

Extended Kalman filter–Koopman operator for tractable stochastic optimal control

Mohammad S. Ramadan, *Member, IEEE*, Mihai Anitescu, *Member, IEEE*

Abstract—The theory of dual control was introduced more than seven decades ago. Although it has provided rich insights to the fields of control, estimation, and system identification, dual control is generally computationally prohibitive. In recent years, however, the use of Koopman operator theory for control applications has been emerging. This paper presents a new reformulation of the stochastic optimal control problem that, employing the Koopman operator, yields a standard LQR problem with the dual control as its solution. We provide a numerical example that demonstrates the effectiveness of the proposed approach compared with certainty equivalence control, when applied to systems with varying observability.

I. INTRODUCTION

Deterministic control theory carries the implicit assumption of complete access to the states, a condition not met in many applications [1]. Stochastic optimal control (SOC) [2], on the other hand, is hindered in practice by its computational complexity, limiting its benefits for the most part to the conceptual level via the introduction of dual control, its solution. Dual control has two key qualitative properties: caution and probing [3]. Caution accounts for uncertainty when achieving safety and improving performance, and it is manifested mostly through constraint tightening in tube-based model predictive control (MPC), or satisfying constraint instances in scenario-based methods [4]. Probing, often in conflict with caution, reflects the online experiment design and information-gathering to regulate the system’s uncertainties.

We target solving SOC of a general differentiable nonlinear system with a quadratic cost. We choose the extended Kalman filter (eKF) as the approximator of the state uncertainty propagation, since (i) the eKF is simple and widespread in navigation [5], robotics [6], power systems [7], and various fields, rendering our developments in this paper immediately beneficial to numerous applications; (ii) the finite-dimensional approximation of the state uncertainty offered by the eKF (mean vector and covariance matrix) is appealing from the computational perspective; and (iii) the quadratic cost function can be written in terms of the state of the eKF.

The intractability of SOC persists even when employing the eKF approximation of uncertainty [3], thus ruling out conventional solution algorithms. Analogous to deterministic optimal control, solution algorithms to SOC have roots in: Dynamic Programming (DP) [8], which suffers from the curse of dimensionality, and its extensions to solve SOC

[9] are generally cumbersome to implement and suffer from scalability limitations; or Pontryagin’s principle, resembled mostly in nonlinear MPC [10] over the stochastic dynamics (uncertainty dynamics), which requires high-dimensional nonconvex programming procedures that can be ill-suited for online computation. Other suboptimal approaches exist [11], typically falling within these two principles, or of limited formulations, e.g. linear Gaussian-parameters systems [12].

The aforementioned nonlinear approaches have in common that their formulations are straightforward, while the challenge is in the solution. In contrast, our approach, which leverages the Koopman operator, switches the challenge from solution finding to problem formulation. That is, upon successful identification of a suitable basis dictionary for the Koopman operator to linearize the entire uncertainty propagation, the solution of SOC is a straightforward linear quadratic regulator (LQR) problem. This discussion of where the challenge / art is located, in the formulation or solution, resembles the discussion in [13, p. 10], explaining the dichotomy between convex and nonconvex programs.

Other works, such as [14], [15], applied the Koopman operator for data-driven Kalman filtering and observer design. Our approach differs from these works in (i) context: ours is non-autonomous, i.e., for control design, hence, the Koopman operator is a parametric map of the control; and (ii) methodology: we solve the SOC problem by linearizing the whole state of the filter, including its uncertainty propagation, which is carried out by the eKF’s Riccati equation.

Different from the typical derivation of SOC [3], [11], which starts from the stochastic DP equation, we use the smoothing theorem [16] in a causality-respecting fashion. We then reformulate the SOC problem so that it is amenable to the application of Koopman operator theory. We discuss the structure of the new formulation and list the assumptions required for the continuity, bilinearity, and robust stability of the resulting new dynamics. The paper concludes with a numerical example with dynamics of varying observability and signal-to-noise ratio.

II. PROBLEM FORMULATION

Consider the dynamic system

$$x_{k+1} = f(x_k, u_k) + w_k, \quad (1a)$$

$$y_k = h(x_k, u_k) + v_k, \quad (1b)$$

where $x_k \in \mathbb{R}^{r_x}$ is the state, $u_k \in \mathbb{R}^{r_u}$ the control input, and $y_k \in \mathbb{R}^{r_y}$ the measured output. The functions f and h are differentiable almost everywhere in their first arguments. The exogenous disturbances $w_k \in \mathbb{R}^{r_x}$, $v_k \in \mathbb{R}^{r_y}$ are

The authors are with the Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439, USA, mramadan@anl.gov, anitescu@mcs.anl.gov.

each independent and identically distributed according to a density function. They are independent from each other and from x_0 , have zero means and have covariances Σ_w and Σ_v , respectively. The initial state x_0 , prior to any measurement (including y_0), has a bounded mean $x_{0|-1}$, a bounded covariance $\mathbf{Cov}(x_0) = \Sigma_{0|-1}$, and a density p_0 .

Assumption 1. (i) Σ_v, Σ_w , and $\Sigma_{0|-1}$ are $\succ 0$ (positive definite) and bounded. (ii)¹ $w_k \in W \subset \mathbb{R}^{r_x}$, $u_k \in \mathbb{U} \subset \mathbb{R}^{r_u}$, such that W and \mathbb{U} yield a compact set \mathbb{X} invariant under (1a), starting from $x_0 \in \mathbb{X}$. (iii) The sequence of events is as follows: at time k , u_k is applied; then y_k becomes available.

The goal is to design a feedback control policy that minimizes the following quadratic cost $J^N(X_0, U_{N-1})$ (arguments suppressed for compactness):

$$J^N = \frac{1}{N} \cdot \mathbb{E} \left\{ x_N^\top Q x_N + \sum_{k=0}^{N-1} x_k^\top Q x_k + u_k^\top R u_k \right\}, \quad (2)$$

where $U_{N-1} = \{u_0, \dots, u_{N-1}\}$, the tuple $X_0 = (x_{0|-1}, \Sigma_{0|-1})$, and $Q \succeq 0$ (positive semi-definite) and $R \succ 0$. The expectation is over the probability space \mathbb{P}_0 , characterizing the random variables (x_0, V_{N-1}, W_{N-1}) , where $V_{N-1} = \{v_0, \dots, v_{N-1}\}$ and $W_{N-1} = \{w_0, \dots, w_{N-1}\}$, and the corresponding product Borel σ -field.

The control input is admissible when it is a causal law. It is restricted to be a function of the accessible data up to the time step of evaluating this law. That is, according to the 3rd point of Assumption 1, $u_k = u_k(Z_{k-1})$, where $Z_{k-1} = \{p_0, Y_{k-1}, U_{k-1}\}$ stores all the past and accessible information up to time k , prior to implementing u_k . Here $Y_{k-1} = \{y_0, \dots, y_{k-1}\}$. For consistency, we denote the prior information by $Z_{-1} = \{p_0\}$.

III. METHODOLOGY

We first show that the cost (2) can be rewritten in terms of the first two moments of x_k .

A. Equivalent description to the cost

Lemma 1. The term $\mathbb{E} x_k^\top Q x_k$ can be expressed by

$$\mathbb{E} \{x_k^\top Q x_k\} = \mathbb{E} \left\{ \text{tr}(Q \Sigma_{k|k-1}) + x_{k|k-1}^\top Q x_{k|k-1} \right\}.$$

The expectation to the left can be expressed in its integral form as $\mathbb{E} \{\cdot\} = \int \cdot p(x_k) dx_k$ with respect to the density function $p(x_k)$ ², while the one to the right can be expressed as $\mathbb{E} \{\cdot\} = \int \cdot p(Y_{k-1}) dY_{k-1}$. The first two conditional moments are

$$\begin{aligned} x_{k|k-1} &= \mathbb{E} \{x_k | Z_{k-1}\}, \\ \Sigma_{k|k-1} &= \mathbb{E} \left\{ [x_k - x_{k|k-1}][x_k - x_{k|k-1}]^\top | Z_{k-1} \right\}. \end{aligned} \quad (3)$$

(The notation $(\cdot)_{k|j}$ strictly follows that in [17, Ch. 3].)

¹This point is typically required in system identification and the Koopman operator approximation methods and can be implied by “nominal robust global asymptotic stability” or “positive invariance with disturbances” as defined in [10, p. 710].

²By the Markov property, $p(x_k) = p(x_0)p(x_1 | x_0) \dots p(x_k, x_{k-1})$, the initial density $p(x_0) = p_0(x_0)$ is given.

Proof. This lemma is a direct consequence of the law of total expectation (smoothing theorem) [16, p. 348], if we condition each additive term in (2) on its corresponding in-time Z_{k-1} (respecting causality). That is,

$$\begin{aligned} \mathbb{E} \{x_k^\top Q x_k\} &= \int x_k^\top Q x_k p(x_k) dx_k, \\ &= \int \int x_k^\top Q x_k p(x_k, Y_{k-1}) dY_{k-1} dx_k, \\ &= \int \left(\int x_k^\top Q x_k p(x_k | Y_{k-1}) dx_k \right) p(Y_{k-1}) dY_{k-1}, \\ &= \mathbb{E} \left\{ \mathbb{E} \{x_k^\top Q x_k | Z_{k-1}\} \right\}. \end{aligned} \quad (4)$$

The state x_k can be decomposed into $x_k = x_{k|k-1} + \tilde{x}_k$, where \tilde{x}_k is the estimation error. The quadratic term $x_k^\top Q x_k = (x_{k|k-1} + \tilde{x}_k)^\top Q (x_{k|k-1} + \tilde{x}_k)$, under the conditional expectation $\mathbb{E} \{\cdot | Z_{k-1}\}$ and, after ignoring the zero mean cross-terms, is equivalent to $x_{k|k-1}^\top Q x_{k|k-1} + \mathbb{E} \{\tilde{x}_k Q \tilde{x}_k^\top | Z_{k-1}\}$. Using the cyclic property of the trace and the linearity of the expectation operator, we have $\mathbb{E} \{\tilde{x}_k Q \tilde{x}_k^\top | Z_{k-1}\} = \text{tr}(Q \Sigma_{k|k-1})$. \square

Proposition 1. The cost function (2) can be represented by

$$\begin{aligned} J^N &= \frac{1}{N} \cdot \mathbb{E} \left[x_{N|N-1}^\top Q x_{N|N-1} + \text{tr}(Q \Sigma_{N|N-1}) \right. \\ &\quad \left. + \sum_{k=0}^{N-1} \left[x_{k|k-1}^\top Q x_{k|k-1} + u_k^\top R u_k + \text{tr}(Q \Sigma_{k|k-1}) \right] \right]. \end{aligned} \quad (5)$$

Proof. Using Lemma 1 on each additive term in (2), we get

$$\begin{aligned} J^N &= \frac{1}{N} \cdot \mathbb{E} \left\{ \mathbb{E} \{x_N^\top Q x_N | Z_{N-1}\} \right. \\ &\quad \left. + \sum_{k=0}^{N-1} \mathbb{E} \{x_k^\top Q x_k + u_k^\top R u_k | Z_{k-1}\} \right\}. \end{aligned}$$

Then we use the derivation subsequent to (4) for each conditional expectation above. \square

B. Evolution of the central moments

In addition to (3), let

$$\begin{aligned} x_{k|k} &= \mathbb{E} \{x_k | Z_{k-1}, y_k\}, \\ \Sigma_{k|k} &= \mathbb{E} \left\{ [x_k - x_{k|k}][x_k - x_{k|k}]^\top | Z_{k-1}, y_k \right\}. \end{aligned}$$

We use the eKF to propagate the central moments $x_{k|k-1}, \Sigma_{k|k-1}$ appearing in (5). At each time step, the eKF is the recursion

$$x_{k+1|k} = f(x_{k|k}, u_k), \quad \Sigma_{k+1|k} = F_k \Sigma_{k|k} F_k^\top + \Sigma_w, \quad (6)$$

where

$$\begin{aligned} x_{k|k} &= x_{k|k-1} + \Omega_k [y_k - h(x_{k|k-1}, u_k)], \\ \Sigma_{k|k} &= [I - \Omega_k H_k] \Sigma_{k|k-1}, \\ \Omega_k &= \Sigma_{k|k-1} H_k^\top [H_k \Sigma_{k|k-1} H_k^\top + \Sigma_v]^{-1}, \\ F_k &= \left. \frac{\partial f(x, u_k)}{\partial x} \right|_{x_{k|k}}, \quad H_k = \left. \frac{\partial h(x, u_k)}{\partial x} \right|_{x_{k|k-1}}. \end{aligned} \quad (7)$$

The eKF extends the Kalman filter to nonlinear systems via employing a first-order approximation formed by the Jacobians F_k and H_k , with Ω_k mimicking the Kalman gain [17], [18]. The recursion is initialized by $x_{0|-1}$, $\Sigma_{0|-1}$, given in Section II.

We note that the cost description in (5) is exact, because the cost is quadratic. However, the central moments $x_{k|k-1}$ and $\Sigma_{k|k-1}$ provided by the eKF above are only approximates and not exact. The reason is that, in general, f and h are nonlinear and the disturbances w_k, v_k are not necessarily Gaussian [17].

Assumption 2. (i) The eKF estimation error $\|x_k - x_{k|k-1}\|_2$ is bounded, and (ii) the covariance $\Sigma_{k|k-1}$ (6) is positive definite and bounded, for all k .³

The variables $x_{k|k-1}, \Sigma_{k|k-1}$ are random since they are functions of the random observation sequence Y_{k-1} (not yet available at $k = 0$). The expectation in (5) averages over Y_{k-1} . Next, we employ an important approximation to omit this expectation.

C. Certainty equivalence of the information state

In the LQG context, the separation principle [1] states that optimal control can be separated into deterministic optimal control and optimal filtering or, equivalently, into LQR and LQE. However, this does not hold for general nonlinear systems. Hence, to omit the expectation in (5), we require some assumptions.

Assumption 3. The measurement correction term $[y_k - h(x_{k|k-1}, u_k)]$, in (7), is independent from Ω_k and is a zero-mean white noise sequence.⁴.

This removes the need for the expectation in (5) and motivates a measurement-free version of the eKF in (6). Let

$$x_{k+1}^p = f(x_k^p, u_k). \quad (8)$$

The state x_k^p is the surrogate of both $x_{k|k-1}$ and $x_{k|k}$ (since the measurement correction is omitted) that evolves solely through prediction.

Remark 1. The surrogate state x_k^p , defined by employing certainty equivalence in the SOC sense (CE-SOC), is deterministic. It will be used offline for learning the Koopman operator and the design of a control law. However, the original eKF state $x_{k|k-1}$ will be used when this control is implemented online in closed-loop.

³The first point is not straightforward to guarantee mathematically [19] but may be justified by the estimation accuracy and the stability of the eKF in various applications. Further discussion of this condition can be found in [18, Sec. 9.6]. The covariance boundedness can be achieved by satisfying the uniform observability condition in [20]. The second point guarantees that the Cholesky factorization is well defined without pivoting/permutations.

⁴This term bears a passing resemblance to the innovation sequence of the Kalman filter. If the system (1) is linear, the eKF reduces to the Kalman filter, and this sequence is therefore white and of zero mean [17, Sec. 5.3]. In the general nonlinear case, the complete whiteness, zero mean, and independence conditions are not guaranteed but are satisfied to some extent in various applications (see [17, Sec. 8.2]).

We define the CE information state dynamics as

$$\pi_{k+1} = T_\pi(\pi_k, u_k), \quad (9)$$

where the tuple $\pi_k = (x_k^p, \Sigma_{k|k-1}^p)$, and $\Sigma_{k|k-1}^p, \Omega_k^p, F_k^p, H_k^p$, are defined as their corresponding in-name variables but with x_k^p replacing both $x_{k|k}$ and $x_{k|k-1}$, that is, with $[y_k - h(x_{k|k-1}, u_k)]$ is set to zero. The initial conditions $x_0^p = x_{0|-1}$ and $\Sigma_{0|-1}^p = \Sigma_{0|-1}$, are given in Section II.

Corresponding to (5), the CE-SOC cost is

$$\hat{J}^N(X_0, U_{N-1}) = \frac{1}{N} \left[x_N^{p\top} Q x_N^p + \text{tr}(Q \Sigma_{N|N-1}^p) + \sum_{k=0}^{N-1} \left[x_k^{p\top} Q x_k^p + u_k^\top R u_k + \text{tr}(Q \Sigma_{k|k-1}^p) \right] \right]. \quad (10)$$

Note that all the involved variables in (10) are deterministic, avoiding the high-dimensional integration over the many stochastic variables as in (5).

The following subsections are geared toward reformulating the dynamic system (9) into a linear realization and then using the Koopman operator theory [21] to put into the LQR context the computation of a feedback control law that minimizes (10).

D. The structure of the dynamics (9)

Note that in (9), $\Sigma_{k+1|k}^p$ is a function of x_k^p and u_k , through H_k^p and F_k^p . This makes $\Sigma_{k+1|k}^p$ a nonlinear function of u_k (it is nonlinear in H_k^p) unless $u_k \mapsto f(x_k^p, u_k)$ is affine in u_k with constant coefficients, for any $x_k^p \in \mathbb{X}$. To be precise, we call a function f_0 defined on $\mathbb{X} \times \mathbb{U}$ *bilinear* if it can be written as $f_0(x, u) = f_1(x) + f_2(x)u$, f_1 and f_2 of the appropriate dimensions, and we call it *bilinear with constant coefficients* if $f_2(x) = B_f$, a constant matrix.

We note that even if f and h are affine in u_k , not necessarily with constant coefficients, $\Sigma_{k+1|k}^p$ is in general a nonlinear function of u_k , since u_k still appears in the Jacobians F_k^p and H_k^p .

Proposition 2. If f and h in (1) are bilinear in u_k with constant coefficients, then: (i) the function T_π is bilinear in u_k with constant coefficients, (ii) T_π is continuous almost everywhere in the elements of the tuple π_k , and (iii) the covariance $\Sigma_{k+1|k}^p$ is a function of x_k^p only and is independent from u_k . \square

This proposition imposes more structure on the dynamics (9), enabling the application of more specialized control algorithms requiring bilinearity in the input, and allowing the discussion of truncation error, convergence, and stability under different model realizations [22], [23]. The third point is important in that it denies the immediate effect of u_k on the state covariance evolution; that is, $\Sigma_{k+1|k}^p$ is independent of u_k . This in turn allows various transformations (for example, Cholesky factorization) of the covariance matrix, without complicating the effect of u_k on the resulting transformed dynamics.

E. Toward the standard LQR form

Let L_k be the Cholesky (lower triangular) factor of $\Sigma_{k|k-1}^p = L_k L_k^\top$. Let ℓ_k be the half-vectorized (the nonzero elements, column by column, from left to right) of L_k . We denote the (invertible) mapping from the tuple $\pi_k = (x_k^p, \Sigma_{k|k-1}^p)$ to the vector $\eta_k = [x_k^p, \ell_k^\top]^\top$,

$$\eta_k = \mathcal{M}(\pi_k). \quad (11)$$

We also denote the evolution of η_k by

$$\eta_{k+1} = T_\eta(\eta_k, u_k).$$

The following describe the well-definedness, the structure of the above reformulation, and their relationship to T_π .

Lemma 2. *The function \mathcal{M} in (11) and its inverse \mathcal{M}^{-1} are well defined and continuous. Moreover, $\eta_{k+1} = T_\eta(\eta_k, u_k) = \mathcal{M}(T_\pi(\mathcal{M}^{-1}(\eta_k), u_k))$, and T_η is continuous almost everywhere in η_k , for any u_k .*

Proof. In the x_k^p portion, \mathcal{M} is simply an identity map.

From Assumptions 1 and 2, the matrix $\Sigma_{k|k-1}^p$ is positive definite [20], which in turn implies that the Cholesky decomposition is continuous [24, p. 295]. The inverse of the decomposition is simply obtained by $\Sigma_{k|k-1}^p = L_k^\top L_k$, which is also continuous. The half-vectorizing map $L_k \mapsto \ell_k$ and its inverse are obviously continuous. The continuity T_η follows from the above and T_π being continuous almost everywhere in η_k . \square

Corollary 1. *For all k , $\eta_k \in \mathbb{H} \subset \mathbb{R}^{r_\eta}$, \mathbb{H} is compact.*

Proof. It follows from Lemma 2 that the image under a continuous function of a compact set is compact, and $\mathbb{X} \ni x_k$ according to Assumption 1 is compact. \square

Now we show that the above transformation turns the cost (10) into the standard LQR form.

Proposition 3. *The cost in (10) is equivalent to*

$$\hat{J}^N = \frac{1}{N} \left[\eta_N^\top Q_{**} \eta_N + \sum_{k=0}^{N-1} [\eta_k^\top Q_{**} \eta_k + u_k^\top R u_k] \right], \quad (12)$$

where $Q_{**} = \text{block-diag}(Q, Q_*)$, and $Q_* = \text{block-diag}(Q_1, Q_2, \dots, Q_{r_x})$, where Q_i is the principal submatrix of Q starting from the element (i, i) of Q to the end (the element (r_x, r_x) , for example, $Q_1 = Q$).

Proof. The term $\text{tr}(Q \Sigma_{k|k-1}^p)$ in (10), using the cyclic property of the trace, can be written as $\text{tr}(L_k^\top Q L_k)$, which equivalently can be described by $\text{tr}(L_k^\top Q L_k) = \ell_k^\top Q_* \ell_k$. By substituting η_k in place of x_k^p and $\Sigma_{k|k-1}^p$, (10) turns into (12). \square

F. Koopman and eDMD for control

We follow the notation and the generalization of the Koopman operator to systems with control inputs as presented in [21]. We let l_u be the space of all control sequences of the

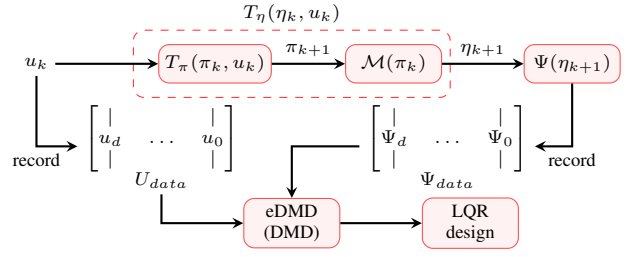


Fig. 1. **(Training Stage):** Block diagram illustrating the steps of the simulation-data collection stage, starting from a randomized initial condition η_0 and an injected, persistently exciting control sequence $\{u_k\}_{k \geq 0}$, next the creation of the data matrices U_{data} and Ψ_{data} , then the application of eDMD, and ending with solving the LQR problem of the lifted system.

form $\{u_k\}_{k=0}^\infty =: \mathbf{u}$, and $\mathbb{H} \times l_u$, where $\mathbb{H} \subset \mathbb{R}^{r_\eta}$ is a compact set. The “uncontrolled dynamics” are

$$\begin{bmatrix} \eta_{k+1} \\ \mathbf{u} \end{bmatrix} =: \chi_{k+1} = \begin{bmatrix} T_\eta(\eta_k, \mathbf{u}(k)) \\ q\mathbf{u} \end{bmatrix} =: \mathcal{T}(\chi_k), \quad (13)$$

where q is the time-shift operator applied elementwise: $q\mathbf{u}(k) = \mathbf{u}(k+1)$, and $\mathbf{u}(k) := u_k$.

The Koopman operator corresponding to (13), assuming its forward invariance, $\mathcal{K} : \mathcal{C}(\mathbb{H} \times l_u) \rightarrow \mathcal{C}(\mathbb{H} \times l_u)$ ($\mathcal{C}(\star)$ is the space of real-valued continuous functions with domain \star), is defined by $\mathcal{K}\phi(\chi_k) = \phi \circ \mathcal{T}(\chi_k)$, for every real-valued function ϕ with domain $\mathbb{H} \times l_u$.

Let Ψ be a finite-dimensional vector of elements of $\mathcal{C}(\mathbb{H})$, such that $\Psi = [\psi_1, \psi_2, \dots, \psi_{N_\Psi}]$. This vector is solely a function of the state η_k . We let $\Psi_k := \Psi(\eta_k)$ and construct Ψ such that the first r_η elements of Ψ_k are η_k and such that it spans a subspace of $\mathcal{C}(\mathbb{H})$.⁵

We now use the extended dynamic mode decomposition (eDMD) procedure, a simulation- or data-driven approach that seeks a finite-dimensional approximation of the Koopman operator [21]. In the data collection stage, as illustrated in Figure 1, a persistently exciting control sequence is injected to the simulation, and the data matrices U_{data} and Ψ_{data} are collected. Next, the eDMD approximation of a linear representation of the Koopman operator can be found through solving the following optimization problem,

$$\min_{A, B} \|\Psi_{data}^+ - A\Psi_{data}^- - BU_{data}\|_F, \quad (14)$$

where $\|\cdot\|_F$ is the Frobenius norm, $\Psi_{data}^+ = [\Psi_d, \dots, \Psi_1]$, and $\Psi_{data}^- = [\Psi_{d-1}, \dots, \Psi_0]$. The unique solution to (14) under the full column rank condition of the data is $[A, B] = \Psi_{data}^+ \begin{bmatrix} \Psi_{data}^- \\ U_{data} \end{bmatrix}^\dagger$. Since η_k forms the first entries of Ψ_k , it can be recovered by the canonical projection $C = [\mathbb{I}_{r_\eta} \times N_\Psi \ 0]$. The predictor is now given by

$$\Psi_{k+1} = A\Psi_k + Bu_k, \quad \eta_k = C\Psi_k.$$

(For the quality of this approximation see [25], [23].)

⁵For technical consideration, we choose $\psi(\eta_k, a) = \psi(\eta_k, b)$ for all $a, b \in l_u$ and suppress the second argument.

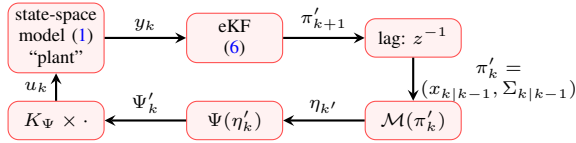


Fig. 2. **(Implementation stage):** A block-diagram illustrating the online implementation of the SOC-LQR control of the lifted system in closed-loop over the original system “plant” (1). The variables with a prime are functions of y_k , since the CE step (8) is omitted in the implementation phase, according to Remark 1.

The cost function (12) can now be described by

$$\hat{J}^N = \frac{1}{N} \left[\Psi_N^\top Q \Psi_N + \sum_{k=0}^{N-1} [\Psi_k^\top Q \Psi_k + u_k^\top R u_k] \right], \quad (15)$$

where $Q = \text{block-diag}(Q_{**}, 0_{N_\Psi - r_\eta \times N_\Psi - r_\eta}) \succeq 0$. We assume (A, B) to be stabilizable and $(A, Q^{\frac{1}{2}})$ detectable (see [26] for imposing these assumptions). Hence, $J^N \rightarrow a < \infty$ as $N \rightarrow \infty$, for a stabilizing feedback law. We denote the resulting LQR control (of the lifted-dynamics) by K_Ψ . Figure 2 illustrates the implementation phase of the control law K_Ψ in closed-loop with the original system. The difference in the used variables between Figures 1 and 2 is explained in Remark 1; since, in the real-time application, measurements are available and one uses the original eKF: (6) and (7), instead of the CE eKF: (9). The latter is used solely for control design and computing K_Ψ offline.

IV. NUMERICAL EXAMPLE

In this section we implement our linearized SOC approach on a system with varying state observability over the state-space. Suppose we have the system

$$x_{k+1} = \begin{bmatrix} .63 & .54 & 0 \\ .74 & .96 & .68 \\ .1 & -.86 & .54 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u_k + w_k, \quad (16)$$

$$y_k = \text{ELU} \left(\sum_{i=1}^3 x_k^i - 3 \right) + v_k,$$

where $x_k = (x_k^1, x_k^2, x_k^3)^\top$, ELU is the exponential linear activation function, widely used in deep learning applications, which is $\text{ELU}(x) = x$ for $x \geq 0$ and plateaus toward $\text{ELU}(x) = -1$ for $x \ll 0$ (in particular, it is $e^x - 1$, $x < 0$). The processes w_k and v_k follow the same assumptions as for (1) in Section II. Furthermore, $w_k \sim \mathcal{N}^3(0, 0.2\mathbb{I}_{3 \times 3})$ ($\mathcal{N}^{\text{trunc}}(\mu, \Sigma)$ is a Gaussian of zero mean and covariance Σ , truncated beyond the Mahalanobis distance $\text{MD} \geq \text{trunc}$ about the mean, then renormalized⁶), $v_k \sim \mathcal{N}^2(0, 0.2)$ and $x_0 \sim \mathcal{N}^3(0, \mathbb{I}_{3 \times 3})$. These truncations are required to satisfy the boundedness in Assumption 1.

This model belongs to an emerging class of models in system identification: the Hammerstein-Wiener family, which consists of linear systems composed in series with algebraic nonlinearities [27]. We pick this example because: (i) the

⁶The covariance after truncation can still be approximated by Σ when the value trunc chosen represents a high percentage confidence region.

corresponding SOC problem, with 9-dimensional state-space, whether through nonconvex MPC or DP, can be cumbersome or even prohibitive, and (ii) it splits the state space into two half-spaces separated by the surface $\sum x^i - 3 = 0$, with one half-space on which the system has significantly higher observability than the other. Hence, the behavior of the dual control derived is immediately interpretable.

Notice that, since $\text{ELU}(x) = x$ for $x \geq 0$, if $\sum_{i=1}^3 x_k^i \geq 3$, y_k has sensitivity w.r.t. x_k , while if $\sum_{i=1}^3 x_k^i \ll 3$, this sensitivity vanishes (compared with the constant y_k ’s sensitivity to v_k , i.e., $\partial y_k / \partial v_k = 1$). This difference in sensitivity between the two half-spaces ($\mathcal{H}_1 = \{\sum_{i=1}^3 x_k^i < 3\}$ and $\mathcal{H}_2 = \{\sum_{i=1}^3 x_k^i \geq 3\}$) affects the observability of x_k in the complete stochastic observability sense defined in [28]. In particular, for any value $\sum_{i=1}^3 x_k^i \ll 3$, $y_k \approx -1 + v_k$, x_k cannot be identified.⁷ Therefore, the quality of estimation varies over the state space. For better observability, it is required to “kick” the state to the half-space \mathcal{H}_2 , opposing stabilization, which, instead, requires the state to stay close to zero, contained in the opposite half-space \mathcal{H}_1 .

In the training stage, and according to Figure 1, we inject a random white noise input $u_k \sim \mathcal{N}^2(0, 0.2)$. We follow Figure 1 for data collection and then use a simple linear DMD model of the state $\Psi_k = [\eta_k^\top, 1]^\top$.

For the control design, we pick $Q = \mathbb{I}_{3 \times 3}$ (Q can be found accordingly) and $R = 1$, then find the LQR control law K_Ψ in (15) (as $N \rightarrow \infty$). We apply K_Ψ in a closed-loop simulation as explained in Figure 2. We also calculate the LQR control law K of the deterministic version of the system (16) ($w_k = 0$ and $y_k = x_k$, i.e., the CE assumption). The comparison between the performance of these two controllers is shown in Figure 3 and Table 1, after implementing each control design (CE-LQR: $K_{x_{k|k-1}}$ and SOC-LQR: $K_\Psi \Psi_k'$, with Ψ_k' as in Figure 2). SOC-LQR achieves a better control cost and a lower estimation error. That is, SOC-LQR not only improves the control performance but also significantly increases the estimation quality (the eKF performance) $\epsilon := \sum_k \|x_{k|k-1} - x_k^{\text{true}}\|_2^2$ ⁸. This can also be seen in Figure 4, where the true state (known in simulation) is wandering around for CE-LQR, while it is estimated well and thus controlled well with SOC-LQR.

| Metric (time-averaged) | CE-LQR: K | SOC-LQR: K_Ψ | reduction |
|------------------------|-------------|-------------------|-----------|
| Achieved cost | 22 | 1.7 | 92% |
| ϵ | 28.9 | 0.52 | 98% |

⁷Different definitions of observability for nonlinear systems exist. For example, the one provided by [29] (Def. 12: Degree of Observability) is related to the magnitude (power) of y_k itself rather than the ability to infer x_k from it. For $\sum_{i=1}^3 x_k^i \ll 3$, this definition returns a nonzero degree of observability, since $y_k \approx -1 + v_k \neq 0$ almost surely, contradicting the lack of identifiability in such regions of the state space. Therefore, the definition of the complete stochastic observability in [28] is more aligned with our intention. The two definitions align in the linear case where the power of y_k is captured in the observability Gramian, which also implies the identifiability of the state.

⁸These results can be reproduced by using our open-source JULIA code found at github.com/msramada/linearizing-uncertainty-for-control.

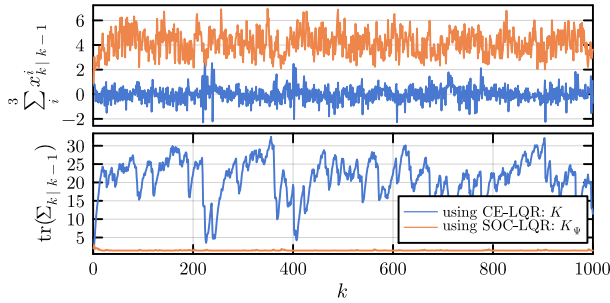


Fig. 3. The SOC-LQR pushes the state estimate to \mathcal{H}_2 (top), resulting in significantly lower estimation covariance (bottom), compared to CE-LQR which insists on keeping the state estimate around zero (in \mathcal{H}_1), leading to huge estimation errors (huge error covariance).

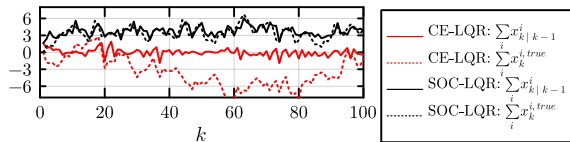


Fig. 4. SOC-LQR, by pushing the state estimate to \mathcal{H}_2 , accurately controls and estimates the true state, whereas CE-LQR drives the state estimate to zero, leading to huge estimation errors and the true state continuously departing far from its estimate.

V. CONCLUSION

The proposed approach requires the eKF to be an accurate Bayesian filter for the problem at hand. For systems with more complex uncertainty descriptions, an alternative filter might be required. Our methodology is also contingent, in its linearization part, on the choice of the dictionary functions used and, hence, on the prediction accuracy of the computed approximate Koopman operator. Therefore, the future of SOC will benefit immediately from advancements in the Koopman operator theory for control.

We seek further investigations toward exploiting the algebraic structure of the eKF and incorporating deep embedding approaches [30] to address large-scale SOC problems.

ACKNOWLEDGMENT

This material was based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR) under Contract DE-AC02-06CH11347.

REFERENCES

- [1] K. J. Åström, *Introduction to stochastic control theory*. Courier Corporation, 2012.
- [2] A. A. Feldbaum, “Dual control theory. i,” *Avtomatika i Telemekhanika*, vol. 21, no. 9, pp. 1240–1249, 1960.
- [3] E. Tse, Y. Bar-Shalom, and L. Meier, “Wide-sense adaptive dual control for nonlinear stochastic systems,” *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 98–108, 1973.
- [4] M. S. Ramadan, M. Alsuwaidan, A. Atallah, and S. Herbert, “A control approach for nonlinear stochastic state uncertain systems with probabilistic safety guarantees,” *arXiv preprint arXiv:2309.08767*, 2023.

- [5] F. Kendoul, “Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems,” *Journal of Field Robotics*, vol. 29, no. 2, pp. 315–378, 2012.
- [6] Z. Jiang, W. Zhou, H. Li, Y. Mo, W. Ni, and Q. Huang, “A new kind of accurate calibration method for robotic kinematic parameters based on the extended Kalman and particle filter algorithm,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 4, pp. 3337–3345, 2017.
- [7] E. Ghahremani and I. Kamwa, “Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements,” *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2556–2566, 2011.
- [8] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1.
- [9] D. Bayard and A. Schumitzky, “Implicit dual control based on particle filtering and forward dynamic programming,” *International Journal of Adaptive Control and Signal Processing*, vol. 24, 2008.
- [10] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model predictive control: theory, computation, and design*, 2nd ed. Nob Hill Publishing Madison, WI, 2017.
- [11] A. Mesbah, “Stochastic model predictive control with active uncertainty learning: A survey on dual control,” *Annual Reviews in Control*, vol. 45, pp. 107–117, 2018.
- [12] T. A. N. Heirung, B. E. Ydstie, and B. Foss, “Dual adaptive model predictive control,” *Automatica*, vol. 80, pp. 340–348, 2017.
- [13] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [14] A. Surana and A. Banaszuk, “Linear observer synthesis for nonlinear systems using Koopman operator framework,” *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 716–723, 2016.
- [15] M. Netto and L. Mili, “Robust Koopman operator-based kalman filter for power systems dynamic state estimation,” in *2018 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2018, pp. 1–5.
- [16] S. Resnick, *A probability path*. Springer, 2019.
- [17] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.
- [18] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.
- [19] B. F. La Scala, R. R. Bitmead, and M. R. James, “Conditions for stability of the extended Kalman filter and their application to the frequency tracking problem,” *Mathematics of Control, Signals and Systems*, vol. 8, pp. 1–26, 1995.
- [20] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, “Stochastic stability of the discrete-time extended Kalman filter,” *IEEE Transactions on Automatic control*, vol. 44, no. 4, pp. 714–728, 1999.
- [21] M. Korda and I. Mezić, “Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control,” *Automatica*, vol. 93, pp. 149–160, 2018.
- [22] D. Bruder, X. Fu, and R. Vasudevan, “Advantages of bilinear Koopman realizations for the modeling and control of systems with unknown dynamics,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4369–4376, 2021.
- [23] L. C. Iacob, R. Tóth, and M. Schoukens, “Koopman form of nonlinear systems with inputs,” *Automatica*, vol. 162, p. 111525, 2024.
- [24] M. Schatzman, *Numerical analysis: a mathematical introduction*. Oxford University Press, USA, 2002.
- [25] M. Haseli and J. Cortés, “Invariance proximity: Closed-form error bounds for finite-dimensional Koopman-based models,” *arXiv preprint arXiv:2311.13033*, 2023.
- [26] G. Mamakoukas, I. Abraham, and T. D. Murphey, “Learning stable models for prediction and control,” *IEEE Transactions on Robotics*, 2023.
- [27] A. Wills, T. B. Schön, L. Ljung, and B. Ninness, “Identification of Hammerstein–Wiener models,” *Automatica*, vol. 49, no. 1, pp. 70–81, 2013.
- [28] A. R. Liu and R. R. Bitmead, “Stochastic observability in network state estimation and control,” *Automatica*, vol. 47, no. 1, pp. 65–78, 2011.
- [29] U. Vaidya, “Observability gramian for nonlinear systems,” in *2007 46th IEEE conference on decision and control*. IEEE, 2007, pp. 3357–3362.
- [30] M. Tiwari, G. Nehma, and B. Lusch, “Computationally efficient data-driven discovery and linear representation of nonlinear systems for control,” *IEEE Control Systems Letters*, 2023.