2015 IEEE International Conference on Robotics and Automation (ICRA)
Washington State Convention Center
Seattle, Washington, May 26-30, 2015

Use of the Nonlinear Observability Rank Condition for Improved Parametric Estimation

Matthew Travers and Howie Choset

***Abstract*— The *correct way* to design controllers for dynamic robots is still very much an open question. This is in a large part due to the complexity and uncertainty in modeling their nonlinear dynamics. In this work, we focus on deriving concise dynamic expressions for a particular class of robots that can be used to better reduce uncertainty with respect to unknown parameters in realtime. We accomplish this by using an extended Kalman filtering framework in conjunction with an online controller that continuously maximizes a local measure of nonlinear observability. The main novel contribution of this work is that we directly use the nonlinear observability rank condition to derive the measure of observability at each time step. We are able to make this extension in part by focusing on serial-chain systems and exploiting the geometric structure in their dynamic models. In particular, we derive concise, closed-form and exact analytical representations for the forward dynamics, linearization, and nonlinear observability rank condition of a fixed-base serial manipulator with actively controlled elastic joints. An example is presented in which the spring constants and damping coefficients for a series-elastic actuated manipulator are estimated using the online observability maximizing techniques we derive.**

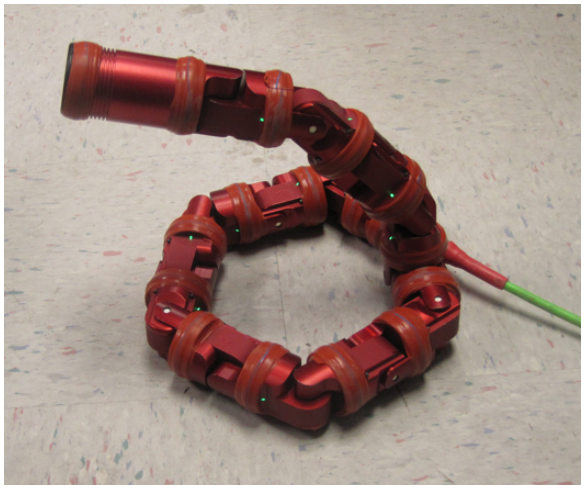

Fig. 1: Quasi-fixed base series-elastic actuated serial-chain manipulator.

I. INTRODUCTION

The inability to precisely model and measure the properties of a robot makes designing controllers for them almost as much of an art as it is a science. Advances in dynamic modeling, sensor development, and nonlinear estimation have helped, but measurement and modeling uncertainties continue to be barriers to the development of effective nonlinear controllers for robots. This work focuses on how advances in modeling and estimation can be combined in a novel way to help address modeling uncertainty; specifically, parametric uncertainty in dynamic systems. We present a method which directly uses properties of the nonlinear observability rank condition to modify an online controller, ultimately enabling a set of unknown system parameters to converge to their true values faster than a standard filtering-based parameter estimation approach.

Observability in both linear and nonlinear systems is traditionally treated as a discrete property; the system either is or is not observable. Recently, several researchers [8], [5] have developed the notion that there is a degree to which a system is observable and that this property is a measurable quantity. Hinson *et. al.* [5] have moved one step further, asking the question whether, for controlled systems, controllers can be designed to maximize observability and subsequently improve online estimation.

In this paper we present an approach which is related to that in [5]. Our main novel extension is that we use the nonlinear observability rank condition in place of the observability gramian as our basis for computing the degree to which a system is observable, or rather its *observability measure*. The advantage of this approach is at this point primarily computational. As we show in the context of serial-chain manipulators with series-elastic joints, it is possible to obtain the nonlinear observability rank condition for relatively complicated systems with an arbitrary number of joints in exact closed form. For moderately sized systems this allows the observability measure to be computed and partiality optimized using an online controller; a major advantage in uncertain environments where desired objectives may be subject to sudden changes. The authors in [8] and [9] do present a method for approximating the observability gramian that grows linearly with the number of states in the system, but requires two numerical integrations per evaluation of the gramian in addition to the calculation of the dynamics required for the method presented in this paper.

Deriving a concise representation for a serial-chain manipulator's dynamics is not novel, but is critical to perform the calculation of the observability rank condition in realtime. A portion of this paper thus reviews results from the literature on computationally efficient forward dynamic calculations. We use a pre-existing batch Newton-Euler formulation to derive analytic forms for both the generalized inertia as well as Coriolis and centripetal terms which appear in the second-order dynamics of an elastic-joint serial manipulator. We use these expressions to analytically linearize the system. We then show that the nonlinear observability rank condition for these systems can be written entirely in terms of the forward

dynamics and linearization.

To demonstrate the benefit of our approach, we simulate a six-link series-elastic actuated (SEA) serial-chain manipulator with rotational joints (originally inspired by the quasi-fixed base modular snake robot configuration shown in Figure 1 [15]). We assume that we are able to measure only the motor- and load-side angular positions, but wish to estimate the positions, velocities, as well as the spring constants and damping coefficients for each of the actuators. We use an extended Kalman filter (EKF) framework to estimate the full state of the system; including the positions, velocities, and spring parameters [10]. We present results which show that the filter converges to the set of originally unknown parameter values faster when the observability measure optimization is incorporated into the controller as opposed to when it is not. A comparison is provided.

II. Problem Definition

Consider the system

$$\begin{aligned} \dot{\bar{x}} &= \bar{f}(\bar{x}, \mathcal{S}, u(t)) \\ y &= \bar{h}(\bar{x}, \mathcal{S}, u(t)), \end{aligned} \tag{1}$$

where $\bar{x} \in \mathbb{R}^{\bar{n}}$, $\mathcal{S} \in \mathbb{R}^p$ is a vector of constant but unknown system parameters, $u(t) \in \mathbb{R}^{n_u}$ is a vector of control inputs, and $y \in \mathbb{R}^m$ is a vector of the measured system outputs. The problem we wish to solve is to estimate the set of parameters $\mathcal{S}$ as fast as possible in the presence of process and measurement uncertainty.

For notational convenience, throughout the remainder of the paper we will rewrite (1) in terms of the extended state $x = (\bar{x}, \mathcal{S}) \in \mathbb{R}^n$ where $n = \bar{n} + p$. The system dynamics thus become

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x, u), \end{aligned} \tag{2}$$

where $f(x, u) = (\bar{f}(\bar{x}, \mathcal{S}, u), 0_p)$.

III. Background

A. Nonlinear Observability

We provide a general description of nonlinear observability. A thorough derivation and more in depth explanation is left to [4], [6], [17]. We start by making several assumptions on the inputs and outputs of the systems we are interested in. In particular, the output function $h(\cdot)$ is assumed to be an explicit function of the input, i.e., $h(x, u) = (h_1(x, u), h_2(x, u), \ldots, h_m(x, u))$. We also require that the input functions be at least jth-order differentiable. In general j will be defined by the particular system; we assume only that $j \geq 1$. We define the following expressions in terms of the *Lie derivative*, $L_\cdot(\cdot)$,

$$\phi_i^0 = h_i$$

$$\phi_i^j = L_f \phi_i^{j-1} = \frac{\partial \phi_i^{j-1}}{\partial x} f + \sum_{k=0}^{j-1} \frac{\partial \phi_i^{j-1}}{\partial u^{(k)}} u^{(k+1)} = y_i^{(j)}.$$

We also define the map $\Phi(x, \mathbf{u}) : \mathbb{R}^n \times \mathbb{R}^{n_u} \to \mathbb{R}^n$

$$\Phi(x, u) = (h_1, \phi_1^1, \ldots, \phi_i^{k_1 - 1}, \ldots, h_m, \phi_m^1, \ldots, \phi_m^{k_m - 1})^T, \tag{3}$$

where, $\sum_{i=1}^m k_i = n$ [3].

Definition 1 The system (2) is locally observable if the mapping $\Phi(x, \mathbf{u})$ is invertible with respect to the state x and its inverse is smooth for all $(x, \mathbf{u})$ in a local subset of $\mathbb{R}^n \times \mathbb{R}^{n_u}$, or rather, if

$$\text{rank}\left(\frac{\partial \Phi(x, \mathbf{u})}{\partial x}\right) = n. \tag{4}$$

Equation (4) is the *nonlinear observability rank condition*.

There are several important things to note about Φ. First, as Hinson *et. al.* point out [5], the derivatives of the components ϕ_i^{j-1} with respect to the controls u and derivatives thereof appear explicitly in Φ and subsequently in the (4). Fundamentally, this reflects the fact that for a nonlinear system the control is implicitly coupled to the estimation. This fact fundamentally explains why the results presented in [5] and in this paper are possible. Secondly, the only requirement on the individual k_i in (3) is that they sum to n. This implies that Φ is in general not unique. We will use this fact in Section IV to facilitate the calculation of (4).

B. Dynamics of Chained Systems

This work was originally inspired by the desire to control a fixed-base series-elastic actuated serial-topology system with rotational joints. Based on this desire we restrict further investigation to this particular class of serial-chain systems. We first combine several previous results which make it possible to concisely express the dynamics of these mechanisms. We then show how these expressions can be used to analytically linearize the associated second-order dynamics. In Section VI we use these results to express both Φ as well as the nonlinear observability rank condition (4) in explicit closed-form for up the N joints.

1) Geometric Preliminaries: The material in this section is drawn from several different sources. Details are provided where necessary to present a clear description, but further explanations are left to [11], [13], [1], [2]. We start by defining a rigid body transformation $g \in \text{SE}(3)$ between frames i and k in terms of the matrix exponential operator

$$\begin{aligned} g_{i,k}(\theta) &= \exp(\xi_i \theta_i) \exp(\xi_{i+1} \theta_{i+1}) \ldots \exp(\xi_k \theta_k) g_{i,k}(0) \\ &= \begin{bmatrix} R & q \\ 0 & 1 \end{bmatrix}, \end{aligned} \tag{5}$$

where $R \in \text{SO}(3)$ is a three-dimensional rotation matrix, $q \in \mathbb{R}^3$, ξ_j is a *twist* [11], and the operator $\hat{\cdot}$ takes the vector $\xi_j = (v_j, \omega_j) \in \mathbb{R}^6$ into the $\mathbb{R}^{4\times 4}$ matrix representation

$$\hat{\xi}_j = \begin{bmatrix} \hat{\omega}_j & v_j \\ 0 & 0 \end{bmatrix}. \tag{6}$$

There is a second $\hat{\cdot}$ operator in (6) which maps $\omega \in \mathbb{R}^3$ into the set of 3×3 skew symmetric matrices, i.e.,

$$\hat{\omega} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \tag{7}$$

We will not explicitly differentiate between these operators in the rest of this paper, but it will be clear from context which instance is being used.

Several of the calculations in the forward dynamics and linearization of a serial-chain mechanism depend on the directional derivatives of the individual link's forward kinematics. The chained matrix exponentials in (5) make calculating these derivatives relatively simple, i.e., for $j \leq k$

$$\frac{\partial}{\partial \theta_j} g_{1,k}(\theta) = \exp(\hat{\xi}_1 \theta_1) \exp(\hat{\xi}_2 \theta_2) \ldots \\ \hat{\xi}_j \exp(\hat{\xi}_j \theta_j) \ldots \exp(\hat{\xi}_k \theta_k) g_{1,k}(0), \tag{8}$$

and $i \leq j \leq k$

$$\frac{\partial}{\partial \theta_i} \frac{\partial}{\partial \theta_j} g_{1,k}(\theta) = \exp(\hat{\xi}_1 \theta_1) \exp(\hat{\xi}_2 \theta_2) \ldots \hat{\xi}_i \exp(\hat{\xi}_i \theta_i) \ldots \\ \hat{\xi}_j \exp(\hat{\xi}_j \theta_j) \ldots \exp(\hat{\xi}_k \theta_k) g_{1,k}(0). \tag{9}$$

The *adjoint transformation* transforms elements ξ of the *Lie algebra* se(3) (which is associated with SE(3)) from one coordinate frame to another. Given some $g \in$ SE(3), the adjoint transformation $\mathrm{Ad}_g : \mathbb{R}^6 \to \mathbb{R}^6$ has the matrix form

$$\mathrm{Ad}_g = \begin{bmatrix} R & \hat{q}R \\ 0 & R \end{bmatrix}$$

There is a second adjoint transformation which maps elements of the *dual Lie algebra* se(3)* between different coordinate frames as well

$$\mathrm{Ad}_g^* = \begin{bmatrix} R^T & -R^T \hat{q} \\ 0 & R^T \end{bmatrix}$$

In addition to adjoint transformations there are *adjoint representations*, which map the action of the Lie algebra onto itself, $\xi \in$ se(3), ad_ξ : se(3) $\to$ se(3), which has the matrix representation

$$\mathrm{ad}_\xi(\cdot) = \begin{bmatrix} \hat{\omega} & \hat{v} \\ 0 & \hat{\omega} \end{bmatrix}. \tag{10}$$

The adjoint representation can also be defined in terms of the *Lie bracket*, $\mathrm{ad}_\xi \eta = [\xi, \eta]$ for $\xi = (v_1, \omega_1)$ and $\eta = (v_2, \omega_2)$ both in se(3). The adjoint representation also has a dual, which maps the action of the Lie algebra onto the dual Lie algebra; the matrix representation for this action is

$$\mathrm{ad}_\xi^*(\cdot) = \begin{bmatrix} -\hat{\omega} & 0 \\ -\hat{v} & -\hat{\omega} \end{bmatrix}.$$

2) Newton-Euler Dynamics: There are a number of different computational methods which use a Newton-Euler framework to efficiently calculate both numerical as well as analytic forward dynamic expressions for serial-chain mechanisms. We use the version found in [13]. Newton-Euler formulations consist of a two-step iterative process. In the first iteration the velocities and accelerations of each link are recursively propagated from the mechanism's base to its end effector. In the second iteration, the forces and joint torques are recursively propagated back from the end effector to the base. More specifically, for an N-link serial chain, we initialize the iterative process with

$$\xi_1^0 = \dot{\xi}_1^0 = F_{N+1} = 0, \tag{11}$$

where, ξ_j^0 is a *body velocity* and F is a *wrench* [11]. The forward recursion proceeds from $i = 1$ to N such that

$$\begin{aligned} g_{i-1,i} &= \exp(\xi_i \theta_i) g_{i-1,i}(0) \\ \xi_i^0 &= \mathrm{Ad}_{g_{i-1,i}^{-1}} \xi_{i-1}^0 + \xi_i \dot{\theta}_i \\ \dot{\xi}_i^0 &= \xi_i \ddot{\theta} + \mathrm{Ad}_{g_{i-1,i}^{-1}} \dot{\xi}_{i-1}^0 + \mathrm{ad}_{\mathrm{Ad}_{g_{i-1,i}^{-1}} \xi_{i-1}^0} \xi_i \dot{\theta}_i. \end{aligned} \tag{12}$$

Following the forward recursion, the backward recursion proceeds from $i = N$ to 1,

$$\begin{aligned} F_i &= \mathrm{Ad}_{g_{i,i+1}^{-1}}^* F_{i+1} + J_i \dot{\xi}_i^0 - \mathrm{ad}_{\xi_i^0}^* J_i \xi_i^0 \\ \tau_i &= \xi_i^T F_i. \end{aligned} \tag{13}$$

Following the results in [12], and the references therein, the expressions in (11), (12), and (13) can be written as a large algebraic system and rearranged using block-row elimination as

$$\begin{bmatrix} M & A^{-T} & 0 \\ A^{-1} & 0 & \xi \\ 0 & 0 & \mathcal{M} \end{bmatrix} \begin{bmatrix} -\dot{\xi}^0 \\ F \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} b \\ -a \\ \tau - \mathcal{C} \end{bmatrix}, \tag{14}$$

where, $\xi^0 = [\xi_1^0 \ \xi_2^0 \ldots \xi_N^0]^T$, $\xi = [\xi_1 \ \xi_2 \ldots \xi_N]^T$, and due to our initialization in (11), $\mathcal{M}$ and $\mathcal{C}$ in the last block row exactly correspond to the generalized inertia and Coriolis and centripetal matrices respectively. Both a and b in (14) are block expressions which can be derived from (12) and (13)

$$a = \begin{bmatrix} \mathrm{ad}_{\xi_1^0}^* J_1 \xi_1^0 \\ \mathrm{ad}_{\xi_2^0}^* J_2 \xi_2^0 \\ \vdots \\ \mathrm{ad}_{\xi_n^0}^* J_n \xi_n^0 \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ \mathrm{ad}_{\mathrm{Ad}_{g_2}^{-1} \xi_1^0} \xi_2 \dot{\theta}_2 \\ \vdots \\ \mathrm{ad}_{\mathrm{Ad}_{g_{n-1,n}}^{-1} \xi_n^0} \xi_n \dot{\theta}_n \end{bmatrix} \tag{15}$$

The expression in (14) leads to the following matrix decompositions for $\mathcal{M}$ and $\mathcal{C}$ [12]

$$\begin{aligned} \mathcal{M} &= \mathrm{di}^6(\xi)^T A^T M A \ \mathrm{di}^6(\xi) \\ \mathcal{C} &= \mathrm{di}^6(\xi)^T A^T (MAa + b), \end{aligned} \tag{16}$$

where, $\mathrm{di}^6 : \mathbb{R}^{6N \times 1} \to \mathbb{R}^{6N \times N}$ is an operator which takes individual ξ_i vectors in and places them on the block diagonal of a $6N \times N$ matrix. Note that much of the related dynamics literature is concerned with manipulating the expressions in (14) to derive numerically efficient algorithms. We are not necessarily interested in directly pursuing

computational efficiency, in this sense, in this work. We are however interested in using the the matrix decompositions in (16) to facilitate the computations of the linearization of the dynamics in (2).

C. Linearization

1) Derivatives of $\mathcal{M}$ and $\mathcal{C}$: The linearization of (2) is found by taking the derivatives of $f(\cdot)$ and $h(\cdot)$ with respect to the extended state vector $x = (\dot{\theta}, \theta, \mathcal{S})$. Using the matrix expressions for $\mathcal{M}$ and $\mathcal{C}$ in (16) it is possible to express these derivative concisely and in closed form. We note that the authors in [7] provide results which linearize the second-order dynamics of serial-chain mechanisms. Their results are ultimately equivalent to those included in this section, except we derive simple closed-form solutions which are intuitively expressed using the geometric notion found in [13] and [11]. Our results make it much easier to, for example, compute the higher-order derivatives of the dynamics.

It is helpful to first define several matrix relationships and operators which appear throughout the linearization calculations. We start by defining the operator $\mathrm{di}^6_{\mathrm{ad}} : \mathbb{R}^{6N\times 1} \to \mathbb{R}^{6N\times 6N}$ s.t.,

$$\mathrm{di}^6_{\mathrm{ad}} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \mathrm{ad}_{\xi_1}(\cdot) & 0 & \dots & 0 \\ 0 & \mathrm{ad}_{\xi_2}(\cdot) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathrm{ad}_{\xi_n}(\cdot) \end{bmatrix}. \tag{17}$$

The dual to (17), $\mathrm{di}^6_{\mathrm{ad}^*}$, has the same diagonal form but replaces ad_ξ with ad^*_ξ. The body velocities ξ^0_i can be written in a single stacked vector using the matrix relationship

$$\begin{bmatrix} \xi^0_1 \\ \xi^0_2 \\ \vdots \\ \vdots \\ \xi^0_n \end{bmatrix} = \begin{bmatrix} \mathrm{I} & 0 & \dots & 0 \\ \mathrm{Ad}_{g_2^{-1}} & \mathrm{I} & \dots & 0 \\ \mathrm{Ad}_{g_{2,3}^{-1}} & \mathrm{Ad}_{g_3^{-1}} & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Ad}_{g_{2,n}^{-1}} & \mathrm{Ad}_{g_{3,n}^{-1}} & \dots & \mathrm{I} \end{bmatrix} \begin{bmatrix} \xi_1\dot{\theta}_1 \\ \xi_2\dot{\theta}_2 \\ \vdots \\ \vdots \\ \xi_n\dot{\theta}_n \end{bmatrix}. \tag{18}$$

Note that (18) can concisely be written $\xi^0 = A\psi$, where $\psi = [\xi_1\dot{\theta}_1\ \xi_2\dot{\theta}_2 \dots \xi_n\dot{\theta}_N]^T$. We also define the block diagonal matrix $\mathcal{I}^j$ whose entries are all zero except for a 6×6 identity matrix in the j-th block of the main diagonal; $\mathcal{I}^j$ serves as a block selection matrix.

The expressions in (17) and (18) can be used to rewrite a and b in (15) as

$$\begin{aligned} a &= \mathrm{di}^6_{\mathrm{ad}^*}(A\psi)MA\psi \\ b &= \mathrm{di}^6_{\mathrm{ad}}((A-\mathrm{I})A\psi)\psi. \end{aligned} \tag{19}$$

Equation 19 allows $\mathcal{M}$ and $\mathcal{C}$ in (16) to be written in terms of ψ, A, and M only. Further, A is a function of θ only, ψ is a function of $\dot{\theta}$ only, and M is a constant matrix. Thus, once the derivatives $\partial A/\partial\theta$ and $\partial\psi/\partial\dot{\theta}$ are determined, the derivatives $\partial\mathcal{M}/\partial\theta$, $\partial\mathcal{M}/\partial\dot{\theta}$, $\partial\mathcal{C}/\partial\theta$, and $\partial\mathcal{C}/\partial\dot{\theta}$ are trivially calculated using the chain rule.

The derivatives $\partial A/\partial\theta$ and $\partial\psi/\partial\dot{\theta}$ can be expressed as tensors of dimension $\mathbb{R}^{6n\times 6n\times n}$ and $\mathbb{R}^{6n\times 1\times n}$ respectively. Calculating $\partial\psi/\partial\dot{\theta}$ is trivial, as each of the n different $6n\times 1$ components has the form

$$\frac{\partial\psi}{\partial\dot{\theta}_i} = [0 \dots \xi_i \dots 0]^T \tag{20}$$

Calculating $\partial A/\partial\theta$ is facilitated by using the following expression [13]

$$\frac{\partial}{\partial\theta_k}\mathrm{Ad}^j_i(\xi_{i-1}) = \begin{cases} \mathrm{Ad}^j_k([\mathrm{Ad}^{k-1}_i\xi_{i-1}, \xi_k]) & i \le k \le j \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Recalling the relationship between the adjoint representation and Lie bracket (and properties of the Lie bracket), the top term on the right hand side of (21) can be rewritten

$$\mathrm{Ad}^j_k([\mathrm{Ad}^{k-1}_i\xi_{i-1}, \xi_k]) = -\mathrm{Ad}^j_k\left(\mathrm{ad}_{\xi_k}(\mathrm{Ad}^{k-1}_i\xi_{i-1})\right). \tag{22}$$

From the expression in (22), the derivative of the adjoint representation can be defined in terms of the operator $\partial\mathrm{Ad}^j_i(\cdot)/\partial\theta_k = -\mathrm{Ad}^j_k\left(\mathrm{ad}_{\xi_k}(\mathrm{Ad}^{k-1}_i)\right)$. This expression appears to be relatively complicated, but it allows each of the $6n\times 6n$ terms $\partial A/\partial\theta_i$ to be written concisely as

$$\partial A/\partial\theta_i = A\mathcal{I}^i\mathrm{di}^6_{\mathrm{ad}}(\xi)(I-A). \tag{23}$$

As stated above, the expressions in (20) and (23) can be used along with the chain rule to derive *explicit* closed-form expressions for the directional derivatives of $\mathcal{M}$ and $\mathcal{C}$.

2) Linearization of elastic joint mechanisms : We assume that the second-order dynamics of a serial chain manipulator with N elastic joints has the form [18]

$$\ddot{\theta}_\lambda = \mathcal{M}^{-1}(\tau_S - \mathcal{C} + \mathcal{G}) \tag{24}$$

$$\ddot{\theta}_\mu = 1/J_\mu(u - \tau_S), \tag{25}$$

$$h(\dot{\theta}, \theta) = \theta, \tag{26}$$

where $\theta = (\theta_\lambda, \theta_\mu)$ and $\dot{\theta} = (\dot{\theta}_\lambda, \dot{\theta}_\mu)$ contain both the load, θ_λ, and motor, θ_μ, angles and velocities respectively. The terms $\mathcal{G}$, τ_S, and u correspond to the dynamic contributions due to gravity, the elastic elements located between joints, and the motor control torques. A linear model with damping is used to model the spring torque, i.e., for each of the N joints $\tau^i_S = k_i(\theta^i_\mu - \theta^i_\lambda) + b_i(\dot{\theta}^i_\mu - \dot{\theta}^i_\lambda)$. The $2N$ second-order terms in (24) and (25) can be written along with the $2N$ first-order link velocities and $2N$ parameter velocities as a $6N$-dimensional first-order system in the form (2). For the purpose of computing the linearization we will treat (24) and (25) independently, then merge these results back into the full state linearization. The linearization of the $4N$ first-order dynamic terms is trivial to compute.

The derivatives of $\ddot{\theta}_\lambda$ and $\ddot{\theta}_\mu$ are defined in terms of the derivatives of $\mathcal{M}$ and $\mathcal{C}$ defined in Section III-C.1 well as the derivatives of τ_S and $\mathcal{G}$ with respect to the full state vector x. The linear spring model makes computing the derivatives of τ_S and subsequently $\ddot{\theta}_\mu$ trivial. These derivatives are not included here due to space considerations. To compute the derivatives of $\mathcal{G}$, we first leverage the structure in (8) and (9),

$$\mathcal{G}^j = \sum_{k=j}^{N} m_j \mathrm{g}\frac{\partial}{\partial\theta_j} g^h_{1,k}(\theta), \tag{27}$$

where m_i is the mass of link i, g is the magnitude of gravity, and the superscript h denotes the position element of $g_{1,j}(\theta)$ corresponding to the height above an inertial frame with zero potential due to gravity. The derivatives of (27) with respect to θ_i, for $j \leq i$, are then defined to be

$$\frac{\partial}{\partial \theta_i}\mathcal{G}^j = \sum_{k=i}^{N} m_j \mathrm{g} \frac{\partial}{\partial \theta_i}\frac{\partial}{\partial \theta_j} g_{1,k}^h(\theta). \tag{28}$$

Note that (28) is only valid for $j \leq i$, but that the derivatives are symmetric, i.e., $\frac{\partial}{\partial \theta_i}\mathcal{G}^j = \frac{\partial}{\partial \theta_j}\mathcal{G}^i$.

One final point is that the inverse of $\mathcal{M}$ appears in (24), where Section III-C.1 only discuss computing the linearization of $\mathcal{M}$. This can be resolved using the simple matrix relationship $\frac{\partial}{\partial x}\mathcal{M}^{-1} = -\mathcal{M}^{-1}\frac{\partial}{\partial x}(\mathcal{M})\mathcal{M}^{-1}$. It is well known that computing the inverse of $\mathcal{M}$ is an $\mathcal{O}(N^3)$ operation, and thus special consideration needs to be applied as the number of joints N gets large. Several methods which make it possible to efficiently invert $\mathcal{M}$ do exist, but for the purpose of this paper are left to the references [14], [16].

IV. Calculation of Nonlinear Observability Rank Condition

Calculating the mapping Φ in (3) can potentially be very difficult to obtain in closed-form. For the serial elastic-joint system described in (24)-(26), the linear expression for the motor-side dynamics in (25) along with the explicit linearizations of θ_μ and θ_λ can be exploited to address this difficulty. We begin our derivation of Φ by defining

$$\begin{aligned}\Phi = \big[&{}_\lambda\phi_1^0 \; {}_\lambda\phi_1^1 \cdots {}_\lambda\phi_N^0 \; {}_\lambda\phi_N^1 \\ &{}_\mu\phi_1^0 \; {}_\mu\phi_1^1 \; {}_\mu\phi_1^2 \; {}_\mu\phi_1^3 \cdots {}_\mu\phi_N^0 \; {}_\mu\phi_N^1 \; {}_\mu\phi_N^2 \; {}_\mu\phi_N^3\big]^T,\end{aligned} \tag{29}$$

where ${}_\lambda\phi_j^0 = \theta_\lambda^j$ and ${}_\mu\phi_j^0 = \theta_\mu^j$. Notice that (29) contains only zeroth and first-order Lie derivatives with respect to the load-side position measurements but has up to third-order derivatives with respect to the motor position measurements. As mentioned in Section III-A this choice of Φ is not unique. This specific form was chosen to take advantage of the relatively simple expressions for $\ddot{\theta}_\mu$. Notice also that most of the components in (29) are trivial to compute, i.e., ${}_{\{\mu,\lambda\}}\phi_i^0 = \theta_{\{\mu,\lambda\}}^i$, ${}_{\{\mu,\lambda\}}\phi_i^1 = \dot{\theta}_{\{\mu,\lambda\}}^i$, and ${}_\mu\phi_i^2 = \ddot{\theta}_\mu^i$. Each of these terms are directly measured or are able to be estimated using the framework described in Section V.

The only expressions in (29) which are not measured or able to be estimated directly are the ${}_\mu\phi_i^3$'s. To calculate these expressions, consider first a single joint system where $x = (\dot{\theta}_\lambda, \dot{\theta}_\mu, \theta_\lambda, \theta_\mu, k_S, b_S)$ and the second-order load and motor dynamics have the forms (24) and (25) respectively. We make the assumption that the motor controller has the form $u = k_P(\theta_\mu^d - \theta_\mu)$, where $k_P \in \mathbb{R}^+$. For this system,

$${}_\mu\phi^3 = D_x\ddot{\theta}_\mu \circ \dot{x} + \dot{u}, \tag{30}$$

where $D_x\ddot{\theta}_\mu \circ \dot{x} = b_S(\ddot{\theta}_\lambda - \ddot{\theta}_\mu) + k_S(\dot{\theta}_\lambda - \dot{\theta}_\mu)$ and $\dot{u} = k_P(\dot{\theta}_\mu^d - \dot{\theta}_\mu)$. The expression in (30) is a sum of components that appear in the extended state dynamics $\dot{x} = f(x, u)$, i.e., (30) contains only terms which can be calculated in closed form using the results of Sections III-C.1 and III-C.2.

The results for this one-dimensional example are directly applicable to systems with more degrees of freedom as long as the load and motor side dynamics have the forms (24) and (25) respectively. This is due to the fact that each link's motor dynamics are decoupled from the rest of the mechanism up to the third-order Lie derivative; ${}_\mu\phi_i^3$ for each link will thus have the exact form as (30).

Using (30), the nonlinear observability rank condition (4) can be computed in closed form using only the results in Sections III-C.1 and III-C.2. In particular, the highest-order terms that appear in Φ are the $\dot{\theta}_\lambda$'s and $\ddot{\theta}_\mu$'s. In computing (4), we need only the closed-form linearizations of these terms to explicitly calculate $\partial\Phi/\partial x$. The calculations necessary to obtain these linearizations are presented in Section III-C.2. Each of the lower-order terms that appear in (4) can directly be measured, estimated, or calculated from the system's forward dynamics (2).

V. Parameter Estimation and Online Control

The main novel contribution of this work is that we use the results from the previous sections of this paper to obtain a closed-form expression for the nonlinear observability rank condition which is then uses to derive a local measure of observability. We subsequently optimize this observability measure at each time step in an online controller to enhance the performance of a recursive filter. The objective of this approach is to make the estimation of a set of unknown spring parameters in a serial-chain elastic-joint mechanism converge as fast as possible to their true values. We accomplish by simultaneously running an EKF to estimate the extended state vector x, which includes the unknown spring parameters k and b, while locally performing gradient descent on a measure of the local observability. There are at least two different *measures of observability* found in the literature [8]. Following [5], we use the ratio of the largest to the least singular value of an observability rank condition matrix; this has previously been referred to as the *local estimation condition number*. This choice was motivated by our goal to estimate all of the the unknown spring parameters simultaneously.

To maximize the observability of the spring parameters, we minimize the *nonlinear estimation condition number* of the matrix block in (4) that corresponds to the spring parameters; for our example system this is the bottom $2N \times 2N$ block on the main diagonal of $\partial\Phi/\partial x$. This local optimization has the effect of forcing the minimum and maximum singular values closer together, i.e., it forces the observability to be more uniform over all the parameters. This objective also forces the minimum singular value away from zero, and thus does implicitly *maximize* observability (using the proximity of the minimum singular value to zero as an objective).

For our simulated results, we make the same assumption on u as we did in (30), i.e., $u = k_P(\theta_\mu^d - \theta_\mu)$, except that here we assume the reference trajectory θ_μ^d is not fixed. We locally minimize the estimation condition number at each time step by first computing its gradient with respect to the motor-side

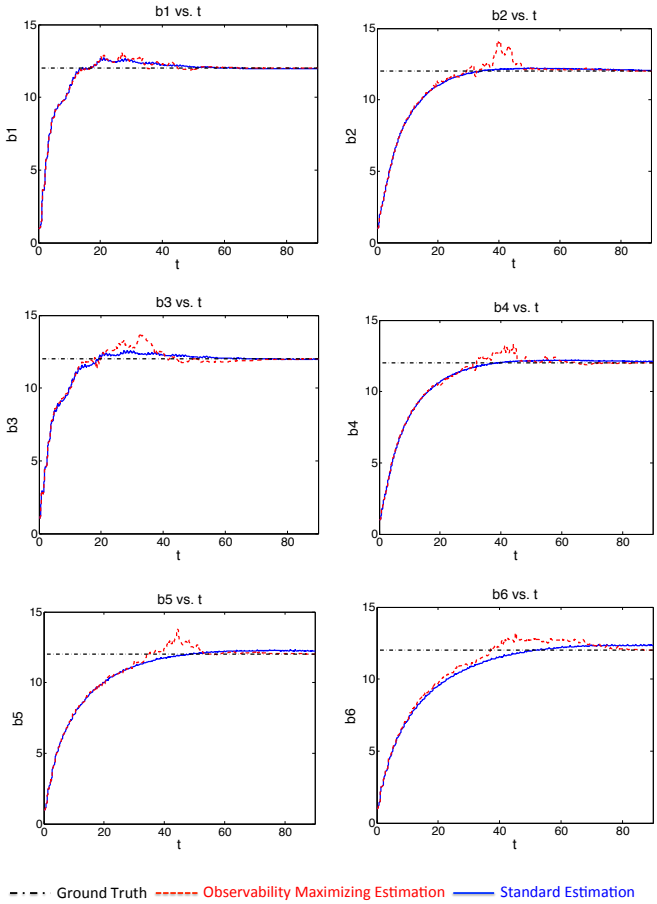

Fig. 2: Damping coefficient (b_S) estimation results for for a six-link series-elastic actuated serial-chain manipulator.

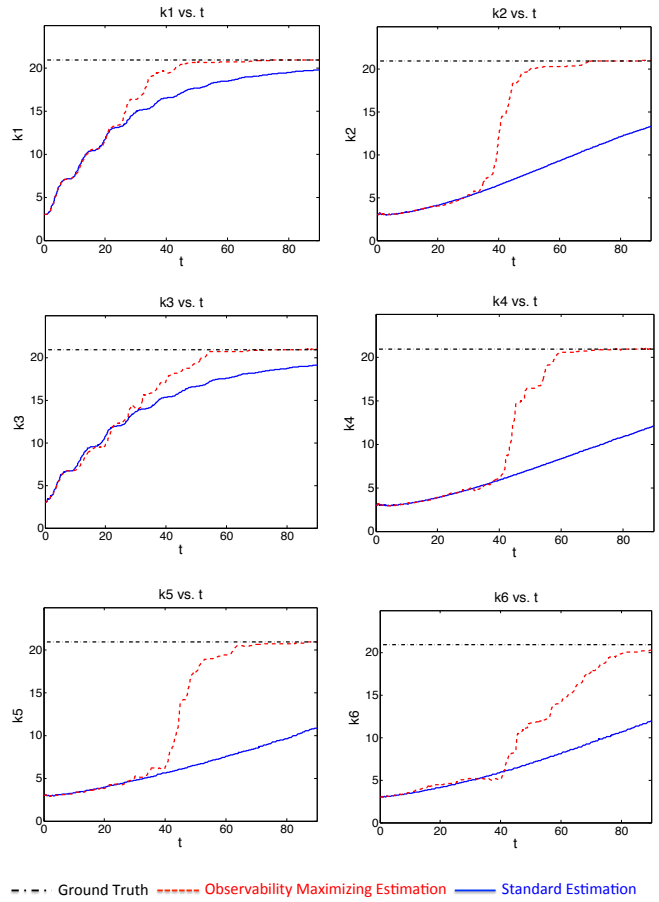

Fig. 3: Spring constant (k_S) estimation results for for a six-link series-elastic actuated serial-chain manipulator.

positions θ_μ. We then use this value to update the motor reference trajectory online by taking a single fixed-length step in the direction opposite that of the condition number gradient, where the gradient is numerically calculated. The step length is set conservatively to ensure descent. More clearly, defining $\sigma_{\mathcal{S}}$ as the local estimation number associated with the spring parameters, $\bar{\theta}^d_\mu$ as the updated motor reference trajectory, and $\alpha \in \mathbb{R}^+$ as the step length,

$$\bar{\theta}^d_\mu = \theta^d_\mu + \alpha \nabla_{\theta_\mu} \sigma_{\mathcal{S}}. \tag{31}$$

The updated motor control law thus becomes

$$u = k_P(\bar{\theta}^d_\mu - \theta_\mu). \tag{32}$$

VI. Results

We demonstrate the use of the observability maximization approach discussed in Section V on a full-dynamic simulation of a six-link series-elastic actuated mechanism. Each link is configured such that it is rotated $\pi/2$ radians relative to the previous link. The link and motor side second-order dynamics are assumed to have the forms (24) and (25) respectively. The link masses and moments are all normalized to one; the motor inertia is also assumed to be one. The physical values of the spring constants and damping coefficients are uniformly set to $k = 21$ and $b = 12$ respectively (represented by the black dotted lines in Figures 2 and 3). The proportional gains in (32) are set to $k_P = 3$, the unmodified reference signal is $\theta^d_\mu = 1.3\sin(6t)$, and the parameter α in (31) is set to $\alpha = 0.1$.

The code for the example in this section was written in C++, using *Eigen*, and interfaced with *Matlab* via *mex* files. The average time to compute the full linearization for the six-link thirty-six dimensional system was on average 0.41×10^{-4} seconds. The average time to compute the singular value decomposition of the nonlinear observability matrix rank condition was 1.3×10^{-3} seconds. We did not use any compile-time optimization in computing these values.

The variance terms that appear in the EKF are set to $\sigma_{\dot{\theta}} = \sigma_\theta = \sigma_z = 0.001$, where $\sigma_{\dot{\theta}}$ corresponds to the variance in the angle velocities, σ_θ the variance in the link angles, and σ_z the variance associated with the measurements. The variance values of the spring constants and damping coefficients are assumed to be $\sigma_{\mathcal{S}} = 0.1$. The process and measurement covariance matrices are both assumed to be diagonal and the initial covariance associated with the state update is set to $P_0 = 0.001I$, where I is a 36×36 identity matrix. The initial estimated values of the spring constants and damping coefficients are set to $k = 3$ and $b = 1$ respectively.

The parameter-value estimate plots are shown in Figures 2 and 3. Figure 2 shows the damping coefficient results

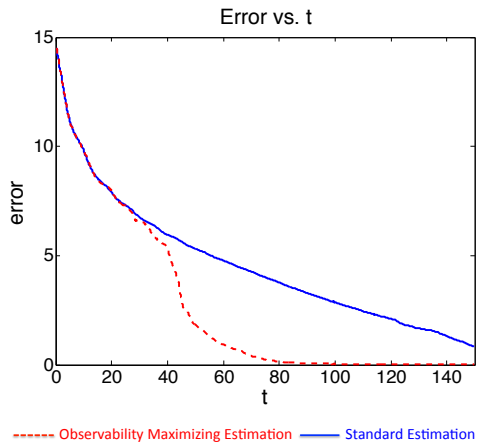

Fig. 4: Total (non-dimensional) average error over all spring parameters, i.e., spring constants and damping coefficients.

and Figure 3 the spring constant estimates. Each of the individual plots in these figures show the estimated parameter values versus time; the ground truth values for each parameter are shown in as black dotted lines. The links are numbered from proximal to distal with respect to the base in ascending order. The plots also show a comparison between the EKF used with (dashed red line) and without (solid blue line) the modified reference trajectory. Figure 3 shows a dramatic improvement in terms of the time taken to converge when using the observability maximizing criteria. Figure 2 shows comparable performance between the two approaches. In fact, the observability-modified convergence results in Figure 2 appear to have slightly worse transient response than the standard filtering approach. The difference in benefits between the results shown in Figures 2 and 3 can intuitively be understood by considering how we measure observability. The fact that we use the estimation condition number has the effect of normalizing the local observability over *all* the parameters simultaneously. Thus some of the parameters which are originally very observable, i.e, have high associated singular values in the observability rank condition, become slightly less observable to allow other parameters, which have low associated singular values, to become much more observable. The overall improvement in performance over the entire set of parameters is quite dramatic, as can be seen in Figure 4, which plots the total non-dimensional average error over all of the spring parameters versus time; the red-dashed line and solid blue line correspond to the observability optimized and standard estimation schemes respectively.

VII. Conclusion

The benefit of the work presented in this paper is shown in Figure 4. We show that using the novel online calculation of the nonlinear estimation condition number, made possible by the exact closed-form forward dynamics and linearization of the series-elastic actuated system, a dramatic improvement in terms of the time it takes to converge to the true values of the unknown spring parameters can clearly be seen.

There are however several directions of future work that we believe can further improve our results. The first is to include a dynamic step sizing algorithm in the calculation of α in (31). Our results show that using a fixed value of α works reasonably well, but taking larger steps along the gradient of the estimation condition number initially and smaller steps eventually can potentially lead to faster rising and better settling times in the parameter estimation signals. Additionally, in this work there is an inherent tradeoff between the control and estimation; the fact that we change the reference signal in the observability-modified estimation scheme makes the comparison between the modified and standard filtering results somewhat unfair. In the future we plan to explore how the non-uniqueness of workspace control solutions for redundant systems can be taken advantage of such that the overall control objective remains unmodified while the observability is simultaneously maximized.

References

[1] A. M. Bloch, J. Baillieul, P. E. Crouch, and J. E. Marsden. *Nonholonomic Mechanics and Control*. Spinger, 2003.

[2] R. W. Brockett. Robotics manipulators and the product of exponentials formula. *Proc. Int. Syp. Math. Theory of Networks and Systems*, 1990.

[3] D. Del Vecchio and R.M. Murray. Observability and local observer construction for unknown parameters in linearly and nonlinearly parameterized systems. *Proceedings of the 2003 American Control Conference, 2003.*, 6(1):4748–4753, 2003.

[4] R. Hermann and A J. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22:728–740, October 1977.

[5] B T. Hinson, M K. Binder, and K A. Morgansen. Path planning to optimize observability in a planar uniform flow field. *2013 American Control Conference*, (1):1392–1399, June 2013.

[6] A. Isidori. *Nonlinear Control Systems*. Springer, London, 1995.

[7] A. Jain and G. Rodriguez. Recursive linearization of manipulator dynamics models. *IEEE International Conference on Systems, Man, and Cybernetics*, pages 475–480, November 1990.

[8] A J. Krener and K Ide. Measures of unobservability. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 6401–6406, December 2009.

[9] S. Lall, J E. Marsden, and S. Glavski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International Journal of Robust Nonlinear Control*, 12:519–535, 2002.

[10] L. Ljung. Asymptotic behavior of the extended Kalman filter as parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1):36–50, February 1979.

[11] R. M. Murray, Z. Li, and S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1994.

[12] D K. Pai and U M. Ascher. Forward dynamics algorithms for multibody chains and contact. *IEEE International Conference on Robotics and Automation*, pages 857–863, April 2000.

[13] F. C. Park, J. E. Bobrow, and S. R. Ploen. A Lie group formulation of robot dynamics. *The International Journal of Robotics Research*, 14:609–618, December 1995.

[14] G. Rodriguez and K. Keutz-Delgado. A spatial operator factorization and inversion of the manipulator mass matrix. *IEEE Transactions on Robotics and Automation*, 1992.

[15] D. Rollinson and H. Choset. Design and Architecture of a Series Elastic Snake Robot. *IEEE International Conference on Intelligent Robots and Systems*, 2014.

[16] S.K. Saha. A decomposition of the manipulator inertia matrix. *IEEE Transactions on Robotics and Automation*, 13(2):301–304, April 1997.

[17] E D. Sontag. *Mathematical Control Theory*. Springer, New York, 1998.

[18] M W. Spong. Modeling and control of elastic joint robots. *Journal of Dynamic Systems Measurement and Control*, 24, 1979.