LINEAR PDES AND EIGENVALUE PROBLEMS CORRESPONDING TO ERGODIC STOCHASTIC OPTIMIZATION PROBLEMS ON COMPACT MANIFOLDS

JORIS BIERKENS, VLADIMIR Y. CHERNYAK, MICHAEL CHERTKOV, AND HILBERT J. KAPPEN

Abstract. We consider long term average or 'ergodic' optimal control poblems with a special structure: Control is exerted in all directions and the control costs are proportional to the square of the norm of the control field with respect to the metric induced by the noise. The long term stochastic dynamics on the manifold will be completely characterized by the long term density ρ and the long term current density J. As such, control problems may be reformulated as variational problems over ρ and J. We discuss several optimization problems: the problem in which both ρ and J are varied freely, the problem in which ρ is fixed and the one in which J is fixed. These problems lead to different kinds of operator problems: linear PDEs in the first two cases and a nonlinear PDE in the latter case. These results are obtained through through variational principle using infinite dimensional Lagrange multipliers. In the case where the initial dynamics are reversible we obtain the result that the optimally controlled diffusion is also symmetrizable. The particular case of constraining the dynamics to be reversible of the optimally controlled process leads to a linear eigenvalue problem for the square root of the density process.

Key words and phrases: Stochastic optimal control, ergodic theory, calculus of variations, differential geometry, flux, current, gauge invariance

Subject Classification: Primary 49K20; Secondary 93E20, 58A25

1. Introduction

In this paper we discuss stochastic, long term average optimal, or 'ergodic' control problems on compact orientable manifolds. The theory about ergodic control theory in continuous spaces has been developed relatively recently; see works by Borkar and Gosh (e.g. [5]) and the recent monograph [1]. To our knowledge no literature is available about this topic in the setting of compact manifolds.

We concentrate on a special case of the control problem, in which control is exerted in all directions and where the control costs are proportional to the square of the norm of the control field with respect to the metric induced by the noise, as discussed in Section 2. As such, our emphasis does not lie on the solution of applied control problems. This setting may however prove relevant for obtaining results in large deviations theory; see e.g. [11], where the connection is made between control problems and large deviations theory. The 'squared control cost' is further motivated by recent results we obtained on stochastic optimal control for finite time horizon problems, with relative entropy determining control cost [4]. This particular setting typically leads to linearized systems [12]. In the ergodic setting it leads typically to operator eigenvalue problems, see e.g. [18] for the diffusion case and [20] for the Markov chain setting.

On a compact manifold, a few phenomena play a special role. The main advantage of this setting is that transient behaviour cannot occur. Therefore, an invariant measure is necessarily unique and ergodicity follows immediately. The long term stochastic dynamics on the manifold

The research at Radboud University, The Netherlands, leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 270327 (CompLACS)..

The research at Wayne State University (MI), USA, has received support from the NSF under grant agreement no. CHE-1111350.

The work at LANL was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396.

will be completely characterized by the long term density ρ and the long term current density J (see Section 3). As such, control problems may be reformulated as variational problems over ρ and J. In the optimization problem, the density ρ is paired with the scalar cost or potential function V, and the current density J is paired with a *vector potential* or *gauge field* A to obtain the cost function. In Section 4 we discuss how to understand the notion of flux as a particular example of choosing a gauge field A.

We then discuss several optimization problems: the problem in which both ρ and J are varied freely (Section 5, the problem in which ρ is fixed (Section 6) and the one in which J is fixed (Section 7). These problems lead to different kinds of operator problems: linear PDEs in the first two cases and a nonlinear PDE in the latter case. These results are obtained through through variational principle using infinite dimensional Lagrange multipliers. This analysis is performed rigourously in Section 5.

In the case where the initial dynamics are reversible, or in other words, in case the diffusion is symmetrizable, we obtain the result that the optimally controlled diffusion is also symmetrizable (Section 5.4). The particular case of insisting $J = 0$ coincides with demanding reversible dynamics of the optimally controlled process. Interestingly, this optimization problem leads to a linear eigenvalue problem for the square root of the density process, just as we see in quantum mechanics (but note that our setting is entirely classical). We conclude this paper with a short discussion (Section 8).

This paper is written for a mathematical audience. The reader interested in the statistical physics interpretation of this material is referred to our related publication [6].

2. Problem setting

We will phrase our setting in terms of diffusion processes on manifolds, in the language of [14, Chapter V]. By smooth we always mean infinitely often differentiable. M will always denote a smooth compact orientable m-dimensional manifold. Let $C^\infty(M)$ denote the space of smooth functions from M into $\mathbb{R}$, let $\mathfrak{X}(M)$ denote the space of smooth vectorfields on M, and let $\Lambda^p(M)$ denote the space of smooth differential forms of order p on M, for $p = 0, 1, \ldots, m$.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ denote a filtered probability space on which is defined a d-dimensional standard Brownian motion. Consider a stochastic process X defined on M by the SDE, given in local coordinates by

$$dX_t^i = g^{ij}(X_t) f_j(t, X_t)\ dt + \sigma_\alpha^i(X_t) \circ dB_t^\alpha, \tag{1}$$

where, for $\alpha = 1, \ldots, d$, $\sigma_\alpha \in \mathfrak{X}(M)$, $g^{ij} := \sum_\alpha \sigma_\alpha^i \sigma_\alpha^j$, for $i, j = 1, \ldots, m$, is a symmetric positive semidefinite bilinear tensorfield on M, f is a differential 1-form on M, denoting force. For any initial condition $x_0 \in M$, let X^{x_0} denote the unique solution to (1) The notation $\circ dB^\alpha$ indicates that we take Stratonovich integrals with respect to the Brownian motion. One can think of f as a force field, resulting from a potential, some external influence, or a combination of both.

We will always assume the following hypothesis.

Hypothesis 2.1 (Ellipticity)**.** g is positive definite on M.

Under this assumption, g defines a Riemannian metric on M and we will use this as the metric of choice without further notice. This Riemannian metric induces a local inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $|| \cdot ||$ on tensors of arbitrary covariant and contravariant orders.

The SDE (1) is referred to as the *uncontrolled dynamics*. These dynamics may be altered by exterting a 'force' or 'control' u in the following way,

$$dX_t^i = g^{ij}(X(t)) \left[f_j(X_t) + u_j(X_t) \right]\ dt + \sum_{\alpha=1}^d \sigma_\alpha^i(X_t) \circ dB_t^\alpha, \tag{2}$$

where $u \in \Lambda^1(M)$. For any $x_0 \in M$ and $u \in \Lambda^1(M)$, the unique solution to (2) will be denoted by $X^{x_0,u}$. The SDE (2) is referred to as the *controlled dynamics*.

Consider the random functional $\mathcal{C} : \Omega \times M \times \Lambda^1(M) \to \mathbb{R}$ denoting pathwise long term average cost,

$$\mathcal{C}(\omega, x_0, u) := \limsup_{T\to\infty} \frac{1}{T}\left[\int_0^T V(X_s^{x_0,u}) + \frac{1}{2\lambda}||u(X_s^{x_0,u})||^2 \, ds + \int_0^T A(X_s^{x_0,u}) \circ dX_s^{x_0,u}\right], \tag{3}$$

where $\lambda > 0$, $V \in C^\infty(M)$ is a *potential* or *state dependent cost function*, $||u(\cdot)||^2$ represents the *(instantaneous) control cost* corresponding to a control field $u \in \Lambda^1(M)$, and $A \in \Lambda^1(M)$. The final term in (2) may represent a *flux*, as explained in Section 4. The differential form A is often called a *gauge field* in physics.

Remark 2.2. The 'lim sup' in (3) is used to avoid any discussion at this point about the existence of the limit. Instead of the *pathwise formulation* in (3) we could alternatively consider the weaker *average formulation*, in which case the cost function would be the long term average of the expectation value $\mathbb{E}^{x_0,u}$ of the integrand in (3). We will see in Section 3 that the limit of (3) exists (and not just the 'lim sup'). Furthermore this limit will turn out to be equal to a deterministic quantity, so that the pathwise formulation and the average formulation may be considered equivalent.

We will consider the following problem, along with some variations which we discuss in Sections 6 and 7.

Problem 2.3. For every $x_0 \in M$, find a differential 1-form $\hat{u} \in \Lambda^1(M)$ such that

$$\mathcal{C}(x_0, \hat{u}) = \inf_{u\in\Lambda^1(M)} \mathcal{C}(x_0, u), \quad \text{almost surely.}$$

3. Ergodic reformulation of the optimization problem

In this section we will derive two equivalent formulations of Problem 2.3. These reformulations, Problem 3.8 and Problem 3.13 below, are better suited to the analysis in the remaining sections. Also some notation will be established that will be used throughout this paper.

Let $\Omega^X = C([0,\infty); M)$ denote the space of sample paths of solutions to (2). We equip Ω^X with the σ-algebra $\mathcal{F}^X$ and filtration $(\mathcal{F}_t^X)_{t\geq 0}$ generated by the cylinder sets of X. Furthermore let probability measures $\mathbb{P}^{x_0,u}$ on Ω^X be defined as the law of $X^{x_0,u}$, for all $x_0 \in M$ and $u \in \Lambda^1(M)$. Note that for all $u \in \Lambda^1(M)$ the collection of probability measures $\mathbb{P}^{\cdot,u}$ defines a Markov process on Ω^X.

For the moment let $u \in \Lambda^1(M)$ be fixed. It will be convenient to use the shorthand notation

$$b_u^i(x) = g^{ij}(x)[f_j(x) + u_j(x)].$$

Recall that associated with the vectorfields σ_α, $\alpha = 1, \dots, d$, there exist first order differential operators also denoted by $\sigma_\alpha : C^\infty(M) \to C^\infty(M)$ defined by

$$\sigma_\alpha f(x) = \sigma_\alpha^i(x)\partial_i f(x), \quad x \in M,$$

for $f \in C^\infty(M)$, where $\partial_i = \frac{\partial}{\partial x^i}$ denotes partial differentation with respect to x^i. Similarly b_u defines a first order differential operator also denoted by b_u. By [14, Theorem V.1.2], the Markov generator corresponding to (2) is given by

$$L_u\phi(x) = \tfrac{1}{2}\sum_{\alpha=1}^d \sigma_\alpha\sigma_\alpha\phi(x) + b_u\phi(x).$$

On $\Lambda^p(M)$ an inner product is defined by $\langle\alpha,\beta\rangle_{\Lambda^p(M)} = \int_M \langle\alpha,\beta\rangle \, dx$, where dx denotes the volume form corresponding to g. The inner product $\langle\cdot,\cdot\rangle_{\Lambda^0(M)}$ is also denoted by $\langle\cdot,\cdot\rangle_{L^2(M)}$. Let $L^2(M,g) = L^2(M)$ denote the usual Hilbert space obtained by completing $C^\infty(M)$ with respect to the $L^2(M)$-inner product.

Lemma 3.1. *L_u may be written as*

$$L_u\Phi = \tfrac{1}{2}\Delta\Phi + \langle u + \widetilde{f}, d\Phi\rangle, \quad \Phi \in C^2(M),$$

where Δ the Laplace-Beltrami operator, $\widetilde{f}_i := f_i + \frac{1}{2} g_{ik} \left(\nabla_{\sigma_\alpha} \sigma_\alpha\right)^k$, and ∇ is the covariant derivative, corresponding to the Levi-Civita connection of the Riemannian metric g. The adjoint of L_u with respect to the $L^2(M)$ inner product is given by

$$L_u^* \Psi = \tfrac{1}{2} \Delta \Psi + \delta \left(\Psi(\widetilde{f} + u) \right), \quad \Psi \in C^2(M).$$

where $\delta : \Lambda^p(M) \to \Lambda^{p-1}(M)$ is the $L^2(M, g)$ adjoint of the exterior derivative operator d, i.e.

$$\int_M \langle d\alpha, \beta \rangle \ dx = \int_M \alpha \delta \beta \ dx, \quad \alpha \in C^\infty(M), \beta \in \Lambda^1(M),$$

where dx denotes the measure induced by the volume form on M.

Proof. The Laplace-Beltrami operator may be expressed as (see [14, p. 285, eqn. (4.32)])

$$\Delta \phi = g^{ij} \partial_i \partial_j \phi - g^{ij} \Gamma_{ij}^k \partial_k \phi, \tag{4}$$

where Γ_{ij}^k denote the Christoffel symbols corresponding to the Levi-Civita connection. Using this expression, we compute

$$\begin{aligned}\sum_\alpha \sigma_\alpha \sigma_\alpha \phi &= \sum_\alpha \sigma_\alpha^i \partial_i \left(\sigma_\alpha^j \partial_j \phi \right) = \sum_\alpha \left(\sigma_\alpha^i \sigma_\alpha^j \partial_i \partial_j \phi + \sigma_\alpha^i \left(\partial_i \sigma_\alpha^j \right) \left(\partial_j \phi \right) \right) = \sum_\alpha \left(g^{ij} \partial_i \partial_j \phi + \sigma_\alpha^i \left(\partial_i \sigma_\alpha^j \right) \left(\partial_j \phi \right) \right) \\ &= \Delta \phi + g^{ij} \Gamma_{ij}^k \partial_k \phi + \sum_\alpha \sigma_\alpha^i \left(\partial_i \sigma_\alpha^j \right) \partial_j \phi = \Delta \phi + \sum_\alpha \left(\nabla_{\sigma_\alpha} \sigma_\alpha \right) \phi,\end{aligned}$$

where the last equality is a result of the definition of the Levi-Civita connection and the corresponding Christoffel symbols. The expression for L_u^* is immediate from its definition. □

In the remainder of this work, we will assume all advection terms are absorbed in the force field f, and thus may omit the tilde in $\widetilde{f}$. This can alternatively be interpreted as assuming $\sum_\alpha \nabla_{\sigma_\alpha} \sigma_\alpha = 0$. This is further equivalent to the following hypothesis.

Hypothesis 3.2. The generator corresponding to X equals $L_u = \frac{1}{2} \Delta + b_u$.

This assumption is justified by the above lemma and the well-known fact that Markov generator L^u determines the law of the diffusion X uniquely.

Lemma 3.3. *Let $x_0 \in M$. The expectation of the trajectory of X over the gauge field may be expressed as*

$$\mathbb{E}^{x_0, u} \int_0^T A(X_t) \circ dX_t = \mathbb{E}^{x_0, u} \int_0^T \left[\langle A, f + u \rangle - \tfrac{1}{2} \delta A \right] (X_t) \ dt,$$

or in local coordinates,

$$\mathbb{E}^{x_0, u} \int_0^T A(X_t) \circ dX_t = \mathbb{E}^{x_0, u} \int_0^T g^{ik} A_i (f_k + u_k)(X_t) + \tfrac{1}{2} \frac{1}{\sqrt{|g|}} \partial_j \left(\sqrt{|g|} g^{ij} A_i \right) (X_t) \ dt.$$

Proof. Using the usual transformation rule between Itô and Stratonovich integrals [14, Equation (1.4), p. 250], we may write

$$dX_t^i = \overline{b}_u^i (X_t) \ dt + \sum_{\alpha=1}^d \sigma_\alpha^i (X_t) dB^\alpha(t), \tag{5}$$

where $\overline{b}_u(x)$ is given by

$$\overline{b}_u^i(x) := b_u^i(x) + \tfrac{1}{2} \sum_{\alpha=1}^d \left(\partial_k \sigma_\alpha^i(x) \right) \sigma_\alpha^k(x).$$

By the definition of the Stratonovich integral, $Z \circ dY = Z \ dY + \frac{1}{2} d[Z, Y]$ for semimartingales Y and Z [14, Equation (1.10), p.100], with $Z \ dY$ denoting the Itô integral. Therefore

$$\begin{aligned} A_i(X_t) \circ dX_t^i &= A_i(X_t) \ dX_t^i + \tfrac{1}{2} d[A_i(X_t), X_t^i] = A_i(X_t) \ dX_t^i + \tfrac{1}{2} \partial_j A_i(X_t) d[X_t^j, X_t^i] \\ &= A_i(X_t) \left[\overline{b}_u^i(X_t) \ dt + \sum_{\alpha=1}^d \sigma_\alpha^i(X_t) \ dB^\alpha(t) \right] + \tfrac{1}{2} \partial_j A_i(X_t) \sum_{\alpha=1}^d \sigma_\alpha^j(X_t) \sigma_\alpha^i(X_t) \ dt. \end{aligned}$$

Integrating over t and taking expectations gives

$$\begin{aligned}
&\mathbb{E}^{x_0,u}\int_0^T A(X_s)\circ dX_s \\
&= \mathbb{E}^{x_0,u}\int_0^T \left\{A_i(X_t)\left[b_u^i(X_t)+\tfrac{1}{2}\left(\partial_k\sigma_\alpha^i(X_t)\right)\sigma_\alpha^k(X_t)\right]+\tfrac{1}{2}\partial_j A_i(X_t)\sigma_\alpha^j(X_t)\sigma_\alpha^i(X_t)\right\}\ dt \\
&= \mathbb{E}^{x_0,u}\int_0^T \left\{g^{ik}A_i(f_k+u_k)(X_t)+\tfrac{1}{2}\nabla_j g^{ij}A_i(X_t)\right\}\ dt.
\end{aligned}$$

In the last expression we recognize the divergence of the vectorfield $g^{ij}A_i$, resulting in the stated expression. □

We recall the notion of an invariant probability distribution. Let $\mathcal{B}(M)$ denote the Borel σ-algebra on M.

Definition 3.4. A probability measure μ_u on M is called an invariant probability distribution for (2), if

$$\int_M \mathbb{P}^{x,u}(X_t\in B)\mu_u(dx)=\mu_u(B),\quad B\in\mathcal{B}(M).$$

The following result on invariant measures for nondegenerate diffusions [14, Proposition V.4.5] is essential for our purposes.

Proposition 3.5 (Existence and uniqueness of invariant probability measure)**.** *Suppose Hypothesis 2.1 is satisfied. Corresponding to any $u\in\Lambda^1(M)$ there exists a unique invariant probability measure μ_u on M corresponding to the diffusion on M defined by (2). Moreover, $\mu_u(dx)$ is given as $\rho_u(x)\ dx$, and $\rho_u\in C^\infty(M)$ is a solution of*

$$L_u^*\rho=0. \tag{6}$$

Furthermore $\rho_u>0$ on M.

We will refer to (6) as the *Fokker-Planck equation*, in agreement with the physics nomenclature. In the remainder of this work let μ_u and ρ_u as defined by Proposition 3.5.

In the physics literature, the *empirical density* and *empirical current density* are defined respectively as (see [7]):

$$\rho_t(x,\omega)=\frac{1}{t}\int_0^t\delta(x-X_s(\omega))\ ds,\quad J_t(x,\omega)=\frac{1}{t}\int_0^t\dot{X}_s\delta(x-X_s(\omega))\ ds.$$

Here (and only here) δ denotes the Dirac delta function. These fields, having a clear heuristic meaning, will be very relevant in the remainder of this work and we will make these precise from a mathematical point of view.

Let $B_b(M)$ denote the set of bounded Borel-measurable functions on M. We will work with the set of empirical average measures $(\nu_t(dx,\omega))_{t>0}$ on $\mathcal{B}(M)\times\Omega^X$, defined by

$$\nu_t(B):=\frac{1}{t}\int_0^t\mathbb{1}_B(X_s)\ ds,\quad t>0,\ B\in\mathcal{B}(M), \tag{7}$$

where $\mathbb{1}_B$ denotes the indicator function of the set B. Note that the measure $\mu_t(B):=\int_0^t\mathbb{1}_B(X_s)\,ds$ is known as the *local time* of X. Our primary interest is in the infinite time horizon limit.

Proposition 3.6. *For all $u\in\Lambda^1(M),\varphi\in L^2(M,\mu_u)$ (in particular for $\varphi\in B_b(M)$) and $x_0\in M$,*

$$\lim_{t\to\infty}\int_M\varphi\ d\nu_t=\int_M\varphi\ d\mu_u,\quad \mathbb{P}^{x_0,u}\text{-almost surely.} \tag{8}$$

Proof. For $u\in\Lambda^1(M)$, we define a stationary probability measure $\mathbb{P}^u$ on Ω^X by

$$\mathbb{P}^u(G)=\int_M\mathbb{P}^{x_0,u}(G)\ \mu_u(dy),\quad G\in\mathcal{F}.$$

For $\varphi \in L^2(M, \mu_u)$ we then have, by the ergodic theorem, see e.g. [8, Theorem 3.3.1], that $\lim_{t\to\infty} \int_M \varphi \, d\nu_t = \int_M \varphi \, d\mu_u$, $\mathbb{P}^u$-almost surely. Since $\rho_u > 0$ on M, this implies that $\lim_{t\to\infty} \int_M \varphi \, d\nu_t = \int_M \varphi \, d\mu_u$, $\mathbb{P}^{x_0,u}$-almost surely for μ_u-almost all $x_0 \in M$. By smooth dependence of the trajectories of X on the initial condition the result extends to all $x_0 \in M$. □

Note that $\lim_{t\to\infty} \nu_t$ is deterministic and does not depend on the choice of the initial condition $x_0 \in M$.

Corollary 3.7. *For $x_0 \in M$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T A(X_t) \circ dX_t = \int_M \left\{ \langle A, f + u \rangle - \tfrac{1}{2} \delta A \right\} \rho_u \; dx, \quad \mathbb{P}^{x_0,u}\text{-a.s.} \tag{9}$$

Proof. This is an immediate corollary of Lemma 3.3, Proposition 3.5 and Proposition 3.6. □

By the above results we may rephrase Problem 2.3 as follows.

Problem 3.8. Minimize

$$\mathcal{C}(\rho, u) := \int_M \left\{ V + \frac{1}{2\lambda} ||u||^2 + \langle A, f + u \rangle - \tfrac{1}{2} \delta A \right\} \rho \; dx. \tag{10}$$

with respect to $(\rho, u) \in C^\infty(M) \times \Lambda^1(M)$ subject to the constraints $L_u^* \rho = 0$ and $\int_M \rho \; dx = 1$.

Remark 3.9. The gauge field A may be completely removed from the problem by redefining f and V to be

$$\widetilde{f} = f - \lambda A, \quad \text{and} \quad \widetilde{V} = V - \tfrac{1}{2} \lambda ||A||^2 - \tfrac{1}{2} \delta A. \tag{11}$$

It may be checked that Problem 3.8 is equivalent to the minimization of

$$C(\rho, \widetilde{u}) = \int_M \left\{ \widetilde{V} + \frac{1}{2\lambda} ||\widetilde{u}||^2 \right\} \rho \; dx,$$

with respect to ρ and $\widetilde{u}$, subject to $\frac{1}{2}\Delta \rho + \delta((\widetilde{f} + \widetilde{u})\rho) = 0$. The control u solving Problem 3.8 for nonzero A may be retrieved by setting $u = \widetilde{u} - \lambda A$. Using this observation would simplify the derivation of the results in subsequent sections, but the process of reintroducing a nonzero gaugefield A in the results would lead to unnecessary confusion, so we will continue to work with a nonzero gaugefield A.

Note that $\dot{X}_s$ is not defined, a.s., so our mathematical analogue of the empirical current density requires more care. In Appendix A, we derive the differential 1-form $J_u \in \Lambda^1(M)$ denoting current density, as

$$J_u = -\tfrac{1}{2} d\rho_u + \rho_u \left(f + u \right). \tag{12}$$

Note that J_u is divergence free: $\delta J_u = L_u^* \rho_u = 0$. Furthermore a control u may be expressed in terms of the corresponding J_u and ρ_u as

$$u = -f + \frac{1}{\rho_u} \left(J_u + \tfrac{1}{2} d\rho_u \right). \tag{13}$$

Lemma 3.10. *Suppose u, ρ_u and J_u are related by (12). Then $\delta J_u = 0$ if and only if $L_u^* \rho_u = 0$, i.e. ρ_u is an invariant measure.*

Proof. This follows immediately from noting that $\delta J_u = L_u^* \rho_u$. □

Recall Lemma 3.3, where the expectation of the gauge field over the trajectory was expressed as an expectation over a Lebesgue integral. For the long term average of the gauge field this leads to the following result.

Lemma 3.11. *For $u \in \Lambda^1(M)$ and $x_0 \in M$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T A(X_t) \circ dX_t = \int_M \langle A, J_u \rangle \; dx, \quad \mathbb{P}^{x_0,u}\text{-a.s.} \tag{14}$$

Proof. From Corollary 3.7 we have (9). By (13), this is equal to

$$\int_M \left(\left\langle A, \frac{1}{\rho_u} \left(J_u + \tfrac{1}{2} d\rho_u \right) \right\rangle - \tfrac{1}{2} \delta A \right) \rho_u \, dx = \int_M \langle A, \left(J_u + \tfrac{1}{2} d\rho_u \right) \rangle - \tfrac{1}{2} \langle A, d\rho_u \rangle \, dx = \int_M \langle A, J_u \rangle \, dx.$$

□

Because of the above observations, instead of varying ρ and u in the optimization problem 3.8, we may as well vary $\rho \in C^\infty$ and $J \in \Lambda^1(M)$, while enforcing $\delta J = 0$ (equivalent to the Fokker-Planck equation for ρ by Lemma 3.10) and $\int_M \rho \, dx = 1$. Because of (13) the control u is then determined explicitly. Combining (13) and (14), we may alternatively express the cost functional (10) as a function of ρ and J, namely

$$\mathcal{C}(\rho, J) = \int_M \left\{ \left(V + \frac{1}{2\lambda} \left\| -f + \frac{1}{\rho} \left(J + \tfrac{1}{2} d\rho \right) \right\|^2 \right) \rho + \langle A, J \rangle \right\} \, dx. \tag{15}$$

Remark 3.12. Strictly speaking the use of $\mathcal{C}$ for different cost functionals is an abuse of notation; we trust this will not lead to confusion.

Problem (2.3) can thus be rephrased as the following problem:

Problem 3.13. Minimize $\mathcal{C}(\rho, J)$ with respect to $\rho \in C^\infty(M)$ and $J \in \Lambda^1(M)$, subject to the constraints $\delta J = 0$ and $\int_M \rho \, dx = 1$, where $\mathcal{C}(\rho, J)$ is given by (15).

4. Flux

In this section we will give a natural interpretation of the term $\int_M \langle A, J \rangle \, dx$, namely as the flux of J through a cross-section α, or equivalently, the long term average intersection index of the stochastic process $(X(t))_{t \geq 0}$ with a cross-section. The section motivates the gauge field A in the cost function. The remainder of this paper does not refer to this section. For background reading in differential geometry, see [2, 21]. See also [6, 7] where the ideas below are described in more detail.

Let M be a compact, oriented, Riemannian manifold of dimension m. Recall the notion of a *singular p-chain in M (with real coefficients)* as a finite linear combination $c = \sum a_i \sigma_i$ of smooth p-simplices σ_i in M where the a_i are real numbers. Let $S_p(M; \mathbb{R})$ denote the real vector space of singular p-chains in M. On $S_p(M; \mathbb{R})$, $p \in \mathbb{Z}$, $p \geq 0$, are defined boundary operators $\partial_p : S_p(M; \mathbb{R}) \to S_{p-1}(M; \mathbb{R})$. The *$p$-th singular homology group of M with real coefficients* is defined by

$$H_p(M; \mathbb{R}) = \ker \partial_p / \operatorname{im} \partial_{p+1}.$$

Elements of $\ker \partial_p$ are called p-cycles, and elements of $\operatorname{im} \partial_{p+1}$ are called p-boundaries. The *deRham cohomology classes* $H^p_{\mathrm{deRham}}(M; \mathbb{R})$ are defined, for $0 \leq p \leq m = \dim(M)$ as

$$H^p_{\mathrm{deRham}}(M) = \ker d_p / \operatorname{im} d_{p-1},$$

where $d_p : \Lambda^p(M) \to \Lambda^{p+1}(M)$ denotes exterior differentiation.

Let α be a p-cycle. The functional $p_\alpha \in \Lambda^p(M) \to \int_\alpha \beta \in \mathbb{R}$ depends, by Stokes' theorem, only on the homology class $[\alpha]$ of α, and the deRham cohomology class $[\beta]$ of β. An element $[\alpha] \in H_p(M; \mathbb{R})$ may therefore be considered an element of $(H^p_{\mathrm{deRham}}(M))^*$. Now let M be oriented and of dimension m. The mapping $q_\beta : [\gamma] \in H^p_{\mathrm{deRham}}(M) \to \int_M \beta \wedge \gamma \in \mathbb{R}$ is an element of $(H^p_{\mathrm{deRham}}(M))^*$. Poincaré duality states that the mapping $[\beta] \in H^{m-p}_{\mathrm{deRham}}(M) \to q_\beta \in (H^p_{\mathrm{deRham}}(M))^*$ is an isomorphism for compact M, i.e. $H^{m-p}_{\mathrm{deRham}}(M) \cong (H^p_{\mathrm{deRham}}(M))^*$. Therefore, for compact, oriented M, we have

$$H_p(M; \mathbb{R}) \cong (H^p_{\mathrm{deRham}}(M))^* \cong H^{m-p}_{\mathrm{deRham}}(M).$$

Since an equivalence class in $H^{m-p}_{\mathrm{deRham}}(M)$ has a unique harmonic representative, we conclude that for a p-cycle α, there exists a unique harmonic $r_\alpha \in \Lambda^{m-p}(M)$ such that

$$\int_\alpha \beta = p_\alpha(\beta) = \int_M r_\alpha \wedge \beta, \quad [\beta] \in H^p_{\mathrm{deRham}}(M).$$

In particular, for $p = m-1$, we may interpret $\int_\alpha \star J$ (with $\star$ the *Hodge star operator*) as the flux of J through α. This quantity may further be interpreted as the long term average intersection index of the stochastic trajectory $(X(t))_{t \geq 0}$ with respect to α, i.e. the long term average of the number of intersections (with ± 1 signs depending on the direction); see e.g. [13, Section 0.4]. Specializing the above result to this situation, we obtain the following proposition.

Proposition 4.1. *For a given $(m-1)$-cycle α, there exists a unique harmonic $A \in \Lambda^1(M)$, which depends only on the singular homology class $[\alpha] \in H_{m-1}(M;\mathbb{R})$ of α, such that $\int_\alpha \star J = \int_M \langle A, J\rangle \, dx$ for all $J \in \Lambda^1(M)$ satisfying $\delta J = 0$.*

Example 4.2 (S^1). The divergence free 1-forms J on S^1 are constant, say $J = J_0 \, d\theta$ for some $J_0 \in \mathbb{R}$. A 0-cycle α of S^1 consists of a collection of points $\theta_1, \ldots, \theta_k \subset [0, 2\pi]$ with multiplicities $\alpha_1, \ldots, \alpha_k$. The flux of J through α is then simply given by $\sum_{i=1}^k \alpha_i J(\theta_i) = \sum_{i=1}^k \alpha_i J_0$. By defining a differential form $A = A_0 \, d\theta$, with constant component $A_0 := \sum_i \alpha_i \frac{1}{2\pi}$, we find that

$$\int_{S^1} \langle A, J\rangle \, d\theta = \int_{S^1} A_0 J_0 \, d\theta = \sum_i \alpha_i J_0 = \int_\alpha \star J.$$

We see that this choice of A is the constant (and therefore harmonic) representative in $H^1_{\mathrm{deRham}}(S^1)$ corresponding to $[\alpha] \in H_0(S^1;\mathbb{R})$.

5. Unconstrained optimization – the HJB equation

In this section we will find necessary conditions for a solution of Problem 3.8 or equivalently Problem 3.13. In fact, for technical reasons we will work with the the formulation in terms of ρ and J, i.e. Problem 3.13. The main reason for this is the simplicity (in particular, the linearity) of the constraint $\delta J = 0$. This should be compared to the equivalent constraint $\frac{1}{2}\Delta\rho + \delta(\rho(f+u)) = 0$, which is nonlinear as a function in (ρ, u).

The approach to Problem 3.8 or Problem 3.13 is to use the method of Lagrange multipliers to enforce the constraints. Since the constraint $\delta J(x) = 0$ needs to be enforced for all $x \in M$, the corresponding Lagrange multiplier is an element of a function space. A purely formal derivation of the necessary conditions using Lagrange multipliers is straightforward, but we wanted to be more rigourous in the derivation of the necessary conditions. In Sections 6 and 7, we will be less rigourous in the derivations, in the comforting knowledge that we can use the machinery outlined in the current section.

5.1. Abstract optimization.

We will relax Problem 3.13 to an optimization problem over Sobolev spaces. In particular, we will rephrase it as the following abstract optimization problem. Let X and Z be Banach spaces and let $\mathcal{U}$ be an open set in X. Let $\mathcal{C} : \mathcal{U} \subset X \to \mathbb{R}$ and $\mathcal{H} : \mathcal{U} \subset X \to Z$.

Problem 5.1. Minimize $\mathcal{C}(x)$ over $\mathcal{U}$ subject to the constraint $\mathcal{H}(x) = 0$.

The *Fréchet derivative* [15] of a mapping $T : D \subset X \to Y$ in $x \in D$ will be denoted by $T'(x) \in L(X;Y)$. We will need the following notion.

Definition 5.2 (Regular point)**.** Let T be a continuously Fréchet differentiable function from an open set D in a Banach space X into a Banach space Y. If $x_0 \in D$ is such that $T'(x_0)$ maps X onto Y, then the point x_0 is said to be a *regular point* of the transformation T.

For the Problem 5.1 the following necessary condition holds for a local extremum [15, Theorem 9.3.1].

Proposition 5.3 (Lagrange multiplier necessary conditions)**.** *Suppose $\mathcal{C}$ and $\mathcal{H}$ are continuously Fréchet differentiable on $\mathcal{U}$. If $\mathcal{C}$ has a local extremum under the constraint $\mathcal{H}(x) = 0$ at the regular point $x_0 \in \mathcal{U}$, then there exists an element $z_0^* \in Z^*$ such that*

$$f'(x_0) + \langle H'(x_0), z_0^*\rangle = 0.$$

Here $\langle \cdot, \cdot \rangle$ denotes the pairing between Z and Z^*.

5.2. **Sobolev spaces.** We will define Sobolev spaces of differential forms as follows. For $k \in \mathbb{N} \cup \{0\}$ and $p \in \{0, \ldots, n\}$, the space $W^{k,p}(M)$ consists of all differential forms of order p which are in $L^2(M)$ together with their covariant derivatives up to order k. (Note the special meaning of the second index p, in contrast with the more common definition.) The norm

$$||\omega||_{k,p}^2 := \int_M \sum_{l=0}^{k} ||\nabla^l \omega||^2 \, dx$$

gives Hilbert space structure to $W^{k,p}(M)$. We will write $W^k(M) := W^{k,0}(M)$.

Note that the Laplace-Beltrami operator $\Delta = -(\delta d + d\delta)$ is a bounded mapping $\Delta : W^{k,p}(M) \to W^{k-2,p}(M)$. Also $\delta : W^{k,p}(M) \to W^{k-1,p-1}(M)$, for $p \geq 1$, and $d : W^{k,p}(M) \to W^{k-1,p+1}(M)$, for $p < n$, are bounded linear mappings.

Recall Sobolev's Lemma [19, Proposition 4.3.3].

Lemma 5.4 (Sobolev embedding)**.** *Suppose $\omega \in H^k(M)$ and suppose $m \in \mathbb{N} \cup \{0\}$ satisfies $k > n/2 + m$. Then $\omega \in C^m(M)$.*

5.3. **Obtaining necessary conditions for the optimization problem.** We let $k \in \mathbb{N}$ such that

$$k > n/2. \tag{16}$$

For this k, we define

$$\begin{aligned}
&X := W^{k+1}(M) \times W^{k,1}(M), \\
&\mathcal{U} := \mathcal{P} \times W^{k,1}(M), && \text{where} \quad \mathcal{P} := \left\{\rho \in W^{k+1}(M) : \rho > 0 \text{ on } M\right\}, \quad \text{and} \\
&Z := Z_1 \times \mathbb{R}, && \text{where} \quad Z_1 := \left\{\psi \in W^{k-1}(M) : \psi = \delta\beta \text{ for some } \beta \in W^{k,1}(M)\right\}.
\end{aligned}$$

By the condition (16) on k, we have by the Sobolev Lemma that $(\rho, J) \in X$ satisfies $\rho \in C(M)$. In particular, the condition $\rho > 0$ defines an open subset $\mathcal{U} \subset X$.

Lemma 5.5. *Suppose* (16) *holds. The mapping $\mathcal{C}(\rho, J)$, as given by* (15)*, may be continuously extended to a mapping $\mathcal{C} : \mathcal{U} \to \mathbb{R}$. Moreover, the mapping $\mathcal{C}$ is continuously differentiable on $\mathcal{U}$ with Fréchet derivative $\mathcal{C}'(\rho, J) \in L(X; \mathbb{R})$ given for $(\rho, J) \in \mathcal{U}$ by*

$$\begin{aligned}
(\zeta, G) \mapsto &\int_M \left\{ V + \frac{1}{2\lambda}||f||^2 - \frac{1}{2\lambda\rho^2}||J + \tfrac{1}{2}d\rho||^2 + \frac{1}{2\lambda}\delta\left(-f + \frac{1}{\rho}\left(J + \tfrac{1}{2}d\rho\right)\right)\right\}\zeta dx \\
&+ \int_M \left\langle \frac{1}{\lambda}\left(-f + \frac{1}{\rho}(J + \tfrac{1}{2}d\rho)\right) + A, G \right\rangle \, dx.
\end{aligned}$$

Proof. We compute $\mathcal{C}'(\rho, u)$ to be the linear functional on X given by

$$\begin{aligned}
(\zeta, G) \mapsto &\int_M \left\{ V + \frac{1}{2\lambda}\left|\left|-f + \frac{1}{\rho}(J + \tfrac{1}{2}d\rho)\right|\right|^2 \right\}\zeta + \frac{1}{\lambda}\left\langle -f + \frac{1}{\rho}(J + \tfrac{1}{2}d\rho), -\frac{1}{\rho^2}\left(J + \tfrac{1}{2}d\rho\right)\zeta + \frac{1}{2\rho}d\zeta \right\rangle \rho dx \\
&+ \int_M \left\{ \frac{1}{\lambda}\left\langle -f + \frac{1}{\rho}(J + \tfrac{1}{2}d\rho), \frac{1}{\rho}G \right\rangle \rho + \langle A, G\rangle \right\} \, dx,
\end{aligned}$$

This is after rearranging, and partial integration of the term containing $d\zeta$, equal to the stated expression. The derivative is a bounded functional on X by the uniform boundedness of $V, 1/\rho, d\rho, f, J$ and A. $\square$

We define the constraint mapping $\mathcal{H} : \mathcal{U} \to Z$ as

$$\mathcal{H}(\rho, J) := \left(\delta J, \int_M \rho \, dx - 1\right), \quad (J, u) \in \mathcal{U}.$$

The following lemma is now immediate.

Lemma 5.6. *The mapping $\mathcal{H}$ is continuously differentiable on X, with Fréchet derivative $\mathcal{H}'(\rho, J) \in L(X; Z)$ given for $(\rho, J) \in X$ by*

$$(\zeta, G) \mapsto \left(\delta G, \int_M \zeta \, dx\right), \quad (\zeta, v) \in X.$$

Every $(\rho, J) \in X$ is regular for $\mathcal{H}$, thanks to our choice of the function space Z:

Lemma 5.7. *Any $(\rho, J) \in X$ is a regular point of $\mathcal{H}$.*

Proof. Let $(\Psi, \kappa) \in Z = Z_1 \times \mathbb{R}$. In particular there exists a $\beta \in W^{k,1}(M)$ such that $\Psi = \delta\beta$. We may pick $G = \beta$, and ζ a constant function such that $\int_M \zeta \, dx = \kappa$. Then $\mathcal{H}'(\rho, J)(\zeta, G) = (\Psi, \kappa)$, showing that $\mathcal{H}'(\rho, J)$ is onto. □

In order to apply Proposition 5.3 in a useful manner, we need to give interpretation to Z^*, and in particular to Z_1^*. First let us recall that the spaces $(H^s(M))^*$, for $s \in \mathbb{N} \cup \{0\}$ may be canonically identified through the $L^2(M)$-inner product with spaces of distributions, denoted by $W^{-s}(M)$ [19, Proposition 4.3.2]. In other words, if $z \in (H^s(M))^*$, then there exists a distribution $\Phi \in W^{-s}(M)$ such that $z(\Psi) = \int_M \Phi\Psi \, dx$. Now in case $\Psi \in Z_1$, i.e. $\Psi = \delta\beta$ for some $\beta \in W^{k,1}(M)$, then $z(\Psi) = \int_M \Phi\delta\beta \, dx = \int_M \langle d\Phi, \beta\rangle \, dx$ for some $\Phi \in W^{-(k-1)}(M)$. Therefore the choice of $\Phi \in W^{-(k-1)}(M)$ representing $z \in Z_1^*$ is fixed up to the addition by a closed form: if $d\gamma = 0$ for $\gamma \in W^{-(k-1)}(M)$, then $\int_M (\Phi + \gamma)\delta\beta \, dx = \int_M \Phi\delta\beta \, dx$. We summarize this in the following lemma.

Lemma 5.8. $Z_1^* \cong W^{-(k-1)}(M)/\left\{\gamma \in W^{-(k-1)}(M) : d\gamma = 0\right\}$.

We may now apply Proposition 5.3 to obtain the following preliminary result.

Lemma 5.9 (Ergodic Hamilton-Jacobi-Bellman equation)**.** *Suppose $(\rho, J) \in \mathcal{U}$ is a local extremum of $\mathcal{C}$, defined by (15), under the constraint that $\mathcal{H}(\rho, J) = 0$. Then there exists $\Phi \in C^1(M)$ and $\gamma \in \mathbb{R}$ such that*

$$\begin{aligned} &V + \langle f, A + d\Phi\rangle - \tfrac{1}{2}\lambda||A + d\Phi||^2 - \tfrac{1}{2}\delta(A + d\Phi) + \gamma = 0, \quad \textit{and} \\ &J = -\tfrac{1}{2}d\rho + \rho f - \lambda\rho(A + d\Phi), \end{aligned} \tag{17}$$

where the first equality holds in weak sense, i.e. in $W^{-1}(M)$. The corresponding control field u is continuous and given by

$$u = -\lambda(A + d\Phi).$$

Proof. Let (ρ, J) as specified. By Proposition 5.3, there exists an element $(\Phi, \gamma) \in W^{-(k-1)}(M) \times \mathbb{R}$ such that the following equations hold.

$$\begin{aligned} V + \frac{1}{2\lambda}||f||^2 - \frac{1}{2\lambda\rho^2}||J + \tfrac{1}{2}d\rho||^2 + \frac{1}{2\lambda}\delta\left(-f + \frac{1}{\rho}\left(J + \tfrac{1}{2}d\rho\right)\right) + \gamma = 0, \quad \text{and} \\ \frac{1}{\lambda}\left(-f + \frac{1}{\rho}(J + \tfrac{1}{2}d\rho)\right) + A + d\Phi = 0. \end{aligned}$$

Here Φ is an arbitrary representative of some equivalence class in $W^{-(k-1)}(M)/\left\{\gamma \in W^{-(k-1)}(M) : d\gamma = 0\right\}$ as motivated by Lemma 5.8. Substituting the second equation into the first, and making some rearrangements, gives the sytem (17). Then $\Phi \in C^1(M)$ as a result of the equation for J in (17) and the continuity of J, ρ and $d\rho$. The expression for u is an immediate result of (13). □

Theorem 5.10 (Linearized HJB equation)**.** *Suppose $(\rho, J) \in \mathcal{U}$ is a local extremum of $\mathcal{C}$, defined by (15), under the constraint that $\mathcal{H}(\rho, J) = 0$. Then there exists a $\psi \in C^\infty(M)$, $\psi > 0$ on M, and $\mu \in \mathbb{R}$ such that*

$$\begin{aligned} &H\psi - W\psi = \mu\psi, \quad \textit{where} \\ &H\psi := L_{u_0}\psi, \quad u_0 = -\lambda A, \quad \textit{and} \qquad W = \lambda V + \lambda\langle f, A\rangle - \tfrac{1}{2}\lambda^2||A||^2 - \tfrac{1}{2}\lambda\delta A. \end{aligned} \tag{18}$$

Furthermore $\rho \in C^\infty(M)$, $J \in \Lambda^1(M)$ and the control field u, related to (ρ, J) as in (13)*, is in $\Lambda^1(M)$ as well, and*

$$u = -\lambda A + d(\ln \psi), \quad J = -\tfrac{1}{2} d\rho + \rho f - \lambda \rho A + \rho \, d(\ln \psi).$$

Proof. Let Φ and γ as in Lemma 5.9 and let $\psi := \exp(-\lambda \Phi)$. We compute

$$d\Phi = -\frac{1}{\lambda \psi} d\psi \quad \text{and} \quad \delta d \Phi = -\frac{1}{\lambda \psi} \delta d \psi - \frac{1}{\lambda \psi^2} ||d\psi||^2.$$

Inserting into (17), and multiplication by $\lambda \psi$ makes the $||d\psi||^2$-terms cancel and results in the equation

$$\lambda \psi V + \lambda \psi \langle f, A \rangle - \langle f, d\psi \rangle - \tfrac{1}{2} \lambda^2 \psi ||A||^2 + \lambda \langle A, d\psi \rangle - \tfrac{1}{2} \lambda \psi \delta A + \tfrac{1}{2} \delta d \psi + \lambda \gamma \psi = 0,$$

or equivalently (18), with $\mu = \lambda \gamma$.

Regularity. Note that, since $\psi \in C^1(M)$, we have $d\psi \in C(M) \subset L^2(M)$, so that $\psi \in H^1(M)$. Note that the first equation of (18) may be rewritten into $\frac{1}{2}\Delta \psi = \varphi$, with $\varphi = -\langle f - \lambda A, d\psi \rangle + W\psi + \mu \psi \in C(M) \subset L^2(M)$. By elliptic regularity ([19, Proposition 5.1.6]), it follows that $\psi \in H^2(M)$. In particular, $\varphi \in H^1(M)$. We may iterate this bootstrapping argument to conclude that $\psi \in H^s(M)$ for any $s \in \mathbb{N}$, and conclude from Sobolev embedding (Lemma 5.4) that $\psi \in C^\infty(M)$.

By (13), the corresponding control field is given by $u = -\lambda A + \frac{1}{\psi} d\psi \in \Lambda^1(M)$. By Proposition 3.5, $\rho \in C^\infty(M)$ and by the equation for J in (17), $J \in \Lambda^1(M)$. □

Since the problem considered in Theorem 5.10 is a relaxed version of Problem 3.13, and smoothness of J and ρ is established in the relaxed case, we immediately have the following corollary.

Corollary 5.11. *Suppose (ρ, u) is a solution of Problem 3.8, or equivalently that (ρ, J) a solution of Problem 3.13. Then the results of Theorem 5.10 hold.*

Remark 5.12. This result should be compared to [9], in which a variational principle is derived for the maximal eigenvalue of an operator L satisfying a maximum principle. Our results give an alternative characterization (as the solution of a control problem) of the largest eigenvalue in the case of an elliptic differential operator $L = L_{u_0}$.

Remark 5.13. An alternative way of deriving the HJB equation is by using the method of vanishing discount, see e.g. [1, Chapter 3].

5.4. **Symmetrizable solution.** In a special case we can represent the optimally controlled invariant measure in terms of ψ and the uncontrolled invariant measure. For this, we recall the notion of a *symmetrizable diffusion* [14, Section V.4]. Other equivalent terminology is that the Markov process *reversible* or that the invariant measure satisfies *detailed balance.*

Definition 5.14. Let $(T(t))_{t \geq 0}$ denote the transition semigroup of a diffusion on M. This diffusion is said to be *symmetrizable* if there exists a Borel measure $\nu(dx)$ such that

$$\int_M (T(t)f)(x) g(x) \; \nu \; dx = \int_M f(x) (T(t)g)(x) \; \nu \; dx \quad \text{for all } f, g \in C(M), t \geq 0.$$

In case a diffusion is symmetrizable with respect to a measure ν, this measure is an invariant measure for the diffusion.

The following results hold for any control field $u \in \Lambda^1(M)$.

Lemma 5.15. *Let X denote a diffusion with generator given by $Lh = \frac{1}{2}\Delta h + \langle f + u, dh \rangle$, $h \in C^2(M)$. The following are equivalent.*

(i) *X is symmetrizable, with invariant density $\rho_u = \exp(-U)$ for some $U \in C^\infty(M)$;*
(ii) *$f + u = -\frac{1}{2} dU$ for some $U \in C^\infty$;*
(iii) *The long term current density J_u, given by* (12)*, vanishes.*

Proof. The equivalence of (i) and (ii) is well known, see e.g. [14, Theorem V.4.6]. The equivalence of (ii) and (iii) is then immediate from (12). □

Proposition 5.16. *Let ρ_0 and $J_0 = -\frac{1}{2} d\rho_0 + \rho_0 f$ denote the density and current corresponding to the uncontrolled dynamics* (1). *The following are equivalent.*

(i) *The diffusion corresponding to the optimal control u is symmetrizable, with density $\rho = \psi^2 \rho_0$, where ψ is as in Theorem 5.10 (normalized such that $\int_M \psi^2 \rho_0 \ dx = 1$);*
(ii) *$J_0 = \lambda \rho_0 A$;*
(iii) *$f = -\frac{1}{2} dU + \lambda A$, for some $U \in C^\infty(M)$.*

In particular, if the uncontrolled diffusion is symmetrizable and $A = 0$, then the controlled difussion is symmetrizable and the density admits the expression given under (i).

Proof. Setting $\rho_u = \psi^2 \rho_0$, we have

$$\begin{aligned} J_u &= -\tfrac{1}{2} d\rho_u + \rho_u \left(f - \lambda A - \frac{1}{\psi} d\psi \right) = -\tfrac{1}{2} \psi^2 d\rho_0 - \rho_0 \psi d\psi + \psi^2 \rho_0 \left(f - \lambda A + \frac{1}{\psi} d\psi \right) \\ &= -\tfrac{1}{2} \psi^2 d\rho_0 + \psi^2 \rho_0 (f - \lambda A) = \psi^2 (J_0 - \lambda \rho_0 A), \end{aligned}$$

which establishes the equivalence of (i) and (ii). Representing the density ρ_0 by $\exp(-U)$ and using (12) with $u = 0$ gives the equivalence between (ii) and (iii). □

5.5. **Gauge invariance.** For a special choice of A, the solution of Problem 3.13 may be related to the solution corresponding to $A = 0$ in a simple way.

Proposition 5.17. *Let $A_0 \in \Lambda^1(M)$. For $\varphi \in C^\infty(M)$ and $A = A_0 + d\varphi$ let $(\rho_\varphi, J_\varphi)$ denote the solution of Problem 3.13 with corresponding solutions ψ_φ and u_φ of* (18). *Then, for $\varphi \in C^\infty(M)$,*

$$\rho_\varphi = \rho, \quad u_\varphi = u, \quad \psi_\varphi = \exp(\lambda \varphi) \psi, \quad \text{and} \quad J_\varphi = J, \tag{19}$$

where ρ, J, ψ and u denote the solution of Problem 3.13 and (18) *corrsponding to $A = A_0$.*

Proof. This is a matter of straightforward computation. □

In other words, the solution of Problem 3.13 depends (essentially) only on the equivalence class of A, under the equivalence relation $A \sim B$ if and only if $A = B + d\varphi$ for some $\varphi \in C^\infty(M)$.

Remark 5.18. A standard way in physics to obtain gauge invariant differential operators is to replace the derivatives with 'long' derivatives. This is illustrated by the observation that the operator H of (18) may be expressed, using the Hodge star operator $\star : \Lambda^p(M) \to \Lambda^{n-p}(M)$, $p = 0, \ldots, n$, as

$$H\psi = \tfrac{1}{2} \star (d - \lambda A \wedge) \star (d - \lambda A \wedge) \psi + \star (f \star (d - \lambda A \wedge)) \psi - \lambda V \psi.$$

The operator $\psi \mapsto d\psi - \lambda A \wedge \psi$ is called a *'long' derivative* operator. This result is easily obtained using the following observations

$$\star(\alpha \wedge \star\beta) = \langle \alpha, \beta \rangle, \quad \star d \star d\phi = \Delta\phi, \quad \star d \star \alpha = -\delta\alpha,$$

for $\alpha, \beta \in \Lambda^1(M)$ and $\phi \in C^\infty(M)$.

6. Fixed density

In this section we consider the problem of fixing the density function ρ, and finding a force u that obtains this density function, at minimum cost. Let $\rho \in C^\infty(M)$ be fixed, with $\rho > 0$ on M and $\int_M \rho \ dx = 1$. Then for some constant c_ρ we have

$$\mathcal{C}(\rho, u) = c_\rho + \int_M \left\{ \frac{1}{2\lambda} ||u||^2 + \langle A, u \rangle \right\} \rho \ dx.$$

Therefore we will consider the following problem.

Problem 6.1. Minimize $\mathcal{C}(u)$ over $\Lambda^1(M)$, subject to the constraint $L_u^* \rho = 0$, where $\mathcal{C}(u)$ is defined by

$$\mathcal{C}(u) = \int_M \left\{ \frac{1}{2\lambda} ||u||^2 + \langle A, u \rangle \right\} \rho \ dx.$$

The corresponding problem in terms of the curent density is the following. As for $\mathcal{C}(u)$, terms that do not depend on J are eliminated from the cost functional.

Problem 6.2. Minimize $\mathcal{C}(J)$ over $\Lambda^1(M)$, subject to the constraint $\delta J = 0$, where $\mathcal{C}(J)$ is defined by

$$\mathcal{C}(J) = \int_M \left(\frac{1}{\lambda} \left\langle -f + \frac{1}{2\rho} d\rho + \lambda A, J \right\rangle + \frac{1}{2\lambda\rho} ||J||^2 \right) \, dx. \tag{20}$$

By an analogous argument as in Section 5 (but less involved, since we are only optimizing over J), we obtain the following result: a relaxed version of Problem 6.2 may be transformed into an elliptic PDE. Essentially, this is obtained through variation of the Lagrangian functional

$$\mathcal{L}(J, \Phi) = \mathcal{C}(J) + \int_M \Phi \delta J \, dx.$$

Theorem 6.3. *Suppose $J \in \Lambda^1(M)$ is a local extremum of $\mathcal{C}(J)$ given by (20) under the constraint that $\delta J = 0$. Then there exists a $\Phi \in C^\infty(M)$ such that*

$$\Delta \Phi + \left\langle \frac{1}{\rho} d\rho, d\Phi \right\rangle = -\frac{1}{\lambda\rho} \left(\tfrac{1}{2} \Delta \rho + \delta(\rho(f - \lambda A)) \right). \tag{21}$$

For this Φ, J is given by

$$J = \rho f - \tfrac{1}{2} d\rho - \rho\lambda(A + d\Phi),$$

and the corresponding control field u is given by

$$u = -\lambda(A + d\Phi).$$

Example 6.4 (Circle). On S^1 every differential 1-form β, and in particular $\beta = f - \lambda A$, may be written as $\beta = -\frac{1}{2} \, dU + \frac{1}{2} k \, d\theta$, where θ represents the polar coordinate function, $U \in C^\infty(S^1)$ and $k \in \mathbb{R}$; see e.g. [21, Example 4.14]. Recall that $\delta f = -\operatorname{div} f$. Equation (21) then reads

$$\Phi''(\theta) + \frac{1}{\rho} \Phi'(\theta) \rho'(\theta) = -\frac{1}{2\lambda\rho} \left(\rho''(\theta) - \frac{d}{d\theta} \left(-\rho(\theta) U'(\theta) + k\rho \right) \right).$$

Based on Remark 6.5, we try a solution of the form $\Phi = -\frac{1}{2\lambda}(\ln \rho + U - \varphi)$. Inserting this into the differential equation, we obtain for ϕ the equation

$$\varphi''(\theta) + \gamma(\theta)\varphi'(\theta) = k\gamma(\theta),$$

where $\gamma(\theta) = \frac{d}{d\theta} \ln \rho(\theta) = \frac{1}{\rho(\theta)} \rho'(\theta)$.

Up to an arbitrary additive constant (which we put to zero), there exists a unique periodic solution φ to this differential equation, given by

$$\varphi(\theta) = \frac{k \left(\theta \int_0^{2\pi} \rho(\xi)^{-1} \, d\xi - 2\pi \int_0^\theta \rho(\xi)^{-1} \, d\xi \right)}{\int_0^{2\pi} \rho(\xi)^{-1} \, d\xi}, \quad \theta \in [0, 2\pi].$$

Remark 6.5 (Solution in the symmetrizable case). If $f - \lambda A = -\frac{1}{2} dU$ for some $U \in C^\infty(M)$, it may be checked that $\Phi = -\frac{1}{2\lambda} (\ln \rho + dU)$ solves (21), so that the optimal control field $u = \frac{1}{2} d(\ln \rho) - f$. In other words, the optimal way to obtain a particular density function ρ if $f - \lambda A$ is in 'gradient form' is by using a control u so that the resulting force field $f + u$ is again in gradient form, $f + u = \frac{1}{2} d(\ln \rho)$, resulting in reversible dynamics; see also Section 5.4.

Remark 6.6 (Gauge invariance). As in Section 5.5, it is straightforward to check that a solution to Problem 6.2 for $A_\varphi = A_0 + d\varphi$ is given by

$$\Phi_\varphi = \Phi - \varphi, \quad \rho_\varphi = \rho, \quad u_\varphi = u, \quad J_\varphi = J,$$

in terms of the solution (Φ, ρ, u, J) corresponding to the gauge field A_0.

7. Fixed current density

In this section we approach the problem of minimizing the average cost, under the constraint that J is fixed. In light of the remark just below (12), it will be necessary to demand that $\delta J = 0$, otherwise we will not be able to obtain a solution. By (13), we may express u in terms of J and ρ by $u = -f + \frac{1}{\rho}\left(J + \frac{1}{2} d\rho\right)$. Note that by Lemma 3.10, the Fokker-Planck equation (6) is satisfied. This leads to the following problem.

In the remainder of this section let $J \in \Lambda^1(M)$ satisfying $\delta J = 0$ be fixed.

Problem 7.1. Minimize $\mathcal{C}(\rho)$ subject to the constraint $\int_M \rho \, dx = 1$, where

$$\mathcal{C}(\rho) = \int_M \left(V + \frac{1}{2\lambda} \left|\left| -f + \frac{1}{\rho}\left(J + \tfrac{1}{2} d\rho\right) \right|\right|^2 \right) \rho \, dx. \tag{22}$$

Remark 7.2. The constraint $\rho \geq 0$ on M does not need to be enforced, since if we find ρ solving Problem 7.1 without this constraint, we may compute u by (13). Then ρ satisfies $L_u^* \rho = \delta J = 0$ by Lemma 3.10, so by Proposition 3.5 and the constraint $\int_M \rho \, dx = 1$, it follows that $\rho > 0$ on M.

Remark 7.3. Note that by Lemma 3.11, the contribution of A is determined once we fix J. Therefore we may put $A = 0$ in the current optimization problem.

Necessary conditions for the solution of Problem 7.1 may be obtained rigourously in a similar manner as in Section 5, to obtain the following result.

Theorem 7.4. *Suppose $\rho \in C^\infty(M)$ minimizes $\mathcal{C}(\rho)$ given by* (22) *under the constraint that $\int_M \rho \, dx = 1$. Then there exists a $\mu \in \mathbb{R}$ such that*

$$\tfrac{1}{2}\Delta\phi - (W + \lambda\mu)\phi = -\frac{||J||^2}{2\phi^3}, \tag{23}$$

holds, where $\phi = \sqrt{\rho}$ and $W = \lambda V + \frac{1}{2}||f||^2 - \frac{1}{2}\delta f$.

Remark 7.5. Equation (23) is known (at least in the one-dimensional case) as *Yermakov's equation* [17].

Instead of proving Theorem 7.4 rigourously (which may be done analogously to Section 5) we provide an informal derivation, which we hope provides more insight to the reader. We introduce the Lagrangian $\mathcal{L} : C^\infty(M) \times \mathbb{R} \to \mathbb{R}$ by

$$\mathcal{L}(\rho, \mu) = \int_M \left(V + \frac{1}{2\lambda} \left|\left| -f + \frac{1}{\rho}\left(J + \tfrac{1}{2} d\rho\right) \right|\right|^2 \right) \rho \, dx + \mu \left(\int_M \rho \, dx - 1 \right).$$

Varying $\mathcal{L}(\rho, \mu)$ with respect to ρ in the direction $\zeta \in C^\infty(M)$ gives

$$\begin{aligned}
&\mathcal{L}'(\rho, \mu)\zeta \\
&= \int_M \left(V + \frac{1}{2\lambda} \left|\left| -f + \frac{1}{\rho}\left(J + \tfrac{1}{2} d\rho\right) \right|\right|^2 + \mu \right) \zeta + \frac{1}{\lambda} \left\langle -f + \frac{1}{\rho}\left(J + \tfrac{1}{2} d\rho\right), -\frac{1}{\rho^2}\left(J + \tfrac{1}{2} d\rho\right)\zeta + \frac{1}{2\rho} d\zeta \right\rangle \rho \, dx \\
&= \int_M \left(\lambda V + \lambda\mu + \tfrac{1}{2}||f||^2 - \frac{1}{2\rho^2}||J||^2 - \tfrac{1}{2}\delta f + \frac{1}{8\rho^2}||d\rho||^2 - \frac{1}{4\rho}\Delta\rho \right) \frac{\zeta}{\lambda} \, dx,
\end{aligned}$$

where we used the identities $\delta J = 0$ and $\delta(h\omega) = -\langle dh, \omega\rangle + h\delta\omega$ for $h \in C^\infty(M)$ and $\omega \in \Lambda^1(M)$. We require that for any direction ζ the above expression equals zero, which is the case if and only if

$$-\frac{1}{4\rho}\Delta\rho + \frac{1}{8\rho^2}||d\rho||^2 - \frac{1}{2\rho^2}||J||^2 + W + \lambda\mu = 0. \tag{24}$$

Note that we need to solve this equation for both ρ and μ, in combination with the constraint that $\rho > 0$ on M and $\int_M \rho \, dx = 1$. By substituting $\phi = \sqrt{\rho}$, equation (24) transforms into the equation (23).

We may then compute the cost corresponding to $\rho = \phi^2$ as

$$
\begin{aligned}
\mathcal{C}(\rho) &= \int_M \left(V + \frac{1}{2\lambda} \left|\left| -f + \frac{1}{\rho}\left(J + \tfrac{1}{2} d\rho\right) \right|\right|^2 \right) \rho \; dx \\
&= \int_M \frac{1}{4\lambda} \Delta \rho - \frac{1}{8\lambda\rho} ||d\rho||^2 + \frac{1}{2\lambda\rho}||J||^2 - \frac{\rho}{2\lambda}||f||^2 + \frac{\rho}{2\lambda}\delta f - \mu\rho + \frac{\rho}{2\lambda} \left|\left| -f + \frac{1}{\rho}\left(J + \tfrac{1}{2} d\rho\right) \right|\right|^2 \; dx \\
&= \int_M \frac{1}{4\lambda}\Delta\rho + \frac{\rho}{2\lambda}\delta f - \mu\rho - \frac{1}{\lambda}\langle f, J\rangle - \frac{\rho}{\lambda}\left\langle f, \frac{1}{2\rho} d\rho \right\rangle + \frac{1}{\lambda\rho}\langle J, \tfrac{1}{2} d\rho\rangle \; dx \\
&= -\frac{1}{\lambda}\left(\mu\lambda + \int_M \langle f, J\rangle \; dx \right),
\end{aligned}
\tag{25}
$$

where we used (24) in the first equality, and the following observations for the last equality:

$$
\int_M \Delta\rho \; dx = -\int \langle d(1), d\rho\rangle \; dx = 0, \qquad \int_M \rho(\delta f) \; dx = \int_M \langle d\rho, f\rangle \; dx,
$$

$$
\int_M \frac{1}{\rho}\langle J, d\rho\rangle \; dx = \int_M \langle J, d(\ln \rho)\rangle \; dx = \int_M (\ln\rho)(\delta J) \; dx = 0, \qquad \int_M \rho \; dx = 1.
$$

We can only influence the first term in (25) by choosing ρ or μ, so we see that minimizing $\mathcal{C}$ therefore corresponds to finding the largest value of μ such that (24), or, equivalently, (23), admits a solution.

7.1. **Symmetrizable solution – time independent Schrödinger equation.** In this section we consider the special case of the above problem for zero current density, $J = 0$. By Lemma 5.15, this is equivalent to $u + f = -\frac{1}{2} d\Psi$ for some unknown $\Psi \in C^\infty(M)$, with $\rho = \exp(-\Psi)$. In other words, we demand the net force field (including the control) to be in gradient form, and the corresponding diffusion to be symmetrizable; see Section 5.4.

In this case (23) transforms into the linear eigenvalue problem,

$$
\tfrac{1}{2}\Delta\phi - W\phi = \lambda\mu\phi. \tag{26}
$$

This is intriguing since this is in fact a time independent Schrödinger equation for the square root of a density function, analogous to quantum mechanics; even though our setting is entirely classical.

By (25), we are interested in the largest value of μ so that (26) has a solution ϕ. The optimal control field is then given by

$$
u = \frac{1}{\phi} d\phi - f. \tag{27}
$$

Remark 7.6. It is straightforward to check that if $f = -\frac{1}{2} dU$ for some $U \in C^\infty(M)$, then $\phi = \exp(-\frac{1}{2}U)$ satisfies (26) with $V = 0$, $\mu = 0$, resulting in $u = 0$. This corresponds to the intuition that, if f is already a gradient, no further control is necessary to obtain a symmetrizable invariant measure.

Remark 7.7. We may also compare the case $f = -\frac{1}{2} dU$ with the result of Section 5.4. There we obtained that, in case $A = 0$ and $f = -\frac{1}{2} dU$, the optimization problem for unconstrained J resulted in a symmetrizable solution. In other words, the constraint $J = 0$ does not need to be enforced, and the solution of this section should equal the solution obtained in Proposition 5.16. Apparently, with ψ as in Proposition 5.16, we have that $\phi^2 = \psi^2 \exp(-U)$.

8. Discussion

In this paper we showed how stationary long term average control problems are related to eigenvalue problems (for the unconstrained problem and the problem constrained to a symmetrizable solution, Sections 5 and 7.1), elliptic PDEs (for the problem with fixed density, Section 6) or a nonlinear PDE (for the problem with fixed current density, Section 7). For this we fruitfully used the representation of an optimal control field u in terms of the density function ρ and the current density J. We showed in detail how an infinite dimensional Lagrange multiplier problem may be transformed into a PDE (Section 5). A striking relation between the classical setting and quantum mechanics was obtained in Section 7.1.

The theory on existence of solutions and spectrum of operators is classical and we refer the interested reader to e.g. [10, 19]. Let us again point out the strong connection of our results with earlier work of Donsker, Varadhan (1975) [9]; see also Remark 5.12. We will further investigate this connection as part of our future research.

One may ask the question whether we may obtain solutions when we constrain a certain flux $\int_M \langle A, J\rangle \ dx$ (see Section 4) to a given value. In this case, one may use $\widetilde{A} = \mu A$ as a Lagrange multiplier and use the results of Section 5 (for constrained flux) and Section 6 (for constrained flux and density) to obtain necessary conditions. As these results did not provide us with profound insight, we aim to report on this topic in a subsequent publication after further analysis of the problem.

Appendix A. Derivation of expression for long term average of current density

In the physics literature [7], the current density is defined formally as

$$J_t^i(x) = \frac{1}{t}\int_0^t \dot{X}_s^i \delta(X_s - x)\ ds, \tag{28}$$

for $x \in M$, where δ is the Dirac delta function. We will derive an alternative expression for this quantity, using the model (2) for the dynamics. Note that (28) formally defines a vector field that acts on functions $f \in C^\infty(M)$ as

$$J_t f(x) = \frac{1}{t}\int_0^t \dot{X}_s^i \delta(X_s - x)\partial_i f(x)\ ds = \frac{1}{t}\int_0^t \dot{X}_s^i \delta(X_s - x)\partial_i f(X_s)\ ds = \frac{1}{t}\int_0^t \delta(X_s - x)\partial_i f(X_s)\circ dX_s^i.$$

The δ-function is still problematic. We may however formally compute the $L^2(M,g)$ inner product of the above expression with any $h \in C^\infty(M)$ with support in a coordinate neighbourhood U containing x. This results in

$$\int_M h(x) J_t f(x)\ dx = \frac{1}{t}\int_U h(x)\int_0^t \delta(X_s - x)\partial_i f(X_s)\circ dX_s^i\ dx = \frac{1}{t}\int_0^t h(X_s)\partial_i f(X_s)\circ dX_s^i.$$

Using the relation $Y \circ dZ = Y\ dZ + \frac{1}{2} d[Y,Z]$ and (5), we compute

$$\begin{aligned}
\int_U h(x) J_t f(x)\ dx &= \frac{1}{t}\int_0^t h(X_s)\partial_i f(X_s)\left(\overline{b}_u^i(X_s) ds + \sigma_\alpha^i(X_s)\ dB_s^\alpha\right) + \tfrac{1}{2}\frac{1}{t}\int_0^t \sigma_\alpha^i \sigma_\alpha^j \partial_j\left(h\partial_i f\right)(X_s)\ ds \\
&\to \int_U \left\{ h(x)(\partial_i f)(x)\left(\overline{b}_u^i(x)\right) + \tfrac{1}{2} g^{ij}\partial_j\left(h\partial_i f\right)(x)\right\}\ \mu_u(dx) \quad \text{(almost surely as } t \to \infty) \\
&= \int_U h(x)\left\{\rho_u(\partial_i f)\left(\overline{b}_u^i\right) - \tfrac{1}{2}\frac{1}{\sqrt{|g|}}(\partial_i f)\,\partial_j\left(\rho_u\sqrt{|g|}g^{ij}\right)\right\}(x)\ dx,
\end{aligned}$$

using Proposition 3.6 and the law of large numbers for martingales ([16, Theorem 3.4], or [3, Section 6.4.1] for a proof). We find that the long term average vector field J has components

$$\begin{aligned}
J^i &= \rho_u \overline{b}_u^i - \tfrac{1}{2}\frac{1}{\sqrt{|g|}}\partial_j\left(\rho_u\sqrt{|g|}g^{ij}\right) = \rho_u\left(b_u^i + \tfrac{1}{2}\sigma_\alpha^k\left(\partial_k \sigma_\alpha^i\right)\right) - \tfrac{1}{2}g^{ij}\left(\partial_j \rho_u\right) - \tfrac{1}{2}\rho_u\frac{1}{\sqrt{|g|}}\partial_j\left(\sqrt{|g|}g^{ij}\right) \\
&= \rho_u b_u^i - \tfrac{1}{2}g^{ij}\left(\partial_j\rho_u\right),
\end{aligned}$$

where the last equality is a result of the identity

$$\left(\nabla_{\sigma_\alpha}\sigma_\alpha\right)^i = \sigma_\alpha^k\left(\partial_k\sigma_\alpha^i\right) - \frac{1}{\sqrt{|g|}}\partial_j\left(\sqrt{|g|}g^{ij}\right),$$

which may be verified by straightforward calculation. Lowering the indices gives the differential form

$$J_u = -\tfrac{1}{2}d\rho + \rho(f+u).$$

Note the abuse of notation: both differential form and vector field are denoted by J. The vector field denoted by J has no relevance in the remainder of this work.

References

[1] Ari Arapostathis, Vivek S. Borkar, and Mrinal K. Ghosh. *Ergodic control of diffusion processes*, volume 143 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2012.

[2] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein.

[3] J. Bierkens. *Long Term Dynamics of Stochastic Evolution Equations*. PhD thesis, Universiteit Leiden, 2009.

[4] J. Bierkens and H.J. Kappen. Explicit solution of relative entropy weighted control. *submitted; arXiv:1205.6946*, 2012.

[5] Vivek S. Borkar and Mrinal K. Ghosh. Ergodic control of multidimensional diffusions. I. The existence results. *SIAM J. Control Optim.*, 26(1):112–126, 1988.

[6] Vladimir Y. Chernyak, Michael Chertkov, J. Bierkens, and H.J. Kappen. Gauge-invariant and current-density controls. *In preparation*.

[7] Vladimir Y. Chernyak, Michael Chertkov, Sergey V. Malinin, and Razvan Teodorescu. Non-equilibrium thermodynamics and topology of currents. *J. Stat. Phys.*, 137(1):109–147, 2009.

[8] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite Dimensional Systems*. Cambridge University Press, 1996.

[9] Monroe D. Donsker and S. R. S. Varadhan. On a variational formula for the principal eigenvalue for operators with maximum principle. *Proc. Nat. Acad. Sci. U.S.A.*, 72:780–783, 1975.

[10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[11] Jin Feng and Thomas G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.

[12] Wendell H. Fleming. Logarithmic transformations and stochastic control. In *Advances in filtering and optimal stochastic control (Cocoyoc, 1982)*, volume 42 of *Lecture Notes in Control and Inform. Sci.*, pages 131–141. Springer, Berlin, 1982.

[13] Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1994. Reprint of the 1978 original.

[14] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989.

[15] David G. Luenberger. *Optimization by vector space methods*. John Wiley & Sons Inc., New York, 1969.

[16] X. Mao. Stochastic Differential Equations and Applications. *Horwood, Chichester*, 1997.

[17] Andrei D. Polyanin and Valentin F. Zaitsev. *Handbook of exact solutions for ordinary differential equations*. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2003.

[18] Per Rutquist, Claes Breitholtz, and Torsten Wik. On the infinite time solution to state-constrained stochastic optimal control problems. *Automatica*, 44(7):1800 – 1805, 2008.

[19] M. E. Taylor. *Partial Differential Equations: Basic Theory*. Springer, 1996.

[20] Emanuel Todorov. Linearly-solvable markov decision problems. In *NIPS*, pages 1369–1376, 2006.

[21] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1983. Corrected reprint of the 1971 edition.

J. Bierkens, Radboud University, Faculty of Science, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

E-mail address: `j.bierkens@science.ru.nl`

V. Y. Chernyak, Wayne State University, Department of Chemistry, Detroit (MI), USA

E-mail address: `chernyak@chem.wayne.edu`

M. Chertkov, Los Alamos National Laboratory, Center for Nonlinear Studies, Los Alamos (NM), USA

E-mail address: `chertkov@lanl.gov`

H. J. Kappen, Radboud University, Faculty of Science, Nijmegen, The Netherlands

E-mail address: `b.kappen@science.ru.nl`