Gradient Structure Of The Ensemble Kalman Flow With Noise

Alfredo Garbuno-Inigo, Franca Hoffmann, Wuchen Li and Andrew M. Stuart

Abstract. Solving inverse problems without the use of derivatives or adjoints of the forward model is highly desirable in many applications arising in science and engineering. In this paper we study a number of variants on Ensemble Kalman Inversion (EKI) algorithms, with goal being the construction of methods which generate approximate samples from the Bayesian posterior distribution that solves the inverse problem. Furthermore we establish a mathematical framework within which to study this question. Our starting point is the continuous time limit of EKI. We introduce a specific form of additive noise to the deterministic flow, leading to a stochastic differential equation (SDE), and consider an associated SDE which approximates the original SDE by replacing function evaluation differences with exact gradients. We demonstrate that the nonlinear Fokker-Planck equation defined by the mean-field limit of the associated SDE has a novel gradient flow structure, built on the Wasserstein metric and the covariance matrix of the noisy flow. Using this structure, we investigate large time properties of the SDE in the case of a linear observation map, and convergence to an invariant measure which coincides with the Bayesian posterior distribution for the underlying inverse problem. We numerically study the effect of the noisy perturbation of the original derivative-free EKI algorithm, illustrating that it gives good approximate samples from the posterior distribution.

Keywords: Ensemble Kalman Inversion; Kalman–Wasserstein metric; Gradient flow; Mean-field Fokker-Planck equation.

Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA. (agarbuno@caltech.edu; fkoh@caltech.edu; astuart@caltech.edu) Department of Mathematics, UCLA, Los Angeles, CA. (wcli@math.ucla.edu)

1. PROBLEM SETTING

1.1 Background

Consider the inverse problem of finding $u \in \mathbb{R}^d$ from $y \in \mathbb{R}^K$ where

$$y = \mathcal{G}(u) + \eta, \tag{1.1}$$

$\mathcal{G} : \mathbb{R}^d \to \mathbb{R}^K$ is a possibly non-linear forward operator and η is the unknown observational noise; we do assume, however, that the distribution of η is known and that it is a centered Gaussian: $\eta \sim \mathsf{N}(0, \Gamma)$ for a known covariance matrix $\Gamma \in \mathbb{R}^{K\times K}$. The objective of the inverse problem is to find information about the truth $u^\dagger$ underlying the data y; the forward map $\mathcal{G}$, the covariance Γ and the data y are all viewed as given.

A key role in any optimization scheme to solve (1.1) is played by a loss function $\ell(y, \mathcal{G}(u))$. For additive Gaussian noise the natural loss function is [1]

$$\ell(y, y') = \frac{1}{2}\|y - y'\|_\Gamma^2,$$

leading to the least squares loss functional

$$\Phi(u) = \frac{1}{2}\|y - \mathcal{G}(u)\|_\Gamma^2. \tag{1.2}$$

In the Bayesian approach to inversion [23] we place a prior distribution on the unknown u, with Lebesgue density $\pi_0(u)$, then the posterior density on $u|y$, denoted $\pi(u)$, is given by

$$\pi(u) \propto \exp\big(-\Phi(u)\big)\pi_0(u). \tag{1.3}$$

In this paper we will concentrate on the case where the prior is a centred Gaussian $\mathsf{N}(0, \Gamma_0)$, assuming throughout that Γ_0 is strictly positive-definite and hence invertible. If we define

$$R(u) = \frac{1}{2}\|u\|_{\Gamma_0}^2 \tag{1.4}$$

and

$$\Phi_R(u) = \Phi(u) + R(u), \tag{1.5}$$

then

$$\pi(u) \propto \exp\big(-\Phi_R(u)\big). \tag{1.6}$$

Note that the regularization R is of Tikhonov-Phillips form [13].

Our focus throughout is on Ensemble Kalman Inversion (EKI), and variants of it, for the Bayesian approach to inversion for u given y. These methods play an important role in large-scale scientific and engineering applications in which it is undesirable, or impossible, to compute derivatives and adjoints defined by the forward map. Our goal is to introduce a framework for the analysis of a noisy version of EKI which may be used to generate approximate samples from (1.6) based only on evaluations of $\mathcal{G}(u)$.

[1] For any positive-definite symmetric matrix A we define $\langle a, a'\rangle_A = \langle a, A^{-1}a'\rangle = \langle A^{-\frac{1}{2}}a, A^{-\frac{1}{2}}a'\rangle$ and $\|a\|_A = \|A^{-\frac{1}{2}}a\|$.

1.2 Literature Review

The Ensemble Kalman filter was originally introduced as a method for state estimation, and later extended as the EKI to the solution of general inverse problems and parameter estimation problems. For a historical development of the subject, the reader may consult the books [15, 29, 36, 41, 52] and the recent review [6].

The Kalman filter itself was derived for linear Gaussian state estimation problems [24, 25]. In the linear setting ensemble Kalman based methods may be viewed as Monte Carlo approximations of the Kalman filter; in the nonlinear case ensemble Kalman based methods do not converge to the filtering or posterior distribution in the large particle limit [14]. Related interacting particle based methodologies of current interest include Stein variational gradient descent [12, 33, 34], the Fokker-Planck particle dynamics of Reich [48, 51] and the consensus-based optimization techniques given a rigorous setting in [7]. There are also other approaches in which optimal transport is used to evolve a sequence of particles through a transportation map [38, 50] to solve probabilistic state estimation or inversion problems as well as interacting particle systems designed to reproduce the solution of the filtering problem [11, 49, 60].

There has been significant activity devoted to the gradient flow structure associated with the Kalman filter itself. A well-known result is that for a constant state process, Kalman Filtering is the gradient flow with respect to the Fisher-Rao metric [19, 28, 42]. It is worth noting that the Fisher-Rao metric connects to the covariance matrix, see details in [3]. On the other hand, optimal transport [59] demonstrates the importance of the L^2-Wasserstein metric in probability density space. The space of densities equipped with this metric introduces an infinite-dimensional Riemannian manifold, called the density manifold [27, 30, 46]. Solutions to the Fokker-Planck equation are gradient flows of the relative entropy in the density manifold [22, 46]. Designing time-stepping methods which preserve gradient structure is also of current interest: see [48] and, within the context of Wasserstein gradient flows, [31, 32, 57]. The subject of discrete gradients for time-integration of gradient and Hamiltonian systems is developed in [16, 18, 21, 39]. Furthermore, the papers [54, 55] study continuous time limits of EKI algorithms and, in the case of linear inverse problems, exhibit a gradient flow structure for the standard least squares loss function, preconditioned by the empirical covariance of the particles; a related structure was highlighted in [5]. Recent interesting work, which has partially inspired this paper, has built on the work in [55] to study the same problem in the mean-field limit [20]; their mean-field perspective brings considerable insight which we build upon in this paper.

In this paper, we study a new noisy version of EKI, and related mean-field limits, the aim being the construction of methods which lead to approximate posterior samples, without the use of adjoints, and overcoming the issue that the standard noisy EKI does not reproduce the posterior distribution, as highlighted in [14]. We emphasize that the practical derivative-free algorithm that we propose rests on a particle-based approximation of a specific preconditioned gradient flow, as described in section 4.3 of

the paper [26]; we add a judiciously chosen noise to this setting and it is this additional noise which enables approximate posterior sampling. Related approximations are also studied in the paper [48] in which the effect of both time-discretization and particle approximation are discussed when applied to various deterministic interacting particle systems with gradient structure. In order to frame the analysis of our methods, we introduce a new metric, named the Kalman-Wasserstein metric, defined through both the covariance matrix of the mean field limit and the Wasserstein metric. The work builds on the novel perspectives introduced in [20] and leads to new algorithms that will be useful within large-scale parameter learning and uncertainty quantification studies, such as those proposed in [56].

1.3 Our Contribution

The contribution of this paper is summarized as follows:

- We introduce a new noisy perturbation of the continuous time ensemble Kalman inversion algorithm, leading to a stochastic differential equation (SDE); we also introduce a related SDE, in which ensemble differences are approximated by gradients.
- We introduce a mean-field limit of the related SDE, and exhibit a novel Kalman–Wasserstein gradient flow structure in the associated nonlinear Fokker-Planck equation.
- Using this Kalman–Wasserstein structure we characterize the steady states of the nonlinear Fokker-Planck equation, and show that one of them is the posterior density (1.6).
- By explicitly solving the nonlinear Fokker-Planck equation in the case of linear $\mathcal{G}$, we demonstrate that the posterior density is a global attractor for all initial densities of finite energy which are not a Dirac measure.
- Motivated by this analysis we numerically study our new noisy perturbation of the continuous time EKI algorithm, showing that it gives good approximate samples from the posterior distribution.

In Section 2 we recap the EKI methodology, and describe the SDE arising in the case when the data is perturbed with noise. Section 3 introduces the new noisy EKI algorithm, which arises from perturbing the particles with noise, rather than perturbing the data. In Section 4 we discuss the theoretical properties underpinning the proposed new methodology and in Section 5 we describe numerical results which demonstrate the value of the proposed new methodology. We conclude in Section 6.

2. ENSEMBLE KALMAN INVERSION AND RELATED FLOWS

Ensemble Kalman Inversion (EKI) can be interpreted as a derivative-free algorithm to invert $\mathcal{G}$. It operates by evolving an interacting set of particles $\{u_n^{(j)}\}_{j=1}^J$ in discrete

time n. The updates are given by

$$u_{n+1}^{(j)} = u_n^{(j)} + \mathsf{C}^{up}(U_n)\left(\mathsf{C}^{pp}(U_n) + h^{-1}\,\Gamma\right)^{-1}\left(y_{n+1}^{(j)} - \mathcal{G}(u_n^{(j)})\right), \tag{2.1}$$

where $y_{n+1}^{(j)} = y + \xi_{n+1}^{(j)}$ with $\xi_{n+1}^{(j)} \sim \mathsf{N}(0, h^{-1}\Sigma)$ for $\Sigma \in \mathbb{R}^{K\times K}$, and $U_n = \{u_n^{(j)}\}_{n=1}^N$ denotes the collection of all ensemble members at iteration n. The scaling of the covariances Γ and Σ by h^{-1} enables natural passage to a continuous time limit as $h \to 0$. The operators C^{up} and C^{pp} are defined as follows

$$\mathsf{C}^{up}(U_n) = \frac{1}{J}\sum_{k=1}^{J}(u_n^{(k)} - \bar{u}_n)\otimes\left(\mathcal{G}(u_n^{(k)}) - \bar{\mathcal{G}}_n\right) \in \mathbb{R}^{d\times K}, \tag{2.2a}$$

$$\mathsf{C}^{pp}(U_n) = \frac{1}{J}\sum_{k=1}^{J}\left(\mathcal{G}(u_n^{(k)}) - \bar{\mathcal{G}}_n\right)\otimes\left(\mathcal{G}(u_n^{(k)}) - \bar{\mathcal{G}}_n\right) \in \mathbb{R}^{K\times K}, \tag{2.2b}$$

$$\mathsf{C}(U_n) = \frac{1}{J}\sum_{k=1}^{J}(u_n^{(k)} - \bar{u}_n)\otimes(u_n^{(k)} - \bar{u}_n) \in \mathbb{R}^{d\times d}, \tag{2.2c}$$

where $\bar{u}_n$ and $\bar{\mathcal{G}}_n$ denote the sample means given by

$$\bar{u}_n = \frac{1}{J}\sum_{k=1}^{J} u_n^{(k)}, \quad \bar{\mathcal{G}}_n = \frac{1}{J}\sum_{k=1}^{J}\mathcal{G}(u_n^{(k)}). \tag{2.3a}$$

Taking limits as $h \to 0$ in equation (2.1) leads to the following continuous time version of the algorithm:

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\langle\mathcal{G}(u^{(k)}) - \bar{\mathcal{G}}, \mathcal{G}(u^{(j)}) - y\rangle_\Gamma\, u^{(k)} \;+\; \mathsf{C}^{up}(U)\,\Gamma^{-1}\sqrt{\Sigma}\,\dot{\mathbf{W}}^{(j)}, \tag{2.4}$$

where the $\{\mathbf{W}^{(j)}\}$ are a collection of i.i.d. standard Brownian motions in the data space $\mathbb{R}^K$.[2] Here $U(t)$ denotes $\{u^{(j)}(t)\}_{n=1}^N$, the collection of all ensemble members, and the operator C^{up} denotes the empirical cross covariance matrix of the ensemble members, as in (2.2), but evaluated at $U(t)$ rather than U_n. We extend use of $\mathsf{C}(\cdot)$ in the same way in what follows below. The sample means of $\{u^{(j)}(t)\}$ and $\{\mathcal{G}(u^{(j)}(t))\}$ are defined analogously to (2.3) but in continuous time. We refer to (2.4) as an Ensemble Kalman Flow (EKF). In the case of a linear forward model where $\mathcal{G}(u) = Au$ and with the choice $\Sigma = \Gamma$ this flow has mean $\mathfrak{m}$ and covariance $\mathfrak{C}$ which satisfy the closed equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{m}(t) = -\mathfrak{C}(t)\left(A^\top\Gamma^{-1}A\mathfrak{m}(t) - r\right) \tag{2.5a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{C}(t) = -\mathfrak{C}(t)A^\top\Gamma^{-1}A\mathfrak{C}(t), \tag{2.5b}$$

[2] In this SDE, and all that follow, the rigorous interpretation is through the Itô integral formulation of the problem.

where $r := A^\top \Gamma^{-1} y \in \mathbb{R}^d$. These results may be established by similar techniques to those used below in Subsection 4.2. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{C}(t)^{-1} = -\mathfrak{C}(t)^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{C}(t)\right)\mathfrak{C}(t)^{-1} = A^\top\Gamma^{-1}A$$

and therefore $\mathfrak{C}(t)^{-1}$ grows linearly in time. The resulting equations for the mean and covariance are simply those which arise from applying the Kalman-Bucy filter [25] to the model

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}u &= 0\\ \frac{\mathrm{d}}{\mathrm{d}t}z := y &= Au + \sqrt{\Gamma}\dot{\mathbf{W}},\end{aligned}$$

where $\mathbf{W}$ denotes a standard unit Brownian motion in the data space $\mathbb{R}^K$. The exact closed form of equations for the first two moments, in the setting of the Kalman-Bucy filter, was established in section 4 of the paper [49] for finite particle approximations, and transfers verbatim to this mean-field setting.

The equations (2.5a), (2.5b) are not consistent with having (1.6) as an invariant measure in the linear setting $\mathcal{G}(u) = Au$. Even if prior information is added via the initial condition on the mean and covariance, the dynamics show that the limiting covariance shrinks to zero: $\mathfrak{C}(t) \to 0$ as $t \to \infty$. A key point to appreciate is that the noise introduced in (2.4) arises from the observation y being perturbed with additional noise. In what follows we instead directly perturb the particles themselves. The benefits of introducing noise on the particles, rather than the data, was demonstrated in [26], although in that setting only optimization, and not Bayesian inversion, is considered.

3. NOISY PARTICLE ENSEMBLE KALMAN FLOW

In this section we develop a sequence of continuous time problems, motivated by the EKF above, with the goal of identifying interacting particle systems of stochastic differential equations which capture the posterior distribution (1.6). We will identify the system (3.2) which has exactly this property, but requires the evaluation of gradients of Φ, something which is undesirable in many applications. We also identify system (3.1) which may be viewed as a derivative-free approximation of (3.2) and which, in the case of linear $\mathcal{G}$, coincides with (3.2). We then take the mean-field limit of (3.2) and identify the corresponding Fokker-Planck equation (3.6).

To this end we modify (2.4) by adding a prior related damping term, as in [10], and changing the noise covariance structure to obtain

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\langle \mathcal{G}(u^{(k)}) - \bar{\mathcal{G}}, \mathcal{G}(u^{(j)}) - y\rangle_\Gamma \, u^{(k)} \; - \; \mathsf{C}(U)\Gamma_0^{-1}u^{(j)} + \sqrt{2\mathsf{C}(U)}\,\dot{\mathbf{W}}^{(j)}. \quad (3.1)$$

Here the $\{\mathbf{W}^{(j)}\}$ are a collection of i.i.d. standard Brownian motions in the space $\mathbb{R}^d$ of the unknown parameter u. This is another EKF. Related to this system of interacting SDEs we will consider the system

$$\dot{u}^{(j)} = -\mathsf{C}(U)\nabla\Phi_R(u^{(j)}) + \sqrt{2\,\mathsf{C}(U)}\,\dot{\mathbf{W}}^{(j)}\,. \tag{3.2}$$

This equation may be re-written as

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\langle d\mathcal{G}(u^{(j)})\big(u^{(k)}-\bar{u}\big), \mathcal{G}(u^{(j)})-y\rangle_\Gamma\, u^{(k)} \;-\; \mathsf{C}(U)\Gamma_0^{-1}u^{(j)} + \sqrt{2\mathsf{C}(U)}\,\dot{\mathbf{W}}^{(j)}. \tag{3.3}$$

(We used the fact that it is possible to replace $u^{(k)}$ by $u^{(k)}-\bar{u}$ after the $\Gamma-$weighted inner-product in (3.1) and (3.3) without changing the equation.) From this it is clear that in the linear case where

$$\mathcal{G}(u) = Au \tag{3.4}$$

the two systems (3.1) and (3.2) are identical. It is also natural to conjecture that if the particles are close to one another then (3.1) and (3.2) will generate similar particle distributions. Based on this exact (in the linear case) and conjectured (in the nonlinear case) relationship we propose (3.1) as a derivative-free algorithm to approximately sample the Bayesian posterior distribution, and we propose (3.2) as a natural object of analysis in order to understand this sampling algorithm.

In order to write down the mean field limit of (3.2), we define the macroscopic mean and covariance:

$$m(\rho) := \int v\rho\,\mathrm{d}v\,, \qquad \mathcal{C}(\rho) := \int \big(v-m(\rho)\big)\otimes\big(v-m(\rho)\big)\,\rho(v)\,\mathrm{d}v\,.$$

Taking the large particle limit leads to the mean field equation

$$\dot{u} = -\,\mathcal{C}(\rho)\nabla\Phi_R(u) + \sqrt{2\,\mathcal{C}(\rho)}\,\dot{W}, \tag{3.5}$$

with corresponding nonlinear Fokker-Planck equation

$$\partial_t\rho = \nabla\cdot\big(\rho\,\mathcal{C}(\rho)\nabla\Phi_R(u)\big) + \mathcal{C}(\rho) : D^2\rho\,. \tag{3.6}$$

Here $A_1 : A_2$ denotes the Frobenius inner-product between matrices A_1 and A_2. The rigorous derivation of the mean-field equations (3.5) and (3.6) is left for future work; for foundational work relating to mean field limits, see [8, 17, 47, 58] and the references therein. The following lemma states the intuitive fact that the covariance, which plays a central role in equation (3.6), vanishes only for Dirac measures.

LEMMA 1. *The only probability densities $\rho \in \mathcal{P}(\mathbb{R}^d)$ at which $\mathcal{C}(\rho)$ vanishes are Diracs,*

$$\rho(u) = \delta_v(u) \text{ for some } v \in \mathbb{R}^d \quad \Leftrightarrow \quad \mathcal{C}(\rho) = 0\,.$$

PROOF. That $\mathcal{C}(\delta_v) = 0$ follows by direct substitution. For the converse, note that $\mathcal{C}(\rho) = 0$ implies $\int |u|^2 \rho \,\mathrm{d}u = \left(\int u\rho \,\mathrm{d}u\right)^2$, which is the equality case of Jensen's inequality, and therefore only holds if ρ is the law of a constant random variable. □

4. THEORETICAL PROPERTIES

In this section we discuss theoretical properties of (3.6) which motivate the use of (3.1) and (3.2) as particle systems to generate approximate samples from the posterior distribution (1.6). In Subsection 4.1 we exhibit a gradient flow structure for (3.6) which shows that solutions evolve towards the posterior distribution (1.6) unless they collapse to a Dirac measure. In Subsection 4.2 we show that in the linear case, collapse to a Dirac does not occur if the initial condition is a Gaussian with non-zero covariance, and instead convergence to the posterior distribution is obtained. In Subsection 4.3 we introduce a novel metric structure which underpins the results in the two preceding sections, and will allow for a rigorous analysis of the long-term behavior of the nonlinear Fokker-Planck equation in future work.

4.1 Nonlinear Problem

Because $\mathcal{C}(\rho)$ is independent of u, we may write equation (3.6) in divergence form, which facilitates the revelation of a gradient structure:

$$\partial_t \rho = \nabla \cdot \left(\rho\, \mathcal{C}(\rho) \nabla \Phi_R(u) + \rho\, \mathcal{C}(\rho) \nabla (\ln \rho)\right), \tag{4.1}$$

where we use the fact $\rho \nabla \ln \rho = \nabla \rho$. Thank to the divergence form, it follows that (4.1) conserves mass along the flow, and so we may assume $\int \rho(t,u)\,\mathrm{d}u = 1$ for all $t \geq 0$. Defining the energy

$$E(\rho) = \int \Big(\rho(u)\Phi_R(u) + \rho(u)\ln \rho(u)\Big)\,\mathrm{d}u\,, \tag{4.2}$$

solutions to (4.1) can be written as a gradient flow:

$$\partial_t \rho = \nabla \cdot \left(\rho\, \mathcal{C}(\rho) \nabla \frac{\delta E}{\delta \rho}\right), \tag{4.3}$$

where $\frac{\delta}{\delta \rho}$ denotes the L^2 first variation. This will be made more explicit in Section 4.3, see Proposition 7. Thanks to the gradient flow structure (4.3), stationary states of (3.6) are given either by critical points of the energy E, or by choices of ρ such that $\mathcal{C}(\rho) = 0$ as characterized in Lemma 1. Critical points of E solve the corresponding Euler-Lagrange condition

$$\frac{\delta E}{\delta \rho} = \Phi_R(u) + \ln \rho(u) = c \qquad \text{on supp } (\rho) \tag{4.4}$$

for some constant c. The unique solution to (4.4) with unit mass is given by the Gibbs measure

$$\rho_\infty(u) := \frac{e^{-\Phi_R(u)}}{\int e^{-\Phi_R(u)}\,\mathrm{d}u}\,. \tag{4.5}$$

Then, up to an additive normalization constant, the energy $E(\rho)$ is exactly the relative entropy of ρ with respect to ρ_∞, also known as the Kullback-Leibler divergence $\mathrm{KL}(\rho(t)\|\rho_\infty)$,

$$\begin{aligned} E(\rho) &= \int (\Phi_R + \ln \rho(t))\,\rho\,\mathrm{d}u \\ &= \int \frac{\rho(t)}{\rho_\infty} \ln\left(\frac{\rho(t)}{\rho_\infty}\right) \rho_\infty\,\mathrm{d}u + \ln\left(\int e^{-\Phi_R(u)}\,\mathrm{d}u\right) \\ &= \mathrm{KL}(\rho(t)\|\rho_\infty) + \ln\left(\int e^{-\Phi_R(u)}\,\mathrm{d}u\right)\,. \end{aligned}$$

Thanks to the gradient flow structure (4.3), we can compute the dissipation of the energy

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\Big\{E(\rho)\Big\} &= \left\langle \frac{\delta E}{\delta \rho}\,,\, \partial_t \rho \right\rangle_{L^2(\mathbb{R}^d)} \\ &= -\int \rho \left\langle \nabla \frac{\delta E}{\delta \rho}, \mathcal{C}(\rho)\nabla \frac{\delta E}{\delta \rho} \right\rangle \mathrm{d}u \\ &= -\int \rho \left| \mathcal{C}(\rho)^{\frac{1}{2}} \nabla(\Phi_R + \ln \rho)\right|^2 \mathrm{d}u\,. \end{aligned} \tag{4.6}$$

As a consequence, the energy E decreases along trajectories until either $\mathcal{C}(\rho)$ approaches zero (collapse to a Dirac measure by Lemma 1) or ρ becomes the Gibbs measure with density ρ_∞.

The dissipation of the energy along the evolution of the classical Fokker-Planck equation is known as the Fisher information [59]. We reformulate equation (4.6) by defining the following generalized Fisher information for any covariance matrix Λ,

$$\mathcal{I}_\Lambda(\rho(t)\|\rho_\infty) := \int \rho \left\langle \nabla \ln\left(\frac{\rho}{\rho_\infty}\right)\,,\, \Lambda \nabla \ln\left(\frac{\rho}{\rho_\infty}\right)\right\rangle \mathrm{d}u\,.$$

One may also refer to $\mathcal{I}_\Lambda$ as a Dirichlet form as it is known in the theory of large particle systems, since we can write

$$\mathcal{I}_\Lambda(\rho(t)\|\rho_\infty) = 4\int \rho_\infty \left\langle \nabla \sqrt{\frac{\rho}{\rho_\infty}}\,,\, \Lambda \nabla \sqrt{\frac{\rho}{\rho_\infty}}\right\rangle \mathrm{d}u\,.$$

For $\Lambda = \mathcal{C}(\rho)$, we name functional $\mathcal{I}_\mathcal{C}$ the relative *Kalman-Fisher information*. We conclude that the following energy dissipation equality holds,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{KL}(\rho(t)\|\rho_\infty) = -\mathcal{I}_\mathcal{C}(\rho(t)\|\rho_\infty)\,.$$

To derive a rate of decay to equilibrium in entropy, we aim to identify conditions on Φ_R such that the following logarithmic Sobolev inequality holds: there exists $\lambda > 0$ such that

$$\mathrm{KL}(\rho(t)\|\rho_\infty) \leq \frac{1}{2\lambda}\mathcal{I}_{I_d}(\rho(t)\|\rho_\infty) \qquad \forall \rho\,. \tag{4.7}$$

By [4], it is enough to impose sufficient convexity on Φ_R, i.e. $D^2\Phi_R \geq \lambda I_d$, where $D^2\Phi_R$ denotes the Hessian of Φ_R. This allows us to deduce convergence to equilibrium as long as $\mathcal{C}(\rho)$ is uniformly bounded from below following standard arguments for the classical Fokker-Planck equation as presented for example in [37].

PROPOSITION 2. *Assume there exists $\alpha > 0$ and $\lambda > 0$ such that*

$$\mathcal{C}(\rho(t)) \geq \alpha I_d\,, \qquad D^2\Phi_R \geq \lambda I_d\,.$$

Then any solution $\rho(t)$ to (4.1) with initial condition ρ_0 satisfying $\mathrm{KL}(\rho_0\|\rho_\infty) < \infty$ decays exponentially fast to equilibrium: there exists a constant $c = c(\rho_0, \Phi_R) > 0$ such that for any $t > 0$,

$$\|\rho(t) - \rho_\infty\|_{L^1(\mathbb{R}^d)} \leq ce^{-\alpha\lambda t}\,.$$

This rate of convergence can most likely be improved using the correct logarithmic Sobolev inequality weighted by the covariance matrix $\mathcal{C}$. However, the above estimate already indicates the effect of having the covariance matrix $\mathcal{C}$ present in the Fokker-Planck equation (4.1). The properties of such inequalities in a more general setting is an interesting future avenue to explore. The weighted logarithmic Sobolev inequality that is well adapted to the setting here depends on the geometric structure of the Kalman-Wasserstein metric, see related studies in [30].

PROOF. Thanks to the assumptions, and using the logarithmic Sobolev inequality (4.7), we obtain decay in entropy,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{KL}(\rho(t)\|\rho_\infty) \leq -\alpha\mathcal{I}_{I_d}(\rho(t)|\rho_\infty) \leq -2\alpha\lambda\mathrm{KL}(\rho(t)\|\rho_\infty)\,.$$

We conclude using the Csiszár-Kullback inequality as it is mainly known to analysts, also referred to as Pinsker inequality in probability (see [2] for more details):

$$\frac{1}{2}\|\rho(t) - \rho_\infty\|^2_{L^1(\mathbb{R}^d)} \leq \mathrm{KL}(\rho(t)\|\rho_\infty) \leq \mathrm{KL}(\rho_0\|\rho_\infty)e^{-2\alpha\lambda t}\,.$$

□

4.2 Linear Problem

Here we show that, in the case of a linear forward operator $\mathcal{G}$, the Fokker-Planck equation (which is still nonlinear) has exact Gaussian solutions. This property holds because the mean field equation (3.5) leads to exact closed equations for the mean

and covariance. Once the covariance is known the nonlinear Fokker-Planck equation (3.6) becomes linear, and is explicitly solvable if $\mathcal{G}$ is linear and the initial condition is Gaussian. Consider equation (3.5) in the context of a linear observation map (3.4). In this case, the misfit equals

$$\Phi_R(u) = \frac{1}{2}\|Au - y\|_\Gamma^2 + \frac{1}{2}\|u\|_{\Gamma_0}^2. \tag{4.8}$$

The corresponding gradient can be written as

$$\nabla\Phi_R(u) = B^{-1}u - r\,, \tag{4.9}$$
$$r := A^\top\Gamma^{-1}y \in \mathbb{R}^d\,, \quad B := \left(A^\top\Gamma^{-1}A + \Gamma_0^{-1}\right)^{-1} \in \mathbb{R}^{d\times d}.$$

Note that since we assume that the covariance matrix Γ_0 is invertible, it is then also strictly positive-definite. Thus it follows that B is strictly positive-definite and hence invertible too. We define $u_0 := Br$ noting that this is the solution of the regularized normal equations defining the minimizer of Φ_R in this linear case; equivalently u_0 maximizes the posterior density. Indeed by completing the square we see that we may write

$$\rho_\infty(u) \propto \exp\Big(-\frac{1}{2}\|u - u_0\|_B^2\Big). \tag{4.10}$$

LEMMA 3. *Let $\rho(t)$ be a solution of (3.6) with $\Phi_R(\cdot)$ given by (4.8). Then the mean $m(\rho)$ and covariance matrix $\mathcal{C}(\rho)$ are determined by $\mathfrak{m}(t)$ and $\mathfrak{C}(t)$ which satisfy the evolution equations*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{m}(t) = -\mathfrak{C}(t)(B^{-1}\mathfrak{m}(t) - r) \tag{4.11a}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{C}(t) = -2\mathfrak{C}(t)B^{-1}\mathfrak{C}(t) + 2\mathfrak{C}(t). \tag{4.11b}$$

In addition, for any $\mathfrak{C}(t)$ satisfying (4.11b), its determinant and inverse solve

$$\frac{\mathrm{d}}{\mathrm{d}t}\det\mathfrak{C}(t) = -2\,(\det\mathfrak{C}(t))\,\mathrm{Tr}\left[B^{-1}\mathfrak{C}(t) - I_d\right]\,, \tag{4.12}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathfrak{C}(t)^{-1}\big) = 2B^{-1} - 2\mathfrak{C}(t)^{-1}. \tag{4.13}$$

As a consequence $\mathfrak{C}(t) \to B$ and $\mathfrak{m}(t) \to u_0$ exponentially as $t \to \infty$.

In fact, solving the ODE (4.13) explicitly and using (4.11a), exponential decay immediately follows:

$$\mathfrak{C}(t)^{-1} = \big(\mathfrak{C}(0)^{-1} - B^{-1}\big)\,e^{-2t} + B^{-1}\,, \tag{4.14}$$

and

$$\|\mathfrak{m}(t) - u_0\|_{\mathfrak{C}(t)} = \|\mathfrak{m}(0) - u_0\|_{\mathfrak{C}(0)}e^{-t}\,. \tag{4.15}$$

PROOF. We begin by deriving the evolution of the first and second moments. This is most easily accomplished by working with the mean-field flow SDE (3.5), using the regularized linear misfit written in (4.8). This yields the update

$$\dot{u} = -\mathcal{C}(\rho)\,(B^{-1}u - r) + \sqrt{2\,\mathcal{C}(\rho)}\,\dot{\mathbf{W}}\,,$$

where $\dot{\mathbf{W}}$ denotes a zero mean random variable. Identical results can be obtained by working directly with the PDE for the density, namely (3.6) with the regularized linear misfit given in (4.8). Taking expectations with respect to ρ results in

$$\dot{m}(\rho) = -\mathcal{C}(\rho)\,(B^{-1}m(\rho) - r).$$

Let us use the following auxiliary variable $e = u - m(\rho)$. By linearity of differentiation we can write

$$\dot{e} = -\mathcal{C}(\rho)\,B^{-1}\,e + \sqrt{2\,\mathcal{C}(\rho)}\,\dot{\mathbf{W}}.$$

By definition of the covariance operator, $\mathcal{C}(\rho) = \mathbb{E}[e \otimes e]$, its derivative with respect to time can be written as

$$\dot{\mathcal{C}}(\rho) = \mathbb{E}[\dot{e} \otimes e + e \otimes \dot{e}].$$

However we must also include the Itô correction, using Itô's formula, and we can write the evolution equation of the covariance operator as

$$\dot{\mathcal{C}}(\rho) = -2\,\mathcal{C}(\rho)\,B^{-1}\,\mathcal{C}(\rho) + 2\,\mathcal{C}(\rho).$$

This concludes the proof of (4.11b). For the evolution of the determinant and inverse, note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\det\mathcal{C}(\rho) = \mathrm{Tr}\left[\det\mathcal{C}(\rho)\,\mathcal{C}(\rho)^{-1}\,\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}(\rho)\right], \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}(\rho)^{-1} = -\mathcal{C}(\rho)^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}(\rho)\right)\mathcal{C}(\rho)^{-1},$$

and so (4.12), (4.13) directly follow. Finally, exponential decay is a consequence of the explicit expressions (4.14) and (4.15). □

Thanks to the evolution of the covariance matrix and its determinant, we can deduce that there is a family of Gaussian initial conditions that stay Gaussian along the flow and converge to the equilibrium ρ_∞.

PROPOSITION 4. *Fix a vector $m_0 \in \mathbb{R}^d$, a matrix $\mathcal{C}_0 \in \mathbb{R}^{d\times d}$ and take as initial density the Gaussian distribution*

$$\rho_0(u) := \frac{1}{(2\pi)^{d/2}}(\det\mathcal{C}_0)^{-1/2}\exp\left(-\frac{1}{2}||u - m_0||^2_{\mathcal{C}_0}\right)$$

with mean m_0 and covariance $\mathcal{C}_0$. Then the Gaussian profile

$$\rho(t,u) := \frac{1}{(2\pi)^{d/2}}(\det \mathfrak{C}(t))^{-1/2}\exp\left(-\frac{1}{2}\big|\big|u-\mathfrak{m}(t)\big|\big|^2_{\mathfrak{C}(t)}\right)$$

solves evolution equation (3.6) with initial condition $\rho(0,u)=\rho_0(u)$, and where $\mathfrak{m}(t)$ and $\mathfrak{C}(t)$ evolve according to (4.11a) and (4.11b) with initial conditions m_0 and $\mathcal{C}_0$. As a consequence, for such initial conditions $\rho_0(u)$, the solution of the Fokker-Planck equation (3.6) converges to $\rho_\infty(u)$ given by (4.10) as $t\to\infty$.

Proof. It is straightforward to see that, for $m(\rho)$ and $\mathcal{C}(\rho)$ given by Lemma 3,

$$\nabla\rho = -\mathcal{C}(\rho)^{-1}(u-m(\rho))\,\rho,$$

since both $m(\rho)$ and $\mathcal{C}(\rho)$ are independent of u. Therefore, substituting the Gaussian ansatz $\rho(t,u)$ into the first term in the right hand side of (3.6), we have

$$\begin{aligned}\nabla\cdot\big(\rho\,\mathcal{C}(\rho)(B^{-1}u-r)\big) &= (\nabla\rho)\cdot\mathcal{C}(\rho)(B^{-1}u-r)+\rho\nabla\cdot(\mathcal{C}(\rho)B^{-1}u)\\ &= \big(-\mathcal{C}(\rho)^{-1}(u-m(\rho))\cdot\mathcal{C}(\rho)(B^{-1}u-r)+\mathrm{Tr}[\mathcal{C}(\rho)B^{-1}]\big)\,\rho\\ &= \Big(-\big|\big|u-m(\rho)\big|\big|^2_B+\Big\langle u-m(\rho),u_0-m(\rho)\Big\rangle_B+\mathrm{Tr}[\mathcal{C}(\rho)B^{-1}]\Big)\,\rho,\end{aligned} \tag{4.16}$$

where $B^{-1}=A^\top\Gamma^{-1}A+\Gamma_0^{-1}$, $r=A^\top\Gamma^{-1}y$ and $u_0=B\,r$. Recall that B^{-1} is invertible. The second term on the right hand side of (3.6) can be simplified, as follows

$$\begin{aligned}\mathcal{C}(\rho):D^2\rho &= \mathcal{C}(\rho):\Big(-\mathcal{C}(\rho)^{-1}+\big(\mathcal{C}(\rho)^{-1}(u-m(\rho))\big)\otimes\big(\mathcal{C}(\rho)^{-1}(u-m(\rho))\big)\Big)\rho\\ &= \big(-\mathrm{Tr}[I_d]+||u-m(\rho)||^2_{\mathcal{C}(\rho)}\big)\,\rho.\end{aligned} \tag{4.17}$$

Thus, combining the previous two equations, the right hand side of (3.6) is given by the following expression

$$\left[\mathrm{Tr}[B^{-1}\mathcal{C}(\rho)-I_d]-||u-m(\rho)||^2_B+\big|\big|u-m(\rho)\big|\big|^2_{\mathcal{C}(\rho)}+\Big\langle u-m(\rho),u_0-m(\rho)\Big\rangle_B\right]\rho. \tag{4.18}$$

For the left-hand side of (3.6), note that by (4.11a) and (4.11b),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big|\big|u-m(\rho)\big|\big|^2_{\mathcal{C}(\rho)} &= 2\left\langle\frac{\mathrm{d}}{\mathrm{d}t}(u-m(\rho))\,,\,\mathcal{C}(\rho)^{-1}(u-m(\rho))\right\rangle\\ &\quad+\left\langle(u-m(\rho))\,,\,\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathcal{C}(\rho)^{-1}\big)(u-m(\rho))\right\rangle\\ &= -2\Big\langle u-m(\rho),u_0-m(\rho)\Big\rangle_B+2\,||u-m(\rho)||^2_B-2\,||u-m(\rho)||^2_{\mathcal{C}(\rho)}\end{aligned}$$

and therefore, combining with (4.12),

$$\begin{aligned}\partial_t \rho &= \left[-\frac{1}{2}(\det \mathcal{C}(\rho))^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\det \mathcal{C}(\rho)\right) - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}||u-u_0||^2_{\mathcal{C}(\rho)}\right]\rho \\ &= \left[\mathrm{Tr}[B^{-1}\mathcal{C}(\rho) - I_d] - ||u-m(\rho)||^2_B + \big|\big|u-m(\rho)\big|\big|^2_{\mathcal{C}(\rho)} + \Big\langle u-m(\rho), u_0-m(\rho)\Big\rangle_B\right]\rho,\end{aligned} \tag{4.19}$$

which concludes the first part of the proof. The second part concerning the large time asymptotics is a straightforward consequence of the asymptotic behaviour of $\mathfrak{m}$ and $\mathfrak{C}$ detailed in Lemma 3. □

In the case of the classical Fokker-Planck equation $\mathfrak{C}(t) = I_d$ with a quadratic confining potential, the result in Proposition 4 follows from the fact that the fundamental solution of (3.6) is a Gaussian, see [9].

COROLLARY 5. *Let ρ_0 be a non-Gaussian initial condition for (3.6) in the case where Φ_R is given by (4.8). Assume that ρ_0 satisfies $\mathrm{KL}(\rho_0\|\rho_\infty) < \infty$. Then any solution of (3.6) converges exponentially fast to ρ_∞ given by (4.5) as $t \to \infty$ both in entropy, and in $L^1(\mathbb{R}^d)$.*

PROOF. Let $a \in \mathbb{R}^d$ have Euclidean norm 1 and define $q(t) := \langle a, \mathfrak{C}(t)^{-1}a\rangle$. From equation (4.13) it follows that

$$\dot{q} \le 2\lambda - 2q$$

where λ is the maximum eigenvalue of B^{-1}. Hence it follows that q is bounded above, independently of a, and that hence $\mathfrak{C}$ is bounded from below as an operator. Together with the fact that the Hessian $D^2\Phi_R = B^{-1}$ is bounded from below, we conclude using Proposition 2. □

4.3 Kalman-Wasserstein Gradient Flow

We introduce an infinite-dimensional Riemannian metric structure, which we name the Kalman-Wasserstein metric, in density space. It allows the interpretation of solutions to equation (3.6) as gradient flows in density space. To this end we denote by $\mathcal{P}$ the space of probability measures on a convex set $\Omega \subseteq \mathbb{R}^d$:

$$\mathcal{P} := \left\{\rho \in L^1(\Omega) \,:\, \rho \ge 0 \text{ a.e. }, \int \rho(x)\,\mathrm{d}x = 1\right\}.$$

The probability simplex $\mathcal{P}$ is a manifold with boundary. For simplicity, we focus on the subset

$$\mathcal{P}_+ := \{\rho \in \mathcal{P} \,:\, \rho > 0 \text{ a.e. }, \rho \in C^\infty(\Omega)\}.$$

The tangent space of $\mathcal{P}_+$ at a point $\rho \in \mathcal{P}_+$ is given by

$$T_\rho \mathcal{P}_+ = \left\{ \frac{\mathrm{d}}{\mathrm{d}t}\rho(t)\Big|_{t=0} : \rho(t) \text{ is a curve in } \mathcal{P}_+ \,,\, \rho(0) = \rho \right\}$$
$$= \left\{ \sigma \in C^\infty(\Omega) \colon \int \sigma dx = 0 \right\} .$$

The second equality follows since for all $\sigma \in T_\rho \mathcal{P}_+$ we have $\int \sigma(x)\,\mathrm{d}x = 0$ as the mass along all curves in $\mathcal{P}_+$ remains constant. For the set $\mathcal{P}_+$, the tangent space $T_\rho \mathcal{P}_+$ is therefore independent of the point $\rho \in \mathcal{P}_+$. Cotangent vectors are elements of the topological dual $T_\rho^* \mathcal{P}_+$ and can be identified with tangent vectors via the action of the *Onsager operator* [35, 40, 43, 44, 45]

$$V_{\rho,\mathcal{C}} : T_\rho^* \mathcal{P}_+ \to T_\rho \mathcal{P}_+ .$$

In this paper, we introduce the following new choice of Onsager operator:

$$V_{\rho,\mathcal{C}}(\phi) = -\nabla \cdot (\rho \mathcal{C}(\rho) \nabla \phi) =: (-\Delta_{\rho,\mathcal{C}})\, \phi . \tag{4.20}$$

By Lemma 1, the weighted elliptic operator $\Delta_{\rho,\mathcal{C}}$ becomes degenerate if ρ is a Dirac. For points ρ in the set $\mathcal{P}_+$ that are bounded away from zero, the operator $\Delta_{\rho,\mathcal{C}}$ is well-defined, non-singular and invertible since $\rho\mathcal{C}(\rho) > 0$. So we can write

$$V_{\rho,\mathcal{C}}^{-1} : T_\rho \mathcal{P}_+ \to T_\rho^* \mathcal{P}_+, \quad \sigma \mapsto (-\Delta_{\rho,\mathcal{C}})^{-1}\, \sigma .$$

This provides a 1-to-1 correspondence between elements $\phi \in T_\rho^* \mathcal{P}_+$ and $\sigma \in T_\rho \mathcal{P}_+$. For general $\rho \in \mathcal{P}_+$, we can instead use the pseudo-inverse $(-\Delta_{\rho,\mathcal{C}})^\dagger$, see [30]. With the above choice of Onsager operator, we can define a generalized Wasserstein metric tensor:

Definition 6 (Kalman-Wasserstein metric tensor). *Define*

$$g_{\rho,\mathcal{C}} \,:\, T_\rho \mathcal{P}_+ \times T_\rho \mathcal{P}_+ \to \mathbb{R}$$

as follows:

$$g_{\rho,\mathcal{C}}(\sigma_1, \sigma_2) = \int_\Omega \langle \nabla \phi_1 \,,\, \mathcal{C}(\rho) \nabla \phi_2 \rangle \; \rho \,\mathrm{d}x,$$

where $\sigma_i = (-\Delta_{\rho,\mathcal{C}})\, \phi_i = -\nabla \cdot (\rho \mathcal{C}(\rho) \nabla \phi_i) \in T_\rho \mathcal{P}_+$ *for* $i = 1, 2$.

With this metric tensor, the Kalman-Wasserstein metric $\mathcal{W}_\mathcal{C} \colon \mathcal{P}_+ \times \mathcal{P}_+ \to \mathbb{R}$ can be represented by the geometric action function. Given two densities ρ^0, $\rho^1 \in \mathcal{P}_+$, consider

$$W_\mathcal{C}(\rho^0, \rho^1)^2 = \inf \int_0^1 \int_\Omega \langle \nabla \phi_t \,,\, \mathcal{C}(\rho_t) \nabla \phi_t \rangle \; \rho_t \,\mathrm{d}x$$
$$\text{subject to} \quad \partial_t \rho_t + \nabla \cdot (\rho_t \mathcal{C}(\rho_t) \nabla \phi_t) = 0, \;\; \rho_0 = \rho^0, \;\; \rho_1 = \rho^1,$$

where the infimum is taken among all continuous density paths $\rho_t := \rho(t,x)$ and potential functions $\phi_t := \phi(t,x)$. The Kalman-Wasserstein metric has several interesting mathematical properties, which will be the focus of future work. In this paper, working in $(\mathcal{P}_+, g_{\rho,\mathcal{C}})$, we derive the gradient flow formulation that underpins the formal calculations given in Subsection 4.1 for the energy functional E defined in (4.2).

PROPOSITION 7. *Given a finite functional $\mathcal{F} : \mathcal{P}_+ \to \mathbb{R}$, the gradient flow of $\mathcal{F}(\rho)$ in $(\mathcal{P}_+, g_{\rho,\mathcal{C}})$ satisfies*

$$\partial_t \rho = \nabla \cdot \left(\rho\, \mathcal{C}(\rho) \nabla \frac{\delta \mathcal{F}}{\delta \rho} \right) .$$

PROOF. The Riemannian gradient operator $\mathrm{grad}\mathcal{F}(\rho)$ is defined via the metric tensor $g_{\rho,\mathcal{C}}$ as follows:

$$g_{\rho,\mathcal{C}}(\sigma, \mathrm{grad}\mathcal{F}(\rho)) = \int_\Omega \frac{\delta}{\delta \rho(u)} \mathcal{F}(\rho) \sigma(u) \mathrm{d}u\,, \qquad \forall \sigma \in T_\rho \mathcal{P}_+\,.$$

Thus, for $\phi := (-\Delta_{\rho,\mathcal{C}})^{-1}\sigma \in T^*_\rho \mathcal{P}_+$, we have

$$\begin{aligned} g_{\rho,\mathcal{C}}(\sigma, \mathrm{grad}\mathcal{F}(\rho)) &= \int \phi(u) \mathrm{grad}\mathcal{F}(\rho)\, \mathrm{d}u = -\int \nabla \cdot (\rho \mathcal{C}(\rho) \nabla \phi) \frac{\delta}{\delta \rho} \mathcal{F}(\rho)\, \mathrm{d}u \\ &= \int \left\langle \nabla \phi\,,\, \mathcal{C}(\rho) \nabla \frac{\delta}{\delta \rho} \mathcal{F}(\rho) \right\rangle \rho\, \mathrm{d}u \\ &= -\int \phi(u) \nabla \cdot (\rho \mathcal{C}(\rho) \nabla \frac{\delta}{\delta \rho} \mathcal{F}(\rho))\, \mathrm{d}u. \end{aligned}$$

Hence

$$\mathrm{grad}\mathcal{F}(\rho) = -\nabla \cdot (\rho \mathcal{C}(\rho) \nabla \frac{\delta}{\delta \rho} \mathcal{F}(\rho)).$$

Thus we derive the gradient flow by

$$\partial_t \rho = -\mathrm{grad}\mathcal{F}(\rho) = \nabla \cdot (\rho \mathcal{C}(\rho) \nabla \frac{\delta}{\delta \rho} \mathcal{F}(\rho)).$$

□

REMARK 8. *Our derivation concerns the gradient flow on the subset $\mathcal{P}_+$ of $\mathcal{P}$ for simplicity of exposition. However, a rigorous analysis of the evolution of the gradient flow (4.3) requires to extend the above arguments to the full set of probabilities $\mathcal{P}$, especially as we want to study Dirac measures in view of Lemma 1. If ρ is an element of the boundary of $\mathcal{P}$, one may consider instead the pseudo inverse of the operator $\Delta_{\rho,\mathcal{C}}$. This will be the focus of future work, also see the more general analysis in [1], e.g. Theorem 11.1.6.*

5. NUMERICAL EXPERIMENTS

In this section we demonstrate that the intuition developed in the previous two sections does indeed translate into useful algorithms for generating approximate posterior samples without computing derivatives of the forward map $\mathcal{G}$. We do this by considering a non-Gaussian inverse problem, defined through a nonlinear forward operator $\mathcal{G}$, showing how numerical solutions of (3.1) are distributed after large time, and comparing them with exact posterior samples found from MCMC, and with the method employed in [20].

The numerical experiment considered here is the example originally presented in [14] and also used in [20]. We start by defining the forward map which is given by the one-dimensional elliptic boundary value problem

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(u_1)\,\frac{\mathrm{d}}{\mathrm{d}x}p(x)\right) = 1, \quad x \in [0,1], \tag{5.1}$$

with boundary conditions $p(0) = 0$ and $p(1) = u_2$. The explicit solution for this problem, [see 20], is given by

$$p(x) = u_2 x + \exp(-u_1)\left(-\frac{x^2}{2} + \frac{x}{2}\right). \tag{5.2}$$

The forward model operator $\mathcal{G}$ is then defined by

$$\mathcal{G}(u) = \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}. \tag{5.3}$$

Here $u = (u_1, u_2)^\top$ is a constant vector that we want to find and we assume that we are given noisy measurements y of $p(\cdot)$ at locations $x_1 = 0.25$ and $x_2 = 0.75$. The precise Bayesian inverse problem considered here is to find the distribution of the unknown u conditioned on the observed data y, assuming additive Gaussian noise $\eta \sim \mathsf{N}(0, \Gamma)$, where $\Gamma = 0.1^2\, I_2$ and $I_2 \in \mathbb{R}^{2\times 2}$ is the identity matrix. We use as prior distribution $\mathsf{N}(0, \Gamma_0)$, $\Gamma_0 = \sigma^2 I_2$ with $\sigma = 10$. The resulting Bayesian inverse problem is then solved, approximately, by the algorithms we now outline and with with observed data $y = (27.5, 79.7)^\top$.

5.1 Derivative-Free

In this subsection we apply (3.1) for the solution of the inverse problem (1.1). We use a linearly implicit split-step discretization scheme given by

$$u_{n+1}^{(*,j)} = u_n^{(j)} - \Delta t_n\, \frac{1}{J}\sum_{k=1}^{J} \langle \mathcal{G}(u_n^{(k)}) - \bar{\mathcal{G}}, \mathcal{G}(u_n^{(j)}) - y\rangle_\Gamma\, u_n^{(k)} \; - \Delta t_n\, \mathsf{C}(U_n)\, \Gamma_0^{-1}\, u_{n+1}^{(*,j)} \tag{5.4a}$$

$$u_{n+1}^{(j)} = u_{n+1}^{(*,j)} + \sqrt{2\,\Delta t_n\, \mathsf{C}(U_n)}\, \xi_n^{(j)}, \tag{5.4b}$$

where $\xi_n^{(j)} \sim \mathsf{N}(0, I)$, $\Gamma_0 = \sigma^2 I$ is the prior covariance and Δt_n is an adaptive timestep computed as in [26]. Following [20], we consider an initial ensemble drawn from $\mathsf{N}(0,1) \times \mathsf{U}(90,110)$. The proposed algorithm was run for 30 iterations with an ensemble size of $J = 10^3$.

5.2 Gold Standard: MCMC

In this subsection we describe the specific Random Walk Metropolis Hastings (RWMH) algorithm used to solve the same Bayesian inverse problem as in the previous subsection; we view the results as gold standard samples from the desired posterior distribution. The proposal distribution is a Gaussian centered at the current state of the Markov chain with covariance given by $\Sigma = \tau \times \mathsf{C}(U^*)$, where $\mathsf{C}(U^*)$ is the covariance computed from the last iteration of the algorithm described in the preceding subsection, and τ is a scaling factor tuned for an acceptance rate of 25% [53]. In our case, $\tau = 4$. The RWMH was used to get $N = 10^5$ samples with the Markov chain starting at an approximate solution given by the mean of the last step of the algorithm from the previous subsection.

5.3 Numerical Results

Figure 1 shows the results for the solution of the Bayesian inverse problem considered above. In addition to implementing the algorithms described in the previous two subsections, we also employ a specific implementation of the EKI formulation introduced in the paper of Herty and Visconti [20], and defined by the numerical discretization shown in (5.4), but with $\mathsf{C}(U)$ replaced by the identity matrix I_2; this corresponds to the algorithm from equation (20) of [20], and in particular the last display of their Section 5, with $\xi \sim \mathsf{N}(0, I_2)$. The blue dots correspond to the output of this algorithm at the last iteration. The red dots correspond to the last ensemble of the EKI optimization algorithm (EKI-OPT) as presented in [26]. The orange dots depict the RWMH gold standard described above. Finally, the green dots shows the ensemble members at the last iteration of the proposed new formulation (3.1) of EKI. All versions of EKI were run with the same ensemble size, adaptive timestep scheme and iterations as in Subsection 5.1.

Consider first the top-left panel. The true distribution, computed by RWMH, is shown in gold. Note that the algorithm of [26] collapses to a point (shown in red), unable to escape overfitting, and relating to a form of consensus formation. In contrast, the proposed algorithm of [20], while avoiding overfitting, overestimates the spread of the ensemble members, relative to the gold standard RWMH; this is exhibited in blue overdispersed points. The proposed formulation of EKI (green points) gives results close to the RWMH gold standard. These issues are further demonstrated in the lower panel which shows the misfit (loss) function as a function of iterations for the three algorithms (excluding RWMH); the red line demonstrates overfitting as the misfit value falls below the noise level, whereas the other two algorithms avoid overfitting.

The mismatch between the RWMH and our formulation can be understood from the

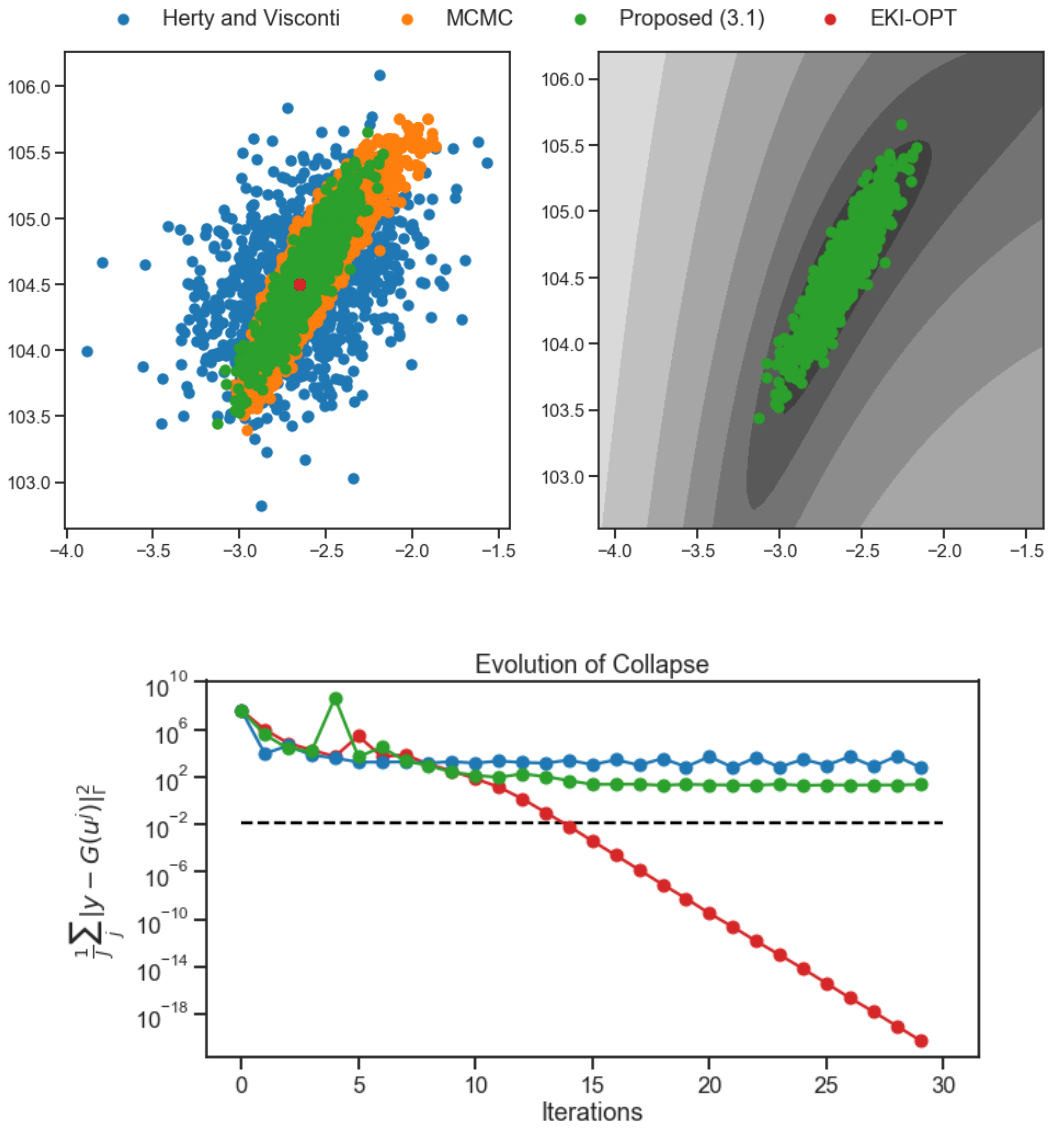

FIG 1. *Results of applying different versions of EKI to the non-linear elliptic boundary problem. For comparison, a Random Walk Metropolis Hastings algorithm is also displayed to provide a gold standard. The proposed formulation and subsequent algorithm captures approximately the true distribution, effectively escaping overfitting or overdispersion shown in the other two implementations. Overfitting is clearly shown from the red line in the lower subfigure. The line in blue, shows the overdispersion effect exhibited by the algorithm proposed in [20]. The upper right subfigure illustrates the approximation to the posterior. Color coding is consistent among the subfigures.*

fact that the use of the ensemble introduces a linear approximation to the curvature of the regularized misfit. This effect is demonstrated in the figure top-right which shows the samples from our proposed algorithm against a background of the level sets of the posterior. However despite this slight mismatch, the key point is that the relatively good set of approximate samples in green is computed without use of the derivative of the forward model $\mathcal{G}$; it thus holds promise as a method for large-scale inverse problems.

6. CONCLUSIONS

In this paper, our goal is to determine how to add noise to the basic EKI algorithm so that it generates approximate samples from the Bayesian posterior distribution. To this end, we propose a new mean-field Fokker-Planck equation which has the desired distribution as an invariant measure. Also, we show how to compute approximate samples from this model by using a particle approximation based on using ensemble difference in place of gradients. Introducing the new Kalman-Wasserstein metric, we propose an appropriate setting for the analysis of such problems. In the future, we will study the properties of the Kalman-Wasserstein metric including its duality, geodesics, and geometric structures. We will investigate the analytical properties of the new metric within Gaussian families. We expect these studies will bring insights to design new numerical algorithms for the approximate solution of inverse problems. Moreover, we will apply the methodology derived here to large-scale inverse problems in which derivative and adjoint computations are not possible.

Acknowledgements: The authors are grateful to José A. Carrillo and Greg Pavliotis for helpful discussions relating to mean field equations and to Sebastian Reich for discussions relating to material at the end of section 2. A.G.I. and A.M.S. are supported by the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program, by Mission Control for Earth, by the Paul G. Allen Family Foundation, and by the National Science Foundation (NSF grant AGS1835860). A.M.S. is also supported by NSF under grant DMS 1818977. F.H. was partially supported by Caltech's von Karman postdoctoral instructorship. W. L. was supported by AFOSR MURI FA9550-18-1-0502.

REFERENCES

[1] L. Ambrosio, N. Gigli, and G. Savaré. Gradient Flows: In Metric Spaces and in the Space of Probability Measures, 2005. 16

[2] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Differential Equations*, 26(1-2):43–100, 2001. 10

[3] N. Ay, J. Jost, H. V. Lê, and L. J. Schwachhöfer. *Information Geometry*. Ergebnisse Der Mathematik Und Ihrer Grenzgebiete A @series of Modern Surveys in Mathematics$l3. Folge, Volume 64. Springer, Cham, 2017. 3

[4] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985. 10

[5] K. Bergemann and S. Reich. A localization technique for ensemble kalman filters. *Quarterly Journal of the Royal Meteorological Society*, 136(648):701–707, 2010. 3

[6] A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e535, 2018. 3

[7] J. A. Carrillo, Y.-P. Choi, C. Totzeck, and O. Tse. An analytical framework for consensus-based global optimization method. *Mathematical Models and Methods in Applied Sciences*, 28(06):1037–1066, 2018. 3

[8] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, Model. Simul. Sci. Eng. Technol., pages 297–336. Birkhäuser Boston, Inc., Boston, MA, 2010. 7

[9] J. A. Carrillo and G. Toscani. Exponential convergence toward equilibrium for homogeneous Fokker-Planck-type equations. *Math. Methods Appl. Sci.*, 21(13):1269–1286, 1998. 14

[10] N. K. Chada, A. M. Stuart, and X. T. Tong. Tikhonov regularization within ensemble kalman inversion. *arXiv preprint arXiv:1901.10382*, 2019. 6

[11] D. Crisan and J. Xiong. Approximate McKean–Vlasov representations for a class of spdes. *Stochastics An International Journal of Probability and Stochastics Processes*, 82(1):53–68, 2010. 3

[12] G. Detommaso, T. Cui, Y. Marzouk, A. Spantini, and R. Scheichl. A stein variational newton method. In *Advances in Neural Information Processing Systems*, pages 9187–9197, 2018. 3

[13] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996. 2

[14] O. G. Ernst, B. Sprungk, and H.-J. Starkloff. Analysis of the ensemble and polynomial chaos kalman filters in bayesian inverse problems. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):823–851, 2015. 3, 17

[15] G. Evensen. *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media, 2009. 3

[16] O. Gonzalez. Time integration and discrete Hamiltonian systems. *Journal of Nonlinear Science*, 6(5):449, 1996. 3

[17] S.-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models*, 1(3):415–435, 2008. 7

[18] E. Hairer and C. Lubich. Energy-diminishing integration of gradient systems. *IMA Journal of Numerical Analysis*, 34(2):452–461, 2013. 3

[19] A. Halder and T. T. Georgiou. Gradient Flows in Filtering and Fisher-Rao Geometry. *arXiv:1710.00064 [cs, math]*, 2017. 3

[20] M. Herty and G. Visconti. Kinetic methods for inverse problems. *arXiv preprint arXiv:1811.09387*, 2018. 3, 4, 17, 18, 19

[21] A. Humphries and A. Stuart. Runge–Kutta methods for dissipative and gradient dynamical systems. *SIAM journal on numerical analysis*, 31(5):1452–1485, 1994. 3

[22] R. Jordan, D. Kinderlehrer, and F. Otto. The Variational Formulation of the Fokker–Planck Equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998. 3

[23] J. Kaipio and E. Somersalo. *Statistical and computational inverse problems*, volume 160. Springer Science & Business Media, 2006. 2

[24] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960. 3

[25] R. E. Kalman and R. S. Bucy. New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83(1):95, 1961. 3, 6

[26] N. B. Kovachki and A. M. Stuart. Ensemble Kalman Inversion: A Derivative-Free Technique For Machine Learning Tasks. *arXiv:1808.03620 [cs, math, stat]*, 2018. 4, 6, 18

[27] J. D. Lafferty. The density manifold and configuration space quantization. *Transactions of the American Mathematical Society*, 305(2):699–741, 1988. 3

[28] R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson's Equation in Nonlinear

Filtering. *SIAM Journal on Control and Optimization*, 53(1):501–525, 2015. 3
[29] K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation: A Mathematical Introduction*. 2015. 3
[30] W. Li. Geometry of probability simplex via optimal transport. *arXiv:1803.06360 [math]*, 2018. 3, 10, 15
[31] W. Li, A. Lin, and G. Montúfar. Affine natural proximal learning. 2019. 3
[32] W. Li and G. Montufar. Natural gradient via optimal transport. *arXiv:1803.07033 [cs, math]*, 2018. 3
[33] Q. Liu and D. Wang. Stein variational gradient descent: A general purpose bayesian inference algorithm. In *Advances In Neural Information Processing Systems*, pages 2378–2386, 2016. 3
[34] J. Lu, Y. Lu, and J. Nolen. Scaling limit of the stein variational gradient descent part i: the mean field regime. *arXiv preprint arXiv:1805.04035*, 2018. 3
[35] S. Machlup and L. Onsager. Fluctuations and irreversible process. II. Systems with kinetic energy. *Physical Rev. (2)*, 91:1512–1515, 1953. 15
[36] A. J. Majda and J. Harlim. *Filtering complex turbulent systems*. Cambridge University Press, 2012. 3
[37] P. A. Markowich and C. Villani. On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis. *Mat. Contemp.*, 19:1–29, 2000. VI Workshop on Partial Differential Equations, Part II (Rio de Janeiro, 1999). 10
[38] Y. Marzouk, T. Moselhy, M. Parno, and A. Spantini. *Sampling via Measure Transport: An Introduction*, pages 1–41. Springer International Publishing, Cham, 2016. 3
[39] R. I. McLachlan, G. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1754):1021–1045, 1999. 3
[40] A. Mielke, M. A. Peletier, and M. Renger. A generalization of onsager's reciprocity relations to gradient flows with nonlinear mobility. *Journal of Non-Equilibrium Thermodynamics*, 41(2), April 2016. SFB 1114 Preprint 10/2015 in arXiv:1510.06219 – accepted for publication in proceedings of the IWNET Workshop. 15
[41] D. S. Oliver, A. C. Reynolds, and N. Liu. *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press, 2008. 3
[42] Y. Ollivier. Online Natural Gradient as a Kalman Filter. *arXiv:1703.00209 [math, stat]*, 2017. 3
[43] L. Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, Feb 1931. 15
[44] L. Onsager. Reciprocal relations in irreversible processes. ii. *Phys. Rev.*, 38:2265–2279, Dec 1931. 15
[45] H. Öttinger. *Beyond Equilibrium Thermodynamics*. Wiley, 2005. 15
[46] F. Otto. The geometry of dissipative evolution equations the porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174, 2001. 3
[47] L. Pareschi and G. Toscani. *Interacting Multiagent Systems: Kinetic equations and Monte Carlo methods*. Number 9780199655465 in OUP Catalogue. Oxford University Press, 2013. 7
[48] S. Pathiraja and S. Reich. Discrete gradients for computational bayesian inference. *arXiv preprint arXiv:1903.00186*, 2019. 3, 4
[49] S. Reich. A dynamical systems framework for intermittent data assimilation. *BIT Numerical Mathematics*, 51(1):235–249, 2011. 3, 6
[50] S. Reich. A nonparametric ensemble transform method for bayesian inference. *SIAM Journal on Scientific Computing*, 35(4):A2013–A2024, 2013. 3
[51] S. Reich. Data assimilation-the Schrödinger perspective. *arXiv preprint arXiv:1807.08351*, 2018. 3
[52] S. Reich and C. Cotter. *Probabilistic forecasting and Bayesian data assimilation*. Cambridge University Press, 2015. 3
[53] G. O. Roberts, A. Gelman, W. R. Gilks, et al. Weak convergence and optimal scaling of random walk metropolis algorithms. *The Annals of Applied Probability*, 7(1):110–120, 1997. 18

[54] C. Schillings and A. M. Stuart. Convergence analysis of ensemble kalman inversion: the linear, noisy case. *Applicable Analysis*, 97(1):107–123. 3

[55] C. Schillings and A. M. Stuart. Analysis of the ensemble kalman filter for inverse problems. *SIAM Journal on Numerical Analysis*, 55(3):1264–1290, 2017. 3

[56] T. Schneider, S. Lan, A. Stuart, and J. Teixeira. Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. *Geophysical Research Letters*, 44(24), 2017. 4

[57] A. Tong Lin, W. Li, S. Osher, and G. Montúfar. Wasserstein proximal of gans. 2018. 3

[58] G. Toscani. Kinetic models of opinion formation. *Commun. Math. Sci.*, 4(3):481–496, 2006. 7

[59] C. Villani. *Optimal Transport: Old and New*. Number 338 in Grundlehren Der Mathematischen Wissenschaften. Springer, Berlin, 2009. 3, 9

[60] T. Yang, P. G. Mehta, and S. P. Meyn. Feedback particle filter. *IEEE transactions on Automatic control*, 58(10):2465–2480, 2013. 3