A Mean-field Optimal Control Formulation for Global Optimization

Chi Zhang, Amirhossein Taghvaei, and Prashant G. Mehta

***Abstract*—A particle filter is introduced to numerically approximate a solution of the global optimization problem. The theoretical significance of this work comes from its variational aspects. Specifically, the particle filter is a controlled interacting particle system where the control input represents the solution of a mean-field type optimal control problem. Its parametric counterpart, obtained when a parametric form of density is known apriori, is shown to be equivalent to the natural gradient algorithm. Explicit formulae for the filter are derived when the objective function is quadratic and the density is Gaussian. The optimal control construction of the particle filter is a significant departure from the classical importance sampling-resampling based approaches.**

I. Introduction

This paper is concerned with the coupled ordinary differential equation-partial differential equation (ODE-PDE) model:

$$\text{ODE:} \qquad \frac{\mathrm{d}X_t^i}{\mathrm{d}t} = -\nabla\phi(X_t^i,t), \quad X_0^i \sim p_0^* \tag{1a}$$

$$\text{PDE:} \quad -\frac{1}{p(x,t)}\nabla\cdot(p(x,t)\nabla\phi(x,t)) = (h(x)-\hat{h}_t) \tag{1b}$$

where $X_t^i \in \mathbb{R}^d$ is the state, at time t, of the i^{th} particle in a population of N particles; $p_0^*(x)$ is a (given) everywhere positive prior density on $\mathbb{R}^d$; and the objective function $h(x)$ is a real-valued function defined on $\mathbb{R}^d$. The control function, $-\nabla\phi(x,t)$, used to define the righthand-side of the ODE (1a), is obtained from solving, at each fixed time $t>0$, the (Poisson equation) PDE (1b) where ∇ and $\nabla\cdot$ denote the gradient and the divergence operators, respectively; $p(x,t)$ is the density of the state X_t^i; and $\hat{h}_t := \int h(x)\,p(x,t)\,\mathrm{d}x$. For the remainder of this paper, the coupled ODE-PDE model (1a)-(1b) is referred to as the particle filter.

The interest in such coupled ODE-PDE models comes from the feedback particle filter (FPF) [25], [24] and related particle flow algorithms for nonlinear filtering [19], [7], [12], [8]. The inspiration for controlling a single particle comes from the mean-field type control formalisms [15], [3], [4], [26] and control methods for optimal transportation [5], [6] where coupled ODE-PDE models also arise.

The particular model (1a)-(1b) is interesting because the Fokker-Planck equation for the density (formally) yields the replicator PDE:

$$\frac{\partial p}{\partial t}(x,t) = \nabla(p(x,t)\nabla\phi(x,t)) = -(h(x)-\hat{h}_t)\,p(x,t)$$

C. Zhang, A. Taghvaei and P. G. Mehta are with the Coordinated Science Laboratory and the Department of Mechanical Science and Engineering at the University of Illinois at Urbana-Champaign (UIUC) czhang54@illinois.edu; taghvae2@illinois.edu; mehtapg@illinois.edu

Financial support from the NSF CMMI grants 1334987 and 1462773 is gratefully acknowledged.

The conference versions of this paper appear in [27].

where the second equality follows from (1b). With initial condition $p(0,t) = p_0^*(x)$, its closed-form solution is given by the Bayes' density:

$$p(x,t) := \frac{p_0^*(x)\exp(-h(x)\,t)}{\int p_0^*(y)\exp(-h(y)\,t)\,\mathrm{d}y} =: p^*(x,t) \tag{2}$$

Consequently, as $t\to\infty$, one would expect the particles X_t^i to converge to the global minimizer of the function h, in the support of the prior p_0^*.

The use of probabilistic models to derive recursive sampling algorithms is by now a standard solution approach to the global optimization problem: The model (2) appears in [23] with closely related variants given in [20], [13]. Importance sampling type schemes based on these and more general stochastic models appear in [29], [17], [18], [21].

The central question for this paper is whether there is an optimal control interpretation for $-\nabla\phi$ term in (1a)? Such an interpretation for the optimization algorithm (1a)-(1b), and more generally for FPF, is of both theoretical and practical interest, e.g., in development of numerical algorithms.

The main contribution of this paper is to describe an optimal control construction for $-\nabla\phi$ (Theorems 1 and 2). The construction is based on the following steps: (i) the density transport (2) is shown to be a gradient flow (steepest descent) in the space of densities, with respect to the Kullback–Leibler divergence pseudo-metric; (ii) a Lagrangian is defined by casting the gradient flow as a dynamic programming recursion; and (iii) the resulting mean-field type optimal control problem is solved to obtain the particle filter (1a)-(1b). The particle filter is shown to be the Hamilton's equation where the Poisson equation is the first order optimality condition obtained from the application of the Pontryagin's minimum principle.

For the case where the objective function h is a quadratic function and the prior p_0^* is a Gaussian density, the Poisson equation (1b) admits a closed-form solution. The resulting control law is affine in the state. The quadratic Gaussian case is an example of the more general parametric case where the density is of a (known) parametrized form. For the parametric case, the filter is shown to be equivalent to the finite-dimensional natural gradient algorithm in the space of parameters.

The outline of the remainder of this paper is as follows: The optimal control construction appears in Sec. II. The quadratic Gaussian and the more general parametric cases are discussed in Sec. III. All the proofs are contained in the Appendix.

Notation: The Euclidean space $\mathbb{R}^d$ is equipped with the Borel σ-algebra denoted as $\mathscr{B}(\mathbb{R}^d)$. The space of Borel probability measures on $\mathbb{R}^d$ with finite second moment is denoted as $\mathscr{P}$:

$$\mathscr{P} \doteq \left\{\rho : \mathbb{R}^d \to [0,\infty)\,\text{meas. density} \,\middle|\, \int |x|^2\rho(x)\,\mathrm{d}x < \infty\right\}$$

The density for a Gaussian random variable with mean m and variance Σ is denoted as $\mathcal{N}(m,\Sigma)$. For vectors $x,y\in\mathbb{R}^d$, the dot product is denoted as $x\cdot y$ and $|x|:=\sqrt{x\cdot x}$; x^T denotes the transpose of the vector. Similarly, for a matrix K, K^T denotes the matrix transpose, and $K\succ 0$ denotes positive-definiteness. C^k is used to denote the space of k-times continuously differentiable functions on $\mathbb{R}^d$. For a function f, $\nabla f=\frac{\partial f}{\partial x_i}$ is used to denote the gradient vector, and $D^2 f=\frac{\partial^2 f}{\partial x_i \partial x_j}$ is used to denote the Hessian matrix. L^∞ denotes the space of bounded functions on $\mathbb{R}^d$ with associated norm denoted as $\|\cdot\|_\infty$. $L^2(\mathbb{R}^d;\rho)$ is the Hilbert space of square integrable functions on $\mathbb{R}^d$ equipped with the inner-product, $\langle\phi,\psi\rangle:=\int\phi(x)\psi(x)\rho(x)\,\mathrm{d}x$. The associated norm is denoted as $\|\phi\|_2^2:=\langle\phi,\phi\rangle$.

The following assumptions are made throughout the paper:

(i) **Assumption A1:** The prior probability density function $p_0^*\in\mathscr{P}$ and is of the form $p_0^*(x)=e^{-V_0(x)}$ where $V_0\in C^2$, $D^2V_0\in L^\infty$, and

$$\liminf_{|x|\to\infty}\ \nabla V_0(x)\cdot\frac{x}{|x|}=\infty$$

(ii) **Assumption A2:** The function $h\in C^2\cap L^2(\mathbb{R}^d;p_0^*)$ with $D^2h\in L^\infty$ and

$$\liminf_{|x|\to\infty}\ \nabla h(x)\cdot\frac{x}{|x|}>-\infty$$

(iii) **Assumption A3:** The function h has a unique minimizer $\bar{x}\in\mathbb{R}^d$ with minimum value $h(\bar{x})=:\bar{h}$. Outside some compact set $D\subset\mathbb{R}^d$, $\exists\, r>0$ such that

$$h(x)>\bar{h}+r \qquad \forall\, x\in\mathbb{R}^d\setminus D$$

Remark 1: Assumptions (A1) and (A2) are important to prove existence, uniqueness and regularity of the solutions of the Poisson equation [24], [16]. (A1) holds for density with Gaussian tails. Assumption (A3) is used to obtain weak convergence of $p^*(x,t)$ to Dirac delta at $\bar{x}$. The uniqueness of the minimizer $\bar{x}$ can be relaxed to obtain weaker conclusions on convergence (See Appendix C).

II. Variational Formulation

A variational formulation of the Bayes formula is the following *time-stepping procedure*: For the discrete-time sequence $\{t_0,t_1,t_2,\ldots,t_{\bar{N}}\}$ with $t_{\bar{N}}=T$ and increments $\Delta t_n:=t_n-t_{n-1}$, set $\rho_0=p_0^*\in\mathscr{P}$ and recursively define $\{\rho_n\}_{n=1}^{\bar{N}}\subset\mathscr{P}$ by taking $\rho_n\in\mathscr{P}$ to minimize the functional

$$\mathsf{I}(\rho|\rho_{n-1}):=\frac{1}{\Delta t_n}\mathsf{D}(\rho\mid\rho_{n-1})+\int h(x)\rho(x)\,\mathrm{d}x \tag{3}$$

where D denotes the relative entropy or Kullback–Leibler divergence,

$$\mathsf{D}(\rho\mid\rho_{n-1}):=\int\rho(x)\ln\Big(\frac{\rho(x)}{\rho_{n-1}(x)}\Big)\,\mathrm{d}x$$

The proof that $\rho_n=p^*(x,t_n)$ (defined in (2)) is the minimizer is straightforward: By Jensen's formula, $\mathsf{I}(\rho|\rho_{n-1})\geq -\ln(\int\rho_{n-1}(y)\exp(-h(y)\Delta t_n)\,\mathrm{d}y)$ with equality if and only if $\rho=\rho_n$.

To obtain an optimal control formulation, the key idea is to view the gradient flow time-stepping procedure as a dynamic programming recursion from time $t_{n-1}\to t_n$:

$$\rho_n=\operatorname*{argmin}_{\rho^{(u)}\in\mathscr{P}}\ \underbrace{\frac{1}{\Delta t_n}\mathsf{D}(\rho^{(u)}|\rho_{n-1})}_{\text{control cost}}+V(\rho^{(u)})$$

where $V(\rho^{(u)}):=\int\rho^{(u)}(x)h(x)\,\mathrm{d}x$ is the cost-to-go. The notation $\rho^{(u)}$ for density corresponds to the following construction: Consider the differential equation

$$\frac{\mathrm{d}X_t^i}{\mathrm{d}t}=u(X_t^i,t)$$

and denote the associated flow from $t_{n-1}\to t_n$ as $x\mapsto s_n(x)$. Under suitable assumptions on u (Lipschitz in x and continuous in t), the flow map s_n is a well-defined diffeomorphism on $\mathbb{R}^d$ and $\rho^{(u)}:=s_n^\#(\rho_{n-1})$, where $s_n^\#$ denotes the push-forward operator. The push-forward of a probability density ρ by a smooth map s is defined through the change-of-variables formula

$$\int f(x)[s^\#(\rho)](x)\,\mathrm{d}x=\int f(s(x))\rho(x)\,\mathrm{d}x$$

for all continuous and bounded test functions f.

Via a formal but straightforward calculation, in the asymptotic limit as $\Delta t_n\to 0$, the control cost is expressed in terms of the control u as

$$\frac{1}{\Delta t_n}\mathsf{D}(\rho^{(u)}|\rho_{n-1})=\frac{\Delta t_n}{2}\int\left|\frac{1}{\rho_{n-1}}\nabla\cdot(\rho_{n-1}u)\right|^2\rho_{n-1}\,\mathrm{d}x+o(\Delta t_n) \tag{4}$$

These considerations help motivate the following optimal control problem:

$$\begin{aligned}\text{Minimize:}\quad & J(u)=\int_0^T\mathsf{L}(\rho_t,u_t)\,\mathrm{d}t+\int h(x)\rho_T(x)\,\mathrm{d}x\\ \text{Constraint:}\quad & \frac{\partial\rho_t}{\partial t}+\nabla\cdot(\rho_t u_t)=0,\quad \rho_0(x)=p_0^*(x)\end{aligned} \tag{5}$$

(the minimization is over u) where the Lagrangian is defined as

$$\begin{aligned}\mathsf{L}(\rho,u):=&\ \frac{1}{2}\int_{\mathbb{R}^d}\left|\frac{1}{\rho(x)}\nabla\cdot(\rho(x)u(x))\right|^2\rho(x)\,\mathrm{d}x\\ &+\ \frac{1}{2}\int_{\mathbb{R}^d}|h(x)-\hat{h}|^2\,\rho(x)\,\mathrm{d}x\end{aligned}$$

where $\hat{h}:=\int h(x)\rho(x)\,\mathrm{d}x$.

The Hamiltonian is defined as

$$\mathsf{H}(\rho,\mathsf{q},u):=\mathsf{L}(\rho,u)-\int\mathsf{q}(x)\nabla\cdot(\rho(x)u(x))\,\mathrm{d}x \tag{6}$$

where q is referred to as the momentum.

Suppose $\rho\in\mathscr{P}$ is the density at time t. The value function is defined as

$$V(\rho,t):=\inf_u\left[\int_t^T\mathsf{L}(\rho_s,u_s)\,\mathrm{d}s\right] \tag{7}$$

The value function is a functional on the space of densities. For a fixed $\rho\in\mathscr{P}$ and time $t\in[0,T)$, the (Gâteaux) derivative of V is a function on $\mathbb{R}^d$, and an element of the function space $L^2(\mathbb{R}^d;\rho)$. This function is denoted as $\frac{\partial V}{\partial\rho}(\rho,t)(x)$ for

$x \in \mathbb{R}^d$. Additional details appear in the Appendix A where the following Theorem is proved.

Theorem 1 (Finite-horizon optimal control): Consider the optimal control problem (5) with the value function defined in (7). Then V solves the following DP equation:

$$\frac{\partial V}{\partial t}(\rho,t) + \inf_{u\in L^2} \mathsf{H}(\rho, \frac{\partial V}{\partial \rho}(\rho,t),u) = 0, \quad t \in [0,T)$$
$$V(\rho,T) = \int h(x)\rho(x)\,\mathrm{d}x$$

The solution of the DP equation is given by

$$V(\rho,t) = \int_{\mathbb{R}^d} h(x)\rho(x)\,\mathrm{d}x$$

and the associated optimal control is a solution of the following pde:

$$\frac{1}{\rho(x)}\nabla\cdot(\rho(x)u(x)) = (h(x)-\hat{h}), \quad \forall\, x\in\mathbb{R}^d \tag{8}$$

It is also useful to consider the following infinite-horizon version of the optimal control problem:

$$\begin{aligned} &\text{Minimize:} \quad J(u) = \int_0^\infty \mathsf{L}(\rho_t,u_t)\,\mathrm{d}t \\ &\text{Constraints:} \quad \begin{cases} \frac{\partial \rho_t}{\partial t} + \nabla\cdot(\rho_t u_t) = 0, \quad \rho_0(x) = p_0^*(x) \\ \lim\limits_{t\to\infty} \int h(x)\rho_t(x) = h(\bar{x}) \end{cases} \end{aligned} \tag{9}$$

For this problem, the value function is defined as

$$V(\rho) = \inf_u J(u) \tag{10}$$

The solution is given by the following Theorem whose proof appears in Appendix A:

Theorem 2 (Infinite-horizon optimal control): Consider the infinite horizon optimal control problem (9) with the value function defined in (10). The value function is given by

$$V(\rho) = \int_{\mathbb{R}^d} h(x)\rho(x)\,\mathrm{d}x - h(\bar{x})$$

and the associated optimal control law is a solution of the pde (8).

The particle filter algorithm (1a)-(1b) in Sec. I is obtained by *additionally* requiring the solution u of (8) to be of the gradient form. One of the advantages of doing so is that the optimizing control law, obtained instead as solution of (1b), is uniquely defined [16]. In part, this choice is guided by the L^2 optimality of the gradient form solution (The proof is omitted on account of space):

Lemma 1 (L^2 optimality): Consider the pde (8) where ρ and h satisfy Assumptions (A1)-(A2). The general solution is given by

$$u = -\nabla\phi + v$$

where ϕ is the solution of (1b), v solves $\nabla\cdot(\rho v) = 0$, and

$$\|u\|_2^2 = \|\nabla\phi\|_2^2 + \|v\|_2^2$$

That is, $u = -\nabla\phi$ is the minimum L^2-norm solution of (8).

Remark 2: In Appendix B, the Pontryagin's minimum principle of optimal control is used to express the particle filter (1a)-(1b) in its Hamilton's form:

$$\boxed{\begin{aligned} \frac{\mathrm{d}X_t^i}{\mathrm{d}t} &= u(X_t^i,t), \quad X_0^i \sim p_0^* \\ 0 &\equiv \mathsf{H}(p(\cdot,t),h,u(\cdot,t)) = \min_{v\in L^2} \mathsf{H}(p(\cdot,t),h,v) \end{aligned}}$$

The Poisson equation (1b) is simply the first order optimality condition to obtain a minimizing control. Under this optimal control, the density $p(x,t)$ is the optimal trajectory. The associated optimal trajectory for the momentum (co-state) is a constant equal to its terminal value $h(x)$.

The following theorem shows that the particle filter implements the Bayes' transport of the density, and establishes the asymptotic convergence for the density (The proof appears in the Appendix (C)). We recall the notation for the two types of density in our analysis:

1) $p(x,t)$: Defines the density of X_t^i.
2) $p^*(x,t)$: The Bayes' density given by (2).

Theorem 3 (Bayes' exactness and convergence): Consider the ODE-PDE model (1a)-(1b). If $p(\cdot,0) = p^*(\cdot,0)$, we have for all $t \geq 0$,

$$p(\cdot,t) = p^*(\cdot,t)$$

As $t\to\infty$, $\int h(x)p(x,t)\,\mathrm{d}x$ decreases monotonically to $h(\bar{x})$ and $X_t^i \to \bar{x}$ in probability.

The hard part of implementing the particle filter is solving the Poisson equation (1b). For the quadratic Gaussian case – where the objective function h is quadratic and the prior p_0^* is Gaussian – the solution can be obtained in an explicit form. This is the subject of the Sec. III-A. The more general parametric case is the subject of Sec. III-B.

III. Parametric Cases

A. Quadratic Gaussian case

For the quadratic Gaussian problem, the solution of the Poisson equation can be obtained in an explicit form as described in the following Lemma. The proof appears in the Appendix D.

Lemma 2: Consider the Poisson equation (1b). Suppose the objective function h is a quadratic function such that $h(x)\to\infty$ as $|x|\to\infty$ and the density ρ is a Gaussian with mean m and variance Σ. Then the control function

$$u(x) = -\nabla\phi(x) = -\mathsf{K}(x-m) - b \tag{11}$$

where the affine constant vector

$$b = \int x(h(x)-\hat{h})\rho(x)\,\mathrm{d}x \tag{12}$$

and the gain matrix $\mathsf{K} = \mathsf{K}^T \succ 0$ is the solution of the Lyapunov equation:

$$\Sigma\mathsf{K} + \mathsf{K}\Sigma = \int (x-m)(x-m)^T(h(x)-\hat{h})\rho(x)\,\mathrm{d}x \tag{13}$$

Using an affine control law (11), it is straightforward to verify that $p(x,t) = p^*(x,t)$ is a Gaussian whose mean $m_t \to \bar{x}$ and variance $\Sigma_t \to 0$. The proofs of the following Proposition and the Corollary appear in the Appendix D:

Proposition 1: Consider the particle filter (1a) with the affine control law (11). Suppose the objective function h is a quadratic function such that $h(x) \to \infty$ as $|x| \to \infty$ and the prior density p_0^* is a Gaussian with mean m_0 and variance Σ_0. Then the posterior density p is a Gaussian whose mean m_t and variance Σ_t evolve according to

$$\begin{aligned} \frac{\mathrm{d}m_t}{\mathrm{d}t} &= -\mathsf{E}\left[X_t^i(h(X_t^i) - \hat{h}_t)\right] \\ \frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} &= -\mathsf{E}\left[(X_t^i - m_t)(X_t^i - m_t)^T(h(X_t^i) - \hat{h}_t)\right] \end{aligned} \tag{14}$$

where $\hat{h}_t := \mathsf{E}[h(X_t^i)]$.

Corollary 1: Under the hypothesis of Proposition 1, with an explicit form for quadratic objective function $h(x) = \frac{1}{2}(x - \bar{x})^T H(x - \bar{x}) + c$ where $H = H^T \succ 0$, the expectations on the righthand-side of (14) are computed in closed-form and the resulting evolution is given by

$$\frac{\mathrm{d}m_t}{\mathrm{d}t} = \Sigma_t H(\bar{x} - m_t) \tag{15a}$$

$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = -\Sigma_t H \Sigma_t \tag{15b}$$

whose explicit solution is given by

$$\begin{aligned} m_t &= m_0 + \Sigma_0 S_t^{-1}(\bar{x} - m_0) \\ \Sigma_t &= \Sigma_0 - \Sigma_0 S_t^{-1}\Sigma_0 \end{aligned}$$

where $S_t := \frac{1}{t}H^{-1} + \Sigma_0$. In particular, $m_t \to \bar{x}$ and $\Sigma_t \to 0$.

In practice, the affine control law (11) is implemented as:

$$\frac{\mathrm{d}X_t^i}{\mathrm{d}t} = -\mathsf{K}_t^{(N)}(X_t^i - m_t^{(N)}) - b_t^{(N)} =: u_t^i \tag{16}$$

where the terms are approximated empirically from the ensemble $\{X_t^i\}_{i=1}^N$. The algorithm appears in Table 1 (the dependence on time t is suppressed).

As $N \to \infty$, the approximations become exact and (11) represents the mean-field limit of the finite-N control in (16). Consequently, the empirical distribution of the ensemble approximates the posterior distribution (density) $p^*(x,t)$.

Remark 3: The finite-dimensional system (14) is the optimization counterpart of the Kalman filter. Likewise the particle filter (16) is the counterpart of the ensemble Kalman filter. While the affine control law (11) is optimal for the quadratic Gaussian case, it can be implemented in the more general non-quadratic non-Gaussian settings – as long as the various approximations can be obtained at each step. The situation is analogous to the filtering setup where the Kalman filter is often used as an approximate algorithm even in nonlinear non-Gaussian settings.

General purpose numerical algorithms for approximating the solution of Poisson equation appear in our earlier papers [22]. These algorithms have been applied to several benchmark global optimization problems. These numerical results, including a comparison with other state-of-the-art algorithms, appear in the PhD thesis of the first author [28].

Remark 4: The particle filter (16) is related to the cross-entropy (CE) and the model reference adaptive search (MRAS) algorithms described in [20], [13]. In these algorithms, one samples recursively from a Gaussian density whose parameters are updated at each time-step via density projection. The update formulae in (14) are the counterparts for similar formulae obtained using the density projection in CE and MRAS. The particle filter (16) is simpler, or at least more direct, because (i) explicit update for the parameters is not necessary, (ii) recursive sampling from the Gaussian density is avoided, and (iii) there is no need to calculate the importance weights.

Algorithm 1 Affine approximation of the control function

Input: $\{X^i\}_{i=1}^N$, $\{h(X^i)\}_{i=1}^N$
Output: $\{u^i\}_{i=1}^N$

1: Calculate $m^{(N)} := \frac{1}{N}\sum_{i=1}^N X^i$,
2: Calculate $\Sigma^{(N)} := \frac{1}{N}\sum_{i=1}^N \left(X^i - m^{(N)}\right)\left(X^i - m^{(N)}\right)^T$
3: Calculate $\hat{h}^{(N)} := \frac{1}{N}\sum_{i=1}^N h(X^i)$
4: Calculate $b^{(N)} := \frac{1}{N}\sum_{i=1}^N X^i \left(h(X^i) - \hat{h}^{(N)}\right)$
5: Calculate

$$C^{(N)} := \frac{1}{N}\sum_{i=1}^N \left(X^i - m^{(N)}\right)\left(X^i - m^{(N)}\right)^T\left(h(X^i) - \hat{h}^{(N)}\right)$$

6: Calculate $\mathsf{K}^{(N)}$ by solving $\Sigma^{(N)}\mathsf{K}^{(N)} + \mathsf{K}^{(N)}\Sigma^{(N)} = C^{(N)}$
7: Calculate $u^i = -\mathsf{K}^{(N)}(X^i - m^{(N)}) - b^{(N)}$

B. Natural Gradient Algorithm

Consider next the case where the density has a known parametric form

$$p(x,t) = \varrho\,(x;\theta_t) \tag{17}$$

where $\theta_t \in \mathbb{R}^M$ is the parameter vector. For example, in the quadratic Gaussian problem, ϱ is a Gaussian with parameters m_t and Σ_t.

For the parametric density, $\frac{\partial}{\partial\vartheta}(\log\varrho\,(x;\vartheta))$ is a $M \times 1$ column vector whose k^{th} entry

$$\left[\frac{\partial}{\partial\vartheta}(\log\varrho\,(x;\vartheta))\right]_k = \frac{\partial}{\partial\vartheta_k}(\log\varrho\,(x;\vartheta))$$

for $k = 1,\ldots,M$.

The Fisher information matrix is a $M \times M$ matrix:

$$G_{(\vartheta)} := \int \frac{\partial}{\partial\vartheta}(\log\varrho\,(x;\vartheta))\left[\frac{\partial}{\partial\vartheta}(\log\varrho\,(x;\vartheta))\right]^T \varrho\,(x;\vartheta)\,\mathrm{d}x \tag{18}$$

By construction, $G_{(\vartheta)}$ is symmetric and positive semidefinite. In the following, it is furthermore assumed that $G_{(\vartheta)}$ is strictly positive definite, and thus invertible, for all $\vartheta \in \mathbb{R}^M$.

In terms of the parameter,

$$e(\vartheta) := \int h(x)\,\varrho\,(x;\vartheta)\,\mathrm{d}x$$

and its gradient is a $M \times 1$ column vector:

$$\nabla e(\vartheta) = \int h(x) \frac{\partial}{\partial \vartheta} (\log \varrho(x;\vartheta)) \, \varrho(x;\vartheta) \, \mathrm{d}x \tag{19}$$

We are now prepared to describe the induced evolution for the parameter vector θ_t. The proof of the following proposition appears in the Appendix E.

Proposition 2: Consider the particle filter (1a)-(1b). Suppose the density admits the parametric form (17) whose Fisher information matrix, defined in (18), is assumed to be invertible. Then the parameter vector θ_t is a solution of the following ordinary differential equation,

$$\frac{\mathrm{d}\theta_t}{\mathrm{d}t} = -G^{-1}_{(\theta_t)} \nabla e(\theta_t) \tag{20}$$

Remark 5: The filter (20) is the well known *natural gradient* algorithm; cf., [1], [11]. It has several variational interpretations:

(i) The filter can be obtained via a time stepping procedure, analogous to (3). The sequence $\{\theta_n\}_{n=1}^N$ is inductively defined as a minimizer of the function

$$\mathsf{I}(\theta|\theta_{n-1}) := \frac{1}{\Delta t_n} \mathsf{D}(\varrho(\cdot;\theta) | \, \varrho(\cdot;\theta_{n-1})) + e(\theta)$$

On taking the limit as $\Delta t_n \to 0$, one arrives at the filter (20).

(ii) The optimal control interpretation of (20) is based on the Pontryagin's minimum principle (see also Remark 2). For the finite-dimensional problem, the Hamiltonian

$$\mathsf{H}(\theta, \mathsf{q}, u) = \mathsf{L}(\theta, u) + \mathsf{q} \cdot u$$

where $\mathsf{q} \in \mathbb{R}^M$ is the momentum. With $\dot{\theta} = u$, the counterpart of (4) is

$$\frac{1}{\Delta t_n} \mathsf{D}(\varrho^{(u)}(\cdot;\theta) | \, \varrho(\cdot;\theta_{n-1})) = \frac{1}{2} u^T G_{(\theta)} u + o(\Delta t_n)$$

With $\frac{1}{2} \, u^T G_{(\theta)} u$ as the control cost component in the Lagrangian, the first order optimality condition gives

$$G_{(\theta)} u = -\mathsf{q} = -\nabla e(\theta)$$

where we have used the fact that $e(\theta)$ is the value function. Note that it was not necessary to write the explicit form of the Lagrangian to obtain the optimal control.

(iii) Finally, the filter (20) represents the gradient flow (in $\mathbb{R}^M$) for the objective function $e(\theta)$ with respect to the Riemannian metric $\langle v, w \rangle_\theta = v^T G_{(\theta)} w$ for all $v, w \in \mathbb{R}^M$.

Example 1: In the quadratic Gaussian case, the natural gradient algorithm (20) with parameters m_t and Σ_t reduces to (14).

Remark 6: While the systems (20) and (14) are finite-dimensional, the righthand-sides will still need to be approximated empirically. The convergence properties of a class of related algorithms is studied using a stochastic approximation framework in [14].

References

[1] S-I. Amari. Natural gradient works efficiently in learning. *Neural Comput.*, 10(2):251–276, 1998.

[2] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*, volume 348. Springer Science & Business Media, 2013.

[3] A. Bensoussan, J. Frehse, and S. C. P. Yam. The master equation in mean field theory. *Journal de Mathématiques Pures et Appliquées*, 103(6):1441–1474, 2015.

[4] R. Brockett. Optimal control of the Liouville equation. *AMS IP Studies in Advanced Mathematics*, 39:23, 2007.

[5] Y. Chen, T. Georgiou, and M. Pavon. On the relation between optimal transport and Schródinger bridges: A stochastic control viewpoint. *J. Optimiz. Theory App.*, 169(2):671–691, 2016.

[6] Y. Chen, T. Georgiou, and M. Pavon. Optimal steering of a linear stochastic system to a final probability distribution, part I. *IEEE Trans. Autom. Control*, 61(5):1158–1169, 2016.

[7] F. Daum and J. Huang. Exact particle flow for nonlinear filters. In *Proc. SPIE 7698*, pages 76980I–1, 2010.

[8] F. Daum, J. Huang, and A. Noushin. Gromov's method for Bayesian stochastic particle flow: A simple exact formula for Q. In *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intel. Syst. (MFI)*, pages 540–545, 2016.

[9] G. E. Dullerud and F. Paganini. *A Course in Robust Control Theory: A Convex Approach*, volume 36. Springer, 2013.

[10] R. Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2010.

[11] M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. S. B*, 73(2):123–214, 2011.

[12] J. Heng, A. Doucet, and Y. Pokern. Gibbs flow for approximate transport with applications to Bayesian computation. arXiv preprint: 1509.08787.

[13] J. Hu, M. C. Fu, and S. I. Marcus. A model reference adaptive search method for global optimization. *Oper. Res.*, 55(3):549–568, 2007.

[14] J. Hu and P. Hu. A stochastic approximation framework for a class of randomized optimization algorithms. *IEEE Trans. Autom. Control*, 57(1):165–178, 2012.

[15] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria. *IEEE Trans. Autom. Control*, 52(9):1560–1571, 2007.

[16] R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson's equation in nonlinear filtering. *SIAM J. Control Optim.*, 53(1):501–525, 2015.

[17] B. Liu. Posterior exploration based sequential Monte Carlo for global optimization. *arXiv preprint:1509.08870*, 2016.

[18] B. Liu, S. Cheng, and Y. Shi. Particle filter optimization: A brief introduction. In *Proc. 7th Int. Conf. Swarm Intel.*, pages 95–104, 2016.

[19] S. Reich. A dynamical systems framework for intermittent data assimilation. *BIT Numerical Mathematics*, 51(1):235–249, 2011.

[20] R. Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodol. Comput. Appl.*, 1(2):127–190, 1999.

[21] P. Stinis. Stochastic global optimization as a filtering problem. *J. Comput. Phys.*, 231(4):2002–2014, 2012.

[22] A. Taghvaei and P. G Mehta. Gain function approximation in the feedback particle filter. In *Proc. 55th IEEE Conf. Decision Control*, pages 5446–5452. IEEE, 2016.

[23] Y. Wang and M. C. Fu. Model-based evolutionary optimization. In *Proc. 2010 Winter Simul. Conf.*, pages 1199–1210, 2010.

[24] T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter. *Automatica*, 71(9):10–23, 2016.

[25] T. Yang, P. G. Mehta, and S. P. Meyn. Feedback particle filter. *IEEE Trans. Autom. Control*, 58(10):2465–2480, 2013.

[26] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Synchronization of coupled oscillators is a game. *IEEE Trans. Autom. Control*, 57(4):920–935, 2012.

[27] C. Zhang. A particle system for global optimization. In *Proc. 52nd IEEE Conf. Decision Control*, pages 1714–1719, 2013.

[28] C. Zhang. *Controlled Particle Systems for Nonlinear Filtering and Global Optimization*. PhD thesis, University of Illinois, Urbana-Champaign, May 2017.

[29] E. Zhou, M. C. Fu, and S. I. Marcus. Particle filtering framework for a class of randomized optimization algorithms. *IEEE Trans. Autom. Control*, 59(4):1025–1030, 2014.

Appendix

A. Optimal control

Preliminaries: Consider a functional $E : \mathscr{P} \to \mathbb{R}$ mapping densities to real numbers. For a fixed $\rho \in \mathscr{P}$, the (Gâteaux) derivative of E is a real-valued function on $\mathbb{R}^d$, and an element of the function space $L^2(\mathbb{R}^d;\rho)$. This function is denoted as $\frac{\partial E}{\partial \rho}(\rho,t)(x)$ for $x \in \mathbb{R}^d$, and defined as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}E(\rho_t)\Big|_{t=0} = -\int_{\mathbb{R}^d} \frac{\partial E}{\partial \rho}(\rho)(x)\nabla\cdot(\rho(x)u(x))\,\mathrm{d}x$$

where ρ_t is a path in $\mathscr{P}$ such that $\frac{\partial \rho_t}{\partial t} = -\nabla\cdot(\rho_t u)$ with $\rho_0 = \rho$, and u is any arbitrary vector-field on $\mathbb{R}^d$. Similarly, $\frac{\partial^2 E}{\partial \rho^2}(\rho) \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$ is the second (Gâteaux) derivative of the functional E if

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial E}{\partial \rho}(\rho_t)(x)\Big|_{t=0} = -\int_{\mathbb{R}^d} \frac{\partial^2 E}{\partial \rho^2}(\rho)(x,y)\nabla\cdot(\rho(y)u(y))\,\mathrm{d}y$$

The optimal control problems (5) and (9) are examples of the mean-field type control problem introduced in [3]. The notation and methodology for the following proofs are based in part on [3].

Proof of Theorem 1: The value function $V(\rho,t)$, defined in (7), is the solution of the DP equation:

$$\begin{aligned}&\frac{\partial V}{\partial t}(\rho,t) + \inf_{u\in L^2} \mathsf{H}(\rho, \frac{\partial V}{\partial \rho}(\rho,t),u) = 0, \quad t \in [0,T)\\ &V(\rho,T) = \int h(x)\rho(x)\,\mathrm{d}x\end{aligned} \tag{21}$$

In the following, we use the notation

$$\Theta = \Theta(\rho,t)(x) := \frac{\partial V}{\partial \rho}(\rho,t)(x)$$

For a fixed $\rho \in \mathscr{P}$ and $t \in [0,T)$, Θ is a function on $\mathbb{R}^d$.

A necessary condition is obtained by considering the first variation of H. Suppose u is a minimizing control function. Then u satisfies the first order optimality condition:

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathsf{H}(\rho,\Theta,u+\varepsilon v)\Big|_{\varepsilon=0} = 0$$

where v is an arbitrary vector field on $\mathbb{R}^d$. Explicitly,

$$\int \nabla(-\frac{1}{\rho}\nabla\cdot(\rho u) + \Theta)\cdot v\,\rho\,\mathrm{d}x = 0$$

or in its strong form

$$-\frac{1}{\rho}\nabla\cdot(\rho u) + \Theta = (\text{constant})$$

Multiplying both sides by ρ and integrating yields the value of the constant as $\int \Theta(\rho,t)(x)\rho(x)\,\mathrm{d}x =: \hat{\Theta}(\rho,t)$. Therefore, the minimizing control solves the pde

$$\frac{1}{\rho}\nabla\cdot(\rho u) = \Theta - \hat{\Theta}$$

On substituting the optimal control law into the DP equation (21), the HJB equation for the V is given by

$$\begin{aligned}&\frac{\partial V}{\partial t}(\rho,t) + \frac{1}{2}\int |h-\hat{h}|^2\rho\,\mathrm{d}x\\ &\qquad - \frac{1}{2}\int |\Theta(\rho,t) - \hat{\Theta}(\rho,t)|^2\rho\,\mathrm{d}x = 0, \quad t\in[0,T)\\ &V(\rho,T) = \int h(x)\rho(x)\,\mathrm{d}x\end{aligned}$$

The equation involves both V and Θ. One obtains the so-called master equation (see [3]) involving only Θ by differentiating with respect to ρ

$$\begin{aligned}&\frac{\partial \Theta}{\partial t}(\rho,t)(x) + \frac{1}{2}|h(x)-\hat{h}|^2 - \frac{1}{2}|\Theta(\rho,t)(x) - \hat{\Theta}(\rho,t)|^2\\ &- \int (\Theta(\rho,t)(y) - \hat{\Theta}(\rho,t))\frac{\partial \Theta}{\partial \rho}(\rho,t)(y,x)\rho(y)\,\mathrm{d}y = 0,\ t\in[0,T)\\ &\Theta(\rho,T) = h\end{aligned}$$

It is easily verified that $\Theta(\rho,t) = h$ solves the master equation. The corresponding value function $V(\rho,t) = \int h\rho\,\mathrm{d}x$.

Sufficiency: The proof that the proposed control law is a minimizer is as follows. Consider any arbitrary control law v_t with the resulting density ρ_t. Taking the time derivative of $-\int h\rho_t\,\mathrm{d}x$:

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}t}\int h\rho_t\,\mathrm{d}x &= \int (h-\hat{h}_t)\,(\frac{1}{\rho_t}\nabla\cdot(\rho_t v_t))\,\rho_t\,\mathrm{d}x\\ &\le \int \left(\frac{1}{2}|h-\hat{h}|^2 + \frac{1}{2}\left|\frac{1}{\rho_t}\nabla\cdot(\rho_t v_t)\right|^2\right)\rho_t\,\mathrm{d}x\\ &= \mathsf{L}(\rho_t,v_t)\end{aligned}$$

On integrating both sides with respect to time,

$$\int_{\mathbb{R}^d} h\rho_0\,\mathrm{d}x \;\le\; \int_0^T \mathsf{L}(\rho_t,v_t)\,\mathrm{d}t \;+\; \int_{\mathbb{R}^d} h\rho_T\,\mathrm{d}x$$

where the equality holds with $v_t = u_t$ (defined as solution of (8)). Therefore,

$$J(u) = \int h\rho_0\,\mathrm{d}x \le J(v)$$

This also confirms that $V(\rho,t) = \int h\rho\,\mathrm{d}x$ is the value function, and completes the proof of Theorem 1. □

The analysis for the infinite horizon optimal control problem (9) is similar and described next.

Proof of Theorem 2: The infinite-horizon value function $V^\infty(\rho) := \inf_u \int_0^\infty \mathsf{L}(\rho_t,u_t)\,\mathrm{d}t$ is a solution of the DP equation:

$$\inf_{u\in L^2} \mathsf{H}(\rho,\Theta^\infty(\rho),u) = 0 \tag{22}$$

where $\Theta^\infty(\rho) := \frac{\partial V^\infty}{\partial \rho}(\rho)$. By carrying out the first order analysis in an identical manner, it is readily verified that:

(i) A minimizing control u is a solution of the pde (8);
(ii) $V^\infty(\rho) = \int h\rho\,\mathrm{d}x - h(\bar{x})$ is a solution of the DP equation (22).

The sufficiency also follows similarly. With any arbitrary control v_t,

$$\int_{\mathbb{R}^d} h\rho_0\,\mathrm{d}x \le \int_0^\infty \mathsf{L}(\rho_t,v_t)\,\mathrm{d}t + \limsup_{t\to\infty}\int_{\mathbb{R}^d} h\rho_t\,\mathrm{d}x$$

with equality if $v_t = u_t$ solves the pde (8). Using the boundary condition, $\limsup_{t\to\infty} \int h\rho_t \, dx = h(\bar{x})$,

$$J(u) = \int h\rho_0 \, dx - h(\bar{x}) \le J(v)$$

$\square$

B. Hamiltonian formulation

The Hamiltonian H is defined in (6). Suppose u_t is the optimal control and ρ_t is the corresponding optimal trajectory. Denote the trajectory for the co-state (momentum) as q_t. Using the Pontryagin's minimum principle, (ρ_t, q_t) satisfy the following Hamilton's equations:

$$\frac{\partial \rho_t}{\partial t} = \frac{\partial \mathsf{H}}{\partial \rho}(\rho_t, \mathsf{q}_t, u_t), \quad \rho_0 = p_0^*$$
$$\frac{\partial \mathsf{q}_t}{\partial t} = -\frac{\partial \mathsf{H}}{\partial \mathsf{q}}(\rho_t, \mathsf{q}_t, u_t), \quad \mathsf{q}_T = \frac{\partial}{\partial \rho}\left(\int h(x)\rho(x)\, dx\right)$$
$$0 = \mathsf{H}(\rho_t, \mathsf{q}_t, u_t) = \min_{v\in L^2} \mathsf{H}(\rho_t, \mathsf{q}_t, v)$$

The calculus of variation argument in the proof of Theorem 1 shows that the minimizing control u_t solves the first order optimality equation

$$\frac{1}{\rho_t}\nabla\cdot(\rho_t u_t) = \mathsf{q}_t - \hat{\mathsf{q}}_t \tag{23}$$

where $\hat{\mathsf{q}}_t := \int \mathsf{q}_t(x)\rho_t(x)\, dx$.

The explicit form of the Hamilton's equations are obtained by explicitly evaluating the derivatives along the optimal trajectory:

$$\frac{\partial \mathsf{H}}{\partial \rho}(\rho_t, \mathsf{q}_t, u_t) = -\nabla\cdot(\rho_t u_t)$$
$$\frac{\partial \mathsf{H}}{\partial \mathsf{q}}(\rho_t, \mathsf{q}_t, u_t) = \frac{1}{2}|h - \hat{h}_t|^2 - \frac{1}{2}|\mathsf{q}_t - \hat{\mathsf{q}}_t|^2$$

It is easy to verify that $\mathsf{q}_t \equiv h(x)$ satisfies both the boundary condition and the evolution equation for the momentum. This results in a simpler form of the Hamilton's equations:

$$\frac{\partial \rho_t}{\partial t} = -\nabla\cdot(\rho_t u_t)$$
$$0 = \mathsf{H}(\rho_t, h, u_t) = \min_{v\in L^2} \mathsf{H}(\rho_t, h, v)$$

In a particle filter implementation, the minimizing control $u_t = -\nabla\phi$ is obtained by solving the first order optimality equation (23) with $\mathsf{q}_t = h$.

C. Bayes' exactness and convergence

Before proving the Theorem 3, we state and prove the following technical Lemma:

Lemma 3: Suppose the prior density $p_0^*(x)$ satisfies Assumption (A1) and the objective function $h(x)$ satisfies Assumption (A2). Then for each fixed time $t \ge 0$:

(i) The posterior density $p^*(x,t)$, defined according to (2), admits a spectral bound;

(ii) The objective function $h \in L^2(\mathbb{R}^d; p^*(\cdot,t))$.

Proof 1: Define $V_t(x) := -\log p^*(x,t) = V_0(x) + th(x) + \gamma_t$ where $\gamma_t := \log(\int e^{-V_0(y) - th(y)}\, dy)$. It is directly verified that $V_t \in C^2$ with $D^2 V_t \in L^\infty$ and $\liminf_{|x|\to\infty} \nabla V_t(x)\cdot \frac{x}{|x|} = \infty$. Therefore, the density $p^*(x,t)$ admits a spectral bound [Thm 4.6.3 in [2]]. The function h is square-integrable because

$$\int |h(x)|^2 p^*(x,t)\, dx \le e^{-t\bar{h} - \gamma_t} \int |h(x)|^2 e^{-V_0(x)}\, dx < \infty$$

$\square$

Proof of Theorem 3: Given any C^1 smooth and compactly supported (test) function f, using the elementary chain rule,

$$df(X_t^i) = -\nabla\phi(X_t^i)\cdot\nabla f(X_t^i)$$

On integrating and taking expectations,

$$\mathsf{E}[f(X_t^i)] = \mathsf{E}[f(X_0^i)] - \int_0^t \mathsf{E}[\nabla\phi(X_s^i)\cdot\nabla f(X_s^i)]\, ds$$
$$= \mathsf{E}[f(X_0^i)] - \int_0^t \mathsf{E}[(h(X_s^i) - \hat{h}_s)f(X_s^i)]\, ds$$

which is the weak form of the replicator pde. Note that the weak form of the Poisson equation is used to obtain the second equality. Since the test function f is arbitrary, the evolution of p and p^* are identical. That the control function is well-defined for each time follows from [16] based on apriori estimates in Lemma 3 for $p^* = p$.

The convergence proof is presented next. The proof here is somewhat more general than needed to prove the Theorem. For a function h, we define the minimizing set:

$$A_0 := \{x \in \mathbb{R}^d \mid h(x) = \bar{h}\}$$

where it is recalled that $\bar{h} = \inf_{x\in\mathbb{R}^d} h(x)$. In the following it is shown that for *any* open neighborhood U of A_0,

$$\liminf_{t\to\infty} \int_U p(x,t)\, dx = 1 \tag{24}$$

It then follows that X_t^i converges in distribution where the limiting distribution is supported on A_0 (Thm. 3.2.5 in [10]). If the minimizer is unique (i.e., $A_0 = \{\bar{x}\}$), X_t^i converges to $\bar{x}$ in probability.

The key to prove the convergence is the following property of the function h:

(P1): For each $\delta > 0$, $\exists\, \varepsilon > 0$ such that:

$$|h(x) - \bar{h}| \le \varepsilon \quad \Rightarrow \quad \text{dist}(x, A_0) \le \delta \quad \forall x \in \mathbb{R}^d$$

where $\text{dist}(x,A_0)$ denotes the Euclidean distance of point x from set A_0. If the minimizer $\bar{x}$ is unique, it equals $|x - \bar{x}|$.

Any lower semi-continuous function satisfying Assumption (A3) also satisfies the property (P1): Suppose $\{x_n\}$ is a sequence such that $h(x_n) \to \bar{h}$. Then $\{x_n\}$ is compact because $h(x) > \bar{h} + r$ outside some compact set. Therefore, the limit set is non-empty and because h is lower semi-continuous, for any limit point z, $\bar{h} \le h(z) \le \liminf_{x_n\to z} h(x_n) = \bar{h}$. That is, $z \in A_0$.

The proof for (24) is based on construction of a Lyapunov function: Denote $A_\varepsilon := \{x \in \mathbb{R}^d \mid h(x) \le \bar{h} + \varepsilon\}$ where $\varepsilon > 0$. By property (P1), given any open neighborhood U containing A_0, $\exists\, \varepsilon > 0$ such that $A_\varepsilon \subset U$. A candidate Lyapunov function $V_{A_\varepsilon}(\mu) := -\log(\mu(A_\varepsilon))$ is defined for measure μ with everywhere positive density. By construction $V_{A_\varepsilon}(\mu) \geq 0$ with equality iff $\mu(A_\varepsilon) = 1$.

Let μ_t be the probability measure associated with $p(x,t)$, i.e, $\mu_t(B) := \int_B p(x,t)\,\mathrm{d}x$ for all Borel measurable set $B \subset \mathbb{R}^d$. Since $p(x,t)$ is a solution of the replicator pde,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} V_{A_\varepsilon}(\mu_t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left[-\log(\mu_t(A_\varepsilon))\right] \\ &= \frac{1}{\mu_t(A_\varepsilon)} \int_{A_\varepsilon} (h(x) - \hat{h}_t)\,\mathrm{d}\mu_t(x) \\ &= (1 - \mu_t(A_\varepsilon)) \left(\frac{\int_{A_\varepsilon} h\,\mathrm{d}\mu_t}{\mu_t(A_\varepsilon)} - \frac{\int_{A_\varepsilon^c} h\,\mathrm{d}\mu_t}{\mu_t(A_\varepsilon^c)} \right) \leq 0\end{aligned}$$

with equality iff $\mu_t(A_\varepsilon) = 1$.

For the objective function h, a direct calculation also shows:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} \int h(x) p(x,t)\,\mathrm{d}x &= -\int h(x)(h(x) - \hat{h}_t) p(x,t)\,\mathrm{d}x \\ &= -\int (h(x) - \hat{h})^2 p(x,t)\,\mathrm{d}x \leq 0\end{aligned}$$

with equality iff $h = \hat{h}$ almost everywhere (with respect to the measure μ_t). □

D. Quadratic Gaussian case

Proof of Lemma 2: We are interested in obtaining an explicit solution of the Poisson equation,

$$-\nabla \cdot (\rho(x) \nabla \phi(x)) = (h(x) - \hat{h})\rho(x) \tag{25}$$

Consider the solution ansatz:

$$\nabla \phi(x) = \mathsf{K}(x - m) + b \tag{26}$$

where the matrix $\mathsf{K} = \mathsf{K}^T \in \mathbb{R}^{d\times d}$ and the vector $b \in \mathbb{R}^d$ are determined as follows:

(i) Multiply both sides of (25) by vector x and integrate (element-by-element) by parts to obtain

$$b = \int x(h(x) - \hat{h})\rho(x)\,\mathrm{d}x \tag{27}$$

(ii) Multiply both sides of (25) by matrix $(x - m)(x - m)^T$ and integrate by parts to obtain

$$\Sigma\mathsf{K} + \mathsf{K}\Sigma = \int (x - m)(x - m)^T (h(x) - \hat{h})\rho(x)\,\mathrm{d}x \tag{28}$$

We have thus far not used the fact that the density ρ is Gaussian and the function h is quadratic. In the following, it is shown that the solution thus defined in fact *solves* the pde (25) under these conditions.

A radially unbounded quadratic function is of the general form:

$$h(x) = \frac{1}{2}(x - \bar{x})^T H (x - \bar{x}) + c$$

where the matrix $H = H^T \succ 0$ and c is some constant. For a Gaussian density ρ with mean m and variance $\Sigma \succ 0$, the integrals (27) and (28) are explicitly evaluated to obtain

$$b = \int x(h(x) - \hat{h})\rho(x)\,\mathrm{d}x = \Sigma H(m - \bar{x}) \tag{29a}$$

$$\Sigma\mathsf{K} + \mathsf{K}\Sigma = \int (x - m)(x - m)^T (h(x) - \hat{h})\rho(x)\,\mathrm{d}x = \Sigma H \Sigma \tag{29b}$$

A unique positive-definite symmetric solution K exists for the Lyapunov equation (29b) because $\Sigma \succ 0$ and $\Sigma H \Sigma \succ 0$ [9].

On substituting the solution (26) into the Poisson equation (25) and dividing through by ρ, the two sides are:

$$\begin{aligned}-\frac{1}{\rho}\nabla \cdot (\rho \nabla \phi) &= (x - m)^T \Sigma^{-1}(\mathsf{K}(x - m) + b) - \mathrm{tr}(\mathsf{K}) \\ h - \hat{h} &= \frac{1}{2}(x - \bar{x})^T H (x - \bar{x}) - \frac{1}{2}(m - \bar{x})^T H (m - \bar{x}) - \frac{1}{2}\mathrm{tr}(H\Sigma)\end{aligned}$$

where $\mathrm{tr}(\cdot)$ denotes the matrix trace. Using formulae (29a)-(29b) for b and K, the two sides are seen to be equal. □

Proof of Proposition 1: Using the affine control law (11), the particle filter is a linear system with a Gaussian prior:

$$\frac{\mathrm{d}X_t^i}{\mathrm{d}t} = -\mathsf{K}_t(X_t^i - m_t) - b_t, \quad X_0^i \sim \mathcal{N}(m_0, \Sigma_0) \tag{30}$$

Therefore, the density of X_t^i is Gaussian for all $t > 0$. The evolution of the mean is obtained by taking an expectation of both sides of the ode (30):

$$\frac{\mathrm{d}m_t}{\mathrm{d}t} = -b_t = -\mathsf{E}[X_t^i(h(X_t^i) - \hat{h}_t)]$$

where (12) is used to obtain the second equality. The equation for the variance Σ_t of X_t^i is similarly obtained:

$$\begin{aligned}\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} &= -(\mathsf{K}_t\Sigma_t + \Sigma_t\mathsf{K}_t) \\ &= -\mathsf{E}\left[(X_t^i - m_t)(X_t^i - m_t)^T (h(X_t^i) - \hat{h}_t)\right]\end{aligned}$$

where (13) has been used. □

Proof of Corollary 1: The closed-form odes (15a) and (15b) are obtained by using explicit formulae (29a) and (29b) for b and K, respectively. □

E. Parametric case

Proof of Theorem 2: The natural gradient ode (20) is obtained by applying the chain rule. In its parameterized form, the density $p(x,t) = \varrho(x;\theta_t)$ evolves according to the replicator pde:

$$\frac{\partial \varrho}{\partial t}(x;\theta_t) = -(h(x) - \hat{h}_t)\, \varrho(x;\theta_t)$$

Now, using the chain rule,

$$\frac{\partial \varrho}{\partial t}(x,\theta_t) = \varrho(x,\theta_t) \left[\frac{\partial}{\partial \vartheta}(\log \varrho(x;\theta_t))\right]^T \frac{\mathrm{d}\theta_t}{\mathrm{d}t}$$

where $\frac{\partial}{\partial \vartheta}(\log \varrho)$ and $\frac{\mathrm{d}\theta_t}{\mathrm{d}t}$ are both $M \times 1$ column vectors. Therefore, the replicator pde is given by

$$\left[\frac{\partial}{\partial \vartheta}(\log \varrho(x;\theta_t))\right]^T \frac{\mathrm{d}\theta_t}{\mathrm{d}t}\, \varrho(x;\theta_t) = -(h(x) - \hat{h}_t)\, \varrho(x;\theta_t)$$

Multiplying both sides by the column vector $\frac{\partial}{\partial \vartheta}(\log \varrho)$, integrating over the domain, and using the definitions (18) of the Fisher information matrix G and (19) for ∇e, one obtains

$$G_{(\theta_t)} \frac{\mathrm{d}\theta_t}{\mathrm{d}t} = -\nabla e(\theta_t)$$

The ode (20) is obtained because G is assumed invertible. □