Inferring System and Optimal Control Parameters of Closed-Loop Systems from Partial Observations

Victor Geadah[1,†], Juncal Arbelaiz[2], Harrison Ritz[3], Nathaniel D. Daw[3],
Jonathan D. Cohen[3], Jonathan W. Pillow[1,3]

Abstract— We consider the joint problem of system identification and inverse optimal control for discrete-time stochastic Linear Quadratic Regulators. We analyze finite and infinite time horizons in a partially observed setting, where the state is observed noisily. To recover closed-loop system parameters, we develop inference methods based on probabilistic state-space model (SSM) techniques. First, we show that the system parameters exhibit non-identifiability in the infinite-horizon from closed-loop measurements, and we provide exact and numerical methods to disentangle the parameters. Second, to improve parameter identifiability, we show that we can further enhance recovery by either (1) incorporating additional partial measurements of the control signals or (2) moving to the finite-horizon setting. We further illustrate the performance of our methodology through numerical examples.

I. INTRODUCTION

An important problem in the analysis of scientific data is accurately inferring the properties of dynamical systems based on partial observations. This is particularly important in neuroscience, where the neural systems that support the highly complex computations of interest are both difficult to observe, which is done for instance through neuro-imaging [1] or electrophysiology, and are often unknown a-priori. Probabilistic state-space models provide a powerful language to model neuroscientific data [2], [3], [4], [5], however, the fact that this systems are naturally operating in closed-loop configurations is often disregarded in the inference.

A growing body of work uses control-theoretic tools to analyze the dynamics of neural systems (see e.g., [6], [7], [8]). For example, recent work used a standard formulation in optimal control, the *Linear Quadratic Regulator* (LQR), to model neural dynamics in motor cortex during movement preparation [9]. While posing neuroscientific questions and analyzing related data from a control-theoretic lens is promising, inferring the properties of controllers from closed-loop neural recordings poses particular statistical challenges [10].

Work in inverse optimal control (IOC)—the problem of inferring the cost function of an optimal controller from recordings with known system dynamics—has historically focused on fully observed systems [11], [12], [13]. Recent work has explored control over partially observed systems [14], [15], but assumes the system parameters are known.

Similarly, inverse optimal control has been proven useful in understanding human sensorimotor control, but existing approaches have depended on full observations and/or a known plant [16], [17], [18], [19]. Refs. [20], [21] explored system parameter inference from a partially-observed neural system, but focused on system identification rather than inferring the parameters of the underlying control process.

Unlike engineered or known sensorimotor systems, for which accurate models of the dynamics might be available, the dynamics of neural systems are often unknown. Furthermore, they entail closed-loop processes. Leveraging control-theoretic perspectives on such un-observed processes recently gained traction in providing strong expressivity in state-space models [4]. Similarly, *model-free* [22], or *direct*, approaches to IOC involve no knowledge of the system and bypass it altogether, and offer discernible strengths for approximate control [23], [24]. In both cases, the study of the underlying joint parameters and identifiability is relatively unexplored.

Motivated by this problem, we focus on identifying both system parameters (e.g., system dynamics) and optimal control parameters (e.g., cost functions) from partial observations, assuming the system is operating in closed-loop with an optimal controller corresponding to the stochastic LQR. Our approach consists of two steps. First, we leverage the analytic and tractable expectation-maximization (EM) algorithm for system identification of linear-Gaussian state space models [25] to estimate a system's closed-loop dynamics. Second, we develop methods for inferring LQR state and action cost functions from these estimates of the closed-loop dynamics.

Our contributions are three-fold. First in Section §III, we provide a characterization of the non-identifiability encountered in the infinite horizon setting, as a necessary and sufficient approach to identifiability. Next, we provide two different alternatives to improve parameter identifiability, namely i) by considering the finite-horizon setting in §IV; and ii) by using partial measurements of the control inputs in §V. In all sections, we provide both theoretical treatments and either exact or numerical methods to disentangle the parameters.

A. Notation

For a sequence of vectors $\{x_t\}_{t=1}^{T}$ where $x_t \in \mathbb{R}^n$, we abbreviate the notation as $x_{1:T}$. We assume that all random variables are defined on some probability space $(\Omega \; \mathcal{B} \; \mathbb{P})$, and for a general random variable $y : \Omega \to \mathbb{R}^n$ we abuse

[1]Program in Applied and Computational Mathematics, Princeton University, Princeton NJ, USA

[2]Department of Mechanical and Aerospace Engineering, Princeton University, Princeton NJ, USA

[3]Princeton Neuroscience Institute, Princeton University, Princeton NJ, USA

[†]Correspondence: victor.geadah@princeton.edu

notation by denoting its instance $y \in \mathbb{R}^n$ and write its probability density function (p.d.f) as $p(y)$. We denote by $\mathbb{E}_{p(x)}[f(x)] = \int f(x)p(x)dx$ the expected value of $f(x)$ for the random variable x. We denote by I_n the $n \times n$ identity matrix. Finally, *identifiability* refers to the following:

Definition 1 (Identifiability, [26, Def. 5.2]). *If X is distributed according to p_θ, then θ is said to be* unidentifiable *on the basis of X if there exist $\theta_1 \neq \theta_2$ for which $p_{\theta_1} = p_{\theta_2}$.*

II. SETTING AND PROBLEM FORMULATION

A. Control Problem and Assumptions

We consider the stochastic linear quadratic regulator (LQR) optimal control problem, which with a finite and known time-horizon $T > 0$ reads

$$\min_{x_{1:T}, u_{0:T-1}} J_T(x_{1:T} \; u_{0:T-1}) =$$
$$\frac{1}{T}\mathbb{E}\left[x_T^\top Q_T x_T + \sum_{t=0}^{T-1} \left(x_t^\top Q x_t + u_t^\top R u_t \right) \right] \quad (1a)$$

subject to

$$x_{t+1} = Ax_t + Bu_t + {}_t \qquad {}_t \sim \mathcal{N}(0 \; W) \quad (1b)$$

for $t \in \{0 \quad T - 1\}$, $x_0 \sim \mathcal{N}(0 \; W_0)$, for the state $x_t \in \mathbb{R}^n$ and control inputs $u_t \in \mathbb{R}^m$, where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$. The system is assumed to be driven by zero-mean Gaussian noise with covariance W. The LQR problem aims to minimize the cost function J_T, with parameters $Q \; Q_T \in \mathbb{R}^{n \times n}$, positive semi-definite matrices, and $R \in \mathbb{R}^{m \times m}$, positive definite. The expectation is taken w.r.t. $\mathbb{P}(x_0 \quad _{0:T-1})$, over the initial condition x_0 and the dynamics noise sequence $_{0:T-1}$.

For simplicity we assume there is not a cross-term between control input and state in the cost function. The system is assumed to have access to *full* and *noiseless* measurements of its own state for feedback control. (That is, there is no need to implement a Kalman filter for control). Furthermore, for well-posedeness and uniqueness of control gain solutions below, we assume that $(A \; B)$ is controllable, $(A \; C)$ is observable [27].

It is well-known [28] that the optimal control inputs u^* for the problem (1) are in *proportional feedback* form

$$u_t^* = -K_t x_t$$
$$:= -(B^\top P_{t+1}B + R)^{-1}(B^\top P_{t+1}A)x_t \quad (2)$$

for the control gain $K_t \in \mathbb{R}^{m \times n}$, determined from the Riccati recursion term

$$P_t = Q + A^\top P_{t+1}A$$
$$- A^\top P_{t+1}B(B^\top P_{t+1}B + R)^{-1}B^\top P_{t+1}A \quad (3)$$

and $P_T = Q_T$. This yields the closed-loop system

$$x_{t+1} = \underbrace{(A - BK_t)}_{F_t}x_t + {}_t \quad (4)$$

with time varying closed-loop linear dynamics $F_t \in \mathbb{R}^{n \times n}$. We depict the block diagram for the closed-loop plant in
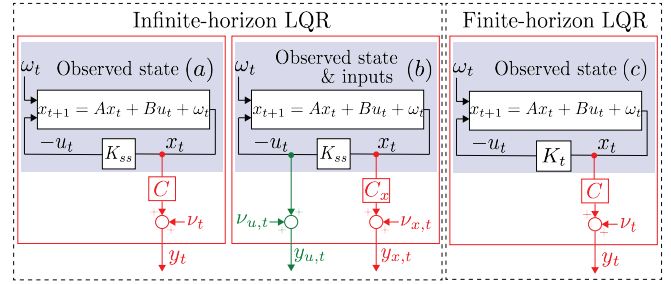

Fig. 1. Models of the optimal closed-loop dynamics and observations in the different scenarios analyzed in this work. The infinite-horizon LQR is displayed in (a)-(b). These panels differ in the observations available for inference to an external observer: in (a) the observer only has access to partial and noisy measurements of the state, while in (b) noisy measurements of the control input are also accessible. (c) Finite-horizon LQR and observation model.

Fig. 1. Furthermore, we will also be considering the infinite horizon setting, where we take $T \to \infty$ in (1a), removing Q_T. In this case, the optimal control signal u_t^* is a linear feedback of the state as well, but with a *static* gain

$$u_t^* = -K_{ss}x_t = -(B^\top P_{ss}B + R)^{-1}(B^\top P_{ss}A)x_t \quad (5)$$

for the control gain K_{ss}, where P_{ss} satisfies the discrete algebraic Riccati equation (DARE)

$$P_{ss} = Q + A^\top P_{ss}A$$
$$- A^\top P_{ss}B(B^\top P_{ss}B + R)^{-1}B^\top P_{ss}A \quad (6)$$

This now yields the closed-loop system

$$x_{t+1} = \underbrace{(A - BK_{ss})}_{F_{ss}}x_t + {}_t \quad (7)$$

with time-invariant linear dynamics $F_{ss} \in \mathbb{R}^{n \times n}$.

Importantly, we consider the setting in which an external observer has access only to partial and noisy measurements $y_t \in \mathbb{R}^d$ of the state x_t for inference. In particular, we consider linear Gaussian observations

$$y_t = Cx_t + \nu_t \qquad \nu_t \sim \mathcal{N}(0 \; V) \quad (8)$$

for $t \in \{0 \quad T\}$, where $C \in \mathbb{R}^{d \times n}$ and for the covariance matrix V. The system (1b)-(8) is sometimes referred to as an *input-driven linear Gaussian state-space model* (LG-SSM [29] (see non-input driven definition in Appendix VI-A), also referred to as *Linear Dynamical System* (LDS) [30], where specifically the control input sequence $u_{0:T-1}$ and corresponding $x_{1:T}$ trajectory are optimal under the stochastic LQR framework.

Having introduced the problem setting, we end with some notation. Let our set of parameters Θ be decomposed into Θ_s the set of system parameters $\theta_s = \{A \; B \; Q \; R\} \in \Theta_s$, Θ_o the set of output parameters $\theta_o = \{C\} \in \Theta_o$, and Θ_n the set of noise parameters $\theta_n = \{W_0 \; W \; V\} \in \Theta_n$. We have $\Theta = \Theta_s + \Theta_o + \Theta_n$. Furthermore, we will denote by $\mathcal{F} : \Theta_s \to \mathbb{R}^{n \times n}$ (resp. $\mathcal{F}_t$) the function that takes our system parameters θ_s and returns the closed-loop dynamics $\mathcal{F}(\theta_s) = F_{ss}$ in (7) (resp. $\mathcal{F}_t(\theta_s) = F_t$ in (4)).

B. Problem Formulation

An external observer has access only to partial observations $y_{0:T}$ of a system under control operating in closed-loop. At a high level, the question we seek to answer in this work is: strictly from observations y, what can be inferred about the system parameters $\{A\ B\ C\ W\ W_0\ V\}$ and the control parameters $\{Q\ R\}$? We will articulate most of our results on the specific interaction between $\{A\ B\ Q\ R\}$. Non-identifiability is expected, thus we further ask: what are additional measurements or assumptions that we can use to reach identifiability?

The basis for identifiability and parameter inference is the *marginal likelihood*,

$$p_\theta(\mathbf{y}) = \prod_{i=1}^{N} p_\theta\left(y_{0:T}^{(i)}\right) \tag{9}$$

for a dataset $\mathbf{y} = \{y_{0:T}^{(i)}\}_{i=1}^{N}$. We perform parameter inference under the closed-loop statistical model p_θ presented in (7)-(8) by seeking $\hat\theta \in \Theta$ maximizing the marginal likelihood.

To shed light into the non-identifiability, we consider the first case of the infinite horizon setting. We plot in Fig. 2**A** what such parameter inference approach can yield from observations in this case. We can infer the closed-loop parameters F_{ss}, but the parameters $\{A\ B\ Q\ R\}$ are not identifiable. This can be explained from the latent dynamics probabilities, which satisfy

$$p_\theta(x_{t+1} \mid x_t\ u_t^*) = p(x_{t+1} \mid x_t\ \theta_s\ \theta_n)$$
$$= p(x_{t+1} \mid x_t\ \mathcal{F}(\theta_s)\ \theta_n) \tag{10}$$

and are *equivalent* to the LG-SSM latent dynamics. In App. A (Prop. 4), we formalize that under the assumption that the true model is indeed p_θ, then parameter inference in the general statistical model defined through (7)-(8) is equivalent to parameter inference in an LG-SSM. Hence performing parameter inference can be reduced to, and is in fact equivalent to, the two following steps:

1) Infer the closed-loop parameters F_{ss} from partial observations $\mathbf{y}$;
2) Investigate the interaction of the system parameters θ_s in the closed-loop dynamics F_{ss}.

In the next section §II-C we address the first step and discuss how in practice we maximize the marginal likelihood in (9) and infer the closed-loop system parameters. The second step will be the topic of Sections §III and §IV for respectively the infinite- and finite-horizon settings. Finally, in Section §V we explore how the introduction of additional control measurements can help relax this identifiability constraint. See Fig. 2**B** for a schematic of our approach.

C. Closed-Loop System Recovery

A general difficulty in fitting parameters to data in state space models is that we must marginalize over latent sequences $x_{0:T}$

$$p_\theta(y_{0:T}) = \int p_\theta(y_{0:T} \mid x_{0:T}) p_\theta(x_{0:T}) dx_{0:T}$$
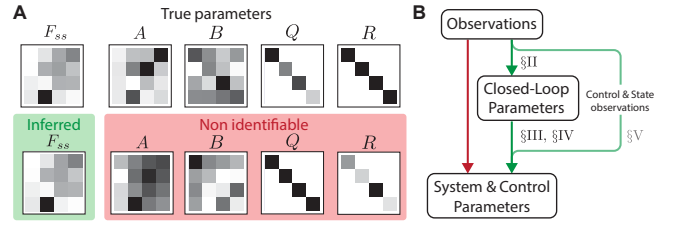$$= \mathbb{E}_{p_\theta(x_{0:T})}\left[p_\theta(y_{0:T} \mid x_{0:T})\right]$$

Fig. 2. (**A**) Recovery from observations of the system $\{A, B\}$ and cost function $\{Q, R\}$ parameters in the infinite horizon setting is limited to identification of the closed-loop dynamics F_{ss}. We plot the inferred F_{ss} using EM, as well as a combination $\{A, B, Q, R\}$ that yields this same F_{ss}. (**B**) Parameter inference can be reduced to closed-loop identification with probabilistic methods, followed by an investigation of the interaction between the different parameters in setting the closed-loop dynamics.

to evaluate the marginal likelihood. For general state space models, this integration cannot be done in closed-form. Furthermore, optimizing this expression requires optimizing parameters of the support distribution of this integral. A popular approach that circumvents the latter consists of using the *expectation-maximization* (EM) algorithm [31], [32], [25] to maximize (9). The EM algorithm is centered around the following lower bound

$$\ln p_\theta(\mathbf{y}) \geq \mathcal{Q}(\theta \mid \theta') := \mathbb{E}_{p(x_{0:T}\mid\mathbf{y},\theta')}\left[\ln p(x_{0:T}\ \mathbf{y} \mid \theta)\right] \tag{11}$$

to the marginal *log*-likelihood, for any $\theta\ \theta' \in \Theta$. The algorithm can then be summarized succinctly as recursively maximizing

$$\theta^{(j+1)} \leftarrow \arg\max_\theta \mathcal{Q}(\theta \mid \theta^{(j)}) \tag{12}$$

over iterates $j = 1\ 2$ until convergence. This algorithm has the convenient property that it guarantees monotonic increase in the lower bound $\mathcal{Q}$, but is generally highly dependent on initialization [33]. We use the EM-algorithm to maximize the likelihood in (9). Further details for our problem are provided in Appendix B. In Fig. 3**B**, we provide estimation errors for the recovery procedures as a function of the number of EM iterations, attesting to its convergence.

In linear Gaussian state-space models, the marginal likelihood in (9) can in fact be evaluated in closed-form using Kalman filtering [29, See §18.3.1.3]. Nevertheless, the EM-algorithm for the LG-SSM has closed-form updates for the parameters θ and remains a standard approach for its efficiency and flexibility [33]. Importantly for us, we will leverage these closed-form updates in §V to showcase how the inclusion of control measurements can help us further in the identification of the parameters θ.

Regardless of the parameter inference methods, since the latent variables x_t are not observed our model is only identifiable up to any invertible transformation of the latent space [10], [11]. Indeed, if we define a new latent state as $\hat{x}_t = Hx_t$ for $H \in \mathbb{R}^{n\times n}$ invertible, then the closed-loop dynamics can be redefined as $\hat{F}_{ss} = HF_{ss}H^{-1}$, with dynamics covariance $HWH^\top$. The topic of invariant metrics under non-identifiability of system and control parameters

dates back to original work by Kalman [11], [34] and is an active area of research. Here, we study the relationship of the parameters $\theta_s = \{A\ B\ Q\ R\}$ in setting F_{ss}. This helps in practice to provide estimates for $\hat{\theta}_s = \{\hat{A}\ \hat{B}\ \hat{Q}\ \hat{R}\}$ from an inferred $\hat{F}_{ss}$. How identifiability considerations impact the estimates $\hat{\theta}$ is out of the scope of this paper. Thus, one must exercise caution when comparing parameters *estimated* (e.g. $\hat{A}$) from this procedure with *true* (e.g. A), generative, parameters.

III. IDENTIFIABILITY IN INFINITE HORIZON

The preceding section highlighted how our higher level goal amounts to disentangling the parameters $\{A\ B\ Q\ R\}$ from measurements of the closed-loop system. The key idea is that by having direct measurements of the closed-loop, we will be able to disentangle our control gain equation (2) (resp. (5)) from our Ricatti term (3) (resp. (6)), yielding a system of two equations.

Let us first consider the infinite time-horizon setting, where we now presume to have access to a measurement of F_{ss}. Strictly from our control gain definition, we have

$$
\begin{aligned}
F_{ss} &= A - BK_{ss} \\
&= \left(\mathrm{I} - B(R + B^\top P_{ss}B)^{-1}B^\top P_{ss}\right)A
\end{aligned}
$$

and in particular

$$
\begin{aligned}
A &= \left(\mathrm{I}_n - B(R + B^\top P_{ss}B)^{-1}B^\top P_{ss}\right)^{-1}F_{ss} \\
&= \left(\left(\mathrm{I}_n + BR^{-1}B^\top P_{ss}\right)^{-1}\right)^{-1}F_{ss} \\
&= \left(\mathrm{I}_n + BR^{-1}B^\top P_{ss}\right)F_{ss} \qquad (13)
\end{aligned}
$$

where we used the Woodbury matrix identity to perform the inverse. Next, shifting to the DARE in (6), we note

$$
\begin{aligned}
P_{ss} &= Q + A^\top P_{ss}A - A^\top P_{ss}B\left(R + B^\top P_{ss}B\right)^{-1}B^\top P_{ss}A \\
&= Q + A^\top P_{ss}A - A^\top P_{ss}BK_{ss} \\
&= Q + A^\top P_{ss}(A - BK_{ss})
\end{aligned}
$$

such that

$$
P_{ss} = Q + A^\top P_{ss}F_{ss} \qquad (14)
$$

These two simple derivations together show that the interaction between the different parameters of interest can be reduced to considering a new system of equations, which we make explicit in the following proposition:

Proposition 1. *Consider the infinite horizon (stochastic) LQR problem with system parameters $\{A\ B\}$ and control parameters $\{Q\ R\}$. Let $(A\ B)$ be stabilizable, and let F_{ss} be the closed-loop dynamics matrix. Then, the unique solution P to the Sylvester equation*

$$
P - A^\top P F_{ss} = Q \qquad (15a)
$$

satisfies the matrix equation

$$
BR^{-1}B^\top P F_{ss} = A - F_{ss} \qquad (15b)
$$

Proof: $(A\ B)$ stabilizable ensures that $P = P_{ss}$ in (15a) is unique, as derived from the DARE. The proof then simply follows from our derivations in (13)-(14).

An interesting implication of the stabilizability of $(A\ B)$, is that since P in (15a) is unique, this in turn implies that F^{-1} and $-A^\top$ (equivalently $-A$) do not share an eigenvalue.

In Prop. 3, F_{ss} is, *a-priori*, the only known parameter. As such, to make any conclusions about one or a combination of parameters without neither resorting to P_{ss} nor constraints on our parameters, we need to assume the others to be known.

a) Solving for A: First and foremost, we may wish to solve for A in (15). This is slightly more challenging, as the system exhibits second-order tendencies in A. To see this, consider the simplification $B = R = \mathrm{I}$. In that case, we can easily solve for P in (15b) and substitute it in (15a), and obtain that A satisfies the system (15) if and only if

$$
Q = (A - F_{ss})F_{ss}^{-1} - A^\top(A - F_{ss}) \qquad (16)
$$

This is a second-order matrix equation in A, combining of a second-order self-adjoint term with a first-order term of Sylvester-transpose form. This combination is unfortunately not (readily) amenable to common second-order matrix equation techniques [35]. Instead, we elect to use an iterative approach that draws from iterative solutions to Ricatti-type second-order equations [36].

Proposition 2. *Let $\{B\ R\ Q\}$ be given, as well as F_{ss}. Let $P_j,\ j = 1\ 2\quad$, be the (unique) positive definite solution to the algebraic Sylvester equation*

$$
P_j - A_j^\top P_j F_{ss} = Q \qquad (17a)
$$

where, recursively,

$$
A_j = \left(\mathrm{I} + BR^{-1}B^\top P_{j-1}\right)F_{ss} \qquad (17b)
$$

where A_0 is chosen such that it does not share any eigenvalue with $-F_{ss}^{-1}$. Then
1) $P_{ss} \leqq P_{j+1} \leqq P_j \leqq \quad, j = 1\ 2$
2) $\lim_{j\to\infty} A_j = A$

Proof: This result follows from a similar iterative procedure in [36]. The key idea is that instead of modifying the F_{ss} at every step holding A constant, we let A be adjusted so that F_{ss} stays constant.
1) By our assumption on A_0, P_0 defined as a solution to the Sylvester equation (17a) is unique and well-defined. Let P_j be well-defined. Then as A_{j+1} is adjusted by the second step (17b), itself derived from the control gain equation, it ensures that there is a control gain K_{j+1} such that $A_{j+1} - BK_{j+1} = F_{ss}$. This in turn ensures that $(A_{j+1}\ B)$ is stabilizable, and that the solution $P_{j+1} \succ 0$ to the DARE (6) of parameters $\{A_{j+1}\ B\ Q\ R\}$ is unique and well defined. Then by induction, the sequence $\{P_j\ K_{j+1}\}$ for $j = 1\ 2\quad$ is well-defined. Furthermore, it satisfies the same recursive form as in [36], and it follows from [36] that $P_j - P_{j+1} \succeq 0$, $P_j - P_{ss} \succeq 0$.
2) Monotonic convergence of positive operators ensures that the limit for P_j exists and the lower bound and

uniqueness of P_{ss} yields $\lim_{j \to \infty} P_j = P_{ss}$. Then by Prop. 3, it holds that $A_j \to A$.

We require A_0 to not share any eigenvalue with $-F_{ss}^{-1}$. Since $(A\ B)$ is stabilizable and F_{ss} is invertible, then F_{ss} has eigenvalues of norm less than 1 and greater than 0. Possible initalizations for A_0 thus include $A_0 \in \{F_{ss}\ \mathrm{I}\ 0\}$.

The above result provides monotonic and quadratic convergence (see Fig. 3**A**) to the true A when we have access to F_{ss}, and know the other parameters $\{B\ R\ Q\}$. In practice, one may have only a-priori estimates or imperfect measurements of any of these parameters. The proof above only holds for the true parameters, thus convergence in this setting is not guaranteed. Nonetheless, we have found this iterative method to fare nearly just as well as a numerical root-finding algorithm based on Newton's method, and the proof above can give insights into the well-posedness of the problem and convergence guarantees. To assess this sensitivity, we investigated the impact of adding a scalar perturbation $\epsilon\,\mathrm{I}$, $\epsilon > 0$, to either F_{ss} or Q on our estimate for A. We show in Fig. 3**C-D** that both iterative and root finding methods handle this perturbation similarly.

b) Solving for B and R: Next, we notice that $\{B\ R\}$ are in direct interaction through $BR^{-1}B^\top$. One can solve for $BR^{-1}B^\top$ given $\{A\ Q\ F_{ss}\}$ by solving for the unique P in (15a) and substituting it into the second equation (15b). As for the specific values of B and R, they are not strictly identifiable. For instance, for any given B and $R \succ 0$, let $\hat{B} = BR^{-\frac{1}{2}}UR^{\frac{1}{2}}$ for $U \in \mathbb{R}^{m \times m}$ unitary, arbitrary. Then

$$
\begin{aligned}
\hat{B}R^{-1}\hat{B}^\top &= \left(BR^{-\frac{1}{2}}UR^{\frac{1}{2}}\right)R^{-1}\left(BR^{-\frac{1}{2}}UR^{\frac{1}{2}}\right)^\top \\
&= BR^{-\frac{1}{2}}UU^\top R^{-\frac{1}{2}}B^\top \\
&= BR^{-1}B^\top
\end{aligned}
$$

making $\hat{B}$ and B non-identifiable.

c) Solving for Q: Finally, solving for our last parameter of interest can be done quite easily if $BR^{-1}B^\top$ is invertible, as it allows to solve for P_{ss} in (15b) and subsitute it in (15a) to readily read of Q. For instance, if $B = R = \mathrm{I}$, then Q is immediately available from (16). If $BR^{-1}B^\top$ is not invertible, for instance, if the control inputs u_t are lower dimensional than the state x_t, then care must be taken by considering the pseudo-inverse and identify its null-space.

IV. PARTIALLY OBSERVED FINITE HORIZON

In the *finite* horizon stochastic LQR setting described in (1), the closed-loop dynamics satisfy (4). We make the simplifying assumption of considering $Q_T = Q$. Now, just as in the infinite-horizon setting, our system transition probabilities satisfy

$$
\begin{aligned}
p_\theta(x_{t-1} \mid x_t\ u_t^*) &= p(x_{t-1} \mid x_t\ A\ B\ Q\ R\ \theta_n) \\
&= p(x_{t-1} \mid x_t\ \mathcal{F}_t(A\ B\ Q\ R) = F_t\ \theta_n)
\end{aligned}
\tag{18}
$$

for $t \in \{1\quad T - 1\}$. Importantly, again, the system parameters $\{A\ B\}$ and performance metric parameters $\{Q\ R\}$ only affect our marginal likelihood through the closed-loop
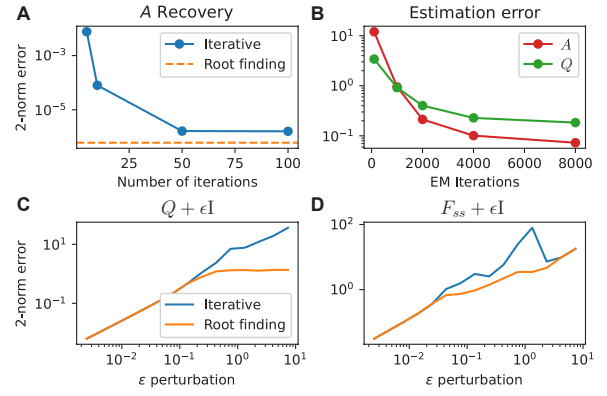
Fig. 3. Parameter estimation convergence and robustness properties. (**A**) Matrix 2-norm error between the true parameter A used for simulation (4×4 unitary rotation) and the estimated A from the iterative procedure in Prop. 2 as a function of the number of iterations. (**B**) Similar estimation error to **A** for A given Q as well as for Q given A, this time using $\hat{F}_{ss}$ estimated from the EM-algorithm, as a function of the number of iterations. We use $C = \mathrm{I}$ for identifiability. (**C-D**) Recovery error for A as a function of Q (resp. F_{ss}) with perturbed entries. See text for further details.

dynamics F_t. In turn, this yields a similar system of equations to Prop. 3 for the finite-horizon setting.

Proposition 3. *Consider the finite horizon (stochastic) LQR problem with system parameters $\{A\ B\}$ and control parameters $\{Q\ R\}$. Let $(A\ B)$ be stabilizable, and let F_t be the closed-loop dynamics for $t \in \{1\quad T - 1\}$. Then,*

$$
P_t - A^\top P_{t+1}F_t = Q
\tag{19a}
$$

satisfies the matrix equation

$$
BR^{-1}B^\top P_{t+1}F_t = A - F_t
\tag{19b}
$$

and $P_T = Q$.

The main difference with the infinite horizon setting in terms of recovery considerations comes down to the terminal condition $P_T = Q$. This means that we can restrict the solution space over Θ_s strictly from our last step. Indeed, for $t = T - 1$, from (19b) we obtain

$$
\begin{aligned}
BR^{-1}B^\top P_T F_{T-1} &= A - F_{T-1} \\
\Longleftrightarrow \quad A &= \left(BR^{-1}B^\top Q + \mathrm{I}\right)F_{T-1}
\end{aligned}
\tag{20}
$$

This provides an essential, supplemental, link between the parameters $\{A\ B\ Q\ R\}$.

To leverage this, consider again $B = R = \mathrm{I}$. Starting from (19a) and substituting in (19b),

$$
\begin{aligned}
P_{T-1} &= Q + A^\top P_T F_{T-1} \\
\Longleftrightarrow (A - F_{T-2})F_{T-2}^{-1} &= Q + A^\top(A - F_{T-1})
\end{aligned}
$$

and now (20) reads as $A = (Q + \mathrm{I})F_{T-1}$, such that

$$
\begin{aligned}
\left(A - F_{T-1}\right)F_{T-1}^{-1} = \\
(A - F_{T-2})F_{T-2}^{-1} - A^\top\left(A - F_{T-1}\right)
\end{aligned}
\tag{21}
$$

This expression is similar to (16) in the infinite horizon setting, but now using our terminal condition to implicitly

solve for Q. With this, we can now solve this expression for A, and substitute in (20) to recover *both A and Q*. We use root-finding methods to solve for A. See Fig. 4(bottom) for a numerical example supporting this.

In the above, we focused on leveraging the terminal condition to identify more parameters than in the infinite horizon setting. We note that generally the now *time-varying* closed-loop dynamics provide us with many more distinct equations.

V. INFINITE HORIZON WITH PARTIALLY OBSERVED STATE AND CONTROL

We use the EM algorithm to identify the dynamical system (1b) operating in closed-loop, under the optimal control input (5) corresponding to the *infinite* horizon stochastic LQR setting. We provide explicit expressions for the E-step and M-step of the EM algorithm, under two different linear Gaussian observation models (8):

$$y_t = Cx_t + \nu_t \tag{22a}$$

$$y_t = \begin{bmatrix} C_x x_t \\ u_t \end{bmatrix} + \begin{bmatrix} \nu_{x,t} \\ \nu_{u,t} \end{bmatrix} = \begin{bmatrix} C_x \\ -K_{ss} \end{bmatrix} x_t + \begin{bmatrix} \nu_{x,t} \\ \nu_{u,t} \end{bmatrix}$$

$$\text{with } \mathbb{E}\left[\begin{bmatrix} \nu_{x,t} \\ \nu_{u,t} \end{bmatrix} \begin{bmatrix} \nu_{x,t} & \nu_{u,t} \end{bmatrix} \right] = \begin{bmatrix} V_x & 0 \\ 0 & V_u \end{bmatrix} \tag{22b}$$

That is, while the observation model (22a) only has access to partial and noisy measurements of the state, model (22b) additionally has access to noisy measurements of the control input. For convenience, we denote $y_{x,t} = C_x x_t + \nu_{x,t}$ and $y_{u,t} = u_t + \nu_{u,t}$ in (22b).

We show that while *model (a)* only allows for the identification of the closed-loop F_{ss}, *model (b)* provides a further relationship between the inner parameters of the closed-loop B and K_{ss}. Technical details are deferred to Appendix B.

VI. CONCLUSIONS AND FUTURE WORKS

In conclusion, we have investigated the identifiability of system and control parameters from partial observations, and provided two extensions that substantially improve identifiability. Future theoretical work should explore how estimation errors during system identification influences the recovery of system and control parameters, as well as data-efficient methods for parameter inference in the finite-horizon setting.

These methods provide a promising path towards domain-specific problems of estimating the control properties of partially observed neural systems. For example, the relationship between MEG sensors and the underlying neuromagnetic sources is approximately linear [37], suggesting that our methods for system and control recovery may be helpful for understanding human neural control processes from noninvasive recordings.
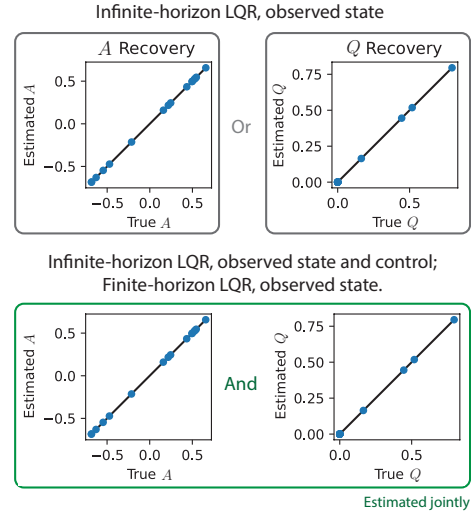
Fig. 4. Overview of the identification possibilities in each setting. In the infinite horizon, we provide exact and numerical estimation procedures for one parameter given the others; with $B = R = I$, we present A given Q and Q given A. In the finite horizon, we can leverage terminal conditions to estimate both A and Q (if $Q_T = Q$) simultaneously. In the infinite horizon with both x_t and u_t partially observed, and given $B = R = I$, we can estimate A and then estimate Q.

APPENDIX

A. Equivalence between models in marginal likelihood

Without loss of generality, we omit the dependence on the noise parameters θ_n in writing our model distributions $p_\theta(x_{0:T} \ y_{0:T})$ below. We start by defining the LG-SSM model class for completeness.

Definition 2. A *linear Gaussian state-space model* (LG-SSM) with parameters $ = \{A \ C \ V \ W_0 \ W\}$ is defined by the joint distribution $q_\varphi(x_{0:T} \ y_{0:T})$ over the latent sequence $x_{0:T}$, $x_t \in \mathbb{R}^n$, and observations $y_{0:T}$, $y_t \in \mathbb{R}^d$, satisfying

$$q(x_0) = \mathcal{N}(0 \ W_0) \quad q(x_{t+1} \mid x_t \ ) = \mathcal{N}(x_{t+1}; Ax_t \ W)$$
$$q(y_t \mid x_t \ ) = \mathcal{N}(y_t; Cx_t \ V)$$

Proposition 4. *Let* $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$ *be the statistical model defined by* (1)-(8) *with parameter space* Θ, *and let* $\mathcal{Q} = \{q_\varphi : \ \in \Phi\}$ *be an LG-SSM with parameter space* Φ. *Let our dataset* $\mathbf{y}$ *be sampled from a model in* $\mathcal{P}$, *then as* $|\mathbf{y}| \to \infty$

$$\hat{\theta} \in \arg\max_{\theta \in \Theta} p_\theta(\mathbf{y}) \iff \{\hat{\theta}_o \ \mathcal{F}(\hat{\theta}_s)\} \in \arg\max_{\varphi \in \Phi} q_\varphi(\mathbf{y})$$

Proof: First, we note that the observation models for the two statistical models coincide, that is $p_{\theta_o}(y_{0:T} \mid x_{0:T}) =$

$q_{\theta_o}(y_{0:T} \mid x_{0:T})$ for all $\theta_o \in \Theta_o$. Next for the dynamics distribution, we expand on (10) and note that

$$p_{\theta_s}(x_{0:T}) = p(x_0) \prod_{t=0}^{T-1} p(x_{t+1} \mid x_t \; \mathcal{F}(\theta_s))$$

$$= q(x_0) \prod_{t=0}^{T-1} q(x_{t+1} \mid x_t \; \mathcal{F}(\theta_s)) = q_{\mathcal{F}(\theta_s)}(x_{0:T})$$

for all $\theta_s \in \Theta_s$. Finally, marginalizing over $x_{0:T}$, we obtain

$$p_\theta(y_{0:T}) = \int p_{\theta_o}(y_{0:T} \mid x_{0:T}) p_{\theta_s}(x_{0:T}) dx_{0:T}$$

$$= \int q_{\theta_o}(y_{0:T} \mid x_{0:T}) q_{\mathcal{F}(\theta_s)}(x_{0:T}) dx_{0:T}$$

$$= q_{\{\theta_o, \mathcal{F}(\theta_s)\}}(y_{0:T})$$

for all $\theta \in \Theta$. We have shown that $\mathcal{P} \subset \mathcal{Q}$ through the mapping $\{\theta_o \; \theta_s\} \mapsto \{\theta_o \; \mathcal{F}(\theta_s)\}$, and $\Theta \subset \Phi$. Now, for the proof,

$\Rightarrow$) By assumption,

$$\max_{\theta \in \Theta} p_\theta(\mathbf{y}) = p_{\hat{\theta}}(\mathbf{y}) = q_{\{\hat{\theta}_o, \mathcal{F}(\hat{\theta}_s)\}}(\mathbf{y})$$

To show this is in the maximum set, let us assume there exists another $\in \Phi$ such that $q_\varphi(\mathbf{y}) > q_{\{\hat{\theta}_o, \mathcal{F}(\hat{\theta}_s)\}}(\mathbf{y})$ holds as $|\mathbf{y}| \to \infty$. Then, since the true model is in $\mathcal{P}$, there exists $\tilde{\theta}$ such that $= \tilde{\theta}_o \cup \tilde{\theta}_s$, which implies $p_{\tilde{\theta}}(\mathbf{y}) > p_{\hat{\theta}}(\mathbf{y})$ — a contradiction.

$\Leftarrow$) This direction follows from from $\Theta \subset \Phi$. If $\hat{\theta}_o \cup \mathcal{F}(\hat{\theta}_s) \in \arg\max_{\varphi \in \Phi} q_\varphi(\mathbf{y})$, then $q_{\hat{\theta}_o, \mathcal{F}(\hat{\theta}_s)} \geq q_{\theta_o, \mathcal{F}(\theta_s)}$ for all $\theta \in \Theta$, which is equivalent to $p_{\hat{\theta}} \geq p_\theta$, as desired.

B. The EM algorithm

The EM algorithm consists of two steps. In the *E-step*, the current parameter values, denoted by θ^{old}, are used to compute $\mathcal{Q}(\theta \mid \theta^{\text{old}})$, as defined in (11). This yields a function of the parameter set θ, which is then maximized in the *M-step* by setting the appropriate partial derivatives to 0, and solving for each parameter. As the expected log-likelihood $\mathcal{Q}(\theta \mid \theta')$ depends on the observation model, we denote by $\mathcal{Q}_{(a)}$ the expected log-likelihood for the observation model (22a), and similarly for $\mathcal{Q}_{(b)}$.

a) Partial state observations: In the *E-step* for the observation model (22a), $y_t = Cx_y + \nu_t$, the expected log-likelihood $\mathcal{Q}_{(a)}(\theta \mid \theta^{\text{old}})$ can be expressed as a function of the parameters of interest in θ and the following *sufficient statistics*

$$M_{(t_1,t_2)}^{(i)} := \sum_{t=t_1}^{t_2} \left(m_t^{(i)} m_t^{(i)\top} + \Sigma_t^{(i)} \right) \tag{23a}$$

$$M_\Delta^{(i)} := \sum_{t=1}^{T} \left(m_{t-1}^{(i)} m_t^{(i)\top} + \Sigma_{t-1,t}^{(i)} \right) \tag{23b}$$

$$Y^{(i)} := \sum_{t=0}^{T} y_t^{(i)} y_t^{(i)\top} \quad \tilde{Y}^{(i)} := \sum_{t=0}^{T} m_t^{(i)} y_t^{(i)\top} \tag{23c}$$

for the i−th sample, with $m_t^{(i)}$ and $\Sigma_t^{(i)}$ denoting the mean and the covariance of the posterior distribution over the latent variable at time t, that is, $p(x_t \mid y_{0:T}^{(i)}) = \mathcal{N}(m_t^{(i)} \; \Sigma_t^{(i)})$. These are computed by running the Kalman filter-smoother [38] in each independent time series $y_{0:T}^{(i)}$ using the parameter values θ^{old}, which can be performed *in parallel across datasets*. We can then pool these sufficient statistics as $\mathcal{M}_{(t_1,t_2)} := \sum_{i=1}^{N} M_{(t_1,t_2)}^{(i)}$, $\mathcal{M}_\Delta := \sum_{i=1}^{N} M_\Delta^{(i)}$, $\mathcal{Y} := \sum_{i=1}^{N} Y^{(i)}$, and $\tilde{\mathcal{Y}} := \sum_{i=1}^{N} \tilde{Y}^{(i)}$.

As for the *M-step*, the explicit parameter updates under the observation model (22a) are then

$$\hat{F}_{ss} = \mathcal{M}_\Delta^\top \mathcal{M}_{(0,T-1)}^{-1} \quad \hat{C} = \tilde{\mathcal{Y}}^\top \mathcal{M}_{(0,T)}^{-1}$$

$$\hat{W}_0 = \frac{1}{N} \mathcal{M}_{(0,0)}$$

$$\hat{V} = \frac{1}{N(T+1)} \left(\mathcal{Y} - \hat{C}\tilde{\mathcal{Y}} - \tilde{\mathcal{Y}}^\top \hat{C}^\top + \hat{C} \mathcal{M}_{(0,T)} \hat{C}^\top \right)$$

$$\hat{W} = \frac{1}{NT} \left(\mathcal{M}_{(1,T)} - \mathcal{M}_\Delta^\top \hat{F}_{ss}^\top \right.$$

$$\left. - \hat{F}_{ss} \mathcal{M}_\Delta + \hat{F}_{ss} \mathcal{M}_{(0,T-1)} \hat{F}_{ss}^\top \right)$$

$$\tag{24}$$

The explicit updates (24) further speed up the computational routine as the numerical approximation of $\nabla_\theta \mathcal{Q}$ is not needed for optimization.

Remark (Closed-loop identification). Under the observation model (22a), $_A\mathcal{Q}_{(a)} = 0$ $_B\mathcal{Q}_{(a)} = 0$ and $_{K_{ss}}\mathcal{Q}_{(a)} = 0$ are redundant. They are all trivially satisfied when $\hat{F}_{ss}$ is as given in (24), highlighting a lack of identifiability of the inner parameters of the closed-loop.

b) Partial state and control observations: The *E-step* for the second observation model (22b) is computed analogously, and we introduce the following partitions of the sufficient statistics (23c)

$$Y^{(i)} = \sum_{t=0}^{T} \begin{bmatrix} y_{x,t}^{(i)} (y_{x,t}^{(i)})^\top & y_{x,t}^{(i)} (y_{u,t}^{(i)})^\top \\ y_{u,t}^{(i)} (y_{x,t}^{(i)})^\top & y_{u,t}^{(i)} (y_{u,t}^{(i)})^\top \end{bmatrix} =: \begin{bmatrix} Y_{xx}^{(i)} & Y_{xu}^{(i)} \\ (Y_{xu}^{(i)})^\top & Y_{uu}^{(i)} \end{bmatrix} \tag{25a}$$

and

$$\tilde{Y}^{(i)} = \sum_{t=0}^{T} m_t^{(i)} \begin{bmatrix} y_{x,t}^{(i)\top} & y_{u,t}^{(i)\top} \end{bmatrix} =: \begin{bmatrix} \tilde{Y}_{xx}^{(i)} & \tilde{Y}_{xu}^{(i)} \end{bmatrix} \tag{25b}$$

and accordingly denote $\mathcal{Y}_{xx} := \sum_{i=1}^{N} Y_{xx}^{(i)}$, $\mathcal{Y}_{xu} := \sum_{i=1}^{N} Y_{xu}^{(i)}$, $\mathcal{Y}_{uu} := \sum_{i=1}^{N} Y_{uu}^{(i)}$, $\tilde{\mathcal{Y}}_{xx} := \sum_{i=1}^{N} \tilde{Y}_{xx}^{(i)}$, and $\tilde{\mathcal{Y}}_{xu} := \sum_{i=1}^{N} \tilde{Y}_{xu}^{(i)}$.

The parameter updates $\hat{F}_{ss}$, $\hat{W}_0$, and $\hat{W}$ are as provided in (24). $\hat{C}_x$ is obtained by replacing $\tilde{\mathcal{Y}} \to \tilde{\mathcal{Y}}_{xx}$ in the $\hat{C}$ update in (24), and $\hat{V}_x$ is obtained by replacing $\mathcal{Y} \to \mathcal{Y}_{xx}$ and $\hat{C} \to \hat{C}_x$ in the $\hat{V}$ update in (24). Finally,

$$\hat{V}_u = \frac{1}{N(T+1)} \left(\hat{K}_{ss} \tilde{\mathcal{Y}}_{xu} + \tilde{\mathcal{Y}}_{xu}^\top \hat{K}_{ss}^\top \right.$$

$$\left. + \hat{K}_{ss} \mathcal{M}_{(0,T)} \hat{K}_{ss}^\top + \mathcal{Y}_{uu} \right) \tag{26a}$$

$$\hat{K}_{ss} = \hat{V}_u \hat{B}^\top \hat{W}^{-1} \left(\hat{F}_{ss} - \mathcal{M}_\Delta^\top \mathcal{M}_{(0,T-1)}^{-1} \right)$$

$$- \tilde{\mathcal{Y}}_{xu}^\top \mathcal{M}_{(0,T-1)}^{-1} \tag{26b}$$

We note that the incorporation of control input observations to the model (22b) makes $_{K_{ss}}\mathcal{Q}_{(b)} = 0$ *not* redundant, providing the additional constraint (26b).

If B is known, the set of equations (26) may be solved for $\hat{V}_u$ and $\hat{K}_{ss}$. This in turn allows to identify $\hat{A}$ through the identified closed-loop as $\hat{A} = \hat{F}_{ss} + B\hat{K}_{ss}$ In this case, the identification of the remaining parameters, namely the pair $\{Q \ R\}$, maps to an infinite-horizon inverse optimal control problem. Given a stabilizing feedback gain $\hat{K}_{ss}$, as far as we know, there is no analytical solution to the problem of finding all possible pairs $\{\hat{Q} \ \hat{R}\}$ that render $\hat{K}_{ss}$ optimal. Indeed, a manifold of possible $\{\hat{Q} \ \hat{R}\}$ pairs yielding $\hat{K}_{ss}$ optimal is expected due to well-established non-uniqueness results [16]. Additional criteria or constraints can be introduced in the problem to uniquely solve for $\{\hat{Q} \ \hat{R}\}$ – see, e.g. [16].

REFERENCES

[1] Thomas R Knösche and Jens Haueisen. *EEG/MEG source reconstruction: textbook for electro-and magnetoencephalography*. Springer, 2022.

[2] Wei Wu, Yun Gao, Elie Bienenstock, John P. Donoghue, and Michael J. Black. Bayesian population decoding of motor cortical activity using a kalman filter. *Neural Computation*, 18(1):80–118, January 2006.

[3] Byron M. Yu, Krishna V Shenoy, and Maneesh Sahani. Expectation propagation for inference in non-linear dynamical models with poisson observations. In *2006 IEEE Nonlinear Statistical Signal Processing Workshop*, pages 83–86, 2006.

[4] Marine Schimel, Ta-Chu Kao, Kristopher T Jensen, and Guillaume Hennequin. iLQR-VAE : control-based learning of input-driven dynamics with applications to neural data. In *International Conference on Learning Representations*, 2022.

[5] Victor Geadah, International Brain Laboratory, and Jonathan W. Pillow. Parsing neural dynamics with infinite recurrent switching linear dynamical systems. In *The Twelfth International Conference on Learning Representations*, 2024.

[6] Evelyn Tang and Danielle S. Bassett. Colloquium: Control of dynamics in brain networks. *Rev. Mod. Phys.*, 90:031003, Aug 2018.

[7] Ta-Chu Kao and Guillaume Hennequin. Neuroscience out of control: control-theoretic perspectives on neural circuit dynamics. *Curr. Opin. Neurobiol.*, 58:122–129, September 2019.

[8] Vivek R Athalye, Jose M Carmena, and Rui M Costa. Neural reinforcement: re-entering and refining neural dynamics leading to desirable outcomes. *Current Opinion in Neurobiology*, 60:145–154, 2020. Neurobiology of Behavior.

[9] Ta-Chu Kao, Mahdieh S Sadabadi, and Guillaume Hennequin. Optimal anticipatory control as a theory of motor preparation: A thalamocortical circuit model. *Neuron*, 109(9):1567–1581, 2021.

[10] Anastasios Tsiamis, Ingvar Ziemann, Nikolai Matni, and George J. Pappas. Statistical learning theory for control: A finite-sample perspective. *IEEE Control Systems Magazine*, 43(6):67–97, 2023.

[11] R. E. Kalman. When is a linear control system optimal? *Journal of Fluids Engineering, Transactions of the ASME*, 86(1):51–60, 1964.

[12] Antony Jameson and Eliezer Kreindler. Inverse Problem of Linear Optimal Control. *SIAM Journal on Control*, 11(1):1–19, 1973.

[13] Han Zhang, Jack Umenberger, and Xiaoming Hu. Inverse optimal control for discrete-time finite-horizon linear quadratic regulators. *Automatica*, 110:108593, December 2019.

[14] Han Zhang, Yibei Li, and Xiaoming Hu. Inverse Optimal Control for Finite-Horizon Discrete-time Linear Quadratic Regulator under Noisy Output. *Proceedings of the IEEE Conference on Decision and Control*, 2019-December(Cdc):6663–6668, 2019.

[15] Han Zhang, Yibei Li, and Xiaoming Hu. Discrete-time inverse linear quadratic optimal control over finite time-horizon under noisy output measurements. *Control Theory and Technology*, 19(4):563–572, 2021.

[16] M. Cody Priess, Richard Conway, Jongeun Choi, John M. Popovich, and Clark Radcliffe. Solutions to the inverse LQR problem with application to biological systems analysis. *IEEE Transactions on Control Systems Technology*, 23(2):770–777, 2015.

[17] Matthias Schultheis, Dominik Straub, and Constantin A. Rothkopf. Inverse Optimal Control Adapted to the Noise Characteristics of the Human Sensorimotor System. *Advances in Neural Information Processing Systems*, 12:9429–9442, 2021.

[18] Dominik Straub and Constantin A. Rothkopf. Putting perception into action with inverse optimal control for continuous psychophysics. *eLife*, 11:1–26, 2022.

[19] Philipp Karg, Simon Stoll, Simon Rothfus, and Soren Hohmann. Inverse Stochastic Optimal Control for Linear-Quadratic Gaussian and Linear-Quadratic Sensorimotor Control Models. *Proceedings of the IEEE Conference on Decision and Control*, 2022-Decem:2801–2808, 2022.

[20] Matthew Golub, Steven Chase, and Byron Yu. Learning an internal dynamics model from control demonstration. In Sanjoy Dasgupta and David McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 606–614, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

[21] Matthew D Golub, Byron M Yu, and Steven M Chase. Internal models for interpreting neural population activity during sensorimotor control. *eLife*, 4:e10015, dec 2015.

[22] Shanelle G. Clarke, Sooyung Byeon, and Inseok Hwang. A low complexity approach to model-free stochastic inverse linear quadratic control. *IEEE Access*, 10:9298–9308, 2022.

[23] Ignasi Clavera, Jonas Rothfuss, John Schulman, Yasuhiro Fujita, Tamim Asfour, and Pieter Abbeel. Model-based reinforcement learning via meta-policy optimization. In Aude Billard, Anca Dragan, Jan Peters, and Jun Morimoto, editors, *Proceedings of The 2nd Conference on Robot Learning*, volume 87 of *Proceedings of Machine Learning Research*, pages 617–629. PMLR, 29–31 Oct 2018.

[24] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, MA, USA, 3rd edition, 2011.

[25] Zoubin Ghahramani and Geoffrey H Hinton. Parameter estimation for linear dynamical systems. *Technical Report CRG-TR-96-2, Dept. Comp. Sci., Univ. Toronto*, pages 6–11, 1996.

[26] Erich L. Lehmann and George Casella. *Theory of Point Estimation*. Springer-Verlag, New York, NY, USA, second edition, 1998.

[27] R.E. Kalman. On the general theory of control systems. *IFAC Proceedings Volumes*, 1(1):491–502, August 1960.

[28] Brian DO Anderson and John B Moore. *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

[29] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[30] Sam Roweis and Zoubin Ghahramani. A unifying review of linear gaussian models. *Neural computation*, 11(2):305–345, 1999.

[31] Robert Shumway and David S. Stoffer. An approach to time series smoothing and forecasting using the em algorithm. *Journal of Time Series Analysis*, 3:253–264, 1982.

[32] A.P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via de EM algorithm. *Journal of the Royal Statistical Society. Series B.*, 39(1):1–38, 1977.

[33] Elizabeth E. Holmes. Derivation of an em algorithm for constrained and unconstrained multivariate autoregressive state-space (marss) models, 2013.

[34] R. E. Kalman. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, 1(2):152–192, January 1963.

[35] Lawrence Crone. Second order adjoint matrix equations. *Linear Algebra and its Applications*, 39:61–71, August 1981.

[36] D. Kleinman. On an iterative technique for riccati equation computations. *IEEE Transactions on Automatic Control*, 13(1):114–115, 1968.

[37] Risto J Ilmoniemi and Jukka Sarvas. *Brain signals: physics and mathematics of MEG and EEG*. Mit Press, 2019.

[38] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. 2006.