

Theory for optimal estimation and control under resource limitations and its applications to biological information processing and decision-making

Takehiro Tottori^{1,2,*} and Tetsuya J. Kobayashi^{2,†}

¹Laboratory for Neural Computation and Adaptation, *RIKEN Center for Brain Science*, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

²Institute of Industrial Science, *The University of Tokyo*, 4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan



(Received 25 September 2024; accepted 16 September 2025; published 14 October 2025)

While the information processing of organisms is evolutionarily optimized, it exhibits remarkable diversity in its complexity and capability. One potential source of this diversity is the limitations of resources available for information processing. However, we still lack a theoretical framework that clarifies the relationship between biological information processing and resource limitations and integrates it with downstream decision-making. Together with the accompanying letter [Tottori and Kobayashi, *Phys. Rev. Res.* **7**, L042012 (2025)], this full paper presents an optimal estimation and control framework that explicitly incorporates resource limitations inherent in biological systems. While the accompanying letter focuses on the nontrivial phase transitions induced by resource limitations, the present full paper establishes the theoretical foundation and provides a comprehensive analysis of these transitions. The proposed framework formulates memory that organisms can store and operate, and optimizes it using optimal control theory. This approach accounts for various resource limitations—such as memory capacity, intrinsic noise, and energy cost—and unifies estimation and control. Applying this theory to minimal models of biological information processing and decision-making, we demonstrate that resource limitations can induce discontinuous and nonmonotonic phase transitions between memoryless and memory-based strategies. These results provide a comprehensive theoretical foundation for studying how organisms process information and make decisions under resource limitations, and suggest that the rich diversity observed in biological systems may emerge from such constraints.

DOI: [10.1103/gvl6-cvby](https://doi.org/10.1103/gvl6-cvby)

I. INTRODUCTION

A wide variety of organisms adapt to dynamic environments through sensory information processing. Recent studies have demonstrated that the information processing is nearly optimal not only in humans [1–6] but also in simpler organisms such as bacteria [7–15]. Despite being optimized, the information processing of organisms is highly diverse, as evidenced by the significant differences between the information processing of humans and that of bacteria. This raises the following fundamental question: What causes the diversity of biological information processing? One potential cause of this diversity could be the limitations of resources available for information processing in organisms. While human brains consume a substantial amount of energy and process information through billions of neurons [16–25], bacteria have access to extremely limited energy and rely on chemical reactions involving only a small number of molecules to process in-

formation [26–49]. The wide variation in resource limitations may provide a unifying perspective on the diversity observed in biological information processing (Fig. 1).

However, a comprehensive theoretical framework for investigating information processing under resource limitations has yet to be developed. Bayesian filtering theory is a well-established method for estimating dynamic environments from noisy observation histories and has numerous engineering applications; see Fig. 2(a) and Refs. [51–54]. As exemplified by the Bayesian brain hypothesis [55–61], this theory is regarded as a fundamental principle of neural information processing, and there is a substantial body of research aimed at explaining brain functions from the perspective of Bayesian filtering [62–67]. The approach used in the Bayesian brain hypothesis has recently applied to bacteria without brains to investigate the problem whether and how bacteria can perform Bayesian filtering through intracellular chemical reactions [13–15, 68–77]. While Bayesian filtering provides statistically optimal estimations, it fails to take into account the resource limitations inherent in biological information processing.

In particular, there are three key resource limitations that Bayesian filtering theory cannot capture. These limitations become especially inevitable in primitive organisms such as bacteria. The first limitation is memory capacity. To perform Bayesian filtering, organisms need to store posterior probabilities. However, this task is challenging for organisms with limited memory capacity, as the posterior probabilities

*Contact author: takehiro.tottori@riken.jp

†Contact author: tetsuya@mail.crmind.net

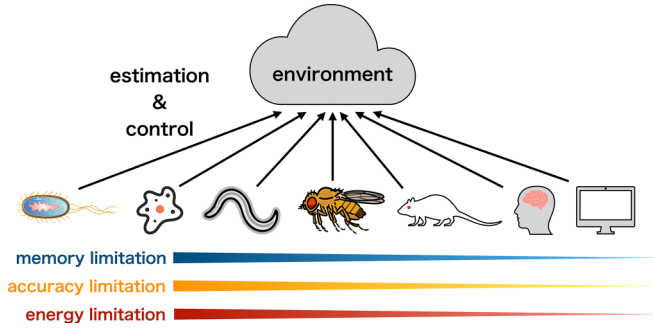


FIG. 1. Schematic diagram of diverse information processing and resource limitations. Resource limitations become more pronounced in primitive organisms such as bacteria. A part of this figure is from TogoTV (© 2016 DBCLS TogoTV, CC-BY-4.0 [50]).

become infinite dimensional even when the environmental states are finite dimensional. While brains can approximate infinite-dimensional posterior probabilities using numerous neurons [55,56,59,78,79], it is far more challenging for bacteria to represent these probabilities with only a small number of molecules. In fact, previous studies implementing Bayesian filtering through intracellular chemical reactions are limited to special cases where the infinite-dimensional posterior probabilities can be reduced to finite-dimensional sufficient statistics to circumvent this issue [13–15,68–77].

The second limitation is the accuracy of information processing. Biological information processing based on either neural firing or intracellular chemical reactions often involves high levels of stochasticity [26–34,80–82], making it diffi-

cult to accurately compute the Bayesian update of posterior probabilities. Enhancing the accuracy of biological information processing requires increasing the number of neurons or molecules involved in Bayesian filtering, which in turn demands greater resource investment. In previous studies, while extrinsic noise arising from environments and observations has been considered, intrinsic noise within internal systems performing Bayesian filtering has often been neglected [13–15,68–77]. However, this assumption is often unrealistic, particularly in the context of cellular information processing.

The third limitation is the cost associated with information processing. Since biological information processing is realized through physical processes such as neural firing and chemical reactions, it inevitably requires costs [16–25,35–49]. Experimental evolution studies indicate that excessive investment of energy or other costs in information processing can reduce the fitness of organisms [83–88]. Therefore, it is crucial for organisms not only to minimize estimation errors, but also to reduce the cost required for these computations. In fact, some studies have discussed the trade-off between estimation errors and the costs implicitly via the information bottleneck method [40,89–95]. Also, other studies considered the energy costs from the viewpoints of thermodynamics [35–39,41–49]. However, their connections to Bayesian filtering theory are still elusive.

In summary, Bayesian filtering theory fails to consider resource limitations inherent in biological systems, such as memory capacity, intrinsic noise, and energy cost. This highlights the need for a theoretical framework to elucidate the relationship between information processing and resource limitations in biological systems.

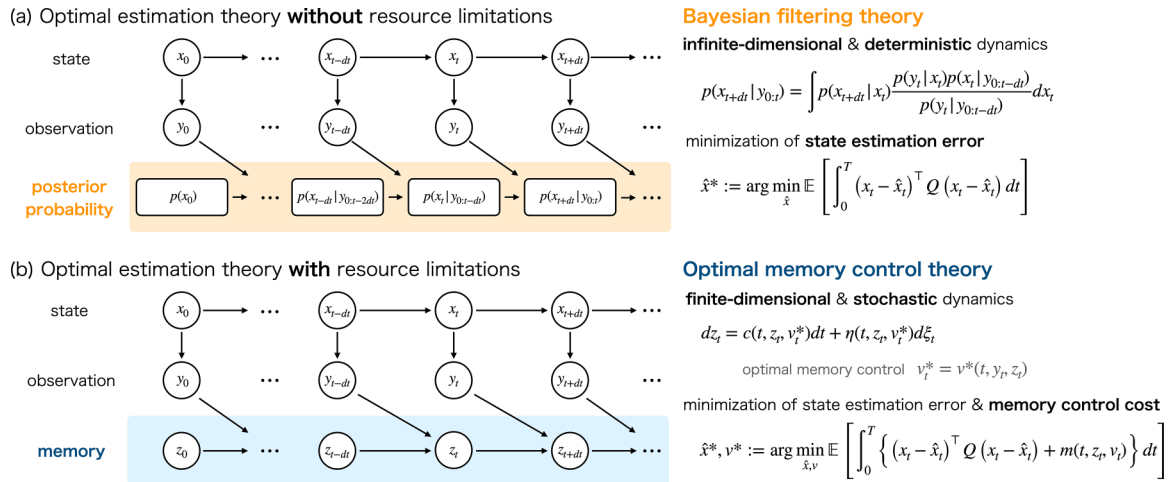


FIG. 2. Schematic diagrams of (a) the conventional optimal estimation theory without resource limitations and (b) the proposed optimal estimation theory with resource limitations. (a) In the conventional Bayesian filtering theory, an agent compresses the information about past observation history $y_{0:t-dt} := \{y_0, \dots, y_{t-dt}\}$ into posterior probability $p(x_t | y_{0:t-dt})$ and estimates environmental state x_t from current observation y_t and posterior probability $p(x_t | y_{0:t-dt})$. To obtain posterior probability $p(x_t | y_{0:t-dt})$, the agent needs to compute infinite-dimensional and deterministic Bayesian update. In addition, optimal state estimator function $\hat{x}$ is determined by minimizing only the state estimation error. These facts ignore the resource limitations of the agent. (b) In the optimal memory control theory, the agent compresses the information about past observation history $y_{0:t-dt}$ into finite-dimensional memory z_t and estimates environmental state x_t from current observation y_t and memory z_t . The memory dynamics includes a noise process, ξ_t , which represents the intrinsic noise of the agent. Furthermore, the memory dynamics is optimized by optimal memory control v_t^* rather than by Bayes's theorem. Optimal state estimator function $\hat{x}^*$ and optimal memory control function v^* are determined by minimizing not only the state estimation error but also the memory control cost. This problem formulation allows us to take into account the resource limitations of the agent.

In addition to estimating environmental states, organisms must control their behaviors based on these estimations. The decision-making of organisms has been extensively studied using optimal control theory [96–100]. The combination of Bayesian filtering and optimal control is one approach to understanding decision-making under uncertain observations [101–107]. However, this combination complicates the optimization process, resulting in studies being limited to simple cases [15,108–111]. Therefore, there is also a need for a theoretical framework that integrates estimation and control in a more comprehensive manner.

To address the issues outlined above, this full paper and the accompanying letter [112] propose an optimal estimation theory that takes into account resource limitations [Fig. 2(b)]. While the accompanying letter [112] focuses on the nontrivial phase transitions induced by resource limitations, the present full paper establishes the theoretical foundation and provides a comprehensive analysis of these transitions. Furthermore, while this theoretical framework builds upon our previously proposed framework [113–115], the formulation is slightly modified for biological plausibility.

The proposed theory explicitly formulates finite-dimensional and stochastic memories that organisms can store and operate, and optimizes the embedding of information from past observation histories into these memories using optimal control theory. This approach accepts a more general objective function that includes not only estimation errors but also energy costs. Therefore, this theory achieves optimal estimations under resource limitations by replacing Bayesian filtering of posterior probabilities with the optimal control of memories.

We apply this theory to a minimal model of biological information processing and investigate optimal estimation strategies under resource limitations. Although our minimal model considers a simple optimal estimation problem with linear Gaussian dynamics and quadratic costs, it exhibits discontinuous and nonmonotonic phase transitions between memoryless and memory-based estimation strategies depending on available resources. Such phase transitions do not occur in the conventional Bayesian filtering that disregards resource limitations. Therefore, our analysis suggests that resource limitations may drive a punctuated and complex evolution of biological information processing, leading to qualitatively diverse behaviors across organisms.

This theory also provides a unified framework for both state estimation and control by utilizing memory control instead of Bayesian filtering. We demonstrate the effectiveness of this approach by applying it to a minimal model of biological information processing and decision-making. This model also exhibits discontinuous phase transitions between memoryless and memory-based control strategies. Yet, nonmonotonic phase transitions do not appear with respect to sensory noise, while they occur with respect to environmental uncertainty. These findings suggest that discontinuity may be a more fundamental characteristic than nonmonotonicity under resource limitations.

This paper is organized as follows: In Sec. II, we briefly review Bayesian filtering theory and explain how this theory fails to capture the resource limitations inherent in biological systems. In Sec. III, we introduce an optimal estimation the-

ory that can take into account these resource limitations. In Sec. IV, we apply this theory to a minimal model of biological information processing and investigate optimal estimation strategies under resource limitations. In Sec. V, we extend the optimal estimation theory to incorporate optimal control. In Sec. VI, we apply this extended theory to a minimal model of biological information processing and decision-making. In Sec. VII, we discuss future directions and extensions of our work.

In Appendixes A and B, we derive the optimal solutions for the optimal estimation and control problems under resource limitations. Specifically, Appendixes A and B address the cases without and with memory, respectively. In Appendix C, we discuss the linear-quadratic-Gaussian (LQG) problem in this theoretical framework. In Appendixes D and E, we further analyze the minimal models of biological information processing and decision-making. In Appendix F, we compare the present work with our previous works [113–115]. In Appendix G, we demonstrate the consistency between our theoretical framework and the Kalman filter, which is a Bayesian filter in the LQG problem [51–54,106,116–118].

II. OPTIMAL ESTIMATION THEORY WITHOUT RESOURCE LIMITATIONS

In this section, we briefly review the conventional optimal theory for estimating dynamic environments, Bayesian filtering theory; see Fig. 2(a) and Refs. [51–54]. This theory considers an environment and an agent. The state of the environment at time $t \in [0, T]$ is denoted by $x_t \in \mathbb{R}^{d_x}$, which evolves by the following stochastic differential equation (SDE):

$$dx_t = b(t, x_t)dt + \sigma(t, x_t)d\omega_t, \quad (1)$$

where initial state x_0 follows $p_0(x_0)$ and $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process. The agent cannot completely observe environmental state x_t . Instead, it obtains noisy observation $y_t \in \mathbb{R}^{d_y}$, which is generated from probability density function $p_t(y_t | x_t)$ as follows:

$$y_t \sim p_t(y_t | x_t). \quad (2)$$

The agent estimates environmental state x_t from noisy observation history $y_t, y_{t-dt}, \dots, y_0$. The state estimator of the agent at time t is generally given by

$$\hat{x}_t = \hat{x}(t, y_t, y_{t-dt}, \dots, y_0). \quad (3)$$

The objective function the agent minimizes is the following state estimation error:

$$J[\hat{x}] := \mathbb{E} \left[\int_0^T (x_t - \hat{x}_t)^\top Q (x_t - \hat{x}_t) dt \right], \quad (4)$$

where $Q \succ O$ is the parameter representing the weight of the state estimation error. Bayesian filtering theory considers a problem to find optimal state estimator function $\hat{x}^*$ that minimizes state estimation error $J[\hat{x}]$:

$$\hat{x}^* := \arg \min_{\hat{x}} J[\hat{x}]. \quad (5)$$

By solving Eq. (5), optimal state estimator function $\hat{x}^*$ is obtained by the conditional expectation of state x_t given

observation history $y_t, y_{t-dt}, \dots, y_0$:

$$\hat{x}^*(t, y_t, y_{t-dt}, \dots, y_0) = \mathbb{E}_{p(x_t|y_t, y_{t-dt}, \dots, y_0)}[x_t]. \quad (6)$$

In order to obtain optimal state estimator $\hat{x}_t^* = \hat{x}^*(t, y_t, y_{t-dt}, \dots, y_0)$, the agent needs to memorize past observation history $y_{t-dt}, \dots, y_0$ and it becomes harder and harder as time progresses. Bayesian filtering theory resolves this problem by compressing past observation history $y_{t-dt}, \dots, y_0$ into the posterior probability density function

$$p(x_t|y_{t-dt}, \dots, y_0) \in \mathcal{P}[\mathbb{R}^{d_x}], \quad (7)$$

where $\mathcal{P}[\mathbb{R}^{d_x}]$ indicates the space of a probability density function on $\mathbb{R}^{d_x}$. The optimal state estimator is given by

$$\hat{x}_t^* = \mathbb{E}_{p(x_t|y_t, p(x_t|y_{t-dt}, \dots, y_0))}[x_t], \quad (8)$$

which means that the agent estimates environmental state x_t from observation y_t and posterior probability density function $p(x_t|y_{t-dt}, \dots, y_0)$. This compression does not impair the performance because $p(x_t|y_{t-dt}, \dots, y_0)$ is the sufficient statistic of $y_{t-dt}, \dots, y_0$ for estimating x_t . $p(x_t|y_{t-dt}, \dots, y_0)$ can be updated sequentially by using Bayes's theorem:

$$p(x_{t+dt}|y_t, y_{t-dt}, \dots, y_0) = \int p(x_{t+dt}|x_t) \frac{p(y_t|x_t)p(x_t|y_{t-dt}, \dots, y_0)}{p(y_t|y_{t-dt}, \dots, y_0)} dx_t. \quad (9)$$

Therefore, in Bayesian filtering theory, the agent obtains optimal state estimator $\hat{x}_t^*$ [Eq. (8)] by computing the Bayesian update of the posterior probability [Eq. (9)].

However, Bayesian filtering is an ideal theory for optimal estimation and overlooks the resource limitations inherent in organisms. This oversight prevents organisms from fully implementing Bayesian filtering. For example, organisms cannot completely store infinite-dimensional posterior probability $p(x_t|y_{t-dt}, \dots, y_0)$ because their memory capacities are finite [13–15, 68–77]. Additionally, organisms find it challenging to accurately compute the Bayesian update of the posterior probability [Eq. (9)] due to the intrinsic noise in their information processing [26–34, 80–82]. Furthermore, organisms need to minimize not only the state estimation error but also the energy consumption associated with the estimation process, as their available energy is severely limited [16–25, 35–49]. These resource limitations become more pronounced and increasingly difficult to ignore in more primitive organisms such as bacteria. Therefore, an optimal estimation theory that takes into account these limitations is necessary to understand the estimation strategies of organisms.

III. OPTIMAL ESTIMATION THEORY WITH RESOURCE LIMITATIONS

In this section, we present an optimal estimation theory that takes into account the resource limitations inherent in biological systems [Fig. 2(b)]. This theory builds upon our previous works [113–115], but the problem formulation is slightly modified, which allows us to discuss biologically plausible situations (see Appendix F). In the following, we first formulate the optimal estimation problem with resource limitations in Sec. III A and then present its optimal solution in Sec. III B.

A. Problem formulation

In this subsection, we formulate the optimal estimation problem with resource limitations. The state x_t and observation y_t are identical to those in the Bayesian filtering problem, which are given by Eqs. (1) and (2), respectively. In contrast, this problem explicitly formulates the finite-dimensional memory that the agent can store and operate,

$$z_t \in \mathbb{R}^{d_z}, \quad (10)$$

where the dimension of the memory, d_z , is determined by the available memory size of the agent. The agent determines state estimator $\hat{x}_t$ based on observation y_t and memory z_t as follows:

$$\hat{x}_t = \hat{x}(t, y_t, z_t). \quad (11)$$

Compared to Eq. (3), infinite-dimensional past observation history $y_{t-dt}, \dots, y_0$ is replaced with finite-dimensional memory z_t , which allows for the implementation in organisms with finite memory capacities.

Memory z_t is assumed to evolve by the following SDE:

$$dz_t = c(t, z_t, v_t)dt + \eta(t, z_t, v_t)d\xi_t, \quad (12)$$

where initial memory z_0 follows $p_0(z_0)$, $v_t \in \mathbb{R}^{d_v}$ is the memory control of the agent, and $\xi_t \in \mathbb{R}^{d_\xi}$ is a standard Wiener process. The agent determines memory control v_t based on observation y_t and memory z_t :

$$v_t = v(t, y_t, z_t). \quad (13)$$

While the dynamics of the posterior probability [Eq. (9)] is deterministic given observation y_t , that of the memory [Eq. (12)] is stochastic due to the intrinsic noise, ξ_t .

A key challenge in this problem lies in how to optimally compress the information about past observation history $y_{t-dt}, \dots, y_0$ into memory z_t . Bayesian filtering theory optimally compresses the past observation information into infinite-dimensional and deterministic posterior probability $p(x_t|y_{t-dt}, \dots, y_0)$ using Bayes's theorem. However, this optimization method cannot be applied to memory z_t because it is finite dimensional and stochastic. We address this issue by using optimal control theory: The agent optimally compresses the past observation information into memory z_t by optimizing memory control v_t . This idea leads to the simultaneous optimization of state estimator function $\hat{x}$ and memory control function v as follows:

$$\hat{x}^*, v^* := \arg \min_{\hat{x}, v} J[\hat{x}, v]. \quad (14)$$

Objective function $J[\hat{x}, v]$ is given by the following expected cumulative cost function:

$$J[\hat{x}, v] := \mathbb{E} \left[\int_0^T f(t, x_t, z_t, \hat{x}_t, v_t) dt + g(x_T, z_T) \right], \quad (15)$$

where $f(t, x_t, z_t, \hat{x}_t, v_t)$ and $g(x_T, z_T)$ are the running and terminal cost functions, respectively.

For comparison with Bayesian filtering theory, we consider the following cost functions as a special case:

$$f := (x_t - \hat{x}_t)^\top Q(x_t - \hat{x}_t) + m(t, z_t, v_t), \quad (16)$$

$$g := 0, \quad (17)$$

where $m(t, z_t, v_t)$ is the cost function for memory z_t and memory control v_t . These cost functions indicate that our theory can minimize not only the state estimation error [the first term of Eq. (16)] but also the energy cost associated with the estimation process [the second term of Eq. (16)].

In summary, this theory achieves optimal estimation under resource limitations, including memory capacity, intrinsic noise, and energy cost, by replacing the Bayesian filtering of the posterior probability with the optimal control of the memory.

B. Optimal solution

The optimal estimation problem with resource limitations can be solved by using Pontryagin's minimum principle on the probability density function space, which has been recently employed in mean-field stochastic optimal control problems [119–122] and our previous works [113–115]. In this subsection, we skip the derivation and only show the main result. The details of the derivation are shown in Appendixes A and B.

Optimal state estimator function $\hat{x}^*$ and optimal memory control function v^* are obtained by minimizing the conditional expected Hamiltonian:

$$\hat{x}^*(t, y, z), v^*(t, y, z) = \arg \min_{\hat{x}, v} \mathbb{E}_{p_t(x|y, z)}[\mathcal{H}(t, x, z, \hat{x}, v, w)]. \quad (18)$$

Hamiltonian $\mathcal{H}$ is defined by

$$\mathcal{H}(t, x, z, \hat{x}, v, w) := f(t, x, z, \hat{x}, v) + \mathcal{L}_v w(t, x, z), \quad (19)$$

where $f(t, x, z, \hat{x}, v)$ is the current cost and $\mathcal{L}_v w(t, x, z)$ represents the future cost. $\mathcal{L}_v$ is the time-backward diffusion operator for state x and memory z , which is defined by

$$\begin{aligned} \mathcal{L}_v w(t, x, z) := & \sum_{i=1}^{d_x} b_i(t, x) \frac{\partial w(t, x, z)}{\partial x_i} \\ & + \sum_{i=1}^{d_z} c_i(t, z, v) \frac{\partial w(t, x, z)}{\partial z_i} \\ & + \frac{1}{2} \sum_{i,j=1}^{d_x} D_{ij}(t, x) \frac{\partial^2 w(t, x, z)}{\partial x_i \partial x_j} \\ & + \frac{1}{2} \sum_{i,j=1}^{d_z} F_{ij}(t, z, v) \frac{\partial^2 w(t, x, z)}{\partial z_i \partial z_j}, \end{aligned} \quad (20)$$

where $D(t, x) := \sigma(t, x)\sigma^\top(t, x)$ and $F(t, z, v) := \eta(t, z, v)\eta^\top(t, z, v)$. $w(t, x, z)$ is the value function, which represents the expected cumulative cost starting from time t with state x and memory z to terminal time T . Value function $w(t, x, z)$ is obtained by solving the following time-backward partial differential equation, Hamilton-Jacobi-Bellman (HJB) equation:

$$-\frac{\partial w(t, x, z)}{\partial t} = \mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, z, \hat{x}^*, v^*, w)], \quad (21)$$

where the terminal condition is given by $w(T, x, z) = g(x, z)$. In the optimal solution [Eq. (18)], conditional expectation

$\mathbb{E}_{p_t(x|y, z)}[\cdot]$ appears, which is consistent with the fact that the agent needs to estimate state x from observation y and memory z . Conditional probability density function $p_t(x|y, z)$ is calculated as follows:

$$p_t(x|y, z) = \frac{p_t(y|x)p_t(x, z)}{\int p_t(y|x)p_t(x, z)dx}. \quad (22)$$

Probability density function $p_t(x, z) := p(t, x, z)$ is obtained by solving the following time-forward partial differential equation, Fokker-Planck (FP) equation:

$$\frac{\partial p(t, x, z)}{\partial t} = \tilde{\mathcal{L}}_v^* p(t, x, z), \quad (23)$$

where the initial condition is given by $p(0, x, z) = p_0(x)p_0(z)$. $\tilde{\mathcal{L}}_v^*$ is the expected time-forward diffusion operator for state x and memory z , which is defined by

$$\begin{aligned} \tilde{\mathcal{L}}_v^* p(t, x, z) := & - \sum_{i=1}^{d_x} \frac{\partial b_i(t, x)p(t, x, z)}{\partial x_i} \\ & - \sum_{i=1}^{d_z} \frac{\partial \mathbb{E}_{p_t(y|x)}[c_i(t, z, v)]p(t, x, z)}{\partial z_i} \\ & + \frac{1}{2} \sum_{i,j=1}^{d_x} \frac{\partial^2 D_{ij}(t, x)p(t, x, z)}{\partial x_i \partial x_j} \\ & + \frac{1}{2} \sum_{i,j=1}^{d_z} \frac{\partial^2 \mathbb{E}_{p_t(y|x)}[F_{ij}(t, z, v)]p(t, x, z)}{\partial z_i \partial z_j}, \end{aligned} \quad (24)$$

where $\mathbb{E}_{p_t(y|x)}[\cdot]$ is necessary because memory control v depends on observation y . Therefore, optimal state estimator function $\hat{x}^*$ and optimal memory control function v^* are obtained by jointly solving HJB equation [Eq. (21)] and FP equation [Eq. (23)].

It should be noted that organisms do not need to solve HJB and FP equations [Eqs. (21) and (23)] because they may obtain optimal solutions $\hat{x}^*$ and v^* as a result of evolution. Therefore, it is not necessary to consider the resource limitations of organisms for the computation of optimal solutions $\hat{x}^*$ and v^* . In contrast, it is necessary to consider their resource limitations associated with the dynamics of memory z_t because memory z_t depends on real-time observation y_t such as olfactory, auditory, and visual sensing information.

If cost function f is given by Eq. (16), optimal state estimator function $\hat{x}^*$ [Eq. (18)] is given by

$$\hat{x}^*(t, y_t, z_t) = \mathbb{E}_{p(x_t|y_t, z_t)}[x_t]. \quad (25)$$

A comparison of Eqs. (6) and (25) clarifies that our theory replaces infinite-dimensional past observation history $y_{t-dt}, \dots, y_0$ with finite-dimensional memory z_t .

IV. APPLICATION TO BIOLOGICAL INFORMATION PROCESSING

In this section, we apply our theory to a minimal model of biological information processing [Fig. 3(a)] to investigate optimal estimation strategies under resource limitations. This

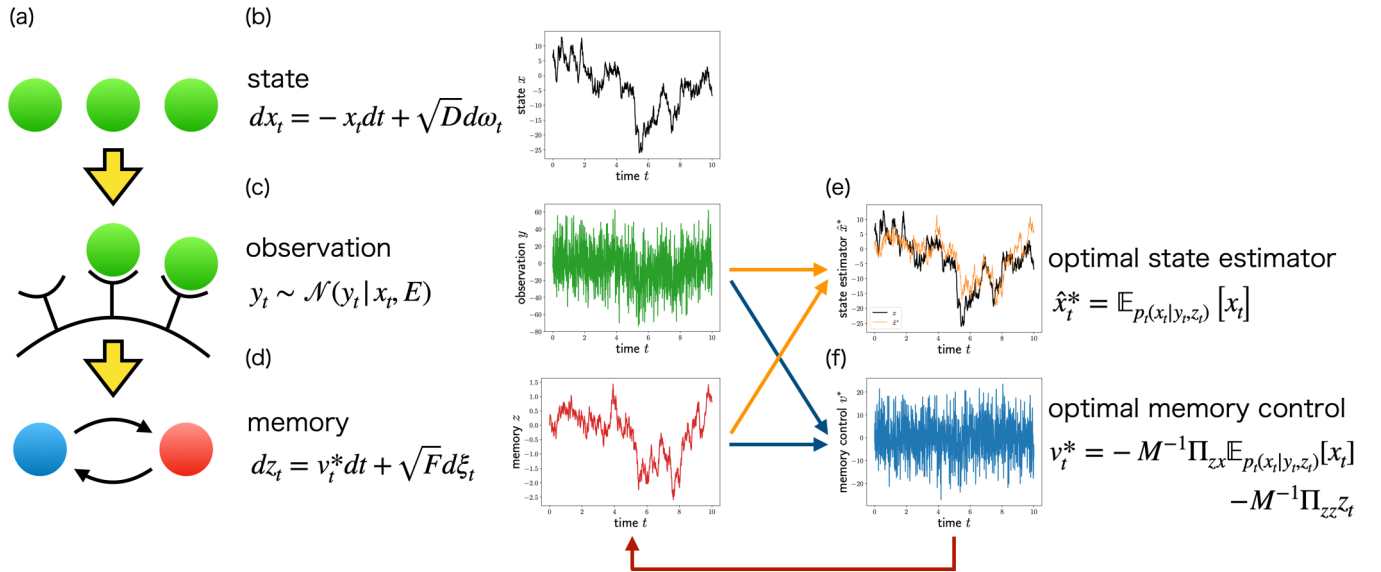


FIG. 3. Schematic diagram of a minimal model of biological information processing. (a) A cell achieves adaptive behavior by accurately perceiving environmental ligands through receptor activities and intracellular chemical reactions. The environmental ligands, receptor activities, and intracellular chemical reactions correspond to the state $x_t \in \mathbb{R}$, observation $y_t \in \mathbb{R}$, and memory $z_t \in \mathbb{R}$, respectively. Stochastic simulation of state x_t [panel (b)], observation y_t [panel (c)], memory z_t [panel (d)], optimal state estimator $\hat{x}_t^*$ [panel (e)], and optimal memory control v_t^* [panel (f)]. Optimal state estimator $\hat{x}_t^*$ and optimal memory control v_t^* are determined based on observation y_t and memory z_t (orange and blue arrows, respectively). Optimal memory control v_t^* shapes the optimal dynamics of memory z_t (red arrow). For comparison, true state x_t (black line) and estimated state $\hat{x}_t^*$ (orange line) are plotted on the same panel [panel (e)]. The parameters are $D = 100$, $E = 500$, $F = 1$, $Q = 10$, and $M = 1$.

section is organized as follows: In Sec. IV A, we formulate a minimal model of biological information processing. In Sec. IV B, we show its optimal solution. In Sec. IV C, we demonstrate the validity of the optimal solution by conducting a stochastic simulation. In Sec. IV D, we investigate the parameter dependence of the optimal solution and find discontinuous phase transitions between memoryless and memory-based estimation strategies. In Sec. IV E, we examine the phase transitions by analyzing the landscapes of the objective function, the state estimation error, and the memory control cost. In Sec. IV F, we further investigate the relationships between parameters in the phase transitions of optimal estimation strategies.

A. Problem formulation

In this subsection, we formulate the minimal model of biological information processing under resource limitations [Fig. 3(a)]. We illustrate this model using cellular information processing, but it should be noted that this model is general and can be applied to the information processing of other systems such as brains and robots.

A cell has to accurately perceive environmental state $x_t \in \mathbb{R}$ to achieve various physiological functions such as chemotaxis, growth, and differentiation. Environmental state $x_t \in \mathbb{R}$ is assumed to evolve by the following Ornstein-Uhlenbeck process:

$$dx_t = -x_t dt + \sqrt{D} d\omega_t, \quad (26)$$

where initial state x_0 follows standard Gaussian distribution $\mathcal{N}(x_0 | 0, 1)$ and $\omega_t \in \mathbb{R}$ is a standard Wiener process. $D \geq 0$ represents the intensity of the state noise.

The cell observes environmental state x_t through receptor activities, which correspond to observation $y_t \in \mathbb{R}$. Observation y_t is generated from the following Gaussian distribution:

$$y_t \sim \mathcal{N}(y_t | x_t, E), \quad (27)$$

where $E \geq 0$ represents the intensity of the observation noise.

The cell memorizes the information about past observation history $y_{t-dt}, \dots, y_0$ through intracellular chemical molecules, which correspond to memory $z_t \in \mathbb{R}$. Memory $z_t \in \mathbb{R}$ evolves by the following SDE:

$$dz_t = v_t dt + \sqrt{F} d\xi_t, \quad (28)$$

where initial memory z_0 follows standard Gaussian distribution $\mathcal{N}(z_0 | 0, 1)$, $v_t \in \mathbb{R}$ is the memory control, and $\xi_t \in \mathbb{R}$ is a standard Wiener process. $F \geq 0$ represents the intensity of the memory noise.

The memory dynamics corresponds to intracellular chemical reactions in cellular information processing. v_t and F represent the intensity and the noise of intracellular chemical reactions, respectively. If a cell could invest a large number of molecules in information processing, intracellular chemical reactions would become almost deterministic, resulting in $F = 0$. However, in most cases, a cell can only invest a small number of molecules in their intracellular chemical reactions due to their resource limitations, which inevitably leads to significant memory noise $F > 0$ [26–34]. While the conventional Bayesian filtering theory takes into account extrinsic noise such as state noise D and observation noise E , it fails to capture intrinsic noise such as memory noise F .

A cell determines state estimator $\hat{x}_t \in \mathbb{R}$ and memory control $v_t \in \mathbb{R}$ based on observation y_t and memory z_t such that

$\hat{x}_t = \hat{x}(t, y_t, z_t)$ and $v_t = v(t, y_t, z_t)$. These functions are simultaneously optimized as follows:

$$\hat{x}^*, v^* := \arg \min_{\hat{x}, v} J[\hat{x}, v]. \quad (29)$$

We use an objective function $J[\hat{x}, v]$ defined as follows:

$$J[\hat{x}, v] := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (Q(x_t - \hat{x}_t)^2 + M v_t^2) dt \right], \quad (30)$$

where $Q > 0$ and $M > 0$ are weighting parameters. The long-time limit $T \rightarrow \infty$ means that this problem focuses on the case where the probability density function of state x and memory z is stationary, but it should be noted that state x and memory z themselves are nonstationary.

The first term of the cost function in Eq. (30) is a quadratic function of the difference between true state x_t and estimated state $\hat{x}_t$, which represents the state estimation error. The second term of the cost function in Eq. (30) is a quadratic function of memory control v_t , representing the memory control cost. As the absolute value of memory control v_t increases, the relative impact of the memory noise decreases in the memory dynamics [Eq. (28)], which enhances the memory effectiveness and improves the state estimation. On the other hand, a larger absolute value of memory control v_t corresponds to more intense intracellular chemical reactions, leading to greater energy consumption. Therefore, the objective function [Eq. (30)] expresses the trade-off between estimation errors and energy costs.

This model considers the linear and Gaussian dynamics of the state, observation, and memory [Eqs. (26)–(28)] and the quadratic cost functions [Eq. (30)], which corresponds to an LQG problem [103–106]. LQG structures approximately emerge when systems fluctuate around their stable fixed points (see Appendix D 1). Therefore, the analysis of LQG models may provide insight into a broad range of biological phenomena.

Furthermore, by applying the linear noise approximation, this model can be described by a push-pull network in cellular information processing [13,36–41,70], where the state, observation, and memory correspond to ligands, receptors, and signaling molecules, respectively (see Appendix D 2). While the thermodynamic cost of intracellular chemical reactions is not inherently quadratic, the quadratic cost function is widely adopted in biological modeling, including not only for chemical reactions [123–125] but also for neural computations [126–129], owing to its mathematical simplicity and analytical tractability. In line with these previous studies, we adopt a quadratic cost function in our minimal model.

B. Optimal solution

Since the minimal model of biological information processing corresponds to the LQG problem, it allows for a semianalytical calculation of optimal state estimator function $\hat{x}^*$ and optimal memory control function v^* [Eq. (18)] as follows (see Appendixes C and D):

$$\hat{x}^*(t, y, z) = \mathbb{E}_{p_t(x|y,z)}[x], \quad (31)$$

$$v^*(t, y, z) = -M^{-1} \Pi_{zx} \mathbb{E}_{p_t(x|y,z)}[x] - M^{-1} \Pi_{zz} z. \quad (32)$$

Optimal state estimator function $\hat{x}^*$ is the conditional expectation of state x given observation y and memory z , which reflects the fact that a cell needs to estimate state x from observation y and memory z . Optimal memory control function v^* is a linear function of conditional expected state $\mathbb{E}_{p_t(x|y,z)}[x]$ and memory z , and there are two control gains: Π_{zx} and Π_{zz} . Control gain Π_{zx} determines how to encode the information from observation y to memory z because $\mathbb{E}_{p_t(x|y,z)}[x]$ contains information about y . In fact, if $\Pi_{zx} = 0$, observation y does not appear in the optimal memory control function [Eq. (32)] and the information is not encoded from observation y to memory z . In this case, since memory z has no information about state x , the optimal state estimator function reduces to $\hat{x}^*(t, y) = \mathbb{E}_{p_t(x|y)}[x]$, resulting in a memoryless estimation. Conversely, if $\Pi_{zx} \neq 0$, memory z accumulates information about state x through observation y , and optimal state estimator function $\hat{x}^*$ depends not only on observation y but also on memory z as in Eq. (31), resulting in a memory-based estimation. The other control gain, Π_{zz} , is a negative feedback control gain for memory z , which suppresses the divergence and the stochasticity of the memory. When the weight of the memory control cost M is higher, the absolute value of optimal memory control function v^* is lower.

In this problem, since $p_t(x|y, z)$ is a Gaussian distribution, $\mathbb{E}_{p_t(x|y,z)}[x]$ becomes a linear function of observation y and memory z as follows:

$$\mathbb{E}_{p_t(x|y,z)}[x] = K_{xy}y + K_{xz}z. \quad (33)$$

The estimation gains, K_{xy} and K_{xz} , and the control gains, Π_{zx} and Π_{zz} , are nonlinear functions of D, E, F, Q , and M , which are obtained from HJB and FP equations (see Appendixes C and D).

C. Stochastic simulation

We conduct a stochastic simulation to validate the behavior of optimal state estimator function $\hat{x}^*$ and memory control function v^* ; see Eqs. (31) and (32) and Figs. 3(b)–3(f). A cell estimates environmental state x_t [Fig. 3(b)] from noisy observation y_t [Fig. 3(c)] and limited memory z_t [Fig. 3(d)]. The dynamics of memory z_t is optimized by optimal memory control v_t^* [Fig. 3(f)]. Despite the high noise in observation y_t , optimal state estimator $\hat{x}_t^*$ accurately recovers true state x_t [Fig. 3(e)]. This result demonstrates that accurate state estimation is achieved with limited memory without relying on Bayesian filtering.

The optimal memory dynamics [Fig. 3(d)] is similar to the state dynamics [Fig. 3(b)]. This is because memory z has a linear relationship with the conditional expected state $\mathbb{E}_{p_t(x|y,z)}[x]$, as shown in Eq. (33). This property of the optimal memory dynamics is similar to that of the Kalman filter, which is a Bayesian filter in the LQG problem [51–54,106,116–118]. In fact, we have proven that the optimal memory dynamics is equivalent to the Kalman filter when the LQG problem has no resource limitations (see Appendix G). However, when the LQG problem has resource limitations, the optimal memory dynamics deviates from the Kalman filter. In particular, as shown in the next subsection, when the observation noise is

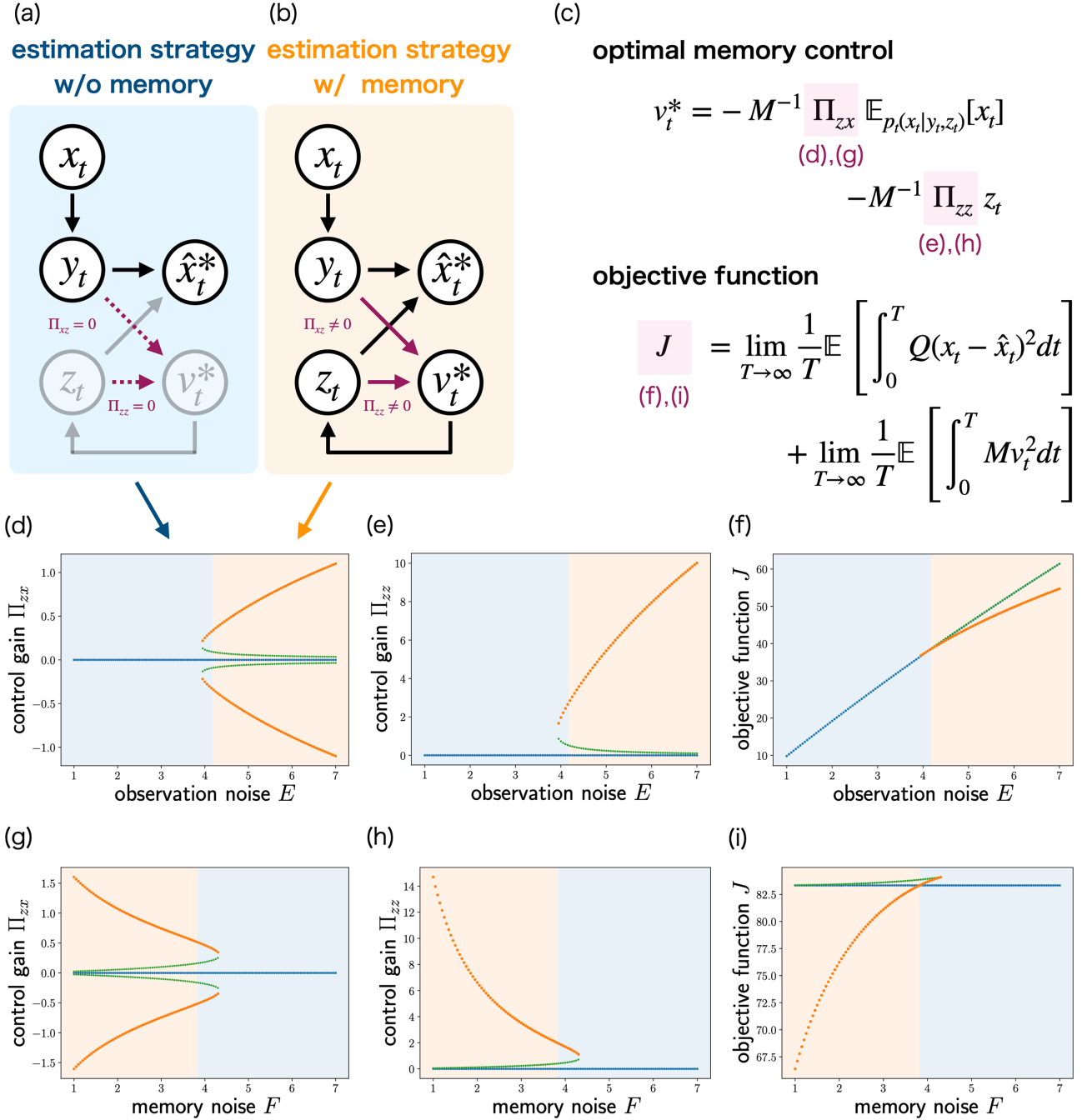


FIG. 4. Phase transition in optimal estimation strategy. Schematic diagrams of estimation strategies without memory [panel (a)] and with memory [panel (b)]. (a) When $\Pi_{xz} = \Pi_{zz} = 0$, the past observation information is not encoded into memory z_t , and state x_t is estimated only from observation y_t . (b) When $\Pi_{xz}, \Pi_{zz} \neq 0$, the past observation information is encoded into memory z_t , and state x_t is estimated from observation y_t and memory z_t . (c) Expressions of optimal memory control v_t^* and objective function J . Memory control gains Π_{zx} and Π_{zz} and associated objective function J with respect to observation noise E [panels (d)–(f)] and memory noise F [panels (g)–(i)]. The blue, green, and orange dots are the solutions of Pontryagin’s minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. There are no regions where the green dots are optimal. The parameters in panels (d)–(f) are $D = 100$, $F = 1$, $Q = 10$, and $M = 1$, while those in panels (g)–(i) are $D = 100$, $E = 10$, $Q = 10$, and $M = 1$.

lower and the memory noise is higher, the optimal memory control becomes zero and the optimal memory dynamics becomes diffusion dynamics, which is entirely different from the Kalman filter.

D. Phase transition

In this subsection, we examine the parameter dependence of the optimal estimation strategy under resource limitations (Fig. 4). The optimal estimation strategy is characterized by

the memory control gains, Π_{zx} and Π_{zz} , in the optimal memory control function, v^* [Fig. 4(c)]. If $\Pi_{xz} = \Pi_{zz} = 0$, the past observation information is not encoded into memory z_t , and state x_t is estimated only from observation y_t , which corresponds to an estimation strategy without memory [Fig. 4(a)]. Conversely, if $\Pi_{xz}, \Pi_{zz} \neq 0$, memory z_t is regulated based on the past observation information, and state x_t is estimated not only from observation y_t but also from memory z_t , which corresponds to an estimation strategy with memory [Fig. 4(b)]. In the following, we investigate the parameter dependence of the optimal memory control gains.

We first analyze the dependence of the optimal memory control gains on observation noise E [Figs. 4(d)–4(f)]. The memory control gains, Π_{zx} and Π_{zz} , are obtained by solving HJB and FP equations (see Appendix D for further details), which correspond to Pontryagin's minimum principle on the probability density function space. Given that Pontryagin's minimum principle is equivalent to the method of Lagrange multipliers (see Appendix A 4), this method identifies not only optimal solutions but also extremal solutions [Figs. 4(d) and 4(e), blue, green, and orange dots]. The optimal memory control gains are identified by the minimizer of the objective function [Fig. 4(f)].

When observation noise E is low, the solution obtained from Pontryagin's minimum principle is only $\Pi_{zx} = \Pi_{zz} = 0$, which is the optimal solution [Figs. 4(d)–4(f), blue dots]. This result indicates that the estimation strategy without memory is optimal when the current observation is sufficiently reliable. The current observation contains enough information for the accurate state estimation, making it optimal to save on the memory control cost. As observation noise E increases, the solutions where $\Pi_{zx}, \Pi_{zz} \neq 0$ discontinuously emerge [Figs. 4(d)–4(f), green and orange dots], while the solution of $\Pi_{zx} = \Pi_{zz} = 0$ remains optimal [Figs. 4(d)–4(f), blue dots]. As observation noise E further increases, the optimal solution switches from $\Pi_{zx} = \Pi_{zz} = 0$ [Figs. 4(d)–4(f), blue dots] to $\Pi_{zx}, \Pi_{zz} \neq 0$ [Figs. 4(d)–4(f), orange dots], demonstrating that the estimation strategy with memory is optimal. This result suggests that the reduction in the state estimation error through the memory control outweighs the memory control cost when the observation is less reliable.

We next analyze the dependence of the optimal memory control gains on memory noise F [Figs. 4(g)–4(i)]. The discontinuous phase transition is also observed for memory noise F , but the direction of the transition is different. While the estimation strategy with memory is optimal when memory noise F is low, that without memory becomes optimal as memory noise F increases. This result indicates that the higher intrinsic noise caused by resource limitations renders memory less useful.

The discontinuous phase transition also occurs for the other parameters, i.e., the intensity of the state noise D , the weight of the state estimation error Q , and the weight of the memory control cost M (see Appendix D, Fig. 15). The phase transition for state noise D and state estimation error Q is similar to that for observation noise E . This result suggests that the estimation strategy with memory is optimal when the state estimation is either challenging or significant. In contrast, the phase transition for memory control cost M is similar to that for memory noise F , indicating that resource limitations, such

as the noise and cost associated with intrinsic information processing, render the estimation strategy without memory optimal.

Our minimal model considers linear Gaussian dynamics and quadratic cost functions, avoiding unnecessary nonlinearities that could be potential sources of complex behaviors. Nevertheless, our model exhibits a discontinuous phase transition in the optimal estimation strategy. This phenomenon is not observed in the conventional Bayesian filtering theory, which always uses past observation information for accurate state estimation. It is consistent with the fact that the Kalman filter corresponds to $E \rightarrow \infty$, $F \rightarrow 0$, and $M \rightarrow 0$ (see Appendix G; we also note that, in the Kalman filter, the memory dimension d_z must be identical to the state dimension d_x , which holds in this problem setting). Therefore, our analysis suggests that resource limitations may drive the punctuated evolution of biological information processing, leading to qualitatively different behaviors across organisms.

We also note that although optimal memory control gains Π_{zy} and Π_{zz} exhibit discontinuous changes [Figs. 4(d), 4(e), 4(g), and 4(h)], optimal objective function J changes continuously [Figs. 4(f) and 4(i)]. Notably, while objective function J corresponding to nonzero memory control emerges discontinuously [Figs. 4(f) and 4(i), orange and green], the minimum value of objective function J remains continuous. This result suggests that even if the improvement in fitness looks gradual, the underlying optimal strategy can change qualitatively.

E. Landscape analysis

In order to examine the phase transition in the optimal estimation strategy from another perspective, we analyze the landscape of objective function J with respect to memory control gains Π_{xz} and Π_{zz} [Figs. 5(a), 5(d), and 5(g)]. When observation noise E is low, $\Pi_{xz} = \Pi_{zz} = 0$ is the global minimum [Fig. 5(a)]. As observation noise E increases, $\Pi_{xz} = \Pi_{zz} = 0$ shifts to a local minimum, and the global minima emerge at $\Pi_{xz}, \Pi_{zz} \neq 0$ [Fig. 5(d)]. Increasing memory noise F further causes the extrema at $\Pi_{xz}, \Pi_{zz} \neq 0$ to disappear, returning $\Pi_{xz} = \Pi_{zz} = 0$ as the global minimum [Fig. 5(g)]. These results are consistent with the previous subsection. We also confirm that the solutions of Pontryagin's minimum principle appear at the extrema [Figs. 5(a), 5(d), and 5(g), blue, green, and orange stars)].

In order to gain additional insights into the phase transition in the optimal estimation strategy, we decompose objective function J into state estimation error J_Q and memory control cost J_M , and analyze their landscapes [Figs. 5(b), 5(e), and 5(h) and 5(c), 5(f), and 5(i)]. State estimation error J_Q and memory control cost J_M are defined as follows:

$$J_Q := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T Q(x_t - \hat{x}_t)^2 dt \right], \quad (34)$$

$$J_M := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T M v_t^2 dt \right]. \quad (35)$$

There is a trade-off between state estimation error J_Q and memory control cost J_M with respect to memory control gains Π_{zx} and Π_{zz} . As memory control gains Π_{zx} and Π_{zz} increase, state estimation error J_Q decreases [Figs. 5(b), 5(e), and 5(h)], whereas memory control cost J_M increases

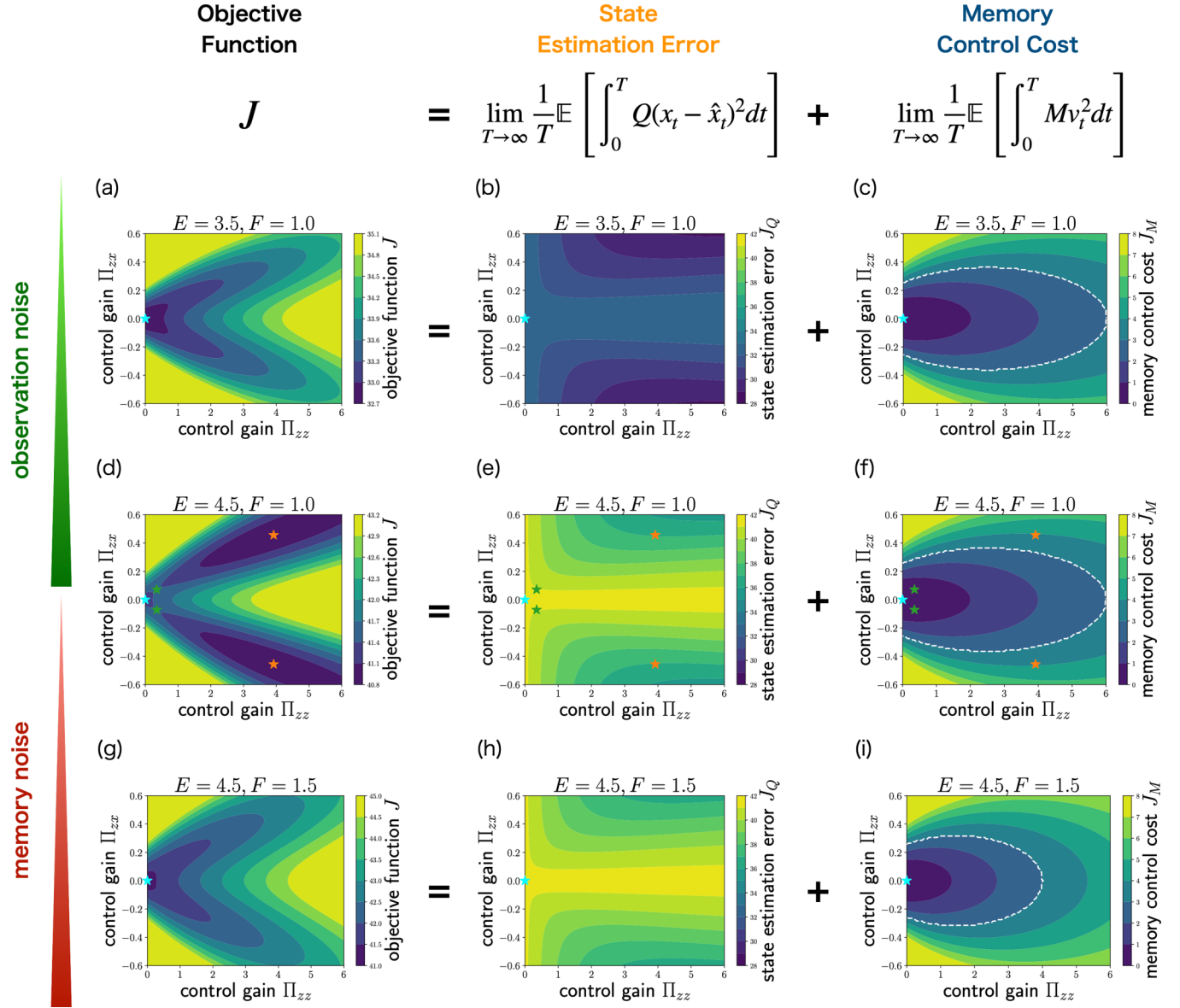


FIG. 5. Objective function J [panels (a), (d), and (g)], state estimation error J_Q [panels (b), (e), and (h)], and memory control cost J_M [panels (c), (f), and (i)] with respect to memory control gains Π_{zx} and Π_{zz} under different observation noise E and memory noise F . The blue, green, and orange stars are the solutions of Pontryagin's minimum principle. The white dashed curves represent $J_M = 3$, which clarify the changes in J_M . The rest of the parameters are $D = 100$, $Q = 10$, and $M = 1$.

[Figs. 5(c), 5(f), and 5(i)]. Therefore, the estimation strategy with memory prioritizes reducing the state estimation error, whereas the estimation strategy without memory prioritizes minimizing the memory control cost.

When observation noise E increases, state estimation error J_Q increases [Figs. 5(b) and 5(e)], while memory control cost J_M remains almost unchanged [Figs. 5(c) and 5(f)], which makes the estimation strategy with memory optimal to reduce the state estimation error. In contrast, when memory noise F increases, memory control cost J_M increases [Figs. 5(f) and 5(i)] compared to state estimation error J_Q [Figs. 5(e) and 5(h)]. This is because a large memory fluctuation leads to a higher memory feedback control. This result renders the estimation strategy without memory optimal for minimizing the memory control cost. Therefore, the phase

transition in the optimal estimation strategy may be caused by the trade-off between the state estimation error and the memory control cost.

F. Relationships among parameters

We further investigate the relationships among parameters in the phase transition of the optimal estimation strategy. In order to quantify the performance improvement induced by the memory, we define the memory effect

$$\chi := \frac{J_{w/o \text{ memory}}^* - J_{w/ \text{ memory}}^*}{J_{w/o \text{ memory}}^*}, \quad (36)$$

where $J_{w/ \text{ memory}}^*$ and $J_{w/o \text{ memory}}^*$ are the minimum values of objective function J in the systems with and without

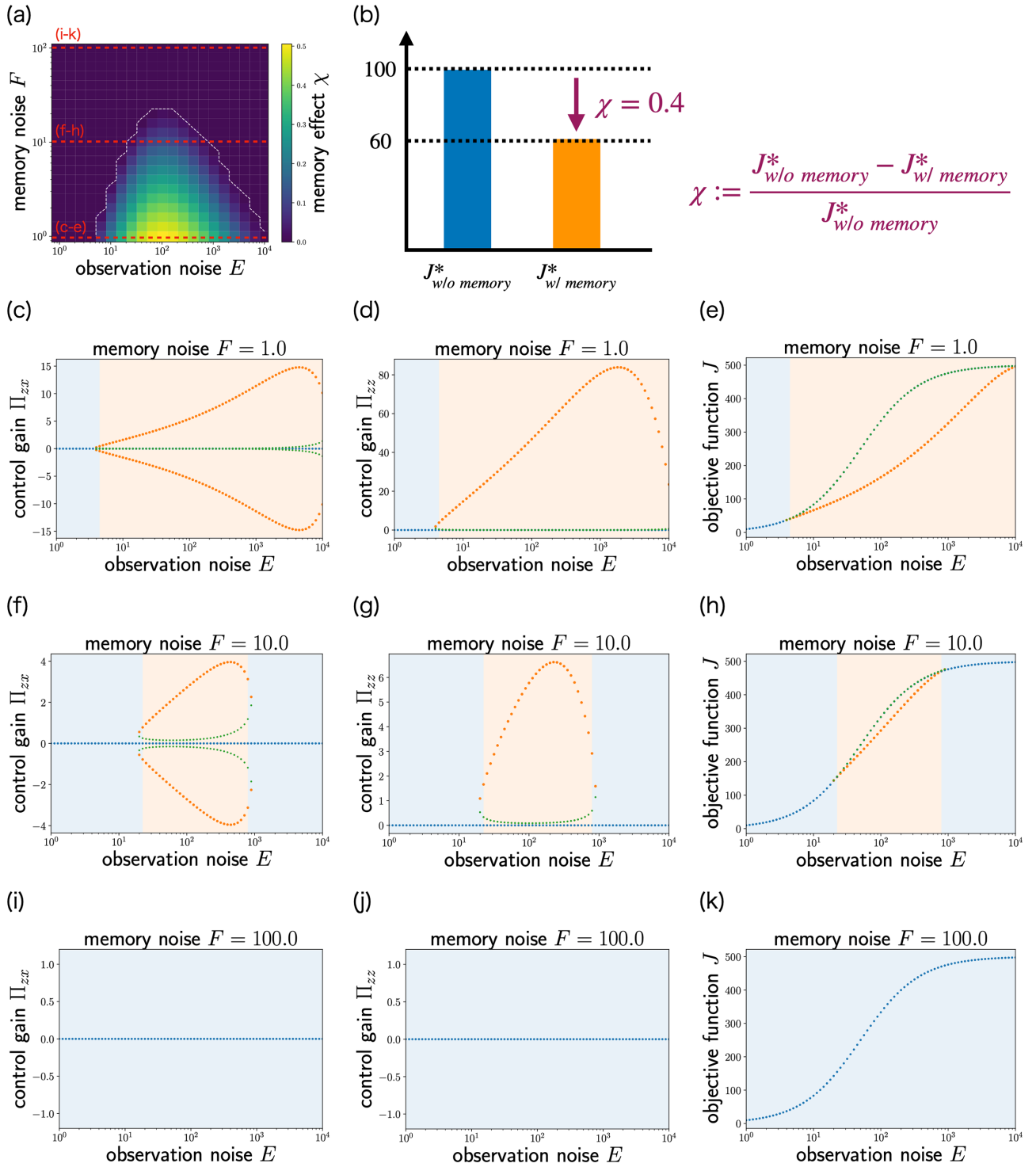


FIG. 6. (a) Memory effect χ as a function of observation noise E and memory noise F . The white dashed curve marks the boundary between $\chi = 0$ and $\chi > 0$. (b) Illustration of the definition of memory effect χ . Memory control gains Π_{zx} and Π_{zz} and associated objective function J with respect to observation noise E under different memory noise F : lower memory noise [panels (c)–(e)], intermediate memory noise [panels (f)–(h)], and higher memory noise [panels (i)–(k)]. The blue, green, and orange dots are the solutions of Pontryagin’s minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. The rest of the parameters are $D = 100$, $Q = 10$, and $M = 1$. We restrict the visualization to $10^0 \leq E \leq 10^4$, since extreme parameter values can lead to numerical instability.

memory, respectively [Fig. 6(b)]. $J^*_{w/ \text{ memory}} \geq 0$ and $J^*_{w/o \text{ memory}} \geq 0$ obviously hold because the cost function is defined as a quadratic function with a minimum of 0.

Moreover, $J^*_{w/ \text{ memory}} \leq J^*_{w/o \text{ memory}}$ always holds because the system with memory becomes equivalent to the system without memory by not controlling the memory.

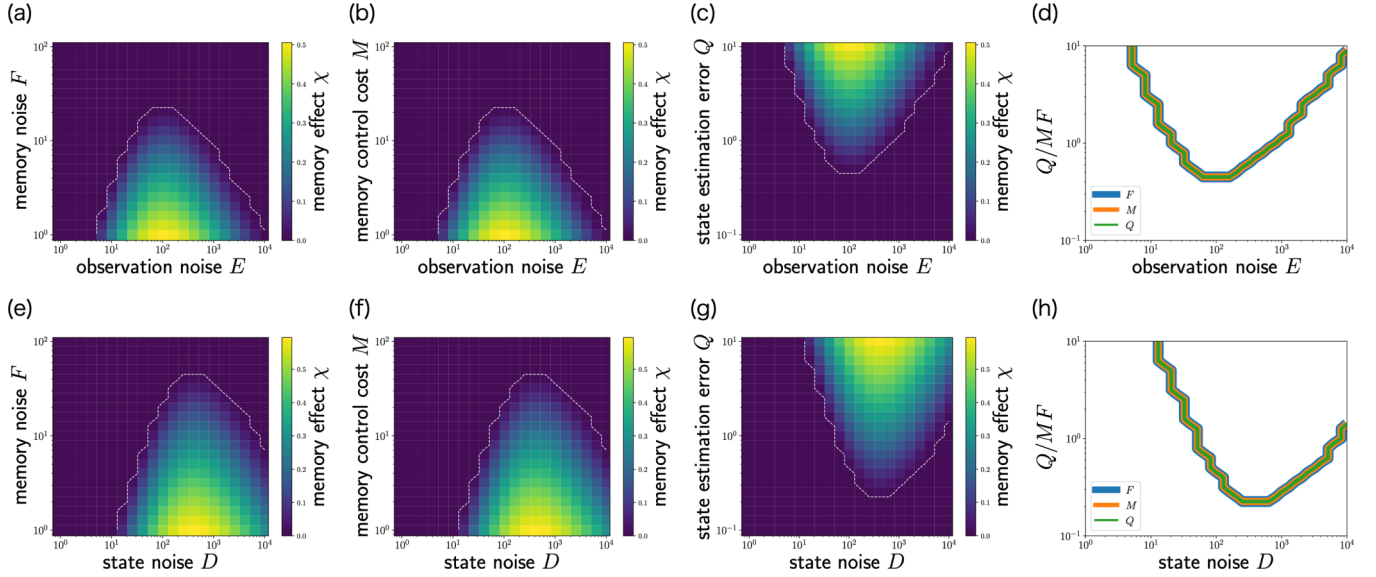


FIG. 7. Memory effect χ as a function of observation noise E and memory noise F [panel (a)], memory control cost M [panel (b)], and state estimation error Q [panel (c)]. (d) Phase boundaries as functions of E and Q/MF . The blue, orange, and green curves correspond to the phase boundaries in panels (a), (b), and (c), respectively. (e)–(h) Similar behaviors are observed for state noise D . The parameters not indicated in the figures are set to $D = 100$, $E = 100$, $F = 1$, $Q = 10$, and $M = 1$.

Consequently, $0 \leq \chi \leq 1$ holds. A higher memory effect χ indicates a greater performance improvement owing to the memory. Additionally, memory effect χ is useful for analyzing the phase transition. The estimation strategy without memory is optimal when $\chi = 0$, whereas that with memory is optimal when $\chi > 0$. The boundary between $\chi = 0$ and $\chi > 0$ corresponds to the phase boundary between the memoryless and memory-based estimation strategies. We note that while optimal memory control gains Π_{zy} and Π_{zz} change discontinuously, memory effect χ itself changes continuously from zero to nonzero values as it consists of the minimum values of objective function J .

Figure 6(a) visualizes memory effect χ with respect to observation noise E and memory noise F . When memory noise F is low, only a phase transition from the memoryless to memory-based estimation strategies occurs as observation noise E increases [Figs. 6(c)–6(e)]. In contrast, when memory noise F is a moderate value, phase transitions occur not only from the memoryless to memory-based estimation strategies, but also back to the memoryless estimation strategy as observation noise E increases [Figs. 6(f)–6(h)]. It is consistent with the fact that memory effect χ is maximized at an intermediate level of observation noise E [Fig. 6(a)]. When memory noise F is high, the phase transition to the memory-based estimation strategy disappears with respect to observation noise E [Figs. 6(i)–6(k)]. Therefore, despite the simplicity of our model, it exhibits not only discontinuous but also nonmonotonic phase transitions due to resource limitations.

We also investigate memory effect χ with respect to the other parameters, i.e., state noise D , state estimation error Q , and memory control cost M , in addition to observation noise E and memory noise F [Figs. 7(a)–7(c) and 7(e)–7(g)]. With respect to observation noise E and state noise D , memory effect χ is maximized at an intermediate level, which exhibits nonmonotonic phase transitions. These results indicate that

memory is not useful when state estimation is too simple or too difficult. In contrast, for memory noise F , state estimation error Q , and memory control cost M , memory effect χ monotonically changes, resulting in monotonic phase transitions. Additionally, we find that the phase boundaries for Q , M , and F are nearly identical when plotted against Q/MF [Figs. 7(d) and 7(h)]. This scaling relationship is consistent with the observation that an increase in state estimation error Q enhances the necessity of memory, while an increase in memory control cost M or memory noise F suppresses its effectiveness.

V. OPTIMAL ESTIMATION AND CONTROL THEORY WITH RESOURCE LIMITATIONS

Our theory achieves optimal state estimation through optimal memory control rather than Bayesian filtering. This approach allows us to extend our theory straightforwardly to include optimal state control [113–115]. In this section, we introduce the unified theory for optimal estimation and control under resource limitations.

In this theory, the agent not only estimates environmental state x_t but also controls it. Thus, the dynamics of environmental state x_t is given by the following SDE:

$$dx_t = b(t, x_t, u_t)dt + \sigma(t, x_t, u_t)d\omega_t, \quad (37)$$

where initial state x_0 follows $p_0(x_0)$, $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process, and $u_t \in \mathbb{R}^{d_u}$ is the state control of the agent. Unlike the state dynamics in Sec. III [Eq. (1)], Eq. (37) depends on state control u_t , which allows the agent to achieve desired state x_t through state control u_t . The agent determines state control u_t based on observation y_t and memory z_t as follows:

$$u_t = u(t, y_t, z_t). \quad (38)$$

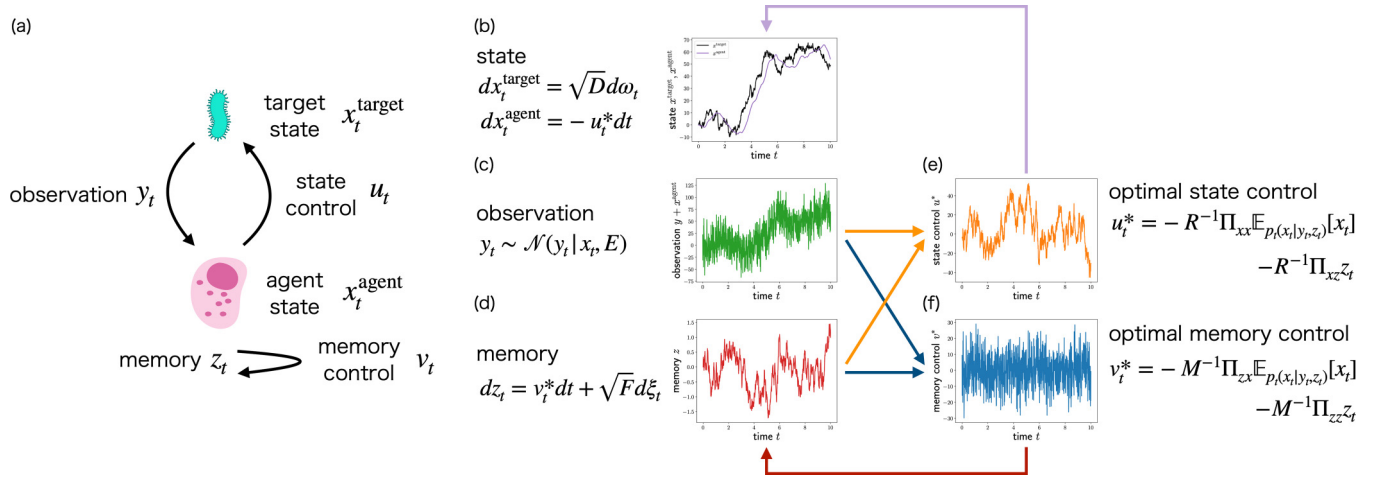


FIG. 8. Schematic diagram of a target tracking problem that is a minimal model of biological information processing and decision-making. (a) In the target tracking problem, the agent estimates the position of the target x_t^{target} from noisy observation y_t and limited memory z_t and tracks it by controlling its own position x_t^{agent} through state control u_t . The agent also controls its own memory z_t through memory control v_t to compress the information from past observation history $y_{t-dt}, \dots, y_0$ to memory z_t . Stochastic simulation of the target and agent states [panel (b)], observation [panel (c)], memory [panel (d)], optimal state control [panel (e)], and optimal memory control [panel (f)]. Optimal state control u_t^* and optimal memory control v_t^* are determined based on observation y_t and memory z_t (orange and blue arrows, respectively) and they shape the optimal dynamics of agent state x_t^{agent} and memory z_t , respectively (purple and red arrows, respectively). For comparison, target state x_t^{target} (black line) and agent state x_t^{agent} (purple line) are plotted on the same panel (b). The parameters are $D = 100$, $E = 500$, $F = 1$, $Q = 10$, $R = 1$, and $M = 1$.

Observation y_t , memory z_t , and memory control v_t are the same as those in Sec. III, given by Eqs. (2), (12), and (13), respectively. The objective function is given by the following expected cumulative cost function:

$$J[u, v] := \mathbb{E} \left[\int_0^T f(t, x_t, z_t, u_t, v_t) dt + g(x_T, z_T) \right], \quad (39)$$

where f and g are the running and terminal cost functions, respectively. Compared with the objective function in Sec. III [Eq. (15)], state estimator function $\hat{x}$ is replaced by state control function u . State control function u and memory control function v are simultaneously optimized as follows:

$$u^*, v^* := \arg \min_{u, v} J[u, v]. \quad (40)$$

This optimal estimation and control problem can be solved in almost the same way as the optimal estimation problem in Sec. III, and the optimal control functions are obtained by minimizing the conditional expected Hamiltonian:

$$u^*(t, y, z), v^*(t, y, z) = \arg \min_{u, v} \mathbb{E}_{p_t(x|y,z)}[\mathcal{H}(t, x, z, u, v, w)], \quad (41)$$

where w and p are the solutions of HJB and FP equations, respectively. The details are shown in Appendixes A and B. This result demonstrates that our theory provides a unified approach to optimal state estimation and optimal state control.

VI. APPLICATION TO BIOLOGICAL INFORMATION PROCESSING AND DECISION-MAKING

In this section, we apply the optimal estimation and control theory with resource limitations to a target tracking problem, which is a minimal model of biological information processing and decision-making [Fig. 8(a)]. This section is organized

as follows: In Sec. VIA, we formulate a target tracking problem. In Sec. VIB, we show its optimal solution and validate it by conducting a stochastic simulation. In Sec. VIC, we investigate the parameter dependence of the optimal solution and find the discontinuous phase transition between memory-less and memory-based control strategies. We also explore the relationships between parameters in this phase transition.

A. Problem formulation

We first formulate a target tracking problem as a minimal model of biological information processing and decision-making [Fig. 8(a)]. In the target tracking problem, an agent estimates the position of a target from a noisy observation and controls its own body to track it. The target tracking problem is prevalent in many biological phenomena, such as cellular chemotaxis [15,76,130,131], insect olfactory search [110,111,132], and animal foraging behavior [133–137].

The positions of the target and the agent are denoted by $x_t^{\text{target}} \in \mathbb{R}$ and $x_t^{\text{agent}} \in \mathbb{R}$, respectively. The position of the target evolves by the following diffusion process:

$$dx_t^{\text{target}} = \sqrt{D}d\omega_t, \quad (42)$$

where initial target state x_0^{target} follows standard Gaussian distribution $\mathcal{N}(x_0^{\text{target}} | 0, 1)$ and $\omega_t \in \mathbb{R}$ is a standard Wiener process. $D \geq 0$ represents the intensity of the state noise. The agent controls its own position as follows:

$$dx_t^{\text{agent}} = -u_t dt, \quad (43)$$

where $x_0^{\text{agent}} = 0$ and $u_t \in \mathbb{R}$ is the state control of the agent. The target tracking problem is concerned only with the difference between target state x_t^{target} and agent state x_t^{agent} . Thus, we define the state by $x_t := x_t^{\text{target}} - x_t^{\text{agent}}$, which evolves by

the following SDE:

$$dx_t = u_t dt + \sqrt{D} d\omega_t, \quad (44)$$

where initial state x_0 follows standard Gaussian distribution $\mathcal{N}(x_0|0, 1)$.

The agent cannot completely observe state x_t and instead obtains noisy observation $y_t \in \mathbb{R}$ that is generated from the following Gaussian distribution:

$$y_t \sim \mathcal{N}(y_t|x_t, E), \quad (45)$$

where $E \geq 0$ represents the intensity of the observation noise.

The memory of the agent, $z_t \in \mathbb{R}$, evolves by the following SDE:

$$dz_t = v_t dt + \sqrt{F} d\xi_t, \quad (46)$$

where initial memory z_0 follows standard Gaussian distribution $\mathcal{N}(z_0|0, 1)$, $v_t \in \mathbb{R}$ is the memory control of the agent, and $\xi_t \in \mathbb{R}$ is a standard Wiener process. $F \geq 0$ represents the intensity of the memory noise.

State control u_t and memory control v_t are determined based on observation y_t and memory z_t as $u_t = u(t, y_t, z_t)$ and $v_t = v(t, y_t, z_t)$, respectively.

The objective function the agent minimizes is given by

$$J[u, v] := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (Qx_t^2 + Ru_t^2 + Mv_t^2) dt \right], \quad (47)$$

where $Q > 0$, $R > 0$, and $M > 0$. The first term of the cost function is the distance between the target and the agent, the second term is the state control cost, and the third term is the memory control cost. Q , R , and M are parameters representing the weights for the state tracking error, the state control cost, and the memory control cost, respectively.

The target tracking problem is to find optimal state control function u^* and optimal memory control function v^* that minimize objective function $J[u, v]$:

$$u^*, v^* := \arg \min_{u, v} J[u, v], \quad (48)$$

which means that the agent minimizes the distance to the target with smaller control costs.

B. Optimal solution

The target tracking problem also corresponds to the LQG problem [103–106] and optimal control functions u^* and v^* [Eq. (41)] are semianalytically calculated as follows (see Appendix E):

$$u^*(t, y, z) = -R^{-1} \Pi_{xx} \mathbb{E}_{p_t(x|y, z)}[x] - R^{-1} \Pi_{xz} z, \quad (49)$$

$$v^*(t, y, z) = -M^{-1} \Pi_{zx} \mathbb{E}_{p_t(x|y, z)}[x] - M^{-1} \Pi_{zz} z. \quad (50)$$

The functional form of optimal memory control function v^* is the same as that in the target estimation problem [Eq. (32)]. (Note that the values of Π_{zx} and Π_{zz} are not necessarily the same as those in the target estimation problem.) In contrast, optimal state estimator function $\hat{x}^*$ is replaced by optimal state control function u^* in the target tracking problem. State control gains Π_{xx} and Π_{xz} represent how to control state x_t based on conditionally expected state $\mathbb{E}_{p_t(x|y, z)}[x]$ and memory z_t , respectively. The higher the weight of the state control cost R , the smaller the absolute value of optimal state control u_t^* .

By using optimal state estimator function $\hat{x}^*(t, y, z) = \mathbb{E}_{p_t(x|y, z)}[x]$ in the target estimation problem [Eq. (31)], optimal control functions u^* and v^* are rewritten as follows:

$$u^*(t, y, z) = -R^{-1} \Pi_{xx} \hat{x}^*(t, y, z) - R^{-1} \Pi_{xz} z, \quad (51)$$

$$v^*(t, y, z) = -M^{-1} \Pi_{zx} \hat{x}^*(t, y, z) - M^{-1} \Pi_{zz} z. \quad (52)$$

These equations clarify that the optimal controls are determined based on the optimal state estimator. In the target tracking problem, $p_t(x|y, z)$ is a Gaussian distribution, and $\mathbb{E}_{p_t(x|y, z)}[x]$ is a linear function of observation y and memory z as follows:

$$\mathbb{E}_{p_t(x|y, z)}[x] = K_{xy} y + K_{xz} z, \quad (53)$$

which is the same as Eq. (33) in the target estimation problem. The estimation gains, K_{xy} and K_{xz} , and the control gains, Π_{xx} , Π_{xz} , Π_{zx} , and Π_{zz} , are nonlinear functions of D , E , F , Q , R , and M , which are obtained from HJB and FP equations. The details are shown in Appendix E. We note that $\Pi_{xz} = \Pi_{zx}$ holds, which means that a part of the optimal control gains between the state and the memory has a symmetry. Nonetheless, it is important to note that state control cost R and memory control cost M are different, which causes an asymmetry between the state control and the memory control.

To verify the performance of the optimal control functions, we conduct a stochastic simulation [Figs. 8(b)–8(f)]. Despite the high observation noise, the agent tracks the target accurately [Fig. 8(b)].

C. Phase transition

We examine the parameter dependence of the optimal estimation and control strategy under resource limitations (Fig. 9). Similarly to the optimal estimation strategy, the optimal estimation and control strategy is also characterized by the memory control gains, Π_{zx} and Π_{zz} . When $\Pi_{zx} = \Pi_{zz} = 0$, the past observation information is not encoded to memory z_t , and the state estimator is reduced from $\mathbb{E}_{p_t(x|y, z)}[x]$ to $\mathbb{E}_{p_t(x|y)}[x]$. Furthermore, with $\Pi_{xz} = \Pi_{zx} = 0$, the direct control from memory z_t to state x_t also disappears. Thus, optimal state control u_t^* is determined based solely on current observation y_t , corresponding to the estimation and control strategy without memory [Fig. 9(a)]. In contrast, when Π_{zx} , $\Pi_{zz} \neq 0$, optimal state control u_t^* is determined based on both observation y_t and memory z_t , resulting in the estimation and control strategy with memory [Fig. 9(b)].

As observation noise E increases, optimal memory control gains Π_{zx} and Π_{zz} undergo a discontinuous shift from zero to nonzero values [Figs. 9(d)–9(g)], indicating that the memory-based control is optimal when the current observation is less reliable [Fig. 9(b)]. Conversely, as memory noise F increases, optimal memory control gains Π_{zx} and Π_{zz} discontinuously shift from nonzero values to zero [Figs. 9(h)–9(k)], indicating that the memoryless control is optimal when the intrinsic information processing is highly stochastic [Fig. 9(a)]. The discontinuous phase transition also occurs for the other parameters, i.e., state noise D , state tracking error Q , state control cost R , and memory control cost M (see Appendix E, Fig. 16). While the phase transition for D , Q , and R is similar to that for E , the phase transition for M is similar to that for F .

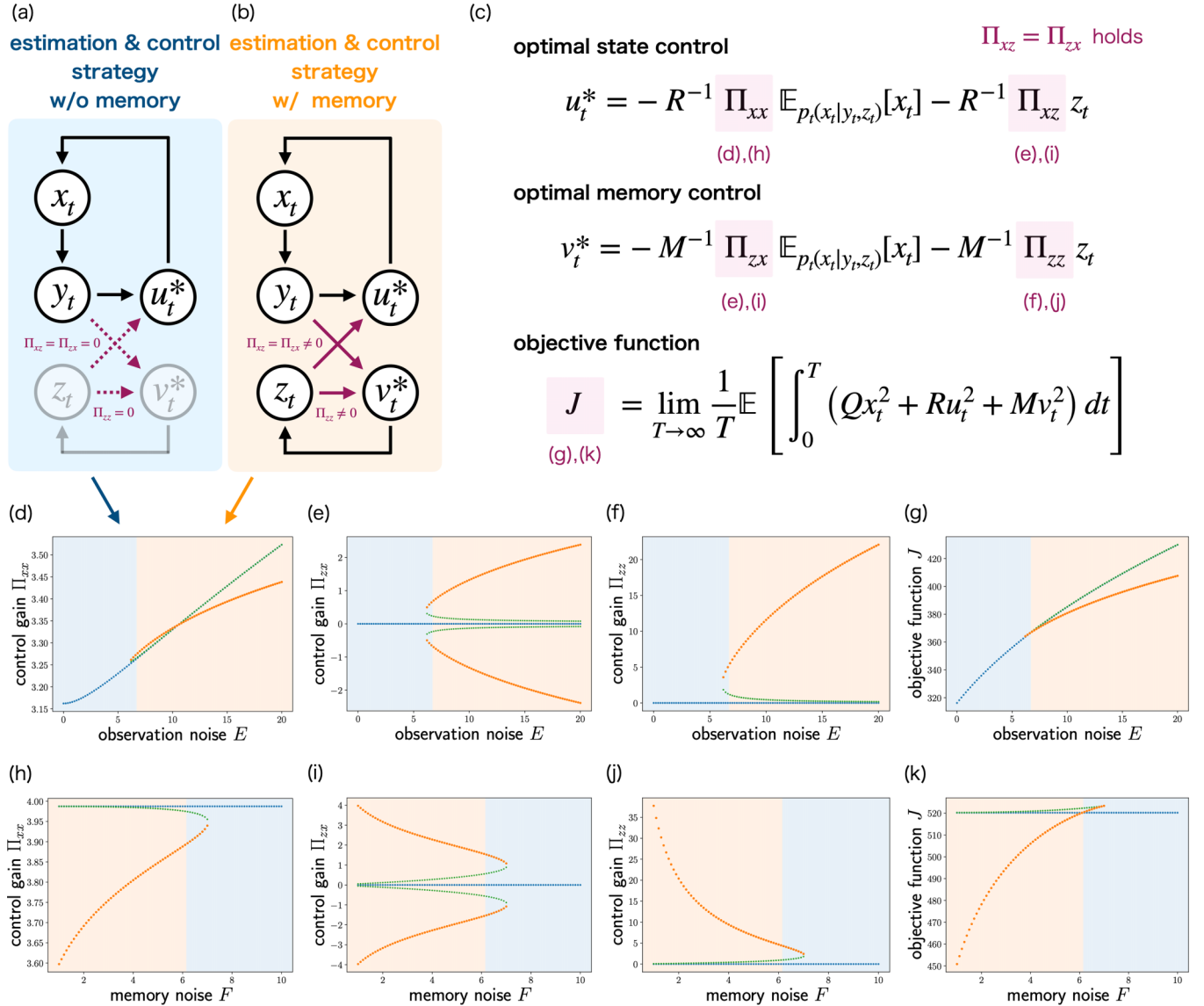


FIG. 9. Phase transition in optimal estimation and control strategy. Schematic diagrams of estimation and control strategies without memory [panel (a)] and with memory [panel (b)]. (a) When $\Pi_{zx} = \Pi_{zz} = 0$, optimal state control u_t is determined based only on current observation y_t . (b) When $\Pi_{zx}, \Pi_{zz} \neq 0$, optimal state control u_t is determined based not only on current observation y_t but also on past memory z_t . (c) Expressions of optimal state control function u^* , optimal memory control function v^* , and objective function J . We note that $\Pi_{xz} = \Pi_{zx}$ holds. Control gains Π_{xx} , Π_{zx} , and Π_{zz} and associated objective function J with respect to observation noise E [panels (d)–(g)] and memory noise F [panels (h)–(k)]. Π_{xz} is not visualized because $\Pi_{xz} = \Pi_{zx}$ holds. The blue, green, and orange dots are the solutions of Pontryagin's minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. There are no regions where the green dots are optimal. The parameters in panels (d)–(g) are $D = 100, F = 1, Q = 10, R = 1$, and $M = 1$, while those in panels (i)–(k) are $D = 100, E = 50, Q = 10, R = 1$, and $M = 1$.

These results suggest that the discontinuous phase transition is an inherent feature of biological information processing and decision-making under resource limitations.

We further investigate the relationship between observation noise E and memory noise F in the optimal estimation and control strategy (Fig. 10). Unlike the target estimation problem, the target tracking problem does not exhibit the non-monotonic phase transition with respect to observation noise E . Memory effect χ monotonically increases with respect to observation noise E [Fig. 10(a)], and the phase transition from memory-based to memoryless control strategies does not occur even when observation noise E is extremely high

and memory noise F is at an intermediate level [Figs. 10(g)–10(j)].

The difference in the phase transitions with respect to observation noise E between the target estimation problem and the target tracking problem may arise from the difference in their state dynamics. In the target estimation problem, the state dynamics is formulated as a trapped diffusion [Eq. (26)], which is stable around the origin. Even if the agent ignores the observation and fails to estimate the state, the resulting state estimation error is bounded due to the finite variance of the state [Figs. 6(e), 6(h), and 6(k)]. Thus, when the observation noise becomes extremely high, the agent can choose to

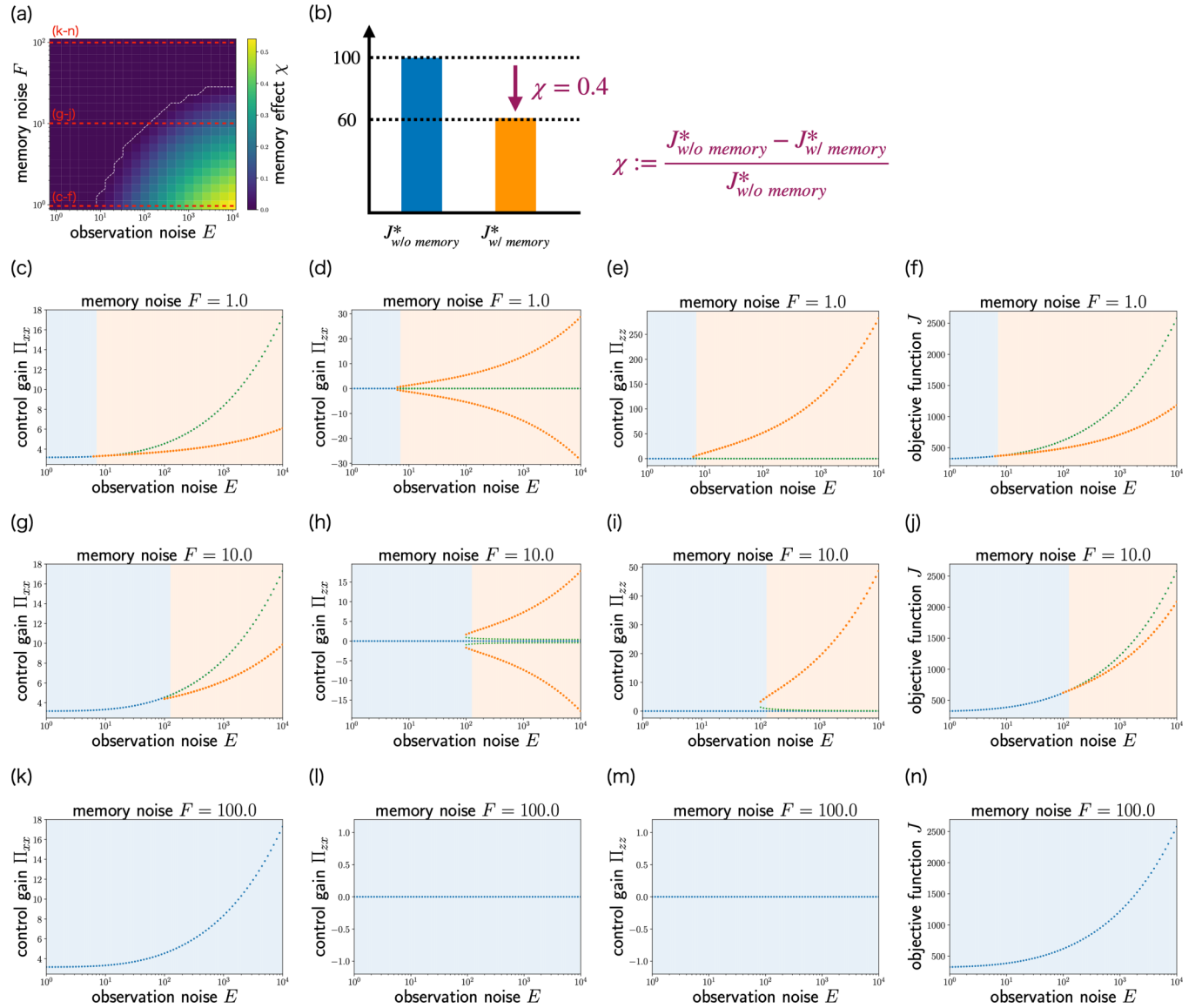


FIG. 10. (a) Memory effect χ as a function of observation noise E and memory noise F . The white dashed curve marks the boundary between $\chi = 0$ and $\chi > 0$. (b) Illustration of the definition of memory effect χ . Control gains Π_{xx} , Π_{xz} , and Π_{zz} and associated objective function J with respect to observation noise E under different memory noise F : lower memory noise [panels (c)–(f)], intermediate memory noise [panels (g)–(j)], and higher memory noise [panels (k)–(n)]. The blue, green, and orange dots are the solutions of Pontryagin’s minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. The rest of the parameters are $D = 100$, $Q = 10$, $R = 1$, and $M = 1$. We restrict the visualization to $10^0 \leq E \leq 10^4$, since extreme parameter values can lead to numerical instability.

give up memory control, leading to the nonmonotonic phase transition with respect to the observation noise. By contrast, in the target tracking problem, the state dynamics is modeled as a free diffusion [Eq. (44)], which diverges to infinity in the absence of state control. If the agent fails in both estimation and control, it completely loses track of the target, resulting in unbounded tracking error [Figs. 10(f), 10(j), and 10(n)]. Thus, the agent cannot give up memory control. Indeed, when the state dynamics in the target tracking problem is modified from free diffusion to trapped diffusion, a nonmonotonic phase transition also emerges (see Appendix E4, Fig. 17).

We also explore memory effect χ with respect to state noise D , state tracking error Q , state control cost R , and memory control cost M , in addition to observation noise E

and memory noise F (Fig. 11). While increasing observation noise E does not lead to a nonmonotonic phase transition [Figs. 11(a)–11(d)], increasing state noise D exhibits such a transition [Figs. 11(h)–11(k)]. This may be because, even if state noise D is extremely high, accurate state estimation can still be achieved solely through current observation y_t as long as observation noise E is low. Moreover, while increasing memory noise F and memory control cost M decreases memory effect χ [Figs. 11(a), 11(b), 11(h), and 11(i)], increasing state tracing error Q and state control cost R has the opposite effect, increasing memory effect χ [Figs. 11(c), 11(d), 11(j), and 11(k)]. The phase boundaries for memory noise F and memory control cost M are similar, suggesting a scaling relationship between them, such as MF [Figs. 11(e)

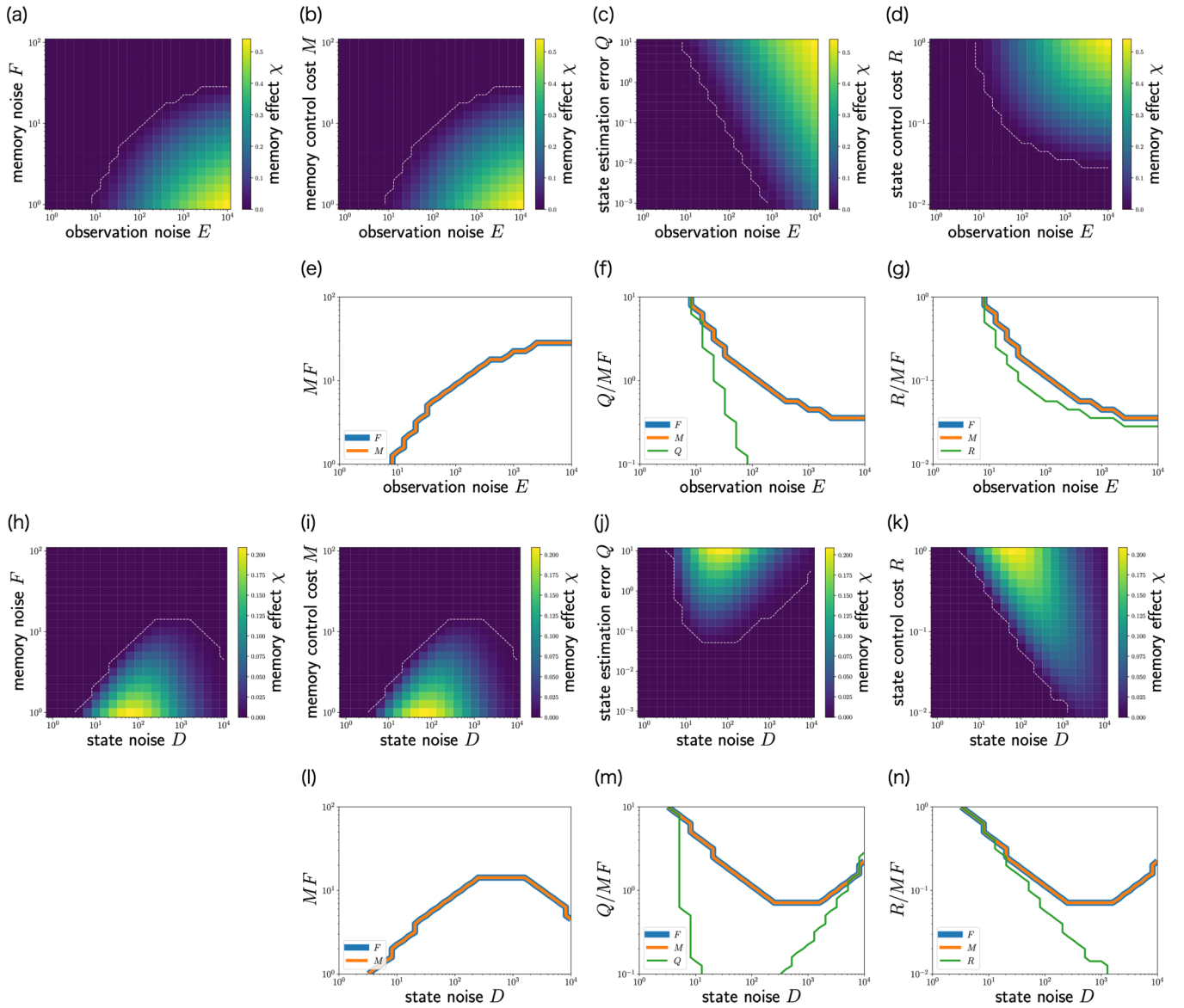


FIG. 11. (a)–(d), (h)–(k) Memory effect χ with respect to the indicated parameters. The white dashed curves mark the boundary between $\chi = 0$ and $\chi > 0$. (e)–(g), (l)–(n) Phase boundaries with respect to the indicated parameters. The blue and orange curves show the result of varying F and M , respectively, while keeping all the other parameters fixed. The green curves show the effect of varying Q in panels (f) and (m) and R in panels (g) and (n), respectively. The parameters not indicated in the figures are set to $D = 100$, $E = 100$, $F = 1$, $Q = 10$, $R = 1$, and $M = 1$.

and 11(l)]. On the other hand, since the phase boundaries for state tracking error Q and state control cost R differ from those for memory noise F and memory control cost M , it suggests that a scaling relationship, such as Q/MF or R/MF , does not hold in the target tracking problem [Figs. 11(f), 11(g), 11(m), and 11(n)]. These results indicate that the discontinuity in phase transitions may be more fundamental characteristic under resource limitations than the nonmonotonicity and scaling relationship.

VII. DISCUSSION

In this paper, we have proposed a comprehensive theoretical framework for addressing biological information processing and decision-making under resource limitations.

Instead of relying on Bayesian filtering of posterior probabilities, our theory employs the optimal control of memories that organisms can store and operate, which achieves optimal estimation under resource limitations. Additionally, since this theory realizes state estimation through memory control, it can be straightforwardly extended to include state control, allowing for an integrated approach toward biological information processing and decision-making. We applied this theory to the minimal models of biological information processing and decision-making and investigated optimal estimation and control strategies under resource limitations. Our analyses reveal that, despite the simplicity of these models, resource limitations can induce discontinuous and nonmonotonic phase transitions between memoryless and memory-based estimations and controls. This result

suggests that resource limitations may be a source of diversity in biological information processing and decision-making.

There are several studies investigating the trade-off between estimation error and energy cost in cellular information processing [35–49]. Among these, the studies conducted by Govern and ten Wolde [36,37] are particularly relevant to our work because they demonstrate how the energy dissipation associated with intracellular chemical reactions affects the effectiveness of memory. When intracellular chemical reactions are at equilibrium, they dissipate no energy. However, under such conditions, cells cannot utilize information about past receptor activities to estimate ligand concentrations. As a result, equilibrium intracellular chemical reactions cannot reduce the lower bound of the estimation error, indicating that estimation based solely on receptor activities reaches this bound. In contrast, when intracellular chemical reactions are at nonequilibrium and dissipate energy, they can exploit past receptor information to improve estimation accuracy. In particular, the lower bound of the estimation error decreases as the energy dissipation of nonequilibrium intracellular chemical reactions increases. These studies therefore highlight that the energy dissipation makes memory effective and enhances estimation performance. While they focus on the idealized setting where environmental ligand concentrations are static, subsequent work has extended this framework to time-varying ligand concentrations [38–41].

Although the present study shares similarities with the previous studies in terms of the trade-off between memory effectiveness and energy cost, it provides three distinct contributions. First, the present study provides a general framework based on optimal control theory. The previous studies have primarily focused on cellular information processing and demonstrated that estimation errors depend on the energy costs associated with intracellular chemical reactions [36–41]. In contrast, our framework addresses more general and abstract biological information processing based on optimal control theory, which could be applied not only to cellular information processing but also to other biological phenomena such as neural computations in the brain [108,109,138–144].

Second, our framework straightforwardly integrates biological information processing with decision-making, as demonstrated in Sec. VI. The integration of information processing and decision-making under resource limitations has not been fully clarified in the previous studies [36–41]. The present study addresses this issue by treating both processes within a unified optimal control framework. Therefore, the proposed framework could contribute to the understanding of biological information processing coupled with decision-making, such as cellular chemotaxis [15,76,130,131], insect olfactory search [110,111,132], animal foraging behavior [133–137], and human motor control [108,109].

Third, the present study explicitly investigates optimal strategies by employing optimal control theory. As a result, we find that resource limitations can induce discontinuous and nonmonotonic phase transitions between memoryless and memory-based strategies, even though the models fall within the standard LQG class. While a few previous studies have reported discontinuous [91,93,94] or nonmonotonic [144,145] phase transitions in information processing strategies, most

others have found only gradual and monotonic changes with increasing resource availability [35–44]. This raises an important open question: Which classes of problems can exhibit such nontrivial resource-induced phase transitions? Identifying these classes will help us understand which biological systems may exhibit discontinuous and nonmonotonic phase transitions in their information processing strategies under resource limitations.

To explore the biological relevance of our theoretical findings, we next present three biological phenomena. Although these examples do not establish a precise quantitative mapping to our theoretical models, they illustrate qualitatively similar phase transitions between simple and complex information processing strategies. Establishing a more rigorous and quantitative correspondence remains a key direction for future research.

One potential example is the circadian rhythms of the cyanobacteria, *Synechococcus elongatus* and *Prochlorococcus marinus* [146–149]. These organisms estimate the time of day by processing sunlight information through intracellular chemical reactions. The circadian rhythm of *S. elongatus* exhibits self-sustained oscillations driven by the KaiA protein, which acts as a negative feedback controller in its intracellular chemical reactions [150–154]. In contrast, *P. marinus* has lost the KaiA protein through evolution [155], and its intracellular chemical reactions exhibit damped oscillations rather than self-sustained ones [156–160]. The KaiA protein regulating the circadian rhythms of cyanobacteria may serve as a biological counterpart to the memory control in our theoretical framework. From this perspective, the self-sustained oscillations of *S. elongatus* may reflect more complex information processing with memory control, whereas the damped oscillations of *P. marinus* may correspond to simpler information processing without memory control. Indeed, *P. marinus* has a much smaller cell volume than *S. elongatus*, leading to higher intrinsic stochasticity due to the limited number of molecules available for information processing [148]. In contrast, *S. elongatus* inhabits a wider range of environments than *P. marinus*, which likely demands more accurate estimation [161]. These experimental observations suggest a potential alignment with our theoretical predictions.

Another potential example can be found in the dual-process theory of human information processing and decision-making [162–165]. This theory posits that human cognition operates through two distinct systems: System 1, which is unconscious, implicit, and intuitive, and System 2, which is conscious, explicit, and logical. Recent experiments have shown that the dominance of these systems can shift depending on the availability of energy resources, such as the level of glucose in the bloodstream [166–168]. System 1 dominates when glucose levels are low, whereas System 2 becomes more active as glucose levels increase. This transition exhibits a qualitative resemblance to the phase transitions observed in our theoretical framework. System 1 corresponds to simpler and low-cost information processing and decision-making, while System 2 reflects more complex processing that enhances performance even at the expense of greater energy consumption. The dual-process theory may thus provide a compelling example of how resource limitations can induce qualitative shifts in the complexity and capacity of cognitive operations.

A third potential example comes from a recent human behavioral experiment [145], which examined how humans infer a latent, time-varying state from noisy observations. The findings suggest that human information processing strategies depend on the level of observation noise: when the observation noise is moderate, subjects appear to adopt complex estimation strategies that resemble memory-based Bayesian filtering. However, when the observation noise is either too low or too high, participants tend to rely on simpler, memory-less strategies. This nonmonotonic dependence on observation noise levels is consistent with the phase transitions predicted by our theoretical models. It implies that humans may flexibly adjust the complexity of their inference strategies depending on the reliability of sensory information due to resource limitations.

To connect our theory with real biological phenomena, such as intracellular chemical reactions, bacterial circadian rhythms, and human decision-making processes, it is necessary to consider more complex and involved problem settings than those addressed in the present study. While our theoretical framework is general enough to accommodate high-dimensional and nonlinear cases, we restricted our analysis here to one-dimensional and linear minimal models for the sake of computational tractability. Applying the framework to more complex systems requires the development of efficient numerical algorithms capable of solving high-dimensional HJB-FP equations. Although this is a challenging task, it might be resolved by employing neural network-based algorithms developed in the context of mean-field games and control [169–172].

By employing these algorithms, our theoretical framework could be applied to more realistic and complex biological scenarios, particularly those in which environmental states are higher dimensional than biological memories [173–176]. The conventional Bayesian filtering struggles to handle such cases because it determines the dimensions of biological memories based on those of environmental states. In contrast, our theory addresses this issue by allowing the dimensions of biological memories to be determined independently of those of environmental states. This property may provide a valuable perspective on how organisms compress high-dimensional environmental information into low-dimensional biological memories. While the dimensionality reduction in machine learning is one approach to this problem [177–179], the potential strength of our theory lies in the explicit consideration of the dynamics of biological memories. This feature would allow us to investigate how complex environmental signals are encoded into diverse dynamical patterns of biological memories, including sustained, transient, and oscillatory dynamics. Such dynamical encoding has been observed in cellular information processing, where different ligand inputs generate distinct temporal concentration profiles in the same downstream molecules [173–176]. Therefore, our framework could serve as a theoretical basis for understanding dynamical information encoding in biological systems.

Another important theoretical direction for future work is to extend our framework from a single-agent system to a multiagent system with communication. In nature, organisms often form groups, allowing for more complex and sophisticated information processing and decision-making through

communication [180–187]. We have previously explored this extension in our earlier work, although it was less biologically plausible [115]. Therefore, the present framework has the potential to provide an effective approach for investigating collective information processing and decision-making.

ACKNOWLEDGMENTS

We thank Kenji Kashima, Simon K. Schnyder, Kento Nakamura, Shuhei A. Horiguchi, Masaki Kato, and Taro Toyozumi for valuable discussions. The first author was supported by JSPS Research Fellowship (Grant No. 22KJ0557), JST ACT-X (Grant No. JPMJAX24LB), and Special Postdoctoral Researcher (SPDR) Program at RIKEN. This research was supported by JST CREST (Grant No. JPMJCR2011) and JSPS KAKENHI (Grants No. 19H05799, No. 24H02148, No. 24H01465, and No. 25H01365).

DATA AVAILABILITY

The sample code is available at [202].

APPENDIX A: OBSERVATION-BASED STOCHASTIC CONTROL

The purpose of Appendixes A and B is to derive the optimal solutions of our theory presented in Sec. III [Eq. (18)] and Sec. V [Eq. (41)]. Appendix A serves as a preliminary section, where we consider the optimal estimation and control problem without memory [Fig. 12(a)]. We refer to this problem as observation-based stochastic control (OSC) because it corresponds to a stochastic optimal control problem where the agent determines the control based only on the current observation. Appendix B is the main section, where we address the optimal estimation and control problem with memory [Fig. 12(b)]. This problem is referred to as observation-and-memory-based stochastic control (OMSC) because the agent determines the controls based on the current observation and memory. OMSC corresponds to the optimal estimation and control problem with resource limitations presented in Sec. V. Additionally, the optimal estimation problem with resource limitations presented in Sec. III is a special case of OMSC. Thus, the purpose of Appendixes A and B reduces to deriving the optimal solution of OMSC. The optimal solution of OMSC can be derived in a similar manner to that of OSC because OMSC can be interpreted as a special case of OSC on an extended space (Fig. 13). Therefore, in this section, we aim to derive the optimal solution of OSC as a preparatory step for deriving that of OMSC. The optimal solution of OMSC is then derived in the following section (Appendix B).

It should be noted that OSC and OMSC build upon our previously proposed theoretical framework, memory-limited partially observable stochastic control (ML-POSC) [113–115]. ML-POSC is an optimal estimation and control theory with resource limitations. While OMSC is similar to ML-POSC, the problem formulation is modified, allowing for more biologically plausible setups. A comparison between OMSC and ML-POSC is provided in Appendix F.

This section is organized as follows: In Appendix A 1, we formulate OSC. In Appendix A 2, we present the optimal

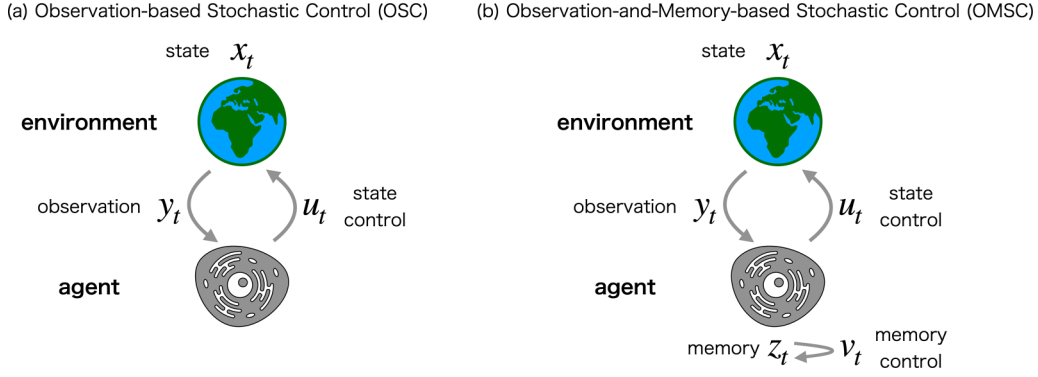


FIG. 12. Schematic diagrams of (a) observation-based stochastic control and (b) observation-and-memory-based stochastic control. observation-based stochastic control (OSC) and observation-and-memory-based stochastic control (OMSC) are optimal estimation and control problems without memory and with memory, respectively. We discuss OSC and OMSC in Appendixes A and B, respectively.

solution of OSC. In Appendix A 3, we derive the optimal solution of OSC using Pontryagin's minimum principle on the probability density function space. In Appendix A 4, we show that Pontryagin's minimum principle corresponds to the method of Lagrange multipliers.

1. Problem formulation

In this subsection, we formulate the optimal estimation and control problem without memory, i.e., OSC [Fig. 12(a)]. OSC considers an environment and an agent. The state of the environment at time $t \in [0, T]$ is denoted by $x_t \in \mathbb{R}^{d_x}$, which evolves by the following SDE:

$$dx_t = b(t, x_t, u_t)dt + \sigma(t, x_t, u_t)d\omega_t, \quad (\text{A1})$$

where initial state x_0 follows $p_0(x_0)$, $u_t \in \mathbb{R}^{d_u}$ is the control of the agent, and $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process. The agent cannot completely observe environmental state x_t and instead obtains noisy observation $y_t \in \mathbb{R}^{d_y}$, which is generated from probability density function $p_t(y_t|x_t)$ as follows:

$$y_t \sim p_t(y_t|x_t). \quad (\text{A2})$$

The agent determines control u_t based on observation y_t :

$$u_t = u(t, y_t). \quad (\text{A3})$$

The objective function is given by the following expected cumulative cost function:

$$J[u] := \mathbb{E} \left[\int_0^T f(t, x_t, u_t)dt + g(x_T) \right], \quad (\text{A4})$$

where f and g are the running and terminal cost functions, respectively. OSC considers a problem to find optimal control function u^* that minimizes objective function $J[u]$:

$$u^* := \arg \min_u J[u]. \quad (\text{A5})$$

We refer to this stochastic optimal control problem as OSC because the agent determines control u_t based on observation y_t [Eq. (A3)]. If observation y_t is identical to state x_t , i.e., $y_t = x_t$, this problem reduces to the most basic stochastic control, where the agent can completely observe state x_t and determines control u_t based on it as $u_t = u(t, x_t)$ [104–106, 188–190]. In contrast, in OSC, the agent cannot completely observe state x_t and needs to estimate it from observation y_t . Additionally, the agent is not allowed to use information from past observation history $y_{t-dt}, \dots, y_0$ to estimate and control

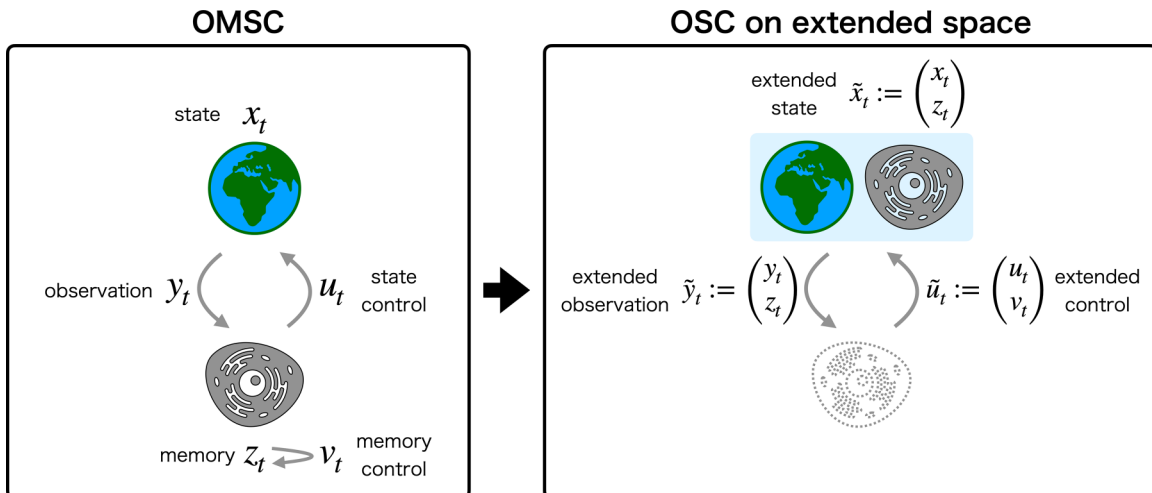


FIG. 13. Schematic diagram of the conversion from OMSC to OSC on the extended space.

state x_t . Therefore, OSC is an optimal estimation and control problem without memory.

2. Optimal solution

In the conventional stochastic control, where the agent can completely observe state x_t , the optimal control function is obtained using Bellman's dynamic programming principle on the state space [104–106,188–190]. However, this method is not applicable to OSC because the agent cannot completely observe state x_t . To address this issue, we propose Pontryagin's minimum principle on the probability density function space, which has recently been used in mean-field stochastic optimal control problems [119–122] and in our previous works [113–115]. In this subsection, we show the main result. The details of the derivation are shown in the following subsection (Appendix A3).

The optimal control function of OSC is given by minimizing the conditional expected Hamiltonian:

$$u^*(t, y) = \arg \min_u \mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, u, w)]. \quad (\text{A6})$$

Hamiltonian $\mathcal{H}$ is defined by

$$\mathcal{H}(t, x, u, w) := f(t, x, u) + \mathcal{L}_u w(t, x), \quad (\text{A7})$$

where $f(t, x, u)$ is the current cost and $\mathcal{L}_u w(t, x)$ represents the future cost. $\mathcal{L}_u$ is the time-backward diffusion operator for state x , which is defined by

$$\begin{aligned} \mathcal{L}_u w(t, x) := & \sum_{i=1}^{d_x} b_i(t, x, u) \frac{\partial w(t, x)}{\partial x_i} \\ & + \frac{1}{2} \sum_{i,j=1}^{d_x} D_{ij}(t, x, u) \frac{\partial^2 w(t, x)}{\partial x_i \partial x_j}, \end{aligned} \quad (\text{A8})$$

where $D(t, x, u) := \sigma(t, x, u)\sigma^\top(t, x, u)$. $w(t, x)$ is the value function, which represents the expected cumulative cost from the starting point at time t and state x to the terminal time T . Value function $w(t, x)$ is obtained by solving the following time-backward partial differential equation, HJB equation:

$$-\frac{\partial w(t, x)}{\partial t} = \mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, u^*, w)], \quad (\text{A9})$$

where the terminal condition is given by $w(T, x) = g(x)$. In the optimal control function of OSC [Eq. (A6)], conditional expectation $\mathbb{E}_{p_t(y|x)}[\cdot]$ appears, which is consistent with the fact that the agent needs to estimate state x from observation y . Conditional probability density function $p_t(x|y)$ is calculated as follows:

$$p_t(x|y) = \frac{p_t(y|x)p_t(x)}{\int p_t(y|x)p_t(x)dx}. \quad (\text{A10})$$

Probability density function $p_t(x) := p(t, x)$ is obtained by solving the following time-forward partial differential equation, FP equation:

$$\frac{\partial p(t, x)}{\partial t} = \check{\mathcal{L}}_u^\dagger p(t, x), \quad (\text{A11})$$

where the initial condition is given by $p(0, x) = p_0(x)$. $\check{\mathcal{L}}_u^\dagger$ is the expected time-forward diffusion operator for state x ,

which is defined by

$$\begin{aligned} \check{\mathcal{L}}_u^\dagger p(t, x) := & - \sum_{i=1}^{d_x} \frac{\partial (\mathbb{E}_{p_t(y|x)}[b_i(t, x, u(t, y))]p(t, x))}{\partial x_i} \\ & + \frac{1}{2} \sum_{i,j=1}^{d_x} \frac{\partial^2 (\mathbb{E}_{p_t(y|x)}[D_{ij}(t, x, u(t, y))]p(t, x))}{\partial x_i \partial x_j}. \end{aligned} \quad (\text{A12})$$

The optimal control function of OSC [Eq. (A6)] is obtained by solving the time-backward HJB equation [Eq. (A9)] and the time-forward FP equation [Eq. (A11)]. Since HJB and FP equations are mutually dependent through the optimal control function, they must be solved simultaneously, which corresponds to a two-point boundary value problem of partial differential equations. This problem can be solved using the forward-backward sweep method, which computes the time-backward HJB equation and the time-forward FP equation alternately [114]. While the convergence of the forward-backward sweep method is not guaranteed in deterministic optimal control problems [191–194] and mean-field stochastic optimal control problems [195–198], it is guaranteed in OSC because the coupling of HJB and FP equations is limited to the optimal control function. This proof is almost the same as our previous work [114].

We note that Pontryagin's minimum principle is a necessary condition for the optimal control function, but not a sufficient condition [114]. This means that while the optimal control function always satisfies Eq. (A6), a control function that satisfies Eq. (A6) is not necessarily the optimal control function. This is evident from the fact that Pontryagin's minimum principle corresponds to the method of Lagrange multipliers, as shown in Appendix A4. In fact, our minimal models yield not only optimal solutions but also nonoptimal solutions based on Pontryagin's minimum principle [see Figs. 4(d)–4(f), blue, green, and orange dots as an example].

If observation y_t is identical to state x_t , i.e., $y_t = x_t$, OSC reduces to the conventional stochastic control, and Eq. (A6) becomes

$$u^*(t, x) = \arg \min_u \mathcal{H}(t, x, u, w). \quad (\text{A13})$$

In this case, the optimal control function is obtained by solving only the time-backward HJB equation [Eq. (A9)], which is consistent with Bellman's dynamic programming principle on the state space [104–106,188–190]. Therefore, Pontryagin's minimum principle on the probability density function space can be interpreted as a generalization of Bellman's dynamic programming principle on the state space to cases of incomplete observation [113–115].

3. Derivation of optimal solution

In this subsection, we derive the optimal control function of OSC [Eq. (A6)] using Pontryagin's minimum principle on the probability density function space. In this paper, we derive this result in three steps, following our previous works [113,115]. First, we convert OSC on the state space into the deterministic control on the probability density function space (Appendix A3a). Second, we apply Bellman's dynamic programming principle on the probability density function space,

not on the state space (Appendix A 3 b). Finally, we convert Bellman's dynamic programming principle on the probability density function space into Pontryagin's minimum principle on the same space (Appendix A 3 c). It is worth noting that Pontryagin's minimum principle is directly applicable without going through Bellman's dynamic programming principle. We omit this proof because it is nearly identical to our previous work [114].

a. Conversion to deterministic control on probability density function space

We first convert OSC on the state space into the deterministic control on the probability density function space. The state SDE [Eq. (A1)] is converted into the following FP equation:

$$\frac{\partial p(t, x)}{\partial t} = \check{\mathcal{L}}_u^\dagger p(t, x), \quad (\text{A14})$$

where $p(0, x) = p_0(x)$ and $\check{\mathcal{L}}_u^\dagger$ is the expected time-forward diffusion operator [Eq. (A12)]. The objective function of OSC [Eq. (A4)] is also calculated as follows:

$$J[u] = \int_0^T \bar{f}(t, p_t, u_t) dt + \bar{g}(p_T), \quad (\text{A15})$$

where $p_t(x) := p(t, x)$, $\bar{f}(t, p, u) := \mathbb{E}_{p(x)}[\mathbb{E}_{p_t(y|x)}[f(t, x, u(t, y))]]$, and $\bar{g}(p) := \mathbb{E}_{p(x)}[g(x)]$. From Eqs. (A14) and (A15), OSC on the state space is converted into the deterministic control on the probability density function space. While OSC on the state space controls state SDE [Eq. (A1)], the deterministic control on the probability density function space controls FP equation [Eq. (A14)].

b. Bellman's dynamic programming principle on probability density function space

While OSC cannot completely observe state x_t , it can obtain complete information about probability density function p_t because probability density function p_t is deterministic. As a result, OSC can be solved by Bellman's dynamic programming principle on the probability density function space, not on the state space. In the following, we show Bellman's dynamic programming principle on the probability density function space.

The minimization of the objective function is calculated as follows:

$$\begin{aligned} \min_u J[u] &= \min_{u_{0:T-dt}} \left[\int_0^T \bar{f}(\tau, p_\tau, u_\tau) d\tau + \bar{g}(p_T) \right] \\ &= \min_{u_{0:T-dt}} \left[\int_0^{t-dt} \bar{f}(\tau, p_\tau, u_\tau) d\tau \right. \\ &\quad \left. + \min_{u_{t:T-dt}} \left[\int_t^T \bar{f}(\tau, p_\tau, u_\tau) d\tau + \bar{g}(p_T) \right] \right] \\ &= \min_{u_{0:t-dt}} \left[\int_0^{t-dt} \bar{f}(\tau, p_\tau, u_\tau) d\tau + V(t, p_t) \right], \end{aligned} \quad (\text{A16})$$

where $V(t, p)$ is the value function, which is defined as follows:

$$V(t, p) := \min_{u_{t:T-dt}} \left[\int_t^T \bar{f}(\tau, p_\tau, u_\tau) d\tau + \bar{g}(p_T) \right], \quad (\text{A17})$$

where $\{p_\tau | \tau \in [t, T]\}$ is the solution of FP equation [Eq. (A14)] with $p_t = p$. The optimal control function is obtained by calculating the value function.

The value function is calculated as follows:

$$\begin{aligned} V(t, p) &= \min_{u_{t:T}} \left[\int_t^T \bar{f}(\tau, p_\tau, u_\tau) d\tau + \bar{g}(p_T) \right] \\ &= \min_u \left[\bar{f}(t, p, u) dt \right. \\ &\quad \left. + \min_{u_{t+dt:T}} \left[\int_{t+dt}^T \bar{f}(\tau, p_\tau, u_\tau) d\tau + \bar{g}(p_T) \right] \right] \\ &= \min_u [\bar{f}(t, p, u) dt + V(t + dt, p + \check{\mathcal{L}}_u^\dagger p dt)] \\ &= \min_u \left[\bar{f}(t, p, u) dt + V(t, p) + \frac{\partial V(t, p)}{\partial t} dt \right. \\ &\quad \left. + \left(\int \frac{\delta V(t, p)}{\delta p}(x) \check{\mathcal{L}}_u^\dagger p(x) dx \right) dt \right]. \end{aligned} \quad (\text{A18})$$

By rearranging the equation, the following equation is obtained:

$$-\frac{\partial V(t, p)}{\partial t} = \min_u \left[\bar{f}(t, p, u) + \int \frac{\delta V(t, p)}{\delta p}(x) \check{\mathcal{L}}_u^\dagger p(x) dx \right], \quad (\text{A19})$$

where $V(T, p) = \bar{g}(p)$. We define expected time-backward diffusion operator $\check{\mathcal{L}}_u$ as follows:

$$\begin{aligned} \check{\mathcal{L}}_u w(t, x) &:= \sum_{i=1}^{d_x} \mathbb{E}_{p_t(y|x)} [b_i(t, x, u(t, y))] \frac{\partial w(t, x)}{\partial x_i} \\ &\quad + \frac{1}{2} \sum_{i,j=1}^{d_x} \mathbb{E}_{p_t(y|x)} [D_{ij}(t, x, u(t, y))] \frac{\partial^2 w(t, x)}{\partial x_i \partial x_j}. \end{aligned} \quad (\text{A20})$$

Since

$$\int \frac{\delta V(t, p)}{\delta p}(x) \check{\mathcal{L}}_u^\dagger p(x) dx = \int p(x) \check{\mathcal{L}}_u \frac{\delta V(t, p)}{\delta p}(x) dx, \quad (\text{A21})$$

the following equation is obtained:

$$\begin{aligned} -\frac{\partial V(t, p)}{\partial t} &= \min_u \left[\bar{f}(t, p, u) + \int p(x) \check{\mathcal{L}}_u \frac{\delta V(t, p)}{\delta p}(x) dx \right] \\ &= \min_u \mathbb{E}_{p_t(y|x)p(x)} \left[f(t, x, u) + \mathcal{L}_u \frac{\delta V(t, p)}{\delta p}(x) \right], \end{aligned} \quad (\text{A22})$$

where $\mathcal{L}_u$ is the time-backward diffusion operator [Eq. (A8)]. From the definition of the Hamiltonian [Eq. (A7)], the following equation is obtained:

$$-\frac{\partial V(t, p)}{\partial t} = \min_u \mathbb{E}_{p_t(y|x)p(x)} \left[\mathcal{H} \left(t, x, u, \frac{\delta V(t, p)}{\delta p}(x) \right) \right], \quad (\text{A23})$$

where $V(T, p) = \mathbb{E}_{p(x)}[g(x)]$. This time-backward functional differential equation is called Bellman equation.

Optimal control function u^* is given by the right-hand side of Bellman equation [Eq. (A23)]. Since control u is determined based on observation y in OSC, the minimization by control u can be exchanged with the expectation by observation y as follows:

$$-\frac{\partial V(t, p)}{\partial t} = \mathbb{E}_{p_t(y)} \left[\min_u \mathbb{E}_{p_t(x|y)} \left[\mathcal{H} \left(t, x, u, \frac{\delta V(t, p)}{\delta p}(x) \right) \right] \right], \quad (\text{A24})$$

where

$$p_t(y) := \int p_t(y|x)p(x)dx, \quad (\text{A25})$$

$$p_t(x|y) := \frac{p_t(y|x)p(x)}{\int p_t(y|x)p(x)dx}. \quad (\text{A26})$$

As a result, optimal control function u^* is given by minimizing the conditional expected Hamiltonian as follows:

$$u^*(t, y, p) = \arg \min_u \mathbb{E}_{p_t(x|y)} \left[\mathcal{H} \left(t, x, u, \frac{\delta V(t, p)}{\delta p}(x) \right) \right], \quad (\text{A27})$$

where $V(t, p)$ is the solution of the following Bellman equation:

$$-\frac{\partial V(t, p)}{\partial t} = \mathbb{E}_{p_t(y|x)p(x)} \left[\mathcal{H} \left(t, x, u^*, \frac{\delta V(t, p)}{\delta p}(x) \right) \right], \quad (\text{A28})$$

where $V(T, p) = \mathbb{E}_{p(x)}[g(x)]$. Therefore, from Bellman's dynamic programming principle on the probability density function space, the optimal control function of OSC [Eq. (A27)] is obtained by solving Bellman equation [Eq. (A28)] backward in time from terminal condition $V(T, p) = \mathbb{E}_{p(x)}[g(x)]$.

In Bellman's dynamic programming principle on the probability density function space, the optimal control function of OSC [Eq. (A27)] depends on probability density function p as $u^*(t, y, p)$, while it is defined by $u^*(t, y)$. In OSC, probability density function p is given by the solution of FP equation [Eq. (A14)], which is denoted by p_t . Therefore, in Bellman's dynamic programming principle on the probability density function space, we first compute Bellman equation [Eq. (A28)] backward in time from terminal condition $V(T, p) = \mathbb{E}_{p(x)}[g(x)]$ and obtain optimal control function $u^*(t, y, p)$. We then compute FP equation [Eq. (A14)] forward in time from initial condition $p(0, x) = p_0(x)$ and obtain optimal control function $u^*(t, y) := u^*(t, y, p_t)$. The agent determines optimal control u_t^* based on real-time observation y_t as $u_t^* := u^*(t, y_t)$.

c. Conversion to Pontryagin's minimum principle on probability density function space

Bellman's dynamic programming principle on the probability density function space is less practical because Bellman equation [Eq. (A28)] is a functional differential equation, not a partial differential equation. We circumvent this problem by converting Bellman's dynamic programming principle on the probability density function space to Pontryagin's minimum principle on the same space. This conversion has also

been used in mean-field stochastic optimal control problems [119–122] and in our previous works [113,115].

We first define

$$W(t, p, x) := \frac{\delta V(t, p)}{\delta p}(x), \quad (\text{A29})$$

which satisfies $W(T, p, x) = g(x)$. Differentiating Bellman equation [Eq. (A28)] with respect to probability density function p , the following equation is obtained:

$$\begin{aligned} -\frac{\partial W(t, p, x)}{\partial t} &= \mathbb{E}_{p_t(y|x)} [\mathcal{H}(t, x, u^*, W)] \\ &+ \mathbb{E}_{p_t(y|x')p(x')} \left[\frac{\partial \mathcal{H}(t, x', u^*, W)}{\partial u} \frac{\delta u^*(t, y, p)}{\delta p}(x) \right] \\ &+ \mathbb{E}_{p_t(y|x')p(x')} \left[\mathcal{L}_{u^*} \frac{\delta W(t, p, x')}{\delta p}(x) \right] \\ &= \mathbb{E}_{p_t(y|x)} [\mathcal{H}(t, x, u^*, W)] \\ &+ \mathbb{E}_{p_t(y)p_t(x'|y)} \left[\frac{\partial \mathcal{H}(t, x', u^*, W)}{\partial u} \frac{\delta u^*(t, y, p)}{\delta p}(x) \right] \\ &+ \mathbb{E}_{p(x')} \left[\check{\mathcal{L}}_{u^*} \frac{\delta W(t, p, x')}{\delta p}(x) \right] \\ &= \mathbb{E}_{p_t(y|x)} [\mathcal{H}(t, x, u^*, W)] \\ &+ \mathbb{E}_{p_t(y)} \left[\frac{\partial \mathbb{E}_{p_t(x'|y)} [\mathcal{H}(t, x', u^*, W)]}{\partial u} \frac{\delta u^*(t, y, p)}{\delta p}(x) \right] \\ &+ \int \frac{\delta W(t, p, x')}{\delta p}(x') \check{\mathcal{L}}_{u^*}^\dagger p(x') dx'. \end{aligned} \quad (\text{A30})$$

From the optimal control function of OSC [Eq. (A27)], the following stationary condition is satisfied:

$$\frac{\partial \mathbb{E}_{p_t(x'|y)} [\mathcal{H}(t, x', u^*, W)]}{\partial u} = 0. \quad (\text{A31})$$

Therefore, the following equation is obtained:

$$\begin{aligned} -\frac{\partial W(t, p, x)}{\partial t} &= \mathbb{E}_{p_t(y|x)} [\mathcal{H}(t, x, u^*, W)] \\ &+ \int \frac{\delta W(t, p, x')}{\delta p}(x') \check{\mathcal{L}}_{u^*}^\dagger p(x') dx'. \end{aligned} \quad (\text{A32})$$

We then define

$$w(t, x) := W(t, p_t, x), \quad (\text{A33})$$

where p_t is the solution of FP equation [Eq. (A14)]. The time derivative of $w(t, x)$ is calculated as follows:

$$\frac{\partial w(t, x)}{\partial t} = \frac{\partial W(t, p_t, x)}{\partial t} + \int \frac{\delta W(t, p_t, x')}{\delta p}(x') \frac{\partial p_t(x')}{\partial t} dx'. \quad (\text{A34})$$

By substituting Eq. (A32) into Eq. (A34), the following equation is obtained:

$$\begin{aligned} -\frac{\partial w(t, x)}{\partial t} &= \mathbb{E}_{p_t(y|x)} [\mathcal{H}(t, x, u^*, w)] - \int \frac{\delta W(t, p_t, x')}{\delta p}(x') \\ &\times \underbrace{\left(\frac{\partial p_t(x')}{\partial t} - \check{\mathcal{L}}_{u^*}^\dagger p_t(x') \right)}_{(*)} dx'. \end{aligned} \quad (\text{A35})$$

From FP equation [Eq. (A14)], $(*) = 0$ holds and Bellman equation [Eq. (A28)] is reduced to the following time-backward partial differential equation, HJB equation:

$$-\frac{\partial w(t, x)}{\partial t} = \mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, u^*, w)], \quad (\text{A36})$$

where $w(T, x) = g(x)$. The optimal control function of OSC [Eq. (A27)] is given as follows:

$$u^*(t, y) = \arg \min_u \mathbb{E}_{p_t(x|y)}[\mathcal{H}(t, x, u, w)], \quad (\text{A37})$$

where $p_t(x|y) := p_t(y|x)p_t(x)/\int p_t(y|x)p_t(x)dx$ and $p_t(x) := p(t, x)$ is the solution of FP equation [Eq. (A14)]. Therefore, the optimal control function of OSC [Eq. (A37)] is obtained by solving the time-backward HJB equation [Eq. (A36)] and the time-forward FP equation [Eq. (A14)]. This optimality condition corresponds to Pontryagin's minimum principle on the probability density function space [Eq. (A6)].

In Pontryagin's minimum principle on the state space, the time-forward state equation and the time-backward adjoint equation can be expressed by the derivatives of the Hamiltonian [199–201]. Similarly, in Pontryagin's minimum principle on the probability density function space, the time-forward FP equation [Eq. (A14)] and the time-backward HJB equation [Eq. (A36)] can be expressed by the variations of expected Hamiltonian

$$\bar{\mathcal{H}}(t, p, u, w) := \mathbb{E}_{p(x)}[\mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, u, w)]] \quad (\text{A38})$$

as follows:

$$\frac{\partial p(t, x)}{\partial t} = \frac{\delta \bar{\mathcal{H}}(t, p, u^*, w)}{\delta w}(x), \quad (\text{A39})$$

$$-\frac{\partial w(t, x)}{\partial t} = \frac{\delta \bar{\mathcal{H}}(t, p, u^*, w)}{\delta p}(x), \quad (\text{A40})$$

where $p(0, x) = p_0(x)$ and $w(T, x) = g(x)$ [114].

While Bellman's dynamic programming principle on the probability density function space is a terminal value problem of a functional differential equation, Bellman equation [Eq. (A28)], Pontryagin's minimum principle on that space is a two-point boundary value problem of partial differential equations, HJB and FP equations [Eqs. (A39) and (A40)], which simplifies computations. We note that while Bellman's dynamic programming principle is a necessary and sufficient condition for the optimal control function, Pontryagin's minimum principle is a necessary condition, not a sufficient condition [114]. Pontryagin's minimum principle becomes a necessary and sufficient condition when expected Hamiltonian $\bar{\mathcal{H}}(t, p, u, w)$ is convex with respect to p and u , which is shown in our previous work [114].

4. Relationship with the Method of Lagrange Multipliers

In this subsection, we show that Pontryagin's minimum principle corresponds to the method of Lagrange multipliers for readers who may be unfamiliar with Pontryagin's minimum principle. The method of Lagrange multipliers is a widely used technique for solving constrained optimization problems. The deterministic control on the probability density

function space can be interpreted as the following constrained optimization problem:

$$\text{minimize } J[u, p] = \int_0^T \bar{f}(t, p_t, u_t)dt + \bar{g}(p_T), \quad (\text{A41})$$

$$\text{subject to } \frac{\partial p(t, x)}{\partial t} = \check{\mathcal{L}}_u^\dagger p(t, x), \quad p(0, x) = p_0(x). \quad (\text{A42})$$

As a result, the method of Lagrange multipliers is an effective approach to solving this problem.

In the method of Lagrange multipliers, we consider the following Lagrange function:

$$\begin{aligned} \tilde{J} := & \int_0^T \bar{f}(t, p_t, u_t)dt + \bar{g}(p_T) \\ & + \int_0^T \int w(t, x) \left(\check{\mathcal{L}}_u^\dagger p(t, x) - \frac{\partial p(t, x)}{\partial t} \right) dx dt, \end{aligned} \quad (\text{A43})$$

where $w(t, x)$ is a Lagrange multiplier. Lagrange function $\tilde{J}$ is calculated as follows:

$$\begin{aligned} \tilde{J} = & \int_0^T \bar{f}(t, p_t, u_t)dt + \bar{g}(p_T) \\ & + \int_0^T \int p(t, x) \check{\mathcal{L}}_u w(t, x) dx dt \\ & - \int (w(T, x)p(T, x) - w(0, x)p(0, x))dx \\ & + \int_0^T \int p(t, x) \frac{\partial w(t, x)}{\partial t} dx dt \\ = & \int_0^T \iint p(t, x)p_t(y|x)\mathcal{H}(t, x, u, w) dx dy dt + \int p(T, x) \\ & \times g(x)dx - \int (p(T, x)w(T, x) - p(0, x)w(0, x))dx \\ & + \int_0^T \int p(t, x) \frac{\partial w(t, x)}{\partial t} dx dt. \end{aligned} \quad (\text{A44})$$

We variate Lagrange function $\tilde{J}$ with respect to u and p as follows:

$$\begin{aligned} \delta \tilde{J} = & \int_0^T \iint p_t(y|x)\mathcal{H}(t, x, u, w)\delta p(t, x) dx dy dt \\ & + \int_0^T \iint p(t, x)p_t(y|x) \frac{\partial \mathcal{H}(t, x, u, w)}{\partial u} \delta u(t, y) dx dy dt \\ & + \int g(x)\delta p(T, x)dx - \int w(T, x)\delta p(T, x)dx \\ & + \int_0^T \int \frac{\partial w(t, x)}{\partial t} \delta p(t, x) dx dt \\ = & \int_0^T \int \left(\mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, u, w)] + \frac{\partial w(t, x)}{\partial t} \right) \\ & \times \delta p(t, x) dx dt + \int (g(x) - w(T, x))\delta p(T, x)dx \\ & + \int_0^T \int p(t, y)\mathbb{E}_{p_t(x|y)} \left[\frac{\partial \mathcal{H}(t, x, u, w)}{\partial u} \right] \delta u(t, y) dy dt. \end{aligned} \quad (\text{A45})$$

From stationary condition $\delta\tilde{J} = 0$, we obtain the following optimality condition:

$$\mathbb{E}_{p_t(x|y)} \left[\frac{\partial \mathcal{H}(t, x, u, w)}{\partial u} \right] = 0, \quad (\text{A46})$$

$$-\frac{\partial w(t, x)}{\partial t} = \mathbb{E}_{p_t(y|x)} [\mathcal{H}(t, x, u, w)], \quad w(T, x) = g(x), \quad (\text{A47})$$

$$\frac{\partial p(t, x)}{\partial t} = \tilde{\mathcal{L}}_u^+ p(t, x), \quad p(0, x) = p_0(x). \quad (\text{A48})$$

Equations (A6), (A9), and (A11) correspond to Eqs. (A46)–(A48), respectively. Therefore, Pontryagin's minimum principle corresponds to the method of Lagrange multipliers.

APPENDIX B: OBSERVATION-AND-MEMORY-BASED STOCHASTIC CONTROL

In this section, we derive the optimal solution of OMSC [Fig. 12(b)]. OMSC corresponds to the optimal estimation and control problem with resource limitations presented in Sec. V. Additionally, the optimal estimation problem with resource limitations presented in Sec. III is a special case of OMSC. Therefore, deriving the optimal solution of OMSC corresponds to deriving the optimal solutions presented in Sec. III [Eq. (18)] and Sec. V [Eq. (41)].

We note that OMSC is similar to our previous works [113–115], but the problem formulation is slightly modified, allowing for more biologically plausible discussions. A comparison between OMSC and our previous works [113–115] is provided in Appendix F.

This section is organized as follows: In Appendix B 1, we formulate OMSC. In Appendix B 2, we show that OMSC can be converted into OSC on an extended space (Fig. 13). In Appendix B 3, we derive the optimal solution of OMSC based on Appendix B 2.

1. Problem formulation

In this subsection, we formulate the optimal estimation and control problem with memory, OMSC [Fig. 12(b)]. OMSC considers the same state $x_t \in \mathbb{R}^{d_x}$ and observation $y_t \in \mathbb{R}^{d_y}$ as OSC, which are given by

$$dx_t = b(t, x_t, u_t)dt + \sigma(t, x_t, u_t)d\omega_t, \quad (\text{B1})$$

$$y_t \sim p_t(y_t|x_t), \quad (\text{B2})$$

where initial state x_0 follows $p_0(x_0)$, $u_t \in \mathbb{R}^{d_u}$ is the state control of the agent, and $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process. In contrast, OMSC formulates the memory of the agent, $z_t \in \mathbb{R}^{d_z}$, which evolves by the following SDE:

$$dz_t = c(t, z_t, v_t)dt + \eta(t, z_t, v_t)d\xi_t, \quad (\text{B3})$$

where initial memory z_0 follows $p_0(z_0)$, $v_t \in \mathbb{R}^{d_v}$ is the memory control of the agent, and $\xi_t \in \mathbb{R}^{d_\xi}$ is a standard Wiener process. OMSC determines state control u_t and memory control v_t based on observation y_t and memory z_t as follows:

$$u_t = u(t, y_t, z_t), \quad (\text{B4})$$

$$v_t = v(t, y_t, z_t). \quad (\text{B5})$$

The objective function is given by the following expected cumulative cost function:

$$J[u] := \mathbb{E} \left[\int_0^T f(t, x_t, z_t, u_t, v_t)dt + g(x_T, z_T) \right], \quad (\text{B6})$$

where f and g are the running and terminal cost functions, respectively. OMSC considers a problem to find optimal state control function u^* and optimal memory control function v^* that minimize objective function $J[u]$:

$$u^* := \arg \min_u J[u]. \quad (\text{B7})$$

OMSC is identical to the optimal estimation and control problem under resource limitations presented in Sec. V. Additionally, the optimal estimation problem under resource limitations presented in Sec. III is a special case of OMSC, where $b(t, x_t, u_t) = b(t, x_t)$ and $\sigma(t, x_t, u_t) = \sigma(t, x_t)$, and state control $u_t = u(t, y_t, z_t)$ is replaced by state estimator $\hat{x}_t = \hat{x}(t, y_t, z_t)$.

2. Conversion to observation-based stochastic control on extended space

In this subsection, we show that OMSC can be converted into OSC on an extended space (Fig. 13) as a preparatory step to obtaining the optimal solution of OMSC. We define extended state $\tilde{x}_t$, extended observation $\tilde{y}_t$, and extended control $\tilde{u}_t$ as follows:

$$\tilde{x}_t := \begin{pmatrix} x_t \\ z_t \end{pmatrix}, \quad \tilde{y}_t := \begin{pmatrix} y_t \\ z'_t \end{pmatrix}, \quad \tilde{u}_t := \begin{pmatrix} u_t \\ v_t \end{pmatrix}. \quad (\text{B8})$$

We denote the memory in extended state $\tilde{x}_t$ by z_t and that in extended observation $\tilde{y}_t$ by z'_t to distinguish between them. In practice, $z_t = z'_t$ holds. Extended state $\tilde{x}_t$ evolves by the following SDE:

$$d\tilde{x}_t = \tilde{b}(t, \tilde{x}_t, \tilde{u}_t)dt + \tilde{\sigma}(t, \tilde{x}_t, \tilde{u}_t)d\tilde{\omega}_t, \quad (\text{B9})$$

where

$$\begin{aligned} \tilde{b}(t, \tilde{x}_t, \tilde{u}_t) &:= \begin{pmatrix} b(t, x_t, u_t) \\ c(t, z_t, v_t) \end{pmatrix}, \\ \tilde{\sigma}(t, \tilde{x}_t, \tilde{u}_t) &:= \begin{pmatrix} \sigma(t, x_t, u_t) & O \\ O & \eta(t, z_t, v_t) \end{pmatrix}, \\ \tilde{\omega}_t &:= \begin{pmatrix} \omega_t \\ \xi_t \end{pmatrix}. \end{aligned} \quad (\text{B10})$$

Extended observation $\tilde{y}_t$ is generated from the following probability density function:

$$\tilde{p}_t(\tilde{y}_t|\tilde{x}_t) := p_t(y_t|x_t)\delta(z'_t - z_t). \quad (\text{B11})$$

Extended control $\tilde{u}_t$ is determined based on extended observation $\tilde{y}_t$ as follows:

$$\tilde{u}_t = \tilde{u}(t, \tilde{y}_t) := \begin{pmatrix} u(t, y_t, z_t) \\ v(t, y_t, z_t) \end{pmatrix}. \quad (\text{B12})$$

The objective function is given by the following expected cumulative cost function:

$$J[\tilde{u}] := \mathbb{E} \left[\int_0^T \tilde{f}(t, \tilde{x}_t, \tilde{u}_t)dt + \tilde{g}(\tilde{x}_T) \right], \quad (\text{B13})$$

where

$$\begin{aligned}\tilde{f}(t, \tilde{x}_t, \tilde{u}_t) &:= f(t, x_t, z_t, u_t, v_t), \\ \tilde{g}(\tilde{x}_T) &:= g(x_T, z_T).\end{aligned}\quad (\text{B14})$$

Therefore, OMSC can be interpreted as OSC on the extended space.

3. Optimal solution

The optimal solution of OMSC can be obtained in a similar way to that of OSC because OMSC can be interpreted as OSC on the extended space. In the following, we present the optimal solution of OMSC, which corresponds to the optimal solutions presented in Sec. III [Eq. (18)] and Sec. V [Eq. (41)]. We note that the following result is identical to the optimal solution of OSC except for $\tilde{\cdot}$.

The optimal extended control function of OMSC is given by minimizing the conditional expectation of the extended Hamiltonian:

$$\tilde{u}^*(t, \tilde{y}) = \arg \min_{\tilde{u}} \mathbb{E}_{\tilde{p}_t(\tilde{x}|\tilde{y})}[\tilde{\mathcal{H}}(t, \tilde{x}, \tilde{u}, \tilde{w})]. \quad (\text{B15})$$

Extended Hamiltonian $\tilde{\mathcal{H}}$ is defined by

$$\tilde{\mathcal{H}}(t, \tilde{x}, \tilde{u}, \tilde{w}) := \tilde{f}(t, \tilde{x}, \tilde{u}) + \tilde{\mathcal{L}}_{\tilde{u}} \tilde{w}(t, \tilde{x}), \quad (\text{B16})$$

where $\tilde{f}(t, \tilde{x}, \tilde{u})$ is the current cost and $\tilde{\mathcal{L}}_{\tilde{u}} \tilde{w}(t, \tilde{x})$ represents the future cost. $\tilde{\mathcal{L}}_{\tilde{u}}$ is the time-backward diffusion operator for extended state $\tilde{x}$, which is defined by

$$\begin{aligned}\tilde{\mathcal{L}}_{\tilde{u}} \tilde{w}(t, \tilde{x}) &:= \sum_{i=1}^{d_{\tilde{x}}} \tilde{b}_i(t, \tilde{x}, \tilde{u}) \frac{\partial \tilde{w}(t, \tilde{x})}{\partial \tilde{x}_i} \\ &+ \frac{1}{2} \sum_{i,j=1}^{d_{\tilde{x}}} \tilde{D}_{ij}(t, \tilde{x}, \tilde{u}) \frac{\partial^2 \tilde{w}(t, \tilde{x})}{\partial \tilde{x}_i \partial \tilde{x}_j},\end{aligned}\quad (\text{B17})$$

where $\tilde{D}(t, \tilde{x}, \tilde{u}) := \tilde{\sigma}(t, \tilde{x}, \tilde{u}) \tilde{\sigma}^\top(t, \tilde{x}, \tilde{u})$. $\tilde{w}(t, \tilde{x})$ is the value function, which represents the expected cumulative cost from the starting point at time t and extended state $\tilde{x}$ to the terminal time T . Value function $\tilde{w}(t, \tilde{x})$ is obtained by solving the following time-backward partial differential equation, HJB equation:

$$-\frac{\partial \tilde{w}(t, \tilde{x})}{\partial t} = \mathbb{E}_{\tilde{p}_t(\tilde{y}|\tilde{x})}[\tilde{\mathcal{H}}(t, \tilde{x}, \tilde{u}^*, \tilde{w})], \quad (\text{B18})$$

where the terminal condition is given by $\tilde{w}(T, \tilde{x}) = \tilde{g}(\tilde{x})$. In the optimal extended control function of OMSC [Eq. (B15)], conditional expectation $\mathbb{E}_{\tilde{p}_t(\tilde{x}|\tilde{y})}[\cdot]$ appears. Conditional probability density function $\tilde{p}_t(\tilde{x}|\tilde{y})$ is calculated as follows:

$$\tilde{p}_t(\tilde{x}|\tilde{y}) = \frac{\tilde{p}_t(\tilde{y}|\tilde{x}) \tilde{p}_t(\tilde{x})}{\int \tilde{p}_t(\tilde{y}|\tilde{x}) \tilde{p}_t(\tilde{x}) d\tilde{x}}. \quad (\text{B19})$$

Probability density function $\tilde{p}_t(\tilde{x}) := \tilde{p}(t, \tilde{x})$ is obtained by solving the following time-forward partial differential equation, FP equation:

$$\frac{\partial \tilde{p}(t, \tilde{x})}{\partial t} = \tilde{\mathcal{L}}_{\tilde{u}^*}^\dagger \tilde{p}(t, \tilde{x}), \quad (\text{B20})$$

where the initial condition is given by $\tilde{p}(0, \tilde{x}) = \tilde{p}_0(\tilde{x})$. $\tilde{\mathcal{L}}_{\tilde{u}}^\dagger$ is the expected time-forward diffusion operator for extended

state $\tilde{x}$, which is defined by

$$\begin{aligned}\tilde{\mathcal{L}}_{\tilde{u}}^\dagger \tilde{p}(t, \tilde{x}) &:= - \sum_{i=1}^{d_{\tilde{x}}} \frac{\partial (\mathbb{E}_{\tilde{p}_t(\tilde{y}|\tilde{x})}[\tilde{b}_i(t, \tilde{x}, \tilde{u}(t, \tilde{y}))] \tilde{p}(t, \tilde{x}))}{\partial \tilde{x}_i} \\ &+ \frac{1}{2} \sum_{i,j=1}^{d_{\tilde{x}}} \frac{\partial^2 (\mathbb{E}_{\tilde{p}_t(\tilde{y}|\tilde{x})}[\tilde{D}_{ij}(t, \tilde{x}, \tilde{u}(t, \tilde{y}))] \tilde{p}(t, \tilde{x}))}{\partial \tilde{x}_i \partial \tilde{x}_j}.\end{aligned}\quad (\text{B21})$$

Therefore, the optimal extended control function of OMSC [Eq. (B15)] is obtained by solving HJB equation [Eq. (B18)] and FP equation [Eq. (B20)].

APPENDIX C: LINEAR-QUADRATIC-GAUSSIAN PROBLEM

In this section, we apply OSC and OMSC to the LQG problem, where the dynamics are linear Gaussian and the costs are quadratic. The LQG problem is a well-known problem in stochastic optimal control theory, which allows for semianalytical calculations [103–106]. We demonstrate that this advantageous property also holds in OSC and OMSC.

The minimal models discussed in Secs. IV and VI correspond to the infinite-horizon LQG problem. The relationship between these models and the infinite-horizon LQG problem is discussed in Appendixes D and E.

This section is organized as follows: In Appendix C 1, we formulate the LQG problem. In Appendix C 2, we present the optimal solution. In Appendix C 3, we provide the derivation of the optimal solution. In Appendix C 4, we extend the LQG problem from a finite-horizon setting to an infinite-horizon setting. We mainly discuss the case of OSC for the notational simplicity, but the same discussion is possible for the case of OMSC by adding $\tilde{\cdot}$, which represents extended variables and parameters.

1. Problem formulation

In this subsection, we formulate the LQG problem in OSC. In the LQG problem, the state of the environment, $x_t \in \mathbb{R}^{d_x}$, evolves by the following linear Gaussian dynamics:

$$dx_t = (Ax_t + Bu_t)dt + \sigma d\omega_t, \quad (\text{C1})$$

where initial state x_0 follows Gaussian distribution $\mathcal{N}(x_0|\mu_0, \Sigma_0)$, $u_t \in \mathbb{R}^{d_u}$ is the state control of the agent, and $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process. The agent receives observation y_t generated from the following Gaussian distribution:

$$y_t \sim \mathcal{N}(y_t|Hx_t, E). \quad (\text{C2})$$

The agent determines control u_t based on observation y_t as $u_t = u(t, y_t)$. The objective function is defined as the following quadratic function:

$$J[u] := \mathbb{E} \left[\int_0^T (x_t^\top Q x_t - 2x_t^\top S u_t + u_t^\top R u_t) dt + x_T^\top P x_T \right],$$

where $Q \geq O$, $R > O$, and $P \geq O$. The LQG problem is to find optimal control function u^* that minimizes objective function $J[u]$ as follows:

$$u^* := \arg \min_u J[u]. \quad (\text{C3})$$

We note that the LQG problem in OMSC is formulated by adding $\tilde{\cdot}$.

2. Optimal solution

In this subsection, we present the optimal solution in the LQG problem of OSC. The derivation of the optimal solution is provided in the following subsection (Appendix C3).

The optimal control function in the LQG problem of OSC is given by

$$u^*(t, y) = -R^{-1}((\Psi B - S)^\top \mu + (\Pi B - S)^\top K(y - H\mu)), \quad (\text{C4})$$

where $K := \Sigma H^\top (E + H \Sigma H^\top)^{-1}$. μ and Σ are the mean vector and the covariance matrix of state x , respectively, which are the solutions of the following time-forward ordinary differential equations:

$$\frac{d\mu}{dt} = (A - BR^{-1}(\Psi B - S)^\top) \mu, \quad (\text{C5})$$

$$\begin{aligned} \frac{d\Sigma}{dt} &= \sigma \sigma^\top + (A - BR^{-1}(\Pi B - S)^\top K H) \Sigma \\ &\quad + \Sigma (A - BR^{-1}(\Pi B - S)^\top K H)^\top, \end{aligned} \quad (\text{C6})$$

where the initial conditions are $\mu(0) = \mu_0$ and $\Sigma(0) = \Sigma_0$. Ψ and Π are the solutions of the following time-backward ordinary differential equations:

$$\begin{aligned} -\frac{d\Psi}{dt} &= Q + A^\top \Psi + \Psi A - (\Psi B - S)R^{-1}(\Psi B - S)^\top, \\ -\frac{d\Pi}{dt} &= Q + A^\top \Pi + \Pi A - (\Pi B - S)R^{-1}(\Pi B - S)^\top \\ &\quad + (I - KH)^\top (\Pi B - S)R^{-1}(\Pi B - S)^\top (I - KH), \end{aligned} \quad (\text{C7}) \quad (\text{C8})$$

where the terminal conditions are $\Psi(T) = \Pi(T) = P$.

Equation (C7) can be solved backward in time independently. Equation (C5) can be solved forward in time based on the solution of Eq. (C7). Equations (C6) and (C8) are mutually dependent and need to be solved simultaneously, which correspond to a two-point boundary value problem of ordinary differential equations. The LQG problem reduces the two-point boundary value problem of partial differential equations, HJB and FP equations, to that of ordinary differential equations, Eqs. (C6) and (C8), allowing for efficient computation. This two-point boundary value problem is solved by using the forward-backward sweep method, which computes Eqs. (C6) and (C8) alternately [114].

The optimal control function in the LQG problem of OSC [Eq. (C4)] is decomposed into a deterministic first term and a stochastic second term. The deterministic first term of Eq. (C4) also appears in conventional optimal control problems, and Eq. (C7) is known as Riccati equation [104–106]. In contrast, the stochastic second term of Eq. (C4) is unique to OSC, and we refer to Eq. (C8) as observation-based Riccati equation. Although Riccati equation and observation-based Riccati equation are similar, the latter includes an additional term that accounts for the observation uncertainty in OSC. In fact, if observation y_t is identical to state x_t , i.e., $y_t = x_t$,

$H = I$ and $E = O$ hold, and the last term of observation-based Riccati equation vanishes, reducing it to Riccati equation.

3. Derivation of optimal solution

In this subsection, we derive the optimal control function in the LQG problem of OSC. From Pontryagin's minimum principle on the probability density function space, the optimal control function of OSC is given by

$$u^*(t, y) = \arg \min_u \mathbb{E}_{p_t(x|y)}[\mathcal{H}(t, x, u, w)]. \quad (\text{C9})$$

In the LQG problem, Hamiltonian $\mathcal{H}(t, x, u, w)$ is given by

$$\begin{aligned} \mathcal{H}(t, x, u, w) &= x^\top Q x - 2x^\top S u + u^\top R u + \left(\frac{\partial w(t, x)}{\partial x} \right)^\top (A x + B u) \\ &\quad + \frac{1}{2} \text{tr} \left(\frac{\partial}{\partial x} \left(\frac{\partial w(t, x)}{\partial x} \right)^\top \sigma \sigma^\top \right). \end{aligned} \quad (\text{C10})$$

From

$$\begin{aligned} \frac{\partial \mathbb{E}_{p_t(x|y)}[\mathcal{H}(t, x, u, w)]}{\partial u} &= -2S^\top \mathbb{E}_{p_t(x|y)}[x] + 2Ru + B^\top \mathbb{E}_{p_t(x|y)} \left[\frac{\partial w(t, x)}{\partial x} \right], \end{aligned} \quad (\text{C11})$$

the optimal control function of OSC is given by

$$u^*(t, y) = -\frac{1}{2} R^{-1} \left(B^\top \mathbb{E}_{p_t(x|y)} \left[\frac{\partial w(t, x)}{\partial x} \right] - 2S^\top \mathbb{E}_{p_t(x|y)}[x] \right). \quad (\text{C12})$$

We first assume that $w(t, x)$ is given by the following quadratic function:

$$w(t, x) = x^\top \Pi(t) x + \alpha^\top(t) x + \beta(t), \quad (\text{C13})$$

where $\Pi(t)$ is a symmetric matrix. Under this assumption, HJB equation

$$-\frac{\partial w(t, x)}{\partial t} = \mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, u^*, w)] \quad (\text{C14})$$

is calculated as follows:

$$\begin{aligned} & - \left(x^\top \frac{d\Pi}{dt} x + \left(\frac{d\alpha}{dt} \right)^\top x + \frac{d\beta}{dt} \right) \\ &= x^\top (Q + A^\top \Pi + \Pi A) x \\ &\quad + ((A - BR^{-1}(\Pi B - S)^\top)^\top \alpha)^\top x \\ &\quad + \text{tr}(\Pi \sigma \sigma^\top) - \frac{1}{4} \alpha^\top B R^{-1} B \alpha \\ &\quad + \mathbb{E}_{p_t(y|x)}[\mathbb{E}_{p_t(x|y)}[x]^\top (\Pi B - S) R (\Pi B - S)^\top \mathbb{E}_{p_t(x|y)}[x]] \\ &\quad - 2x^\top (\Pi B - S) R (\Pi B - S)^\top \mathbb{E}_{p_t(y|x)}[\mathbb{E}_{p_t(x|y)}[x]], \end{aligned} \quad (\text{C15})$$

where $\Pi(T) = P$, $\alpha(T) = 0$, and $\beta(T) = 0$. We next assume that $p(t, x)$ is given by the following Gaussian distribution:

$$p(t, x) = \mathcal{N}(x|\mu(t), \Sigma(t)), \quad (\text{C16})$$

where $\mu(t)$ and $\Sigma(t)$ are the mean vector and the covariance matrix of state x , respectively. Under this assumption, $\mathbb{E}_{p_t(x|y)}[x]$ is calculated as follows:

$$\mathbb{E}_{p_t(x|y)}[x] = \mu + K(y - H\mu), \quad (\text{C17})$$

where K is defined by

$$K := \Sigma H^\top (E + H \Sigma H^\top)^{-1}. \quad (\text{C18})$$

Similarly, the following equations hold:

$$\mathbb{E}_{p_t(y|x)}[\mathbb{E}_{p_t(x|y)}[x]] = \mu + KH(x - \mu), \quad (\text{C19})$$

$$\begin{aligned} & \mathbb{E}_{p_t(y|x)}[\mathbb{E}_{p_t(x|y)}[x]^\top (\Pi B - S) R (\Pi B - S)^\top \mathbb{E}_{p_t(x|y)}[x]] \\ &= \text{tr}(K^\top (\Pi B - S) R (\Pi B - S)^\top K E) \\ &+ (\mu + KH(x - \mu))^\top (\Pi B - S) R (\Pi B - S)^\top \\ &\times (\mu + KH(x - \mu)). \end{aligned} \quad (\text{C20})$$

As a result, HJB equation [Eq. (C15)] is reduced to the following ordinary differential equations:

$$\begin{aligned} -\frac{d\Pi}{dt} &= Q + A^\top \Pi + \Pi A - (\Pi B - S) R^{-1} (\Pi B - S)^\top \\ &+ (I - KH)^\top (\Pi B - S) R^{-1} (\Pi B - S)^\top (I - KH), \end{aligned} \quad (\text{C21})$$

$$\begin{aligned} -\frac{d\alpha}{dt} &= (A - BR^{-1}(\Pi B - S)^\top)^\top \alpha \\ &+ 2(I - KH)^\top (\Pi B - S) R^{-1} (\Pi B - S)^\top (I - KH) \mu, \end{aligned} \quad (\text{C22})$$

$$\begin{aligned} -\frac{d\beta}{dt} &= \text{tr}(\Pi \sigma \sigma^\top + K^\top (\Pi B - S) R (\Pi B - S)^\top K E) \\ &- \frac{1}{4} \alpha^\top BR^{-1} B^\top \alpha + \mu^\top (I - KH)^\top \\ &\times (\Pi B - S) R^{-1} (\Pi B - S)^\top (I - KH) \mu, \end{aligned} \quad (\text{C23})$$

where $\Pi(T) = P$, $\alpha(T) = 0$, and $\beta(T) = 0$. If $\Pi(t)$, $\alpha(t)$, and $\beta(t)$ satisfy Eqs. (C21)–(C23), respectively, HJB equation [Eq. (C14)] is satisfied, which indicates that our first assumption [Eq. (C13)] is consistent. The optimal control function of OSC [Eq. (C12)] is calculated as follows:

$$u^*(t, y) = -\frac{1}{2} R^{-1} (2(\Pi B - S)^\top (\mu + K(y - H\mu)) + B^\top \alpha). \quad (\text{C24})$$

Since the optimal control function of OSC [Eq. (C24)] is linear with respect to y , probability density function $p(t, x)$ becomes a Gaussian distribution, which is consistent with our second assumption [Eq. (C16)].

We assume that $\alpha(t) = 2\Upsilon(t)\mu(t)$ holds, where $\Upsilon(t)$ is a symmetric matrix. Under this assumption, the optimal control function of OSC [Eq. (C24)] is calculated as follows:

$$\begin{aligned} u^*(t, y) &= -R^{-1} (((\Pi + \Upsilon)B - S)^\top \mu \\ &+ (\Pi B - S)^\top K(y - H\mu)). \end{aligned} \quad (\text{C25})$$

From

$$\begin{aligned} \frac{d\alpha}{dt} &= 2\frac{d\Upsilon}{dt}\mu + 2\Upsilon\frac{d\mu}{dt}, \\ \frac{d\mu}{dt} &= (A - BR^{-1}((\Pi + \Upsilon)B - S)^\top)\mu \end{aligned} \quad (\text{C26})$$

and Eq. (C22), $\Upsilon(t)$ evolves by the following ordinary differential equation:

$$\begin{aligned} -\frac{d\Upsilon}{dt} &= (A - BR^{-1}(\Pi B - S))^\top \Upsilon + \Upsilon(A - BR^{-1}(\Pi B - S)) \\ &- \Upsilon BR^{-1} B^\top \Upsilon - (I - KH)^\top \Upsilon BR^{-1} B^\top (I - KH), \end{aligned} \quad (\text{C27})$$

where $\Upsilon(T) = O$. From Eq. (C27), $\Upsilon(t)$ is a symmetric matrix, which is consistent with our assumption. By defining $\Psi(t) := \Pi(t) + \Upsilon(t)$, the optimal control function of OSC [Eq. (C25)] is calculated as follows:

$$u^*(t, y) = -R^{-1} ((\Psi B - S)^\top \mu + (\Pi B - S)^\top K(y - H\mu)). \quad (\text{C28})$$

From Eqs. (C21) and (C27), $\Psi(t)$ evolves by the following ordinary differential equation:

$$-\frac{d\Psi}{dt} = Q + A^\top \Psi + \Psi A - (\Psi B - S) R^{-1} (\Psi B - S)^\top, \quad (\text{C29})$$

where $\Psi(T) = P$. Equations. (C28), (C29), and (C21) are Eqs. (C4), (C7), and (C8), respectively. Given the optimal control function [Eq. (C28)], the mean vector and the covariance matrix of state x , $\mu(t)$ and $\Sigma(t)$, are given by the solutions of Eqs. (C5) and (C6), respectively.

4. Infinite-horizon case

In Appendixes C 1–C 3, we discuss the finite-horizon case, in which the objective function is defined by

$$J[u] := \mathbb{E} \left[\int_0^T (x_t^\top Q x_t - 2x_t^\top S u_t + u_t^\top R u_t) dt + x_T^\top P x_T \right]. \quad (\text{C30})$$

In this subsection, we discuss the infinite-horizon case, where the objective function is defined by

$$J[u] := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (x_t^\top Q x_t - 2x_t^\top S u_t + u_t^\top R u_t) dt \right]. \quad (\text{C31})$$

Additionally, since the infinite-horizon case is related with our minimal models in Secs. IV and VI, we mainly consider OMSC rather than OSC by adding $\tilde{\tau}$ as follows:

$$J[\tilde{u}] := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (\tilde{x}_t^\top \tilde{Q} \tilde{x}_t - 2\tilde{x}_t^\top \tilde{S} \tilde{u}_t + \tilde{u}_t^\top \tilde{R} \tilde{u}_t) dt \right], \quad (\text{C32})$$

where $\tilde{\tau}$ indicates the extended variables and parameters. We note that we can make the same discussion for OSC by removing $\tilde{\tau}$.

The optimal control function in the infinite-horizon case is the same as that in the finite-horizon case, which is given by

$$\tilde{u}^*(t, \tilde{y}) = -\tilde{R}^{-1} ((\tilde{\Psi} \tilde{B} - \tilde{S})^\top \tilde{\mu} + (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K}(\tilde{y} - \tilde{H} \tilde{\mu})), \quad (\text{C33})$$

where $\tilde{K} := \tilde{\Sigma} \tilde{H}^\top (\tilde{E} + \tilde{H} \tilde{\Sigma} \tilde{H}^\top)^{-1}$. In contrast, in the infinite-horizon case, the ordinary differential equations of $\tilde{\mu}$, $\tilde{\Sigma}$, $\tilde{\Psi}$, and $\tilde{\Pi}$ become the steady-state equations as follows:

$$0 = (\tilde{A} - \tilde{B} \tilde{R}^{-1} (\tilde{\Psi} \tilde{B} - \tilde{S})^\top) \tilde{\mu}, \quad (\text{C34})$$

$$0 = \tilde{\sigma} \tilde{\sigma}^\top + (\tilde{A} - \tilde{B} \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K} \tilde{H}) \tilde{\Sigma} \\ + \tilde{\Sigma} (\tilde{A} - \tilde{B} \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K} \tilde{H})^\top, \quad (\text{C35})$$

$$0 = \tilde{Q} + \tilde{A}^\top \tilde{\Psi} + \tilde{\Psi} \tilde{A} - (\tilde{\Psi} \tilde{B} - \tilde{S}) \tilde{R}^{-1} (\tilde{\Psi} \tilde{B} - \tilde{S})^\top, \quad (\text{C36})$$

$$0 = \tilde{Q} + \tilde{A}^\top \tilde{\Pi} + \tilde{\Pi} \tilde{A} - (\tilde{\Pi} \tilde{B} - \tilde{S}) \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top + (I - \tilde{K} \tilde{H})^\top \\ \times (\tilde{\Pi} \tilde{B} - \tilde{S}) \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top (I - \tilde{K} \tilde{H}). \quad (\text{C37})$$

$\tilde{\Psi}$ is obtained by solving Eq. (C36). Given $\tilde{\mu}$, $\tilde{\mu}$ is obtained by solving Eq. (C34). $\tilde{\Sigma}$ and $\tilde{\Pi}$ are obtained by solving Eqs. (C35) and (C37) simultaneously.

However, Eq. (C35) is impractical for our discussions in Secs. IV and VI. When $\Pi_{zx} = \Pi_{zz} = 0$, the variance of the memory diverges to infinity as $\Sigma_{zz} \rightarrow \infty$ and Eq. (C35) does not hold. In order to avoid this problem, we convert the steady-state equation for covariance matrix $\tilde{\Sigma}$ [Eq. (C35)] to that for precision matrix $\tilde{\Lambda} := \tilde{\Sigma}^{-1}$ as follows:

$$0 = -(\tilde{A} - \tilde{B} \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K} \tilde{H})^\top \tilde{\Lambda} \\ - \tilde{\Lambda} (\tilde{A} - \tilde{B} \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K} \tilde{H}) - \tilde{\Lambda} \tilde{\sigma} \tilde{\sigma}^\top \tilde{\Lambda}. \quad (\text{C38})$$

When $\Sigma_{zz} \rightarrow \infty$, $\Lambda_{zz} \rightarrow 0$. Therefore, Eq. (C38) holds even when $\Pi_{zx} = \Pi_{zz} = 0$. In Secs. IV and VI, we obtain control matrix $\tilde{\Pi}$ and precision matrix $\tilde{\Lambda}$ by numerically solving Eqs. (C37) and (C38). To solve these equations, we use the NLSolve.jl package in Julia, which implements a trust-region method for solving nonlinear equations. Since Eqs. (C37) and (C38) are derived from Pontryagin's minimum principle, which corresponds to the method of Lagrange multipliers, they can admit multiple solutions, including not only optimal solutions but also extremal solutions. To explore all possible solutions, we solve these equations 1000 times using randomly chosen initial values for the trust-region method. As Eqs. (C37) and (C38) are steady-state equations and do not take into account the stability of the fixed points, they have some unphysical solutions, such as precision matrix Λ is not positive semidefinite, or corresponding objective function J takes negative values. We discard such solutions from our analysis. The optimal solutions minimizing J are identified by directly evaluating and comparing the values of J for each solution of Eqs. (C37) and (C38). In the infinite-horizon case, given $\tilde{\mu}$, $\tilde{\Sigma}$, $\tilde{\Psi}$, and $\tilde{\Pi}$, objective function J is calculated as follows:

$$J[\tilde{u}^*] = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (\tilde{x}_t^\top \tilde{Q} \tilde{x}_t - 2 \tilde{x}_t^\top \tilde{S} \tilde{u}_t^* + (\tilde{u}_t^*)^\top \tilde{R} \tilde{u}_t^*) dt \right] \\ = \mathbb{E}[\tilde{x}^\top \tilde{Q} \tilde{x} - 2 \tilde{x}^\top \tilde{S} \tilde{u}^* + (\tilde{u}^*)^\top \tilde{R} \tilde{u}^*] \\ = \tilde{\mu}^\top \tilde{Q} \tilde{\mu} + 2 \tilde{\mu}^\top \tilde{S} \tilde{R}^{-1} (\tilde{\Psi} \tilde{B} - \tilde{S})^\top \tilde{\mu} \\ + \tilde{\mu}^\top (\tilde{\Psi} \tilde{B} - \tilde{S}) \tilde{R} (\tilde{\Psi} \tilde{B} - \tilde{S})^\top \tilde{\mu} \\ + \text{tr}\{(\tilde{Q} + 2 \tilde{S} \tilde{R}^{-1} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K} \tilde{H} \\ + \tilde{K}^\top (\tilde{\Pi} \tilde{B} - \tilde{S}) \tilde{R} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K}) \tilde{\Sigma}\} \\ + \text{tr}\{\tilde{K}^\top (\tilde{\Pi} \tilde{B} - \tilde{S}) \tilde{R} (\tilde{\Pi} \tilde{B} - \tilde{S})^\top \tilde{K} \tilde{E}\}. \quad (\text{C39})$$

We use this expression to compute J corresponding to $\tilde{\Pi}$ and $\tilde{\Sigma} = \tilde{\Lambda}^{-1}$. We note that $\tilde{\mu} = 0$ and $\tilde{\Psi} = 0$ hold in Secs. IV and VI. To ensure reproducibility, we have made the sample code available on Github [202].

APPENDIX D: DETAILS OF APPLICATION TO BIOLOGICAL INFORMATION PROCESSING

In this section, we provide a detailed explanation of the minimal model of biological information processing, the target estimation problem, presented in Sec. IV. This section is organized as follows: In Appendix D1, we show that LQG structures approximately emerge if systems fluctuate near stable fixed points. In Appendix D2, we show that the minimal model of biological information processing can be approximately realized by a push-pull network in cellular information processing [13,36–41,70]. In Appendix D3, we show that the target estimation problem corresponds to the infinite-horizon LQG problem of OMSC. In Appendix D4, we derive the optimal solution of the target estimation problem [Eqs. (31) and (32)]. In Appendix D5, we show the discontinuous phase transition with respect to the other parameters.

1. Linear-quadratic-Gaussian approximation

LQG structures approximately emerge when systems possess stable fixed points and their fluctuations are sufficiently small. In this subsection, we illustrate this point by focusing on the minimal model of biological information processing introduced in Sec. IV. The same reasoning can also be applied to the target tracking problem discussed in Sec. VI.

In the general case, the state dynamics is given by

$$dx_t = b(x_t)dt + \sigma(x_t)d\omega_t, \quad (\text{D1})$$

where ω_t is a standard Wiener process. For simplicity, we consider the case where x_t is one dimensional, and both b and σ are time invariant. We assume that x_t has a stable fixed point $\bar{x}$, satisfying $b(\bar{x}) = 0$ and $\partial b(\bar{x})/\partial x < 0$. We also assume that fluctuations around $\bar{x}$ are sufficiently small. Under these assumptions, x_t remains close to $\bar{x}$, and we can apply the linear noise approximation [203,204]. This yields the following linear and Gaussian dynamics:

$$dx_t = \frac{\partial b(\bar{x})}{\partial x} (x_t - \bar{x})dt + \sigma(\bar{x})d\omega_t. \quad (\text{D2})$$

By defining

$$\tilde{x}_t := x_t - \bar{x}, \quad \tilde{t} := -\frac{\partial b(\bar{x})}{\partial x} t, \\ d\tilde{\omega}_t := \sqrt{-\frac{\partial b(\bar{x})}{\partial x}} d\omega_t, \quad \tilde{\sigma} := \frac{\sigma(\bar{x})}{\sqrt{-\frac{\partial b(\bar{x})}{\partial x}}}, \quad (\text{D3})$$

we obtain the rescaled dynamics

$$d\tilde{x}_t = -\tilde{x}_t d\tilde{t} + \tilde{\sigma} d\tilde{\omega}_t, \quad (\text{D4})$$

which is equivalent to the state dynamics in our minimal model [Eq. (26)]. For notational simplicity, we hereafter denote $\tilde{t}$ as t .

Next, the observation dynamics is given by

$$dy_t = h(x_t, y_t)dt + \gamma(x_t, y_t)d\nu_t, \quad (\text{D5})$$

where ν_t is a standard Wiener process. For simplicity, we consider the case where y_t is one dimensional, and both h and γ are time invariant. Assuming that y_t has a stable fixed point $\bar{y}$ at $x_t = \bar{x}$, and the fluctuations around $\bar{x}$ and $\bar{y}$ are sufficiently small, we obtain the following linear and Gaussian dynamics from the linear noise approximation [203,204]:

$$dy_t = \left\{ \frac{\partial h(\bar{x}, \bar{y})}{\partial x} \tilde{x} + \frac{\partial h(\bar{x}, \bar{y})}{\partial y} (y_t - \bar{y}) \right\} dt + \gamma(\bar{x}, \bar{y})d\nu_t. \quad (\text{D6})$$

We note that $h(\bar{x}, \bar{y}) = 0$ and $\partial h(\bar{x}, \bar{y})/\partial y < 0$. We also assume that the dynamics of y_t is much faster than those of x_t and z_t . This assumption is natural, since environmental dynamics are typically much slower than sensory dynamics. Moreover, for z_t to serve as memory, its dynamics must be slower than that of y_t . Under this assumption, we obtain the following stationary distribution of y_t :

$$y_t \sim \mathcal{N} \left(y_t \middle| \bar{y} - \left(\frac{\partial h(\bar{x}, \bar{y})}{\partial y} \right)^{-1} \frac{\partial h(\bar{x}, \bar{y})}{\partial x} \tilde{x}, \frac{1}{2} \left(\frac{\partial h(\bar{x}, \bar{y})}{\partial y} \right)^{-1} (\gamma(\bar{x}, \bar{y}))^2 \right). \quad (\text{D7})$$

By defining

$$\begin{aligned} \tilde{y}_t &:= \left(\frac{\partial h(\bar{x}, \bar{y})}{\partial x} \right)^{-1} \frac{\partial h(\bar{x}, \bar{y})}{\partial y} (y_t - \bar{y}), \\ E &:= \frac{1}{2} \left(\frac{\partial h(\bar{x}, \bar{y})}{\partial x} \right)^{-2} \frac{\partial h(\bar{x}, \bar{y})}{\partial y} (\gamma(\bar{x}, \bar{y}))^2, \end{aligned} \quad (\text{D8})$$

we obtain

$$\tilde{y}_t \sim \mathcal{N}(\tilde{y}_t | \tilde{x}, E), \quad (\text{D9})$$

which is equivalent to the observation distribution in our minimal model [Eq. (27)].

Finally, the memory dynamics is given by

$$dz_t = c(z_t, v_t)dt + \eta(z_t, v_t)d\xi_t, \quad (\text{D10})$$

where ξ_t is a standard Wiener process. For simplicity, we consider the case where z_t and v_t are one dimensional, and both c and η are time invariant. Assuming that z_t fluctuates around $\bar{z}$ and the absolute value of v_t is sufficiently small, the above dynamics can be approximated by

$$\begin{aligned} dz_t &= \left\{ c(\bar{z}, 0) + \frac{\partial c(\bar{z}, 0)}{\partial z} (z_t - \bar{z}) + \frac{\partial c(\bar{z}, 0)}{\partial v} v_t \right\} dt \\ &+ \eta(\bar{z}, 0)d\xi_t. \end{aligned} \quad (\text{D11})$$

We note that $c(\bar{z}, 0) = 0$ and $\partial c(\bar{z}, 0)/\partial z < 0$ do not necessarily hold, since the stabilization of z_t can be achieved by control v_t . When $c(\bar{z}, 0) = 0$ and $\partial c(\bar{z}, 0)/\partial z = 0$, this dynamics reduces to

$$dz_t = v_t dt + \tilde{\eta} d\xi_t, \quad (\text{D12})$$

where

$$\tilde{z}_t := \left(\frac{\partial c(\bar{z}, 0)}{\partial v} \right)^{-1} z_t, \quad \tilde{\eta} := \left(\frac{\partial c(\bar{z}, 0)}{\partial v} \right)^{-1} \eta(\bar{z}, 0). \quad (\text{D13})$$

This is the simplest form of memory dynamics, where the drift coefficient depends solely on control v_t . Hence, we adopt this formulation in our minimal model [Eq. (28)].

In the above, we have explained the linear and Gaussian structures of the state, observation, and memory dynamics. In the following, we shift our focus to the quadratic structure of the objective function, specifically in terms of the state estimation error and memory control cost.

We first address the quadratic structure of the state estimation error. Let $f_Q(\hat{x})$ denote the state estimation error, where $\hat{x}$ is the estimator of state x . The state estimation error $f_Q(\hat{x})$ must be minimum and zero when $\hat{x}$ is equal to x . Accordingly, the following conditions must hold:

$$f_Q(x) = 0, \quad \frac{\partial f_Q(x)}{\partial \hat{x}} = 0, \quad \frac{\partial^2 f_Q(x)}{\partial \hat{x}^2} > 0. \quad (\text{D14})$$

Assuming that $\hat{x}$ is sufficiently close to x , we can approximate the state estimation error using a second-order Taylor expansion:

$$f_Q(\hat{x}) \approx Q(\hat{x} - x)^2, \quad (\text{D15})$$

where $Q := \frac{1}{2} \partial^2 f_Q(x) / \partial \hat{x}^2 > 0$. This form corresponds to the state estimation error used in our minimal model.

A similar argument applies to the memory control cost. Let $f_M(v)$ denote the cost of memory control v . In general, no memory control ($v = 0$) is assumed to incur the minimal cost. Hence, the cost function must satisfy

$$\frac{\partial f_M(0)}{\partial v} = 0, \quad \frac{\partial^2 f_M(0)}{\partial v^2} > 0. \quad (\text{D16})$$

Assuming that the absolute value of v is sufficiently small, we can approximate the memory control cost as

$$f_M(v) \approx f_M(0) + Mv^2, \quad (\text{D17})$$

where $M := \frac{1}{2} \partial^2 f_M(0) / \partial v^2 > 0$. The constant term $f_M(0)$ simply sets the reference level of the memory control cost and does not affect the optimization of v . Therefore, we set $f_M(0) = 0$ in our minimal model.

In summary, LQG structures tend to emerge when systems fluctuate around their stable fixed points, indicating that our analysis based on LQG models is not limited to specific cases but may have broad applicability. At the same time, we note that LQG models cannot account for intrinsically nonlinear phenomena such as bistability, limit cycles, or chaos. Extending our analysis to such complex cases remains an important direction for future work.

2. Correspondence to push-pull network in cellular information processing

In this subsection, we show that the minimal model of biological information processing, presented in Sec. IV, can be approximately realized by a push-pull network in cellular

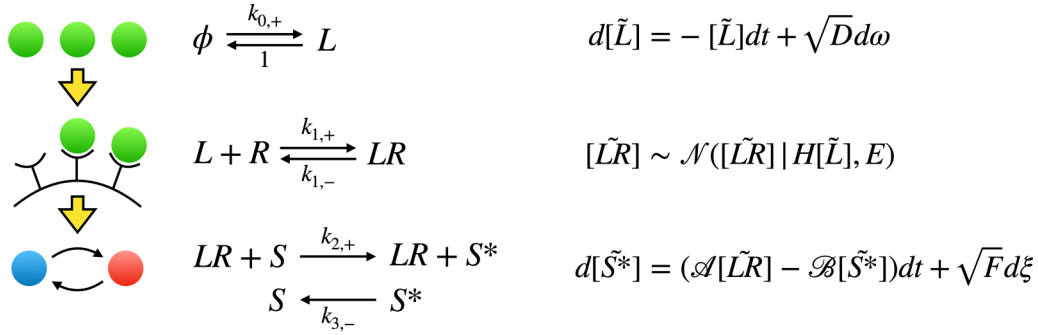
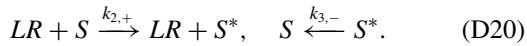


FIG. 14. Schematic diagram of a push-pull network in cellular information processing. L is a ligand, R is a ligand-unbound receptor, LR is a ligand-bound receptor, S is an inactive signaling molecule, and S^* is its active form. In the push-pull network, the ligand-bound receptors promote the activation of the signaling molecules, while the active signaling molecules are spontaneously inactivated. $[\tilde{L}]$, $[\tilde{LR}]$, and $[\tilde{S}^*]$ are the deviations of the concentrations of ligands, ligand-bound receptors, and active signaling molecules from their means, respectively. By applying the linear noise approximation, $[\tilde{L}]$, $[\tilde{LR}]$, and $[\tilde{S}^*]$ can be described as the linear and Gaussian dynamics, which correspond to the state x , observation y , and memory z , respectively, in the minimal model of biological information processing presented in Sec. IV.

information processing (Fig. 14) [13,36–41,70]. The push-pull network is given by the following chemical reactions:



Equation (D18) represents the dynamics of a ligand concentration $[L]$. The ligand concentration fluctuates stochastically, which corresponds to the state dynamics in our framework [13,38–40,72].

Equation (D19) represents the binding and unbinding reactions between ligands and receptors. The concentration of ligand-bound receptors $[LR]$ reflects that of ligands $[L]$, which corresponds to the observation. If the number of receptors is infinitely large, $[LR]$ provides perfect information about $[L]$. However, in real cellular systems, the number of receptors is finite and small, leading to observation noise [8,9,13,36–41,69,72].

Equation (D20) represents the activation and inactivation reactions of signaling molecules, which is called the push-pull network. The concentration of ligand-bound receptors $[LR]$ activates signaling molecules, whereas activated signaling molecules are spontaneously inactivated. The push-pull network acts as memory: The activation reaction accumulates the past information about $[LR]$, and the inactivation reaction erases it. If the number of signaling molecules is infinitely large, memory updating can be deterministic [13,69,72]. However, producing the infinite number of signaling molecules is restricted due to resource limitations, which cause the intrinsic stochasticity of memory dynamics [33,36–41]. Moreover, these reactions are irreversible, which dissipate energy [36–41]. Increasing the reactions rates $k_{2,+}$ and $k_{3,-}$ enhances energy dissipation, which corresponds to the memory control cost.

In summary, the concentrations of ligands $[L]$, ligand-bound receptors $[LR]$, and active signaling molecules $[S^*]$ in the push-pull network correspond to the state x , observation y , and memory z , respectively. In the following, we show a

more detailed correspondence between the push-pull network [Eqs. (D18)–(D20)] and our minimal model [Eqs. (26)–(28)] using the linear noise approximation. The linear noise approximation in the push-pull network is also provided in Ref. [40].

a. Relationship between ligand and state

First, we show the relationship between the ligands and the state. We denote the volume of the environment by V , the number of ligands by L , and the ligand concentration $[L] := L/V$. If V is sufficiently large, the dynamics of ligand concentration $[L]$ can be described by the following chemical Langevin equation:

$$d[L] = (k_{0,+} - [L])dt + \sqrt{\frac{k_{0,+} + [L]}{V}}d\omega, \quad (D21)$$

where $\omega \in \mathbb{R}$ is a standard Wiener process. We assume that the number of ligands is much larger than that of receptors, so the ligand concentration remains unchanged by ligand-receptor binding reactions. Under the assumption of large volume V , we can apply the linear noise approximation and obtain the following Ornstein-Uhlenbeck process:

$$d[\tilde{L}] = -[\tilde{L}]dt + \sqrt{\frac{2k_{0,+}}{V}}d\omega, \quad (D22)$$

where $[\tilde{L}] := [L] - \bar{L}$ and $\bar{L} := k_{0,+}$. This ligand concentration dynamics is equivalent to the state dynamics in the minimal model [Eq. (26)]:

$$dx = -xdt + \sqrt{D}d\omega, \quad (D23)$$

where $x := [\tilde{L}]$ and $D = 2k_{0,+}/V$.

b. Relationship between receptors and observation

Next, we show the relationship between the ligand-bound receptors and the observation. We denote the number of ligand-unbound receptors by R , the number of ligand-bound receptors by LR , and the total number of receptors by $R_T := R + LR$. We also denote the concentration of ligand-unbound receptors by $[R] := R/R_T$ and the concentration of ligand-bound receptors by $[LR] := LR/R_T$. If R_T is sufficiently large,

the dynamics of ligand-receptor binding reactions can be described by the following chemical Langevin equation:

$$d[LR] = (k_{1,+}[L][R] - k_{1,-}[LR])dt + \sqrt{\frac{k_{1,+}[L][R] + k_{1,-}[LR]}{R_T}}d\nu, \quad (D24)$$

where $\nu \in \mathbb{R}$ is a standard Wiener process. From $[R] = 1 - [LR]$, we obtain

$$d[LR] = \{k_{1,+}(1 - [LR])[L] - k_{1,-}[LR]\}dt + \sqrt{\frac{k_{1,+}(1 - [LR])[L] + k_{1,-}[LR]}{R_T}}d\nu. \quad (D25)$$

Under the assumption of large R_T , we can apply the linear noise approximation and obtain the following linear and Gaussian dynamics:

$$d[\tilde{LR}] = \{k_{1,+}(1 - [\tilde{LR}])[\tilde{L}] - (k_{1,+}[\tilde{L}] + k_{1,-})[\tilde{LR}]\}dt + \sqrt{\frac{k_{1,+}(1 - [\tilde{LR}])[\tilde{L}] + k_{1,-}[\tilde{LR}]}{R_T}}d\nu, \quad (D26)$$

where $[\tilde{LR}] := [LR] - [\bar{LR}]$ and $[\tilde{L}] := k_{1,+}[\bar{L}]/(k_{1,+}[\bar{L}] + k_{1,-})$. Assuming that the dynamics of $[\tilde{LR}]$ is much faster than that of $[\tilde{L}]$, we obtain the following stationary distribution:

$$\begin{aligned} [\tilde{LR}] &\sim \mathcal{N}\left([\tilde{LR}] \left| \frac{k_{1,+}(1 - [\tilde{LR}])}{(k_{1,+}[\tilde{L}] + k_{1,-})}[\tilde{L}], \right. \right. \\ &\quad \left. \frac{k_{1,+}(1 - [\tilde{LR}])[\tilde{L}] + k_{1,-}[\tilde{LR}]}{2(k_{1,+}[\tilde{L}] + k_{1,-})R_T} \right) \\ &\sim \mathcal{N}\left([\tilde{LR}] \left| \frac{k_{1,+}k_{1,-}}{(k_{1,+}[\tilde{L}] + k_{1,-})^2}[\tilde{L}], \right. \right. \\ &\quad \left. \frac{k_{1,+}k_{1,-}[\tilde{L}]}{(k_{1,+}[\tilde{L}] + k_{1,-})^2R_T} \right). \end{aligned} \quad (D27)$$

By defining

$$\begin{aligned} y &:= \frac{(k_{1,+}[\tilde{L}] + k_{1,-})^2}{k_{1,+}k_{1,-}}[\tilde{LR}], \\ x &:= [\tilde{L}], \quad E := \frac{(k_{1,+}[\tilde{L}] + k_{1,-})^2[\tilde{L}]}{k_{1,+}k_{1,-}R_T}, \end{aligned} \quad (D28)$$

it reduces to the observation distribution in the minimal model [Eq. (27)]:

$$y \sim \mathcal{N}(y|x, E). \quad (D29)$$

From $E \propto R_T^{-1}$, the observation noise E becomes 0 when the number of receptors R_T is infinitely large. However, the number of receptors R_T is limited in real cells, causing nonzero observation noise E [8,9,13,36–41,69,72].

c. Relationship between signaling molecules and memory

Finally, we show that the push-pull network corresponds to the memory dynamics. We denote the number of inactive signaling molecules by S , the number of active signaling molecules by S^* , and the total number of signaling molecules by $S_T := S + S^*$. We also denote the concentration of inactive

signaling molecules by $[S] := S/S_T$ and the concentration of active signaling molecules by $[S^*] := S^*/S_T$.

If S_T is sufficiently large, the dynamics of the push-pull network can be described by the following chemical Langevin equation:

$$d[S^*] = (k_{2,+}[LR][S] - k_{3,-}[S^*])dt + \sqrt{\frac{k_{2,+}[LR][S] + k_{3,-}[S^*]}{S_T}}d\xi, \quad (D30)$$

where $\xi \in \mathbb{R}$ is a standard Wiener process. From $[S] = 1 - [S^*]$, we obtain

$$d[S^*] = \{k_{2,+}[LR](1 - [S^*]) - k_{3,-}[S^*]\}dt + \sqrt{\frac{k_{2,+}[LR](1 - [S^*]) + k_{3,-}[S^*]}{S_T}}d\xi. \quad (D31)$$

Under the assumption of large S_T , we can apply the linear noise approximation and obtain the following linear and Gaussian dynamics:

$$d[\tilde{S}^*] = \{k_{2,+}(1 - [\tilde{S}^*])[\tilde{LR}] - (k_{2,+}[\tilde{LR}] + k_{3,-})[\tilde{S}^*]\}dt + \sqrt{\frac{k_{2,+}[\tilde{LR}](1 - [\tilde{S}^*]) + k_{3,-}[\tilde{S}^*]}{S_T}}d\xi, \quad (D32)$$

where $[\tilde{S}^*] := [S^*] - [\bar{S}^*]$ and $[\tilde{S}^*] := k_{2,+}[\bar{LR}]/(k_{2,+}[\bar{LR}] + k_{3,-})$. The dynamics of $[\tilde{S}^*]$ corresponds to the memory dynamics in the minimal model [Eq. (28)]:

$$dz = v(y, z)dt + \sqrt{F}d\xi, \quad (D33)$$

where

$$z := [\tilde{S}^*], \quad (D34)$$

$$y := \frac{(k_{1,+}[\tilde{L}] + k_{1,-})^2}{k_{1,+}k_{1,-}}[\tilde{LR}], \quad (D35)$$

$$\begin{aligned} v(y, z) &:= k_{2,+}(1 - [\tilde{S}^*])\frac{k_{1,+}k_{1,-}}{(k_{1,+}[\tilde{L}] + k_{1,-})^2}y \\ &\quad - (k_{2,+}[\tilde{LR}] + k_{3,-})z, \end{aligned} \quad (D36)$$

$$F := \frac{k_{2,+}[\tilde{LR}](1 - [\tilde{S}^*]) + k_{3,-}[\tilde{S}^*]}{S_T}. \quad (D37)$$

The memory control $v(y, z)$ is a linear function with respect to observation y and memory z , which is consistent with the linear optimal control function in the minimal model [Eq. (32)]. Furthermore, $F \propto S_T^{-1}$ indicates that reducing memory noise requires increased production of signaling molecules. However, due to resource limitations, this production is constrained, resulting in inevitable intrinsic noise [33,36–41].

The optimization of memory control $v(y, z)$ corresponds to optimizing the reaction rates $k_{2,+}$ and $k_{3,-}$ in the push-pull network. However, these parameters cannot be optimized within the LQG framework, as they influence not only memory control $v(y, z)$ [Eq. (D36)] but also diffusion constant F [Eq. (D37)]. A simple approximation to avoid this problem is to fix the values of $k_{2,+}$ and $k_{3,-}$ in F by constants $\bar{k}_{2,+}$ and $\bar{k}_{3,-}$, assuming that $|k_{2,+} - \bar{k}_{2,+}|$ and $|k_{3,-} - \bar{k}_{3,-}|$ are

sufficiently small. To investigate the optimal parameters of chemical reaction networks more rigorously, it is necessary to go beyond the LQG framework, which is an important future direction.

3. Correspondence to infinite-horizon linear-quadratic-Gaussian problem

The target estimation problem and the LQG problem are discussed in Sec. IV and Appendix C, respectively. In this subsection, we show that the target estimation problem corresponds to the infinite-horizon LQG problem of OMSC. We define extended state $\tilde{x}_t$, extended observation $\tilde{y}_t$, extended control $\tilde{u}_t$, and extended standard Wiener process $\tilde{\omega}_t$ as follows:

$$\tilde{x}_t := \begin{pmatrix} x_t \\ z_t \end{pmatrix}, \quad \tilde{y}_t := \begin{pmatrix} y_t \\ z_t \end{pmatrix}, \quad \tilde{u}_t := \begin{pmatrix} \hat{x}_t \\ v_t \end{pmatrix}, \quad \tilde{\omega}_t := \begin{pmatrix} \omega_t \\ \xi_t \end{pmatrix}. \quad (\text{D38})$$

Extended state $\tilde{x}_t$ evolves by the following SDE:

$$d\tilde{x}_t = (\tilde{A}\tilde{x}_t + \tilde{B}\tilde{u}_t)dt + \tilde{\sigma}d\tilde{\omega}_t,$$

where

$$\tilde{A} := \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{B} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{\sigma} := \begin{pmatrix} \sqrt{D} & 0 \\ 0 & \sqrt{F} \end{pmatrix}. \quad (\text{D39})$$

Extended observation $\tilde{y}_t$ is generated from the following Gaussian distribution:

$$\tilde{p}(\tilde{y}_t|\tilde{x}_t) = \mathcal{N}(\tilde{y}_t|\tilde{H}\tilde{x}_t, \tilde{E}), \quad (\text{D40})$$

where

$$\tilde{H} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{E} := \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{D41})$$

The objective function is given by the following cost function:

$$J[\tilde{u}] := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (\tilde{x}_t^\top \tilde{Q}\tilde{x}_t - 2\tilde{x}_t^\top \tilde{S}\tilde{u}_t + \tilde{u}_t^\top \tilde{R}\tilde{u}_t) dt \right], \quad (\text{D42})$$

where

$$\tilde{Q} := \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{S} := \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{R} := \begin{pmatrix} Q & 0 \\ 0 & M \end{pmatrix}. \quad (\text{D43})$$

Therefore, the target estimation problem corresponds to the infinite-horizon LQG problem of OMSC.

4. Derivation of optimal solution

In this subsection, we derive the optimal solution of the target estimation problem, which is given by Eqs. (31) and (32). Since the target estimation problem corresponds to the infinite-horizon LQG problem of OMSC, the optimal solution of the target estimation problem is given by

$$\tilde{u}^*(t, \tilde{y}) = -\tilde{R}^{-1}((\tilde{\Psi}\tilde{B} - \tilde{S})^\top \tilde{\mu} + (\tilde{\Pi}\tilde{B} - \tilde{S})^\top \tilde{K}(\tilde{y} - \tilde{H}\tilde{\mu})), \quad (\text{D44})$$

where $\tilde{K} := \tilde{\Lambda}^{-1}\tilde{H}^\top(\tilde{E} + \tilde{H}\tilde{\Lambda}^{-1}\tilde{H}^\top)^{-1}$, $\tilde{\mu}$, $\tilde{\Psi}$, $\tilde{\Pi}$, and $\tilde{\Lambda}$ are the solutions of Eqs. (C34) and (C36)–(C38) (see Appendix C 4).

Since $\tilde{\mu} = 0$, the optimal solution of the target estimation problem is calculated as follows:

$$\tilde{u}^*(t, \tilde{y}) = -\tilde{R}^{-1}(\tilde{\Pi}\tilde{B} - \tilde{S})^\top \tilde{K}\tilde{y}, \quad (\text{D45})$$

where $\tilde{K} := \tilde{\Lambda}^{-1}\tilde{H}^\top(\tilde{E} + \tilde{H}\tilde{\Lambda}^{-1}\tilde{H}^\top)^{-1}$. $\tilde{\Pi}$ and $\tilde{\Lambda}$ are the solutions of Eqs. (C37) and (C38).

Since $\mathbb{E}_{\tilde{p}_t(\tilde{x}|\tilde{y})}[\tilde{x}] = \tilde{\mu} + \tilde{K}(\tilde{y} - \tilde{H}\tilde{\mu})$ [Eq. (C17)] and $\tilde{\mu} = 0$,

$$\tilde{K}\tilde{y} = \mathbb{E}_{\tilde{p}_t(\tilde{x}|\tilde{y})}[\tilde{x}] = \begin{pmatrix} \mathbb{E}_{p_t(x|y,z)}[x] \\ z \end{pmatrix} \quad (\text{D46})$$

holds. Thus, the optimal solution of the target estimation problem is calculated as follows:

$$\begin{aligned} \tilde{u}^*(t, \tilde{y}) &= -\begin{pmatrix} Q & 0 \\ 0 & M \end{pmatrix}^{-1} \begin{pmatrix} \Pi_{xx} & \Pi_{xz} \\ \Pi_{zx} & \Pi_{zz} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad - \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}^\top \begin{pmatrix} \mathbb{E}_{p_t(x|y,z)}[x] \\ z \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{E}_{p_t(x|y,z)}[x] \\ -M^{-1}\Pi_{zx}\mathbb{E}_{p_t(x|y,z)}[x] - M^{-1}\Pi_{zz}z \end{pmatrix}. \end{aligned} \quad (\text{D47})$$

Therefore, Eqs. (31) and (32) are obtained.

5. Phase transition with respect to other parameters

The discontinuous phase transition also arises with other parameters: the intensity of the state noise D , the weight of the state estimation error Q , and the weight of the memory control cost M (Fig. 15). Notably, the phase transitions for D [Figs. 15(a)–15(c)] and Q [Figs. 15(d)–15(f)] resemble that for observation noise E [Figs. 4(d)–4(f)], whereas the phase transition for M [Figs. 15(g)–15(i)] parallels that for memory noise F [Figs. 4(g)–4(i)].

When D is low, $\Pi_{zx} = \Pi_{zz} = 0$ is the unique and optimal solution, indicating that the estimation strategy without memory is optimal when the state dynamics is less uncertain [Figs. 15(a)–15(c), blue dots]. In contrast, as D increases, $\Pi_{zx} \neq 0$ and $\Pi_{zz} \neq 0$ discontinuously emerge [Figs. 15(a)–15(c), green and orange dots], while $\Pi_{zx} = \Pi_{zz} = 0$ remains optimal [Figs. 15(a)–15(c), blue dots]. With further increases in D , the optimal solution switches from $\Pi_{zx} = \Pi_{zz} = 0$ [Figs. 15(a)–15(c), blue dots] to $\Pi_{zx}, \Pi_{zz} \neq 0$ [Figs. 15(a)–15(c), orange dots], showing that memory becomes beneficial when the state dynamics is highly uncertain.

A similar transition is observed for Q [Figs. 15(d)–15(f)]: the memoryless strategy is optimal for small Q [Figs. 15(d)–15(f), blue dots], whereas the memory-based strategy becomes optimal for larger Q [Figs. 15(d)–15(f), orange dots]. This indicates that memory is effective when minimizing the state estimation error is prioritized, which is consistent with the landscape analysis in Sec. IV E [Figs. 5(a)–5(f)].

In contrast, the phase transition for M proceeds in the opposite direction [Figs. 15(g)–15(i)]. When M is small, the memory-based strategy is optimal [Figs. 15(g)–15(i), orange dots], but as M increases, the optimal strategy shifts to memoryless [Figs. 15(g)–15(i), blue dots]. This reversal highlights

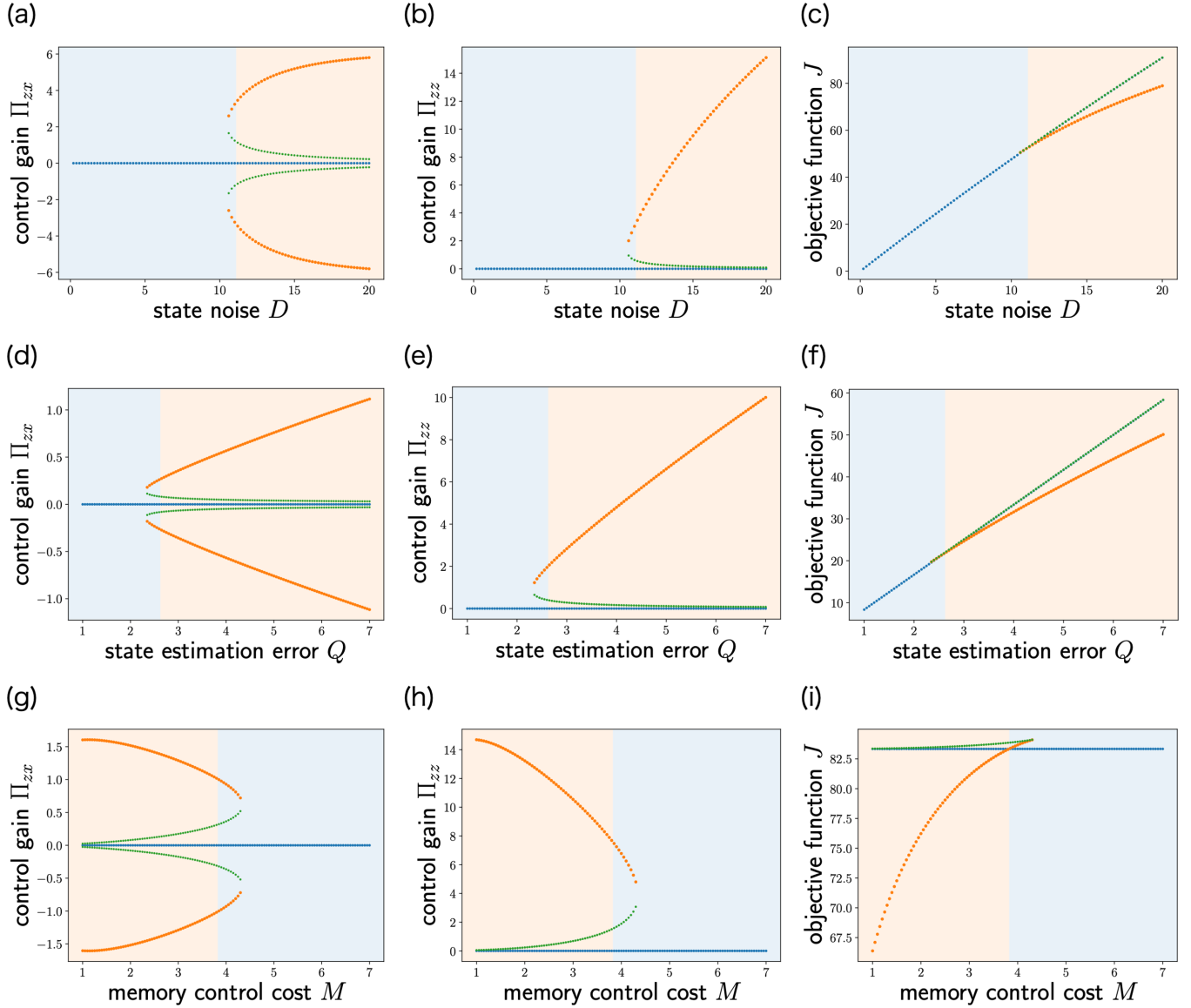


FIG. 15. Memory control gains Π_{zx} and Π_{zz} and associated objective function J with respect to state noise D [panels (a)–(c)], state estimation error Q [panels (d)–(f)], and memory control cost M [panels (g)–(i)] in the target estimation problem. The blue, green, and orange dots are the solutions of Pontryagin’s minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. The parameters are $E = 100$, $F = 1$, $Q = 10$, and $M = 1$ [panels (a)–(c)], $D = 100$, $E = 1$, $F = 1$, and $M = 1$ [panels (d)–(f)], and $D = 100$, $E = 10$, $F = 1$, and $Q = 10$ [panels (g)–(i)].

that memory becomes ineffective when the memory control cost dominates, again consistent with the analysis in Sec. IV E [Figs. 5(d)–5(i)].

APPENDIX E: DETAILS OF APPLICATION TO BIOLOGICAL INFORMATION PROCESSING AND DECISION-MAKING

In this section, we provide a detailed explanation of the minimal model of biological information processing and decision-making, the target tracking problem, presented in Sec. VI. This section is organized as follows: In Appendix E 1, we show that the target tracking problem corresponds to the infinite-horizon LQG problem of OMSC. In Appendix E 2,

we derive the optimal solution of the target tracking problem [Eqs. (49) and (50)]. In Appendix E 3, we show the discontinuous phase transition with respect to the other parameters. In Appendix E 4, we discuss the target tracking problem where the state dynamics is modified from a free diffusion to a trapped diffusion.

1. Correspondence to infinite-horizon linear-quadratic-Gaussian problem

The target tracking problem and the LQG problem are discussed in Sec. VI and Appendix C, respectively. In this subsection, we show that the target tracking problem corresponds to the infinite-horizon LQG problem of OMSC. We define

extended state $\tilde{x}_t$, extended observation $\tilde{y}_t$, extended control $\tilde{u}_t$, and extended standard Wiener process $\tilde{\omega}_t$ as follows:

$$\tilde{x}_t := \begin{pmatrix} x_t \\ z_t \end{pmatrix}, \quad \tilde{y}_t := \begin{pmatrix} y_t \\ z_t \end{pmatrix}, \quad \tilde{u}_t := \begin{pmatrix} u_t \\ v_t \end{pmatrix}, \quad \tilde{\omega}_t := \begin{pmatrix} \omega_t \\ \xi_t \end{pmatrix}. \quad (\text{E1})$$

Extended state $\tilde{x}_t$ evolves by the following SDE:

$$d\tilde{x}_t = (\tilde{A}\tilde{x}_t + \tilde{B}\tilde{u}_t)dt + \tilde{\sigma}d\tilde{\omega}_t,$$

where

$$\tilde{A} := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{B} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{\sigma} := \begin{pmatrix} \sqrt{D} & 0 \\ 0 & \sqrt{F} \end{pmatrix}. \quad (\text{E2})$$

Extended observation $\tilde{y}_t$ is generated from the following Gaussian distribution:

$$\tilde{p}(\tilde{y}_t | \tilde{x}_t) = \mathcal{N}(\tilde{y}_t | \tilde{H}\tilde{x}_t, \tilde{E}), \quad (\text{E3})$$

where

$$\tilde{H} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{E} := \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{E4})$$

The objective function is given by the following cost function:

$$J[\tilde{u}] := \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (\tilde{x}_t^\top \tilde{Q}\tilde{x}_t - 2\tilde{x}_t^\top \tilde{S}\tilde{u}_t + \tilde{u}_t^\top \tilde{R}\tilde{u}_t) dt \right], \quad (\text{E5})$$

where

$$\tilde{Q} := \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{S} := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{R} := \begin{pmatrix} R & 0 \\ 0 & M \end{pmatrix}. \quad (\text{E6})$$

Therefore, the target tracking problem corresponds to the infinite-horizon LQG problem of OMSC.

2. Derivation of optimal solution

In this subsection, we derive the optimal solution of the target tracking problem, which is given by Eqs. (49) and (50). Since the target tracking problem corresponds to the infinite-horizon LQG problem of OMSC, the optimal solution of the target tracking problem is given by

$$\tilde{u}^*(t, \tilde{y}) = -\tilde{R}^{-1}((\tilde{\Psi}\tilde{B} - \tilde{S})^\top \tilde{\mu} + (\tilde{\Pi}\tilde{B} - \tilde{S})^\top \tilde{K}(\tilde{y} - \tilde{H}\tilde{\mu})), \quad (\text{E7})$$

where $\tilde{K} := \tilde{\Lambda}^{-1}\tilde{H}^\top(\tilde{E} + \tilde{H}\tilde{\Lambda}^{-1}\tilde{H}^\top)^{-1}$, $\tilde{\mu}$, $\tilde{\Psi}$, $\tilde{\Pi}$, and $\tilde{\Lambda}$ are the solutions of Eqs. (C34) and (C36)–(C38) (see Appendix C 4).

Since $\tilde{\mu} = 0$, the optimal solution of the target estimation problem is calculated as follows:

$$\tilde{u}^*(t, \tilde{y}) = -\tilde{R}^{-1}(\tilde{\Pi}\tilde{B} - \tilde{S})^\top \tilde{K}\tilde{y}, \quad (\text{E8})$$

where $\tilde{K} := \tilde{\Lambda}^{-1}\tilde{H}^\top(\tilde{E} + \tilde{H}\tilde{\Lambda}^{-1}\tilde{H}^\top)^{-1}$. $\tilde{\Pi}$ and $\tilde{\Lambda}$ are the solutions of Eqs. (C37) and (C38).

Since $\mathbb{E}_{\tilde{p}_t(\tilde{x}|\tilde{y})}[\tilde{x}] = \tilde{\mu} + \tilde{K}(\tilde{y} - \tilde{H}\tilde{\mu})$ [Eq. (C17)] and $\tilde{\mu} = 0$,

$$\tilde{K}\tilde{y} = \mathbb{E}_{\tilde{p}_t(\tilde{x}|\tilde{y})}[\tilde{x}] = \begin{pmatrix} \mathbb{E}_{p_t(x|y,z)}[x] \\ z \end{pmatrix} \quad (\text{E9})$$

holds. Thus, the optimal solution of the target tracking problem is calculated as follows:

$$\begin{aligned} \tilde{u}^*(t, \tilde{y}) &= -\begin{pmatrix} R & 0 \\ 0 & M \end{pmatrix}^{-1} \left(\begin{pmatrix} \Pi_{xx} & \Pi_{xz} \\ \Pi_{zx} & \Pi_{zz} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right. \\ &\quad \left. - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}^\top \begin{pmatrix} \mathbb{E}_{p_t(x|y,z)}[x] \\ z \end{pmatrix} \right) \\ &= \begin{pmatrix} -R^{-1}\Pi_{xx}\mathbb{E}_{p_t(x|y,z)}[x] - R^{-1}\Pi_{xz}z \\ -M^{-1}\Pi_{zx}\mathbb{E}_{p_t(x|y,z)}[x] - M^{-1}\Pi_{zz}z \end{pmatrix}. \end{aligned} \quad (\text{E10})$$

Therefore, Eqs. (49) and (50) are obtained.

3. Phase transition with respect to other parameters

The discontinuous phase transition also occurs for the other parameters: the intensity of the state noise D , the weight of the state tracking error Q , the weight of the state control cost R , and the weight of the memory control cost M (Fig. 16). As D , Q , or R increases, the optimal strategy switches discontinuously from the memoryless strategy [$\Pi_{zx} = \Pi_{zz} = 0$; Figs. 16(a)–16(l), blue dots] to the memory-based strategy [$\Pi_{zx}, \Pi_{zz} \neq 0$; Figs. 16(a)–16(l), orange dots]. This behavior is similar to the phase transition for observation noise E [Figs. 9(d)–9(g)]. In contrast, as M increases, the optimal strategy switches in the opposite direction—from the memory-based strategy [$\Pi_{zx}, \Pi_{zz} \neq 0$; Figs. 16(m)–16(p), orange dots] to the memoryless strategy [$\Pi_{zx} = \Pi_{zz} = 0$; Figs. 16(m)–16(p), blue dots]—which resembles the phase transition for memory noise F [Figs. 9(h)–9(k)].

4. Trapped diffusion case

While the phase transition with respect to observation noise E is nonmonotonic in the target estimation problem (Fig. 6), it is monotonic in the target tracking problem [Figs. 10 and 17(a)–17(e)]. In Sec. VI C, we hypothesized that this difference arises from the distinct state dynamics. In the target estimation problem, the state dynamics evolves according to a trapped diffusion process [Eq. (26)], which ensures that the estimation error remains bounded due to the finite state variance. Thus, when the observation noise is extremely high, the agent can afford to stop controlling memory without incurring unbounded error. In contrast, in the target tracking problem, the state dynamics follows a free diffusion process [Eq. (44)], where the absence of state control leads to unbounded tracking error. As a result, the agent must continue controlling memory, even under extremely high observation noise.

To test this hypothesis, we consider a modified version of the target tracking problem in which the state dynamics is changed from a free diffusion [Eq. (44)] to a trapped diffusion:

$$dx_t = (-x_t + u_t)dt + \sqrt{D}d\omega_t. \quad (\text{E11})$$

All the other aspects of the problem formulation are identical to those in Sec. VI A.

The results are shown in Figs. 17(f)–17(j). When observation noise E is low, $\Pi_{zx} = \Pi_{zz} = 0$ is the unique and optimal solution and the memoryless strategy is optimal [Figs. 17(g)–17(j), blue dots]. As E increases, the optimal solution switches

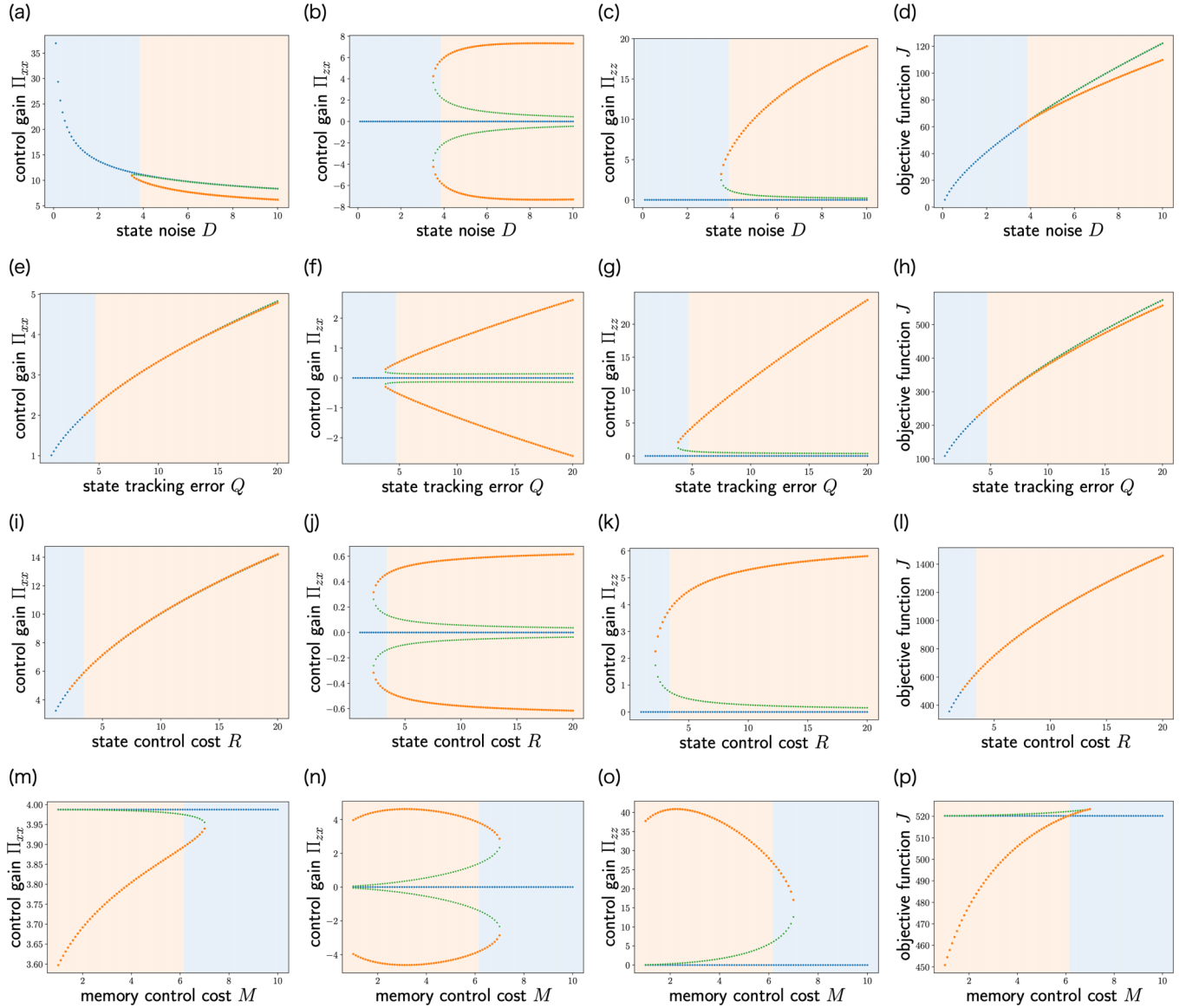


FIG. 16. Control gains Π_{xx} , Π_{zx} , and Π_{zz} and associated objective function J with respect to state noise D [panels (a)–(d)], state tracking error Q [panels (e)–(h)], state control cost R [panels (i)–(l)], and memory control cost M [panels (m)–(p)] in the target tracking problem. Control gain Π_{xz} is not visualized because $\Pi_{xz} = \Pi_{zx}$ holds. The blue, green, and orange dots are the solutions of Pontryagin’s minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. The parameters are $E = 100, F = 1, Q = 10, R = 1$, and $M = 1$ [panels (a)–(d)], $D = 100, E = 1, F = 1, R = 1$, and $M = 1$ [panels (e)–(h)], $D = 100, E = 5, F = 1, Q = 10$, and $M = 1$ [panels (i)–(l)], and $D = 100, E = 50, F = 1, Q = 10$, and $R = 1$ [panels (m)–(p)].

from $\Pi_{zx} = \Pi_{zz} = 0$ to $\Pi_{zx}, \Pi_{zz} \neq 0$ and the memory-based strategy is optimal [Figs. 17(g)–17(j), orange dots]. As E further increases, the solutions realizing $\Pi_{zx}, \Pi_{zz} \neq 0$ disappear and the memoryless strategy becomes optimal again [Figs. 17(g)–17(j), blue dots]. This result indicates that the nonmonotonic phase transition with respect to E also occurs in the target tracking problem if the state dynamics is modified from a free diffusion [Eq. (44)] to a trapped diffusion [Eq. (E11)]. Therefore, the nonmonotonicity of the phase transitions depends on the state dynamics and may not be universal.

APPENDIX F: MEMORY-LIMITED PARTIALLY OBSERVABLE STOCHASTIC CONTROL

In this section, we briefly review our theoretical framework proposed previously, ML-POSC [113–115], and compare it with OSC and OMSC introduced in this work. The details of OSC and OMSC are explained in Appendixes A and B, respectively. This section is organized as follows: In Appendix F1, we formulate ML-POSC, comparing it with OMSC. In Appendix F2, we show that ML-POSC can be converted into OSC on the extended space, similar to OMSC.

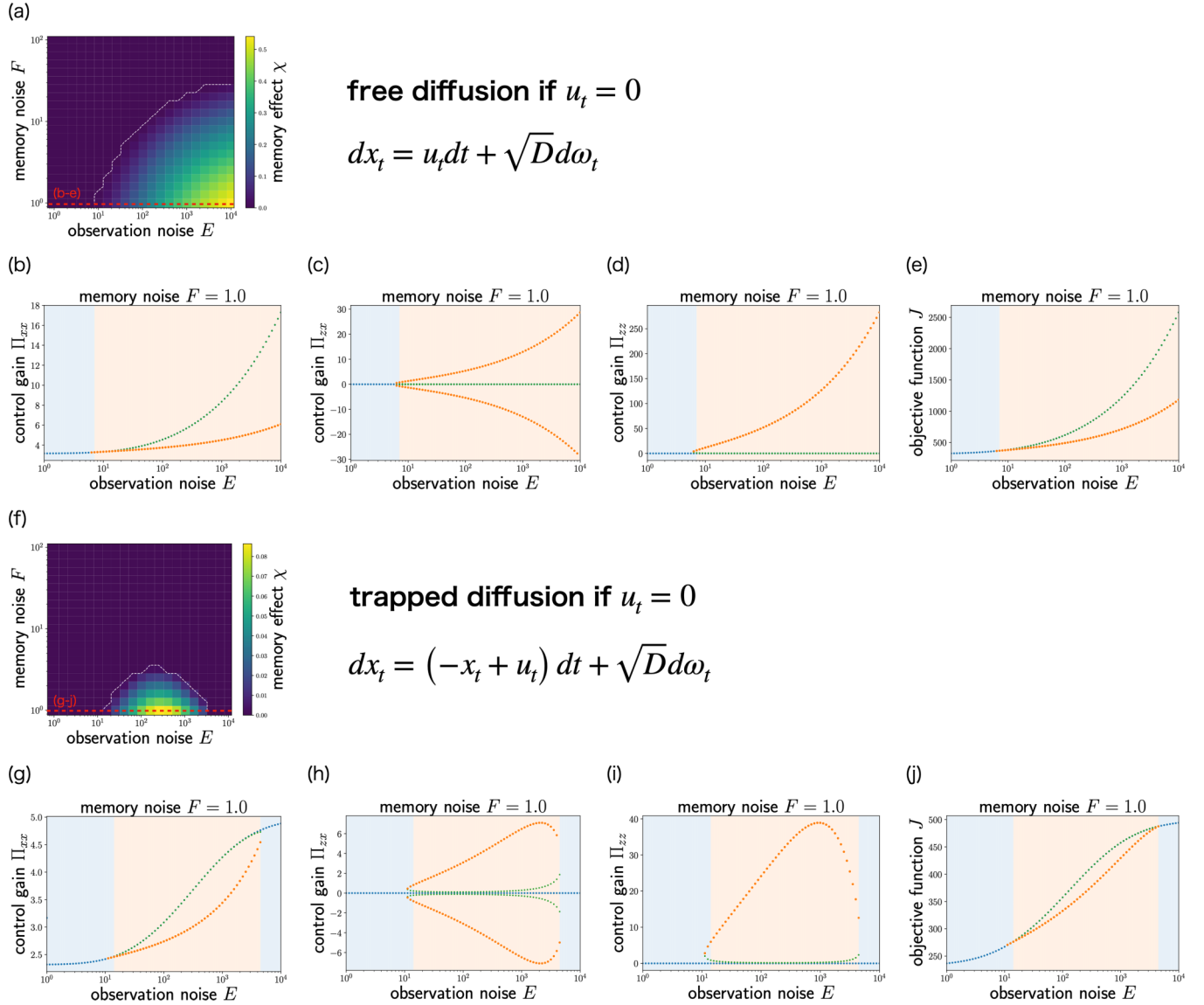


FIG. 17. The difference in the phase transitions between the free diffusion [panels (a)–(e)] and the trapped diffusion [panels (f)–(j)] in the target tracking problem. (a), (f) Memory effect χ with respect to observation noise E and memory noise F . (b)–(e), (g)–(j) Control gains Π_{xx} , Π_{xz} , and Π_{zz} and associated objective function J with respect to observation noise E . Control gain Π_{xz} is not visualized because $\Pi_{xz} = \Pi_{zx}$ holds. The blue, green, and orange dots are the solutions of Pontryagin’s minimum principle, corresponding to zero, intermediate, and high memory controls, respectively. The blue dots are optimal in the blue region, while the orange dots are optimal in the orange region. The rest of the parameters are $D = 100$, $Q = 10$, $R = 1$, and $M = 1$.

1. Problem formulation

In this subsection, we formulate ML-POSC [113–115], comparing it with OMSC. The dynamics of state $x_t \in \mathbb{R}^{d_x}$ in ML-POSC is the same with that in OMSC [Eq. (B1)], which is given by the following SDE:

$$dx_t = b(t, x_t, u_t)dt + \sigma(t, x_t, u_t)d\omega_t, \quad (\text{F1})$$

where initial state x_0 follows $p_0(x_0)$, $u_t \in \mathbb{R}^{d_u}$ is the control for the state, and $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process. In contrast, the generation rule of observation $y_t \in \mathbb{R}^{d_y}$ in ML-POSC is different from that in OMSC. While observation y_t is generated from probability density function $p_t(y_t|x_t)$ as $y_t \sim p_t(y_t|x_t)$ in OMSC [Eq. (B2)], it is generated from the

following SDE in ML-POSC:

$$dy_t = h(t, x_t)dt + \gamma(t)dv_t, \quad (\text{F2})$$

where $v_t \in \mathbb{R}^{d_v}$ is a standard Wiener process. From this difference, the dynamics of memory $z_t \in \mathbb{R}^{d_z}$ in ML-POSC is formulated by

$$dz_t = c(t, z_t, v_t)dt + \kappa(t, z_t, v_t)dy_t + \eta(t, z_t, v_t)d\xi_t, \quad (\text{F3})$$

where initial memory z_0 follows $p_0(z_0)$, $v_t \in \mathbb{R}^{d_v}$ is the control for the memory, and $\xi_t \in \mathbb{R}^{d_\xi}$ is a standard Wiener process. Compared with the memory dynamics in OMSC [Eq. (B3)], $\kappa(t, z_t, v_t)dy_t$ is an additional term in ML-POSC. While OMSC determines state control u_t and memory control v_t based on observation y_t and memory z_t [Eqs. (B4) and (B5)],

ML-POSC determines them based solely on memory z_t as follows:

$$u_t = u(t, z_t), \quad (\text{F4})$$

$$v_t = v(t, z_t). \quad (\text{F5})$$

The rest of the formulation in ML-POSC is the same as that in OMSC. The objective function is given by the following expected cumulative cost function:

$$J[u, v] := \mathbb{E} \left[\int_0^T f(t, x_t, z_t, u_t, v_t) dt + g(x_T, z_T) \right], \quad (\text{F6})$$

where f and g are the running and terminal cost functions, respectively. ML-POSC is a problem to find optimal state control function u^* and optimal memory control function v^* that minimize objective function $J[u, v]$ as follows:

$$u^*, v^* := \arg \min_{u, v} J[u, v]. \quad (\text{F7})$$

While OMSC encodes observation information into memory through memory control $v_t = v(t, y_t, z_t)$, ML-POSC encodes it through $\kappa(t, z_t, v_t) dy_t$. Since the optimization problem of $\kappa(t, z_t, v_t) dy_t$ lies outside the scope of the LQG problem, ML-POSC cannot explicitly and efficiently investigate how the agent optimally encodes observation information into memory.

For instance, the minimal model of biological information processing in Sec. IV is modified in ML-POSC as follows:

$$dx_t = -x_t dt + \sqrt{D} d\omega_t, \quad (\text{F8})$$

$$dy_t = x_t dt + \sqrt{E} dv_t, \quad (\text{F9})$$

$$dz_t = v_t dt + \kappa_t dy_t + \sqrt{F} d\xi_t, \quad (\text{F10})$$

$$v_t = v(t, z_t), \quad (\text{F11})$$

$$\kappa_t = \kappa(t, z_t). \quad (\text{F12})$$

The state dynamics remains the same. The generation rule of observation y_t is modified from the probability density function [Eq. (27)] to the stochastic differential equation [Eq. (F9)]. The memory dynamics has $\kappa_t dy_t$ as an additional term, where κ_t is an additional memory control. Memory control κ_t determines how the agent encodes observation information into memory. The memory controls, v_t and κ_t , depend on memory z_t , but do not depend on observation y_t [Eqs. (F11) and (F12)]. Substituting Eq. (F9) into Eq. (F10) yields the following equation:

$$dz_t = v_t dt + \kappa_t x_t dt + \kappa_t \sqrt{E} dv_t + \sqrt{F} d\xi_t. \quad (\text{F13})$$

Since $\kappa_t x_t dt$ and $\kappa_t \sqrt{E} dv_t$ lie outside the scope of the LQG problem, this problem cannot be investigated efficiently. OMSC addresses this issue by encoding observation information into memory through memory control $v_t = v(t, y_t, z_t)$, not $\kappa(t, z_t, v_t) dy_t$.

2. Conversion to observation-based stochastic control on extended space

In this subsection, we show that ML-POSC can be converted into OSC on the extended space, which is similar to OMSC. We define extended state $\tilde{x}_t$, extended observation $\tilde{y}_t$, and extended control $\tilde{u}_t$ as follows:

$$\tilde{x}_t := \begin{pmatrix} x_t \\ z_t \end{pmatrix}, \quad \tilde{y}_t := z'_t, \quad \tilde{u}_t := \begin{pmatrix} u_t \\ v_t \end{pmatrix}. \quad (\text{F14})$$

We denote the memory in extended state $\tilde{x}_t$ by z_t and that in extended observation $\tilde{y}_t$ by z'_t to distinguish between them. In practice, $z_t = z'_t$ holds. Extended state $\tilde{x}_t$ evolves by the following SDE:

$$d\tilde{x}_t = \tilde{b}(t, \tilde{x}_t, \tilde{u}_t) dt + \tilde{\sigma}(t, \tilde{x}_t, \tilde{u}_t) d\tilde{\omega}_t, \quad (\text{F15})$$

where

$$\tilde{b}(t, \tilde{x}_t, \tilde{u}_t) := \begin{pmatrix} b(t, x_t, u_t) \\ c(t, z_t, v_t) + \kappa(t, z_t, v_t) h(t, x_t) \end{pmatrix}, \quad (\text{F16})$$

$$\tilde{\sigma}(t, \tilde{x}_t, \tilde{u}_t) := \begin{pmatrix} \sigma(t, x_t, u_t) & O & O \\ O & \kappa(t, z_t, v_t) \gamma(t) & \eta(t, z_t, v_t) \end{pmatrix}, \quad (\text{F17})$$

$$\tilde{\omega}_t := \begin{pmatrix} \omega_t \\ v_t \\ \xi_t \end{pmatrix}. \quad (\text{F18})$$

Extended observation $\tilde{y}_t$ is generated from the following probability:

$$\tilde{p}_t(\tilde{y}_t | \tilde{x}_t) := \delta(z'_t - z_t). \quad (\text{F19})$$

Extended control $\tilde{u}_t$ is determined based on extended observation $\tilde{y}_t$ as follows:

$$\tilde{u}_t = \tilde{u}(t, \tilde{y}_t) := \begin{pmatrix} u(t, z_t) \\ v(t, z_t) \end{pmatrix}. \quad (\text{F20})$$

The objective function is given by the following expected cumulative cost function:

$$J[\tilde{u}] := \mathbb{E} \left[\int_0^T \tilde{f}(t, \tilde{x}_t, \tilde{u}_t) dt + \tilde{g}(\tilde{x}_T) \right], \quad (\text{F21})$$

where

$$\tilde{f}(t, \tilde{x}_t, \tilde{u}_t) := f(t, x_t, z_t, u_t, v_t), \quad \tilde{g}(\tilde{x}_T) := g(x_T, z_T). \quad (\text{F22})$$

Therefore, ML-POSC can be converted into OSC on the extended space, and its optimal solution can be obtained through OSC, similar to OMSC. This result suggests that OSC provides a convenient theoretical framework that allows us to treat OMSC and ML-POSC in a unified manner.

APPENDIX G: CONNECTION TO KALMAN FILTER

While Bayesian filtering theory encodes information about past observation history into an infinite-dimensional posterior probability density function, optimal memory control theory proposed in this study compresses it into a finite-dimensional memory. This fundamental difference makes a direct connection between the two theories nontrivial. However, in the

LQG problem, the posterior probability density function becomes Gaussian, allowing the infinite-dimensional Bayesian filter to be reduced to the finite-dimensional Kalman filter [51–54,106,116–118]. Thus, the Kalman filter serves as a bridge between Bayesian filtering theory and optimal memory control theory. In this section, we demonstrate that optimal memory control theory aligns with the Kalman filter in the LQG problem without computational resource limitations.

This section is organized as follows: In Appendix G 1, we briefly review the Kalman filter. In Appendix G 2, we show that optimal memory control theory is consistent with the Kalman filter.

1. Brief review of Kalman filter

In this subsection, we briefly review the Kalman filter [51–54,106,116–118]. The Kalman filter estimates the state of an environment from a noisy observation history in the LQG problem. The dynamics of environmental state $x_t \in \mathbb{R}^{d_x}$ is given by the following linear and Gaussian dynamics:

$$dx_t = Ax_t dt + \sigma d\omega_t, \quad (G1)$$

where initial state x_0 follows Gaussian distribution $\mathcal{N}(x_0|\mu_{0,x}, \Sigma_{0,xx})$ and $\omega_t \in \mathbb{R}^{d_\omega}$ is a standard Wiener process. The noisy observation of environmental state x_t is given by $dy'_t \in \mathbb{R}^{d_{y'}}$, which is generated from the following linear and Gaussian dynamics:

$$dy'_t = H'x_t dt + \gamma' dv_t, \quad (G2)$$

where $v_t \in \mathbb{R}^{d_v}$ is a standard Wiener process. We note that observation dy'_t is given by a stochastic differential equation, not a probability density function such as Eq. (2). In this problem, the posterior probability density function of the environmental state given the noisy observation history is represented by the following Gaussian distribution:

$$p(x_t|dy'_t, dy'_{t-dt}, \dots, dy'_0) = \mathcal{N}(x_t|\check{\mu}_t, \check{\Sigma}_t), \quad (G3)$$

where mean vector $\check{\mu}_t$ and covariance matrix $\check{\Sigma}_t$ are the solutions of the following equation:

$$d\check{\mu}_t = A\check{\mu}_t dt + \check{\Sigma}_t H'^\top E'^{-1}(dy'_t - H'\check{\mu}_t dt), \quad (G4)$$

$$\frac{d\check{\Sigma}_t}{dt} = D + A\check{\Sigma}_t + \check{\Sigma}_t A^\top - \check{\Sigma}_t H'^\top E'^{-1} H' \check{\Sigma}_t, \quad (G5)$$

where $D := \sigma\sigma^\top$, $E' := \gamma'\gamma'^\top$, $\check{\mu}_0 = \mu_{0,x}$, and $\check{\Sigma}_0 = \Sigma_{0,xx}$. Equations (G4) and (G5) are called the Kalman filter, in which the estimator of x_t is given by $\check{\mu}_t$.

2. Connection to optimal memory control theory

In this subsection, we demonstrate that optimal memory control theory is consistent with the Kalman filter.

First, optimal memory control theory considers that observation $y_t \in \mathbb{R}^{d_y}$ is given by a probability density function, not a stochastic differential equation, as follows:

$$y_t \sim \mathcal{N}(y_t|Hx_t, E). \quad (G6)$$

We note that our previous work [113–115] considers the observation given by the stochastic differential equation [see Eq. (F2)] and shows the connection to the Kalman filter more

straightforwardly [113]. On the other hand, the present work mainly considers the observation given by the probability density function such as Eq. (G6) to efficiently investigate how the agent optimally encodes observation information into memory (see Appendix F). To demonstrate the consistency between the Kalman filter and the present work, we need to clarify the connection between the observation given by Eq. (G2) and that given by Eq. (G6). The observation given by Eq. (G2) is rewritten as follows:

$$dy'_t \sim \mathcal{N}(dy'_t|H'x_t dt, E' dt). \quad (G7)$$

To match the order of the expected value of dy'_t with that of y_t , we divide dy'_t by dt as follows:

$$\frac{dy'_t}{dt} \sim \mathcal{N}\left(\frac{dy'_t}{dt}\middle|H'x_t, \frac{E'}{dt}\right). \quad (G8)$$

Hence, the observation given by Eq. (G2) corresponds to that given by Eq. (G6) as follows:

$$y_t = \frac{dy'_t}{dt}, \quad H = H', \quad E = \frac{E'}{dt}. \quad (G9)$$

Since the stochastic differential equation considers $dt \rightarrow 0$, the noise of the observation corresponding to the stochastic differential equation is infinitely large.

Next, optimal memory control theory explicitly formulates finite-dimensional memory $z_t \in \mathbb{R}^{d_z}$, whose dynamics is given by the following stochastic differential equation:

$$dz_t = v_t dt + \eta d\xi_t, \quad (G10)$$

where initial memory z_0 follows Gaussian distribution $\mathcal{N}(z_0|\mu_{0,z}, \Sigma_{0,zz})$, $v_t \in \mathbb{R}^{d_z}$ is the memory control, and $\xi_t \in \mathbb{R}^\xi$ is a standard Wiener process. The diffusion constant of the memory dynamics is given by $F := \eta\eta^\top$. Since the Kalman filter does not account for the resource limitations of the agent, the dimension of the memory must be equal to that of the state, and the dynamics of the memory must be deterministic. These constraints are expressed by the following equations:

$$d_z = d_x, \quad F := \eta\eta^\top = O. \quad (G11)$$

To show the connection to the Kalman filter, we also assume

$$\mu_{0,z} = \mu_{0,x}, \quad \Sigma_{0,zz} = O. \quad (G12)$$

In optimal memory control theory, state estimator $\hat{x}_t \in \mathbb{R}^{d_x}$ and memory control $v_t \in \mathbb{R}^{d_z}$ are determined based on observation $y_t (= dy'_t/dt)$ and memory z_t as follows:

$$\hat{x}_t = \hat{x}(t, y_t, z_t), \quad v_t = v(t, y_t, z_t). \quad (G13)$$

Optimal state estimator function $\hat{x}^*$ and memory control function v^* are defined by

$$\hat{x}^*, v^* := \arg \min_{\hat{x}, v} J[\hat{x}, v], \quad (G14)$$

where objective function $J[\hat{x}, v]$ is given by the summation of the state estimation error and the memory control cost:

$$J[\hat{x}, v] := \mathbb{E} \left[\int_0^T ((x_t - \hat{x}_t)^\top Q (x_t - \hat{x}_t) + v_t^\top M v_t) dt \right], \quad (G15)$$

where $Q \succeq O$ and $M \succeq O$ are weighting parameters. The Kalman filter does not account for the memory control cost, which corresponds to

$$M = O. \quad (\text{G16})$$

From optimal memory control theory, the optimal solution is given by the minimization of the conditional expected Hamiltonian (see Sec. III B):

$$\hat{x}^*(t, y, z), v^*(t, y, z) = \arg \min_{\hat{x}, v} \mathbb{E}_{p_t(x|y,z)}[\mathcal{H}(t, x, z, \hat{x}, v, w)], \quad (\text{G17})$$

where Hamiltonian $\mathcal{H}$ is given by

$$\mathcal{H}(t, x, z, \hat{x}, v, w) = (x - \hat{x})^\top Q(x - \hat{x}) + \left(\frac{\partial w(t, x, z)}{\partial x} \right)^\top A x + \left(\frac{\partial w(t, x, z)}{\partial z} \right)^\top v + \frac{1}{2} \text{tr} \left\{ \frac{\partial}{\partial x} \left(\frac{\partial w(t, x, z)}{\partial x} \right)^\top D \right\}. \quad (\text{G18})$$

From Eq. (G17), the optimal solution is given by

$$\hat{x}^*(t, y, z) = \mathbb{E}_{p_t(x|y,z)}[x_t], \quad (\text{G19})$$

$$v_i^*(t, y, z) = \begin{cases} +\infty, & \mathbb{E}_{p_t(x|y,z)} \left[\frac{\partial w(t, x, z)}{\partial z_i} \right] < 0, \\ \text{arbitrary}, & \mathbb{E}_{p_t(x|y,z)} \left[\frac{\partial w(t, x, z)}{\partial z_i} \right] = 0, \\ -\infty, & \mathbb{E}_{p_t(x|y,z)} \left[\frac{\partial w(t, x, z)}{\partial z_i} \right] > 0, \end{cases} \quad (\text{G20})$$

where $w(t, x, z)$ is the value function, which is the solution of the following HJB equation:

$$\begin{aligned} -\frac{\partial w(t, x, z)}{\partial t} &= \mathbb{E}_{p_t(y|x)}[\mathcal{H}(t, x, z, \hat{x}^*, v^*, w)] \\ &= \mathbb{E}_{p_t(y|x)}[(x - \hat{x}^*(t, y, z))^\top Q(x - \hat{x}^*(t, y, z))] + \left(\frac{\partial w(t, x, z)}{\partial x} \right)^\top A x \\ &\quad + \left(\frac{\partial w(t, x, z)}{\partial z} \right)^\top \mathbb{E}_{p_t(y|x)}[v^*(t, y, z)] + \frac{1}{2} \text{tr} \left\{ \frac{\partial}{\partial x} \left(\frac{\partial w(t, x, z)}{\partial x} \right)^\top D \right\}, \end{aligned} \quad (\text{G21})$$

where $w(T, x, z) = 0$.

When $\mathbb{E}_{p_t(x|y,z)}[\partial w(t, x, z)/\partial z_i] \neq 0$, $v_i^*(t, y, z) \rightarrow \pm\infty$. The absolute value of the memory control could diverge to infinity in the limit where the resource limitations are absent such as $M \rightarrow O$. This result is consistent with the minimal model of biological information processing in Sec. IV. In this model, the optimal memory control function is given by $v^*(t, y, z) = -M^{-1}(\Pi_{zx}\mathbb{E}_{p_t(x|y,z)}[x] + \Pi_{zz}z)$ [Eq. (32)] and it could diverge to infinity when $\Pi_{zx} \neq 0$, $\Pi_{zz} \neq 0$, and $M \rightarrow 0$. In this solution, the agent actively controls its memory, which could correspond to the orange solutions in Figs. 4 and 15. When $w(t, x, z)/\partial z = 0$, $w(t, x, z)$ does not depend on z and can be reduced to $w(t, x)$. In this case, HJB equation is calculated as follows:

$$-\frac{\partial w(t, x)}{\partial t} = \mathbb{E}_{p_t(y|x)}[(x - \hat{x}^*(t, y, z))^\top Q(x - \hat{x}^*(t, y, z))] + \left(\frac{\partial w(t, x)}{\partial x} \right)^\top A x + \frac{1}{2} \text{tr} \left\{ \frac{\partial}{\partial x} \left(\frac{\partial w(t, x)}{\partial x} \right)^\top D \right\}, \quad (\text{G22})$$

where $w(T, x) = 0$. In order to satisfy this equation, $\hat{x}^*$ must not depend on z , which corresponds to the estimation strategy without memory (Figs. 4 and 15, blue dots). However, these solutions obviously do not correspond to the Kalman filter. Therefore, the solution corresponding to the Kalman filter should satisfy the following condition:

$$\frac{\partial w(t, x, z)}{\partial z} \neq 0, \quad \mathbb{E}_{p_t(x|y,z)} \left[\frac{\partial w(t, x, z)}{\partial z} \right] = 0. \quad (\text{G23})$$

If the optimal memory control function is given by

$$v^*(t, y, z) = \check{\Sigma}_t H'^\top E'^{-1} y + (A - \check{\Sigma}_t H'^\top E'^{-1} H') z, \quad (\text{G24})$$

the optimal memory dynamics is given by

$$dz_t = v_t^* dt = (\check{\Sigma}_t H'^\top E'^{-1} y_t + (A - \check{\Sigma}_t H'^\top E'^{-1} H') z_t) dt = A z_t dt + \check{\Sigma}_t H'^\top E'^{-1} (dy_t' - H' z_t dt), \quad (\text{G25})$$

where $z_0 = \mu_{0,x}$. Comparing Eq. (G4) with Eq. (G25), $\check{\mu}_t = z_t$ holds. Therefore, Eq. (G24) is the solution corresponding to the Kalman filter. In the following, we show that Eq. (G24) satisfies Eq. (G23).

First, the stochastic differential equation of state x_t and memory z_t is given by

$$\begin{aligned} d\begin{pmatrix} x_t \\ z_t \end{pmatrix} &= \begin{pmatrix} Ax_t dt + \sigma d\omega_t \\ Az_t dt + \tilde{\Sigma}_t H'^\top E'^{-1} (dy'_t - H' z_t dt) \end{pmatrix} \\ &= \begin{pmatrix} Ax_t dt + \sigma d\omega_t \\ Az_t dt + \tilde{\Sigma}_t H'^\top E'^{-1} (H' x_t dt + \gamma' dv_t - H' z_t dt) \end{pmatrix} \\ &= \begin{pmatrix} A & O \\ \tilde{\Sigma}_t H'^\top E'^{-1} H' & A - \tilde{\Sigma}_t H'^\top E'^{-1} H' \end{pmatrix} \begin{pmatrix} x_t \\ z_t \end{pmatrix} dt + \begin{pmatrix} \sigma & O \\ O & \tilde{\Sigma}_t H'^\top E'^{-1} \gamma' \end{pmatrix} \begin{pmatrix} d\omega_t \\ dv_t \end{pmatrix}, \end{aligned} \quad (\text{G26})$$

where the initial condition is given by

$$\begin{pmatrix} x_0 \\ z_0 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} x_0 \\ z_0 \end{pmatrix} \middle| \begin{pmatrix} \mu_{0,x} \\ \mu_{0,z} \end{pmatrix}, \begin{pmatrix} \Sigma_{0,xx} & O \\ O & O \end{pmatrix}\right). \quad (\text{G27})$$

Since this dynamics is linear and Gaussian, the probability density function of state x and memory z at time t is given by the Gaussian distribution:

$$p(t, x, z) = \mathcal{N}(\tilde{x} | \tilde{\mu}_t, \tilde{\Sigma}_t), \quad (\text{G28})$$

where $\tilde{x}$, $\tilde{\mu}_t$, and $\tilde{\Sigma}_t$ are defined by

$$\tilde{x} = \begin{pmatrix} x \\ z \end{pmatrix}, \quad \tilde{\mu}_t = \begin{pmatrix} \mu_{t,x} \\ \mu_{t,z} \end{pmatrix}, \quad \tilde{\Sigma}_t = \begin{pmatrix} \Sigma_{t,xx} & \Sigma_{t,xz} \\ \Sigma_{t,zx} & \Sigma_{t,zz} \end{pmatrix}. \quad (\text{G29})$$

From Eq. (G26), $\tilde{\mu}_t$ and $\tilde{\Sigma}_t$ evolve by the following ordinary differential equations:

$$\frac{d\tilde{\mu}_t}{dt} = \tilde{A}\tilde{\mu}_t, \quad (\text{G30})$$

$$\frac{d\tilde{\Sigma}_t}{dt} = \tilde{D} + \tilde{A}\tilde{\Sigma}_t + \tilde{\Sigma}_t\tilde{A}^\top, \quad (\text{G31})$$

where

$$\tilde{\mu}_0 = \begin{pmatrix} \mu_{0,x} \\ \mu_{0,z} \end{pmatrix}, \quad (\text{G32})$$

$$\tilde{\Sigma}_0 = \begin{pmatrix} \Sigma_{0,xx} & O \\ O & O \end{pmatrix}, \quad (\text{G33})$$

$$\tilde{A} := \begin{pmatrix} A & O \\ \tilde{\Sigma}_t H'^\top E'^{-1} H' & A - \tilde{\Sigma}_t H'^\top E'^{-1} H' \end{pmatrix}, \quad \tilde{D} := \begin{pmatrix} D & O \\ O & \tilde{\Sigma}_t H'^\top E'^{-1} H' \tilde{\Sigma}_t \end{pmatrix}. \quad (\text{G34})$$

Equations (G30) and (G31) are further calculated as follows:

$$\frac{d\mu_{t,x}}{dt} = A\mu_{t,x}, \quad (\text{G35})$$

$$\begin{aligned} \frac{d\mu_{t,z}}{dt} &= \tilde{\Sigma}_t H'^\top E'^{-1} H' \mu_{t,x} + (A - \tilde{\Sigma}_t H'^\top E'^{-1} H') \mu_{t,z} \\ &= A\mu_{t,z} + \tilde{\Sigma}_t H'^\top E'^{-1} H' (\mu_{t,x} - \mu_{t,z}), \end{aligned} \quad (\text{G36})$$

$$\frac{d\Sigma_{t,xx}}{dt} = D + A\Sigma_{t,xx} + \Sigma_{t,xx}A^\top, \quad (\text{G37})$$

$$\begin{aligned} \frac{d\Sigma_{t,xz}}{dt} &= A\Sigma_{t,xz} + \Sigma_{t,xx}H'^\top E'^{-1} H' \tilde{\Sigma}_t + \Sigma_{t,xz}(A - \tilde{\Sigma}_t H'^\top E'^{-1} H')^\top \\ &= A\Sigma_{t,xz} + \Sigma_{t,xz}A^\top + (\Sigma_{t,xx} - \Sigma_{t,xz})H'^\top E'^{-1} H' \tilde{\Sigma}_t, \end{aligned} \quad (\text{G38})$$

$$\begin{aligned} \frac{d\Sigma_{t,zz}}{dt} &= \tilde{\Sigma}_t H'^\top E'^{-1} H' \tilde{\Sigma}_t + \tilde{\Sigma}_t H'^\top E'^{-1} H' \Sigma_{t,xz} + (A - \tilde{\Sigma}_t H'^\top E'^{-1} H') \Sigma_{t,zz} + \Sigma_{t,zx} H'^\top E'^{-1} H' \tilde{\Sigma}_t + \Sigma_{t,zz}(A - \tilde{\Sigma}_t H'^\top E'^{-1} H')^\top \\ &= A\Sigma_{t,zz} + \Sigma_{t,zz}A^\top + \tilde{\Sigma}_t H'^\top E'^{-1} H' \tilde{\Sigma}_t + \tilde{\Sigma}_t H'^\top E'^{-1} H' (\Sigma_{t,xz} - \Sigma_{t,zz}) + (\Sigma_{t,zx} - \Sigma_{t,zz})H'^\top E'^{-1} H' \tilde{\Sigma}_t. \end{aligned} \quad (\text{G39})$$

We note that $\Sigma_{t,zx} = \Sigma_{t,xz}^\top$. From Eqs. (G32), (G35), and (G36),

$$\mu_{t,x} = \mu_{t,z} \quad (\text{G40})$$

holds. In addition, we assume that

$$\Sigma_{t,xz} = \Sigma_{t,zz} \quad (\text{G41})$$

holds, which is consistent with the initial condition, $\Sigma_{0,xz} = \Sigma_{0,zz} = O$. Under this assumption, Eq. (G39) is calculated as follows:

$$\frac{d\Sigma_{t,zz}}{dt} = A\Sigma_{t,zz} + \Sigma_{t,zz}A^\top + \check{\Sigma}_t H'^\top E'^{-1} H' \check{\Sigma}_t. \quad (\text{G42})$$

In addition, the conditional covariance matrix of state x given memory z , $\Sigma_{t,x|z} := \Sigma_{t,xx} - \Sigma_{t,xz} \Sigma_{t,zz}^{-1} \Sigma_{t,zx}$, is calculated as follows:

$$\Sigma_{t,x|z} = \Sigma_{t,xx} - \Sigma_{t,xz}. \quad (\text{G43})$$

From Eqs. (G37) and (G38), the dynamics of $\Sigma_{t,x|z}$ is given by

$$\begin{aligned} \frac{d\Sigma_{t,x|z}}{dt} &= \frac{d\Sigma_{t,xx}}{dt} - \frac{d\Sigma_{t,xz}}{dt} \\ &= D + A\Sigma_{t,xx} + \Sigma_{t,xx}A^\top - [A\Sigma_{t,xz} + \Sigma_{t,xz}A^\top + (\Sigma_{t,xx} - \Sigma_{t,xz})H'^\top E'^{-1} H' \check{\Sigma}_t] \\ &= D + A(\Sigma_{t,xx} - \Sigma_{t,xz}) + (\Sigma_{t,xx} - \Sigma_{t,xz})A^\top - (\Sigma_{t,xx} - \Sigma_{t,xz})H'^\top E'^{-1} H' \check{\Sigma}_t \\ &= D + A\Sigma_{t,x|z} + \Sigma_{t,x|z}A^\top - \Sigma_{t,x|z}H'^\top E'^{-1} H' \check{\Sigma}_t, \end{aligned} \quad (\text{G44})$$

where $\Sigma_{0,x|z} = \Sigma_{0,xx}$. Comparing Eq. (G44) with Eq. (G5),

$$\check{\Sigma}_t = \Sigma_{t,x|z} \quad (\text{G45})$$

hold. Hence, Eq. (G42) is equivalent to Eq. (G38), which justifies the assumption of Eq. (G41).

Next, we discuss optimal state estimator function $\hat{x}^*(t, y, z) = \mathbb{E}_{p_t(x|y,z)}[x]$. Since

$$p_t(\tilde{x}) = \mathcal{N}(\tilde{x}|\tilde{\mu}, \tilde{\Sigma}) = \mathcal{N}\left(\begin{pmatrix} x \\ z \end{pmatrix} \middle| \begin{pmatrix} \mu_x \\ \mu_z \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xz} \\ \Sigma_{zx} & \Sigma_{zz} \end{pmatrix}\right), \quad (\text{G46})$$

$$p_t(\tilde{y}|\tilde{x}) = \mathcal{N}(\tilde{y}|\tilde{H}\tilde{x}, \tilde{E}) := \mathcal{N}\left(\begin{pmatrix} y \\ z \end{pmatrix} \middle| \begin{pmatrix} H & O \\ O & I \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix}, \begin{pmatrix} E & O \\ O & O \end{pmatrix}\right), \quad (\text{G47})$$

the conditional expectation of extended state $\tilde{x}$ (state x and memory z) given extended observation $\tilde{y}$ (observation y and memory z) is provided by

$$\mathbb{E}_{p_t(\tilde{x}|\tilde{y})}[\tilde{x}] = \tilde{\mu} + \tilde{K}(\tilde{y} - \tilde{H}\tilde{\mu}), \quad (\text{G48})$$

where estimation gain $\tilde{K}$ is given by

$$\tilde{K} = \tilde{\Sigma}\tilde{H}^\top(\tilde{E} + \tilde{H}\tilde{\Sigma}\tilde{H}^\top)^{-1}. \quad (\text{G49})$$

Estimation gain $\tilde{K}$ is further calculated as follows:

$$\begin{aligned} \tilde{K} &= \begin{pmatrix} K_{xy} & K_{xz} \\ K_{zy} & K_{zz} \end{pmatrix} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xz} \\ \Sigma_{zx} & \Sigma_{zz} \end{pmatrix} \begin{pmatrix} H & O \\ O & I \end{pmatrix}^\top \left(\begin{pmatrix} E & O \\ O & O \end{pmatrix} + \begin{pmatrix} H & O \\ O & I \end{pmatrix} \begin{pmatrix} \Sigma_{xx} & \Sigma_{xz} \\ \Sigma_{zx} & \Sigma_{zz} \end{pmatrix} \begin{pmatrix} H & O \\ O & I \end{pmatrix}^\top \right)^{-1} \\ &= \begin{pmatrix} \Sigma_{x|z}H^\top(E + H\Sigma_{x|z}H^\top)^{-1} & (I - \Sigma_{x|z}H^\top(E + H\Sigma_{x|z}H^\top)^{-1}H)\Sigma_{xz}\Sigma_{zz}^{-1} \\ O & I \end{pmatrix}. \end{aligned} \quad (\text{G50})$$

Since

$$\mathbb{E}_{p_t(\tilde{x}|\tilde{y})}[\tilde{x}] = \begin{pmatrix} \mathbb{E}_{p_t(x|y,z)}[x] \\ z \end{pmatrix}, \quad (\text{G51})$$

optimal state estimator function $\hat{x}^*(t, y, z) = \mathbb{E}_{p_t(x|y,z)}[x]$ is given by

$$\hat{x}^*(t, y, z) = \mathbb{E}_{p_t(x|y,z)}[x] = \mu_x + K_{xy}(y - H\mu_x) + K_{xz}(z - \mu_z), \quad (\text{G52})$$

where

$$K_{xy} = \Sigma_{x|z}H^\top(E + H\Sigma_{x|z}H^\top)^{-1}, \quad (\text{G53})$$

$$K_{xz} = (I - \Sigma_{x|z}H^\top(E + H\Sigma_{x|z}H^\top)^{-1}H)\Sigma_{xz}\Sigma_{zz}^{-1}. \quad (\text{G54})$$

Since $E = E'/(dt)$ is infinitely large and $\Sigma_{xz} = \Sigma_{zz}$, estimation gains K_{xy} and K_{xz} are calculated as follows:

$$K_{xy} = O, \quad (G55)$$

$$K_{xz} = I. \quad (G56)$$

In addition, as $\mu_z = \mu_x$, optimal state estimator function $\hat{x}^*(t, y, z)$ is given by

$$\hat{x}^*(t, y, z) = z. \quad (G57)$$

Therefore, the memory is equivalent to the optimal state estimator in this setting. Considering that z corresponds to $\check{x}$, this result is consistent with the Kalman filter.

Finally, we examine value function $w(t, x, z)$ and verify the conditions in Eq. (G23). We assume that value function $w(t, x, z)$ is given by the following quadratic function:

$$w(t, x, z) = \tilde{x}^\top \tilde{\Pi}(t) \tilde{x} + \tilde{\alpha}^\top(t) \tilde{x} + \tilde{\beta}(t), \quad (G58)$$

where $\tilde{\Pi}$, $\tilde{\alpha}$, and $\tilde{\beta}$ are defined by

$$\tilde{\Pi} = \begin{pmatrix} \Pi_{xx} & \Pi_{xz} \\ \Pi_{zx} & \Pi_{zz} \end{pmatrix}, \quad \tilde{\alpha} = \begin{pmatrix} \alpha_x \\ \alpha_z \end{pmatrix}, \quad \tilde{\beta} = \beta. \quad (G59)$$

This assumption is also used in Appendix C. Under this assumption, the conditions in Eq. (G23) are calculated as follows:

$$0 \neq \frac{\partial w(t, x, z)}{\partial z} = 2(\Pi_{zx}x + \Pi_{zz}z) + \alpha_z, \quad (G60)$$

$$0 = \mathbb{E}_{p_t(x|y,z)} \left[\frac{\partial w(t, x, z)}{\partial z} \right] = 2(\Pi_{zx} \mathbb{E}_{p_t(x|y,z)}[x] + \Pi_{zz}z) + \alpha_z = 2(\Pi_{zx} + \Pi_{zz})z + \alpha_z. \quad (G61)$$

This result indicates that the solution corresponding to the Kalman filter should satisfy the following conditions:

$$\Pi_{zx} \neq O, \quad (G62)$$

$$\Pi_{zz} \neq O, \quad (G63)$$

$$\Pi_{zx} + \Pi_{zz} = O, \quad (G64)$$

$$\alpha_z = 0. \quad (G65)$$

We verify these conditions by calculating HJB equation [Eq. (G21)]. By using Eq. (G58), HJB equation [Eq. (G21)] is calculated as follows:

$$\begin{aligned} -\left(\tilde{x}^\top \frac{d\tilde{\Pi}}{dt} \tilde{x} + \left(\frac{d\tilde{\alpha}}{dt} \right)^\top \tilde{x} + \frac{d\tilde{\beta}}{dt} \right) &= \mathbb{E}_{p_t(y|x)} [(x - \hat{x}^*(t, y, z))^\top Q (x - \hat{x}^*(t, y, z))] + \left(\frac{\partial w(t, x, z)}{\partial x} \right)^\top A x \\ &\quad + \left(\frac{\partial w(t, x, z)}{\partial z} \right)^\top \mathbb{E}_{p_t(y|x)} [v^*(t, y, z)] + \frac{1}{2} \text{tr} \left\{ \frac{\partial}{\partial x} \left(\frac{\partial w(t, x, z)}{\partial x} \right)^\top D \right\} \\ &= (x - z)^\top Q (x - z) + (2(\Pi_{xx}x + \Pi_{xz}z) + \alpha_x)^\top A x \\ &\quad + (2(\Pi_{zx}x + \Pi_{zz}z) + \alpha_z)^\top (\check{\Sigma}_t H'^\top E'^{-1} H' x + (A - \check{\Sigma}_t H'^\top E'^{-1} H') z) + \text{tr}(\Pi_{zz} D) \\ &= x^\top (Q + A^\top \Pi_{xx} + \Pi_{xx} A + H'^\top E'^{-1} H' \check{\Sigma}_t \Pi_{zx} + \Pi_{xz} \check{\Sigma}_t H'^\top E'^{-1} H') x \\ &\quad + 2z^\top (-Q + \Pi_{zx} A + \Pi_{zz} \check{\Sigma}_t H'^\top E'^{-1} H' + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \Pi_{zx}) x \\ &\quad + z^\top (Q + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \Pi_{zz} + \Pi_{zz} (A - \check{\Sigma}_t H'^\top E'^{-1} H')) z \\ &\quad + (A^\top \alpha_x + H'^\top E'^{-1} H' \check{\Sigma}_t \alpha_z)^\top x + ((A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \alpha_z)^\top z + \text{tr}(\Pi_{zz} D). \end{aligned} \quad (G66)$$

As a result, HJB equation [Eq. (G21)] is reduced to the following ordinary differential equations:

$$-\frac{d\Pi_{xx}}{dt} = Q + A^\top \Pi_{xx} + \Pi_{xx} A + H'^\top E'^{-1} H' \check{\Sigma}_t \Pi_{zx} + \Pi_{xz} \check{\Sigma}_t H'^\top E'^{-1} H', \quad (G67)$$

$$-\frac{d\Pi_{zx}}{dt} = -Q + \Pi_{zx} A + \Pi_{zz} \check{\Sigma}_t H'^\top E'^{-1} H' + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \Pi_{zx}, \quad (G68)$$

$$-\frac{d\Pi_{zz}}{dt} = Q + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \Pi_{zz} + \Pi_{zz} (A - \check{\Sigma}_t H'^\top E'^{-1} H'), \quad (\text{G69})$$

$$-\frac{d\alpha_x}{dt} = A^\top \alpha_x + H'^\top E'^{-1} H' \check{\Sigma}_t \alpha_z, \quad (\text{G70})$$

$$-\frac{d\alpha_z}{dt} = (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \alpha_z, \quad (\text{G71})$$

$$-\frac{d\beta}{dt} = \text{tr}(\Pi_{zz} D), \quad (\text{G72})$$

where $\Pi_{xx}(T) = \Pi_{zx}(T) = \Pi_{zz}(T) = O$, $\alpha_x(T) = \alpha_z(T) = 0$, and $\beta(T) = 0$. Obviously, $\Pi_{zx}(t) = O$ and $\Pi_{zz}(t) = O$ do not hold except for $t = T$, which satisfies Eqs. (G62) and (G63). In addition, $\alpha_z(t) = 0$ holds, which satisfies Eq. (G65). The remaining condition is Eq. (G64). The dynamics of $\Pi_{zx} + \Pi_{zz}$ is calculated as follows:

$$\begin{aligned} -\frac{d(\Pi_{zx} + \Pi_{zz})}{dt} &= -\frac{d\Pi_{zx}}{dt} - \frac{d\Pi_{zz}}{dt} \\ &= -Q + \Pi_{zx} A + \Pi_{zz} \check{\Sigma}_t H'^\top E'^{-1} H' + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \Pi_{zx} \\ &\quad + Q + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top \Pi_{zz} + \Pi_{zz} (A - \check{\Sigma}_t H'^\top E'^{-1} H') \\ &= (\Pi_{zx} + \Pi_{zz}) A + (A - \check{\Sigma}_t H'^\top E'^{-1} H')^\top (\Pi_{zx} + \Pi_{zz}), \end{aligned} \quad (\text{G73})$$

where $\Pi_{zx}(T) + \Pi_{zz}(T) = O$. The solution of this equation is given by $\Pi_{zx}(t) + \Pi_{zz}(t) = O$, which satisfies Eq. (G64). Therefore, the solution corresponding to the Kalman filter is consistent with optimal memory control theory.

-
- [1] M. O. Ernst and M. S. Banks, Humans integrate visual and haptic information in a statistically optimal fashion, *Nature (London)* **415**, 429 (2002).
 - [2] Y. Weiss, E. P. Simoncelli, and E. H. Adelson, Motion illusions as optimal percepts, *Nat. Neurosci.* **5**, 598 (2002).
 - [3] K. P. Körding and D. M. Wolpert, Bayesian integration in sensorimotor learning, *Nature (London)* **427**, 244 (2004).
 - [4] A. A. Stocker and E. P. Simoncelli, Noise characteristics and prior expectations in human visual speed perception, *Nat. Neurosci.* **9**, 578 (2006).
 - [5] T. E. J. Behrens, M. W. Woolrich, M. E. Walton, and M. F. S. Rushworth, Learning the value of information in an uncertain world, *Nat. Neurosci.* **10**, 1214 (2007).
 - [6] J. Burge, M. O. Ernst, and M. S. Banks, The statistical determinants of adaptation rate in human reaching, *J. Vision* **8**, 20 (2008).
 - [7] H. C. Berg and E. M. Purcell, Physics of chemoreception, *Biophys. J.* **20**, 193 (1977).
 - [8] W. Bialek and S. Setayeshgar, Physical limits to biochemical signaling, *Proc. Natl. Acad. Sci. USA* **102**, 10040 (2005).
 - [9] R. G. Endres and N. S. Wingreen, Maximum likelihood and the single receptor, *Phys. Rev. Lett.* **103**, 158101 (2009).
 - [10] F. Tostevin and P. R. ten Wolde, Mutual information between input and output trajectories of biochemical networks, *Phys. Rev. Lett.* **102**, 218101 (2009).
 - [11] A. Celani and M. Vergassola, Bacterial strategies for chemotaxis response, *Proc. Natl. Acad. Sci. USA* **107**, 1391 (2010).
 - [12] H. H. Mattingly, K. Kamino, B. B. Machta, and T. Emonet, *Escherichia coli* chemotaxis is information limited, *Nat. Phys.* **17**, 1426 (2021).
 - [13] N. B. Becker, A. Mugler, and P. R. ten Wolde, Optimal prediction by cellular signaling networks, *Phys. Rev. Lett.* **115**, 258103 (2015).
 - [14] K. Nakamura and T. J. Kobayashi, Connection between the bacterial chemotactic network and optimal filtering, *Phys. Rev. Lett.* **126**, 128102 (2021).
 - [15] K. Nakamura and T. J. Kobayashi, Optimal sensing and control of run-and-tumble chemotaxis, *Phys. Rev. Res.* **4**, 013120 (2022).
 - [16] S. B. Laughlin, R. R. de Ruyter van Steveninck, and J. C. Anderson, The metabolic cost of neural information, *Nat. Neurosci.* **1**, 36 (1998).
 - [17] S. B. Laughlin, Energy as a constraint on the coding and processing of sensory information, *Curr. Opin. Neurobiol.* **11**, 475 (2001).
 - [18] P. Lennie, The cost of cortical computation, *Curr. Biol.* **13**, 493 (2003).
 - [19] J. E. Niven and S. B. Laughlin, Energy limitation as a selective pressure on the evolution of sensory systems, *J. Exp. Biol.* **211**, 1792 (2008).
 - [20] D. F. Rolfe and G. C. Brown, Cellular energy utilization and molecular origin of standard metabolic rate in mammals, *Physiol. Rev.* **77**, 731 (1997).
 - [21] A. Ames, CNS energy metabolism as related to function, *Brain Res. Rev.* **34**, 42 (2000).
 - [22] L. C. Aiello and P. Wheeler, The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution, *Curr. Anthropol.* **36**, 199 (1995).
 - [23] K. Isler and C. P. van Schaik, Metabolic costs of brain size evolution, *Biol. Lett.* **2**, 557 (2006).
 - [24] K. Fonseca-Azevedo and S. Herculano-Houzel, Metabolic constraint imposes tradeoff between body size and number of brain neurons in human evolution, *Proc. Natl. Acad. Sci. USA* **109**, 18571 (2012).
 - [25] V. Balasubramanian, Brain power, *Proc. Natl. Acad. Sci. USA* **118**, e2107022118 (2021).

- [26] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Stochastic gene expression in a single cell, *Science* **297**, 1183 (2002).
- [27] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, *Proc. Natl. Acad. Sci. USA* **99**, 12795 (2002).
- [28] J. Paulsson, Summing up the noise in gene networks, *Nature (London)* **427**, 415 (2004).
- [29] I. Lestas, G. Vinnicombe, and J. Paulsson, Fundamental limits on the suppression of molecular fluctuations, *Nature (London)* **467**, 174 (2010).
- [30] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells, *Science* **329**, 533 (2010).
- [31] T. Shibata and K. Fujimoto, Noisy signal amplification in ultrasensitive signal transduction, *Proc. Natl. Acad. Sci. USA* **102**, 331 (2005).
- [32] M. Ueda and T. Shibata, Stochastic signal processing and transduction in chemotactic response of eukaryotic cells, *Biophys. J.* **93**, 11 (2007).
- [33] T. J. Kobayashi and A. Kamimura, Dynamics of intracellular information decoding, *Phys. Biol.* **8**, 055007 (2011).
- [34] H. H. Mattingly, K. Kamino, J. Ong, R. Kottou, T. Emonet, and B. B. Machta, Chemotaxing *E. coli* do not count single molecules, [arXiv:2407.07264](https://arxiv.org/abs/2407.07264).
- [35] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, The energy-speed-accuracy trade-off in sensory adaptation, *Nat. Phys.* **8**, 422 (2012).
- [36] C. C. Govern and P. R. ten Wolde, Optimal resource allocation in cellular sensing systems, *Proc. Natl. Acad. Sci. USA* **111**, 17486 (2014).
- [37] C. C. Govern and P. R. ten Wolde, Energy dissipation and noise correlations in biochemical sensing, *Phys. Rev. Lett.* **113**, 258102 (2014).
- [38] G. Malaguti and P. R. ten Wolde, Theory for the optimal detection of time-varying signals in cellular sensing systems, *eLife* **10**, e62574 (2021).
- [39] G. Malaguti and P. R. ten Wolde, Receptor time integration via discrete sampling, *Phys. Rev. E* **105**, 054406 (2022).
- [40] A. J. Tjalma, V. Galstyan, J. Goedhart, L. Slim, N. B. Becker, and P. R. ten Wolde, Trade-offs between cost and information in cellular prediction, *Proc. Natl. Acad. Sci. USA* **120**, e2303078120 (2023).
- [41] A. J. Tjalma and P. R. t. Wolde, Predicting concentration changes via discrete receptor sampling, *Phys. Rev. Res.* **6**, 033049 (2024).
- [42] P. Mehta and D. J. Schwab, Energetic costs of cellular computation, *Proc. Natl. Acad. Sci. USA* **109**, 17978 (2012).
- [43] A. H. Lang, C. K. Fisher, T. Mora, and P. Mehta, Thermodynamics of statistical inference by cells, *Phys. Rev. Lett.* **113**, 148103 (2014).
- [44] S. E. Harvey, S. Lahiri, and S. Ganguli, Universal energy-accuracy tradeoffs in nonequilibrium cellular sensing, *Phys. Rev. E* **108**, 014403 (2023).
- [45] D. Hartich, A. C. Barato, and U. Seifert, Sensory capacity: An information theoretical measure of the performance of a sensor, *Phys. Rev. E* **93**, 022116 (2016).
- [46] S. J. Bryant and B. B. Machta, Physical constraints in intracellular signaling: The cost of sending a bit, *Phys. Rev. Lett.* **131**, 068401 (2023).
- [47] S. Ito and T. Sagawa, Maxwell's demon in biochemical signal transduction with feedback loop, *Nat. Commun.* **6**, 7498 (2015).
- [48] T. Matsumoto and T. Sagawa, Role of sufficient statistics in stochastic thermodynamics and its implication to sensory adaptation, *Phys. Rev. E* **97**, 042103 (2018).
- [49] N. Ohga, S. Ito, and A. Kolchinsky, Thermodynamic bound on the asymmetry of cross-correlations, *Phys. Rev. Lett.* **131**, 077101 (2023).
- [50] <https://creativecommons.org/licenses/by/4.0/>.
- [51] G. Kallianpur, *Stochastic Filtering Theory, Stochastic Modelling and Applied Probability*, edited by A. V. Balakrishnan (Springer New York, New York, NY, 1980), Vol. 13.
- [52] Z. Chen, Bayesian filtering: From Kalman filters to particle filters, and beyond, *Statistics* **182**, 1 (2003).
- [53] A. H. Jazwinski, *Stochastic Processes and Filtering Theory* (Courier Corporation, Mineola, New York, 2007).
- [54] A. Bain and D. Crisan, in *Fundamentals of Stochastic Filtering, Stochastic Modelling and Applied Probability*, edited by B. Rozovski and G. Grimmett (Springer, New York, NY, 2009), Vol. 60.
- [55] A. Pouget, P. Dayan, and R. Zemel, Information processing with population codes, *Nat. Rev. Neurosci.* **1**, 125 (2000).
- [56] D. C. Knill and A. Pouget, The Bayesian brain: The role of uncertainty in neural coding and computation, *Trends Neurosci.* **27**, 712 (2004).
- [57] K. Doya, S. Ishii, A. Pouget, and R. P. N. Rao, *Bayesian Brain: Probabilistic Approaches to Neural Coding* (The MIT Press, Cambridge, MA, 2006).
- [58] K. P. Körding and D. M. Wolpert, Bayesian decision theory in sensorimotor control, *Trends Cognit. Sci.*, **10**, 319 (2006).
- [59] A. Pouget, J. M. Beck, W. J. Ma, and P. E. Latham, Probabilistic brains: Knowns and unknowns, *Nat. Neurosci.* **16**, 1170 (2013).
- [60] H. Sohn and D. Narain, Neural implementations of Bayesian inference, *Curr. Opin. Neurobiol.* **70**, 121 (2021).
- [61] R. D. Lange, S. Shivkumar, A. Chatteraj, and R. M. Haefner, Bayesian encoding and decoding as distinct perspectives on neural coding, *Nat. Neurosci.* **26**, 2063 (2023).
- [62] S. Denève, J.-R. Duhamel, and A. Pouget, Optimal sensorimotor integration in recurrent cortical networks: A neural implementation of kalman filters, *J. Neurosci.* **27**, 5744 (2007).
- [63] R. Wilson and L. Finkel, A neural Implementation of the Kalman Filter, *Advances in Neural Information Processing Systems* (Curran Associates, Red Hook, New York, 2009), Vol. 22.
- [64] N. Hiratani and T. Fukai, Redundancy in synaptic connections enables neurons to learn optimally, *Proc. Natl. Acad. Sci. USA* **115**, E6871 (2018).
- [65] N. Hiratani and P. E. Latham, Rapid Bayesian learning in the mammalian olfactory system, *Nat. Commun.* **11**, 3845 (2020).
- [66] L. Aitchison, J. Jegminat, J. A. Menendez, J.-P. Pfister, A. Pouget, and P. E. Latham, Synaptic plasticity as Bayesian inference, *Nat. Neurosci.* **24**, 565 (2021).
- [67] A. Kutschireiter, M. A. Basnak, R. I. Wilson, and J. Drugowitsch, Bayesian inference in ring attractor networks, *Proc. Natl. Acad. Sci. USA* **120**, e2210622120 (2023).

- [68] B. W. Andrews, T.-M. Yi, and P. A. Iglesias, Optimal noise filtering in the chemotactic response of *Escherichia coli*, *PLoS Comput. Biol.* **2**, e154 (2006).
- [69] T. J. Kobayashi, Implementation of dynamic bayesian decision making by intracellular kinetics, *Phys. Rev. Lett.* **104**, 228104 (2010).
- [70] M. Hinczewski and D. Thirumalai, Cellular signaling networks function as generalized wiener-kolmogorov filters to suppress noise, *Phys. Rev. X* **4**, 041017 (2014).
- [71] C. Zechner, G. Seelig, M. Rullan, and M. Khammash, Molecular circuits for dynamic noise filtering, *Proc. Natl. Acad. Sci USA* **113**, 4729 (2016).
- [72] T. Mora and I. Nemenman, Physical limit to concentration sensing in a changing environment, *Phys. Rev. Lett.* **123**, 198101 (2019).
- [73] K. Husain, W. Pittayakanchit, G. Pattanayak, M. J. Rust, and A. Murugan, Kalman-like self-tuned sensitivity in biophysical sensing, *Cell Syst.* **9**, 459 (2019).
- [74] M. Novak and B. M. Friedrich, Bayesian gradient sensing in the presence of rotational diffusion, *New J. Phys.* **23**, 043026 (2021).
- [75] A. Auconi, M. Novak, and B. M. Friedrich, Gradient sensing in Bayesian chemotaxis, *Europhys. Lett.* **138**, 12001 (2022).
- [76] J. Rode, M. Novak, and B. M. Friedrich, Information theory of chemotactic agents using both spatial and temporal gradient sensing, *PRX Life* **2**, 023012 (2024).
- [77] K. Nakamura and T. J. Kobayashi, Gradient sensing limit of an elongated cell with orientational control, *Phys. Rev. E* **110**, 064407 (2024).
- [78] R. S. Zemel, P. Dayan, and A. Pouget, Probabilistic interpretation of population codes, *Neural Comput.* **10**, 403 (1998).
- [79] W. J. Ma, J. M. Beck, P. E. Latham, and A. Pouget, Bayesian inference with probabilistic population codes, *Nat. Neurosci.* **9**, 1432 (2006).
- [80] A. A. Faisal, L. P. J. Selen, and D. M. Wolpert, Noise in the nervous system, *Nat. Rev. Neurosci.* **9**, 292 (2008).
- [81] C. Ribault, K. Sekimoto, and A. Triller, From the stochasticity of molecular processes to the variability of synaptic transmission, *Nat. Rev. Neurosci.* **12**, 375 (2011).
- [82] M. D. McDonnell and L. M. Ward, The benefits of noise in neural systems: bridging theory and experiment, *Nat. Rev. Neurosci.* **12**, 415 (2011).
- [83] R. Dukas, Costs of memory: Ideas and predictions, *J. Theor. Biol.* **197**, 41 (1999).
- [84] F. Mery and T. J. Kawecki, A fitness cost of learning ability in *Drosophila melanogaster*, *Proc. R. Soc. London B* **270**, 2465 (2003).
- [85] F. Mery and T. J. Kawecki, A cost of long-term memory in *Drosophila*, *Science* **308**, 1148 (2005).
- [86] J. M. S. Burger, M. Kolss, J. Pont, and T. J. Kawecki, Learning ability and longevity: A symmetrical evolutionary trade-off in *Drosophila*, *Evolution* **62**, 1294 (2008).
- [87] M. Kolss and T. J. Kawecki, Reduced learning ability as a consequence of evolutionary adaptation to nutritional stress in *Drosophila melanogaster*, *Ecol. Entomol.* **33**, 583 (2008).
- [88] J. G. Burns, J. Foucaud, and F. Mery, Costs of memory: Lessons from “mini” brains, *Proc. R. Soc. B* **278**, 923 (2011).
- [89] N. Tishby, F. C. Pereira, and W. Bialek, The information bottleneck method, [arXiv:physics/0004057](https://arxiv.org/abs/physics/0004057).
- [90] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, Information bottleneck for gaussian variables, *J. Mach. Learn. Res.* **6**, 165 (2005).
- [91] F. Creutzig, A. Globerson, and N. Tishby, Past-future information bottleneck in dynamical systems, *Phys. Rev. E* **79**, 041925 (2009).
- [92] R. Fox and N. Tishby, Minimum-information LQG control part I: Memoryless controllers, in *2016 IEEE 55th Conference on Decision and Control (CDC)* (IEEE, Piscataway, NJ, 2016), pp. 5610–5616.
- [93] R. Fox and N. Tishby, Minimum-information LQG control Part II: Retentive controllers, in *2016 IEEE 55th Conference on Decision and Control (CDC)* (IEEE, Piscataway, NJ, 2016), pp. 5603–5609.
- [94] V. Sachdeva, T. Mora, A. M. Walczak, and S. E. Palmer, Optimal prediction with resource constraints using the information bottleneck, *PLoS Comput. Biol.* **17**, e1008743 (2021).
- [95] M. Bauer and W. Bialek, Information bottleneck in molecular sensing, *PRX Life* **1**, 023005 (2023).
- [96] E. Todorov and M. I. Jordan, Optimal feedback control as a theory of motor coordination, *Nat. Neurosci.* **5**, 1226 (2002).
- [97] E. Todorov, Optimality principles in sensorimotor control, *Nat. Neurosci.* **7**, 907 (2004).
- [98] E. Todorov, Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system, *Neural Comput.* **17**, 1084 (2005).
- [99] J. A. Kromer, S. Mrcker, S. Lange, C. Baier, and B. M. Friedrich, Decision making improves sperm chemotaxis in the presence of noise, *PLoS Comput. Biol.* **14**, e1006109 (2018).
- [100] C. Calascibetta, L. Biferale, F. Borra, A. Celani, and M. Cencini, Optimal tracking strategies in a turbulent flow, *Commun. Phys.* **6**, 256 (2023).
- [101] W. M. Wonham, On the separation theorem of stochastic control, *SIAM J. Control* **6**, 312 (1968).
- [102] M. H. A. Davis and P. Varaiya, Dynamic programming conditions for partially observable stochastic systems, *SIAM J. Control* **11**, 226 (1973).
- [103] A. Bensoussan, *Stochastic Control of Partially Observable Systems* (Cambridge University Press, Cambridge, 1992).
- [104] J. Yong and X. Y. Zhou, *Stochastic Controls* (Springer New York, New York, NY, 1999).
- [105] M. Nisio, *Stochastic Control Theory, Probability Theory and Stochastic Modelling* (Springer, Tokyo, Japan, 2015), Vol. 72.
- [106] A. Bensoussan, *Estimation and Control of Dynamical Systems, Interdisciplinary Applied Mathematics* (Springer International, Cham, 2018), Vol. 48.
- [107] G. Wang, Z. Wu, and J. Xiong, *An Introduction to Optimal Control of FBSDE with Incomplete Information, Springer-Briefs in Mathematics* (Springer International, Cham, 2018).
- [108] W. Li and E. Todorov, An iterative optimal control and estimation design for nonlinear stochastic system, in *Proceedings of the 45th IEEE Conference on Decision and Control* (IEEE, Piscataway, NJ, 2006), pp. 3242–3247.
- [109] W. Li and E. Todorov, Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system, *Int. J. Control* **80**, 1439 (2007).
- [110] R. A. Heinonen, L. Biferale, A. Celani, and M. Vergassola, Optimal policies for Bayesian olfactory search in turbulent flows, *Phys. Rev. E* **107**, 055105 (2023).

- [111] K. V. B. Verano, E. Panizon, and A. Celani, Olfactory search with finite-state controllers, *Proc. Natl. Acad. Sci. USA* **120**, e2304230120 (2023).
- [112] T. Tottori and T. J. Kobayashi, companion paper, Resource limitations induce phase transitions in biological information processing, *Phys. Rev. Res.* **7**, L042012 (2025).
- [113] T. Tottori and T. J. Kobayashi, Memory-limited partially observable stochastic control and its mean-field control approach, *Entropy* **24**, 1599 (2022).
- [114] T. Tottori and T. J. Kobayashi, Forward-backward sweep method for the system of HJB-FP equations in memory-limited partially observable stochastic control, *Entropy* **25**, 208 (2023).
- [115] T. Tottori and T. J. Kobayashi, Decentralized stochastic control with finite-dimensional memories: A memory limitation approach, *Entropy* **25**, 791 (2023).
- [116] R. E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng.* **82**, 35 (1960).
- [117] R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory, *J. Basic Eng.* **83**, 95 (1961).
- [118] C. K. Chui and G. Chen, *Kalman Filtering* (Springer International, Cham, 2017).
- [119] A. Bensoussan, J. Frehse, and S. C. P. Yam, The Master equation in mean field theory, *J. Math. Pures Appl.* **103**, 1441 (2015).
- [120] A. Bensoussan, J. Frehse, and S. C. P. Yam, On the interpretation of the master equation, *Stoch. Process. their Appl.* **127**, 2093 (2017).
- [121] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I, Probability theory and Stochastic Modelling* (Springer Nature, Cham, 2018), Vol. 83.
- [122] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications II, Probability Theory and Stochastic Modelling* (Springer International Publishing, Cham, 2018), Vol. 84.
- [123] C. Briat and M. Khammash, Continuous and sampled-data h control of linear stochastic reaction networks, in *2022 IEEE 61st Conference on Decision and Control (CDC)* (IEEE, Piscataway, NJ, 2022), pp. 1924–1929.
- [124] C. Briat and M. Khammash, Continuous-time and sampled-data optimal control of linear stochastic reaction networks, in *2022 IEEE 61st Conference on Decision and Control (CDC)* (IEEE, Piscataway, NJ, 2022), pp. 3865–3870.
- [125] C. Briat and M. Khammash, Optimal and H_∞ control of stochastic reaction networks, [arXiv:2111.14754](https://arxiv.org/abs/2111.14754) [cs, eess, math, q-bio].
- [126] T. Toyozumi, J.-P. Pfister, K. Aihara, and W. Gerstner, Optimality model of unsupervised spike-timing-dependent plasticity: Synaptic memory and weight distribution, *Neural Comput.* **19**, 639 (2007).
- [127] K. Yoshida and T. Toyozumi, Information maximization explains state-dependent synaptic plasticity and memory reorganization during non-rapid eye movement sleep, *PNAS Nexus* **2**, pgac286 (2023).
- [128] S. Gu, R. F. Betzel, M. G. Mattar, M. Cieslak, P. R. Delio, S. T. Grafton, F. Pasqualetti, and D. S. Bassett, Optimal trajectories of brain state transitions, *NeuroImage* **148**, 305 (2017).
- [129] J. Z. Kim, J. M. Soffer, A. E. Kahn, J. M. Vettel, F. Pasqualetti, and D. S. Bassett, Role of graph architecture in controlling dynamical networks with applications to neural systems, *Nat. Phys.* **14**, 91 (2018).
- [130] M. F. Krummel, F. Bartumeus, and A. Grard, T cell migration, search strategies and mechanisms, *Nat. Rev. Immunol.* **16**, 193 (2016).
- [131] D. R. Sowinski, J. Carroll-Nellenback, R. N. Markwick, J. Piero, M. Gleiser, A. Kolchinsky, G. Ghoshal, and A. Frank, Semantic information in a model of resource gathering agents, *PRX Life* **1**, 023003 (2023).
- [132] A. Celani and E. Panizon, Olfactory search, [arXiv:2405.03374](https://arxiv.org/abs/2405.03374) [physics].
- [133] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley, Optimizing the success of random searches, *Nature (London)* **401**, 911 (1999).
- [134] O. Bénichou, M. Coppey, M. Moreau, P.-H. Suet, and R. Voituriez, Optimal search strategies for hidden targets, *Phys. Rev. Lett.* **94**, 198101 (2005).
- [135] G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz, Lévy flights and superdiffusion in the context of biological encounters and random searches, *Phys. Life Rev.* **5**, 133 (2008).
- [136] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, Intermittent search strategies, *Rev. Mod. Phys.* **83**, 81 (2011).
- [137] A. M. Hein, F. Carrara, D. R. Brumley, R. Stocker, and S. A. Levin, Natural search algorithms as a bridge between organisms, evolution, and ecology, *Proc. Natl. Acad. Sci. USA* **113**, 9413 (2016).
- [138] K. Friston, J. Kilner, and L. Harrison, A free energy principle for the brain, *J. Physiol. Paris* **100**, 70 (2006).
- [139] K. Friston, The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **11**, 127 (2010).
- [140] C. L. Buckley, C. S. Kim, S. McGregor, and A. K. Seth, The free energy principle for action and perception: A mathematical review, *J. Math. Psychol.* **81**, 55 (2017).
- [141] Y. Konaka and H. Naoki, Decoding rewardcuriosity conflict in decision-making from irrational behaviors, *Nat. Comput. Sci.* **3**, 418 (2023).
- [142] J. S. Li, A. A. Sarma, T. J. Sejnowski, and J. C. Doyle, Internal feedback in the cortical perceptionaction loop enables fast and accurate behavior, *Proc. Natl. Acad. Sci. USA* **120**, e2300445120 (2023).
- [143] J. J. Moore, A. Genkin, M. Tournoy, J. L. Pughe-Sanford, R. R. de Ruyter van Steveninck, and D. B. Chklovskii, The neuron as a direct data-driven controller, *Proc. Natl. Acad. Sci. USA* **121**, e2311893121 (2024).
- [144] L. Boominathan and X. Pitkow, Phase transitions in when feedback is useful, *Adv. Neural Inf. Process. Syst.* **35**, 10849 (2022).
- [145] G. Tavoni, T. Doi, C. Pizzica, V. Balasubramanian, and J. I. Gold, Human inference reflects a normative balance of complexity and accuracy, *Nat. Hum. Behav.* **6**, 1153 (2022).
- [146] C. H. Johnson, C. Zhao, Y. Xu, and T. Mori, Timing the day: What makes bacterial clocks tick? *Nat. Rev. Microbiol.* **15**, 232 (2017).
- [147] M. Monti, D. K. Lubensky, and P. R. ten Wolde, Robustness of clocks to input noise, *Phys. Rev. Lett.* **121**, 078101 (2018).
- [148] W. Pittayakanchit, Z. Lu, J. Chew, M. J. Rust, and A. Murugan, Biophysical clocks face a trade-off between internal and external noise resistance, *eLife* **7**, e37624 (2018).

- [149] M. Seki and H. Ito, Evolution of self-sustained circadian rhythms is facilitated by seasonal change of daylight, *Proc. R. Soc. B: Biol. Sci.* **289**, 20220577 (2022).
- [150] M. Ishiura, S. Kutsuna, S. Aoki, H. Iwasaki, C. R. Andersson, A. Tanabe, S. S. Golden, C. H. Johnson, and T. Kondo, Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria, *Science* **281**, 1519 (1998).
- [151] H. Iwasaki, T. Nishiwaki, Y. Kitayama, M. Nakajima, and T. Kondo, KaiA-stimulated KaiC phosphorylation in circadian timing loops in cyanobacteria, *Proc. Natl. Acad. Sci. USA* **99**, 15788 (2002).
- [152] M. Nakajima, K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama, and T. Kondo, Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro, *Science* **308**, 414 (2005).
- [153] J. Tomita, M. Nakajima, T. Kondo, and H. Iwasaki, No transcription-translation feedback in circadian Rhythm of KaiC phosphorylation, *Science* **307**, 251 (2005).
- [154] N. Kawamoto, H. Ito, I. T. Tokuda, and H. Iwasaki, Damped circadian oscillation in the absence of KaiA in *Synechococcus*, *Nat. Commun.* **11**, 2242 (2020).
- [155] V. Dvornyk, O. Vinogradova, and E. Nevo, Origin and evolution of circadian clock genes in prokaryotes, *Proc. Natl. Acad. Sci. USA* **100**, 2495 (2003).
- [156] J. Holtzendorff, F. Partensky, D. Mella, J.-F. Lennon, W. R. Hess, and L. Garczarek, Genome streamlining results in loss of robustness of the circadian clock in the marine cyanobacterium *Prochlorococcus marinus* PCC 9511, *J. Biol. Rhythms* **23**, 187 (2008).
- [157] I. M. Axmann, U. Dhring, L. Seeliger, A. Arnold, J. T. Vanselow, A. Kramer, and A. Wilde, Biochemical evidence for a timing mechanism in *Prochlorococcus*, *J. Bacteriol.* **191**, 5342 (2009).
- [158] C. W. Mullineaux and R. Stanewsky, The rolex and the hourglass: A simplified circadian clock in *Prochlorococcus*? *J. Bacteriol.* **191**, 5333 (2009).
- [159] P. Ma, T. Mori, C. Zhao, T. Thiel, and C. H. Johnson, Evolution of KaiC-dependent timekeepers: A proto-circadian timing mechanism confers adaptive fitness in the purple bacterium *Rhodospseudomonas palustris*, *PLoS Genet.* **12**, e1005922 (2016).
- [160] J. Chew, E. Leypunskiy, J. Lin, A. Murugan, and M. J. Rust, High protein copy number is required to suppress stochasticity in the cyanobacterial circadian clock, *Nat. Commun.* **9**, 3004 (2018).
- [161] P. Flombaum, J. L. Gallegos, R. A. Gordillo, J. Rincn, L. L. Zabala, N. Jiao, D. M. Karl, W. K. W. Li, M. W. Lomas, D. Veneziano, C. S. Vera, J. A. Vrugt, and A. C. Martiny, Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*, *Proc. Natl. Acad. Sci. USA* **110**, 9824 (2013).
- [162] S. Epstein, Integration of the cognitive and the psychodynamic unconscious, *Am. Psychol.* **49**, 709 (1994).
- [163] M. Muraven, D. M. Tice, and R. F. Baumeister, Self-control as a limited resource: Regulatory depletion patterns, *J. Pers. Soc. Psychol.* **74**, 774 (1998).
- [164] D. Kahneman, A perspective on judgment and choice: Mapping bounded rationality, *Am. Psychol.* **58**, 697 (2003).
- [165] K. E. Stanovich, *Who Is Rational?: Studies of individual Differences in Reasoning* (Psychology Press, New York, 1999).
- [166] E. Masicampo and R. F. Baumeister, Toward a physiology of dual-process reasoning and judgment: Lemonade, willpower, and expensive rule-based analysis, *Psychol. Sci.* **19**, 255 (2008).
- [167] A. J. McMahon and M. H. Scheel, Glucose promotes controlled processing: Matching, maximizing, and root beer, *Judgm. Decis. Mak.* **5**, 450 (2010).
- [168] X. Wang and R. D. Dvorak, Sweet future: Fluctuating blood glucose levels affect future discounting, *Psychol. Sci.* **21**, 183 (2010).
- [169] L. Ruthotto, S. J. Osher, W. Li, L. Nurbekyan, and S. W. Fung, A machine learning framework for solving high-dimensional mean field game and mean field control problems, *Proc. Natl. Acad. Sci. USA* **117**, 9183 (2020).
- [170] A. T. Lin, S. W. Fung, W. Li, L. Nurbekyan, and S. J. Osher, Alternating the population and control neural networks to solve high-dimensional stochastic mean-field games, *Proc. Natl. Acad. Sci. USA* **118**, e2024713118 (2021).
- [171] H. Pham and X. Warin, Mean-field neural networks: Learning mappings on Wasserstein space, *Neural Networks* **168**, 380 (2023).
- [172] H. Cao, X. Guo, and M. Laurire, Connecting GANs, mean-field games, and optimal transport, *SIAM J. Appl. Math.* **84**, 1255 (2024).
- [173] S. Sasagawa, Y.-I. Ozaki, K. Fujita, and S. Kuroda, Prediction and validation of the distinct dynamics of transient and sustained ERK activation, *Nat. Cell Biol.* **7**, 365 (2005).
- [174] M. W. Covert, T. H. Leung, J. E. Gaston, and D. Baltimore, Achieving stability of lipopolysaccharide-induced NF- κ B activation, *Science* **309**, 1854 (2005).
- [175] Y. Hasegawa, Multidimensional biochemical information processing of dynamical patterns, *Phys. Rev. E* **97**, 022401 (2018).
- [176] L. Hahn, A. M. Walczak, and T. Mora, Dynamical information synergy in biochemical signaling networks, *Phys. Rev. Lett.* **131**, 128401 (2023).
- [177] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006), 1st ed.
- [178] L. van der Maaten, E. Postma, and H. Herik, Dimensionality reduction: A comparative review, *J. Mach. Learn. Res.* **10**, 66 (2007).
- [179] J. A. Lee and M. Verleysen, *Nonlinear Dimensionality Reduction* (Springer, New York, 2007).
- [180] S. Fancher and A. Mugler, Fundamental limits to collective concentration sensing in cell populations, *Phys. Rev. Lett.* **118**, 078101 (2017).
- [181] J. Varennes, S. Fancher, B. Han, and A. Mugler, Emergent versus individual-based multicellular chemotaxis, *Phys. Rev. Lett.* **119**, 188101 (2017).
- [182] A. Erez, T. A. Byrd, M. Vennettilli, and A. Mugler, Cell-to-cell information at a feedback-induced bifurcation point, *Phys. Rev. Lett.* **125**, 048103 (2020).
- [183] M. Vennettilli, A. Erez, and A. Mugler, Multicellular sensing at a feedback-induced critical point, *Phys. Rev. E* **102**, 052411 (2020).
- [184] L. Tweedy, P. A. Thomason, P. I. Paschke, K. Martin, L. M. Machesky, M. Zagnoni, and R. H. Insall, Seeing around

- corners: Cells solve mazes and respond at a distance using attractant breakdown, *Science* **369**, eaay9792 (2020).
- [185] A. Pezzotta, M. Adorisio, and A. Celani, Chemotaxis emerges as the optimal solution to cooperative search games, *Phys. Rev. E* **98**, 042401 (2018).
- [186] M. Durve, L. Piro, M. Cencini, L. Biferale, and A. Celani, Collective olfactory search in a turbulent environment, *Phys. Rev. E* **102**, 012402 (2020).
- [187] F. Borra, M. Cencini, and A. Celani, Optimal collision avoidance in swarms of active Brownian particles, *J. Stat. Mech.* (2021) 083401.
- [188] R. Bellman, Dynamic programming and stochastic control processes, *Inf. Control* **1**, 228 (1958).
- [189] H. Kushner, Optimal stochastic control, *IRE Trans. Autom. Control* **7**, 120 (1962).
- [190] W. Fleming and R. Rishel, *Deterministic and Stochastic Optimal Control* (Springer New York, New York, NY, 1975).
- [191] W. Hackbusch, A numerical method for solving parabolic equations with opposite orientations, *Computing* **20**, 229 (1978).
- [192] M. McAsey, L. Mou, and W. Han, Convergence of the forward-backward sweep method in optimal control, *Comput. Optim. Appl.* **53**, 207 (2012).
- [193] S. Lenhart and J. T. Workman, *Optimal Control Applied to Biological Models* (Chapman and Hall/CRC, New York, 2007).
- [194] J. A. Sharp, K. Burrage, and M. J. Simpson, Implementation and acceleration of optimal control for systems biology, *J. R. Soc. Interface.* **18**, 20210241 (2021).
- [195] E. Carlini and F. J. Silva, Semi-Lagrangian schemes for mean field game models, in *52nd IEEE Conference on Decision and Control* (IEEE, Piscataway, NJ, 2013), pp. 3115–3120.
- [196] E. Carlini and F. J. Silva, A fully discrete semi-lagrangian scheme for a first order mean field game problem, *SIAM J. Numer. Anal.* **52**, 45 (2014).
- [197] E. Carlini and F. J. Silva, A semi-Lagrangian scheme for a degenerate second order mean field game system, *Discrete Contin. Dyn. Syst.* **35**, 4269 (2015).
- [198] M. Lauriere, Numerical methods for mean field games and mean field type control, [arXiv:2106.06231](https://arxiv.org/abs/2106.06231) [cs, math].
- [199] L. S. Pontryagin, *Mathematical Theory of Optimal Processes* (CRC Press, Boca Raton, Florida, 1987), [google-Books-ID: kwzq0F4cBVAC](https://books.google.com/books?id=kwzq0F4cBVAC).
- [200] R. Vinter, *Optimal Control* (Birkhuser, Boston, 2010).
- [201] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control* (John Wiley & Sons, Hoboken, New Jersey, 2012).
- [202] <https://github.com/takehirotottori/Tottori2025-OMSC>
- [203] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2009), 4th ed.
- [204] N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, 2011).