



## Optimization Based Methods for Partially Observed Chaotic Systems

Daniel Paulin<sup>1,2</sup> · Ajay Jasra<sup>1</sup> · Dan Crisan<sup>3</sup> ·  
Alexandros Beskos<sup>4</sup>

Received: 15 February 2017 / Revised: 9 March 2018 / Accepted: 9 March 2018  
© The Author(s) 2018

**Abstract** In this paper we consider filtering and smoothing of partially observed chaotic dynamical systems that are discretely observed, with an additive Gaussian noise in the observation. These models are found in a wide variety of real applications and include the Lorenz 96' model. In the context of a fixed observation interval  $T$ , observation time step  $h$  and Gaussian observation variance  $\sigma_Z^2$ , we show under assumptions that the filter and smoother are well approximated by a Gaussian with high probability when  $h$  and  $\sigma_Z^2 h$  are sufficiently small. Based on this result we show that the maximum a posteriori (MAP) estimators are asymptotically optimal in mean square error as  $\sigma_Z^2 h$  tends to 0. Given these results, we provide a batch algorithm for the smoother and filter, based on Newton's method, to obtain the MAP. In particular, we show that if the initial point is close enough to the MAP, then Newton's method

---

Communicated by Christian Lubich.

---

✉ Daniel Paulin  
paulindani@gmail.com

Ajay Jasra  
staja@nus.edu.sg

Dan Crisan  
d.crisan@ic.ac.uk

Alexandros Beskos  
a.beskos@ucl.ac.uk

<sup>1</sup> Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546, Singapore

<sup>2</sup> Department of Statistics, University of Oxford, Oxford OX1 3LB, UK

<sup>3</sup> Department of Mathematics, Imperial College London, London SW7 2AZ, UK

<sup>4</sup> Department of Statistical Science, University College London, London WC1E 6BT, UK

converges to it at a fast rate. We also provide a method for computing such an initial point. These results contribute to the theoretical understanding of widely used 4D-Var data assimilation method. Our approach is illustrated numerically on the Lorenz 96' model with state vector up to 1 million dimensions, with code running in the order of minutes. To our knowledge the results in this paper are the first of their type for this class of models.

**Keywords** Filtering · Smoothing · Chaotic dynamical systems · Gaussian approximation · Newton's method · Concentration inequalities · 4D-Var

**Mathematics Subject Classification** 37D45 · 65K10

## 1 Introduction

Filtering and smoothing are among the most important problems for several applications, featuring contributions from mathematics, statistics, engineering and many more fields; see, for instance, [13] and the references therein. The basic notion of such models is the idea of an unobserved stochastic process that is observed indirectly by data. The most typical model is perhaps where the unobserved stochastic process is a Markov chain or a diffusion process. In this paper, we are mainly concerned with the scenario when the unobserved dynamics are deterministic and, moreover, chaotic. The only randomness in the unobserved system is uncertainty in the initial condition, and it is this quantity that we wish to infer, on the basis of discretely and sequentially observed data; we explain the difference between filtering and smoothing in this context below. This class of problems has slowly become more important in the literature, particularly in the area of data assimilation [27]. The model itself has a substantial number of practical applications, including weather prediction, oceanography and oil reservoir simulation, see, for instance, [24].

In this paper we consider smoothing and filtering for partially observed deterministic dynamical systems of the general form

$$\frac{du}{dt} = -Au - B(u, u) + f, \quad (1.1)$$

where  $u : \mathbb{R}^+ \rightarrow \mathbb{R}^d$  is a dynamical system in  $\mathbb{R}^d$  for some  $d \in \mathbb{Z}_+$ ,  $A$  is linear operator in  $\mathbb{R}^d$  (i.e.  $A$  is a  $d \times d$  matrix),  $f \in \mathbb{R}^d$  is a constant vector, and  $B(u, u)$  is a bilinear form corresponding to the nonlinearity (i.e.  $B$  is a  $d \times d \times d$  array). We denote the solution of Eq. (1.1) with initial condition  $u(0) := v$  for  $t \geq 0$  by  $v(t)$ . The derivatives of the solution  $v(t)$  at time  $t = 0$  will be denoted by

$$D^i v := \left. \frac{d^i v(t)}{dt^i} \right|_{t=0} \quad \text{for } i \in \mathbb{N}, \quad (1.2)$$

in particular,  $D^0 v = v$ ,  $Dv := D^1 v = -Av - B(v, v) + f$  (the right-hand side of (1.1)), and  $D^2 v = -AD^1 v - B(D^1 v, v) - B(v, D^1 v)$ .

In order to ensure the existence of a solution to Eq. (1.1) for every  $t \geq 0$ , we assume that there are constants  $R > 0$  and  $\delta > 0$  such that

$$\langle \mathbf{D}\mathbf{v}, \mathbf{v} \rangle \leq 0 \text{ for every } \mathbf{v} \in \mathbb{R}^d \text{ with } \|\mathbf{v}\| \in [R, R + \delta]. \tag{1.3}$$

We call this the *trapping ball* assumption. Let  $\mathcal{B}_R := \{\mathbf{v} \in \mathbb{R}^d : \|\mathbf{v}\| \leq R\}$  be the ball of radius  $R$ . Using the fact that  $\left\langle \frac{d}{dt} \mathbf{v}(t), \mathbf{v}(t) \right\rangle = \frac{1}{2} \frac{d}{dt} \|\mathbf{v}(t)\|^2$ , one can show that the solution to (1.1) exists for  $t \geq 0$  for every  $\mathbf{v} \in \mathcal{B}_R$  and satisfies that  $\mathbf{v}(t) \in \mathcal{B}_R$  for  $t \geq 0$ .

Equation (1.1) was shown in Sanz-Alonso and Stuart [42] and Law et al. [27] to be applicable to three chaotic dynamical systems: the Lorenz 63’ model, the Lorenz 96’ model and the Navier–Stokes equation on the torus; such models have many applications. We note that instead of the trapping ball assumption, these papers have considered different assumptions on  $\mathbf{A}$  and  $\mathbf{B}(\mathbf{v}, \mathbf{v})$ . As we shall explain in Sect. 1.1, their assumptions imply (1.3); thus, the trapping ball assumption is more general.

We assume that the system is observed at time points  $t_j = jh$  for  $j = 0, 1, \dots$ , with observations

$$\mathbf{Y}_j := \mathbf{H}\mathbf{u}(t_j) + \mathbf{Z}_j,$$

where  $\mathbf{H} : \mathbb{R}^d \rightarrow \mathbb{R}^{d_o}$  is a linear operator and  $(\mathbf{Z}_j)_{j \geq 0}$  are i.i.d. centred random vectors taking values in  $\mathbb{R}^{d_o}$  describing the noise. We assume that these vectors have distribution  $\eta$  that is Gaussian with i.i.d. components of variance  $\sigma_Z^2$ .<sup>1</sup>

The contributions of this article are as follows. In the context of a fixed observation interval  $T$ , we show under assumptions that the filter and smoother are well approximated by a Gaussian law when  $\sigma_Z^2 h$  is sufficiently small. Our next result, using the ideas of the first one, shows that the maximum a posteriori (MAP) estimators (of the filter and smoother) are asymptotically optimal in mean square error when  $\sigma_Z^2 h$  tends to 0. The main practical implication of these mathematical results is that we can then provide a batch algorithm for the smoother and filter, based on Newton’s method, to obtain the MAP. In particular, we prove that if the initial point is close enough to the MAP, then Newton’s method converges to it at a fast rate. We also provide a method for computing such an initial point and prove error bounds for it. Our approach is illustrated numerically on the Lorenz 96’ model with state vector up to 1 million dimensions. We believe that the method of this paper has a wide range of potential applications in meteorology, but we only include one example due to space considerations.

We note that in this paper, we consider finite-dimensional models. There is a substantial interest in the statistics literature in recent years in nonparametric inference for infinite-dimensional PDE models, see [15] for an overview and references, and Giné and [21] for a comprehensive monograph on the mathematical foundations of infinite-dimensional statistical models. This approach can result in MCMC algorithms

---

<sup>1</sup> We believe that our results in this paper hold for non-Gaussian noise distributions as well, but proving this would be technically complex.

that are robust with respect to the refinement of the discretisation level, see, for example, [11, 12, 14, 39, 50, 52]. There are also other randomization and optimization based methods that have been recently proposed in the literature, see, for example, [2, 51].

A key property of these methods is that the prior is defined on the function space, and the discretisations automatically define corresponding prior distributions with desirable statistical properties in a principled manner. This is related to modern Tikhonov–Phillips regularisation methods widely used in applied mathematics, see [4] for a comprehensive overview. In the context of infinite-dimensional models, MAP estimators are non-trivial to define in a mathematically precise way on the infinite-dimensional function space, but several definitions of MAP estimators, various weak consistency results under the small noise limit, and posterior contraction rates have been shown in recent years, see, for example, [10, 16, 19, 23, 25, 34, 36, 49]. Some other important works on similar models and/or associated filtering/smoothing algorithms include [6, 22, 28]. These results are very interesting from a mathematical and statistical point of view; however, the intuitive meaning of some of the necessary conditions, and their algorithmic implications are difficult to grasp.

In contrast to these works, our results in this paper concern the finite-dimensional setting that is the most frequently used one in the data assimilation community. By working in finite dimensions, we are able to show consistency results and convergence rates for the MAP estimators under small observation noise/high observation frequency limits under rather weak assumptions. (In particular, in Section 4.4 of Paulin et al. [38] our key assumption on the dynamics was verified in 100 trials when  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{f}$  were randomly chosen, and only the first component of the system was observed, and they were always satisfied.) Moreover, previous work in the literature has not said anything about the computational complexity of actually finding the MAP estimators, which is a non-trivial problem in nonlinear setting due to the existence of local maxima for the log-likelihood. In our paper we propose appropriate initial estimators and show that Newton's method started from them converges to the true MAP with high probability in the small noise/high observation frequency scenario when started from this initial estimator.

It is important to mention that the MAP estimator forms the basis of the 4D-Var method introduced in Le Dimet and Talagrand [29] and Talagrand and Courtier [44] that is widely used in weather forecasting. A key methodological innovation of this method is that the gradients of the log-likelihood are computed via the adjoint equations, so that each gradient evaluation takes a similar amount of computation effort as a single run of the model. This has allowed the application of the method on large-scale models with up to  $d = 10^9$  dimensions. See Dimet and Shutyaev [17] for some theoretical results, and Navon [35], Bannister [1] for an overview of some recent advances. The present paper offers rigorous statistical foundations for this method for the class of nonlinear systems defined by (1.1).

The structure of the paper is as follows. In Sect. 1.1, we state some preliminary results for systems of the type (1.1). Section 2 contains our main results: Gaussian approximations, asymptotic optimality of MAP estimators and approximation of MAP estimators via Newton's method with precision guarantees. In Sect. 3 we apply our algorithm to the Lorenz 96' model. Section 4 contains some preliminary results, and

Sect. 5 contains the proofs of our main results. Finally, ‘‘Appendix’’ contains the proofs of our preliminary results based on concentration inequalities for empirical processes.

### 1.1 Preliminaries

Some notations and basic properties of systems of the form (1.1) are now detailed below. The one-parameter solution semigroup will be denoted by  $\Psi_t$ ; thus, for a starting point  $\mathbf{v} = (v_1, \dots, v_d) \in \mathbb{R}^d$ , the solution of (1.1) will be denoted by  $\Psi_t(\mathbf{v})$ , or equivalently,  $\mathbf{v}(t)$ . Sanz-Alonso and Stuart [42] and Law et al. [27] have assumed that the nonlinearity is energy conserving, i.e.  $\langle \mathbf{B}(\mathbf{v}, \mathbf{v}), \mathbf{v} \rangle = 0$  for every  $\mathbf{v} \in \mathbb{R}^d$ . They also assume that the linear operator  $\mathbf{A}$  is positive definite, i.e. there is a  $\lambda_A > 0$  such that  $\langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle \geq \lambda_A \langle \mathbf{v}, \mathbf{v} \rangle$  for every  $\mathbf{v} \in \mathbb{R}^d$ . As explained on page 50 of Law et al. [27], (1.1) together with these assumptions above implies that for every  $\mathbf{v} \in \mathbb{R}^d$ ,

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{v}(t)\|^2 \Big|_{t=0} \leq \frac{1}{2\lambda_A} \|\mathbf{f}\|^2 - \frac{\lambda_A}{2} \|\mathbf{v}\|^2. \tag{1.4}$$

From (1.4) one can show that  $\mathcal{B}_R$  is an absorbing set for any

$$R \geq \frac{\|\mathbf{f}\|}{\lambda_A}; \tag{1.5}$$

thus, all paths enter into this set, and they cannot escape from it once they have reached it. This in turn implies the existence of a global attractor (see, for example, [45], or Chapter 2 of Stuart and Humphries [43]). Moreover, the trapping ball assumption (1.3) holds.

For  $t \geq 0$ , let  $\mathbf{v}(t)$  and  $\mathbf{w}(t)$  denote the solutions of (1.1) started from some points  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^d$ . Based on (1.1), we have that for any two points  $\mathbf{v}, \mathbf{w} \in \mathcal{B}_R$ , any  $t \geq 0$ ,

$$\frac{d}{dt} \langle \mathbf{v}(t) - \mathbf{w}(t), \mathbf{v}(t) - \mathbf{w}(t) \rangle = -\langle \mathbf{A}(\mathbf{v}(t) - \mathbf{w}(t)), \mathbf{v}(t) - \mathbf{w}(t) \rangle - \langle \mathbf{B}(\mathbf{v}(t), \mathbf{v}(t) - \mathbf{w}(t)) - \mathbf{B}(\mathbf{w}(t) - \mathbf{v}(t), \mathbf{w}(t)), \mathbf{v}(t) - \mathbf{w}(t) \rangle,$$

and therefore by Grönwall’s lemma, we have that for any  $t \geq 0$ ,

$$\exp(-Gt) \|\mathbf{v} - \mathbf{w}\| \leq \|\mathbf{v}(t) - \mathbf{w}(t)\| \leq \exp(Gt) \|\mathbf{v} - \mathbf{w}\|, \tag{1.6}$$

for a constant  $G := \|\mathbf{A}\| + 2\|\mathbf{B}\|R$ , where

$$\|\mathbf{A}\| := \sup_{\mathbf{v} \in \mathbb{R}^d: \|\mathbf{v}\|=1} \|\mathbf{A}\mathbf{v}\| \quad \text{and} \quad \|\mathbf{B}\| := \sup_{\mathbf{v}, \mathbf{w} \in \mathbb{R}^d: \|\mathbf{v}\|=1, \|\mathbf{w}\|=1} \|\mathbf{B}(\mathbf{v}, \mathbf{w})\|.$$

For  $t \geq 0$ , let  $\Psi_t(\mathcal{B}_R) := \{\Psi_t(\mathbf{v}) : \mathbf{v} \in \mathcal{B}_R\}$ , then by (1.6), it follows that  $\Psi_t : \mathcal{B}_R \rightarrow \Psi_t(\mathcal{B}_R)$  is a one-to-one mapping, which has an inverse that we denote as  $\Psi_{-t} : \Psi_t(\mathcal{B}_R) \rightarrow \mathcal{B}_R$ .

The main quantities of interest of this paper are the smoothing and filtering distributions corresponding to the conditional distribution of  $\mathbf{u}(t_0)$  and  $\mathbf{u}(t_k)$ , respectively,

given the observations  $\mathbf{Y}_{0:k} := \{\mathbf{Y}_0, \dots, \mathbf{Y}_k\}$ . The densities of these distributions will be denoted by  $\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})$  and  $\mu^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})$ . To make our notation more concise, we define the observed part of the dynamics as

$$\Phi_t(\mathbf{v}) := \mathbf{H}\Psi_t(\mathbf{v}), \tag{1.7}$$

for any  $t \in \mathbb{R}$  and  $\mathbf{v} \in \mathcal{B}_R$ . Using these notations, the densities of the smoothing and filtering distributions can be expressed as

$$\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) = \left[ \prod_{i=0}^k \eta(\mathbf{Y}_i - \Phi_{t_i}(\mathbf{v})) \right] \frac{q(\mathbf{v})}{Z_k^{\text{sm}}} \text{ for } \mathbf{v} \in \mathcal{B}_R, \text{ and } 0 \text{ for } \mathbf{v} \notin \mathcal{B}_R \tag{1.8}$$

$$\mu^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) = \left[ \prod_{i=0}^k \eta(\mathbf{Y}_i - \Phi_{t_i - t_k}(\mathbf{v})) \right] |\det(\mathbf{J}\Psi_{-t_k}(\mathbf{v}))| \frac{q(\Psi_{-t_k}(\mathbf{v}))}{Z_k^{\text{fi}}} \text{ for } \mathbf{v} \in \Psi_{t_k}(\mathcal{B}_R),$$

and 0 for  $\mathbf{v} \notin \Psi_{t_k}(\mathcal{B}_R)$ , (1.9)

where  $\det$  stands for determinant, and  $Z_k^{\text{sm}}, Z_k^{\text{fi}}$  are normalising constants independent of  $\mathbf{v}$ . Since the determinant of the inverse of a matrix is the inverse of its determinant, we have the equivalent formulation

$$\det(\mathbf{J}\Psi_{-t_k}(\mathbf{v})) = \left( \det(\mathbf{J}\Psi_{-t_k}(\mathbf{v})\Psi_{t_k}) \right)^{-1}. \tag{1.10}$$

We assume a prior  $q$  on the initial condition that is absolutely continuous with respect to the Lebesgue measure, and zero outside the ball  $\mathcal{B}_R$  [where the value of  $R$  is determined by the trapping ball assumption (1.3)].

For  $k \geq 1$ , we define the  $k$ th Jacobian of a function  $g : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$  at point  $\mathbf{v}$  as a  $k + 1$ -dimensional array, denoted by  $\mathbf{J}^k g(\mathbf{v})$  or equivalently  $\mathbf{J}_{\mathbf{v}}^k g$ , with elements

$$(\mathbf{J}^k g(\mathbf{v}))_{i_1, \dots, i_{k+1}} := \frac{\partial^k}{\partial v_{i_1} \dots \partial v_{i_k}} g_{i_{k+1}}(\mathbf{v}), \quad 1 \leq i_1, \dots, i_k \leq d_1, \quad 1 \leq i_{k+1} \leq d_2.$$

We define the norm of this  $k$ th Jacobian as

$$\|\mathbf{J}^k g(\mathbf{v})\| := \sup_{\mathbf{v}^{(1)} \in \mathbb{R}^{d_1}, \dots, \mathbf{v}^{(k)} \in \mathbb{R}^{d_1}, \mathbf{v}^{(k+1)} \in \mathbb{R}^{d_2} : \|\mathbf{v}^{(j)}\| \leq 1, 1 \leq j \leq k+1} (\mathbf{J}^k g(\mathbf{v})) \left[ \mathbf{v}^{(1)}, \dots, \mathbf{v}^{(k+1)} \right],$$

where for a  $k + 1$ -dimensional  $d_1 \times \dots \times d_{k+1}$ -sized array  $\mathbf{M}$ , we denote

$$\mathbf{M} \left[ \mathbf{v}^{(1)}, \dots, \mathbf{v}^{(k+1)} \right] := \sum_{1 \leq i_1 \leq d_1, \dots, 1 \leq i_{k+1} \leq d_{k+1}} M_{i_1, \dots, i_{k+1}} \cdot v_{i_1}^{(1)} \cdot \dots \cdot v_{i_{k+1}}^{(k+1)}.$$

Using (1.1) and (1.3), we have that

$$\sup_{\mathbf{v} \in \mathcal{B}_R, t \geq 0} \left\| \frac{d\mathbf{v}(t)}{dt} \right\| \leq v_{\max} := \|\mathbf{A}\|R + \|\mathbf{B}\|R^2 + \|\mathbf{f}\|, \tag{1.11}$$

$$\sup_{\mathbf{v} \in \mathcal{B}_R, t \geq 0} \left\| \mathbf{J}_{\mathbf{v}(t)} \left( \frac{d\mathbf{v}(t)}{dt} \right) \right\| \leq a_{\max} := \|\mathbf{A}\| + 2\|\mathbf{B}\|R. \tag{1.12}$$

By induction, we can show that for any  $i \geq 2$ , and any  $\mathbf{v} \in \mathbb{R}^d$ , we have

$$\mathbf{D}^i \mathbf{v} = -\mathbf{A} \cdot \mathbf{D}^{i-1} \mathbf{v} - \sum_{j=0}^{i-1} \binom{i-1}{j} \mathbf{B} \left( \mathbf{D}^j \mathbf{v}, \mathbf{D}^{i-1-j} \mathbf{v} \right). \tag{1.13}$$

From this, the following bounds follow (see Section A.1 of ‘‘Appendix’’ for a proof).

**Lemma 1.1** *For any  $i \geq 0, k \geq 1, \mathbf{v} \in \mathcal{B}_R$ , we have*

$$\|\mathbf{D}^i \mathbf{v}\| \leq C_0 (C_{\text{der}})^i \cdot i!, \tag{1.14}$$

$$\|\mathbf{J}_{\mathbf{v}}^k (\mathbf{D}^i \mathbf{v})\| \leq \left( C_{\mathbf{J}}^{(k)} \right)^i \cdot i!, \text{ where} \tag{1.15}$$

$$C_0 := R + \frac{\|\mathbf{f}\|}{\|\mathbf{A}\|}, \quad C_{\text{der}} := \|\mathbf{A}\| + \|\mathbf{B}\|R + \frac{\|\mathbf{B}\|}{\|\mathbf{A}\|} \|\mathbf{f}\|, \quad C_{\mathbf{J}}^{(k)} := 2^k (C_{\text{der}} + \|\mathbf{B}\|), \quad k \geq 1. \tag{1.16}$$

In some of our arguments we are going to use the multivariate Taylor expansion for vector-valued functions. Let  $g : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$  be  $k + 1$  times differentiable for some  $k \in \mathbb{N}$ . Then using the one-dimensional Taylor expansion of the functions  $g_i(\mathbf{a} + t\mathbf{h})$  in  $t$  (where  $g_i$  denotes the  $i$ th component of  $g$ ), one can show that for any  $\mathbf{a}, \mathbf{h} \in \mathbb{R}^{d_1}$ , we have

$$g(\mathbf{a} + \mathbf{h}) = g(\mathbf{a}) + \sum_{1 \leq j \leq k} \frac{1}{j!} \cdot \left( \mathbf{J}^j g(\mathbf{a})[\mathbf{h}^j, \cdot] \right) + \mathbf{R}_{k+1}(\mathbf{a}, \mathbf{h}), \tag{1.17}$$

where  $\mathbf{h}^j := (\mathbf{h}, \dots, \mathbf{h})$  denotes the  $j$  times repetition of  $\mathbf{h}$  and the error term  $\mathbf{R}_{k+1}(\mathbf{a}, \mathbf{h})$  is of the form

$$\mathbf{R}_{k+1}(\mathbf{a}, \mathbf{h}) := \frac{k+1}{(k+1)!} \cdot \int_{t=0}^1 (1-t)^k \mathbf{J}^{k+1} g(\mathbf{a} + t\mathbf{h})[\mathbf{h}^{k+1}, \cdot] dt, \tag{1.18}$$

whose norm can be bounded using the fact that  $\int_{t=0}^1 (1-t)^k dt = \frac{1}{k+1}$  as

$$\|\mathbf{R}_{k+1}(\mathbf{a}, \mathbf{h})\| \leq \frac{\|\mathbf{h}\|^{k+1}}{(k+1)!} \cdot \sup_{0 \leq t \leq 1} \|\mathbf{J}^{k+1} g(\mathbf{a} + t\mathbf{h})\|. \tag{1.19}$$

In order to be able to use such multivariate Taylor expansions in our setting, the existence and finiteness of  $\mathbf{J}^k \Psi_t(\mathbf{v})$  can be shown rigorously in the following way. Firstly, for  $0 < t < (C_{\mathbf{J}}^{(k)})^{-1}$ , one has

$$\mathbf{J}^k \Psi_t(\mathbf{v}) = \mathbf{J}_{\mathbf{v}}^k \left( \sum_{i=0}^{\infty} \mathbf{D}^i \mathbf{v} \cdot \frac{t^i}{i!} \right),$$

and using inequality (1.15), we can show that

$$\mathbf{J}^k \Psi_t(\mathbf{v}) = \sum_{i=0}^{\infty} \mathbf{J}_{\mathbf{v}}^k (\mathbf{D}^i \mathbf{v}) \cdot \frac{t^i}{i!}$$

is convergent and finite. For  $t \geq (C_{\mathbf{J}}^{(k)})^{-1}$ , we can express  $\Psi_t(\mathbf{v})$  as a composition  $\Psi_{t_1}(\dots(\Psi_{t_m}(\mathbf{v})))$  for  $t_1 + \dots + t_m = t$  and establish the existence of the partial derivatives by the chain rule.

After establishing the existence of the partial derivatives  $\mathbf{J}^k \Psi_t(\mathbf{v})$ , we are going to bound their norm in the following lemma (proven in Section A.1 of ‘‘Appendix’’).

**Lemma 1.2** *For any  $k \geq 1$ , let*

$$D_{\mathbf{J}}^{(k)} := 2^k (\|\mathbf{A}\| + \|\mathbf{B}\| + 2\|\mathbf{B}\|R). \tag{1.20}$$

*Then for any  $k \geq 1, T \geq 0$ , we have*

$$M_k(T) := \sup_{\mathbf{v} \in \mathcal{B}_R} \sup_{0 \leq t \leq T} \|\mathbf{J}^k \Psi_t(\mathbf{v})\| \leq \exp(D_{\mathbf{J}}^{(k)} T), \text{ and} \tag{1.21}$$

$$\widehat{M}_k(T) := \sup_{\mathbf{v} \in \mathcal{B}_R} \sup_{0 \leq t \leq T} \|\mathbf{J}^k \Phi_t(\mathbf{v})\| \leq \|\mathbf{H}\| M_k(T) \leq \|\mathbf{H}\| \exp(D_{\mathbf{J}}^{(k)} T). \tag{1.22}$$

## 2 Main Results

In this section, we present our main results. We start by introducing our assumptions. In Sect. 2.1 we show that the smoother and the filter can be well approximated by Gaussian distributions when  $\sigma_Z^2 h$  is sufficiently small. This is followed by Sect. 2.2 where based on the Gaussian approximation result we show that the maximum a posteriori (MAP) estimators are asymptotically optimal in mean square error in the  $\sigma_Z^2 h \rightarrow 0$  limit. We also show that Newton’s method can be used for calculating the MAP estimators if the initial point  $\mathbf{x}_0$  can be chosen sufficiently close to the true starting position  $\mathbf{u}$ . Finally, in Sect. 2.3 we propose estimators to use as initial point  $\mathbf{x}_0$  that satisfy this criteria when  $\sigma_Z^2 h$  and  $h$  are sufficiently small.

We start with an assumption that will be used in these results.

**Assumption 2.1** Let  $T > 0$  be fixed, and suppose that  $T = kh$ , where  $k \in \mathbb{N}$ . Suppose that  $\|\mathbf{u}\| < R$ , and that there exist constants  $h_{\max}(\mathbf{u}, T) > 0$  and  $c(\mathbf{u}, T) > 0$  such that for every  $\mathbf{v} \in \mathcal{B}_R$ , for every  $h \leq h_{\max}(\mathbf{u}, T)$  (or equivalently, every  $k \geq T/h_{\max}(\mathbf{u}, T)$ ), we have

$$\sum_{i=0}^k \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 \geq \frac{c(\mathbf{u}, T)}{h} \|\mathbf{u} - \mathbf{v}\|^2. \tag{2.1}$$

As we shall see in Proposition 2.1, this assumption follows from the following assumption on the derivatives (introduced in Paulin et al. [38]).

**Assumption 2.2** Suppose that  $\|\mathbf{u}\| < R$ , and there is an index  $j \in \mathbb{N}$  such that the system of equations in  $\mathbf{v}$  defined as

$$\mathbf{H}\mathbf{D}^i \mathbf{u} = \mathbf{H}\mathbf{D}^i \mathbf{v} \text{ for every } 0 \leq i \leq j \tag{2.2}$$

has a unique solution  $\mathbf{v} := \mathbf{u}$  in  $\mathcal{B}_R$ , and

$$\text{span} \left\{ \nabla \left( \mathbf{H}\mathbf{D}^i \mathbf{u} \right)_k : 0 \leq i \leq j, 1 \leq k \leq d_o \right\} = \mathbb{R}^d, \tag{2.3}$$

where  $(\mathbf{H}\mathbf{D}^i \mathbf{u})_k$  refers to coordinate  $k$  of the vector  $\mathbf{H}\mathbf{D}^i \mathbf{u} \in \mathbb{R}^{d_o}$  and  $\nabla$  denotes the gradient of the function in  $\mathbf{u}$ .

**Proposition 2.1** Assumption 2.2 implies Assumption 2.1.

The proof is given in Section A.1 of ‘‘Appendix’’. Assumption 2.2 was verified for the Lorenz 63’ and 96’ models in Paulin et al. [38] (for certain choices of the observation matrix  $\mathbf{H}$ ); thus, Assumption 2.1 is also valid for these models.

We denote  $\text{int}(\mathcal{B}_R) := \{\mathbf{v} \in \mathbb{R}^d : \|\mathbf{v}\| < R\}$  the interior of  $\mathcal{B}_R$ . In most of our results, we will make the following assumption about the prior  $q$ .

**Assumption 2.3** The prior distribution  $q$  is assumed to be absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d$ , supported on  $\mathcal{B}_R$ . We assume that  $v \rightarrow q(\mathbf{v})$  is strictly positive and continuous on  $\mathcal{B}_R$  and that it is 3 times continuously differentiable at every interior point of  $\mathcal{B}_R$ . Let

$$C_q^{(i)} := \sup_{\mathbf{v} \in \text{int}(\mathcal{B}_R)} \|\mathbf{J}^i \log q(\mathbf{v})\| \text{ for } i = 1, 2, 3.$$

We assume that these are finite.

After some simple algebra, the smoothing distribution for an initial point  $\mathbf{u}$  and the filtering distribution for the current position  $\mathbf{u}(T)$  can be expressed as

$$\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) = \exp \left[ -\frac{1}{2\sigma_Z^2} \sum_{i=0}^k \left( \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 + 2 \langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}), \mathbf{Z}_i \rangle \right) \right] \cdot q(\mathbf{v}) / C_k^{\text{sm}}, \tag{2.4}$$

$$\mu^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) = 1_{[\mathbf{v} \in \Psi_T(\mathcal{B}_R)]} \cdot \mu^{\text{sm}}(\Psi_{-T}(\mathbf{v})|\mathbf{Y}_{0:k}) \cdot |\det(\mathbf{J}\Psi_{-T}(\mathbf{v}))| \tag{2.5}$$

$$= 1_{[\mathbf{v} \in \Psi_T(\mathcal{B}_R)]} \cdot q(\Psi_{-T}(\mathbf{v})) \cdot |\det(\mathbf{J}\Psi_{-T}(\mathbf{v}))| / C_k^{\text{sm}} \cdot \exp \left[ -\frac{1}{2\sigma_Z^2} \sum_{i=0}^k \left( \|\Phi_{t_i}(\Psi_{-T}(\mathbf{v})) - \Phi_{t_i}(\mathbf{u})\|^2 + 2 \langle \Phi_{t_i}(\Psi_{-T}(\mathbf{v})) - \Phi_{t_i}(\mathbf{u}), \mathbf{Z}_i \rangle \right) \right], \tag{2.6}$$

where  $C_k^{\text{sm}}$  is a normalising constant independent of  $\mathbf{v}$  (but depending on  $(\mathbf{Z}_j)_{j \geq 0}$ ).

In the following sections, we will present our main results for the smoother and the filter. First, in Sect. 2.1 we are going to state Gaussian approximation results, then in Sect. 2.2 we state various results about the MAP estimators, and in Sect. 2.3 we propose an initial estimator for  $\mathbf{u}$  based on the observations  $\mathbf{Y}_0, \dots, \mathbf{Y}_k$ , to be used as a starting point for Newton’s method.

### 2.1 Gaussian Approximation

We define the matrix  $\mathbf{A}_k \in \mathbb{R}^{d \times d}$  and vector  $\mathbf{B}_k \in \mathbb{R}^d$  as

$$\mathbf{A}_k := \sum_{i=0}^k \left( \mathbf{J}\Phi_{t_i}(\mathbf{u})' \mathbf{J}\Phi_{t_i}(\mathbf{u}) + \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i] \right), \tag{2.7}$$

$$\mathbf{B}_k := \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \cdot \mathbf{Z}_i, \tag{2.8}$$

where  $\mathbf{J}\Phi_{t_i}$  and  $\mathbf{J}^2\Phi_{t_i}$  denote the first and second Jacobian of  $\Phi_{t_i}$ , respectively, and  $\mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i]$  denotes the  $d \times d$  matrix with elements

$$\left[ \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i] \right]_{i_1, i_2} = \sum_{j=1}^{d_o} (\mathbf{J}^2\Phi_{t_i}(\mathbf{u}))_{i_1, i_2, j} \mathbf{Z}_i^j \text{ for } 1 \leq i_1, i_2 \leq d.$$

If  $\mathbf{A}_k$  is positive definite, then we define the centre of the Gaussian approximation of the smoother as

$$\mathbf{u}^{\mathcal{G}} := \mathbf{u} - \mathbf{A}_k^{-1} \mathbf{B}_k \tag{2.9}$$

and define the Gaussian approximation of the smoother as

$$\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) := \frac{\det(\mathbf{A}_k)^{1/2}}{(2\pi)^{d/2} \cdot \sigma_Z^d} \cdot \exp \left[ -\frac{(\mathbf{v} - \mathbf{u}^{\mathcal{G}})' \mathbf{A}_k (\mathbf{v} - \mathbf{u}^{\mathcal{G}})}{2\sigma_Z^2} \right]. \tag{2.10}$$

If  $\mathbf{A}_k$  is not positive definite, then we define the Gaussian approximation of the smoother  $\mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  to be the  $d$ -dimensional standard normal distribution (an arbitrary choice), and  $\mathbf{u}^{\mathcal{G}} := \mathbf{0}$ . If  $\mathbf{A}_k$  is positive definite, and  $\mathbf{u}^{\mathcal{G}} \in \mathcal{B}_R$ , then we define the Gaussian approximation of the filter as

$$\begin{aligned} \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) := & \frac{\det(\mathbf{A}_k)^{1/2}}{|\det(\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))|} \cdot \frac{1}{(2\pi)^{d/2} \cdot \sigma_Z^d} \\ & \cdot \exp \left[ -\frac{(\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}}))' \left( (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} \right)' \mathbf{A}_k (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}}))}{2\sigma_Z^2} \right]. \end{aligned} \tag{2.11}$$

Alternatively, if  $\mathbf{A}_k$  is not positive definite, or  $\mathbf{u}^{\mathcal{G}} \notin \mathcal{B}_R$ , then we define the Gaussian approximation of the smoother  $\mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$  to be the  $d$ -dimensional standard normal distribution.

In order to compare the closeness between the target distributions and their Gaussian approximation, we are going to use two types of distance between distributions. The *total variation distance* of two distributions  $\mu_1, \mu_2$  on  $\mathbb{R}^d$  that are absolutely continuous with respect to the Lebesgue measure is defined as

$$d_{\text{TV}}(\mu_1, \mu_2) := \frac{1}{2} \int_{x \in \mathbb{R}^d} |\mu_1(x) - \mu_2(x)| dx, \tag{2.12}$$

where  $\mu_1(x)$  and  $\mu_2(x)$  denote the densities of the distributions.

The *Wasserstein distance* (also called first Wasserstein distance) of two distributions  $\mu_1, \mu_2$  on  $\mathbb{R}^d$  (with respect to the Euclidean distance) is defined as

$$d_{\text{W}}(\mu_1, \mu_2) := \inf_{\gamma \in \Gamma(\mu_1, \mu_2)} \int_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^d} \|\mathbf{x} - \mathbf{y}\| d\gamma(\mathbf{x}, \mathbf{y}), \tag{2.13}$$

where  $\Gamma(\mu_1, \mu_2)$  is the set of all measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\mu_1$  and  $\mu_2$ .

The following two theorems bound the total variation and Wasserstein distances between the smoother, the filter and their Gaussian approximations. In some of our bounds, the quantity  $T + h_{\max}(\mathbf{u}, T)$  appears. For brevity, we denote this as

$$\bar{T}(\mathbf{u}) := T + h_{\max}(\mathbf{u}, T). \tag{2.14}$$

We are also going to use the constant  $C_{\|A\|}$  defined as

$$C_{\|A\|} := \widehat{M}_1(T)^2 \cdot \bar{T}(\mathbf{u}) + \frac{c(\mathbf{u}, T)}{2}. \tag{2.15}$$

**Theorem 2.1** (Gaussian approximation of the smoother) *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $\mathbf{u}$  and the prior  $q$ . Then there are constants  $C_{TV}^{(1)}(\mathbf{u}, T)$ ,  $C_{TV}^{(2)}(\mathbf{u}, T)$ ,  $C_W^{(1)}(\mathbf{u}, T)$ , and  $C_W^{(2)}(\mathbf{u}, T)$  independent of  $\sigma_Z$ ,  $h$  and  $\varepsilon$  such that for any  $0 < \varepsilon \leq 1$ ,  $\sigma_Z > 0$ , and  $0 \leq h \leq h_{\max}(\mathbf{u}, T)$  satisfying that  $\sigma_Z \sqrt{h} \leq \frac{1}{2} C_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , we have*

$$\mathbb{P} \left[ \frac{c(\mathbf{u}, T)}{2h} \mathbf{I}_d < \mathbf{A}_k < \frac{C_{\|A\|}}{h} \mathbf{I}_d \right. \\ \left. \text{and } d_{TV}(\mu^{\text{sm}}(\cdot | \mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot | \mathbf{Y}_{0:k})) \leq C_{TV}(\mathbf{u}, T, \varepsilon) \sigma_Z \sqrt{h} \right] \tag{2.16}$$

$$\text{and } d_W(\mu^{\text{sm}}(\cdot | \mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot | \mathbf{Y}_{0:k})) \leq C_W(\mathbf{u}, T, \varepsilon) \sigma_Z^2 h \left| \mathbf{u} \right| \geq 1 - \varepsilon, \text{ where} \tag{2.17}$$

$$C_{TV}(\mathbf{u}, T, \varepsilon) := C_{TV}^{(1)}(\mathbf{u}, T) + C_{TV}^{(2)}(\mathbf{u}, T) \left( \log \left( \frac{1}{\varepsilon} \right) \right)^2, \text{ and} \tag{2.18}$$

$$C_W(\mathbf{u}, T, \varepsilon) := C_W^{(1)}(\mathbf{u}, T) + C_W^{(2)}(\mathbf{u}, T) \left( \log \left( \frac{1}{\varepsilon} \right) \right)^2. \tag{2.19}$$

**Theorem 2.2** (Gaussian approximation of the filter) *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $\mathbf{u}$  and the prior  $q$ . Then there are constants  $D_{TV}^{(1)}(\mathbf{u}, T)$ ,  $D_{TV}^{(2)}(\mathbf{u}, T)$ ,  $D_W^{(1)}(\mathbf{u}, T)$ , and  $D_W^{(2)}(\mathbf{u}, T)$  independent of  $\sigma_Z$ ,  $h$  and  $\varepsilon$  such that for any  $0 < \varepsilon \leq 1$ ,  $\sigma_Z > 0$ , and  $0 \leq h \leq h_{\max}(\mathbf{u}, T)$  satisfying that  $\sigma_Z \sqrt{h} \leq \frac{1}{2} D_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , we have*

$$\mathbb{P} \left[ \frac{c(\mathbf{u}, T)}{2h} \mathbf{I}_d < \mathbf{A}_k < \frac{C_{\|A\|}}{h} \mathbf{I}_d \text{ and } \mathbf{u}^{\mathcal{G}} \in \mathcal{B}_R \text{ and} \right. \\ \left. d_{TV}(\mu^{\text{fi}}(\cdot | \mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot | \mathbf{Y}_{0:k})) \leq D_{TV}(\mathbf{u}, T, \varepsilon) \sigma_Z \sqrt{h} \text{ and} \right. \\ \left. d_W(\mu^{\text{fi}}(\cdot | \mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot | \mathbf{Y}_{0:k})) \leq D_W(\mathbf{u}, T, \varepsilon) \sigma_Z^2 h \left| \mathbf{u} \right| \geq 1 - \varepsilon, \text{ where} \right] \tag{2.20}$$

$$D_{TV}(\mathbf{u}, T, \varepsilon) := D_{TV}^{(1)}(\mathbf{u}, T) + D_{TV}^{(2)}(\mathbf{u}, T) \left( \log \left( \frac{1}{\varepsilon} \right) \right)^2, \text{ and} \tag{2.21}$$

$$D_W(\mathbf{u}, T, \varepsilon) := D_W^{(1)}(\mathbf{u}, T) + D_W^{(2)}(\mathbf{u}, T) \left( \log \left( \frac{1}{\varepsilon} \right) \right)^2. \tag{2.22}$$

Note that the Gaussian approximations  $\mu_{\mathcal{G}}^{\text{sm}}$  and  $\mu_{\mathcal{G}}^{\text{fi}}$  as defined above are not directly computable based on the observations  $\mathbf{Y}_{0:k}$ , since they involve the true initial position  $\mathbf{u}$  in both their mean and covariance matrix. However, we believe that with some additional straightforward calculations one could show that results similar to Theorems 2.1

and 2.2 also hold for the Laplace approximations of the smoothing and filtering distributions (i.e. when the mean and covariance of the normal approximation is replaced by the MAP and the inverse Hessian of the log-likelihood at the MAP, respectively), which are directly computable based on the observations  $Y_{0:k}$ .

### 2.2 MAP Estimators

Let  $\bar{u}^{sm}$  be the mean of the smoothing distribution,  $\bar{u}^{fi}$  be the mean of the filtering distribution, and  $\hat{u}_{MAP}^{sm}$  be the maximum a posteriori of the smoothing distribution, i.e.

$$\hat{u}_{MAP}^{sm} := \arg \max_{v \in \mathcal{B}_R} \mu^{sm}(v | Y_{0:k}). \tag{2.23}$$

In case there are multiple maxima, we choose any of them. For the filter, we will use the push-forward MAP estimator

$$\hat{u}^{fi} := \Psi_T(\hat{u}_{MAP}^{sm}). \tag{2.24}$$

Based on the Gaussian approximation results, we prove the following two theorems about these estimators.

**Theorem 2.3** (Comparison of mean square error of MAP and posterior mean for smoother) *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $u$  and the prior  $q$ . Then there is a constant  $S_{max}^{sm}(u, T) > 0$  independent of  $\sigma_Z$  and  $h$  such that for  $0 < h < h_{max}(u, T)$ ,  $\sigma_Z \sqrt{h} \leq S_{max}^{sm}(u, T)$ , we have that*

$$\underline{C}^{sm}(u, T) \leq \frac{\mathbb{E} [\|\bar{u}^{sm} - u\|^2 | u]}{\sigma_Z^2 h} \leq \overline{C}^{sm}(u, T), \text{ and} \tag{2.25}$$

$$\left| \mathbb{E} [\|\hat{u}_{MAP}^{sm} - u\|^2 | u] - \mathbb{E} [\|\bar{u}^{sm} - u\|^2 | u] \right| \leq C_{MAP}^{sm}(u, T) (\sigma_Z^2 h)^{\frac{3}{2}}, \tag{2.26}$$

where the expectations are taken with respect to the random observations, and  $\underline{C}^{sm}(u, T)$ ,  $\overline{C}^{sm}(u, T)$ , and  $C_{MAP}^{sm}(u, T)$  are finite positive constants independent of  $\sigma_Z$  and  $h$ .

**Theorem 2.4** (Comparison of mean square error of MAP and posterior mean for filter) *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $u$  and the prior  $q$ . Then there is a constant  $S_{max}^{fi}(u, T) > 0$  independent of  $\sigma_Z$  and  $h$  such that for  $0 < h < h_{max}(u, T)$ ,  $\sigma_Z \sqrt{h} \leq S_{max}^{fi}(u, T)$ , we have that*

$$\underline{C}^{fi}(u, T) \leq \frac{\mathbb{E} [\|\bar{u}^{fi} - u(T)\|^2 | u]}{\sigma_Z^2 h} \leq \overline{C}^{fi}(u, T), \text{ and} \tag{2.27}$$

$$\left| \mathbb{E} [\|\hat{u}^{fi} - u(T)\|^2 | u] - \mathbb{E} [\|\bar{u}^{fi} - u(T)\|^2 | u] \right| \leq C_{MAP}^{fi}(u, T) (\sigma_Z^2 h)^{\frac{3}{2}}, \tag{2.28}$$

where the expectations are taken with respect to the random observations, and  $\underline{C}^{fi}(u, T)$ ,  $\overline{C}^{fi}(u, T)$ , and  $C_{MAP}^{fi}(u, T)$  are finite positive constants independent of  $\sigma_Z$  and  $h$ .

*Remark 2.1* Theorems 2.3 and 2.4 in particular imply that when  $T$  is fixed, and  $\sigma_Z\sqrt{h}$  tends to 0, the ratio between the mean square errors of the posterior mean and MAP estimators conditioned on the initial position  $\mathbf{u}$  tends to 1. Since the mean of the posterior distributions,  $\bar{\mathbf{u}}^{\text{sm}}$  (or  $\bar{\mathbf{u}}^{\text{fi}}$  for the filter), is the estimator  $U(\mathbf{Y}_{0:k})$  that minimises  $\mathbb{E}(\|U(\mathbf{Y}_{0:k}) - \mathbf{u}\|^2)$  (or  $\mathbb{E}(\|U(\mathbf{Y}_{0:k}) - \mathbf{u}(T)\|^2)$  for the filter), our results imply that the mean square error of the MAP estimators is close to optimal when  $\sigma_Z\sqrt{h}$  is sufficiently small.

Next we propose a method to compute the MAP estimators. Let  $g^{\text{sm}} : \mathcal{B}_R \rightarrow \mathbb{R}$  be

$$g^{\text{sm}}(\mathbf{v}) := -\log(q(\mathbf{v})) + \frac{1}{2\sigma_Z^2} \sum_{i=0}^k \|\mathbf{Y}_i - \Phi_{t_i}(\mathbf{v})\|^2. \tag{2.29}$$

Then  $-g^{\text{sm}}(\mathbf{v})$  is the log-likelihood of the smoother, except that it does not contain the normalising constant term.

The following theorem shows that Newton’s method can be used to compute  $\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}$  to arbitrary precision if it is initiated from a starting point  $\mathbf{x}_0$  that is sufficiently close to the initial position  $\mathbf{u}$ . The proof is based on the concavity properties of the log-likelihood near  $\mathbf{u}$ . Based on this, an approximation for the push-forward MAP estimator  $\hat{\mathbf{u}}^{\text{fi}}$  can be then computed by moving forward the approximation of  $\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}$  by time  $T$  according to the dynamics  $\Psi_T$ . [This will not increase the error by more than a factor of  $\exp(GT)$  according to (1.6)].

**Theorem 2.5** (Convergence of Newton’s method to the MAP) *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $\mathbf{u}$  and the prior  $q$ . Then for every  $0 < \varepsilon \leq 1$ , there exist finite constants  $S_{\text{max}}^{\text{sm}}(\mathbf{u}, T, \varepsilon)$ ,  $N^{\text{sm}}(\mathbf{u}, T)$  and  $D_{\text{max}}^{\text{sm}}(\mathbf{u}, T) \in (0, N^{\text{sm}}(\mathbf{u}, T)]$  (defined in (5.62) and (5.63)) such that the following holds. If  $\sigma_Z\sqrt{h} \leq S_{\text{max}}^{\text{sm}}(\mathbf{u}, T, \varepsilon)$ , and the initial point  $\mathbf{x}_0 \in \mathcal{B}_R$  satisfies that  $\|\mathbf{x}_0 - \mathbf{u}\| < D_{\text{max}}^{\text{sm}}(\mathbf{u}, T)$ , then the iterates of Newton’s method defined recursively as*

$$\mathbf{x}_{i+1} := \mathbf{x}_i - (\nabla^2 g^{\text{sm}}(\mathbf{x}_i))^{-1} \cdot \nabla g^{\text{sm}}(\mathbf{x}_i) \text{ for } i \in \mathbb{N} \tag{2.30}$$

satisfy that

$$\mathbf{P}\left(\text{for every } i \in \mathbb{N}, \mathbf{x}_i \text{ is well defined and } \|\mathbf{x}_i - \hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}\| \leq N^{\text{sm}}(\mathbf{u}, T) \left(\frac{\|\mathbf{x}_0 - \mathbf{u}\|}{N^{\text{sm}}(\mathbf{u}, T)}\right)^{2^i} \mid \mathbf{u}\right) \geq 1 - \varepsilon. \tag{2.31}$$

*Remark 2.2* The bound (2.31) means that the number of digits of precision essentially doubles in each iteration. In other words, only a few iterations are needed to approximate the MAP estimator with high precision if  $\mathbf{x}_0$  is sufficiently close to  $\mathbf{u}$ .

### 2.3 Initial Estimator

First, we are going to estimate the derivatives  $HD^l u$  for  $l \in \mathbb{N}$  based on observations  $Y_{0:k}$ . For technical reasons, the estimators will depend on  $Y_{0:\hat{k}}$  for some  $0 \leq \hat{k} \leq k$  (which will be chosen depending on  $l$ ). For any  $j \in \mathbb{N}$ , we define  $v^{(j|\hat{k})} \in \mathbb{R}^{\hat{k}+1}$  as

$$v^{(j|\hat{k})} := \left\{ \binom{i}{\hat{k}}^j \right\}_{0 \leq i \leq \hat{k}}, \text{ with the convention that } 0^0 := 1. \tag{2.32}$$

For  $j_{\max} \in \mathbb{N}$ , we define  $M^{(j_{\max}|\hat{k})} \in \mathbb{R}^{(j_{\max}+1) \times (\hat{k}+1)}$  as a matrix with rows  $v^{(0|\hat{k})}, \dots, v^{(j_{\max}|\hat{k})}$ . We denote by  $I_{j_{\max}+1}$  the identity matrix of dimension  $j_{\max} + 1$ , and by  $e^{(l|j_{\max})}$  a column vector in  $\mathbb{R}^{j_{\max}+1}$  whose every component is zero except the  $l + 1$ th one which is 1. For any  $l \in \mathbb{N}$ ,  $j_{\max} \geq l$ ,  $\hat{k} \geq j_{\max}$ , we define the vector

$$c^{(l|j_{\max}|\hat{k})} := \frac{l!}{(\hat{k}h)^l} (M^{(j_{\max}|\hat{k})})' \left( M^{(j_{\max}|\hat{k})} (M^{(j_{\max}|\hat{k})})' \right)^{-1} \cdot e^{(l|j_{\max})}. \tag{2.33}$$

Then

$$\hat{\Phi}^{(l|j_{\max})}(Y_{0:\hat{k}}) := \sum_{i=0}^{\hat{k}} c_i^{(l|j_{\max}|\hat{k})} \cdot Y_i \tag{2.34}$$

is an estimator of  $HD^l u$ . The fact that the matrix  $M^{(j_{\max}|\hat{k})} (M^{(j_{\max}|\hat{k})})'$  is invertible follows from the fact that  $v^{(0|\hat{k})}, \dots, v^{(j_{\max}|\hat{k})}$  are linearly independent (since the matrix with rows  $v^{(0|\hat{k})}, \dots, v^{(j_{\max}|\hat{k})}$  is the so-called Vandermonde matrix whose determinant is nonzero). From (2.33), it follows that the norm of  $c^{(l|j_{\max}|\hat{k})}$  can be expressed as

$$\|c^{(l|j_{\max}|\hat{k})}\| = \frac{l!}{(\hat{k}h)^l} \sqrt{\left[ \left( M^{(j_{\max}|\hat{k})} (M^{(j_{\max}|\hat{k})})' \right)^{-1} \right]_{l+1,l+1}}. \tag{2.35}$$

To lighten the notation, for  $j_{\max} \geq l$  and  $\hat{k} \geq j_{\max}$ , we will denote

$$C_M^{(l|j_{\max}|\hat{k})} := \sqrt{\hat{k} \cdot \left[ \left( M^{(j_{\max}|\hat{k})} (M^{(j_{\max}|\hat{k})})' \right)^{-1} \right]_{l+1,l+1}}. \tag{2.36}$$

The next proposition gives an error bound for this estimator, which we will use for choosing the values  $\hat{k}$  and  $j_{\max}$  given  $l$ .

**Proposition 2.2** *Suppose that  $j_{\max} \geq l$  and  $\hat{k} \geq 2j_{\max} + 3$ . Then for any  $0 < \varepsilon \leq 1$ ,*

$$\begin{aligned} & \mathbb{P} \left[ \left\| \hat{\Phi}^{(l|j_{\max})}(Y_{0:\hat{k}}) - \mathbf{H} \mathbf{D}^l \mathbf{u} \right\| \right] \\ & \geq C_M^{(l|j_{\max}|\hat{k})} \cdot l! \cdot g(l, j_{\max}, \hat{k}) \cdot \sqrt{1 + \frac{\log(1/\varepsilon)}{\log(d_o + 1)} \|\mathbf{u}\|} \leq \varepsilon, \end{aligned}$$

where

$$g(l, j_{\max}, \hat{k}) := \frac{C_0 \|\mathbf{H}\| C_{\text{der}}^{j_{\max}+1}}{\sqrt{j_{\max} + 3/2}} \cdot (\hat{k}h)^{j_{\max}+1-l} + (\hat{k}h)^{-l-1/2} \sigma_Z \sqrt{h} \sqrt{2d_o \log(d_o + 1)}. \tag{2.37}$$

The following lemma shows that as  $\hat{k} \rightarrow \infty$ , the constant  $C_M^{(l|j_{\max}|\hat{k})}$  tends to a limit.

**Lemma 2.6** *Let  $\mathbf{K}^{(j_{\max})} \in \mathbb{R}^{j_{\max}+1 \times j_{\max}+1}$  be a matrix with elements  $\mathbf{K}_{i,j}^{(j_{\max})} := \frac{1}{i+j-1}$  for  $1 \leq i, j \leq j_{\max} + 1$ . Then for any  $l \in \mathbb{N}$ ,  $j_{\max} \geq l$ , the matrix  $\mathbf{K}^{(j_{\max})}$  is invertible, and*

$$\lim_{\hat{k} \rightarrow \infty} C_M^{(l|j_{\max}|\hat{k})} = \left[ \left( \mathbf{K}^{(j_{\max})} \right)^{-1} \right]_{l+1, l+1}. \tag{2.38}$$

The proofs of the above two results are included in Sect. 5.4. Based on these results, we choose  $\hat{k} \in \{2j_{\max} + 3, \dots, k\}$  such that the function  $g(l, j_{\max}, \hat{k})$  is minimised. We denote this choice of  $\hat{k}$  by  $\hat{k}_{\text{opt}}(l, j_{\max})$ . (If  $g(l, j_{\max}, \hat{k})$  takes the same value for several  $\hat{k}$ , then we choose the smallest of them.) By taking the derivative of  $g(l, j_{\max}, \hat{k})$  in  $\hat{k}$ , it is easy to see that it has a single minimum among positive real numbers achieved at

$$\hat{k}_{\min}(l, j_{\max}) := \frac{1}{h} \cdot \left( \frac{\sigma_Z \sqrt{h} \sqrt{d_o \log(d_o + 1)} (j_{\max} + 3/2) (l + 1/2)}{(j_{\max} + 1 - l) C_0 \|\mathbf{H}\| C_{\text{der}}^{j_{\max}+1}} \right)^{1/(j_{\max}+3/2)}. \tag{2.39}$$

Based on this, we have

$$\begin{aligned} \hat{k}_{\text{opt}}(l, j_{\max}) & := 1_{\hat{k}_{\min}(l, j_{\max}) \leq 2j_{\max}+3} \cdot (2j_{\max} + 3) + 1_{\hat{k}_{\min}(l, j_{\max}) \geq k} \cdot k \\ & + 1_{2j_{\max}+3 < \hat{k}_{\min}(l, j_{\max}) < k} \cdot \arg \min_{\hat{k} \in \{\lfloor \hat{k}_{\min}(l, j_{\max}) \rfloor, \lceil \hat{k}_{\min}(l, j_{\max}) \rceil\}} g(l, j_{\max}, \hat{k}). \end{aligned} \tag{2.40}$$

Finally, based on the definition of  $\hat{k}_{\text{opt}}(l, j_{\max})$ , we choose  $j_{\max}^{\text{opt}}(l)$  as

$$j_{\max}^{\text{opt}}(l) := \arg \min_{l \leq j_{\max} \leq J_{\max}^{(l)}} \left( C_M^{(l|j_{\max}|\hat{k}_{\text{opt}}(l, j_{\max}))} \cdot g(l, j_{\max}, \hat{k}_{\text{opt}}(l, j_{\max})) \right), \tag{2.41}$$

where  $J_{\max}^{(l)} \in \{l, l + 1, \dots, \lfloor (k - 3)/2 \rfloor\}$  is a parameter to be tuned. We choose the smallest possible  $J_{\max}$  where the minimum is taken. Based on these notations, we define our estimator for  $\mathbf{HD}^l \mathbf{u}$  as

$$\hat{\Phi}^{(l)} := \hat{\Phi}^{(l|J_{\max}^{\text{opt}}(l))}(\mathbf{Y}_{0:\hat{k}_{\text{opt}}(l, J_{\max}^{\text{opt}})}). \tag{2.42}$$

The following theorem bounds the error of this estimator.

**Theorem 2.7** *Suppose that  $\mathbf{u} \in \mathcal{B}_R$ , and  $T = kh$ . Then for any  $l \in \mathbb{N}$ , there exist some positive constants  $h_{\max}^{(l)}, s_{\max}^{(l)}(T)$  and  $S_{\max}^{(l)}(T)$  such that for any choice of the parameter  $J_{\max}^{(l)} \in \{l, l + 1, \dots, \lfloor (k - 3)/2 \rfloor\}$ , any  $\varepsilon > 0, 0 < s \leq s_{\max}^{(l)}(T), 0 < h \leq \frac{h_{\max}^{(l)} \cdot s}{\sqrt{1 + \frac{\log(1/\varepsilon)}{\log(d_o + 1)}}}$ ,*

$$0 \leq \sigma_Z \sqrt{h} \leq S_{\max}^{(l)}(T) \cdot \left( \frac{s}{\sqrt{1 + \frac{\log(1/\varepsilon)}{\log(d_o + 1)}}} \right)^{l+3/2},$$

$$\mathbb{P} \left( \left\| \mathbf{HD}^l \mathbf{u} - \hat{\Phi}^{(l)} \right\| \geq s \mid \mathbf{u} \right) \leq \varepsilon.$$

The following theorem proposes a way of estimating  $\mathbf{u}$  from estimates for the derivatives  $(\mathbf{HD}^i \mathbf{u})_{0 \leq i \leq j}$ . This will be used as our initial estimator for Newton’s method.

**Theorem 2.8** *Suppose that for some  $j \in \mathbb{N}$  there is a function  $F : (\mathbb{R}^{d_o})^{j+1} \rightarrow \mathbb{R}^d$  independent of  $\mathbf{u}$  such that*

$$F(\mathbf{Hu}, \dots, \mathbf{HD}^j \mathbf{u}) = \mathbf{u}, \text{ and} \tag{2.43}$$

$$\|F(\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(j)}) - \mathbf{u}\| \leq C_F(\mathbf{u}) \cdot \left( \sum_{i=0}^j \|\mathbf{HD}^i \mathbf{u} - \mathbf{x}^{(i)}\|^2 \right)^{1/2} \tag{2.44}$$

for  $\sum_{i=0}^j \|\mathbf{HD}^i \mathbf{u} - \mathbf{x}^{(i)}\|^2 \leq D_F(\mathbf{u})$ , for some positive constants  $C_F(\mathbf{u}), D_F(\mathbf{u})$ . Then  $F(\hat{\Phi}^{(0)}, \dots, \hat{\Phi}^{(j)})$  satisfies that if  $\sum_{i=0}^j \|\mathbf{HD}^i \mathbf{u} - \hat{\Phi}^{(i)}\|^2 \leq D_F(\mathbf{u})$ , then

$$\|F(\hat{\Phi}^{(0)}, \dots, \hat{\Phi}^{(j)}) - \mathbf{u}\| \leq C_F(\mathbf{u}) \cdot \left( \sum_{i=0}^j \|\mathbf{HD}^i \mathbf{u} - \hat{\Phi}^{(i)}\|^2 \right)^{1/2}. \tag{2.45}$$

In particular, under Assumption 2.2 and for  $j$  as determined therein, function  $F$  defined as

$$F(\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(j)}) := \arg \min_{\mathbf{v} \in \mathcal{B}_R} \sum_{i=0}^j \|\mathbf{HD}^i \mathbf{v} - \mathbf{x}^{(i)}\|^2 \tag{2.46}$$

satisfies conditions (2.43) and (2.44).

Thus, the initial estimator can simply be chosen as

$$x_0 := \arg \min_{v \in \mathcal{B}_R} \sum_{i=0}^j \left\| \mathbf{H} \mathbf{D}^i v - \hat{\phi}^{(i)} \right\|^2, \tag{2.47}$$

with the above two theorems implying that the estimate gets close to  $u$  for decreasing  $\sigma_Z \sqrt{h}$ . Solving polynomial sum of squares minimisation problems of this type is a well-studied problem in optimization theory (see [26] for a theoretical overview), and several toolboxes are available (see [40,47]). Besides (2.46), other problem-specific choices of  $F$  satisfying conditions (2.43) and (2.44) can also be used, as we explain in Sect. 3.2 for the Lorenz 96’ model.

### 2.4 Optimization Based Smoothing and Filtering

The following algorithm provides an estimator of  $u$  given  $Y_{0:k}$ . We assume that either there is a problem-specific  $F$  satisfying conditions (2.43) and (2.44) for some  $j \in \mathbb{N}$ , or we suppose that Assumption 2.2 is satisfied for the true initial point  $u$ , and use  $F$  as defined in (2.46).

---

#### Algorithm 1 Optimization based smoothing

**Input:**  $k \in \mathbb{N}$  (window size parameter),  $\Delta_{\min} > 0$  (minimum step size parameter),  $Y_{0:k}$  (observations).

**Step 1:** We compute the estimators  $\hat{\phi}^{(0)}, \dots, \hat{\phi}^{(j)}$  based on (2.42), and set the initial point as  $x_0 := F(\hat{\phi}^{(0)}, \dots, \hat{\phi}^{(j)})$ .

**Step 2:** We compute the iterates  $x_i$  for  $i \geq 1$  based on (2.30) recursively until  $\|x_i - x_{i-1}\|$  becomes smaller than  $\Delta_{\min}$ , and return  $\hat{u} = x_n$  for  $n := \min_{i \in \mathbb{Z}_+} \|x_i - x_{i-1}\| < \Delta_{\min}$ .

---

The following algorithm returns an online estimator of  $(u(t_i))_{i \geq 0}$  given  $Y_{0:i}$  at time  $t_i$ .

---

#### Algorithm 2 Optimization based filtering

**Input:**  $k \in \mathbb{N}$  (window size parameter),  $\Delta_{\min} > 0$  (minimum step size parameter),  $(Y_i)_{i \geq 0}$  (observations come consecutively in time).

**Step 1:** For  $i < k$ , return the estimate  $\widehat{u}(t_i) = 0$ .

**Step 2:** For  $i \geq k$ , we first compute the estimate  $\hat{u}^{(i-k)}$  of  $u(t_{i-k})$  based on Algorithm 1 applied on  $Y_{i-k:i}$ , and then return  $\widehat{u}(t_i) = \psi_T(\hat{u}^{(i-k)})$ .

---

This algorithm can be modified to run Step 2 only at every  $K$  step for some  $K \in \mathbb{Z}_+$  (i.e. for  $i = k + lK$  for  $l \in \mathbb{N}$ ) and propagate forward the estimate of the previous time we ran Step 2 at the intermediate time points. This increases the execution speed at the cost of the loss of some precision (depending on the choice of  $K$ ).

Based on our results in the previous sections, we can see that if the assumptions of the results hold and  $\Delta_{\min}$  is chosen sufficiently small then the estimation errors are

of  $O(\sigma_Z\sqrt{h})$  with high probability for both algorithms (in the second algorithm, for  $i \geq k$ ).

### 3 Application to the Lorenz 96' Model

The Lorenz 96' model is a  $d$ -dimensional chaotic dynamical system that was introduced in Lorenz [31]. In its original form, it is written as

$$\frac{d}{dt}u_i = -u_{i-1}u_{i-2} + u_{i-1}u_{i+1} - u_i + f, \tag{3.1}$$

where the indices are understood modulo  $d$ , and  $f$  is the so-called forcing constant. In this paper we are going to fix this as  $f = 8$ . (This is a commonly used value that is experimentally known to cause chaotic behaviour, see [32,33].) As shown on page 16 of Sanz-Alonso and Stuart [42], this system can be written in the form (1.1), and the bilinear form  $\mathbf{B}(\mathbf{u}, \mathbf{u})$  satisfies the energy-conserving property (i.e.  $\langle \mathbf{B}(\mathbf{v}, \mathbf{v}), \mathbf{v} \rangle = 0$  for any  $\mathbf{v} \in \mathbb{R}^d$ ).

We consider 2 observation scenarios for this model. In the first scenario, we assume that  $d$  is divisible by 6 and choose  $\mathbf{H}$  such that coordinates 1, 2, 3, 7, 8, 9, . . . ,  $d - 5, d - 4, d - 3$  are observed directly, i.e. each observed batch of 3 is followed by a non-observed batch of 3. In this case, the computational speed is fast, and we are able to obtain simulation results for high dimensions. We consider first a small-dimensional case ( $d = 12$ ) to show the dependence of the MSE of the MAP estimator on the parameter  $k$  (the amount of observations), when the parameters  $\sigma_Z$  and  $h$  are fixed. After this, we consider a high-dimensional case ( $d = 1,000,002$ ) and look at the dependence of the MSE of the MAP estimator on  $\sigma_Z\sqrt{h}$ .

In the second scenario, we choose  $\mathbf{H}$  such that we observe the first 3 coordinates directly. We present some simulation results for  $d = 60$  dimensions for this scenario.

In Paulin et al. [38], we have shown that in the second scenario, the system satisfies Assumption 2.2 for Lebesgue-almost every initial point  $\mathbf{u} \in \mathbb{R}^d$ . A simple modification of that argument shows that Assumption 2.2 holds for Lebesgue-almost every initial point  $\mathbf{u} \in \mathbb{R}^d$  in the first observation scenario, too.

In each case, we have set the initial point as  $\mathbf{u} = (\frac{d+1}{2d}, \frac{d+2}{2d}, \dots, 1)$ . (We have tried different randomly chosen initial points and obtained similar results.) Figures 1, 2 and 3 show the simulation results when applying Algorithm 1 (optimization based smoother) to each of these cases. Note that Algorithm 2 (optimization based filter) applied to this setting yields very similar results.

In Fig. 1, we can see that the MAP estimators RMSE (root-mean-square error) does not seem to decrease significantly after a certain amount of observations. This is consistent with the non-concentration of the smoother due to the existence of leaf sets, described in Paulin et al. [38]. Moreover, by increasing  $k$  above 100, we have observed that the Newton's method often failed to improve significantly over the initial estimator, and the RMSE of the estimator became of order  $10^{-1}$ , significantly worse than for smaller values of  $k$ . We believe that this is due to the fact that as we increase  $k$ , while keeping  $\sigma_Z$  and  $h$  fixed, the normal approximation of the smoother breaks down, and the smoother becomes more and more multimodal. Due to this, we are unable to

find the true MAP when starting from the initial estimator and settle down at another mode. To conclude, for optimal performance, it is important to tune the parameter  $k$  of the algorithm.

In Fig. 2, we present results for a  $d = 1,000,002$ -dimensional Lorenz 96' system, with half of the coordinates observed. The observation time is  $T = 10^{-5}$ . The circles correspond to data points with  $h = 10^{-6}$  (so  $k = 10$ ), while the triangles correspond to data points with  $h = 2 \cdot 10^{-7}$  (so  $k = 50$ ). The plots show that the method works as expected for this high-dimensional system and that the RMSE of the estimator is proportional to  $\sigma_Z \sqrt{h}$ .

Finally, in Fig. 3, we present results for a  $d = 60$ -dimensional system with the first 3 coordinates observed. The observation time is  $T = 10^{-3}$ . The circles correspond to data points with  $h = 5 \cdot 10^{-5}$ , while the triangles correspond to data points with  $h = 2.5 \cdot 10^{-5}$ . We can see that Algorithm 1 is able to handle a system which has only a small fraction of its coordinates observed. Note that the calculations are done with numbers having 360 decimal digits of precision. Such high precision is necessary because the interaction between the 3 observed coordinates of the system and some of the non-observed coordinates is weak.

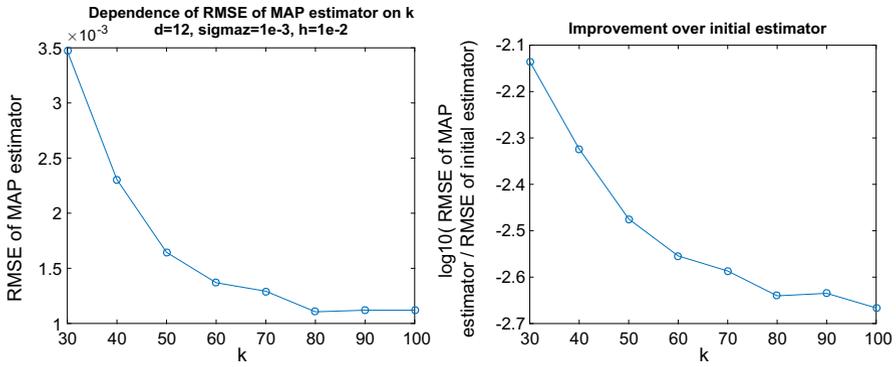
The requirements on the observations noise  $\sigma_Z$  and observation time step  $h$  for the applicability of our method depend heavily on the parameters of the model (such as the dimension  $d$ ) and on the observation matrix  $\mathbf{H}$ . In the first simulation (Fig. 1), relatively large noise ( $\sigma_Z = 10^{-3}$ ) and large time step ( $h = 10^{-2}$ ) were possible. For the second simulation (Fig. 1), due to the high dimensionality of the model ( $d = 1,000,002$ ), and the sparse approximation used in the solver, we had to choose smaller time step ( $h \leq 10^{-6}$ ) and smaller observation noise ( $\sigma_Z \leq 10^{-6}$ ). Finally, in the third simulation (Fig. 3), since only a very small fraction of the  $d = 60$  coordinates is observed, the observation noise has to be very small in order for us to be able to recover the unobserved coordinates ( $\sigma_Z \leq 10^{-120}$ ).

In all of the above cases, the error of the MAP estimator is several orders of magnitude less than the error of the initial estimator. When comparing these results with the simulation results of Law et al. [28] using the 3D-Var and extended Kalman filter methods for the Lorenz 96' model, it seems that our method improves upon them, since it allows for larger dimensions and smaller fraction of coordinates observed.

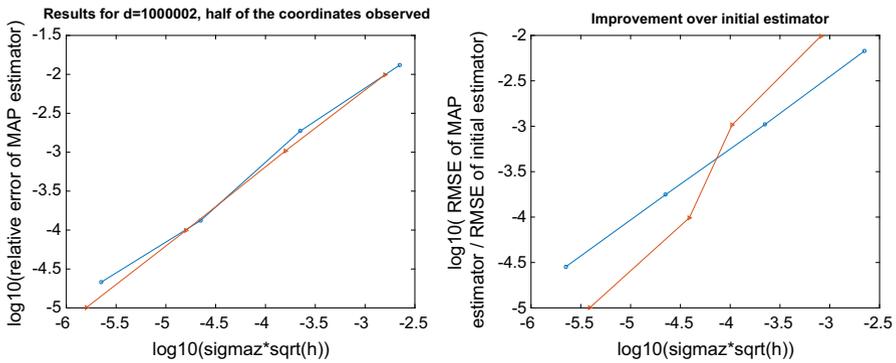
In the following sections, we describe the theoretical and technical details of these simulations. First, in Sect. 3.1, we bound some constants in our theoretical results for the Lorenz 96' model. In Sect. 3.2, we explain the choice of the function  $F$  in our initial estimator (see Theorem 2.8) in the two observation scenarios. In Sect. 3.3, we adapt the Taylor expansion method for numerically solving ODEs to our setting. Finally, based on these preliminary results, we give the technical details of the simulations in Sect. 3.4.

### 3.1 Some Bounds for Lorenz 96' Model

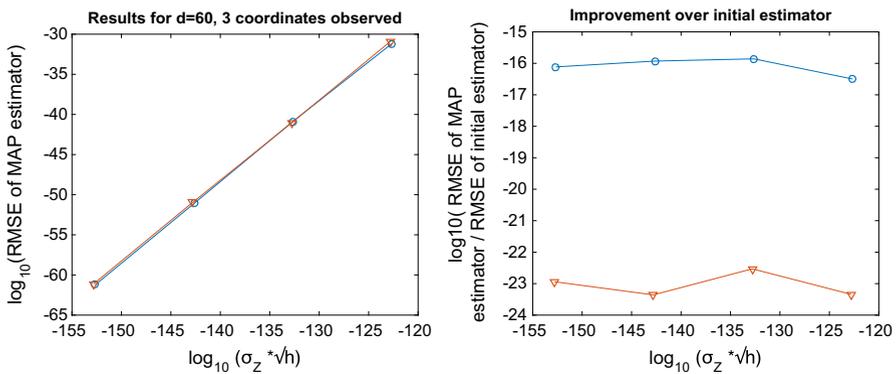
In this section, we will bound some of the constants in Sect. 1.1 for the Lorenz 96' model. Since  $\mathbf{A}$  is the identity matrix, we have  $\|\mathbf{A}\| = 1$  which satisfies that  $\langle \mathbf{v}, \mathbf{A}\mathbf{v} \rangle \geq \lambda_A \mathbf{v}$  for every  $\mathbf{v} \in \mathbb{R}^d$  for  $\lambda_A = 1$ . The condition that  $\langle \mathbf{B}(\mathbf{v}, \mathbf{v}), \mathbf{v} \rangle = 0$  for every



**Fig. 1** Dependence of RMSE of estimator on  $k$  for  $d = 12$



**Fig. 2** Dependence of RMSE of estimator on  $\sigma_Z$  and  $h$  for  $d = 1,000,002$



**Fig. 3** Dependence of RMSE of estimator on  $\sigma_Z$  and  $h$  for  $d = 60$ , 3 coordinates observed

$v \in \mathbb{R}^d$  was verified for the Lorenz 96' model, see Property 3.1 of Law et al. [28]. For our choice  $f = 8$ , we have  $\|f\| = 8\sqrt{d}$ . Thus, based on (1.5), the trapping ball assumption (1.3) is satisfied for the choice  $R := \frac{\|f\|}{\lambda_A} = 8\sqrt{d}$ .

For  $B$ , given any  $u, v \in \mathbb{R}^d$  such that  $\|u\|, \|v\| \leq 1$ , by the arithmetic mean–root-mean-square inequality, and the inequality  $(ab)^2 \leq \frac{a^2+b^2}{2}$ , we have

$$\begin{aligned} \|B(u, v)\|^2 &= \frac{1}{4} \sum_{i=1}^d (v_{i-1}u_{i+1} + u_{i-1}v_{i+1} - v_{i-2}u_{i-1} - u_{i-2}v_{i-1})^2 \\ &\leq (v_{i-1}u_{i+1})^2 + (u_{i-1}v_{i+1})^2 + (v_{i-2}u_{i-1})^2 + (u_{i-2}v_{i-1})^2 \leq 4; \end{aligned}$$

thus,  $\|B\| \leq 2$ . If  $d$  is divisible by 2, then the choice  $u_i = v_i = \frac{(-1)^i}{\sqrt{d}}$  shows that this bound is sharp, and  $\|B\| = 2$ . For simplicity, we have chosen the prior  $q$  as the uniform distribution on  $\mathcal{B}_R$ . Based on these, and definitions (1.16) and (1.6), we have

$$C_0 := 16\sqrt{d}, \quad C_{\text{der}} \leq 1 + 32\sqrt{d}, \quad \text{and} \quad G \leq 1 + 32\sqrt{d}. \tag{3.2}$$

### 3.2 Choice of the Function $F$ in the Initial Estimator

In this section, we will construct a computationally simple function  $F$  satisfying the conditions of Theorem 2.8 for the two observation scenarios.

First, we look at the second scenario, when only the first 3 coordinates are observed. We are going to show that for  $j = \lceil \frac{d-3}{3} \rceil$ , it is possible to construct a function  $F : (\mathbb{R}^{d_0})^{j+1} \rightarrow \mathbb{R}^d$  such that  $F$  is computationally simple, Lipschitz in a neighbourhood of  $u$ , and  $F(Hu, \dots, HD^j v) = u$ , and thus satisfies the conditions of Theorem 2.8. Notice that

$$Du_{i-1} = -u_{i-2}u_{i-3} + u_{i-2}u_i - u_{i-1} + f, \tag{3.3}$$

so for  $m = 0$ , we have

$$u_i = D^0 u_i = (Du_{i-1} - f + u_{i-1} + u_{i-2}u_{i-3}) / u_{i-2}. \tag{3.4}$$

In general, for  $m \geq 1$ , by differentiating (3.3)  $m$  times, we obtain that

$$D^{m+1}u_{i-1} = -D^m u_{i-1} - \sum_{l=0}^m \binom{m}{l} D^l u_{i-2} \cdot D^{m-l} u_{i-3} + \sum_{l=0}^m \binom{m}{l} D^l u_i \cdot D^{m-l} u_{i-2}; \tag{3.5}$$

thus, for any  $m \geq 1$ ,

$$\begin{aligned} D^m u_i &= \left( D^{m+1}u_{i-1} + D^m u_{i-1} + \sum_{l=0}^m \binom{m}{l} D^l u_{i-2} \cdot D^{m-l} u_{i-3} \right. \\ &\quad \left. - \sum_{l=0}^{m-1} \binom{m}{l} D^l u_i \cdot D^{m-l} u_{i-2} \right) / u_{i-2}. \end{aligned} \tag{3.6}$$

Thus, for any  $m \in \mathbb{N}$ , we have a recursion for the  $m$ th derivative of  $u_i$  based on the first  $m + 1$  derivatives of  $u_{i-1}$  and the first  $m$  derivatives of  $u_{i-2}$  and  $u_{i-3}$ . Based on this recursion, and the knowledge of the first  $j$  derivatives of  $u_1, u_2$  and  $u_3$ , we can compute the first  $j - 1$  derivatives of  $u_4$ , then the first  $j - 2$  derivatives of  $u_5$ , etc., and finally the zeroth derivative of  $u_{3+j}$  (i.e.  $u_{3+j}$  itself).

In the other direction,

$$D u_{i+2} = f - u_{i+2} - u_{i+1}u_i + u_{i+1}u_{i+3}; \tag{3.7}$$

therefore, for  $m = 0$ , we have

$$u_i = D^0 u_i = (f - D u_{i+2} - u_{i+2} + u_{i+1}u_{i+3}) / u_{i+1}. \tag{3.8}$$

By differentiating (3.7)  $m$  times, we obtain that

$$D^{m+1} u_{i+2} = -D^m u_{i+2} + \sum_{l=0}^m \binom{m}{l} D^l u_{i+1} \cdot D^{m-l} u_{i+3} - \sum_{l=0}^m \binom{m}{l} D^l u_i \cdot D^{m-l} u_{i+1}; \tag{3.9}$$

thus,

$$D^m u_i = \left( -D^{m+1} u_{i+2} - D^m u_{i+2} + \sum_{l=0}^m \binom{m}{l} D^l u_{i+1} \cdot D^{m-l} u_{i+3} - \sum_{l=0}^{m-1} \binom{m}{l} D^l u_i \cdot D^{m-l} u_{i+1} \right) / u_{i+1}. \tag{3.10}$$

Thus, for any  $m \in \mathbb{N}$ , we have a recursion allowing us to compute the first  $m$  derivatives of  $u_i$  based on the first  $m + 1$  derivatives of  $u_{i+2}$  and the first  $m$  derivatives of  $u_{i+1}$  and  $u_{i+3}$  (with indices considered modulo  $d$ ). This means that given the first  $j$  derivatives of  $u_1, u_2$  and  $u_3$ , we can compute the first  $j - 1$  derivatives of  $u_d$  and  $u_{d-1}$ , then the first  $j - 2$  derivatives of  $u_{d-2}$  and  $u_{d-3}$ , etc., and finally the zeroth derivatives of  $u_{d+2-2j}, u_{d+1-2j}$ .

Based on the choice  $j := \lceil \frac{d-3}{3} \rceil$ , these recursions together define a function  $F$  for this case. From the recursion formulas, and the boundedness of  $\mathbf{u}$ , it follows that  $F$  is Lipschitz in a neighbourhood of  $(H\mathbf{u}, \dots, HD^j \mathbf{u})$  as long as none of the coordinates of  $\mathbf{u}$  is 0 (thus for Lebesgue-almost every  $\mathbf{u} \in \mathcal{B}_R$ ).

Now we look at the first observation scenario, i.e. suppose that  $d$  is divisible by 6, and we observe coordinates  $(6i + 1, 6i + 2, 6i + 3)_{0 \leq i \leq d/6-1}$ . In this case, we choose  $j = 1$  and define  $F(H\mathbf{u}, HD\mathbf{u})$  based on formulas (3.4) and (3.8), so that we can express  $u_{6i+4}, u_{6i-1}, u_{6i-2}$  based on  $u_{6i+1}, u_{6i+2}, u_{6i+3}$  and  $Du_{6i+1}, Du_{6i+2}, Du_{6i+3}$  for  $0 \leq i \leq d/6 - 1$  (with indices counted modulo  $d$ ). Based on Eqs. (3.4) and (3.8), we can see that  $F$  defined as above satisfies the conditions of Theorem 2.8 as long as none of the coordinates of  $u$  is 0 (thus for Lebesgue-almost every  $\mathbf{u} \in \mathcal{B}_R$ ). In both scenarios,  $F$  is computationally simple.

We note that in the above argument,  $F$  might be not defined if some of the components of  $\mathbf{u}$  are 0. (And the proof of Assumption 2.2 in Paulin et al. [38] also requires that none of the components are 0.) Moreover, due to the above formulas, some numerical instability might arise when some of the components of  $\mathbf{u}$  are very small in absolute value. In Section A.2 of ‘‘Appendix’’, we state a simple modification of the initial estimator of Theorem 2.8 based on the above  $F$  that is applicable even when some of the components of  $\mathbf{u}$  are zero.

### 3.3 Numerical Solution of Chaotic ODEs Based on Taylor Expansion

Let  $\mathbf{v} \in \mathcal{B}_R$  and  $i_{\max} \in \mathbb{N}$ . The following lemma provides some simple bounds that allow us to approximate the quantities  $\mathbf{v}(t) = \Psi_t(\mathbf{v})$  for sufficiently small values of  $t$  by the sum of the first  $i_{\max}$  terms in their Taylor expansion. These bounds will be used to simulate system (1.1) and to implement Newton’s method as described in (2.30).

**Lemma 3.1** *For any  $\mathbf{v} \in \mathcal{B}_R$ , we have*

$$\left\| \mathbf{v}(t) - \sum_{i=0}^{i_{\max}} \frac{t^i}{i!} \mathbf{D}^i \mathbf{v} \right\| \leq C_0 (C_{\text{der}} t)^{i_{\max}+1}, \tag{3.11}$$

where  $\mathbf{D}^0 = \mathbf{v}$ , and for  $i \geq 1$ , we have the recursion

$$\mathbf{D}^i \mathbf{v} = -\mathbf{A} \mathbf{D}^{i-1} \mathbf{v} - \sum_{j=0}^{i-1} \binom{i-1}{j} \mathbf{B} \left( \mathbf{D}^j \mathbf{v}, \mathbf{D}^{i-1-j} \mathbf{v} \right). \tag{3.12}$$

*Proof* The bounds on the error in the Taylor expansion follow from (1.14). The recursion equation is just (1.13). □

The following proposition shows a simple way of estimating  $\mathbf{v}(t)$  for larger values  $t$ . (The bound (3.11) is not useful for  $t \geq \frac{1}{C_{\text{der}}}$ .) For  $\mathbf{v} \in \mathbb{R}^d$ , we let

$$P_{\mathcal{B}_R}(\mathbf{v}) := \mathbf{v} \cdot 1_{[\mathbf{v} \in \mathcal{B}_R]} + R \frac{\mathbf{v}}{\|\mathbf{v}\|} \cdot 1_{[\mathbf{v} \notin \mathcal{B}_R]} \tag{3.13}$$

be the projection of  $\mathbf{v}$  on  $\mathcal{B}_R$ .

**Proposition 3.1** *Let  $\mathbf{v} \in \mathcal{B}_R$ , and suppose that  $\Delta < \frac{1}{C_{\text{der}}}$ . Let  $\widehat{\mathbf{v}}(0) = \mathbf{v}$ ,  $\widehat{\mathbf{v}}(\Delta) := P_{\mathcal{B}_R} \left( \sum_{i=0}^{i_{\max}} \frac{\Delta^i}{i!} \mathbf{D}^i \mathbf{v} \right)$ , and similarly, given  $\widehat{\mathbf{v}}(j\Delta)$ , define*

$$\widehat{\mathbf{v}}((j+1)\Delta) := P_{\mathcal{B}_R} \left( \sum_{i=0}^{i_{\max}} \frac{\Delta^i}{i!} \mathbf{D}^i \left( \widehat{\mathbf{v}}(j\Delta) \right) \right).$$

Finally, let  $\delta := t - \lfloor \frac{t}{\Delta} \rfloor \Delta$  and  $\widehat{\mathbf{v}}(t) := P_{\mathcal{B}_R} \left( \sum_{i=0}^{i_{\max}} \frac{\delta^i}{i!} \mathbf{D}^i(\widehat{\mathbf{v}}(\lfloor \frac{t}{\Delta} \rfloor \Delta)) \right)$ . Then the error is bounded as

$$\|\widehat{\mathbf{v}}(t) - \mathbf{v}(t)\| \leq (t + \Delta) \exp(Gt) C_0 C_{\text{der}} \cdot (C_{\text{der}} \Delta)^{i_{\max}}.$$

*Proof* Using (3.11) and the fact that the projection  $P_{\mathcal{B}_R}$  decreases distances we know that for every  $j \in \{0, 1, \dots, \lceil t/\Delta \rceil - 1\}$ ,

$$\|\widehat{\mathbf{v}}((j + 1)\Delta) - \Psi_{\Delta}(\widehat{\mathbf{v}}(j\Delta))\| \leq C_0 (C_{\text{der}} \Delta)^{i_{\max} + 1}.$$

By inequality (1.6), it follows that

$$\|\Psi_{t-j\Delta}(\widehat{\mathbf{v}}(j\Delta)) - \Psi_{t-(j+1)\Delta}(\widehat{\mathbf{v}}((j + 1)\Delta))\| \leq \exp(Gt) C_0 (C_{\text{der}} \Delta)^{i_{\max} + 1},$$

and using the triangle inequality, we have

$$\|\mathbf{v}(t) - \widehat{\mathbf{v}}(t)\| \leq \sum_{j=0}^{\lceil t/\Delta \rceil - 1} \|\Psi_{t-j\Delta}(\widehat{\mathbf{v}}(j\Delta)) - \Psi_{t-(j+1)\Delta}(\widehat{\mathbf{v}}((j + 1)\Delta))\|,$$

so the claim follows. □

### 3.4 Simulation Details

The algorithms were implemented in Julia and ran on a computer with a 2.5Ghz Intel Core i5 CPU. In all cases, the convergence of Newton’s method up to the required precision (chosen to be much smaller than the RMSE) occurred in typically 3–8 steps.

In the case of Fig. 1 ( $d = 12$ , half of the coordinates observed) the observation time  $T$  is much larger than  $\frac{1}{C_{\text{der}}}$ , so we have used the method of Proposition 3.1 to simulate from the system. The gradient and Hessian of the function  $g^{\text{sm}}$  were approximated numerically based on finite difference formulas (requiring  $O(d^2)$  simulations from the ODE). We have noticed that in this case, the Hessian has elements with significantly large absolute value even far away from the diagonal. The running time of Algorithm 1 was approximately 1 s. The RMSEs were numerically approximated from 20 parallel runs. The parameters  $J_{\text{max}}^{(0)}$  and  $J_{\text{max}}^{(1)}$  of the initial estimator were chosen as 1.

In the case of Fig. 1 ( $d = 1,000,002$ , half of the coordinates observed), we could not use the same simulation technique as previously (finite difference approximation of the gradient and Hessian of  $g^{\text{sm}}$ ) because of the huge computational and memory requirements. Instead, we have computed the Newton’s method iterations described in (2.30) based on preconditioned conjugate gradient solver, with the gradient and the product of the Hessian with a vector evaluated based on adjoint methods as described by Eqs. (3.5)–(3.7) and Section 3.2.1 of Paulin et al. [37] (see also Le Dimet et al. [30]). This means that the Hessians were approximated using products of Jacobian matrices that were stored in sparse format due to the local dependency of Eq. (3.1). This efficient storage has allowed us to run Algorithm 1 in approximately 20–40 min

in the simulations. We made 4 parallel runs to estimate the MSEs. The parameters  $J_{\max}^{(0)}$  and  $J_{\max}^{(1)}$  of the initial estimator were chosen as 2.

Finally, in the case of Fig. 3 ( $d = 60$ , first 3 coordinates observed), we used the same method as in the first example (finite difference approximation of the gradient and Hessian of  $g^{\text{sm}}$ ). The running time of Algorithm 1 was approximately 1 hour (in part due to the need of using arbitrary precision arithmetics with hundreds of digits of precision). The MSEs were numerically approximated from 2 parallel runs. The parameters  $J_{\max}^{(0)}, \dots, J_{\max}^{(19)}$  of the initial estimator were chosen as 24.

### 4 Preliminary Results

The proof of our main theorems are based on several preliminary results. Let

$$l^{\text{sm}}(\mathbf{v}) := \sum_{i=0}^k \left( \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 + 2 \langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}), \mathbf{Z}_i \rangle \right), \text{ and} \tag{4.1}$$

$$l_{\mathcal{G}}^{\text{sm}}(\mathbf{v}) := (\mathbf{v} - \mathbf{u})' \mathbf{A}_k (\mathbf{v} - \mathbf{u}) + 2 \langle \mathbf{v} - \mathbf{u}, \mathbf{B}_k \rangle. \tag{4.2}$$

These quantities are related to the log-likelihoods of the smoothing distribution and its Gaussian approximation as

$$\mu^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) = \frac{q(\mathbf{v})}{C_k^{\text{sm}}} \exp \left[ -\frac{l^{\text{sm}}(\mathbf{v})}{2\sigma_Z^2} \right], \text{ and if } \mathbf{A}_k \succ \mathbf{0}, \text{ then} \tag{4.3}$$

$$\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) = \frac{\det(\mathbf{A}_k)^{1/2}}{(2\pi)^{d/2} \cdot \sigma_Z^d} \cdot \exp \left[ -\frac{\mathbf{B}_k \mathbf{A}_k^{-1} \mathbf{B}_k}{2\sigma_Z^2} \right] \cdot \exp \left[ -\frac{l_{\mathcal{G}}^{\text{sm}}(\mathbf{v})}{2\sigma_Z^2} \right], \tag{4.4}$$

where  $\succ$  denotes the positive definite order (i.e.  $\mathbf{A} \succ \mathbf{B}$  if and only if  $\mathbf{A} - \mathbf{B}$  is positive definite). Similarly,  $\succeq$  denotes the positive semidefinite order.

The next three propositions show various bounds on the log-likelihood-related quantity  $l^{\text{sm}}(\mathbf{v})$ . Their proof is included in Section A.1 of ‘‘Appendix’’.

**Proposition 4.1** (A lower bound on the tails of  $l^{\text{sm}}(\mathbf{v})$ ) *Suppose that Assumption 2.1 holds for  $\mathbf{u}$ , and then for any  $0 < \varepsilon \leq 1$ , for every  $\sigma_Z > 0, h \leq h_{\max}(\mathbf{u}, T)$ , we have*

$$\mathbb{P} \left( l^{\text{sm}}(\mathbf{v}) \geq \frac{c(\mathbf{u}, T)}{h} \|\mathbf{v} - \mathbf{u}\|^2 - \frac{C_1(\mathbf{u}, T, \varepsilon)\sigma_Z}{\sqrt{h}} \cdot \|\mathbf{v} - \mathbf{u}\| \text{ for every } \mathbf{v} \in \mathcal{B}_R \left| \mathbf{u} \right. \right) \geq 1 - \varepsilon,$$

where

$$C_1(\mathbf{u}, T, \varepsilon) := 44(\widehat{M}_2(T)R + \widehat{M}_1(T))\sqrt{\overline{T}(\mathbf{u})(d + 1)d_o} + 2\sqrt{2\overline{T}(\mathbf{u})d_o}\widehat{M}_1(T) \log \left( \frac{1}{\varepsilon} \right).$$

**Proposition 4.2** (A bound on the difference between  $I^{\text{sm}}(\mathbf{v})$  and  $I_{\mathcal{G}}^{\text{sm}}(\mathbf{v})$ ) *If Assumption 2.1 holds for  $\mathbf{u}$ , then for any  $0 < \varepsilon \leq 1$ ,  $\sigma_Z > 0$ ,  $0 < h \leq h_{\max}(\mathbf{u}, T)$ , we have*

$$\mathbb{P} \left( \left| I^{\text{sm}}(\mathbf{v}) - I_{\mathcal{G}}^{\text{sm}}(\mathbf{v}) \right| \leq \|\mathbf{v} - \mathbf{u}\|^3 \cdot \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h}}{h} \text{ for every } \mathbf{v} \in \mathcal{B}_R \mid \mathbf{u} \right) \geq 1 - \varepsilon,$$

where

$$C_2(\mathbf{u}, T) := \bar{T}(\mathbf{u})\widehat{M}_1(T)\widehat{M}_2(T) \text{ and}$$

$$C_3(\mathbf{u}, T, \varepsilon) := 22(\widehat{M}_3(T) + \widehat{M}_4(T)R)\sqrt{(d+1)d_o\bar{T}(\mathbf{u})} + \sqrt{\frac{4}{3}\bar{T}(\mathbf{u})\widehat{M}_3(T)d_o \log\left(\frac{2}{\varepsilon}\right)}.$$

**Proposition 4.3** (A bound on the difference between  $\nabla I^{\text{sm}}(\mathbf{v})$  and  $\nabla I_{\mathcal{G}}^{\text{sm}}(\mathbf{v})$ ) *If Assumption 2.1 holds for  $\mathbf{u}$ , then for any  $0 < \varepsilon \leq 1$ ,  $\sigma_Z > 0$ ,  $0 < h \leq h_{\max}(\mathbf{u}, T)$ , we have*

$$\mathbb{P} \left( \|\nabla I^{\text{sm}}(\mathbf{v}) - \nabla I_{\mathcal{G}}^{\text{sm}}(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{u}\|^2 \cdot \frac{C_4(\mathbf{u}, T) + C_5(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h}}{h} \text{ for every } \mathbf{v} \in \mathcal{B}_R \mid \mathbf{u} \right) \geq 1 - \varepsilon,$$

where

$$C_4(\mathbf{u}, T) := 4\bar{T}(\mathbf{u})\widehat{M}_1(T)\widehat{M}_2(T), \text{ and}$$

$$C_5(\mathbf{u}, T, \varepsilon) := 66(\widehat{M}_3(T) + \widehat{M}_4(T)R)\sqrt{\bar{T}(\mathbf{u})(2d+1)d_o} + 2\sqrt{\bar{T}(\mathbf{u})\widehat{M}_3(T)d_o \log\left(\frac{1}{\varepsilon}\right)}.$$

The following lemma is useful for controlling the total variation distance of two distributions that are only known up to normalising constants.

**Lemma 4.1** *Let  $f$  and  $g$  be two probability distributions which have densities on  $\mathbb{R}^d$  with respect to the Lebesgue measure. Then their total variation distance satisfies that for any constant  $c > 0$ ,*

$$d_{\text{TV}}(f, g) \leq \int_{\mathbf{x} \in \mathbb{R}^d} |f(\mathbf{x}) - cg(\mathbf{x})| d\mathbf{x}.$$

*Proof* We are going to use the following characterisation of the total variation distance,

$$\begin{aligned} d_{\text{TV}}(f, g) &= \frac{1}{2} \int_{\mathbf{x} \in \mathbb{R}^d} |f(\mathbf{x}) - g(\mathbf{x})| d\mathbf{x} \\ &= \int_{\mathbf{x} \in \mathbb{R}^d} (f(\mathbf{x}) - g(\mathbf{x}))_+ d\mathbf{x} = \int_{\mathbf{x} \in \mathbb{R}^d} (f(\mathbf{x}) - g(\mathbf{x}))_- d\mathbf{x}. \end{aligned}$$

Based on this, the result is trivial for  $c = 1$ . If  $c > 1$ , then we have

$$\int_{\mathbf{x} \in \mathbb{R}^d} (f(\mathbf{x}) - g(\mathbf{x}))_- d\mathbf{x} \leq \int_{\mathbf{x} \in \mathbb{R}^d} (f(\mathbf{x}) - cg(\mathbf{x}))_- d\mathbf{x} \leq \int_{\mathbf{x} \in \mathbb{R}^d} |f(\mathbf{x}) - cg(\mathbf{x})| d\mathbf{x},$$

and the  $c < 1$  case is similar. □

The following lemma is useful for controlling the Wasserstein distance of two distributions.

**Lemma 4.2** *Let  $f$  and  $g$  be two probability distributions which have densities on  $\mathbb{R}^d$  with respect to the Lebesgue measure, also denoted by  $f(\mathbf{x})$  and  $g(\mathbf{x})$ . Then their Wasserstein distance (defined as in (2.13)) satisfies that for any  $y \in \mathbb{R}^d$ ,*

$$d_W(f, g) \leq \int_{\mathbf{x} \in \mathbb{R}^d} |f(\mathbf{x}) - g(\mathbf{x})| \cdot \|\mathbf{x} - y\| d\mathbf{x}.$$

*Proof* Let  $m(\mathbf{x}) := \min(f(\mathbf{x}), g(\mathbf{x}))$ ,  $\gamma := \int_{\mathbf{x} \in \mathbb{R}^d} m(\mathbf{x}) d\mathbf{x}$ ,  $\hat{f}(\mathbf{x}) := f(\mathbf{x}) - m(\mathbf{x})$ ,  $\hat{g}(\mathbf{x}) := g(\mathbf{x}) - m(\mathbf{x})$ . Suppose first that  $\gamma \neq 0$  and  $1 - \gamma \neq 0$ . Let  $v_{f,g}^{(1)}$  denote the distribution of a random vector  $(\mathbf{X}, \mathbf{X})$  on  $\mathbb{R}^d \times \mathbb{R}^d$  for a random variable  $\mathbf{X}$  with distribution with density  $m(\mathbf{x})/\gamma$ . (The two components are equal.) Let  $v_{f,g}^{(2)}$  be a distribution on  $\mathbb{R}^d \times \mathbb{R}^d$  with density  $v_{f,g}^{(2)}(\mathbf{x}_1, \mathbf{x}_2) := \frac{\hat{f}(\mathbf{x}_1)}{1-\gamma} \cdot \frac{\hat{g}(\mathbf{x}_2)}{1-\gamma}$ . (The two components are independent.)

We define the optimal coupling of  $f$  and  $g$  as a probability distribution  $v_{f,g}$  on  $\mathbb{R}^d \times \mathbb{R}^d$  as a mixture of  $v_{f,g}^{(1)}$  and  $v_{f,g}^{(2)}$ , that is, for any Borel-measurable  $E \in \mathbb{R}^d \times \mathbb{R}^d$ ,

$$v_{f,g}(E) := \gamma v_{f,g}^{(1)}(E) + (1 - \gamma) v_{f,g}^{(2)}(E). \tag{4.5}$$

It is easy to check that this distribution has marginals  $f$  and  $g$ , so by (2.13), and the fact that  $\int_{\mathbf{x}_1 \in \mathbb{R}^d} \hat{f}(\mathbf{x}_1) d\mathbf{x}_1 = \int_{\mathbf{x}_2 \in \mathbb{R}^d} \hat{g}(\mathbf{x}_2) d\mathbf{x}_2 = 1 - \gamma$ , we have

$$\begin{aligned} d_W(f, g) &\leq \int_{\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d} \|\mathbf{x}_1 - \mathbf{x}_2\| d v_{f,g}(\mathbf{x}_1, \mathbf{x}_2) \\ &= \int_{\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d} \frac{\hat{f}(\mathbf{x}_1) \hat{g}(\mathbf{x}_2)}{1 - \gamma} \cdot \|\mathbf{x}_1 - \mathbf{x}_2\| d\mathbf{x}_1 d\mathbf{x}_2 \\ &\leq \int_{\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d} \frac{\hat{f}(\mathbf{x}_1) \hat{g}(\mathbf{x}_2)}{1 - \gamma} \cdot (\|\mathbf{x}_1 - \mathbf{y}\| + \|\mathbf{x}_2 - \mathbf{y}\|) d\mathbf{x}_1 d\mathbf{x}_2 \\ &= \int_{\mathbf{x}_1 \in \mathbb{R}^d} \hat{f}(\mathbf{x}_1) \|\mathbf{x}_1 - \mathbf{y}\| d\mathbf{x}_1 + \int_{\mathbf{x}_2 \in \mathbb{R}^d} \hat{g}(\mathbf{x}_2) \|\mathbf{x}_2 - \mathbf{y}\| d\mathbf{x}_2 \\ &= \int_{\mathbf{x} \in \mathbb{R}^d} (\hat{f}(\mathbf{x}) + \hat{g}(\mathbf{x})) \|\mathbf{x} - \mathbf{y}\| d\mathbf{x}, \end{aligned}$$

and the result follows from the fact that  $\hat{f}(\mathbf{x}) + \hat{g}(\mathbf{x}) = |f(\mathbf{x}) - g(\mathbf{x})|$ . Finally, if  $\gamma = 0$  then we can set  $v_{f,g}$  as  $v_{f,g}^{(2)}$ , and the same argument works, while if  $\gamma = 1$ , then both sides of the claim are zero (since  $f(\mathbf{x}) = g(\mathbf{x})$  Lebesgue-almost surely).  $\square$

The following two lemmas show concentration bounds for the norm of  $\mathbf{B}_k$  and the smallest eigenvalue of  $\mathbf{A}_k$ . The proofs are included in Section A.1 of ‘‘Appendix’’. (They are based on matrix concentration inequalities.)

**Lemma 4.3** *Suppose that Assumption 2.1 holds for  $\mathbf{u}$ , then the random vector  $\mathbf{B}_k$  defined in (2.8) satisfies that for any  $t \geq 0$ ,*

$$\mathbb{P}(\|\mathbf{B}_k\| \geq t \mid \mathbf{u}) \leq (d + 1) \exp\left(-\frac{t^2}{(k + 1)d_o \widehat{M}_1(T)^2 \sigma_Z^2}\right). \tag{4.6}$$

Thus, for any  $0 < \varepsilon \leq 1$ , we have

$$\mathbb{P}\left(\|\mathbf{B}_k\| \geq C_B(\varepsilon) \frac{\sigma_Z}{\sqrt{h}} \mid \mathbf{u}\right) \leq \varepsilon \text{ for } C_B(\varepsilon) := \sqrt{\log\left(\frac{d + 1}{\varepsilon}\right) \cdot \widehat{M}_1(T)^2 \bar{T}(\mathbf{u}) d_o}. \tag{4.7}$$

**Lemma 4.4** *Suppose that Assumption 2.1 holds for  $\mathbf{u}$ , then the random matrix  $\mathbf{A}_k$  defined in (2.7) satisfies that for any  $t \geq 0$ ,*

$$\begin{aligned} &\mathbb{P}\left(\lambda_{\min}(\mathbf{A}_k) \leq \frac{c(\mathbf{u}, T)}{h} - t \text{ or } \|\mathbf{A}_k\| \geq \widehat{M}_1(T)^2 \cdot \frac{\bar{T}(\mathbf{u})}{h} + t \mid \mathbf{u}\right) \\ &\leq 2d \exp\left(-\frac{t^2}{(k + 1)\sigma_Z^2 \widehat{M}_2(T)^2 d_o}\right); \end{aligned} \tag{4.8}$$

thus, for any  $0 < \varepsilon \leq 1$ , we have

$$\begin{aligned} &\mathbb{P}\left(\lambda_{\min}(\mathbf{A}_k) > \frac{c(\mathbf{u}, T) - C_A(\varepsilon)\sigma_Z\sqrt{h}}{h} \right. \\ &\quad \left. \text{and } \|\mathbf{A}_k\| < \widehat{M}_1(T)^2 \cdot \frac{\bar{T}(\mathbf{u})}{h} + \frac{C_A(\varepsilon)\sigma_Z}{\sqrt{h}}\right) \leq \varepsilon \text{ for} \\ &C_A(\varepsilon) := \sqrt{\log\left(\frac{2d}{\varepsilon}\right) \cdot \widehat{M}_2(T)^2 \bar{T}(\mathbf{u}) d_o}. \end{aligned} \tag{4.9}$$

The following proposition bounds the derivatives of the log-determinant of the Jacobian.

**Proposition 4.4** *The function  $\log \det \mathbf{J}\Psi_T : \mathcal{B}_R \rightarrow \mathbb{R}$  is continuously differentiable on  $\text{int}(\mathcal{B}_R)$ , and its derivative can be bounded as*

$$\sup_{\mathbf{v} \in \text{int}(\mathcal{B}_R)} \|\nabla \log \det \mathbf{J}\Psi_T(\mathbf{v})\| \leq M_1(T)M_2(T)d.$$

*Proof* Let  $\mathbb{M}^d$  denote the space of  $d \times d$  complex matrices. By the chain rule, we can write the derivatives of the log-determinant as functions of the derivatives of the determinant, which were shown to exist in Bhatia and Jain [5]. Following their notation, we define the  $k$ th derivative of  $\det$  at a point  $\mathbf{M} \in \mathbb{M}^d$  as a map  $D^k \det \mathbf{M}$  from  $(\mathbb{M}^d)^k$  to  $\mathbb{C}$  with value

$$D^k \det \mathbf{M}(X^1, \dots, X^k) := \frac{\partial^k}{\partial t_1 \dots \partial t_k} \Big|_{t_1=\dots=t_k=0} \det \left( \mathbf{M} + t_1 X^1 + \dots + t_k X^k \right).$$

They have defined the norm of  $k$ th derivative of the determinant as

$$\|D^k \det \mathbf{M}\| := \sup_{\{X^i\}_{1 \leq i \leq k} : \|X^i\|=1 \text{ for } 1 \leq i \leq k} \left| D^k \det \mathbf{M}(X^1, \dots, X^k) \right|.$$

From Theorem 4 of Bhatia and Jain [5], it follows that for any  $k \geq 1$ , the norm of the  $k$ th derivative can be bounded as

$$\|D^k \det \mathbf{M}\| \leq \|\mathbf{M}\|^k d^k \cdot |\det \mathbf{M}|. \tag{4.10}$$

Based on the chain rule, the norm first derivative of the log-determinant can be bounded as

$$\|\nabla \log \det \mathbf{J}\Psi_T(\mathbf{v})\| = \left| \frac{\nabla \det \mathbf{J}\Psi_T(\mathbf{v})}{\det \mathbf{J}\Psi_T(\mathbf{v})} \right| \leq \frac{\|D \det \mathbf{J}\Psi_T(\mathbf{v})\| \|\mathbf{J}^2 \Psi_T(\mathbf{v})\|}{|\det \mathbf{J}\Psi_T(\mathbf{v})|},$$

and the result follows by (1.21) and (4.10) for  $k = 1$  and  $\mathbf{M} = \mathbf{J}\Psi_T(\mathbf{v})$ . □

## 5 Proof of the Main Results

### 5.1 Gaussian Approximation

In the following two subsections, we are going to prove our Gaussian approximation results for the smoother and the filter, respectively.

#### 5.1.1 Gaussian Approximation for the Smoother

In this section, we are going to describe the proof of Theorem 2.1. First, we will show that the result can be obtained by bounding 3 separate terms, then bounds these in 3 lemmas and finally, combine them.

By choosing the constants  $C_{TV}^{(1)}(\mathbf{u}, T)$  and  $C_{TV}^{(2)}(\mathbf{u}, T)$  sufficiently large, we can assume that  $C_{TV}(\mathbf{u}, T, \varepsilon)$  satisfies the following bounds,

$$C_{TV}(\mathbf{u}, T, \varepsilon) \geq \frac{2C_A(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \tag{5.1}$$

$$C_{TV}(\mathbf{u}, T, \varepsilon) \geq \frac{C_3(\mathbf{u}, T, \frac{\varepsilon}{4})}{C_2(\mathbf{u}, T)} \tag{5.2}$$

$$C_{TV}(\mathbf{u}, T, \varepsilon) \geq \left( 2C_2(\mathbf{u}, T) \min \left( \frac{1}{2C_q^{(1)}}, R - \|\mathbf{u}\| \right) \right)^{-3/2} \tag{5.3}$$

$$C_{TV}(\mathbf{u}, T, \varepsilon) \geq 512 \left( \frac{C_2(\mathbf{u}, T)C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \right)^3 \tag{5.4}$$

$$C_{TV}(\mathbf{u}, T, \varepsilon) \geq 64 \left( \frac{C_1(\mathbf{u}, T, \varepsilon)C_2(\mathbf{u}, T)}{c(\mathbf{u}, T)} \right)^3 \tag{5.5}$$

Based on the assumption that  $\sigma_Z\sqrt{h} \leq \frac{1}{2}C_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , (5.1), and Lemma 4.4, we have

$$\mathbb{P} \left( \frac{c(\mathbf{u}, T)}{2h} \mathbf{I}_d \prec \mathbf{A}_k \prec \frac{C_{\|\mathbf{A}\|}}{h} \mathbf{I}_d \mid \mathbf{u} \right) \geq 1 - \frac{\varepsilon}{4}, \tag{5.6}$$

where  $C_{\|\mathbf{A}\|}$  is defined as in (2.15). The event  $\lambda_{\min}(\mathbf{A}_k) > \frac{c(\mathbf{u}, T)}{2h}$  implies in particular that  $\mathbf{A}_k$  is positive definite. From Lemma 4.3, we know that

$$\mathbb{P} \left( \|\mathbf{B}_k\| < C_B \left( \frac{\varepsilon}{4} \right) \cdot \frac{\sigma_Z}{\sqrt{h}} \mid \mathbf{u} \right) \geq 1 - \frac{\varepsilon}{4}. \tag{5.7}$$

From Proposition 4.1, we know that

$$\begin{aligned} \mathbb{P} \left( l^{\text{sm}}(\mathbf{v}) \geq \frac{c(\mathbf{u}, T)}{h} \|\mathbf{v} - \mathbf{u}\|^2 - \frac{C_1(\mathbf{u}, T, \frac{\varepsilon}{4})\sigma_Z}{\sqrt{h}} \cdot \|\mathbf{v} - \mathbf{u}\| \text{ for every } \mathbf{v} \in \mathcal{B}_R \mid \mathbf{u} \right) \\ \geq 1 - \frac{\varepsilon}{4}. \end{aligned} \tag{5.8}$$

Finally, from Proposition 4.2, it follows that

$$\begin{aligned} \mathbb{P} \left( |l^{\text{sm}}(\mathbf{v}) - l_{\mathcal{G}}^{\text{sm}}(\mathbf{v})| \leq \|\mathbf{v} - \mathbf{u}\|^3 \cdot \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4})\sigma_Z\sqrt{h}}{h} \right. \\ \left. \text{ for every } \mathbf{v} \in \mathcal{B}_R \mid \mathbf{u} \right) \geq 1 - \frac{\varepsilon}{4}. \end{aligned} \tag{5.9}$$

In the rest of the proof, we are going to assume that all four of the events in Eqs. (5.6), (5.7), (5.8) and (5.9) hold. From the above bounds, we know that this happens with probability at least  $1 - \varepsilon$ .

Let  $\mathbf{Z}$  be a  $d$ -dimensional standard normal random vector, then from the definition of  $\mu_{\mathcal{G}}^{\text{sm}}$ , it follows that when conditioned on  $\mathbf{B}_k$  and  $\mathbf{A}_k$ ,

$$\mathbf{W} := \sigma_Z \cdot \mathbf{A}_k^{-1/2} \cdot \mathbf{Z} + \mathbf{u} - \mathbf{A}_k^{-1} \mathbf{B}_k \tag{5.10}$$

has distribution  $\mu_{\mathcal{G}}^{\text{sm}}(\cdot | \mathbf{Y}_{0:k})$ . This fact will be used in the proof several times.

Since the normalising constant of the smoother  $C_k^{\text{sm}}$  of (2.4) is not known, it is not easy to bound the total variation distance of the two distributions directly. Lemma 4.1 allows us to deal with this problem by rescaling the smoothing distribution suitably. We define the rescaled smoothing distribution (which is not a probability distribution in general) as

$$\tilde{\mu}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) := \frac{\det(\mathbf{A}_k)^{1/2}}{(2\pi)^{d/2} \cdot \sigma_Z^d} \cdot \exp\left[-\frac{\mathbf{B}_k \mathbf{A}_k^{-1} \mathbf{B}_k}{2\sigma_Z^2}\right] \cdot \frac{q(\mathbf{v})}{q(\mathbf{u})} \cdot \exp\left[-\frac{l^{\text{sm}}(\mathbf{v})}{2\sigma_Z^2}\right], \tag{5.11}$$

which is of similar form as the Gaussian approximation

$$\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) = \frac{\det(\mathbf{A}_k)^{1/2}}{(2\pi)^{d/2} \cdot \sigma_Z^d} \cdot \exp\left[-\frac{\mathbf{B}_k \mathbf{A}_k^{-1} \mathbf{B}_k}{2\sigma_Z^2}\right] \cdot \exp\left[-\frac{l_{\mathcal{G}}^{\text{sm}}(\mathbf{v})}{2\sigma_Z^2}\right]. \tag{5.12}$$

Let

$$\rho(h, \sigma_Z) := \frac{(h\sigma_Z^2)^{1/3}}{2C_2(\mathbf{u}, T)}, \tag{5.13}$$

then based on the assumption that  $\sigma_Z \sqrt{h} \leq C_{\text{TV}}(\mathbf{u}, T, \varepsilon)^{-1}$ , (5.2) and (5.3), it follows that

$$\rho(h, \sigma_Z) \leq \min\left[\left(\frac{h\sigma_Z^2}{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4})\sigma_Z \sqrt{h}}\right)^{1/3}, \frac{1}{2C_q^{(1)}}, R - \|\mathbf{u}\|\right]. \tag{5.14}$$

Let  $B_\rho := \{\mathbf{v} \in \mathbb{R}^d : \|\mathbf{v} - \mathbf{u}\| \leq \rho(h, \sigma_Z)\}$  and denote by  $B_\rho^c$  its complement in  $\mathbb{R}^d$ . Then by Lemma 4.1, we have

$$d_{\text{TV}}(\mu^{\text{sm}}(\cdot | \mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot | \mathbf{Y}_{0:k})) \leq \int_{\mathbf{v} \in \mathbb{R}^d} |\tilde{\mu}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k})| d\mathbf{v} \tag{5.15}$$

$$\leq \tilde{\mu}^{\text{sm}}(B_\rho^c | \mathbf{Y}_{0:k}) + \mu_{\mathcal{G}}^{\text{sm}}(B_\rho^c | \mathbf{Y}_{0:k}) + \int_{\mathbf{v} \in B_\rho} |\tilde{\mu}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k})| d\mathbf{v}. \tag{5.16}$$

By Lemma 4.2, we can bound the Wasserstein distance of  $\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  and  $\mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  as

$$\begin{aligned}
 d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) &\leq \int_{\mathbf{v} \in \mathbb{R}^d} \|\mathbf{v} - \mathbf{u}\| \cdot |\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} \\
 &\leq \int_{\mathbf{v} \in B_\rho} \|\mathbf{v} - \mathbf{u}\| \cdot |\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} \\
 &\quad + \int_{\mathbf{v} \in B_\rho^c} \|\mathbf{v} - \mathbf{u}\| \mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) d\mathbf{v} + \int_{\mathbf{v} \in B_\rho^c} \|\mathbf{v} - \mathbf{u}\| \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) d\mathbf{v}. \tag{5.17}
 \end{aligned}$$

In the following six lemmas, we bound the three terms in inequalities (5.16) and (5.17).

**Lemma 5.1** *Using the notations and assumptions of this section, we have*

$$\begin{aligned}
 \int_{\mathbf{v} \in B_\rho} |\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} &\leq D_1(\mathbf{u}, T, \varepsilon) \sigma_Z \sqrt{h} + D_2(\mathbf{u}, T, \varepsilon) \sigma_Z^2 h \text{ for} \\
 D_1(\mathbf{u}, T, \varepsilon) &:= \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} + \sqrt{\frac{2d}{c(\mathbf{u}, T)}} + C_2(\mathbf{u}, T) \left( 6 \left( \frac{2d}{c(\mathbf{u}, T)} \right)^{3/2} + 2 \left( \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \right)^3 \right), \\
 D_2(\mathbf{u}, T, \varepsilon) &:= C_3 \left( \mathbf{u}, T, \frac{\varepsilon}{4} \right) \left( 6 \left( \frac{2d}{c(\mathbf{u}, T)} \right)^{3/2} + 2 \left( \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \right)^3 \right).
 \end{aligned}$$

*Proof* Note that by (5.14), we know that  $B_\rho \subset \mathcal{B}_R$ , and

$$\begin{aligned}
 \int_{\mathbf{v} \in B_\rho} |\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} &= \int_{\mathbf{v} \in B_\rho} \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) \left| 1 - \frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} \right| d\mathbf{v} \\
 &= \int_{\mathbf{v} \in B_\rho} \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) \left| 1 - \exp \left( \log \left( \frac{q(\mathbf{v})}{q(\mathbf{u})} \right) - \frac{I^{\text{sm}}(\mathbf{v}) - I^{\text{sm}}(\mathbf{u})}{2\sigma_Z^2} \right) \right| d\mathbf{v}.
 \end{aligned}$$

Now using (5.14), we can see that  $\sup_{\mathbf{v} \in B_\rho} \left| \log \left( \frac{q(\mathbf{v})}{q(\mathbf{u})} \right) \right| \leq C_q^{(1)} \cdot \frac{1}{2C_q^{(1)}} \leq \frac{1}{2}$ , and using (5.9), we have  $\left| \frac{I^{\text{sm}}(\mathbf{v}) - I^{\text{sm}}(\mathbf{u})}{2\sigma_Z^2} \right| \leq \frac{1}{2}$ . Using the fact that  $|1 - \exp(x)| \leq 2|x|$  for  $-1 \leq x \leq 1$ , and the bounds  $|\log(q(\mathbf{v})/q(\mathbf{u}))| \leq C_q^{(1)} \|\mathbf{v} - \mathbf{u}\|$  and (5.9), we can see that for every  $\mathbf{v} \in B_\rho$ ,

$$\left| 1 - \frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} \right| \leq 2 \left( C_q^{(1)} \|\mathbf{v} - \mathbf{u}\| + \frac{C_2(\mathbf{u}, T) + C_3 \left( \mathbf{u}, T, \frac{\varepsilon}{4} \right) \sigma_Z \sqrt{h}}{h} \cdot \frac{\|\mathbf{v} - \mathbf{u}\|^3}{2\sigma_Z^2} \right); \tag{5.18}$$

therefore,

$$\int_{\mathbf{v} \in B_\rho} |\tilde{\mu}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k})| \, d\mathbf{v} \leq 2 \int_{\mathbf{v} \in \mathbb{R}^d} \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) \left( C_q^{(1)} \|\mathbf{v} - \mathbf{u}\| + \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z \sqrt{h}}{h} \cdot \frac{\|\mathbf{v} - \mathbf{u}\|^3}{2\sigma_Z^2} \right).$$

Let  $\mathbf{Z}$  denote a  $d$ -dimensional standard normal random vector, then it is easy to see that  $\mathbb{E}(\|\mathbf{Z}\|) \leq (\mathbb{E}(\|\mathbf{Z}\|^2))^{1/2} \leq d^{1/2}$  and  $\mathbb{E}(\|\mathbf{Z}\|^3) \leq (\mathbb{E}(\|\mathbf{Z}\|^4))^{3/4} \leq (3d^2)^{3/4} \leq 3d^{3/2}$ . Since we have assumed that the events in (5.6) and (5.7) hold, we know that

$$\|\mathbf{A}_k^{-1/2}\| \leq \sqrt{\frac{2h}{c(\mathbf{u}, T)}}, \text{ and } \|\mathbf{A}_k^{-1} \mathbf{B}_k\| \leq \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \cdot \sigma_Z \sqrt{h}. \tag{5.19}$$

Finally, it is not difficult to show that for any  $a, b \geq 0$ ,  $(a + b)^3 \leq 4(a^3 + b^3)$ . Therefore,

$$\begin{aligned} & \int_{\mathbf{v} \in B_\rho} |\tilde{\mu}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k})| \, d\mathbf{v} \\ & \leq \mathbb{E} \left[ C_q^{(1)} \|\mathbf{W} - \mathbf{u}\| + \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z \sqrt{h}}{h} \cdot \frac{\|\mathbf{W} - \mathbf{u}\|^3}{2\sigma_Z^2} \middle| \mathbf{A}_k, \mathbf{B}_k \right] \\ & \leq C_q^{(1)} \left( \|\mathbf{A}_k^{-1} \mathbf{B}_k\| + \sigma_Z \|\mathbf{A}_k^{-1/2}\| \sqrt{d} \right) \\ & \quad + \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z \sqrt{h}}{2\sigma_Z^2 h} \cdot 4 \cdot \left( 3 \|\mathbf{A}_k^{-1/2}\|^3 \sigma_Z^3 d^{3/2} + \|\mathbf{A}_k^{-1} \mathbf{B}_k\|^3 \right) \\ & \leq \left( \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} + \sqrt{\frac{2d}{c(\mathbf{u}, T)}} \right) \sigma_Z \sqrt{h} + \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z \sqrt{h}}{2\sigma_Z^2 h} \\ & \quad \cdot 4 \cdot \left( 3 \left( \frac{2h}{c(\mathbf{u}, T)} \right)^{3/2} \cdot \sigma_Z^3 d^{3/2} + \left( \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \cdot \sigma_Z \sqrt{h} \right)^3 \right); \end{aligned}$$

thus, the result follows. □

**Lemma 5.2** *Using the notations and assumptions of this section, we have*

$$\mu_{\mathcal{G}}^{\text{sm}}(B_\rho^c | \mathbf{Y}_{0:k}) \leq (d + 1) \exp \left( - \frac{c(\mathbf{u}, T) \cdot (\sigma_Z \sqrt{h})^{-2/3}}{64d C_2(\mathbf{u}, T)^2} \right).$$

*Proof* Note that if  $\mathbf{Z}$  is a  $d$ -dimensional standard normal random vector, then by Theorem 4.1.1 of Tropp [46], we have

$$\mathbb{P}(\|\mathbf{Z}\| \geq t) \leq (d + 1) \exp \left( - \frac{t^2}{2d} \right) \text{ for any } t \geq 0. \tag{5.20}$$

Since the random variable  $\mathbf{W}$  defined in (5.10) is distributed as  $\mu_{\mathcal{G}}^{\text{sm}}(\cdot | \mathbf{Y}_{0:k})$  when conditioned on  $\mathbf{A}_k, \mathbf{B}_k$ , we have

$$\begin{aligned} \mu_{\mathcal{G}}^{\text{sm}}(B_{\rho}^c | \mathbf{Y}_{0:k}) &= \mathbb{P}(\|\mathbf{W} - \mathbf{u}\| > \rho(h, \sigma_Z) | \mathbf{A}_k, \mathbf{B}_k) \\ &\leq \mathbb{P}\left(\|\mathbf{Z}\| > \frac{\rho(h, \sigma_Z) - \|\mathbf{A}_k^{-1} \mathbf{B}_k\|}{\sigma_Z \cdot \|\mathbf{A}_k^{-1/2}\|}\right) \leq \mathbb{P}\left(\|\mathbf{Z}\| > \frac{\rho(h, \sigma_Z) - \frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \cdot \sigma_Z \sqrt{h}}{\sigma_Z \cdot \sqrt{\frac{2h}{c(\mathbf{u}, T)}}}\right). \end{aligned}$$

Based on the assumption that  $\sigma_Z \sqrt{h} \leq C_{\text{TV}}(\mathbf{u}, T, \varepsilon)^{-1}$ , and (5.4), we have  $\frac{2C_B(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \cdot \sigma_Z \sqrt{h} \leq \frac{\rho(h, \sigma_Z)}{2}$ , and

$$\mu_{\mathcal{G}}^{\text{sm}}(B_{\rho}^c | \mathbf{Y}_{0:k}) \leq \mathbb{P}\left(\|\mathbf{Z}\| \geq \frac{(\sigma_Z \sqrt{h})^{-1/3} \cdot \sqrt{c(\mathbf{u}, T)}}{4\sqrt{2}C_2(\mathbf{u}, T)}\right),$$

and the result follows by (5.20). □

**Lemma 5.3** *Using the notations and assumptions of this section, we have*

$$\begin{aligned} \tilde{\mu}^{\text{sm}}(B_{\rho}^c | \mathbf{Y}_{0:k}) &\leq D_3(\mathbf{u}, T) \cdot \exp\left(-\frac{(\sigma_Z \sqrt{h})^{-2/3}}{D_4(\mathbf{u}, T)}\right), \text{ with} \\ D_3(\mathbf{u}, T) &:= C_{\|\mathbf{A}\|}^{d/2} \cdot \frac{\sqrt{2} \sup_{\mathbf{v} \in \mathcal{B}_R} q(\mathbf{v})}{\sqrt{c(\mathbf{u}, T)}} \cdot (d + 1) \text{ and} \\ D_4(\mathbf{u}, T) &:= \frac{16d \cdot (C_2(\mathbf{u}, T))^2}{c(\mathbf{u}, T)}. \end{aligned}$$

*Proof* Let  $q_{\max} := \sup_{\mathbf{v} \in \mathcal{B}_R} q(\mathbf{v})$ . By our assumption that the event in (5.8) holds, we have

$$\begin{aligned} l^{\text{sm}}(\mathbf{v}) &\geq \frac{c(\mathbf{u}, T)}{h} \|\mathbf{v} - \mathbf{u}\|^2 - \frac{C_1(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z}{\sqrt{h}} \cdot \|\mathbf{v} - \mathbf{u}\| \text{ for every } \mathbf{v} \in \mathcal{B}_R, \text{ and thus} \\ \tilde{\mu}^{\text{sm}}(B_{\rho}^c | \mathbf{Y}_{0:k}) &\leq \frac{\det(\mathbf{A}_k)^{1/2}}{(2\pi)^{d/2} \cdot \sigma_Z^d} \exp\left[-\frac{\mathbf{B}_k \mathbf{A}_k^{-1} \mathbf{B}_k}{2\sigma_Z^2}\right] \\ &\quad \int_{\mathbf{v} \in B_{\rho}^c} \frac{q(\mathbf{v})}{q(\mathbf{u})} \cdot \exp\left[-\frac{\left(\frac{c(\mathbf{u}, T)}{h} \|\mathbf{v} - \mathbf{u}\|^2 - \frac{C_1(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z}{\sqrt{h}} \cdot \|\mathbf{v} - \mathbf{u}\|\right)}{2\sigma_Z^2}\right] d\mathbf{v} \\ &\leq C_{\|\mathbf{A}\|}^{d/2} \cdot \frac{q_{\max}}{(2\pi)^{d/2} \cdot (\sigma_Z \sqrt{h})^d} \\ &\quad \cdot \int_{\mathbf{v} \in B_{\rho}^c} \exp\left[-\frac{\left(c(\mathbf{u}, T) \|\mathbf{v} - \mathbf{u}\|^2 - C_1(\mathbf{u}, T, \frac{\varepsilon}{4}) \sigma_Z \sqrt{h} \cdot \|\mathbf{v} - \mathbf{u}\|\right)}{2\sigma_Z^2 h}\right] d\mathbf{v}, \end{aligned}$$

where in the last step we have used the fact that  $\det(\mathbf{A}_k)^{1/2} \leq \|\mathbf{A}_k\|^{d/2}$ . Based on the assumption that  $\sigma_Z\sqrt{h} \leq C_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , and (5.5), we have

$$\frac{1}{2}c(\mathbf{u}, T)\|\mathbf{v} - \mathbf{u}\|^2 \geq -C_1\left(\mathbf{u}, T, \frac{\varepsilon}{4}\right)\sigma_Z\sqrt{h} \cdot \|\mathbf{v} - \mathbf{u}\| \text{ for } \|\mathbf{v} - \mathbf{u}\| \geq \rho(h, \sigma_Z),$$

therefore by (5.20), we have

$$\begin{aligned} \tilde{\mu}^{\text{sm}}(B_\rho^c|\mathbf{Y}_{0:k}) &\leq C_{\|\mathbf{A}\|}^{d/2} \cdot \frac{q_{\max}}{(2\pi)^{d/2} \cdot (\sigma_Z\sqrt{h})^d} \cdot \int_{\mathbf{v} \in B_\rho^c} \exp\left[-\frac{c(\mathbf{u}, T)\|\mathbf{v} - \mathbf{u}\|^2}{4\sigma_Z^2h}\right] d\mathbf{v} \\ &\leq C_{\|\mathbf{A}\|}^{d/2} \cdot \frac{q_{\max}\sqrt{2}}{\sqrt{c(\mathbf{u}, T)}} \cdot \mathbb{P}\left(\|\mathbf{Z}\| \geq \rho(h, \sigma_Z) \cdot \frac{\sqrt{c(\mathbf{u}, T)}}{\sqrt{2}\sigma_Z\sqrt{h}}\right), \end{aligned}$$

and the claim of the lemma follows by (5.20). □

From inequality (5.16) and Lemmas 5.1, 5.2 and 5.3, it follows that under the assumptions of this section, we can set  $C_{TV}^{(1)}(\mathbf{u}, T)$  and  $C_{TV}^{(2)}(\mathbf{u}, T)$  sufficiently large such that we have

$$\begin{aligned} d_{TV}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) &\leq \int_{\mathbf{v} \in \mathbb{R}^d} |\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} \\ &\leq C_{TV}(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h}. \end{aligned} \tag{5.21}$$

Now we bound the three terms needed for the Wasserstein distance.

**Lemma 5.4** *Using the notations and assumptions of this section, we have*

$$\begin{aligned} &\int_{\mathbf{v} \in B_\rho} \|\mathbf{v} - \mathbf{u}\| \cdot |\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} \\ &\leq \sigma_Z^2h \cdot \left(C_1^*(\mathbf{u}, T) + C_2^*(\mathbf{u}, T) \left(\log\left(\frac{1}{\varepsilon}\right)\right)^2\right), \end{aligned}$$

for some finite positive constants  $C_1^*(\mathbf{u}, T)$ ,  $C_2^*(\mathbf{u}, T)$ .

*Proof* Note that for any  $\mathbf{v} \in B_\rho$ , we have

$$\begin{aligned} \left|\frac{\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1\right| &= \left|\frac{\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} \cdot \frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1\right| \\ &\leq \left|\frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1\right| + \left|\frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} \cdot \frac{\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1\right|. \end{aligned}$$

From (5.21), and the fact that  $\tilde{\mu}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  is a rescaled version of  $\tilde{\mu}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$ , it follows that for  $\sigma_Z\sqrt{h} \leq \frac{1}{2}C_{\text{TV}}(\mathbf{u}, T, \varepsilon)^{-1}$ , we have

$$\left| \frac{\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right| \leq 2C_{\text{TV}}(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h}. \tag{5.22}$$

By (5.18) and (5.14), it follows that  $\frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} \leq 3$  for every  $\mathbf{v} \in B_\rho$ . By using these and bounding  $\left| \frac{\tilde{\mu}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right|$  via (5.18), we obtain that

$$\begin{aligned} & \int_{\mathbf{v} \in B_\rho} \|\mathbf{v} - \mathbf{u}\| \cdot \left| \mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) \right| d\mathbf{v} \\ &= \int_{\mathbf{v} \in B_\rho} \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) \|\mathbf{v} - \mathbf{u}\| \times \left| \frac{\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right| d\mathbf{v} \\ &\leq \int_{\mathbf{v} \in \mathbb{R}^d} \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) \|\mathbf{v} - \mathbf{u}\| \\ &\quad \cdot \left( 2 \left( C_q^{(1)} \|\mathbf{v} - \mathbf{u}\| + \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4})\sigma_Z\sqrt{h}}{h} \cdot \frac{\|\mathbf{v} - \mathbf{u}\|^3}{2\sigma_Z^2} \right) \right. \\ &\quad \left. + 6C_{\text{TV}}(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h} \right) d\mathbf{v} \\ &= \mathbb{E} \left( \|\mathbf{W} - \mathbf{u}\| \cdot 6C_{\text{TV}}(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h} + \|\mathbf{W} - \mathbf{u}\|^2 \cdot 2C_q^{(1)} \right. \\ &\quad \left. + \|\mathbf{W} - \mathbf{u}\|^4 \cdot \frac{C_2(\mathbf{u}, T) + C_3(\mathbf{u}, T, \frac{\varepsilon}{4})\sigma_Z\sqrt{h}}{\sigma_Z^2 h} \Big| \mathbf{A}_k, \mathbf{B}_k \right). \end{aligned}$$

Based on (5.19), we have

$$\|\mathbf{W} - \mathbf{u}\| = \|\sigma_Z \cdot \mathbf{A}_k^{-1/2} \cdot \mathbf{Z} - \mathbf{A}_k^{-1} \mathbf{B}_k\| \leq \sigma_Z\sqrt{h} \left( \sqrt{\frac{2}{c(\mathbf{u}, T)}} \|\mathbf{Z}\| + \frac{2C_{\mathbf{B}}(\frac{\varepsilon}{4})}{c(\mathbf{u}, T)} \right),$$

where  $\mathbf{Z}$  is a  $d$ -dimensional standard normal random vector. The claimed result now follows using the fact that  $\mathbb{E}(\|\mathbf{Z}\|) \leq \sqrt{d}$ ,  $\mathbb{E}(\|\mathbf{Z}\|^2) \leq d$ , and  $\mathbb{E}(\|\mathbf{Z}\|^4) \leq 3d^2$ .  $\square$

**Lemma 5.5** *Using the notations and assumptions of this section, we have*

$$\int_{\mathbf{v} \in B_\rho^c} \|\mathbf{v} - \mathbf{u}\| \mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) d\mathbf{v} \leq 4RD_3(\mathbf{u}, T) \cdot \exp \left( -\frac{(\sigma_Z\sqrt{h})^{-2/3}}{D_4(\mathbf{u}, T)} \right).$$

*Proof* Using (5.22), Lemma 5.3, and the fact that  $\sigma_Z\sqrt{h} \leq \frac{1}{2}C_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , we have

$$\mu^{\text{sm}}(B_\rho^c|\mathbf{Y}_{0:k}) \leq 2\tilde{\mu}^{\text{sm}}(B_\rho^c|\mathbf{Y}_{0:k}) \leq 2D_3(\mathbf{u}, T) \cdot \exp\left(-\frac{(\sigma_Z\sqrt{h})^{-2/3}}{D_4(\mathbf{u}, T)}\right),$$

and the result follows from  $\int_{\mathbf{v} \in B_\rho^c} \|\mathbf{v} - \mathbf{u}\| \mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) d\mathbf{v} \leq 2R \cdot \mu^{\text{sm}}(B_\rho^c|\mathbf{Y}_{0:k})$ .  $\square$

**Lemma 5.6** *Using the notations and assumptions of this section, we have*

$$\int_{\mathbf{v} \in B_\rho^c} \|\mathbf{v} - \mathbf{u}\| \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) d\mathbf{v} \leq C_3^*(\mathbf{u}, T) \exp\left(-C_4^*(\mathbf{u}, T) \cdot (\sigma_Z\sqrt{h})^{-2/3}\right),$$

for some finite positive constants  $C_3^*(\mathbf{u}, T)$ ,  $C_4^*(\mathbf{u}, T)$ .

*Proof* Let  $\mathbf{W}$  be defined as in (5.10) and  $\mathbf{Z}$  be  $d$ -dimensional standard normal. Using the fact that for a nonnegative-valued random variable  $X$ , we have  $\mathbb{E}(X) = \int_{t=0}^\infty \mathbb{P}(X \geq t) dt$ , it follows that

$$\begin{aligned} \int_{\mathbf{v} \in B_\rho^c} \|\mathbf{v} - \mathbf{u}\| \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) d\mathbf{v} &= \mathbb{E}\left(\|\mathbf{W} - \mathbf{u}\| \mathbb{1}_{[\|\mathbf{W} - \mathbf{u}\| \geq \rho(\sigma_Z, h)]} \mid \mathbf{A}_k, \mathbf{B}_k\right) \\ &= (\rho(\sigma_Z, h)) \mu_{\mathcal{G}}^{\text{sm}}(B_\rho^c) + \int_{t=\rho(\sigma_Z, h)}^\infty \mathbb{P}(\|\mathbf{W} - \mathbf{u}\| \geq t) dt \\ &\leq (\rho(\sigma_Z, h)) \mu_{\mathcal{G}}^{\text{sm}}(B_\rho^c) + \int_{t=\rho(\sigma_Z, h)}^\infty \mathbb{P}\left(\|\mathbf{Z}\| \geq \frac{t}{\sigma_Z\sqrt{h}} \cdot \sqrt{\frac{c(\mathbf{u}, T)}{2}}\right) dt \\ &\leq (\rho(\sigma_Z, h)) \mu_{\mathcal{G}}^{\text{sm}}(B_\rho^c) + (d+1) \int_{t=\rho(\sigma_Z, h)}^\infty \exp\left(-\frac{t^2}{4d\sigma_Z^2 h/c(\mathbf{u}, T)}\right) dt \\ &\leq (\rho(\sigma_Z, h))(d+1) \exp\left(-\frac{c(\mathbf{u}, T) \cdot (\sigma_Z\sqrt{h})^{-2/3}}{64dC_2(\mathbf{u}, T)^2}\right) \\ &\quad + (d+1) \sqrt{\frac{4\pi d\sigma_Z^2 h}{c(\mathbf{u}, T)}} \cdot \exp\left(-\frac{\rho(\sigma_Z, h)^2}{4d\sigma_Z^2 h/c(\mathbf{u}, T)}\right), \end{aligned}$$

and the claim of the lemma follows. (We have used Lemma 5.2 in the last step.)  $\square$

From inequality (5.17) and Lemmas 5.1, 5.2 and 5.3, it follows that under the assumptions of this section, for some appropriate choice of constants  $C_W^{(1)}(\mathbf{u}, T)$  and  $C_W^{(2)}(\mathbf{u}, T)$ , we have

$$d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) \leq C_W(\mathbf{u}, T, \varepsilon)\sigma_Z^2 h. \tag{5.23}$$

*Proof of Theorem 2.1* The claim of the theorem follows from inequalities (5.21) and (5.23), and the fact that the assumption that all four of the events in Eqs. (5.6), (5.7), (5.8), and (5.9) hold happens with probability at least  $1 - \varepsilon$ .  $\square$

5.1.2 Gaussian Approximation for the Filter

In this section, we are going to describe the proof of Theorem 2.2. We start by some notation. We define the restriction of  $\mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  to  $\mathcal{B}_R$ , denoted by  $\mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  as

$$\mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(S|\mathbf{Y}_{0:k}) = \frac{\mu_{\mathcal{G}}^{\text{sm}}(S \cap \mathcal{B}_R|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R|\mathbf{Y}_{0:k})} \text{ for any Borel-measurable } S \subset \mathbb{R}^d. \quad (5.24)$$

This is a probability distribution which is supported on  $\mathcal{B}_R$ . We denote its push-forward map by  $\Psi_T$  as  $\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$ , i.e. if a random vector  $\mathbf{X}$  is distributed as  $\mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$ , then  $\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$  denotes the distribution of  $\Psi_T(\mathbf{X})$ .

The proof uses a coupling argument stated in the next two lemmas that allows us to deduce the results based on the Gaussian approximation of the smoother (Theorem 2.1).

**Lemma 5.7** (Coupling argument for total variation distance bound) *The total variation distance of the filtering distribution and its Gaussian approximation can be bounded as follows,*

$$d_{\text{TV}}(\mu^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})) \leq d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) + d_{\text{TV}}(\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})).$$

*Proof* First, notice that by Proposition 3(f) of Roberts and Rosenthal [41], we have

$$\begin{aligned} d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) &= \int_{\mathbf{v} \in \mathcal{B}_R} (\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}))_+ \\ &\leq \int_{\mathbf{v} \in \mathbb{R}^d} (\mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k}))_+ = d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})). \end{aligned}$$

By Proposition 3(g) of Roberts and Rosenthal [41], there is a coupling  $(\mathbf{X}_1, \mathbf{X}_2)$  of random vectors such that  $\mathbf{X}_1 \sim \mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$ ,  $\mathbf{X}_2 \sim \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$ , and  $\mathbb{P}(\mathbf{X}_1 \neq \mathbf{X}_2|\mathbf{Y}_{0:k}) = d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}))$ . Given this coupling, we look at the coupling of the transformed random variables  $(\Psi_T(\mathbf{X}_1), \Psi_T(\mathbf{X}_2))$ . This obviously satisfies that  $\mathbb{P}(\Psi_T(\mathbf{X}_1) \neq \Psi_T(\mathbf{X}_2)|\mathbf{Y}_{0:k}) \leq d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}))$ . Moreover, we have  $\Psi_T(\mathbf{X}_1) \sim \mu^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$  and  $\Psi_T(\mathbf{X}_2) \sim \eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$ ; thus,

$$\begin{aligned} d_{\text{TV}}(\mu^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})) &\leq d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) \\ &\leq d_{\text{TV}}(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})). \end{aligned}$$

The statement of the lemma now follows by the triangle inequality. □

**Lemma 5.8** (Coupling argument for Wasserstein distance bound) *The Wasserstein distance of the filtering distribution and its Gaussian approximation satisfies that*

$$\begin{aligned}
 & d_W(\mu^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})) \\
 & \leq \exp(GT) \cdot [d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) + 2R\mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R^c|\mathbf{Y}_{0:k})] \\
 & \quad + d_W(\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})). \tag{5.25}
 \end{aligned}$$

*Proof* By Theorem 4.1 of Villani [48], there exists a coupling of random variables  $(X_1, X_2)$  (called the optimal coupling) such that  $X_1 \sim \mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$ ,  $X_2 \sim \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$ , and

$$\mathbb{E}(\|X_1 - X_2\||\mathbf{Y}_{0:k}) = d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})).$$

Let  $\hat{X}_2 := X_2 \cdot 1_{[\|X_2\| \leq R]} + \frac{X_2}{\|X_2\|} \cdot R \cdot 1_{[\|X_2\| > R]}$  denote the projection of  $X_2$  on the ball  $\mathcal{B}_R$ . Then using the fact that  $X_1 \in \mathcal{B}_R$ , it is easy to see that  $\|\hat{X}_2 - X_1\| \leq \|X_2 - X_1\|$ , and therefore,

$$d_W(\mathcal{L}(\hat{X}_2|\mathbf{Y}_{0:k}), \mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) \leq d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})),$$

where  $\mathcal{L}(\hat{X}_2|\mathbf{Y}_{0:k})$  denotes the distribution of  $\hat{X}_2$  conditioned on  $\mathbf{Y}_{0:k}$ .

Moreover, by the definitions, for a given  $\mathbf{Y}_{0:k}$ , it is easy to see we can couple random variables  $\hat{X}_2$  and  $\tilde{X}_2 \sim \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})$  such that they are the same with probability at least  $1 - \mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R^c)$ . Since the maximum distance between two points in  $\mathcal{B}_R$  is at most  $2R$ , it follows that

$$d_W(\mathcal{L}(\hat{X}_2|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) \leq 2R\mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R^c).$$

By the triangle inequality, we obtain that

$$d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}|\mathcal{B}_R}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) \leq d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) + 2R\mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R^c),$$

and by (1.6), it follows that

$$d_W(\mu^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})) \leq \exp(GT)[d_W(\mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k})) + 2R\mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R^c)].$$

The claim of the lemma now follows by the triangle inequality. □

As we can see, the above results still require us to bound the total variation and Wasserstein distances between the distributions  $\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$  and  $\mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$ . Let

$$\mathbf{A}_k^{\text{fi}} := ((\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1})' \cdot \mathbf{A}_k \cdot (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1}, \tag{5.26}$$

then the density of  $\mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$  can be written as

$$\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) := \frac{(\det(\mathbf{A}_k))^{\frac{1}{2}}}{|\det(\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))|} \cdot \frac{1}{(2\pi)^{d/2} \cdot \sigma_Z^d} \cdot \exp\left[-\frac{(\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}}))' \mathbf{A}_k^{\text{fi}} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}}))}{2\sigma_Z^2}\right]. \tag{5.27}$$

Since the normalising constant is not known for the case of  $\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$ , we define a rescaled version  $\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})$  with density

$$\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) := 1_{[\mathbf{v} \in \Psi_T(\mathcal{B}_R)]} \cdot \frac{(\det(\mathbf{A}_k))^{\frac{1}{2}}}{(2\pi)^{d/2} \cdot \sigma_Z^d} \cdot \exp\left[-\frac{(\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}})' \mathbf{A}_k (\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}})}{2\sigma_Z^2}\right]. \tag{5.28}$$

The following lemma bounds the difference between the logarithms of  $\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})$  and  $\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})$ .

**Lemma 5.9** *For any  $\mathbf{v} \in \Psi_T(\mathcal{B}_R)$ , we have*

$$\begin{aligned} & |\log(\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})) - \log(\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}))| \\ & \leq \frac{M_2(T)\|\mathbf{A}_k\| \exp(4GT)\|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|^3}{2\sigma_Z^2} + M_1(T)M_2(T)d\|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|. \end{aligned}$$

*Proof of Lemma 5.9* By (5.27) and (5.28), we have

$$\begin{aligned} & \log(\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})) - \log(\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})) = \log|\det(\mathbf{J}\Psi_{-T}(\mathbf{v}))| - \log|\det(\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}})\Psi_{-T})| \\ & + \frac{1}{2\sigma_Z^2} \cdot \left[ (\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}})' \mathbf{A}_k (\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}}) \right. \\ & \left. - \left( (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right)' \mathbf{A}_k (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right]. \end{aligned}$$

The absolute value of the first difference can be bounded by Proposition 4.4 as

$$|\log|\det(\mathbf{J}\Psi_{-T}(\mathbf{v}))| - \log|\det(\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}})\Psi_{-T})|| \leq \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| \cdot M_1(T)M_2(T)d.$$

For any two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ , we have

$$\begin{aligned} & |\mathbf{x}' \mathbf{A}_k \mathbf{x} - \mathbf{y}' \mathbf{A}_k \mathbf{y}| \\ & = |\mathbf{x}' \mathbf{A}_k \mathbf{x} - \mathbf{x}' \mathbf{A}_k \mathbf{y} + \mathbf{x}' \mathbf{A}_k \mathbf{y} - \mathbf{y}' \mathbf{A}_k \mathbf{y}| \leq \|\mathbf{A}_k\| \|\mathbf{x} - \mathbf{y}\| (\|\mathbf{x}\| + \|\mathbf{y}\|), \end{aligned}$$

so the second difference can be bounded as

$$\begin{aligned} & |(\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}})' \mathbf{A}_k (\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}}) \\ & \quad - \left( (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right)' \mathbf{A}_k (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}}))| \\ & \leq \|\mathbf{A}_k\| \cdot \left\| \Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}} - (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right\| \\ & \quad \cdot \left( \left\| \Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}} \right\| + \left\| (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right\| \right). \end{aligned}$$

Using (1.6), we have

$$\left\| \Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}} \right\| + \left\| (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right\| \leq 2 \exp(GT) \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|.$$

By (1.19), we have

$$\begin{aligned} & \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}}) - \mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}})(\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}})\| \\ & = \|\Psi_T(\Psi_{-T}(\mathbf{v})) - \Psi_T(\mathbf{u}^{\mathcal{G}}) - \mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}})(\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}})\| \\ & \leq \frac{1}{2} M_2(T) \|\Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}}\|^2, \end{aligned}$$

so by (1.6), it follows that

$$\begin{aligned} & \left\| \Psi_{-T}(\mathbf{v}) - \mathbf{u}^{\mathcal{G}} - (\mathbf{J}\Psi_T(\mathbf{u}^{\mathcal{G}}))^{-1} (\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})) \right\| \\ & \leq \frac{1}{2} M_2(T) \exp(3GT) \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|^2. \end{aligned}$$

We obtain the claim of the lemma by combining the stated bounds. □

Now we are ready to prove our Gaussian approximation result for the filter.

*Proof of Theorem 2.2* We suppose that  $D_{\text{TV}}^{(1)}(\mathbf{u}, T) \geq C_{\text{TV}}^{(1)}(\mathbf{u}, T)$  and  $D_{\text{TV}}^{(2)}(\mathbf{u}, T) \geq C_{\text{TV}}^{(2)}(\mathbf{u}, T)$ , and thus  $D_{\text{TV}}(\mathbf{u}, T, \varepsilon) \geq C_{\text{TV}}(\mathbf{u}, T, \varepsilon)$ . We also assume that  $D_{\text{TV}}^{(1)}(\mathbf{u}, T)$  satisfies that

$$D_{\text{TV}}^{(1)}(\mathbf{u}, T) \geq \frac{2^{\frac{5}{2}} d^{\frac{3}{2}} M_2(T) \exp(4GT)}{\sqrt{C_{\|\mathbf{A}\|}}}, \tag{5.29}$$

$$D_{\text{TV}}^{(1)}(\mathbf{u}, T) \geq \frac{2\sqrt{M_2(T) \exp(4GT) C_{\|\mathbf{A}\|}}}{(R - \|\mathbf{u}\|)^{\frac{3}{2}}}. \tag{5.30}$$

Based on these assumptions on  $D_{TV}(\mathbf{u}, T, \varepsilon)$ , and the assumption that  $\sigma_Z\sqrt{h} \leq \frac{1}{2}D_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , it follows that the probability that all the four events in Eqs. (5.6), (5.7), (5.8) and (5.9) hold is at least  $1 - \varepsilon$ . We are going to assume that this is the case for the rest of the proof. We define

$$\begin{aligned} \rho'(\sigma_Z, h) &:= \frac{(4\sigma_Z^2h)^{\frac{1}{3}}}{(M_2(T) \exp(4GT)C_{\|A\|})^{\frac{1}{3}}} \text{ and} \\ \mathcal{B}_{\rho'} &:= \{\mathbf{v} \in \mathbb{R}^d : \|\mathbf{v} - \Psi_T(\mathbf{u})\| \leq \rho'(\sigma_Z, h)\}. \end{aligned} \tag{5.31}$$

Based on (5.29), (5.30), and the assumption that  $\sigma_Z\sqrt{h} \leq \frac{1}{2}D_{TV}(\mathbf{u}, T, \varepsilon)^{-1}$ , it follows that

$$\begin{aligned} &\rho'(\sigma_Z, h) \\ &\leq \min \left( \frac{(4\sigma_Z^2h)^{\frac{1}{3}}}{(M_2(T) \exp(4GT)C_{\|A\|})^{\frac{1}{3}}}, \frac{1}{2dM_2(T) \exp(4GT)}, (R - \|\mathbf{u}\|) \exp(-GT) \right). \end{aligned} \tag{5.32}$$

The surface of a ball of radius  $R - \|\mathbf{u}\|$  centred at  $\mathbf{u}$  is contained in  $\mathcal{B}_R$ , and it will be transformed by  $\Psi_T$  to a closed continuous manifold whose points are at least  $(R - \|\mathbf{u}\|) \exp(-GT)$  away from  $\Psi_T(\mathbf{u})$  (based on (1.6)). This implies that the ball of radius  $(R - \|\mathbf{u}\|) \exp(-GT)$  centred at  $\Psi_T(\mathbf{u})$  is contained in  $\Psi_T(\mathcal{B}_R)$ , and thus by (5.32),  $\mathcal{B}_{\rho'} \subset \Psi_T(\mathcal{B}_R)$ .

By Lemma 4.1, we have

$$\begin{aligned} d_{TV} \left( \tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}) \right) &\leq \int_{\mathbf{v} \in \mathcal{B}_{\rho'}} |\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} \\ &\quad + \tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathcal{B}_{\rho'}^c|\mathbf{Y}_{0:k}) + \mu_{\mathcal{G}}^{\text{fi}}(\mathcal{B}_{\rho'}^c|\mathbf{Y}_{0:k}). \end{aligned} \tag{5.33}$$

By Lemma 5.9, (5.32), and the fact that  $|\exp(x) - 1| \leq 2|x|$  for  $x \in [-1, 1]$ , it follows that for  $\mathbf{v} \in \mathcal{B}_{\rho'}$ , we have

$$\begin{aligned} \left| \frac{\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right| &\leq \frac{M_2(T)C_{\|A\|} \exp(4GT)\|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|^3}{\sigma_Z^2h} \\ &\quad + 2M_1(T)M_2(T)d\|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|. \end{aligned}$$

Therefore, the first term of (5.33) can be bounded as

$$\begin{aligned} & \int_{\mathbf{v} \in \mathcal{B}_{\rho'}} |\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})| d\mathbf{v} = \int_{\mathbf{v} \in \mathcal{B}_{\rho'}} \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \left| \frac{\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right| d\mathbf{v} \\ & \leq \int_{\mathbf{v} \in \mathbb{R}^d} \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \left( \frac{M_2(T)C_{\|A\|} \exp(4GT) \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|^3}{\sigma_Z^2 h} \right. \\ & \quad \left. + 2M_1(T)M_2(T)d\|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| \right) d\mathbf{v}. \end{aligned}$$

This in turn can be bounded as in Lemma 5.1. The terms  $\tilde{\eta}_{\mathcal{G}}^{\text{fi}}(\mathcal{B}_{\rho'}^c|\mathbf{Y}_{0:k})$  and  $\mu_{\mathcal{G}}^{\text{fi}}(\mathcal{B}_{\rho'}^c|\mathbf{Y}_{0:k})$  can be bounded in a similar way as in Lemmas 5.2 and 5.3. Therefore, by Lemma 5.7 we obtain that under the assumptions of this section, there are some finite constants  $D_{\text{TV}}^{(1)}(\mathbf{u}, T)$  and  $D_{\text{TV}}^{(2)}(\mathbf{u}, T)$  such that

$$d_{\text{TV}} \left( \mu^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}) \right) \leq D_{\text{TV}}(\mathbf{u}, T, \varepsilon) \sigma_Z \sqrt{h}. \tag{5.34}$$

For the Wasserstein distance bound, the proof is based on Lemma 5.8. Note that by the proof of Theorem 2.1, under the assumptions on this section, we have

$$d_W \left( \mu^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot|\mathbf{Y}_{0:k}) \right) \leq C_W(\mathbf{u}, T, \varepsilon) \sigma_Z^2 h.$$

Therefore, we only need to bound the last two terms of (5.25). The fact that

$$\mu_{\mathcal{G}}^{\text{sm}}(\mathcal{B}_R^c|\mathbf{Y}_{0:k}) = o(\sigma_Z^2 h)$$

can be shown similarly to the proof of Lemma 5.3. Finally, the last term can be bounded by applying Lemma 4.2 for  $\mathbf{y} := \Psi_T(\mathbf{u}^{\mathcal{G}})$ . This implies that

$$\begin{aligned} d_W(\eta_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{fi}}(\cdot|\mathbf{Y}_{0:k})) & \leq \int_{\mathbf{v} \in \mathbb{R}^d} \left| \eta_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) - \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \right| \cdot \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| d\mathbf{v} \\ & = \int_{\mathbf{v} \in \mathbb{R}^d} \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \left| \frac{\eta_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right| \cdot \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| d\mathbf{v} \\ & \leq \int_{\mathbf{v} \in \mathcal{B}_{\rho'}} \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \left| \frac{\eta_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})}{\mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k})} - 1 \right| \cdot \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| d\mathbf{v} \\ & \quad + \int_{\mathbf{v} \in \mathcal{B}_{\rho'}^c} \eta_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \cdot \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| d\mathbf{v} + \int_{\mathbf{v} \in \mathcal{B}_{\rho'}^c} \mu_{\mathcal{G}}^{\text{fi}}(\mathbf{v}|\mathbf{Y}_{0:k}) \cdot \|\mathbf{v} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| d\mathbf{v}. \end{aligned}$$

These terms can be bounded in a similar way as in Lemmas 5.4, 5.5 and 5.6, and the claim of the theorem follows. □

## 5.2 Comparison of Mean Square Error of MAP and Posterior Mean

In the following two subsections, we are going to prove our results concerning the mean square error of the MAP estimator for the smoother, and the filter, respectively.

### 5.2.1 Comparison of MAP and Posterior Mean for the Smoother

In this section, we are going to prove Theorem 2.3. First, we introduce some notations. Let  $\mathbf{u}^{\mathcal{G}} := \mathbf{u} - \mathbf{A}_k^{-1} \mathbf{B}_k$  denote the centre of the Gaussian approximation  $\mu_{\mathcal{G}}^{\text{sm}}$  (defined when  $\mathbf{A}_k$  is positive definite). Let  $\rho(h, \sigma_Z)$  be as in (5.13),  $B_\rho := \{\mathbf{v} \in \mathbb{R}^d : \|\mathbf{v} - \mathbf{u}\| \leq \rho(h, \sigma_Z)\}$  and  $B_\rho^c$  be the complement of  $B_\rho$ . The proof is based on several lemmas which are described as follows. All of them implicitly assume that the assumptions of Theorem 2.3 hold.

**Lemma 5.10** (A bound on  $\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|$ ) *There are some finite constants  $D_5(\mathbf{u}, T)$  and  $D_6(\mathbf{u}, T)$  such that for any  $0 < \varepsilon \leq 1$ , for  $\sigma_Z \sqrt{h} \leq \frac{1}{2} \cdot C_{\text{TV}}(\mathbf{u}, T, \varepsilon)^{-1}$ , we have*

$$\mathbb{P}\left(\frac{c(\mathbf{u}, T)}{2h} \mathbf{I}_d \prec \mathbf{A}_k \prec \frac{C_{\|\mathbf{A}\|}}{h} \mathbf{I}_d\right)$$

and  $\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| \leq \left(D_5(\mathbf{u}, T) + D_6(\mathbf{u}, T) \left(\log\left(\frac{1}{\varepsilon}\right)\right)^2\right) \sigma_Z^2 h \Big| \mathbf{u} \geq 1 - \varepsilon.$

*Proof* This is a direct consequence of the Wasserstein distance bound of Theorem 2.1, since

$$\begin{aligned} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| &= \left\| \int_{\mathbf{x} \in \mathbb{R}^d} \mathbf{x} \cdot \mu^{\text{sm}}(\mathbf{x} | \mathbf{Y}_{0:k}) d\mathbf{x} - \int_{\mathbf{y} \in \mathbb{R}^d} \mathbf{y} \cdot \mu_{\mathcal{G}}^{\text{sm}}(\mathbf{y} | \mathbf{Y}_{0:k}) d\mathbf{y} \right\| \\ &\leq d_W(\mu^{\text{sm}}(\cdot | \mathbf{Y}_{0:k}), \mu_{\mathcal{G}}^{\text{sm}}(\cdot | \mathbf{Y}_{0:k})). \end{aligned}$$

□

**Lemma 5.11** (A bound on  $\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\|$ ) *For any  $0 < \varepsilon \leq 1$ , we have*

$$\mathbb{P}\left(\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\| \leq \frac{C_1(\mathbf{u}, T, \varepsilon) \sigma_Z \sqrt{h} + 2C_q^{(1)} \sigma_Z^2 h}{c(\mathbf{u}, T)} \Big| \mathbf{u}\right) \geq 1 - \varepsilon. \tag{5.35}$$

*Proof* From Proposition 4.1, and (4.3), it follows that for any  $0 < \varepsilon \leq 1$ ,

$$\begin{aligned} \mathbb{P}\left(\log \mu^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) - \log \mu^{\text{sm}}(\mathbf{u} | \mathbf{Y}_{0:k}) \leq -\frac{1}{2\sigma_Z^2 h} \cdot \left(c(\mathbf{u}, T) \|\mathbf{v} - \mathbf{u}\|^2\right.\right. \\ \left.\left. - \left(C_1(\mathbf{u}, T, \varepsilon) \sigma_Z \sqrt{h} + 2C_q^{(1)} \sigma_Z^2 h\right) \|\mathbf{v} - \mathbf{u}\|\right) \text{ for every } \mathbf{v} \in \mathcal{B}_R\right) \geq 1 - \varepsilon. \end{aligned}$$

Since  $\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}$  is the maximiser of  $\log \mu^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k})$  on  $\mathcal{B}_R$ , our claim follows. □

**Lemma 5.12** (A bound on  $\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|$ ) *There are finite constants  $S_{\text{MAP}}^{(1)} > 0$ ,  $S_{\text{MAP}}^{(2)}$ ,  $D_7(\mathbf{u}, T)$  and  $D_8(\mathbf{u}, T)$  such that for any  $0 < \varepsilon \leq 1$ , for*

$$\sigma_Z \sqrt{h} < \left( S_{\text{MAP}}^{(1)} + S_{\text{MAP}}^{(2)} \left( \log \left( \frac{1}{\varepsilon} \right) \right)^{1/2} \right)^{-1},$$

we have

$$\mathbb{P} \left( \mathbf{A}_k \succ \mathbf{0} \text{ and } \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| \leq \left( D_7(\mathbf{u}, T) + D_8(\mathbf{u}, T) \left( \log \left( \frac{1}{\varepsilon} \right) \right)^{3/2} \right) \sigma_Z^2 h \middle| \mathbf{u} \right) \geq 1 - \varepsilon. \tag{5.36}$$

*Proof* By choosing  $S_{\text{MAP}}^{(1)}$  and  $S_{\text{MAP}}^{(2)}$  sufficiently large, we can assume that

$$\sigma_Z \sqrt{h} < \min \left( \frac{c(\mathbf{u}, T)}{2} \cdot \frac{R - \|\mathbf{u}\|}{C_1(\mathbf{u}, T, \frac{\varepsilon}{3})}, \sqrt{\frac{(R - \|\mathbf{u}\|)c(\mathbf{u}, T)}{C_q^{(1)}}}, \frac{c(\mathbf{u}, T)}{2C_A(\frac{\varepsilon}{3})} \right). \tag{5.37}$$

From Lemma 5.11, we know that

$$\mathbb{P} \left( \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\| \leq \frac{C_1(\mathbf{u}, T, \frac{\varepsilon}{3})\sigma_Z \sqrt{h} + 2C_q^{(1)}\sigma_Z^2 h}{c(\mathbf{u}, T)} \middle| \mathbf{u} \right) \geq 1 - \frac{\varepsilon}{3}. \tag{5.38}$$

Using (5.37), it follows that if the above event happens, then  $\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}\| < R$ , and thus

$$\nabla \log(\mu^{\text{sm}}(\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} | \mathbf{Y}_{0:k})) = \nabla l^{\text{sm}}(\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}) - 2\sigma_Z^2 \nabla \log q(\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}) = 0. \tag{5.39}$$

Using the fact that  $\nabla l_{\mathcal{G}}^{\text{sm}}(\mathbf{v}) = 2\mathbf{A}_k(\mathbf{v} - \mathbf{u}^{\mathcal{G}})$ , and Proposition 4.3, it follows that

$$\mathbb{P} \left( \|\nabla l^{\text{sm}}(\mathbf{v})\| \geq \|2\mathbf{A}_k(\mathbf{v} - \mathbf{u}^{\mathcal{G}})\| - \|\mathbf{v} - \mathbf{u}\|^2 \cdot \frac{C_4(\mathbf{u}, T) + C_5(\mathbf{u}, T, \frac{\varepsilon}{3})\sigma_Z \sqrt{h}}{h} \middle| \mathbf{u} \right) \geq 1 - \frac{\varepsilon}{3}. \tag{5.40}$$

Moreover, by Lemma 4.4, we know that for any  $0 < \varepsilon \leq 1$ , we have

$$\mathbb{P} \left( \lambda_{\min}(\mathbf{A}_k) > \frac{c(\mathbf{u}, T)}{2h} \middle| \mathbf{u} \right) \geq 1 - \frac{\varepsilon}{3} \text{ for } \sigma_Z \sqrt{h} \leq \frac{c(\mathbf{u}, T)}{2C_A(\frac{\varepsilon}{3})}. \tag{5.41}$$

By combining the four Eqs. (5.38), (5.39), (5.40) and (5.41), it follows that with probability at least  $1 - \varepsilon$ , we have

$$2\sigma_Z^2 C_q^{(1)} \geq \frac{c(\mathbf{u}, T)}{h} \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| - \left( \frac{C_1(\mathbf{u}, T, \frac{\varepsilon}{3})\sigma_Z\sqrt{h} + 2C_q^{(1)}\sigma_Z^2 h}{c(\mathbf{u}, T)} \right)^2 \cdot \frac{C_4(\mathbf{u}, T) + C_5(\mathbf{u}, T, \frac{\varepsilon}{3})\sigma_Z\sqrt{h}}{h},$$

and the claim of the lemma follows by rearrangement. □

**Lemma 5.13** (A lower bound on  $\mathbb{E}(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 | \mathbf{u})$ ) *There are positive constants  $D_9(\mathbf{u}, T)$  and  $D_{10}(\mathbf{u}, T)$  such that for  $\sigma_Z\sqrt{h} \leq D_{10}(\mathbf{u}, T)$ , we have*

$$\mathbb{E}(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 | \mathbf{u}) \geq D_9(\mathbf{u}, T) \cdot \sigma_Z^2 h. \tag{5.42}$$

*Proof* By applying Lemma 5.10 for  $\varepsilon = 0.1$ , we obtain that for

$$\sigma_Z\sqrt{h} \leq \frac{1}{2} \cdot C_{\text{TV}}(\mathbf{u}, T, 0.1)^{-1},$$

we have

$$\mathbb{P}\left(\lambda_{\min}(\mathbf{A}_k) > \frac{c(\mathbf{u}, T)}{2h} \text{ and } \|\mathbf{A}_k\| < \frac{C_{\|\mathbf{A}\|}}{h}\right) \text{ and } \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| \leq \left(D_5(\mathbf{u}, T) + D_6(\mathbf{u}, T) (\log(10))^2\right) \sigma_Z^2 h | \mathbf{u} \geq 0.9. \tag{5.43}$$

If this event happens, then in particular, we have

$$\|\mathbf{u}^{\mathcal{G}} - \mathbf{u}\| = \|\mathbf{A}_k^{-1} \mathbf{B}_k\| \geq \frac{h}{C_{\|\mathbf{A}\|}} \cdot \|\mathbf{B}_k\|. \tag{5.44}$$

By the definition of  $\mathbf{B}_k$  in (2.8), it follows that conditioned on  $\mathbf{u}$ ,  $\mathbf{B}_k$  has  $d$ -dimensional multivariate normal distribution with covariance matrix

$$\Sigma_{\mathbf{B}_k} := \sigma_Z^2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \mathbf{J}\Phi_{t_i}(\mathbf{u}).$$

This means that if  $\mathbf{Z}$  is a  $d$ -dimensional standard normal random vector, then  $(\Sigma_{\mathbf{B}_k})^{1/2} \cdot \mathbf{Z}$  has the same distribution as  $\mathbf{B}_k$  (conditioned on  $\mathbf{u}$ ). By Assumption 2.1, we have  $\lambda_{\min}(\Sigma_{\mathbf{B}_k}) \geq \sigma_Z^2 \cdot \frac{c(\mathbf{u}, T)}{h}$ , and thus  $\|(\Sigma_{\mathbf{B}_k})^{1/2} \cdot \mathbf{Z}\| \geq \sigma_Z \cdot \sqrt{\frac{c(\mathbf{u}, T)}{h}} \cdot \|\mathbf{Z}\|$ .

It is not difficult to show that for any  $d \geq 1$ ,  $\mathbb{P}\left(\|\mathbf{Z}\| \geq \frac{\sqrt{d}}{2}\right) \geq \frac{1}{4}$ . (Indeed, if  $(Z_{(i)})_{1 \leq i \leq d}$  are i.i.d. standard normal random variables, then for any  $\lambda > 0$ ,  $\mathbb{E}(e^{-\lambda Z_1^2}) = \sqrt{\frac{1}{1+2\lambda}}$ , so  $\mathbb{E}(e^{-\lambda(\|\mathbf{Z}\|^2-d)}) = \mathbb{E}(e^{-\lambda \sum_{i=1}^d (Z_{(i)}^2-1)}) = (1+2\lambda)^{-d/2} \cdot e^{\lambda d}$ , and the claim follows by applying Markov's inequality  $\mathbb{P}(\|\mathbf{Z}\|^2 - d \leq -t) \leq \mathbb{E}(e^{-\lambda(\|\mathbf{Z}\|^2-d)}) \cdot e^{-\lambda t}$  for  $t = \frac{3}{4}d$  and  $\lambda = 1$ .) Therefore, we have

$$\mathbb{P}\left(\|\mathbf{B}_k\| \geq \sigma_Z \cdot \sqrt{\frac{c(\mathbf{u}, T)}{h}} \cdot \frac{\sqrt{d}}{2} \mid \mathbf{u}\right) \geq \frac{1}{4},$$

and thus by (5.43) and (5.44), it follows that

$$\begin{aligned} \mathbb{P}\left(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\| \geq \frac{\sigma_Z \sqrt{h} \cdot \sqrt{dc(\mathbf{u}, T)}}{2C_{\|\mathbf{A}\|}} - \left(D_5(\mathbf{u}, T) + D_6(\mathbf{u}, T) (\log(10))^2\right) \sigma_Z^2 h\right) \\ \geq 0.15. \end{aligned}$$

By choosing  $D_{10}(\mathbf{u}, T)$  sufficiently small, we have that for  $\sigma_Z \sqrt{h} \leq D_{10}(\mathbf{u}, T)$ ,

$$\left(D_5(\mathbf{u}, T) + D_6(\mathbf{u}, T) (\log(10))^2\right) \sigma_Z^2 h \leq \frac{1}{2} \cdot \frac{\sigma_Z \sqrt{h} \cdot \sqrt{dc(\mathbf{u}, T)}}{2C_{\|\mathbf{A}\|}},$$

and the result follows. □

**Lemma 5.14** (Bounding  $|\mathbb{E}(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u}) - \mathbb{E}(\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u})|$ ) *There are some finite constants  $D_{11}(\mathbf{u}, T)$  and  $D_{12}(\mathbf{u}, T) > 0$  such that for  $\sigma_Z \sqrt{h} \leq D_{12}(\mathbf{u}, T)$ , we have*

$$\left|\mathbb{E}\left(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u}\right) - \mathbb{E}\left(\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u}\right)\right| \leq D_{11}(\mathbf{u}, T) \cdot (\sigma_Z^2 h)^{\frac{3}{2}}. \tag{5.45}$$

*Proof* We define the event  $E_k$  as

$$E_k := \left\{ \frac{c(\mathbf{u}, T)}{2h} \cdot \mathbf{I}_d < \mathbf{A}_k < \frac{C_{\|\mathbf{A}\|}}{h} \cdot \mathbf{I}_d, \|\mathbf{A}_k^{-1} \mathbf{B}_k\| < R - \|\mathbf{u}\| \right\}. \tag{5.46}$$

Under this event, we have, in particular,  $A_k > \mathbf{0}$  and  $\|\mathbf{u}^G\| < R$ . Let  $E_k^c$  denote the complement of  $E_k$ . Then the difference in the variances can be bounded as

$$\begin{aligned}
 & \left| \mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u} \right) - \mathbb{E} \left( \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u} \right) \right| \\
 & \leq \mathbb{E} \left( 4R^2 1_{E_k^c} \right. \\
 & \quad \left. + 1_{E_k} \left( \left| \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 - \|\mathbf{u}^G - \mathbf{u}\|^2 \right| + \left| \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\|^2 - \|\mathbf{u}^G - \mathbf{u}\|^2 \right| \right) \mid \mathbf{u} \right) \\
 & \leq 4R^2 \mathbb{P} \left( E_k^c \mid \mathbf{u} \right) + \mathbb{E} \left( 1_{E_k} \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^G\| \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\| + 2\|\mathbf{u}^G - \mathbf{u}\| \right) \right. \right. \\
 & \quad \left. \left. + \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^G\| \left( \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^G\| + 2\|\mathbf{u}^G - \mathbf{u}\| \right) \right) \mid \mathbf{u} \right) \\
 & \leq 4R^2 \mathbb{P} \left( E_k^c \mid \mathbf{u} \right) + \mathbb{E} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^G\|^2 \mid \mathbf{u} \right) + \mathbb{E} \left( 1_{E_k} \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^G\|^2 \mid \mathbf{u} \right) \\
 & \quad + 2\sqrt{\mathbb{E} \left( 1_{E_k} \|\mathbf{u}^G - \mathbf{u}\|^2 \mid \mathbf{u} \right)} \\
 & \quad \cdot \left( \sqrt{\mathbb{E} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^G\|^2 \mid \mathbf{u} \right)} + \sqrt{\mathbb{E} \left( 1_{E_k} \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^G\|^2 \mid \mathbf{u} \right)} \right), \tag{5.47}
 \end{aligned}$$

where in the last step we have used the Cauchy–Schwarz inequality. The above terms can be further bounded as follows. By Lemmas 4.4 and 4.3 it follows that

$$\begin{aligned}
 & \mathbb{P} \left( E_k^c \mid \mathbf{u} \right) \\
 & = \mathbb{P} \left( \lambda_{\min}(\mathbf{A}_k) \leq \frac{c(\mathbf{u}, T)}{2h} \text{ or } \|\mathbf{A}_k\| \leq \frac{C_{\|\mathbf{A}\|}}{h} \text{ or } \|\mathbf{A}_k^{-1} \mathbf{B}_k\| \geq R - \|\mathbf{u}\| \mid \mathbf{u} \right) \\
 & \leq 2d \exp \left( -\frac{c(\mathbf{u}, T)^2}{4\bar{T}(\mathbf{u}) \widehat{M}_2(T)^2 d_o \sigma_Z^2 h} \right) + (d + 1) \exp \left( -\frac{(R - \|\mathbf{u}\|)^2 c(\mathbf{u}, T)^2}{\bar{T}(\mathbf{u}) \widehat{M}_1(T) d_o \sigma_Z^2 h} \right) \\
 & \leq C_E(\mathbf{u}, T) (\sigma_Z^2 h)^2, \tag{5.48}
 \end{aligned}$$

for some finite constant  $C_E(\mathbf{u}, T)$  independent of  $h$  and  $\sigma_Z$ .

The term  $\mathbb{E} \left( 1_{E_k} \|\mathbf{u}^G - \mathbf{u}\|^2 \mid \mathbf{u} \right)$  can be bounded as

$$\begin{aligned}
 \mathbb{E} \left( 1_{E_k} \|\mathbf{u}^G - \mathbf{u}\|^2 \mid \mathbf{u} \right) & = \mathbb{E} \left( 1_{E_k} \|\mathbf{A}_k^{-1} \mathbf{B}_k\|^2 \mid \mathbf{u} \right) \leq \left( \frac{2h}{c(\mathbf{u}, T)} \right)^2 \cdot \mathbb{E} \left( 1_{E_k} \|\mathbf{B}_k\|^2 \mid \mathbf{u} \right) \\
 & \leq \left( \frac{2h}{c(\mathbf{u}, T)} \right)^2 \widehat{M}_1(T)^2 d_o \frac{\bar{T}(\mathbf{u})}{h} \sigma_Z^2 = \frac{4\bar{T}(\mathbf{u}) \widehat{M}_1(T)^2 d_o}{c(\mathbf{u}, T)^2} \cdot \sigma_Z^2 h. \tag{5.49}
 \end{aligned}$$

For bounding the term  $\mathbb{E} ( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 | \mathbf{u} )$ , we define

$$t_{\min} := D_5(\mathbf{u}, T)\sigma_Z^2 h, \text{ and}$$

$$t_{\max} := \left( D_5(\mathbf{u}, T) + D_6(\mathbf{u}, T) \cdot \left( \frac{(2\sigma_Z\sqrt{h})^{-1} - C_{\text{TV}}^{(1)}(\mathbf{u}, T)}{C_{\text{TV}}^{(2)}(\mathbf{u}, T)} \right)^{\frac{5}{4}} \right) \sigma_Z^2 h,$$

and then by Lemma 5.10, it follows that for  $\sigma_Z\sqrt{h} < \frac{1}{2}(C_{\text{TV}}^{(1)})^{-1}$ , for  $t \in [t_{\min}, t_{\max}]$ , we have

$$\mathbb{P} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| \geq t \mid \mathbf{u} \right) \leq \exp \left( - \left( \frac{t/(\sigma_Z^2 h) - D_5(\mathbf{u}, T)}{D_6(\mathbf{u}, T)} \right)^{\frac{2}{5}} \right). \tag{5.50}$$

By writing

$$\begin{aligned} \mathbb{E} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 \mid \mathbf{u} \right) &= \int_{t=0}^{\infty} \mathbb{P}(1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 > t) dt \\ &= \int_{t=0}^{\infty} \mathbb{P}(1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| > \sqrt{t}) dt, \end{aligned}$$

and using the fact that  $1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 < 4R^2$ , one can show that for  $\sigma_Z\sqrt{h} < \bar{S}(\mathbf{u}, T)$ ,

$$\mathbb{E} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 \mid \mathbf{u} \right) \leq \bar{C}(\mathbf{u}, T)(\sigma_Z^2 h)^2, \tag{5.51}$$

for some constants  $\bar{S}(\mathbf{u}, T) > 0$ ,  $\bar{C}(\mathbf{u}, T) < \infty$ , that are independent of  $\sigma_Z$  and  $h$ .

Finally, for bounding the term  $\mathbb{E} ( 1_{E_k} \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 | \mathbf{u} )$ , we define

$$t'_{\min} := D_7(\mathbf{u}, T)\sigma_Z^2 h, \text{ and}$$

$$t'_{\max} := \left( D_7(\mathbf{u}, T) + D_8(\mathbf{u}, T) \left( \frac{(\sigma_Z\sqrt{h})^{-1} - S_{\text{MAP}}^{(1)}}{S_{\text{MAP}}^{(2)}} \right)^3 \right) \cdot \sigma_Z^2 h.$$

By Lemma 5.12, it follows that for  $\sigma_Z\sqrt{h} < (S_{\text{MAP}}^{(1)})^{-1}$ , for  $t \in [t'_{\min}, t'_{\max}]$ , we have

$$\mathbb{P} \left( \|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\| > t \mid \mathbf{u} \right) \leq \exp \left( - \left( \frac{t/(\sigma_Z^2 h) - D_7(\mathbf{u}, T)}{D_8(\mathbf{u}, T)} \right)^{\frac{2}{3}} \right), \tag{5.52}$$

which implies that for  $\sigma_Z\sqrt{h} < S_{\text{M}}(\mathbf{u}, T)$ ,

$$\mathbb{E} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 \mid \mathbf{u} \right) \leq C_{\text{M}}(\mathbf{u}, T)(\sigma_Z^2 h)^2, \tag{5.53}$$

for some constants  $S_M(\mathbf{u}, T) > 0, C_M(\mathbf{u}, T) > 0$ .

The result now follows by (5.35) and the bounds (5.48), (5.49), (5.51) and (5.53).  $\square$

*Proof of Theorem 2.3* The lower bound on  $\frac{\mathbb{E}(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 | \mathbf{u})}{\sigma_Z^2 h}$  follows by Lemma 5.13. Let  $E_k$  be defined as in (5.46), then we have

$$\begin{aligned} \mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 \mid \mathbf{u} \right) &\leq \mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 \cdot 1_{E_k^c} \mid \mathbf{u} \right) \\ &+ 2\mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 \cdot 1_{E_k} \mid \mathbf{u} \right) + 2\mathbb{E} \left( \|\mathbf{u}^{\mathcal{G}} - \mathbf{u}\|^2 \cdot 1_{E_k} \mid \mathbf{u} \right) \\ &\leq 4R^2 \mathbb{P} \left( 1_{E_k^c} \mid \mathbf{u} \right) + 2\mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}^{\mathcal{G}}\|^2 \cdot 1_{E_k} \mid \mathbf{u} \right) + 2\mathbb{E} \left( \|\mathbf{u}^{\mathcal{G}} - \mathbf{u}\|^2 \cdot 1_{E_k} \mid \mathbf{u} \right), \end{aligned}$$

so the upper bound on  $\frac{\mathbb{E}(\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 | \mathbf{u})}{\sigma_Z^2 h}$  follows by (5.48), (5.49) and (5.51). Finally, the bound on  $|\mathbb{E}[\|\hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}} - \mathbf{u}\|^2 | \mathbf{u}] - \mathbb{E}[\|\bar{\mathbf{u}}^{\text{sm}} - \mathbf{u}\|^2 | \mathbf{u}]|$  follows directly from Lemma 5.14.  $\square$

### 5.2.2 Comparison of Push-Forward MAP and Posterior Mean for the Filter

The main idea of proof is similar to the proof of Theorem 2.3. We are going to use the following Lemmas (variants of Lemmas 5.10–5.14).

**Lemma 5.15** (A bound on  $\|\bar{\mathbf{u}}^{\text{fi}} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|$ ) *There are some finite constants  $D'_5(\mathbf{u}, T)$  and  $D'_6(\mathbf{u}, T)$  such that for any  $0 < \varepsilon \leq 1$ , for  $\sigma_Z \sqrt{h} \leq \frac{1}{2} \cdot D_{\text{TV}}(\mathbf{u}, T, \varepsilon)^{-1}$ , we have*

$$\begin{aligned} \mathbb{P} \left( \frac{c(\mathbf{u}, T)}{2h} \mathbf{I}_d < \mathbf{A}_k < \frac{C\|\mathbf{A}\|}{h} \mathbf{I}_d \text{ and } \mathbf{u}^{\mathcal{G}} \in \mathcal{B}_R \text{ and} \right. \\ \left. \|\bar{\mathbf{u}}^{\text{fi}} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| \leq \left( D'_5(\mathbf{u}, T) + D'_6(\mathbf{u}, T) \left( \log \left( \frac{1}{\varepsilon} \right) \right)^2 \right) \sigma_Z^2 h \mid \mathbf{u} \right) \geq 1 - \varepsilon. \end{aligned}$$

*Proof* This is a direct consequence of the Wasserstein distance bound of Theorem 2.2.  $\square$

**Lemma 5.16** (A bound on  $\|\hat{\mathbf{u}}^{\text{fi}} - \Psi_T(\mathbf{u}^{\mathcal{G}})\|$ ) *There are finite constants  $S_{\text{MAP}}^{(1')} > 0, S_{\text{MAP}}^{(2')}, D'_7(\mathbf{u}, T)$  and  $D'_8(\mathbf{u}, T)$  such that for any  $0 < \varepsilon \leq 1$ , for*

$$\sigma_Z \sqrt{h} < \left( S_{\text{MAP}}^{(1')} + S_{\text{MAP}}^{(2')} \left( \log \left( \frac{1}{\varepsilon} \right) \right)^{1/2} \right)^{-1},$$

we have

$$\mathbb{P}\left(\mathbf{A}_k \succ \mathbf{0} \text{ and } \mathbf{u}^{\mathcal{G}} \in \mathcal{B}_R \text{ and } \|\hat{\mathbf{u}}^{\text{fi}} - \Psi_T(\mathbf{u}^{\mathcal{G}})\| \leq \left(D'_7(\mathbf{u}, T) + D'_8(\mathbf{u}, T) \left(\log\left(\frac{1}{\varepsilon}\right)\right)^{\frac{3}{2}}\right) \sigma_Z^2 h \mid \mathbf{u}\right) \geq 1 - \varepsilon. \tag{5.54}$$

*Proof* The result follows by Lemma 5.12, Theorem 2.2, and (1.6). □

**Lemma 5.17** (A lower bound on  $\mathbb{E}(\|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u})$ ) *There are positive constants  $D'_9(\mathbf{u}, T)$  and  $D'_{10}(\mathbf{u}, T)$  such that for  $\sigma_Z \sqrt{h} \leq D'_{10}(\mathbf{u}, T)$ , we have*

$$\mathbb{E}\left(\|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u}\right) \geq D'_9(\mathbf{u}, T) \cdot \sigma_Z^2 h. \tag{5.55}$$

*Proof* The proof is similar to the proof of Lemma 5.13. By applying Lemma 5.15 for  $\varepsilon = 0.1$ , we obtain that for  $\sigma_Z \sqrt{h} \leq \frac{1}{2} \cdot D_{\text{TV}}(\mathbf{u}, T, 0.1)^{-1}$ , we have

$$\mathbb{P}\left(\lambda_{\min}(\mathbf{A}_k) > \frac{c(\mathbf{u}, T)}{2h} \text{ and } \|\mathbf{A}_k\| < \frac{C\|\mathbf{A}\|}{h} \text{ and } \|\Psi_{-T}(\bar{\mathbf{u}}^{\text{fi}}) - \mathbf{u}^{\mathcal{G}}\| \leq \left(D'_5(\mathbf{u}, T) + D'_6(\mathbf{u}, T) (\log(10))^2\right) \sigma_Z^2 h \mid \mathbf{u}\right) \geq 0.9. \tag{5.56}$$

If this event happens, then by (1.6), we have

$$\begin{aligned} \|\Psi_T(\mathbf{u}^{\mathcal{G}}) - \mathbf{u}(T)\| &\geq \exp(-GT) \|\mathbf{u}^{\mathcal{G}} - \mathbf{u}\| \geq \exp(-GT) \|\mathbf{A}_k^{-1} \mathbf{B}_k\| \\ &\geq \exp(-GT) \frac{h}{C\|\mathbf{A}\|} \cdot \|\mathbf{B}_k\|. \end{aligned}$$

The rest of the argument is the same as in the proof of Lemma 5.13, so it is omitted. □

**Lemma 5.18** (A bound on  $\left|\mathbb{E}(\|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}\|^2 \mid \mathbf{u}) - \mathbb{E}(\|\hat{\mathbf{u}}^{\text{fi}} - \mathbf{u}\|^2 \mid \mathbf{u})\right|$ ) *There are some finite constants  $D'_{11}(\mathbf{u}, T)$  and  $D'_{12}(\mathbf{u}, T) > 0$  such that for  $\sigma_Z \sqrt{h} \leq D'_{12}(\mathbf{u}, T)$ , we have*

$$\left|\mathbb{E}\left(\|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u}\right) - \mathbb{E}\left(\|\hat{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u}\right)\right| \leq D'_{11}(\mathbf{u}, T) \cdot (\sigma_Z^2 h)^{\frac{3}{2}}. \tag{5.57}$$

*Proof* We define the event  $E_k$  as in (5.46). Under this event,  $\mathbf{A}_k \succ \mathbf{0}$  and  $\|\mathbf{u}^{\mathcal{G}}\| < R$ , so  $\mu_{\mathcal{G}}^{\text{fi}}(\cdot \mid \mathbf{Y}_{0:k})$  is defined according to (2.11). The proof of the claim of the lemma follows the same lines as the proof of Lemma 5.14. In particular, we obtain from (5.49) and (1.6) that

$$\mathbb{E} \left( 1_{E_k} \|\Psi_T(\mathbf{u}^G) - \mathbf{u}(T)\|^2 \mid \mathbf{u} \right) \leq \frac{4\bar{T}(\mathbf{u})\widehat{M}_1(T)^2 d_o \exp(GT)}{c(\mathbf{u}, T)^2} \cdot \sigma_Z^2 h. \tag{5.58}$$

Based on Lemma 5.15, we obtain that for  $\sigma_Z\sqrt{h} < \bar{S}'(\mathbf{u}, T)$ ,

$$\mathbb{E} \left( 1_{E_k} \|\bar{\mathbf{u}}^{\text{fi}} - \Psi_T(\mathbf{u}^G)\|^2 \mid \mathbf{u} \right) \leq \bar{C}'(\mathbf{u}, T)(\sigma_Z^2 h)^2, \tag{5.59}$$

for some constants  $\bar{S}'(\mathbf{u}, T) > 0, \bar{C}'(\mathbf{u}, T) < \infty$ , that are independent of  $\sigma_Z$  and  $h$ . We omit the details.  $\square$

*Proof of Theorem 2.4* The lower bound on  $\frac{\mathbb{E}(\|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u})}{\sigma_Z^2 h}$  follows by Lemma 5.17. Similarly to the case of the smoother, we have

$$\begin{aligned} \mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u} \right) &\leq 4R^2 \mathbb{P} \left( 1_{E_k^c} \mid \mathbf{u} \right) + 2\mathbb{E} \left( \|\bar{\mathbf{u}}^{\text{fi}} - \Psi_T(\mathbf{u}^G)\|^2 \cdot 1_{E_k} \mid \mathbf{u} \right) \\ &\quad + 2\mathbb{E} \left( \|\Psi_T(\mathbf{u}^G) - \mathbf{u}(T)\|^2 \cdot 1_{E_k} \mid \mathbf{u} \right), \end{aligned}$$

which can be further bounded by (5.48), (5.58) and (5.59) to yield the upper bound on  $\frac{\mathbb{E}(\|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u})}{\sigma_Z^2 h}$ . Finally, the bound on  $\left| \mathbb{E} \left[ \|\hat{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u} \right] - \mathbb{E} \left[ \|\bar{\mathbf{u}}^{\text{fi}} - \mathbf{u}(T)\|^2 \mid \mathbf{u} \right] \right|$  follows directly from Lemma 5.18.  $\square$

### 5.3 Convergence of Newton’s Method to the MAP

In this section, we will prove Theorem 2.5. The following proposition shows a classical bound on the convergence of Newton’s method. (This is a reformulation of Theorem 5.3 of Bubeck [9] to our setting.) For  $\mathbf{v} \in \mathbb{R}^d, r > 0$ , we denote the ball of radius  $r$  centred at  $\mathbf{v}$  by  $B(\mathbf{v}, r) := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{v} - \mathbf{x}\| \leq r\}$

**Proposition 5.1** *Suppose that  $\Omega \subset \mathbb{R}^d$  is an open set, and  $g : \Omega \rightarrow \mathbb{R}$  is a 3 times continuously differentiable function satisfying that*

1.  *$g$  has a local minimum at a point  $\mathbf{x}^* \in \Omega$ ,*
2. *there exists a radius  $r^* > 0$  and constants  $C_H > 0, L_H < \infty$  such that  $B(\mathbf{x}^*, r^*) \subset \Omega, \nabla^2 g(\mathbf{x}) \succeq C_H \cdot \mathbf{I}_d$  for every  $\mathbf{x} \in B(\mathbf{x}^*, r^*)$ , and  $\nabla^2 g(\mathbf{x})$  is  $L_H$ -Lipschitz on  $B(\mathbf{x}^*, r^*)$ .*

*Suppose that the starting point  $\mathbf{x}_0 \in \Omega$  satisfies that  $\|\mathbf{x}_0 - \mathbf{x}^*\| < \min\left(r^*, 2\frac{C_H}{L_H}\right)$ . Then the iterates of Newton’s method defined recursively for every  $i \in \mathbb{N}$  as*

$$\mathbf{x}_{i+1} := \mathbf{x}_i - (\nabla^2 g(\mathbf{x}_i))^{-1} \cdot \nabla g(\mathbf{x}_i)$$

*always stay in  $B(\mathbf{x}^*, r^*)$  (thus they are well defined) and satisfy that*

$$\|\mathbf{x}_i - \mathbf{x}^*\| \leq \frac{2C_H}{L_H} \cdot \left( \frac{L_H}{2C_H} \|\mathbf{x}_0 - \mathbf{x}^*\| \right)^{2^i} \text{ for every } i \in \mathbb{N}. \tag{5.60}$$

*Proof* We will show that  $\mathbf{x}_i \in B(\mathbf{x}^*, r^*)$  by induction in  $i$ . This is true for  $i = 0$ . Assuming that  $\mathbf{x}_i \in B(\mathbf{x}^*, r^*)$ , we can write the gradient  $\nabla g(\mathbf{x}_i)$  as

$$\begin{aligned} \nabla g(\mathbf{x}_i) &= \int_{t=0}^1 \nabla^2 g(\mathbf{x}^* + t(\mathbf{x}_i - \mathbf{x}^*)) \cdot (\mathbf{x}_i - \mathbf{x}^*) dt, \text{ therefore} \\ \mathbf{x}_{i+1} - \mathbf{x}^* &= \mathbf{x}_i - \mathbf{x}^* - (\nabla^2 g(\mathbf{x}_i))^{-1} \cdot \nabla g(\mathbf{x}_i) \\ &= \mathbf{x}_i - \mathbf{x}^* - (\nabla^2 g(\mathbf{x}_i))^{-1} \cdot \int_{t=0}^1 \nabla^2 g(\mathbf{x}^* + t(\mathbf{x}_i - \mathbf{x}^*)) \cdot (\mathbf{x}_i - \mathbf{x}^*) dt \\ &= (\nabla^2 g(\mathbf{x}_i))^{-1} \int_{t=0}^1 \left[ \nabla^2 g(\mathbf{x}_i) - \nabla^2 g(\mathbf{x}^* + t(\mathbf{x}_i - \mathbf{x}^*)) \right] \cdot (\mathbf{x}_i - \mathbf{x}^*) dt. \end{aligned}$$

By the  $L_H$ -Lipschitz property of  $\nabla^2 g(\mathbf{x})$  on  $B(\mathbf{x}^*, r^*)$ , we have

$$\int_{t=0}^1 \left\| \nabla^2 g(\mathbf{x}_i) - \nabla^2 g(\mathbf{x}^* + t(\mathbf{x}_i - \mathbf{x}^*)) \right\| dt \leq \frac{L_H}{2} \|\mathbf{x}_i - \mathbf{x}^*\|.$$

By combining this with the fact that  $\|(\nabla^2 g(\mathbf{x}_i))^{-1}\| \leq \frac{1}{C_H}$ , we obtain that  $\|\mathbf{x}_{i+1} - \mathbf{x}^*\| \leq \frac{L_H}{2C_H} \|\mathbf{x}_i - \mathbf{x}^*\|^2$  for every  $i \in \mathbb{N}$ , and by rearrangement, it follows that

$$\log \left( \frac{L_H}{2C_H} \|\mathbf{x}_{i+1} - \mathbf{x}^*\| \right) \leq 2 \log \left( \frac{L_H}{2C_H} \|\mathbf{x}_i - \mathbf{x}^*\| \right)$$

for every  $i \in \mathbb{N}$ , hence the result. □

The following proposition gives a lower bound on the Hessian near  $\mathbf{u}$ . The proof is included in Section A.1 of ‘‘Appendix’’.

**Proposition 5.2** *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $\mathbf{u}$  and the prior  $q$ . Let*

$$r_H(\mathbf{u}, T) := \min \left( \frac{c(\mathbf{u}, T)}{8\bar{T}(\mathbf{u})\widehat{M}_1(T)\widehat{M}_2(T)}, R - \|\mathbf{u}\| \right).$$

Then for any  $0 < \varepsilon \leq 1$ ,  $\sigma_Z > 0$ ,  $0 < h \leq h_{\max}(\mathbf{u}, T)$ , we have

$$\mathbb{P} \left( \nabla^2 \log \mu^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) \leq \frac{-\frac{3}{4}c(\mathbf{u}, T) + C_6(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h} + C_q^{(2)}\sigma_Z^2 h}{\sigma_Z^2 h} \cdot \mathbf{I}_d \right.$$

*for every  $\mathbf{v} \in B(\mathbf{u}, r_H(\mathbf{u}, T))$ )  $\geq 1 - \varepsilon$ , where*

$$C_6(\mathbf{u}, T, \varepsilon) := 33\widehat{M}_2(T)(R + 1)\sqrt{\bar{T}(\mathbf{u})(2d + 1)d_o} + \sqrt{2\bar{T}(\mathbf{u})\widehat{M}_2(T)d_o \log \left( \frac{1}{\varepsilon} \right)}. \tag{5.61}$$

The following proposition bounds the Lipschitz coefficient of the Hessian. The proof is included in Section A.1 of “Appendix”.

**Proposition 5.3** *Suppose that Assumptions 2.1 and 2.3 hold for the initial point  $\mathbf{u}$  and the prior  $q$ . Then for any  $0 < \varepsilon \leq 1$ ,  $\sigma_Z > 0$ ,  $0 < h \leq h_{\max}(\mathbf{u}, T)$ , we have*

$$\begin{aligned} &\mathbb{P}\left(\|\nabla^3 \log \mu^{\text{sm}}(\mathbf{v}|\mathbf{Y}_{0:k})\| \leq C_q^{(3)}\right. \\ &\quad \left.+ \frac{\bar{T}(\mathbf{u})}{\sigma_Z^2 h} \left(C_7(\mathbf{u}, T) + C_8(\mathbf{u}, T, \varepsilon)\sigma_Z\sqrt{h}\right) \text{ for every } \mathbf{v} \in \mathcal{B}_R \Big| \mathbf{u}\right) \geq 1 - \varepsilon, \text{ where} \\ &C_7(\mathbf{u}, T) := 3\widehat{M}_1(T)\widehat{M}_2(T) + 2\widehat{M}_1(T)\widehat{M}_3(T)R, \text{ and} \\ &C_8(\mathbf{u}, T, \varepsilon) := 44(\widehat{M}_4(T)R + \widehat{M}_3(T))\sqrt{3\bar{T}(\mathbf{u})(d+1)d_o} \\ &\quad + \sqrt{2\bar{T}(\mathbf{u})\widehat{M}_3(T)d_o} \log\left(\frac{1}{\varepsilon}\right). \end{aligned}$$

Now we are ready to prove Theorem 2.5.

*Proof of Theorem 2.5* Let

$$\begin{aligned} S_{\max}^{\text{sm}}(\mathbf{u}, T, \varepsilon) &:= \min\left(\frac{c(\mathbf{u}, T)}{8C_6(\mathbf{u}, T, \frac{\varepsilon}{3})}, \sqrt{\frac{c(\mathbf{u}, T)}{8C_q^{(2)}}}, \frac{C_7(\mathbf{u}, T)}{8C_8(\mathbf{u}, T, \frac{\varepsilon}{3})}, \sqrt{\frac{C_7(\mathbf{u}, T)\bar{T}(\mathbf{u})}{C_q^{(3)}}}, \right. \\ &\quad \left. \frac{c(\mathbf{u}, T) \min\left(r_H(\mathbf{u}, T), \frac{c(\mathbf{u}, T)}{C_7(\mathbf{u}, T)}\right)}{8C_1(\mathbf{u}, T, \frac{\varepsilon}{3})}, \sqrt{\frac{c(\mathbf{u}, T) \min\left(r_H(\mathbf{u}, T), \frac{c(\mathbf{u}, T)}{C_7(\mathbf{u}, T)}\right)}{8C_q^{(1)}}}\right). \end{aligned} \tag{5.62}$$

Then by the assumption that  $\sigma_Z\sqrt{h} \leq S_{\max}^{\text{sm}}(\mathbf{u}, T, \varepsilon)$ , using Propositions 5.2, 5.3 and inequality (5.35) we know that with probability at least  $1 - \varepsilon$ , all three of the following events hold at the same time,

1.  $\nabla^2 g^{\text{sm}}(\mathbf{v}) \succeq C_g \cdot \mathbf{I}_d$  for every  $\mathbf{v} \in B(\mathbf{u}, r_H(\mathbf{u}, T))$  for  $C_g := \frac{1}{2} \frac{c(\mathbf{u}, T)}{\sigma_Z^2 h}$ ,
2.  $\nabla^2 g^{\text{sm}}$  is  $L$ -Lipschitz in  $\mathcal{B}_R$  for  $L := \frac{2C_7(\mathbf{u}, T)}{\sigma_Z^2 h}$ ,
3.  $\|\mathbf{u} - \hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}\| \leq \min\left(\frac{r_H(\mathbf{u}, T)}{4}, \frac{1}{4} \frac{c(\mathbf{u}, T)}{C_7(\mathbf{u}, T)}\right)$ .

If these events hold, then the conditions of Proposition 5.1 are satisfied for the function  $g$  with  $\mathbf{x}^* = \hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}$ ,  $r^* := \frac{3}{4}r_H(\mathbf{u}, T)$ ,  $C_H := \frac{1}{2} \frac{c(\mathbf{u}, T)}{\sigma_Z^2 h}$  and  $L_H := \frac{2C_7(\mathbf{u}, T)}{\sigma_Z^2 h}$ . Therefore, (5.60) holds if  $\|\mathbf{x}_0 - \hat{\mathbf{u}}_{\text{MAP}}^{\text{sm}}\| \leq \min\left(\frac{3}{4}r_H(\mathbf{u}, T), \frac{1}{2} \frac{c(\mathbf{u}, T)}{C_7(\mathbf{u}, T)}\right)$ . By the triangle inequality, and the third event above, this is satisfied if  $\|\mathbf{x}_0 - \mathbf{u}\| \leq \min\left(\frac{1}{2}r_H(\mathbf{u}, T), \frac{1}{4} \frac{c(\mathbf{u}, T)}{C_7(\mathbf{u}, T)}\right)$ .

Thus, the claim of the theorem follows for

$$D_{\max}^{\text{sm}}(\mathbf{u}, T) := \min\left(\frac{1}{2}r_H(\mathbf{u}, T), \frac{1}{4} \frac{c(\mathbf{u}, T)}{C_7(\mathbf{u}, T)}\right), \text{ and } N^{\text{sm}}(\mathbf{u}, T) := \frac{c(\mathbf{u}, T)}{2C_7(\mathbf{u}, T)}. \tag{5.63}$$

□

### 5.4 Initial Estimator

In this section, we will prove our results about the initial estimator that we have proposed in Sect. 2.3.

*Proof of Proposition 2.2* By Taylor series expansion of  $\hat{\Phi}^{(l|j_{\max})}(\mathbf{Y}_{0:\hat{k}})$  with remainder term of order  $j_{\max} + 1$ , we obtain that

$$\begin{aligned} \mathbb{E}\left(\hat{\Phi}^{(l|j_{\max})}(\mathbf{Y}_{0:\hat{k}}) \mid \mathbf{u}\right) - \mathbf{H} \frac{d^l \mathbf{u}}{dt^l} &= \sum_{i=0}^{\hat{k}} c_i^{(l|j_{\max}|\hat{k})} \mathbf{H} \mathbf{u}(ih) - \mathbf{H} \frac{d^l \mathbf{u}}{dt^l} \\ &= \left(\sum_{j=0}^{j_{\max}} \frac{(\hat{k}h)^j}{j!} \mathbf{H} \frac{d^j \mathbf{u}}{dt^j} \langle \mathbf{e}^{(l|j_{\max}|\hat{k})}, \mathbf{v}^{(j|\hat{k})} \rangle\right) - \mathbf{H} \frac{d^l \mathbf{u}}{dt^l} + R_{l, j_{\max}+1}, \end{aligned}$$

where by (1.14) the remainder term  $R_{l, j_{\max}+1}$  can be bounded using the Cauchy–Schwarz inequality as

$$\begin{aligned} \|R_{l, j_{\max}+1}\| &\leq C_0 \|\mathbf{H}\| \sum_{i=0}^{\hat{k}} \left| c_i^{(l|j_{\max}|\hat{k})} \right| \frac{(ih)^{j_{\max}+1}}{(j_{\max} + 1)!} \cdot (j_{\max} + 1)! (C_{\text{der}})^{j_{\max}+1} \\ &\leq C_0 \|\mathbf{H}\| (C_{\text{der}} \hat{k}h)^{j_{\max}+1} \left(\sum_{i=0}^{\hat{k}} \left(\frac{i}{\hat{k}}\right)^{2(j_{\max}+1)}\right)^{1/2} \cdot \|\mathbf{e}^{(l|j_{\max}|\hat{k})}\| \\ &\leq C_0 \|\mathbf{H}\| (C_{\text{der}} \hat{k}h)^{j_{\max}+1} \left(1 + \hat{k} \int_{x=0}^1 x^{2j_{\max}+2} dx\right)^{1/2} \cdot \|\mathbf{e}^{(l|j_{\max}|\hat{k})}\| \\ &\leq C_0 \|\mathbf{H}\| (C_{\text{der}} \hat{k}h)^{j_{\max}+1} \left(1 + \frac{\hat{k}}{2j_{\max} + 3}\right)^{1/2} \cdot \|\mathbf{e}^{(l|j_{\max}|\hat{k})}\|. \end{aligned}$$

Due to the particular choice of the coefficients of  $\mathbf{e}^{(l|j_{\max}|\hat{k})}$ , we can see that all the terms up to order  $j_{\max}$  disappear, and we are left with the remainder term that can be

bounded as

$$\begin{aligned} & \left\| \mathbb{E} \left( \hat{\Phi}^{(l|j_{\max})}(\mathbf{Y}_{0:\hat{k}}) \mid \mathbf{u} \right) - \mathbf{H} \frac{d^l \mathbf{u}}{dt^l} \right\| \\ & \leq C_0 \|\mathbf{H}\| \sqrt{\frac{\hat{k}}{j_{\max} + 3/2}} (C_{\text{der}} \hat{k} h)^{j_{\max} + 1} \left\| \mathbf{c}^{(l|j_{\max}|\hat{k})} \right\|, \end{aligned}$$

using the assumption that  $\hat{k} \geq 2j_{\max} + 3$ . The concentration bound now follows directly from this bias bound, (5.20) and the fact that the estimator  $\hat{\Phi}^{(l|j_{\max})}(\mathbf{Y}_{0:\hat{k}})$  has Gaussian distribution with covariance matrix  $\left\| \mathbf{c}^{(l|j_{\max}|\hat{k})} \right\| \cdot \sigma_Z \cdot \mathbf{I}_{d_o}$ .  $\square$

*Proof of Lemma 2.6* Let  $\mathcal{P}$  denote the space of finite degree polynomials with real coefficients on  $[0, 1]$ . For  $a, b \in \mathcal{P}$ , we let  $\langle a, b \rangle_{\mathcal{P}} = \int_{x=0}^1 a(x)b(x)dx$ . Then the elements of the matrix  $\mathbf{K}^{j_{\max}}$  can be written as

$$K_{i,j}^{j_{\max}} = \frac{1}{i + j - 1} = \left\langle x^{i-1}, x^{j-1} \right\rangle_{\mathcal{P}}.$$

If  $\mathbf{K}^{j_{\max}}$  would not be invertible, then its rows would be linearly dependent, that is, there would exist a nonzero vector  $\alpha \in \mathbb{R}^{j_{\max} + 1}$  such that  $\sum_{i=1}^{j_{\max} + 1} K_{i,j}^{j_{\max}} = 0$  for every  $1 \leq j \leq j_{\max} + 1$ . This would imply that  $\left\langle \sum_{i=1}^{j_{\max} + 1} \alpha_i x^{i-1}, x^{j-1} \right\rangle_{\mathcal{P}} = 0$  for every  $1 \leq j \leq j_{\max} + 1$ , and thus,

$$\int_{x=0}^1 \left( \sum_{i=1}^{j_{\max} + 1} \alpha_i x^{i-1} \right)^2 dx = \left\langle \sum_{i=1}^{j_{\max} + 1} \alpha_i x^{i-1}, \sum_{i=1}^{j_{\max} + 1} \alpha_i x^{i-1} \right\rangle_{\mathcal{P}} = 0.$$

However, this is not possible, since by the fundamental theorem of algebra, the polynomial  $\sum_{i=1}^{j_{\max} + 1} \alpha_i x^{i-1}$  can have at most  $j_{\max}$  roots, so it cannot be zero Lebesgue almost everywhere in  $[0, 1]$ . Therefore,  $\mathbf{K}^{j_{\max}}$  is invertible. The result now follows from the continuity of the matrix inverse and the fact that for any  $1 \leq i, j \leq j_{\max} + 1$ ,

$$\begin{aligned} \lim_{\hat{k} \rightarrow \infty} \left( \frac{\mathbf{M}^{(j_{\max}|\hat{k})} (\mathbf{M}^{(j_{\max}|\hat{k})})'}{\hat{k}} \right)_{i,j} &= \lim_{\hat{k} \rightarrow \infty} \frac{1}{\hat{k}} \sum_{m=0}^{\hat{k}} \binom{m}{\hat{k}}^{i-1} \binom{m}{\hat{k}}^{j-1} \\ &= \int_{x=0}^1 x^{i+j-2} dx = \frac{1}{i + j - 1}. \end{aligned}$$

$\square$

*Proof of Theorem 2.7* Let

$$B_{\mathbf{M}}^{(l|j_{\max})} := \sup_{k \in \mathbb{N}, k \geq 2j_{\max} + 3} C_{\mathbf{M}}^{(l|j_{\max}|\hat{k})}.$$

Based on Lemma 2.6, this is finite. With the choice  $j_{\max} = l$ , by (2.39), we obtain that

$$\hat{k}_{\min}(l, l) = \frac{1}{h} \cdot (\sigma_Z \sqrt{h})^{\frac{1}{l+3/2}} \cdot C_{\text{der}}^{-\frac{l+1}{l+3/2}} \cdot \left( \frac{\sqrt{d_o \log(d_o + 1)(l + 3/2)(l + 1/2)}}{C_0 \|\mathbf{H}\|} \right)^{\frac{1}{l+3/2}}.$$

By choosing  $s_{\max}^{(l)}$  sufficiently small, we can ensure that for  $\sigma_Z \sqrt{h} \leq s_{\max}^{(l)}$ ,  $h \hat{k}_{\min}(l, l) \leq T$ , and thus by definition (2.40), we have

$$\left| \hat{k}_{\text{opt}}(l, l) - \max \left( 2l + 3, \frac{1}{h} \cdot (\sigma_Z \sqrt{h})^{\frac{1}{l+3/2}} \cdot C_{\text{der}}^{-\frac{l+1}{l+3/2}} \cdot \left( \frac{\sqrt{d_o \log(d_o + 1)(l + 3/2)(l + 1/2)}}{C_0 \|\mathbf{H}\|} \right)^{\frac{1}{l+3/2}} \right) \right| < 1.$$

By substituting this into (2.37), and applying some algebra, we obtain that

$$\begin{aligned} g(l, l, \hat{k}_{\text{opt}}(l, l)) &= \frac{C_0 \|\mathbf{H}\| C_{\text{der}}^{l+1}}{\sqrt{l + 3/2}} \cdot (\hat{k}_{\text{opt}}(l, l)h) + (\hat{k}_{\text{opt}}(l, l)h)^{-l-1/2} \sigma_Z \sqrt{h} \sqrt{2d_o \log(d_o + 1)} \\ &\leq \frac{C_0 \|\mathbf{H}\| C_{\text{der}}^{l+1}}{\sqrt{l + 3/2}} \\ &\quad \cdot \left[ (2l + 4)h + (\sigma_Z \sqrt{h})^{\frac{1}{l+3/2}} \cdot C_{\text{der}}^{-\frac{l+1}{l+3/2}} \cdot \left( \frac{\sqrt{d_o \log(d_o + 1)(l + 3/2)(l + 1/2)}}{C_0 \|\mathbf{H}\|} \right)^{\frac{1}{l+3/2}} \right] \\ &\quad + 2C_{\text{der}}^{\frac{(l+1)(l+1/2)}{l+3/2}} \cdot \left( \frac{\sqrt{d_o \log(d_o + 1)(l + 3/2)(l + 1/2)}}{C_0 \|\mathbf{H}\|} \right)^{\frac{-l-1/2}{l+3/2}} (\sigma_Z \sqrt{h})^{\frac{1}{l+3/2}} \\ &\quad \cdot \sqrt{2d_o \log(d_o + 1)}, \end{aligned}$$

and the claim of the theorem now follows by substituting this into Proposition 2.2.  $\square$

## 6 Conclusion

In this paper, we have proved the consistency and asymptotic efficiency results for MAP estimators for smoothing and filtering a class of partially observed nonlinear dynamical systems. We have also shown that the smoothing and filtering distributions are approximately Gaussian in the low observation noise/high observation frequency regime when the length of the assimilation window is fixed. These results contribute to the statistical understanding of the widely used 4D-Var data assimilation method [29, 44]. The precise size of the observation noise  $\sigma_Z$  and assimilation step  $h$  under which the Gaussian approximation approximately holds, and the MAP estimator is close to the posterior mean, is strongly dependent on the model parameters and the size of the assimilation window. However, we have found in simulations shown in Fig. 2

that even for relatively large values of  $\sigma_Z$  and  $h$ , for large dimensions, and not very short assimilation windows, these approximations seem to be working reasonably well. Besides theoretical importance, the Gaussian approximation of the smoother can be also used to construct the prior (background) distributions for the subsequent intervals in a flow-dependent way, as we have shown in Paulin et al. [37] for the nonlinear shallow-water equations, even for realistic values of  $\sigma_Z$  and  $h$ . These flow-dependent prior distributions can considerably improve filtering accuracy. Going beyond the approximately Gaussian case (for example when  $\sigma_Z$ ,  $h$  and  $T$  are large, or the system is highly nonlinear) in a computationally efficient way is a challenging problem for future research (see Bocquet et al. [7] for some examples where this situation arises).

**Acknowledgements** DP and AJ were supported by an AcRF tier 2 Grant: R-155-000-161-112. AJ is affiliated with the Risk Management Institute, the Center for Quantitative Finance and the OR and Analytics cluster at NUS. DC was partially supported by the EPSRC Grant: EP/N023781/1. We thank the anonymous referees for their valuable comments and suggestions that have substantially improved the paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## A: Appendix

### A.1: Proof of Preliminary Results

*Proof of Lemma 1.1* To prove (1.14), it suffices to first verify (1.14) and for  $i = 0$  and  $i = 1$ , and then use induction and the recursion formula (1.13) for  $i \geq 2$ . For (1.15), by taking the  $k$ th derivative of (1.13), we obtain that

$$\begin{aligned}
 J_v^k(D^i v) &= -A \cdot J_v^k(D^{i-1} v) \\
 &\quad - \sum_{j=0}^{i-1} \binom{i-1}{j} \sum_{l=0}^k \binom{k}{l} B \left( J_v^l(D^j v), J_v^{k-l}(D^{i-1-j} v) \right). \quad (\text{A.1})
 \end{aligned}$$

For  $k = 1$ , (1.15) can be verified by first checking it for  $i = 0$  and  $i = 1$ , and then using mathematical induction and (1.13) for  $i \geq 2$ . Suppose that (1.15) holds for  $k = 1 \dots, k' - 1$ , then by mathematical induction and (A.1), we only need to show that

$$\begin{aligned} (C_{\mathbf{J}}^{(k')})^i \cdot i! &\geq \|A\| (C_{\mathbf{J}}^{(k')})^{i-1} (i-1)! \\ &+ \sum_{j=0}^{i-1} \binom{i-1}{j} \sum_{l=1}^{k'-1} \binom{k'}{l} j!(i-1-j)! \|B\| (C_{\mathbf{J}}^{(l)})^j (C_{\mathbf{J}}^{(k'-l)})^{i-1-j} \\ &+ \sum_{j=0}^{i-1} \binom{i-1}{j} j!(i-1-j)! \cdot 2\|B\| C_0 (C_{\text{der}})^j (C_{\mathbf{J}}^{(k')})^{i-1-j}, \end{aligned}$$

which is straightforward to check since  $\sum_{l=0}^{k'} \binom{k'}{l} = 2^{k'}$  and  $\binom{i-1}{j} j!(i-1-j)! = (i-1)!$ . □

*Proof of Lemma 1.2* Notice that the derivative  $\frac{d}{dt}(\mathbf{J}^k \Psi_t(\mathbf{v}))$  can be rewritten by exchanging the order of derivation (which can be justified by the Taylor series expansion and the bounds (1.15)) as

$$\begin{aligned} \frac{d}{dt}(\mathbf{J}^k \Psi_t(\mathbf{v})) &= \mathbf{J}_{\mathbf{v}}^k (-A \Psi_t(\mathbf{v}) - \mathbf{B}(\Psi_t(\mathbf{v}), \Psi_t(\mathbf{v})) + \mathbf{f}) \\ &= -A \mathbf{J}^k \Psi_t(\mathbf{v}) - \sum_{l=0}^k \binom{k}{l} \mathbf{B}(\mathbf{J}^l \Psi_t(\mathbf{v}), \mathbf{J}^{k-l} \Psi_t(\mathbf{v})). \end{aligned} \tag{A.2}$$

For  $k = 1$ , the above equation implies that

$$\frac{d}{dt} \|\mathbf{J} \Psi_t(\mathbf{v})\| \leq (\|A\| + 2\|B\|R) \|\mathbf{J} \Psi_t(\mathbf{v})\|;$$

thus, using the fact that  $\|\mathbf{J} \Psi_0(\mathbf{v})\| = \|\mathbf{J}_{\mathbf{v}}(\mathbf{v})\| = 1$ , by Grönwall’s lemma, we have

$$\|\mathbf{J} \Psi_t(\mathbf{v})\| \leq \exp((\|A\| + 2\|B\|R)t) \text{ for any } t \geq 0.$$

Now we are going to show that

$$\|\mathbf{J}^k \Psi_t(\mathbf{v})\| \leq \exp\left(D_{\mathbf{J}}^{(k)} t\right) \text{ for any } \mathbf{v} \in \mathcal{B}_R, t \geq 0, k \in \mathbb{Z}_+. \tag{A.3}$$

Indeed, this was shown for  $k = 1$  above, and for  $k \geq 2$ , from (A.2), it follows that

$$\begin{aligned} \frac{d}{dt} \|\mathbf{J}^k \Psi_t(\mathbf{v})\| &\leq (\|A\| + 2\|B\|R) \|\mathbf{J}^k \Psi_t(\mathbf{v})\| + \|B\| \sum_{l=1}^{k-1} \binom{k}{l} \|\mathbf{J}^l \Psi_t(\mathbf{v})\| \|\mathbf{J}^{k-l} \Psi_t(\mathbf{v})\|; \end{aligned}$$

thus, (A.3) can be proven by mathematical induction and Grönwall’s lemma. This implies in particular our first claim, (1.21). Our second claim (1.22) follows by the fact that  $\Phi_t(\mathbf{v}) = \mathbf{H} \Psi_t(\mathbf{v})$  is a linear transformation. □

*Proof of Lemma 4.3* Let  $(\mathbf{J}\Phi_{t_i}(\mathbf{u}))_{\cdot,j}$  denote the  $j$ th column of the Jacobian, and let  $\tilde{\mathbf{Z}}_i^j := \mathbf{Z}_i^j/\sigma_Z$  (which is a standard normal random variable). Then we can write  $\mathbf{B}_k$  as

$$\mathbf{B}_k = \sum_{i=0}^k \sum_{j=1}^{d_o} \sigma_Z \cdot (\mathbf{J}\Phi_{t_i}(\mathbf{u}))_{\cdot,j} \cdot \tilde{\mathbf{Z}}_i^j.$$

This is of the same form as in equation (4.1.2) of Theorem 4.1.1 of Tropp [46]. Since  $\|\mathbf{J}\Phi_{t_i}(\mathbf{u})\| \leq \widehat{M}_1(T)$ , one can see that we also have  $\|(\mathbf{J}\Phi_{t_i}(\mathbf{u}))_{\cdot,j}\| \leq \widehat{M}_1(T)$ , and thus the variance statistics  $v(\mathbf{Z})$  of Theorem 4.1.1 can be bounded as  $v(\mathbf{Z}) \leq \sigma_Z^2 \widehat{M}_1(T)^2(k+1)d_o$ . The result now follows from equation (4.1.6) of Tropp [46] (with  $d_1 = 1$  and  $d_2 = d$ ).  $\square$

*Proof of Lemma 4.4* First, note that from Assumption 2.1 and looking at the Taylor expansion near  $\mathbf{u}$  it follows that the first term in the definition of  $\mathbf{A}_k$  satisfies that

$$\lambda_{\min} \left( \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \mathbf{J}\Phi_{t_i}(\mathbf{u}) \right) \geq \frac{c(\mathbf{u}, T)}{h}.$$

We can rewrite the second term in definition (2.7) as

$$\mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i] = \sum_{j=1}^{d_o} H\Phi_{t_i}^j(\mathbf{u}) \cdot \sigma_Z \cdot \tilde{\mathbf{Z}}_i^j, \tag{A.4}$$

where  $H\Phi_{t_i}^j(\mathbf{u})$  denotes the Hessian of the function  $\Phi_{t_i}^j$  at point  $\mathbf{u}$ . This is of the same form as in equation (4.1.2) of Theorem 4.1.1 of Tropp [46]. Since  $\|\mathbf{J}^2\Phi_{t_i}(\mathbf{u})\| \leq \widehat{M}_2(T)$ , one can see that we also have  $\|H\Phi_{t_i}^j(\mathbf{u})\| \leq \widehat{M}_2(T)$ , and thus the variance statistics  $v(\mathbf{Z})$  of Theorem 4.1.1 can be bounded as  $v(\mathbf{Z}) \leq \sigma_Z^2 \widehat{M}_2(T)^2(k+1)d_o$ , and thus for any  $t \geq 0$ , we have

$$\mathbb{P} \left( \left\| \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i] \right\| \geq t \right) \leq 2d \exp \left( -\frac{t^2}{\sigma_Z^2 \widehat{M}_2(T)^2(k+1)d_o} \right). \tag{A.5}$$

By the Bauer–Fike theorem (see [3]), we have

$$\lambda_{\min}(\mathbf{A}_k) \geq \lambda_{\min} \left( \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \mathbf{J}\Phi_{t_i}(\mathbf{u}) \right) - \left\| \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i] \right\|,$$

and

$$\|\mathbf{A}_k\| \leq \widehat{M}_1(T)^2 \cdot \frac{\bar{T}(\mathbf{u})}{h} + \left\| \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \cdot, \mathbf{Z}_i] \right\|,$$

so (4.8) follows by bounds (A.4) and (A.5), and (4.9) follows by rearrangement.  $\square$

*Proof of Proposition 2.1* The proof is quite similar to the proof of Proposition 4.1 of Paulin et al. [38]. With a slight modification of that argument, we will show that for any  $\delta \in [0, \tilde{h})$ ,

$$\sum_{i=0}^j \|\Phi_{i\tilde{h}+\delta}(\mathbf{v}) - \Phi_{i\tilde{h}+\delta}(\mathbf{u})\|^2 \geq c'(\mathbf{u}, \tilde{h})\|\mathbf{v} - \mathbf{u}\|^2, \tag{A.6}$$

for some constant  $c'(\mathbf{u}, \tilde{h}) > 0$  independent of  $\delta$ , that is monotone increasing in  $\tilde{h}$  for  $0 < \tilde{h} \leq \tilde{h}_{\max}$ . This result allows us to decouple the summation in (2.1) into sets of size  $j + 1$  as follows. Let  $h_0 := \min\left(\frac{T}{j+1}, \tilde{h}_{\max}\right)$ , and set  $\tilde{h} := \lfloor h_0/h \rfloor \cdot h$ . Then for  $h \leq h_0$ , we have  $\lfloor h_0/h \rfloor > \frac{h_0}{2h}$  and  $\tilde{h} > h_0/2$ , so using (A.6), we obtain that

$$\begin{aligned} \sum_{i=0}^k \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 &\geq \sum_{l=0}^{\lfloor h_0/h \rfloor - 1} \sum_{i=0}^j \|\Phi_{i\tilde{h}+lh}(\mathbf{v}) - \Phi_{i\tilde{h}+lh}(\mathbf{u})\|^2 \\ &\geq \frac{h_0}{2h} \cdot c'(\mathbf{u}, h_0/2)\|\mathbf{v} - \mathbf{u}\|^2. \end{aligned}$$

Thus, Assumption 2.1 holds with  $h_{\min}(\mathbf{u}, T) := h_0/2$  and  $c(\mathbf{u}, T) := c'(\mathbf{u}, h_0/2) \cdot h_0/2$ .

To complete the proof, we will now show (A.6). Using inequality (1.14), we can see that the Taylor expansion

$$\Phi_t(\mathbf{v}) = \sum_{i=0}^{\infty} \frac{\mathbf{H}\mathbf{D}^i \mathbf{v} \cdot t^i}{i!}$$

is valid for times  $0 \leq t < C_{\text{der}}^{-1}$ . Based on this expansion, assuming that  $i\tilde{h} + \delta < C_{\text{der}}^{-1}$ ,  $\mathbf{H}\mathbf{D}^i \mathbf{v}$  can be approximated by a finite difference formula depending on the values of  $\Phi_{\delta}(\mathbf{v}), \Phi_{\tilde{h}+\delta}(\mathbf{v}), \dots, \Phi_{i\tilde{h}+\delta}$ , with error of  $O(\tilde{h})$ . This finite difference formula will be denoted as

$$\hat{\Phi}^{(i,\delta)}(\mathbf{v}) := \frac{\sum_{l=0}^i a_l^{(i,\delta)} \Phi_{l\tilde{h}+\delta}(\mathbf{v})}{\tilde{h}^i}. \tag{A.7}$$

The coefficients  $a_l^{(i,\delta)}$  are explicitly defined in Fornberg [20], and they only depend on  $i$  and the ratio  $\delta/\tilde{h}$ . Based on the definition of these coefficients on page 700 of Fornberg [20], we can see that

$$\bar{a} := \sup_{\tilde{h} > 0, \delta \in [0, \tilde{h})} \max_{0 \leq i \leq j, 0 \leq l \leq i} \left| a_l^{(i,\delta)} \right| < \infty, \tag{A.8}$$

i.e. they can be bounded by a finite constant independently of  $\delta$ . By Taylor’s expansion of the terms  $\Phi_{l\tilde{h}+\delta}(\mathbf{v})$  around time point 0, for  $l\tilde{h} + \delta < C_{\text{der}}^{-1}$ , we have

$$\Phi_{l\tilde{h}+\delta}(\mathbf{v}) = \sum_{m=0}^{\infty} \mathbf{H}D^m \mathbf{v} \cdot \frac{(l\tilde{h} + \delta)^m}{m!}, \text{ and thus}$$

$$\hat{\Phi}^{(i,\delta)}(\mathbf{v}) = \frac{1}{\tilde{h}^i} \cdot \sum_{l=0}^i a_l^{(i,\delta)} \sum_{m=0}^{\infty} \mathbf{H}D^m \mathbf{v} \cdot \frac{(l\tilde{h} + \delta)^m}{m!} = \frac{1}{\tilde{h}^i} \sum_{m=0}^{\infty} \tilde{h}^m b_m^{(i,\delta)} \mathbf{H}D^m \mathbf{v},$$

with  $b_m^{(i,\delta)} := \frac{1}{m!} \cdot \sum_{l=0}^i a_l^{(i,\delta)} (l + \delta/\tilde{h})^m$ . Due to the particular choice of the constants  $a_l^{(i,\delta)}$ , we have  $b_m^{(i,\delta)} = 0$  for  $0 \leq m < i$  and  $b_m^{(i,\delta)} = 1$  for  $m = i$ . Based on this, we can write the difference between approximation (A.7) and the derivative explicitly as

$$\hat{\Phi}^{(i,\delta)}(\mathbf{v}) - \mathbf{H}D^i \mathbf{v} = \tilde{h} \left( \sum_{m=i+1}^{\infty} \tilde{h}^{m-i-1} \cdot b_m^{(i,\delta)} \cdot \mathbf{H}D^m \mathbf{v} \right).$$

Let us denote  $\tilde{\Phi}^{(i,\delta)}(\mathbf{v}, \tilde{h}) := \sum_{m=i+1}^{\infty} \tilde{h}^{m-i-1} \cdot b_m^{(i,\delta)} \cdot \mathbf{H}D^m \mathbf{v}$ . Using inequality (1.15), and the bound  $|b_m^{(i,\delta)}| \leq \frac{\bar{a} \cdot (i+1)^{m+1}}{m!}$ , we have that for  $0 \leq i \leq j, \tilde{h} \leq \frac{1}{2(j+1)C_J^{(1)}}$ ,

$$\begin{aligned} \|\mathbf{J}_v \tilde{\Phi}^{(i,\delta)}(\mathbf{v}, \tilde{h})\| &\leq \|\mathbf{H}\| \cdot \sum_{m=i+1}^{\infty} \tilde{h}^{m-i-1} |b_m^{(i,\delta)}| \left(C_J^{(1)}\right)^m m! \\ &\leq \|\mathbf{H}\| \cdot \frac{\bar{a}(i+1)}{\tilde{h}^{i+1}} \sum_{m=i+1}^{\infty} \left((i+1)\tilde{h}C_J^{(1)}\right)^m \\ &\leq 2\|\mathbf{H}\|\bar{a}(i+1) \left((i+1)C_J^{(1)}\right)^{i+1}. \end{aligned}$$

Denote  $C_{\text{Lip}} := 2\|\mathbf{H}\|\bar{a} \cdot \max_{0 \leq i \leq j} \left( (i+1) \cdot \left( (i+1)C_J^{(1)} \right)^{i+1} \right)$ , then we know that for every  $0 \leq i \leq j, \tilde{h} < \frac{1}{2(j+1)C_J^{(1)}}$ , the functions  $\tilde{\Phi}^{(i)}(\mathbf{v}, \tilde{h})$  are  $C_{\text{Lip}}$ -Lipschitz in  $\mathbf{v}$  with respect to the  $\|\cdot\|$  norm, and thus for every  $\mathbf{v} \in \mathcal{B}_R$ ,

$$\left\| \left[ \mathbf{H}D^i \mathbf{v} - \mathbf{H}D^i \mathbf{u} \right] - \left[ \hat{\Phi}^{(i,\delta)}(\mathbf{v}) - \hat{\Phi}^{(i,\delta)}(\mathbf{u}) \right] \right\| \leq \tilde{h} C_{\text{Lip}} \|\mathbf{v} - \mathbf{u}\|. \tag{A.9}$$

By Assumption 2.2, and the boundedness of  $\mathcal{B}_R$ , it follows that there is constant  $C_D(\mathbf{u}, T) > 0$  such that for every  $\mathbf{v} \in \mathcal{B}_R$ ,

$$\sum_{i=0}^j \left\| \mathbf{H}D^i \mathbf{v} - \mathbf{H}D^i \mathbf{u} \right\|^2 \geq C_D(\mathbf{u}, T) \|\mathbf{v} - \mathbf{u}\|^2. \tag{A.10}$$

From equations (A.9), (A.10), and the boundedness of  $\mathcal{B}_R$ , it follows that there is a constant  $C_j(\mathbf{u}, T, \tilde{h}) > 0$  that is non-decreasing in  $\tilde{h}$  such that for every  $\mathbf{v} \in \mathcal{B}_R$ ,

$$\sum_{i=0}^j \left\| \hat{\Phi}^{(i,\delta)}(\mathbf{v}) - \hat{\Phi}^{(i,\delta)}(\mathbf{u}) \right\|^2 \geq C_j(\mathbf{u}, T, \tilde{h}) \|\mathbf{v} - \mathbf{u}\|^2. \tag{A.11}$$

By definitions (A.7) and (A.8), it follows that

$$\bar{a} \cdot \max\left(\frac{1}{\tilde{h}^j}, 1\right) \cdot \sum_{i=0}^j \left\| \Phi_{i\tilde{h}+\delta}(\mathbf{v}) - \Phi_{i\tilde{h}+\delta}(\mathbf{u}) \right\|^2 \geq \sum_{i=0}^j \left\| \hat{\Phi}^{(i,\delta)}(\mathbf{v}) - \hat{\Phi}^{(i,\delta)}(\mathbf{u}) \right\|^2, \tag{A.12}$$

and thus (A.6) follows by rearrangement.  $\square$

The following lemma bounds the number of balls of radius  $\delta$  required to cover a  $d$ -dimensional unit ball. It will be used in the proofs of Propositions 4.1, 4.2 and 4.3.

**Lemma A.1** *For any  $d \geq 1$ ,  $0 < \delta \leq 1$ , a  $d$ -dimensional unit ball can be fully covered by the union of  $c(d) \cdot (\frac{1}{\delta})^d$  balls of radius  $\delta$ , where*

$$c(1) := 2, \quad c(2) := 6, \quad \text{and} \\ c(d) := d \log(d) \cdot \left( \frac{1}{2} + \frac{2 \log(\log(d))}{\log(d)} + \frac{5}{\log(d)} \right) \text{ for } d \geq 3.$$

*Proof* For  $d \geq 3$ , this follows from Theorem 1 of Dumer [18]. For  $d = 1$ , the result follows from the fact that  $\frac{1}{\delta} + 1$  intervals suffice, and  $\frac{1}{\delta} + 1 \leq \frac{2}{\delta}$ . For  $d = 2$ , we know that the circles of radius  $\delta$  contain a square of edge length  $\sqrt{2}\delta$ , and in order to cover a square of edge length 2 containing the unit ball, it suffices to use  $\lceil 2/(\sqrt{2}\delta) \rceil^2 \leq \left(\frac{\sqrt{2}}{\delta} + 1\right)^2 < \frac{6}{\delta^2}$  squares of edge length  $\sqrt{2}\delta$ ; thus, the result follows.  $\square$

Our next lemma shows some concentration inequalities that will be used in the proof of our propositions. It is a reformulation of Corollary 13.2 and Theorem 5.8 of Boucheron et al. [8] to our setting.

**Lemma A.2** *For every  $l \in \mathbb{N}$ , define the sets*

$$\mathcal{T}_l := \{(r, s_1, \dots, s_l) \in [0, 2R] \times \mathcal{B}_1^l : \mathbf{u} + r\mathbf{s}_1 \in \mathcal{B}_R\}, \quad \bar{\mathcal{T}}_l := \mathcal{B}_R \times \mathcal{B}_1^l. \tag{A.13}$$

*For any two elements  $(r, s_1, \dots, s_l), (r', s'_1, \dots, s'_l) \in \mathcal{T}_l$ , we define the distance*

$$d_l((r, s_1, \dots, s_l), (r', s'_1, \dots, s'_l)) := \frac{|r - r'|}{2R} + \sum_{i=0}^l \|s_i - s'_i\|. \tag{A.14}$$

Similarly, for any two elements  $(\mathbf{v}, s_1, \dots, s_l), (\mathbf{v}', s'_1, \dots, s'_l) \in \bar{\mathcal{T}}_l$ , we define

$$\bar{d}_l((\mathbf{v}, s_1, \dots, s_l), (\mathbf{v}', s'_1, \dots, s'_l)) := \frac{\|\mathbf{v} - \mathbf{v}'\|}{R} + \sum_{i=0}^l \|s_i - s'_i\|. \tag{A.15}$$

Suppose that  $\mathbf{Z}_0, \dots, \mathbf{Z}_k$  are i.i.d.  $d_o$ -dimensional standard normal random vectors, and  $\varphi_0, \dots, \varphi_k : \mathcal{T}_l \rightarrow \mathbb{R}^{d_o}$  are functions that are  $L$ -Lipschitz with respect to the distance  $d_l$  on  $\mathcal{T}_l$ , and satisfy that  $\|\varphi_i(r, s_1, \dots, s_l)\| \leq M$  for any  $0 \leq i \leq k, (r, s_1, \dots, s_l) \in \mathcal{T}_l$ . (The constants  $M$  and  $L$  can depend on  $l$ .) Then  $W_l := \sup_{(r, s_1, \dots, s_l) \in \mathcal{T}_l} \sum_{i=0}^k \langle \varphi_i(r, s_1, \dots, s_l), \mathbf{Z}_i \rangle$  satisfies that for any  $0 < \varepsilon \leq 1$ ,

$$\mathbb{P}(W_l \geq C^{(l)}(\mathbf{u}, k, \varepsilon)) \leq \varepsilon \text{ for}$$

$$C^{(l)}(\mathbf{u}, k, \varepsilon) := 11(l+1)L\sqrt{(k+1)(ld+1)d_o} + \sqrt{2(k+1)Md_o \log\left(\frac{1}{\varepsilon}\right)}. \tag{A.16}$$

Similarly, if  $\bar{\varphi}_0, \dots, \bar{\varphi}_k : \bar{\mathcal{T}}_l \rightarrow \mathbb{R}^{d_o}$  are  $L$ -Lipschitz with respect to  $\bar{d}_l$ , and satisfy that  $\|\bar{\varphi}_i(\mathbf{v}, s_1, \dots, s_l)\| \leq M$  for any  $0 \leq i \leq k, (\mathbf{v}, s_1, \dots, s_l) \in \bar{\mathcal{T}}_l$ , then the quantity  $\bar{W}_l := \sup_{(\mathbf{v}, s_1, \dots, s_l) \in \bar{\mathcal{T}}_l} \sum_{i=0}^k \langle \bar{\varphi}_i(\mathbf{v}, s_1, \dots, s_l), \mathbf{Z}_i \rangle$  satisfies that for any  $0 < \varepsilon \leq 1$ ,

$$\mathbb{P}(\bar{W}_l \geq \bar{C}^{(l)}(\mathbf{u}, k, \varepsilon)) \leq \varepsilon \text{ for}$$

$$\bar{C}^{(l)}(\mathbf{u}, k, \varepsilon) := 11(l+1)L\sqrt{(k+1)l(d+1)d_o} + \sqrt{2(k+1)Md_o \log\left(\frac{1}{\varepsilon}\right)}. \tag{A.17}$$

*Proof* For  $(r, s_1, \dots, s_l) \in \mathcal{T}_l$ , let us denote

$$W'_l(r, s_1, \dots, s_l) := \left( \sum_{i=0}^k \langle \varphi_i(r, s_1, \dots, s_l), \mathbf{Z}_i \rangle \right) / (L\sqrt{(k+1)d_o}).$$

Then we have  $W'_l(0, \mathbf{0}, \dots, \mathbf{0}) = 0$ , and by the  $L$ -Lipschitz assumption on  $\varphi_i$ , we can see that  $W'_l(r, s_1, \dots, s_l) - W'_l(0, \mathbf{0}, \dots, \mathbf{0})$  is a one-dimensional Gaussian random variable whose variance is bounded by  $(d_l((r, s_1, \dots, s_l), (0, \mathbf{0}, \dots, \mathbf{0})))^2$ . Therefore, its moment-generating function can be bounded as

$$\log \mathbb{E}(e^{\lambda(W'_l(r, s_1, \dots, s_l) - W'_l(0, \mathbf{0}, \dots, \mathbf{0}))}) \leq \frac{\lambda^2 [d_l((r, s_1, \dots, s_l), (0, \mathbf{0}, \dots, \mathbf{0}))]^2}{2}.$$

This means that Dudley’s entropy integral expectation bound (Corollary 13.2 of Boucheron et al. [8]) is applicable here. To apply that result, we first need to upper-bound the packing number  $N(\delta, \mathcal{T}_l)$ , which is the maximum number of points that can be selected in  $\mathcal{T}_l$  such that all of them are further away from each other than  $\delta$  in

$d_l$  distance. It is easy to show that  $N(\delta, \mathcal{T}_l) \leq N'(\delta/2, \mathcal{T}_l)$ , where  $N'(\delta/2, \mathcal{T}_l)$  is the number of spheres of radius  $\frac{\delta}{2}$  in  $d_l$  distance needed to cover  $\mathcal{T}_l$ .

Since  $\mathcal{T}_l \subset [0, 2R] \times \mathcal{B}_1^l$ , it follows that  $N'(\delta/2, \mathcal{T}_l) \leq N'(\delta/2, [0, 2R] \times \mathcal{B}_1^l)$ . Moreover, due to the product nature of the space  $[0, 2R] \times \mathcal{B}_1^l$  and the definition of the distance  $d_l$ , if we first cover  $[0, 2R]$  with intervals of length  $R\delta/(l + 1)$  (in one-dimensional Euclidean distance), and then cover each sphere  $\mathcal{B}_1$  in  $\mathcal{B}_1^l$  with spheres of radius  $\delta/(2l + 2)$  (in  $d$ -dimensional Euclidean distance), then the product of any such interval and spheres will be contained in a sphere of  $d_l$ -radius less than or equal to  $\delta/2$ , and the union of all such spheres will cover  $[0, 2R] \times \mathcal{B}_1^l$ . Therefore, using Lemma A.1, we obtain that for  $0 < \delta \leq 2l + 2$ ,

$$N(\delta, \mathcal{T}_l) \leq N'(\delta/2, \mathcal{T}_l) \leq \frac{4}{\delta/(l + 1)} \cdot \left( \frac{c(d)}{(\delta/(2l + 2))^d} \right)^l,$$

and using the fact that  $\log(c(d)) \leq d$  for any  $d \in \mathbb{Z}_+$ , we have

$$H(\delta, \mathcal{T}) := \log(N(\delta, \mathcal{T})) \leq (ld + 1) + (ld + 1)\log(2(l + 1)/\delta).$$

Using this, and the fact that the maximum  $d_l$  distance of any point in  $\mathcal{T}_l$  from  $(0, \mathbf{0}, \dots, \mathbf{0})$  is bounded by  $l + 1$ , by Corollary 13.2 of Boucheron, Lugosi and Massart [8], we obtain that

$$\begin{aligned} \mathbb{E} \sup_{(r, s_1, \dots, s_l) \in \mathcal{T}_l} W'_l(r, s_1, \dots, s_l) &\leq 12 \int_{\delta=0}^{(l+1)/2} \sqrt{H(\delta, \mathcal{T})} d\delta \\ &\leq 12\sqrt{ld + 1} \int_{\delta=0}^{(l+1)/2} \sqrt{1 + \log(2(l + 1)/\delta)} d\delta \\ &= 12(l + 1)\sqrt{ld + 1} \int_{x=0}^{1/2} \sqrt{1 + \log(2/x)} dx \leq 11(l + 1)\sqrt{ld + 1}. \end{aligned}$$

By the definition of  $W'_l$ , this implies that  $\mathbb{E}(W_l) \leq 11(l + 1)\sqrt{ld + 1} \cdot L\sqrt{k + 1}$ . Moreover, it is easy to check that the conditions of Theorem 5.8 of Boucheron et al. [8] hold for  $W_l$ , and thus for any  $t \geq 0$ ,

$$\mathbb{P}(W_l \geq \mathbb{E}(W_l) + t) \leq \exp\left(-\frac{t^2}{2\sigma_{W_l}^2}\right), \tag{A.18}$$

with

$$\begin{aligned} \sigma_{W_l}^2 &:= \sup_{(r, s_1, \dots, s_l) \in \mathcal{T}_l} \mathbb{E} \left[ \left( \sum_{i=0}^k \langle \varphi_i(r, s_1, \dots, s_l), \mathbf{Z}_i \rangle \right)^2 \right] \\ &\leq \sum_{i=0}^k M \mathbb{E} \left[ \|\mathbf{Z}_i\|^2 \right] = (k + 1)Md_o\sigma_Z^2, \end{aligned}$$

and (A.16) follows. The proof of (A.17) is similar. □

*Proof of Proposition 4.1* Using Assumption 2.1, we know that

$$\sum_{i=0}^k \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 \geq \frac{c(\mathbf{u}, T)}{h} \|\mathbf{v} - \mathbf{u}\|^2; \tag{A.19}$$

thus, it suffices to lower-bound the terms  $2 \sum_{i=0}^k \langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}), \mathbf{Z}_i \rangle$ .

Let  $\varphi_i(r, s) := \frac{\Phi_{t_i}(\mathbf{u}) - \Phi_{t_i}(\mathbf{u} + r\mathbf{s})}{r}$  for  $(r, s) \in \mathcal{T}_1, r > 0$ . ( $\mathcal{T}_1$  was defined as in Lemma A.2.) We continuously extend it to  $r = 0$  as

$$\varphi_i(0, s) := \lim_{r \rightarrow 0} \frac{\Phi_{t_i}(\mathbf{u}) - \Phi_{t_i}(\mathbf{u} + r\mathbf{s})}{r} = \mathbf{J}\Phi_{t_i}(\mathbf{u})\mathbf{s}.$$

Based on Lemma A.2, the lower bound of the random part can be obtained based on the upper bound on the quantity  $W_1 := \sup_{(r,s) \in \mathcal{T}} \sum_{i=0}^k \langle \varphi_i(r, s), \mathbf{Z}_i \rangle$ , since

$$2 \sum_{i=0}^k \langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}), \mathbf{Z}_i \rangle \geq -2W_1 \|\mathbf{v} - \mathbf{u}\|. \tag{A.20}$$

Now we are going to obtain bounds on the constants  $L$  and  $M$  of Lemma A.2. We have

$$\begin{aligned} \|\mathbf{J}_s \varphi_i(r, s)\| &= \|\mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s})\| \leq \widehat{M}_1(T), \text{ and} \\ \left\| \frac{\partial}{\partial r} \varphi_i(r, s) \right\| &= \left\| \frac{-\mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s}) \cdot r\mathbf{s} + (\Phi_{t_i}(\mathbf{u} + r\mathbf{s}) - \Phi_{t_i}(\mathbf{u}))}{r^2} \right\| \leq \frac{\widehat{M}_2(T)}{2}; \end{aligned}$$

thus, the  $\varphi_i(r, s)$  is  $L$ -Lipschitz with respect to the  $d_1$  distance for  $L := \widehat{M}_2(T)R + \widehat{M}_1(T)$ . Moreover, from the definition of  $\varphi_i(r, s)$ , by (1.19), it follows that  $\|\varphi_i(r, s)\| \leq M$  for  $M := \widehat{M}_1(T)$ . The claim of the proposition now follows by Lemma A.2. □

*Proof of Proposition 4.2* From the definitions, we have

$$\begin{aligned} l^{\text{sm}}(\mathbf{v}) &= \sum_{i=0}^k \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 + 2 \sum_{i=0}^k \langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}), \mathbf{Z}_i \rangle, \text{ and} \\ l_{\mathcal{G}}^{\text{sm}}(\mathbf{v}) &= \sum_{i=0}^k \|\mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})\|^2 + \sum_{i=0}^k \mathbf{J}^2 \Phi_{t_i}(\mathbf{u})[\mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u}, \mathbf{Z}_i] \\ &\quad + 2 \sum_{i=0}^k \langle \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u}), \mathbf{Z}_i \rangle; \end{aligned}$$

thus,

$$\begin{aligned}
 |I^{\text{sm}}(\mathbf{v}) - I_{\mathcal{G}}^{\text{sm}}(\mathbf{v})| &\leq \sum_{i=0}^k \left| \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 - \|\mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})\|^2 \right| \\
 &+ 2 \left| \sum_{i=0}^k \left\langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u}) - \frac{1}{2} \mathbf{J}^2 \Phi_{t_i}(\mathbf{u})[\mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u}, \cdot], \mathbf{Z}_i \right\rangle \right|.
 \end{aligned}
 \tag{A.21}$$

The first term in the right-hand side of the above inequality can be upper-bounded as

$$\begin{aligned}
 &\sum_{i=0}^k \left| \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\|^2 - \|\mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})\|^2 \right| \\
 &\leq \sum_{i=0}^k \|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})\| \\
 &\quad \cdot (\|\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})\| + \|\mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})\|) \\
 &\leq (k + 1) \frac{1}{2} \widehat{M}_2(T) \|\mathbf{v} - \mathbf{u}\|^2 \cdot 2\widehat{M}_1(T) \|\mathbf{v} - \mathbf{u}\| \leq \frac{\overline{T}(\mathbf{u}) \widehat{M}_1(T) \widehat{M}_2(T)}{h} \|\mathbf{v} - \mathbf{u}\|^3,
 \end{aligned}
 \tag{A.22}$$

where we have used the multivariate Taylor’s expansion bound (1.19).

For the second term in (A.21), for  $(r, s) \in \mathcal{T}_1$  (defined as in Lemma A.2),  $r > 0$ , let

$$\varphi_i(r, s) := \left( \Phi_{t_i}(\mathbf{u} + r\mathbf{s}) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot r\mathbf{s} - \frac{1}{2} \mathbf{J}^2 \Phi_{t_i}(\mathbf{u})[r\mathbf{s}, r\mathbf{s}, \cdot] \right) / r^3.$$

For  $r = 0$ , this can be continuously extended as

$$\varphi_i(0, s) := \lim_{r \rightarrow 0} \varphi_i(r, s) = \frac{1}{6} \mathbf{J}^3 \Phi_{t_i}(\mathbf{u})[\mathbf{s}, \mathbf{s}, \mathbf{s}, \cdot].$$

Similarly to Lemma A.2, we define  $W_1 := \sup_{(r,s) \in \mathcal{T}_1} \sum_{i=0}^k \langle \varphi_i(r, s), \mathbf{Z}_i \rangle$ , and  $W'_1 := \sup_{(r,s) \in \mathcal{T}_1} \sum_{i=0}^k \langle -\varphi_i(r, s), \mathbf{Z}_i \rangle$ , then the second term in (A.21) can be bounded as

$$\begin{aligned}
 &2 \left| \sum_{i=0}^k \left\langle \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u}) - \frac{1}{2} \mathbf{J}^2 \Phi_{t_i}(\mathbf{u})[\mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u}, \cdot], \mathbf{Z}_i \right\rangle \right| \\
 &\leq 2 \max(W_1, W'_1) \|\mathbf{v} - \mathbf{u}\|^3.
 \end{aligned}$$

Based on (1.19), the partial derivatives of  $\varphi_i(r, s)$  satisfy that

$$\|\mathbf{J}_s \varphi_i(r, s)\| = \frac{\|r \mathbf{J} \Phi_{t_i}(\mathbf{u} + r\mathbf{s}) - r \mathbf{J} \Phi_{t_i}(\mathbf{u}) - r \mathbf{J}^2 \Phi_{t_i}(\mathbf{u})[r\mathbf{s}, \mathbf{s}, \cdot]\|}{r^3} \leq \frac{1}{2} \widehat{M}_3(T), \text{ and}$$

$$\begin{aligned} \left\| \frac{\partial}{\partial r} \varphi_i(r, s) \right\| &= \left\| \left( \mathbf{J}\Phi_{t_i}(\mathbf{u} + rs) \cdot s - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot s - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[rs, s, \cdot] \right) r^3 \right. \\ &\quad \left. - \left( \Phi_{t_i}(\mathbf{u} + rs) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot sr - \frac{1}{2}\mathbf{J}^2\Phi_{t_i}(\mathbf{u})[rs, rs, \cdot] \right) 3r^2 \right\| / r^6 \\ &= -3 \left\| \Phi_{t_i}(\mathbf{u} + rs)\Phi_{t_i}(\mathbf{u}) - \frac{1}{3}\mathbf{J}\Phi_{t_i}(\mathbf{u} + rs) \cdot sr - \frac{2}{3}\mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot sr \right. \\ &\quad \left. - \frac{1}{6}\mathbf{J}^2\Phi_{t_i}(\mathbf{u})[rs, rs, \cdot] \right\| / r^4 \\ &\leq \frac{1}{4}\widehat{M}_4(T); \end{aligned}$$

therefore,  $\varphi_i$  is  $L$ -Lipschitz with respect to the distance  $d_1$  for  $L := \frac{1}{2}\widehat{M}_3(T) + \frac{1}{2}\widehat{M}_4(T)R$ , and we have  $\|\varphi_i(r, s)\| \leq M$  for  $M := \frac{1}{6}\widehat{M}_3(T)$ . The claim of the proposition now follows by applying Lemma A.2 to  $W_1$  and  $W'_1$  separately (for  $\varepsilon/2$  instead of  $\varepsilon$ ), and then using the union bound.  $\square$

*Proof of Proposition 4.3* From the definitions, we have

$$\begin{aligned} \nabla l^{\text{sm}}(\mathbf{v}) &= 2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{v})' \cdot (\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})) + 2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \cdot \mathbf{Z}_i, \text{ and} \\ \nabla l_{\mathcal{G}}^{\text{sm}}(\mathbf{v}) &= 2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u}) + 2 \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\cdot, \mathbf{v} - \mathbf{u}, \mathbf{Z}_i] \\ &\quad + 2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \cdot \mathbf{Z}_i; \end{aligned}$$

thus,

$$\begin{aligned} \|\nabla l^{\text{sm}}(\mathbf{v}) - \nabla l_{\mathcal{G}}^{\text{sm}}(\mathbf{v})\| &\leq 2 \left\| \sum_{i=0}^k \left( \mathbf{J}\Phi_{t_i}(\mathbf{v})' - \mathbf{J}\Phi_{t_i}(\mathbf{u})' \right) (\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})) \right\| \\ &\quad + 2 \left\| \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' (\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})) \right\| \\ &\quad + 2 \left\| \sum_{i=0}^k \left( \mathbf{J}\Phi_{t_i}(\mathbf{v})' - \mathbf{J}\Phi_{t_i}(\mathbf{u})' - \left( \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\mathbf{v} - \mathbf{u}, \cdot, \cdot] \right)' \right) \cdot \mathbf{Z}_i \right\|. \quad (\text{A.23}) \end{aligned}$$

Using the multivariate Taylor’s expansion bound (1.19), the first two in the right-hand side of the above inequality can be upper-bounded as

$$\begin{aligned}
 & 2 \left\| \sum_{i=0}^k (\mathbf{J}\Phi_{t_i}(\mathbf{v})' - \mathbf{J}\Phi_{t_i}(\mathbf{u})') (\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})) \right\| \\
 & + 2 \left\| \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' (\Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u}) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})) \right\| \\
 & \leq 4 \frac{\overline{T}(\mathbf{u})}{h} \widehat{M}_1(T) \widehat{M}_2(T) \|\mathbf{v} - \mathbf{u}\|^2.
 \end{aligned}
 \tag{A.24}$$

For the last term in (A.23), for  $(r, s_1, s_2) \in \mathcal{T}_2$  (defined as in Lemma A.2),  $r > 0$ , we let

$$\varphi_i(r, s_1, s_2) := \left( \mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1)[s_2, \cdot] - \mathbf{J}\Phi_{t_i}(\mathbf{u})[s_2, \cdot] - r\mathbf{J}^2\Phi_{t_i}(\mathbf{u})[s_1, s_2, \cdot] \right) / r^2.$$

We extend this continuously to  $r = 0$  as

$$\varphi_i(0, s_1, s_2) := \lim_{r \rightarrow 0} \varphi_i(r, s_1, s_2) = \frac{1}{2} \mathbf{J}^3\Phi_{t_i}(\mathbf{u})[s_1, s_1, s_2, \cdot].$$

We define  $W_2$  as in Lemma A.2 as  $W_2 := \sup_{(r, s_1, s_2) \in \mathcal{T}_2} \sum_{i=0}^k \langle \varphi_i(r, s_1, s_2), \mathbf{Z}_i \rangle$ , and then the last term of (A.23) can be bounded as

$$2 \left\| \sum_{i=0}^k \left( \mathbf{J}\Phi_{t_i}(\mathbf{v})' - \mathbf{J}\Phi_{t_i}(\mathbf{u})' - \left( \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[\mathbf{v} - \mathbf{u}, \cdot, \cdot] \right)' \right) \cdot \mathbf{Z}_i \right\| \leq 2W_2 \|\mathbf{v} - \mathbf{u}\|^2.
 \tag{A.25}$$

By (1.19), for any  $(r, s_1, s_2) \in \mathcal{T}_2$ , the partial derivatives of  $\varphi_i$  satisfy that

$$\begin{aligned}
 \|\mathbf{J}_{s_1}\varphi_i(r, s_1, s_2)\| & \leq \|r\mathbf{J}^2\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1) - r\mathbf{J}^2\Phi_{t_i}(\mathbf{u})\| / r^2 \leq \widehat{M}_3(T), \\
 \|\mathbf{J}_{s_2}\varphi_i(r, s_1, s_2)\| & = \|\mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[r\mathbf{s}_1, \cdot, \cdot]\| / r^2 \leq \frac{1}{2} \widehat{M}_3(T), \\
 \|\mathbf{J}_r\varphi_i(r, s_1, s_2)\| & \leq \left\| \frac{\partial}{\partial r} \left( \frac{\mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[r\mathbf{s}_1, \cdot, \cdot]}{r^2} \right) \right\| \\
 & = \left\| r^2(\mathbf{J}^2\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1)[s_1, \cdot, \cdot] - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[s_1, \cdot, \cdot]) \right. \\
 & \quad \left. - 2r(\mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[r\mathbf{s}_1, \cdot, \cdot]) \right\| / r^4 \\
 & = \left\| -2 \left[ \mathbf{J}\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1) - \mathbf{J}\Phi_{t_i}(\mathbf{u}) - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[r\mathbf{s}_1, \cdot, \cdot] \right] \right. \\
 & \quad \left. - \frac{1}{2}(\mathbf{J}^2\Phi_{t_i}(\mathbf{u} + r\mathbf{s}_1)[s_1, \cdot, \cdot] - \mathbf{J}^2\Phi_{t_i}(\mathbf{u})[s_1, \cdot, \cdot]) \right\| / r^3 \leq \widehat{M}_4(T).
 \end{aligned}$$

Based on these bounds, we can see that  $\varphi_i$  is  $L$ -Lipschitz with respect to the  $d_2$  distance for  $L := 2R\widehat{M}_4(T) + \widehat{M}_3(T)$ , and it satisfies that  $\|\varphi_i(r, s_1, s_2)\| \leq M$  for  $M := \frac{1}{2}\widehat{M}_3(T)$ . The claim of the proposition now follows by Lemma A.2.  $\square$

*Proof of Proposition 5.2* By (4.1) and (4.3), we have

$$\begin{aligned} \nabla^2 \log \mu^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) &= \nabla^2 \log q(\mathbf{v}) - \frac{1}{\sigma_Z^2} \nabla^2 l^{\text{sm}}(\mathbf{v}) = \nabla^2 \log q(\mathbf{v}) - \frac{1}{2\sigma_Z^2} \\ &\cdot \left( 2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{v})' \cdot \mathbf{J}\Phi_{t_i}(\mathbf{v}) + 2 \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})] \right. \\ &\quad \left. + 2 \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \mathbf{Z}_i] \right). \end{aligned}$$

We first study the deterministic terms. Notice that

$$\left\| \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{v})' \cdot \mathbf{J}\Phi_{t_i}(\mathbf{v}) - \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \cdot \mathbf{J}\Phi_{t_i}(\mathbf{u}) \right\| \leq \frac{2\widehat{M}_1(T)\widehat{M}_2(T)\overline{T}(\mathbf{u})\|\mathbf{v} - \mathbf{u}\|}{h}.$$

By Assumption 2.1, it follows that  $\sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{u})' \cdot \mathbf{J}\Phi_{t_i}(\mathbf{u}) \geq \frac{c(\mathbf{u}, T)}{h} \cdot \mathbf{I}_d$ ; thus, for every  $\mathbf{v} \in B(\mathbf{u}, r_H(\mathbf{u}, T))$ , we have

$$\begin{aligned} &\nabla^2 \log q(\mathbf{v}) - \frac{1}{2\sigma_Z^2} \left( 2 \sum_{i=0}^k \mathbf{J}\Phi_{t_i}(\mathbf{v})' \cdot \mathbf{J}\Phi_{t_i}(\mathbf{v}) + 2 \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})] \right) \\ &\leq \left( C_q^{(2)} - \frac{3}{4} \cdot \frac{c(\mathbf{u}, T)}{\sigma_Z^2 h} \right) \cdot \mathbf{I}_d. \end{aligned}$$

For the random terms, we first define  $\overline{\varphi}_i : \overline{\mathcal{T}}_2 \rightarrow \mathbb{R}$  ( $\overline{\mathcal{T}}_2$  was defined in Lemma A.2) for  $0 \leq i \leq k$  as

$$\overline{\varphi}_i(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2) := \mathbf{J}^2\Phi_{t_i}(\mathbf{v})[\mathbf{s}_1, \mathbf{s}_2, \cdot].$$

Based on these, we let  $\overline{W}_2 := \sup_{(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2) \in \overline{\mathcal{T}}_2} \sum_{i=0}^k \langle \overline{\varphi}_i(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2), \mathbf{Z}_i \rangle$ , and then the random terms can be bounded as

$$\left\| -\frac{1}{2\sigma_Z^2} \cdot 2 \sum_{i=0}^k \mathbf{J}^2\Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \mathbf{Z}_i] \right\| \leq \frac{\overline{W}_2}{\sigma_Z^2}.$$

By its definition, it is easy to see that for every  $0 \leq i \leq k$ ,  $\overline{\varphi}_i$  is  $L$ -Lipschitz with respect to the  $\overline{d}_2$  distance for  $L := \widehat{M}_2(T) + \widehat{M}_3(T)R$ , and that  $\|\overline{\varphi}_i(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2)\| \leq M$  for  $M := \widehat{M}_2(T)$  for every  $(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2) \in \overline{\mathcal{T}}_2$ . The claim of the proposition now follows from Lemma A.2.  $\square$

*Proof of Proposition 5.3* By (4.1) and (4.3), we have

$$\begin{aligned} \nabla^3 \log \mu^{\text{sm}}(\mathbf{v} | \mathbf{Y}_{0:k}) &= \nabla^3 \log(q(\mathbf{v})) - \frac{1}{2\sigma_Z^2} \nabla^3 l^{\text{sm}}(\mathbf{v}) = \nabla^3 \log(q(\mathbf{v})) - \frac{1}{2\sigma_Z^2} \\ &\cdot \left( 6 \sum_{i=0}^k \mathbf{J}^2 \Phi_{t_i}(\mathbf{v})' \cdot \mathbf{J} \Phi_{t_i}(\mathbf{v}) + 2 \sum_{i=0}^k \mathbf{J}^3 \Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})] \right. \\ &\quad \left. + 2 \sum_{i=0}^k \mathbf{J}^3 \Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \cdot, \mathbf{Z}_i] \right). \end{aligned}$$

Based on the assumption on  $q$ , and (1.19), the deterministic terms can be bounded as

$$\begin{aligned} &\left\| \nabla^3 \log(q(\mathbf{v})) - \frac{1}{2\sigma_Z^2} \left( 6 \sum_{i=0}^k \mathbf{J}^2 \Phi_{t_i}(\mathbf{v})' \cdot \mathbf{J} \Phi_{t_i}(\mathbf{v}) + 2 \sum_{i=0}^k \mathbf{J}^3 \Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \Phi_{t_i}(\mathbf{v}) - \Phi_{t_i}(\mathbf{u})] \right) \right\| \\ &\leq C_q^{(3)} + \frac{\bar{T}(\mathbf{u})}{\sigma_Z^2 h} (3\widehat{M}_1(T)\widehat{M}_2(T) + 2\widehat{M}_1(T)\widehat{M}_3(T)R). \end{aligned}$$

For the random terms, we first define  $\bar{\varphi}_i : \bar{\mathcal{T}}_3 \rightarrow \mathbb{R}$  ( $\bar{\mathcal{T}}_3$  was defined in Lemma A.2) for  $0 \leq i \leq k$  as

$$\bar{\varphi}_i(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3) := \mathbf{J}^3 \Phi_{t_i}(\mathbf{v})[\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \cdot].$$

Based on this, we let  $\bar{W}_3 := \sup_{(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3) \in \bar{\mathcal{T}}_3} \sum_{i=0}^k \langle \bar{\varphi}_i(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3), \mathbf{Z}_i \rangle$ , and then one can see that the random terms can be bounded as

$$\left\| -\frac{1}{2\sigma_Z^2} \cdot 2 \sum_{i=0}^k \mathbf{J}^3 \Phi_{t_i}(\mathbf{v})[\cdot, \cdot, \cdot, \mathbf{Z}_i] \right\| \leq \frac{\bar{W}_3}{\sigma_Z^2}.$$

By its definition, it is easy to see that for every  $0 \leq i \leq k$ ,  $\bar{\varphi}_i$  is  $L$ -Lipschitz with respect to the  $\bar{d}_3$  distance for  $L := \widehat{M}_3(T) + \widehat{M}_4(T)R$ , and that  $\|\bar{\varphi}_i(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3)\| \leq M$  for  $M := \widehat{M}_3(T)$  for every  $(\mathbf{v}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3) \in \bar{\mathcal{T}}_3$ . The claim of the proposition now follows from Lemma A.2. □

### A.2 Initial Estimator When Some of the Components are Zero

In Sect. 3.2, we have proposed a function  $F$  that allows us to express the unobserved coordinates of  $\mathbf{u}$  from the observed coordinates and their derivatives (in the two observation scenarios described in Sect. 3). By substituting appropriate estimators of the derivatives, we obtained an initial estimator based on Theorem 2.8. Unfortunately, this function  $F$  was not defined when some of coordinates of  $\mathbf{u}$  are 0. In this section we propose a modified version of this estimator that overcomes this difficulty.

We start by a lemma allowing us to run the ODE (1.1) backwards in time (for a while).

**Lemma A.3** Suppose that  $\mathbf{v} \in \mathcal{B}_R$  and the trapping ball assumption (1.3) holds. Then for any  $0 \leq t < C_{\text{der}}^{-1}$ , the series

$$\Psi_{-t}(\mathbf{v}) := \sum_{i=0}^{\infty} \mathbf{D}^i \mathbf{v} \cdot \frac{(-t)^i}{i!} \tag{A.26}$$

is convergent, is well defined and satisfies that  $\Psi_t(\Psi_{-t}(\mathbf{v})) = \mathbf{v}$  and that for any  $i_{\text{max}} \in \mathbb{N}$ ,

$$\left\| \Psi_{-t}(\mathbf{v}) - \sum_{i=0}^{i_{\text{max}}} \mathbf{D}^i \mathbf{v} \cdot \frac{(-t)^i}{i!} \right\| \leq \frac{C_0 C_{\text{der}}^{i_{\text{max}}+1}}{1 - C_{\text{der}} t}. \tag{A.27}$$

*Proof* The result follows from bounds (1.14), and the definition of Eq. (1.1). □

Based on this lemma, given the observations  $\mathbf{Y}_{0:k}$ , we propose the following initial estimator. First, select some intermediate indices  $0 = i_1 < i_2 < \dots < i_m < k$  satisfying that  $i_m \cdot h < C_{\text{der}}^{-1}$ . For each index  $i_r$ , we compute the derivative estimates  $\hat{\boldsymbol{\phi}}^{(l)}$  of  $\mathbf{D}^l(\mathbf{u}(t_{i_r}))$  and then use the function  $F$  described in Sect. 3.2 to obtain initial estimators  $\widehat{\mathbf{u}}(t_{i_r})$  of  $\mathbf{u}(t_{i_r})$  for  $0 \leq r \leq m$ . After this, we project these estimators to  $\mathcal{B}_R$  (see (3.13)), and run them backwards by  $t_{i_r}$  time units via approximation (A.27), and project them back to  $\mathcal{B}_R$ , that is, for some sufficiently large  $i_{\text{max}} \in \mathbb{N}$ , we let

$$\hat{\mathbf{u}}^r := P_{\mathcal{B}_R} \left( \sum_{i=0}^{i_{\text{max}}} \mathbf{D}^i \left( P_{\mathcal{B}_R} \left( \widehat{\mathbf{u}}(t_{i_r}) \right) \right) \cdot \frac{(-t_{i_r})^i}{i!} \right) \text{ for } 0 \leq r \leq m. \tag{A.28}$$

The final initial estimator  $\hat{\mathbf{u}}$  is then chosen as the one among  $(\hat{\mathbf{u}}^r)_{0 \leq r \leq m}$  that has the largest a posteriori probability  $\mu^{\text{sm}}(\hat{\mathbf{u}}^r | \mathbf{Y}_{0:k})$  (see (2.4)). Based on (2.8), and some algebra, one can show that this estimator will satisfy the conditions required for the convergence of Newton’s method (Theorem 2.5) if  $\sigma_Z \sqrt{h}$  and  $h$  are sufficiently small, and  $i_{\text{max}}$  is sufficiently large, as long as at least one of the vectors  $(\mathbf{u}(t_{i_r}))_{0 \leq r \leq m}$  has no zero coefficients. Moreover, by a continuity argument, it is possible to show that Assumption 2.2 holds as long as there is a  $t \in [0, T]$  such that none of the coefficients of  $\mathbf{u}(t)$  are 0.

## References

1. R. Bannister. A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 2016.
2. J. M. Bardsley, A. Solonen, H. Haario, and M. Laine. Randomize-then-optimize: a method for sampling from posterior distributions in nonlinear inverse problems. *SIAM J. Sci. Comput.*, 36(4): A1895–A1910, 2014. ISSN 1064-8275.
3. F. L. Bauer and C. T. Fike. Norms and exclusion theorems. *Numer. Math.*, 2: 137–141, 1960.
4. M. Benning and M. Burger. Modern regularization methods for inverse problems. *arXiv preprint arXiv:1801.09922*, 2018.
5. R. Bhatia and T. Jain. Higher order derivatives and perturbation bounds for determinants. *Linear Algebra Appl.*, 431(11):2102–2108, 2009.

6. D. Blomker, K. Law, A. Stuart, and K. Zygalkis. Accuracy and stability of the continuous time 3DVAR filter for the Navier-Stokes equation. *Nonlinearity*, 26:2193–2227, 2013.
7. M. Bocquet, C. A. Pires, and L. Wu. Beyond Gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Review*, 138(8):2997–3023, 2010.
8. S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.
9. S. Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.
10. S. L. Cotter, M. Dashti, J. C. Robinson, and A. M. Stuart. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems*, 25(11):115008, 43, 2009. ISSN 0266-5611.
11. S. L. Cotter, M. Dashti, and A. M. Stuart. Variational data assimilation using targetted random walks. *Internat. J. Numer. Methods Fluids*, 68(4):403–421, 2012. ISSN 0271-2091.
12. S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. MCMC methods for functions: modifying old algorithms to make them faster. *Statist. Sci.*, 28(3):424–446, 2013. ISSN 0883-4237.
13. D. Crisan and B. Rozovskii. *The Oxford Handbook of Nonlinear Filtering*. OUP, Oxford, 2011.
14. T. Cui, K. J. Law, and Y. M. Marzouk. Dimension-independent likelihood-informed mcmc. *Journal of Computational Physics*, 304:109–137, 2016.
15. M. Dashti and A. M. Stuart. The Bayesian approach to inverse problems. *Handbook of Uncertainty Quantification*, pages 311–428, 2017.
16. M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss. MAP estimators and their consistency in Bayesian nonparametric inverse problems. *Inverse Problems*, 29(9):095017, 27, 2013.
17. F.-X. L. Dimet and V. Shutyaev. On deterministic error analysis in variational data assimilation. *Nonlinear Processes in Geophysics*, 12(4):481–490, 2005.
18. I. Dumer. Covering spheres with spheres. *Discrete Comput. Geom.*, 38(4):665–679, 2007.
19. M. M. Dunlop and A. M. Stuart. Map estimators for piecewise continuous inversion. *Inverse Problems*, 32(10):105003, 2016.
20. B. Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Math. Comp.*, 51(184):699–706, 1988.
21. E. Giné and R. Nickl. *Mathematical foundations of infinite-dimensional statistical models*. Cambridge Series in Statistical and Probabilistic Mathematics, [40]. Cambridge University Press, New York, 2016. ISBN 978-1-107-04316-9.
22. K. Hayden, E. Olson, and E. S. Titi. Discrete data assimilation in the Lorenz and 2D Navier-Stokes equations. *Phys. D*, 240(18):1416–1425, 2011. ISSN 0167-2789.
23. T. Helin and M. Burger. Maximum *a posteriori* probability estimates in infinite-dimensional Bayesian inverse problems. *Inverse Problems*, 31(8):085009, 22, 2015. ISSN 0266-5611.
24. E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, 2003.
25. H. Kekkonen, M. Lassas, and S. Siltanen. Posterior consistency and convergence rates for Bayesian inversion with hypoelliptic operators. *Inverse Problems*, 32(8):085005, 31, 2016. ISSN 0266-5611.
26. J. B. Lasserre. *Moments, positive polynomials and their applications*, volume 1 of *Imperial College Press Optimization Series*. Imperial College Press, London, 2010.
27. K. Law, A. Stuart, and K. Zygalkis. *Data assimilation*, volume 62 of *Texts in Applied Mathematics*. Springer, Cham, 2015. A mathematical introduction.
28. K. Law, D. Sanz-Alonso, A. Shukla, and A. Stuart. Filter accuracy for the lorenz 96 model: Fixed versus adaptive observation operators. *Physica D: Nonlinear Phenomena*, 325:1–13, 2016.
29. F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A: Dynamic Meteorology and Oceanography*, 38(2):97–110, 1986.
30. F.-X. Le Dimet, I. M. Navon, and D. N. Daescu. Second-order information in data assimilation. *Monthly Weather Review*, 130(3):629–648, 2002.
31. E. N. Lorenz. Predictability: A problem partly solved. In *Proc. Seminar on predictability*, volume 1, 1996.
32. A. J. Majda and J. Harlim. *Filtering complex turbulent systems*. Cambridge University Press, Cambridge, 2012. ISBN 978-1-107-01666-8.
33. A. J. Majda, J. Harlim, and B. Gershgorin. Mathematical strategies for filtering turbulent dynamical systems. *Discrete Contin. Dyn. Syst.*, 27(2):441–486, 2010. ISSN 1078-0947.

34. F. Monard, R. Nickl, and G. P. Paternain. Efficient nonparametric Bayesian inference for X-Ray transforms. *arXiv preprint arXiv:1708.06332*, 2017.
35. I. M. Navon. Data assimilation for numerical weather prediction: a review. In *Data assimilation for atmospheric, oceanic and hydrologic applications*, pages 21–65. Springer, 2009.
36. R. Nickl. Bernstein-von Mises theorems for statistical inverse problems i: Schrödinger equation. *arXiv preprint arXiv:1707.01764*, 2017.
37. D. Paulin, A. Jasra, A. Beskos, and D. Crisan. A 4D-Var method with flow-dependent background covariances for the shallow-water equations. *arXiv preprint arXiv:1710.11529*, 2017.
38. D. Paulin, A. Jasra, D. Crisan, and A. Beskos. On concentration properties of partially observed chaotic systems. *Advances in Applied Probability*, to appear, 2018.
39. N. S. Pillai, A. M. Stuart, and A. H. Thiéry. Noisy gradient flow from a random walk in Hilbert space. *Stoch. Partial Differ. Equ. Anal. Comput.*, 2(2):196–232, 2014. ISSN 2194-0401.
40. S. Prajna, A. Papachristodoulou, and P. A. Parrilo. Introducing sostools: A general purpose sum of squares programming solver. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 1, pages 741–746. IEEE, 2002.
41. G. O. Roberts and J. S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probab. Surv.*, 1:20–71, 2004.
42. D. Sanz-Alonso and A. M. Stuart. Long-time asymptotics of the filtering distribution for partially observed chaotic dynamical systems. *SIAM/ASA J. Uncertain. Quantif.*, 3(1):1200–1220, 2015.
43. A. M. Stuart and A. R. Humphries. *Dynamical systems and numerical analysis*, volume 2 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 1996.
44. O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478):1311–1328, 1987.
45. R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1997.
46. J. A. Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8(1-2):1–230, 2015.
47. R. H. Tütüncü, K.-C. Toh, and M. J. Todd. Solving semidefinite-quadratic-linear programs using sdpt3. *Mathematical programming*, 95(2):189–217, 2003.
48. C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
49. S. J. Vollmer. Posterior consistency for Bayesian inverse problems through stability and regression results. *Inverse Problems*, 29(12):125011, 32, 2013. ISSN 0266-5611.
50. S. J. Vollmer. Dimension-independent MCMC sampling for inverse problems with non-Gaussian priors. *SIAM/ASA J. Uncertain. Quantif.*, 3(1):535–561, 2015. ISSN 2166-2525.
51. K. Wang, T. Bui-Thanh, and O. Ghattas. A Randomized Maximum A Posteriori Method for Posterior Sampling of High Dimensional Nonlinear Bayesian Inverse Problems. *SIAM J. Sci. Comput.*, 40(1):A142–A171, 2018. ISSN 1064-8275.
52. Z. Yao, Z. Hu, and J. Li. A TV-Gaussian prior for infinite-dimensional Bayesian inverse problems and its numerical implementations. *Inverse Problems*, 32(7):075006, 19, 2016. ISSN 0266-5611.