Variational Integrators for Structure-Preserving **Filtering**

Jarvis Schultz PhD. Member of ASME Department of Mechanical Engineering Department of Mechanical Engineering Northwestern University Evanston, Illinois, 60208

Kathrin Flaßkamp

Dr., Postdoctoral Fellow Northwestern University Evanston, Illinois, 60208 Email: jschultz@northwestern.edu Email: kathrin.flasskamp@northwestern.edu

> Todd D. Murphey Professor, Member of ASME Department of Mechanical Engineering Northwestern University Evanston, Illinois, 60208 Email: t-murphey@northwestern.edu

Estimation and filtering are important tasks in most modern control systems. These methods rely on accurate discretetime approximations of the system dynamics. We present filtering algorithms that are based on discrete mechanics techniques (variational integrators), which are known to preserve system structures (momentum, symplecticity, constraints, for instance) and have stable long-term energy behavior. These filtering methods show increased performance in simulations and experiments on a real digital control system. The particle filter as well as the extended Kalman filter benefit from the statistics-preserving properties of a variational integrator discretization, especially in low bandwidth applications. Moreover, it is shown how the optimality of the Kalman filter can be preserved through discretization by means of modified discrete-time Riccati equations for the covariance updates. This leads to further improvement in filter accuracy, even in a simple test example.

1 Introduction

Estimation and filtering techniques for unknown system states based on noisy sensor data are crucial in many control systems. These methods are implemented in discrete time on digital embedded systems. Thus, they rely on accurate approximations of continuous-time system behavior; in low-bandwidth applications, predictions over longer time-horizons are required. Traditionally, explicit Euler discretizations of the system dynamics or high-order Runge-Kutta methods are chosen. Due to computational restrictions and real-time constraints the first-order explicit Euler method is often the standard choice. However, low-order variational integration methods provide an alternative that allows efficient and accurate predictions. This paper studies the benefits of structure-preserving integration in two commonly used filtering approaches: (extended) Kalman filters and particle filters.

Kalman filters The Kalman filter is likely the most wellknown and widely used estimator [1]. When the Kalman filter is applied to a linear system with additive Gaussian noise, it provides the optimal, maximum-likelihood (minimumvariance) state estimator [2]. The Kalman filter has been successfully applied to a wide variety of problems in image processing, wireless communication, aerospace, robotics, and more. While the Kalman filter is simple to compute, the requirement of linear dynamics limits its applicability to nontrivial systems. The extended Kalman filter (EKF) relaxes the linearity requirement by utilizing local linearizations to approximate posterior distributions as Gaussian. It is only slightly more complex than the Kalman filter, and it is applicable to a larger class of systems. One drawback of the EKF is that its performance is strongly influenced by the severity of any nonlinearities in the system. Nevertheless, the EKF is commonly employed and has demonstrated reasonable performance in a wide range of applications.

Particle filters In recent years particle filters have become a popular technique for estimation [3, 4], likely due to their generality. In a particle filter, the uncertainty posterior distribution is represented by a finite collection of parameters referred to as particles. These particles are drawn from a distribution representing the current belief of the system's state, and then each particle is mapped forward in time using a model of the system. As the number of particles approaches infinity, the distribution of the particles approaches the solution to the Fokker-Plank equation associated with the system [5]. A primary feature of the particle filter is its wide applicability. Particle filters work with arbitrary noise model, and have no issues with nonlinear system dynamics. A strong disadvantage of particle filter lies in their computational complexity. Its reliability is primarily governed by the number of particles used, and in some problems, the number of particles required to approximate the true posterior distribution may be quite large (10^4-10^6) .

Discrete mechanics Discrete mechanics derives models for mechanical systems via discrete variational principles [6]. These models define numerical integrators (variational integrators (VIs), also called symplectic integrators) which possess beneficial properties: They conserve symplecticity and symmetries (momenta) of the continuous-time system, and they have stable energy behavior even in long-term simulations, which is important in astrodynamics or molecular dynamics simulations [7]. furthermore, VIs exactly satisfy holonomic constraints common in multibody systems. VIs can be used to design an indirect or a direct optimal control method (cf. e.g. [8]) that inherits the advantages of forward integration. Additionally, the problem formulation requires fewer optimization variables and leverages structured linearizations for efficient optimization algorithm design (cf. [9]). recently, the discrete mechanics concept has been extended to stochastic integrators (see [10, 11]) which almost surely preserve the symplectic structure. In [12], stochastic variational integrators have been successfully applied for estimation problems on Lie groups using a method based on uncertainty ellipsoids. This paper extends results of the authors' previous works [13, 14].

Embedded systems One of the primary motivations for this work is to develop and explore filter implementations that preserve geometric structure in order to improve performance while also facilitating implementation on an experimental embedded system. In recent years, much work has been done on leveraging a variety of numerical methods for improving the performance of estimation routines [15–17], but, these methods typically achieve increased performance through higher-order, computationally expensive methods that may not be appropriate for embedded estimation. For an arbitrary choice of integrator for use in an embedded system one would hope that the computational cost was low to deal with limited computational resources, and that the formulation could be cast as a one-step map that is easily linearized to fit the form of standard discrete control and estimation routines. One may also hope that the integration method satisfy holonomic constraints as often encountered in mechanical systems. Moreover, the satisfaction of holonomic constraints should not result in loss of the ability to linearize the discrete map across a timestep. The variational integrator (VI) presented in [9] satisfies all of these requirements while also a providing a symplectic integrator with stable long-term energy behavior. The linearizations of that VI are explicit and come with fixed computational cost

further aiding in implementation on an embedded system. While the VI from [9] is not the only integrator used herein it provides motivation for much of this work. This integrator is a drop-in replacement for linearizable one-step maps, such as the explicit Euler scheme, in standard forms of particle filters and extended Kalman filters. While there are a multitude of other integration schemes available, it is difficult to find another technique that doesn't compromise at least one of these requirements. For example, many integrators that satisfy holonomic constraints are not easily linearizable. Higher-order integrators often result in multi-step methods, and the increased order results in increased computational cost. For several experimental examples presented in this paper, we compare VI-based estimation routines to corresponding routines based on explicit Euler integration. While other choices could be made for this comparison, due to our interest in estimation for embedded systems, we constrain our choice to low-order, one-step maps, and in this context, the explicit Euler integrator is a standard choice.¹

Contributions To begin, we illustrate the advantageous statistical properties of first- and second-order variational integrators for simulating a stochastic harmonic oscillator. We then present particle filters and extended Kalman filters based on VI discretization and demonstrate their performance in simulated and experimental examples. It is demonstrated that both particle filters and extended Kalman filters may be highly sensitive to the choice of integration scheme and that VI-based methods outperform the discretetime filter obtained from zero-order-hold and explicit Euler discretization - a combination that fits well within the embedded systems requirements discussed in the previous section. A second important contribution of this work is the derivation of structure-preserving covariance updates for a Kalman filter. We prove that the modified update equations additionally capture the Hamiltonian/symplectic structure resulting from the optimality property of the Kalman filter. We then apply a traditional Kalman filter and a Kalman filter with the structure-preserving covariance update equations to a simple example. It is shown that a Kalman filter with the structure-preserving update equations outperforms the filter with standard covariance updates.

In Section 2, we give an introduction to variational/symplectic integrators for deterministic and stochastic systems. Section 3 provides an introduction to particle filters and illustrates how symplectic integration can be used in particle filtering in order to reduce the common issue of particle deprivation. In Section 4, we recall the definition of discretetime Kalman filter and show how the one-step maps can be derived from variational integrators. The increased performance of these filters is experimentally demonstrated for a planar crane system in Section 5. We extend our approach by deriving structure-preserving covariance updates in Section 6. Finally, we discuss the results and provide concluding

¹In the conclusion of this document we mention some recent work [18] that describes the generation arbitrary-order VIs that are still one-step maps. These VIs may provide another choice of integrator that satisfies the aforementioned requirements while also yielding a higher order method.

remarks in Section 7.

2 Variational Integrators

This section provides a short overview of variational integrators for deterministic/stochastic simulation, optimal control, and estimation. For more details, see [6, 7, 7-11].

2.1 Variational Mechanics

A mechanical system can be described by its *Lagrangian L*, a real-valued function that depends on the system's configuration q(t) and velocity $\dot{q}(t)$ at any time t. The Lagrangian is typically the difference of kinetic and potential energy, whereas the system's *Hamiltonian H* is typically the sum of all energies. In addition, there may be forces $f(q, \dot{q}, u)$ such as friction or a control input force that influence the system's dynamics. The system dynamics are described by the Lagrange-d'Alembert principle, such that they satisfy the variational equation (cf. [6])

$$\delta \int_0^T L(q,\dot{q}) dt + \int_0^T f(q,\dot{q},u) \cdot \delta q \, dt = 0. \tag{1}$$

Equivalently, (q, \dot{q}) (for a given control input u on [0, T]) satisfy the forced Euler-Lagrange equations

$$\frac{\partial L}{\partial q}(q,\dot{q}) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q,\dot{q}) \right) + f(q,\dot{q},u) = 0.$$
(2)

Under certain regularity assumption (cf. [6] for details), these equations can be equivalently written as Hamilton equations in configuration q and momentum variables $p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q})$ with Hamiltonian $H(q, p) = p \cdot \dot{q} - L(q, \dot{q})$ and Hamiltonian forcing f_H :

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q} + f_H(q, p, u).$$
 (3)

In the absence of forces, the flow of a Hamiltonian system (and thus, of an unforced Lagrangian system) is $symplectic^2$.

2.2 Discrete Mechanics

In order to simulate the system dynamics, most numerical integration schemes, e.g. Runge-Kutta methods, would be applied to either the Euler-Lagrange equations (2) or the Hamilton equations (3). This differs from *variational integrators*, which instead discretize the variational equation (1). This process generates an iterative numerical integration scheme, and the approach guarantees that structures of the original continuous system are preserved, e.g. symplecticity and symmetries in terms of system invariances (cf. [6]). Furthermore, variational integrators have exponentially stable energy behavior over long time horizons (cf. [7]).

To derive a discrete variational integration scheme, the action map, i.e. the first term in Eq. (1), is approximated over each time step [kh, (k+1)h] by

$$L_d(q_k, q_{k+1}, h) \approx \int_{kh}^{(k+1)h} L(q(t), \dot{q}(t)) dt$$
 (4)

and the forcing term is discretized in a similar manner with discrete forces f_k^- , f_k^+ . The term $L_d(q_k, q_{k+1}, h)$ is referred to as the discrete Lagrangian, and f_k^- , f_k^+ are the left and right discrete forces, respectively [6]. Taking variations of the discrete configurations q_k for k = 1, ..., N - 1 leads to the *discrete forced Euler-Lagrange equations* (cf. [6, 8]) for k = 1, ..., N - 1,

$$D_1L_d(q_k, q_{k+1}) + D_2L_d(q_{k-1}, q_k) + f_k^- + f_{k-1}^+ = 0$$

where D_nL_d is the *slot derivative*³ of L_d . Discrete momenta are given by the relationships $p_k = -D_1L_d(q_k, q_{k+1}) - f_k^$ and $p_{k+1} = D_2L_d(q_k, q_{k+1}) + f_k^+$ which implicitly define a symplectic one-step integration map $\Psi_{k+1} : (q_k, p_k) \mapsto$ (q_{k+1}, p_{k+1}) . It is shown in [7] that (in the unforced case) a large class of variational integrators can be written as symplectic partitioned Runge-Kutta methods. In [8], the definition of symplecticity in VI/symplectic partitioned Runge-Kutta methods is extended to external forces. Thus, we will use both VI and symplectic partitioned Runge-Kutta schemes throughout this paper. In particular, will use the VI midpoint rule for the planar crane system (Section 5) and the symplectic Euler method for Kalman filtering in Section 6.

2.3 Stochastic Symplectic Integrators

In [11], the Hamilton-Pontryagin variational principle is extended to stochastic dynamical systems. It is shown that a stochastic mechanical system possesses an almost surely variational structure and thus, is almost surely symplectic and momentum map preserving. These properties carry over to the stochastic variational integrator that [11] derives from analogous discrete variations. While their first order integrator is derived from Hamilton-Pontryagin principle in configuration, momentum, and velocity coordinates, the stochastic symplectic Euler method used in this work is based on Hamilton equations in (q, p) coordinates.

In [10] stochastic integrators for Hamiltonian systems are developed with deterministic dynamics (as in Eq. (3)) and additive noise terms driven by independent Wiener processes. Their methods are built on symplectic partitioned Runge-Kutta schemes as introduced in [7] for deterministic

is symplectic if $A^T J A = J$ with $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

²In general, a map is symplectic if it preserves a *symplectic form* ([6,7]). To get an intuition for symplecticity, it is helpful to consider the fact that symplectic transformations are area-preserving. A linear map $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$

³The slot derivative $D_nL(A_1, A_2, ...)$ represents the derivative of the function *L* with respect to the *n*-th argument, A_n . At times, the arguments to the function *L* may be dropped for compactness – in that case the slot derivative still applies to the argument order from the function definition.

systems. In our numerical examples, we use one of the firstorder methods presented in [10, Eq. (3.6)],

$$q_{k+1} = q_k + h \frac{\partial H}{\partial p}(p_k, q_{k+1}) + \sigma \Delta_k \omega_1$$

$$p_{k+1} = p_k - h \frac{\partial H}{\partial q}(p_k, q_{k+1}) + \gamma \Delta_k \omega_2,$$
(5)

for k = 0, 1, ..., N - 1, where $\Delta_k \omega_{1,2}$ are discretized independent Wiener processes and $\sigma, \gamma \in \mathbb{R}$. In the following, we refer to this method as the *symplectic Euler scheme*. Note that this method simplifies to an explicit symplectic method for separable Hamiltonian systems (H = T(p) + V(q)).

Additionally, we approximate the deterministic mechanics by a second-order *variational midpoint integrator*, which, applied to stochastic mechanical systems, leads to

$$q_{k+1} = q_k + h \frac{\partial H}{\partial p} \left(\frac{p_k + p_{k+1}}{2}, \frac{q_k + q_{k+1}}{2} \right) + \sigma \Delta_k \omega_1$$

$$p_{k+1} = p_k - h \frac{\partial H}{\partial q} \left(\frac{p_k + p_{k+1}}{2}, \frac{q_k + q_{k+1}}{2} \right) + \gamma \Delta_k \omega_2.$$
(6)

Alternatively, the midpoint integrator can be derived variationally from $L_d(q_k, q_{k+1}, h) = h \cdot L\left(\frac{q_{k+1}+q_k}{2}, \frac{q_{k+1}-q_k}{h}\right)$ which we will use in the examples presented in Section 5. The stochastic processes are then included as shown in [11].

The following example of a stochastic harmonic oscillator is adapted from [10], and we use it to illustrate the crucial role of the integration scheme in stochastic simulations.

Example 2.1 (Stochastic Harmonic Oscillator). The stochastic differential equations for the harmonic oscillator with configuration $q \in \mathbb{R}$, momentum $p \in \mathbb{R}$, Hamiltonian $H(q,p) = \frac{1}{2} (q^2 + p^2)$, and Wiener processes $w_{1,2}$, are given by

$$\begin{pmatrix} dq \\ dp \end{pmatrix} = \begin{pmatrix} 0 \ 1 \\ -1 \ 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} dt + \begin{pmatrix} \sigma \ dw_1(t) \\ \gamma \ dw_2(t) \end{pmatrix}.$$
(7)

We use the analytic solution presented in [10] to investigate the flow from a set of initial conditions (circle about (0,0)in Fig. 1) and to compare the approximation quality of the different integrators. The first integration method is the standard Euler-Maruyama integrator (see e.g. [5]). As can be seen in Fig. 1, its flow artificially expands the radius of the circle (blue dots), and the mean location of the circle drifts significantly from the true solution. This is unlike the behavior of the two symplectic methods (2nd order midpoint in red; points mostly hidden by 1st order symplectic Euler results depicted in green), which both follow the exact flow very closely and retain the circle area. In particular, we cannot observe a clear superiority of the 2nd order method over the symplectic Euler scheme in this example. Both symplectic integrators have statistics-preserving⁴ properties that are



Fig. 1. Approximations of the stochastic harmonic oscillator flow for a set of initial conditions. All integrators use the same step-size h = 0.03125 and an identical sample path of the Wiener process with $\sigma = \gamma = 1.0$. The results from the Euler-Maruyama integrator are depicted in blue, the VI midpoint in red and the symplectic Euler in green. The analytic solution (black) from [10] is given for comparison.

crucially important for estimator performance, as we illustrate in the remainder of this paper.

3 Particle Filters

This section will provide a brief introduction to particle filters and it will provide a numerical example illustrating how a symplectic integrator can be leveraged to alleviate a common issue with particle filters.

3.1 Basic Particle Filtering

The key concept behind a particle filter is that uncertainty distributions of the system's state will be represented by a finite collections of particles. No analytical expression for the probability density function associated with these distributions is required; rather, we assume that the particles are drawn from this distribution. In this way, the collection of particles itself represents an approximation of the distribution. This is opposed to other filters, e.g. the Kalman filter, where the uncertainty distribution is parameterized and the filter predicts the evolution of the parameters. We will now sketch the steps required for a particle filter following the terminology and syntax of [3].

Initialization: A particle filter is initialized by generating a set of *M* particles X_0 where each particle x_0^m with $1 \le m \le M$ represents a possible initial state of the system. This initial particle collection may be drawn from a known distribution or even assumed to all have the same value.

Proposal Distribution: In each step of the particle filter, each particle from the previous timestep is mapped forward

⁴By statistics-preserving, we mean that these symplectic integrators are

capable of both accurately predicting statistical quantities such as temperature [19] and of almost surely preserving symmetries of stochastic systems [11].

to the current time using a noisy system model and knowledge of the controls used during the previous timestep. This new collection of particles is referred to as a *proposal distribution*. This step is also the step where the integration techniques of this work are implemented.

Importance Weights: In order to incorporate measurements an *importance weight* for each particle is calculated. The importance weight uses a noisy measurement model to evaluate the probability of a measurement z_t given a particle x_t^m i.e. the weight of particle m at time t is $w_t^m = p(z_t | x_t^m)$.

Resampling: In most implementations of the particle filter, the next step is to incorporate the measurement z_t into the system's uncertainty distribution through a resampling process. In this process the particles representing a sampling of the posterior distribution for the current timestep is generated by randomly drawing with replacement a new set of M particles from the proposal distribution. This random drawing is biased by the importance weights calculated in the previous step. Particles from the proposal distribution with higher measurement probability have a higher probability of being drawn for inclusion in the posterior distribution. In this way particles with low probability are culled from the collection, and the total particle distribution is drawn to the measurement. This process allows the statistic of the posterior distribution to include knowledge of the measurements and it helps to ensure that the process of generating the proposal distribution is not wasting time computational effort on integrating particles that have very low posterior probability.

Iterate: To continue the particle filter the posterior particle set from timestep t is used as the initial set for the Proposal Distribution step of timestep t + 1.

3.2 Particle Filters and Choice of Integrator

A particle filter inherits the statistical properties of the integration scheme that is used for propagating the collection of particles. We demonstrate this by revisiting the stochastic harmonic oscillator example. In the applications presented in Section 5 we show an example of how a common issue with particle filters (particle deprivation) can be significantly reduced by using variational integrators.

In particle filtering algorithms, a collection of particles is mapped forward a single timestep using a noisy process model in order to predict a proposal distribution. Then a measurement of the system is taken, and a new set of particles is produced by resampling the proposal distribution to produce an approximation of the posterior distribution. Particles with higher probability given the measurement value (and its corresponding probability density function) have a higher probability of being selected for the posterior distribution. This resampling process incorporates the measurement in the filter state by removing particles with low probability.

Typically, if measurements are collected regularly from a reliable sensor, the incorporation of the measurements will help to regulate errors caused by inaccurate integration of the particle states. However, there are many situations where one may encounter unreliable measurements, for example, in high-dimensional data association problems, systems with



Fig. 2. Image of the particle filter covariance propagation for the harmonic oscillator without resampling. Note that the system's covariance matrix has two eigenvalues, thus there are two eigenvalues plotted for each integrator.

unreliable communication networks, or tracking systems suffering from occlusions. In these situations, the particle filter must predict the state of the system over long time horizons without measurements. As a result, symplectic integrators may significantly increase the performance of the particle filter. An illustration of this can be seen in the following example.

Example 3.1 (Stochastic Harmonic Oscillator). In Fig. 2, the statistical properties of the harmonic oscillator of Eq. (7) are simulated by integrating a collection of 1,000 particles for 60 seconds with a timestep of 0.0625 s using a midpoint VI and an explicit Euler integrator (RK1). The plot shows the eigenvalues of the covariance of the particle distribution as a function of time. It can be seen that the RK1 scheme adds significantly more noise to the system than the variational integrator. This is true even though the two collections of particles were driven by the same set of sample paths of the Wiener process. Even in this simple system, with no measurements to regulate the integration errors the non-symplectic RK1 integrator significantly over-predicts the uncertainty in the system state.

3.3 Particle Deprivation

A common issue encountered in particle filters is that of *particle deprivation* [3]. In a filter suffering from particle deprivation, there are too few particles in the vicinity of the true state of the system. When this occurs, the resampling process may eventually drive the number of unique particles down to a single particle. While there are known heuristics for preventing this situation [3], we point out that an integrator adding artificial diffusion to the collection of particles will increase the filter's probability of suffering from particle deprivation when measurements are incorporated. Results illustrating this point are in Section 5.

4 Kalman Filter

As a starting point, we consider the following formulation of a Kalman filter; originally developed in [1]. For a textbook reference, we refer to [20].

Proposition 4.1 (Discrete-time Kalman filter, [1]). Let a model of the system and the measurement be given by

$$\begin{aligned} x_{k+1} &= A_k x_k + G_k w_k, \quad x_0 \sim \mathcal{N}(x^0, P_0), w_k \sim \mathcal{N}(0, Q_k), \\ z_k &= C_k x_k + v_k, \qquad \qquad v_k \sim \mathcal{N}(0, R_k), \end{aligned}$$

with $R_k > 0$ and $\{w_k\}_{k=0}^N$, $\{v_k\}_{k=0}^N$ uncorrelated Gaussian white noise processes with covariance Q_k and R_k respectively. The initial state x_0 is assumed to be drawn from a Gaussian centered around a nominal initial state x^0 with covariance the same as the filter's initial covariance P_0 . A measurement at timestep k is denoted z_k . Then, the optimal estimation of the state, \hat{x}_{k+1} , given measurements up to step k, is given by the following update equations:

Prediction:

$$\begin{split} P_{k+1}^{-} &= A_k P_k A_k^T + G_k Q_k G_k^T, \\ \hat{x}_{k+1}^{-} &= A_k \hat{x}_k, \end{split}$$

Measurement update:

$$K_{k+1} = P_{k+1}^{-} C_{k+1}^{T} (C_{k+1} P_{k+1}^{-} C_{k+1}^{T} + R_{k+1})^{-1},$$

$$P_{k+1} = (\mathbb{1} - K_{k+1} C_{k+1}) P_{k+1}^{-},$$

$$\hat{x}_{k+1} = \hat{x}_{k+1}^{-} + K_{k+1} (z_{k+1} - C_{k+1} \hat{x}_{k+1}^{-}).$$

In Kalman's original work [1], the state transition matrix of the continuous-time system is used for A_k in Prop. 4.1. In general, this transition matrix cannot be computed analytically, but it can be approximated by an integration scheme applied to the continuous differential equation.

To give an example: When applying the explicit Euler scheme to the dynamical system $\dot{x} = Ax$ one obtains the discrete update matrix $A_k = (\mathbb{1} + hA)$, whereas the symplectic Euler scheme leads to the one-step map

$$A_{k} = \begin{pmatrix} (\mathbb{1} - hA^{11})^{-1} & h(\mathbb{1} - hA^{11})^{-1}A^{12} \\ hA^{21}(\mathbb{1} - hA^{11})^{-1} & h^{2}A^{21}(\mathbb{1} - hA^{11})^{-1}A^{12} + \mathbb{1} + hA^{22} \end{pmatrix}$$

4.1 Linearizations for Extended Kalman Filters

The midpoint variational integrator has a unique characteristic referred to as a *structured linearization* [21, 22]. Many estimation algorithms, such as the EKF, rely on local linearizations to adapt linear system tools to nonlinear systems. Typically, in the continuous time setting, to obtain a local linearization at a point one would use a Taylor expansion of $\dot{x} = f(x, u)$ about that point. This linearization is an infinitesimal linearization, but the discrete-time linearization is what is required. In other words, one needs to linearize the discrete map $x_{k+1} = f_k(x_k, u_k)$. In a standard implementation of an EKF [3], the discrete-time Kalman filter from Proposition 4.1 is adapted to a nonlinear system using the following



Fig. 3. Schematic of planar crane system including relevant geometric parameters.

linearizations

$$A_{k} = \frac{\partial f_{k}}{\partial x} \Big|_{\hat{x}_{k}, u_{k}} \quad \text{and} \quad C_{k} = \frac{\partial h_{k}}{\partial x} \Big|_{\hat{x}_{k+1}^{-}}$$
(8)

where the measurement model is assumed to be of the form $z_k = h(x_k) + v_k$.

As the presented VI is a one step method, it admits a simple linearization and even though the method is an implicit method, one can calculate an explicit linearization of the DEL equations [22]. Furthermore, since the VI is *exactly* sampling a nearby mechanical system, the VI linearization is an *exact* linearization of the nearby mechanical system. It is thus called a *structured linearization* [21, 22]. This restriction on the behavior of the linearization greatly improves its local accuracy. In Section 5 we show an example of how the improved accuracy of this linearization leads to better EKF estimator performance.

In principle, a one-step map can be derived and linearized for implicitly defined RK schemes in the same manner, but higher order RK methods lead to complex Taylor series expansions due to the intermediate evaluation points. For that reason, implementations are often restricted to first order methods. While this paper does not explicitly consider results related to estimation for constrained systems, we note that an additional interesting feature of the midpoint VI integrator is that it exactly satisfies holonomic constraints at each timestep while still admitting the explicit, structured linearization.

5 Application to Planar Crane System

This section presents experimental and simulated examples demonstrating how the choice of discretization scheme may dramatically affect the performance of standard particle filters and extended Kalman filters. Further, we demonstrate improved performance of both filters when leveraging the structure-preserving properties of variational integrators.

In the following two sub-sections we apply particle filters and EKFs to the planar crane system shown in Fig. 3. This system and corresponding modeling strategy have been discussed by the authors in [21] and [23]. In those works, a modeling strategy was presented where the position of the winch system x_r , and length of the string r were treated as kinematic inputs; i.e., inputs where the system has sufficient control authority to perfectly track any desired trajectory [24]. With this modeling strategy, the Lagrangian for the system is only a function of the dynamic configuration variables (x, y) and is given by

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy.$$
(9)

The corresponding midpoint discrete Lagrangian (Eq. (4)) for the planar crane system is then

$$L_d(q_k, q_{k+1}) = \frac{m}{2\Delta t} \left((x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2 \right) - \frac{mg}{2\Delta t} (y_{k+1} + y_k).$$
(10)

As a result of the kinematic input assumption, all damping in the crane portion of the system is assumed to be negligible. Moreover, through experimentation, we found that the damping in the swing dynamics of the payload was low enough to also be considered negligible. Thus, the deterministic models of the dynamics of this system have no damping.

For all results in this section, the noise models are assumed to be additive Gaussian noise with zero mean. The covariances parameterizing these distributions are all assumed to be diagonal, and the experimental system only measures relevant configuration variables i.e. there are no measurements of velocity or discrete generalized momentum.

5.1 Numerical Results

Figure 4 shows simulated effects of the choice of integration scheme on performance of the EKF algorithm at a range of timesteps. Each point on the plot was generated by running 1000 trials of the filter on a nominal feasible trajectory for the system described in Fig. 3. In the "prediction" step of the EKF, actual samples from the nominal trajectory were used i.e. the predictions without additive noise corresponded exactly with the nominal trajectory. Measurements were simulated by adding Gaussian noise to the nominal trajectory. For each trial, the L_2 error between the filtered signal and the reference signal was determined at each timestep. These errors were then averaged to produce an "error norm" for each trial. The error norms were then averaged, and their standard deviations calculated to produce the points and error bars on Fig. 4. For the upper, solid curve the local linearization was performed by evaluating the infinitesimal derivatives of the continuous dynamics about the current best estimate, and then using an explicit Euler approximation to convert this into a discrete linearization. For the bottom curve, the structured linearization of the midpoint variational integrator discussed in Section 4.1 was used. It is easily seen that as the filter frequency decreases, the performance of the discrete- mechanics-based variational integrator representation significantly outperforms the explicit Euler approximation.



Fig. 4. Plot illustrating variations in EKF filter performance using two different discrete representations of the continuous dynamics. Both curves are simulated from 1000 trials, with Gaussian noise added to produce "measurements".

5.2 Experimental Results

In this section we use experimental data from the robotic planar crane system described in [21] and [23]. This system utilizes digital encoders to close control loops around x_r and r. Thus, the kinematic-input modeling strategy mentioned in the previous section is used again. Since these inputs are assumed to be perfectly controlled, the dynamic configuration variables (x, y) are the only state variables estimated.

Fig. 5 shows the parametric evolution of estimates of the dynamic configuration variables using particle filters and EKFs at two different frequencies using both a midpoint VI and an RK1 integrator to represent the system. The strength of the VI integrator can be seen by noting that the particle filter and the EKF estimates of the system's uncertainty are nearly identical even at frequencies as low as 6 Hz. The two filter estimates are not only in agreement with each other, but they are also in agreement between the frequencies. We emphasize that at 6 Hz the timestep used for integration, measuring, linearization, and estimation is ≈ 0.167 s. With this large timestep the RK1 particle filter is useless.

Fig. 6 shows the time evolution of the eigenvalues of the covariances from each of the estimators in Fig. 5. This figure further demonstrates that the VI covariance predictions are stable even at large timesteps. Additionally, it demonstrates that in this particular system the structured linearization allows the EKF to perform nearly identically to the particle filter. In this case the increased accuracy of the structured linearization avoids the increased computational expense of the particle filter while achieving similar performance.

In Section 3.3 the issue of particle deprivation in particle filters was discussed. It was explained that the VI reduces the likelihood of this issue by removing artificial noise injected by a traditional integrator; this is especially true at low frequencies. Fig. 7 shows the number of unique particles as a function of time for the 6 Hz particle filters using the VI and RK1 integrator. In several cases, the particle distribution from the RK1 integrator is artificially spread to the point that resampling produces only 2 unique particles (out



Fig. 5. This set of figures illustrates the results of utilizing a particle filter (PF) and an EKF to estimate the dynamic configuration variables of the system shown in Fig. 3. For each filter, both VI and RK1 integrators are used. In the left plots, measurements, controls and integrations happen at 30 Hz, while in the right plots they all occur at 6 Hz. To generate the data for a given frequency, an experimental system was sent a set of commands at the desired frequency; the same commands were used to step the integrators for predictions. A Microsoft Kinect[®] configured to provide data at the target frequency was used to measure both of the dynamic configuration variables (x, y). For the 30 Hz data, only some of the measurements/ estimator updates are shown to avoid overcrowding the figure. The particle filters used 1000 particles, and the "low variance sampler" algorithm from page 110 of [3] is used to resample. The ellipses shown represent the local covariance estimates for each filter. The eigenvalues and eigenvectors of the covariance are used to define the size and orientation of the ellipses. Note that the VI-based filters outperform the RK1-based filters at both frequencies, and that the VI filter performance is similar for both frequencies. Additionally note that the VI covariance estimates are in excellent agreement between the filters at both frequencies; this is not true for the RK1 filters.

of 1000) which implies uncertainty estimates from the filter are essentially useless.

6 Kalman Filter with Structure-Preserving Covariance Updates

In Section 4, we discussed the difference between symplectic and non-symplectic discrete system matrices A_k in the Kalman filter and showed benefits of using symplectic discretization schemes for filtering applications in Section 5. However, the Kalman filter provides further structure due to its optimality property. In this section, we derive a modification to the previously-used Kalman filter that preserves this optimality structure through discretization and thereby further improves the filter performance.

To this aim, we start with the *continuous-time* Kalman filter, as introduced in [25] (see Prop. 6.1 below). The continuous-time Kalman filtering problem is dual to a lin-

ear quadratic regulator problem: While an LQR controller minimizes a quadratic cost functional, the Kalman filter determines state estimates that minimize the expected squared error to the process' true state (cf. [25]). In particular, both problems exhibit an additional *Hamiltonian structure* (independent of the mechanical Hamiltonian structure). This can be seen by examining the state-adjoint system that belongs to the respective problem's optimal solutions.

Our approach aims at additionally preserving this property through numerical discretization. Therefore, we set up a symplectic discretization of the state-adjoint system for the Kalman filtering problem. It is not necessary to explicitly solve for the adjoint equations, though. Instead, we directly derive modified discrete-time Riccati equations which are independent of the states and adjoints (as previously presented in [14]). Thus, our Kalman filter algorithm only consists of a single covariance update step and an all-in-one state update afterwards (see Algorithm 6.1). The performance of



Fig. 6. This figure shows the time evolution of the eigenvalues of the covariance prediction for particle filters (PF) and EKFs applied to the planar crane problem of Fig. 3 with both a 30 Hz (left plots) and 6 Hz (right plots) measurement, estimator and integrator update rate. These eigenvalues are the same eigenvalues used to define the major and minor axes of the uncertainty ellipses plotted in Fig. 5. Thus, there are two eigenvalues for each filter and integrator. Note that the VI-based PF and EKF estimates are in excellent agreement with each other even when compared across frequencies. The two 6 Hz RK1 filters predict very different covariance propagation, and they both disagree with the corresponding 30 Hz filters.



Fig. 7. Illustration of the VI improving particle deprivation. The red circles represent the 6 Hz VI particle filter, and the noisy, dashed blue line is the 6 Hz RK1 particle filter.

this algorithm is studied in Example 6.1 and compared to the discrete-time Kalman filter methods with standard covariance updates.

Proposition 6.1 (From [25]). Let the model of the system

and the measurement be given by

$$\dot{x} = Ax + Gw, \quad z = Cx + v$$

where $x(t_0) \sim \mathcal{N}(x^0, P_0), w \sim \mathcal{N}(0, Q), v \sim \mathcal{N}(0, R)$ with white noise processes $\{w(t)\}, \{v(t)\}$ being uncorrelated, also with x^0 . Then, the optimal filter is given by

$$\dot{x} = A\hat{x} + K(z - C\hat{x}), \quad \hat{x}(0) = x^0, \text{ with } K = PC^T R^{-1},$$

 $\dot{P} = AP + PA^T - PC^T R^{-1} CP + GQG^T, P(t_0) = P_0.$

In [25], the duality relation to the LQR problem is used to state a Hamiltonian for the filtering problem,

$$\mathcal{H}(x,w) = -\frac{1}{2} \|G^T x\|_Q^2 - w^T A^T x + \frac{1}{2} \|Cw\|_{R^{-1}}^2,$$

such that the state-adjoint equations are given by

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial w} = -A^T x + C^T R^{-1} C w,$$

$$\dot{w} = -\frac{\partial \mathcal{H}}{\partial x} = G Q G^T x + A w$$
(11)

where *w* is the adjoint vector. The covariance matrix P(t) with $P(t_0) = P_0 \ge 0$ (P_0 symmetric) is determined by

 $P(t) = \left(\Theta^{21}(t,t_0) + \Theta^{22}(t,t_0)P_0\right) \left(\Theta^{11}(t,t_0) + \Theta^{12}(t,t_0)P_0\right)^{-1}$ (12)

with

$$\Theta(t,t_0) = \begin{pmatrix} \Theta^{11}(t,t_0) \ \Theta^{12}(t,t_0) \\ \Theta^{21}(t,t_0) \ \Theta^{22}(t,t_0) \end{pmatrix}$$
(13)

being the transition matrix of the continuous-time stateadjoint equations. As shown in [25, Sect. 10], the continuous state-adjoint system in (11) has a unique matrix solution pair (X(t), W(t)) that satisfies initial conditions $X(t_0) =$ $\mathbb{1}, W(t_0) = P_0$. The solution pair satisfies the identity

$$W(t) = P(t)X(t), \quad \forall t \ge t_0, \tag{14}$$

from which the covariance equation (12) can be derived.

Lemma 6.1. The state-adjoint system (11) is Hamiltonian and the transition matrix $\Theta(t, t_0)$ is symplectic.

Proof. The system matrix of the combined state adjoint equations is $S := \begin{pmatrix} -A^T & C^T R^{-1} C \\ GQG^T & A \end{pmatrix}$ and since $C^T R^{-1} C$ and GQG^T are symmetric, it satisfies $(JS)^T = JS$ with $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. So we have a (linear) Hamiltonian system, for which is the flow is defined by the transition matrix. Thus, it follows that $\Theta(t,t_0)$ is a symplectic matrix for all $t \ge t_0$. \Box

In our discrete-time approximation of the Riccati equation for the covariance matrix, we derive discrete-time covariance updates that exactly preserve the following two properties: the symplectic structure of the state-adjoint differential equations and their linear relationship defined by P(t). As in the continuous-time Kalman filter, the update equation for the covariance matrix P_k does not depend on the state or on the adjoint. For known system and error covariance matrices at all time nodes, $\{P_k\}_{k=1}^N$ can be computed beforehand.

Theorem 6.1. Let a system and measurement model be given as in Proposition 6.1 for a 2*n*-dimensional mechanical system with state $(q, p)^T$. Let $t_d = \{t_k\}_{k=0}^N$ denote a set of discrete time points, with $t_k = kh$ for step size *h*, and assume P_0 symmetric is given. We denote by Ψ_k the one-step map of the symplectic Euler⁵ discretization of the 4-dimensional

state-adjoint system (11),

$$\begin{pmatrix} v_{k+1} \\ y_{k+1} \\ \lambda_{k+1} \\ \mu_{k+1} \end{pmatrix} = \begin{pmatrix} v_k \\ y_k \\ \lambda_k \\ \mu_k \end{pmatrix} + h \begin{pmatrix} -A^T & C^T R^{-1} C \\ GQG^T & A \end{pmatrix} \begin{pmatrix} v_{k+1} \\ y_k \\ \lambda_k \\ \mu_{k+1} \end{pmatrix},$$
(15)

with coordinates $x_k = (v_k, y_k)^T$ and $w_k = (\lambda_k, \mu_k)^T$. Note that $v_k, y_k, \lambda_k, \mu_k \in \mathbb{R}^n$. Then, Ψ_k is symplectic by construction. Further, if the discrete covariance updates for $k = 0, 1, \dots, N-1$ are computed by

$$P_{k+1} = (\Psi_k^{21} + \Psi_k^{22} P_k) \cdot (\Psi_k^{11} + \Psi_k^{12} P_k)^{-1}, \qquad (16)$$

with the linear one-step map Ψ_k being identified with its representing matrix and $\Psi_k = \begin{pmatrix} \Psi_k^{11} & \Psi_k^{12} \\ \Psi_k^{21} & \Psi_k^{22} \end{pmatrix}$, then we have that $w_k = P_k x_k$ for all $k = 0, 1, \dots, N$.

Proof. Since $\Theta(t,0)$ is a transition matrix, it satisfies the property

$$\Theta((k+1)h,0) = \Theta((k+1)h,kh) \cdot \Theta(kh,0).$$

The symplectic Euler discretization of (11) approximates the solution on the time grid $t_d = \{t_k\}_{k=0}^N$, $x_k \approx x(hk)$, $w_k \approx w(hk)$. Thus, we have Ψ_k as an approximation of $\Theta((k+1)h, kh) = \Theta(t_{k+1}, t_k)$,

$$\begin{pmatrix} x_{k+1} \\ w_{k+1} \end{pmatrix} = \Psi_k \begin{pmatrix} x_k \\ w_k \end{pmatrix} \approx \Theta((k+1)h, kh) \begin{pmatrix} x_k \\ w_k \end{pmatrix},$$

and by construction, Ψ_k is symplectic (the symplectic Euler scheme indeed leads to a symplectic one-step map, see [7] for a proof).

The discrete state-adjoint relations, $(x_{k+1}, w_{k+1})^T = \Psi_k(x_k, w_k)^T$, and the discrete covariance updates as defined in equation (16) are consistent in the way that a linear relationship between state and adjoints via the covariance matrices is ensured, i.e. $w_k = P_k x_k$ for all k = 0, 1, ..., N (compare to equation (14) in the continuous-time setting):

$$P_{k+1}x_{k+1} = w_{k+1}$$

$$\Leftrightarrow P_{k+1} \left(\Psi_k^{11}x_k + \Psi_k^{12}w_k \right) = \Psi_k^{21}x_k + \Psi_k^{22}w_k$$

$$\Leftrightarrow P_{k+1} \left(\Psi_k^{11} + \Psi_k^{12}P_k \right) x_k = (\Psi_k^{21} + \Psi_k^{22}P_k)x_k$$

$$\Leftrightarrow P_{k+1} = \left(\Psi_k^{21} + \Psi_k^{22}P_k \right) \left(\Psi_k^{11} + \Psi_k^{12}P_k \right)^{-1}. \square$$

Note that as long as we have a time-invariant estimation problem, Ψ_k is a constant matrix that can be obtained from solving the set of equations (15) for $(v_{k+1}, y_{k+1}, \lambda_{k+1}, \mu_{k+1})$ and generating the matrix entries just once. If the system and measurement model matrices are time-varying though, this step would have to be performed in every iteration.

⁵Note that the states of the Hamiltonian system do not coincide with the filter state. Since the Hamiltonian structure of the mechanical system matrix *A* suggests a splitting of the state and adjoint variables, though, we *choose* a symplectic Euler that treats half of the states and adjoints as implicit and the other half as explicit. However, we point out that other splittings between the set of coordinates define a first order symplectic integration scheme, as well.

For estimation problems of mechanical systems, it is desirable to also preserve the mechanical structure (i.e. Hamiltonian matrix A, or Hamiltonian plus forcing), as demonstrated in Section 4. Therefore, we combine the structurepreserving covariance update of Theorem 6.1 with discrete symplectic state updates. This leads to the following algorithm, that resembles the continuous-time filter (see Prop. 6.1) in that it does not split into prediction and measurement update steps as the discrete Kalman filter does (cf. Prop. 4.1).

Algorithm 6.1. For the implementation of the structurepreserving Kalman filter with symplectic state and covariance updates using symplectic Euler integration, we modify the continuous-time Kalman filter algorithm of [20] to

- 1. take measurement⁶ z_{k+1} ,
- 2. update covariance matrix by

$$P_{k+1} = \left(\Psi_k^{21} + \Psi_k^{22} P_k\right) \left(\Psi_k^{11} + \Psi_k^{12} P_k\right)^{-1},$$

3. compute Kalman gain as $K_{k+1} = P_{k+1} C^T R^{-1}.$

5. compute Kaiman gain as
$$\kappa_{k+1} = P_{k+1}C$$

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} q_k \\ p_k \end{pmatrix} + h \left[(A - K_{k+1}C) \begin{pmatrix} q_{k+1} \\ p_k \end{pmatrix} + K_{k+1}z_{k+1} \right]$$

Step 4 is again an implicit, but linear update, so it can be analytically inverted beforehand and then directly evaluated during filtering. While Algorithm 6.1 may not initially appear to follow the same procedure of the standard discrete Kalman filter (Proposition 4.1) in that there is no *prediction* or *update* step, we point out that one could rewrite the standard discrete Kalman filter to give a single covariance update equation analogous to the update rule in Step 2 of Algorithm 6.1. In general the covariance update from Algorithm 6.1 and from Proposition 4.1 will not result in the same covariance evolution.

The previous results specifically use a first-order symplectic scheme for two reasons: We want to give practitioners a good idea how to implement the filter and we want to have a fair comparison to the standard approach that is based on non-symplectic, first-order, explicit Euler discretizations (see Example 6.1). However, if higher order symplectic discretization schemes, e.g. a second order midpoint rule, are used, only the calculation of the one-step map Ψ_k based on a modified equation (15) and the state update in Step 4 of the algorithm need to be adjusted.

Example 6.1 (Stochastic Harmonic Oscillator). We revisit the example of a stochastic harmonic oscillator as introduced in Example 2.1. The structure-preserving Kalman filter method as presented in Algorithm 6.1 is compared to the discrete-time variants introduced in Section 4: the standard covariance updates with either an explicit Euler discretization for the discrete filter state update or a symplectic discretization by midpoint rule. Figure 8 shows the resulting filter behavior in state space (top row), as well as the covariance

matrix eigenvalues (bottom left) and the mechanical energy of the oscillator (bottom right). All integrators use a step-size of h = 0.03125 and identical measurements generated from one sample path. Unlike the examples presented in Section 4, the symplectic state updates alone do not improve the filter; likely due to the simplistic nature of the harmonic oscillator. However, the novel variant of a structure-preserving filter has smaller covariances, as can be seen from the ellipses in the top plot and the covariance eigenvalues. In particular, the mechanical energy of the filtered state, which would ideally — in the absence of noise and without numerical errors be a constant, is much more accurately preserved than in the other two cases.

7 Conclusion

This paper presents three structure-preserving filtering methods for mechanical control systems: 1) Particle filter with symplectic state predictions, 2) Extended Kalman filter with symplectic state predictions and structured linearizations of variational integrators, and 3) Kalman filter with symplectic state update and additional symplectic covariance update. The filtering techniques have been tested in simulated examples and in a real-time experiment with a planar crane system. Symplectic discrete-time predictions of the states increased the performance of both the particle and the extended Kalman filter compared to the standard choice of explicit Euler updates. Replacing the covariance updates in the linear Kalman filter can improve its performance even further. We derive a modified discrete-time Riccati equation for the covariance matrices that preserves the Hamiltonian/symplectic structure of the system of optimality conditions that belongs to the continuous-time Kalman filter. This novel filter variant can be extended to nonlinear systems by linearizing the discretized state-adjoint system of the optimality conditions. The one-step map of the symplectic state-adjoint scheme, which is used for the covariance Riccati equations, will then become time-dependent. Thus, one would expect an increased computational effort. However, this will be balanced by the observed higher accuracy of the method which allows a coarser time-grid than the nonsymplectic first order methods. Further theoretical studies of the symplectic filtering approach can be performed by backwards error analysis [7] applied to the state-adjoint system. We also note that recent work [18] has shown it may be possible to construct arbitrarily high-order variational integrators that are still linearizable one-step maps. These integrators are developed through backward error analysis [7] by constructing surrogate Lagrangians. High-order VI-based filtering techniques is another area for future investigation. In many real-world applications, the filtering methods presented herein would be combined with optimal control techniques. First results of a combination with the presented filters are given in [13]; a detailed study is left for future work.

Acknowledgements

This material is based upon work supported by Army Research Office grant W911NF-14-1-0461, and by the Na-

⁶In simulated examples we generate measurements by $z_{k+1} = Cx((k+1)h) + w_{k+1}$ where x(t) is some ground truth trajectory, e.g. obtained from a high-order integration with small step sizes and w_k is a noise vector drawn from the assumed Gaussian distribution.



Fig. 8. Comparison of the Kalman filter variants for the harmonic oscillator. Analytic deterministic solution is given in black, standard discrete covariance updates with explicit Euler state updates in blue, VI midpoint state updates with standard discrete covariance updates in green, and symplectic Euler state update with symplectic covariance update in red. Top left figure shows the filtered states and some covariance ellipses in the phase plane. The top right figure is a zoomed-in view of the data in the top left figure. The symplectic covariance updates increase the filter performance. Bottom left figure shows the time evolution of the eigenvalues of the covariance matrices (as in Fig. 2 there are two eigenvalues per filter). The symplectic method leads to eigenvalues that are an order of magnitude smaller. The bottom right figure shows the mechanical energy of the oscillator's filtered states for each of the filter variants; the symplectic covariance update yields a better approximation that the two other filters with standard covariance updates.

tional Science Foundation under Grant CMMI 1334609. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- Kalman, R. E., 1960. "A new approach to linear filtering and prediction problems". *J. of Basic Engineering*, 82(1), pp. 35–45.
- [2] Stengel, R. F., 1986. *Optimal Control and Estimation*. Dover Publications, Mineola, NY.
- [3] Thrun, S., Burgard, W., and Fox, D., 2005. Probabilistic Robotics. Intelligent Robotics and Autonomous Agents. The MIT Press.
- [4] Thrun, S., 2002. "Particle filters in robotics". In Proc. of the 17th Annual Conf. on Uncertainty in AI (UAI).
- [5] Chirikjian, G. S., 2009. Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods. Applied and Numerical

Harmonic Analysis. Birkhäuser Boston, Sept.

- [6] Marsden, J. E., and West, M., 2001. "Discrete mechanics and variational integrators". *Acta Numerica*, 10, pp. 357–514.
- [7] Hairer, E., Lubich, C., and Wanner, G., 2006. Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd ed., Vol. 31 of Springer Series in Computational Mathematics. Springer.
- [8] Ober-Blöbaum, S., Junge, O., and Marsden, J. E., 2011. "Discrete mechanics and optimal control: an analysis". *Control, Optimisation and Calculus of Variations*, 17(2), pp. 322–352.
- [9] Johnson, E., Schultz, J., and Murphey, T., 2014. "Structured linearization of discrete mechanical systems for analysis and optimal control". *IEEE Trans. on Automation Science and Engineering*, **12**(1), July, pp. 140–152.
- [10] Milstein, G. N., Repin, Y. M., and Tretyakov, M. V., 2002. "Symplectic integration of hamiltonian systems with additive noise". SIAM J. on Numerical Analysis,

39(6), pp. 2066-2088.

- [11] Bou-Rabee, N., and Owhadi, H., 2009. "Stochastic variational integrators". *IMA J. of Numerical Analysis*, 29(2), pp. 421–443.
- [12] Sanyal, A. K., Lee, T., Leok, M., and McClamroch, N. H., 2008. "Global optimal attitude estimation using uncertainty ellipsoids". *System & Control Letters*, 57, pp. 236–245.
- [13] Schultz, J., and Murphey, T., 2014. "Extending filter performance through structured integration". In American Controls Conf. (ACC), pp. 430–436.
- [14] Flaßkamp, K., and Murphey, T., 2015. "Variational integrators in linear optimal filtering". In American Controls Conf. (ACC), pp. 5140–5145.
- [15] Mazzoni, T., 2007. "Computational aspects of continuous–discrete extended Kalman-filtering". *Computational Statistics*, 23(4), pp. 519–539.
- [16] Patel, H. G., and Sharma, S. N., 2014. "Third-order continuous-discrete filtering for a nonlinear dynamical system". *J. of Computational and Nonlinear Dynamics*, 9(3), Feb.
- [17] Emran, B. J., Al-Omari, M., Abdel-Hafez, M. F., and Jaradat, M. A., 2015. "Hybrid low-cost approach for quadrotor attitude estimation". *J. of Computational and Nonlinear Dynamics*, **10**(3), May.
- [18] De La Torre, G., and Murphey, T., 2016. "On the benefits of surrogate lagrangians in optimal control and

planning algorithms". In IEEE Int. Conf. on Decision and Control (CDC). accepted for publication.

- [19] Lew, A., Marsden, J. E., Ortiz, M., and West, M., 2004.
 "Variational time integrators". *Int. J. for Numerical Methods in Engineering*, **60**(1), pp. 153–212.
- [20] Lewis, F. L., Xie, L., and Popa, D., 2008. *Optimal and Robust Estimation*, 2nd ed. CRC Press.
- [21] Schultz, J., and Murphey, T. D., 2013. "Embedded control synthesis using one-step methods in discrete mechanics". In American Controls Conf. (ACC), pp. 5393–5298.
- [22] Murphey, T., and Johnson, E., 2011. "Control aesthetics in software architecture for robotic marionettes". In American Controls Conf. (ACC), pp. 3825–3830.
- [23] Schultz, J., and Murphey, T., 2012. "Trajectory generation for underactuated control of a suspended mass". In IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 123–129.
- [24] Johnson, E., and Murphey, T. D., 2007. "Dynamic modeling and motion planning for marionettes: Rigid bodies articulated by massless strings". In IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 330– 335.
- [25] Kalman, R. E., and Bucy, E. S., 1961. "New results in linear filtering and prediction theory". *J. of Basic Engineering*.