Extending Filter Performance Through Structured Integration

Jarvis Schultz and Todd D. Murphey

Abstract— **Estimation and filtering form an important component of most modern control systems. Techniques such as extended Kalman filters and particle filters have been successfully utilized for estimation in many different applications. Integrators derived from discrete mechanics possess desirable numerical properties such as stable long-time energy behavior, exact constraint satisfaction, and accurate statistical calculations. In the present work, we leverage these features by utilizing a variational integrator derived from discrete mechanics within extended Kalman filters and particle filters. By filtering real experimental data from the nonlinear, underactuated planar crane problem we demonstrate that the linearizations available through the discrete mechanics framework increase the accuracy of uncertainty estimates provided by an extended Kalman filter, especially when operating at low frequencies. Additionally, we illustrate situations where particle filter performance is increased through the statistics-preserving properties provided by the variational integrator.**

I. Introduction

In control systems, measurements of a system's state are utilized in a feedback loop to regulate and control the behavior of the system, and in a real-world application these measurements are subject to sensor noise. Additionally, there may be components of the system state that go unmeasured for a variety of reasons; an unreliable system model is then used to fill in information about these unmeasured states. Dealing with these challenges is the primary goal of estimation and filtering.

Likely the most well-known and widely used estimator is the Kalman filter [1]. When the Kalman filter is applied to a linear Gaussian system, it provides the optimal, maximum-likelihood (minimum-variance) 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. It is simple to compute, but limited because in its most common form, it is restricted to system models that are linear in their arguments and represent uncertainty as additive Gaussian noise. 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 because most nontrivial real-world systems are nonlinear, it is applicable to a much wider range of problems. One drawback of the EKF is that its performance is strongly dictated by the severity of any nonlinearities and the levels of uncertainty inherent in the system. That said, the EKF is quite common and it has been successfully applied to a wide range of problems.

More recently, particle filters have become a popular technique for estimation [3], [4]. In particle filters, the uncertainty posterior distribution is represented by a finite collection of parameters. These parameters, called *particles*, are drawn from a distribution representing the current belief of the system's state. Each particle is independently mapped to a future time using a noisy system model. In this way, 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]. In principle, the strength of the particle filter lies in its generality; it can handle any noise model, and it is still applicable when the system exhibits nonlinear dynamics. The main disadvantage of the particle filter lies in its computational complexity. Its reliability is primarily governed by the number of particles used, and in some problems, the number of particles required to begin approximating the true posterior distribution is quite large (10^4–10^6).

Recently, much work has been done in developing the theory of discrete mechanics and corresponding numerical methods produced by the theory. Discrete mechanics lies at the intersection of classical mechanics, computer science, and applied mathematics. The integrators that are derived through the discrete mechanics approach possess long-term energy stability, they provide exact momentum conservation, they are symplectic, and they are able to accurately predict statistics of systems [6]–[8]. While these properties are well established, the implementation of these integrators in real-world systems is limited to a small number of examples [9]–[12]. In this work we investigate the application of a structured integration technique called a variational integrator, which is derived from discrete mechanics theory, in place of a classical integrator in the estimation of mechanical systems. Several examples are presented where the numerical properties of this variational integrator lead to improved estimation performance for both EKFs and particle filters.

In Section II an overview of discrete mechanics, and the derivation of the particular integrator used in the present study is provided. Section III discusses several important numerical features related to estimator performance of this integrator, as well as discrete mechanics in general. Section IV utilizes this integrator with particle filters and ex-

J. Schultz jschultz@u.northwestern.edu
T.D. Murphey t-murphey@northwestern
Department of Mechanical Engineering, Northwestern University, Evanston, IL

This material is based upon work supported by the National Science Foundation under award IIS-1018167. 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.

tended Kalman filters applied to estimation problems of both simulated and experimental data. Since standard particle filter and EKF algorithms are used the details of these algorithms are omitted. It is recommended to consult [3] for the specifics of the implemented algorithms — EKF algorithm on page 59 and particle filter algorithm on page 98. Finally, Section V discusses the conclusions that may be drawn based on these results.

II. Discrete Mechanics

In the discrete mechanics framework, one attempts to find a sequence $\{(t_0,q_0),(t_1,q_1),\ldots,(t_n,q_n)\}$ that approximates a continuous time trajectory of a system i.e. $q_k \approx q(t_k)$ where t is time, and $q \in Q$, the configuration space of the system. In traditional variational mechanics, Hamilton's principle is used to derive governing differential equations for a system. In practice, numerical integration techniques must be used to approximate their solutions. In discrete mechanics, approximations are instead applied to the underlying physical principle, and then variational methods produce governing difference equations. To begin this derivation, one approximates a system's Lagrangian using an arbitrary quadrature rule over a timestep $\Delta t = t_{k+1} - t_k$

$$L_d(q_k,q_{k+1}) \approx \int_{t_k}^{t_{k+1}} L(q(\tau),\dot{q}(\tau))d\tau \tag{1}$$

(this quantity is referred to as the discrete Lagrangian). Next the action integral is approximated with an action sum as

$$S(q[t_0,t_f]) = \int_{t_0}^{t_f} L(q(\tau),\dot{q}(\tau))d\tau \approx \sum_{k=0}^{n-1} L_d(q_k,q_{k+1}). \tag{2}$$

Hamilton's principle states that the evolution of a mechanical system is a stationary point in the action. Taking the first variation of Eq. (2), and invoking the fundamental lemma of the calculus of variations [13], one can derive the unforced, unconstrained Discrete Euler-Lagrange (DEL) equations[1]

$$D_1L_d(q_k,q_{k+1}) + D_2L_d(q_{k-1},q_k) = 0. \tag{3}$$

In this form the DEL equations provide a discrete map $(q_{k-1},q_k) \to (q_k,q_{k+1})$ that is implicitly solved using a numerical root-finding algorithm [7]. The precise quadrature rule chosen to define the discrete Lagrangian in Eq. (1) will determine the order of the integrator derived and whether it is an explicit or an implicit method. Any quadrature rule used to define a discrete Lagrangian produces a corresponding set of DEL equations; the integrator used to solve these DEL equations is referred to as a *variational integrator* (VI). In this work, we use the midpoint rule to define a VI yielding the following discrete Lagrangian

$$L_d(q_k,q_{k+1}) = L\left(\frac{q_k+q_{k+1}}{2}, \frac{q_{k+1}-q_k}{\Delta t}\right)\Delta t. \tag{4}$$

Differentiating Eq. (4) with respect to its arguments allows computation of all terms in Eq. (3) in terms of the original continuous Lagrangian.

[1]Here we have used the *slot derivative* notation where $D_i f(\cdot)$ represents a derivative of f with respect to its ith argument.

In the given form, the DEL equations provide a two-step map from $Q \times Q \to Q \times Q$. In Section III we discuss the advantages of one-step maps in relation to linearizations. To convert Eq. (3) to a one-step map, invoke the discrete Legendre transform to define the discrete generalized momentum as

$$p_k = -D_1L_d(q_k,q_{k+1}) = D_2L_d(q_{k-1},q_k). \tag{5}$$

Eq. (3) can now be viewed as the following set of equations

$$p_k = -D_1L_d(q_k,q_{k+1}) \tag{6a}$$

$$p_{k+1} = D_2L_d(q_{k-1},q_k) \tag{6b}$$

To integrate this form of the DEL equations, start with the given the pair (q_k,p_k) then implicitly solve Eq. (6a) to obtain q_{k+1}. Next evaluate Eq. (6b) to obtain p_{k+1}. Thus the DEL equations are now the one-step map $T^*Q \to T^*Q$ [7]. In this work we use Eq. (4) to define the discrete Lagrangian and we use the Newton-Raphson method to implicitly solve Eq. (6a).

As a final point, note that the theory presented in this section is easily extended to include external forcing, holonomic constraints, and kinematic inputs. For more information on these extensions see [7], [8], [14], [15].

III. Discrete Mechanics Applied to Estimation

In this section we discuss two key reasons why the variational integrator discussed in Section II should improve estimator performance.

A. Numerical Properties of the Integrator

The aforementioned desirable numerical properties of discrete mechanics integrators are one of the primary reasons they improve estimator performance. Exploring these numerical properties has been one of the main academic thrusts of the discrete mechanics and structured integration communities in the recent years. As a result, we only briefly mention some of the important results related to the present study, and then illustrate how these properties impact estimator performance using a numerical experiment of a simple system.

The first advantageous property of the present VI is that it is a symplectic integrator[2]. The key consequence of symplecticity is that for a given Hamiltonian system, the VI will provide the *exact* solution to a nearby Hamiltonian system called the *modified system.* The discrete trajectory produced by the VI is an exact sampling of the nearby Hamiltonian system, and thus exactly conserves the energy and momentum of the modified system. This is a key characteristic of the VI, and indeed all symplectic integrators. For explanations of the underlying mathematics, see [7], [16], or [17].

Another numerical benefit of the VI involves systems with constraints. A set of holonomic constraints of the form $h_i(q) = 0$, is incorporated in the integrator at the endpoints as $h_i(q_{k+1}) = 0$ [18]. As the DEL equations are only in terms

[2]We point out that there are many symplectic integrators including symplectic Runge-Kutta methods [16] and other variational integrators.

of configurations, the constrained DEL equations ensure that holonomic constraints are *exactly* satisfied at every time step. With traditional integrators applied to constrained mechanical systems one may directly integrate index-3 Differential Algebraic Equations (DAEs), in which configuration-level constraint satisfaction may introduce artificial noise in the simulated state trajectories [19]. Alternatively, one can use index-reduction techniques to obtain a set of index-1 DAEs where constraints are enforced at the acceleration level. When applying this technique, care must be taken to ensure that the simulation does not exhibit excessive constraint drift [20]. The VI inherently satisfies constraints and artificial noise is regulated by the dynamics of the modified system.

In the context of state estimation one benefit of the VI, and symplectic integrators in general, is related to the their statistics-preserving properties. Recent work has developed special symplectic methods for stochastic systems [6], [21], [22]. These integrators have been shown to better predict several statistical quantities than higher-order RK schemes [23], [24]. To illustrate these effects with a simple system, we borrow the example system from [21], but implement the results with the present integrator rather than the symplectic integrator they described. The example system is a linear harmonic oscillator with state-independent additive noise driven by a Wiener process. The governing stochastic differential equations are given by

$$dx^1 = x^2 dt + \sigma dw_1(t), \qquad x^1(0) = x_0^1 \tag{7a}$$
$$dx^2 = -x^1 dt + \gamma dw_2(t), \qquad x^2(0) = x_0^2. \tag{7b}$$

Using Ito calculus [5], one may determine an analytical solution for an individual sample path of this system. The effects of the present symplectic integrator on the statistics of this system may be visualized in Fig. 1. In Fig. 1, domains of three different discrete phase plane flows are produced using the exact solution, standard Euler-Maruyama integration [5], and the VI described in Section II. In the exact flow of the initial unit circle domain, the domain remains circular and is driven away from the origin by Wiener process forcing. Analyzing Fig. 1, one may see that the discrete flow produced by the VI follows the exact flow very closely. The flow produced by Euler-Maruyama integration artificially expands the radius of the domain, and the mean location of the domain drifts significantly from the true solution. This artificial noise injection degrades estimator performance; examples are shown in Section IV.

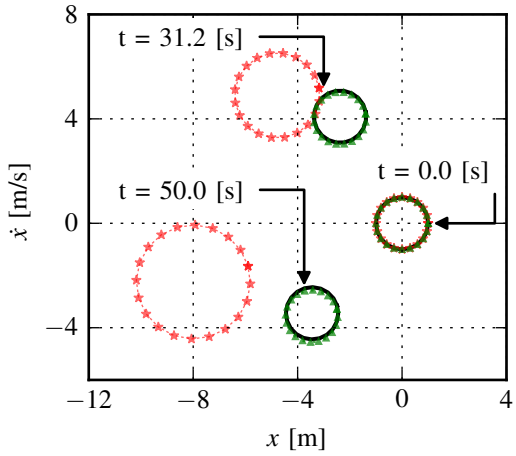

Fig. 1. Plot showing the effects of symplectic integrators for stochastic differential equations. Shown are the time evolutions of a domain of the phase plane of Eq. (7) using three different schemes – the exact mapping [21], Euler-Maruyama integration, and the VI described in Section II. At $t = 0.0$ s the domain is represented by a collection of points distributed on the unit circle. Each point's time evolution is then calculated with each scheme using an *identical* sample path of the Wiener process for all points and schemes. The solid black circles are generated using the exact mapping, the red stars with dashed lines are produced by Euler-Maruyama integrations, and the green triangles are produced by the present VI. The noise parameters are given by $\sigma = 1$, $\gamma = 1$, and the timestep is $dt = 0.031$ s.

B. Linearizations Provided by Discrete Mechanics

The VI presented in Section II has a unique characteristic referred to as a *structured linearization* [11], [25]. Many estimation algorithms, such as the EKF, rely on local linearizations to adapt linear system tools to nonlinear systems. The form of the discrete map provided by VIs admits particularly nice linearizations. 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

$$x_{k+1} = f_k(x_k, u_k). \tag{8}$$

One issue with this is that for most higher-order integrators the discrete map f_k from t_k to t_{k+1} involves evaluating the continuous function f at intermediate points between (x_k, u_k) and (x_{k+1}, u_{k+1}). This significantly complicates the Taylor expansion required to define the linearization. This is true for all Runge-Kutta methods above first order, but is not true for implicit or explicit Euler integration as they are first-order Runge-Kutta methods. In theory, more accurate linearizations could be obtained through proper expansions of higher-order integrators, or numerical techniques could be used to approximate the solution to the state transition matrix for higher-order integrators [25]. But in practice, if a linearization about (x_k, u_k) is required, and a "simple form" linearization is desirable, one is restricted to one-step methods.

As the present 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 [25]. Furthermore, since the VI is *exactly* sampling a nearby mechanical system, the VI linearization is an *exact* linearization of the nearby mechanical system. This restriction on the behavior of the linearization greatly improves its local accuracy. In Section IV we show an example of how the improved accuracy of this linearization leads to better EKF estimator performance.

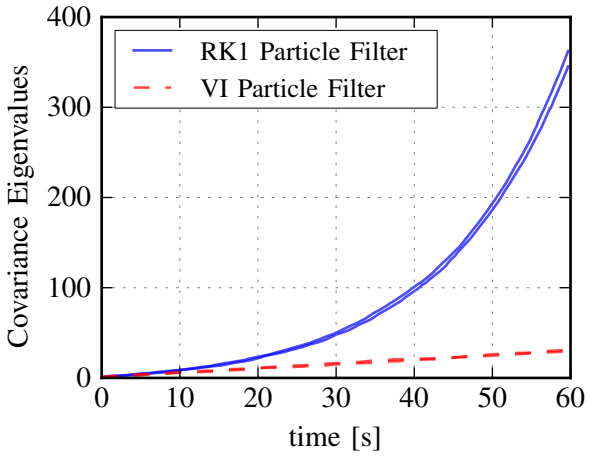

Fig. 2. Image of the particle filter covariance propagation for the harmonic oscillator without resampling.

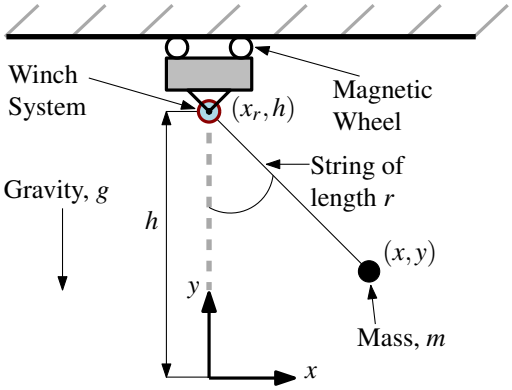

Fig. 3. Schematic of planar crane system including relevant geometric parameters.

IV. Results

A. Structured Integration for Particle Filters

In particle filtering algorithms, a collection of particles is mapped forward a single timestep using a noisy process model to represent 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 that have a low probability. In principle, as long as enough particles are used, and the measurement frequency is high enough this process works well. However, there exist many situations where measurement frequency may be unreliable, and one cannot count on the measurements to regulate the filter performance. Examples of this situation include high-dimensional data association problems, systems with unreliable communication networks, or tracking systems suffering from occlusions.

In situations where a particle filter must predict the state of the system over long time horizons without measurements symplectic integrators can significantly increase the performance of the particle filter. An illustration of this can be seen in Fig. 2. 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 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.

Note that even though a relatively large timestep was used for the simulation in Fig. 2, the overprediction of noise by the RK1 integrator would still be happening regardless of how small the timestep was. The specific system dynamics, the levels of noise in the system, and timestep size are all important factors in determining the horizon over which the unmeasured, unresampled RK1 particle filter will return meaningful predictions of the system's uncertainty. This is still true even for higher-order, non-symplectic integrators i.e. eventually, additional noise introduced by the integrator itself may significantly impact the quality of uncertainty estimates. The same is not true for a symplectic integrator; the bounds on the statistics guaranteed by symplecticity alleviate this issue. Thus given a system with sparse or unreliable measurements the use of symplectic integration techniques will increase the reliability of particle filter estimators.

One common issue with particle filters is referred to as *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 methods preventing this situation [3], we simply point out that an integrator adding inappropriate amounts of noise to an unregulated particle filter will increase the filter's probability of suffering from particle deprivation when measurements are incorporated. Results illustrating this point will be shown in the next section.

B. Estimation for the Planar Crane

In this section 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 [11] and [26]. 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 [14]. With this modeling strategy, the Lagrangian for the system is only a function of the dynamic configuration variables (x,y) and is given by

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

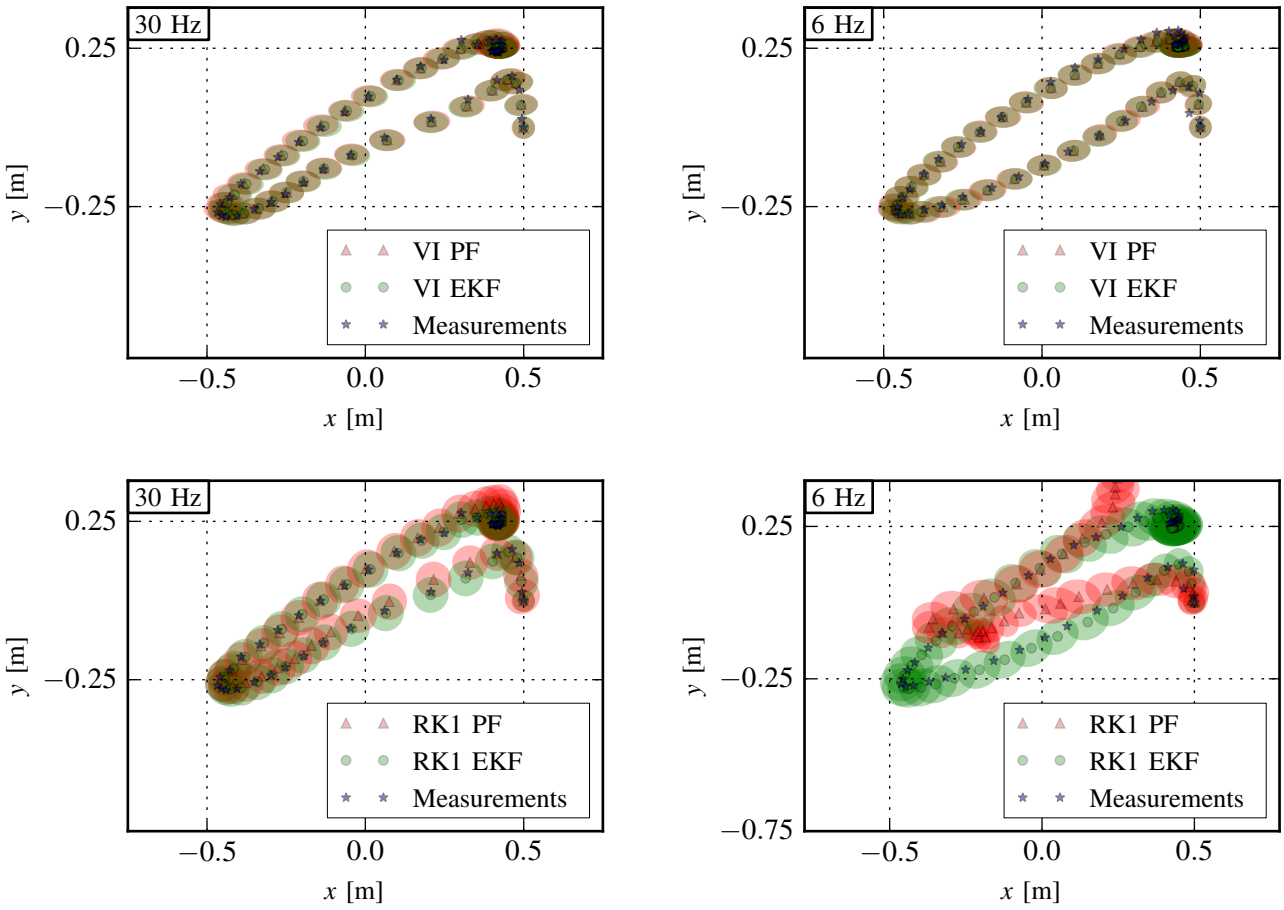

Fig. 4. This set of figures illustrates the results of utilizing a particle filter (PF) and an EKF to estimate the dynamic configuration variables of the system shown in Fig. 3. For each filter, both VI and RK1 integrators are used. In the left plots, measurements, controls and integrations happen at 30 Hz, while in the right plots they all occur at 6 Hz. To generate the data for a given frequency, an *actual* robotic system was sent a set of commands at the desired frequency and those commands were used to map both integrators forward. 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.

The corresponding midpoint discrete Lagrangian, as defined by Eq. (4), is given by

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

In order to emphasize the practical aspects of integration techniques discussed in the present work, *actual experimental data* from the robotic system described in [11] and [26] was used. This system utilizes digital encoders to close its control loops around x_r and r. Thus the kinematic-input modeling strategy was employed in the present work. Since these inputs are assumed to be perfectly controlled, the dynamic configuration variables (x,y) are the only state variables that are estimated.

Fig. 4 shows the parametric evolution of estimates of the dynamic configuration variables using particle filters and EKFs at two different frequencies using both a VI and an RK1 integrator to represent the system. The strength of the present 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 excellent 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 as the amount of noise introduced by the integrator is of the same order as the noise in the system model.

Fig. 5 shows the time evolution of the eigenvalues of the covariances from each of the estimators in Fig. 4. This figure further demonstrates that the VI covariance predictions are remarkably stable even at large timesteps. Additionally, it demonstrates that in this particular system the structured

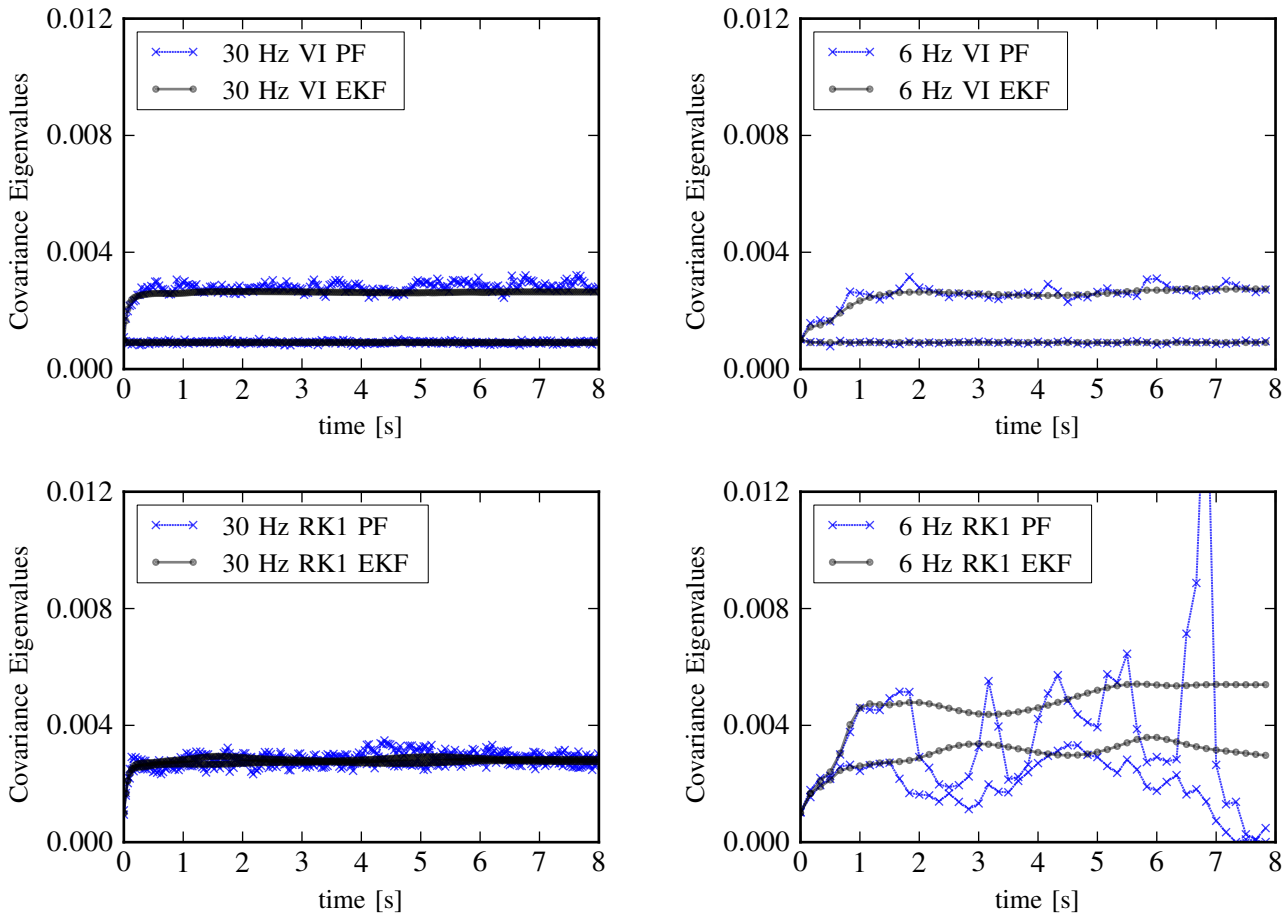

Fig. 5. This figure shows the time evolution of the eigenvalues of the covariance prediction for particle filters (PF) and EKFs applied to the planar crane problem of Fig. 3 with both a 30 Hz (left plots) and 6 Hz (right plots) measurement, estimator and integrator update rate. These eigenvalues are the same eigenvalues used to define the major and minor axes of the uncertainty ellipses plotted in Fig. 4. 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.

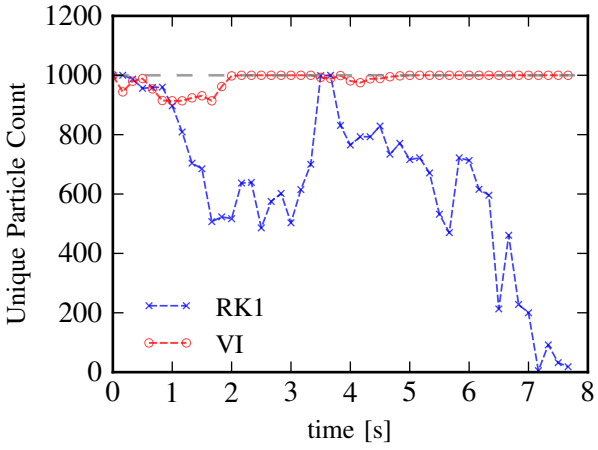

Fig. 6. Illustration of the VI improving particle deprivation. The red circles represent the 6 Hz VI particle filter, and the noisy, dashed blue line is the 6 Hz RK1 particle filter.

linearization allows the EKF to perform nearly identically to the particle filter. While this may not be true for arbitrary systems, it is noteworthy that it is true for *some* systems. In this case the increased accuracy of the structured linearization avoids the increased computational expense of the particle filter while achieving similar performance. It is also important to note that for high-dimensional systems, the number of particles required to achieve reliable particle filter performance prohibitively high due to computational costs. Thanks to the scalability of the VI [15], the VI-based EKF computational cost allows for real-time estimator computation even for systems of high state dimension.

In Section IV-A 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. 6 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 which implies uncertainty estimates from the filter are essentially useless.

V. Conclusions

The primary contribution of this work is the demonstration of improved particle filters and extended Kalman filters through the use of a structured integrator. We presented the basics of discrete mechanics, and used the theory to describe a midpoint variational integrator (VI). We then discussed several key numerical features of this VI, namely symplecticity and structured linearizations. We then applied standard estimation algorithms to several systems using traditional low-order Runge-Kutta integrators and the VI. We emphasize that as the integrator takes the same effective form as traditional integrators, no modification to standard estimation algorithms is necessary. Even though we only present results for several sample systems, the benefits of the VI can be expected to be seen for other systems. The only caveat is that in order to derive the VI structure, the system's governing equations must be derivable from a variational principle.

One downside to the VI is that it is an implicit integrator, and thus requires the use of a numerical root solving algorithm to advance the integrator. The added computation induced by the implicit nature of the VI is minimal and it has successfully been used for real-time estimators of complex, nonlinear systems using EKFs and particle filters. However the structure of the computation is problematic in particle filters where explicit schemes more-easily leverage parallel computation. Because the VI is less amenable to parallel computation it is challenging to implement real-time particle filters requiring high particle counts; this is not true for the EKF.

In some cases, the additional accuracy provided by the structured linearization elevates EKF performance to match that of the particle filter while saving significant computation. Particle filters benefit from the VI in situations where a filter may need to run without the aid of measurements over an appreciable time span. Additionally, as the VI helps prevent particle deprivation, VI-based particle filters can be expected to achieve similar performance with fewer particles as compared to traditional particle filters. Finally, the VI exhibits excellent stability of predictions for both particle filters and extended Kalman filters over a range of timesteps.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. of the ASME, J. of Basic Eng.*, vol. 82 (Series D), pp. 35–45, 1960.

[2] R. F. Stengel, *Optimal Control and Estimation*. Mineola, NY: Dover Publications, 1986.

[3] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, ser. Intelligent Robotics and Autonomous Agents. The MIT Press, 2005.

[4] S. Thrun, "Particle filters in robotics," in *Proc. of the 17th Annual Conf. on Uncertainty in AI (UAI)*, 2002.

[5] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods*, ser. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Sept. 2009.

[6] N. Bou-Rabee and H. Owhadi, "Stochastic variational integrators," *IMA J. Numerical Anal.*, vol. 48, no. 2, pp. 421–443, Apr. 2009.

[7] J. E. Marsen and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[8] O. Junge, J. E. Marsden, and S. Ober-blöbaum, "Discrete mechanics and optimal control," in *In Proc. of the 16th IFAC World Congress*, July 2005.

[9] T. Kai and K. Bito, "Swing-up control of the cart-pendulum system based on discrete mechanics," in *2011 Proc. of SICE Annual Conf. (SICE)*, Sept. 2011, pp. 1914 –1919.

[10] ——, "A new discrete mechanics approach to swing-up control of the cart-pendulum system," *Commun. in Nonlinear Sci and Numerical Simulation*, vol. 19, no. 1, pp. 230–244, Jan. 2014.

[11] J. Schultz and T. D. Murphey, "Embedded control synthesis using one-step methods in discrete mechanics," in *American Controls Conf. (ACC)*, Washington, D.C., 2013, pp. 5393–5298.

[12] J. Timmermann, S. Khattab, S. Ober-blöbaum, and A. Trächtler, "Discrete mechanics and optimal control and its application to a double pendulum on a cart," in *In Proc. of the 18th IFAC World Congress*, Aug. 2011, pp. 10 199–10 206.

[13] D. E. Kirk, *Optimal Control Theory: An Introduction*. Dover Publications, Apr. 2004.

[14] E. Johnson and T. D. Murphey, "Dynamic modeling and motion planning for marionettes: Rigid bodies articulated by massless strings," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, Roma, Italy, Apr. 2007, pp. 330–335.

[15] ——, "Scalable variational integrators for constrained mechanical systems in generalized coordinates," *IEEE Trans. on Robotics*, vol. 25, pp. 1249–1261, Oct. 2009.

[16] E. Hairer, C. Lubich, and G. Wanner, "Geometric numerical integration illustrated by the Störmer/Verlet method," *Acta Numerica*, vol. 12, pp. 399–450, 2003.

[17] S. Reich, "Backward error analysis for numerical integrators," *SIAM J. of Numerical Anal.*, vol. 36, pp. 1549–1570, 1996.

[18] E. Johnson and T. Murphey, "Dangers of two-point holonomic constraints for variational integrators," in *American Controls Conf. (ACC)*, St. Louis, MO, June 2009, pp. 4723–4728.

[19] O. A. Bauchau, C. L. Bottasso, and A. Epple, "Scaling of constraints and augmented lagrangian formulations in multibody dynamics simulations," *J. Computational Nonlinear Dynamics*, vol. 4, no. 2, Mar. 2009.

[20] G. Fábián, D. Van Beek, and J. Rooda, "Index reduction and discontinuity handling using substitute equations," *Math. and Comput. Modelling of Dynamical Syst.*, vol. 7, no. 2, pp. 173–187, 2001.

[21] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov, "Symplectic integration of hamiltonian systems with additive noise," *SIAM J. of Numerical Anal.*, vol. 39, no. 6, pp. 2066–2088, Jan. 2002.

[22] ——, "Numerical methods for stochastic systems preserving symplectic structure," *SIAM J. of Numerical Anal.*, vol. 40, no. 4, pp. 1583–1604, Jan. 2003.

[23] N. Bou-Rabee and J. E. Marsden, "Hamilton-pontryagin integrators on lie groups part I: Introduction and structure-preserving properties," *Found. Computational Math.*, vol. 9, no. 2, pp. 197–219, Mar. 2009.

[24] A. Lew, J. E. Marsden, M. Ortiz, and M. West, "Variational time integrators," *Int. J. for Numerical Methods in Eng.*, vol. 60, no. 1, pp. 153–212, 2004.

[25] T. Murphey and E. Johnson, "Control aesthetics in software architecture for robotic marionettes," in *American Controls Conf. (ACC)*, July 2011, pp. 3825–3830.

[26] J. Schultz and T. Murphey, "Trajectory generation for underactuated control of a suspended mass," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, May 2012, pp. 123–129.