Stochastic Variational Integrators

Jeremy Schmitt

Variational integrators are numerical geometric integrator's derived from discretizing Hamilton's principle. They are symplectic integrators that exhibit excellent long-term stability, near-energy preservation and preserve many invariants of the underlying dynamical system. In particular, it has been shown that variational integrators perform very well when computing statistics of the system such as Poincare sections or instantaneous temperature. These reasons and more spurred the extension of variational integrators to stochastic differential equations. Symplectic integrators have already played an important role in stochastic simulations via the Hybrid/Hamiltonian Monte Carlo method, but have only recently been extended as direct stochastic integrators. In this paper we examine the performance of stochastic variational integrators through numerical experimentation, and compare them with standard methods for stochastic differential equations.

DISCRETE MECHANICS

Discrete mechanics is a discretization of classical mechanics that can be used to approximate continuous systems or even to inform the continuous theory. First, we review and define a continuous framework for classical mechanics, then using this framework we define discrete mechanics and variational integrators.

Let Q be a differentiable manifold, hereafter called the **configuration manifold**. The configuration manifold along with a differentiable function $L : TQ \to \mathbb{R}$ defines a **Lagrangian system**. Define the **path space** as the set of C^2 curves $\mathcal{C}(Q) = \{q : [0,T] \to Q | q \text{ is a } C^2 \text{ curve }\}$. Then the dynamics of the Lagrangian system are determined by $q \in \mathcal{C}(Q)$ such that q is a critical point of the **action** functional,

$$\mathbf{S}(q) = \int_0^T L(q(t), \dot{q}(t))dt.$$

All such q must satisfy the Euler–Lagrange equations,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0.$$

However, this second-order differential equation, under the right assumptions, can be written as a particular system of first-order equations exhibiting many wonderful features, not the least of which is it's symmetry. To make this precise, we first define the fiber derivative $\mathbb{F}L : TQ \to T^*Q$ by,

$$\mathbb{F}L(v) \cdot w = \frac{d}{ds}|_{s=0} L(v + sw),$$

where $v, w \in T_qQ$. We call $\mathbb{F}L$ the **Legendre Transform**. For a finite-dimensional configuration manifold this has the convienient local coordinate representation as,

$$\mathbb{F}L : (q, \dot{q}) \to (q, \frac{\partial L}{\partial \dot{q}}) = (q, p).$$

We define $p = \frac{\partial L}{\partial \dot{q}}$ as the **momentum**. Assuming $\mathbb{F}L$ is a diffeomorphism, define the **Hamiltonian** $H : T^*Q \to \mathbb{R}$ as,

$$H(q,p) = p\dot{q} - L(q,\dot{q})|_{p=\frac{\partial L}{\partial \dot{q}}}.$$

Under this assumption, the Euler–Lagrange equations are equivalent to the system of first-order differential equations known as Hamilton's equations,

$$\dot{q} = \frac{\partial H}{\partial p} \quad \dot{p} = -\frac{\partial H}{\partial q}.$$

If the Lagrangian is not explicitly a function of time, then the Hamiltonian is an invariant of the system(i.e. $\frac{dH}{dt} = 0$) . In specific mechanical cases this exactly represents the conservation of energy of the underlying system. Standard numerical methods fail to capture such qualitative aspects of the dynamical system, generally at the cost of stability of the numerical algorithm. As an example of this phenomenon, consider the simple pendulum with Hamiltonian,

$$H(p,q) = \frac{1}{2}p^2 - 9.8cos(q)$$

and corresponding Euler–Lagrange equation

$$\ddot{q} = -9.8sin(q).$$

Figure 1 compares the level sets of the Hamiltonian for the exact simple pendulum model, a simulation using a variational integrator, and a simulation using Taylor's method. It is clear that Taylor's method completely fails to accurately approximate the true trajectory in phase space. On the other hand the variational integrator shows an excellent simulation of the true trajectory over a long time interval. To discuss variational integrators we first need to define discrete mechanics.

Discrete mechanics is defined using the configuration space, Q, but instead of the state space T^*Q define the **discrete State Space** as $Q \times Q$(which is locally isomorphic to T^*Q). The **discrete Lagrangian** function is given by $L_d : Q \times Q \to \mathbb{R}$. Given an increasing sequence of times $\{t_k = kh | k = 0, \ldots, N\} \subset \mathbb{R}$ define the

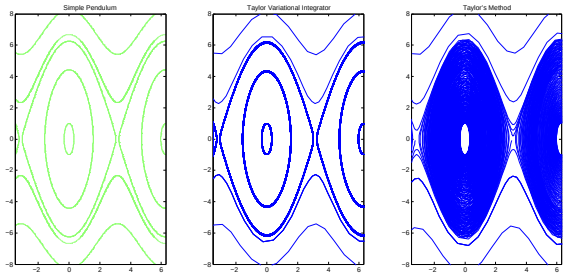

FIG. 1. The first plot on the left shows the level sets of the Hamiltonian for the Simple Pendulum. The middle plot shows the simulated level sets using a 2^{nd} order Taylor variational integrator and the right plot was simulated using the 2^{nd} order Taylor's method. Both simulations used a step size of $h = 0.1$ over a time interval [0,2000].

discrete Path Space to be $\mathcal{C}_d(Q) = \{q_d : \{t_k\}_{k=0}^N \to Q\}$. The **discrete Action Sum**, $\mathbb{S}_d : \mathcal{C}_d(Q) \to \mathbb{R}$, is defined as $\mathbb{S}_d(q_0, q_1, \ldots, q_N) = \sum_{i=0}^{N-1} L_d(q_i, q_{i+1})$. The **discrete Hamilton's principle** defines $q \in \mathcal{C}_d(Q)$ as a **discrete motion** if it satisfies $\delta\mathbb{S}_d = 0$. This leads to the **discrete Euler − Lagrange** equations, $D_2L_d(q_{k-1}, q_k) + D_1L_d(q_k, q_{k+1}) = 0$, which are equivalent to the **implicit discrete Euler − Lagrange** equations, $p_k = -D_1L_d(q_k, q_{k+1})$ and $p_{k+1} = D_2L_d(q_k, q_{k+1})$. Using these implicit equations and given (q_k, p_k), we can implicitly solve for (q_{k+1}, p_{k+1}), which gives rise to the **discrete Hamiltonian map** $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$. This map defines a one-step method, but the question remains, how is this map related to a particular Hamiltonian flow map? The answer comes from an object that connects continuous mechanics to discrete mechanics. We define the **exact discrete Lagrangian** as $L_d^E(q_0, q_1; h) = \int_0^h L(q_{01}(t), \dot{q}_{01}(t))dt$, where $q_{01}(0) = q_0$, $q_{01}(h) = q_1$, and $q_{01}(t)$ satisfies the Euler–Lagrange equations for $t \in (0, h)$. This object is also known as a type one generating function, since $\frac{-\partial L_d^E(q_0,q_1;h)}{\partial q_0} = p_0$ and $\frac{\partial L_d^E(q_0,q_1;h)}{\partial q_1} = p_1$(where we assume q_0 and q_1 are independent coordinates). Thus the Hamiltonian flow map can be implicitly defined via these equations. It is a well known result that maps resulting from a generating function are symplectic. In fact, since any discrete Lagrangian is a generating function, this implies that the discrete Hamiltonian map is symplectic and thus can be viewed as a symplectic integrator. However, symplecticity alone doesn't guarantee a good approximation to the continuous Hamiltonian flow map. For that we need the following error analysis result from [1], which connects the exact discrete Lagrangian to the discrete Lagrangian, and thus connects the Hamiltonian flow map to the discrete Hamiltonian map.

Theorem 1 *If a discrete Lagrangian, $L_d : Q \times Q \to \mathbb{R}$, approximates the exact discrete Lagrangian, $L_d^E : Q \times Q \to \mathbb{R}$ to order r, i.e.,*

$$L_d(q_0, q_1; h) = L_d^E(q_0, q_1; h) + \mathcal{O}(h^{r+1}),$$

then the discrete Hamiltonian map, $\tilde{F}_{L_d} : (q_k, p_k) \mapsto (q_{k+1}, p_{k+1})$, viewed as a one-step method, is order r accurate.

The upshot is that if we build our discrete Lagrangian to approximate a particular exact discrete Lagrangian, we are guaranteed to have a similar approximation to the Hamiltonian flow map given by the discrete Hamiltonian map, and furthermore, both maps will have the desirable property of symplecticity. In fact, it can be shown that the discrete Hamiltonian map will have long-term near-energy preservation, by viewing the discrete map as coming from a modified Hamiltonian function that is closely related to the true Hamiltonian function.

STOCHASTIC HAMILTONIAN DYNAMICS

How can we leverage this theory to construct stochastic integrators for stochastic differential equations? In [2] stochastic geometric mechanics is extended based on the work of [3]. Given a paracompact configuration manifold, Q, the stochastic Hamiltonian system is specified by a Hamiltonian $H : T^*Q \to \mathbb{R}$ and m deterministic function $\gamma_i : Q \to \mathbb{R}$ for $i = 1, \ldots, m$. Define the Lagrangian $L : TQ \to \mathbb{R}$ via the usual Legendre transform of H. The stochastic perturbations are given by a probability space $(\Omega, \mathcal{F}, P)$ and $(W_i(t), \mathcal{F}_t)_{t\in[a,b]}$, where $\{W_i\}_{i=1}^m$ are independent real-valued Wiener processes and $\mathcal{F}_t$ is the filtration generated by these processes. The Hamilton-Pontryagin principle, introduced in [4] is a generalization of Hamilton's principle and Hamilton's phase space principle, which unifies Lagrangian and Hamiltonian mechanics, as well naturally accommodating Legendre transforms that fail to be isomorphisms from TQ to T^*Q. It is in this context that Bou-Rabee and Owhadi introduce a stochastic variational principle. Define the Pontrayagin bundle of Q as $PQ = TQ \bigoplus T^*Q$, then the stochastic Hamilton-Pontrayagin action integral is given by,

$$\mathbb{S}(q, v, p) = \int_a^b [L(q, v)dt + \sum_{i=1}^m \gamma_i(q) \circ dW_i(t) + \langle p, \frac{dq}{dt} - v\rangle dt],$$

where (q, v, p) is in the corresponding path space and $\gamma_i(q) \circ dW_i(t)$ represents the Stratonovich integral. We then have a stochastic variational principle, which states that the curve (q, v, p) minimizes the stochastic action integral if and only if it satisfies the stochastic Hamilton-

Pontryagin equations,

$$dq = vdt$$
$$dp = \frac{\partial L}{\partial q}dt + \sum_{i=1}^{m} \frac{\partial \gamma_i}{\partial q} \circ dW_i$$
$$p = \frac{\partial L}{\partial v}.$$

In [2], it is shown that this framework allows for the natural extension to the stochastic setting of momentum-maps, symplecticity, Noether's theorem, holonomic constraints, non-conservative forces, and reduction. Furthermore, a discrete mechanics was developed in this stochastic setting. Given its similarity with the above deterministic setting and for the sake of brevity we now proceed directly to stochastic variational integrators. As before the variational integrator is constructed by discretizing the stochastic action integral, and then applying the discrete Hamiltonian map. While the theoretical groundwork was clearly layed out in [2], numerical demonstrations were not presented, so we have attempted to fill this gap.

NUMERICAL SIMULATIONS

Consider the stochastic linear oscillator given by the equations,

$$dq = vdt$$
$$dp = -qdt + \sigma dW(t)$$

which are the stochastic Hamilton-Pontryagin equations with $H(q,p) = \frac{1}{2}(q^2+p^2)$ and $\gamma_1(q) = -\sigma q$. We will run our simulation with intial conditions $(q(0), p(0)) = (1, 0)$ and $\sigma = 1$. The solution is given by $q(t) = cos(t) + \int_0^1 sin(t-s)dW(s)$. Using this it can be shown that $\mathbb{E}(q(t)^2 + p(t)^2) = 1 + \sigma^2 t$, thus we expect the Hamiltonian to grow linearly with time, on average. We construct the variational integrator from the stochastic action integral by approximating the velocity with the backwards difference formula $\frac{q_{k+1}-q_k}{h}$ and the integral by the rectangular quadrature rule about the endpoint q_{k+1}. After applying the discrete Hamiltonian map and rearranging terms we have the one-step method given by,

$$q_{k+1} = q_k + hp_{k+1}$$
$$p_{k+1} = p_k - hq_k + \sigma \Delta W_k.$$

This may rightfully be called the symplectic Euler-Maruyama method. It bears close resemblance to the partitioned Euler-Maruyama method derived in [5] as,

$$q_{k+1} = q_k + hp_k$$
$$p_{k+1} = p_k - hq_{k+1} + \sigma \Delta W_k.$$

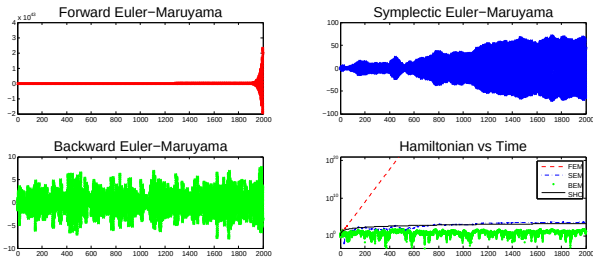

FIG. 2. Numerical simulations of the stochastic linear oscillator for a random realization with $\sigma = 1$, and intial conditions $(q_0, p_0) = (1, 0)$. The first column consists of the forward Euler-Maruyama(FEM) and the backward Euler-Maruyama(BEM), which show artificial blow-up and dissipation respectively. The last column shows the symplectic Euler-Maruyama(SEM) and a plot of the Hamiltonian vs time of all methods. All simulations used a step size of $h = 0.1$ over a time interval [0,2000].

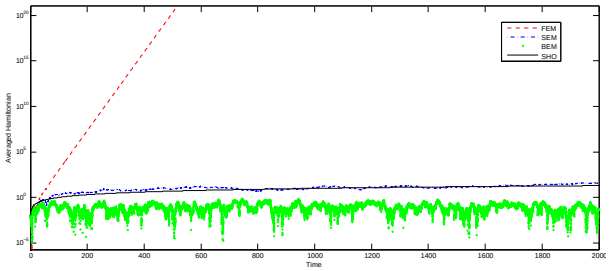

FIG. 3. Here we plot the averaged Hamiltonian vs time generated by running 100 simulations. The symplectic Euler-Maruyama method does an excellent job of approximating the true value of the expected Hamiltonian. The true line is denoted by SHO.

In fact, this is also a symplectic method, and it can be derived in the same variational framework by approximating the velocity with a forwards difference formula $\frac{q_{k+1}-q_k}{h}$ and the rectangular quadrature rule about the initial point q_k. For comparison we also have implemented the explicit forward Euler-Maruyama method and the implicit backward Euler-Maruyama method(see figure 2 and 3). As to be expected the symplectic Euler-Maruyama method clearly approximates the expected Hamiltonian far better than the other methods.

While we are able to nimbly model additive noise under the previous framework, another stochastic framework naturally allows multiplicative noise. The main idea, due to [6], is to consider random effects or noise as a non-conservative force. In the deterministic setting, given a hamiltonian system with a non-conservative force $\mathbf{F}$ and associated position vector $\mathbf{r} = \mathbf{r}(q,t)$, Hamilton's

equations become,

$$\dot{p} = -\frac{\partial H^T}{\partial q} + \frac{\partial \mathbf{r}^T}{\partial q}\mathbf{F}$$
$$\dot{q} = \frac{\partial H^T}{\partial p} - \frac{\partial \mathbf{r}^T}{\partial p}\mathbf{F}$$

Consider a stochastic Hamiltonian system with one noise coming from a Wiener process, $dW(t)$. We introduce the notation $\dot{W}(t)dt = dW(t)$, acknowledging that $W(t)$ is nowhere differentiable, and we refer to $\dot{W}(T)$ as white noise. The stochastic Hamilton's equations then become,

$$\dot{p} = -\frac{\partial H^T}{\partial q} - \frac{\partial \mathbf{H_1}^T}{\partial q} \circ \dot{W}$$
$$\dot{q} = \frac{\partial H^T}{\partial p} + \frac{\partial \mathbf{H_1}^T}{\partial p} \circ \dot{W}.$$

The stratonovich integral allows us to use the standard chain rule, as a result we can mimic much of the same derivation process used for the deterministic Hamiltonian case. The stochastic action integral associated to this system is given by,

$$\int_{t_0}^{t_1} L dt - \int_{t_0}^{t_1} H_1 \circ dW(t).$$

As expected the stochastic Lagrange equations of motion minimize the stochastic action integral and give a stochastic Hamilton's principle. For a concrete example, consider the Kubo oscillator with corresponding stochastic Hamilton's equations,

$$\dot{p} = -aq - \sigma q \circ \dot{W}(t)$$
$$\dot{q} = ap + \sigma p \circ \dot{W}(t).$$

Letting $H(p,q) = \frac{a}{2}(p^2+q^2)$ and $H_1(p,q) = \frac{\sigma}{2}(p^2+q^2)$, then the associated stochastic action integral over a time interval $[0,h]$ is given by,

$$S(q_0, q_1 : h) = \int_0^h [p\dot{q} - \frac{a}{2}(p^2+q^2)]dt - \int_0^h H_1 \circ dW(t)$$

We can construct a stochastic variational integrator by approximating the stochastic action integral, and applying the discrete Hamiltonian map. We apply the midpoint quadrature rule combined with a first order finite difference. Applying the discrete Hamiltonian map and rearranging terms yields,

$$p_1 = p_0 - ah\frac{q_1+q_0}{2} - \sigma\Delta W_0 \frac{q_1+q_0}{2}$$
$$q_1 = q_0 + ah\frac{p_1+p_0}{2} + \sigma\Delta W_0 \frac{p_1+p_0}{2}$$

This can be recognized as the symplectic midpoint method proposed by [7]. See figure 2 for a comparison of this method versus the standard Euler-Maruyama method.

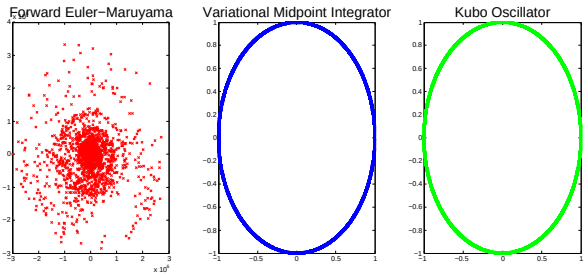

FIG. 4. The first plot on the right shows the true phase plot of the Kubo Oscillator for a random realization with $a = 1$, $\sigma = 1$, and intial conditions $(q_0, p_0) = (1, 0)$. The middle plot is the phase plot corresponding to the variational method, and the left plot correspons to the Euler-Maruyama method. Clearly, the variational integrator provides the better qualitative and quantitative solution. Both simulations used a step size of $h = 0.1$ over a time interval [0,2000].

CONCLUSION

In [2], a rigorous framework was developed that extended many of the deep geometric concepts underlying mechanics into a stochastic realm. This leads to a cohesive theory for stochastic variational integrators that can be developed further. Other frameworks may not be as pleasing, but allow for more flexibility in other areas, such as multiplicative noise. Either way, the numerical results show the effectiveness of stochastic variational integrators when applied to appropriate problems. It has been shown(see [8]) that the Feynman-Kac formulas for non-linear parabolic partial differential equations can be of 2^{nd} order, and perhaps even stochasic Hamiltonian. It would be interesting to see if applying stochastic variational integrators to the Feynman-Kac formula can yield numerical solutions of the pde's that benefit from the underlying structure preservation of the stochastic differential equation.

[1] Jerrold E. Marsden and Matthew West, "Discrete mechanics and variational integrators," Acta Numer. **10**, 357–514 (2001).

[2] Nawaf Bou-Rabee and Houman Owhadi, "Stochastic variational integrators," IMA Journal of Numerical Analysis **29**, 421–443 (2009).

[3] J.M. Bismut, *Mecanique Aleatoire* (Springer, Berlin, 1981).

[4] H. Yoshimura and J. E. Marsden, "Dirac structures in Lagrangian mechanics Part II: Variational structures," Journal of Geometry and Physics **57**, 209–250 (2006).

[5] A. H. Strmmen Melb and D. J. Higham, .

[6] Lijin Wang, Hong Jialin, Rudolph Scherer, and Fengshan Bai, "Dynamics and variational integrators of stochastic hamiltonian systems," International Journal of Numerical

Analysis and Modelling **6**, 586–602 (2009).

[7] G.N. Mistein, Y.U.M. Repi, and M.V. Tretyakov, "Numerical methods for stochastic systems preserving symplectic structure," SIAM J. Numer. Anal..

[8] Patrick Cheridito, H. Mete Soner, Nizar Touzi, and Nicolas Victoir, "Second-order backward stochastic differential equations and fully nonlinear parabolic pdes," Communications on Pure and Applied Mathematics **60**, 1–30 (2007).