PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY
DEVELOPMENTAL BIOLOGY OPEN ACCESS

Invariant nonequilibrium dynamics in gene regulation optimize information flow

Benjamin Zoller[a], Alexis Bénichou[b], Thomas Gregor[a,c], and Gašper Tkačik[b,1]

Affiliations are included on p. 11.

Edited by Rosa Martinez-Corral, Centre de Regulacio Genomica, Barcelona, Spain; received September 8, 2025; accepted June 5, 2026 by Editorial Board Member Mehran Kardar

Eukaryotic gene regulation relies on stochastic yet controlled promoter switching, in which genes transition between transcriptionally active and inactive states. Despite the molecular complexity of this process, recent studies have revealed a surprising invariance of the "switching correlation time" (T_C)—the characteristic decay time of the autocorrelation function of promoter activity fluctuations—across gene expression levels in multiple genes and organisms. A biophysically plausible explanation for this invariance has so far been lacking. Here, we show that this empirical constraint imposes stringent requirements on minimal yet realistic models of transcriptional regulation. Specifically, reproducing T_C–invariance requires regulatory architectures with at least four internal states and nonequilibrium dynamics that break detailed balance. Using Bayesian inference on *Drosophila* gap gene expression data, we demonstrate that such models i) quantitatively reproduce the observed T_C–invariance, ii) remain robust to parameter perturbations, and iii) maximize information transmission from transcription factor concentration to gene expression. Remarkably, the T_C-invariant modulation strategy we identify as optimal closely parallels contemporary control-theoretic results on the modulation of stochastic switching systems. Taken together, our results suggest that eukaryotic transcriptional regulation operates in a nonequilibrium regime to balance precision, reaction-rate limitations, and energy dissipation, thereby achieving near-optimal information transmission under fundamental physical constraints.

transcriptional regulation | promoter switching | Markov models | information theory

Significance

Eukaryotic genes switch on and off stochastically, yet cells regulate gene expression with remarkable precision by biasing the probability of switching. Recent experiments revealed a surprising and unexplained constraint: The characteristic timescale of promoter switching remains nearly constant across expression levels, genes, and perturbations. Here we show that this invariance places strong limits on models of gene regulation. Regulatory architectures consistent with the data must operate out of thermodynamic equilibrium, cycle through at least four internal states, and dissipate chemical energy. The inferred models quantitatively reproduce high-precision fruit fly measurements. This invariance is a signature of optimal information transmission during gene expression control, suggesting that evolution favors regulatory strategies that are precise, robust, and energy-aware.

Transcriptional regulation enables cells to control gene expression in response to environmental and developmental cues. At the single-cell level, this process is inherently stochastic, with transcription occurring in random bursts as promoters switch between inactive and active states (1, 2). The frequency and duration of these bursts are influenced by transcription factors (TFs), cis-regulatory elements (e.g., enhancers), and chromatin accessibility, shaping gene expression dynamics in ways essential for cellular function but challenging to analyze (3, 4). Advances in live-cell imaging and single-molecule tracking have provided unprecedented insight into these dynamics, enabling increasingly stringent tests of canonical mathematical models of transcription, such as the "telegraph model" (5), and motivating the development of new, mechanistically grounded alternatives (6–11).

Recent precise measurements of gene expression during early *Drosophila* development enabled stringent quantitative tests of the effective two-state telegraph model of transcription. In this model, the gene switches stochastically between ON periods (of average duration T_{ON}), during which mRNA is produced at a constant rate, and OFF periods (of average duration T_{OFF}) during which transcription is inactive. Surprisingly, both inference from static measurements (12) and direct estimates from time-resolved data (10) consistently revealed that the sum of the reciprocal mean dwell times in ON and OFF states, $T_C^{-1} = T_{OFF}^{-1} + T_{ON}^{-1}$, remains approximately constant across gene activity levels, a phenomenon we refer to as T_C–invariance.

Whenever promoter activity can be described as stochastically switching between two observable states (ON and OFF) in a memoryless (Markovian) fashion, with exponentially distributed dwell times, T_C can be interpreted as the decay timescale of the autocorrelation function of the binary promoter-state time series. More generally, T_C is a key dynamical observable that characterizes the timescale over which promoters fluctuate between ON and OFF states. It is related to the relaxation timescale toward steady state and the window over which gene expression fluctuations can be effectively averaged or buffered (13–15). Further analyses of published data suggested that this invariance is conserved across multiple genes and organisms (10).

The authors declare no competing interest.

This article is a PNAS Direct Submission. R.M.-C. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: gtkacik@ist.ac.at.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2524855123/-/DCSupplemental.

Published July 6, 2026.

The T_C–invariance is unexpected for a number of reasons. First, transcriptional regulation involves many molecular species coupled through a large number of chemical reactions driven by time-dependent inputs. One might therefore expect that quantitative agreement with data requires comparably complex models. In contrast, precise measurements have shown that simple, coarse-grained descriptions can be successful despite discarding most molecular detail. In particular, the two-state "telegraph model" (1, 5) provides an excellent effective phenomenological description in several settings.

For *Drosophila* gap genes, the telegraph model not only accounts well for the static features of the data (12), but also captures key dynamical aspects (10). Notably, the autocorrelation function of promoter state switching can be reliably estimated and is dominated by a single exponential decay with characteristic timescale T_C, once measurement systematics are taken into account. In contrast, dwell time distributions in the ON and OFF states are much less diagnostic in this system, as gap gene expression is not truly in steady state; nevertheless, for times $t \gtrsim T_C$, they are consistent with exponential behavior.

Taken together, these observations suggest that an effective two-state model is remarkably predictive, raising the theoretical question of why such a minimal model succeeds so well despite the underlying molecular complexity.

Second, even if such a two-state approximation holds, explaining T_C–invariance across expression levels requires a coordinated dependence of reaction rates on the regulatory input. This coordination lacks an obvious mechanistic basis (2). Within the telegraph model framework, regulation is often described as modulating either the burst frequency (T_{OFF}) or burst duration (T_{ON}) (16); in particular, burst frequency modulation has been supported experimentally (17–20). However, a constant T_C is inconsistent with exclusive modulation of either T_{ON} or T_{OFF}, and instead requires a distinct mode in which both timescales adjust in a coordinated manner to achieve a target expression level.

Third, standard biochemical models of transcription do not predict such invariance, raising the questions of *how* it can arise mechanistically and *why* it might be selectively favored for biological regulation.

Unlike many nonliving systems, whose macroscopic behavior is often well captured by typical parameter values or average responses, biological regulatory systems are shaped by evolutionary selection. Selection can bias evolutionary outcomes toward narrow regions of parameter space and favor architectures that realize specific functional phenotypes, even when these architectures are rare under neutral sampling (21). Evolved regulatory systems can therefore be driven far from "typical behaviors" expected from various ensembles of random or generic biochemical reaction networks. This distinction between neutral ensembles and selected ensembles has been formalized at the interface of evolutionary theory, statistical physics, and information theory (22, 23).

These considerations imply that understanding the origins and implications of the observed T_C–invariance requires examining not only the molecular mechanisms that regulate transcription (the "how"), but also the potential functional benefits for which such an invariance could have been selected (the "why"). Therefore, in this manuscript, we first address the mechanistic constraints imposed by T_C–invariance, and then turn to its functional interpretation.

Recent studies have begun to integrate mechanistic models inferred from data with theoretical frameworks that describe regulatory architectures in terms of optimization principles, quantitative phenotypes, and physical constraints (24–26). In this context, we view transcriptional regulation as balancing trade-offs between molecular resource availability, reaction speeds, and energy dissipation (27–29) to achieve rapid yet accurate gene expression control. Such trade-offs may necessitate energy dissipation and favor nonequilibrium regulatory architectures.

Theoretical models suggest that nonequilibrium dynamics can enhance regulatory precision and robustness, and mitigate deleterious effects of conflicting signals, such as those arising from crosstalk (25, 29–35). Yet, whether eukaryotic transcription actively exploits these advantages remains an open empirical question (36–40). Similarly, it is unclear whether T_C–invariance necessarily implies nonequilibrium function, or whether the two properties are causally linked.

Here, we identify the minimal mechanistic model required to reproduce T_C–invariance and reveal its functional significance. The resulting model takes the form of a regulatory cycle, for which a constant T_C necessarily implies irreversible dynamics operating out of thermodynamic equilibrium. Mechanistically, this regime corresponds to a balance between enhancer-mediated facilitation and stabilization of transcription, which together maintain invariant switching dynamics at the promoter. We show that T_C–invariance is strongly nongeneric and instead emerges as a signature of selection for near-maximal information flow under constraints on promoter switching speed.

Results

Empirical Constraints for Models of Transcriptional Regulation. Motivated by the surprising invariance of T_C and its potential functional significance, we aim to identify the simplest regulatory architectures that can account for this phenomenon. These architectures should be consistent with three core experimental findings, which we summarize below (Fig. 1*A*), and refer to as "empirical constraints" throughout the manuscript.

First, transcription occurs in bursts: Promoters stochastically alternate between active (ON) and inactive (OFF) states, with average durations T_{ON} and T_{OFF}, respectively. Second, the average promoter activity, $P_{ON} = T_{ON}/(T_{ON} + T_{OFF})$, is modulated by TF concentrations across a large dynamic range, $E = (\max P_{ON})/(\min P_{ON})$. Third, the switching correlation time, $T_C = T_{OFF}P_{ON} \sim 1.5$ min remains approximately constant over the full range of P_{ON} (10), and is typically much longer than the TF residence time on the DNA (3).

Whether the separation between T_C and the typical TF residence time, b^{-1}, holds generally across eukaryotic TFs remains an open question. Recent analyses suggest that specific TF binding events can persist for ∼200 s for certain TFs (41), whereas measurements in the *Drosophila* gap gene system, directly relevant here, report specific residence times for Bicoid (the primary TF activator of the gap genes) below 10 s (42). In this work, we adopt the latter regime and focus on models in which TF binding/unbinding is fast relative to promoter switching.

The third constraint—the T_C–invariance—is particularly striking. Unlike a bound on a parameter, it specifies an entire functional relationship between T_C and P_{ON} that holds robustly across conditions and many genetic perturbations. Published models currently neither explain the mechanistic origin of this invariance, nor do they clarify the conditions under which it can arise.

To evaluate a wide range of candidate models consistent with the empirical constraints (Fig. 1*A*), we represent transcriptional regulation as a continuous-time discrete-state Markov process.

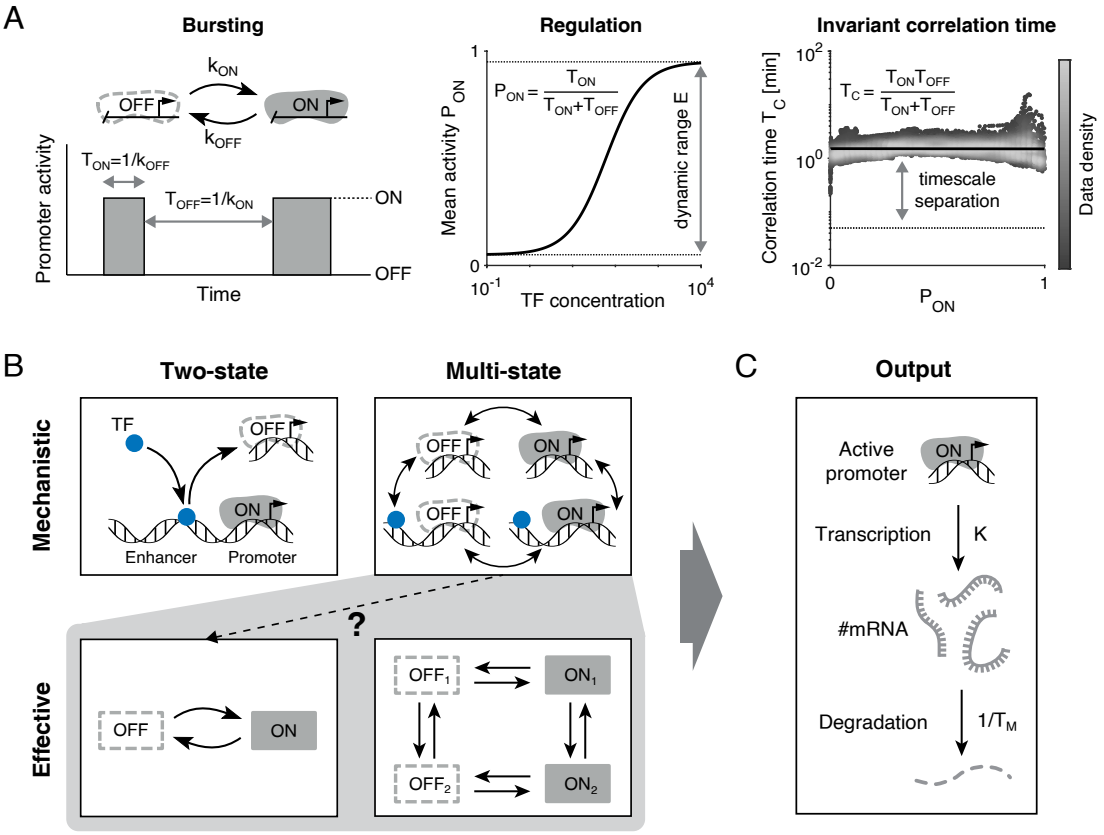

Fig. 1. Empirical constraints inform the structure of minimal transcriptional regulatory models. (*A*) Experimental constraints include: *i)* bursting, i.e., stochastic switching between transcriptionally inactive (OFF) and active (ON) promoter states, with mean durations T_{OFF} and T_{ON} (*Left*); *ii)* regulation, quantified by the induction curve relating the mean activity $P_{ON} = T_{ON}/(T_{ON} + T_{OFF})$ to TF concentration, spanning a large dynamic range E (*Middle*); and *iii)* a nearly constant promoter switching correlation time $T_C = T_{OFF}P_{ON}$, invariant across expression levels and significantly slower than the typical (dotted line) TF residence time (*Right*). (*B*) Conceptual classification of regulatory models: two-state vs. multistate (horizontal), and mechanistic vs. effective (vertical). Mechanistic two-state models (*Top Left*) cannot reproduce invariant T_C [panel (*A*), *Right*]. Effective two-state models (*Bottom Left*) accurately describe data; our goal is to find mechanistic multistate models (*Top Right*) that coarse-grain to such effective models and reproduce T_C–invariance (dashed arrow). (*C*) ON/OFF promoter switching can be coupled to transcriptional output via a birth–death process, generating experimentally measurable mRNA distributions in single cells.

Our general approach is to consider both the "effective" models, in which the internal states remain abstract and switching rates can be modulated without restriction, and "mechanistic" models, in which the Markov states correspond to the microstates of a chemical reaction network and state transitions represent elementary reactions (Fig. 1*B*).

In either case, the system states can be partitioned into two disjoint sets corresponding to the ON and OFF observable macrostates of an effective (and possibly non-Markovian) two-state description, in which transcription does or does not take place. These macrostates are the only features that can be directly inferred from the data; the promoter activity P_{ON} corresponds to the stationary probability of occupying any transcriptionally active configuration, i.e., any microstate from which transcription can proceed (Fig. 1*C*). The mathematical mapping between microstates and such observable macrostates is called coarse-graining (or lumping). While coarse-graining is an active theoretical topic in its own right, it also provides a key link between data-driven phenomenology and mechanistic transcription models (40).

As our first step, we consider minimal models built from elementary reactions, such as unimolecular decays and bimolecular interactions (e.g., TF–DNA binding/unbinding; see *SI Appendix*, Fig. S1). These models can reproduce transcriptional bursts and tunable P_{ON}, but not an invariant T_C (*SI Appendix*, section 3.1). In these models only one rate (e.g., k_{ON} or k_{OFF}) depends on TF concentration. Achieving constant T_C would require finely tuned, anticorrelated adjustments to both $T_{ON} = 1/k_{OFF}$ and $T_{OFF} = 1/k_{ON}$ simultaneously, a scenario which is biologically implausible.

We then consider mechanistic multistate models, particularly those based on cooperative TF binding (31, 34), which introduce additional intermediate microstates (*SI Appendix*, Fig. S2; see *SI Appendix*, section 3.2). These models can approximate constant T_C over part of the expression range, but they fall short in two important ways. First, they predict $T_C = 1/b$, where b is the TF unbinding rate, thereby tying promoter switching to TF residence time on the DNA. This contradicts observations that switching occurs on a much slower, TF-independent timescale (31, 34). Second, they give rise to multiple promoter switching correlation timescales, while data support a single dominant T_C. In particular, as $(T_{ON}^{-1} + T_{OFF}^{-1})^{-1}$ becomes increasingly invariant in this model class, deviations of the promoter autocorrelation function from a single exponential become more pronounced. Thus, while T_C can be formally calculated, it lacks a clear physical meaning: It arises from a superposition of latent dynamics that cannot be faithfully coarse-grained into an effective two-state model.

To overcome these issues, we adopt a systematic, constraint-based approach. Rather than carefully hand-crafting a single model built out of plausible molecular interactions, we explore a broad space of multistate reaction networks to identify those consistent with empirical constraints (25). We organize these

models along two independent axes (Fig. 1*B*). Horizontally, we distinguish two-state from multistate models. Vertically, we separate mechanistic models—built from explicit chemical microstates—from effective models, which provide coarse-grained descriptions and abstract away molecular detail. Effective models can often be inferred from data but lack physical interpretability.

Previous studies show that two-state effective models provide excellent functional and statistical descriptions of transcriptional bursting. In the *Drosophila* gap gene system, in particular, the telegraph model accounted well for both the static and dynamic aspects of transcriptional regulation (10, 12). Thus, our goal is to identify minimal mechanistic multistate models that can be coarse-grained into effective two-state behavior while satisfying the T_C–invariance and other empirical constraints. In *SI Appendix*, section 2.2, we derive a quantitative criterion for when such coarse-graining succeeds in recapitulating both the static and dynamic features of multistate models, building on our own and previously published analytical work (15, 43, 44).

Finally, to relate regulatory mechanisms to measurable outputs, we couple the ON/OFF promoter dynamics to a simple birth–death process for mRNA production and degradation (Fig. 1*C*). This allows us to predict the full steady-state mRNA distribution in single cells for any TF input. More broadly, this framework enables connections between physical models and functional phenotypes, such as information flow (27, 45–48) and specificity (25, 29). It also bridges theory and experiment in a way that exposes the constraints and trade-offs of transcriptional regulation.

A Minimal Nonequilibrium Cycle Model Captures T_C–Invariance. To gain intuition, we started by looking for the simplest mechanistically plausible model built from elementary reactions that maintains a constant switching correlation time T_C (*SI Appendix*, Fig. S3*A*). We constructed this model incrementally. First, we identified two separate two-state models, each exhibiting identical T_C values, but differing in their mean promoter activity (high and low). Next, we coupled the two models by introducing a regulatory element whose TF binding drives transitions between the two (high and low) promoter activity regimes. These transitions are rapid and concentration-dependent, capturing TF association/dissociation dynamics. The resulting four-state mechanistic model allows tunable activity levels (*SI Appendix*, Fig. S3*B*), maintains an invariant switching correlation time (*SI Appendix*, Fig. S3*C*), and remains effectively two-state from a coarse-grained perspective (*SI Appendix*, Fig. S3*D* and section 2.2). It naturally separates two timescales: fast TF (un)binding and slow promoter switching dynamics (*SI Appendix*, Fig. S3*E*), a key feature observed experimentally.

To explore a broader range of behaviors compatible with T_C–invariance, we extended the resulting reaction scheme into a more versatile four-state "cycle model." This model can parametrically accommodate a broad space of behaviors, enabling both functional tuning (e.g., of the $T_C(P_{ON})$ curve) and thermodynamic control via dissipation. Similar cycle-based architectures have been proposed previously in the context of regulatory precision and nonequilibrium control (29, 33, 49, 50).

A possible mechanistic interpretation of the cycle model is shown in Fig. 2*A*: promoter activation and deactivation occur either basally (states 00 and 01, with rates q and r, respectively) or under the influence of an active enhancer (states 10 and 11). Enhancer activation is driven by TF binding (transitions between states 00 and 10, with a concentration-dependent rate f and unbinding rate b). Once the enhancer is active, it can both facilitate promoter activation (via $\alpha_1 > 1$) and stabilize transcriptional activity (via $\beta_1 > 1$), thereby promoting entry into, as well as residence in, the doubly active state 11.

The cycle model generates typical sigmoidal induction curves (*SI Appendix*, Fig. S4*A*), where the probability of being transcriptionally ON, $P_{ON} = P(01) + P(11)$, increases monotonically with the TF binding rate, f. The ON probability ranges from $P_{min} = q/(q + r)$ at $f = 0$ to $P_{max} = \alpha_1\beta_1 q/(\alpha_1\beta_1 q + r)$ as $f \to \infty$. To ensure a wide dynamic range of expression consistent with experimental observations (10), we define the range as $E = P_{max}/P_{min}$ and choose $E = 10^3$. This constraint implies specific relationships among parameters: $r = Eq$ and $\alpha_1\beta_1 = E^2$.

The model's reaction rates can be reparameterized to make the model more amenable to analysis and interpretation. Below, we introduce the new parameter set r, η, γ, and ν, which provides direct control over promoter responsiveness and nonequilibrium behavior (see also Fig. 2*A* and caption for definitions).

To relate enhancer facilitation and stabilization to the dynamic range E, we introduce a parameter $\eta \in [0, 1]$, with $\alpha_1 = E^{2\eta}$ and $\beta_1 = E^{2(1-\eta)}$. Intuitively, η governs the balance between two regulatory mechanisms: When $\eta = 0$, the system relies entirely on stabilization of the doubly active state 11 ($\alpha_1 = 1$ and $\beta_1 = E^2$), whereas $\eta = 1$ corresponds to pure facilitation of the transition $10 \to 11$ into the doubly active state ($\alpha_1 = E^2$ and $\beta_1 = 1$). Intermediate η values tune the relative contributions of each mechanism (*SI Appendix*, section 5.1).

The (non)equilibrium dynamics of the cycle model can be examined by analyzing the stationary probability currents. Detailed balance holds when $\alpha_1\beta_1 = \alpha_2\beta_2$, with cycle reversibility controlled by α_2 and β_2 (*SI Appendix*, section 5.3). To systematically explore departures from equilibrium, we reparameterize these rates using two new variables: $\alpha_2 = \alpha_1 \exp(\gamma + \nu)$ and $\beta_2 = \beta_1 \exp(\gamma - \nu)$. Here, $\gamma \in (-\infty, +\infty)$ controls the system's dissipation and directionality, while $\nu \in (-\infty, +\infty)$ sets the asymmetry between the two reaction branches. Together, these parameters allow a systematic exploration of departures from the equilibrium ($\gamma = 0$) regime.

In sum, the model's behavior depends on four free control parameters: r, η, γ, ν. Other parameters are fixed by empirical constraints. For example, the TF-dependent rate f must remain tunable to dial the mean promoter activity across the entire dynamic range, $E = 10^3$. TF unbinding rate b is fixed to $b = 60\,\mathrm{min}^{-1}$, in line with experimentally measured TF residence times on the DNA (3) (a ten-fold lower b leads to higher ν but leaves other inferred parameters as well as all our conclusions unchanged). b also sets the absolute timescales for the cycle model.

Fitting the cycle model to the empirical T_C function extracted from *Drosophila* transcriptional data (10) yields well-constrained parameters (Fig. 2*B*) (51). The best-fit parameter values indicate clear nonequilibrium behavior, with $\gamma < 0$. Systematic variation of the model's parameters shows that r primarily controls the scale of T_C (its average value), while η modulates its shape (Fig. 2*C*). Invariance of T_C—i.e., a constant, flat T_C as a function of P_{ON}—is only maintained within a narrow parameter range: $\eta \sim 0.5$ and $r \lesssim b$ (Fig. 2*D*). These constraints are also necessary for the observed timescale separation between TF binding and enhancer switching; they lead to a precise balance between enhancer-mediated facilitation and stabilization of the promoter. At $\eta = 0.5$, the enhancer mirrors the reverse

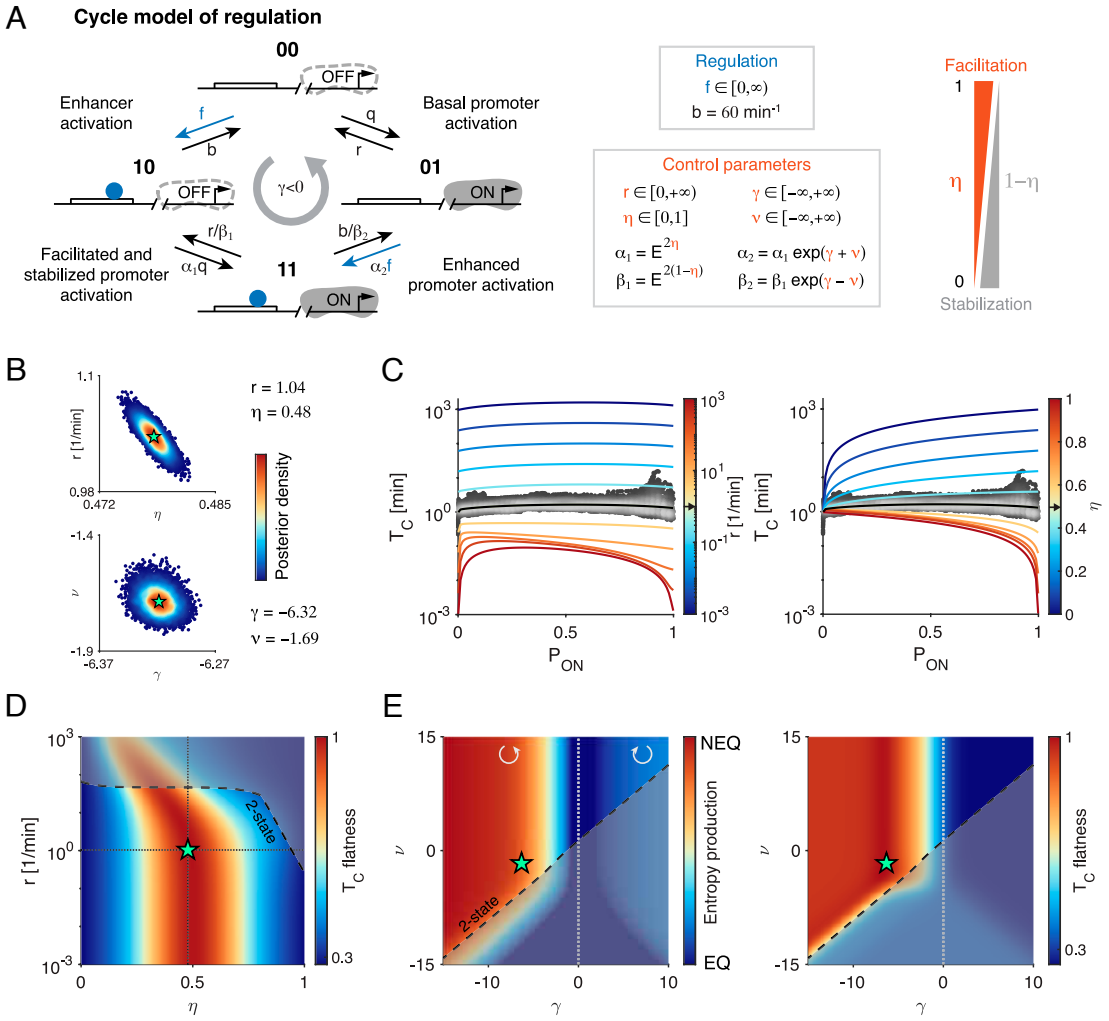

Fig. 2. A minimal nonequilibrium model captures T_C–invariance and reveals functional constraints. (*A*) Schematic of the four-state "cycle model" of transcriptional regulation. A possible mechanistic interpretation of the model shown here assumes that TF binding (blue, rate f) modulates enhancer activation to affect promoter state. States are labeled by enhancer (first binary digit) and promoter (second binary digit) activation. Transcription occurs in states with promoter ON (gray). Reaction rates combine compactly into the following control parameters: b (TF unbinding rate), r (basal promoter deactivation rate), η (regulatory tradeoff between stabilization and facilitation), γ (cycle directionality), and ν (activation branch asymmetry). (*B*) Posterior distributions of model parameters obtained by fitting the cycle model $T_C(P_{ON})$ measurements in *Drosophila* embryos (10). Two 2D projections of the 4D posterior are shown, with a cyan star marking the best-fit values (posterior mean). We used Gaussian likelihood in log-space, uninformative priors, and empirical errors from Fig. 1. (*C*) Systematic parameter variations (color bar) around the best-fit value: T_C shape is insensitive to changes in r (*Left*), but strongly deformed by small changes in η (*Right*). Experimental T_C data from Fig. 1*A* shown in gray for reference. (*D*) Landscape of T_C flatness across (r, η) plane (*SI Appendix*, section 5.6). T_C–invariance (i.e., flatness $\approx$1) emerges only within a narrow regime of $\eta \approx 0.5$ and $r \lesssim b$, where enhancer facilitation and stabilization are balanced and time-scale separation is observed. Cyan star: best-fit values from panel (*B*). Semitransparent region: models that (unlike data) cannot be coarse-grained to a two-state description (*SI Appendix*, section 2.2). (*E*) Effect of (γ, ν) on dissipation (*Left*) and T_C flatness (*Right*). $\gamma = 0$ (vertical dotted line) defines the thermodynamic equilibrium manifold; models with $\gamma \neq 0$ are nonequilibrium. White circular arrows: probability flux direction in the cycle. High T_C flatness is confined to the $\gamma < 0$ region within the coarse-grainable regime. Models with $\gamma \gg 0$ produce less bursty dynamics and fail to match experimental constraints.

kinetics of basal promoter activation, with $\alpha_1 = \beta_1 = E$, so that $\alpha_1 q = r$ and $r/\beta_1 = q$. While the best-fit model remains accurately coarse-grainable into an effective two-state description, moderate deviations in parameters can break this property, leading to qualitatively different behavior where T_C loses its interpretability (semitransparent region in Fig. 2*D*, see also *SI Appendix*, Fig. S4*G*).

We analytically identify conditions under which the cycle model can be approximately or exactly coarse-grained into an effective two-state model (*SI Appendix*, section 5.8). Surprisingly, we demonstrate that exact coarse-graining into an effective two-state model is possible precisely when T_C is strictly invariant and equal to $1/(r+q)$, consistent with the plateau observed in Fig. 2*C*. This analysis clarifies the mechanistic requirements for T_C–invariance and links them to the elementary rates of the underlying reaction network.

The cycle model can be extended to include cooperative TF binding (*SI Appendix*, Fig. S5 and section 5.8), but this does not change our qualitative conclusions. Attempts to reduce the four-state cycle model to a simpler, three-state system fail to reproduce the desired T_C–invariance (*SI Appendix*, Fig. S6). Consequently, the four-state cycle model is the simplest mechanistically plausible reaction scheme consistent with empirical constraints.

Does the observed T_C–invariance require nonequilibrium dynamics? To address this question, we varied γ and ν over a broad range around their best-fit values, thereby scanning through both the equilibrium ($\gamma = 0$, or $\nu \to -\infty$) and the nonequilibrium regimes. As γ is displaced from zero, dissipation

increases (Fig. 2 *E*, *Left*), and the net probability flux reverses direction upon crossing the equilibrium boundary (*SI Appendix*, sections 5.3 and 5.4). Importantly, T_C–invariance cannot be achieved anywhere along the equilibrium parameter manifold (Fig. 2 *E*, *Right*), a result we prove analytically for all r and η within the four-state cycle model class (*SI Appendix*, section 5.7). The converse does not hold: Not all nonequilibrium models exhibit T_C–invariance. Taken together, our results rule out thermodynamic equilibrium as a viable operating regime for models exhibiting invariant T_C. This motivates us to ask whether there are any functional benefits that justify the energetic cost for maintaining the system out of equilibrium.

Invariant T_C Is a Signature of Optimized Information Flow. Promoters stochastically switch between ON and OFF states, thereby controlling mRNA levels. The characteristic timescale of promoter switching is, by definition, T_C, which one might a priori expect to depend on gene activity or other regulatory parameters. In contrast, the mRNA lifetime T_M is typically fixed and independent of regulation. The ratio of these timescales determines the steady-state mRNA distribution and therefore the precision of gene expression through temporal noise averaging (15). This suggests a simple intuition: If T_M is fixed, then maintaining comparable precision across the full dynamic range of expression could be achieved if T_C were also approximately invariant. Indeed, while mean expression is set by the fraction of time spent in the ON state, expression noise depends on how many switching events occur per mRNA lifetime, i.e., on $\sim T_C/T_M$. Does this qualitative picture hold quantitatively, and could it provide a selective advantage for T_C–invariance?

To examine this idea quantitatively, we computed information flow through different regulatory models. Information theory provides a formal language to quantify how reliably an input signal modulates an output in the presence of intrinsic noise and physical constraints. Importantly, mutual information (and, in particular, its maximal possible value, known as channel capacity) evaluates input–output relationships in absolute and interpretable units (bits), without committing to a particular decoding strategy or specific downstream objective (26). For these reasons, we used information as a functional metric to compare various regulatory models.

In our case, the regulatory input was the control parameter P_{ON}, which reflects promoter activation probability and can be mapped directly to changes in the kinetic parameter f or, equivalently, the TF concentration. (Formally, P_{ON} is the conditional expectation of a stochastic binary promoter state X, where $X \in \{0, 1\}$ and $X = 0(1)$ corresponds to promoter OFF(ON), given the TF-concentration-dependent binding rate f. Then, $P_{ON} = \mathbb{E}(X|f)$; but since f is a random variable, drawn from a distribution of TF concentrations sampled over the cell's lifetime, P_{ON} itself is also a random variable drawn from some distribution $p(P_{ON})$.) The output was the mRNA copy number M, a random variable with a steady-state distribution given by the corresponding birth–death process (Fig. 1*C* and *SI Appendix*, Fig. S7*A*, assuming a constant transcription initiation rate K from the active states, and a fixed mRNA lifetime T_M). Together, the input and output define an information channel, $p(M|P_{ON}(f))$, which can be fully characterized by solving the corresponding Master Equation (*SI Appendix*, sections 6.1 and 6.2). "Information flow," the functional metric for our regulatory models, can then be defined for each model as the channel capacity, i.e., the maximum mutual information $I(M; P_{ON})$, achievable by optimizing the input distribution $p^*(P_{ON})$ (52).

Within the class of effective two-state models, the shape of the $T_C(P_{ON})$ function fully determines the model's behavior. We can therefore ask which functional forms of T_C optimize information flow. Unsurprisingly, we find that information is maximized as $\langle T_C \rangle \to 0$, where promoter switching becomes infinitely fast. In this so-called "Poisson limit," burstiness (and therefore promoter switching noise) vanishes (45), and the promoter effectively behaves as if it were continuously transcribing at rate KP_{ON}. The only remaining source of noise in the channel is then the stochastic birth–death dynamics of mRNA.

Physical reaction rates in any real system are necessarily finite, ruling out the Poisson limit and imposing a constraint on promoter switching speed. To capture this constraint quantitatively, we define the effective switching speed as $V = \langle 1/T_C \rangle = \int dP_{ON}\ p^*(P_{ON})/T_C(P_{ON})$. As expected, V diverges in the Poisson limit, highlighting its idealized and unphysical nature.

For all other effective two-state models, we broadly sample the space of $T_C(P_{ON})$ functions (*SI Appendix*, section 6.3), and for each sampled model, we compute two key phenotypes: the information flow I and the switching speed V (*SI Appendix*, Fig. S7*B*). Plotting these quantities against one another reveals a structured manifold bounded from above and below (Fig. 3*A*). The lower boundary corresponds to inefficient models with minimal information flow given a fixed switching speed: This boundary is not biologically relevant. In contrast, the upper boundary defines a bona fide Pareto front: Information flow can only be increased by allowing faster switching. As $V \to \infty$, I approaches the Poisson limit. Strikingly, models on this optimal front exhibit increasingly invariant T_C functions as information flow increases, providing a quantitative link between information flow optimality and T_C–invariance. While broadly consistent with the intuition that invariant T_C ensures uniform temporal averaging over promoter noise across the full induction range, the mathematical emergence of this connection is both nontrivial and highly unexpected.

What region of the optimal manifold is accessible to our mechanistic cycle models, and can they approach the Pareto front? To address this, we broadly sampled the four independent model parameters (r, η, γ, ν), retaining only those configurations that can be reliably coarse-grained into an effective two-state description (concretely, for which $\delta_1 < 0.1$ in *SI Appendix*, section 2.2). The resulting cycle models densely populate the phenotype space and tightly approach the Pareto front defined by effective two-state models (Fig. 3*B*). Importantly, cycle models that maximize information flow along the Pareto front consistently exhibit invariant T_C across the full range of V; in contrast, T_C–invariance disappears away from the front. These findings suggest that T_C–invariance is not a generic feature, but rather a specific consequence of optimizing information flow under kinetic constraints.

We next examined how the information flow I and switching speed V in cycle models relate to entropy production (or, equivalently, energy dissipation) $\dot{S}$. As shown in Fig. 3*C*, the entire equilibrium manifold, comprising models with $\gamma = 0$, falls substantially below the Pareto front and, by extension, well below the best performing nonequilibrium models (*SI Appendix*, Fig. S8*A*). Compared to their equilibrium counterparts, nonequilibrium models can support significantly higher information transmission while operating at slower switching speeds. These differences are far from negligible: In some cases, gains in information can reach nearly 1 bit, highlighting a substantial functional benefit of nonequilibrium regulation. Importantly, the best-fit nonequilibrium cycle model from Fig. 2*B* lies very close to the Pareto front, providing clear support for the information

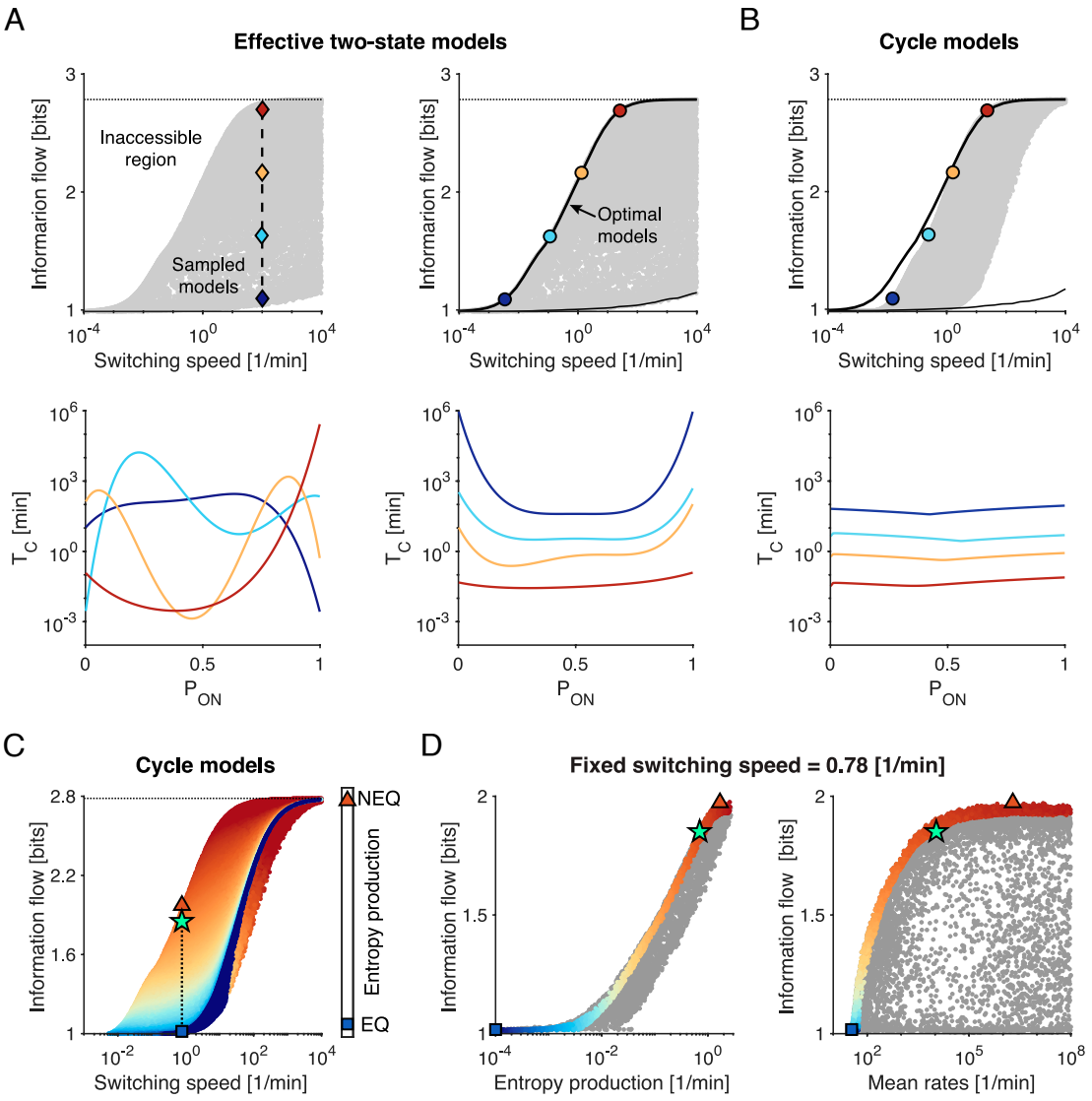

Fig. 3. Optimizing information flow under a switching speed constraint predicts invariant T_C. (*A*) *Top*: Information flow I as a function of switching speed V, sampled for effective two-state models with diverse $T_C(P_{ON})$ functions (gray dots). Maximal information flow (dotted line) is achieved in the Poisson limit ($V \to \infty$). Four representative models at fixed V are highlighted (colored diamonds; *Top Left*). Models that maximize I at fixed V (colored circles; *Top Right*) lie on the Pareto front (upper black line). *Bottom*: Corresponding T_C functions for the highlighted models. (*B*) *Top*: Manifold of coarse-grainable four-state cycle models (gray dots) plotted in the I–V plane. Cycle models, particularly at high I, closely approach the effective two-state Pareto front [black line, reproduced from (*A*)]. *Bottom*: Cycle models on the Pareto front exhibit nearly invariant T_C, in contrast to off-the-front models. (*C*) Information flow vs. switching speed V for cycle models, colored by minimal dissipation (entropy production) required to achieve the corresponding (V, I) combination. Equilibrium models (dark blue) lie far below the Pareto front, whereas nonequilibrium models (brighter colors) achieve higher I at lower V. The best-fit model from Fig. 2*B* (cyan star, $V^* \sim 0.78$ min^{-1}) lies near the front. Matched V^* equilibrium and nonequilibrium models are highlighted (blue square and red triangle, respectively). (*D*) Sections of the regulatory phenotype manifold for cycle models at fixed switching speed $V^* \sim 0.78$ min^{-1} [gray dots, as in (*B*)]. The Pareto front [colored dots, as in (*C*)] tracks trade-offs between information flow, energy dissipation, and mean reaction rates. The same three models from (*C*) are highlighted.

optimization hypothesis in the context of *Drosophila* gap gene regulation.

The analysis of constraints and trade-offs can be extended further (*SI Appendix*, Fig. S9*A*). We expect that the true Pareto front for cycle models lives in a high-dimensional space of regulatory phenotypes. While Fig. 3*A* explored the projection of these phenotypes onto the I vs. V plane, other sections of the solution space may yield complementary insights.

To illustrate this, Fig. 3*D* fixes the switching speed to the value inferred from data ($V^* \sim 0.78\,\mathrm{min}^{-1}$) and displays two Pareto front cross-sections. Along the front, the cycle model can be continuously "deformed" from best equilibrium model (no dissipation, $I = 1$ bit) to the best nonequilibrium model (maximal dissipation, $I \approx 2$ bits). This increase in performance, however, comes at a steep cost: Achieving maximal information flow requires the mean reaction rates in the cycle to increase by nearly four orders of magnitude.

Interestingly, while the best-fit system lies close to the optimal nonequilibrium model in terms of information flow and energy dissipation (requiring a ~2.4 fold increase in entropy production to close an information gap of ≈0.12 bits), it achieves this performance with mean rates that are two orders of magnitude lower. This contrast suggests that biological systems may face important constraints not only on energy usage but also on the magnitude of individual reaction rates, a consideration that has received surprisingly little attention to date.

In line with the perspective laid out in the introduction, our analysis integrates prior empirical observations, data-driven inference, physico-chemical constraints, and optimization theory to delineate the space of regulatory models that are both admissible and functional. Within this space, a simple class of four-state cycle models emerges as particularly instructive. These models not only reproduce key features of the data but also reveal nontrivial trade-offs that biological systems must navigate. Our results suggest that selection for regulatory precision may have driven the emergence of invariant T_C. This invariance, near the Pareto front of functional performance, is a hallmark of optimal information transmission given physical constraints.

Cycle Models Are Robust to Perturbations. Perturbation experiments in *Drosophila* gap genes revealed striking robustness of T_C–invariance (10). Changes in TF concentrations or mutations/deletions in enhancer sequences primarily affected the mean gene activity (P_{ON}) but only had a limited impact on the magnitude or shape of T_C. Many such cis- and trans-perturbations, ranging from altered TF levels and chromatin modifications to mutations in DNA binding sites or enhancer deletions, are expected to primarily affect the "enhancer activation" step of the regulatory cycle (Fig. 2*A*). In our model, these perturbations can be effectively parameterized as changes in the TF unbinding rate (b). When b is perturbed tenfold around its optimal value, the induction curves $P_{ON}(f)$ exhibit substantial shifts, yet T_C remains largely unchanged (Fig. 4*A*), consistent with the experimental observations (10).

To systematically evaluate whether this robustness is recapitulated by the cycle model, we perturbed each of its elementary reaction rate parameters ($b, q, r, \alpha_1, \beta_1, \alpha_2, \beta_2$; see Fig. 2*A*), one at a time, around their best-fit values, using multiplicative log-normal noise. This approach, equivalent to additive Gaussian noise in log-space, is well justified, as kinetic rates are strictly positive and naturally defined on a multiplicative scale. We chose to perturb elementary reaction rates rather than the derived control parameters (r, η, γ, ν), which are more convenient for analytical calculations, because the former have a clearer mechanistic interpretation and can be more directly linked to mutational effects. For each perturbed model, we quantified changes in the induction and the T_C curves, by computing the SD (across perturbations) of the mean logarithmic change (*SI Appendix*, section 7.1). The largest effects on induction were caused by perturbing b (TF unbinding), r and q (promoter (de)activation), and α_1 (enhancer facilitation), as shown in Fig. 4*B* and *SI Appendix*, Fig. S10*A*. In contrast, T_C was more robust overall, with only moderate sensitivity to changes in r, q, α_1, and β_2, as shown in Fig. 4B and *SI Appendix*, Fig. S10*B*. These model predictions qualitatively mirror the robustness patterns observed experimentally.

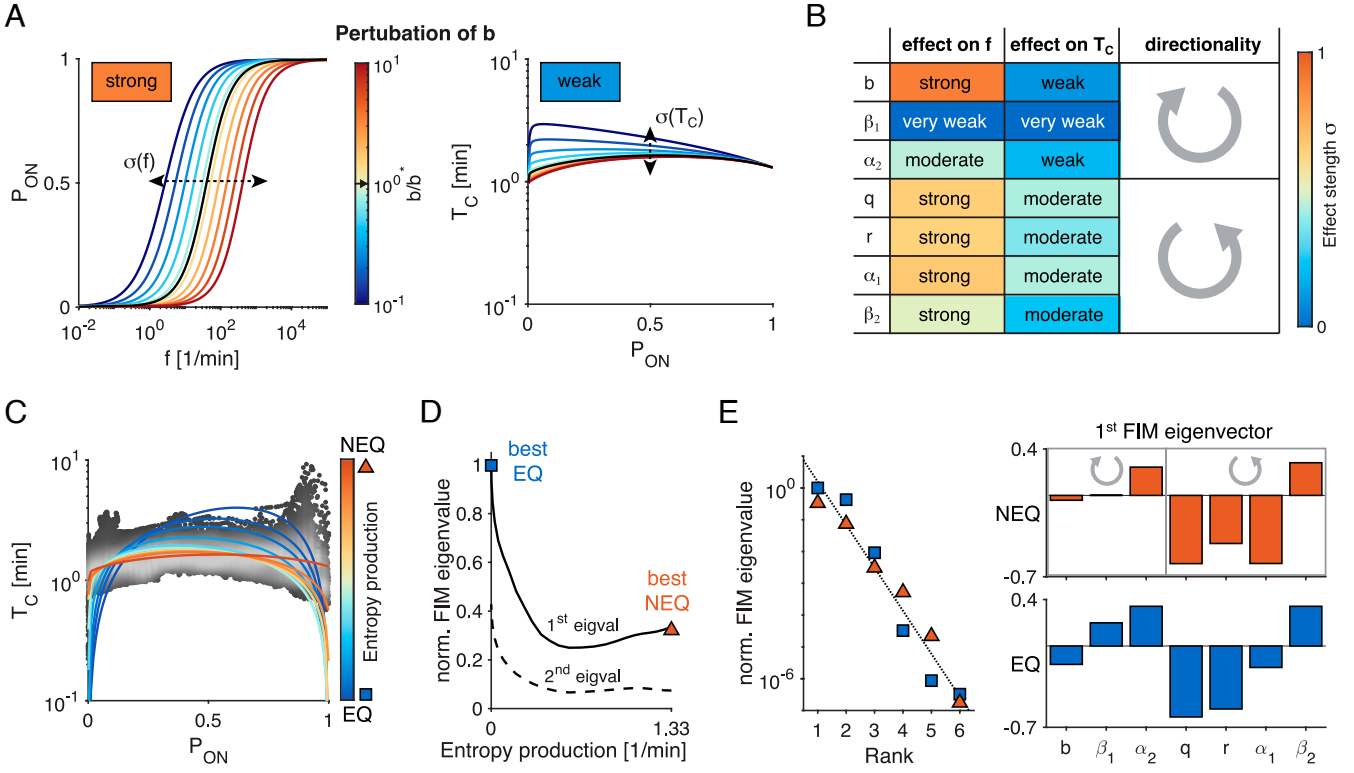

Fig. 4. Best-fit cycle model is robust to perturbations. (*A*) Perturbing the TF unbinding rate b by ±10-fold (log-normal noise around the best-fit value b^*) has a strong effect on the induction curves $P_{ON}(f)$ (*Left*), but only a weak effect on the switching correlation time $T_C(P_{ON})$, consistent with experimental findings. (*B*) Systematic analysis of individual parameter perturbations reveals that rates aligned with the natural counterclockwise direction of the transcription cycle (e.g., r, q, and α_1) have strong effects on induction and moderate effects on T_C, whereas reverse-direction rates have minimal impact on T_C. Color scale quantifies effect strength (SD across perturbations); gray arrows indicate cycle directionality aligned with the perturbed rates. See also *SI Appendix*, Fig. S10 *A* and *B* for quantitative effect sizes. (*C*) T_C functions for a family of interpolating models, generated by varying the entropy production (color scale, controlled by γ) between the best-fit NEQ model ($\gamma = -6.32$, dark orange; cyan star in Figs. 2 and 3) and the best-fit equilibrium model (EQ) ($\gamma = 0$, dark blue). For each γ, all remaining parameters were refit to the observed T_C. Fit quality steadily deteriorates toward equilibrium. See also *SI Appendix*, Fig. S10 *C*–*E* for parameter values and refitting performance. (*D*) First two eigenvalues of the Fisher Information Matrix (FIM) as a function of entropy production. Eigenvalues are normalized to the largest eigenvalue for the best-fit EQ model. (*E*) Normalized eigenvalue spectra (*Left*) of the best-fit NEQ and EQ models reveal "sloppiness," a logarithmic hierarchy of sensitivities. The dominant eigenvector (*Right*) for the NEQ model shows that sensitivity is concentrated in parameters aligned with cycle direction (q, r, and α_1), whereas the EQ model displays more diffuse and unfocused sensitivity.

Could robustness be a general feature of nonequilibrium dynamics? The parameters that most strongly influence T_C in our perturbation analysis (q, r, α_1, β_2) correspond to reaction steps aligned with the natural, counterclockwise progression of the transcriptional cycle (Fig. 4*B*). In contrast, reactions directed opposite to this direction have negligible influence on T_C. This directional asymmetry in sensitivity suggests that nonequilibrium dynamics—by favoring a unidirectional flux through the transcriptional cycle—may inherently buffer the system against a broad class of local perturbations.

To investigate this idea more systematically, we assessed how robustness varies across the model's parameter space as a function of entropy production. Starting with our best-fit nonequilibrium model (NEQ; with $\gamma = -6.32$), we generated a family of interpolating models by gradually increasing γ toward zero, thereby reducing entropy production and approaching equilibrium (*SI Appendix*, Fig. S10 *C–E*). For each value of γ, all remaining parameters were refit to match the invariant T_C function (Fig. 4*C*). As expected, fit quality deteriorated steadily with decreasing dissipation. Nevertheless, this family of models provided a consistent basis for computing the Fisher Information Matrix (FIM), which quantifies the local sensitivity of the T_C fit to changes in the elementary reaction rates (*SI Appendix*, section 7.2).

First, we observed that the leading eigenvalue of the FIM decreased as the best-fit NEQ model was approached and entropy production increased. Specifically, the dominant eigenvalue was approximately threefold higher for the best-fit EQ model than for the best-fit NEQ model. This indicates that the T_C function for the EQ model is significantly more sensitive, and thus less robust, to coordinated parameter perturbations, reinforcing the single-parameter perturbation results shown in Fig. 4*B*.

Second, we analyzed the full eigenvalue spectrum of the FIM, for both the best NEQ and EQ models. In both cases, the spectrum displayed a characteristic hierarchy, with eigenvalues approximately equally spaced on a logarithmic scale: a hallmark of the so-called "sloppy" models (Fig. 4 *E*, *Left*) (53). This indicates that only a few effective parameter combinations are tightly constrained by the T_C data. For the best NEQ model, the eigenvector associated with the leading eigenvalue revealed that three parameters (q, r, α_1) dominate the sensitivity of the T_C function (Fig. 4 *E*, *Right*). These parameters align with the natural counterclockwise progression of the transcriptional cycle. In contrast, the best EQ model displayed more broadly distributed and less focused sensitivity to parameter variation.

Taken together, these results highlight a key feature of the nonequilibrium regime: It enables the transcriptional cycle to maintain high information flow while being robust to multiple perturbation directions in the parameter space. Interestingly, this robustness is not a generic consequence of cyclic topology per se, but rather a consequence of a specific parameter regime within which the cycle model operates.

Discussion

This study provides both a mechanistic and functional account of an empirically robust yet theoretically unexplained observation: the invariance of promoter switching correlation time, T_C, across a wide range of gene expression levels in early *Drosophila* development. This invariance has now been observed in multiple genes and appears conserved across species, though with varying degrees of measurement precision. Canonical models of transcriptional regulation cannot naturally account for this property. Using analytical theory and inference, we demonstrate that reproducing T_C–invariance requires regulatory architectures with at least four promoter states and broken detailed balance, i.e., systems operating out of thermodynamic equilibrium. These nonequilibrium dynamics unfold on timescales slower than those of individual TF binding and unbinding events (29), in agreement with experimental observations. We note that while our analyses ruled out equilibrium models of comparable complexity to the four-state cycle model (*SI Appendix*, Fig. S2), we cannot currently exclude *all possible*, even arbitrarily complex, equilibrium models as inconsistent with empirical constraints; this remains an open question for the future. Given the growing interest in the equilibrium vs. nonequilibrium nature of gene regulation (25, 39), our identification of T_C–invariance as a defining experimental signature provides a concrete, quantitative foothold for further theoretical and empirical exploration.

Four-state transcriptional models are not new: They appeared in early work on gene expression noise in yeast (54) and are now being revisited using time-resolved single-cell methods (40). Theoretical studies have explored the rich phenotypic behaviors that such models enable, including nonequilibrium control of gene regulation (25, 29, 55). Related models have also been linked to mammalian transcriptional data, albeit from different conceptual perspectives (56). This convergence—across systems, methods, and motivations—suggests that a unifying quantitative framework for eukaryotic transcriptional regulation may be emerging, potentially playing a role analogous to that of thermodynamic models in prokaryotic systems.

Of particular note are strong parallels to a recent study by Shelansky et al. (40), who performed live imaging of gene expression in yeast. Much like in our work, the authors find evidence for two distinct transcriptionally active (ON) states that are not directly resolved observationally and are therefore coarse-grained into a single effective ON state, along with an effective OFF state whose internal structure, if present, cannot be inferred. They further provide rigorous theoretical arguments that the measured dwell-time distributions of these coarse-grained states require cyclic, irreversible dynamics, and therefore regulation out of thermodynamic equilibrium.

While the biological systems differ substantially between yeast and *Drosophila*, and while differences in experimental access and timescale separation may plausibly account for the apparent three- vs. four-state identifiability differences, we highlight three points. *i)* The qualitative biological conclusions of the two studies are in strong agreement: Both support the view that transcriptional regulation is organized into an irreversible cycle with at least two ON states. *ii)* Both theoretical approaches motivate the continued development of coarse-graining, identifiability, and inference frameworks for Markov chains and chemical reaction networks; specifically, the search for complementary signatures (e.g., dwell-time distributions, T_C dependence on gene activity) of out-of-equilibrium operation. *iii)* It would be of particular interest to quantify the promoter switching autocorrelation time T_C in the yeast system studied by Shelansky et al. and to test whether it exhibits invariance comparable to that observed here.

From an inference standpoint, extensive previous work also studied how and when multistate or mechanistic models can be discriminated from the effective two-state (telegraph) model (46, 57, 58). Given limited data, this is a challenging and sometimes nonidentifiable problem if one has experimental access to steady-state mRNA distributions alone. Dynamical observables derived from the sampled single-cell ON/OFF state switching trajectories (such as their autocorrelation time, T_C, in this study), or the dependence on mRNA distribution on an

external control parameter (such as the P_{ON} in this study), can provide substantially more statistical discriminatory power to rule in or out individual model classes.

Cycle models tuned to reproduce T_C–invariance exhibit multiple desirable properties. **First**, they can be accurately coarse-grained into effective two-state models that reproduce experimental data. This is of practical importance: Two-state models are easier to infer from data, while inference of full mechanistic multistate models remains challenging and often ill-posed. Yet, when two-state models fit well, they can be interpreted as statistical summaries that compress more complex underlying mechanisms. In this setting, theoretical analysis, not inference, can bridge effective and mechanistic descriptions. We provide an example of such mapping by showing how T_C–invariance emerges from coarse-grainable models with broken detailed balance, enabling concrete inferences about nonequilibrium regulatory mechanisms. This mechanistic-to-effective connection also opens directions for analyzing other biological signal-processing systems with defined input and output states, extending beyond transcription (26, 59).

Second, nonequilibrium cycle models exhibit increased robustness to parameter perturbations relative to equilibrium models. Cycle models are particularly robust to changes in parameters controlling reactions that run against the probability flux in the transcriptional cycle. While robustness might be desirable on general grounds, it is especially relevant here, since previously reported genetic perturbations in *Drosophila* failed to disrupt the T_C–invariance despite affecting expression levels and induction curves (10). We also report on the broad eigenvalue spectrum, or sloppiness, of the Fisher Information Matrix (53, 60). In our case, sloppiness arises in a model class where both data fitting and optimization for a functional phenotype (information transmission) are tightly interlinked. Whether the observed sloppiness results primarily from data-fitting-related or from optimality-related degeneracy remains unclear. Moreover, the relevance of sloppiness may depend on whether one analyzes elementary reaction rates or coarse-grained combinations. Systems, such as ours, where optimization, sloppiness, and inference can be analyzed within a single framework, offer an exciting platform for further explorations (24, 61, 62).

Third, we find that T_C–invariance in constrained cycle models is a direct signature of optimal information transmission. Biologically, this is a statement about intrinsic noise control: Maximal information transmission implies the most reliable propagation of signals from transcription factor concentrations to downstream gene expression, given fundamental sources of noise such as promoter switching and stochastic mRNA transcription (45, 46). This connects to classical problems in chemoreception (28, 63), with the key distinction here being the constraint of T_C–invariance. We find that four-state cycle models maximizing steady-state information flow inevitably exhibit this invariance. While we support this numerically and provide intuitive arguments, a general analytical proof remains elusive at the moment. Recent progress in the information theory community, motivated by analyzing telegraph models as communication channels relevant to gene expression, holds early promise for understanding our numerical results at a deeper level (47, 48).

Importantly, T_C invariance is not observed when cycle model parameters are sampled randomly; it only emerges as a result of information flow optimization at a fixed switching speed V. While the value of V sets the timescale of T_C, it does not need to be precisely tuned to generate invariance. Without such a constraint, optimization favors infinitely fast switching and Poisson mRNA statistics, biologically implausible scenarios. Although physical and energetic limits cap V, the appropriate mathematical form of this constraint remains unclear: Should the bound be placed on the average switching rate of the promoter, the maximum reaction rate in the network, or some other network-level feature? Moreover, switching speed constraints are distinct from energy dissipation limits. For instance, equilibrium models with fast kinetics can suppress bursting noise and approach the Poisson limit (Fig. 3*C*), yet still preserve detailed balance. Furthermore, molecular reaction schemes for gene regulation could implement kinetic proofreading (29, 30, 64) to ensure that only cognate TF signals activate or repress a given gene, thereby rejecting "transcriptional crosstalk" due to nonspecific or noncognate binding (65). Such schemes generically require extra time before committing to a change in ON/OFF gene state, thereby providing an alternative reason—functional and not purely physical—for the constraint on the promoter switching speed. In sum, these considerations suggest that switching speed, independent of entropy production, is an underappreciated constraint that crucially shapes the landscape of feasible regulatory dynamics.

Our work opens several intriguing links to contemporary control theory. Classical control theory has long relied on pulse-width modulation (PWM)—a technique underpinning modern power electronics and digital control implementations (66)—as a canonical approach for regulating systems with binary or saturating actuators. In deterministic PWM, a fixed clock period T is imposed, and regulation is achieved by varying the duty cycle D, i.e., the fraction of time spent in the active state. In our notation, this would correspond to fixing $T = T_{ON} + T_{OFF}$ and modulating $D = T_{ON}/T$, which maps directly onto gene activity, $P_{ON} = D$.

Clocked PWM strategies can minimize variance around a desired setpoint and thus implement the tightest possible control in idealized engineered settings. However, they rely on a precise global timing mechanism that enforces a fixed period T within which the duty cycle D can be modulated. In contrast, many natural and stochastic systems—particularly in chemical and biological contexts— regulate activity through random switching between discrete states, for which implementing a global clock is intrinsically challenging. Stochastic (as opposed to clocked and thus deterministic) PWM switching schemes have also been studied in more recent control engineering literature, including randomized PWM in power converters to reduce electromagnetic interference, and random access or service scheduling in communication and queuing systems (67–69). Stochastic switching dynamics and stochastic PWM control may therefore be a paradigm more applicable to the biological setting.

Stochastic switching dynamics are naturally described by continuous-time discrete-space Markov processes and, when control is applied through the choice of transition rates, by continuous-time Markov decision processes (CTMDPs). A central result in CTMDP theory is that for long-run average objectives, optimal control can be achieved by the modulation of transition rates, with system performance determined by the induced stationary distribution (70, 71). For a binary (ON/OFF) system with exponentially distributed dwell times—corresponding to the telegraph model and our effective two-state description of transcription—the mean activity depends on the ratio of transition rates, while the T_C sets the overall tempo of stochastic switching. Within this framework, T_C–invariance can arise as a natural stochastic analogue of fixed frequency PWM under broad conditions, for example *i)* when switching

incurs a cost proportional to transition frequency (71, 72); *ii)* when open-loop control favors minimally structured (maximum-entropy) rate-modulation strategies that admit fully memoryless implementations (70, 73); or *iii)* when performance depends primarily on mean activity rather than fine-grained temporal correlations.

Beyond these specific highlights and connections to control theory, our study reveals a complex trade-off landscape among information flow, switching speed, energy dissipation, and reaction rate constraints. These objectives often cannot be simultaneously optimized, either due to fundamental physical limits or evolutionary constraints. Rather than assume the existence of a single optimal model, we advocate for exploring ensembles of near-optimal solutions. In such ensembles, no model achieves the best value for every objective, but many models might simultaneously approach optimal values for multiple objectives. These ensembles are computationally accessible and can be statistically characterized thanks to recent theoretical advances (24, 25, 61).

Computationally, such ensembles can be generated by unbiased sampling (as in Fig. 3 *A* and *B*) or biased toward functional phenotypes. Biologically, this bias mimics selective preference for functional states that are reachable through evolutionary adaptation (62). Deep links between statistical physics, information theory, and population genetics connect these ensembles to the strength of natural selection (22, 23, 74), supporting the view that optimality in biology is better understood as a landscape of trade-offs rather than a single global maximum.

As an illustration, our analysis (Fig. 3*D*) shows that *near-maximal* information transmission can be achieved with reaction rates two orders of magnitude slower than those required for *maximal* information flow. Given the limits to selection in finite populations, such marginal phenotypic gains in information may be neither resolvable nor selectively advantageous, especially if they come at the cost of reduced robustness or higher energetic demands. These trade-offs suggest that evolution might favor "good enough" solutions over theoretically optimal ones. Quantitatively linking physics-style constrained optimization to evolutionary fitness models thus offers an exciting direction for future research.

In sum, our findings suggest that eukaryotic transcription leverages nonequilibrium regulation to enhance robustness against perturbations and noise. Energy dissipation plays a central role in enabling high information transmission at biologically realistic switching speeds. In the complex regulatory trade-off landscape, moderate dissipation allows systems to approach a performance plateau, beyond which further gains in precision would require disproportionate increases in energy consumption and reaction speed (25). The widespread relevance of transcription, the availability of precise measurements, the presence of fundamental biophysical constraints, and the existence of information-theoretic proxies for function together make gene regulation a particularly fertile ground for integrative studies. As such, gene regulation offers a uniquely tractable and meaningful setting for connecting biological function, physical law, and evolutionary dynamics.

Materials and Methods

For each regulatory model, we use the analytical solutions of the corresponding chemical master equation based on the matrix exponential framework to compute the steady state probabilities, induction curve, autocorrelation function, and other regulatory phenotypes. Channel capacities are computed using a version of Blahut–Arimoto algorithm. Pareto fronts are obtained by initial uniform/log-uniform (depending on the parameter in question) sampling of the parameter space to identify the approximate boundaries, followed by targeted refined model space sampling close to the boundary. Please refer to *SI Appendix* for the detailed model description, theoretical derivations, proofs, and computational details.

Data, Materials, and Software Availability. Software code data have been deposited in Institute Pasteur GitHub (https://gitlab.pasteur.fr/tglab/invariantpromoterdynamicspaper) (51).

ACKNOWLEDGMENTS. This work was supported by the French National Research Agency (ANR-20-CE12-0028 "ChroDynE" and ANR-23-CE13-0021 "GastruCyp" and ANR-10 LABX-73 "Revive;" all T.G.), and by funding from the European Research Council (ERC-2023-SyG, "Dynatrans," 101118866, T.G. and G.T.). This work was also supported in part by the U.S. NSF, through the Center for the Physics of Biological Function (PHY-1734030, T.G.), and by NIH Grants R01GM097275, U01DA047730, and U01DK127429 (T.G.).

Author affiliations: [a]Department of Stem Cell and Developmental Biology, CNRS UMR3738 Paris Cité, Institut Pasteur, Paris FR-75015, France; [b]Institute of Science and Technology Austria, Klosterneuburg AT-3400, Austria; and [c]Joseph Henry Laboratories of Physics, Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544

Author contributions: B.Z., T.G., and G.T. designed research; B.Z. and A.B. performed research; B.Z. and G.T. analyzed data; and B.Z., A.B., T.G., and G.T. wrote the paper.

1. J. Rodriguez, D. R. Larson, Transcription in living cells: Molecular mechanisms of bursting. *Annu. Rev. Biochem.* **89**, 1–24 (2020).
2. J. V. Meeussen, T. L. Lenstra, Time will tell: Comparing timescales to gain insight into transcriptional bursting. *Trends Genet* **40**, 160–174 (2024).
3. W. J. d. Jonge, H. P. Patel, J. V. Meeussen, T. L. Lenstra, Following the tracks: How transcription factor binding dynamics control transcription. *Biophys. J.* **121**, 1583–1592 (2022).
4. K. Wagh, D. A. Stavreva, A. Upadhyaya, G. L. Hager, Transcription factor dynamics: One molecule at a time. *Annu. Rev. Cell Dev. Biol.* **39**, 277–305 (2023).
5. J. Peccoud, B. Ycart, Markovian modeling of gene-product synthesis. *Theor. Popul. Biol.* **48**, 222–234 (1995).
6. J. Rodriguez *et al.*, Intrinsic dynamics of a human gene reveal the basis of expression heterogeneity. *Cell* **176**, 213–226.e18 (2019).
7. Y. Wan *et al.*, Dynamic imaging of nascent RNA reveals general principles of transcription dynamics and stochastic splice site selection. *Cell* **184**, 2878–2895.e20 (2021).
8. K. Tantale *et al.*, Stochastic pausing at latent HIV-1 promoters generates transcriptional bursting. *Nat. Commun.* **12**, 4503 (2021).
9. V. L. Pimmett *et al.*, Quantitative imaging of transcription in living Drosophila embryos reveals the impact of core promoter motifs on promoter state dynamics. *Nat. Commun.* **12**, 4504 (2021).
10. P. T. Chen, M. Levo, B. Zoller, T. Gregor, A conserved coupling of transcriptional ON and OFF periods underlies bursting dynamics. *Nat. Struct. Mol. Biol.* **32**, 1959–1971 (2025).
11. P. Trzaskoma *et al.*, 3D chromatin architecture, BRD4, and Mediator have distinct roles in regulating genome-wide transcriptional bursting and gene network. *Sci. Adv.* **10**, eadl4893 (2024).
12. B. Zoller, S. C. Little, T. Gregor, Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting. *Cell* **175**, 835–847.e25 (2018).
13. J. Paulsson, Models of stochastic gene expression. *Phys. Life Rev.* **2**, 157–175 (2005).
14. A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, S. Tyagi, Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).
15. G. Tkačik, T. Gregor, W. Bialek, The role of input noise in transcriptional regulation. *PLoS One* **3**, e2774 (2008).
16. D. Nicolas, N. E. Phillips, F. Naef, What shapes eukaryotic transcriptional bursting? *Mol. BioSyst.* **13**, 1280–1290 (2017).
17. D. R. Larson *et al.*, Direct observation of frequency modulated transcription in single cells using light activation. *eLife* **2**, e00750 (2013).
18. A. Senecal *et al.*, Transcription factors modulate c-Fos transcriptional bursts. *Cell Reports* **8**, 75–83 (2014).
19. T. Fukaya, B. Lim, Levine M. Enhancer control of transcriptional bursting. *Cell*. **166**, 358–368 (2016).
20. D. Nicolas, B. Zoller, D. M. Suter, F. Naef, Modulation of transcriptional burst frequency by histone acetylation. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7153–7158 (2018).
21. J. M. Smith, Optimization theory in evolution. *Annu. Rev. Ecol. System.* **9**, 31–56 (1978).
22. G. Sella, A. E. Hirsh, The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9541–9546 (2005).
23. M. Hledík, N. Barton, G. Tkačik, Accumulation and maintenance of information in evolution. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2123152119 (2022).
24. W. Młynarski, M. Hledík, T. R. Sokolowski, G. Tkačik, Statistical analysis and optimality of neural systems. *Neuron* **109**, 1227–1241.e5 (2021).
25. B. Zoller, T. Gregor, G. Tkačik, Eukaryotic gene regulation at equilibrium, or non? *Curr. Opin. Syst. Biol.* **31**, 100435 (2022).

26. G. Tkačik, P. R. t. Wolde, Information processing in biochemical networks. *Annu. Rev. Biophys.* **54**, 249–274 (2025).
27. G. Tkačik, C. G. Callan, W. Bialek, Information flow and optimization in transcriptional regulation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 12265–12270 (2008).
28. C. C. Govern, P. R. t. Wolde, Optimal resource allocation in cellular sensing systems. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17486–17491 (2014).
29. R. Grah, B. Zoller, Nonequilibrium models of optimal enhancer function. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 31614–31622 (2020).
30. S. A. Cepeda-Humerez, G. Rieckh, G. Tkačik, Stochastic proofreading mechanism alleviates crosstalk in transcriptional regulation. *Phys. Rev. Lett.* **115**, 248101 (2015).
31. J. Estrada, F. Wong, A. DePace, J. Gunawardena, Information integration and energy expenditure in gene regulation. *Cell* **166**, 234–244 (2016).
32. C. Scholes, A. H. DePace, A. Sanchez, Combinatorial gene regulation through kinetic control of the transcription cycle. *Cell Syst.* **4**, 97–108.e9 (2017).
33. N. C. Lammers, A. I. Flamholz, H. G. Garcia, Competing constraints shape the nonequilibrium limits of cellular decision-making. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2211203120 (2023).
34. R. Martinez-Corral, K. M. Nam, A. H. DePace, J. Gunawardena, The Hill function is the universal Hopfield barrier for sharpness of input-output responses. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2318329121 (2024).
35. M. L. Perkins, J. Crocker, G. Tkačik, Chromatin enables precise and scalable gene regulation with factors of limited specificity. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2411887121 (2025).
36. D. M. Suter *et al*., Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472–474 (2011).
37. B. Zoller, D. Nicolas, N. Molina, F. Naef, Structure of silent transcription intervals and noise characteristics of mammalian genes. *Mol. Syst. Biol.* **11**, 823 (2015).
38. C. Li, F. Cesbron, M. Oehler, M. Brunner, T. Höfer, Frequency modulation of transcriptional bursting enables sensitive and rapid gene regulation. *Cell Syst.* **6**, 409–423.e11 (2018).
39. F. Wong, J. Gunawardena, Gene regulation in and out of equilibrium. *Annu. Rev. Biophys.* **49**, 199–226 (2020).
40. R. Shelansky *et al*., Single gene analysis in yeast suggests nonequilibrium regulatory dynamics for transcription. *Nat. Commun.* **15**, 6226 (2024).
41. D. Huynh *et al*., Effective in vivo binding energy landscape illustrates kinetic stability of RBPJ-DNA binding. *Nat. Commun.* **16**, 1259 (2025).
42. M. Mir *et al*., Dynamic multifactor hubs interact transiently with sites of active transcription in drosophila embryos. *eLife* **7**, e40497 (2018).
43. P. B. Warren, S. Tănase-Nicola, P. R. t. Wolde, Exact results for noise power spectra in linear biochemical reaction networks. *J. Chem. Phys.* **125**, 144904 (2006).
44. I. Lestas, J. Paulsson, N. E. Ross, G. Vinnicombe, Noise in gene regulatory networks. *IEEE Trans. Autom. Control.* **53**, 189–200 (2008).
45. G. Tkačik, C. G. Callan Jr., W. Bialek, Information capacity of genetic regulatory elements. *Phys. Rev. E* **78**, 011910 (2008).
46. G. Rieckh, G. Tkačik, Noise and information transmission in promoters with multiple internal states. *Biophys. J.* **106**, 1194–1204 (2014).
47. M. Sinzger, M. Gehri, H. Koeppl, "Poisson channel with binary Markov input and average sojourn time constraint" in *2020 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2020), pp. 2873–2878.
48. M. Gehri, N. Engelmann, H. Koeppl, "Mutual information of a class of poisson-type channels using Markov renewal theory" in 2024 *IEEE International Symposium on Information Theory* (ISIT, 2024), pp. 1931–1936.
49. L. A. Mirny, Nucleosome-mediated cooperativity between transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22534–22539 (2010).
50. G. Tkačik, A. M. Walczak, Information transmission in genetic regulatory networks: A review. *J. Phys. Condens. Matter* **23**, 153102 (2011).
51. B. Zoller, Code for invariant non-equilibrium dynamics in gene regulation optimize information transmission. Gitlab. https://gitlab.pasteur.fr/tglab/invariantpromoterdynamicspaper. Accessed 22 June 2026.
52. R. Blahut, Computation of channel capacity and rate-distortion functions. *IEEE Trans. Inf. Theory* **18**, 460–473 (1972).
53. M. K. Transtrum *et al*., Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *J. Chem. Phys.* **143**, 010901 (2015).
54. W. J. Blake, M. Kærn, C. R. Cantor, J. J. Collins, Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
55. S. D. Mahdavi, G. L. Salmon, P. Daghlian, H. G. Garcia, R. Phillips, Flexibility and sensitivity in gene regulation out of equilibrium. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2411395121 (2024).
56. J. Tünnermann, G. Roth, J. Cramard, L. Giorgetti, Enhancer control of promoter activity and variability via frequency modulation of clustered transcriptional bursts. bioRxiv [Preprint] (2025). https://doi.org/10.1101/2025.03.26.645410 (Accessed 22 June 2026).
57. S. Braichenko, J. Holehouse, R. Grima, Distinguishing between models of mammalian gene expression: Telegraph-like models versus mechanistic models. *J. R. Soc. Interface* **18**, 20210510 (2021).
58. F. Jiao *et al*., What can we learn when fitting a simple telegraph model to a complex gene expression model?. *PLoS Comput. Biol.* **20**, e1012118 (2024).
59. M. Reinhardt, G. Tkačik, P. R. Ten Wolde, Path weight sampling: Exact Monte Carlo computation of the mutual information between stochastic trajectories. *Phys. Rev. X* **13**, 041017 (2023).
60. B. B. Machta, R. Chachra, M. K. Transtrum, J. P. Sethna, Parameter space compression underlies emergent theories and predictive models. *Science* **342**, 604–607 (2013).
61. M. Bauer *et al*., Optimization and variability can coexist. arXiv [Preprint] (2025). http://arxiv.org/abs/2505.23398 (Accessed 22 June 2026).
62. T. R. Sokolowski, T. Gregor, W. Bialek, G. Tkačik, Deriving a genetic regulatory network from an optimization principle. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2402925121 (2025).
63. G. Malaguti, P. R. Ten Wolde, Theory for the optimal detection of time-varying signals in cellular sensing systems. *eLife* **10**, e62574 (2021).
64. R. Shelansky, H. Boeger, Nucleosomal proofreading of activator-promoter interactions. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2456–2461 (2020).
65. T. Friedlander, R. Prizak, C. C. Guet, N. H. Barton, G. Tkačik, Intrinsic limits to gene regulation by global crosstalk. *Nat. Commun.* **7**, 12307 (2016).
66. K. J. Åström, B. Wittenmark, *Computer-Controlled Systems: Theory and Design* (Dover Publications, Mineola, NY, ed. 3, 2011).
67. T. M. Habetler, D. M. Divan, Acoustic noise reduction in sinusoidal PWM drives using a randomly modulated carrier. *IEEE Trans. Power Electron.* **6**, 356–363 (1991).
68. S. K. Mazumder, A. H. Nayfeh, D. Boroyevich, Randomized pulse-width modulation for power electronic converters. *IEEE Trans. Power Electron.* **16**, 701–709 (2001).
69. R. J. Williams, On dynamic scheduling of a stochastic processing network. *Ann. Probab.* **24**, 936–974 (1996).
70. X. Guo, U. Rieder, Average optimality for continuous-time Markov decision processes. *Ann. Appl. Probab.* **16**, 730–756 (2006).
71. E. A. Feinberg, M. E. Lewis, A. G. Piunovskiy, *Continuous-Time Markov Decision Processes: Theory and Applications* (Springer, New York, 2012).
72. E. A. Feinberg, X. Zhang, Optimal switching on and off the entire service capacity of a parallel queue. *Oper. Res.* **63**, 839–852 (2015).
73. S. P. Meyn, *Control Techniques for Complex Networks* (Cambridge University Press, Cambridge, 2020).
74. N. H. Barton, G. Tkacik, Evolution and information content of optimal gene regulatory architectures. bioRxiv [Preprint] (2025). https://doi.org/10.1101/2025.06.10.657849 (Accessed 22 June 2026).