

# Coupled Soliton Dynamics in Quadratic Mean Field Games

Kaivalya Bakshi

June 26, 2018

## Abstract

In this report we discuss coupled dynamics of solitons introduced in [1] modeling Quadratic Mean Field Games (MFG) and introduce an action functional for admitting nonlinear Langevin dynamics in the soliton model.

## 1 Coupled Soliton Dynamics

### 1.1 Dynamics

From pp 49, [1] plugging the forward backward PDE ansatz into the general variational ansatz in the 1D dynamics case, we obtain the Lagrangian expression

$$S[X_t, P_t, \Sigma_t, \Lambda_t] = \int_0^T L(X_t, P_t, \Sigma_t, \Lambda_t) dt = \int_0^T \dot{P}_t X_t - \frac{\Lambda_t \dot{\Sigma}_t}{2\Sigma_t} + \frac{P_t^2}{\mu} + \frac{\Lambda_t^2 - \mu^2 \sigma^4}{8\mu \Sigma_t^2} + \langle U_0 \rangle + \frac{g}{(\alpha+1)\sqrt{(\alpha+1)(2\pi)^\alpha}} \frac{1}{\Sigma_t^\alpha} dt \quad (1)$$

wherein one can obtain as in appendix B [1]

$$\int_0^T \delta \left( \dot{P}_t X_t + \frac{P_t^2}{\mu} \right) dt = \int_0^T \dot{P}_t \delta X_t + \left( \frac{P_t}{\mu} - \dot{X}_t \right) \delta P_t dt \quad (2)$$

$$\int_0^T \delta \left( -\frac{\Lambda_t \dot{\Sigma}_t}{2\Sigma_t} \right) dt = \int_0^T \frac{\dot{\Lambda}_t}{2\Sigma_t} \delta \Sigma_t - \frac{\dot{\Sigma}_t}{2\Sigma_t} \delta \Lambda_t dt \quad (3)$$

$$\int_0^T \delta \left( -\frac{\Lambda_t \dot{\Sigma}_t}{2\Sigma_t} \right) dt = \int_0^T -\frac{\Lambda_t}{4\mu \Sigma_t^2} \delta \Lambda_t - \frac{\Lambda_t^2 - \mu^2 \sigma^4}{4\mu \Sigma_t^3} \delta \Sigma_t dt \quad (4)$$

$$\int_0^T \delta \left( \frac{g}{(\alpha+1)\sqrt{(\alpha+1)(2\pi)^\alpha}} \frac{1}{\Sigma_t^\alpha} \right) dt = \int_0^T \frac{-g\alpha}{(\alpha+1)\sqrt{(\alpha+1)(2\pi)^\alpha}} \frac{1}{\Sigma_t^{\alpha+1}} \delta \Sigma_t dt. \quad (5)$$

Coupling between the ODEs on  $(X_t, \Sigma_t)$  can occur only due to the potential energy  $\langle U_0 \rangle(X_t, \Sigma_t)$ .

We note some steps in computing the variation before stating the result. Since  $\langle U_0 \rangle = \int U_0(x) \frac{e^{-\frac{(x-X_t)^2}{2\Sigma_t^2}}}{\sqrt{2\pi\Sigma_t}} dx$ , we need to compute

$$\exp\left(\frac{-(x - X_t - \delta X_t)^2}{2(\Sigma_t + \delta \Sigma_t)^2}\right) \stackrel{O(\delta X_t, \delta \Sigma_t)}{=} e^{-\frac{(x-X_t)^2}{2\Sigma_t^2}} \left( 1 + \frac{(x - X_t)}{\Sigma_t^2} \delta X_t + \frac{(x - X_t)^2}{\Sigma_t^3} \delta \Sigma_t \right) \quad (6)$$

and further

$$\begin{aligned} \delta \left( \frac{e^{\frac{-(x-X_t)^2}{2\Sigma_t^2}}}{\Sigma_t} \right) &\stackrel{O(\delta X_t, \delta \Sigma_t)}{=} \frac{e^{\frac{-(x-X_t)^2}{2\Sigma_t^2}}}{\Sigma_t} \left( \frac{(x-X_t)}{\Sigma_t^2} \delta X_t + \frac{(x-X_t)^2}{\Sigma_t^3} \delta \Sigma_t - \frac{1}{\Sigma_t} \delta \Sigma_t \right) \\ &= \frac{e^{\frac{-(x-X_t)^2}{2\Sigma_t^2}}}{\Sigma_t} \left( \frac{(x-X_t)}{\Sigma_t^2} \delta X_t + \frac{(x-X_t)^2 - \Sigma_t^2}{\Sigma_t^3} \delta \Sigma_t \right). \end{aligned} \quad (7)$$

Integrals of relevance are

$$\int U_0(y + X_t) y \mathcal{N}_y(0, \Sigma_t) dy = - \int \nabla U_0(y + X_t) \Sigma_t^2 \cdot -\mathcal{N}_y(0, \Sigma_t) dy = \Sigma_t^2 \langle \nabla U_0(y + X_t) \rangle \quad (8)$$

$$\int U_0(y + X_t) y^2 \mathcal{N}_y(0, \Sigma_t) dy = U_0(y + X_t) \Sigma_t^2 \Big|_{-\infty}^{+\infty} \quad (9)$$

$$- \int \nabla U_0(y + X_t) \left( \frac{\Sigma_t^2}{2} \operatorname{erf} \left( \frac{y}{\Sigma_t \sqrt{2}} \right) - \Sigma_t^2 y \mathcal{N}_y(0, \Sigma_t) \right) dy \quad (10)$$

$$- \int U_0(y + X_t) \Sigma_t^2 \mathcal{N}_y(0, \Sigma_t) dy = - U_0(y + X_t) \Sigma_t^2 \Big|_{-\infty}^{+\infty} + \int \nabla U_0(y + X_t) \left( \frac{\Sigma_t^2}{2} \operatorname{erf} \left( \frac{y}{\Sigma_t \sqrt{2}} \right) \right) dy. \quad (11)$$

The first order variation may be obtained as

$$\int_0^T \delta \langle U_0 \rangle dt = \int_0^T \langle \nabla U_0 \rangle \delta X_t + \Sigma_t \langle \nabla^2 U_0 \rangle \delta \Sigma_t dt. \quad (12)$$

The required coupled forward backward ODE system is therefore

$$\dot{X}_t = \frac{P_t}{\mu} \quad (13)$$

$$\dot{P}_t = - \langle \nabla U_0 \rangle \quad (14)$$

$$\dot{\Sigma}_t = \frac{\Lambda_t}{2\mu\Sigma_t} \quad (15)$$

$$\dot{\Lambda}_t = \frac{\Lambda_t^2 - \mu^2 \sigma^4}{2\mu\Sigma_t^2} + \frac{2g\alpha}{(\alpha+1)\sqrt{(\alpha+1)(2\pi)^\alpha}} \frac{1}{\Sigma_t^\alpha} - \Sigma_t^2 \langle \nabla^2 U_0 \rangle. \quad (16)$$

The decoupled soliton dynamics maybe obtained by observing under the delta Dirac distribution assumption, that  $\langle \nabla U_0 \rangle = \nabla U_0(X_t)$  and  $\Sigma_t \rightarrow 0$ .

Particularly,  $U_0(x) = -\frac{x^4}{4} - \frac{x^2}{2}$  implies  $\nabla U_0 = -x^3 - x$ ,  $\nabla^2 U_0 = -3x^2 - 1$ ,  $\nabla^3 U_0 = -6x$ ,  $\nabla^4 U_0 = -6$  so that  $\langle U_0 \rangle = -\frac{X_t^4 + 6X_t^2\Sigma_t^2 + 3\Sigma_t^4}{4} - \frac{X_t^2 + \Sigma_t^2}{2}$ ,  $\langle \nabla U_0 \rangle = -X_t^3 - 3X_t\Sigma_t^2 - X_t$ ,  $\langle \nabla^2 U_0 \rangle = -3(X_t^2 + \Sigma_t^2) - 1$ ,  $\langle \nabla^3 U_0 \rangle = -6X_t$  and  $\langle \nabla^4 U_0 \rangle = -6$ .

## 1.2 Hamiltonian mechanics

Modifying the first term in the Lagrangian as  $\int_0^T \dot{P}_t X_t = X_t P_t \Big|_0^T - \int_0^T \dot{X}_t P_t$  we arrive at a Hamiltonian system for canonical variables defined as follows. Using the canonical variables

$(q_1, p_1, q_2, p_2) = (X_t, P_t, \Sigma_t, \frac{\Lambda_t}{2\Sigma_t} = \frac{\Lambda_t}{2q_2})$  we have the Langrangian

$$L(t) = -p_2\dot{q}_1 - p_2\dot{q}_2 + \frac{P_t^2}{2\mu} + \frac{p_2^2}{2\mu} - \frac{\mu\sigma^4}{8q_2^2} + \frac{k}{q_2^\alpha} + \langle U_0 \rangle \quad (17)$$

where  $k := \frac{g}{\alpha+1} \frac{1}{\sqrt{(\alpha+1)(2\pi)^\alpha}}$ . Since  $\delta L$  must be set equal to zero to obtain dynamical equations, we may equivalently consider the Lagrangian

$$L(t) = p_2\dot{q}_1 + p_2\dot{q}_2 - \left( \frac{p_1^2}{2\mu} + \frac{p_2^2}{2\mu} - \frac{\mu\sigma^4}{8q_2^2} + \frac{k}{q_2^\alpha} + \langle U_0 \rangle \right) = p_2\dot{q}_1 + p_2\dot{q}_2 - H(q, p) \quad (18)$$

with the Hamiltonian, recognized as the conserved total energy  $E_{\text{tot}}(t)$ ,

$$H(q, p) \triangleq \frac{p_1^2}{2\mu} + \frac{p_2^2}{2\mu} - \frac{\mu\sigma^4}{8q_2^2} + \frac{k}{q_2^\alpha} + \langle U_0 \rangle(q_1, q_2). \quad (19)$$

In this standard Hamiltonian system, identifying the Poisson conjugate momenta  $(p_1, p_2) = (P_t, \frac{\Lambda_t}{2\Sigma_t})$  and position variables  $(q_1, q_2) = (X_t, \Sigma_t)$ , the equations of motion are

$$\begin{bmatrix} \dot{q}_1 \\ \dot{p}_1 \\ \dot{q}_2 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} \nabla_{p_1} H \\ -\nabla_{q_1} H \\ \nabla_{p_2} H \\ -\nabla_{q_2} H \end{bmatrix} = \begin{bmatrix} \frac{p_1}{\mu} \\ -\langle \nabla U_0 \rangle \\ \frac{p_2}{\mu} \\ -\frac{\mu\sigma^4}{4q_2^3} + \frac{k\alpha}{q_2^{\alpha+1}} - q_2 \langle \nabla^2 U_0 \rangle \end{bmatrix} \quad (20)$$

because  $\frac{\partial \langle f \rangle}{\partial X_t}(X_t, \Sigma_t) = \langle \nabla f \rangle$ ,  $\frac{\partial \langle f \rangle}{\partial \Sigma_t}(X_t, \Sigma_t) = \Sigma_t \langle \nabla^2 f \rangle$ . If  $U_0(x) = -\frac{x^4}{4} - \frac{x^2}{2}$  then the equilibrium point is  $(q_{1*}, p_{1*}, q_{2*}, p_{2*}) = (0, 0, q_{2*}, 0)$  where

$$-\frac{\mu\sigma^4}{4q_{2*}^3} + \frac{k\alpha}{q_{2*}^{\alpha+1}} - q_{2*} \langle \nabla^2 U_0 \rangle(0, q_{2*}) = -\frac{\mu\sigma^4}{4q_{2*}^3} + \frac{k\alpha}{q_{2*}^{\alpha+1}} + q_{2*}(3q_{2*}^2 + 1) = 0 \quad (21)$$

which leads to the polynomial equation

$$-\frac{1}{4}\mu\sigma^4 q_{2*}^{\alpha+1} + k\alpha q_{2*}^3 + q_{2*}^{\alpha+5} + 3q_{2*}^{\alpha+7} = 0 \quad (22)$$

and for  $\alpha = 1$  leads to the polynomial equation

$$-\frac{1}{4}\mu\sigma^4 q_{2*}^2 + k\alpha q_{2*}^3 + q_{2*}^6 + 3q_{2*}^8 = 0. \quad (23)$$

The Jacobian of the dynamics above with  $U_0(x) = -\frac{x^4}{4} - \frac{x^2}{2}$ , at  $(X_*, P_*, q_*, p_*)$  is

$$J_* = \begin{bmatrix} 0 & \frac{1}{\mu} & 0 & 0 \\ -\langle \nabla^2 U_0 \rangle & 0 & -q_{2*} \langle \nabla^3 U_0 \rangle & 0 \\ 0 & 0 & 0 & \frac{1}{\mu} \\ -q_{2*} \langle \nabla^3 U_0 \rangle & 0 & -\frac{\partial^2 H}{\partial q^2} & 0 \end{bmatrix} \quad (24)$$

where  $\langle \nabla^3 U_0 \rangle(0, q_{2*}) = 0$  and

$$-\frac{\partial^2 H}{\partial q^2}(0, q_{2*}) = \frac{3\mu\sigma^4}{4q_{2*}^4} - \frac{k\alpha(\alpha+1)}{q_{2*}^{\alpha+2}} - \langle \nabla^2 U_0 \rangle(0, q_{2*}) - q_{2*}^2 \langle \nabla^4 U_0 \rangle(0, q_{2*}). \quad (25)$$

When  $\alpha = 1$ , numerical values for equilibrium  $q_2$  and Jacobian eigenvalues and eigenvectors are  $q_{2*}(\in \mathbb{R}^+) = 0.5844$  and  $\text{spec}(J_*) = 1.6267, -1.6267, 3.2053, -3.2053$ ,

$$v(J_*) = \begin{bmatrix} 0.5237 & -0.5237 & 0 & 0 \\ 0.8519 & 0.8519 & 0 & 0 \\ 0 & 0 & 0.2978 & -0.2978 \\ 0 & 0 & 0.9546 & 0.9546 \end{bmatrix}. \quad (26)$$

## 2 An Action Functional for Nonlinear Quadratic MFGs

Quadratic MFGs with nonlinear Langevin dynamics

$$dx_s = (b(x_s) + u(x_s))ds + \sigma dw_s, \quad x_0 = x \quad (27)$$

are associated with the forward backward coupled HJB-FP MF optimality system, with control cost matrix  $\mu$

$$p_t = -\partial_x((b - \frac{v_x}{\mu})p) + \frac{\sigma^2}{2}p_{xx}, \quad p(0, x) = p_0(x) \quad (28)$$

$$-v_t = q(x; p) - \frac{(v_x)^2}{2\mu} + v_x b + \frac{\sigma^2}{2}v_{xx}, \quad v(T, x) = \Phi(x) \quad (29)$$

which under the (exponential) value and (proportional) density transformed variables  $\phi = \exp(\frac{-v}{\sigma^2\mu})$ ,  $\Gamma = \frac{p}{\phi}$  introduced in [2] yield the PDE system

$$-\Gamma_t = \frac{q\Gamma}{\sigma^2\mu} + (\Gamma b)_x - \frac{\sigma^2}{2}\Gamma_{xx} \quad (30)$$

$$\phi_t = \frac{q\Gamma}{\sigma^2\mu} - \phi_x b - \frac{\sigma^2}{2}\phi_{xx}. \quad (31)$$

Note that in the absence of the passive drift, the PDEs are identical except for a negative sign, thus allowing one to use the nonlinear imaginary time Schrodinger PDE model along with the associated fixed point solution. The variational ansatz for the PDE system maybe plugged into a satisfactory action functional to yield an ODE optimality system on parameters in the ansatz. We therefore need to modify the proposed action in equation (2.28), pp 10, [1] to accomodate the additional terms due to the nonlinear drift. We propose including the additional the term  $\int_0^T dt \int dx \mu\sigma^2\phi_x b\Gamma$  for this purpose. The variation of this quantity is

$$\begin{aligned} \mu\sigma^2 \int_0^T dt \int dx (\phi + \delta\phi)_x b(\Gamma + \delta\Gamma) &\stackrel{O(\delta\phi, \delta\Gamma)}{=} \mu\sigma^2 \int_0^T dt \int dx \phi_x b\Gamma + \delta\phi_x b\Gamma + \phi_x b \delta\Gamma \\ &= \mu\sigma^2 \int_0^T dt \int dx \phi_x b\Gamma - (\Gamma b)_x \delta\phi + \phi_x b \delta\Gamma \end{aligned} \quad (32)$$

by restricting the class of variations  $\delta\phi(t, \cdot)$  such that  $\delta\phi(\cdot, \pm\infty) = 0$ . The required advection terms in the PDE system above are yielded. We note that there is no adhoc assumption made on the nonlinear drift (other than the usual Lipschitz and almost linear growth requirements for existence, uniqueness of SDE solutions) upto this point.

## 3 Integrator Systems theory for Nonlinear Stochastic Systems

Consider the optimal control problem (OCP)

$$\min_u J(u) := \mathbb{E} \left[ \int_0^T q(x_s, p(s, x_s))ds + \frac{R}{2}u^2 + \Phi(T, x_T) \right] \quad (33)$$

subject to the controlled nonlinear diffusion dynamics

$$dx_s = b(x_s)ds + u(s)ds + \sigma dw_s. \quad (34)$$

The associated semilinear Hamilton Jacobi Bellman (HJB) equation and Linear Fokker Planck (FP) equations are

$$-v_t = q - \frac{v_x^2}{2R} + v_x b + \frac{\sigma^2}{2} v_{xx} \quad (35)$$

$$p_t = -\partial_x \left( \left( b - \frac{\partial_x v}{R} \right) p \right) + \frac{\sigma^2}{2} p_{xx} \quad (36)$$

with the terminal time boundary condition  $v(T, x) = \Phi(T, x)$ . The HJB PDE has a linear representation, obtained using a Hopf Cole transform  $\phi := \exp(-v/\sigma^2 R)$  and the ad-hoc assumption  $\frac{1}{R} = \sigma^2$ . It is as follows

$$-\phi_t = -\frac{q\phi}{\sigma^2 R} + \phi_x b + \frac{\sigma^2}{2} \phi_{xx} \quad (37)$$

with the terminal time boundary condition  $\phi(T, x) = \exp(-v(T, x)/\sigma^2 R)$ . This PDE has a path integral representation [3] which is useful in computing the control [4–6]. However, the original paper on this control application titled 'A linear control theory for nonlinear stochastic systems' [4] computes path integral solutions to the above equation by *sampling from the nonlinear passive dynamics*. In what follows we will improve upon this innovation. It will be shown that the above equation may be further reduced from its advection-diffusion form to a pure diffusion PDE.

### 3.1 A new Hopf Cole transform

Let the primitive or indefinite integral of the passive dynamics be denoted  $\int b$ . We introduce a modified Hopf Cole transform  $\tilde{\phi} := \exp(-(v - R \int b)/\sigma^2 R)$  combined with the assumption  $\frac{1}{R} = \sigma^2$ . This leads to the following PDE representation of equation (35)

$$-\tilde{\phi}_t = -\frac{\tilde{q}\tilde{\phi}}{\sigma^2 R} + \frac{\sigma^2}{2} \tilde{\phi}_{xx} \quad (38)$$

with the terminal time boundary condition  $\tilde{\phi}(T, x) = \exp\left(-\left(v(T, x) - R(\int b)(T, x)\right)/\sigma^2 R\right)$  and where we denote the modified cost  $\tilde{q} := q + (R/2)b^2 - (\sigma^2 R/2)b_x$ . A verification proof is easily demonstrated by substitution. The calculations  $\tilde{\phi}_t = -\frac{\tilde{\phi}}{\sigma^2 R} v_t$ ,  $\tilde{\phi}_x = -\frac{\tilde{\phi}}{\sigma^2 R} (v_x + R\nu_x)$ ,  $\tilde{\phi}_{xx} = -\frac{\tilde{\phi}_x}{\sigma^2 R} (v_x + R\nu_x) - \frac{\tilde{\phi}}{\sigma^2 R} (v_{xx} + R\nu_{xx})$  and  $\frac{v_x^2}{2R} = \left(\frac{\sigma^4 R}{2} \left(\frac{\tilde{\phi}_x}{\tilde{\phi}}\right)^2 + \sigma^2 R \frac{\tilde{\phi}_x}{\tilde{\phi}} \nu_x + \frac{R}{2} \nu_x^2\right)$  are helpful.

### 3.2 Interpretation

Let us observe the case where the dynamics (34) are of overdamped Langevin type, that is,  $b = -\nu_x$ . It can be observed that the introduced transform is connected to an ergodic density in that case. If the value function is time independent, then the exponent in the proposed transform is recognizable in Gibbs distribution term  $\exp(-(v + R\nu)/\sigma^2 R)$ . The modification in the cost,  $\tilde{q} - q = (R/2)\nu_x^2 - (\sigma^2 R/2)\nu_{xx}$ , could be interpreted as some sort of Lagrangian [7], that is the difference between kinetic and potential energies associated with the nonlinear passive dynamics. However, in what follows, for the case of Langevin type dynamics (34), we provide a rigorous interpretation in terms of an alternate, fictitious, integrator system-OCP which is equivalent to the original nonlinear OCP (33), (34),  $b = -\nu_x$ .

The OCP (33), (34) is equivalent to the PDE optimality system given by equations (35), (36). Due to the transform introduced in the previous section, this system is equivalent to an optimality

system given by equations (38), (36). The way to interpret this formulation is to ask the question: *what is the integrator system-OCP directly associated with the system (38), (36)?* We answer this question by constructing this integrator system-OCP below.

Let us denote  $\tilde{\Phi} := \Phi(T, x) + R\nu(x)$ . Consider the OCP

$$\min_u J(u) := \mathbb{E} \left[ \int_0^T \tilde{q}(x_s, p(s, x_s)) ds + \frac{R}{2} u^2 + \tilde{\Phi}(T, x_T) \right] \quad (39)$$

subject to the controlled linear diffusion dynamics

$$dx_s = u(s) ds + \sigma dw_s. \quad (40)$$

The associated semilinear Hamilton Jacobi Bellman (HJB) equation and Linear Fokker Planck (FP) equations are

$$-\tilde{v}_t = \tilde{q} - \frac{\tilde{v}_x^2}{2R} + \frac{\sigma^2}{2} \tilde{v}_{xx} \quad (41)$$

$$p_t = - \left( -\frac{\tilde{v}_x}{R} p \right) + \frac{\sigma^2}{2} p_{xx} \quad (42)$$

with the terminal time boundary condition  $\tilde{v}(T, x) = \tilde{\Phi}(T, x)$ . It can then be seen from equations (35) and (41) that  $\tilde{v} = v + R\nu$  is a solution to equation (41). If we assume uniqueness of solutions to the system (41), (42) by a monotonicity assumption on  $\tilde{q}$ , then we can conclude by substituting  $\tilde{v}_x = v_x + R\nu_x$  into equation (42), that the PDE systems given by (35), (36) and (41), (42) respectively are equivalent. Therefore, **the nonlinear system OCP given by (33), (34),  $b = -\nu_x$  is equivalent to the integrator system-OCP (39), (40).** (We need to verify that this is not previously known) Finally, (1) this fact is true for the case of multidimensional dynamics (replace:  $\int b$  instead of  $-\nu_x$  and  $\text{div} \cdot b$  instead of  $b_x$ ) and

(2) the assumption of Langevin dynamics for this fact can be discarded by proving this fact for second order (mechanical/electrical engineering) systems (extend by incorporating the second order dynamics,  $dx_p = x_v dt$ , which will appear via the Liouville operator in the PDE system,  $\tilde{\phi} := \exp(-(x_v^2/2 + v - R \int b)/\sigma^2 R)$ ).

### 3.3 Applications

#### 3.3.1 Nonlinear Control: Sampling from Brownian Motion

Consider a robotic system such as an autonomous car in an aggressive driving environment [8]. The path integral control has been demonstrated with some degree of success in this prior work as well as others. The critical quality of all sampling based algorithms related to the PDE formulation of the nonlinear OCP problems is that the PDE solution and hence the feedback optimal control is computed only locally. This reduces unsolved issues related to the *curse of dimensionality*.

Since we have constructed an integrator system-OCP which is equivalent to a given nonlinear OCP, as in section 3.2, our new PDE fomulation can find direct applications for nonlinear stochastic problems in robotics in two ways. In order to explain this simply, let use assume that the cost  $q$  is independent of the density. This assumption decouples the HJB and FP PDEs and therefore allows us to compute the optimal control at  $(t, x)$  by solving only the HJB equation.

(1) The first application of the fictitious OCP representing the nonlinear OCP is that the related path integral solution to equation (38) can be solved by sampling from an integrator system. This means that we may solve a nonlinear OCP by *sampling from a passive linear system instead of a*

nonlinear system, thus saving on the computational cost. *The purport is that importance sampling can be eliminated in the path integral control of nonlinear stochastic systems.*

(2) Another approach to solve this problem would be to use the structure of equation (38), which is a simple heat PDE to solve the control problem. Spectral methods could allow reduction of this problem into a linear systems problem involving back propagation of linear ODEs to solve the PDE over a domain instead of at just a point (the sampling approach). This could be a step towards solving the *curse of dimensionality* problem in nonlinear stochastic OCPs.

### 3.3.2 Mathematical Physics: Modeling MF Systems

A challenging problem in mathematical physics is to build MF models for multiagent systems in which there are *rational, infinitely many anonymous agents who base decisions on information about a collection of the agents*. These key properties constitute the theory of MFGs and are useful in modeling and therefore control of (why is this important?) large scale, self organizing systems and the notion of *emergent behaviour*.

Flocking of agents with nonlinear mobilities is such an example of self organizing system [9, 10]. In this section we show how the fact presented in section 3.1 and its interpretation 3.2 can be used to obtain an optimal control model of the uncontrolled flocking model presented in [9]. Let us denote the nonlinear Langevin dynamics of the uncontrolled dynamics by  $b(p)$ . The question we would like to address is: *can we obtain an OCP given by (39), (40) which will result in closed loop dynamics of the original system which has the passive drift  $b(p)$ ?*

An educated guess is to set  $\tilde{q} := (1/2)b^2 - (\sigma^2/2)b_x$  and set a high value for  $R$  which will cause  $v_x/R$  to be a small number at all  $(t, x)$ .  $v$  here is governed by equation (35). Due to this, we conclude from our interpretation of the fictitious OCP being constructed in section 3.2, the closed loop control will be  $-\tilde{v}_x/R = b - v_x/R \approx b$ . (Needs more work, maybe  $b(p) = b_1(x) + b_2(p) = -\tilde{v}_x/R = b_1(x) - v_x/R$  and get  $b_2 \equiv -v_x/R$ )

### 3.3.3 Learning Control: MF Control Computation from Ergodic samples

The work by Manfred, O. [11] on learning control using ergodic controlled samples can be immediately extended to MFGs using the fact presented in this section.

### 3.3.4 Soliton theory: MFGs

In the next section we will use the PDE representation in section 3.1 in this section to create a soliton model for highly attractive agents in MFGs.

## 4 Soliton theory for MFGs of agents with Nonlinear Dynamics

The MF approach to analyze populations of agents with nonlinear passive dynamics has been previously used to model self organized or controlled flocking in [9, 10] and crowds in [12]. We will construct a theory for solitons for a class of nonlinear dynamics appearing in such problems. Consider the OCP (33) under the dynamics (34),  $b = -\nu_x$ . The MF optimality system is comprised by equations (35), (36). We introduced the transform  $\tilde{\phi} := \exp(-(v + R\nu)/\sigma^2 R)$  leading to the PDE representation (38) of the equation (35). To obtain consistent PDE equations on variables  $(\tilde{\phi}, p)$  in the optimality system we may replace  $v$  in (36) in terms of  $\tilde{\phi}$  from the introduced transform. However, we introduce yet another transform on the density which will lead to identical form for

both representative PDEs in the optimality system. We *hermitize* [1] the density by  $\tilde{\Gamma} := \frac{p}{\tilde{\phi}}$ . From equation (36) we have

$$-\Gamma_t = \frac{\tilde{q}\tilde{\Gamma}}{\sigma^2 R} - \frac{\sigma^2}{2}\tilde{\Gamma}_{xx} \quad (43)$$

with the initial time boundary condition  $\tilde{\Gamma}(0, x) = \frac{p}{\tilde{\phi}}(0, x)$ . The derivatives  $p_t = \tilde{\Gamma}_t \tilde{\phi} + \tilde{\Gamma} \tilde{\phi}_t$ ,  $p_x = \tilde{\Gamma}_x \tilde{\phi} + \tilde{\Gamma} \tilde{\phi}_x$ ,  $p_{xx} = \tilde{\Gamma}_{xx} \tilde{\phi} + 2\tilde{\Gamma}_x \tilde{\phi}_x + \tilde{\Gamma} \tilde{\phi}_{xx}$ , the calculation  $(\sigma^2 \ln \tilde{\phi})_x p = \sigma^2 \frac{\tilde{\phi}_x}{\tilde{\phi}} \tilde{\Gamma} \tilde{\phi} = \sigma^2 \tilde{\phi}_x \tilde{\Gamma}$  and equation (38) are useful in obtaining the above equation. Thus, the MFG given by the nonlinear OCP (33) subject to dynamics (34),  $b = -\nu$  with the optimality system (35), (36) has an equivalent optimality system given by the Poisson PDEs (38), (43). In what follows we will exploit the identical PDE forms in this new representation in order to construct soliton solutions to the optimality system.

## 4.1 Variational Representation of MF Optimality system

The action or functional

$$A(\tilde{\phi}, \tilde{\Gamma}) := \int_0^T dt \int_{\mathbb{R}^d} dx \left[ -\frac{R\sigma^2}{2} (\tilde{\phi}(\tilde{\Gamma}_t) - (\tilde{\phi}_t)\tilde{\Gamma}) - \frac{R\sigma^4}{2} \tilde{\phi}_x \cdot \tilde{\Gamma}_x - \tilde{\phi} \tilde{q} \tilde{\Gamma} \right] \quad (44)$$

introduced in [1], has the variational property that  $\delta A / \delta \tilde{\Gamma} = 0$  and  $\delta A / \delta \tilde{\phi} = 0$  are equivalent to the equations (38), (43). Therefore this action provides a formal representation of the optimality system in the transient state. Denoting  $q_\nu := (R/2)\nu_x^2 - (\sigma^2 R/2)$  we have  $\tilde{q} = q + q_\nu$ . We now assume  $q = U(x) + f(p)$  with  $f$  being a nonlinear function of the density. Then the simplified form of the action is

$$A(\tilde{\phi}, \tilde{\Gamma}) := \int_0^T dt \int_{\mathbb{R}^d} dx \left[ -\frac{R\sigma^2}{2} (\tilde{\phi}(\tilde{\Gamma}_t) - (\tilde{\phi}_t)\tilde{\Gamma}) - \frac{R\sigma^4}{2} \tilde{\phi}_x \cdot \tilde{\Gamma}_x - \tilde{\phi}(U + q_\nu)\tilde{\Gamma} - F(\tilde{\phi}\tilde{\Gamma}) \right] \quad (45)$$

where  $F(p) = \int^p f(p) dp$ . Remember that  $f(p)$  will be chosen to be a strictly decreasing function later. We identify the following terms as kinetic, potential and interaction energies

$$E_{\text{kin}} = -\frac{R\sigma^4}{2} \tilde{\phi}_x \cdot \tilde{\Gamma}_x \quad (46)$$

$$E_{\text{pot}} = -\langle U + q_\nu \rangle \quad (47)$$

$$E_{\text{int}} = -\int_{\mathbb{R}^d} dx F(p) \quad (48)$$

with the total energy defined as  $E_{\text{tot}} = E_{\text{kin}} + E_{\text{pot}} + E_{\text{int}}$ . The negative signs in the last two expressions have been chosen so because of our monotonicity assumptions on  $F(p)$ . The argument using Noether's theorem [13] in [1] applies to the part of the integrand in the action which is not explicitly time dependent, so that  $E_{\text{tot}}$  is conserved (which I do not understand) implying

$$A(\tilde{\phi}, \tilde{\Gamma}) = \int_0^T dt \int_{\mathbb{R}^d} dx \left[ -\frac{R\sigma^2}{2} (\tilde{\phi}(\tilde{\Gamma}_t) - (\tilde{\phi}_t)\tilde{\Gamma}) \right] + E_{\text{tot}} T. \quad (49)$$

## 4.2 Gaussian Ansatz and Soliton Dynamics

We revert to the notation  $\mu = R$  for the control cost in this section. Since equations (38), (43) have an identical form with an asymmetry only in the direction of time, we introduced the following



Gaussian ansatz [1]

$$\tilde{\phi}(t, x) = \exp\left(\frac{-\gamma_t + P_t \cdot x}{\mu\sigma^2}\right) \frac{1}{(2\pi\Sigma_t^2)^{\frac{1}{4}}} \exp\left(-\frac{(x - X_t)^2}{(2\Sigma_t^2)^2} \left(1 - \frac{\Lambda_t}{\mu\sigma^2}\right)\right) \quad (50)$$

$$\tilde{\Gamma}(t, x) = \exp\left(\frac{+\gamma_t + P_t \cdot x}{\mu\sigma^2}\right) \frac{1}{(2\pi\Sigma_t^2)^{\frac{1}{4}}} \exp\left(-\frac{(x - X_t)^2}{(2\Sigma_t^2)^2} \left(1 + \frac{\Lambda_t}{\mu\sigma^2}\right)\right) \quad (51)$$

by assuming that the population of agents in the MFG rapidly concentrates around its mean value,  $X_t$  and moves as a static Gaussian shaped formation with covariance  $\Sigma_t$ . The Gaussian density can be therefore readily obtained as  $p(t, x) = \frac{1}{(2\pi\Sigma_t^2)^{\frac{1}{2}}} \exp\left(-\frac{(x - X_t)^2}{(2\Sigma_t^2)^2}\right)$ .

We will now consider exclusively the case  $d = 1$ . Applying this ansatz in the action will provide the ODE soliton dynamics on the variables  $(X_t, P_t, \Sigma_t, \Lambda_t)$ . This gives us the Lagrangian form of the action

$$A[\tilde{\phi}, \tilde{\Gamma}] = \int_0^T dt L(t) \quad (52)$$

where  $L(t) = L_\tau(t) + E_{\text{tot}}$  and the expressions for  $L_\tau$ ,  $E_{\text{kin}}$  are given in equations (B.3), (B.4) in [1].

#### 4.2.1 Polynomial cost

Now, we choose  $f(\cdot) = gp^\alpha$ ,  $\alpha > 0$ . Modifying the first term in the Lagrangian as  $\int_0^T \dot{P}_t X_t = X_t P_t \Big|_0^T - \int_0^T \dot{X}_t P_t$  we arrive at a Hamiltonian system for canonical variables defined as follows. Using the canonical variables  $(q_1, p_1, q_2, p_2) = (X_t, P_t, \Sigma_t, \frac{\Lambda_t}{2\Sigma_t} = \frac{\Lambda_t}{2q_2})$  we have the Lagrangian

$$L(t) = -p_2 \dot{q}_1 - p_2 \dot{q}_2 + \frac{P_t^2}{2\mu} + \frac{p_2^2}{2\mu} - \frac{\mu\sigma^4}{8q_2^2} + \frac{k}{q_2^\alpha} - \langle U + q_\nu \rangle \quad (53)$$

where  $k := \frac{g}{\alpha+1} \frac{1}{\sqrt{(\alpha+1)(2\pi)^\alpha}}$ . Since  $\delta L$  must be set equal to zero to obtain dynamical equations, we may equivalently consider the Lagrangian

$$L(t) = p_2 \dot{q}_1 + p_2 \dot{q}_2 - \left( \frac{p_1^2}{2\mu} + \frac{p_2^2}{2\mu} - \frac{\mu\sigma^4}{8q_2^2} + \frac{k}{q_2^\alpha} - \langle U + q_\nu \rangle \right) = p_2 \dot{q}_1 + p_2 \dot{q}_2 - H(q, p) \quad (54)$$

with the Hamiltonian, recognized as the conserved total energy  $E_{\text{tot}}(t)$ ,

$$H(q, p) \triangleq \frac{p_1^2}{2\mu} + \frac{p_2^2}{2\mu} - \frac{\mu\sigma^4}{8q_2^2} + \frac{k}{q_2^\alpha} - \langle U + q_\nu \rangle(q_1, q_2). \quad (55)$$

In this standard Hamiltonian system, identifying the Poisson conjugate momenta  $(p_1, p_2) = (P_t, \frac{\Lambda_t}{2\Sigma_t})$  and position variables  $(q_1, q_2) = (X_t, \Sigma_t)$ , the equations of motion are

$$\begin{bmatrix} \dot{q}_1 \\ \dot{p}_1 \\ \dot{q}_2 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} \nabla_{p_1} H \\ -\nabla_{q_1} H \\ \nabla_{p_2} H \\ -\nabla_{q_2} H \end{bmatrix} = \begin{bmatrix} \frac{p_1}{\mu} \\ \langle \nabla(U + q_\nu) \rangle \\ \frac{p_2}{\mu} \\ -\frac{\mu\sigma^4}{4q_2^3} + \frac{k\alpha}{q_2^{\alpha+1}} - q_2 \langle \nabla^2(U + q_\nu) \rangle \end{bmatrix} \quad (56)$$

because  $\frac{\partial \langle f \rangle}{\partial X_t}(X_t, \Sigma_t) = \langle \nabla f \rangle$ ,  $\frac{\partial \langle f \rangle}{\partial \Sigma_t}(X_t, \Sigma_t) = \Sigma_t \langle \nabla^2 f \rangle$  for any smooth function  $f$ , due to the Gaussian solution ansatz. The uncoupled dynamics can be obtained by applying the Dirac delta

limit of the Gaussian density to the unevaluated expectations in the above dynamical equations of motion [how?](#).

Specifically, if the nonlinear dynamics are occurring under a passive bistable potential  $\nu = \frac{x^4}{4} - \frac{x^2}{2}$  with an additional potential  $U_0 = \frac{(x-1)^2}{2}$  then we compute the coupled soliton dynamics here. Note that  $q_\nu = \frac{x^6}{2} - x^4 + x^2 \left( \frac{1-3\sigma^2}{2} \right) + \frac{\sigma^2}{2}$ ,  $\nabla q_\nu = 3x^5 - 4x^3 + x(1-3\sigma^2)$ ,  $\nabla^2 q_\nu = 15x^4 - 12x^2 + (1-3\sigma^2)$  and  $\nabla U_0 = (x-1)^2$ ,  $\nabla^2 U_0 = 1$ .

### 4.3 Verification of Mean Dynamics by Ito Stochastic Calculus

I can do this. This will show that the mean dynamics obtained directly from the optimally controlled SDE (34) result in the mean dynamics given by our Gaussian ansatz. Is this necessary? This is the equivalent of using Ehrenfest relations for verification, which had no further use in [1]. Further, it is excessively complicated to thus verify the variance dynamics, which will display inconsistency in the nature of the presented results

## References

- [1] Ullmo D., Swiecicki I., and Gobron T. Quadratic mean field games. *arXiv:1708.07730*, 2017.
- [2] O. Guéant. A reference case for mean field games models. *Journal de Mathématiques Pures et Appliquées*, 92(3):276–294, 2009.
- [3] W.H. Fleming. Exit probabilities and optimal stochastic control. *Applied Math. Optim*, 9:329–346, 1978.
- [4] H. J. Kappen. Linear theory for control of nonlinear stochastic systems. *Phys. Rev. Lett.*, 95:200201, Nov 2005.
- [5] H. J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, 11:P11011, 2005.
- [6] E. Theodorou. *Iterative Path Integral Stochastic Optimal Control: Theory and Applications to Motor Control*. PhD thesis, University of Southern California, Los Angeles, CA, USA, 2011.
- [7] L. Elsgolc. *Calculus of Variations*. Dover, Mineola, New York, 2007.
- [8] Grady Williams, Brian Goldfain, Paul Drews, James M. Rehg, and Evangelos A. Theodorou. Autonomous racing with autorally vehicles and differential games. *CoRR*, abs/1707.04540, 2017.
- [9] Alethea BT Barbaro, José A Cañizo, José A Carrillo, and Pierre Degond. Phase transitions in a kinetic flocking model of cucker–smale type. *Multiscale Modeling & Simulation*, 14(3):1063–1088, 2016.
- [10] P. Grover, K. Bakshi, and Theodorou E. A. A homogeneous model for flocking. *arXiv:1803.05250v3*, 2018.
- [11] Manfred Opper. An estimator for the relative entropy rate of path measures for stochastic differential equations. *Journal of Computational Physics*, 330:127 – 133, 2017.

- [12] Martin Burger, Marco Di Francesco, Peter A. Markowich, and Marie-Therese Wolfram. Mean field games with nonlinear mobilities in pedestrian dynamics. *Discrete & Continuous Dynamical Systems - B*, 19(1531-3492-2014-5-1311):1311, 2014.
- [13] J. Papastavridis. *Analytical Mechanics*. World Scientific, reprint edition, 2014.