# An adaptive high-order minimum action method

Xiaoliang Wan

Department of Mathematics and Center for Technology & Computation, Louisiana State University

## Abstract

In this work, we present an adaptive high-order minimum action method for dynamical systems perturbed by small noise. We use the $hp$ finite element method to approximate the minimal action path and nonlinear conjugate gradient method to solve the optimization problem given by the Freidlin-Wentzell least action principle. The gradient of the discrete action functional is obtained through the functional derivative and the moving mesh technique is employed to enhance the approximation accuracy. Numerical examples are given to demonstrate the efficiency and accuracy of the proposed numerical method.

## Problem description

We consider random perturbations of dynamical systems. Let the random process $X_t = X(t) : \mathbb{R}_+ \to \mathbb{R}^n$ defined by the following stochastic ordinary differential equation (SODE):

$$dX_t = b(X_t)dt + \sqrt{\varepsilon}dW_t, \tag{1}$$

where $W_t$ is a standard Wiener process in $\mathbb{R}^n$ and $\varepsilon$ is a small positive parameter. Let $\phi(t) \in \mathbb{R}^n$ be an absolutely continuous function defined for $t \in [0, T]$. The Wentzell-Freidlin theory tells us that the probability of $X(t)$ passing through the $\delta$-tube about $\phi$ on $[0, T]$ is

$$\Pr(\rho(X, \phi) < \delta) \approx \exp(-\frac{1}{\varepsilon}S_T(\phi)), \tag{2}$$

with $\rho(\phi, \varphi) = \sup_{t\in[0,T]} |\phi(t) - \varphi(t)|$, $|\cdot|$ indicates the $\ell_2$ norm in $\mathbb{R}^n$, and $S_T(\phi)$ is the action functional of $\phi$ on $[0, T]$, defined as

$$S_T(\phi) = \frac{1}{2}\int_0^T L(\dot{\phi}, \phi)dt, \tag{3}$$

where $L(\dot{\phi}, \phi) = |\dot{\phi} - b(\phi)|^2$. In general, we have the following large deviation principle

$$\lim_{\varepsilon\to 0} \varepsilon \log \Pr(X \in A) = -\min_{\phi\in A} S_T(\phi), \tag{4}$$

where $A$ is a particular set of random events. Thus the basic contribution to $\Pr(X \in A)$ is given by the neighborhood of the minimum of $S_T(\phi)$ when $\varepsilon$ is small enough. The minimizer $\phi^*$, which satisfies $S_T(\phi^*) = \min_{\phi\in A} S_T(\phi)$ is also called the "minimal action path" (MAP).

## Available minimum action methods

We here only focus on numerical methods for general (non-gradient) dynamical systems, which are usually called minimum action methods (MAM).

- **The original MAM**. The original MAM was proposed in [1] coupling the finite difference discretization in time and an L-BFGS optimization solver.
- **The adaptive MAM**. The adaptive MAM was proposed in [5]. The authors observed that a sufficiently large time interval can resolve the MAP defined on an infinite time interval and the main reason that the original MAM may converge poorly is due to the slow dynamics in the transition regions which makes the uniform temporal discretization dramatically skewed with respect to the arc length. A moving mesh technique is then used to redistribute the grid points to make them more uniform according to the arc length, which improves significantly both accuracy and efficiency.
- **The geometric MAM**. The geometric MAM was proposed in [2]. The authors directly considered discretization with respect to a parametrization variable (the arc length) instead of time. However, the problem needs to be reformulated with respect to the parametrization variable, i.e., in the space of curves.

All available MAMs are focused on finite difference discretization without addressing parallel computing, which make them not straightforward to couple with a finite element solver of the original dynamical system modeled by a PDE and not appropriate for large-scale simulations.

## A variational approach based on the $hp$ finite element approximation

Suppose that we choose a finite dimensional approximation space spanned by $\{\psi_i(t)\}_{i=1}^M$ such that

$$\phi_h(t) = \sum_{i=1}^M \phi_i\psi_i(t), \quad \delta\phi_h(t) = \sum_{i=1}^M \delta\phi_i\psi_i(t), \tag{5}$$

where $\phi_i, \delta\phi_i \in \mathbb{R}^n$. We then have

$$\delta S_T(\phi_h) = \langle\frac{\delta S_T}{\delta\phi_h}, \delta\phi_h\rangle_t = \sum_{i=1}^M \langle\frac{\delta S_T}{\delta\phi_h}, \delta\phi_i\psi_i(t)\rangle_t. \tag{6}$$

Consider the particular choice of $\delta\phi_h$, whose coefficients are equal to zero except the $j$-th component $\delta\phi_{i,j}$ of $\delta\phi_i$. We then obtain

$$\delta S_T(\phi_h) = \langle\frac{\delta S_T}{\delta\phi_h}, \psi_i(t)\mathbf{e}_j\rangle_t\delta\phi_{i,j}, \tag{7}$$

which implies that

$$(\nabla S_T(\phi_h))_{k(i,j)} = \frac{\partial S_T}{\partial\phi_{i,j}} = \langle\frac{\delta S_T}{\delta\phi_h}, \psi_i(t)\mathbf{e}_j\rangle_t, \tag{8}$$

where $\mathbf{e}_j \in \mathbb{R}^n$ is the unit Euclidean vector such that its $j$th component is 1 and the rest components are zero, $k(i, j)$ is a global index uniquely determined by $i = 1, \ldots, M$ and $j = 1, \ldots, n$. Note that $\nabla S_T$ is usually required by an efficient optimization algorithm, such as nonlinear conjugate gradient method, L-BFGS, etc. For the stochastic ODE system, we have

$$(\nabla S_T(\phi_h))_{k(i,j)} = \langle\dot{\phi}_h - b(\phi_h), \dot{\psi}_i(t)\mathbf{e}_j - \hat{b}(\phi_h)\psi_i(t)\mathbf{e}_j\rangle_t, \tag{9}$$

where

$$b(\phi + \delta\phi) = b(\phi) + \hat{b}(\phi)\delta\phi + O(\delta^2\phi).$$

It is observed that the computation of $\partial S_T/\partial\phi_{i,j}$ is an integration problem, which can be easily dealt with by using Gauss-type quadrature formulas. Such a strategy can also be easily generalized to deal with stochastic PDE. Once the gradient of the action functional is obtained, we use the nonlinear conjugate gradient (CG) method to solve the optimization problem to get the MAP $\phi_h^*$.

## Parallelism of the algorithm

On the reference element $R$, we assume that the approximation space consists of linear combinations of the following basis functions:

$$\hat{\psi}_i(\tau) = \begin{cases} \frac{1-\tau}{2} & i = 0, \\ \frac{1-\tau}{2}\frac{1+\tau}{2}P_{i-1}^{1,1}(\tau), & 0 < i < p, \\ \frac{1+\tau}{2} & i = p, \end{cases} \tag{10}$$

where $P_i^{1,1}$ denote orthogonal Jacobi polynomials of degree $i$ with respect to the weight function $(1-\tau)(1+\tau)$. $\hat{\psi}_0(\tau)$, and $\hat{\psi}_p(\tau)$ are consistent with linear finite element basis, and $\hat{\psi}_i(\tau)$, $0 < i < p$, are introduced for high-order approximation. Note that $\hat{\psi}_i(\pm 1) = 0$ for $0 < i < p$. We call $\hat{\psi}_0(\tau)$ and $\hat{\psi}_p(\tau)$ boundary modes, and $\hat{\psi}_i(\tau)$, $0 < i < p$, interior modes. We then identify the following two-level parallelism:

**Level one** - Along the time direction, we use MPI for parallelization, where only the information about boundary modes needs to be exchanged between two adjacent computation nodes.

**Level two** - On each computation node, we use OpenMP to parallelize the computation of the gradient, which is more important for a stochastic PDE.

Thus parallel MAM is consistent with the current high performance computing (HPC) architecture.

## Numerical results

We consider the following example, for which the MAP can be obtained explicitly:

$$\begin{cases} dx = -\partial_x V(x, y)dt + \sqrt{\varepsilon}dW_t^x \\ dy = -\partial_y V(x, y)dt + \sqrt{\varepsilon}dW_t^y \end{cases} \tag{11}$$

where the potential $V(x, y)$ is

$$V(x, y) = (1 - x^2 - y^2)^2 + y^2/(x^2 + y^2). \tag{12}$$

The dynamical system has two stable fixed points $a_1 = (-1, 0)$ and $a_2 = (1, 0)$, which are local minima of the potential $V(x, y)$. We consider the MAP in the upper half-plane connecting $a_1$ and $a_2$ through the saddle point $a_3 = (0, 1)$. Then the explicit form of this MAP is the upper branch of the unit circle: $x^2 + y^2 = 1$. The exact action functional is $2 \times (V(a_3) - V(a_1)) = 2$. In figure 1 we demonstrate the $h$- and $p$-converge of our algorithm, where the moving mesh technique is used to make the time mesh nearly uniform with respect to the arc length.

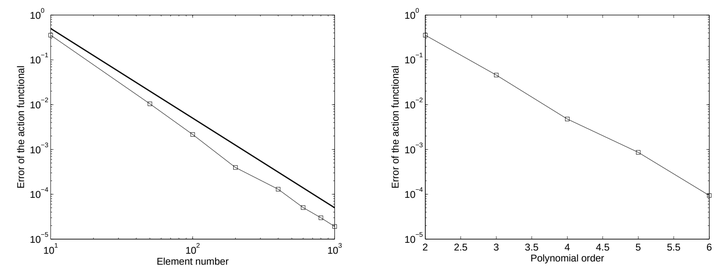


**Figure 1: Left: $h$-convergence. Right: $p$-convergence.**

We subsequently look at the random perturbations of Kuramoto-Sivashinsky (K-S) equation

$$\begin{cases} u_t + 4u_{xxxx} + \alpha\left[u_{xx} + \frac{1}{2}(u_x)^2\right] = \sqrt{\varepsilon}\dot{W}(x, t), \\ u(x, 0) = u_0(x), \quad u(x, t) = u(x + 2\pi, t), \end{cases} \tag{13}$$

where $\alpha$ is a bifurcation parameter. The Euler-Lagrange equation of the action functional

$$(-\partial_t^2 + \alpha^2\partial_x^4 + 8\alpha\partial_x^6 + 16\partial_x^8)v + \mathcal{N}(v) = 0, \tag{14}$$

where $\mathcal{N}(v)$ is the non-linear part. From the linear part of the Euler-Lagrange equation (14), we see that the ratio between the largest and smallest eigenvalues is $C_k = (\alpha k^2 - 4k^4)^2/(\alpha - 4)^2$, where $k \in \mathbb{Z}$ and the smallest absolute value of $k$ is equal to 1 since the mean mode is taken out. When 32 Fourier modes are employed, $C_{16} = O(10^7)$, which implies that the condition number of the Hessian can be very large. Thus an efficient preconditioner is necessary to improve the convergence of the nonlinear CG iteration. We use the inverse of the linear part of the Euler-Lagrange equation, i.e. $(-\partial_t^2 + \alpha^2\partial_x^4 + 8\alpha\partial_x^6 + 16\partial_x^8)^{-1}$, as the preconditioner for the nonlinear CG method, and the effectiveness is shown in figure 2.

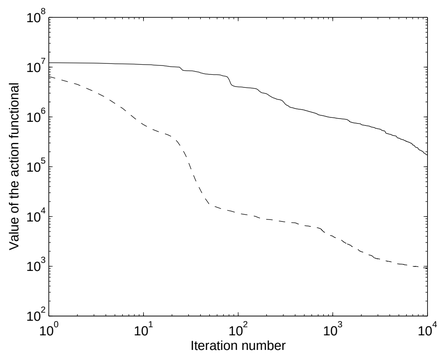


**Figure 2: Effectiveness of the pre-conditioner.**

## References

[1] W. E, W. Ren and E. Vanden-Eijnden, *Minimum action method for the study of rare events*, Commun. Pure Appl. Math., **57** (2004), 637-565.

[2] M. Heymann and E. Vanden-Eijnden, *The geometric minimum action method: A least action principle on the space of curves*, Commun. Pure Appl. Math., **61** (2008), 1052-1117.

[3] X. Wan, *An adaptive high-order minimum action method*, J. Compt. Phys., **230** (2011), 8669-8682.

[4] X. Wan, X. Zhou and W. E, *Study of the noise-induced transition and the exploration of the configuration space for the Kuramoto-Sivashinsky equation using the minimum action method*, Nonlinearity, **23** (2010), 475-493.

[5] X. Zhou, W. Ren and W. E, *Adaptive minimum action method for the study of rare events*, J. Chem. Phys., **128** (2008), 104111.

†**Address for correspondence: xiaolian@math.princeton.edu**