Generalized end-product feedback circuit can sense high dimensional environmental fluctuations

Fang Yu and Mikhail Tikhonov
Department of Physics, Washington University in St. Louis, St. Louis, Missouri, USA

(Dated: October 29, 2024)

Understanding computational capabilities of simple biological circuits, such as the regulatory circuits of single-cell organisms, remains an active area of research. Recent theoretical work has shown that a simple cross-talk architecture based on end-product inhibition can exhibit predictive behavior by learning fluctuation statistics of one or two environmental parameters. Here we extend this analysis to higher dimensions, i.e., a large number of fluctuating inputs. We show that a generalized version of the cross-talk architecture can learn not only the dominant direction of fluctuations, as shown previously, but also the subdominant modes, orienting its responsiveness spectrum to the fluctuation eigenmodes. We comment on the relevance of our results to living systems at other scales of organization, such as ecosystems of species competing for fluctuating resources.

I. INTRODUCTION

As organisms evolve to better survive in changing environments, they develop adaptations that allow them to respond to change, but also to predict change. Characterizing such predictive (anticipatory [1–3]) behavior in microorganisms, whose regulatory circuits are far less complex than what can be achieved by a neuron-based brain [4], revealed many examples of evolutionary ingenuity attaining complex objectives with minimal ingredients [5, 6] (e.g., robust circadian clocks in photosynthetic algae, which allow them to reorganize their metabolism in preparation for sunrise [7]).

Theoretical computer science has long established that even the simplest building blocks, if used in sufficient numbers, can support complex computations: very simple instruction sets can already be Turing-complete [8–10]. The biologically relevant sister question—how simple of a *circuit* can perform how complex a task?—is understood less well. Some well-studied examples include the chemotaxis circuit achieving perfect adaptation [11, 12], mechanisms of temperature compensation in circadian clocks [7, 13], or the bistable genetic regulatory network storing and retrieving associative memories [14]. Still, understanding the computational capabilities of simple circuits remains an active area of research.

Recent theoretical work explored the ability of a simple circuit to learn complex statistical features of a fluctuating input [15] (as experienced, for example, by a bacterium faced with a fluctuating environment). Specifically, that work considered the case of two fluctuating environmental parameters, and demonstrated that their variances and correlations may, in principle, be both learned and usefully "recalled" by a simple circuit based on the end-product inhibition motif [15].

Here, we explore a high-dimensional generalization of this circuit. We ask whether the three ingredients identified in Ref. [15]—nonlinearity, an excess of regulators, and cross-talk between them—are sufficient to learn the fluctuation structure of high-dimensional environments. This extension is interesting, because the complexity of the task grows dramatically with dimension. Indeed, in the two-dimensional problem considered previously [15], the difficult part of learning the input fluctuation structure amounted to learning a single number: the direction of a single "dominant" direction of fluctuation. In contrast, in higher dimensions, even the simplest Gaussian-structured fluctuations include sub-dominant modes, as explained below. However, the cross-talk architecture solving the 2d problem is naturally generalizable to arbitrary dimension, prompting us to test its performance.

We find that, as in the low-dimensional case, the cross-talk architecture can upregulate its reactivity to respond faster in epochs when environment fluctuations are larger. We further show that even a small excess of regulators already makes the cross-talk architecture responsive to changes in fluctuation structure, and that the state adopted by the cross-talk architecture encodes both the dominant and subdominant fluctuation modes of environmental parameters.

Our analysis demonstrates how a highly complex task can be approximately solved by a simple circuit. We are cautious at drawing conclusions pertaining to biological regulatory circuits, as our abstract model ignores stochasticity of transcriptional regulation. However, the circuit ingredients considered here are ubiquitous across several contexts such as organelle dynamics or ecological interactions. As a result, our results may be relevant for appreciating the complexity of the behavior that the collective dynamics of such simple elements may be able to achieve [16–18].

II. THE MODEL

Our approach builds directly on that of Landmann *et al.* [15], but this section provides enough details to be self-contained.

Specific adaptation problems faced by real organisms are highly diverse. Following Ref. [15], here we distill the general problem of physiological learning to a minimal model. Specifically, we consider a scenario where a set of internal quantities $\vec{P} = (P_1, \ldots, P_N)$ (which a cell can regulate) must track a set of fluctuating external factors, $\vec{D} = (D_1, \ldots, D_N)$. For the sake of concreteness, we will think of this problem in metabolic terms, with P_i representing the rates of production of metabolites x_i. In our model, the cell seeks to match these production rates $\vec{P}$ to the (time-dependent) demands $\vec{D}(t)$ imposed by the external conditions. As an example, environmental conditions that trigger biofilm formation in bacteria require a different stoichiometry of synthesis than the condition of fast planktonic growth.

If the fluctuations of demands are slow, the organism could sense them and directly match $\vec{P}$ to $\vec{D}$ at all times. But if fluctuations are too fast to be followed precisely, the organism must instead rely on the "statistical structure" of $\vec{D}(t)$, such as the mean value or correlations between its components $D_i(t)$ [15]. If this statistical structure remains constant over a very long timescale, the optimal behavior (given this structure) could be hardwired into the cross-talk architecture by evolution. But if the structure itself occasionally changes, the organism would need to learn it from recent observations via physiological mechanisms. This is the regime where this problem can serve as a minimal model for the task of physiological learning.

There are different levels of statistical structure to be learned. Under our tracking problem, the simplest form of learning would be to set the production rates P_i to match the average demand in the recent past. Beyond that, the subtler statistics include the variances and correlations among fluctuations. To model $D(t)$ in a way where both means and correlations can be tuned, we consider a multi-dimensional random walk in a quadratic potential [15].

$$\vec{D}(t + \Delta t) = \vec{D}(t) - M\Delta t \cdot (\vec{D}(t) - \langle\vec{D}\rangle) + \sqrt{2\Gamma\Delta t}\,\vec{\eta} \quad (1)$$

Here $\langle\vec{D}\rangle$ denotes the average demand, Γ denotes the fluctuation strength, and $\vec{\eta}$ is a series of independent Gaussian random variables with zero mean and unit variance. The matrix M determines the correlation among fluctuations of different components of D. If M is isotropic (all of its eigenvalues are the same), the fluctuations of individual components of D will be decoupled.

Our approach will be as follows. To probe whether a given regulatory architecture successfully learns the statistical structure of the fluctuating environment, we expose it in simulations to several environmental epochs that differ by statistical structure, and allow the system to reach a steady state. To say that the system successfully "learns" its environment, we require two criteria. First, the system should be sensitive to the change of statistics, i.e. we expect the steady-state regulator activity to be reorganized between epochs. Second, we should be able to exhibit the "rule" by which the statistical feature of interest is encoded in the regulator activity.

The simplest form of statistical structure is the average demand $\langle\vec{D}\rangle$. This average demand can be learned already by the simple end-product inhibition (SEPI) circuit, where the production P_i of each metabolite x_i is placed under control of a single dedicated regulator a_i inhibited by x_i itself (Fig. 1A). In epochs of low demand, the unused x_i accumulates and decreases production until it balances the demand. The average demand over a recent past is stored in the activity of the regulator a_i (Fig. 1B, C).

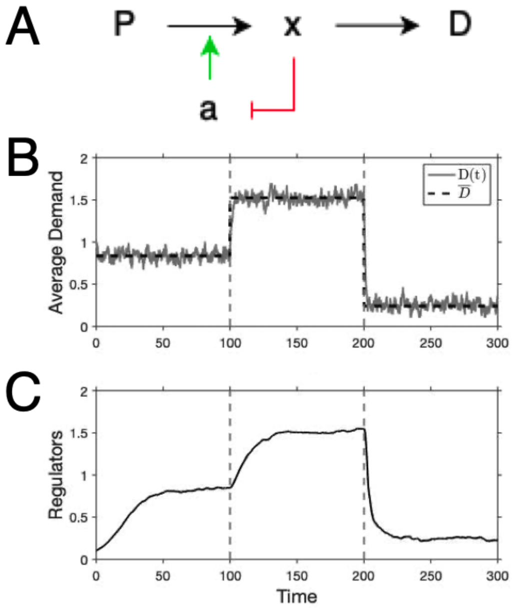

FIG. 1. The Simple end-product inhibition circuit (SEPI) can be seen as learning signal mean. (A) In the simple end-product inhibition architecture, the production P of a metabolite x is placed under control of a single dedicated regulator a inhibited by x itself. (B) We expose the SEPI architecture to 3 environmental epochs (solid line) that differ by the average demand $\langle\vec{D}\rangle$ (dashed line). (C) The expression level of the regulator a encodes the average demand.

The regulatory motif of end-product inhibition is not only simple, but has been shown to be remarkably effective. For example, under certain assumptions, this motif alone can not only "solve" the problem of proteome reallocation after a change of environmental conditions, but do so in an optimal time [19]. However, the effectiveness of SEPI necessarily applies only when dealing with states that, in the language of our model, differ by the signal mean. Indeed, at steady state, the internal degrees of freedom (the regulators a, serving as memory) can store only one value per metabolite x_i. To be sensitive to additional statistics, additional degrees of freedom would necessarily be required. (It is worth noting that real cells can transiently circumvent the "one regulator encodes one number" intuition, using temporally patterned signaling to transmit more information through a single regulator than one might naively expect [20]. However, even in this case, downstream readout circuitry must decode the information carried by dynamic features of the signal into steady-state expression levels of other genes, reducing to the same problem considered here.) Thus, from now on, we will allow the number of regulators N_a to exceed the number of metabolites N_x, and label regulators using Greek indices μ, running from 1 to N_a (while Roman indices i, labeling metabolites, run from 1 to N_x).

Landmann *et al.* showed that a generalized end-product feedback architecture can learn the variances and correlations of D_i [15]. Their architecture takes three ingredients: an excess of regulators ($N_a > N_x$), non-linear activation/repression of the regulators a_μ by the metabolite concentrations x_i, and cross-talk among different regulatory pathways. Specifically, they considered the following dynamics:

$$\begin{aligned} x_i &= \frac{P_i}{D_i} \\ P_i &= \sum_\mu \sigma_{\mu i} a_\mu \\ \tau_a \dot{a}_\mu &= a_\mu \max\Big(d, \sum_i \sigma_{\mu i}(1 - x_i)\Big) - \kappa a_\mu \end{aligned} \tag{2}$$

Here $\sigma_{\mu i}$ describes how the activities of regulators a_μ control the synthesis of metabolites x_i; d parameterizes nonlinearity; τ_a sets the timescale of regulator dynamics, and κ/τ_a is the degradation rate. (This form assumes timescale separation with dynamics of metabolites x_i being faster than that of regulators a_μ; for details, see derivation and discussion in Ref. [15].)

We should note that this simplified model ignores the intrinsic stochasticity of transcriptional regulation and is a poor representation of the complexity of real biological dynamics. Here, we use it as a model context to explore an abstract topic, namely the complexity of a task that a small number of simple regulatory elements can in principle achieve.

In two dimensions ($N_x = 2$), this cross-talk architecture can sense, store and usefully "recall" the information on second-order input statistics, such as variances and correlations, and do so near-optimally [15]. Here, we extend this architecture to higher dimensions (Fig. 2A). We choose regulators σ to be minimally redundant (see Supplemental Material [31] section B). Briefly, the elements of σ, normalized as $|\sigma| = 1$, can be seen as N_a points on an N_x-dimensional sphere, and we pick them to be spread out as far away from each other as possible by treating them as repelling charges on a sphere's surface (Fig. 2B). The charge-repelling procedure is initialized with random initial conditions, such that the residual rotational symmetry is fixed randomly in each simulation run. We expose the generalized end-product feedback architecture to environmental epochs that differ in fluctuation structure M only. To guarantee that any restructuring of regulator activity between exposure epochs is due to the changes of M, we keep the mean demand $\bar{D}$ the same in all epochs. For concreteness, we pick M to be a random rotation of $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 100 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 100 \end{pmatrix}$, so that the environmental fluctuations have one preferred direction (the direction of the dominant eigenvector of M^{-1}), where the restoring force is the weakest, and the fluctuations are thus the largest. When shifting from one epoch to another, we reorient M by applying a random rotation (Fig. 2C), and observe how the regulator expression levels reorganize in response (Fig. 2D).

Note that changing the direction of the dominant eigenvector is only one way to change the environment. For example, Ref. [15] also considered environments with different extent of correlation among fluctuations of D_i's, which the cross-talk architecture was also able to learn. Here, we will use the former approach, because the existence of a preferred direction of fluctuations allows for more intuitive metrics quantifying circuit responsiveness, as we describe below. We will show that the architecture of Fig. 2A is indeed responsive to higher-dimensional rotations of M, and will quantify this sensitivity.

III. RESULTS

A. The generalized architecture can outperform SEPI, but is costly

When Landmann *et al.* considered the cross-talk architecture (Eq 2) in the two-dimensional case, their focus was not just learning, but also the benefit of learning. Specifically, the primary readout used in that work was the 'tracking performance' $\mathscr{P}$ defined by $\mathscr{P} \equiv -\sqrt{\sum_i (P_i - D_i)^2}$, intended as a proxy for organism fitness, and they showed that the learning-capable circuit can enhance tracking performance over SEPI. In this section, we demonstrate that this observation continues to hold in higher dimensions: namely, the cross-talk architecture of Fig. 2A can achieve better tracking performance $\mathscr{P}$ than the SEPI architecture. However, we will also show that this performance increase is very costly.

We will use two metrics of cost. One is Control Input Power (CIP), a concept rooted in control theory and defined here by $\int ||\dot{P}||^2 dt$. Measuring cost in this way has the advantage that the family of optimal strategies on the performance-CIP plane can be derived analytically, but CIP is difficult to interpret in biological terms. For this reason, we will also consider a more biologically relevant measure of cost, namely the total expression of all regulators combined: $\sum_\mu a_\mu$.

Fig. 3 confirms that both cost metrics yield similar results. As expected, increasing the number of regulators increases performance $\mathscr{P}$. Performance can also be improved by reducing the degradation rate κ, since a higher expression of regulators (with activators and repressors active simultaneously, known as paradoxical regulation [21]) allows P_i to change faster [15]. As a result, the performance $\mathscr{P}$ of the cross-talk architecture can ex-

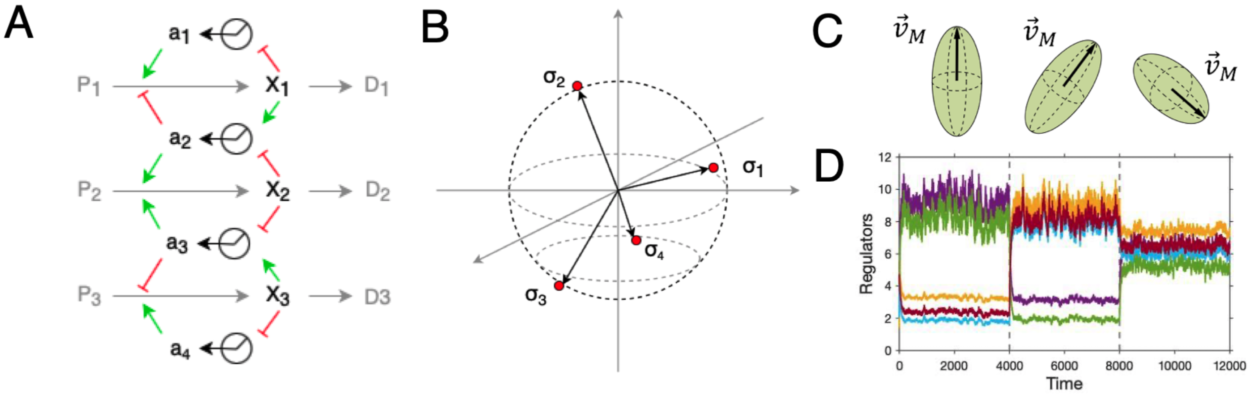

FIG. 2. The generalized architecture is built to learn the variances and correlations of demand $\vec{D}$. (A) An example of generalized end-product feedback architecture in the case where $N_x = 3$ and $N_a = 4$. (B) 4 regulatory pathways on 3 resources are shown as 4 vectors in 3d space. Regulation vectors are simulated as charges repelling each other to distribute evenly. (C) In this work, we model environmental fluctuations as $\vec{D}$ executing a random walk in a quadratic potential M. Such fluctuation structure can be visually represented as an ellipsoid, indicating the volume where a particle executing such random walk would typically be found: each axis of the ellipsoid is the direction of the corresponding eigenvector of M, and the length of each semiaxis is the inverse of the eigenvalue. The direction with the weakest restoring force (the dominant eigenvector of M^{-1}) is the direction of largest fluctuations, and is denoted $\vec{v}_M$. To test the ability of the system to learn, we expose it consecutively to environmental epochs differing in the fluctuation structure M, visually represented here as ellipsoids of different orientations. (D) In response to changes of M, the system dynamically adjusts the expression level of regulators.

ceed that of SEPI. However, Fig. 3 also shows the remarkable effectiveness of SEPI, which lies closest to the optimal curve at minimal circuit complexity.

The ability to invest resources into improving precision at an important task can be useful; such mechanisms are known to be employed by cells in other contexts, e.g. investing energy to improve the accuracy of sensing [22, 23] or copying its DNA [24]. However, in our context it seems implausible that this marginal performance increase alone would be sufficient to offset the cost of a significant increase in protein expression levels and circuit complexity, particularly since our simple model ignores the intrinsic stochasticity of regulatory dynamics, known to be substantial. Thus, from here on, we will no longer consider tracking performance $\mathscr{P}$ as our readout. Instead, we will assume that an ability to sense subtle changes in environmental statistics may itself be of value to the organism (e.g. as an early cue indicative of some other upcoming change), and investigate the ability of this cross-talk architecture to learn the environmental state and react to its changes.

B. The regulatory state adopted by the cross-talk architecture tracks the dominant eigenvector of the fluctuation structure M^{-1}

Fig. 2C, D provides an example showing that the cross-talk architecture is sensitive to the statistics of environmental fluctuations. After an environmental change, when v_M is reoriented, the expression level of regulators is seen to change as well. We will now show that the regulation state adopted by the system is not random or idiosyncratic, but encodes information about the fluctuation structure M in a simple way. We follow Landmann *et al.* to define system *responsiveness* $R_{ij} = \frac{d\dot{P}_i}{dD_j}$, where $\dot{P}_i$ represents the time derivative of production P_i (see Eq. (2)), and will show that the system preferentially aligns its eigenvectors to the dominant eigenvector of M^{-1}. This behavior is, in fact, the "smart" thing to do: it can be shown that the optimal strategy (in the sense of control theory, with CIP as cost metric) would be similarly anisotropic, with the dominant direction of fluctuations eliciting the strongest response [15, 25]. (This analytical result is what motivates defining R_{ij} as above.)

To quantify this degree of alignment, we define

$$\gamma = \mathrm{Prob}_{\|u\|=1}\Big(\|R\cdot\vec{u}\| < \|R\cdot\vec{v}_M\|\Big)$$

where $\vec{v}_M$ denotes the dominant eigenvector of M^{-1}, normalized to unit length. The intuition behind this definition is as follows. By definition of the responsiveness matrix R, the norm $||R\cdot\vec{v}_M||$ is the strength of the system's response following a fluctuation of D in the direction $\vec{v}_M$. Thus, γ quantifies how unlikely it is that a randomly drawn unitary vector $\vec{u}$ would elicit a response as strong, or stronger, than v_M. A large γ indicates the projection of $\vec{v}_M$ on the dominant eigenvectors of R is atypically large, and we will colloquially refer to this quantity as measuring "alignment of $\vec{v}_M$ to R". The larger the value of γ, the stronger the evidence that our generalized end-product feedback architecture adopted a state with a responsiveness matrix preferentially aligned to v_M. Heatmap of γ (Fig. 4A) demonstrates that the cross-talk architecture succeeds at aligning its responsiveness with the input fluctuations, and that just two extra regulators are sufficient to enable such alignment.

To evaluate the cost of such a regulatory strategy, Fig. 4B shows the correspondence between the system's ability to align with the dominant direction of fluctuations (v_M), and the control input power it incurs. The analogous plot for the second measure of cost, total expression level of regulators, looks similar and is shown in the Supplemental Material [31] (see Fig. S2).

Curiously, in contrast to Fig. 3A, Fig. 4B shows that the learning ability of the cross-talk architecture is controlled primarily by the number of regulators, and can be modulated without incurring an expression or CIP cost. One may also notice that the dependence of the incurred

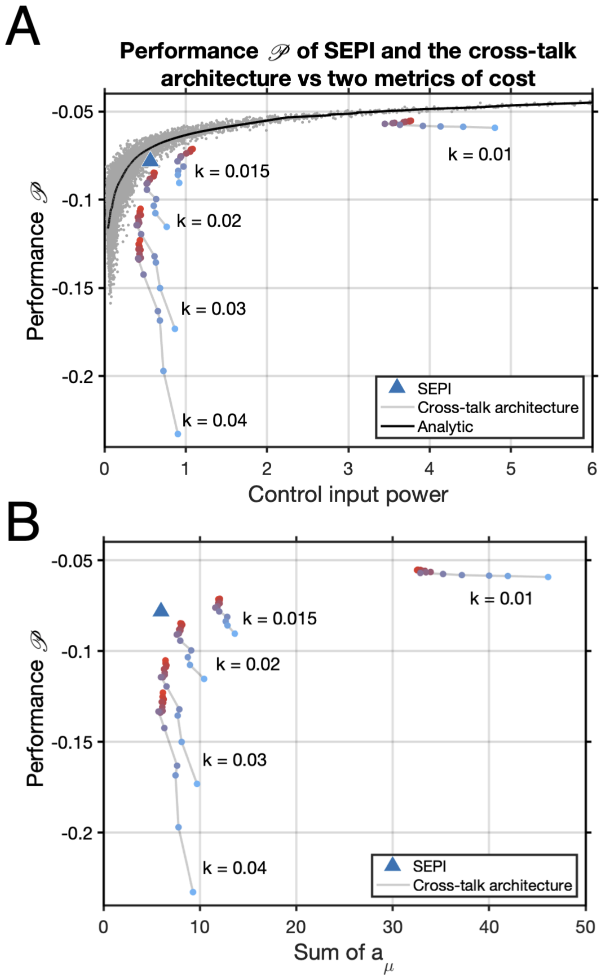

FIG. 3. Improving performance $\mathscr{P}$ beyond SEPI is very costly. (A) Tracking performance of different architectures when $N_x = 6$, shown against Control Input Power (CIP), which is a measure of cost (see text). Gray dots show simulation results of the analytically-derived optimal strategy (defined in [15]), averaged over 10 replicate simulations (see Supplemental Material [31] for more details); the black curve is a smoothed guide for the eye. The blue triangle indicates performance of SEPI, and the grey lines show the performance of the cross-talk architecture for different values of the degradation constant κ. The dots' color changing from blue to red indicates an increasing number of regulators from 6 to 20. (B) Same as A, replotted using a more biologically relevant measure of cost (the total expression of all regulators $\sum_\mu a_\mu$).

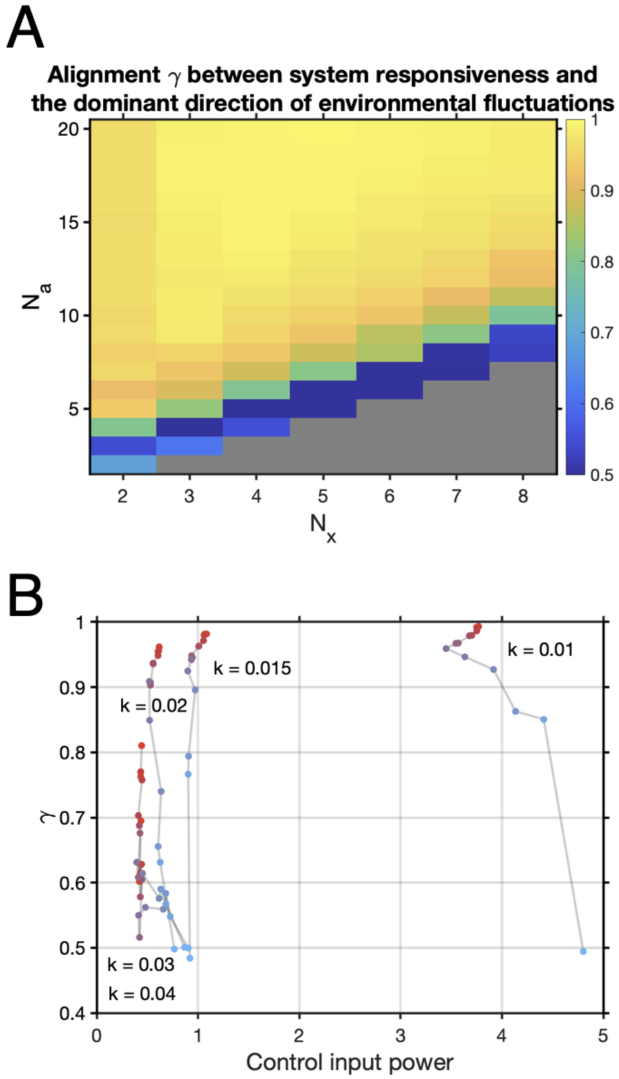

FIG. 4. A modest excess of regulators allows the cross-talk architecture to learn the dominant direction of environmental fluctuations. (A) Heatmap of the alignment γ between responsiveness R and the dominant eigenvector of M^{-1}. The panel shows that even a small excess of regulators is sufficient for the cross-talk architecture to track the dominant eigenvector of the fluctuations of its input. The $N_a < N_x$ region is grayed out, as we require the number of regulators N_a to be at least as large as the number of inputs N_x. (B) More regulators enable better alignmnet without incurring extra cost of control. Panel shows the alignment γ for $N_x = 6$ and varying κ and N_a, plotted against CIP. Dots changing from blue to red indicate the number of regulators increasing from 6 to 20.

cost, by either cost metric, on the number of regulators N_a is non-monotonic. Intuitively, a larger N_a means a more thorough tiling of the sphere of possible directions (Fig. 2B), so that any given direction in fluctuation space can be approximated more efficiently. However, if N_a is too large, this benefit is undermined by the fact that the expression of *any* regulator is positive and has non-zero fluctuations, incurring some cost even if “unused.”

C. The architecture also tracks non-dominant statistics

So far we have been considering the case where M^{-1} has only one dominant direction, but what happens when the fluctuation structure is more complex? To test this, we consider M with a sub-dominant eigenvector—specifically, a randomly rotated version of $M_0 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 100 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 100 \end{pmatrix}$—and restrict environment changes to those that change only the sub-dominant direction, keeping the dominant eigenvector fixed. For example, in a three-dimensional case, one can intuitively think of this as rotating an anisotropic ellipsoid around its dominant axis (Fig. 5A).

Fig. 5B confirms that in this regime, the alignment of system responsiveness R to the dominant eigenvector of M^{-1} remains significant (better than random), but becomes worse than we observed in Fig. 4A. This is, of course, expected: the structure of fluctuations no longer reduces to a single dominant direction. To fully assess the alignment between M and the responsiveness R, looking at only the dominant eigenvector is insufficient. To also take non-dominant eigenvectors of M into consideration, consider the quantity ϕ defined as

$$\phi_{[M,R]} = \frac{\|[M,R]\|}{\|\{M,R\}\|},$$

where $\|\ldots\|$ denotes the Frobenius norm of a matrix [26], $[M,R]$ denotes the commutator of M and R and $\{M,R\}$ denotes their anti-commutator. Note that if M and R are jointly diagonalizable (share the eigenbasis), then M and R commute and $\phi_{[M,R]}$ would be 0. Thus, a non-zero $\phi_{[M,R]}$ can be seen as a measure of misalignment between the eigenvectors of M and R. Dividing $[M,R]$ by $\{M,R\}$ yields a quantity invariant under rescaling of M or R by constant factors.

Similar to the trick we used when defining γ, to enable meaningful comparisons across dimensions, instead of focusing on the raw value of ϕ, we compute the probability

$$\varphi = \mathrm{Prob}_{R'}(\phi_{[M,R']} > \phi_{[M,R]}).$$

Here R' is a random ‘subtle rotation’ of R, which we define as a rotation that preserves its dominant eigenvector. We refer to φ as the alignment of non-dominant eigenvectors. If φ is close to 1, it means that M is better aligned to R than to almost any of its rotated versions, even when the rotations only change subdominant eigenvectors.

The results are shown in Fig. 5C. The $N_x = 2$ column is grayed out: in two dimensions, the only rotation that preserves the dominant eigenvector is the identity matrix. We see that the cross-talk architecture (Eq. 2) is not only responsive to changes in the subdominant direction of fluctuations, but succeeds at realigning its responsiveness matrix accordingly, particularly in dimensions 3 and 4 (at the price of a somewhat worse alignment to the dominant direction; compare with panel B). However, as the dimension N_x increases, the number of regulators required to achieve a good non-dominant alignment $\varphi > 0.9$ quickly becomes unreasonably large ($N_a > 20$).

IV. DISCUSSION

The regulatory circuit we considered in this work generalizes simple end-product inhibition by including three additional ingredients: nonlinearity, an excess of regulators, and cross-talk between them. Previous work has shown that these ingredients can endow the circuit with an ability to learn time-dependent fluctuation statistics of its inputs through a form of associative learning, at least in the low-dimensional scenarios (with one or two inputs) [15]. Here, we generalized this circuit to the higher-dimensional case and presented two results. Just like in lower dimensions, this architecture can show an improved performance $\mathscr{P}$ at the task of tracking environmental fluctuations. This small performance gain comes at a significant complexity cost. However, if sensing changes in environmental statistics is of value to the organism, then this architecture is quite interesting as it offers a sensitivity to subtle changes, sensing not only the dominant direction of fluctuations, but also the subdominant fluctuation modes.

How relevant is this high-dimensional case for real cells? It is easy to imagine that a specific pair of resources might be correlated at some point of an organism’s lifecycle but not at another; thus, it is clear that dynamically learning and unlearning a *single* correlation could be useful. By comparison, the “universal learning” capacity described here—an ability to pick out any preferred direction of correlations in a high-dimensional space—seems rather more abstract. However, this could be the relevant regime for the regulatory architecture of a cell as a whole, seen as a high-dimensional learning circuit. Some intriguing recent ideas propose a possible common ground between evolvability of regulatory circuits, their ability to solve complex problems, and the success of overparameterized models in machine learning [27, 28]. So far, these parallels remain speculative; but the fact that simple elements can enable regulatory circuits to perform a form of associative learning could be a valuable piece of this puzzle. In this work, we have shown that one previously proposed mechanism successfully generalizes to higher dimensions.

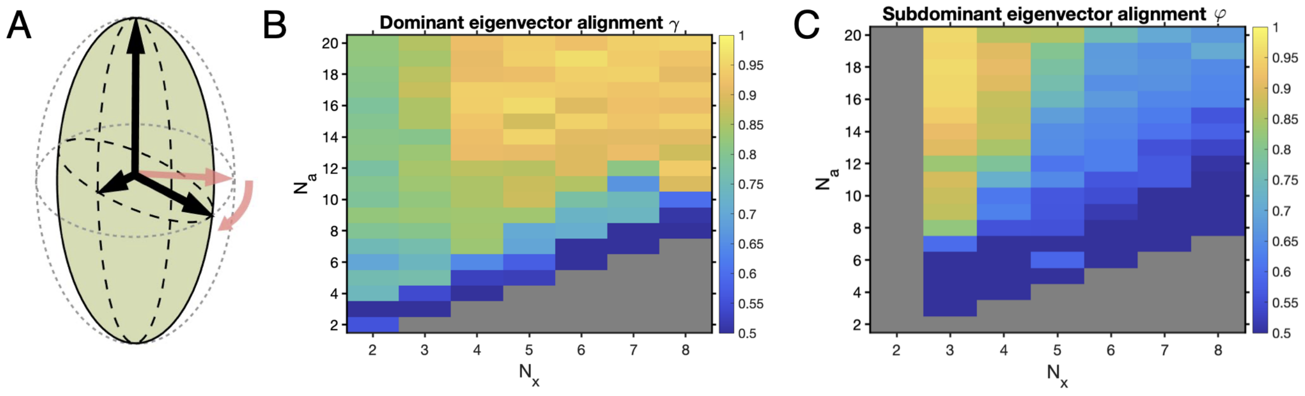

FIG. 5. The generalized architecture learns to track the changing sub-dominant eigenvector of the input fluctuation. (A) We rotate M with its dominant eigenvector fixed. Note that such rotations exist only in dimension 3 and above. (B) Heatmap of the alignment γ between responsiveness R and the dominant eigenvector of M. The alignment remains significant, indicating the system continues to track the dominant eigenvector. (C) Additionally, the system is sensitive to the sub-dominant fluctuation directions, as indicated by this heatmap of the alignment of non-dominant eigenvectors $\varphi(R, M)$ (see text), which is better than random for all allowed values of N_x ($N_x \geq 3$; see panel A).

In considering the biological implications of our analysis, it is important to remember that our goal was not to accurately model real regulatory circuits. Rather, our simple model helps us explore the computational ability of a simple circuit composed of biologically plausible elements.

This level of abstraction has some advantages. Circuit elements similar to those discussed here appear also in several other contexts, such as organelle regulation [29, 30] or resource competition dynamics of an ecosystem (where a species' resource exploitation inhibits its own growth). In fact, in the ecological context, the ingredients required by the cross-talk architecture are arguably more natural. Indeed, in the gene regulatory context, the autocatalytic aspect is nontrivial to implement [15], while in the ecological context, the "self-activation" is automatically ensured by replicator dynamics. This parallel suggests that under some minimal assumptions, the commonly considered resource competition dynamics may cause the ecosystem to align its responsiveness to the spectrum of environmental fluctuations. Confirming this behavior, and exploring its relevance for ecosystem dynamics is an exciting avenue for future work.

At the same time, the level of abstraction adopted here also has clear limitations. As mentioned above, our simplified model ignores the intrinsic stochasticity of transcriptional regulation. Thus, any interpretation in terms of real regulatory circuits requires caution.

All simulations were performed using MATLAB (Mathworks, Inc). The code reproducing all figures from scratch is provided as Supplementary File 1.

[1] I. Tagkopoulos, Y.-C. Liu, and S. Tavazoie, Predictive behavior within microbial genetic networks, Science 320, 1313 (2008).

[2] M. A. Savageau, Escherichia coli habitats, cell types, and molecular mechanisms of gene control, Am. Nat. 122, 732 (1983).

[3] M. Sims, Many paths to anticipatory behavior: Anticipatory model acquisition across phylogenetic and ontogenetic timescales, Biol. Theory 18, 114 (2023).

[4] S. J. Gershman, The molecular memory code and synaptic plasticity: A synthesis, Biosystems 224, 104825 (2023).

[5] J. P. Dexter, S. Prabakaran, and J. Gunawardena, A complex hierarchy of avoidance behaviors in a single-cell eukaryote, Curr. Biol. 29, 4323 (2019).

[6] S. J. Gershman, P. E. Balbi, C. R. Gallistel, and J. Gunawardena, Reconsidering the evidence for learning in single cells, eLife 10, e61907 (2021).

[7] D. Bell-Pedersen, V. M. Cassone, D. J. Earnest, S. S. Golden, P. E. Hardin, T. L. Thomas, and M. J. Zoran, Circadian rhythms from multiple oscillators: lessons from diverse organisms, Nat. Rev. Genet. 6, 544 (2005).

[8] A. M. Turing, On Computable Numbers, with an Application to the Entscheidungsproblem, Proc. Lond. Math. Soc. 42, 230 (1936).

[9] W. S. Brainerd and L. H. Landweber, Theory of Computation (Wiley, 1974).

[10] J. Pérez, J. Marinković, and P. Barceló, On the turing completeness of modern neural network architectures, arXiv:1901.03429 (2019).

[11] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle, Robust perfect adaptation in bacterial chemotaxis through integral feedback control, Proc. Natl. Acad. Sci. U.S.A. 97, 4649 (2000).

[12] H. Chang and A. Levchenko, Adaptive molecular networks controlling chemotactic migration: dynamic inputs

and selection of the network architecture, Philos. Trans. R. Soc. B Biol. Sci. 368, 20130117 (2013).

[13] 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).

[14] M. Sorek, N. Q. Balaban, and Y. Loewenstein, Stochasticity, bistability and the wisdom of crowds: a model for associative learning in genetic regulatory networks, PLoS Comput. Biol. 9, e1003179 (2013).

[15] S. Landmann, C. M. Holmes, and M. Tikhonov, A simple regulatory architecture allows learning the statistical structure of a changing environment, eLife 10, e67455 (2021).

[16] J. G. Sanders, H. Akl, S. J. Hagen, and B. Xue, Crosstalk enables mutual activation of coupled quorum sensing pathways through "jump-start" and "push-start" mechanisms, Sci. Rep. 13, 19230 (2023).

[17] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, Statistical mechanics for natural flocks of birds, Proc. Natl. Acad. Sci. U.S.A. 109, 4786 (2012).

[18] M. Aldana, H. Larralde, and B. Vázquez, On the emergence of collective order in swarming systems: a recent debate, Int. J. Mod. Phys. B 23, 3661 (2009).

[19] M. Y. Pavlov and M. Ehrenberg, Optimal control of gene expression for fast proteome adaptation to environmental change, Proc. Natl. Acad. Sci. U.S.A. 110, 20527 (2013).

[20] A. S. Hansen and E. K. O'Shea, Limits on information transduction through amplitude and frequency regulation of transcription factor activity, eLife 4, e06559 (2015).

[21] Y. Hart, Y. E. Antebi, A. E. Mayo, N. Friedman, and U. Alon, Design principles of cell circuits with paradoxical components, Proc. Natl. Acad. Sci. U.S.A. 109, 8346 (2012).

[22] S. B. Laughlin, Energy as a constraint on the coding and processing of sensory information, Curr. Opin. Neurobiol. 11, 475 (2001).

[23] 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).

[24] A. C. Olson, J. N. Patro, M. Urban, and R. D. Kuchta, The energetic difference between synthesis of correct and incorrect base pairs accounts for highly accurate DNA replication, J. Am. Chem. Soc. 135, 1205 (2013).

[25] D. Liberzon, Calculus of Variations and Optimal Control Theory: A Concise Introduction (Princeton University Press, 2011).

[26] C. F. Van Loan and G. Golub, Matrix computations (Johns Hopkins Studies in Mathematical Sciences, 1996).

[27] O. Howell, S. Patterson, J. F. Green, and C. J. Lane, Machine learning as ecology, J. Phys. A: Math. Theor. 53, 334001 (2020).

[28] J. W. Rocks and P. Mehta, The geometry of overparameterized regression and adversarial perturbations, arXiv:2103.14108 (2021).

[29] K. P. Amiri, A. Kalish, and S. Mukherji. Robustness and universality in organelle size control. Physical review letters, 130(1), 018401 (2023).

[30] W.F. Marshall. The flagellar length control system: exploring the physical biology of organelle size. Physical biology, 20(2), 10.1088/1478-3975/acb18d (2023).

[31] See Supplemental Material at [URL will be inserted by publisher] for more details on the model, simulation parameters, and figures using an alternative definition of expression 'cost'.