Data-Based Modeling and Control of Dynamical Systems: Parameter Estimation

Damien Gueho[1], Manoranjan Majji[2], and Puneet Singla[3]

***Abstract*— Parameter estimation methods to provide data-based models to control complex dynamical systems are reviewed. Starting from least square minimization of the equation error, the tutorial provides an overview of how different perspectives of parameter estimation lead to various algorithms that are used in diverse contexts. Both statistical and deterministic approaches are discussed, and the utility of model inferences are explained. The discussions provide a context and review relevant background with respect to three application papers involving recent advances in Gaussian Process Regression (GPR), state estimation approaches and data-driven modeling.**

I. Introduction

Advances in materials science and manufacturing fuel the explosive growth of next generation autonomous systems. Finding fundamental ways to integrate these innovations with appropriate sensing, actuation and embedded computing is necessary to engineer effective systems that meet ever increasing performance goals. Emerging frontiers in robotics and autonomous systems (e.g., micro air vehicles, humanoid robots, and soft robotics) represent the technology pushes [19], [36], [39], [52], [56], [57] to integrate multiphysical models, while the commercial pulls associated with applications such as self driving cars, autonomous aerial vehicles, humanoid and personal robotic platforms typify the path breaking societal needs for this integration [11], [45], [58], [62]. Mathematical modeling of physical systems is at the core of effective integration of autonomous systems. Actuators, sensors and the dynamical processes governing the subsystem components need to be represented mathematically such that the signals and system managing the interfaces between the subsystems that constitute the autonomous operations are engineered in a compatible fashion. Most often, engineering approximations are made in simplifying the infinitely complex component technology models. After integration however, appropriate calibration and parameter estimation of the simplified model is essential to assign the correct values of the model parameters in order to make it relevant and representative for its effective integration into other subsystems.

The aim of this tutorial paper in parameter estimation methods is to summarize various approaches to estimate parameters that are most often applied in engineering practice and discuss their broader utility in modeling and control of dynamical systems. In particular, the parameter estimation approaches are broadly organized into direct approaches for algebraic and dynamical systems and indirect approaches that use other approaches like state estimation to estimate parameters. The tutorial session has dedicated papers associated with each of these topics. This paper aims at providing a historical context to the three papers that focus on application of parameter estimation approaches of various types to estimate parameters in complex dynamical system applications. The first companion paper to this tutorial involves an application of a Gaussian Process Regresion (GPR) to model reduction of hypersonic dynamical systems for control applications by Huang [20]. A tutorial on the use of recent advances in Data-Driven modeling system called SINDy is presented by Fasel et al [14]. An extended Kalman filter is applied to estimate an unknown input associated with the plasma dynamical system is shown in the third paper [17] to demonstrate the application of state estimation approaches. Quantitative measures of sensitivity of the states

II. Parameter Estimation Methods: Direct Approaches

Linear least squares approach is among the most popular approaches for parameter estimation, and forms the cornerstone of data-based modeling [10], [48]. At the outset, the analyst is given a set of m measurements $\mathbf{y} \in \mathbb{R}^{m\times 1}$ that are related to the n unknown parameters $\boldsymbol{\theta} \in \mathbb{R}^{n\times 1}$ using a linear model given in Eq. 1.

$$\mathbf{y} = \mathbf{H}(\omega)\boldsymbol{\theta} + \boldsymbol{\nu} \tag{1}$$

where ω is a function of the state dynamics or other independent variables such as time or frequency. Frequently, number of measurements exceeds the number of parameters and an exact solution is not possible, therefore, following Gauss, the analyst provides a best estimates of these parameters by minimizing the sum squares of the equation errors of the m conditions specified by the measurements provided (primarily founded on the underlying model of Eq. 1) facilitates a closed form solution to the parameter estimation problem, written as

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}, \tag{2}$$

where the matrix $\mathbf{R}$ is a weight matrix associated with the measurements involved. While the equations of 1 can be solved by minimizing other measures (sum of absolute values or the minimization of the equations with the largest residual error), least squares solution is the most prevalent approach,

[1]PhD Candidate, Department of Aerospace Engineering, Penn State University, University Park, PA 77840. Email: *djg79@psu.edu*

[2]Assistant Professor, Department of Aerospace Engineering, Texas A&M University, College Station, TX 77840. Email: *mmajji@tamu.edu*

[3]Professor, Department of Aerospace Engineering, Penn State University, University Park, PA, 77840. Email: *psingla@psu.edu*

owing to the fact that a compact formula of Eq. 2 provides a solution to the system of equations. The minimization of sum of absolute values of the equation error denoted by the one norm, $\|\mathbf{y}-\mathbf{H}(\omega)\boldsymbol{\theta}\|_1$ or the minimization of the worst case measurements, denoted by the infinity norm of the equation error, $\|\mathbf{y}-\mathbf{H}\boldsymbol{\theta}\|_\infty$ are equally well posed. Both of these problems are convex in nature, and therefore unique solutions are guaranteed for each of these problems [7], [35]. In fact, any p norm can be considered. However, it can be shown that the parameter estimation in each of these cases leads to a mathematical linear program. Since the solution to such a problem involves computations, they are not very popular. However, literature shows that these solutions are robust in nature to certain types of uncertainties associated with the model error. That is to say when the parameter estimation error $\boldsymbol{\nu}$ is non-Gaussian, the likelihood functional formed using the characteristics of the probability density function (PDF) of the equation error $\boldsymbol{\nu}$, denoted by $p(\boldsymbol{\nu})$ will provide a better approach to solve for the parameter $\hat{\boldsymbol{\theta}}$ [10], [49], [50]. The fundamental principles of Hilbert space approximations involved in solving this system of equations various vector spaces form an integral part of engineering systems analysis [35]. Recently, these approaches are forming an important element of solution of high dimensional partial differential equations [1], [2]. Since the cost functionals are convex, the deterministic approaches that minimize an appropriate measure of the equation used gradient descent type approaches early on in systems and control [4], [5].

Earliest use of parameter estimation in control of practical dynamical systems is seen in the development of regulators using the frequency response functions obtained from empirical data [42]. Earliest input-output approach for compensator design involves the use of parameter estimation methods to *identify* a transfer function between the input $u(t)$ to a dynamical system and the output signals sensed by the sensors measuring $y(t)$. A dynamical relationship between the outputs and the inputs can be established in the Laplace domain by looking at the frequency response function that examines the ratio of the output signal to the input. In a physical problem, the frequency response function, $g(s)$ is typically determined using the ratio of correlation functions, written as

$$\frac{y(s)}{u(s)} = g(s) \approx \frac{\mathcal{F}[y(t)]\mathcal{F}[u(t)]*}{\mathcal{F}[u(t)]\mathcal{F}[u(t)]*}, \tag{3}$$

where s is the Laplace variable, and $\mathcal{F}(.)$ denotes the Fourier transform of a signal. Also known as the transfer function, an estimate for the frequency response function is facilitated better in discrete time using an application of the Discrete Fourier Transform (DFT) on sampled empirical data [43], [51]. By utilizing the approximation $z \approx e^{sdT}$ where dT is the sample time, the discrete time frequency response function is a rational polynomial given by

$$g_d(z) = \frac{n(z,\mathbf{p})}{d(z,\mathbf{p})}, \tag{4}$$

where $n(z,\mathbf{p})$ and $d(z,\mathbf{p})$ are the numerator and denominator polynomials respectively. The fact that the coefficients defining these parameters appear linearly in the transfer function description of Eq. 4, makes the dynamic modeling problem a parameter estimation problem. From spectral response data, typically the measurements of $g_d z_i$ are available at certain choices of the sampling frequency $z_i = e^{j\omega_i dT}$, with $j^2 = -1$ from the correlation functions (or the composite Frequency Response Function) of Eq. 3. Then the parameter estimation problem for data-based modeling is written as

$$\min_{\mathbf{p}} J = \|g_d(z_i) - \frac{n(z_i,\mathbf{p})}{d(z,\mathbf{p})}\|_\omega \tag{5}$$

where $\|\|_\omega$ is a norm chosen to represent the weights assigned for participating frequencies. Frequently, the minimization in Eq. 5 is a non-convex program, so engineering applications minimize an equivalent function $\|g_d(z_i)d(z,\mathbf{p}) - n(z_i,\mathbf{p})\|_\omega$ such that a linear least squares solution can be used to estimate the model parameters [24], [34]. However, this practical solution causes some problems, as it interferes with the weights involved in the norm used for parameter estimation. In frequency domain, weighting is an important part of the modeling process [48]. This is because of the fact that the spectral weighting frequently determines the degree to which the approximated transfer function captures the physics important for modeling. It is well known that for control applications, modeling typically only requires a few of the Markov parameters of the system [33], [47], [59]. Therefore, the control design also gets impacted by this choice. Furthermore, since the measured data $g_d(z_i)$ are noisy, the estimated parameters are random variables with a mean and variance [54], [55]. Appropriate modifications of the parameter estimation problem like a total least squares formulation [43], and various information theoretic formulations have been studied in the literature [3], [60], [61].

Among the most interesting recent advances in parameter estimation is the Gaussian Process Regression (GPR) approach for both modeling and classification [44]. GPR regression builds on the principles of Bayesian inference and uses measurement data to build a likelihood functional based upon mean and covariance functions as opposed to constants based on statistics of the measurement data. This key difference accommodates a wide variety of covariance equivalent models that explain the measurements and also for a non-parametric method for statistical parameter estimation. Hyper-parameters associated with the mean and covariance process dictate the performance of the approach. In this tutorial, a GPR approach to facilitate dynamic modeling of complex models used for guidance, navigation and control of hypersonic vehicles is provided in a separate detailed article by Huang [20]. We dedicate the fundamentals of GPR and other non-parametric approaches for modeling dynamical system to the companion paper of this session.

III. ESTIMATION OF DYNAMICAL SYSTEM PARAMETERS: DIRECT APPROACHES

If the objective of modeling extends beyond the needs specified by the control system, parameter estimation approaches need to incorporate more sophisticated approaches to accomplish these objectives. One of the key applications for parameter estimation in dynamical systems is to achieve closure with physics based models [27], [37], [38]. This is particularly true in several physical applications that are non-minimal. Flexible space structures are particular applications where the closure between physics based models and data driven approaches is necessary. Since the dynamical system is essentially uncontrollable, model closure between the physics based design reference model and the empirical data-driven model is essential to design a control system that ensures that no energy is spent on exciting the uncontrollable modes of the structure [27], [28]. It is natural to carryout the model closure activities in time domain using state space models [48]. To this end, mechanical systems have specialized structure in the dynamics that is quite useful to assess the system properties of the matrix second order systems [21]. The physics-based differential equation model for the state $\boldsymbol{\omega}$ can be written for a general dynamical system as

$$\dot{\boldsymbol{\omega}} = \mathbf{g}(\boldsymbol{\omega}, \mathbf{u}, \mathbf{p}) \tag{6}$$

where $\mathbf{u}$ is the control function, $\mathbf{p}$ is the parameter of the dynamical system. In mechanical systems, these parameters can represent the mass, stiffness or damping ratio of the system. Adaptive identification and control has an illustrious past. Identification of parameters of dynamical systems with simple optimization strategies like gradient descent were used in earlier parts of the adaptive identification [4], [6], [15], [16], [23], [30]. Later works studied the imposition of stability constraints on the control system has lead to excellent sources of scholarship [22], [40], [46]. In addition to using physics based models for online parameter identification, the formulations of stable and robust adaptive control laid the foundations for machine learning using arbitrary basis function models in the dynamics [41]. The unique aspect of the adaptive identification control is the infusion with Lyapunov's stability theory [32].

If the parameter vector $\mathbf{p}$ appears linearly in the dynamic model function $\mathbf{g}(\boldsymbol{\omega}, \mathbf{u}, \mathbf{p})$, all the approaches, including the least squares approaches discussed in the previous parameter estimation section are directly applicable. Recent textbooks in adaptive control theory have outlined these formulations very clearly [22], [40], [46]. Various formulations of the parametric models can be used to transform the differential equation into algebraic regression models [41]. Stable filters and realization theory is exploited to realize the regression models in time domain, and excellent results like the Yakubovitch lemmas are used to prove stability of the identification approach in a certainty equivalence formulation of nonlinear systems [16]. Parameter convergence proofs and sufficient conditions for the signals in the regression are also derived in the literature [6]. To demonstrate a key implementation detail in the regression vector, let us consider the dynamics of Eq 6 and assume that the state $\boldsymbol{\omega}$ is available for measurement. Further, let us assume that the parameter vector is linear, that is to say that $\mathbf{g}(\boldsymbol{\omega}, \mathbf{u}, \mathbf{p}) = G(\boldsymbol{\omega}, \mathbf{u})\mathbf{p}$. The transformation of the dynamical system into a linear parametric model for parameter estimation is still not straightforward. To facilitate this, a filter is applied on both sides of Eq. 6. This process leads to an algebraic system of equations $\mathbf{z} - \boldsymbol{\phi}(\boldsymbol{\omega}, \mathbf{u})\mathbf{p}$ that can be minimized at each instant of time. Note that the signal $\mathbf{z}(t)$ realized as $\boldsymbol{\omega} - W(s)\boldsymbol{\psi}$ and $\boldsymbol{\phi}$ is realized as $W(s)G(\boldsymbol{\omega}, \mathbf{u})$, where $W(s)$ is a strictly passive real (SPR) filter that is implemented using realization theory of linear systems [29]. In the most simplistic case when $W(s) = \frac{1}{s+\lambda}$, $\boldsymbol{\psi}$ is the solution of the differential equation $\dot{\boldsymbol{\psi}} = -\lambda\boldsymbol{\psi} + \boldsymbol{\omega}$. Let us now consider an example that demonstrates the use of system theoretical concepts to solve a dynamic parameter estimation problem using parameter estimation techniques developed for algebraic systems.

To consider an interesting parameter estimation process, let us consider a nonlinear dynamical system that is typically used to model multiphysical processes such as the plasma dynamical systems [18]. The same application will be demonstrated using a state estimation approach in a dedicated paper in the tutorial session [17]. This particular example illustrates the fact that the approach for parameter estimation can be used directly to also track slowly varying parameters as identified in related work [53].

One of the key physical models of plasma discharge is provided by the ordinary differential equations governing the number of ion and neutral atoms in a plasma environment, given by n_i and n_n, and the ion bulk velocities denoted by u_i. The global model that provides a simplistic description of the change in the numbers of atoms, as a function of time can be written in the so-called 0-D discharge ordinary differential equation written as

$$\dot{n}_i = -\frac{1}{L} n_n u_i + n_n n_i \zeta(T_e), \tag{7}$$

$$\dot{n}_n = -n_n \frac{(u_n - u_{int})}{L} - n_n n_i \zeta(T_e), \tag{8}$$

where n_i and n_n are the ion and neutral atom number densities, u_i and u_n are the bulk velocity of the ions and neutral atoms, u_{int} is the injection velocity of the neutral atoms from the anode, L is the characteristic length of the ionization region, $\zeta(T_e)$ is the ionization rate coefficient, and T_e is the electron temperature. In physical problems, the electron temperature is an implicit function of time, and acts as an excitation input, affecting the concentration of the ions and neutrals, making the plasma a dynamic environment. Based on the predator - prey model, this model captures the oscillations of the plasma ions that get energized by the ionization source. To this end, $\phi(t)$ (sometimes also denoted by k_{ion}) is the ionization rate coefficient, that is a function of electron temperature T_e, and the u_{int} is the neutral atom density at the inlet (anode), while L denotes the characteristic length of the ionization region.

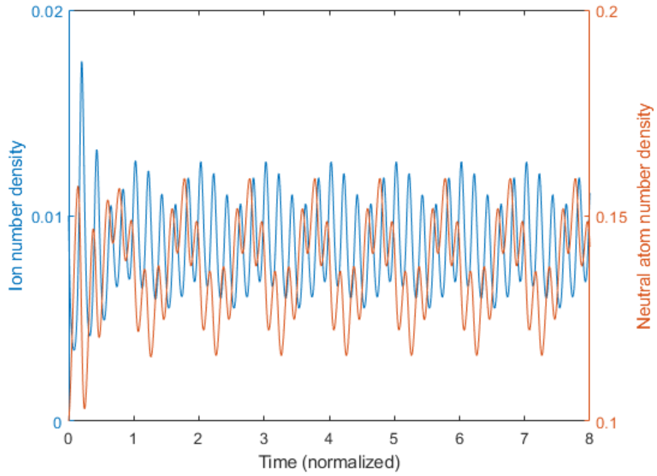

Fig. 1. The nonlinear state time histories associated with the estimation of unknown input.

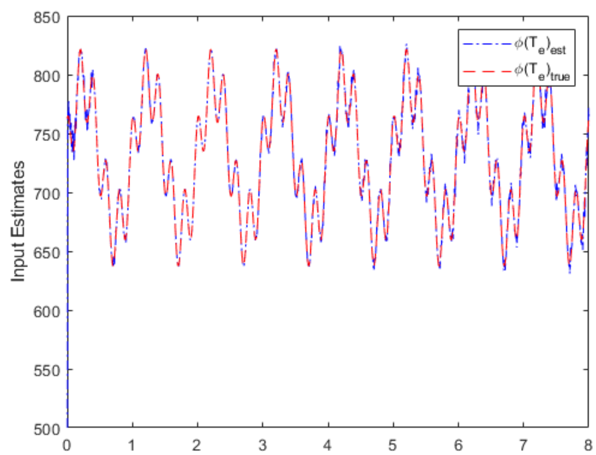

Fig. 2. Parameter convergence to estimate unknown input.

While Fig 1 plots the state time histories, Fig 2 shows the true and estimated input to the system $\zeta(T_e)$. While the theory of parameter estimation assumes that this parameter is constant, appropriate use of instantaneous estimates and fading memory of the filter formulations used in the regression process enable us to estimate the time varying unknown input effectively. This is an interesting example that demonstrates the broad utility of the parameter estimation approaches. Specifically this example illustrates how linear system theory can be used to transform problems of dynamic parameter estimation to problems involving parameter estimation in algebraic systems.

Recent advances in sparse approximation approaches [12], [13] enable extensions of the Hilbert space parameter estimation tools mainly due to the convexity of the formulations involved in the optimization process [7]. Brunton et al [8] and others have worked on building upon these tools for parsimonious modeling of complex physical phenomena from simulation and experimental data. In this tutorial, an implementation of their data-driven modeling tool, SINDy is discussed with illustrative examples [14].

IV. State Estimation Techniques for Parameter Estimation

Yet another approach for parameter estimation is to make use of statistical approaches to estimate the state of dynamical systems [31]. Mostly following the Kalman filter approach, Bucy devised statistical approach to estimate the parameters of a dynamical system by augmenting the state evolution equation with parameter dynamics $\dot{\mathbf{p}} = 0$, as shown in Eq. 10 [9].

$$\dot{\boldsymbol{\omega}} = \mathbf{g}(\boldsymbol{\omega}, \mathbf{u}, \mathbf{p}) \tag{9}$$

$$\dot{\mathbf{p}} = 0 \tag{10}$$

Recent methods for particle filtering and unscented transformation can also be utilized for this process [25], [26]. High order quadrature methods can also be used [2]. The primary difference between deterministic parameter estimation approaches and the use of statistical approaches is that a measure of uncertainty associated with the parameterr estimate is generated by the approaches that use state estimation fundamentals. Since the augmented state dynamics is inherently nonlinear in nature, the statistical version of the parameter estimation problem is fundamentally non Gaussian even when the dynamics of the problem are linear. To identify this more clearly, consider the application of an undamped harmonic oscillator, written as $\ddot{x} = -\frac{k}{m}x + \frac{1}{m}u$. Even if we assume the particle is of unit mass, (i.e., setting $m = 1$) and simplify the problem with the stiffness parameter used as the uncertain parameter, the zero input response given by $x(t) = x_0 \cos\sqrt{k}t + \frac{\dot{x}(0)}{\sqrt{k}} \sin\sqrt{k}t$ clearly shows a nonlinear dependence of the solution on the unknown parameter. Furthermore, if one assumes a uniform distribution for this stiffness parameter, even with no uncertainty in the process model, the state probabiliity density function evolves in a non-Gaussian manner. Particle filters and other approaches can be used to solve this problem [10]. A useful byproduct of the application of state estimation methods and information theory is that a *posterior* measure of the parameter density function can be derived. This measure is not directly available when one considers parameter estimation approaches. An application of state estimation tools on a slowly varying unknown input in plasma dynamical systems is shown in the companion paper [17]. Application specific challenges and the challenges associated with the use of an extended Kalman filter that simplifies the generic problem by assuming Gaussian distributions for propagation and update are considered in the paper, outlining key aspects of the parameter estimation methods that build upon state estimation tools.

V. Sensitivity Analyses for Parameter Estimation

In parameter estimation applications, one of the most interesting aspects is the sensitivity of the parameters to the changes in the states $\boldsymbol{\omega}(\boldsymbol{\omega}_0, \mathbf{p}, \mathbf{u})$. In the fundamental parameter estimation problem of Eq. 1, the regressor matrix $\boldsymbol{H}$ is important in solving the inverse problem of estimating the parameter $\boldsymbol{\theta}$. While it is more difficult to develop a direct sensitivity measure, it is certainly easier to use parameter influence matrix and use it to develop an identifiability measure. This theory is now briefly developed. Parameter influence matrix $\psi(t, t_0) := \frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{\theta}}$ models the ability of the analyst to estimate the parameter $\boldsymbol{\theta}$ from the system of equations. This can be seen by looking at the analytic dependence of the states on the parameters of interest by expanding the state using a Taylor series expansion: $\boldsymbol{\omega} = \boldsymbol{\omega}_{\hat{\boldsymbol{\theta}}} + \psi(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \ldots$. The parameter influence matrix however can be calculated directly using the differential equation $\dot{\psi} = \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\omega}}\right] \psi + \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}\right]$. The influence matrix is a function of the trajectory, i.e., an implicit function of the initial state $\boldsymbol{\omega}_0$, $\mathbf{u}$, the time elapsed $t - t_0$ and the true parameter values $\boldsymbol{\theta}$. The linear independence of the columns of the parameter influence matrix is necessary for the regressor matrix to be well conditioned to solve for the parameters from the states. A measure of the linear independence can be derived from the observability grammian, which in turn is derived from the linear independence of time varying functions [48]. Since the identifiability measure is a positive definite matrix, it can be compared between different sets of trajectories, which in-turn are implicit functions of distinct initial conditions, parameter estimates and excitation inputs. A comparison of relative measures of identifiability is shown in Fig. 4, where different state trajectories shown in Fig. 3 are used for identification. The dynamics of the problem is that of a rigid body, whose attitude dynamics are excited by torque rods. Magnitude differences of the eigenvalues of the identifiability grammian show that some trajectories are more exciting than others in identification process. This is an important topic of future research in system identification.

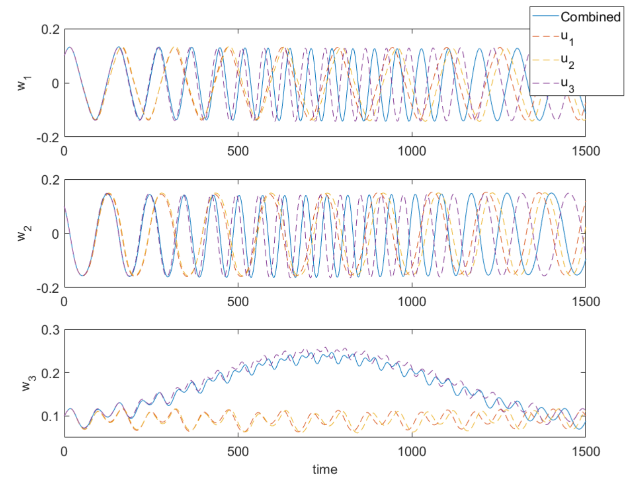

Fig. 3. Distinct state profiles for different types of inputs applied to a rigid body.

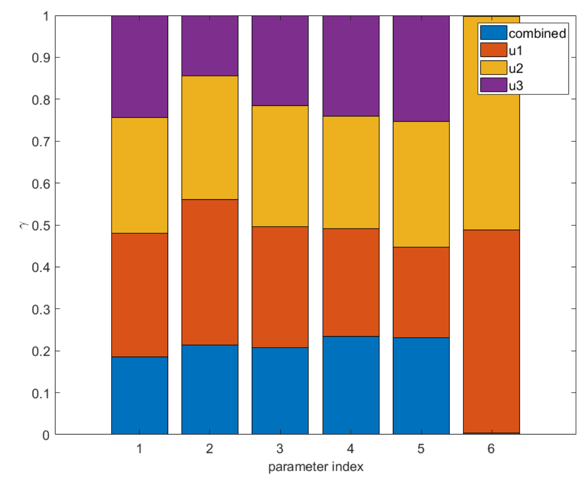

Fig. 4. Relative measures of identifiability of the inertia matrix parameters.

Acknowledgment

The third author wishes to acknowledge Texas A & M Institute of Data Sciences (TAMIDS) Career Initiation Fellowship program for their support. Partial support of the National Geospatial Intelligence Agency (Grant No. HM0476-19-1-2015) and the Office of Naval Research (Grant No. N00014-19-1-2435) is gratefully acknowledged. The results reported in sensitivity analysis techniques are part of collaborative research with Sandia National Laboratories. Mr. Michael Wang of Texas A & M University and Dr. Anup Parikh of Sandia National Laboratories are acknowledged for posing the parameter sensitivity problem and helping the second author develop the approaches discussed here.

References

[1] Nagavenkat Adurthi, Puneet Singla, and Tarunraj Singh. The conjugate unscented transform - an approach to evaluate multidimensional expectation integrals. In *Proceedings of the American Control Conference*, pages 5556–5561, Chicago, IL, 2012. IFAC.

[2] Nagavenkat Adurthi, Puneet Singla, and Tarunraj Singh. Conjugate unscented transformation: Applications to estimation and control. *Journal of Dynamic Systems, Measurement and Control*, 140(3):030907, 2018.

[3] M. Aoki and P. C. Yue. On certain convergence questions in system identification. *SIAM Journal of Control*, 8:239–256, 1970.

[4] Richard E Bellman. *Adaptive Control Processes - A Guided Tour*. Princeton University Press, Princeton, NJ, 1961.

[5] R. R. Bitmead. Persistence of excitaation conditions and the convergence of adaptive systems. *IEEE Transactionss on Information Theory*, 30(3):183–191, 1984.

[6] Robert Bitmead, M. Gevers, and V. Wertz. *Adaptive Optimal Control*. Prentice-Hall, Englewood Cliffs, NJ, 1990.

[7] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, United Kingdom, 2004.

[8] Steven L. Brunton and J. N. Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Contrrol*. Cambridge University Press, 2019.

[9] R. S. Bucy and P. D. Joseph. *Filtering for Stochastic Processes with Applications to Guidance*. John Wiley & Sons, New York, NY, USA, 1968.

[10] John Crassidis and John L. Junkins. *Optimal Estimation of Dynamical*

Systems. Nonlinear Systems Series. Chapman-Hall/CRC Press, Boca Raton, FL, 2nd edition, 2016.
[11] Megan Dickey. Uber's aerial taxi play, May 2018.
[12] D. L. Donoho. For most large underdetermined systems of linear equations the minimal l1-norm solution iss also the sparsest solution. *Communications on Pure and Applied Mathematics*, 56(6):797–829, 2006.
[13] D. L. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *Proceedings of the National Academy of Sciences*, 100(5):2197–2202, 2003.
[14] Urban Fasel, Eurika Kaiser, J. Nathan Kutx, Bingni W. Brunton, and Steven L. Brunton. Sindy with control: A tutorial. In *Proceedings of the IFAC CDC Conference 2021*, page submitted. International Federation of Automatic Control, 2021.
[15] G. C. Goodwin and D. Q. Mayne. A parameter estimation perspective of continuous time adaptive control. *Automatica*, 23, 1987.
[16] G. C. Goodwin and K. C. Sin. *Adaptive Filtering, Prediction and Control*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
[17] Christine Greve, Kentaro Hara, and Manoranjan Majji. Application of state estimation methods to low-temperature plasma dynamics. In *Proceedings of the IFAC CDC Conference 2021*, page submitted. International Federation of Automatic Control, 2021.
[18] Christine Greve, A. Thomas, Manoranjan Majji, and Kentaro Hara. Real-time estimation of electron dynamics in hall effect thrusters using an extended kalman filter. In *Proceedingss of the AIAA Propulsion and Energy Forum*, number AIAA 2020-3621, 2020.
[19] Vikram Hrishikeshavan, Moble Benedict, and Inderjit Chopra. Identification of flight dynamics of a cyclocopter micro air vehicle in hover. *Journal of Aircraft*, 52(1):116–129, 2014.
[20] Daning Huang. Applications of gaussian process regression in the aero-thermo-servo-elastic analysis towardss integrated hypersonic flight dynamic analysis. In *Proceedings of the IFAC CDC Conference 2021*, page submitted. International Federation of Automatic Control, 2021.
[21] P. C. Hughes and R. E. Skelton. Controllability and observability of linear matrix second order systems. *Journal of Applied Mechanics*, 47:415–420, 1980.
[22] Petros Ioannou and Jing Sun. *Robust Adaptive Control*. Adaptive and Learning Systems. Prentice-Hall, Englewood Cliffs, NJ, 1996.
[23] C. R. Johnson. *Lectures on Adaptive Parameter Estimation*. Prentice-Hall, Englewood Cliffs, NJ, 1988.
[24] Jer-Nan Juang. *Applied System Identification*. Prentice-Hall, Englewood Cliffs, NJ, 1994.
[25] Simon Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3), 2004.
[26] Simon Julier and Jeffrey Uhlmann. A new method for the nonlinear transformation of means and covariances in nonlinear filters. *IEEE Transactions on Automatic Control*, 45(3), 2000.
[27] J. L. Junkins and Youdan Kim. *Introduction to Dynamics and Control of Flexible Structures*. AIAA Education Series, 1992.
[28] J. L. Junkins, Z. H. Rahman, and H. Bang. Near minimum time control of distributed parameter systems: Analytical and experimental results. *Journal of Guidance, Control and Dynamics*, 14(2):406–415, 1991.
[29] Thomas Kailath. *Linear Systems*. Prentice Hall, Upper Saddle River, NJ, 1980.
[30] R. E. Kalman. Design of a self optimizing control system. *Transaction of the ASME*, 80:468–478, 1958.
[31] R. E. Kalman. A new approach for linear filtering and prediction problems. *Journal of Basic Engineering*, 82:35–45, 1960.
[32] M. Kristic, I. Kanellakopoulos, and P. Kokotovic. *Nonlinear and Adaptive Control Design*. Wiley, 1995.
[33] K. Liu and R. E. Skelton. The q-markov cover and its applications to flexible structure identification. *Journal of Guidance, Control and Dynamics*, 57(2), 1992.
[34] Lennart Ljung. *System Identification: Theory for the User*. Information and System Science Series. Prentice Hall, Upper Saddle River, NJ, 1999.
[35] David G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, NY, USA, 1969.
[36] Kevin Y. Ma, Pakpong Chirarattananon, Sawyer B. Fuller, and Robert J. Wood. Controlled flight of a biologically inspired, insect-scale robot. *Science*, 340(6132):603–607, 2013.
[37] Robert Marino and P. Tomei. Dynamic output-feedback linearization and global stabilization. *Systems and Control Letters*, 17(3), 1991.
[38] Leonard Meirovitch. *Elements of Vibration Analysis*. McGraw-Hill Book Company, 2nd edition, 1986.
[39] M. P. Murphy, A. Saunders, C. Moriera, Alfred A. Rizzi, and Mark Raibert. The littledog robot. *The International Journal of Robotics Research*, 30(2):145–149, 2011.
[40] K. S. Narendra and A. M. Annaswamy. *Stable Adaptive Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1989.
[41] K. S. Narendra and K. Parthasarathy. Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1(1):4–27, 1990.
[42] Harry Nyquist. Regeneration theory. *Bell System Technical Journal*, 11(1):126–147, 1932.
[43] Rik Pintelton and Johan Schoukens. *System Identification: A Frequency Domain Approach*. IEEE Press, Piscataway, NJ, 2001.
[44] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006.
[45] Alfred A. Rizzi and Mark Raibert. Spot: Programming diverse dynamic robot behavior. In Kumpati S. Narendra, editor, *Proceedings of the 18th Yale Workshop on Adaptive and Learning Systems*, volume 18, 2017.
[46] S. Sastry and Marc Bodson. *Adaptive Systems: Stability, Convergence and Robustness*. Prentice-Hall Information and System Sciences Series. Prentice-Hall, 1989.
[47] Guaming Shi and R. E. Skelton. Iterative identification and control using a weighted q-markov cover with measurement noise. *Signal Processing*, 52(2):217–234, 1996.
[48] Robert E. Skelton. *Dynamic Systems Control: Linear Systems Analysis and Synthesis*. John Wiley & Sons, New York, 1988.
[49] H. W. Sorenson. Least squares estimation from gauss to kalman. *IEEE Spectrum*, 7:63–68, 1970.
[50] H. W. Sorenson. *Parameter Estimation: Principles and Problems*. Mercel Dekker Publishing Co, New York, 1980.
[51] William Terryl Thomson. *Theory of Vibration with Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1972.
[52] Ryan L. Truby, Michael Wehner, Abigail K. Grosskopf, Daniel Vogt, Robert J. Wood, and Jennifer A. Lewis. Soft sematosensitive actuators via embedded 3d printing. *Advanced Materials*, 1706383, 2018.
[53] Kostas Tsakalis and Petros Ioannou. Adaptive control of linear time-varying plants. *Automatica*, 23(4):459–468, 1987.
[54] Dong-Huei Tseng. *Variance and Bias Computation for Improved Modal Identification*. PhD thesis, Columbia University, New York, NY, USA, 1992.
[55] Dong-Huei Tseng, Richard W. Longman, Jer-Nan Juang, and C.-W. Chen. Variance and bias confidence criteria for system and observer/kalman filter identification. In *Proceedings of the Thirteeth Annual Allerton Conference on Communication, Control and Computing*, pages 98–109, September 1992.
[56] A. D. Valentine, T. A. Busbee, and Jennifer A. Lewis. Hybrid 3d printing of soft electronics. *Advanced Materials*, 1703817, 2017.
[57] Ranjan Vepa. *Dynamics of Smart Structures*. John Wiley & Sons, 2010.
[58] James Vincent. Robots with legs are getting ready to walk among us, February 2017.
[59] D. A. Wagie, A. Yousuff, and R. E. Skelton. Linear system approximation via covariance equivalent realization. *Journal of Mathematical Analysis and Applications*, 106(1):91–115, 1985.
[60] H. L. Weidemann. Entropy analysis of feedback control systems. *Advances in Control Systems*, 7:225–255, 1969.
[61] J. Zaborsky. An information theory viewpoint for the general identification problem. *IEEE Transactions on Automatic Control*, 11:130–131, 1966.
[62] Huihua Zhao, Jonathan Horn, Jacob Reher, Victor Paredes, and Aaron D. Ames. First steps toward translating robotic walking to presthesis: a nonlinear opitimization based control approach. *Autonomous Robots*, 41(3):725–742, 2017.