Observer Design for Stochastic Nonlinear Systems via Contraction-Based Incremental Stability

Ashwin P. Dani, *Member, IEEE*, Soon-Jo Chung, *Senior Member, IEEE*, and Seth Hutchinson, *Fellow, IEEE*

Abstract—This paper presents a new design approach to nonlinear observers for Itô stochastic nonlinear systems with guaranteed stability. A stochastic contraction lemma is presented which is used to analyze incremental stability of the observer. A bound on the mean-squared distance between the trajectories of original dynamics and the observer dynamics is obtained as a function of the contraction rate and maximum noise intensity. The observer design is based on a non-unique state-dependent coefficient (SDC) form, which parametrizes the nonlinearity in an extended linear form. The observer gain synthesis algorithm, called linear matrix inequality state-dependent algebraic Riccati equation (LMI-SDARE), is presented. The LMI-SDARE uses a convex combination of multiple SDC parametrizations. An optimization problem with state-dependent linear matrix inequality (SDLMI) constraints is formulated to select the coefficients of the convex combination for maximizing the convergence rate and robustness against disturbances. Two variations of LMI-SDARE algorithm are also proposed. One of them named convex state-dependent Riccati equation (CSDRE) uses a chosen convex combination of multiple SDC matrices; and the other named Fixed-SDARE uses constant SDC matrices that are pre-computed by using conservative bounds of the system states while using constant coefficients of the convex combination pre-computed by a convex LMI optimization problem. A connection between contraction analysis and $\mathcal{L}_2$ gain of the nonlinear system is established in the presence of noise and disturbances. Results of simulation show superiority of the LMI-SDARE algorithm to the extended Kalman filter (EKF) and state-dependent differential Riccati equation (SDDRE) filter.

Index Terms—Estimation theory, state estimation, stochastic systems, observers, optimization methods.

I. INTRODUCTION

THE present paper is motivated by the fact that the state estimation for many engineering and robotics applications, such as simultaneous localization and mapping (SLAM) [1], [2], has to overcome issues with nonlinearities and stochastic uncertainty. Itô stochastic nonlinear dynamic systems, driven by white noise, exhibit a non-Gaussian probability density, whose time-evolution is characterized by the Fokker-Planck equation—a partial differential equation [3], [4]. Both nonlinearity and non-Gaussian distribution of the probability density of the states make the optimal estimation problem very challenging. A state estimator for nonlinear systems driven by Cauchy noise is presented in [5]. Popular filtering approaches for nonlinear systems include the extended Kalman filter (EKF) [6], the unscented Kalman filter [7], particle filters [8], and the set membership filter [9]. Conventional nonlinear observer designs are based on deterministic nonlinear systems (e.g., Lipschitz nonlinear systems [10], [11], monotone nonlinearities [12], and high gain observers [13]).

The observer design methods based on deterministic systems neglect the stochastic measurement and process noise in the system. The main objective of such observers is to design (possibly globally stable) observers for various classes of nonlinearities using a nonlinear transformation of the original system into a pseudo-linear form or using a dominant linear time invariant (LTI) term in the dynamics (i.e., $\dot{x} = Ax + g(x)$). A nonlinear observer is designed in a recent work [14] for deterministic systems with a special class of nonlinearities that satisfy incremental quadratic inequalities. In [15], a state estimation algorithm for stochastic nonlinear systems is presented for the nonlinearities satisfying the integral quadratic constraint (IQC). In [16], EKF algorithms are developed based on a Carleman approximation (see [17]), which transforms the original nonlinear system into a polynomial form. The dimension of the transformed state space is higher than the original system and increases with the degree of the Carleman approximation. Note that its first degree is equivalent to the Jacobian used in the EKF. In [18], [19], a differential mean-value theorem (DMVT) is used to transform the nonlinearity into a linear parameter varying (LPV) system which leads to an LMI feasibility problem. In [20], a simultaneous input and state estimation method is proposed based on the Gauss-Newton optimization method.

In contrast with the nonlinear transformations used in a deterministic case and the linearization approach used in EKF-like observers, an alternative approach is based on parametrization of nonlinear systems into an "extended linear" form or so called state-dependent coefficient (SDC) form [21], [22]. This paper addresses the problem of observer design in the presence of nonlinearities and stochastic noise by rewriting an Itô stochastic nonlinear system in an SDC form. The SDC parametrization is not unique and there exist multiple choices of such parametrization. An algorithm called "linear matrix inequality state-dependent algebraic Riccati equation" (LMI-SDARE) is proposed which uses the degree of freedom of the SDC form to compute the observer gain. A block diagram describing
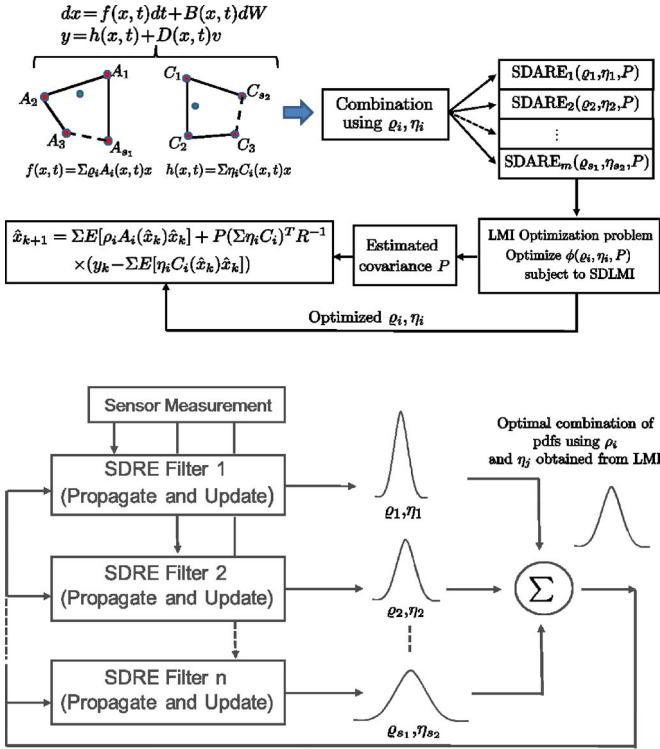
Fig. 1. System diagrams describing the LMI-SDARE algorithm: the flow chart of the LMI-SDARE (top) and the conceptual design of the LMI-SDARE (bottom).

the LMI-SDARE algorithm is shown in Fig. 1. The observer gain design problem is cast into a state-dependent linear matrix inequality (SDLMI) feasibility problem. An optimization problem is formulated with the SDLMI constraints so that the optimal convex combination of the multiple SDC forms can be selected to compute the observer gain that achieves desirable estimator properties, such as the improved convergence rate and the minimum mean-squared estimation error. The optimization problem is converted to a convex optimization problem and can be solved using the polynomial-time interior point methods [23], [55]. Two variations of the LMI-SDARE algorithm, called "convex state-dependent Riccati equation" (CSDRE) filter and "fixed-state-dependent algebraic Riccati equation" (Fixed-SDARE) filter, are also presented. For the CSDRE algorithm, a differential Riccati equation is integrated from a positive definite initial condition or a state-dependent algebraic Riccati equation is solved using an *a priori* selected convex combination of the multiple state-dependent parametrizations A and C to compute the observer gain. For the Fixed-SDARE filter, constant (A, C) matrices computed based on the conservative upper bounds on the states of the system are used to solve a linear matrix inequality for computing parameters of the convex combination of constant (A, C) matrices. The observer gain for the Fixed-SDARE filter is a constant matrix which can be computed *a priori*. Incremental stability of the proposed observer algorithm is analyzed by using contraction analysis [24], [25]. Compared to our preliminary work [26], the present paper shows a more systematic process of the observer design and provides rigorous stability proofs.

Contribution 1: Stochastic Contraction and Observer Design: Nonlinear stability is analyzed with respect to an equilibrium point by using Lyapunov-like approaches. In contrast, incremental stability (e.g., see [27]) analyzes the behavior of the system trajectories with respect to each other and is very useful for observer design and synchronization problems [28], [29]. In this paper, a new lemma is presented which characterizes stochastic incremental stability of the solutions of two Itô stochastic differential equations driven by different Wiener processes. The lemma shows that the distance between the solution trajectories of two systems, with respect to a state-dependent Riemannian metric, exponentially converges to a bound. Furthermore, an exponentially stabilizing observer for Itô stochastic differential equations is designed, which computes the observer gain using state-dependent linear matrix inequalities (SDLMIs). Stochastic incremental stability of the observer is analyzed by using the result of the stochastic contraction lemma.

In [24], contraction analysis is developed as a tool to study incremental stability of deterministic systems. The incremental stability approaches developed in [30]–[32] are similar to contraction analysis for analyzing stability of differential equations. Recently, in [25], contraction analysis has been extended to stochastic nonlinear systems using a state-*independent* metric. The results in this paper are more general than the observer example presented in [25], because of the following reasons: (1) incremental stability of stochastic systems is shown with respect to a state-dependent metric; (2) explicit algorithms for observer gain computation are presented; and (3) the observer can be designed for a nonlinear measurement model. Since a contracting system is globally exponentially stable, the system exhibits robustness properties and follows the connection between $\mathcal{L}_2$ gain of the system and time-domain H_∞ norm [33], [34]. A connection between contraction theory and robustness of the stochastic nonlinear systems, characterized by $\mathcal{L}_2$ gain of the system is established. It is shown that the estimation error has a finite $\mathcal{L}_2$ gain with respect to the noise and disturbance acting on the system. This provides a property of robustness against uncertainties in the dynamics.

Contribution 2: Optimal Choice of the Convex SDC Forms for Observer Design: As pointed out earlier, SDC formulations are not unique. In general, the choice of parametrization primarily depends on the observability property. In this paper, a convex combination of such parametrizations is used to achieve desirable estimator properties in addition to avoiding loss of observability. Hence, this paper presents the first result that fully uses the flexibility of SDC form for observer gain design, thereby providing a systematic algorithm based on SDLMI constraints solved by convex optimization. The effectiveness of this algorithm is demonstrated by the results of two simulation examples; the LMI-SDARE algorithm outperforms the EKF and conventional SDRE filter.

SDC-based filters for deterministic systems are presented in [35]–[37]. Most existing approaches to SDC-based filters solve a state-dependent algebraic Riccati equation (SDARE) at every time instant to obtain a positive definite (PD) solution used to construct the observer gain. Some observer design approaches use a state-dependent differential Riccati equation (SDDRE)

formulation that propagates a solution to the differential Riccati equation by integration [36]. The conventional SDARE and SDDRE filters use a single state-dependent parametrization (A, C) to solve the Riccati equation. In contrast, the algorithms presented in this paper use a convex combination of multiple parametrizations to achieve a better estimation performance as shown with the help of numerical simulations in Section VI. If the uniform observability property is satisfied, the SDDRE has a PD solution. In [38], an algorithm is proposed to re-select the parametrization at a particular time-instant if the observability of parametrization is lost.

Notation: Throughout the paper, we adopt the following notation. For a vector $x \in \mathbb{R}^n$, $\|x\|$ denotes the Euclidean norm, x_i denotes the ith component of vector x. For a real matrix $A \in \mathbb{R}^{n \times m}$, $\|A\|$ denotes the induced 2-norm; i.e., the maximum singular value of A, $\|A\|_F$ denotes the Frobenius norm, A_{ij} denotes the component in ith row and jth column, $(A_{ij})_a$ denotes the partial derivative with respect to vector a, $(A_{ij})_{a_i a_j}$ denotes the double partial derivative with respect to components of vector a, $(A_{ij})_t$ denotes the partial derivative with respect to t; if A is square, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote minimum and maximum eigenvalues, and $\mathrm{tr}(A)$ denotes the trace of A; $I \in \mathbb{R}^{n \times n}$ denotes the identity matrix; $\mathrm{Co}\{A_1, \ldots, A_n\}$ denotes the convex hull of matrices $A_1, \ldots, A_n$; $E[\cdot]$ denotes the expected value operator, $E_{y_0}[y(t)]$ denotes the expected value of y at the time instant t given the initial value $y(t_0) = y_0$; and the $\mathcal{L}_2$ norm of a vector is defined in an $\mathcal{L}_2$−extended space.

II. PRELIMINARIES

A. Brief Review of Contraction Analysis

In this section, contraction analysis [24] for analyzing exponential stability of nonlinear systems is briefly reviewed. Consider a nonlinear, non-autonomous system of the form

$$\dot{x} = f(x, t) \qquad (1)$$

where $x(t) \in \mathbb{R}^n$ is a state vector and $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is a continuously differentiable nonlinear function. With the assumed properties of (1), the exact relation $\delta \dot{x} = (\partial f(x, t)/\partial x)\delta x$ holds, where δx is an infinitesimal virtual displacement in fixed time. The squared virtual displacement between two trajectories of (1) with a symmetric, uniformly positive definite metric $M(x, t) \in \mathbb{R}^{n \times n}$ is given by $\delta x^T M(x, t)\delta x$ (cf. Riemannian metric [39]). Its time derivative is given by

$$\frac{d}{dt}\left(\delta x^T M(x, t)\delta x\right)$$
$$= \delta x^T \left(\frac{\partial f}{\partial x}^T M(x, t) + \dot{M}(x, t) + M(x, t)\frac{\partial f}{\partial x}\right)\delta x. \qquad (2)$$

If the following inequality is satisfied

$$\frac{\partial f}{\partial x}^T M(x, t) + \dot{M}(x, t) + M(x, t)\frac{\partial f}{\partial x} \leq -2\gamma M(x, t) \forall t, \forall x \qquad (3)$$

for a strictly positive constant γ, then the system (1) is said to be contracting with the rate γ and all the system trajectories

exponentially converge to a single trajectory irrespective of the initial conditions (hence, globally exponentially stable).

Now consider a perturbed system of (1)

$$\dot{x} = f(x, t) + d(x, t) \qquad (4)$$

such that the deterministic disturbance $\|d(x, t)\|$ is bounded. The following lemma shows that the distance between the trajectory of the perturbed system and the trajectory of the globally exponentially stable nominal system remains bounded.

Lemma 1: (Robustness of Contracting Dynamics) [24], [29] Let $T_1(t)$ be a trajectory of the globally contracting system (1) and $T_2(t)$ be a trajectory of a perturbed system (4). The smallest distance between $T_1(t)$ and $T_2(t)$ is defined by $S(t) \triangleq \int_{T_1}^{T_2} \|\delta z\|$, where $\delta z \triangleq \Theta(x, t)\delta x$ and $\Theta^T(x, t)\Theta(x, t) = M(x, t)$ satisfies

$$S(t) \leq S(t_0)e^{-\gamma(t-t_0)} + \frac{1 - e^{-\gamma(t-t_0)}}{\gamma} \sup_{x,t} \|\Theta d\| \quad \forall t \geq t_0. \quad (5)$$

As $t \to \infty$, $S(t) \leq \sup_{x,t} \|\Theta d\|/\gamma$.

Proof: Differentiating the distance $S(t)$, we obtain $\dot{S} + \gamma S \leq \|\Theta d\|$ [24]. For bounded $\|\Theta d\|$, the estimate in (5) can be obtained by using the comparison lemma (cf. [40, Lemma 3.4]). ∎

If the unperturbed system (1) is globally contracting, the perturbed system (4) is finite-gain $\mathcal{L}_p$ stable with $p \in [1, \infty]$ in the sense of the bounded output function $y = h(x, d, t)$ with $\int_{Y_1}^{Y_2} \|\delta y\| \leq \eta_0 \int_{T_1}^{T_2} \|\delta x\| + \eta_1 \|d\|$, $\forall \eta_0, \eta_1 > 0$ (see [29]) where $h : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$, $Y_1(t)$ and $Y_2(t)$ denote the output trajectories of the globally contracting system (1) and its perturbed system (4).

III. STOCHASTIC CONTRACTION LEMMA

Consider a stochastically perturbed system of the nominal system (1) represented using an Itô stochastic differential equation

$$dx = f(x, t)dt + B(x, t)dW, \qquad x(0) = x_0 \qquad (6)$$

and the conditions for existence and uniqueness of a solution to (6)

$$\exists L_1 > 0, \ \forall t, \ \forall x_1, x_2 \in \mathbb{R}^n :$$
$$\|f(x_1, t) - f(x_2, t)\| + \|B(x_1, t) - B(x_2, t)\|_F \leq L_1 \|x_1 - x_2\|,$$
$$\exists L_2 > 0, \ \forall t, \ \forall x_1 \in \mathbb{R}^n :$$
$$\|f(x_1, t)\|^2 + \|B(x_1, t)\|_F^2 \leq L_2 \left(1 + \|x_1\|^2\right) \qquad (7)$$

where $B : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n \times d}$ is a matrix-valued function, $W(t)$ is a d-dimensional Wiener process, and x_0 is a random variable independent of W [41].

Consider any two systems with trajectories $a(t)$ and $b(t)$ obtained by the same function $f(\cdot)$ in (6) but driven by independent Wiener processes $\bar{W}_1$ and $\bar{W}_2$

$$dz = \begin{pmatrix} f(a, t) \\ f(b, t) \end{pmatrix} dt + \begin{pmatrix} B_1(a, t) & 0 \\ 0 & B_2(b, t) \end{pmatrix} \begin{pmatrix} d\bar{W}_1 \\ d\bar{W}_2 \end{pmatrix}$$
$$= f_s(z, t)dt + B_s(z, t)d\bar{W} \qquad (8)$$

where $z(t) = (a(t)^T, b(t)^T)^T \in \mathbb{R}^{2n}$.

We present the so-called stochastic contraction lemma that uses a state-dependent metric, thereby generalizing the main result presented in [25]. The following lemma analyzes stochastic incremental stability of the two trajectories $a(t)$ and $b(t)$ with respect to each other in the presence of noise where the system without noise $\dot{x} = f(x,t)$ is contracting in a state-dependent metric $M(x(\mu,t),t)$, for $\mu \in [0,1]$. The trajectories of (6) are parametrized as $x(0,t) = a$ and $x(1,t) = b$, and $B_1(a,t)$ and $B_2(b,t)$ are defined as $B(x(0,t),t) = B_1(a,t)$, and $B(x(1,t),t) = B_2(b,t)$, respectively.

Assumption 1: $\operatorname{tr}(B_1(a,t)^T M(x(a,t),t)B_1(a,t)) \leq C_1$, $\operatorname{tr}(B_2(b,t)^T M(x(b,t),t)B_2(b,t)) \leq C_2$, $\bar{m}_x = \sup\limits_{t \geq 0, i,j} \|(M_{ij}(x,t))_x\|$, and $\bar{m}_{x^2} = \sup\limits_{t \geq 0, i,j} \|\partial^2(M_{ij}(x,t))/\partial x^2\|$, where C_1, C_2, $\bar{m}_x$, and $\bar{m}_{x^2}$ are constants.

Assumption 2: The nominal deterministic system (1) is contracting in a metric $M(x(\mu,t),t)$ in the sense that (3) is satisfied and $M(x(\mu,t),t)$ satisfies the bound $\underline{m} \triangleq \inf\limits_{t \geq 0}(\lambda_{\min}M)$. The function f and the metric M are the same as in (1) and (3).

Lemma 2. (Stochastic Contraction Lemma): Consider the generalized squared length with respect to a Riemannian metric $M(x(\mu,t),t)$ defined by $V(x,\delta x,t) = \int_0^1 (\partial x/\partial \mu)^T M(x(\mu,t),t)(\partial x/\partial \mu)d\mu$ such that $\underline{m}\|a-b\|^2 \leq V(x,\delta x,t)$. If Assumptions 1 and 2 are satisfied then the trajectories $a(t)$ and $b(t)$ of (8), whose initial conditions, given by a probability distribution $p(a_0,b_0)$, are independent of $d\hat{W}_1$ and $d\hat{W}_2$, satisfy the bound

$$E\left[\|a(t)-b(t)\|^2\right] \leq \frac{1}{\underline{m}}\left(\frac{C}{2\gamma_1} + E\left[V\left(x(0),\delta x(0),0\right)\right]e^{-2\gamma_1 t}\right) \tag{9}$$

where $\exists \varepsilon > 0$ such that $\gamma_1 \triangleq \gamma - ((\beta_1^2 + \beta_2^2)/2\underline{m})(\varepsilon\bar{m}_x + (\bar{m}_{x^2}/2)) > 0$, γ is the contraction rate defined in (3), $C = C_1 + C_2 + (\bar{m}_x/\varepsilon)(\beta_1^2 + \beta_2^2)$, $\beta_1 = \|B_1\|_F$, and $\beta_2 = \|B_2\|_F$.

Proof: By using the Itô formula [3], [41], the stochastic derivative of the Lyapunov function $V(x,\delta x,t)$ is given by $dV(x,\delta x,t) = \mathcal{L}V(x,\delta x,t)dt + \sum_{i=1}^n \sum_{j=1}^d V_{x_i}(x,\delta x,t)(B(x,t))_{ij}dW_j + V_{\delta x_i}(x,\delta x,t)(\delta B(x,t))_{ij}dW_j$, where $\mathcal{L}$ is an infinitesimal differential generator such that

$$\mathcal{L}V = V_t + \sum_{i=1}^n \left(V_{x_i}f_i + V_{\delta x_i}\frac{\partial f}{\partial x}\delta x\right)$$

$$+ \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \left[V_{x_i x_j}\left(B(x,t)B^T(x,t)\right)_{ij}\right.$$

$$+ V_{\delta x_i \delta x_j}\left(\delta B(x,t)\delta B^T(x,t)\right)_{ij}$$

$$\left. + 2V_{x_i \delta x_j}\left(B(x,t)\delta B^T(x,t)\right)_{ij}\right]. \tag{10}$$

Using (6), (10) can be written as

$$\mathcal{L}V = \int_0^1 \left(\frac{\partial x}{\partial \mu}\right)^T dM\left(x(\mu,t),t\right)\left(\frac{\partial x}{\partial \mu}\right)d\mu$$

$$+ \int_0^1 \left(\frac{\partial x}{\partial \mu}\right)^T \left(M\frac{\partial f}{\partial x} + \frac{\partial f}{\partial x}^T M\right)\left(\frac{\partial x}{\partial \mu}\right)d\mu + V_b \tag{11}$$

such that $dM_{ij}(x(\mu,t),t) = (\partial M_{ij}(x(\mu,t),t)/\partial t) + (M_{ij})_x f(x(\mu,t),t)$ and

$$V_b = \int_0^1 \sum_{i=1}^n \sum_{j=1}^n M_{ij}\left(\frac{\partial B(x,t)}{\partial \mu}\frac{\partial B(x,t)}{\partial \mu}^T\right)_{ij}d\mu$$

$$+ \int_0^1 \left[2\sum_{i=1}^n \sum_{j=1}^n (M_i)_{x_j}\left(\frac{\partial x}{\partial \mu}\right)\left(B(x,t)\frac{\partial B(x,t)}{\partial \mu}^T\right)_{ij}\right]d\mu$$

$$+ \frac{1}{2}\int_0^1 \left(\sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n \sum_{l=1}^n (M_{kl}(x(\mu,t),t))_{x_i x_j}\right.\right.$$

$$\left.\left. \times \frac{\partial x_k}{\partial \mu}\frac{\partial x_l}{\partial \mu}\right)(B(x,t)B(x,t)^T)_{ij}\right)d\mu \tag{12}$$

where x is a function of μ such that $x(0,t) = a$ and $x(1,t) = b$, and M_i is the ith row of M. The following bounds can be computed:

$$\int_0^1 \sum_{i=1}^n \sum_{j=1}^n M_{ij}\left(\frac{\partial B(x,t)}{\partial \mu}\frac{\partial B(x,t)}{\partial \mu}^T\right)_{ij}d\mu$$

$$\leq \operatorname{tr}\left(M(a,t)B_1 B_1^T\right) + \operatorname{tr}\left(M(b,t)B_2 B_2^T\right) \tag{13}$$

$$\frac{1}{2}\int_0^1 \left(\sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n \sum_{l=1}^n (M_{kl}(x,t))_{x_i x_j}\right.\right.$$

$$\left.\left. \times \frac{\partial x_k}{\partial \mu}\frac{\partial x_l}{\partial \mu}\right)(BB^T)_{ij}\right)d\mu$$

$$\leq \frac{1}{2}\bar{m}_{x^2}\left(\beta_1^2 + \beta_2^2\right)\int_0^1 \left\|\frac{\partial x}{\partial \mu}\right\|^2 d\mu \tag{14}$$

$$\int_0^1 \left[2\sum_{i=1}^n \sum_{j=1}^n (M_i)_{x_j}\left(\frac{\partial x}{\partial \mu}\right)\left(B(x,t)\frac{\partial B(x,t)}{\partial \mu}^T\right)_{ij}\right]d\mu$$

$$\leq 2\bar{m}_x\left(\beta_1^2 + \beta_2^2\right)\int_0^1 \left\|\frac{\partial x}{\partial \mu}\right\|d\mu$$

$$\leq \bar{m}_x\left(\beta_1^2 + \beta_2^2\right)\left(\int_0^1 \varepsilon\left\|\frac{\partial x}{\partial \mu}\right\|^2 d\mu + \frac{1}{\varepsilon}\right) \tag{15}$$

where the inequality $2a'b' \leq \varepsilon^{-1}a'^2 + \varepsilon b'^2$, for scalars a' and b' with $\varepsilon > 0$, is used. Using (3) and the bounds in (13), (14), and (15), the differential generator in (11) can be bounded above as follows

$$\mathcal{L}V \leq -2\gamma_1 \int_0^1 \left(\frac{\partial x}{\partial \mu}\right)^T M\left(x(\mu,t),t\right)\left(\frac{\partial x}{\partial \mu}\right)d\mu$$

$$+ \frac{\bar{m}_x}{\varepsilon}\left(\beta_1^2 + \beta_2^2\right) + \operatorname{tr}\left(M(a,t)B_1 B_1^T\right)$$

$$+ \operatorname{tr}\left(M(b,t)B_2 B_2^T\right) \tag{16}$$

where $\gamma_1 \triangleq \gamma - ((\beta_1^2 + \beta_2^2)/2\underline{m})(\varepsilon\bar{m}_x + (\bar{m}_{x^2}/2)) > 0$. Assumptions 1–2 and (16) yield

$$\mathcal{L}V \leq -2\gamma_1 V + C. \tag{17}$$

Using the stopping time argument, the integral of the last term in $dV = \mathcal{L}V dt + \sum_{i=1}^{n} \sum_{j=1}^{d} V_{x_i}(x, \delta x, t)(B(x,t))_{ij} dW_j + V_{\delta x_i}(x, \delta x, t)(\delta B(x,t))_{ij} dW_j$ is a martingale (cf. Theorem 4.1 of [42]). Taking the expectation operator on both the sides of dV and using (17) along with Dynkin's formula (pp. 10 of [3]) $\forall u, t, 0 \leq u \leq t < \infty$ yields

$$E_{x_0}[V(x(\mu,t), \delta x(t), t)] - E_{x_0}[V(x(\mu,u), \delta x(u), u)]$$
$$\leq \int_u^t (-2\gamma_1 E_{x_0}[V(x(\mu,s), \delta x(s), s)] + C) \, ds \tag{18}$$

where Fubini's theorem is used for changing the order of integration [3]. By using the Gronwall-type lemma (see Appendix A), the following inequality can be developed:

$$E_{x_0}[V(x(\mu,t), \delta x, t)]$$
$$\leq \left[V(x(0), \delta x(0), 0) - \frac{C}{2\gamma_1} \right]^+ e^{-2\gamma_1 t} + \frac{C}{2\gamma_1} \tag{19}$$

where $[\cdot]^+ = \max(0, \cdot)$. Integrating (19) with respect to x_0, and using $\underline{m}E[\|a-b\|^2] \leq E[V(x, \delta x, t)]$, $[V(x(0), \delta x(0), 0) - (C/2\gamma_1)]^+ \leq V(x(0), \delta x(0), 0)$, and $E[V(x(0), \delta x(0), 0)] = \int V(x(0), \delta x(0), 0)dp(x_0)$, the bound on the mean-squared estimation error given in (9) is obtained. Hence, the mean-squared estimation error is exponentially bounded. ∎

A. Choice of ε for Optimal Bound in (9)

The contraction rate γ_1 and the uncertainty bound C in (9) depend on the choice of ε. To derive an optimal choice of ε so that $C/(2\underline{m}\gamma_1)$ in (9) is minimized, consider $F(\varepsilon) = C/(2\underline{m}\gamma_1) = (1/(2\underline{m})(\gamma - ((\beta_1^2 + \beta_2^2)/2\underline{m})(\varepsilon\bar{m}_x + (\bar{m}_{x^2}/2))))(C_1 + C_2 + (L/\varepsilon))$, where $L = \bar{m}_x(\beta_1^2 + \beta_2^2)$. Computing $dF/d\varepsilon = 0$ yields $(C_1+C_2)(\beta_1^2+\beta_2^2)\bar{m}_x\varepsilon^2 + 2L(\beta_1^2+\beta_2^2)\bar{m}_x\varepsilon - 2L\underline{m}\gamma + L(\beta_1^2 + \beta_2^2)(\bar{m}_{x^2}/2) = 0$, whose solution minimizes the bound $C/(2\underline{m}\gamma_1)$ in (9).

Remark 1: In certain cases, such as the observer design in Section V, the convergence of the trajectories of the solutions of two stochastic systems is difficult to verify by using a state-independent metric, such as $M(t)$ (cf. [43]). Hence, generalization of the stochastic contraction result with a state-dependent metric is important. The bound in (9) reduces to the one obtained in [25] when $M = M(t)$ or $M = $ constant. It is easy to verify from (12) that for $M = M(t)$ or $M = $ constant, the terms related to $(M_i)_{x_j}$ and $(M_{kl})_{x_i x_j}$ vanish because they are independent of x.

IV. System Formulation for Observer Design

Consider a dynamic system represented by an Itô stochastic differential equation with a measurement equation

$$dx = f(x,t)dt + d(x,t)dt + B(x,t)dW_1(t) \tag{20}$$
$$y = h(x,t) + D(x,t)\nu(t) \tag{21}$$

where $x(t) \in \mathbb{R}^n$ is the state; $y(t) \in \mathbb{R}^m$ is the measurement; $f(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$; $h(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$; $B(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n \times n}$; $D(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{m \times m}$; $\nu(t)$ is m dimensional white noise formally defined as $dW_2 = \nu(t)dt$; $W_1(t)$ and $W_2(t)$ are standard n and m dimensional independent Wiener processes; and $d(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is an unknown, bounded, deterministic disturbance. The $d(x,t)$ term is not considered in Theorems 1–2 but is used in Theorem 3 to show the robustness result using $\mathcal{L}_2$ stability.

Problem Statement: Given the stochastic system in (20) and the measurement (21), the objective is to design a state observer that estimates the state $x(t)$ using noisy measurements $y(t)$; i.e., given all the noisy measurements up to time $t > 0$, compute $\hat{x}(t)$ such that

$$E\left[\|x(t) - \hat{x}(t)\|^2\right] \leq \zeta_1 E\left[\|x(0) - \hat{x}(0)\|^2\right] e^{-\zeta_2 t} + \epsilon \tag{22}$$

for positive constants ζ_1, ζ_2, and ϵ. The observer is another stochastic differential equation that uses sensor measurements and exponentially forgets the initial conditions to follow the behavior of the original system (20).

The nonlinear system (20) can also be expressed in a state-dependent coefficient (SDC) form

$$dx = A(x,t)xdt + d(x,t)dt + B(x,t)dW_1(t) \tag{23}$$
$$y = C(x,t)x + D(x,t)\nu(t) \tag{24}$$

where $f(x,t) = A(x,t)x$ and $h(x,t) = C(x,t)x$ are parametrized using nonlinear matrix functions $A(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n \times n}$ and $C(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{m \times n}$. The choice of such parametrization is not unique for $n > 1$. If there exists $A(x,t) \in \text{Co}\{A_1(x,t), A_2(x,t), \ldots, A_{s_1}(x,t)\}$ such that

$$f(x,t) = A_1(x,t)x = A_2(x,t)x = \ldots = A_{s_1}(x,t)x \tag{25}$$

then there exist an infinite number of parametrizations

$$f(x,t) = A(\varrho,x,t)x = \varrho_1 A_1(x,t)x + \ldots + \varrho_{s_1} A_{s_1}(x,t)x \tag{26}$$

where $\varrho = \{\varrho_i | i = 1, \ldots, s_1\}$, $\varrho_i \geq 0$ and $\sum_{i=1}^{s_1} \varrho_i = 1$. Similarly, $h(x,t)$ can be parametrized using $C(x,t) \in \text{Co}\{C_1(x,t), C_2(x,t), \ldots, C_{s_2}(x,t)\}$ as

$$h(x,t) = C(\eta,x,t)x = \eta_1 C_1(x,t)x + \ldots + \eta_{s_2} C_{s_2}(x,t)x \tag{27}$$

where $\eta = \{\eta_i | i = 1, \ldots, s_2\}$, $\eta_i \geq 0$ and $\sum_{i=1}^{s_2} \eta_i = 1$. In the subsequent sections, a convex-optimization problem along with an LMI constraint is presented to optimize a suitable cost metric over the coefficients ϱ and η.

Assumption 3: The convex combination of the matrices $(A(\varrho,x,t), C(\eta,x,t))$ is selected such that the pair is uniformly observable.

Remark 2: Even though the individual parametrization $(A_i(x,t), C_i(x,t))$ is not observable at some points of the state space, the parameters ϱ and η can be used to preserve the uniform observability of $(A(\varrho,x,t), C(\eta,x,t))$. This fact is illustrated using Example 1 below.

Let

$$\frac{\partial}{\partial x}\left(A(\varrho,x,t)x\right) = A(\varrho,x,t) + \Delta_1(\varrho,x,t) \qquad (28)$$

$$\frac{\partial}{\partial x}\left(C(\eta,x,t)x\right) = C(\eta,x,t) + \Delta_2(\eta,x,t) \qquad (29)$$

The matrices $\Delta_1(\varrho,x,t)$ and $\Delta_2(\eta,x,t)$ can be evaluated by multiplying the rows of $A(\varrho,x,t)$ and $C(\eta,x,t)$ by $x(t)$ and computing the partial derivatives of the entries of $A(\varrho,x,t)$ and $C(\eta,x,t)$ matrices with respect to $x(t)$.

Assumption 4: The Euclidean norm of the state vector $\|x(t)\|$ is upper bounded by a constant [36], [44], [45]; i.e., $x(t) \in \mathcal{D}$ where $\mathcal{D} \subset \mathbb{R}^n$ is a compact domain. The SDC parametrization $A(\varrho,x,t)$ and $C(\eta,x,t)$ is such that the following inequalities hold [44]:

$$\|\Delta_1(\varrho,x,t)\| \leq \delta_1, \ \|\Delta_2(\eta,x,t)\| \leq \delta_2, \ \underline{\delta}_3 \leq \|C(\eta,x,t)\| \leq \bar{\delta}_3 \tag{30}$$

where δ_1, δ_2, $\underline{\delta}_3$ and $\bar{\delta}_3$ are positive scalars.

Remark 3: Assumption 4 is satisfied by many engineering applications, e.g., pose estimation of a robot moving on earth's surface, state estimation for aircraft guidance and control, etc. Assumption 4 does not imply any assumption on the stability of the system (20) or the incremental stability of the observer design. It is only assumed that the trajectories of the original system remain bounded within an arbitrarily large compact set.

Example 1: We illustrate how the parameters ϱ_i can be used to minimize the uncertainty in the SDC parametrization and avoid loss of observability of the pair (A, C). Consider a nonlinear function $f(x) = (x_1 x_2, -x_2)^T$, where $x = (x_1, x_2)^T \in \mathcal{D} \subset \mathbb{R}^2$, and an observation matrix $C = (1, 0)$. Let two parametrizations be selected as $A_1 = \begin{pmatrix} 0 & x_1 \\ 0 & -1 \end{pmatrix}$, $A_2 = \begin{pmatrix} x_2/2 & x_1/2 \\ 0 & -1 \end{pmatrix}$. Consider a convex combination $f(x) = \varrho_1 A_1 x + \varrho_2 A_2 x$ such that $\varrho_1 + \varrho_2 = 1$. The corresponding matrix $\triangle_1(\varrho_i, x)$ is $\Delta_1 = \begin{pmatrix} (\varrho_2 x_1/2) + \varrho_1 x_2 & \varrho_2(x_1/2) \\ 0 & 0 \end{pmatrix}$, whose Frobenius norm is bounded inside the compact domain $\mathcal{D}$. The parameters ϱ can be used to preserve the observability of the pair $(\varrho_1 A_1 + \varrho_2 A_2, C)$. The pair $(\varrho_1 A_1 + \varrho_2 A_2, C)$ is unobservable if $\varrho_1 x_1 + (\varrho_2 x_2/2) = 0$. To avoid the loss of observability, a nonzero ϱ_1 is selected for $x_2 = 0$, a nonzero ϱ_2 is selected for $x_1 = 0$, and ϱ_1, ϱ_2 are selected such that $\varrho_1 x_1 + (\varrho_2 x_2/2) \neq 0$ for $x_1 \neq 0$, and $x_2 \neq 0$. Note that the pairs (A_1, C) and (A_2, C) are not observable for $x_1 = 0$, and $x_2 = 0$, respectively. The problem of loss of observability of an individual pair can be avoided by suitably choosing the coefficients ϱ. In Section V-B, a linear matrix inequality (LMI) is formulated which includes constraints on ϱ and η to avoid loss of observability of the pair (A, C).

V. OBSERVER DESIGN AND STABILITY ANALYSIS

In this section, an observer is designed for the stochastic system in (20) and (23) to estimate the state $x(t)$. The estimate

is denoted by $\hat{x}(t) \in \mathbb{R}^n$. A stochastic observer for the system in (23) is designed as

$$d\hat{x} = A(\varrho,\hat{x},t)\hat{x}dt + K(\hat{x},t)\left(y - C(\eta,\hat{x},t)\hat{x}\right)dt \tag{31}$$

which can be written in the following form using (24):

$$d\hat{x} = [A(\varrho,\hat{x},t)\hat{x} + K(\hat{x},t)\left(C(\eta,x,t)x - C(\eta,\hat{x},t)\hat{x}\right)]\,dt \\ + K(\hat{x},t)D(x,t)dW_2 \quad (32)$$

where dW_2 is defined below (21). The observer gain $K(\hat{x},t)$ is given by

$$K(\hat{x},t) = P(\hat{x},t)C^T(\eta,\hat{x},t)R^{-1}(\hat{x},t) \tag{33}$$

where $R(\hat{x},t) = D(\hat{x},t)D^T(\hat{x},t)$ is a positive definite approximation of the measurement noise covariance matrix and the positive definite (PD) symmetric matrix $P(\hat{x},t)$ is a solution to

$$\begin{aligned} dP(\hat{x},t) = &\left(A(\varrho,\hat{x},t)P(\hat{x},t) + P(\hat{x},t)A^T(\varrho,\hat{x},t) \right. \\ &+ 2\alpha P(\hat{x},t) - P(\hat{x},t) \\ &\times \left(-2\kappa I + C^T(\eta,\hat{x},t)R^{-1}(\hat{x},t)C(\eta,\hat{x},t) \right) \\ &\left. \times P(\hat{x},t) \right) dt \end{aligned} \tag{34}$$

where $\alpha > 0$ and $\kappa > 0$. The matrices $A(\varrho,\hat{x},t)$ and $C(\eta,\hat{x},t)$ are obtained via SDC parametrizations in (26) and (27). Equation (34) can be integrated with a positive definite symmetric initial condition $P(0) > 0$.

If the matrices $A(\varrho,\hat{x},t)$ and $C(\eta,\hat{x},t)$ are not state-dependent, then the exact solution of the differential Riccati equation can be derived, which gives an optimal gain of the Kalman-Bucy filter with $\kappa = 0$.

Assumption 5: There exist two time-varying scalar functions $p_u(t)$ and $p_l(t)$ such that the positive definite solution $P(\hat{x},t)$ of the differential Riccati (34) satisfies the bound

$$p_l(t)I \leq P^{-1}(\hat{x},t) \leq p_u(t)I, \ \forall t \geq 0. \tag{35}$$

The time-varying bounds in (35) can be replaced by constants using $\bar{p} \triangleq \sup_t p_u(t)$ and $\underline{p} \triangleq \inf_t p_l(t)$.

Remark 4: If the pair $(A(\varrho,x,t), C(\eta,x,t))$ is uniformly observable by Assumption 3, the solution to (34) satisfies the bound in (35) (cf. [6], [43], [46, Theorem 7], [47, Lemma 2]).

A. Observer Stability by Contraction Analysis

In this section, stochastic incremental stability of the observer is studied using the results of Lemma 2. The analysis is performed in two steps. First, global exponential convergence (contraction) of noise-free trajectories of the observer in (31) towards noise-free trajectories of the original system (20) is proved in Theorem 1 using partial contraction theory [48]. Second, the bound on the mean-squared distance between the trajectories of the original system with process noise and the trajectories of the observer with measurement noise is computed in Theorem 2.

Notice that noise-free and disturbance-free trajectories of the systems (23) and (31) can be represented by the following

virtual system of $q \in \mathbb{R}^n$:

$$\dot{q} = f_v(q, t) = A(\varrho, q, t)q + K(\hat{x}, t)\left(C(\eta, x, t)x - C(\eta, q, t)q\right). \quad (36)$$

For $q = x$, (36) reduces to a noise-free and disturbance-free version of (20) and for $q = \hat{x}$, (36) yields noise-free version of (31). Hence, $q = x$ and $q = \hat{x}$ are particular solutions of the virtual system (36). Consider a stochastic system

$$dq = f_v\left(q(\mu, t), t\right)dt + B_q\left(q(\mu, t), t\right)dW \quad (37)$$

such that $q(0, t) = x$ and $q(1, t) = \hat{x}$, and $B_q(0, t) = B(x, t)$ and $B_q(1, t) = K(\hat{x}, t)D(x, t)$. For the development of the results in Theorems 1–2, the deterministic disturbance term $d(x, t)$ in (20) is not considered in (37). Theorem 3 shows that the $\mathcal{L}_2$ norm of the estimation error is bounded in the presence of noise and bounded deterministic disturbance $d(x, t)$.

Theorem 1: (Deterministic Stability) Under Assumptions 3–5, the virtual system (36) is contracting if there exists a uniformly positive metric $P(q, t)$ that satisfies

$$\begin{aligned}
\dot{P}(q, t) &= A(\varrho, q, t)P(q, t) + P(q, t)A^T(\varrho, q, t) \\
&\quad + 2\alpha P(q, t) - P(q, t) \\
&\quad \times \left(-2\kappa I + C^T(\eta, q, t)R^{-1}(q, t)C(\eta, q, t)\right)P(q, t) \\
&\quad + (K(\hat{x}, t) - K(q, t))R(q, t)(K(\hat{x}, t) - K(q, t))^T.
\end{aligned}$$
$$(38)$$

Proof: We show that the virtual auxiliary system (36) is contracting. Note again that $\hat{x}(t)$ of (31) without noise and $x(t)$ of (20) without noise and disturbance are particular solutions of (36). Using (28) and (29), the virtual dynamics of (36) can be expressed as

$$\delta\dot{q} = (A(\varrho, q, t) - K(\hat{x}, t)C(\eta, q, t))\delta q + \phi(\varrho, \eta, q, \hat{x}, t)\delta q \quad (39)$$

where $\phi(\varrho, \eta, q, \hat{x}, t) = \Delta_1(\varrho, q, t) - K(\hat{x}, t)\Delta_2(\eta, q, t)$. In the virtual dynamics (39), the term $\phi(\varrho, \eta, q, \hat{x}, t)$ can be seen as an uncertainty term, which is bounded due to Assumptions 3–5. To analyze the contraction of the infinitesimal virtual displacement vector δq, consider the rate of change of the squared length in the metric $P^{-1}(q, t)$

$$\begin{aligned}
&\frac{d}{dt}\left(\delta q^T P^{-1}(q, t)\delta q\right) \\
&= \delta q^T P^{-1}(q, t) \\
&\quad \times \Big[A(\varrho, q, t)P(q, t) + P(q, t)A^T(\varrho, q, t) - \dot{P}(q, t) \\
&\quad + (K(\hat{x}, t) - K(q, t))R(q, t)(K(\hat{x}, t) - K(q, t))^T \\
&\quad - K(\hat{x}, t)R(q, t)K^T(\hat{x}, t) - K(q, t)R(q, t)K^T(q, t) \\
&\quad + \phi P(q, t) + P(q, t)\phi^T\Big]P^{-1}(q, t)\delta q \quad (40)
\end{aligned}$$

where $\dot{P}^{-1} = -P^{-1}\dot{P}P^{-1}$, and $-K(\hat{x}, t)R(q, t)K^T(\hat{x}, t) - K(q, t)R(q, t)K^T(q, t)$ is added and subtracted. Using the

bounds $\underline{r}I \geq R \geq \bar{r}I$, and $-P^{-1}(q, t)K(\hat{x}, t)R(q, t)K^T(\hat{x}, t)P^{-1}(q, t) \leq -(\underline{p}^2\underline{\delta}_3^2\bar{r}/\bar{p}^2\underline{r}^2)I = -2\kappa_2 I$, and substituting (33), and (38), (40) can be expressed as

$$\begin{aligned}
\frac{d}{dt}\left(\delta q^T P^{-1}(q, t)\delta q\right) &\leq \delta q^T \Big(-2\alpha P^{-1}(q, t) - 2\kappa I - 2\kappa_2 I \\
&\quad + P^{-1}(q, t)\phi + \phi^T P^{-1}(q, t)\Big)\delta q. \quad (41)
\end{aligned}$$

As stated previously, $\phi(\varrho, \eta, q, \hat{x}, t)$ is norm-bounded according to Assumptions 3–5. The following bound can be established using Assumptions 4 and 5:

$$\left\| P^{-1}(q, t)\phi + \phi^T P^{-1}(q, t) \right\| \leq 2\kappa_1 \quad (42)$$

where $\kappa_1 \triangleq \bar{p}\delta_1 + (\bar{p}/\bar{r}\underline{p})\bar{\delta}_3\delta_2$. Using (42), for sufficiently large α and κ, (41) satisfies

$$\frac{d}{dt}\left(\delta q^T P^{-1}(q, t)\delta q\right) \leq -2\alpha_1 \delta q^T P^{-1}(q, t)\delta q \quad (43)$$

where $\alpha > \alpha_1 > 0$, and

$$\kappa_1 - \kappa_2 \leq \kappa + (\alpha - \alpha_1)\underline{p}. \quad (44)$$

Using (43), the bound on the squared length can be established

$$\delta q^T P^{-1}(q, t)\delta q \leq \delta q^T(0)P^{-1}(q(0), 0)\delta q(0)e^{-2\alpha_1 t} \quad (45)$$

which reduces to

$$\|\delta q(t)\| \leq \sqrt{\frac{\bar{p}}{\underline{p}}}\|\delta q(0)\|e^{-\alpha_1 t}. \quad (46)$$

Using the equivalence of the norm of the distance between x and $\hat{x}$, $\|x - \hat{x}\| = \|\int_x^{\hat{x}}\delta q\| \leq \int_x^{\hat{x}}\|\delta q\| = \int_0^1\|\partial q(\mu, t)/\partial\mu\|d\mu$ and (46), it can be seen that the system (36) is globally exponentially contracting in the absence of noise. Thus, $\hat{x}(t)$ converges to $x(t)$ exponentially and globally. ∎

Remark 5: Note that (38) is evaluated at an auxiliary variable q. For the observer implementation, (34) is used. Theorem 1 shows that any trajectory of q converges; i.e., trajectories of q and $\hat{x}(t)$ converge to each other and hence the noise-free solution of (34) converges towards the solution of (38).

Remark 6: If the Jacobian of the nonlinear function $h(x, t)$ is used instead of the SDC parametrization, the bound κ_1 becomes smaller because $\Delta_2(\eta, q, t) = 0$.

Assumption 6: Let $\bar{p}_x = \sup\limits_{t \geq 0, i, j}\|(P_{ij}^{-1})_q\|$, $\bar{p}_{x^2} = \sup\limits_{t \geq 0, i, j}\|\partial^2(P_{ij}^{-1})/\partial q_i \partial q_j\|$, $\|B(x, t)\|_F \leq \bar{b}$.

Remark 7: The state estimates can be bounded within the compact domain using tools, such as saturation or projection algorithms (cf. [13], [49], [50]).

Theorem 2: (Stochastic Stability) If Assumptions 3–6 are satisfied, the mean-squared estimation error of the observer in (31) is exponentially bounded with the bound

$$E\left[\|x - \hat{x}\|^2\right] \leq \frac{1}{\underline{p}}\left(E[V(q(0), \delta q(0), 0)]e^{-2\alpha_3 t} + \frac{\delta_4}{2\alpha_3}\right) \quad (47)$$

where

$$\delta_4 \geq \frac{\bar{p}_x}{\varepsilon_1}\left(\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2}\mathrm{tr}\left(P^2(\hat{x},t)\right)\right) + \bar{p}\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2}\mathrm{tr}\left(P(\hat{x},t)\right) \quad (48)$$

where $\alpha_3 \triangleq \alpha_1 - (\kappa_P/2p)(\varepsilon_1 \bar{p}_x + (\bar{p}_{x^2}/2))$ such that $\alpha_1 > (\kappa_P/2p)(\varepsilon_1 \bar{p}_x + (\bar{p}_{x^2}/2))$, $\exists \varepsilon_1 > 0$, $\alpha_1 > 0$ is defined in (43) and $\kappa_P = \bar{b}^2 + (\bar{\delta}_3^2 \underline{r}/\bar{r}^2)\mathrm{tr}(P^2(\hat{x},t))$.

Proof: We prove this theorem as a special case of Lemma 2. To prove that the mean-squared estimation error $E[\|x - \hat{x}\|^2]$ is exponentially bounded, consider a Lyapunov-like function defined by $V(q, \delta q, t) = \int_0^1 ((\partial q/\partial \mu))^T P^{-1}(q(\mu, t), t)((\partial q/\partial \mu))d\mu$, where $q(\mu = 0) = x$ and $q(\mu = 1) = \hat{x}$. For conciseness of the presentation, we only compute the differential generator of $V(q, \delta q, t)$ by following the development in the proof of Lemma 2 for the system in (37)

$$\mathcal{L}V(q, \delta q, t) = \int_0^1 \left(\frac{\partial q}{\partial \mu}\right)^T \frac{d}{dt} P^{-1}\left(q(\mu, t), t\right)\left(\frac{\partial q}{\partial \mu}\right) d\mu$$

$$+ \int_0^1 \left(\frac{\partial q}{\partial \mu}\right)^T \left(P^{-1}\frac{\partial f_v}{\partial q} + \frac{\partial f_v}{\partial q}^T P^{-1}\right)\left(\frac{\partial q}{\partial \mu}\right) d\mu + V_2 \quad (49)$$

where f_v is defined in (36), note that $q = q(\mu, t)$ with $q(0, t) = x$ and $q(1, t) = \hat{x}$, computation of V_2 is shown in Appendix B, and the upper bound of V_2 is given by

$$\bar{V}_2 = \mathrm{tr}\left(B(x,t)^T P^{-1}(x,t)B(x,t)\right.$$
$$\left. + (K(\hat{x},t)D(x,t))^T P^{-1}(\hat{x},t)K(\hat{x},t)D(x,t)\right)$$
$$+ \bar{p}_x\left(\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2}\mathrm{tr}\left(P^2(\hat{x},t)\right)\right)$$
$$\times \left(\int_0^1 \varepsilon_1 \left\|\frac{\partial q}{\partial \mu}\right\|^2 d\mu + \frac{1}{\varepsilon_1}\right)$$
$$+ \frac{1}{2}\bar{p}_{x^2}\left(\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2}\mathrm{tr}\left(P^2(\hat{x},t)\right)\right)\int_0^1 \left\|\frac{\partial q}{\partial \mu}\right\|^2 d\mu. \quad (50)$$

The derivative of f_v with respect to q is given by

$$\frac{\partial f_v}{\partial q} = A(\varrho, q, t) - P(\hat{x}, t)C^T(\eta, \hat{x}, t)R^{-1}(\hat{x}, t)$$
$$\times C(\eta, q, t) + \phi(\varrho, \eta, q, t). \quad (51)$$

Substituting (51) into (49), and using $-P^{-1}(q, t)K(\hat{x}, t)$ $R(q, t)K^T(\hat{x}, t)P^{-1}(q, t) \leq -2\kappa_2 I$ yields

$$\mathcal{L}V \leq \int_0^1 \left(\frac{\partial q}{\partial \mu}\right)^T \frac{d}{dt} P^{-1}(q, t)\left(\frac{\partial q}{\partial \mu}\right) d\mu + \int_0^1 \left(\frac{\partial q}{\partial \mu}\right)^T P^{-1}(q, t)$$
$$\times \left[A(\varrho, q, t)P(q, t) + P(q, t)A^T(\varrho, q, t)\right.$$
$$- P(q, t)C^T(\eta, q, t)R^{-1}(q, t)C(\eta, q, t)P(q, t)$$
$$+ (K(\hat{x},t) - K(q,t))R(q,t)(K(\hat{x},t) - K(q,t))^T$$
$$\left. - 2\kappa_2 I + \phi P(q, t) + P(q, t)\phi^T\right]$$
$$\times P^{-1}(q, t)\left(\frac{\partial q}{\partial \mu}\right) d\mu + \bar{V}_2. \quad (52)$$

Using (42), $(d/dt)P^{-1} = -P^{-1}((d/dt)P)P^{-1}$, and the Riccati (38), and the results of Theorem 1, the differential generator satisfies the following inequality:

$$\mathcal{L}V \leq -2\alpha_1 \int_0^1 \left(\frac{\partial q}{\partial \mu}\right)^T P^{-1}(q, t)\left(\frac{\partial q}{\partial \mu}\right) d\mu + \bar{V}_2$$

$$= -2\alpha_1 V(q, t) + \bar{V}_2 \quad (53)$$

where α_1 is defined below (43). Using Assumption 6, and derivations in (11)–(16), the following upper bound is obtained:

$$\mathcal{L}V(q, \delta q, t) \leq -2\alpha_3 V(q, \delta q, t) + \delta_4 \quad (54)$$

where δ_4 is defined in (48), and α_3 is defined below (48). The properties of trace operator: $\mathrm{tr}(XY) \leq \mathrm{tr}(X)\mathrm{tr}(Y)$, $\mathrm{tr}(XY) \leq \|Y\|\mathrm{tr}(X)$, and $\sqrt{\mathrm{tr}(Y^T Y)} = \|Y\|_F$, for positive semi-definite matrices X and Y, and $\mathrm{tr}((K(\hat{x},t)D(x,t))^T P^{-1}(\hat{x},t)K(\hat{x},t)D(x,t)) \leq (\bar{\delta}_3^2 r/\bar{r}^2)\mathrm{tr}(P(\hat{x},t))$ are used to derive the δ_4 bound. Using the development in the proof of Lemma 2, the bound in (47) is obtained. ∎

Remark 8: In general, the process and measurement noise terms do not vanish. Hence, the constant δ_4, which is a function of the process and measurement noise intensities, will not be zero. This implies that $\mathcal{L}V(q(\mu, t), \delta q(t), t)$ may not always be non-positive and $V(q(\mu, t), \delta q(t), t)$ may sometimes be increasing. Thus, $E[V(q(\mu, t), \delta q(t), t)] \leq V(q(\mu, s), \delta q(s), s) \, \forall 0 \leq s \leq t < \infty$, may not always be true; i.e., $V(q(\mu, t), \delta q(t), t)$ may not be a supermartingale. The supermartingale inequality [3], [41] cannot be used to prove stability in an almost-sure sense [25].

Remark 9: Theorems 1–2 can be used to analyze stability of extended Kalman filter (EKF) using $\kappa = 0$, $\alpha = 0$ in (34).

The observer (31)–(34) is shown to be robust against the disturbances and noise in the sense of finite expected value of the $\mathcal{L}_2$ norm of the estimation error with respect to the disturbances and noise acting on the system. The $\mathcal{L}_2$ norm bound is derived for the estimation error of a generalized state $g(t) \triangleq L(t)x(t)$, where $L(t) \in \mathbb{R}^{m \times n}$ satisfies $L^T(t)L(t) \leq \bar{\ell}I$, where $\bar{\ell}$ is a positive constant.

Corollary 1: ($\mathcal{L}_2$ robustness) If Assumptions 3–6 are satisfied, the observer in (31)–(34) is robust against the external disturbances and satisfies the following $\mathcal{L}_2$ norm bound on the estimation error:

$$E_{q_0}\left[\int_0^t \|g(\tau) - \hat{g}(\tau)\|^2 d\tau\right] \leq \frac{\bar{\ell}}{\xi_1}\|x(0) - \hat{x}(0)\|_{P^{-1}(0)}^2$$
$$+ \frac{\bar{\ell}}{\xi_1}E_{q_0}\left[\int_0^t \left(\xi_2\|d(x, \tau)\|^2 + \delta_4\right) d\tau\right] \quad (55)$$

where δ_4 is defined in (48), $\|\cdot\|_{P^{-1}(0)}^2$ is the Euclidean vector norm square with respect to weight $P^{-1}(0)$, $\xi_1 = (1 - \theta)2\alpha_3 p$, $\xi_2 = \bar{p}/\varepsilon_2$, $\exists \varepsilon_2 > 0$, $q_0 = q(\mu = 0, t = 0)$, and $0 < \theta = (\varepsilon_2 \bar{p})/(2\alpha_3 p) < 1$.

Proof: See Appendix C. ∎

B. Computation of P and LMI Formulation

To compute the estimator gain, a PD solution $P(\hat{x}, t)$ of (34) is required. In this section, algorithms are presented to compute the matrix $P(\hat{x}, t)$.

1) Main Algorithm (LMI-SDARE): In this section, the optimal observer gain design algorithm, called the linear matrix inequality state-dependent algebraic Riccati equation (LMI-SDARE), is presented. In the LMI-SDARE algorithm, $\dot{P}(\hat{x}, t)$ is approximated to its steady state value; i.e., zero [51]. The differential Riccati (34) can be converted into the following algebraic Riccati inequality (cf. [23, pp. 114–115])

$$A(\varrho, \hat{x}, t)P + PA^T(\varrho, \hat{x}, t) + 2\alpha P - PC^T(\eta, \hat{x}, t)\times$$
$$R^{-1}(\hat{x}, t)C(\eta, \hat{x}, t)P + 2\kappa PP \leq 0. \quad (56)$$

Proposition 1: The inequality (56) can be converted to the following linear matrix inequality (LMI) by Shor's relaxation [52] in terms of variables $Q = P^{-1}$, $Q_{\varrho_i} = \varrho_i Q$, ϱ_i, $\forall i = \{1, \ldots, s_1\}$, $\eta_i, \forall i = \{1, \ldots, s_2\}$, and $\eta_{l_{ij}}$

$$\begin{bmatrix} \sum\limits_{i=1}^{s_1} Q_{\varrho_i} A_i(\hat{x}, t) + \sum\limits_{i=1}^{s_1} A_i^T(\hat{x}, t)Q_{\varrho_i} + 2\alpha Q - \Upsilon & I \\ I & -\frac{1}{2\kappa}I \end{bmatrix} \leq 0 \quad (57)$$

where $\Upsilon = \sum_{i,j=1}^{s_2} \eta_{l_{ij}} C_i^T(\hat{x}, t)R^{-1}(\hat{x}, t)C_j(\hat{x}, t)$, and

$$Q > 0, \sum_{i=1}^{s_1} Q_{\varrho_i} A_i(\hat{x}, t)\hat{x} = Qf(\hat{x}), \quad (58)$$

$$\text{sym}\begin{bmatrix} I & Q \\ \varrho_i I & Q_{\varrho_i} \end{bmatrix} \geq 0, \sum_{i=1}^{s_1} \varrho_i = 1, \varrho_i \in [0, 1], \quad (59)$$

$$\sum_{i=1}^{s_2} \eta_j = 1, \eta_j \in [0, 1], oc_k(\varrho, \eta, \hat{x}) < 0, \forall k = 1, \ldots, n_o$$
$$(60)$$

$$W_i = \begin{bmatrix} 1 & \eta_i \\ \eta_i & \eta_{l_{ii}} \end{bmatrix} \geq 0, \eta_{l_{ij}} \in [0, 1] \quad (61)$$

$$\sum_{i=1}^{s_2} \eta_{l_{ii}} + \sum_{i,j=1, i\neq j}^{s_2} \eta_{l_{ij}} = 1 \quad (62)$$

where $\text{sym}(\cdot)$ is a symmetric part of a matrix, $oc_j(\varrho, \eta, \hat{x}) < 0$, $\forall j = 1, \ldots, n_o$ denotes n_o number of convex constraints to maintain the observability of the pair (A, C). Note that $oc_j(\varrho, \eta, \hat{x}) < 0$ might impose $\varrho_j \geq \chi_j > 0$, and $\eta_j \geq \psi_j > 0$ for some js, where χ_j and ψ_j are small constants.

Proof: First multiplying (56) by $Q = P^{-1}$ from left and right side yields

$$QA(\varrho, \hat{x}, t) + A^T(\varrho, \hat{x}, t)Q + 2\alpha Q - C^T(\eta, \hat{x}, t)$$
$$\times R^{-1}(\hat{x}, t)C(\eta, \hat{x}, t) + 2\kappa I \leq 0. \quad (63)$$

By applying the Schur complement lemma to (56), we obtain the bilinear matrix inequality (BMI) [52], [53]

$$\begin{bmatrix} \sum\limits_{i=1}^{s_1} Q_{\varrho_i} A_i(\hat{x}, t) + \sum\limits_{i=1}^{s_1} A_i^T(\hat{x}, t)Q_{\varrho_i} + 2\alpha Q - \Upsilon_1 & I \\ I & -\frac{1}{2\kappa}I \end{bmatrix} \leq 0 \quad (64)$$

with the constraints in (58) where Υ_1 is given by

$$\Upsilon_1 = \sum_{i,j=1}^{s_2} \eta_i \eta_j C_i^T(\hat{x}, t)R^{-1}(\hat{x}, t)C_j(\hat{x}, t). \quad (65)$$

The multiplication of components of η in Υ_1 makes (64) a BMI. To convert the BMI to an LMI, new lifting variables $\eta_{l_{ij}}$ are defined. Using the lifting variables and the Shor's relaxation [52], the BMI in (64) is converted into an LMI in terms of the variables Q, Q_{ϱ_i}, η_i, and $\eta_{l_{ij}}$, by writing Υ_1 as $\Upsilon = \sum_{i,j=1}^{s_2} \eta_{l_{ij}} C_i^T(\hat{x}, t)R^{-1}(\hat{x}, t)C_j(\hat{x}, t)$ with the constraints (61), (62). The Shor's relaxation is applied for each individual constraint $\eta_{l_{ii}} = \eta_i \eta_i$ $\forall i \in \{1, \ldots, s_2\}$. Note that due to Shor's relaxation the rank 1 constraints on W_is are ignored. For deriving the constraints on the cross terms $\eta_{l_{ij}}$, the equality $(\sum_{i=1}^{s_2} \eta_i)^2 = 1$ is used, which can be written as (62) in terms of $\eta_{l_{ij}}$ by using $\eta_i \eta_j = \eta_{l_{ij}}$. A similar constraint is recently derived in [54]. To take care of the $\varrho_i Q = Q_{\varrho_i}$ constraint, new constraints on Q and Q_{ϱ_i} are given in (59).

Additional constraints in the form $oc_k(\varrho, \eta, \hat{x}) < 0$ can be formulated so that the pair (A, C) is observable. These constraints can be obtained by symbolically computing the observability matrix and formulating state-dependent constraints for which the observability matrix is full rank. The observability constraints make sure that Assumption 3 is satisfied. See Example 1 for an example of the observability constraint. If the observability constraints are not convex, relaxation methods can be used to approximate the observability constraints to convex constraints [55]. The formulation in (57), (58) is useful for implementation purposes when there are multiple $A_i(\hat{x}, t)$. If the observability constraint $oc_k(\cdot)$ is an affine function of ϱ only then those terms can be added as $\sum_{i=1}^{s_1} oc_{k_i}(\hat{x})\varrho_i Q < 0$, which is equivalent to $\sum_{i=1}^{s_1} oc_{k_i}(\hat{x})Q_{\varrho_i} < 0$. If $oc_k(\cdot)$ is an affine function of η only then those terms can be added as $\sum_{j=1}^{s_2} oc_{k_j}(\hat{x})\eta_j < 0$. ∎

The solution to the LMI can be used to minimize the mean-squared bound $\delta_4/2\underline{p}\alpha_3$. Towards this goal, three conditions are derived and briefly explained. First, based on (48), minimizing $\text{tr}(P^2(\hat{x}))$ minimizes δ_4 and $1/2\underline{p}\alpha_3 = 1/(2\alpha_1\underline{p} - \kappa_P(\varepsilon_1\bar{p}_x + (\bar{p}_{x^2}/2)))$, and minimizing and $\text{tr}(P(\hat{x}))$ minimizes δ_4. The function $\text{tr}(P^2(\hat{x}))$ is not a convex function of $P(\hat{x})$, but we can minimize a convex upper bound $\text{tr}(P(\hat{x}))^2$. Second, the constraint (44) should be satisfied for deterministic contraction. For given α, κ_1, and κ_2 maximizing $\kappa + \alpha\lambda_{\min}(Q)$ maximizes $\alpha_1\lambda_{\min}(Q)$, which minimizes $1/(2\alpha_1\underline{p} - \kappa_P(\varepsilon_1\bar{p}_x + (\bar{p}_{x^2}/2)))$. Third, minimizing $\lambda_{\max}(Q)$ reduces δ_4. A convex objective function summarizing all three conditions can be formulated in terms of Q and κ as[1]

$$\min \left(\Lambda_1 \text{tr}(Q^{-1})^2 - \Lambda_2\kappa - \Lambda_3\alpha\lambda_{\min}(Q) + \Lambda_4\lambda_{\max}(Q) \right)$$
subject to (57)–(62) $\quad (66)$

where Λ_1, Λ_2, Λ_3 and Λ_4 are the normalized weight parameters selected by the user. The decision variables for the optimization are Q, Q_{ϱ_i}, ϱ_i, $\eta_{l_{ij}}$, η_i and κ. The variable α is usually selected by the user.

---

[1]This cost function is one of many choices a user can select.

In addition to the constraints in (66), an inequality constraint (44) can be included in the LMI formulation as $\kappa_1 - \kappa_2 \leq \kappa + \alpha\lambda_{\min}(Q) - \alpha_2$ where α and κ_1 are selected by the user. To make (44) a linear constraint, a new variable $\alpha_2 = \alpha_1\lambda_{\min}(Q) > 0$ is formed with $(\alpha - \alpha_1)\lambda_{\min}(Q) = \alpha\lambda_{\min}(Q) - \alpha_2 > 0$. The constraint makes sure that the condition (44) is satisfied by the estimator gain.

The LMI contains state-dependent matrices which can be solved at each time instant using the polynomial-time interior point methods [56]. The solution of the LMI optimization problem (66) returns the optimal feasible values of the decision variables to improve the convergence rate or to reduce the mean-squared estimation error with respect to the external disturbances acting on the system. The solution to the LMI problem provides a suboptimal criterion to select the estimator gain.

2) Variations of LMI-SDARE: In this section, two variations of the LMI-SDARE algorithm are presented, which may be computationally less expensive during real-time computation than the LMI-SDARE algorithm at the cost of reduced performance in terms of the mean-squared estimation error.

1) **CSDRE Algorithm**: The differential Riccati equation (34), which uses the convex combination of multiple SDC forms of A and C, can be integrated with some PD initial condition $P(\hat{x}(0),0)$ to compute $P(\hat{x},t)$ matrix at each filter update step. An alternate approach is to solve (56) with equality, which forms an ARE. It is shown in practice that a solution to a state-dependent ARE can be obtained effectively in real-time by using various numerical techniques, such as Schur decomposition of the Hamiltonian matrix [57], the Kleinman algorithm [58], [59], the Chandrasekhar algorithm [60], the square root algorithm, spectral factorization, the information filter algorithm, or the matrix sign function algorithm (see [61]). The parameters ϱ and η can be set to some values *a priori*, which may preserve the observability of the parametrization. The SDC matrices A and C are state-dependent and are evaluated at the state estimate for each filter update step to compute $P(\hat{x},t)$. Since the parameters ϱ and η are selected *a priori*, they may not be optimal in the sense of the best convergence rate or minimizing the trace of the covariance.

2) **Fixed-SDARE Algorithm**: This algorithm pre-computes a constant solution P by solving (66) offline using the constant parametrizations of A and C, which belong to the convex polytope (convex hull of matrices); i.e., $A(\varrho,x,t) \in \text{Co}\{A_1(\varrho),\ldots,A_{s_1}(\varrho)\}$, $C(\eta,x,t) \in \text{Co}\{C_1(\eta),\ldots,C_{s_2}(\eta)\}$, where $A_i(\varrho)$ and $C_i(\eta)$ are vertices of the convex hulls. The convex hull of A and C are computed using the upper and lower bounds of the individual entries inside the domain $\mathcal{D}$. The difference between CSDRE and Fixed-SDARE algorithms is that the CSDRE algorithm uses fixed ϱ and η whereas the Fixed-SDARE algorithm uses fixed (A,C), and ϱ and η are computed using a set of LMIs. A similar approach for the Lipschitz nonlinear systems is developed in [18] by assuming that the Jacobian of the nonlinear function belongs to a convex polytope. The LMI in (66) can be

solved for each vertex of the polytope. The observer gain can be computed using the common feasible solution Q to the LMIs (57)–(62). The convergence of the estimator can be easily shown with the results of Theorem 1 (cf. Section 4.3 of [53]). Although the Fixed-SDARE algorithm permits offline computation of the gains, a new vertex of the convex hull adds another LMI constraint to the feasibility problem (66). Moreover, the convex hull may use conservative bounds of the parametrization [53].

Remark 10: For a standard differential inclusion (DI) method (cf. [23, p. 62]) where a common solution Q is computed that satisfies multiple LMI inequalities formed using A_i and C_i, a common solution Q may not exist if one of the parameterizations (A_i, C_i) is not observable. Similar to the Fixed-SDARE algorithm, a standard DI approach is very conservative since it has to satisfy all the LMIs. In a standard DI method, observability constraints cannot be exploited as can be done in the proposed LMI-SDARE method.

VI. NUMERICAL SIMULATIONS

In this section, the performance of the observer (31)–(34) along with (66) is evaluated using two examples.

A. 2D Robot Pose and Landmark Position Estimation Example

In this example, the filter estimates the robot position and orientation, and 2D landmark positions in the world frame W using the landmark positions measured in the robot body frame B. Let $r(t) \in \mathbb{R}^2$ be robot's position, $\theta(t) \in [0, 2\pi)$ be robot's orientation in W, $v(t) \in \mathbb{R}$ denote linear velocity of the robot in the world frame, and $\omega(t) \in \mathbb{R}$ denote the body angular velocity of the robot. Let the state vector be defined by $x = (x_v^T(t), x_l^T(t))^T$, where $x_v(t) = (r^T, \theta)^T$. The positions for n_l landmarks in the world frame are $x_l(t) = (l_1^T(t),\ldots,l_2^T(t),l_{n_l}^T(t))^T$, where $l_i \in \mathbb{R}^2, \forall i = \{1, 2, \ldots, n_l\}$ is the position of ith landmark in W. The state dynamics are given by $\dot{x}_v = (v\cos(\theta), v\sin(\theta), \omega)^T$, $\dot{l}_i = 0$, $\forall i = \{1, 2, \ldots, n_l\}$ with the measurement model $y_i = \bar{R}^T(\theta)(r - l_i)$, $\forall i = 1, \ldots, n_l$, where $y_i(t) \in \mathbb{R}^2$ denotes the measurement of each landmark in the robot's coordinate frame, and $\bar{R}(\theta)$ is a rotation matrix. The robot motion parameters are selected as $v = 1$ m/s, $\omega = 0.01$ rad/s, and $x_v = (0\ 0\ 0)^T$. The process and measurements noise are zero-mean Gaussian with variance of 0.1 and 1, respectively. The nonlinear state dynamics and the measurement model are parametrized in the form (24). The $A(x)$ and $C(x)$ are given by

$$A = \begin{bmatrix} 0_{2\times2} & a_{\text{param}}^T & 0_{2\times2n} \\ 0_{(2n+1)\times2} & 0_{(2n+1)\times1} & 0_{(2n+1)\times2n} \end{bmatrix},$$

$$C = \begin{bmatrix} -N_1(\theta) & N_2(x_v)-N_2(\theta,l_1) & N_1(\theta) & 0_{2\times2} & 0_{2\times2} \\ -N_1(\theta) & N_2(x_v)-N_2(\theta,l_2) & 0_{2\times2} & N_1(\theta) & 0_{2\times2} \\ -N_1(\theta) & N_2(x_v)-N_2(\theta,l_3) & 0_{2\times2} & 0_{2\times2} & N_1(\theta) \\ -N_1(\theta) & N_2(x_v)-N_2(\theta,l_4) & 0_{2\times2} & 0_{2\times2} & 0_{2\times2} \\ -N_1(\theta) & N_2(x_v)-N_2(\theta,l_5) & 0_{2\times2} & 0_{2\times2} & 0_{2\times2} \end{bmatrix}$$
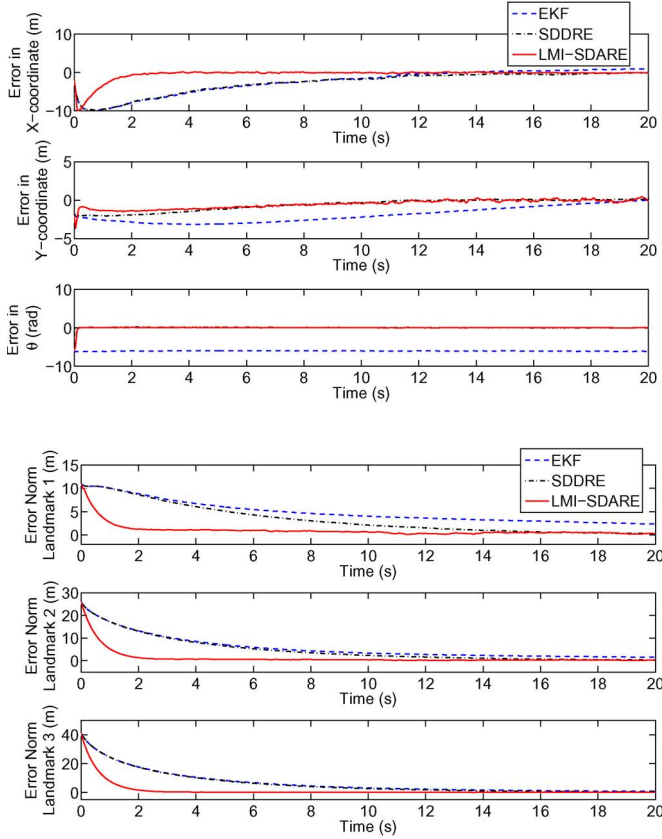
$$(67)$$

Fig. 2. Comparison of the robot pose estimation errors (top plot) and landmark estimation error norms for three landmarks (bottom plot) computed using the extended Kalman filter (dotted blue line), the SDDRE algorithm (dashed green line), and the LMI-SDARE algorithm (solid red line).
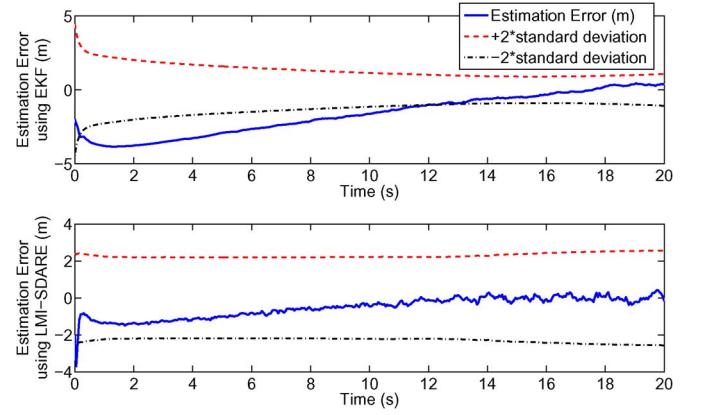

Fig. 3. Comparison of the mean estimation error and estimated $+/-95\%$ confidence interval using the extended Kalman filter (subplot 1) and the LMI-SDARE algorithm (subplot 2).

where $a_{\mathrm{param}} = [v((\cos(\theta)-1)/\theta) \quad v\mathrm{sinc}(\theta)], \quad \mathrm{sinc}(\theta) \triangleq \sin(\theta)/\theta, \quad n=3, \quad N_1(\cdot) \in \mathbb{R}^{2\times 2}$ and $N_2(\cdot) \in \mathbb{R}^{2\times 1}$ are nonlinear functions.

In the first simulation, the proposed observer using the LMI-SDARE algorithm (Section V-B1) is compared with the conventional SDDRE algorithm [21] and the EKF. We use $\hat{x}(0) = (2,2,6,16,2,36,2,56,4)^T$ to test the performance of the observers in the presence of large initial uncertainty. The pair $(A(x), C(\eta, x))$ in (67) satisfies Assumption 3 and Remark 4. Hence, the solutions of the Riccati equations for the EKF and SDDRE remain positive definite and bounded.

For the SDDRE algorithm, the filter parameters are selected as, $R=I$, $\alpha=0.15$, and the initial error covariance $P(0) = [x(0)-\hat{x}(0)][x(0)-\hat{x}(0)]^T$. The matrix $P(\hat{x},t)$ is computed by integrating (34) using a single parametrization of (A,C) and the observer gain is computed using (33). For the EKF noise covariance and process covariance are $R_{\mathrm{EKF}}=I$ and $Q_{\mathrm{EKF}}=0.1I$. For the LMI-SDARE algorithm, a convex combination of two different C matrix parametrizations is used with free parameters η_1 and η_2, represented by $C(\eta,x)=\eta_1 C_1 + \eta_2 C_2$. Using the parametrization of C, the optimization problem in (66) is implemented using the cvx toolbox in Matlab [62], [56].[2] Filter parameters are selected as $R=I$, $\alpha=0.15$, $\kappa_1=0.1$.

---

[2]Since the LMI-SDARE algorithm requires solving the LMI (66) at each time-instant, the sampling rate needs to be carefully chosen.

Since there is only one parametrization of A, Q_{ϱ_i} parameters are not be formed and constraints (58) and (59) are removed from the LMI implementation. The optimized parameters η_1 and η_2 are obtained for $\alpha=1.5$. In Fig. 2, a comparison of the pose estimation errors and a comparison of the estimation error norm of the landmark state using all three filters are presented. From Fig. 2, it is observed that the EKF estimates fail to converge to the *true* value for a large initial uncertainty while the SDDRE and LMI-SDARE algorithms converge to the *true* value; i.e., the estimation error tends nearly to zero with the desired convergence property. The LMI-SDARE algorithm shows better performance than SDDRE and EKF, and provides a more systematic way to choose optimal parameters of the convex hull. In Fig. 3, the mean estimation error and $\pm 95\%$ estimated confidence interval (± 2 standard deviation) for state 1 are plotted. The EKF overestimates its performance in terms of estimated standard deviation; i.e., the estimation errors may lie outside the estimated standard deviation. This causes statistical inconsistency in the EKF predictions and possible divergence as seen from the simulation. The predicted standard deviation of the LMI-SDARE is larger than that of the EKF, but the estimation error mean lies within the predicted standard deviation during the steady state; i.e., the filter is statistically consistent.

In the second simulation, the optimization problem (66) is solved for the LMI-SDARE algorithm by keeping the simulation parameters (i.e., the robot velocity, initial conditions of the estimator, and measurement covariance matrix) the same as the simulation case 1. The following sets of values for $\mathbf{\Lambda} = [\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4]$ are selected, $\mathbf{\Lambda}_a = [0, 0.5, 0.5, 0]$, $\mathbf{\Lambda}_b = [0, 1, 0, 0]$, and $\mathbf{\Lambda}_c = [0.25, 0.25, 0.25, 0.25]$ for the objective function in (66). The multi-objective optimization in (66) is formulated using scalarization for finding Pareto optimal points [55]. In Fig. 4, a comparison of pose estimation errors is shown. From Fig. 4, it is observed that a larger Λ_3 corrresponds to a faster convergence rate, Λ_4 tends to reduce the effects of noise, and a larger κ with a larger Λ_2 corresponds to robustness against noise in the estimation error steady-state response.

Fifty simulations are performed using the EKF, the conventional SDDRE, and the LMI-SDARE. The measurement and process noise, and initial state are selected as Gaussian random variables with zero-mean for noises and the mean initial state
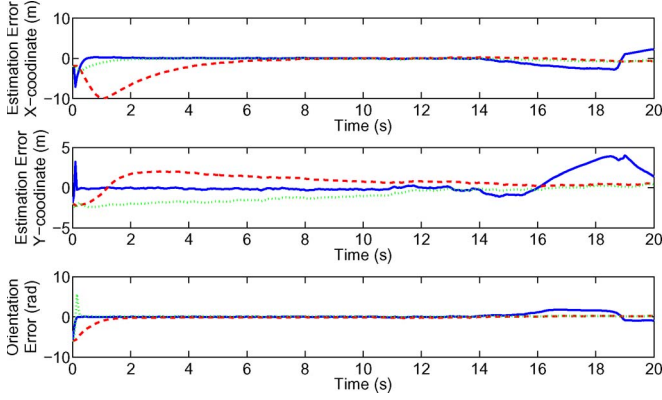
Fig. 4. Comparison of robot pose estimation using the LMI-SDARE estimator for $\mathbf{\Lambda}_a = [0, 0.5, 0.5, 0]$ (solid blue), $\mathbf{\Lambda}_b = [0, 1, 0, 0]$ (dashed red), and $\mathbf{\Lambda}_c = [0.25, 0.25, 0.25, 0.25]$ (dotted green).

TABLE I
COMPARISON OF THE ROOT MEAN SQUARE ERROR (RMSE) AND PEAK ESTIMATION ERROR (PE) FOR THE EKF, THE CONVENTIONAL SDDRE AND THE PROPOSED LMI-SDARE ALGORITHMS AVERAGED OVER 50 MONTE CARLO RUNS

	EKF	Conventional SDDRE	LMI-SDARE
RMSE State 1	0.4058	0.3500	0.2338
PE State 1	0.8475	0.8074	0.3323
RMSE State 2	0.4577	0.5468	0.2868
PE State 2	0.8181	0.9982	0.8118
RMSE State 3	0.1111	0.1124	0.0917
PE State 3	0.0979	0.0751	0.1227

$\hat{x}(0) = (0.5, 0.5, 0.5, 4, 0.5, 9, 0.5, 14, 1)^T$. The results of average root mean-squared errors (RMSE) and average worst case peak errors (PE) for state 1 and state 2 are tabulated in Table I. The LMI-SDARE outperforms the EKF and the conventional SDDRE filter in terms of RMSE and PE as seen from Table I.

B. Lorentz Oscillator

We consider the problem of state estimation for the Lorentz oscillator. The dynamics of the state $x(t) = (x_1, x_2, x_3)^T$ are described by

$$\dot{x}_1 = \sigma_L(x_2 - x_1), \quad \dot{x}_2 = -\rho_L x_1 - x_2 - x_1 x_3,$$
$$\dot{x}_3 = -\beta_L x_3 + x_1 x_2. \qquad (68)$$

The measurement equation is given by $y = Cx + \nu$, where $C = [1, 0, 0]$. The parameters of the simulation are chosen as: $\sigma_L = 10$, $\rho_L = 28$, $\beta_L = 8/3$, $x(0) = (0, 2, 0)^T$, and $\hat{x}(0) = (0, 1.8, 0)^T$. A zero-mean Gaussian white noise with variance of 0.1 is used as the measurement noise. Two different parametrizations A_1 and A_2 are selected and the optimization objective in (66) is used. A comparison of the state estimation errors computed using the LMI-SDARE algorithm and the deterministic observer presented in [14] is shown in Fig. 5. Although the observer in [14] is computationally simpler, the LMI-SDARE observer presented in this paper shows improved performance over the observer in [14] in terms of estimation error. Note that the algorithms in [10], [12] cannot be used for this model because the nonlinearity does not satisfy the required constraints.
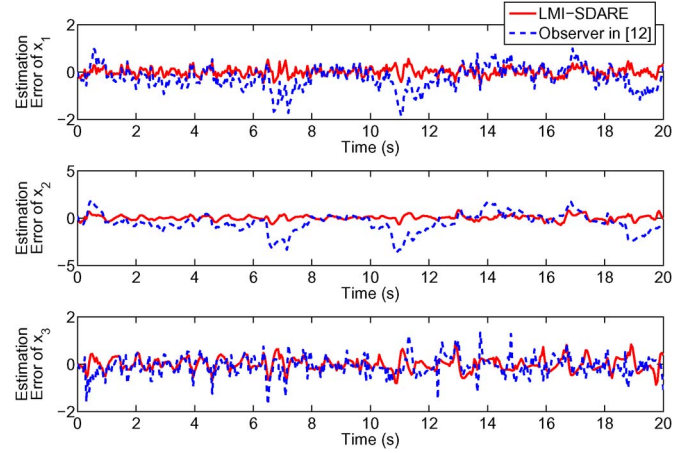

Fig. 5. Comparison of the estimation errors computed using the LMI-SDARE observer and observer proposed in [14].

VII. CONCLUSION

In this paper, a new exponentially converging observer based on a convex combination of multiple SDC parametrizations is presented for a class of Itô stochastic nonlinear systems perturbed by process and measurement noise. Stochastic incremental stability of the observer is studied with respect to a state-dependent metric $M(x, t)$. It is shown that the mean-squared estimation error is exponentially bounded with the bound proportional to the measurement and process noise. The flexibility of non-uniqueness of the SDC form is utilized to obtain the improved convergence rate and disturbance-attenuation property by computing the observer gain via an LMI problem. The observer gain design problem is straightforward and can also handle state constraints related to preserving the observability of the SDC parametrization.

The performance comparison of the observer with the EKF and the conventional SDDRE filter is shown by robot navigation and Lorentz oscillator examples. Statistical inconsistency due to linearization is one of the reasons for filter divergence of the EKF. It is observed from the simulation examples that the LMI-SDARE filter is statistically consistent and yields smaller estimation errors. From a set of multiple numerical simulations, it is concluded that the LMI-SDARE filter outperforms the EKF and conventional SDDRE filters in terms of RMSE. For the LMI-SDARE algorithm, a solution to a SDLMI problem is required at each time instant to compute the gain, which can be efficiently computed using the interior-point methods.

APPENDIX A
GRONWALL-TYPE LEMMA

Lemma 3: Let $g : [0, \infty) \to \mathbb{R}$ be a continuous function, and real numbers C and $\lambda > 0$. If

$$\forall u, t \quad 0 \le u \le t \quad g(t) - g(u) \le \int_u^t (-\lambda g(s) + C) ds \quad (69)$$

then

$$\forall t \ge 0 \quad g(t) \le \frac{C}{\lambda} + \left[g(0) - \frac{C}{\lambda} \right]^+ e^{-\lambda t} \qquad (70)$$

where $[\cdot]^+ = max(0, \cdot)$.

Proof: See [63]. ∎

APPENDIX B
COMPUTATION OF DIFFERENTIAL GENERATOR

The derivative of V_2 of the Lyapunov generator of $V(q(\mu,t),\delta q,t)$, defined in the proof of Theorem 2, can be computed as follows:

$$V_2 = \int_0^1 \sum_{i=1}^n \sum_{j=1}^n P_{ij}^{-1} \left(\frac{\partial B_q(x,t)}{\partial \mu} \frac{\partial B_q(x,t)}{\partial \mu}^T \right)_{ij} d\mu$$

$$+ \int_0^1 \left[2 \sum_{i=1}^n \sum_{j=1}^n (P_i^{-1})_{q_j} \left(\frac{\partial q}{\partial \mu} \right) \left(B_q(q,t) \frac{\partial B_q(q,t)}{\partial \mu}^T \right)_{ij} \right] d\mu$$

$$+ \frac{1}{2} \int_0^1 \left(\sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n \sum_{l=1}^n (P_{kl}^{-1}(q(\mu,t),t))_{q_i q_j} \right. \right.$$
$$\left. \left. \times \frac{\partial q_k}{\partial \mu} \frac{\partial q_l}{\partial \mu} \right) (B_q(q,t) B_q(q,t)^T)_{ij} \right) d\mu \tag{71}$$

where $q = q(\mu,t)$ such that $q(0,t) = x$ and $q(1,t) = \hat{x}$. The integrals of (71) are bounded above as follows:

$$\int_0^1 \sum_{i=1}^n \sum_{j=1}^n P_{ij}^{-1} \left(\frac{\partial B_q(q,t)}{\partial \mu} \frac{\partial B_q(q,t)}{\partial \mu}^T \right)_{ij} d\mu$$

$$\leq \operatorname{tr}\left(P^{-1}(x,t) B(x,t) B(x,t)^T \right)$$
$$+ \operatorname{tr}\left(P^{-1}(\hat{x},t) \left(K(\hat{x},t) D(x,t) \right) \left(K(\hat{x},t) D(x,t) \right)^T \right) \tag{72}$$

$$\int_0^1 \left[2 \sum_{i=1}^n \sum_{j=1}^n (P_i^{-1})_{q_j} \left(\frac{\partial q}{\partial \mu} \right) \left(B_q(q,t) \frac{\partial B_q(q,t)}{\partial \mu}^T \right)_{ij} \right] d\mu$$

$$\leq 2\bar{p}_x \left(\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2} \operatorname{tr}\left(P^2(\hat{x},t) \right) \right) \int_0^1 \left\| \frac{\partial q}{\partial \mu} \right\| d\mu$$

$$\leq \bar{p}_x \left(\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2} \operatorname{tr}\left(P^2(\hat{x},t) \right) \right) \left(\int_0^1 \varepsilon_1 \left\| \frac{\partial q}{\partial \mu} \right\|^2 d\mu + \frac{1}{\varepsilon_1} \right) \tag{73}$$

where $2a'b' \leq \varepsilon_1^{-1} a'^2 + \varepsilon_1 b'^2$, for an $\varepsilon_1 > 0$, for scalars a' and b' is used, $\bar{p}_x$ and $\bar{b}$ are defined in Assumption 6

$$\frac{1}{2} \int_0^1 \left(\sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n \sum_{l=1}^n (P_{kl}^{-1}(q,t))_{q_i q_j} \frac{\partial q_k}{\partial \mu} \frac{\partial q_l}{\partial \mu} \right) (B_q B_q^T)_{ij} \right) d\mu$$

$$\leq \frac{1}{2} \bar{p}_{x^2} \left(\bar{b}^2 + \frac{\bar{\delta}_3^2 r}{\bar{r}^2} \operatorname{tr}\left(P^2(\hat{x},t) \right) \right) \int_0^1 \left\| \frac{\partial q}{\partial \mu} \right\|^2 d\mu \tag{74}$$

where $\bar{p}_{x^2}$ is defined in Assumption 6.

APPENDIX C
PROOF OF COROLLARY 1

Proof: The system (20) along with (31) can be written in the form (37)

$$dq = f_v(q(\mu,t),t)\, dt + d_q(q(\mu,t),t)\, dt + B_q(q(\mu,t),t)\, dW \tag{75}$$

such that $q(0,t) = x$ and $q(1,t) = \hat{x}$, $B_q(0,t) = B(x,t)$ and $B_q(1,t) = K(\hat{x},t) D(x,t)$, and $d_q(0,t) = d(x,t)$ and $d_q(1,t) = 0$. Consider the Lyapunov function used in Theorem 2 $V(q,\delta q,t) = \int_0^1 (\partial q/\partial \mu)^T P^{-1}(q,t)(\partial q/\partial \mu)\, d\mu$. Following the development in Theorem 2 and using (53), the differential generator (49) for the system (75) with respect to the Lyapunov function $V(q,\delta q,t)$ is given by

$$\mathcal{L}V \leq -2\alpha_3 \int_0^1 \left(\frac{\partial q}{\partial \mu} \right)^T P^{-1}(q,t) \left(\frac{\partial q}{\partial \mu} \right) d\mu$$

$$+ \int_0^1 \left(\frac{\partial q}{\partial \mu} \right)^T P^{-1} \left(\frac{\partial d_q}{\partial \mu} \right) d\mu$$

$$+ \int_0^1 \left(\frac{\partial d_q}{\partial \mu} \right)^T P^{-1} \left(\frac{\partial q}{\partial \mu} \right) d\mu + \bar{V}_2 \tag{76}$$

where $\bar{V}_2$ is defined in (50). Using the bound

$$\int_0^1 \left(\frac{\partial q}{\partial \mu} \right)^T P^{-1} \left(\frac{\partial d_q}{\partial \mu} \right) + \left(\frac{\partial d_q}{\partial \mu} \right)^T P^{-1} \left(\frac{\partial q}{\partial \mu} \right) d\mu$$

$$\leq \int_0^1 2 \left\| P^{-\frac{1}{2}} \left(\frac{\partial d_q}{\partial \mu} \right) \right\| \left\| P^{-\frac{1}{2}} \left(\frac{\partial q}{\partial \mu} \right) \right\| d\mu$$

$$\leq \int_0^1 \frac{1}{\varepsilon_2} \left\| P^{-\frac{1}{2}} \left(\frac{\partial d_q}{\partial \mu} \right) \right\|^2 d\mu + \int_0^1 \varepsilon_2 \left\| P^{-\frac{1}{2}} \left(\frac{\partial q}{\partial \mu} \right) \right\|^2 d\mu$$

the following upper bound can be obtained:

$$\mathcal{L}V \leq -2\alpha_3 \underline{p} \int_0^1 \left\| \frac{\partial q}{\partial \mu} \right\|^2 d\mu$$

$$+ \frac{1}{\varepsilon_2} d^T P^{-1} d + \varepsilon_2 \bar{p} \int_0^1 \left\| \frac{\partial q}{\partial \mu} \right\|^2 d\mu + \bar{V}_2$$

$$\leq -(1-\theta) 2\alpha_3 \underline{p} \int_0^1 \left\| \frac{\partial q}{\partial \mu} \right\|^2 d\mu$$

$$+ \frac{\bar{p}}{\varepsilon_2} \|d(x,t)\|^2 + \bar{V}_2 \tag{77}$$

where $0 < \theta = (\varepsilon_2 \bar{p} / 2\alpha_3 \underline{p}) < 1$. By using Dynkin's formula [3] and (77)

$$E_{q_0}\left[V(q(t),\delta q,t) \right] - V(q(0),\delta q(0),0)$$

$$\leq E_{q_0}\left[\left(-\int_0^t (1-\theta) 2\alpha_3 \underline{p} \|x(\tau) - \hat{x}(\tau)\|^2 \right. \right.$$

$$\left. \left. + \frac{\bar{p}}{\varepsilon_2} \|d(x,\tau)\|^2 + \delta_4 \right) d\tau \right]. \tag{78}$$

Using $V(q(0), \delta q(0), 0) = \|x(0) - \hat{x}(0)\|^2_{P^{-1}(0)}$ and $E_{q_0}[V(q(t), \delta q, t)] \geq 0$ yield

$$E_{q_0}\left[\int_0^t \|x(\tau) - \hat{x}(\tau)\|^2 d\tau\right] \leq \frac{1}{(1-\theta)2\alpha_3 \underline{p}}$$

$$\times \left(\|x(0) - \hat{x}(0)\|^2_{P^{-1}(0)} + E_{q_0}\left[\int_0^t \left(\frac{\bar{p}}{\varepsilon_2}\|d(x, \tau)\|^2 + \delta_4\right) d\tau\right]\right). \tag{79}$$

Using the inequality $E_{q_0}[\int_0^t \|(g(\tau) - \hat{g}(\tau))\|^2 d\tau] \leq \bar{\ell} E_{q_0}[\int_0^t \|(x(\tau) - \hat{x}(\tau))\|^2 d\tau]$, where $\bar{\ell}$ is a constant, the following inequality can be obtained

$$E_{q_0}\left[\int_0^t \|(g(\tau) - \hat{g}(\tau))\|^2 d\tau\right] \leq \frac{\bar{\ell}}{\xi_1}\left(\|x(0) - \hat{x}(0)\|^2_{P^{-1}(0)}\right.$$

$$\left. + E_{q_0}\left[\int_0^t \left(\xi_2 \|d(x, \tau)\|^2 + \delta_4\right) d\tau\right]\right). \tag{80}$$

∎

ACKNOWLEDGMENT

REFERENCES

[1] A. Davison, I. Reid, N. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 1052–1067, 2007.

[2] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Observability-based rules for designing consistent EKF SLAM estimators," *Int. J. Robot. Res.*, vol. 29, no. 5, pp. 502–528, 2010.

[3] H. Kushner, *Stochastic Stability and Control*. New York: Academic, 1967.

[4] B. F. Spencer and L. A. Bergman, "On the numerical solution of the Fokker-Planck equation for nonlinear stochastic systems," *Nonlin.Dynam.*, vol. 4, no. 4, pp. 357–372, 1993.

[5] M. Idan and J. L. Speyer, "Cauchy estimation for linear scalar systems," *IEEE Trans Autom. Control*, vol. 55, no. 6, pp. 1329–1342, Jun. 2010.

[6] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, "Stochastic stability of the descrete-time extended Kalman filter," *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.

[7] E. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *IEEE Adaptive Sys. for Signal Processing, Comm and Control Symp.*, 2000, pp. 153–158.

[8] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[9] E. Scholte and M. Campbell, "A nonlinear set-membership filter for online applications," *Int. J. Robust Nonlin. Control*, vol. 13, pp. 1337–1358, 2003.

[10] M. Arcak and P. Kokotovic, "Nonlinear observers: A circle criterion design and robustness analysis," *Automatica*, vol. 37, no. 12, pp. 1923–1930, 2001.

[11] M.-S. Chen and C.-C. Chen, "Robust nonlinear observer for Lipschitz nonlinear systems subject to disturbances," *IEEE Trans. Autom. Control*, vol. 52, no. 12, pp. 2365–2369, Dec. 2007.

[12] X. Fan and M. Arcak, "Observer design for systems with multivariable monotone nonlinearities," *Syst. Control Lett.*, vol. 50, no. 4, pp. 319–330, 2003.

[13] J. P. Gauthier, H. Hammouri, and S. Othman, "A simple observer for nonlinear systems: Applications to bioreactors," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 875–880, Jun. 1992.

[14] B. Acikmese and M. Corless, "Observers for systems with nonlinearities satisfying incremental quadratic constraints," *Automatica*, vol. 47, no. 7, pp. 1339–1348, 2011.

[15] I. R. Petersen, "Robust guaranteed cost state estimation for nonlinear systems via an IQC approach," *Syst. Control Lett.*, vol. 58, pp. 865–870, 2009.

[16] A. Germani, C. Manes, and P. Palumbo, "Filtering of stochastic nonlinear differential systems via a Carleman approximation approach," *IEEE Trans. Autom. Control*, vol. 52, no. 11, pp. 2166–2172, Nov. 2007.

[17] S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*. New York: Springer-Verlag, 1999.

[18] A. Zemouche, M. Boutayeb, and G. I. Bara, "Observers for a class of Lipschitz systems with extension to $\mathcal{H}_\infty$ performance," *Syst. Contol Lett.*, vol. 57, pp. 18–27, 2008.

[19] A. Zemouche and M. Boutayeb, "On LMI conditions to design observers for lipschitz nonlinear systems," *Automatica*, vol. 49, no. 2, pp. 585–591, 2013.

[20] H. Fang, R. A. de Callafon, and J. Cortes, "Simultaneous input and state estimation for nonlinear systems with applications to flow field estimation," *Automatica*, vol. 49, no. 9, pp. 2805–2812, 2013.

[21] J. R. Cloutier, "State dependent Riccati equation technique: An overview," in *Proc. Amer. Contr. Conf.*, 1997, pp. 932–936.

[22] T. Cimen, "Survey of state-dependent Riccati equation in nonlinear optimal feedback control synthesis," *J. Guidance, Control, Dynamics*, vol. 35, no. 4, 2012.

[23] S. Boyd, L. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994, ser. Studies in Applied Mathematics.

[24] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[25] Q. Pham, N. Tabareau, and J.-J. E. Slotine, "A contraction theory approach to stochastic incremental stability," *IEEE Trans. Autom. Control*, vol. 54, no. 4, pp. 816–820, Apr. 2009.

[26] A. P. Dani, S.-J. Chung, and S. Hutchinson, "Observer design for stochastic nonlinear systems using contraction analysis," in *Proc. IEEE Conf. Decision Control*, Maui, HI, 2012, pp. 6028–6035.

[27] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 410–421, Mar. 2002.

[28] S.-J. Chung and J.-J. E. Slotine, "Cooperative robot control and concurrent synchronization of Lagrangian systems," *IEEE Trans. Robot.*, vol. 25, no. 3, pp. 686–700, 2009.

[29] S.-J. Chung, S. Bandyopadhyay, I. Chang, and F. Y. Hadaegh, "Phase synchronization control of complex networks of Lagrangian systems on adaptive digraphs," *Automaica*, vol. 49, no. 5, pp. 1148–1161, 2013.

[30] B. P. Demidovich, "Dissipativity of a nonlinear system of differential equations," *Vestnik Moscow State University, Ser. Mat. Mekh., Part I*, no. 6, pp. 19–27, 1961.

[31] B. P. Demidovich, "Dissipativity of a nonlinear system of differential equations," *Vestnik Moscow State University, Ser. Mat. Mekh., Part I*, no. 1, pp. 3–8, 1962.

[32] D. C. Lewis, "Metric properties of differential equations," *Amer. J. Math.*, vol. 71, pp. 294–312, 1949.

[33] A. J. Van der Schaft, " $\mathcal{L}_2$ gain analysis of nonlinear systems and nonlinear state feedback $\mathcal{H}_\infty$ control," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 770–784, Jun. 1992.

[34] T. Basar and P. Bernhard, H_∞ *Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Cambridge, MA: Birkhauser, 1995.

[35] C. P. Mracek, J. R. Cloutier, and C. A. D'Souza, "A new technique for nonlinear estimation," in *Proc. EEE Int. Conf. Control Appl.*, 1996, pp. 338–343.

[36] C. Jaganath, A. Ridley, and D. Bernstein, "A SDRE-based asymptotic observer for nonlinear discrete-time systems," in *Proc. Amer. Control Conf.*, 2005, pp. 3630–3635.

[37] T. Cimen, "Asymptotically optimal nonlinear filtering: Theory and examples with application to target state estimation," in *IFAC World Congress*, 2008, pp. 8611–8617.

[38] Y.-W. Liang and L.-G. Lin, "On factorization of the nonlinear drift term for SDRE approach," in *IFAC World Congr.*, 2011, pp. 9607–9612.

[39] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems*. New York: Springer, 2004, ser. Texts in Applied Mathematics.

[40] H. K. Khalil, *Nonlinear Systems*, 3rd ed: Prentice Hall, 2002.

[41] L. Arnold, *Stochastic Differential Equations: Theory and Applications*: Wiley, 1974.

[42] H. Deng, M. Krstic, and R. Williams, "Stabilization of stochastic nonlinear systems driven by noise of unknown covariance," *IEEE Trans. Autom. Control*, vol. 46, no. 8, pp. 1237–1253, Aug. 2001.

[43] R. Sanfelice and L. Praly, "Convergence of nonlinear observers on $\mathbb{R}^n$ with a riemannian metric (Part I)," *IEEE Trans Autom. Control*, vol. 57, no. 7, pp. 1709–1722, Jul. 2012.

[44] H. T. Banks, B. M. Lewis, and H. T. Tran, "Nonlinear feedback controllers and compensators: A state-dependent Riccati equation approach," *Comput. Optim. and Appl.*, vol. 37, no. 2, pp. 177–218, 2007.

[45] T. Jennawasin, M. Kawanishi, and T. Narikiyo, "Optimal control of polynomial systems with performance bounds: A convex optimization approach," in *Proc. Amer. Control Conf.*, 2011, pp. 281–286.

[46] J. S. Baras, A. Bensoussan, and M. R. James, "Dynamic observers as asymptotic limits of recursive filters: Special cases," *SIAM J. Appl. Math.*, vol. 48, no. 5, pp. 1147–1158, 1988.

[47] A. P. Aguiar and J. Hespanha, "Robust filtering for deterministic systems with implicit outputs," *Syst. Contol Lett.*, vol. 58, no. 4, pp. 263–270, 2009.

[48] W. Wang and J.-J. E. Slotine, "On partial contraction analysis for coupled nonlinear oscillators," *Biolog. Cybern.*, vol. 92, no. 1, 2005.

[49] H. K. Khalil and F. Esfandiari, "Semi-global stabilization of a class of nonlinear systems using output feedback," *IEEE Trans. Autom. Control*, vol. 38, no. 9, pp. 1412–1425, Sep. 1993.

[50] M. Farza, M. M'Saad, and L. Rossignol, "Observer design for a class of MIMO nonlinear systems," *Automaica*, vol. 40, pp. 135–143, 2004.

[51] I. Chang and S.-J. Chung, "Exponential stability region estimates for the state-dependent Riccati equation controllers," in *Proc. IEEE Conf. Decision Control*, 2009, pp. 1974–1979.

[52] S. Boyd and L. Vandenberghe, "Semidefinite programming relaxations of non-convex problems in control and combinatorial optimization," in *Communications, Computation, Control and Signal Processing: A Tribute to Thomas Kailath*.   Boston, MA, USA: Kluwer, 1997, pp. 279–287.

[53] J. G. VanAntwerp and R. D. Braatz, "A tutorial on linear and bilinear matrix inequalities," *J. Process Control*, vol. 10, pp. 363–385, 2000.

[54] J. E. Mitchelle, J.-S. Pang, and B. Yu, "Convex quadratic relaxations of nonconvex quadratically constrained quadratic programs," *Optimiz. Methods and Softw.*, vol. 29, no. 1, pp. 120–136, 2014.

[55] S. Boyd and L. Vandenberghe, *Convex Optimization*.   Cambridge, U.K.: Cambridge University Press, 2004.

[56] J. Mattingely and S. Boyd, "Real-time convex optimization in signal processing," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 50–61, 2010.

[57] A. J. Laub, "A Schur method for solving algebraic Riccati equations," *IEEE Trans Autom. Control*, vol. AC-24, pp. 913–921, 1979.

[58] D. L. Kleinman, "On an iterative technique for Riccati equation computation," *IEEE Trans. Autom. Control*, vol. 13, no. 1, pp. 114–115, 1968.

[59] D. L. Kleinman, "Stabilizing a discrete, constant, linear system with application to iterative methods for solving the Riccati equation," *IEEE Trans. Autom. Control*, vol. AC-19, pp. 252–254, 1974.

[60] D. G. Lainiotis, "Generalized Chandrasekhar algorithms: Time-varying models," *IEEE Trans. Autom. Control*, vol. AC-20, p. 728, 1976.

[61] B. D. O. Anderson and J. B. Moore, *Optimal Control: Linear Quadratic Methods*.   Upper Saddle River, NJ: Prentice-Hall, 1990.

[62] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, Version 1.21, Apr. 2011.

[63] Q.-C. Pham, "A variation of Gronwall's lemma," arXiv:0704.0922v3, Tech. Rep., 2009.

Soon-Jo Chung (M'06–SM'12) received the B.S. degree summa cum laude from Korea Advanced Institute of Science and Technology, Daejeon, in 1998, the S.M. degree in aeronautics and astronautics, and the Sc.D. degree in estimation and control from the Massachusetts Institute of Technology, Cambridge, in 2002 and 2007, respectively.

He is currently an Assistant Professor in the Department of Aerospace Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign (UIUC). He was a Summer Faculty Fellow with the Jet Propulsion Laboratory, working on guidance and control of spacecraft swarms, during the summers of 2010–2014. His research areas include nonlinear control and estimation theory and flight guidance and control, with application to aerial robotics, distributed spacecraft systems, and vision-based navigation.

Dr. Chung is a recipient of the UIUC Engineering Dean's Award for Excellence in Research, the Beckman Fellowship of the UIUC Center for Advanced Study, the U.S. Air Force Office of Scientific Research Young Investigator Award, the National Science Foundation Faculty Early Career Development Award, and two Best Conference Paper awards from the IEEE and the American Institute of Aeronautics and Astronautics.

Seth Hutchinson (F'06) received the Ph.D. degree from Purdue University, West Lafayette, IN, in 1988.

In 1990, he joined the faculty at the University of Illinois at Urbana-Champaign, where he is currently a Professor in the Department of Electrical and Computer Engineering, the Coordinated Science Laboratory, and the Beckman Institute for Advanced Science and Technology. He served as Associate Department Head of ECE from 2001 to 2007. He has published more than 200 papers on the topics of robotics and computer vision, and is coauthor of the books Principles of Robot Motion: Theory, Algorithms, and Implementations (Cambridge, MA: MIT Press, 2005) and Robot Modeling and Control (New York: Wiley, 2005).

Dr. Hutchinson currently serves on the editorial boards of the *International Journal of Robotics Research* and the *Journal of Intelligent Service Robotics*. He was Founding Editor-in-Chief of the IEEE Robotics and Automation Society's Conference Editorial Board (2006–2008) and Editor-in-Chief of the IEEE TRANSACTIONS ON ROBOTICS (2008–2013).

Ashwin P. Dani (M'11) received the M.S. degree in 2008 and the Ph.D. degree in 2011 from the University of Florida (UF), Gainesville.

He is currently an Assistant Professor in the Department of Electrical and Computer Engineering, University of Connecticut, Storrs. He was a post-doctoral research associate at the University of Illinois at Urbana-Champaign from 2011–2013. His research interests are in the areas of nonlinear estimation and control, stochastic nonlinear estimation, vision-based estimation and control, autonomous navigation, and human-robot interaction. He is a co-author of two patents on single-camera ranging of stationary and moving objects.

Dr. Dani is a recipient of the Best Dissertation Award for Dynamics, Systems and Control for the year 2011 from the Department of Mechanical and Aerospace Engineering, UF, and the 2012 Technology Innovator Award from the Office of Research/Office of Licensing Technology of the UF.