift function are the main contribution of this paper. Such rates have been studied before for stationary non-parametric Bayesian inverse problems, and here we demonstrate the consistency of our time-dependent formulation with these previous results building upon scale separation and a slow manifold approximation.

1. Introduction. In this paper, we consider the combined state and drift estimation problem for the stochastic heat equation

\[ du_t = \partial_x^2 u_t dt + f^* dt + \gamma^{1/2} dw_t \]  

over the domain \( \Omega = (0, \pi) \) and for \( t \geq 0 \), with zero Dirichlet boundary conditions \( u_t(0) = u_t(\pi) = 0 \) and zero initial condition \( u_0(x) = 0 \) for all \( x \in \Omega \), without loss of generality. We denote by \( w_t \) a cylindrical Wiener process where \( \gamma > 0 \) characterises the strength of the model error. The unknown drift function \( f^* \) is assumed to belong to a Sobolev space of appropriate regularity. We specify the properties of \( w_t \) and \( f^* \) more precisely in Section 2 below.

Linear stochastic partial differential equations (SPDEs) of the form (1) are well understood from analytical and numerical perspectives. See, for example, [5, 10, 14]. In this paper, we specifically focus on the estimation of the states \( u_t, t > 0 \), as well as the unknown drift function \( f^* \) from noisy state measurements given by

\[ dy_t = u_t dt + \rho^{1/2} dv_t, \]  

where \( y_0(x) = 0 \) for all \( x \in \Omega \), \( \rho > 0 \) denotes the strength of the measurement noise, and \( v_t \) is another cylindrical Wiener process independent of the model error \( w_t \).

Our focus is on asymptotic posterior contraction rates, which have been widely studied in the context of non-parametric stationary inverse problems. See, for example, [13, 8, 7]. We note that pure parameter estimation problems for SPDEs given
exact observations of the states are also well-studied. See, for example, [3] for a recent survey. First steps towards non-parametric time-dependent inverse problems have been taken in [20]. However, state estimation for a given drift function \( f^* \) is framed as an inference problem over a fixed time interval resulting in a smoothing problem, whereas filtering problems are the focus of this paper. Conversely, estimation of a drift function \( f^* \) is analysed under the assumption of perfectly observed model states. Similarly, parameter estimation for a class of linear SPDEs from noiseless local state measurements has been considered in [1], whereas we assume noisy observations of the full states throughout this paper.

In this paper, the infinite-dimensional, time-dependent combined state and drift estimation problem is formulated and analysed in terms of the well-known Kalman–Bucy filter equations for the posterior mean and covariance operator in an augmented state space. See, for example, [11, 17, 4] for an introduction to the Kalman–Bucy filter in finite and infinite dimensions. More recently, the Kalman–Bucy filter has been reformulated as a set of mean-field equations termed the ensemble Kalman–Bucy filter. See, for example [2, 6, 15]. These mean-field equations allow for a concise formulation of our time-dependent estimation problem. Furthermore, although not considered in this paper, the Kalman–Bucy mean-field equations can be generalised to non-Gaussian estimation problems [21, 18]. A key observation arising from the analysis of the Kalman–Bucy mean-field equations is the time-scale separation in the dynamics of the state and parameter variables, allowing us to apply the concept of slow manifolds [19].

The remainder of this paper is structured as follows. A mathematical formulation of the combined state and drift estimation problem in terms of Fourier modes is provided in Section 2. Furthermore, the time-dependent estimation problem is formulated in terms of Kalman–Bucy mean-field equations. This formulation is applied to a stationary drift estimation problem in Section 3. Asymptotic posterior contraction rates are then derived and compared to existing results for a closely related non-parametric Bayesian inference problem. These results are extended to the combined state and drift estimation problem in Section 4. The essential analysis of the single-mode Kalman–Bucy filter equations is carried out in Section 4.1. A numerical exploration of the combined state and parameter estimation problem is carried out in Section 5. The paper concludes with Section 6.

2. Mathematical problem formulation. In this paper, we analyse the state and drift estimation problem using a spectral representation in terms of Fourier modes. We provide the essential background in this section. It is well known that solutions of (1) can be expanded in Fourier modes \( \{U_t(k)\}_{k \geq 1} \), that is,

\[
u_t(x) = \sum_{k \geq 1} U_t(k) \sin(kx). \tag{3}\]

Because of (1), the Fourier modes obey the stochastic differential equations (SDEs)

\[
dU_t(k) = -k^2 U_t(k) dt + F^*(k) dt + \gamma^{1/2} dW_t(k), \quad k \geq 1, \tag{4}\]

where

\[
w_t(x) = \sum_{k \geq 1} W_t(k) \sin(kx) \tag{5}\]

denotes the Fourier representation of the cylindrical Wiener process \( w_t \). Here \( F^*(k) \), \( k \geq 1 \), denote the Fourier coefficients of \( f^* \), and \( W_t(k) \) independent standard Brownian motions. The initial conditions are \( U_0(k) = 0 \) for all \( k \geq 1 \). Therefore, (1)
can be viewed as a stochastic evolution equation on the separable Hilbert space of square integrable functions \( u : \Omega \rightarrow \mathbb{R} \) of sufficient spatial regularity with zero boundary conditions.

The measurement model (2) can also be transformed into Fourier space, yielding
\[
dY_t(k) = U_t(k)dt + \rho^{1/2}dV_t(k), \quad k \geq 1,
\]
with \( V_t(k) \) representing independent standard Brownian motions, independent of the model errors \( W_t(k') \) for all \( k' \), that is,
\[
v_t(x) = \sum_{k \geq 1} V_t(k) \sin(kx).
\]
Please also note that measurement processes defined by (2) and (6), respectively, are related by
\[
y_t(x) = \sum_{k \geq 1} Y_t(k) \sin(kx).
\]

Throughout this paper, we will exclusively work with the formulation (4) and (6) of our state and drift estimation problem in Fourier space. We now proceed to formulate a continuous-time Bayesian inference framework for this problem, which is used to derive asymptotic posterior contraction rates in Section 4.

Let us denote the joint posterior distribution in the state \( U_t(k) \) and the drift \( F_t \), conditioned on the observed data \( Y_{(0,t]}(k) \) up to time \( t \), \( k \geq 1 \), by \( \pi_t^{(k)} \). Here \( Y_{(0,t]}(k) \) represents a path of the observation process up to time \( t > 0 \), whereas \( Y_t(k) \) in (7) denotes its instantaneous value at time \( t \). By construction, the equivalent process in function space is given by
\[
y_{(0,t]}(x) = \sum_{k \geq 1} Y_{(0,t]}(k) \sin(kx).
\]
We emphasize that this joint posterior distribution is Gaussian due to the linearity of (4) and (6), respectively.

We next reformulate this Bayesian inference problem in terms of pairs of new random variables \( (U_t(k), F_t(k)) \), \( k \geq 1 \), which satisfy the mean-field Kalman–Bucy equations [6, 15]
\[
dU_t(k) = -k^2U_t(k)dt + F_t(k)dt + \gamma^{1/2}dW_t(k) - K_t^u(k)dI_t(k),
\]
\[
dF_t(k) = -K_t^f(k)dI_t(k),
\]
for a given observation path \( Y_{(0,t]}(k) \) with the innovation \( I_t(k) \) that solves
\[
dI_t(k) := \frac{1}{2} \left( U_t(k) + \mathbb{E}[U_t(k)|Y_{(0,t]}(k)] \right) dt - dY_t(k),
\]
and with Kalman gains
\[
K_t^u(k) := \rho^{-1}\mathbb{E}[\Delta U_t(k)^2|Y_{(0,t]}(k)], \quad K_t^f(k) := \rho^{-1}\mathbb{E}[\Delta F_t(k)\Delta F_t(k)|Y_{(0,t]}(k)],
\]
where
\[
\Delta U_t(k) := U_t(k) - \mathbb{E}[U_t(k)|Y_{(0,t]}(k)]
\]
and
\[
\Delta F_t(k) := F_t(k) - \mathbb{E}[F_t(k)|Y_{(0,t]}(k)].
\]
The stochastic process defined by the Kalman–Bucy equations (10) satisfies
\[
(U_t(k), F_t(k)) \sim \pi_t^{(k)}.
\]
In other words, upon defining the conditional posterior expectation value

\[ \pi_t^{(k)}[g] = \int_{\mathbb{R}^2} g(u, f) \pi_t^{(k)}(u, f) \, du \, df \]  

for any measurable function \( g(u, f) \), it holds, for example, that the mean of \( U_t(k) \), denoted by \( \mathbb{E}[U_t(k) \mid Y_{0,t}(k)] \), is equal to \( \pi_t^{(k)}[u] \), and

\[ K_t^u(k) = \rho^{-1} \pi_t^{(k)} \left[ \left( u - \pi_t^{(k)}[u] \right)^2 \right] = \frac{\sigma_i^2(k)}{\rho} \]

where \( \sigma_i^2(k) \) is the variance of \( U_i(k) \). We will use the shorthand \( \mathbb{E}[U_t(k)] \) instead of \( \mathbb{E}[U_t(k) \mid Y_{0,t}(k)] \) from now on in order to simplify notation.

The prior assumptions are that \( U_0(k) = 0 \) almost surely, and that the \( F_0(k), k \geq 1 \), are independent Gaussian random variables with mean zero and variance

\[ \sigma_i^2(k) = k^{-2\alpha - 1} \]

for a suitable \( \alpha > 0 \). Since (4) and (6) are linear in the unknowns, \( U_t(k) \) and \( F_t(k) \) remain Gaussian under the evolution equations (10), and we investigate their behaviour in terms of their mean and covariance for \( t \gg 1 \). We further assume that the true drift function \( f^* \) has Sobolev regularity \( \beta > 0 \), that is

\[ \sum_{k \geq 1} F^*(k) k^{-2\beta} < \infty. \]  

These choices correspond to a standard setting for Bayesian non-parametric inference [13, 8, 7]. For example, (19) is satisfied if

\[ F^*(k) = c_k k^{-\beta - 1/2 - \delta} \]

for any \( \delta > 0 \) and any sequence of coefficients \( c_k \) with bounded \( |c_k| \) as \( k \to \infty \). These coefficients can, for example, be realisations of i.i.d. uniform random variables from the interval \([-1,1]\).

In addition to an analysis of the continuous-time Bayesian inference problem (10), we also study the (frequentist) dependence of the estimators (Bayesian posterior means)

\[ U_t(k) := \mathbb{E}[U_t(k)] = \pi_t^{(k)}[u], \quad F_t(k) := \mathbb{E}[F_t(k)] = \pi_t^{(k)}[f], \]

on the random observation process \( Y_{0,t}(k) \) in Section 4. We denote the expectation values of (21) with respect to \( Y_{0,t}(k) \) by

\[ m_t^*(k) := \mathbb{E}^* [U_t(k)] \quad \text{and} \quad m_t^f(k) := \mathbb{E}^* [F_t(k)]. \]

These two quantities characterise the systematic bias in the estimators (21), while the implied variances

\[ p_t^u(k) := \mathbb{E}^* \left[ (U_t(k) - m_t^u(k))^2 \right], \quad p_t^f(k) := \mathbb{E}^* \left[ (F_t(k) - m_t^f(k))^2 \right] \]

are a measure of the ‘frequentist’ uncertainty of the estimators (21).

According to Lemma 8.2 from [7], the posterior contraction rate \( \phi_t(k) \) in the \( k \)th Fourier mode of the drift estimator \( F_t(k) \), that is, \( \phi_t(k) \) in

\[ \mathbb{E}^* \left[ \mathbb{P} \{ F_t(k) : |F_t(k) - F^*(k)| \geq M_t \phi_t(k) \mid Y_{0,t}(k) \} \right] \xrightarrow{t \to \infty} 0, \]

is provided by

\[ \phi_t(k)^2 = \sigma_i^2(k) + p_t^f(k) + (m_t^f(k) - F^*(k))^2. \]
Equation (24) holds for any nonnegative function $M_t$ with $M_t \to \infty$. Here, $\sigma^2_t(k)$ denotes the posterior variance of $F_t(k)$.

The main contributions from our subsequent analysis consist of providing bounds for the three quantities $\sigma^2_t(k)$, $p^2_t(k)$, and $m^2_t(k) - F^*$ in terms of the statistical models (4) and (6), and additionally studying the asymptotic behaviour of the $l_2$ asymptotic contraction rate $\phi_t$, defined by

$$\phi^2_t := \sum_{k \geq 1} \phi_t(k)^2,$$

in terms of the Sobolev regularity $\beta$ of the true drift function $f^*$ and the variance of the prior characterised by $\alpha > 0$ in (18). Here, the $l_2$ asymptotic contraction rate $\phi_t$ is to be understood in the sense that

$$\mathbb{E}_* \left[ P \left\{ F_t : \sum_{k \geq 1} (F_t(k) - F^*(k))^2 \geq M^2_t \phi^2_t \left\{ \mathcal{Y}_{[0,t]}(k) \right\}_{k \geq 1} \right\} \right] \xrightarrow{t \to \infty} 0$$

for any nonnegative function $M_t$ with $M_t \to \infty$. In physical space,

$$\mathbb{E}_* \left[ P \left\{ f_t : \frac{1}{\pi} \int_{\Omega} \left( f_t(x) - f^*(x) \right)^2 dx \geq M^2_t \phi^2_t \left\{ y_{[0,t]} \right\} \right\} \right] \xrightarrow{t \to \infty} 0,$$

where $f_t$ denotes the inverse Fourier transform of $F_t$, that is,

$$f_t(x) = \sum_{k \geq 1} F_t(k) \sin(kx).$$

We first discuss a stationary formulation of the drift estimation problem in Section 3, which is closely linked to available results for non-parametric Bayesian inverse problems [13, 8, 7]. The combined state and drift estimation problem is analysed in Section 4.

**Remark 2.1.** A more general class of time-scaled prior variances over the unknown set of static parameters $F^*(k)$ has been considered in [13]. This more general class of prior distributions could be incorporated into our analysis as well. However, we restrict the subsequent analysis to the simpler case (18).

3. **Stationary drift estimation problem.** Before addressing the time-dependent problem (4), we first consider the stationary drift estimation problem in which $\gamma = 0$ and $d\mathcal{U}_t(k) = 0$ in (4), giving rise to the forward model

$$\mathcal{U}^*(k) = k^{-2} F^*(k)$$

and observations

$$d\mathcal{Y}_t(k) = \mathcal{U}^*(k) dt + \rho^{1/2} d\mathcal{V}_t(k)$$

$$= k^{-2} F^*(k) dt + \rho^{1/2} d\mathcal{V}_t(k)$$

for $t \geq 0$. The forward model (31) leads to a mildly ill-posed inverse problem. We follow a Bayesian approach by placing a Gaussian mean-zero prior with variance given by (18) on the Fourier coefficients of the unknown drift function. Let us summarise our assumptions for ease of reference in the following:

**Assumption 3.1.** We assume that the Fourier coefficients of the true drift function satisfy (19). The prior coefficients, denoted by $F_0(k)$, are independent Gaussian random variables with mean zero, that is, $\mathcal{F}_0(k) = 0$, and with variance $\sigma^2_0(k)$ satisfying (18).
Let us denote the Bayesian posterior estimate of the drift at time $t > 0$ by $F_t(k)$, $k \geq 1$. We recall that $F_t(k)$ is Gaussian for all $t > 0$ provided $F_0(k)$ is Gaussian. Additionally, we denote the mean of $F_t(k)$ by $\mathcal{F}_t(k)$ and the variance by $\sigma^2_t(k)$. The Kalman–Bucy evolution equations (10) reduce to the following equations for the mean and the variance of the $k$th Fourier mode:

$$d\mathcal{F}_t(k) = -K_t^f(k) \left( k^{-2} \mathcal{F}_t(k) dt - d\mathcal{Y}_t(k) \right),$$  \hfill (32a)  

$$\frac{d}{dt} \sigma^2_t(k) = -K_t^f(k) \left( \sigma^2_t(k) \frac{\sigma^2_0(k)}{k^2} \right),$$  \hfill (32b)

where $K_t^f(k)$ denotes the Kalman gain factor

$$K_t^f(k) := \frac{\sigma^2_t(k)}{\rho k^2}. \quad (33)$$

The initial conditions are provided by Assumption 3.1.

**Remark 3.2.** Alternatively, the Kalman–Bucy equations (32) for the mean and the variance can be reformulated as mean-field equations in $F_t(k)$ directly, that is,

$$dF_t(k) = -K_t^f(k) \left( \frac{1}{2k^2} \left( F_t(k) + \mathcal{F}_t(k) \right) dt - d\mathcal{Y}_t(k) \right).$$  \hfill (34)

These equations are a special case of (10), and provide the starting point for numerical implementation in the form of the ensemble Kalman–Bucy filter [2], as well as extensions to nonlinear and non-Gaussian Bayesian inference problems [21, 18].

The evolution equation (32b) for the posterior variance has the closed form solution

$$\sigma^2_t(k) = \frac{\sigma^2_0(k)}{\rho^{-1}k^{-4}\sigma^2_0(k)t + 1}. \quad (35)$$

The trace of the covariance of the full joint posterior process thus satisfies the asymptotic estimate

$$\sum_{k \geq 1} \sigma^2_t(k) \asymp t^{-2\alpha/(2\alpha+5)}$$  \hfill (36)

with the initial variances satisfying (18). This result follows from asymptotic bounds for infinite sums. See Lemma K.7 in [7] in particular.

Next, we carry out a ‘frequentist’ analysis of the mean $\mathcal{F}_t(k)$ in terms of its bias with respect to the true $\mathcal{F}^*(k)$, and its variance $p_t^f(k)$ with respect to the measurement noise $\mathcal{V}_t(k)$. We rewrite (32a) as

$$d\mathcal{F}_t(k) = -K_t^f(k) \left( k^{-2}(\mathcal{F}_t(k) - F^*(k)) dt - \rho^{1/2}d\mathcal{V}_t(k) \right).$$  \hfill (37)

Denoting the expectation value of $\mathcal{F}_t(k)$ with respect to the measurement noise $\mathcal{V}_t(k)$ by $m_t^f(k)$, we obtain

$$\frac{d}{dt}m_t^f(k) = -k^{-2}K_t^f(k) \left( m_t^f(k) - F^*(k) \right)$$  \hfill (38)

with initial condition $m_0^f(k) = 0$. Equation (38) has the closed-form solution

$$m_t^f(k) = F^*(k) - \frac{\sigma_t(k)}{\sigma_0(k)}F^*(k).$$  \hfill (39)
Hence, based on (20), the $l_2$-norm of the frequentist bias satisfies the asymptotic estimate
\[
\sum_{k \geq 1} \left( m^f_t(k) - \mathcal{F}^*(k) \right)^2 \lesssim t^{-2\beta/(2\alpha+5)}.
\] (40)
This result follows again from asymptotic bounds for infinite sums.

We finally investigate the time evolution of the frequentist variance $\sigma^f_t(k)$. Starting from the stochastic differential equation
\[
d \left( \mathcal{F}_t(k) - m^f_t(k) \right) = -K^f_t(k) \left( k^{-2} \left( \mathcal{F}_t(k) - m^f_t(k) \right) dt - \rho^{3/2} d\mathcal{V}_t(k) \right),
\] (41)
using Itô’s lemma yields
\[
\frac{d}{dt} \sigma^f_t(k) = -2 \frac{\sigma^f_t(k)}{\rho^k} \sigma^f_t(k) + \rho \left( K^f_t(k) \right)^2
\] (42a)
\[
= -2 \frac{\sigma^f_t(k)}{\rho^k} \sigma^f_t(k) + \frac{\left( \sigma^f_t(k) \right)^2}{\rho^k}.
\] (42b)
The initial variance is $\sigma^f_0(k) = 0$ for all $k \geq 1$. In order to analyse the solution behaviour of (42), we introduce $\Delta \sigma^f_t(k) := \sigma^f_t(k) - \sigma^f_t(k)$ and find the associated evolution equation
\[
\frac{d}{dt} \Delta \sigma^f_t(k) = -2 \frac{\sigma^f_t(k)}{\rho^k} \Delta \sigma^f_t(k)
\] (43)
from which we conclude that $\sigma^f_t(k) < \sigma^f_t(k)$, and
\[
\lim_{t \to \infty} \sigma^f_t(k) = \sigma^f_t(k).
\] (44)
In fact, the explicit solution of (42) is
\[
\sigma^f_t(k) = \frac{\rho^k \sigma^f_0(k)^2 t}{\left( \rho^k + \sigma^f_0(k)^2 k t \right)^2}.
\] (45)

The ‘frequentist’ uncertainty is characterised by
\[
\mathcal{F}_t(k) - \mathcal{F}^*(k) = \mathcal{O}_\mathcal{P}(\Delta_\mathcal{V}_t(k)), \quad \Delta_\mathcal{V}_t(k) \sim N(m^f_t(k) - \mathcal{F}^*(k), \sigma^f_t(k)),
\] (46)
so the probability of the event
\[
|\mathcal{F}_t(k) - \mathcal{F}^*(k)| > M_t|\Delta_\mathcal{V}_t(k)|
\] (47)
has probability tending to zero for any nonnegative function $M_t$ with $M_t \to \infty$. Here $N(m, \sigma)$ denotes the Gaussian distribution with mean $m$ and variance $\sigma$.

**Theorem 3.3.** Under Assumption 3.1, the posterior contraction rate (26) is given by
\[
\phi_t^2 = \sum_{k \geq 1} \left( \sigma^f_t(k) + \sigma^\mathcal{V}_t(k) + \left( m^f_t(k) - \mathcal{F}^*(k) \right)^2 \right) \lesssim t^{-2\min\{\alpha, \beta\}/(2\alpha+5)}
\] (48)
for $t \to \infty$.

**Proof.** According to Lemma 8.2 from [7], the posterior contraction rate $\phi_t(k)$ in each Fourier mode is provided by
\[
\phi_t(k)^2 = \sigma^f_t(k) + \sigma^\mathcal{V}_t(k) + \left( m^f_t(k) - \mathcal{F}^*(k) \right)^2.
\] (49)
Because of (36) and (40) together with $\sigma^f_t(k) < \sigma^f_t(k)$, the result (48) follows. $\square$
Remark 3.4. The following white noise forward model has been investigated in [13]:

\[ Y_n(k) = k^{-p} F^*(k) + \frac{1}{n^{1/2}} \Xi_n(k), \quad n \in \mathbb{Z}^+, \tag{50} \]

where \( \Xi_n(k) \) are i.i.d. Gaussian random variables with mean zero and variance one. The associated inference problem corresponds to a sequence of measurements with measurement error decreasing as \( 1/n \). In our continuous-time problem, we obtain the same asymptotic rates as in the case considered above with \( p = 2 \) under the formal equivalence \( t = n \to \infty \).

4. Time-dependent state and drift estimation. We now return to the full dynamic model (4) subject to observations (6). We primarily wish to estimate the drift function \( F^* \). However, because of the stochastic model errors in (4), we also need to estimate the states \( U_t \). We start with a careful analysis of the single mode system for both small- and large- \( k \) Fourier modes.

4.1. Analysis of the single-mode filtering problem. In this section, we conduct a careful analysis of the single-mode filtering and parameter estimation problem. We suppress the dependence on the mode number \( k \), and introduce the parameter \( \epsilon = k^{-2} \). The signal process is therefore given by

\[ dU_t = -\epsilon^{-1} U_t dt + F^* dt + \gamma^{1/2} dW_t \tag{51} \]

with given initial condition \( U_0 = 0 \) almost surely. Observations of the process are given by

\[ dY_t = U_t dt + \rho^{1/2} dV_t \tag{52} \]

with \( Y_0 = 0 \) almost surely. The complete observation path up to time \( t \) is denoted by \( Y_{(0,t]} \).

Recall that the mean-field Kalman–Bucy equations are given by (10) and that their solutions \( (U_t, F_t) \) are Gaussian distributed. We again drop the dependence on the Fourier mode number \( k \) in the subsequent analysis. Taking conditional expectations in the mean-field equations (10) gives rise to the following evolution equations in the conditional means \( (\overline{U}_t, \overline{F}_t) \):

\[ d\overline{U}_t = -\epsilon^{-1} \overline{U}_t dt + F_t dt - K_{\mu} \overline{U}_t dt - dY_t, \tag{53a} \]
\[ d\overline{F}_t = -K_{\mu} \overline{U}_t dt - dY_t. \tag{53b} \]

The deviations \( \Delta U_t := U_t - \overline{U}_t \) and \( \Delta F_t := F_t - \overline{F}_t \) thus satisfy

\[ d\Delta U_t = -\epsilon^{-1} \Delta U_t dt + \Delta F_t dt + \gamma^{1/2} dW_t - \frac{1}{2} K_{\mu} \Delta U_t dt, \tag{54a} \]
\[ d\Delta F_t = -\frac{1}{2} K_{\mu} \Delta U_t dt. \tag{54b} \]

We can further decompose (53) by introducing the ‘frequentist’ expectation values \( \mu_t^\prime = E^*[\overline{U}_t] \) and \( \mu_t^\prime = E^*[\overline{F}_t] \) with respect to the observation process (52), and the deviations \( \Delta \overline{U}_t := \overline{U}_t - \mu_t^\prime \) and \( \Delta \overline{F}_t := \overline{F}_t - \mu_t^\prime \). Taking this expectation in (6) yields

\[ E^*[dY_t] = E^*[U_t] dt. \tag{55} \]
Introducing the shorthand $\mu_t = \mathbb{E}^*[U_t]$, one is left with the ordinary differential equations
\[
\frac{d}{dt} m_t^u = -\epsilon^{-1} m_t^u + m_t^f - K_t^u (m_t^u - \mu_t),
\]
\[
\frac{d}{dt} m_t^f = -K_t^f (m_t^u - \mu_t),
\]
for the mean values $(m_t^u, m_t^f)$. Furthermore, defining $\Delta U_t := U_t - \mu_t$,
\[
\frac{d}{dt} \Delta U_t = -\epsilon^{-1} \Delta U_t dt + \Delta F_t dt - K_t^u \left( (\Delta U_t - \Delta U_t) dt - \rho^{1/2} d\mathcal{W}_t \right),
\]
\[
\frac{d}{dt} \Delta F_t = -K_t^f \left( (\Delta U_t - \Delta U_t) dt - \rho^{1/2} d\mathcal{W}_t \right)
\]
for the deviations $(\Delta U_t, \Delta F_t)$. It follows from (51) that the expectation value $\mu_t$ satisfies the evolution equation
\[
\frac{d}{dt} \mu_t = -\epsilon^{-1} \mu_t + \mathcal{F}^*,
\]
while the deviation $\Delta U_t$ satisfies
\[
\frac{d}{dt} \Delta U_t = -\epsilon^{-1} \Delta U_t dt + \gamma^{1/2} d\mathcal{W}_t.
\]
Both equations follow from (51). We finally combine (57a) and (59) into a single equation for the new variable $\hat{U}_t := \Delta U_t - \Delta U$, and replace (57) by
\[
\frac{d}{dt} \hat{U}_t = -\epsilon^{-1} \hat{U}_t dt + \Delta F_t dt - \gamma^{1/2} d\mathcal{W}_t - K_t^u \left( \hat{U}_t dt - \rho^{1/2} d\mathcal{W}_t \right),
\]
\[
\frac{d}{dt} \Delta F_t = -K_t^f \left( \hat{U}_t dt - \rho^{1/2} d\mathcal{W}_t \right).
\]
We have thus decomposed the mean-field Kalman–Bucy equations into the analysis of the three subsystems (54), (56) and (60). Here, (56) describes the systematic bias in the Bayesian mean estimator, while (60) and (54) characterise the ‘frequentist’ and ‘Bayesian’ uncertainties, respectively.

We note that the equations in (54) do not depend on the observation process $\mathcal{Y}_{(0,t)}$, and are decoupled from (56) and (60). In fact, since $(\Delta U_t, \Delta F_t)$ are centred Gaussian random variables, it is sufficient to look at the (deterministic) time evolution equations for the posterior variances
\[
\sigma_t^u := \mathbb{E}[\Delta U_t^2], \quad \sigma_t^f := \mathbb{E}[\Delta F_t^2],
\]
and the posterior covariance
\[
\sigma_t^{uf} := \mathbb{E}[\Delta U_t \Delta F_t],
\]
which are given by
\[
\frac{d}{dt} \sigma_t^u = -2\epsilon^{-1} \sigma_t^u + 2\sigma_t^{uf} + \gamma - \rho^{-1} (\sigma_t^u)^2,
\]
\[
\frac{d}{dt} \sigma_t^{uf} = -\epsilon^{-1} \sigma_t^{uf} + \sigma_t^f - \rho^{-1} \sigma_t^u \sigma_t^{uf},
\]
\[
\frac{d}{dt} \sigma_t^f = -\rho^{-1} (\sigma_t^{uf})^2.
\]
Since $(\hat{U}_t, \Delta F_t)$ are also centred Gaussian random variables, the time-dependent linear SDEs (60) can again be analysed in terms of the variances
\[
p_t^u := \mathbb{E}*[\hat{U}_t^2], \quad p_t^f := \mathbb{E}*[\Delta F_t^2],
\]
and the covariance
\[ p^{uf}_t := E^* [\hat{U}_t \Delta F_t]. \]  

(65)

These quantities satisfy the linear time-dependent ordinary differential equations
\[ \frac{d}{dt} p^u_t = -2\epsilon^-1 p^u_t + 2p^{uf}_t + \gamma - 2K^u_t p^u_t + \rho(K^u_t)^2, \]  

(66a)
\[ \frac{d}{dt} p^{uf}_t = -\epsilon^-1 p^{uf}_t + p^u_t - K^u_t p^{uf}_t - K^f_t p^u_t + \rho K^u_t K^f_t, \]  

(66b)
\[ \frac{d}{dt} p^f_t = -2K^f_t p^{uf}_t + \rho(K^f_t)^2. \]  

(66c)

Equations (63) and (66) are combined to to yield
\[ \frac{d}{dt} \Delta p^u_t = -2(\epsilon^-1 + K^u_t) \Delta p^u_t + 2\Delta p^{uf}_t, \]  

(67a)
\[ \frac{d}{dt} \Delta p^{uf}_t = -(\epsilon^-1 + K^u_t) \Delta p^{uf}_t + \Delta p^f_t - K^f_t \Delta p^u_t, \]  

(67b)
\[ \frac{d}{dt} \Delta p^f_t = -2K^f_t \Delta p^{uf}_t \]  

(67c)

in the variables
\[ \Delta p^{uf}_t := p^{uf}_t - \sigma^{uf}_t, \quad \Delta p^u_t := p^u_t - \sigma^u_t, \quad \Delta p^f_t := p^f_t - \sigma^f_t. \]  

(68)

The initial conditions are \( p^u_0 = p^f_0 = p^{uf}_0 = \sigma^u_0 = \sigma^{uf}_0 = 0, \) and \( \sigma^f_0 = \epsilon^{\alpha+1/2}. \)

**Lemma 4.1.** It holds that
\[ p^u_t \to \sigma^u_t, \quad p^f_t \to \sigma^f_t, \quad p^{uf}_t \to \sigma^{uf}_t \]  

as \( t \to \infty, \) that is, \( \Delta p^u_t = \Delta p^{uf}_t = \Delta p^f_t = 0 \) is a stable equilibrium point of (67).

**Proof.** The lemma follows from the decay property of
\[ V_t(\Delta p^u_t, \Delta p^{uf}_t, \Delta p^f_t) := \frac{1}{2} (\Delta p^u_t)^2 + K^f_t (\Delta p^{uf}_t)^2 + \frac{(K^f_t)^2}{2} (\Delta p^f_t)^2 \]  

(70)

and the fact that \( K^f_t \sim c/t \) for \( t \) sufficiently large with an appropriate constant \( c > 0. \) In particular,
\[ \partial_t V_t = \left\{ (\Delta p^{uf}_t)^2 + K^f_t (\Delta p^u_t)^2 \right\} \frac{d}{dt} K^f_t \leq 0, \]  

(71)

and for the gradient \( \nabla V_t(\Delta p^u_t, \Delta p^{uf}_t, \Delta p^f_t) \in \mathbb{R}^3: \)
\[ \nabla V_t \cdot \frac{d}{dt} \begin{pmatrix} \Delta p^u_t \\ \Delta p^{uf}_t \\ \Delta p^f_t \end{pmatrix} = -2(\epsilon^-1 + K^u_t) \left\{ K^f_t (\Delta p^{uf}_t)^2 + (\Delta p^u_t)^2 \right\} K^f_t \leq 0. \]  

(72)

Therefore, the total time derivative satisfies
\[ \frac{D}{Dt} V_t(\Delta p^u_t, \Delta p^{uf}_t, \Delta p^f_t) \leq 0. \]  

(73)

Furthermore, for large enough \( t, \)
\[ p^u_t \approx (\epsilon^-1 + K^u_t)^{-1} p^f_t, \quad p^f_t \approx (\epsilon^-1 + K^u_t)^{-1} p^{uf}_t \approx (\epsilon^-1 + K^u_t)^{-2} p^f_t, \]  

(74)

and \( \partial_t V_t \) becomes small relative to (72) since \( \frac{d}{dt} K^f_t \sim -c/t^2. \) Hence, for \( t \) again sufficiently large and with an appropriate constant \( C > 0, \)
\[ \frac{d}{dt} (\Delta p^f_t)^2 \approx -\frac{C}{t} (\Delta p^f_t)^2, \]  

(75)
and the desired convergence result follows.

Lemma 4.1 implies that, as for the pure drift estimation problem, the Bayesian (filtering) variance asymptotically covers the ‘frequentist’ (data) variance.

4.1.1. The large-\( k \) Fourier mode case. We now investigate the case \( \epsilon \ll 1 \) in more detail. The equations (63) possess a slow manifold for \( \epsilon > 0 \) sufficiently small [19, 16]. To leading order, solutions on the slow manifold satisfy

\[
\sigma_t^u = \epsilon \gamma / 2, \quad \sigma_t^u = \epsilon \sigma_t^f.
\] (76)

In other words, the long-time variance of \( U_t \) is governed by \( \sigma_t^u = \epsilon \gamma / 2 \), and of \( F_t \) by

\[
\frac{d}{dt} \sigma_t^f = -K_t \sigma_t^f
\] (77)

up to higher-order terms in \( \epsilon \). Here we have introduced the new Kalman gain factor

\[
K_t := \frac{\epsilon^2}{\rho} \sigma_t^f.
\] (78)

We also notice that the Kalman gain factor \( K_t^u \) satisfies \( K_t^u = O(\epsilon) \). See the Appendix for more details on the derivation of the reduced system (77).

The combined time-dependent linear equations (58) and (56) also give rise to a slow manifold, with the slow dynamics governed by \( \mu_t = \epsilon F^* \), \( m_t^u = \epsilon m_t^f \), and

\[
\frac{d}{dt} m_t^f = -K_t (m_t^u - F^*).
\] (79)

The initial condition is given by \( m_0^f = 0 \).

The solution of (77) is explicitly given by

\[
\sigma_t^f = \frac{\rho \sigma_0^f}{\epsilon^2 \sigma_0^f t + \rho}.
\] (80)

We also find that the solution of (79) satisfies

\[
m_t^f = \left(1 - \frac{\sigma_t^f}{\sigma_0^f}\right) F^*.
\] (81)

The large-time limit yields that \( m_t^f - F^* \) decays with rate \( t^{-1} \). We have already shown that \( p_t^f \to \sigma_t^f \) as \( t \to \infty \).

**Remark 4.2.** We note that the combined state and drift estimation problem under the slow manifold approximation behaves exactly like the stationary drift estimation problem from Section 3. In particular, compare equations (35) and (39). The numerical experiments in Section 5 reveal that the slow manifold approximation already holds for rather small-\( k \) Fourier modes.

In summary, we have shown that

\[
F_t - F^* = O_t(\Delta_t)
\] (82)

with

\[
\Delta_t \sim N(m_t^f - F^*, p_t^f).
\] (83)

Hence, according to Lemma 8.2. from [7], the posterior contraction rate \( \phi_t \) in

\[
\mathbb{E}^* \mathbb{P} \{ F_t : |F_t - F^*| \geq M_t \phi_t | \mathcal{Y}_{(0,t]} \} \overset{t \to \infty}{\longrightarrow} 0
\] (84)
is provided by
\[\phi_t^2 = \sigma_t^f + p_t^f + (m_t^f - F^*)^2 \approx t^{-1}\] (85)
since \(p_t^f < \sigma_t^f\).

**Remark 4.3.** While the posterior contraction rate is \(\phi_t \approx t^{-1/2}\) for any fixed \(k \gg 1\), the analysis of the infinite-dimensional state and parameter estimation problem is more complicated since one takes the limit \(k \to \infty\) first, and then the limit \(t \to \infty\).

4.1.2. *The small-\(k\) Fourier mode case.* The above analysis holds for \(\epsilon\) sufficiently small, that is, for sufficiently large Fourier modes. We now investigate the behaviour of the filtering and estimation problem for \(\epsilon = O(1)\). We set \(\epsilon = 1\), that is, \(k = 1\) for simplicity. We start again with (63) and note that there is a finite time \(t^*\) such that \(\sigma_t^u\) has become sufficiently small relative to \(\gamma\), and we conclude from (63a) that
\[\sigma_t^u \approx C_0 \quad \text{for all} \quad t \geq t^*,\] (86)
where \(C_0 > 0\) satisfies the quadratic equation
\[2C_0 - \gamma + \rho^{-1}C_0^2 = 0.\] (87)

**Remark 4.4.** We note that \((\sigma_t^{uf})^2 \leq \sigma_t^u \sigma_t^f\). We also know that \(\sigma_t^u\) remains bounded since the system is fully observed, and that \(\sigma_t^f\) decays. Therefore, the assumption that \(\sigma_t^{uf}\) becomes small relative to \(\gamma\) is justified for \(t \geq t^*\).

Upon substituting (86) into (63b), we can conclude an exponential decay of \(\sigma_t^{uf}\) towards \(\sigma_t^f\) and, hence, we can choose \(t^*\) large enough such that
\[\sigma_t^{uf} \approx \left(1 + \frac{C_0}{\rho}\right)^{-1} \sigma_t^f \quad \text{for all} \quad t \geq t^*.\] (88)
In other words, the asymptotic dynamics in \(\sigma_t^f\) is again governed by (77) with Kalman gain factor
\[K_t := C_1^2 \frac{\sigma_t^f}{\rho}\] (89)
and constant
\[C_1 := \left(1 + \frac{C_0}{\rho}\right)^{-1}.\] (90)
Therefore, \(\sigma_t^f\) behaves asymptotically like \(t^{-1}\).

We next analyse the long-time dynamics of (56). We use \(K_t^u \approx C_0/\rho\) and find that (56a) gives rise to the quasi-equilibrium
\[0 \approx -m_t^u + m_t^f - \frac{C_0}{\rho} (m_t^u - F^*) = -C_1^{-1} (m_t^u - F^*) + m_t^f - F^*\] (91)
for \(t \geq t^*\) and \(t^*\) sufficiently large. Upon substituting this relation into (56b) and using
\[K_t^f \approx C_1^2 \frac{\sigma_t^f}{\rho} = C_1^{-1} K_t,\] (92)
(89) as well as \(\mu_t \approx F^*\), we arrive at (79) with gain factor \(K_t\) given by (89). Repeating the analysis from Section 4.1.1, one finds again that \(F_t - F^*\) satisfies (82) with \(|m_t^f - F^*| \approx t^{-1}\) and \(p_t^u < \sigma_t^u\).

Summarising our findings, we conclude that the contraction rate \(\phi_t\) in (84) is of order \(t^{-1/2}\) for \(t\) sufficiently large.
4.2. Asymptotic rates for the stochastic heat equation. We recall our Assumption 3.1 on the asymptotic behaviour of the true drift function \( F^*(k) \) and on the prior random variables \( F_0(k), k \geq 1 \). These assumptions together with the results from Sections 4.1.1 and 4.1.2, and in particular (80) and (81) with \( \epsilon = k^{-2} \), imply that

\[
\sigma_f^2(k) \sim \frac{\rho k^{-(2\alpha+1)}}{k^{-(2\alpha+5)}t + \rho}, \tag{93}
\]

and

\[
|m_f^2(k) - F^*(k)| \sim \frac{\rho k^{-(\beta+1/2)}}{k^{-(2\alpha+5)}t + \rho}, \tag{94}
\]

for \( k \geq k_* \), and

\[
\sigma_f^2(k) \approx t^{-1}, \quad m_f^2(k) - F^*(k) \approx t^{-1}, \tag{95}
\]

for all \( k \leq k_* \), and \( t \geq t_* \) sufficiently large. Here \( k_* \) and \( t_* \) are chosen sufficiently large such that the analysis of Sections 4.1.1 and 4.1.2, respectively, applies.

**Theorem 4.5.** Under Assumption 3.1 and the regularity of the true drift function \( F^*(k) \) and the prior \( F_0(k), k \geq 1 \), the asymptotic contraction rate (26) in the \( l_2 \)-norm becomes

\[
\phi(t)^2 = \sum_{k \geq 1} \left( (m_f^2(k) - F^*(k))^2 + p_f^2(k) + \sigma_f^2(k) \right) \approx t^{-2 \min\{\alpha, \beta\}/(2\alpha+5)}. \tag{96}
\]

**Proof.** The theorem follows from (93) and (94) together with \( p_f^2(k) < \sigma_f^2(k) \) and standard asymptotic estimates for infinite sums. See, for example, again Lemma K.7 from [7].

We note that the rate in Theorem 4.5 is the same as in (48) for the stationary drift estimation problem. In other words, the need for estimating the states as well as the drift does not deteriorate the asymptotic contraction rates. In fact, the data does not affect the posterior uncertainty in the states as \( k \to \infty \), which is entirely determined by the equilibrium distribution of the associated Ornstein–Uhlenbeck process. We verify this behaviour through a simple numerical experiment in the following section.

5. Numerical exploration. We numerically implemented the evolution equations (56), (63), and (66) for the Bayesian means, the Bayesian variances, and the frequentist variances, respectively, for different Fourier modes \( k \). We also implemented the corresponding equations from Section 3 for the pure drift estimation problem. Simulations were run with \( \rho = \gamma = 1 \). The true reference value was set to \( F^*(k) = k^{-\beta-1/2} \) with \( \beta = 1/2 \), and we used \( \alpha = 1/2 \) for the prior variance (18).

The results can be found in Figure 1. They reveal a very similar behaviour for the combined state and drift estimation and the pure drift estimation problems. The delay in the onset of the asymptotic \( t^{-1} \) regime as \( k \) increases can also be seen. Overall, this simple numerical experiment confirms our theoretical investigations with regard to the dynamical behaviour in each Fourier mode from Section 4.

6. Conclusions. We have provided an analysis of the infinite-dimensional Kalman–Bucy filter mean-field equations (10) for a combined state and drift estimation problem defined in spectral space by (4) and (6). The derived asymptotic posterior contraction rates in the unknown drift function \( f^* \) from Theorem 4.5 agree with those derived for an associated stationary problem formulation in Theorem
We display the time evolution of the bias $|m_t^k - F^*_t(k)|$ and the variances $\sigma_t^k(k)$ and $p_t^k(k)$ for increasing values of $k$. We compare the results from the combined state and parameter estimation problem (labelled ‘dynamic’) with those from the associated direct inference problem in the parameter alone (labelled ‘stationary’). The delay in the onset of the asymptotic $t^{-1}$ regime as $k$ increases can be clearly seen as well as that $p_t^k(k) < \sigma_t^k(k)$ for all $t > 0$. Furthermore, as theoretically predicted, the dynamic and stationary problem formulations behave almost identically in terms of their drift estimates.

3.3. These theoretical findings imply that the required additional estimation of the states does not lead to a deterioration of the contraction rates, which has been confirmed numerically in Section 5. An extension to nonlinear and non-Gaussian estimation problems for SPDEs and different types of observation processes can be envisioned through nonlinear extensions of the Kalman–Bucy mean-field equations. See, for example, [15]. Furthermore, our results can be combined with those from [9] to study the coverage probabilities of Bayesian credible sets in a non-asymptotic regime. One could also discuss adaptive choices for the prior $F_0(k)$. See, for example, [12].
Acknowledgments. This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) - SFB1294/1 - 318763901.

Appendix: Slow manifold approximation. Given a system of differential equations of the form
\[ \frac{d}{dt} x_t = -\epsilon^{-1} A x_t + f(x_t, y_t; \epsilon), \]  
\[ \frac{d}{dt} y_t = g(x_t, y_t; \epsilon), \]  
with \( \epsilon > 0 \) sufficiently small and \( A \) a symmetric positive definite matrix, there exists a smooth manifold \( M_\epsilon \) which is exponentially attractive and invariant under the dynamics of (97). See, for example, [19, 16] for details.

Approximations of the slow manifold can be found by utilising the principle of bounded derivatives, that is,
\[ \frac{d^n}{dt^n} x_t = O(\epsilon^0) \]  
for \( n \geq 1 \). More specifically, (98) with \( n = 1 \) implies that
\[ x_t^{(0)} = O(\epsilon), \]  
while the same equation with \( n = 2 \) leads to
\[ \frac{d}{dt} x_t^{(1)} = O(\epsilon) \]  
and, hence,
\[ x_t^{(1)} = \epsilon A^{-1} f(0, y_t; 0) + O(\epsilon^2). \]  
Here the superscript indicates the order of the approximation. Substituting the leading order term of (101) into (97b) provides a reduced dynamics in the \( y \)-variable alone. The next order correction is obtained from
\[ \frac{d^2}{dt^2} x_t^{(2)} = -\epsilon^{-1} A \frac{d}{dt} x_t^{(2)} + D_1 f(x_t^{(2)}, y_t; \epsilon) \frac{d}{dt} x_t^{(2)} + D_2 f(x_t^{(2)}, y_t; \epsilon) \frac{d}{dt} y_t = O(\epsilon) \]  
and yields
\[ x_t^{(2)} = \epsilon A^{-1} f(x_t^{(1)}, y_t; \epsilon) + \epsilon^2 A^{-2} D_2 f(0, y_t; 0) \frac{d}{dt} y_t + O(\epsilon^3). \]  
Higher-order approximations to \( M_\epsilon \) can be obtained by exploiting (98) for \( n > 3 \).

Note that (63), for example, fits into the framework (97) with \( x = (\sigma^u, \sigma^v)^T \), \( y = \sigma^f \), and
\[ A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \]  
Hence (101) results in (76) and the evolution equation (77) in the slow variable \( y = \sigma^f \). The induced approximation error remains of order \( \epsilon \) over time intervals of order \( \epsilon^{-2} \).

What about \( t \gg \epsilon^{-2} \)? In this large time limit, the solution (80) of (77) satisfies
\[ \frac{d}{dt} \sigma_t^f = O(\epsilon^{-2} t^{-2}) \]  
implying
\[ \frac{d}{dt} \sigma_t^{uf} = O(\epsilon^{-1} t^{-2}) \]
along solutions on the slow manifold. In other words, it follows from (63b) that

\[(1 + \epsilon \rho^{-1} \sigma_t^u) \sigma_t^u f + \epsilon \sigma_t^f = O(t^{-2}) \tag{107}\]

for \(t \gg \epsilon^{-2}\). The resulting correction terms to the first-order balance relation \(\sigma_t^u f = \epsilon \sigma_t^f\) do not alter the qualitative long-time solution behaviour of the reduced slow equations (77) as verified by the numerical results from Section 5, where a \(t^{-1}\) long-time decay rate in \(\sigma_t^f\) has been observed.

REFERENCES

[1] R. Altmeyer and M. Reiß, Nonparametric estimation for linear SPDEs from local measurements, *Annals of Appl. Probability*, preprint, [arXiv:1903.06984](http://arxiv.org/abs/1903.06984).

[2] K. Bergemann and S. Reich, An ensemble Kalman–Bucy filter for continuous data assimilation, *Meteorolog. Zeitschrift*, 21 (2012), 213–219.

[3] I. Clainco, Statistical inference for SPDEs: An overview, *Stat. Inference Stoch. Process.*, 21 (2018), 309–329.

[4] R. Curtain, A survey of infinite-dimensional filtering, *SIAM Review*, 17 (1975), 305–411.

[5] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, 1st edition, Encyclopedia of Mathematics and its Applications, 45, Cambridge University Press, Cambridge, 1992.

[6] J. de Wiljes, S. Reich and W. Stannat, Long-time stability and accuracy of the ensemble Kalman–Bucy filter for fully observed processes and small measurement noise, *SIAM J. Appl. Dyn. Syst.*, 17 (2018), 1152–1181.

[7] S. Ghosal and A. van der Vaart, *Fundamentals of Nonparametric Bayesian Inference*, Cambridge Series in Statistical and Probabilistic Mathematics, 44, Cambridge University Press, Cambridge, 2017.

[8] E. Giné and R. Nickl, *Mathematical Foundations of Infinite-Dimensional Statistical Models*, Cambridge Series in Statistical and Probabilistic Mathematics, 40, Cambridge University Press, Cambridge, 2015.

[9] F. Götze, A. Naumov, V. Spokoiny and V. Ulyanov, Large ball probabilities, Gaussian comparison and anti-concentration, *Bernoulli*, 25 (2019), 2538–2563.

[10] M. Hairer, An introduction to stochastic PDEs, unpublished lecture notes, [arXiv:0907.4178](http://arxiv.org/abs/0907.4178).

[11] A. Jazwinski, *Stochastic Processes and Filtering Theory*, Mathematics in Science and Engineering, 64, Academic Press, New York, 1970.

[12] B. T. Knapik, B. T. Szabó, A. W. van der Vaart and J. H. van Zanten, Bayes procedures for adaptive inference in inverse problems for the white noise model, *Probab. Theory Relat. Fields*, 164 (2016), 771–813.

[13] B. T. Knapik, A. W. van der Vaart and J. H. van Zanten, Bayesian inverse problems with Gaussian priors, *Annals of Statistics*, 39 (2011), 2626–2657.

[14] G. J. Lord, C. E. Powell and T. Shardlow, *An Introduction to Computational Stochastic PDEs*, 1st edition, Cambridge Texts in Applied Mathematics, 50, Cambridge University Press, Cambridge, 2014.

[15] N. Nüsken, S. Reich and P. J. Rozdeba, State and parameter estimation from observed signal increments, *Entropy*, 21 (2019), 505.

[16] E. Shchepakina, V. Sobolev and M. P. Mortell, *Singular Perturbations: Introduction to System Order Reduction Methods with Applications*, Lecture Notes in Mathematics, 2114, Springer International Publishing, 2014.

[17] D. Simon, *Optimal State Estimation*, Wiley, Hoboken, 2006.

[18] A. Taghvaei, J. de Wiljes, P. Mehta and S. Reich, Kalman filter and its modern extensions for the continuous-time nonlinear filtering problem, *J. Dyn. Sys., Meas., Control.*, 140 (2017), 030904.

[19] F. Verhulst, Singular perturbation methods for slow-fast dynamics, *Nonlinear Dynamics*, 50 (2007), 747–753.

[20] D. Yan, *Bayesian inference for Gaussian models: Inverse problems and evolution equations*, PhD thesis, Universiteit Leiden, Leiden, 2020.

[21] T. Yang, P. Mehta and S. Meyn, Feedback particle filter, *IEEE Trans. Automatic Control*, 58 (2013), 2465–2480.
Received March 2020; revised August 2020.

E-mail address: sebastian.reich@uni-potsdam.de
E-mail address: paul.rozdeba@uni-potsdam.de