Variational Integrators in Linear Optimal Filtering

Kathrin Flaßkamp[1] and Todd D. Murphey[1]

Abstract— **Discrete-time estimation and control techniques play a crucial role in digital control architectures. These methods rely on accurate approximations of continuous-time system behavior. For mechanical systems, this includes not only the system state, but also mechanical properties such as symplecticity or the long-term energy behavior. Additionally, we aim to preserve the Hamiltonian structure of optimally controlled or filtered systems. In this contribution, it is discussed how these requirements can be met when replacing the standard discretization schemes by variational integrators. We show that if one chooses a symplectic discretization scheme for a Kalman filtering problem, the discretization inherits the Hamiltonian structure of the continuous-time linear quadratic problem. Numerical experiments with this filter show better results than obtained with standard discretization.**

I. INTRODUCTION

Estimation and filtering techniques play a crucial role in signal processing tasks in robotics, aerodynamics or automotive engineering. A widely used concept in digital control architectures is the discrete-time Kalman filter. Thus, numerical discretization techniques are required for approximating the continuous-time dynamics of the system. To meet the various requirements on a filter in real-time applications, in particular, the numerical discretization scheme has to be chosen carefully.

Variational integrators (VI) (cf. [1] and Section II for a short introduction) provide a simulation method for mechanical systems that has shown great benefits in applications (cf. [2], [3], [4], [5], for instance). Since VI are derived from mechanical variational principles, they preserve specific properties of the continuous-time system in the discrete solution and show a good long-term energy behavior. This is important not only for forward simulation, but also for control and estimation problems for uncertain mechanical systems. The theory of variational integrators and symplectic integrators has been extended to stochastic mechanical systems in [6] and in [7]. In [8], stochastic variational integrators have been successfully applied for estimation problems on Lie groups using a method based on uncertainty ellipsoids instead of the Kalman filtering approach. Both numerical and experimental evidence for a benefit of VI in comparison to standard integration schemes is given in [5], [9].

[1]Kathrin Flaßkamp and Todd D. Murphey are with the Neuroscience and Robotics Laboratory, McCormick School of Engineering and Applied Sciences, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208 kathrin.flasskamp@northwestern.edu
This material is partially based upon work supported by the United States National Science Foundation under Grant CMMI 1200321. 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.

In this paper, we focus on the (linear) *Kalman filtering (KF) problem*, addressing it by transferring results on its dual, the *linear quadratic optimal control (LQ) problem*. As is particularly clear in the extended Kalman filter method, many optimal control and filtering algorithms for *nonlinear systems* are actually based on time-varying linearizations of the nonlinear system dynamics in combination with the linear methods.

Variational filters (see e.g. [10]), a class of Bayesian filtering schemes which use variational methods in the sampling step, are not considered in this work.

The contributions of this paper can be found in Section III and Section IV. In Section III, we derive discrete state-adjoint equations for the LQ problem which preserve the original structure of the time-continuous problem by applying theory of optimal control methods (cf. [11], [2]). By this, it is guaranteed that the Hamiltonian structure of the optimality conditions is mapped to the discretization in terms of a nearby Hamiltonian (in the sense of backwards error analysis, cf. [12]).

This type of *symplectic state-adjoint discretization schemes* is then transferred to the linear filtering problem, which is the major focus of this work (cf. Section IV). In real applications, digital control interfaces apply the discrete-time Kalman filter (cf. [13]). This requires a discretization of the continuous-time system dynamics. The choice of discretization scheme has an impact on the filter performance (see e.g. Section IV-A), but it is not clear from the discrete-time filter definition which discretization should be chosen. For this reason, we take the continuous-time Kalman filter (see [14]) as the starting point and derive a discrete-time version which provides a clear structure-preserving relation to the original problem (cf. Section IV-B). As a consequence of duality (cf. [13], [14]), the optimal filter of a continuous-time system can be also determined by a state-adjoint system with Hamiltonian structure. Thus, we choose a symplectic state-adjoint discretization scheme as developed for LQ problems in Section III. The resulting new version of a discrete Kalman filter is structure-preserving w.r.t. i) the Hamiltonian structure of the mechanical system and ii) the symplectic relationship between state and adjoint equations of an optimal filter in the discrete update equations for the estimated state and its covariance matrix. This is numerically illustrated with the stochastic harmonic oscillator example. Providing a comparison to standard implementations which are typically based on an explicit Euler method, we use a symplectic scheme of the same order, namely the *symplectic Euler method* (cf. [12]). From computational results, we found that for the symplectic Euler method, state estimates with better statistical properties and better approximation of the system energy can be obtained.

II. VARIATIONAL INTEGRATORS

We give a short introduction to variational mechanics and discuss VI advantages in applications.

Consider a mechanical system with configuration $q(t)$ and velocity $\dot{q}(t)$, for time $t \in [0, T]$, whose dynamical behavior is described by a Lagrangian L. This is typically the difference of kinetic and potential energy, whereas the Hamiltonian H is the sum of all energies. In addition, there are forces $f(q, \dot{q}, u)$ such as friction or a control input force that influence the system's dynamics. These dynamics are described by the variational equation (cf. [1], [2])

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

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

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

Under certain regularity assumption (cf. [1] for details), the equations can be equivalently written as Hamilton equations in configuration and momentum variables $p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q})$ with Hamiltonian $H(q, p) = p \cdot \dot{q} - L(q, \dot{q})$:

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

In the absence of forces, the flow of a Hamiltonian system is *symplectic*[1]. 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

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

and the forcing term is discretized in a similar manner with discrete forces f_k^+, f_k^-. Taking variations of the discrete configurations q_k for $k = 1, \ldots, N-1$ leads to *discrete forced Euler-Lagrange equations* ([1], [2]) for $k = 1, \ldots, N-1$,

$$D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) + f_k^- + f_{k-1}^+ = 0.$$

Discrete momenta are given by the relationships $p_k = -D_1 L_d(q_k, q_{k+1}) - f_k^-$ and $p_{k+1} = D_2 L_d(q_k, q_{k+1}) + f_k^+$ 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 [12] that (in the unforced case) a big class of variational integrators has a symplectic map Ψ_k and it falls into the category of symplectic partitioned Runge-Kutta methods. In [2], the theory is extended to external forces.

Example 1: Throughout this work, we will consider the *symplectic Euler integrator* which is a symplectic partitioned Runge-Kutta scheme of order one. For a system given in Hamilton equations, it is implicitly defined by

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} q_k \\ p_k \end{pmatrix} + \begin{pmatrix} \frac{\partial}{\partial p} H(q_{k+1}, p_k) \\ -\frac{\partial}{\partial q} H(q_{k+1}, p_k) + f(q_{k+1}, p_k, u_k) \end{pmatrix}.$$

In applications, VI have shown numerous advantages:

- *(long-term) forward integration:* Besides conservation of symplecticity and symmetries in terms of invariances, VI have an excellent energy behavior even in long-term simulations, which is important in astrodynamics or molecular dynamics (cf. [12]). Furthermore, VI exactly satisfy constraints of e.g. multibody systems.
- *optimal control:* VI can be used to design an indirect or a direct optimal control method (cf. e.g. [2]) that takes over the advantages of forward integration. Additionally, it needs fewer optimization variables and provides structured linearizations for an efficient optimization (cf. [3]).
- *estimation:* VI can be extended to stochastic integrators (see [7], [6]) which almost surely preserve the symplectic structure. They have great statistical properties and show further benefits in real-time filtering (cf. [9]).

III. VI FOR LINEAR QUADRATIC OPTIMAL CONTROL

We consider a standard optimal control problem for a linear mechanical system and a quadratic cost criterion.

Problem 1:

$$\min_{x,u} J = \frac{1}{2} \int_0^T \left(x^T Q x + u^T R u \right) dt + x(T)^T P_0 x(T)$$

$$\text{w.r.t. } \dot{x} = Ax + Bu,\ x(0) = x^0, \tag{2}$$

$$\text{with } x = \begin{pmatrix} q \\ p \end{pmatrix} \text{ and } A \text{ a Hamiltonian matrix.}$$

Further, we assume symmetric weighting matrices satisfying $Q \geq 0$, $P_0 \geq 0$, and $R > 0$. We emphasize that all matrices are allowed to be time-varying, but we neglect the time-dependence in the notation for simpler reading. An optimal solution (x, u) to Problem 1 satisfies the necessary optimality conditions

$$(i)\ \frac{\partial \mathcal{H}}{\partial u} = 0, \quad (ii)\ \dot{x} = \frac{\partial \mathcal{H}}{\partial \lambda},\ \dot{\lambda} = -\frac{\partial \mathcal{H}}{\partial x},$$
$$(iii)\ \begin{pmatrix} x(0) \\ \lambda(T) \end{pmatrix} = \begin{pmatrix} x^0 \\ P_0 x(T) \end{pmatrix},$$

where the *problem's Hamiltonian* $\mathcal{H}$ is given by $\mathcal{H}(x, u, \lambda) = \frac{1}{2}\left(x^T Q x + u^T R u\right) + \lambda^T (Ax + Bu)$ and λ denotes the *adjoint variable*.

By applying a numerical integration scheme to the state and discretizing the control input, we obtain a discrete-time version of Problem 1.

Problem 2: Fix a time grid $\{t_0 = 0, t_1, \ldots, t_N = T\}$, and, for simplicity, $t_k - t_{k-1} = h$ for $k = 1, \ldots, N$. Then, a discrete-time LQ problem for $x_d = \{x_k\}_{k=0}^N$, $u_d =$

[1]In general, a map is symplectic if it preserves a *symplectic form* ([1], [12]). To get an intuition for symplecticity, it is helpful to consider the fact that symplectic transformations are area-preserving. Linear symplectic maps $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ satisfy $A^T J A = J$ with $J = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}$.

$\{u_k\}_{k=0}^{N-1}$ is given by

$$\min_{x_d,u_d} = \sum_{k=0}^{N-1} \frac{1}{2}\left[x_k^T Q_k x_k + u_k^T R_k u_k\right] + x_N^T P_0 x_N$$

w.r.t. $x_{k+1} = \Psi_k\left(x_k,\, u_k\right)$, for $k = 0, \ldots, N-1$, $x_0 = x^0$.

The linear one-step map Ψ_k can be obtained from an integration scheme extended by the forcing term and it approximates the continuous-time state transition matrix between t_k and t_{k+1}.

Example 2: (a) The discrete map Ψ_k of a *forced explicit Euler integrator* with step size h for a system (2) with control input is given by

$$\Psi_k(x_k, u_k) = \left(\mathbb{I} + hA_k,\, hB_k\right) \cdot \begin{pmatrix} x_k \\ u_k \end{pmatrix}$$

since $x_{k+1} = x_k + h(A_k x_k + B_k u_k)$.

(b) The update matrix for the discrete map Ψ_k of a *forced symplectic Euler* with step size h and with $x = \left(q,\, p\right)^T$ is given by

$$\begin{pmatrix} S_k & S_k h A_k^{12} & S_k h B_k^1 \\ hA_k^{21} S_k & h^2 A_k^{21} S_k A_k^{12} + \mathbb{I} + hA_k^{22} & h(B_k^2 + hA_k^{21} S_k B_k^1) \end{pmatrix},$$

with $S_k = (\mathbb{I} - hA_k^{11})^{-1}$. This can be derived from the linear update equation

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} q_k \\ p_k \end{pmatrix} + h\left[\begin{pmatrix} A_k^{11} & A_k^{12} \\ A_k^{21} & A_k^{22} \end{pmatrix}\begin{pmatrix} q_{k+1} \\ p_k \end{pmatrix} + \begin{pmatrix} B_k^1 \\ B_k^2 \end{pmatrix} u_k\right].$$

Here and in the following, when applying a symplectic Euler to similar systems, we assume the matrix $(\mathbb{I} - hA_k^{11})$ to be invertible.

Proposition 1 (Symplectic discretization of Problem 1): Assume a symplectic Euler integration scheme is used for the state map Ψ_k in Problem 2. Then, we subsume from [2] and [12]:

(i) Discrete adjoints can be derived from the discrete Hamiltonian of Problem 2 which satisfy the equations of the symplectic Euler scheme applied to the continuous time adjoint equations.

(ii) The combined discrete state-adjoint trajectories generate a symplectic Euler scheme for the continuous-time Hamiltonian of Problem 1.

(iii) A symplectic Euler scheme applied to the state-adjoint system exactly samples a modified Hamiltonian.

Namely, the symplectic adjoint scheme for $\lambda = \left(\mu,\, \rho\right)$ is

$$\begin{pmatrix} \mu_{k+1} \\ \rho_{k+1} \end{pmatrix} = \begin{pmatrix} \mu_k \\ \rho_k \end{pmatrix} - h\left[A_k^T \begin{pmatrix} \mu_k \\ \rho_{k+1} \end{pmatrix} + Q_k \begin{pmatrix} q_{k+1} \\ p_k \end{pmatrix}\right];$$

a proof with detailed derivation has to be left for a future publication though.

Directing towards real-time implementations in robotics, we restrict to the symplectic Euler as the simplest, lowest order VI, but it has to be mentioned that analogous results can be obtained for general nonlinear systems, and for higher order integration schemes for the state and higher order control approximations (cf. [2]). Furthermore, it has to be pointed out that starting with a non-symplectic scheme for the state system also results in a symplectic scheme for the state-adjoint discretization (cf. [11]). Then, however, these combined schemes may be of lower order and do not take into account the Hamiltonian structure of the mechanical system.

In the continuous LQ problem, the *optimal* state trajectory together with its adjoint obey a specific structure, namely they are generated by a symplectic flow defined by the problem's Hamiltonian. Underlying, we have the system dynamics, which are also Hamiltonian w.r.t. to the mechanical Hamiltonian augmented by the forcing term.

It is our aim to preserve both these structures in a discrete-time setting used for numerical computations. As we see from Proposition 1, starting with a symplectic Euler discretization (more generally, any symplectic Runge-Kutta scheme) for the state differential equations and deriving the discrete adjoint equations as described above meets this goal.

In Section IV, it is shown how these requirements can be met in the dual problem, i.e. the optimal linear filtering problem. Then, we go a step further and replace the discrete state-adjoint update by an update equation of the covariance matrix which is based on the linear relationship between states and adjoints, but does not directly depend either on the discrete state or on the adjoint equations. Numerical experiments illustrate the improvements compared to standard explicit Euler discretization, in particular regarding an approximation of the mechanical system's energy.

IV. VI FOR KALMAN FILTERING

The goal of this section is to derive a *structure-preserving optimal linear filter for stochastic Hamiltonian systems*. As a starting point, we consider the following formulations of discrete-time and continuous-time Kalman filters, which have been originally developed by Kalman in [13] and Kalman and Bucy in [14]. For a textbook reference, we refer to [15].

Proposition 2 (Discrete-time Kalman Filter, [13]): Let a model of the system and the measurement be given by

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

with $R_k > 0$ and $\{w_k\}_{k=0}^N$, $\{v_k\}_{k=0}^N$ uncorrelated white noise processes. 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: $P_{k+1}^- = A_k P_k A_k^T + G_k Q_k G_k^T$, $\hat{x}_{k+1}^- = A_k \hat{x}_k$

Measurement update:

$$\begin{aligned} K_{k+1} &= P_{k+1}^- C_{k+1}^T (C_{k+1} P_{k+1}^- C_{k+1}^T + R_{k+1})^{-1}, \\ P_{k+1} &= (\mathbb{I} - K_{k+1} C_{k+1}) P_{k+1}^-, \\ \hat{x}_{k+1} &= \hat{x}_{k+1}^- + K_{k+1}(z_{k+1} - C_{k+1}\hat{x}_{k+1}^-). \end{aligned}$$

Proposition 3 (Continuous-time Kalman Filter, [14]): Let the model of the system and the measurement be given by

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

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

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

Assume a (linear) mechanical system, modeled with continuous time differential equations, and an estimation problem to be given. In order to apply the discrete Kalman filter (Prop. 2), a discrete-time system matrix A_k has to be provided which can be generated by an integration scheme applied to the differential equation. In Section IV-A, we study the difference between symplectic and non-symplectic system matrices A_k. However, this has not taken into account the Hamiltonian structure of the state-adjoint system, yet. Therefore, in Section IV-B, we start with the continuous filter of Prop. 3 and approximate the full state-adjoint equations of this problem by a symplectic integration scheme. This results in a discrete filter which preserves the symplectic structure of the continuous-time problem, but it is not necessary to explicitly solve for the adjoint equations. Instead, a discrete Riccati equation is derived.

A. Discrete Kalman filter for discretized system dynamics

In Kalman's original work [13], the state transition matrix of the continuous-time system is used for A_k. In general, this transition matrix cannot be computed analytically, but it can be approximated by an integration scheme applied to the continuous differential equation.

Example 3: Analogous to Example 2, when applying the explicit Euler scheme to the dynamical system $\dot{x} = Ax$ one obtains the discrete update matrix $A_k = (\mathbb{I} + hA(kh))$, whereas the symplectic Euler scheme leads to the one-step map

$$A_k = \begin{pmatrix} (\mathbb{I} - hA_k^{11})^{-1} & (\mathbb{I} - hA_k^{11})^{-1} A_k^{12} \\ hA_k^{21}(\mathbb{I} - hA_k^{11})^{-1} & hA_k^{21}(\mathbb{I} - hA_k^{11})^{-1} A_k^{12} + \mathbb{I} + hA_k^{22} \end{pmatrix}.$$

a) The stochastic harmonic oscillator: The differential equations of a harmonic oscillator with configuration $q \in \mathbb{R}$ and momentum $p \in \mathbb{R}$ can be derived from its Hamiltonian $H(q,p) = \frac{1}{2}\left(q^2 + p^2\right)$ as

$$\frac{d}{dt}q = \frac{\partial H}{\partial p} = p, \qquad \frac{d}{dt}p = -\frac{\partial H}{\partial q} = -q$$
$$\Leftrightarrow \dot{x} = \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}.$$

We assume that the full state can be measured, $C_k = \mathbb{I}$, and also $G_k = \mathbb{I}$. We simulate a measurement by taking the exact state from the continuous-time system at the specific time and add a random value drawn from a Gaussian distribution, $w_k \sim \mathcal{N}(0, R_k)$.

b) Numerical results: We apply the Kalman filter as described in Proposition 2 with A_k being either the discrete update matrix from an explicit Euler or from a symplectic Euler scheme, as derived in Example 3. We perform 100 runs with $N = 60$ time steps each and a step size $h = 0.5$. The error variances are chosen to be $Q_k = \operatorname{diag}(1,1)$, $R_k = \operatorname{diag}(1,100)$ and we start from state $x^0 = (5,1)^T$

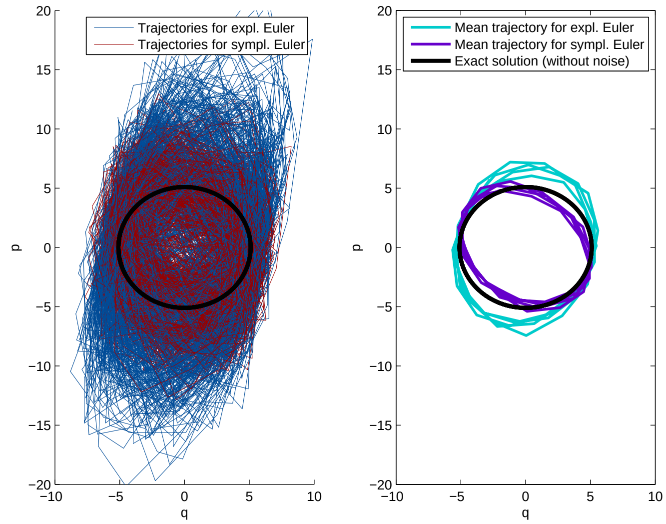

Fig. 1. Estimated states (configuration and momentum coordinates) from 100 runs of the discrete Kalman filter (cf. Proposition 2) applied to the harmonic oscillator with system matrix A_k obtained by symplectic Euler map (red, purple) and explicit Euler map (blue, cyan). In the left figure, it can be seen that many runs of the explicit Euler shows a larger deviation than the symplectic Euler runs. From the right figure we see that also the mean of the symplectic Euler trajectories is closer to the original solution and approximates the volume of the exact periodical solution.

with $P_0 = \operatorname{diag}(1,1)$. In Fig. 1, the individual runs as well as their means are shown for the explicit Euler in blue and cyan, respectively and for the symplectic Euler in red and purple, respectively. For comparison, the exact solution, i.e. in continuous-time and without noise, is given in black. When the explicit Euler is used, there are more runs with high deviation in the estimated states and also the mean trajectory is not as good as the one obtained for the symplectic Euler. This has an important influence on the estimated energy of the oscillator, which is computed by $\hat{E}_k = \hat{q}_k^2 + \hat{p}_k^2$.

In Fig. 2, we show the energy mean over the explicit Euler runs (blue) and the symplectic Euler runs (red), respectively, as well as a moving average (for a window length of 5) that smoothes the estimates over time (purple and cyan). For comparison, the exact energy value corresponding to $x^0 = (5,1)^T$ is given in black. The deviation of the energy value approximated by explicit Euler from the exact energy value is three to four times higher than that of the symplectic Euler. The reason why the symplectic Euler's energy differs from the true value might be that even in the non-stochastic case, a symplectic integrator exactly preserves the energy of a nearby Hamiltonian system and not exactly the continuous-time system's energy.

B. Structure-preserving discretization of the continuous-time Kalman filter

In [14], the duality relation to the LQ problem is used to state a Hamiltonian for the filtering problem,

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

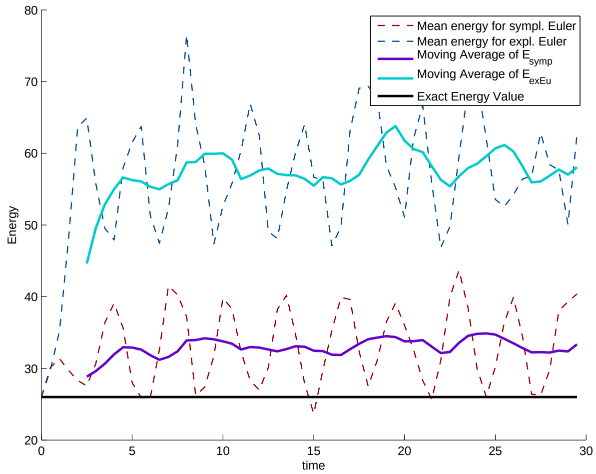

Fig. 2. Estimated energy for the harmonic oscillator corresponding to the state estimates in Figure 1. First, the mean of the energy values is taken over all runs and for each time-step leading to discrete energy curves E_{symp} and E_{exEu}. Then, a moving average with window length 5 is computed. The error of the approximated energy when using the explicit Euler is much higher than for the symplectic Euler.

The corresponding state-adjoint equations are

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

and the covariance matrix $P(t)$ with $P(0) = P_0 \geq 0$ (P_0 symmetric) is determined by

$$P(t)=[\Theta_{21}(t,0) + \Theta_{22}(t,0)P_0][\Theta_{11}(t,0) + \Theta_{12}(t,0)P_0]^{-1}$$

with $\Theta(t,0) = \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}$ being the transition matrix of the state-adjoint equations.

Proposition 4: A structure-preserving discretization of the continuous-time Kalman filter (cf. Prop. 3) is obtained by

$$\begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} = \Psi_k \begin{pmatrix} x_{k-1} \\ \lambda_{k-1} \end{pmatrix}$$

with Ψ_k being a symplectic integration map. Then, the linear relation P_k between discrete states and adjoints, i.e. $P_k x_k = \lambda_k$, is recursively determined by

$$P_{k+1} = (\Psi^{21}_{k+1} + \Psi^{22}_{k+1} P_k) \cdot (\Psi^{11}_{k+1} + \Psi^{12}_{k+1} P_k)^{-1}$$

for $k = 0, 1, \ldots, N$, where $\Psi_k = \begin{pmatrix} \Psi^{11}_k & \Psi^{12}_k \\ \Psi^{21}_k & \Psi^{22}_k \end{pmatrix}$.

Proof: Since the continuous-time state-adjoint differential equations have been derived from the Hamiltonian $\mathcal{H}(q,p)$, $\Theta(t,0)$ is a symplectic map. Thus, we want to derive a discrete update equation that preserves this property. For this reason, we apply a symplectic integration map Ψ_k to approximate the exact flow:

$$\Theta(kh,0) \begin{pmatrix} x(0) \\ \lambda(0) \end{pmatrix} = \begin{pmatrix} x(kh) \\ \lambda(kh) \end{pmatrix} \approx \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} =: \Psi_k \begin{pmatrix} x_{k-1} \\ \lambda_{k-1} \end{pmatrix}.$$

Note that it is now not necessary to compute the full discrete state-adjoint trajectory. Instead, we derive a recursive update equation for the covariance. Starting from the linear relationship for P at node $k+1$, we obtain

$$\begin{aligned} & P_{k+1} x_{k+1} = \lambda_{k+1} \\ \Leftrightarrow\ & P_{k+1} \left(\Psi^{11}_{k+1} x_k + \Psi^{12}_{k+1} \lambda_k\right) = \Psi^{21}_{k+1} x_k + \Psi^{22}_{k+1} \lambda_k \\ \Leftrightarrow\ & P_{k+1} \left(\Psi^{11}_{k+1} x_k + \Psi^{12}_{k+1} P_k x_k\right) = \Psi^{21}_{k+1} x_k + \Psi^{22}_{k+1} P_k x_k \\ \Leftrightarrow\ & P_{k+1} = \left(\Psi^{21}_{k+1} + \Psi^{22}_{k+1} P_k\right) \left(\Psi^{11}_{k+1} + \Psi^{12}_{k+1} P_k\right)^{-1}. \end{aligned}$$

■

As in the continuous-time and standard discrete-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, $\{P_k\}_{k=1}^N$ could be computed beforehand. However, note that our update equation for P_k differs from the discrete-time Kalman update. Also, the update for the estimated state is different because we choose a matching symplectic discretization, e.g. a symplectic Euler, as well. An implementation of the VI-discretized Kalman Filter using symplectic Euler integration is comprised of the following steps

1) take measurement $z_{k+1} = C_{k+1} x((k+1)h) + w_{k+1}$,
2) update covariance matrix $P_{k+1} = \left(\Psi^{21}_{k+1} + \Psi^{22}_{k+1} P_{k+1}\right)\left(\Psi^{11}_{k+1} + \Psi^{12}_{k+1} P_k\right)^{-1}$,
3) compute Kalman gain $K_{k+1} = P_{k+1} C^T_{k+1} R^{-1}_{k+1}$,
4) update state $x_{k+1} = \Phi_{k+1} \begin{pmatrix} x_k \\ z_{k+1} \end{pmatrix}$.

The state update 4) is defined by a linear symplectic Euler map Φ_k which is augmented by the measurement term. Concretely, we use the discrete scheme

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

c) Numerical results for the harmonic oscillator: We run the algorithm sketched above for the harmonic oscillator example for 3000 steps, with step size $h = 0.5$ and initial value $x^0 = (5,1)^T$. The variances are set to $Q = \text{diag}(1,1)$ and $R = \text{diag}(1,100)$.

For comparison, we then replace both symplectic Euler matrices, Ψ_k for the covariance update and Φ_k for the state update, by the corresponding maps for an explicit Euler integrator and rerun the algorithm.

In Fig. 3, the estimated state trajectories[2] are shown as a phase plane plot with red for the explicit Euler and blue for the symplectic Euler solution. The corresponding energy values of the harmonic oscillator are shown in Fig. 4.

Again, we smooth the stochastic values over time by a moving average with window length 50 (green for explicit Euler, cyan for symplectic Euler). Influenced by the step size h and the error covariance for the measurement R, the advantage of the symplectic Euler can be clearly seen in both plots. While the estimated states from explicit Euler show a considerable offset from the true solution, the states generated by the symplectic Euler have smaller variations and, on average over time, approximately preserve the volume of the cyclic exact solution. This corresponds to

[2] In contrast to the previous example, we now take a single, but long-term ($T = 1500$) simulation.

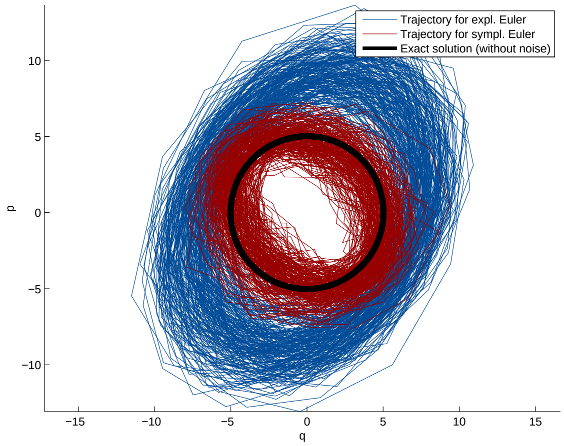

Fig. 3. Estimated long-term trajectory from the discretized version of the continuous-time Kalman Filter for the harmonic oscillator. The solution from the structure-preserving filter based on the symplectic Euler scheme is shown in red and the trajectory obtained with the explicit Euler scheme in blue. While the symplectic method gives good approximations of the true solution (shown in black), the explicit Euler leads to high deviations.

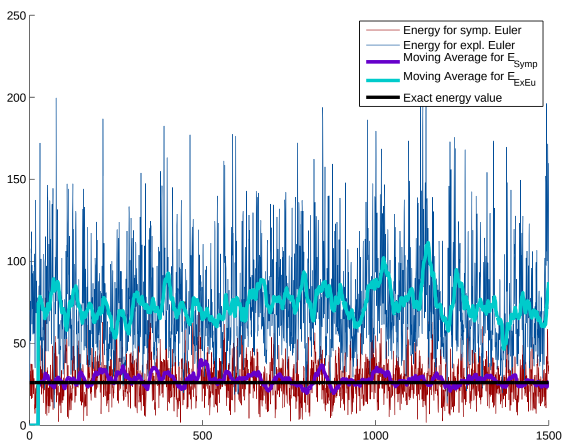

Fig. 4. Estimated energy for the stochastic trajectories of the harmonic oscillator given in Figure 3. Again, the blue curve refers to the filter with explicit Euler scheme and the red curve to the symplectic Euler. The moving average for the energy from the symplectic filter (purple) is close to the true energy value (black).

the good energy values compared to the highly inaccurate and thus useless approximation of the true energy value of $E = 26.0$ by the explicit Euler. Note that after transitioning from the standard discrete Kalman filter formulation to the structure-preserving discretization of the continuous-time Kalman filter, the numerical approximation of the energy, in particular, has improved reasonably. One reason for this, which has to be examined further in future work, might be that we now choose matching update equations for the state and the covariance matrix which are both based on symplectic integration schemes.

V. CONCLUSIONS AND OUTLOOK

Recent experimental results showing the excellent real-time performance of variational integrators in filtering and control problems (cf. [5]) motivated us to study the interplay between symplectic variational integration and Kalman filtering solutions in this work. To this aim we derive discrete filter equations for the estimated state and for the covariance matrix based on variational integrator discretizations. Thus, the discrete filter inherits the symplecticity of the continuous-time Kalman filter. As we are able to illustrate even by the simple example of a stochastic harmonic oscillator, a structure-preserving discretization of the Kalman filter leads to a better estimate of the mechanical energy and to better estimates of filter states.

In future work, we would like to tie together the results for LQ and KF problems to address the linear quadratic Gaussian problem. To reveal the relationship between the standard discrete Kalman filter and the developed structure-preserving variant, further theoretical studies are necessary, which can be based on backward error analysis (cf. [12]). Then, the next step should be to address nonlinear, possibly constrained, mechanical systems, for which linearizations have to be included in the Kalman filter. As it has been shown both numerically and experimentally in [5], [3], the structured linearizations of variational integration schemes provide further benefit compared to standard explicit Euler integrations.

REFERENCES

[1] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.
[2] S. Ober-Blöbaum, O. Junge, and J. E. Marsden, "Discrete mechanics and optimal control: an analysis," *Control, Optimisation and Calculus of Variations*, vol. 17, no. 2, pp. 322–352, 2011.
[3] E. Johnson, J. Schultz, and T. Murphey, "Structured linearization of discrete mechanical systems for analysis and optimal control," *IEEE Transactions on Automation Science and Engineering*, 2014.
[4] K. Flaßkamp, J. Timmermann, S. Ober-Blöbaum, and A. Trächtler, "Control strategies on stable manifolds for energy-efficient swing-ups of double pendula," *International Journal of Control*, vol. 87, no. 9, pp. 1886–1905, 2013.
[5] J. Schultz and T. Murphey, "Extending filter performance through structured integration," in *Proceedings of the American Control Conference*, 2014.
[6] N. Bou-Rabee and H. Owhadi, "Stochastic variational integrators," *IMA Journal of Numerical Analysis*, vol. 29, no. 2, pp. 421–443, 2009.
[7] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov, "Symplectic integration of Hamiltonian systems with additive noise," *SIAM Journal on Numerical Analysis*, vol. 39, no. 6, pp. 2066–2088, 2002.
[8] A. K. Sanyal, T. Lee, M. Leok, and N. H. McClamroch, "Global optimal attitude estimation using uncertainty ellipsoids," *System & Control Letters*, vol. 57, pp. 236–245, 2008.
[9] J. Schultz, "Discrete mechanics computation for real-time embedded control," Ph.D. dissertation, Northwestern University, 2014.
[10] K. J. Friston, "Variational filtering," *NeuroImage*, vol. 41, pp. 747–766, 2008.
[11] W. W. Hager, "Runge-Kutta methods in optimal control and the transformed adjoint system," *Numerische Mathematik*, vol. 87, pp. 247–282, 2000.
[12] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed., ser. Springer Series in Computational Mathematics. Springer, 2006, vol. 31.
[13] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
[14] R. E. Kalman and E. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, 1961.
[15] F. L. Lewis, L. Xie, and D. Popa, *Optimal and Robust Estimation*, 2nd ed. CRC Press, 2008.