THE UNIVERSITY OF CHICAGO

NONLINEAR DYNAMICAL SYSTEMS DRIVEN AT INTERMEDIATE TIMESCALES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
WEERAPAT PITTAYAKANCHIT

CHICAGO, ILLINOIS
JUNE 2021

To my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to thank my partner Panithi Permsirivanich who is always there listening for me. Without our everyday conversations, I would feel much more isolated in this foreign land. I thanks my mother Chanida for supporting me through difficult times. Her love gives meaning to my existence and the place I feel at home. Even though they are on the other side of the world, I always feel as if they are with me every day.

I would like to thank all my colleagues: especially Aaron, Jackson, Jiayi, Kabir, Martin, Nachi, Ryo, Vedant, Wayne Weng, Weishun, and Zhiyue. The best part about my Ph.D. is to be together with friends who love to discuss science and interesting potential research ideas. I would lose interest and not finish my Ph.D. if it is not for them.

Next, I would like to thank my advisor, Prof. Arvind Murugan, for being a great mentor, teaching most of what I want from a Ph.D. "Research is about the storytelling" remains my most memorable phrase from Arvind. The complexity of scientific research is just the beginning of the challenge. The main obstacle is how we should convince the usefulness of our ideas to other people. I am grateful for Arvind's support both intellectually and financially in following my own research interests.

I thank Prof. Michael Rust, for helping me throughout the biological clock project. Without the welcome supports from Eugene, Justin, Gopal, Soo, and Yi, I would not able to complete the publication and perform experiments on cyanobacteria.

Finally, I would like to thank Prof. Ishanu Chattopadhyay for supporting my work on the influenza data. Even though the time is short, our project gives me crucial data science skills and experience to help me in my future career.

ABSTRACT

Driven nonlinear dynamical systems are everywhere in biology. Powerful mathematical tools like Fourier analysis or the principal of superposition are not available in nonlinear systems. This thesis investigates three different dynamical systems drawn from biology and machine learning as a function of the ratio of the system's response timescale to the signal's timescale.

Chapter 2 analyzes driven circadian rhythms. Many organisms have free running 24 hour internal clocks to estimate the time of the day. These biochemical clocks are self-sustained limit cycles that are 'entrained' by 24 hour external signals such as temperature or light changes. The quality of internal clock time is thus affected by two kinds of noise: fluctuations intrinsic to the chemical dynamics that implement a limit cycle, and fluctuations in the external driving signal such as weather-induced fluctuations in light intensity. We examine all possible clock architectures from a dynamical systems perspective and find that robustness against external noise implies vulnerability to internal noise and vice-versa. We are able to explain the differing clock architectures in different species of cyanobacteria through this mechanism.

Chapter 3 analyzes a stochastic resonance-like phenomenon in mammalian cell signalling. Our experimental collaborators found that the NF-kB signaling pathway in mammalian cells is activated by fluctuations of an external signal (here, concentration of a molecule TNF). In contrast, even 20 times the same total amount of TNF is unable to activate the cells in the absence of fluctuations. We explained these observations using a phenomenological model of rectification; using this insight, we were able to predict NF-kB activation in terms of the timescale of external fluctuations and the timescale of internal rectification processes in the cell.

Chapter 4 introduces a new learning paradigm by viewing the training of neural networks with different streams of data as driven dynamical systems. A core problem in machine learning involves generalizing from training data to test data, i.e., to ignore idiosyncratic

features in the training data that may not be informative in any test data set. Researchers have proposed multiple regularization methods to modify the loss landscape and solve such overfitting problems. However, no universal regularization can solve every data and neural network. This chapter suggests an alternative regularization without modifying the loss function but by simply switching between different data streams during training. We find that generalization is optimal when data streams are switched at a timescale comparable to the timescale of gradient descent.

We have analyzed three separate driven nonlinear systems in regimes where the timescale of the external driving signal is comparable to the timescale of internal responses. The methods of analyses here may not solve every driven nonlinear problem because no universal method likely exists for strongly driven non-linear systems. Nevertheless, the results will help solve future similar challenges.

CHAPTER 1

INTRODUCTION

The field of dynamical systems is simply the study of how things change over time. It covers a broad range of phenomena: from the grand motions of stars in the universe to the minuscule chemical reactions inside bacterias. Understanding how the stars move was the beginning of classical mechanics in the 17th century. A better grasp of chemical reactions inside cells helps scientists understand enzymes and gene functionalities. The study of dynamical systems gives insight into how we should mathematically model these phenomena based on past observations. The field of dynamical systems will continue to be applicable as long as people still want to know how any systems will evolve.

Most systems in nature are nonlinear. Each part of the system often interferes with the others, so we cannot decompose the solution into smaller bits. The second bowl of ice cream does not give us the same pleasure as the first because it already made us less appetized. Since the superposition principle does not apply, critical mathematical tools like Fourier analysis or Laplace transform becomes useless. We can use numerical simulations to obtain an accurate quantitative prediction of the systems. However, we often want to have a qualitative understanding, which leads us to the geometric way of thinking about dynamical systems.

The geometric way of thinking is about interpreting differential equations as a vector field [19]. Even though nonlinear differential equations are rarely solvable, the positions of "fixed points" where the vector fields become zero give valuable qualitative solutions. Each fixed point can be classified into several classes using linear stability analysis. For instance, if all the eigenvalues from the stability analysis are negative, then it is a stable fixed point where all nearby vector fields point inward. It will asymptotically pull in all nearby particles. On the other hand, if all the eigenvalues are positive, then it is an unstable fixed point where all nearby vector fields point outward. The positions and classes of these fixed points allow

researchers to gain qualitative information about the nonlinear system.

However, many dynamical systems do not have fixed equilibrium positions as we described. The classic example of these is the Kuramoto model [20]:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \zeta_i, i = 1..N$$

where N oscillators are coupled together with the coupling constant K. $Theta_i$ and $zeta_i$ are the phase and the fluctuation of the oscillator i. In the simple case that ω_is are the same, when all oscillators are synchronized, $theta_i$ are the same for all i. The system moves in a circle with a steady angular velocity omega, but it is in stable equilibrium because small phase displacements tend to decay back to zero.

Driven dynamical systems

Many dynamical systems relevant to biology are externally driven. Organisms need to process the signals from their surrounding in real-time to adapt to changes in the environment. For example, cells need to measure the chemical concentration of their surroundings, so they have surface receptors that can bind to specific ligand molecules for a certain amount of time. These receptors must compute the average binding time over the irregular arrival of diffusing molecules [21]. A simple setup of this mechanism is:

$$\frac{d\Delta X}{dt} = \frac{1}{\tau_0}(\Delta X - s) \tag{1.1}$$

where $s(t)$ is the concentration of the external ligand and X tracks the moving average of $s(t)$ in the timescale τ_0. τ_0 depends on the receptors' average binding time. Larger τ_0 allows for better noise filtering, but the system is less responsive to the concentration change.

Another example is motivated by a circadian rhythm. A plant has an internal clock that can anticipate the time of sunrise. However, the weather can delay the necessary sunlight

for several hours. So, the plant needs to correct its clock with the Sun as much as possible. A circadian rhythm can be modelled with the Stuart-Landau equation:

$$\frac{dz}{dt} = Az - bz^2 + f(t)z \tag{1.2}$$

where $f(t)$ is the external driving signal and z is the position in a complex plane [22, 23]. This equation captures the behavior of any non-linear oscillator in the weakly driven limit. The force amplitude and the period difference dictate the speed of approaching synchronization.

This thesis presents the progress on three different driven dynamical systems using our analytical process: comparing the system's response time to the signal's time required for the change. Suppose the system is in an equilibrium state when the driven signal is F_0. When the signal changes to F_1, the 'response time' τ_0 is the time the system takes to return close to the equilibrium. The 'signal time' τ_f is the amount of the time the signal takes to change from F_0 to F_1. Surprising nonlinear behavior often occurs when $\tau_0/\tau_f \sim 1$, defined as the intermediate timescale. Our analytical process begins by first analyzing the system at the two opposite limits: $\tau_0/\tau_f \ll 1$ and $\tau_0/\tau_f \gg 1$. These analyses hint at the qualitative solution when $\tau_0/\tau_f \sim 1$. Our method may not apply to every nonlinear system, but it is practical because driven nonlinear systems are challenging and do not have a universal mathematical tool.

Chapter 2 serves as the first application of our method to solve the driven dynamical system. Here, we study the dynamical system model of the clock inside cyanobacteria. The model consists of a limit cycle attractor, driven by a time-dependent force. The crucial response time τ_0 is the time it takes for the phase difference between the system and the periodic input to go down by a constant factor, or the entrainment time. The higher the input force, the smaller the entrainment time.

There are two kinds of noises affecting the circadian rhythms: internal and external noise. Internal noise comes from the small copy numbers of the clock protein, while external noise

comes from the daylight's fluctuation. From rationalizing on different timescales, we show that the clock is vulnerable to internal noise when the entrainment time is high and vulnerable to external noise when the entrainment time is low. The robustness of the biological clock depends on the ratio of the internal and external noise. This chapter is the reprint of [24].

Chapter 3 is about the noise correlation time in a noise-induced activation system. The significant timescale in this system is the timescale T of the adaptive circuit. If the noise correlation time is much shorter than T, then the signal will be perceived as a constant signal. On the other hand, if this correlation time is much larger than T, then the system is fast enough to adapt and filter out the noise. Intriguing phenomena of noise-induced activation happens when the correlation time is about the same as the adaptation time. This chapter is the reprint of [25].

Chapter 4 is about the gradient descent of a neural network. The classification of texts and images has been advancing rapidly and becoming a part of our everyday technology. For example, driverless cars have to distinguish a road from a building. Also, the email services become safer as they can often separate normal texts from scams. Neural networks play an important role to the classification problems. The biggest problem of using data to train neural networks is overfitting. Most popular methods apply static regularization to the loss function such as L1 and L2 norms. The training will repeat on the same training dataset until reaching the desirable cross-validation accuracy. In this nonlinear system, the number of passes on the entire dataset (the epoch) is the response time of the neural network. Our research finds that swapping the stream of training data at around the response time can help train the parameters toward the minimum, common among all datasets.

CHAPTER 2

BIOPHYSICAL CLOCKS FACE A TRADE-OFF BETWEEN INTERNAL AND EXTERNAL NOISE RESISTANCE

Many organisms use free running circadian clocks to anticipate the day night cycle. However, others organisms use simple stimulus-response strategies ('hourglass clocks') and it is not clear when such strategies are sufficient or even preferable to free running clocks. Here, we find that free running clocks, such as those found in the cyanobacterium *Synechococcus elongatus* and humans, can efficiently project out light intensity fluctuations due to weather patterns ('external noise') by exploiting their limit cycle attractor. However, such limit cycles are necessarily vulnerable to 'internal noise'. Hence, at sufficiently high internal noise, point attractor-based 'hourglass' clocks, such as those found in a smaller cyanobacterium with low protein copy number, *Prochlorococcus marinus*, outperform free running clocks. By interpolating between these two regimes in a diverse range of oscillators drawn from across biology, we demonstrate biochemical clock architectures that are best suited to different relative strengths of external and internal noise.

Extracting information from a noisy external signal is fundamental to the survival of organisms in dynamic environments [26]. From yeast anticipating the length of starvation [27] and bacteria estimating sugar availability[28], to dictyostelium counting cAMP pulses [29], organisms must often infer properties of the environment that are masked by noisy irregular fluctuations in order to be well-adapted [30, 31].

A striking example of regularity in environmental stimuli is the daily day-night cycle of light on earth; organisms from all kingdoms of life use circadian clocks to synchronize - or 'entrain' - in phase to these 24-hour periodic signals in order to anticipate and prepare for future changes in light [12]. The most remarkable and well-studied examples of clocks are free running circadian clocks, found in organisms ranging from the cyanobacterium *S. elongatus* to insects, plants and humans. Such clocks use non-linear dynamics to generate

self-sustained 24-hr rhythms of a preferred amplitude even in the absence of external driving. Many salient properties have been ascribed to such free running internal rhythms [32, 12].

However, several organisms have only damped clocks or 'hourglass clocks'; their response to daily changes in light is not a self-sustaining oscillation, but rather a physiological program that decays to a steady state over a day. For example, some strains of *P. Marinus*, a smaller $0.5\mu m$ cyanobacterium with an estimated $50\times$ smaller protein copy number than *S. elongatus* [33, 34, 35, 36, 37], appear to have such a damped 'hourglass' clock, despite the clock being constituted from Kai proteins similar to those in *S. elgonatus*.

The potential benefits and drawbacks of these timing systems are not immediately obvious. In particular, it is unclear when an 'hourglass' clock might be sufficient or even preferred over free running clocks.

Here, we compare such classes of clocks when driven by the day-night cycle of light in fluctuating conditions. One source of fluctuations are amplitude fluctuations in the external day-night signal due to weather patterns[14] or other environmental disturbances. Phase entrainment to such fluctuating environmental signals is a challenge because while amplitude fluctuations are uninformative of phase, an entrainment mechanism looking for dawn-dusk transitions might conflate such amplitude fluctuations with true variations in phase. Biomolecular clocks also face an internal source of fluctuations [38], due to various causes like finite copy number effects [39], bursty transcription, interactions with the cell cycle and cell division [40]. It is clear that the inability to deal with either of these fluctuations will lead to poor phase entrainment, with a host of associated fitness costs in cyanobacteria [41], plants, rodents and humans[42]. However, it is not clear what kinds of clock architecture are best at dealing with internal and external fluctuations and whether these demands are compatible.

We find that free running clocks, based on limit cycle attractors, are a double edged sword when subject to such internally and externally fluctuating conditions. The flat direc-

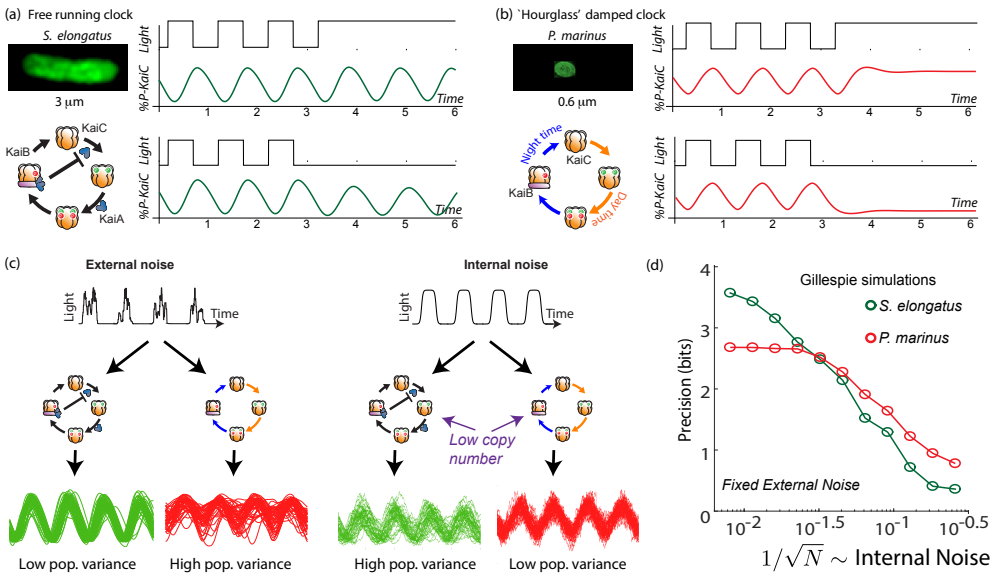

Figure 2.1: Free running clocks and damped 'hourglass' clocks are equally good time-keepers in noiseless conditions but internal and external fluctuations reveal significant differences. (a) Free running circadian clocks, such as the KaiABC protein clock in *S. elongatus*, show rhythms in both oscillating and constant light (top) or dark (bottom) conditions. (b) In contrast, damped circadian clocks, such as that in *P. marinus* which lacks Kai A, show rhythms only in changing light conditions and decay to a fixed state in constant conditions. (c) When subject to external noise (i.e., weather-related amplitude fluctuations in light), simulations of the free running clock show low population variance while the damped clock shows high variance. In contrast, Gillespie simulations with high internal noise due to low copy number of Kai molecule reveals that damped clocks are much more robust than free running clocks. (d) A systematic study of clock precision (i.e., mutual information between clock state and time) at fixed external noise but decreasing Kai protein copy number N reveals that free running clocks are preferred at low internal noise but damped clocks are preferable at sufficiently high internal noise.

tion along such continuous limit cycle attractors can selectively project out external amplitude fluctuations while retaining phase information. However, the flat direction along the attractor makes these continuous attractor-based clocks susceptible to internal fluctuations (e.g. low protein copy number [43]). In contrast, point attractor-based damped clocks are relatively resistant to internal fluctuations because they have no flat directions. Hence such 'hourglass' clocks out-perform free running clocks at sufficiently high internal noise.

We first demonstrate our results in diverse biochemical oscillators, drawn from the literature [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 44] on clocks in cyanobacteria, plants and mammals to

cell cycle and synthetic oscillators. We complement this detailed network-based study with dynamical systems theory that explains the same trade-off in terms of the broad features common to the diverse models studied here. In all cases, our approach involves systematically deforming the dynamics to interpolate between free running and 'hourglass' clocks and using information theoretic measures to quantify clock performance in the presence of fluctuations.

By continuously interpolating between these clock architectures, our work predicts that a survey of clock systems in different environmental niches will reveal that clock architecture vary systematically with the relative strength of external and internal fluctuations[45]. Further, our work suggests intriguing forward evolution experiments in the lab where the same structured external environment can nevertheless result in distinct regulatory systems, depending on the size of internal fluctuations. Finally, the existence of 'hourglass' clocks are easier to overlook experimentally than free running oscillations. Hence our theoretical demonstration that 'hourglass' clocks have functional benefits in specific conditions highlights the importance of experiments that specifically look for such damped clocks. More broadly, our work highlights the need to experimentally probe regulatory strategies by varying different kinds of noise independently when possible, since the strategies to deal with different kinds of noise are not equivalent and can be in conflict.

2.1 Free running clocks and damped 'hourglass' clocks

Many organisms like humans and rodents have free running clocks that show self-sustained 24 hr rhythms even in constant dark or light conditions. A particularly simple and well-characterized free running clock is that found in *S. elongatus* where the clock dynamics can be reproduced by the post-translational dynamics of Kai ABC *in vivo* as well. Measuring the phosophorylation state at any one of several sites on KaiC reveals an orderly phosophorylation reaction with a period of 24 hours. As shown in Fig 2.1a, oscillations of a characteristic

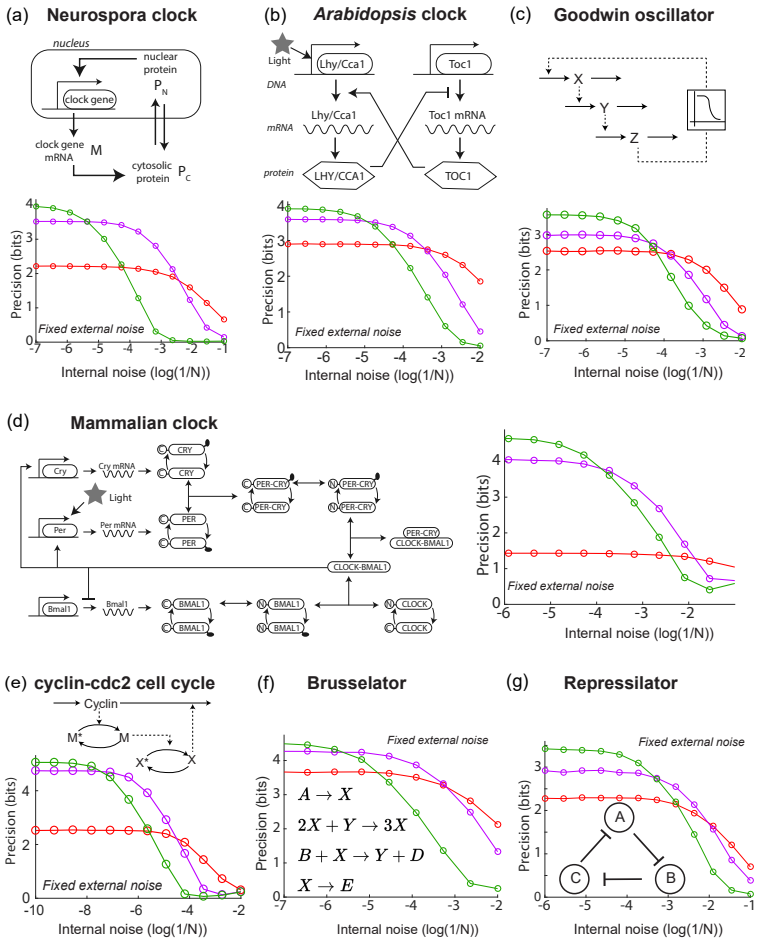

Figure 2.2: A diverse range of biochemical oscillators show the trade-off between resistance to external and internal noise. For each oscillator, the regime (green) of largest free running amplitude relative to the driving strength is most robust to external fluctuations but is most fragile to internal noise. In contrast, damped oscillations (red) are robust to internal noise and thus preferable at sufficiently high internal noise. Regimes (purple) of intermediate free running amplitude are preferred at intermediate internal noise levels. (a-g) Diverse biochemical oscillators from the literature were simulated with increasing internal noise $\epsilon_{int} = 1/\sqrt{N}$ while driven by a periodic square wave light signal with fixed strength external noise, using the external coupling and parameters specified in the original publications [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. Clock precision is defined as mutual information between outputs and time. The original publications identified a Hopf bifurcation parameter in these models, with free running oscillations on one side and damped oscillations on the other. Green and purple data correspond to parameter regimes with large and smaller amplitude free running oscillations relative to driving amplitude while the red data corresponds to a damped oscillator. Details in Appendix A.1.

amplitude are sustained even in constant darkness or constant light, i.e., in the absence of a periodic external drive.

Not all organisms have a free-running clock; e.g., many insects [13] have damped 'hour-glass' clocks that decay to a fixed point under constant light or constant dark conditions but show oscillatory dynamics under day-night cycling (see Fig 2.1b). In fact, a sister cyanobacterial species *P. marinus* has a KaiBC-protein based clock. While the details of this clock are not fully characterized, the clock lacks the KaiA-mediated negative feedback [36, 35] loop that enables free running oscillations in *S. elongatus*. Consequently, in constant light or dark conditions, the clock's state decays to a distinct day or a night state respectively [35].

Thus, both classes of clock show regular oscillations when externally driven. With cloudless day-night cycling, both systems can synchronize in phase with the external signal (i.e., 'entrain') and show distinct clock states at distinct times of the day. In this way, the clock state provides the rest of the cell with an estimate of the time of the day. However, while the free running clock has a natural amplitude relatively independent of the external drive, the damped clock's amplitude is directly set by the external drive.

External fluctuations:

The day-night pattern of light on earth does not resemble the clean square wave shown in Fig.2.1a but is rather subject to large amplitude fluctuations during the day due to weather patterns. Such amplitude fluctuations and their spectrum have been quantified [14] and also identified as playing a critical role in several studies on the evolution and performance of circadian clocks [46, 32]. The biological impact of such changes in light intensity in cyanobacteria have been quantified recently [40]. The clock must entrain in phase to the external signal while ignoring such amplitude fluctuations.

Internal fluctuations In addition to external fluctuations, circadian clocks also deal with the intrinsically noisy nature of biochemical reactions[47]. Sources of internal noise for clocks include finite copy number effects, bursty transcription, cell division and other

sources[39]. In particular, based on their relative sizes[36, 35, 33], *P. marinus* is thought to have far fewer copies of the Kai clock proteins (e.g., $\sim$ 500 of KaiC) than *S. elongatus* ($\sim O(10000)$ copies of KaiC [34, 37]). Such finite numbers of molecules is known to create significant stochasticity in oscillators in the absence of an external signal [43].

Noise resistance of Kai-based clocks:

We tested the impact of such external and internal fluctuations on the contrasting clock architectures in *S. Elongatus* and *P. Marinus* through simulations. We set up explicit Gillespie simulations[48] of explicit biomolecular models of the post-translational Kai clock that captures the known biochemistry[15] of *S. elongatus*'s clock and the putative KaiBC clock[33, 35] in *P. marinus* (Fig.2.1). We do not include transcriptional coupling [49] of the clock here and focus on the core post-translational oscillator. See Appendix 5 for details. The ATP levels in these models [50] were were coupled to an external square wave input of period 24 hours, representing the day-night cycle of light. To model external fluctuations, we modulated the amplitude of the input square wave over a broad range of frequencies, reflecting the broad frequency spectrum quantified by the Harvard Forest database [51]. To model internal fluctuations, we varied the copy number in these Gillespie simulations.

With only external fluctuations but suppressing internal fluctuations using high copy numbers, we find that the damped oscillator develops a much larger population variance than the free running clock. In contrast, at low copy number (i.e., high internal noise) but with a noiseless external signal, we find the situation is reversed; the free running clock has significantly higher population variance. See Fig.2.1c.

To study this effect quantitatively, we fixed the strength of amplitude fluctuations and increased the internal noise by decreasing the copy number of all Kai proteins in the Gillespie simulation. We measured the resulting mutual information between clock state and objective time of day. (Mutual information is intuitively a measure of population variance along the most informative directions; see Appendix A.3 for more.)

We see that the free running clock has higher precision than the damped clock at low internal noise (high copy number). However, as the internal noise is increased, the precision of the free running clock drops earlier and consequently, the damped oscillator has higher precision at sufficiently high internal noise (low copy number). This is shown in Fig. 2.1d, where the precision measures the mutual information between the clock state and the time. For a fair comparison, in undriven conditions, different clocks are assumed to lose information at the same rate.

2.2 Noise resistance in other biochemical clocks

While our study here was motivated by the contrasting Kai protein-based clocks in the two cyanobacterial species *S. Elongatus* and *P. Marinus*, we sought to test the broader validity of our results. Hence we analyzed the internal and external noise resistance in a range of eight well-studied biochemical oscillators in the literature.

These models range from circadian clocks in numerous organisms - *Neurospora*[1], *Arabidopsis*[2, 3], mammalian cells [4]) - to other oscillators such as cell cycle models [5], the Goodwin [6, 7] oscillator, the Brusselator [8] and the synthetic repressilator [9, 10] - see Fig.2.2. While the internal noise properties of these oscillators in undriven conditions have been studied before [52], here we analyzed the contrasting internal and external noise properties of these oscillators under externally driven conditions. The results are shown in Fig.2.2.

In each case, we set all kinetic parameters to values specified in the original publications and coupled the external driving signal in the way specified in those original publications. As in the Kai clock simulations, the external signal was a square wave with amplitude fluctuations of fixed strength. Finally, we add Langevin noise to the equations to simulate internal noise; when available, we followed the finite volume prescription for rates in the original publications or related papers to set the size of Langevin noise for each reaction. Simulation and model details are in the SI.

These models here all exhibit a Hopf bifurcation as kinetic parameters are tuned. The publications [1, 2, 3, 4, 5, 6, 7, 8, 10] identified a parameter which when tuned leads to a Hopf bifurcation; i.e., on side of the bifurcation, we find damped oscillations while on the other side, we find free running oscillations of increasing amplitude. We picked three points along this parameter; the green and purple data correspond to free running oscillations of large and smaller natural amplitude relative to the size of the external drive. The red data corresponds to a choice of parameters on the other side of the Hopf bifurcation, i.e., to damped oscillations. For the red data, we chose μ such that the relaxation timescale was comparable to the period of the external driving force, much as in the Kai model of *P. Marinus*. The damped oscillator is a useful clock only when the relaxation timescale is comparable to the period.

In each case, we observed the same trade-off as seen in the Kai system; free running oscillations of large amplitude relative to the external drive (green) were most precise when only subject to external noise but are most fragile to internal noise. Damped oscillations in the same oscillator models are more robust and thus are preferable at sufficiently high internal noise. We find that intermediate amplitude free running oscillations show intermediate noise properties. Consequently, we can continuously trade-off resistance to internal noise for resistance to external noise by changing the amplitude of free running oscillations relative to the strength of the external drive.

2.3 Dynamical systems theory of noise resistance

We have demonstrated a trade-off between external and internal noise resistance in diverse clocks. While it is possible to trace the origin of this trade-off to specific features of each clock, here, we take a different approach based on dynamical systems theory. Dynamical systems theory has been use to make fruitful general predictions about biological clocks since Winfree's analysis of phase singularities [12]. In a similar vein, we use dynamical systems

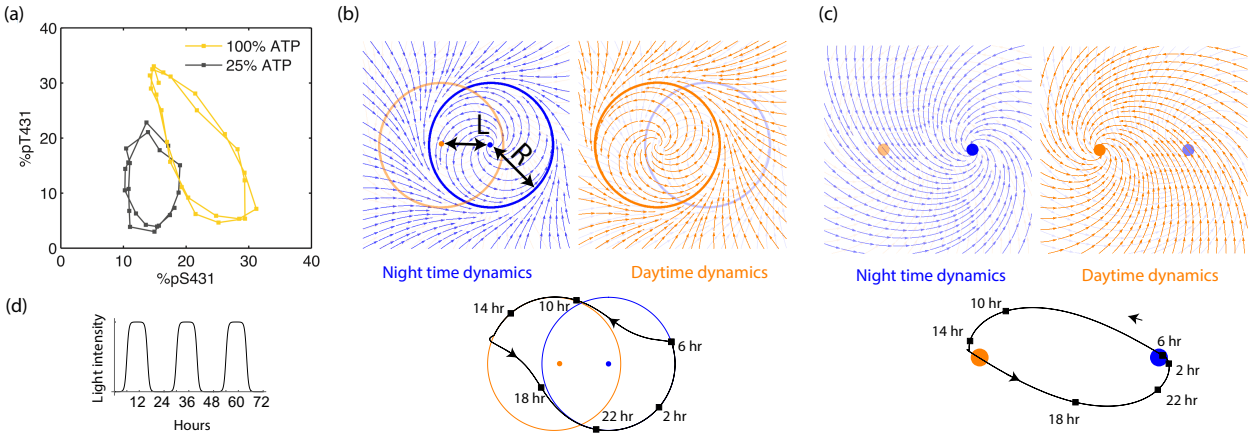

Figure 2.3: Experiments and models of biological clocks show that external driving can be viewed as a switch between distinct day-time and night-time dynamics. (a) Experiments on the Kai system at distinct ATP levels corresponding to day and night conditions reveal limit cycles shifted relative to each other in a phosphorylation space for Kai (reproduced from [11]). Similar behavior[12] is seen in models of diverse biochemical oscillators studied in Fig.2.2. (b) We build a minimal model of such driven clocks as a limit cycle of radius R whose center is shifted by a distance L between day and night. In cycling conditions (see signal in (d)), an entrained clock's state executes a trajectory that encompasses both limit cycles as shown (bottom). (c) For damped clocks [13], phenomenology suggests that the day and night limit cycle dynamics are replaced by a point attractor whose position changes between day and night. The relaxation time between the day and night attractors is comparable to $\sim$ 12 hours, giving rise to the trajectory shown in cycling conditions. (d) The plot shows cycling conditions of light intensities that couple to (b) and (c).

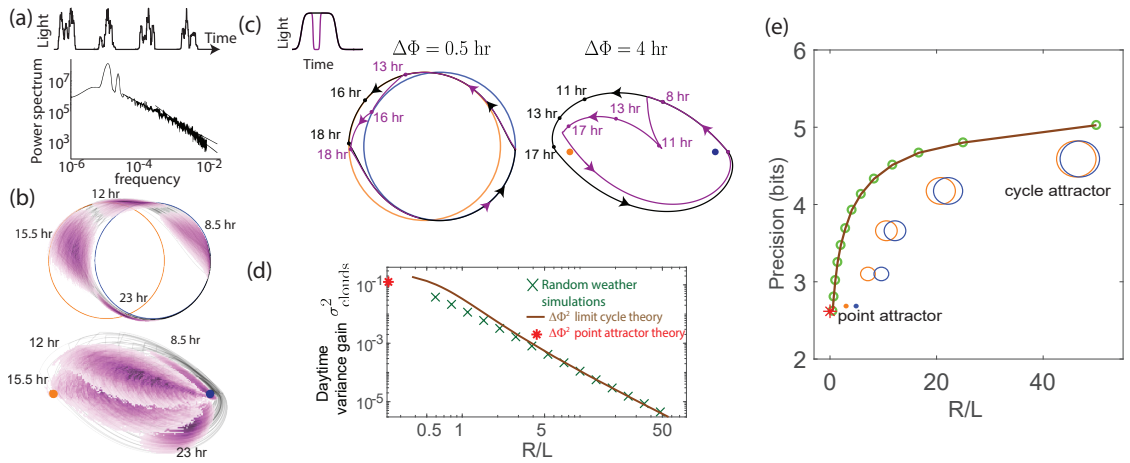

Figure 2.4: External weather-related light fluctuations are filtered out by limit cycle attractors but not by point attractors. (a) Light intensity levels fluctuate on a range of time scales due to weather (power spectrum reproduced from [14]). (b) A population of limit cycle clocks of identical fixed geometry, subject to different realizations of weather conditions, show non-overlapping distributions (purple blobs) at different times of the day. Point attractor clocks form larger and more overlapping distributions. (c) A single representative dark pulse of ~ 2.4 duration causes only a $\Delta\Phi \sim 30$ min phase lag in limit cycles since the trajectory's deviation (purple) is fundamentally bounded by the circular attractor. In contrast, $\Delta\Phi \sim 4$ hr for the point attractor since the trajectory is in free-fall towards the blue night-time attractor. (d) The geometrically computed $\Delta\Phi^2$ phase shift for a dark pulse of any fixed duration and time of occurrence (see Appendix A.4) drops rapidly as $(R/L)^{-2}$ for large-R/L limit cycles; this theoretical prediction agrees well with the population variance gain over a day in simulations. (e) Consequently, weakly driven limit cycles (i.e., high R/L) can tell time with high precision.

theory to show this trade-off emerges due to basic features of free running and damped clock dynamics and can thus be expected to hold broadly.

2.3.1 Limit cycle clocks and point attractor clocks

Free running clocks are phenomenologically well-described by a limit cycle attractor, a nonlinear oscillator of fixed amplitude[12]. While such descriptions have been used for numerous biochemical oscillators, limit cycle dynamics can be experimentally seen in molecular detail for the KaiABC clock in *S. elongatus* as shown in Fig.2.3a (reproduced from [11]). The clock follows distinct limit cycle dynamics during the day (orange data) and night (black

data) [11, 50], with the day cycle positioned at higher phosphorylation levels due to a higher ATP/ADP ratio.

The Kai model and indeed the diverse range of biochemical oscillators in Fig.2.2 show such a change in the limit cycle between day and night conditions. Here, we build a minimal model of such free-running clocks using circular day and night limit cycles of radius R in a plane. The limit cycle is defined by the dynamics $\tau_{relax}\dot{r} = r - r^3/R^2, \dot{\theta} = \omega$ about its own center; but the center of the limit cycle itself moves along the x axis in Fig.2.3b as $(\rho(t)L, 0)$ where $\rho(t) \in [0, 1]$ is the normalized light intensity level at time t and L is a measure of the physiological changes between day and night (e.g., ATP/ADP ratio change in *S. elongatus*). Thus, e.g. in Fig.2.3b, the system follows the blue dynamics at night and then after dawn it relaxes to the orange day attractor on a time scale τ_{relax}. Note that R represents the amplitude of free-running oscillations while L represents the strength or amplitude of the external driving signal.

In contrast, damped clocks are phenomenologically well-described by a day-time and a night-time point attractor with slow relaxation dynamics between them (Fig. 2.3c), modeled as $\dot{r} = -r/\tau_{relax}, \dot{\theta} = \omega$ about an attractor point whose location varies with current light levels as $(-\rho(t)L, 0)$. Here we assume $2\tau_{relax} \sim 24$ hrs as in *P. marinus* [35]; if relaxation were faster and completed before the day is over, the clock cannot resolve all times of the day.

Here, we will also consider a family of limit cycle clocks of varying R/L to interpolate between large-R/L limit cycles and point attractors (approximated by $R/L = 0$).

2.3.2 External noise

We begin with the performance of different clocks in the presence of external intensity fluctuations. Weather patterns cause large fluctuations in the intensity of light over a wide range of time-scales as shown in Fig.2.4a. Much like with biochemical circuits, we subject an *in*

silico population of dynamical system clock models to different realizations of such noisy weather patterns.

When subject to weather fluctuations, we see in Fig.2.4b that the population variance of clock states for limit cycles at given times (purple) is fundamentally limited by the spacing between the day and night limit cycles. Point attractors develop larger overlapping population distributions at different times.

We can geometrically understand the daytime phase variance increase σ^2_{clouds} in terms of the phase lag $\Delta\Phi$ due to a single, say 2.4 hr dark pulse, administered during the day. Fig.2.4c shows that the deviation in trajectory for limit cycle clocks (purple) is fundamentally limited by the presence of a continuous attractor. In contrast, for the point attractor, the trajectory is in free fall towards the night point attractor, with no limit cycle to arrest such a fall. Consequently, the geometrically computed phase shift $\Delta\Phi$ due to the particular dark pulse shown in Fig.2.4c is much smaller for limit cycles ($\Delta\Phi \sim 0.5$ hr for the R, L geometry shown) than for point attractors ($\Delta\Phi \sim 4$ hr) (see Appendix A.4).

In fact, this contrast in $\Delta\Phi$ between limit cycles and point attractors holds for dark pulses of any duration and time of occurrence. The contrast is even greater at small L/R since $(\Delta\Phi)^2 \sim (L/R)^2$ for small L/R, as shown geometrically in the Appendix A.4 and confirmed in simulations that average over random weather conditions (Fig.2.4d). Hence, limit cycles are more resistant to external fluctuations than point attractors.

To complete the analysis, we note that phase variance increases additively during the day and falls multiplicatively at dusk (and dawn), i.e.,

$$\sigma^2 \xrightarrow{day} \sigma^2+\sigma^2_{clouds} \xrightarrow{dusk} (\sigma^2+\sigma^2_{clouds})/s^2 \xrightarrow{night} (\sigma^2+\sigma^2_{clouds})/s^2 \xrightarrow{dawn} (\sigma^2+\sigma^2_{clouds})/s^4$$

. Solving for steady state phase variance ($\sigma^2 = (\sigma^2 + \sigma^2_{clouds})/s^4$), we obtain

$$\sigma^{2,ext}_{limit\ cycle} \sim \Delta\Phi^2/(s^4 - 1). \tag{2.1}$$

where we have equated σ^2_{clouds} to $\Delta\Phi^2$ for a typical dark pulse Here, s^2 represents the variance drop during a dawn/dusk entrainment. As shown in the Appendix A.4 for external noise (and in Fig.2.5 for internal noise), this factor s, can be geometrically explained by the slope of the circle map relating the two cycles [11]; we find that $s^2 - 1 \sim L/R$ for large-R/L limit cycles. Plugging this and $\Delta\Phi^2 \sim (L/R)^2$ into Eq.2.1, we see that $\sigma^2 \to L/R \to 0$ for large-R/L cycles.

Fig.2.4e shows that the precision (i.e., mutual information between clock state and time) computed from random weather simulations agrees with this theory; clock precision drops as we interpolate from limit cycles to point attractors by changing L (with equivalent results for changing R).

2.3.3 Internal noise

Internal noise due to finite copy number effects in biochemical networks can be modeled exactly using the Gillespie method used in Fig.2.1. In the context of our dynamical systems model, we follow [48] and add Langevin noise to all dynamical variables of the system of strength $\epsilon_{int} \sim 1/\sqrt{N}$, where N is the overall copy number, with the ratios of different species assumed fixed (see Appendix A.2.7). Such a Langevin approach is an approximation [48] to the exact Gillespie method used in Fig.2.1.

We simulated a population of clocks in externally noiseless day-night light cycles but with internal Langevin noise. We see in Fig.2.5b that limit cycle populations have significantly higher variance of clock state due to internal noise than point attractors, in contrast to Fig.2.4b with external noise alone.

We can understand the weakness of limit cycle attractor relative to the point attractor in terms of diffusion during day/night balanced by dawn/dusk transitions. The flat direction along the limit cycle attractor cannot contain diffusion caused by the Langevin noise during the day/night; hence the phase variance increases by $\sigma^2 \to \sigma^2 + \epsilon^2_{int} T_{day}$ during a day of

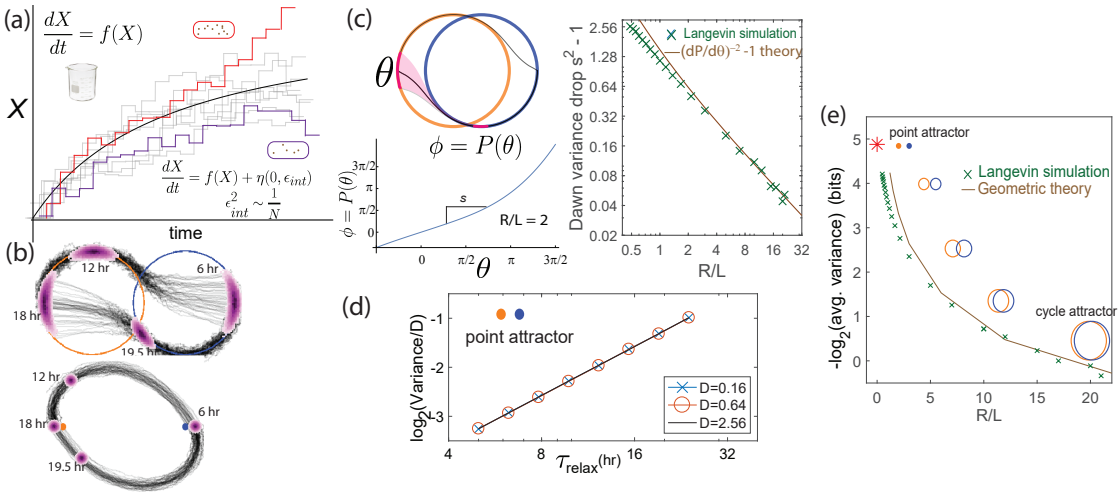

Figure 2.5: Internal fluctuations severely affect continuous attractors but not point attractors. (a) We model fluctuations due to finite copy number N as Langevin noise with mean zero and standard deviation ϵ_{int}, resulting in a diffusion constant $\epsilon_{int}^2 \sim 1/N$ for the clock state. (b) The flat direction of limit cycles cannot contain diffusion, leading to large increases $\epsilon_{int}^2 T_{day}$ in population variance of clock state during each day (and night). In contrast, point attractor dynamics have constant curvature at all times, leading to a constant population variance over time. (c) The variance drops $\sigma^2 \to \sigma^2/s^2$ at dawn and dusk for limit cycles during the off-attractor dynamics between the day and night cycles. As with external noise, the variance drop is predicted by the slope $dP(\theta)/d\theta$ of the circle map between the cycles. This dawn/dusk drop goes to zero for large R/L limit cycles but variance still increases during the day and night. (d) The variance for point attractors is $D\tau_{relax}$, a constant determined by the curvature τ_{relax}^{-1} of the harmonic potential. (e) Thus, with only internal noise present, the precision of limit cycle clocks increases with increasing separation L/R, asymptotically approaching the performance of point attractors.

length T_{day} (and similarly at night).

Dawn and dusk times reduce the phase variance $\sigma^2 \to \sigma^2/s^2$ as the trajectories originating on, say, the day cycle converge on the night cycle (see Fig. 2.5c and [11, 53]). In fact, we can compute this variance drop s^2 entirely through geometric considerations. We define the circle map $\phi = P(\theta)$ as relating originating points θ near dusk on the day cycle to final points on the night cycle ϕ after relaxation (experimentally characterized in [11]). Then $s^{-1} = dP(\theta)/d\theta$. Fig.2.5c shows that this slope $s^{-1} = dP(\theta)/d\theta$, geometrically computed in the SI, agrees with the dawn/dusk variance drop in Langevin simulations and scales as $s^2 - 1 \sim L/R$ for large R/L.

Thus, the population phase variance changes as $\sigma^2 \xrightarrow{Day} \sigma^2 + \epsilon_{int}^2 T_{day} \xrightarrow{Dusk} (\sigma^2 + \epsilon_{int}^2 T_{day})/s^2 \xrightarrow{Night} \ldots$ where the night adds another $+\epsilon_{int}^2 T_{night}$ and so on. Assuming $T = T_{day} = T_{night}$ and solving for steady-state phase variance ($\sigma^2 = ((\sigma^2 + \epsilon_{int}^2 T_{day})/s^2 + \epsilon_{int}^2 T_{day})/s^2$), we obtain

$$\sigma_{cycle}^{2,int} \sim \frac{\epsilon_{int}^2 T}{s^2 - 1} \tag{2.2}$$

Consequently, as the cycles become large (large R/L), the dawn/dusk variance drop vanishes as $s^2 - 1 \sim L/R \to 0$ while diffusion along the flat direction still adds $\epsilon_{int}^2 T$ to the variance during each day and each night; hence large-R/L limit cycles have large $\sigma_{cycle}^{2,int}$ and thus low precision. (Unlike with external noise, internal noise introduces a diffusion length scale and hence changing L and R are not equivalent. To make a fair comparison, we fix R and internal noise while changing L in Fig. 2.5e; see Appendix A.2.7 for more detail about other equivalent choices).

Note that Eqn.2.2 is invalid for strictly undriven clocks (i.e., $s = 1$); such clocks show a variance that increases indefinitely with time. Here, we focus on driven clocks, which always settle to the finite variance given by Eqn.2.2.

In contrast, for point attractors, the population variance stays constant during the day-night cycle and is shown to be

$$\sigma_{point}^{2,int} \sim \epsilon_{int}^2 \tau_{relax}$$

in the SI, which matches Langevin simulations (Fig.2.5d). Since $\tau_{relax} \sim T_{day}$ to have distinct clock states through the day (Fig.2.3)), we find $\sigma_{cycle}^{2,int} \geq \sigma_{point}^{2,int}$.

In summary, in both cases, population variance is reduced by the geometric 'curvature' of the dynamics, i.e., convergence of nearby trajectories. Point attractor trajectories experience a constant curvature of $1/\tau_{relax}$. But limit cycle clocks experience such 'curved' off-attractor dynamics only at dawn and dusk, which is offset by dephasing[54, 52] during long periods of zero curvature on the limit cycle (day/night). Hence limit cycles underperform point

attractors under high internal noise.

2.3.4 Combination of external and internal noise

We now subject the clock systems to both internal and external noise at the same time. We find results (see Fig.2.6a) that parallel those for explicit molecular models of biochemical oscillators studied in Fig.2.2. Large-R/L limit cycles outperform other clocks in filtering out external noise when internal noise is low but their precision degrades more rapidly than other clocks as internal noise $\epsilon_{int}^2 \sim 1/N$ is increased. Point attractors have poor precision with only external noise but do not significantly degrade with internal noise and outperform all other clocks at high internal noise. At comparable strengths of internal and external noise, limit cycles with an intermediate value of R/L are most precise. In the SI, we show that the optimal geometry is set by the ratio of internal and external noise strength,

$$(L/R)_{optimal} = \frac{\epsilon_{int}}{\epsilon_{ext}}. \tag{2.3}$$

In the SI, also we show that, under certain simplifying assumptions, Eqns. 2.1 and 2.2 can be combined to give an explicit trade-off relationship,

$$\sigma_{int}^2 \sigma_{ext}^2 \sim Q \tag{2.4}$$

where $Q = \epsilon_{int}^2 \epsilon_{ext}^2$ and where σ_{int}^2 is the population angular variance of the clock state due to internal noise when driven by a noiseless external signal and σ_{ext}^2 is the population angular variance in the absence of internal noise due to amplitude fluctuations of the external signal. Note that angular variance is a better indicator than variance because we want to know how well the system can tell time.

Eqn.2.4 makes our trade-off explicit and also clarifies which parameters are varied and which parameters are held fixed in this trade-off. As long as Q is held fixed, we allow all

other parameters to vary – e.g., the overall strength of the external drive L, the size of the cycle R, and as discussed in the SI, all other parameters characterizing the normal form of limit cycles near a Hopf bifurcation.

However, in holding Q fixed, our trade-off does assume that the strength of the external fluctuations ϵ_{ext} – i.e., the fractional size of amplitude fluctuations in the external signal – is held fixed. Similarly, we hold ϵ_{int}^2, the phase diffusion constant, fixed – i.e., we are comparing clocks that would show the same population phase variance (in units of radians) over the same time in undriven conditions. See Appendix A.2.7 for alternative comparisons and other details.

2.3.5 Speed-precision trade-off

Another measure of clock quality is the entrainment speed, i.e., the time taken to reach steady state population variance, starting from a population uniformly distributed in clock phase. In Fig.2.6b, we see that with external noise only, the most precise clocks (i.e., small-L/R limit cycles) are the slowest to entrain because they retain a longer history of the external signal, allowing them to average out external noise better.

But strikingly, such a speed-precision trade-off is absent if internal noise is high. In Fig.2.6c, only internal noise is present and the external signal has no fluctuations. We see that clocks most robust to internal noise are also the fastest to entrain. Intuitively, the phase of slow entraining clocks is affected by the cumulative effect of internal fluctuations over a longer period of time. With both external and internal noise present, clocks with intermediate entraining speed - i.e., intermediate $(L/R)_{optimal} = \epsilon_{int}/\epsilon_{ext}$ - will have the highest precision.

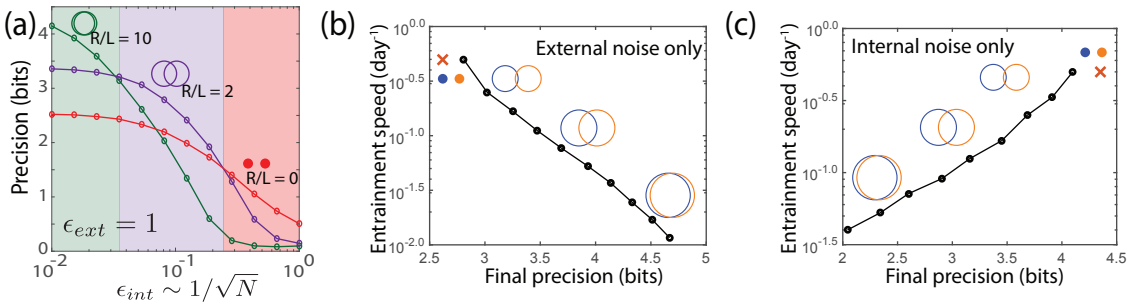

Figure 2.6: Large-R/L limit cycle attractors, which correspond to large amplitude free running clocks, outperform all other oscillators in projecting out external noise but are least robust to internal noise. (a) Point attractors and smaller R/L limit cycles (red and purple curves) show low precision (i.e., low mutual information) but do not degrade as much as large-R/L limit cycles with increasing internal noise ϵ_{int}. Thus this simple dynamical systems model of clocks reproduces and explains the trade-off seen in the complex biochemical clocks shown in Fig.2.1,2.2. (b,c) Speed-precision trade-off. (b) With external noise alone, the most precise clocks (i.e., large R/L limit cycles) average over longer signal history and are thus the slowest to entrain, i.e, slow to transform a population with uniform phase distribution to the steady state distribution. (c) However, with internal noise alone, there is no trade-off between speed and precision; faster entraining clocks (i.e., point attractors) are more accurate since slow clocks are exposed to more internal noise.

2.4 Discussion

Free running circadian clocks are a remarkable result of evolution in a changing but predictable environment and are thought to provide numerous benefits[32]. Here, we showed that the limit cycle attractor underlying such a clock is able to effectively project out weather-related amplitude changes that are perpendicular to the flat direction of the attractor. Similar roles for the flat direction of continuous attractors in projecting out external (or input) fluctuations have been explored in neuroscience [55, 56], e.g., for head and eye motor control and spatial navigation. However, we also found that the same flat direction becomes a vulnerability with internal fluctuations since such fluctuations cannot be restricted to be perpendicular to the attractor.

We confirmed the trade-off between resistance to external and internal noise in diverse models of biochemical clocks and oscillators, using parameters inferred from experimental data by the original publications [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. The trade-off in each of these

models can be given explanations that are specific to those models; for example, one can identify specific bottlenecks for external and internal noise in these models[57]. However, we have provided an alternative kind of analysis based on the geometry of the dynamical systems involved. Such an explanation misses aspects specific to each clock - e.g., how specific biologically tuneable parameters in each model affect internal and external noise resistance. However, the dynamical systems picture has the advantage in that it identifies the common origin of the trade-off across these systems. Such a dynamical systems picture has been fruitful in making other general but falsifiable predictions in biology[58, 11, 59], going back to Winfree's phase singularity [12].

Our dynamical systems theory shows that the critical parameter for noise resistance is the strength of the external driving relative to the amplitude of free running oscillations, captured by the geometric ratio L/R in our analysis. Weak driving provides resistance to external noise while strong driving provides resistance to internal noise. While our dynamical systems theory involve planar circular limit cycles, the models in Fig.2.2 have complex non-planar non-circular limit cycles and yet reproduce our trade-off. Finally, while the internal noise discussed here is set by finite copy number, this dependence is not essential to the results here. Any source of disturbance (e.g., bursty transcription) that perturbs the phase of the oscillator in constant light conditions is equivalent to internal noise. Similarly, external noise can involve any kind of fluctuation (e.g., multiplicative fluctuations, phase fluctuations) of the external signal that does not result in a persistent phase shift of the external signal itself.

Our work suggests that the damped oscillators are not merely poor cousins of the remarkable free running oscillators found e.g., in *S. elongatus*. At the low protein copy numbers such as those found in *P. marinus*, damped clocks keep time more reliably than free running clocks. (Low copy number has been linked to a trend towards reduced genome size and cell size in *P. marinus* [33].) In addition to the noisy internal environment of *P. marinus*, the

external environment might also play a role in selecting a damped clock; *P. marinus* is typically found in the open ocean, where the external environment may be more regular than the fresh water habitat of *S. elongatus*. In addition to *P. marinus*, damped oscillators are found elsewhere in biology, often in specific physiological conditions [13, 60]. Understanding the benefits and drawbacks of such damped oscillators in different conditions is critical since such oscillations are easily overlooked experimentally, in comparison to free running oscillations.

While numerous upstream and downstream considerations can modify [61, 46] the ultimate biological impact of fluctuations, we find that the core oscillator's geometry in itself can continuously trade off protection against external fluctuations for protection against internal fluctuations in the diverse range of models studied here.

CHAPTER 3

ENVIRONMENTAL NOISE ENABLES SENSITIVE DETECTION AND TRANSCRIPTIONAL DECODING OF CYTOKINE INPUTS

Cells communicate effectively in noisy biochemical environments despite the detrimental effect of noise on information transfer. Understanding how cells deal with environmental noise – chemical fluctuations- when detecting and transcriptionally decoding external signaling inputs has immense importance for studies of healthy and dysregulated signaling. Here, we use microfluidic live-cell experiments and mathematical modeling to study NF-κB pathway's ability for detecting, discriminating and transcriptionally interpreting noisy immune inputs. Surprisingly, we find that single-cell NF-κB activation is dramatically enhanced through the addition of noisy fluctuations to an otherwise constant cytokine signal. Theoretical analysis shows and experiments confirm that the noise-enhanced sensitivity is explained by nonlinear adaptive dynamics – rectified adaptation – in the NF-κB network. Our results show a surprising role for external fluctuations in enabling the detection of otherwise undetectable signals and demonstrate how cells exploit non-trivial molecular mechanisms to thrive in noisy environmental conditions.

Cells are exposed to randomly fluctuating levels of signaling molecules in their natural environment [62, 63, 64, 65, 66]. Noisy transcription, burst-like protein secretion, thermal fluctuations and interference from other cells all contribute to the noisiness of the extracellular signaling environment [67, 68]. Noise is considered detrimental to information transfer and signaling, both in biological and physical systems. Despite constant influence of noise on cellular inputs, signaling network somehow create robust and finely tuned gene expression profiles [69]. How cells deal with external (environmental) fluctuations and create appropriate signaling responses is not well understood, and it is central to many signaling problems

in health and disease. The well curated cellular responses can be the result of noise dampening features of gene regulatory network motifs [18]. These network motifs can potentially reduce the detrimental effects of noise and limit the response to input fluctuations through mechanisms like time-averaging [18, 70]. However, such features rely on reducing the responsiveness (sensitivity) and speed of signaling systems through time averaging. On the other hand, many cellular signaling pathways manage to combine high sensitivity and speed with noise tolerance, which are seemingly conflicting properties to exist in the same signaling system. Here, we carefully examine such a sensitive and dynamic signaling system central to immunity, the NF-κB system, which we show to take advantage of noisy fluctuations in a cytokine signal to enhance its sensitivity and detects an otherwise undetectable signal in the environment. NF-κB is an important example of a dynamically regulated transcriptional network that responds to a range of signaling inputs [69]. NF-κB controls the expression of hundreds of genes in response to a wide range of immune stimuli including signaling molecules secreted by host cells and from pathogens such as bacteria and viruses11. NF-κB coordinates many basic functions like innate immunity, immune development, and is involved in pathophysiological outcomes including autoimmune disorders and cancer [71]. Cells in the resting (unstimulated) state contain cytoplasmic NF-κB transcription factors, including p65. Upon cellular stimulation, NF-κB transcription factors rapidly shuttle to the nucleus and activate downstream gene expression. One of the transcriptional targets of NF-κB is its own inhibitor IκB. Newly synthesized IκB causes NF-κB to shuttle back to the cytoplasm, thereby creating oscillations in nuclear NF-κB and in target gene expression. NF-κB activation is digital at the single cell level: cells that are exposed to TNF or LPS activate in an all-or-none fashion [69, 72], resulting in total localization of NF-κB to the nucleus in responding cells. Increasing signal dose causes more cells in the population to activate and in a faster way compared to lower doses. NF-κB has been a model system for understanding the role of protein dynamics in signal specific gene regulation [73, 74, 16], and the influence

of cellular variability and internal transcriptional fluctuations on dynamic NF-κB signaling has been studied [69, 75, 17]. However, the effect of noisy external inputs to an immune signaling pathway like NF-κB has not been studied to date, and how cells robustly operate in noisy signaling environments remains an important open question (Figure 3.1a).

The addition of noise to a signal typically degrades the transfer of information to the output, both for physical and biological signaling systems [67, 76, 77]. Here, we ask how biological systems deal with signal input noise and address whether signaling pathways like NF-κB can actually utilize noise in the input to improve signaling fidelity. We investigate how environmental noise influences the detection and transcriptional decoding of cytokine inputs received by cells. Theoretically, the sensitivity of a system to small signals can be improved by addition of white noise through a range of effects collectively called stochastic resonance [78]. In stochastic resonance, frequencies in the input noise spectrum amplify similar frequencies in the responding system and create greater output [77]. The precise mechanism of stochastic resonance varies, but in general, noise induces a state change (Figure 3.1b) or excitation [77] that would not happen in the absence of noise. Further, stochastic resonance requires a strongly non-linear system. We reasoned that biological signaling systems like NF-κB could exhibit stochastic resonance since they are non-linear and exhibit dynamics of different time scales. In particular, NF-κB has multiple nested feedback loops leading to oscillations, and exhibits digital activation [69, 72]: an obvious example of nonlinearity that could benefit from noise. Despite the favorable theoretical arguments, stochastic resonance has not been experimentally seen in living cells, and whether it plays a role in immune signaling is not known.

3.1 Dynamic modeling predicts that input noise enhances NF-κB activation in single cells

We hypothesized that the multiple feedback time-scales in the NF-κB network could resonate with various frequencies in random noise in biochemical inputs and create a more robust activation effect in live cells (Figure 3.1c). To explore this possibility, we first simulated the mathematical model of the NF-κB network in single cells under both constant and fluctuating (noisy) concentration of the inflammatory cytokine TNF (Figure 3.1d). The simulated input signal, when supplied at a constant (noise-free) concentration below the activation threshold, does not result in NF-κB translocation and cellular activation (Figure 3.1e). Surprisingly, a noisy signal at the same mean dose causes cells to robustly activate, with 100% of cells in the population activating the NF-κB pathway. To confirm that cells were not responding to the maximum TNF values in the fluctuating noisy signal, we modeled a noisy TNF signal at an even lower level than the constant signal and observed that the cells still activated (Figure 3.1f). Our simulations predicted that a noisy input will activate a much larger fraction of cells in the population than a constant input dose, even at the maximum value of the noisy input. Thus, our simulations suggest that despite being below the all-or-none threshold for digital NF-κB activation, a low dose input can activate cells with the addition of chemical noise, and that this enhanced response cannot be explained by the maximum dose alone. Such simulations show subthreshold signal detection with the addition of noise, a characteristic signature of stochastic resonance, and imply that noise can help increase the sensitivity of NF-κB response in single cells.

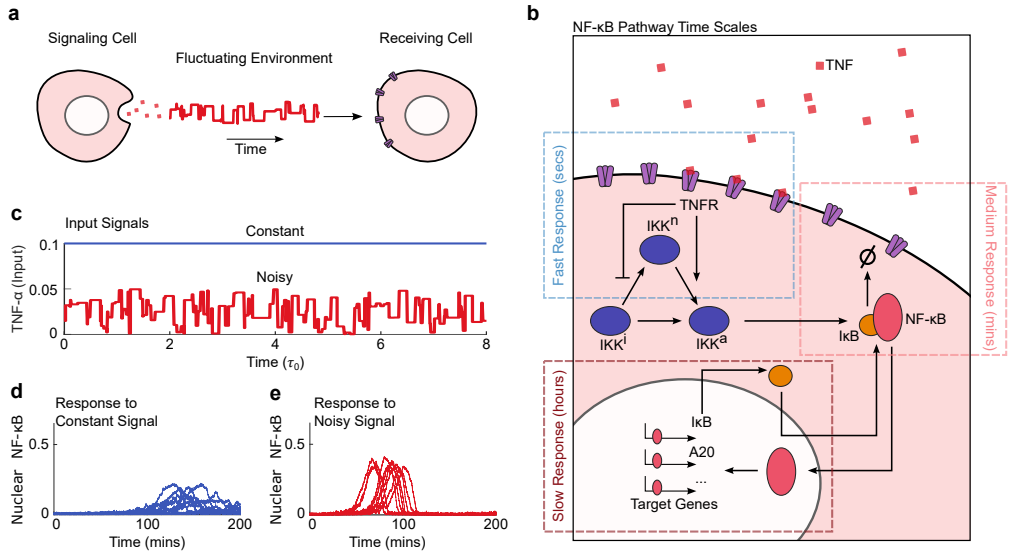

Figure 3.1: Simulations predict that weak (sub-threshold) cytokine inputs may evoke a strong NF-κB response upon addition of chemical noise to those inputs. (a) Extra-cellular environments are subject to chemical fluctuations, which make signals between cells to become noisy. Signal receiving cells process such noisy inputs. (b) NF-κB pathway has complex nonlinear dynamics on a wide range of timescales, from fast receptor dynamics to the slower pulse-like nuclear translocation of NF-κB. (c) Dynamic input profiles used for simulations with the mathematical model of NF-κB pathway, under both constant (blue) and fluctuating (red) input levels from the cytokine TNF. (d, e) In simulations, fluctuating TNF levels cause strong activation of NF-κB nuclear translocation in all cells. However, constant TNF inputs induce lower NF-κB activity and in fewer cells, despite the fact that TNF dose they see exceed maximum of the fluctuating input.

3.2 Live-cell experiments reveal that noisy cytokine inputs create increased NF-κB activation in single cells

To experimentally test whether addition of chemical white noise to a cytokine input would indeed enhance the activation of NF-kB in live cells, we studied nuclear localization dynamics of NF-κB in response to real-time fluctuations of TNF. We cultured 3T3 mouse fibroblast cells with p65-dsRed reporter in an automated microfluidic cell culture platform, and quantitatively measured single cell NF-κB nuclear translocation in live-cell imaging experiments under a range of time-varying TNF concentrations (Figure 3.2a). Our automated microfluidic device can deliver time-varying dynamic chemical signals to cell culture cham-

bers, including signals that contain predetermined levels of chemical fluctuations. The device creates 154 pre-determined concentration levels between 0.01 ng/mL and 0.11 ng/mL TNF with 0.5 second timesteps[79] (Figure 3.2b), which is much faster than the typical timescales of NF-κB activation dynamics measured in live cells (tens of minutes). Cells were incubated in this device and exposed to constant TNF signals whose concentrations range from 0.001 ng/ml to 1ng/mL (Supplementary Video 1; Supplementary Figure 1; Supplementary Figure 3.2) and a noisy signal (Gaussian white noise) with mean concentration of 0.05ng/mL and Signal to Noise Ratio (SNR) of 5/1 (Figure 3.2c; Supplementary Video 2; Supplementary Figure 3). During stimulation experiments, cells were imaged with time-lapse fluorescence microscopy, and individual cells were tracked using custom image processing software[80]. We found that the cells responded much more sensitively to the noisy input: The fraction of responding cells in the population as well as the nuclear NF-κB amplitude of individual cells showed significant enhancement under the noisy signal when compared with an equivalent dose constant signal (constant dose at 0.05ng/mL) (Figure 3.2d; Supplementary Video 3; Supplementary Figure 3). A randomly selected set of activated single-cell traces show this dramatic increase in NF-κB activity (Figure 3.2e and 2f). In control experiments with constant TNF at 0.05 ng/mL, only 33% of cells are activated, but under the Gaussian white noise stimulation at the same time-integrated TNF dose, 85% of the single cells in the population are activated (Figure 3.2i). Single cell NF-κB traces measured under the noisy signal are similar to those stimulated with much higher constant doses of TNF (mean integrated TNF dose for noisy signal is 0.05ng/mL). Furthermore, the fraction of responding cells is also significantly higher under the noisy input. To achieve 80% activated fraction we needed to use 1 ng/mL constant TNF, which is a 20-fold dose increase over the mean profile and 10-fold increase over the max value of the noisy signal (Figure 3.2g). These results clearly show that individual cells perceive a noisy signal to be more than 10-fold more potent and respond accordingly. There are several components of the single cell NF-κB localization dy-

namical profiles that are relevant when comparing NF-κB activation. First peak height and integrated area of nuclear fluorescence are linked to target gene expression; different levels of NF-κB in the nucleus are correlated with different phenotypic gene response [69, 81]. For a noisy stimulus, while the normalized NF-κB peak height and integrated area is consistent with a dramatically higher dose (Figure 3.2j and 3.2k), we found that the response time to first peak is similar to the noise-free stimulation. Both noisy and noiseless signals show the same response time, with the same single-cell distribution and mean around 65 mins, which is consistent with activation at a constant dose of 0.05ng/mL (Figure 3.2h and 3.2l). This result shows that while noisy fluctuation of TNF signal increases the perceived dose for NF-κB activation and amplitude, the speed of NF-κB activation is not increased due to presence of noise. Overall, our experiments confirmed the theoretical prediction of dramatically increased noise-induced activation of NF-κB in living cells; further, the activation arises as a cellular response to input fluctuations and cannot be simply explained as a perception of the maximum dose. To our knowledge this is the first experimental demonstration of noise enhanced immune signaling activation in living cells, and highlights how biological systems can exploit non-trivial physical effects to take advantage of noisy environments.

3.3 Gene expression response to cytokine noise shows switching of transcriptional programs to anti-cancer responses

To evaluate the functional effects of input noise, we probed gene expression response in cells under both noisy and constant TNF stimuli. We stimulated cells again through our automated microfluidic cell culture platform, then extracted cells at 90min, 180min and 300min timepoints. Approximately 150 cells were lysed and frozen for pooling in RNA sequencing (SMART-sequencing) for each timepoint in triplicate (Figure 3.3a). Analyzing the gene expression differences between the two conditions reveals many differences across a range of transcriptional programs. We looked at the all genes that were significantly upregulated

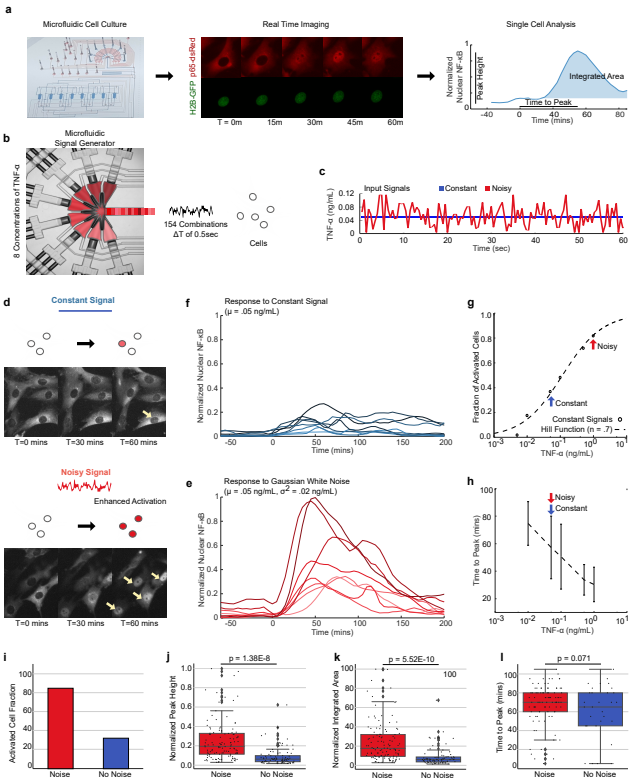

Figure 3.2: Live-cell stimulation experiments show that addition of white noise to a weak TNF input causes NF-κB to respond very sensitively to that input. (a) Microfluidic live cell stimulation experiments allow analysis of single cells in precisely controlled dynamic environments. Cells cultured under dynamic chemical signals are automatically tracked via live cell microscopy, and NF-κB nuclear localization is analyzed in individual cells over time. (b) Cells are exposed to signals generated by the combination of several TNF concentrations, producing 154 variable exposures at 0.5 second timesteps. Actual picture of the microfluidic device is shown on the left. Colors indicate controllable valves that regulate input media. (c) Gaussian white noise with 5:1 signal to noise ratio (red) is added to the control TNF signal with a mean at 0.05 ng/mL (blue), and this chemical signal is delivered to hundreds of live cells. (d) Fluorescent images of mouse fibroblast cells expressing dsRed-p65 reporter that are exposed to noisy (red) or noise-free (blue) signals. Activated cells, showing nuclear NF-κB localization, are indicated with arrows. Noisy signal causes strong NF-κB activation in individual cells. (e) Activated single cell traces measured under noisy stimulus. (f) Single cell traces under constant TNF exposure show significantly less NF-κB translocation to the nucleus. (g) Population fraction of activated cells under different constant doses of TNF, measured in the same device. Blue and red arrows indicate the dose perceived by cells in stimulation experiments with constant and noisy inputs in (e) and (f). The noisy input is perceived by cells to be at a much high dose than the constant input. (h) NF-κB response time at different constant doses of TNF. Both noisy and noise-free inputs lead to the same response time. (i) Comparison of activation fraction for noisy vs. constant signals. Noisy input induces more cells to activate in the population, despite that fact that both inputs have the same mean TNF dose. (j) Comparison of peak height normalized to maximum peak value for noisy vs. constant exposure. (k) Comparison of integrated area of fluorescent readout for nuclear translocation normalized to maximum integrated area for noisy vs. constant exposure. (l) Peak timings for activated cells remain the same despite variation in peak height and integrated area.

and downregulated (pval < 0.01, and foldchange > 0.5) through noisy stimulation compared to noise-free stimulation, and found 2540 genes that had significant variation (Supplementary Table 1) of which a majority are upregulated in response to noise. This indicates that cytokine noise largely enhances or broadens gene expression response compared to a constant signal. We narrow down our analysis to the two broad categories of noise sensitive gene expression response using hierarchical clustering with two different time scales (early and late). Gene ontology (GO) of these significant genes show upregulation of cell cycle, ribonucleotide binding, anti-cancer and chromosome modifying genes across both early and late response under noisy stimulation. It is of note that at the late timepoint, there is also an increase in transcription regulator and protein phosphorylation and modification activity (Figure 3.3b). At the early timepoint there is upregulation of NF-κB coregulator binding: transcription factors with $> 15\%$ of total upregulated target genes include Jdp2, Mta1, Mzf1, Sall4, Hdac3, e2f and Klf13 [82, 83, 84, 85, 86]. This transcription factor response alongside broadly enhanced DREAM complex and p53 program response indicate a cellular effort to limit differentiation and proliferation while preparing to address DNA damage within the cell [87]. This coordinated response at the early 90min timepoint suggest that a noisy environment could induce a programmed response to prevent cancer and proliferation of damaged cells [88].

While some of the early transcription factors continue to be upregulated highly at the 300m timepoint (Hdac3, Klf13, Sal4, JDP2) several new transcription factors emerge including Lef1from the WNT pathway (Figure 3.3c). The chromatin regulating protein Morc2 also appears upregulated in both early and late noise upregulated genes (Supplementary Table 2, Supplementary Table 3). Morc2 is responsible for many gene silencing effects via modification of chromatin through interaction with the HUSH [89]. While many of the cellular programs point towards increased cell cycle regulation, there is also upregulation in anti-cancer markers to limit proliferation as well [90, 91, 92] (Figure 3.3d). Exposure to

noisy stimuli thus evoke enhanced expression of anti-proliferation and anti-differentiation gene expression programs and likely also prime these cells for immune cell interference.

These programs combined suggest that noise in the signaling environment may play a role in creating a more robust anti-cancer response for an inflammatory event. The significant difference in phenotype of cells when comparing different noise levels in the environment is consistent with the pleotropic response of NF-κB under different physiological conditions [93]. It has been shown that different temporal dynamics create different gene expression responses , so this noise-responsive phenomenon is an extension of NF-κB gene expression studies and shows the power and adaptability of innate immune response to different environmental signals.

3.4 Noisy cytokine stimulation induces non-canonical alternatives to NF-κB signaling

Next we analyzed expression differences between noisy and constant stimulus on genes with NF-κB specific binding sites [94, 95] (Supplementary Table 4). We used hierarchical clustering to cluster genes that have expression levels that are amplified by and dampened by noise (Figure 3.4a). Interestingly, genes that typically respond within the first 90 minutes show similar expression profile between noisy and constant stimulation (Junb, Tnfaip3, Stat5a) which differs from a broad analysis of gene expression. At the later time points, however, the NF-κB transcriptional response diverges between the noisy and constant stimulation groups at 180 minutes and 300 minutes. This behavior suggests that the broad gene expression activity that occurs in the early 90 min timepoint, primes the specific NF-κB response for later time points. Classic NF-κB target genes show upregulation upon noisy stimulation at later time points (Supplementary Figure 4). Although the TNF dosage that cells perceive is increasing with the addition of noise (Figure 2g), not all NF-κB responsive genes show increases given a noisy signal. Instead, it appears that the pathway is accessing different pro-

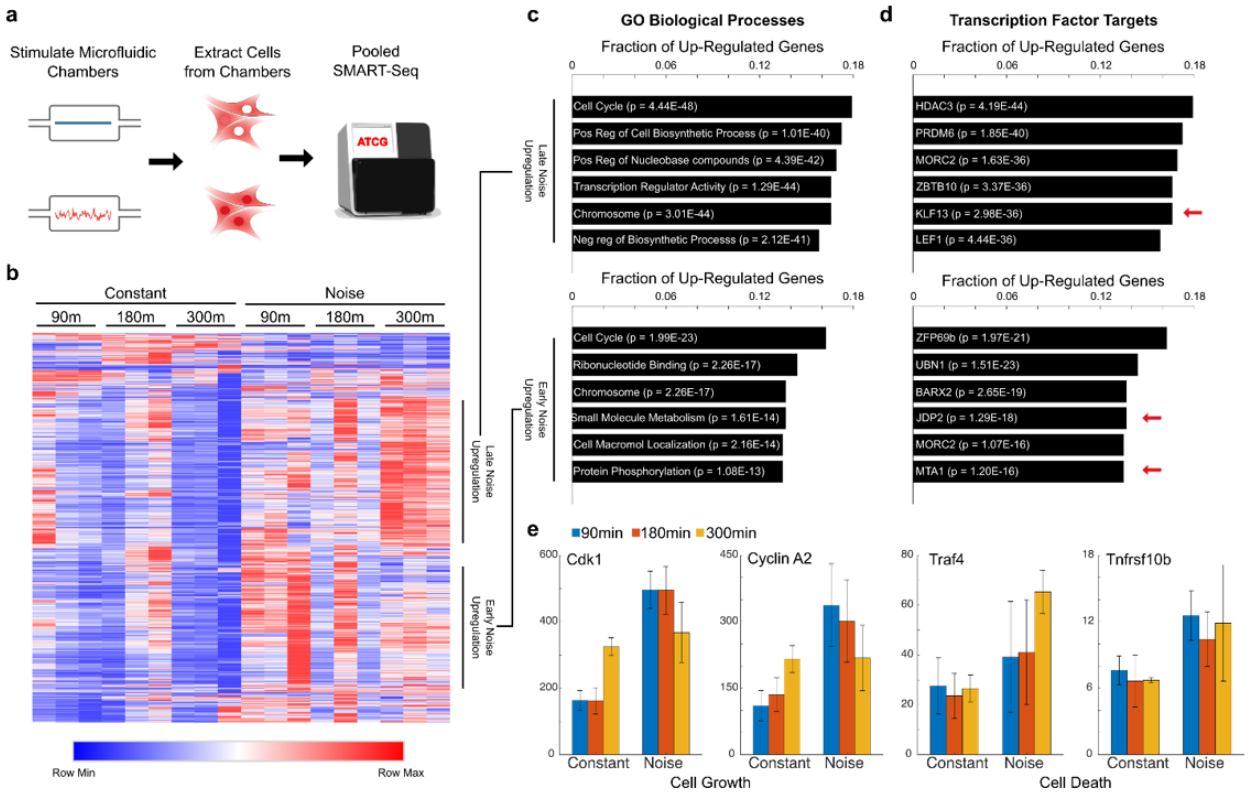

Figure 3.3: Sequencing of stimulated cells shows noise dependent gene expression. (a) Cells were loaded and stimulated in microfluidic chambers, then extracted and pooled for SMART-sequencing. (b) Heatmap of significantly different gene expression responses for each time-point shows that noise results in different gene expression response (p ¡ .01, fold-change ¿.5) (c) Gene Ontology for early and late noise-upregulated genes show cell cycle and chromosome related gene expression. Late noise responsive genes show increased transcription regulator activity (d) Transcription factors that were identified as the expression effectors include NF-κB coregulators (shown with red arrow) as well as crosstalk transcription factors lef1 and Hdac3. (e) Expression of cell growth and cell death gene markers are both upregulated under noisy stimulation.

grams leading to a different NF-κB gene expression outcome at different times in a manner similar to our whole transcriptome analysis [96, 97]. The chromatin is being primed by early transcription factor response via chromatin remodeling and is shifting the NF-κB-specific response [98]. To understand where NF-κB specific response functionally changes with the addition of noise, we mapped the interactions of the response genes to each other using STRING[99]. We then look at the differences between noise dampened and noise induced genes in the NF-κB pathway. High confidence protein interactions are shown in a map of NF-κB response proteins and we then use the clusters generated through hierarchical clustering to assign each protein in the interaction network a grouping based on expression behavior given noise (Figure 3.4b). Graphing the network of these response genes shows that the noise amplified genes stem from Cxcl cytokines and Nfkb1, whereas noise dampened genes are affiliated with canonical NF-κB response given TNF and p53 inhibition (Myc, Bcl2). Our analysis of the general whole transcriptome programs given noisy stimulation add insight on this result as well. Release of these Cxcl cytokines is used in neutrophil recruitment [100, 101] and lef1 is also known for mediating cell-cell communication. This programmed behavior suggests that a noisy TNF signal can be used to prepare a microenvironment for an inflammatory response through immune cell engagement while limiting the growth potential of the effected cell. The overexpression of p53, e2f related genes and Nfkb1/ CXCL related proteins interestingly also hints at a branch point in NF-κB noise regulation. The IKK family proteins are known mediators for p53 and e2f as well as activators for non-canonical Nfkb1 signaling [102, 103, 104, 105].

3.5 Rectified adaptation is the likely mechanism for noise enhanced sensitivity in NF-κB

To understand the mechanisms behind stochastic resonance in NF-κB, we theoretically evaluated how fundamental components of the NF-κB signaling network respond to fluctuating

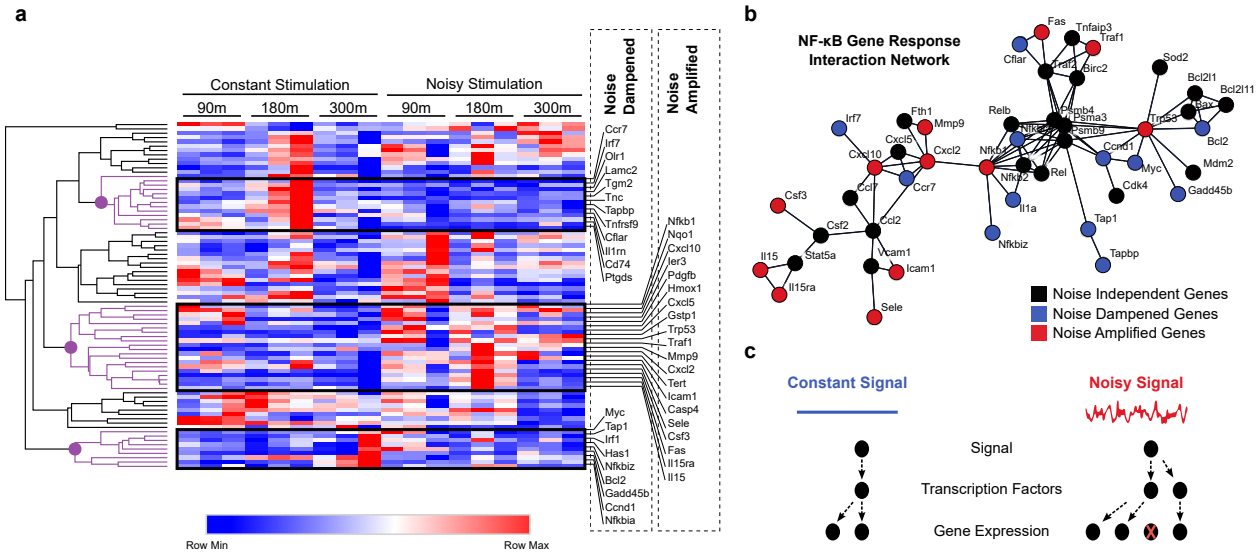

Figure 3.4: NF-κB responsive genes show divergent noise sensitivity across different timescales. (a) NF-κB associated genes were isolated and compared across stimulation patterns and time. Genes that respond within the first 90 minutes show similar expression profile between noisy and constant stimulation. Gene expression diverges between noise amplified and noise dampened genes at 180 minutes and 300 minutes (b) Mapping the highest confidence interactions between NF-κB responsive genes shows that the noise amplified genes stem from Cxcl2 interacting partners, whereas noise dampened genes are affiliated with p53 inhibition and canonical NF-κB response.

inputs. In particular, adaptive dynamics [106, 107, 18, 70], known to be present in the NF-κB pathway [108, 109] and also widely across biology, show a large response to a change in input but the output soon returns to its prior resting value. Hence, the output of adaptive dynamics is naturally more sensitive to input fluctuations than to the steady state level of that input. Adaptive dynamics are commonly implemented by either Incoherent Feed-Forward Loops (iFFLs) or negative feedback loops. Since iFFLs are known to occur in the NF-κB pathway [108, 109, 110], we simulated a general iFFL in a simple input-output model upstream of IKK (Figure 3.5a). We found that upward and downward fluctuations in the input signal cause equal and opposite changes in the response, thus cancelling each other. Consequently, noise in the input has no net impact on the output for a regular iFFL network (Figure 3.5c). However, we found that a simple modification of the iFFL led to a dramatically different conclusion (Figure 3.5b); if protein concentration (Y) in the network has a natural floor, the response to positive and negative input changes is asymmetric. Such a node in the network acts much like a diode or a 'rectifier' in an electrical circuit, allowing responses to only upward but not downward fluctuations in the input. Such rectification in the presence of noise leads to counterintuitive behaviors in physical systems, from molecular motors to Feynman's ratchet [111]. In the NF-κB signaling context, we find that such a floor-rectified iFFL responds predominantly to upward fluctuations in the input, and thus can lead to a higher sensitivity to noise as seen in experiments (Figure 3.5d).

3.6 Sinusoidal input dynamics can decouple noise sensitivity timescales

To evaluate the timescales involved in noise sensitivity, we derived a prediction for NF-κB response to periodic single frequency stimulation. In stochastic resonance, the response is highest when the input timescale matches the corresponding timescale in the responding system. Our simulation suggested that NF-κB activation should be the highest for inputs

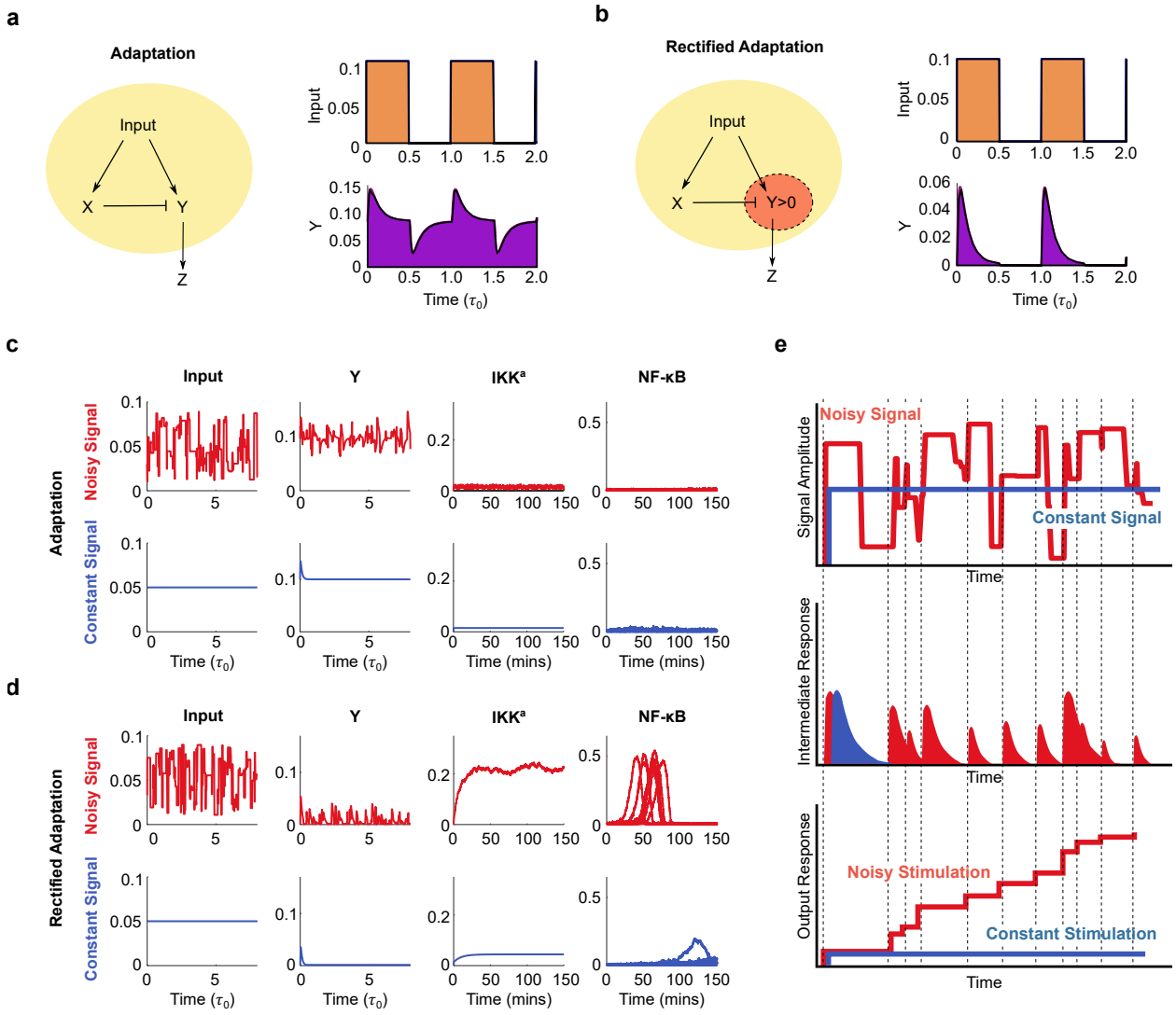

Figure 3.5: Rectified adaptation leads to stochastic resonance. (a) The Incoherent Feed Forward Loop, shown in the schematic on the left, demonstrates adaptation. A step-like change in the input produces a transient response in Y that dies out and the system returns to the resting level. Such motifs are found in the NF-κB pathway and molecules like IKK are known to show an adaptive response to step changes in TNF. (b) Rectified adaptation, where step ups produce a response, but step downs produce no response because the output molecule Y is already at near-zero levels in steady state. (c, d) When subject to fluctuating input signals, the response of the conventional adaptive circuit to step ups and downs cancel each other out. However, the output of the rectified adaptation circuit builds up over time, since it only responds to step ups of TNF input. Consequently, active IKK levels rise with time, causing nuclear translocation of NF-κB in simulations.

whose period match the timescale of the adaptive response. For longer input periods, cells are expected to behave like they are simply experiencing periodic stimulation at the maximum seen dose (Figure 3.6a). The model also predicts that cell activation will saturate at frequencies higher than the adaptation timescale (Figure 3.6b). To experimentally test our predictions, we measured single cell activation probabilities using a periodic cosine TNF stimulus at different periods, ranging from 2 s to 30 s (Figure 3.6c; Supplemental Video 4). These chemical stimuli are generated using our microfluidic device, and cells were imaged by time-lapse microscopy during stimulation. Activated single-cell traces show higher peak heights for stimulation periods under 30 s (Figure 3.6d; Supplemental Figure 4) and faster response times with shorter periods (Supplemental Figure 5). For 30s and higher, cells fail to show the significant activation seen under stimuli with shorter periods. When comparing activation profiles under different time-scales, we find that cells perceive fast time scales (¡15 sec period) at an effective constant dose greater than 1 ng/mL TNF, which is much higher than the actual used mean dose of 0.05 ng/mL in these experiments. These results show that NF-κB resonated at the input periods shorter than 15 seconds. As we increase the timescale to above 15 sec, cells perceive the stimulus at an effective dose of approximately 0.1 ng/mL TNF (Figure 3.6e and f). Similar misperception of oscillatory signals on a minutes timescale has been previously observed in yeast [112, 113]. Our experimental results here confirm the predictions from the model and suggest that noise perception at fast time scales (5 s to 15 s) via a rectified Incoherent FeedForward Loop is contributing to the increased NF-κB activation observed in experiments with gaussian white noise.

3.7 Discussion

In summary, our results show that input noise can enhance the strength of NF-κB response to weak environmental (cytokine) signals via a mechanism that resembles stochastic resonance, and offer a novel mode of cellular regulation for this important transcription factor. We find

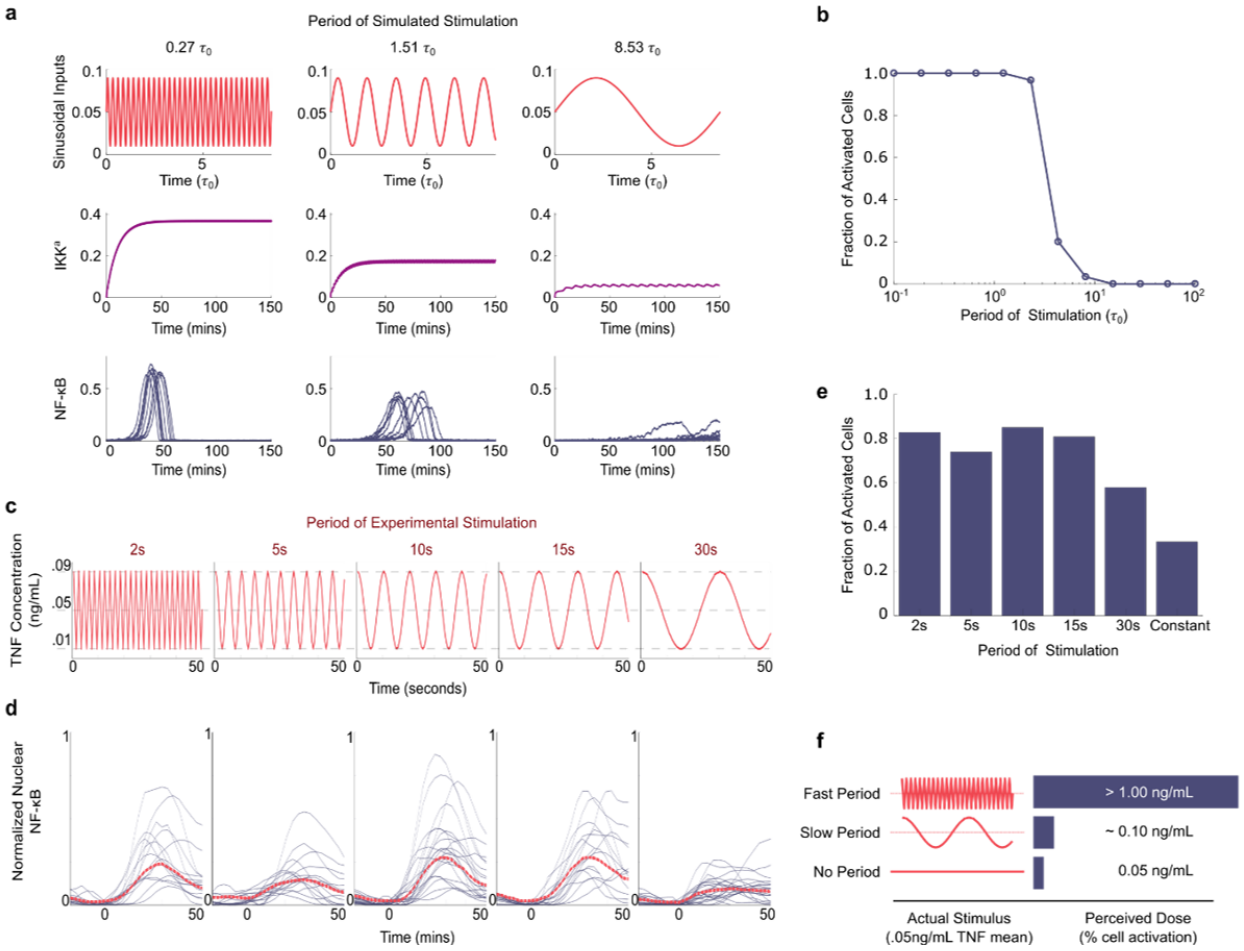

Figure 3.6: Periodic stimulation uncovers time-scales behind stochastic resonance. NF-κB shows resonance under fast oscillating TNF inputs. (a) Model simulations allows deriving predictions for response to different periodic inputs. Simulations show NF-κB activation only for periods comparable of faster than the adaptation timescale τ_0. For longer periods, the IKK response to different cycles of TNF input do not build on each other and there is no NF-κB activation. (b) The model also predicts that the fraction of cells activated saturates at a high fraction for frequencies higher than the adaptation timescale. (c) Experimental stimulation profiles for cosine exposure at different periods ranging from 2s to 30s. (d) Activated single cell traces in live cell imaging experiments show similar profiles for exposures under 30second period, but dramatically reduced peak height for 30second period (mean of populations in red). (e) Low period cosine stimulation increases fraction of activated cells in the population but increasing period of cosine stimulation after cutoff reduces activated fraction. (f) Cells exposed to slow periods behave as an intermediate between cells exposed to fast periods and constant signals. There is a 10-fold increase in the perceived dose with the addition of fast fluctuations to the signal.

that NF-κB is sensitive to the dynamics of a pleiotropic factor like TNF, and temporally random modulation of the TNF signal can create significantly different phenotypic responses. This raises the intriguing possibility that biological noise can help discriminate regulatory signals. For example, the noise level in TNF could serve as a proxy for the number of signaling cells in the environment, where higher noise corresponds to higher number or density of surrounding cells. In such a scenario, the NF-κB signaling properties uncovered here would serve as an quorum sensing-like mechanism72, with distinct activation programs only at high cell density for weak signals. Or more generally, a noisy environmental signal could mimic an intrinsic biological phenomenon. For example, immune regulatory cells such as natural killers have cell membrane associated TNF73, and it is possible that attracted natural killers could reproduce such noisy signaling. Input noise perception as an unexplored regime in cellular regulation may play an important physiological role in communication and signaling in fluctuating environments for many additional regulatory systems. The gene expression differences from NF-κB specific transcription of noise-stimulated cells show how adaptable the NF-κB pathway is to dynamic stimuli. Not only does the signal enhance NF-κB activation, but also leads to alternative response upon a secondary stimulation. Changes due to transcription of proteins involved in regulation of the pathway will lead to alternative outcomes given a secondary stimulation. More generally, we found that cytokine noise induces a broad anti-cancer transcriptional response by limiting the growth potential of cells and activating anti-cancer transcriptional programs. It also prepares cells for the further recruitment of neutrophils. Interestingly, the behavior and expression patterns that we find is also a signature in tumorigenesis74,75. It is possible that key oncogenic mutations could lead to the dysregulation of this program and as a result, induce cancer instead protecting against it. Transcription factor dynamics in gene network regulation and signaling is an important element in mammalian cell response to many agonists - many of these transcription factors do behave non-linearly and rely on surpassing concentration thresholds of ligand molecules to

produce all-or-none responses76. We found that noisy fluctuations of an inflammatory signal causes cells to perceive this signal to be much more potent. Our discovery of noise enhanced signal sensitivity in living cells, which was theoretically predicted in many settings but was not experimentally observed. This behavior emerges from the interaction of noisy cytokine inputs, NF-κB pathway resonance at fast time-scales, and rectified adaptation in the NF-κB pathway. These results thus demonstrate how biological systems can exploit highly non-trivial and interesting physics to turn a limitation such as environmental fluctuations into a strength.

CHAPTER 4

DOMAIN GENERALIZATION USING TEMPORAL DYNAMICS DURING LEARNING

Generalizing from training to testing data is fundamental to the success of any machine learning algorithm. This success often assumes that both training and testing data come from the same distribution. However, in some data-limited problems,the distribution of the testing data is systematically different from all available training data because training data comes from a fundamentally different domain. For example, the training domain might involve hand-drawn sketches or watercolor paintings of cats and dogs while the testing domain might involve photographs of cats and dogs. Carelessly optimizing training error can lead the algorithm to learn "specialist" features with predictive power on training domains but ineffective on test domains. Instead, training should emphasize learning "generalist" features with potentially lower predictive power in training domains but with more *consistent* (i.e., less variable) predictive power across training domains. Our work focuses on the supervised classification problem where multiple sources of training data with distinct distributions are available. We present a training algorithm suggesting that switching between data domains during training increases the probability of the algorithm learning generalist features over specialists.

In machine learning, a "domain" is a set of data whose underlying distributions are the same. For example, we may generate a dataset of real numbers consisting of two different domains. One domain comes from the normal distribution, but the other comes from the uniform distribution between 0 and 1. Machine learning techniques work best when the training data and the testing data come from the same domain. Nonetheless, this assumption may not always hold. One of the reasons is that the cost of collecting the target data can be prohibitive. For example, the number of patients with newly discovered cancer subtypes is much smaller than the number of patients with other common subtypes. Domain

generalization is a field in machine learning, aiming to create algorithms trained on a multi-domain dataset that can maintain high predictive power over the target data from different domains.

Previous papers show that regularizing feature spaces can improve generalization. Adam Kalai shows that features with the lowest variance of correlations in multi-domain datasets tend to have the highest performance on unseen testing data [114]. Another group put particular regularization (Global Class Alignment and Local Sample Clustering objectives) on the feature spaces to encourage the universal features and suppress idiosyncratic ones [115].

Other researches show that training the model while temporally hiding some parts of the dataset can improve the generalization. A standard training algorithm uses a fixed amount of data repeatedly over multiple epochs. Da Li splits each batch of data randomly into two parts: meta-train and meta-test [116]. The algorithm optimizes the objective function of the meta-train and the meta-test alternatively. This process stimulates the domain shift and helps increase generalization, as demonstrated in the paper.

Another research on the time-varying training methods reveals that dynamical loss function can lead to better generalization. Parameters in the loss functions are often static for a conventional approach. Miguel Ruiz-Garcia periodically changes the loss function to emphasize one class at a time [117]. This algorithm increases the validation accuracy from 0.73 to 0.79 on the benchmark CIFAR10 dataset.

We propose a model-agnostic training on the multi-domain dataset that increases generalization. At any particular epoch, the model is trained only by the data from one domain. The algorithm periodically rotate the domain of the training data to lead the loss function toward discovering the features common between all domains. This paper show the applications on a convolutional neural network and a Support Vector Machine (SVM).

4.1 Convolutional Neural Network on a text dataset

4.1.1 Dataset

Many universities have webpages for their faculties and students. Human can distinguish between the faculty and the student webpages by looking at important clues, such as the graduation date, the title, etc. Nevertheless, a machine learning algorithm may instead focus on idiosyncratic clues useful only in some universities. For example, in the webpages dataset gathered by [114][1], the word '19' consistently appears in the student webpages, because the web headers include the download time and the student data were coincidentally downloaded at 7pm for many universities. We want to avoid this type of features and only select features common in most universities (or domains).

Inspired by the above dataset, we design artificial text data with two domains. One domain features the word 'UCLA,' and the other features 'NEWS'. The common feature in both is 'PROF' because the word "professor" in a webpage is a universal indicator. 'UCLA' comes from the acronym for the University of California, Los Angeles, and the word "news" could be included only in some universities' webpages.

Each text data consists of 400 consecutive English alphabets. The Student texts are just random letters, while the Faculty texts include the above feature words with some probabilities. The specialist features 'UCLA' and 'NEWS' occur with a 90% chance in its specific domain and a 0% chance in the other. On the opposite, 'PROF' is included with a 45% chance in both domains. Thus, in a combined dataset, all three features are equally predictive of Faculty webpages.

1. avaliable at http://www.cs.cmu.edu/afs/cs/project/theo-20/www/data/

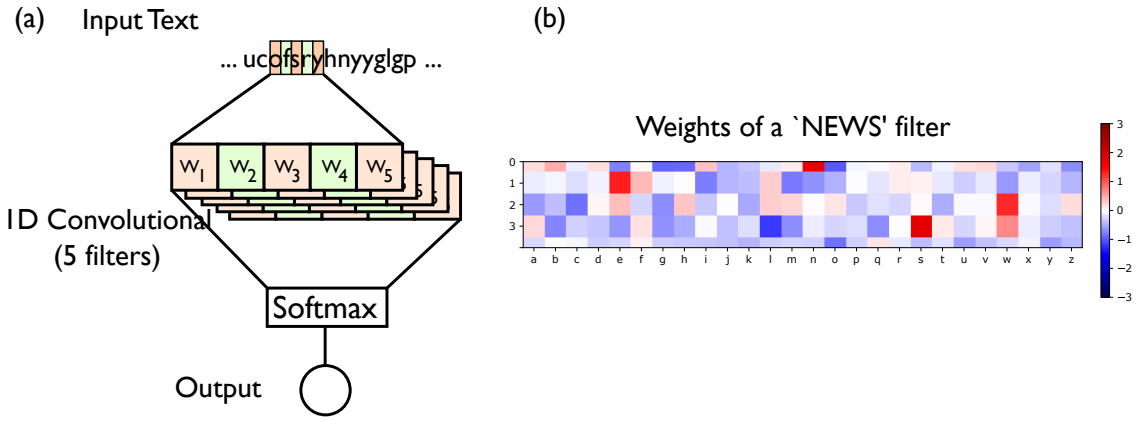

Figure 4.1: (a) The model takes in an input text and outputs one number, predicting the class of the text. The 1D convolutional network takes in five letters at a time, and the softmax layer summarizes the outputs from all the positions and filters to one number. (b) The weight vectors (W_i) reveal the learned features. Each row is a vector W_i consisting of parameters for each English alphabet. The colors indicate the relative responses toward each letter (x-axis) at each position (y-axis) of the kernel. High values signify that the filter is scanning through the texts for these letters. In particular, this filter responds to the letters 'NEWS'.

4.1.2 Model

We consider a 2-layer neural network where the first layer is the 1-dimensional convolutional network and the second layer is the softmax (see Fig. 4.1). The algorithm converts the input text of length L into a hot-vector of size (L, N_α) where N_α is the number of English alphabets. The kernel of the convolutional network checks the existence of a word of length 5 (see 4.1)b.

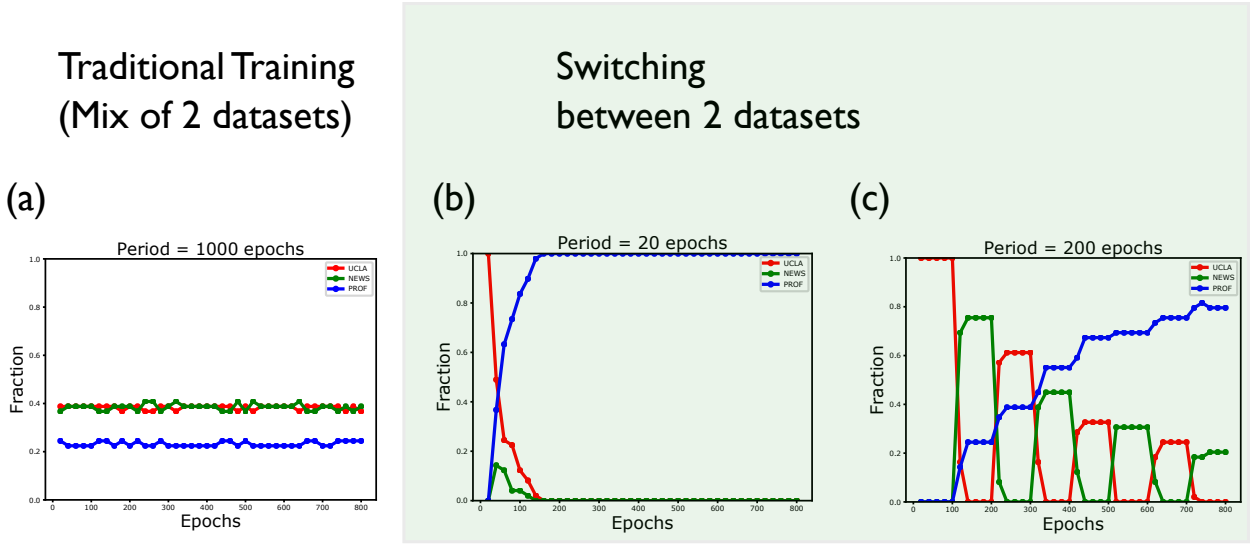

Figure 4.2: (a) Our algorithm trains the network with data from two domains. Each text of the teacher class has the same 45% probability of containing the word 'UCLA,' 'NEWS,' and 'PROF,' so the chance that the network learns each keyword is around 33%. On the right, the algorithm trains the networks with the data only from one domain at a time. The algorithm switches the training data's domains every half period. At period = 20 epochs, the network mostly learns the 'PROF' keyword, which is the common keyword among the two domains. At period = 200 epochs, the network learned the 'PROF' keyword but at a slower rate.

4.1.3 Result

In each plot of Fig 4.2, we trained 49 convolutional neural networks separately and categorized each of them according to its responses toward the keywords 'UCLA,' 'NEWS,' and 'PROF.'

In the traditional training, we combine the two domains. As a result, each feature appears at the same probability, so it is impossible to distinguish between specialist and generalist features. As demonstrated in Fig. 4.2a. On the other hand, the switching training increases the likelihood to learn the common feature. In the appendix, we simulate and show that the model takes about two epochs to learn a new features. When the domain is shifted, the models are encouraged to forget about the specialist features. As shown in the blue lines in Fig. 4.2b, all models are able to learn the common feature within ten periods of switching.

In Fig. 4.2c, the period is much higher than the response time of two epochs (see Appendix). The models show gradual progress toward learning the generalist feature. Nevertheless, the rates are much slower than those in Fig. 4.2b. The plot shows that the fraction is already in a plateau when each switching happens.

4.2 The support-vector machine on a 3-dimensional dataset

We want to demonstrate the application of this switching algorithm in a model as simple as the Support-vector machine (SVM). Thus, we abstract the Student and Faculty dataset by [114] from texts into three numbers because texts can be converted to one-hot vectors. We represents the three feature words as three binaries.

4.2.1 Dataset

Each data point has three binaries and belongs to one of the two classes: Student or Faculty. Similarly, the dataset has two domains. The first feature only appears in one domain. The second feature only appears in the other, but the third appears in both.

All three binaries of Student data are zeros, while some binaries of Faculty are ones. The two specialist binaries in the first and second dimensions have 90% chance to be one in its specific domain and 0% chance in the other. The common binary in the third dimension becomes one with 45% chance in both domains. Thus, in a combined dataset, all three binaries are equally predictive of Faculty webpages.

4.2.2 Model

We restrict our SVM model to only one dimension. In addition to the weights of the SVM, the model needs the "Trust" variables T_i to select which dimension to fit (see Fig4.3). If the selected dimension gives accuracy higher than a random model, then the corresponding

Trust increases. Algorithm 1 describes the mathematical details of the training.

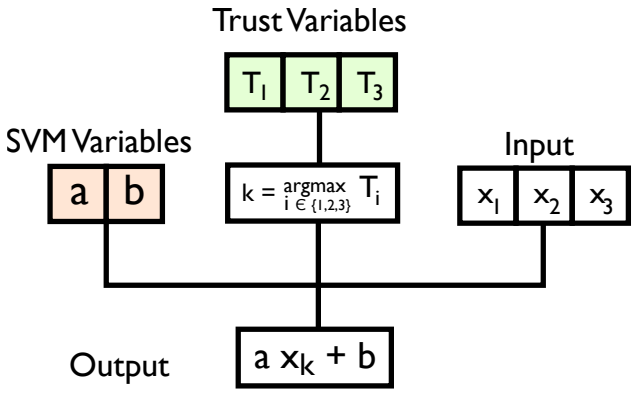

Figure 4.3: This SVM model has 5 parameters: a, b, T_1, T_2, and T_3. The input is (x_1, x_2, x_3) and the outputs is $ax_k + b$, where k is the maximum index of T_i. This argmax function for k forces the nonlinearity into the model, allowing a simple loss landscape with 3 minima, one for each $k = 1$, 2, and 3.

Algorithm 1: Training while switching with period P between 2 datasets

Initialize a, b randomly;

Initialize Trust $\mathbf{T} \equiv \{T_0, T_1, T_2\}$ randomly;

for *iteration* i **do**

  if $(i \mod P) < P/2$ **then**

    $D \leftarrow$ train1;

  else

    $D \leftarrow$ train2;

  Obtain m pairs of (data $\mathbf{X}$, targets y) from D;

  $j \leftarrow$ the maximum index of $\mathbf{T}$;

  Compute $M(j, \mathbf{X}) = a\mathbf{X}_j + b$;

  Update T_j by descending its stochastic gradient:

$$\frac{1}{m}\sum_{k=0}^{m}\Big(\text{loss}(M(j, \mathbf{X}^{(\mathbf{k})}), y^{(k)}) - \log(2)\Big);$$

where the loss function is simply the binary cross entropy loss

$$\text{loss}(x, y) = y\log(x) + (1-y)\log(1-x)),$$

and the $\log(2)$ is loss if the model M is random.

4.2.3 Result

In each plot of Fig 4.4, we trained 400 models and categorized each of them according to its responses toward the specialist and generalist features. With a mixed dataset, the models learn specialists about 66% of the time because there are 2 specialist features while there is only one generalist feature. With a period of 20 epochs, almost all models learn the generalist feature.

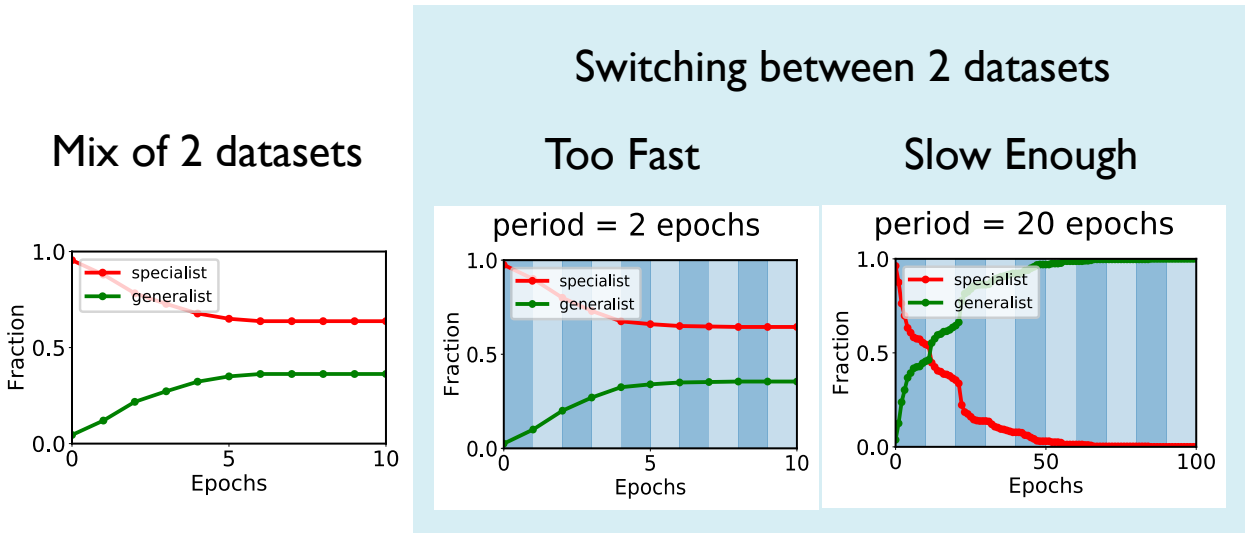

Figure 4.4: In each plot, I trained 400 SVMs. Then, I classified each SVM based on which of the three features it responds most strongly with. Switching between two datasets in the right time scale helps evolve the solutions toward the generalist feature. 'Mix of 2 datasets' represents the traditional method of sampling from the entire training data. As expected, the model only learns the generalist 1/3 of a time. The figures in the blue background represent the switching methods described above. With the slow enough time scale of switching, the model eventually converges to learn the generalist.

4.3 Discussion

Swapping between domains of the training data helped our algorithm avoid domain specific features and learn the common ones. Two study cases are presented in this chapter. One is inspired from the classification problem between teachers and students webpages. The other is a classification of binary data in three dimensions. The challenge of this regularization techniques is in tuning the swapping time to be close to the response time of the neural network. The response time can be inferred from the amount of epochs a network required to learn a new feature, but this value is sensitive to changes in the learning rate and the parameters of the neural networks (see Appendix).

When the domain of the training data shifts, the model is forced to learn new features. If the domain switch back too quickly, the model may not have enough time to acquire these features. If the time of switching is more than the time required to learn new features, then

the model can gradually acquire the generalist features. The more periods have passed, the better chance that the final features in the model are generalists, because they work well in all domains.

The results in this chapter may be limited to artificial data. Given more time and resources, investigation of this technique on real dataset for the Domain Generalization problem like PACS or rotated MNIST are required [118, 119].

CHAPTER 5

DISCUSSION AND OUTLOOK

We have investigated diverse dynamic non-linear phenomena in biology and machine learning by focusing on two parameters: the timescale of changes in an external signal and the timescale of the system's internal response. In the regime where the signal time is much larger than the response time, the system may change the equilibrium state. On the other hand, if the response time is much larger than the signal time, the system often remains in the same equilibrium state. The insights on the two different regimes of parameters offer clues toward the intermediate regime when the ratio of the times is close to one. For the biological clocks in chapter 2, we learn the trade-off principle for the robustness against internal and external noise. For the noise activation phenomena in chapter 3, we can explain why the cells only respond to noise with a particular correlation time. For the gradient descent dynamics in chapter 4, we obtain a novel regularization method by tuning the signal and the response time. Since the systems we consider are highly non-linear, both timescales need to be measured at the specific signal change because the behavior of driven nonlinear systems depends on the strength of the external signal.

Our analytical process of considering the ratio of the response time to the signal time may not apply to every system. If the system does not have a stable solution, for example, a chaotic system like the Lorentz attractor, then the response time is ill-defined. Besides, the response times sometimes need to be computed numerically for different values of the signal changes. Therefore, this process is not always analytical.

Future works involve finding more applications of these time regime analyses. More applications would give better insights about comparing the signal and the system time. Furthermore, a mathematical study on the necessary conditions for this time regime analysis will the application of the analysis more credible.

APPENDIX A

BIOPHYSICAL CLOCKS

A.1 Trade-off in Kai-based clocks

We demonstrate our trade-off using Gillespie simulations of an explicit biomolecular KaiABC model of the post-translational clocks in *S. elongatus* and *P. marinus*.

S . elongatus clock - hexamers with collective KaiA feedback

The *S. elongatus* clock has been well-characterized experimentally [33, 34, 36, 37] - see Fig.A.1a. The clock is fundamentally based on the ordered phosphorylation and dephosphorylation of KaiC [15]. Phosphorylation of KaiC is KaiA-dependent which allows for feedback that enables collective coherent oscillations in a cell. After complete phosphorylation of KaiA-C complexes (usually by the end of the day), KaiC forms a KaiB-C complex which then dephosphorylates in an ordered manner. Crucially, the KaiB-C complex also sequesters KaiA in a KaiABC complex, reducing the pool of available KaiA for phosphorylation of other KaiC hexamers. This negative feedback enables coherent oscillations of the population of KaiC molcules in a single cell[15].

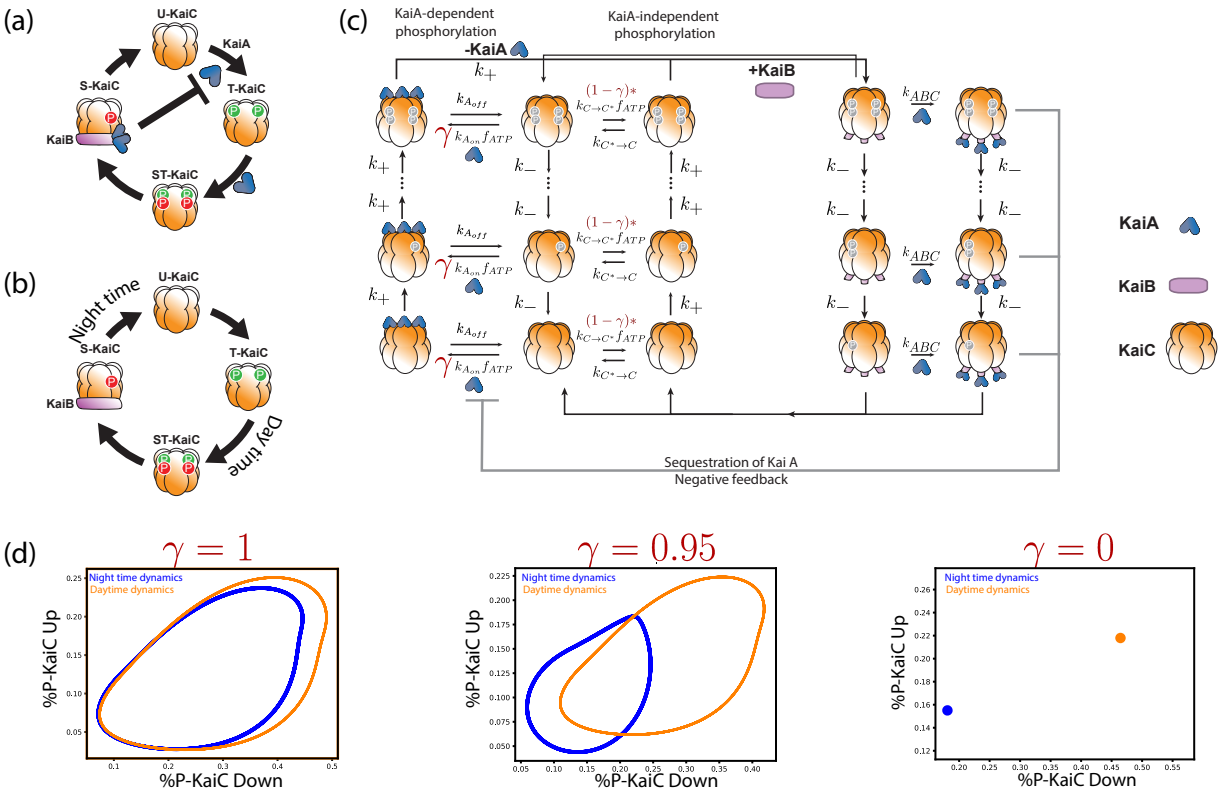

Figure A.1: Explicit biochemical KaiABC model simulated using the Gillespie algorithm. (a) The experimentally well-characterized clock in *S. elongatus* consists of a negative feedback-enabled self-sutained oscillator. KaiBC complexes sequester KaiA, preventing runaway KaiC molecules from going through the cycle independently. (b) The genome of *P. marinus* lacks *kaiA*. We assume a minimal model consistent with known facts [15] about this clock; KaiC phosphorylation proceeds without KaiA and hence different KaiC hexamers can proceed independently through the cycle. (c) We combine both clocks in one model with an interpolating parameter γ that selects between an *S. elongatus*-like KaiA-dependent pathway and an *P. marins*-like KaiA-independent pathway. All reactions shown are assumed to be first order mass-action kinetics. We simulate such a system at different overall copy numbers N using the Gillespie algorithm. (d) We find limit cycles for $\gamma > 0.9$. The resulting limit cycles for $\gamma = 1, 0.95$ violate the simplifying assumptions used in our dynamical systems (e.g., non-circular cycles of different size); and yet our results are qualitatively validated by this model (Fig.1d from the main text).

P. marinus model - independent hexamers

P. marinus lacks the *kaiA* gene but possesses and expresses *kaiB* and *kaiC*. While the details of the protein clock dynamics are not as fully known as with *S. elongatus*, gene expression shows cycling in cycling conditions but decays in constant conditions [35]. A

conservative model, consistent with all these known facts about *P. marinus*, is shown in Fig.A.1b; without KaiA feedback, different hexamer units phosphorylate independently and settle to a hyperphosphorylated state at the end of the day. At night, they dephosphorylate along a distinct pathway (homologous to that used by *S. elongatus* but without KaiA) and reach a hypophosphorylated state by dawn.

Hybrid model

We created the following hybrid model that includes *S. elongatus* and *P. marinus* models as different limits. In our model, shown in Fig.A.1c, KaiC has a KaiA-dependent phosphorylation pathway, much like in *S. elongatus*, that is used during the day and driven forward by ATP.

But to also include *P. marinus*-like behavior in the model, we allow for a second parallel phosphorylation pathway for KaiC that is independent of KaiA. The relative access of these two pathways is controlled by a parameter γ. When $\gamma = 1$, only the *S. elongatus*-like KaiA dependent pathway is accessible. When $\gamma = 0$, only the *P. marinus*-like KaiA independent pathway is accessible. Collectively, we call these states along these phosphorylation pathways, the UP states of KaiC - phosphorylation are going UP along these pathways which are usually used during the day.

After maximum phosphorylation (usually at dusk), KaiA unbinds (if present) and a KaiB-based dephosphorylation pathway takes over (common to both systems). We call these states the DOWN states of KaiC.

Critically, KaiA is assumed to be sequestered through the formation of KaiABC complexes during this dephosphorylation stage. In *S. elongatus*, reduced KaiA availability prevents other KaiC hexamers from proceeding independently through the UP stage while most of the population is in the DOWN state. Such negative feedback is critical in maintaining free-running limit cycle oscillations in *S. elongatus*.

However, as $\gamma \to 0$, the KaiA-independent pathway is more active and thus the system effectively has no feedback. In fact, we find that at about $\gamma \approx 0.82$, sustained oscillations disappear (for kinetic parameters used here and reported below). Hence we chose $\gamma = 1, 0.95, 0$ as representative of two limit cycle-based free running clocks and one point-attractor based damped clock respectively. In this way, we can view the clock dynamics of *S. elongatus* and *P. Marinus* can be viewed as being on either side of the Hopf bifurcation that occurs at $\gamma \approx 0.82$.

Gillespie simulations

We ran explicit Gillespie simulations corresponding to the deterministic equations above at different overall copy number N with fixed stoichiometric ratios of the molecules KaiA,B,C.

We simulated external input noise by varying the ATP levels during the day. External noise in these simulations were implemented by changing ATP levels in the following way: we fluctuated the ATP levels $f_{ATP} = ATP/(ATP + ADP)$ during the day between the f_{ATP}^{day} and $f_{ATP}^{night} + (f_{ATP}^{day} - f_{ATP}^{night})/3$, where f_{ATP}^{day}, f_{ATP}^{night} are the ATP values during a cloudless day and night respectively. We used different day and night ATP levels for different γ that ensure that the limit cycles had periods comparable to 24 hours. For $\gamma = 1$, we used ATP/ADP ratios of $f_{ATP}^{day} = 0.55, f_{ATP}^{night} = 0.45$. For $\gamma = 0.95$, we used $f_{ATP}^{day} = 0.57, f_{ATP}^{night} = 0.17$ and for $\gamma = 0$, $f_{ATP}^{day} = 0.8, f_{ATP}^{night} = 0.2$. The corresponding limit cycles and point attractors are shown in Fig.A.1d.

We used the following kinetic parameters in all simulations: $dt = 0.01\ hr, k_+ = k_- = 2m \cdot 0.04932\ hr^{-1}, k_{Aon} = 0.2466\ \mu M^{-1} hr^{-1}, k_{Aoff} = 0.02466\ hr^{-1}, k_{C \to C*} = 0.2466\ hr^{-1}$, $k_{C* \to C} = 0.1\ k_{C \to C*}, k_{ABC} = 123.30\ hr^{-1}, m = 18$. We set up Kai C and Kai A in a $1:1$ stoichiometric ratio, each present at a copy number N where N was varied systematically. These rates are consistent with those measured in [120, 121, 122].

Much like with Langevin simulations of dynamical systems performed in this paper, we

run the Gillespie simulation until equilibration of the population. However, the system appears to reach the equilibrium state much faster (only over 5 light-dark cycles of 12h:12h). We extracted one day of such a trajectory on day 6 and repeated the simulation 100-400 times. We repeat 400 times when the copy number is low (< 1200) since the spread will be big and we found that the probability distribution is not smooth. We run only 100 times for the high copy number (> 1200). Pooling together these trajectories, we computed the mutual information between clock state (i.e., (u, d) where u is the net phosphorylation state of KaiC in the up-pathways and d is the net phosphorylation state of KaiC in the KaiB-bound 'down' pathways in Fig.A.1c) and time of day. The (u, d) space was binned using bins of fixed size of dimension $(0.05, 0.05)$ while the 24 hr time-of-day was binned with bins of size 0.5 hrs.

Phase portrait

With these choices, we see in Fig.A.1d, that this model has limit cycles of different position during the day and night. The corresponding experimental data, reproduced from [11], is presented in the main paper.

A.2 Other oscillators

Here, we study the effect of internal and external noise on a diverse array of biochemical oscillator models from the literature in the parameter regimes described in the original publications. We confirm the same trade-off described in the paper in these models; a summary of our results is presented in the main paper.

In all of the models described below, we set all parameters to values used in the original or cited papers with only two exceptions: (a) the parameter identified as coupling to external signals in these publications is varied over time as a square wave with amplitude fluctuations added, (b) the parameter designated by the relevant original publication as controlling the

distance from the Hopf bifurcation was used to simulate a point attractor-based 'hourglass' oscillator (red lines in Fig. 2.2 of the main paper) and limit cycles of different free running oscillation amplitude (green and purple lines in Fig. 2.2 of the main paper). This latter parameter roughly corresponds to R, the size of limit cycle, while the amplitude of square wave coupled to the former parameter corresponds to L, the separation of the limit cycles, in our dynamical systems theory, i.e., the separation of the 'day' and 'night' limit cycles. (In several papers, these two are the same parameter, in which case the day-night difference reflects L while the mean value reflects R.) Finally, we add Langevin noise to the equations to simulate internal noise; when available, we followed the finite volume prescription for rates in these papers to set the size of Langevin noise for each reaction.

We keep the strength of external noise ϵ_{ext}, defined as the noise-to-signal ratio of the amplitude fluctuations in the external signal, fixed. We varied internal noise ϵ_{int} along the x axis of plots in Fig. 2.2 of the main paper. Here, ϵ_{int} is defined as the *phase* diffusion constant of a clock in undriven conditions (see how we define internal noise in the section on Neurospora and Drosophilia below); this normalization, which depends on the Hopf bifurcation parameter in (b) above, allows us to make a fair comparison between different clocks since they develop the same phase variance over the same time in undriven conditions.

As seen in Fig. 2.2 of the main paper, these diverse models agree with the trends found in our analysis of dynamical systems and with simulations of the KaiABC system, showing that our results are not tied to any particular molecular model.

A.2.1 Neurospora and Drosophilla circadian clocks by Goldbeter

The circadian clock in *Neurospora* has been modeled [1] as arising from interactions between mRNA (M) and a protein that can shuttle in and out of a nucleus (P_N, P_C). The equations

used in [123] to model this are,

$$\begin{aligned}
\frac{dM}{dt} &= v_s\Omega\frac{(K_I\Omega)^n}{(K_I\Omega)^n + P_N^n} + v_m\Omega\frac{M}{K_m\Omega + M} \\
\frac{dP_C}{dt} &= k_s M - v_d\Omega\frac{P_C}{K_d\Omega + P_C} - k_1 P_C + k_2 P_N \\
\frac{dP_N}{dt} &= k_1 P_C - k_2 P_N
\end{aligned} \tag{A.1}$$

where ν_s is an mRNA transcription rate that is modulated by external signals [1, 123], and Ω is the volume of the system which in turn determines the strength of stochastic noise. (A model with very similar equations has also been suggested as a model of the *Drosophila* circadian clock [1].)

We use the same parameters used in the Ref. [1, 123]: $K_I = 1$ nM, $n = 4$, $v_m = 0.505$ nM/h, $K_m = 0.5$ nM, k_s= 0.5 1/h, $v_d = 1.4$ nM/h, K_d= 0.13 nM, $k_1 = 0.5$ nM/h, $k_2 =$ 0.6 nM/h, and assume the volume Ω dependence of these parameters to be exactly as used in [123]. We add internal stochasticity by adding Langevin noise with a diffusion matrix [123] :

$$\frac{dX}{dt} = \mu(x,t) + \Sigma(x,t)\eta(0,1) \tag{A.2}$$

where $\mu(x,t)$ is the RHS of Eqn.A.1, $\eta(0,1)$ is a vector whose entries are independent standardized Gaussian noise (mean 0, variance 1), and

$$\Sigma(x,t) = \begin{bmatrix} \sqrt{A} & \sqrt{B} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{k_s M} & -\sqrt{k_1 P_C} & \sqrt{k_2 P_N} \\ 0 & 0 & 0 & \sqrt{k_1 P_C} & -\sqrt{k_2 P_N} \end{bmatrix} \tag{A.3}$$

$$= \begin{bmatrix} \sqrt{A} & \sqrt{B} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{k_s M} & \sqrt{k_1 P_C} & \sqrt{k_2 P_N} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{k_1 P_C} & \sqrt{k_2 P_N} \end{bmatrix} \tag{A.4}$$

where $A = v_s\Omega\frac{(K_I\Omega)^n}{(K_I\Omega)^n+P_N^n}$ and $B = v_m\Omega\frac{M}{K_m\Omega+M}$. This is how internal noise get added into other oscillators models as well. However, for the system of equations that use concentration instead of the number of molecules, the equation has to be modified to $\frac{dX}{dt} = \mu(x,t) + \frac{1}{\sqrt{\Omega}}\Sigma(x,t)\eta(0,1)$.

As in [1, 123], we take ν_s to be modulated by the external signal (light). As shown in [1, 123], a Hopf bifurcation occurs at $\nu_s = 0.57$ nM/h. Hence, in Fig. 2.2 from the main text, we used $\nu_s^{Day} = 0.55$ nM/h,$\nu_s^{Night} = 0.05$ nM/h for the point attractor (red). For the two limit cycles, we used $\nu_s^{Day} = 0.9$ nM/h,$\nu_s^{Night} = 0.6$ nM/h (green), and $\nu_s^{Day} = 0.705$ nM/h,$\nu_s^{Night} = 0.695$ nM/h (purple). The driving period is 18 h, similar to the driving period of the system at $v_1 = 0.7$ nM/h

A.2.2 Arabidopsis circadian clock by Millar et al

A model of the circadian clock in *Arabidopsis thaliana* was introduced in [3]. While many biologically important features have been added in the years since, the original model was based on a single negative feedback loop and involves two transcription factors (LHY and CCA1) that inhibit their activator $TOC1$. A reduced model with the same phenomenology was presented in [2], in which LHY and CCA1 are combined into one variable, representing their mRNA and protein levels by $M_L(t)$ and $P_L(t)$ respectively. Denoting the mRNA and

protein levels of $TOC1$ by $M_T(t)$ and $P_T(t)$, [2] presents a reduced version of the model in [3] as:

$$\begin{aligned}
\frac{dM_L}{dt} &= L(t) + v_1 \frac{P_T^a}{g_1^a + P_T^a} - m_1 \frac{M_L}{k_1 + M_L} \\
\frac{dP_L}{dt} &= p_1 M_L - m_2 \frac{P_L}{k_2 + P_L} \\
\frac{dM_T}{dt} &= v_2 \frac{g_2^b}{g_2^b + P_L^b} - m_3 \frac{M_T}{k_3 + M_T} \\
\frac{dP_N}{dt} &= p_2 M_T - m_4 \frac{P_T}{k_4 + P_T}
\end{aligned} \tag{A.5}$$

where $L(t)$ is a light input function, and other parameters except the variables specified on the left hand sides are constant.

To simulate this system, we use the parameters used in [2]: a = 2, b = 2, $g_1 = 0.5$ nM, $g_2 = 0.1$ nM, $m_1 = 0.4$ nM/h, $m_2 = 0.6$ nM/h, $m_3 = 0.6$ nM/h, $m_4 = 0.3$ nM/h, $k_1 = 1$ nM, $k_2 = 0.5$ nM, $k_3 = 1$ nM, $k_3 = 1$ nM, $p_1 = 0.5$ 1/h, $p_2 = 0.3$ 1/h, $v_2 = 0.6$ nM/h. With other parameters fixed, the system undergoes Hopf bifurcation at $v_1 = 0.194$ nM/h We use $v_1 = 0.26$ nM/h for limit cycles and $v_1 = 0.05$ nM/h for point attractor. $L(t)$ is a light input function. For the two limit cycles in Fig. 2.2 in the main text, we set $L^{Day} = 0.05$ nM/h and $L^{Night} = 0$ (green data) and we set $L^{Day} = 0.01$ nM/h and $L^{Night} = 0$ (purple data). For point attractor, we set $L^{Day} = 0.2$ nM/h and $L^{Night} = 0$ (red data). The period of the driving signal is 24 h, which is around the natural period of the system when $v_1 = 0.26$ nM/h and $L = 0$.

A.2.3 Mammalian Per-Cry circadian clock by Leloup et al

The circadian clock in mammalian cells was modeled in detail [4] by LeLoup and Goldbeter, using 19 equations representing the interactions between Per, Cry and other genes. We simulate this entire system explicitly with the parameter values specified in the original

publication [4]. To introduce Langevin noise, we use a simplified diagonal diffusion matrix with entry $\sqrt{DX_i}$ for species X_i. We do not reproduce these 19 equations or parameter values used from [4] here in interest of space; the only modification we made is to introduce Langevin noise to each of the 19 equations.

LeLoup and Goldbeter [4] identified parameter v_{sP} (a transcriptional rate) to be the light input function. We use $v_{sP}^{Day} = 1.09$ nM/h and $v_{sP}^{Night} = 1.07$ nM/h for the purple limit cycle data in Fig 2.2, $v_{sP}^{Day} = 1.15$ nM/h and $v_{sP}^{Night} = 1.07$ nM/h for the green limit cycle data. For the point attactor data (red), we set $v_{sP}^{Day} = 1.5$ nM/h and $v_{sP}^{Night} = 0$. In addition, [4] identified parameters KAC, vmB as controlling the distance from the Hopf bifurcation. For the point attractor, we used KAC = 0.4 nM, and vmB = 0.9 nM/h (also used in [4]). For the limit cycles, we used KAC = 0.6 nM, and vmB = 0.8 nM/h, which lies on the other side of the Hopf bifurcation. The period of the input signal is at 21.5 h, which is around the natural period of the system when $v_{sP} = 1.07$ nM/h.

A.2.4 cdc2-cyclin cell cycle by Goldbeter

A classic model of the cell cycle was proposed by [5]. While many additional details have been added on since, the model captures the essential mechanism behind the self-sustained nature of cell cycles.

$$\begin{aligned}
\frac{dC}{dt} &= v_i\Omega - k_dC - v_dX\Omega\frac{C}{K_d\Omega + C} \\
\frac{dM}{dt} &= v_1\frac{C}{K_C\Omega + C}\frac{1-M}{K_1 + (1-M)} - V_2\frac{M}{K_2 + M} \\
\frac{dX}{dt} &= v_3M\frac{(1-X)}{K_3 + (1-X)} - V_4\frac{X}{K_4 + X}
\end{aligned} \tag{A.6}$$

where Ω is the size of the system and other parameters are constants. The three variables are the cyclin concentration C, the fraction of active *cdc2* kinase M, and the fraction of active cyclin protease X. For C, the internal noise is proportional to the square root of

the rates, but for M and X, it is proportional to the square root of the rates divided by Ω because they are fractions and not concentrations (following the prescription in [123] for a similar model). Parameter values were taken from [5]: $K_i = 0.1$ (i = 1-4), $V_{M1} = 0.5$ min^{-1}, $V_2 = 0.167$ min$^{-}1$, $V_{M3} = 0.2$ min^{-1}, $V_4 = 0.1$ min^{-1}, $v_d = 0.1$ μMmin^{-1}, $K_C = 0.3\mu M$, $K_d = 0.02$ μM, $k_d = 3.33 \times 10^{-3}$ min^{-1}.

Goldbeter [5] suggested that v_i is modulated by external signals. So, we use $v_i^{Day} = 0.0106$ μMmin^{-1} and $v_i^{Night} = 0.0105$ μMmin^{-1} for small L/R limit cycle, $v_i^{Day} = 0.0111$ μMmin^{-1} and $v_i^{Night} = 0.0105$ μMmin^{-1} for large L/R limit cycle, and $v_i^{Day} = 0.009$ μMmin^{-1} and $v_i^{Night} = 0$ for point attractor. The bifurcation from point attractor to limit cycle happen around $v_i = 0.01$ μMmin^{-1}. The period of the driving signal is 35 minutes.

A.2.5 Goodwin oscillator

One of the earliest models of biochemical oscillators was proposed by Goodwin [6] (later corrected). We use the simplest widely-studied version of such a Goodwin oscillator [7, 124],

$$\begin{aligned} \frac{dX}{dt} &= \frac{\alpha(t)}{1+Z^n} - X \\ \frac{dY}{dt} &= X - Y \\ \frac{dZ}{dt} &= Y - Z \end{aligned} \tag{A.7}$$

When $n = 9$, the limit cycles disappear at a Hopf bifurcation found at $\alpha \approx 7$. As is commonly done [124], we couple the external signal to the bifurcation parameter $\alpha(t)$. We use $\alpha^{Day} = 120, \alpha^{Night} = 80$ for the green limit cycle in Fig.2.2c of the main paper, $\alpha^{Day} = 108, \alpha^{Night} = 92$ for purple limit cycle data, and $\alpha^{Day} = 2.5, \alpha^{Night} = 1$ for the red point attractor data. The input signal has a period of 4, which is roughly the natural period of the limit cycle at $\alpha = 100$. X, Y are taken to the output of the clock for computing MI.

A.2.6 Repressilator

The repressilator is a model of an early synthetic biology system [9] that demonstrated oscillations in a synthetically wired gene regulatory circuit. While resembling the Goodwin oscillator in topology, the network has the total non-linearity spread equally amongst all three reactions, lowering the requisite Hill coefficient of any one reaction to a biochemically realistic $n = 3$. Repressilator circuits are not usually driven by an external signal, except in a few theoretical analyses (e.g., [125, 126]). We use the simplest version of these, with the driving signal modulating the transcription rate of only one of the proteins

$$\begin{aligned} \frac{dX}{dt} &= \frac{\alpha * (1 + s(t))}{1 + Y^n} - X \\ \frac{dY}{dt} &= \frac{\alpha}{1 + Z^n} - Y \\ \frac{dZ}{dt} &= \frac{\alpha}{1 + X^n} - Z \end{aligned} \tag{A.8}$$

where α is the bifurcation parameter and $s(t)$ is the variable coupled to the input signal. In a non-driven repressilator, $s(t) = 0$. When $n = 3$, the Hopf bifurcation occur at $\alpha \approx 2.5$, so for limit cycles, we use $\alpha = 5.2$ and for point attractor we use $\alpha = 1.9$.

We use $s^{Day} = 0.7/5.2, s^{Night} = -0.7/5.2$ for the green limit cycle in Fig.2.2g in the main paper, $s^{Day} = 4.8/5.2, s^{Night} = -1.7/5.2$ for the purple limit cycle data, and $s^{Day} = 0.5/1.9, s^{Night} = -1.9/1.9$ for the point attractor (red). The input signal has a period of 4, which is roughly the the natural period of the limit cycle at $\alpha = 5.2$ and $s(t) = 0$. X, Y are taken to the output of the clock for computing MI.

A.2.7 Brusselator

The Brusselator is a model of autocatalytic reactions that show limit cycle oscillations. This model has been extensively studied over the years; while the explicit biochemical reactions

can be found in [8], these reactions are modeled by the ODEs:

$$\begin{aligned} \frac{dX}{dt} &= 1 - (1 + b(t))X + X^2Y \\ \frac{dY}{dt} &= b(t)X - X^2Y \end{aligned} \tag{A.9}$$

where b has been identified as a bifurcation parameter [8]. Most studies do not consider driven Brusselator models; we follow the driving prescriptions of the Goodwin model and couple the external light to the bifurcation parameter, converting the constant b into $b(t)$. The bifurcation point is at $b = 2$. For $b < 2$, we have a point attractor and for $b > 2$ we have a limit cycle. We use $b^{Day} = 2.25$ and $b^{Night} = 2.2$ for the purple limit cycle data in Fig.2.2f from main text. We use $b^{Day} = 2.8$ and $b^{Night} = 2.2$ for the green limit cycle data. Lastly, we use $b^{Day} = 1.8$ and $b^{Night} = 0.5$ for the point attractor (red). The signal has a period of 6.4 which is around the natural period of the system when $b = 2.2$.

A.3 Optimal Dynamical system and trade-off

To complement our study of detailed biochemical implementations of such systems, we study two kinds of dynamical systems in this paper; limit cycles and point attractors. The minimal model of limit cycles and point attractors is given by the 'normal form' near a Hopf bifurcation:

$$\dot{r} = \alpha \left(r - \frac{r^3}{R^2} \right) \tag{A.10}$$

$$\dot{\theta} = \omega \tag{A.11}$$

For $\alpha > 0$, the above equation describes a circular limit cycle of radius R and frequency ω. This equation undergoes a Hopf bifurcation at $\alpha = 0$, where the limit cycle shrinks to zero and resulting in point attractor for $\alpha < 0$. The 'normal form' can be seen as the

universal simple form – e.g., circular limit cycles of radius R – that any limit cycle and point attractor will reduce to in the neighborhood of a Hopf bifurcation. We add White noise in the Cartesian space representation of the Dynamical equations to represent the internal noise as follows:

$$\mathrm{d}x = \left(\alpha\left(1 - \frac{x^2+y^2}{R^2}\right)x - \omega y\right)\mathrm{d}t + \sqrt{2D}\ \mathrm{d}W \tag{A.12}$$

$$\mathrm{d}y = \left(\alpha\left(1 - \frac{x^2+y^2}{R^2}\right)y + \omega x\right)\mathrm{d}t + \sqrt{2D}\ \mathrm{d}W \tag{A.13}$$

where $D \sim R^2\epsilon_{int}^2$ is the diffusion constant, $\mathrm{d}W$ is a Wiener process, and here we assume that the internal noise is a homogeneous white-noise in the 2-dimensional space. (Similar assumptions are made in reference [43].)

While we assume this simple form here as a minimal model, we do not assume that the oscillator is weakly driven. Instead, based on experimental observations of the Kai clock [11] and models of numerous other clocks [12], we assume that the origin of the limit cycle or point attractor equations above moves by a finite amount L as the external light signal switches between day and night values. In fact, we move the origin along the x-axis as a function of time as $(Ls(t), 0)$ where $s(t)$ is the external light signal, assumed to be of amplitude 1. Thus we are assuming a simple circular form of limit cycles and point attractors but do not restrict to weak driving. (In the limit of weak driving, i.e., small L/R, our model can be shown to reduce to the universal Stuart-Landau model of weakly driven oscillators as a special case.)

In Eqn. A.11, $\tau_{\mathrm{relax}} \sim \frac{1}{|\alpha|}$ is the relaxation time for perturbations away from the limit cycle or point attractor. For limit cycles, perturbations away from the limit cycle tend to decay fast relative to the period $2\pi/\omega$, typically on the order of hours [11].

In contrast, the point attractor in damped 'hourglass' clocks *P. marinus* needs to have relaxation $\tau_{\mathrm{relax}} \sim \frac{1}{|\alpha|} \sim 2\pi/\omega$ comparable to the period of the day-night cycle itself. As explained in the main paper, if relaxation were much faster, the clock state would decay to a fixed point before the end of the day (or night) and thus not show distinct states at distinct

times of the day-night cycle.

Simulations

For both limit cycles and point attractors, we simulate a population of clocks, each represented by a particle in the given dynamical system, subject to external and/or internal noise.

We use $\alpha = 5$ for limit cycle system and $\alpha = -5$ for point attractor system where $\omega = 2\pi$ in these units. For point attractors, we set $R = 1000L$, where L is the separation of the day and night attractor. In such a limit, the point attractors are quadratic potentials with linear restoring forces since $\frac{r^3}{R^2}$ is small. The center of the cycle and point attractors during the day are assumed to be at $(-L, 0)$ and at $(0, 0)$ at night; or more generally at $(-Ls(t), 0)$ were $s(t)$ the light signal (assumed to be of amplitude 1).

To simulate external noise, we use a square wave signal $s(t)$ of amplitude 1 with amplitude fluctuations set by ϵ_{ext}. As explained in the Appendix A.3 on external noise, we take the fluctuations in $s(t)$ to have a correlation time of 2.4 hours.

To simulate internal noise, we add Langevin noise to Eqns. A.11 as described in the Appendix A.6 on Langevin noise. As explained in that Appendix A.6, our measure of internal noise ϵ_{int}^2 is a measure of *phase* diffusion, independent of limit cycle size. In other words, ϵ_{int}^2 is a measure of the population phase variance (i.e., variance in θ) developed by limit cycles of any size in undriven conditions in a given period of time.

To interpolate between limit cycles and point attractors, we systematically change L holding R fixed. For limit cycle simulations, changing L and R are equivalent. To see this for external noise simulations, note that L/R is the only dimensionless parameter. For internal noise, our definition of ϵ_{int}^2 above as the phase diffusion in undriven conditions for limit cycles of any size, ensures that changing L and R are equivalent for limit cycles.

For point attractors, we set the separation L be the diameter of the limit cycles simulated

in the same plots to keep the size of the resolvable chemical spaces roughly comparable for limit cycles and point attractors. While this precise choice is arbitrary to some extent, note that the point attractor results for external noise do not depend on this parameter at all since L is the only relevant length scale in external noise simulations. The separation L does affect the clock precision with internal noise (of fixed absolute strength ϵ_{int}) but the mutual information changes only logarithmically with L.

We evolve our dynamical system using the Euler method with time step $dt = 0.001$ days until the value of mutual information from one day to the next does not change by more than 2-3% - i.e,. the system has reached steady state. Reaching steady-state usually takes around 200 days, but if the ratio of L/R is smaller than 0.1, then we may need to run the simulation until day 500 to reach an equilibrium (See speed-error tradeoff in Fig. 2.6b and c in the main text).

For limit cycles, we initialize the population of 10^4 particles by uniformly distributing them along the perimeter of the night cycle. In the point attractor system, we initialize a population of 10^5 at the night-time point attractor.

We use a larger population with point attractors since the particles tend to be distributed over a larger area of the dynamical system. Note that we bin the population by position to compute mututal information between position in the 2d state space and time. Doing so reliably requires a smooth distribution after binning. For limit cycles, the particles usually stay close to attractor and thus provide sufficient count in each bin. However, for the point attractor, the population is usually spread over the entire 2d area between the two point attractors. Therefore, we need 10^5 particles to get an accurate value of mutual information of point attractor system.

Trade-off

To find the optimal dynamical system geometry that operates with best accuracy when both internal noise and external noises are present, recall that we derived the following equations for strongly-driven limit cycles,

$$\sigma_{int}^2 \sim \frac{\epsilon_{int}^2 T}{s^2 - 1} \tag{A.14}$$

$$\sigma_{ext}^2 \sim \frac{\Delta\Phi^2}{s^4 - 1} \tag{A.15}$$

For a small L/R, we had found that $\Delta\Phi \sim \epsilon_{ext} L/R$ where ϵ_{ext}^2 is a measure of the variance of light during the day. Further, we showed that $s^2 - 1 \sim L/R$ in this limit. Hence, in the small L/R ('Stuart-Landau') regime, the above equations reduce to,

$$\sigma_{int}^2 \sim \epsilon_{int}^2 R/L \tag{A.16}$$

$$\sigma_{ext}^2 \sim \epsilon_{ext}^2 L/R \tag{A.17}$$

The population variance when both noises are present is approximately given by $\max(\epsilon_{int}^2 R/L, \epsilon_{ext}^2 L/R)$. This variance is minimized when the two terms are equal, giving

$$\left(\frac{L}{R}\right)_{optimal} \sim \frac{\epsilon_{int}}{\epsilon_{ext}},$$

which defines the optimal geometry of the dynamical system for given strength of internal and external noise.

In contrast, by taking the product of the equations above, we find the trade-off relationship,

$$\sigma_{ext}^2 \sigma_{int}^2 \sim Q \equiv \epsilon_{ext}^2 \epsilon_{int}^2 \tag{A.18}$$

The trade-off relationship above clarifies which parameters are held fixed and which ones are varied in our trade-off. If Q is held fixed, this trade-off relationship holds under variations of all the parameters of the normal form of limit cycles (Eqn.A.11). (While L/R allows us to navigate the trade-off by increasing one of $\sigma^2_{ext}, \sigma^2_{int}$ and decreasing the other, other parameters such as the relaxation time leave both $\sigma^2_{ext}, \sigma^2_{int}$ relatively unaffected.)

Holding Q fixed does involve holding the strength of external ϵ_{ext} and internal ϵ_{int} noise fixed. In all the models studied here, ϵ_{ext} is simply defined as the size of the amplitude fluctuations in the external signal relative to the amplitude of the external signal itself - i.e., the noise-to-signal ratio of the external signal - with no reference to the clock dynamics. Changing L/R and other parameters can strength or weaken the coupling of this noisy external signal to the clock but do not affect the signal-to-noise ratio of the external signal itself.

Analogously, the strength of internal noise ϵ^2_{int} is defined as the diffusion constant for the *phase* of a clock (e.g., in radians^2) in the absence of an external driving signal. As discussed in Appendix A.2.7, this definition ensures that limit cycles of different sizes develop the same phase variance over the same time when subject to the same ϵ_{int}.

For some purposes, it may make sense to hold the dimensionful diffusion constant $D_{int} = \epsilon^2_{int}R^2$ fixed while making comparisons. In this case, in addition to the trade-off effect discussed in this paper, large limit cycles are given an additional robustness to internal noise, trivially by virtue of their size, since the diffusion constant D_{int} in chemical space is held fixed (instead of the dimensionless diffusion constant ϵ^2_{int} for clock phase). In this case, it is insightful to re-write $Q = \epsilon^2_{ext}\epsilon^2_{int} = \epsilon^2_{ext}D_{int}/R^2$ and recognize that $P = Q^{-1}$ is a measure of the power needed to maintain free running clock oscillations [127] - larger cycles cost more energy per cycle to maintain. Thus, in this case, our trade-off should be understood as one at fixed power.

A.4 Modeling noise

Modeling external noise (weather fluctuations) We generate a square wave of period 24 hours to model the day-night cycle of light on Earth with the day length of 12 hours. However, such a square wave is modulated by weather fluctuations, e.g. periods of reduced intensity due to passing clouds during the daytime. We model such fluctuating intensity as follows. We assume each weather condition lasts a random interval of time drawn from an exponential distribution of mean 2.4 hrs (1/10 of a day). During a given weather condition, we set the intensity of light to a random value, drawn uniformly from $[1 - \text{noise}_{\text{ext}}, 1]$ where $\text{noise}_{\text{ext}}$ is the strength of the external noise: 0 means no external noise and 1 means full external noise. This random value will range from 0 to 1 where 1 represents the maximum intensity during the day. (At night, the intensity is held at zero with no fluctuations). In the simulation of our limit cycle model, we set $\text{noise}_{\text{ext}}$ to 1. However, in our simulations for 8 different models of biological clocks, $\text{noise}_{\text{ext}}$ ranges from 0.5 to 1 because when $\text{noise}_{\text{ext}}$ is too high, the system may not get entrained due to the difference in the natural and driving frequencies.

This noisy external signal is coupled to the diverse range of systems studied here in different ways as described in the respective sections. For each system, we simulate a population of organisms where each individual is subject to a different realization of the weather conditions described above. (a) In the Kai clock, the light signal is taken to affect the cellular ATP levels. (b) In the other eight diverse oscillators in the main paper, we coupled the light signal to the parameter specified as coupled to external signals in the original publications. (c) For the dynamical systems model, we assume that the position of the limit cycle is moved by the light signal. When the light intensity is reduced during the day to a value $\rho \in [0, 1]$, we switch the dynamics to an alternative limit cycle (or point attractor) at a fractional distance ρ between the ideal day and night cycles. For example, assume the night cycle is centered at $(0, 0)$ and the day cycle is centered at $(-L, 0)$. During a weather condition with intensity

$\rho \in [0, 1]$, we follow dynamics due to a limit cycle located at $(-\rho L, 0)$. We follow the same rules for the point attractor.

Note that we have used a square wave to approximate the natural cycle of light on earth. The square wave also allows for an intuitive derivation of Eqns. 2.1, 2.2 by dividing up the day-night cycle into four parts: diffusion during the day and during the night, contracting variance during dawn and dusk. For other waveforms, such a clear separation is not possible and all these processes occur concurrently. However, Eqns.2.1, 2.2 are expected to still hold up to $O(1)$ prefactors. Numerically, we tested sinusoidal inputs and verified our trade-off relationship.

Modeling internal noise The internal noise represents any source of stochasticity intrinsic to a single cell that would exist even in constant conditions. Such noise could be due to finite copy numbers of molecules, bursty of transcription etc.

We model internal noise in the Kai clock using explicit Gillespie simulations at finite copy number N as described in the section on Kai clocks. For the diverse other biochemical clocks studied here, we add Langevin noise to the dynamical equations, following the prescriptions laid out in the original publications when available. In the dynamical systems models, we model internal noise by adding Langevin noise to the dynamical equations as described in the section on Langevin noise. Each individual particle in our simulation is subject to a independent random realizations of such Langevin noise. In order to ensure an apples-apples comparison between different clocks, we define the strength of the internal noise ϵ_{int} to be the phase diffusion constant in undriven condition. See Appendix A.2.7 on Dynamical systems simulations for more.

Measures of clock time-telling quality

We develop and use two distinct measures of performance of noisy clocks driven by noisy inputs.

Mutual information: The performance of the clock is quantified by the mutual information between the clock state $\vec{c}$ and the time t,

$$MI(C;T) = \sum_{\vec{c}\in C, t\in T} p(\vec{c},t)\log_2\left(\frac{p(\vec{c},t)}{p(\vec{c})p(t)}\right) \tag{A.19}$$

for all $\vec{c}$ in the set of available positions C and all t in the available time bins T. (In the dynamical systems model, $\vec{c}$ represents the position in the 2d r,t plane. For the explicit KaiABC biomolecular model, $\vec{c}$ represents the phosophorylation state of KaiC.) We simulate a population of clocks, where each clock is subject to a different realization of input signals, representing different weather conditions and also subject to different realizations of internal Langevin noise (or Gillespie fluctuations). We then collect the trajectories of each clock on the last day of the simulations and calculate the probability distribution $p(\vec{c}|t)$ of clock states at a given (objective) time $t \in [0, 24]$ hrs of the last day in the simulation. The probability function $p(\vec{c})$ is calculated by accumulating the distribution of $p(\vec{c}|t)$ over time $t \in [0, 24]$ hrs of the last day. The position $\vec{c}$ and time t are binned into different bins depending on their values. We start the minimum and maximum values of the bins to the minimum and maximum values of the variables. The bin size in the time dimension is 0.48 hrs or 28.8 minutes, while The bin size in the x and y dimensions are both 0.01.

We refer to this mutual information measure as 'Precision' in Fig. A.1d, 2.2, 2.6a from the main text.

Population variance along direction of motion: Mutual information is a good indicative of how well the clock encodes information about time. However, it is calculated for the entire day. Often, we want to see how the time-telling ability of a clock changes during the day (e.g., day vs night or before and after dusk). Hence we develop a new measure, closely related to mutual information, but can be computed at specific times of day.

Intuitively, the mutual information quantifies how much the population distributions of clock states at different times t overlap. If these distributions are not overlapping, the clock state is a good readout of the time t. Such distributions are shown in Fig. 2.4b and 2.5b (purple) in the main text.

We argue that only the spread of the clock distribution along the direction of motion of the clock in state space affects mutual information. The spread of the distribution in orthogonal directions does not affect mutual information as much.

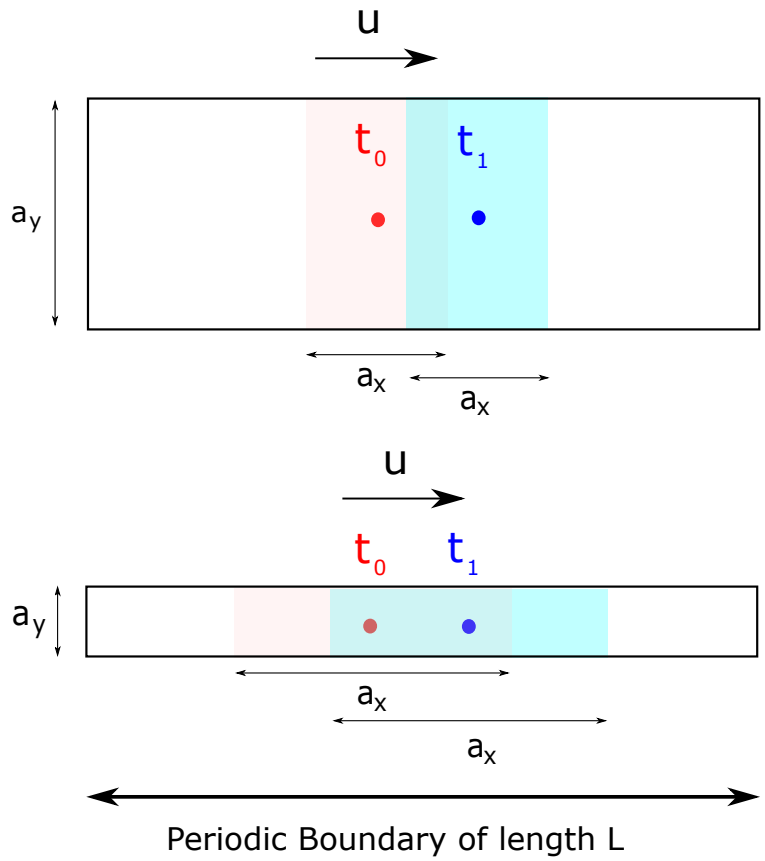

Figure A.2: Mutual information $MI(\vec{c}, t)$ between clock state $\vec{c}$ and time t is only affected by the variance of the clock state distribution $p(\vec{c}|t)$ at a given time t along the direction of motion and not orthogonal to it. In this toy example, we assume the distribution $p(\vec{c}|t)$ to be supported on a rectangle of size a_x and a_y in a 2d clock state space. The clock state moves at a speed u in the x-direction. Time telling quality is affected by how much the population at different times overlap with each other. Consequently, clocks with large a_x and small a_y (*bottom*) have lower mutual information $MI(\vec{c}, t)$ relative to clocks with small a_x and large a_y (*top*). Consequently, we use the population variance along the direction of motion as an instantaneous measure of time-telling ability in the paper.

To see this, we write mutual information between clock state $\vec{c}$ and time t as,

$$MI(C;T) = H(T) - H(T|C). \tag{A.20}$$

Here $H(T)$ is a constant, independent of the clock mechanism. Thus, MI depends entirely on the entropy of the distribution $p(t|c)$ of real times given clock state c, averaged over different clock states,

$$\begin{aligned} H(T|C) &= \int p(c)dcH(T|c) && \text{(A.21)} \\ &= -\int p(c)dc\left[\int dtp(t|c)\log p(t|c)\right] && \text{(A.22)} \end{aligned}$$

Consider a clock whose state-space is two dimensional with a periodic x-axis as shown in Fig.A.2. Further, assume that the distribution $p(\vec{c}|t)$ of clock states at a given time is supported on a rectangle of size $a_x \times a_y$ as shown in Fig.A.2 and that the clock states move along the x-axis at a uniform velocity u. This situation implies that

$$p(t|c) = \begin{cases} 0 & \text{for } |c_x - ut| > a_x \\ \frac{u}{2a_x} & \text{for } |c_x - ut| \leq a_x \end{cases}$$

So,

$$\begin{aligned} H(T|C) &= -\int p(c)dc \int_{t=(c_x-a_x)/u}^{(a_x+c_x)/u} dt \frac{u}{2a_x}\log\left(\frac{u}{2a_x}\right) \\ &= \log\left(\frac{2a_x}{u}\right) \end{aligned}$$

Since $MI(C;T) = H(T) - H(T|C)$, MI depends on $-\log a_x$ and is independent of a_y, meaning that only the spread in the direction of motion a_x affect the mutual information. Consequently, to understand the quality of time-telling at different times of the day, we

project the population variance of $p(\vec{c}|t)$ to the direction of the instantaneous velocity of the center of mass of $p(\vec{c}|t)$.

Cramer-Rao bounds

Cramer-Rao (CR) bounds quantify the total available information about phase in a given length of history of the signal. Any estimator working with that length of history must necessarily have higher variance (i.e., higher error) than the Cramer- lower bound corresponding to that length of history. In the limit of infinitely long histories, the CR bound in this context corresponds to zero error; with any finite binning in time, the upper bound on MI is simply set by the number of bins in time. In our case, this bound is given by $log_2 50 = 5.64$ bits. As shown in the main paper, as $L/R \to 0$, limit cycles process longer and longer histories of the external signal. Consequently, the mutual information for such cycles approaches the upper bound in the limit $L/R \to 0$ (assuming no internal noise) when computed with the same number of temporal bins (50 in this case).

A.5 Circle Map - Dark pulse phase shift

During the daytime, sunlight intensity fluctuates because of cloud cover and we have referred to these fluctuations as external noise. In our simulations, we subject each individual in a population to a different realization of these weather conditions and compute the resulting population variation of clock state. Such variation limits the ability of the cell to read out the objective time from the clock state.

Here, we relate the population phase variance caused by random cloud cover in our dynamical systems model to the geometrically computed Phase Response Curve (PRC) due to a single dark pulse administered during the day. Using this geometric method, we will find that the ability of limit cycle to withstand external intensity fluctuations increases with R/L, the size R of limit cycles relative to their separation L. In particular, we will show

geometrically that the gain in phase variance during the day $\sigma^2 \Rightarrow \sigma^2 + \sigma^2_{clouds}$ scales as $(L/R)^2$, in perfect agree with stochastic weather simulations.

To compute the scaling relationship of σ^2_{clouds}, we compute the phase shift $\Delta\Phi$ caused by a single dark pulse with width τ on the limit cycles with angular speed ω (i.e., the Phase Response Curve (PRC) corresponding to such a dark pulse). Fig.A.4a shows an example of a dark pulse in the signal and how it affects the trajectory. Consider a clock at state θ on the day cycle. A dark pulse of length τ administered just then will change the dynamics to that of the night cycle. This clock has state $\phi = P(\theta)$ with respect to the night cycle and will evolve for a time τ according to the night cycle dynamics, reaching a new state $\phi + \omega\tau$, at a radial position determined by R, L. At the end of the dark pulse, we use the night-day circle map, $\theta = Q(\phi)$, to find the clock state back on the day cycle. Note that all these shifts depend on the limit cycle geometry, i.e., on R and L, as shown in Fig.A.4. We can write each mapping using simple trigonometry:

$$\phi = P(\theta) = \arctan\left(\frac{L + R\sin\theta}{R\cos\theta}\right) \tag{A.23}$$

and

$$\theta^* = Q(\phi) = \arctan\left(\frac{-L + R\sin(\phi + \omega\tau)}{R\cos(\phi + \omega\tau)}\right). \tag{A.24}$$

Notice the mapping Q only differs from P by changing L to $-L$. We also include the diagram showing the transition due to dark pulse in Fig.A.4. The process "1" corresponds to $\phi = P(\theta)$, "2" corresponds to the rotation on the night cycle $\phi \to \phi + \omega\tau$, and "3" corresponds to the transition back to the day cycle $\theta^* = Q(\phi + \omega\tau)$. Combining this 3 processes, we write θ^* as $\theta^*(\theta, \tau, L/R)$ and expand it in the limit that $L/R \Rightarrow 0$ to obtain that

$$\Delta\Phi = -\frac{L}{R}\left(\cos(\theta + \omega\tau) - \cos(\theta)\right) + \mathcal{O}\left(\frac{L}{R}\right)^2 \tag{A.25}$$

where $\Delta\Phi = \theta^* - (\theta + \omega\tau)$ because $\theta + \omega\tau$ is the phase of the clock if it did not experience the dark pulse.

This expression $\Delta\Phi$ indicates the amount of phase shifted that the cloud causes. With different clocks experiencing different weather conditions, the variance gained among the population due to the fluctuation of sunlight grows like $|\Delta\Phi|^2 \sim (L/R)^2$. We see good agreement between stochastic weather simulations and this geometric computation as shown in Fig.A.4d.

In this calculation, we focused on dark pulses administered at a fixed generic time (8 AM in Fig.A.4d). However, the PRC $\Delta\Phi(\theta)$ for dark pulses has a zero at a specific time of the day (see Fig.A.4c). That is, for each dark pulse of width τ, there exists a time of administration such that $\Delta\Phi = 0$! In fact, such a dark pulse has an entraining effect, reducing the population variance. We leave experimental and theoretical investigation of the counter-intuitive effects of such specially time dark pulses to future work.

Here, we show that even if we include such dark pulses with an entraining effect, the variance gained at the end of the day is still proportional to $(L/R)^2$ in the limit that L/R goes to zero. To simplify our derivation but retain the essence of what dark pulses do during the daytime, let's us consider dark pulses coming at three times: in the morning ($\theta = -\pi/2$), around noon ($\theta = -\omega\tau/2$ with small $\omega\tau$), and in the evening ($\theta = \pi/2$). Starting the day with variance σ_0^2, by the end of the day the variance becomes

$$\sigma^2 = \frac{\sigma_0^2 + (\Delta\Phi)^2_{\theta=-\pi/2}}{(1 + \left(\frac{d\Delta\Phi}{d\theta}\right)_{\theta=-\omega\tau/2})^2} + (\Delta\Phi)^2_{\theta=\pi/2} \tag{A.26}$$

$$\approx \frac{\sigma_0^2 + \left(\frac{L}{R}\sin\omega\tau\right)^2}{\left(1 + \frac{2L}{R}\sin\left(\frac{\omega\tau}{2}\right)\right)^2} + \left(\frac{L}{R}\sin\omega\tau\right)^2 \tag{A.27}$$

$$\sigma^2 \approx \sigma_0^2 + 2\left(\frac{L}{R}\sin\omega\tau\right)^2 + \mathcal{O}\left(\frac{L}{R}\right)^2. \tag{A.28}$$

Thus, the variance gained due to fluctuation, $\sigma^2 - \sigma_0^2 = \sigma^2_{clouds}$, is proportional to $(L/R)^2$. This simple derivation may not rigorously reflect the correct constant in front of $(L/R)^2$ term, but the full rigorous derivation, concerning the dark pulses coming randomly at random time during the day, should yield the same power law dependent on L/R. Fig.A.4d shows that averaging $\Delta\Phi^2$ over pulses administered at different times numerically (dashed line) results in the same power law as for single pulses and as seen in stochastic weather simulations.

A.6 Circle Map - Step Response Curve

In our main paper, we claim that the variance of the clock state across a population drops $\sigma^2 \Rightarrow \sigma^2/s^2$ at dusk where $s^2 - 1 \sim L/R$ as $L/R \Rightarrow 0$. Data from Langevin simulations was presented. Here we will derive this result using a simple geometric argument about circle maps.

We define $\phi = P_T(\theta)$ to be the phase on the night cycle that a clock evolves to, after time a time T, if the lights were suddenly turned off when the clock is at state θ on the day cycle. See Fig.A.3a,b. In principle, with complex relaxation dynamics between the limit cycles, $P_T(\theta)$ could show complex dependence on T. However, we work in a simplified model where the angular frequency of the clock is independent of the amplitude of oscillations. In this limit, T only causes an overall shift in $\phi = P_T(\theta)$; i.e., we can write $P_T(\theta) = P(\theta) + \omega T$ where ω is the angular frequency of the clock. In what follows, we will be interested in the derivative of $\partial_\theta P_T(\theta)$; hence we will work with $P(\theta)$ instead of $P_T(\theta)$.

This circle map, $\phi = P(\theta)$, is important since it determines whether two differing daytime clock states are brought closer or taken further at dusk and thus determines the rate of entrainment of a population to the external signal. Consider two organisms that have nearby but distinct clock states θ_0, $\theta_0 + \Delta\theta$ at dusk. After dusk, these two clocks will be mapped to $P(\theta_0)$ and $P(\theta_0 + \Delta\theta) \approx P(\theta_0) + \Delta\theta dP(\theta)d\theta|_{\theta=\theta_0}$ respectively. Thus, dusk changes the

difference between the clock states from $\Delta\theta$ to $\Delta\phi$ where,

$$\Delta\phi \approx \Delta\theta \left.\frac{dP(\theta)}{d\theta}\right|_{\theta=\theta_0} \tag{A.29}$$

By a similar argument, if the phase variance of clock states across a population is σ^2 before dusk, it will be reduced by,

$$\sigma^2 \xrightarrow{dusk} \sigma^2 \left(\left.\frac{dP(\theta)}{d\theta}\right|_{\theta=\theta_0}\right)^2 \tag{A.30}$$

This expression is valid in the regime where the population variance σ^2 is small enough to linearize the circle map $P(\theta)$. Similar considerations apply to the dawn transition between the night and day cycle as well. Both circle maps were recently experimentally characterized for *S. elongatus* in [11].

In our simple theoretical model where clock frequency does not change with amplitude (i.e. the radial coordinate), we can easily compute $P(\theta)$ from geometry. In Fig.A.3, we draw a diagram of the transition from a particle on the day cycle at the phase θ to the night cycle at the phase ϕ. By trigonometry, we write

$$\phi = P(\theta) = \arctan\left(\frac{L + R\sin\theta}{R\cos\theta}\right), \tag{A.31}$$

and derive

$$s^2 - 1 = \left(\frac{dP(\theta)}{d\theta}\right)^{-2} - 1 \tag{A.32}$$

$$= \frac{L(2L^3 + 7LR^2 - 3LR^2\cos(2\theta) + 4R(2L^2 + R^2)\sin\theta)}{2R^2(R + L\sin\theta)^2} \tag{A.33}$$

$$= 2\sin(\theta)\frac{L}{R} + \mathcal{O}\left(\frac{L}{R}\right)^2, \tag{A.34}$$

where θ corresponds to the angle on the day cycle at dusk, which is at $\pi/2$ in Fig.A.3a. This

equation implies that as the day and night limit cycle gets closer, the geometric focusing effect s converges to one. This asymptotic behavior is intuitive because if $L = 0$, meaning no transition, then the variance should remain the same ($s = 1$, so $\sigma^2 \to \sigma^2/1^2$ at the transition).

Remarkably, our geometric derivation of $s^2 - 1$ matches the variance drop $\sigma^2 \to \sigma^2/s^2$ seen in stochastic simulations of weather conditions; see Fig.A.3d. The variance gain during the day is the result of the fluctuation of sunlight, simulated as random dark pulses of random intervals, amplitude and time of delivery. Such variance is accumulated during the day and the drop over dusk time is measured (green Xs).

Fig.A.3e shows the variance drop seen in simulations with internal noise in Langevin simulations. While the cause of variance increase during the day is different (finite copy number effects), the variance drop at dusk agrees well with the geometric computation of s^2 and thus with the external noise simulations as well. In both cases, the simulations and geometric theory show that $s^2 - 1 \sim L/R$ as $L/R \Rightarrow 0$.

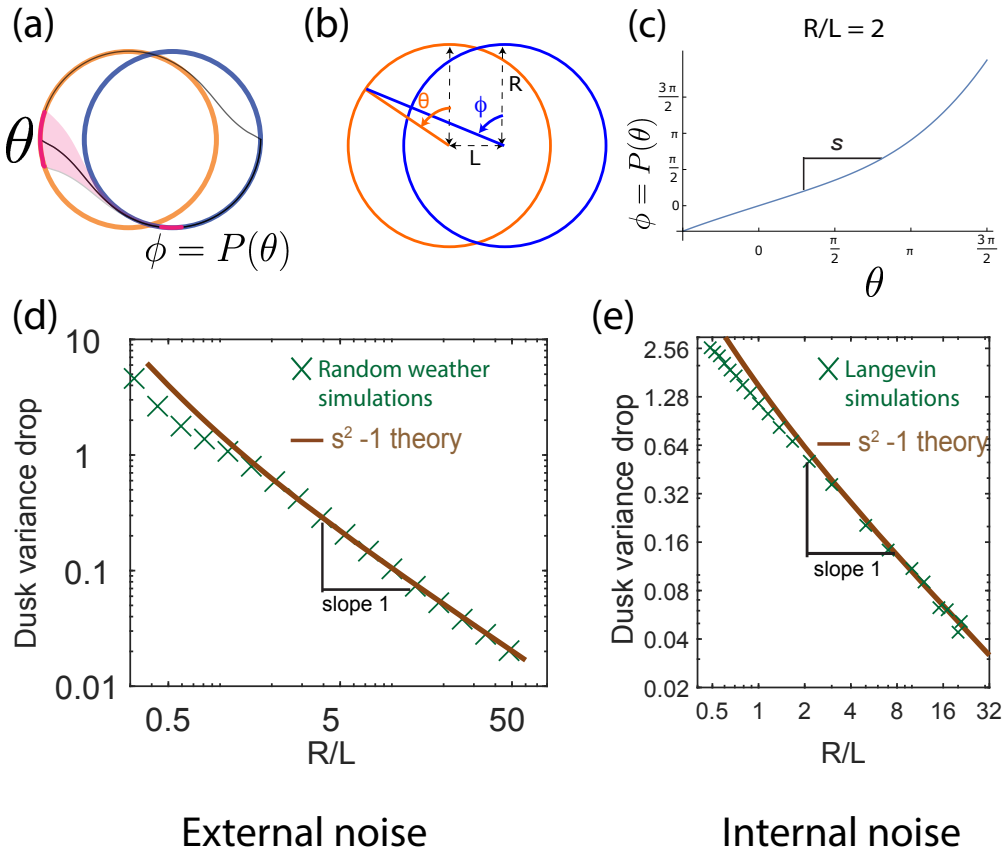

Figure A.3: The population variance of clock states is reduced by dusk and can be computed geometrically. (a) A population of clocks near state θ on the day cycle is mapped to the neighborhood of state ϕ on the night cycle by the dusk transition. We define $\phi = P(\theta)$ to be the map relating the clock state θ on the day cycle just before dusk to its eventual position ϕ on the night cycle after dusk (assumed greater than the relaxation time). (b) This map can be analytically computed for circles of size R with centers separated by length L. (c) For a given R/L = 2 , we obtain $P(\theta)$ shown here. Since $\theta = \pi/2$ corresponds to the dusk time of the entrained trajectory, the slope $s^{-1} = dP/d\theta$ at $\theta = \pi/2$ determines the change in population variance of clock states at dusk. (d,e) The variance drop s^2 at dusk, defined as $\sigma^2 \to \sigma^2/s^2$ at dusk, seen in both the external (averaging over weather) and internal noise (averaging over Langevin noise) simulations agree well with the geometrically computed $s(R/L)$, especially at large R/L. We find that $s^2 - 1 \sim L/R$ for large-R/L limit cycles.

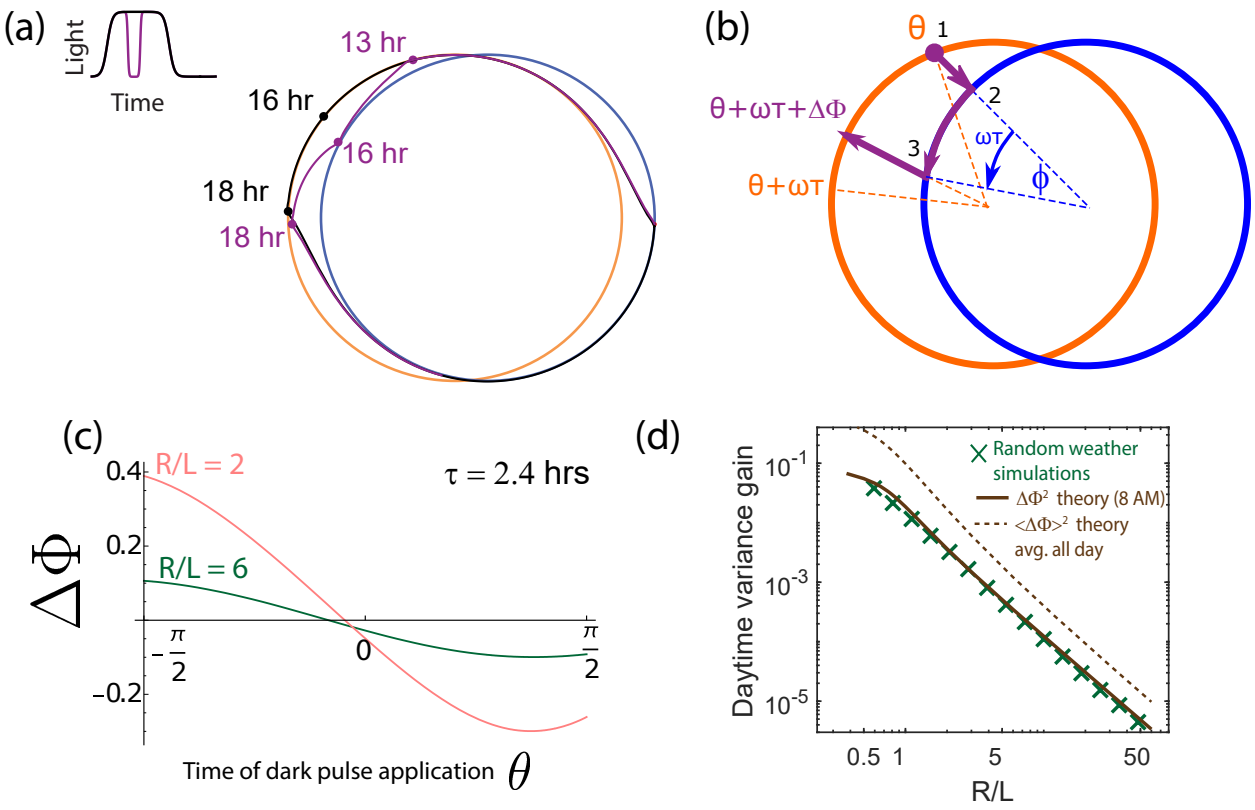

Figure A.4: Increase in population variance due to random weather conditions can be estimated from the phase shifts $\Delta\Phi$ due to dark pulses (i.e., the Phase Response Curve). (a) A single dark pulse administered during the day shifts the phase of a clock (purple) relative to a clock that experiences no such dark pulse (black). (b) We can compute the phase shift $\Delta\Phi$ due to such a dark pulse geometrically by computing the deviation in trajectory. Assuming a dark pulse of length τ, the clock evolves for a time τ according to the night cycle dynamics. At the end of such a pulse, we switch back to the day limit cycle and compute the resulting phase shift $\Delta\Phi$. (c) The resulting phase shift $\Delta\Phi$ due to a pulse of length $\tau = 2.4$ hrs, depends on the time θ when it is administered but is generally smaller for larger R/L. (d) We find that $\Delta\Phi^2$ for a specific $\tau = 2.4$ hrs dark pulse administered at the same time (8 AM) falls as $(L/R)^2$ for large-R/L limit cycles. This trend matches the variance gain σ^2_{clouds} seen in stochastic simulations that average over random weather conditions (pulses of different length, intensity and time of application). The broken brown curve shows a theoretical prediction for such an average $\langle\Delta\Phi^2\rangle$, obtained by sampling the curve shown in (c) at different points of application and differing intensity. Despite the presence of a variance-reducing zero around mid-day in (c), σ^2_{clouds} drops as $(L/R)^2$, much as $\Delta\Phi^2$ for any particular pulse. (Brown theory curves translated together using one fitting parameter.).

A.7 Langevin model of finite copy number fluctuations

Chemical reactions that occur in the bulk of a homogeneous solution can be described by a set of ordinary differential equations. However, within a single cell the copy number of molecule is limited and thus the reaction carries internal noise from the stochastic fluctuations. Gillespie showed that chemical reactions under finite copy number can be approximated by a Langevin dynamics using the following argument [48],

Consider an elementary reaction

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{C} + \mathrm{D} \tag{A.35}$$

with the forward rate constant k_+, during each infinitesimal time δt, the probability of the occurrence of this reaction follows a Poisson distribution whose mean and variance both equals to $R_+\delta t = k_+ \cdot N_A \cdot N_B \cdot \delta t$. Integration over a larger time step, the Poisson distribution can be approximated into a Gaussian form, resulting in Langevin dynamics,

$$\mathrm{d}N_A = -k_+ \cdot N_A \cdot N_B \cdot \mathrm{d}t + \sqrt{R_+}\mathrm{d}W \tag{A.36}$$

where W is a standard Wiener process of mean 0 and autocorrelation function $\langle W(t_1)W(t_2)\rangle = \delta(t_1 - t_2)$.

To describe a chemical reaction network, the Langevin equation for each species consists of contributions to the noise from each reaction where the species is involved. Now consider adding another reaction

$$\mathrm{C} + \mathrm{D} \longrightarrow \mathrm{A} + \mathrm{B} \tag{A.37}$$

with the rate constant k_-, then the Langevin equation for species A becomes,

$$\mathrm{d}N_A = -k_+ \cdot N_A \cdot N_B \cdot \mathrm{d}t + k_- \cdot N_C \cdot N_D \cdot \mathrm{d}t + \sqrt{R_+}\mathrm{d}W_1 + \sqrt{R_-}\mathrm{d}W_2 \tag{A.38}$$

where $R_+ = k_+ \cdot N_A \cdot N_B$ and $R_- = k_- \cdot N_C \cdot N_D$ respectively denote the number rates of the forward and the backward reaction; $\mathrm{d}W_1$ and $\mathrm{d}W_2$ are identical independent standard Wiener processes.

To fully determine the effect of the noise using the Langevin dynamics for a chemical reaction network, one needs to consider all of the reactions corresponding to the species of interest; the noise term usually becomes time-dependent and multiplicative. To simplify the description of internal noise in our phenomenological model of limit cycle/ point attractor, we take a first order approximation that the diffusion coefficient in the reaction coordinate space is homogeneous in both space and time. (See similar treatments of another biological system in [43]. In contrast, our explicit KaiABC simulations, as well as numerical simulations on the other types of bio-oscillators, presented later, do not make this simplifying assumption of homogeneous diffusion.) This allows us to write a 2-dimension phenomenological stochastic differential equation

$$\mathrm{d}\vec{z} = f(\vec{z}, t) \cdot \mathrm{d}t + \sqrt{2D} \cdot \mathrm{d}\vec{W} \tag{A.39}$$

where the $f(\vec{z}, t)$ denotes the deterministic dynamics driven by day-night cycles and the diffusion constant D is assumed to be inversely proportional to the total number of Kai-C molecules within the cell. For limit cycles of radius R, we set $D \sim R^2 \epsilon_{int}^2$. Then, ϵ_{int}^2 is the diffusion constant for the *phase* of the oscillator. We hold ϵ_{int}^2 fixed while changing R to make a fair comparison across systems of different size.

Population variance

For the cell to carry out a reliable computation, the population variance from the internal noise needs to be reduced. Such noise reduction comes from the dynamics of the attractor. In the limit cycle attractor mechanism, the internal noise reduction is performed only along the radial axis but not along the flat attractor direction.

In contrast, the point attractor mechanism is able to limit population variance due to

internal noise in all directions due to the effective 'curvature' of the dynamics. Here we analytically estimate the steady-state population variance for a point attractor mechanism. The population variance is together determined by the diffusive term $\sqrt{2D} \cdot \mathrm{d}\vec{W}$, and the noise reduction effect from the restoring force of the point attractor's harmonic well. During each infinitesimal time δt, the internal noise increase the variance by

$$\sigma^2(t + \delta t) = \sigma^2(t) + 2D\delta t. \tag{A.40}$$

In contrast, the overdamped deterministic motion within a harmonic well provides a focusing effect that reduces the variance exponentially with time. To quantify this focusing effect, consider a 1-d overdamped dynamics of a particle within a harmonic energy well of $V(r) = k \cdot r^2$. The solution to the equation of motion is $r(t) = r_0 \cdot e^{-2kt}$, with initial position $r(0) = r_0$. Consider an ensemble of points with a mean initial position μ_0 and a initial variance of σ_0^2, one can solve the dynamics of the mean as

$$\mu(t) = \mu_0 \cdot e^{-2kt} \tag{A.41}$$

and the dynamics of the variance as

$$\sigma^2(t) = \sigma_0^2 \cdot e^{-4kt} \tag{A.42}$$

Thus, per infinitesimal time δt, the geometric focusing effect of the energy well of the point attractor reduces the population variance by

$$\sigma^2(t + \delta t) = \sigma^2(t)/g \tag{A.43}$$

where $g = e^{4k\delta t}$.

Under the competition between the spreading effect from the internal noise and the

geometrical focusing effect from the deterministic dynamics, the population variance reaches a steady value solved by

$$\sigma_{st}^2 = \frac{\sigma_{st}^2 + 2D\delta t}{g} = \frac{\sigma_{st}^2 + 2D\delta t}{e^{4k\delta t}} \tag{A.44}$$

and by taking the limit of δt goes to 0, we have $\sigma_{st}^2 = D/2k$.

APPENDIX B

STOCHASTIC RESONANCE IN LIVING CELLS

B.1 TNF-α Stimulation Using Microfluidic Cell Culture

We use the signal generator chip described previously [79]. PDMS chambers coated with fibronectin (FC010-10MG) and cells were seeded at constant density 20,000 cells/cm^2. Cells were cultured for 5 hours before stimulation; they were taken at 100% confluence and incubated with .25% Trypsin-EDTA for 5 mins (25200-056) prior to loading. Standard culture conditions of % CO2 and 37°C were maintained using an incubation chamber for imaging and cell culture. Mouse TNF-α (PMC3014_3671982503) was diluted in Fluorobrite DMEM media (A1896701) with 2x glutamax (35050061), pen/strep (15140-122) and FBS (16140071) in vials pressured with 5% CO2 at 5psi and kept on ice. Microbore tubing (PEEK, Vici) was used to connect the TNF-α supply to the chip. For continuous pumping input, the on-chip peristaltic pump was operated at a sampling rate of .5 seconds of a combination (3,4, or 5 concurrently open) of the following inputs: [.2 .1 .05 .025 .0125 .00625 .003125 .001625] ng/mL TNF-α.

B.1.1 Image Acquisition and Data Processing

We use an automated Nikon eclipse ti2 microscope and capture fluorescence images (red and green channels for p65 and H2B reporters) at 20× magnification via a Hamamatsu ORCA-Flash4.0 V3 Camera (C13440) every 3-5 min for 1-10 hr. Microfluidic device is mounted on the microscope. Custom Matlab scripts were used for image processing (available on request). NF-κB activation was quantified as sum of nuclear fluorescence intensity. For peak analysis, data were smoothed followed by peak detection using a combination of integrated area, first derivative and peak height.

B.1.2 NF-κB Reporter Cell Line

Creation of mouse (3T3) fibroblasts displaying near-endogenous p65 levels was previously described in [69]. Knockout p65-/- mouse 3T3 fibroblasts were engineered to display p65-DsRed under the 1.5 kb p65 promoter (4). A minimum fluorescence clone was selected to achieve near-endogenous expression level to represent NF-κB wild-type dynamics. Addition of a ubiquitin-promoter driven H2B-GFP cassette provides a nuclear marker to facilitate automated image processing.

B.1.3 RNA Sequencing

Cells were removed from microfluidic chambers using .25% Trypsin-EDTA for 5mins (25200-056) to free the cells from microfluidic surface, then flushed with water to remove from device. Cells were placed in lysis buffer containing 0.2% (vol/vol) Triton X-100, RNase inhibitor, oligoDT primer and dntps. Protocol for SMART seq was followed thereafter. Samples were pooled then sequenced on NovaSeq platform.

B.2 Mathematical modeling of NF-κB pathway

A comprehensive model for the NF-κB system with 26 variables was first elaborated by Hoffman in 2002 [81]. Krishna et al. reduced the model to only five non-linear coupling equations [16, 17] while still retaining the essential dynamics of the pathway:

$$\frac{dN_n}{dt} = \frac{K_I}{K_I + I} k_{Nin} \left(N_{tot} - N_n\right) - k_{lin} I \frac{N_n}{K_N + N_n}$$
$$\frac{dI_m}{dt} = k_t N_n^2 - \gamma_m I_m$$
$$\frac{dI}{dt} = k_{tl} I_m - \alpha [IKK]_a \left(N_{tot} - N_n\right) \frac{I}{K_I + I}$$
$$\frac{d[IKK]_a}{dt} = k_a [TNF] \left([IKK]_{tot} - [IKK]_a - [IKK]_i\right) - k_i [IKK]_a$$
$$\frac{d[IKK]_i}{dt} = k_i [IKK]_a - k_p [IKK]_i \frac{k_{A20}}{k_{A20} + [A20][TNF]}$$

where N_n is the nuclear NF-κB concentration, I_m is the IκB mRNA level, I is the concentration of cytoplasmic IκB protein, $[IKK]_a$ is the level of active IκB kinase, and $[IKK]_i$ is the level of inactive IκB kinase. This simple model reproduces spiky NF-κB oscillations seen in experiments and entrainment of such oscillations by time-varying TNF signals. The last two equations above are a simplified representation of signal transduction from external TNF levels to IKK levels [16, 17]. Note the existence of both positive and negative regulation of IKK by TNF, characteristic of an incoherent feedforward loop [18]: TNF increases the amount of active IKK but also decreases active IKK indirectly through inactive IKK.

In reality, signal transduction between external TNF and IKK is known to be more complex, involving receptor dynamics and numerous intermediate complexes. These details include incoherent feedforward loops of different timescales in addition to the one modelled here [16]. Such features may be revealed by the response to high frequency TNF signals used in this study but cannot be predicted by the last two equations in this model since they are based on experiments involving slowly changing TNF (timescale of hours).

In this paper we modify the last two equations to capture faster signaling processes that might be visible to fast TNF signals (timescale of seconds). We continue to model the pathway between TNF and IKK as having general positive and negative regulation but with

an unspecified timescale τ_0. This approach gives a simple coarse-grained model of signaling without making assumptions about the exact molecular identity of the molecules involved in these fast processes:

$$\begin{aligned}
\frac{dN_n}{dt} &= \frac{K_I}{K_I + I} k_{Nin} \left(N_{tot} - N_n\right) - k_{lin} I \frac{N_n}{K_N + N_n} \\
\frac{dI_m}{dt} &= k_t N_n^2 - \gamma_m I_m \\
\frac{dI}{dt} &= k_{tl} I_m - \alpha [IKK]_a \left(N_{tot} - N_n\right) \frac{I}{K_I + I} \\
\tau_0 \frac{dX}{dt} &= -\alpha_x (X - TNF) \\
\tau_0 \frac{dY}{dt} &= -\beta_y \left(Y - c_0\right) - \beta_x X \left(\frac{Y^n}{Y_0^n + Y^n}\right) + \beta_{TNF} TNF \\
\frac{d[IKK]_a}{dt} &= -\gamma_{IKK} [IKK]_a + \gamma_y \left(Y - c_0\right)
\end{aligned}$$

Here TNF is the input signal of interest, X is promoted by TNF, and Y is promoted by TNF but suppressed by X. The dynamics downstream of IKK (i.e., NF-κB and IκB) are exactly as in the model of Krishna et al. [17]. The value of the parameters we use in these equations are summarized in Table B.1.

We have two sources of noise in our simulation, one from TNF, and the other from the stochasticity inside the NF-κB network. Additive noise in the TNF signal is simulated from two random distributions. The amplitude of the noise is drawn from a uniform distribution $[-\sigma, \sigma]$, and the time difference between the change of amplitudes of noise is drawn from an exponential distribution with mean equals to τ_{noise} . Meanwhile, the stochasticity inside the NF-κB networks (only the first three equations) is simulated from the Langevin noise $\eta(0, \epsilon)$ where $\epsilon = 0.01\tau_0^{-1/2}$.

Parameters for simplified NF-κB model [16, 17]			
Parameter	Value	Parameter	Value
k_{Nin}	5.4 min^{-1}	α_x	10
k_{lin}	0.018 min^{-1}	β_y	60
k_l	1.03μM^{-1} min^{-1}	β_x	60
k_{tl}	0.24 min^{-1}	β_{TNF}	60.6
K_I	0.035μM	n	10
K_N	0.029μM	y_0	0.001
γ_m	0.017 min^{-1}	γ_{IKK}	1
α	1.05μM^{-1} min^{-1}	γ_y	30
N_{tot}	1μM	τ_0	1 min
		τ_{noise}	0.0667τ_0

Table B.1: The first two columns give parameters of the simplified NF-kB system (k_{lin} to N_{tot}), taken directly from [16, 17]. The last columns give parameters for the rectified adaptation model, following design principles laid out in [18] for Incoherent Feed-Forward Loop-based adaptive circuits.

B.2.1 Rectified adaptation

The existence of both positive and negative interactions (i.e., an Incoherent Feed Forward Loop), leads to adaptation [18]. Adaptation in a chemical circuit refers to a transient response to a step change in the input after which the circuit output eventually settles down to its initial resting value. In our model, the resting value of Y is c_0. For non-zero $c_0 = 0.1$, we find that Y responds adaptively to both step ups and step downs in the input (TNF).

However, if $c_0 = 0, \text{Y}$ only responds to step ups in TNF and not step downs, a phenomenon we term 'rectified adaptation'. Such a circuit will perceive noise as an ever-increasing signal, and trigger an activation threshold much larger than the standard deviation of the noise. Thus, we find that a simple limit of the commonly found adaptation motif in biology explains stochastic resonance seen in NF-κB activation.

B.2.2 Details of simulations in each figure

In Figure 3.1d-f, NF-κB simulated using the rectified adaptative circuit (i.e., $c_0 = 0$) activates from noisy signal of low moving average. For the constant signal, we set the TNF level to be 0.1. For the noisy signal, we set the average value of TNF to be 0.05 and σ to be 0.04.

In Figure 3.5c and d, NF-κB only responds to noisy signal if the adaptation is rectified (i.e., if $c_0 = 0$) but not with conventional adaptation ($c_0 = 0.1$). Similar to Figure 3.1d-f, for constant signal, we set the TNF level to be 0.05. For the noisy signal, we set the average value of TNF to be 0.05 and σ to be 0.04.

In Figure 3.6a, NF-κB in the rectified adaptive circuit ($c_0 = 0$) shows decreasing responses as the period of the input signal increases. In creating Figure 3.6a, we use noiseless sinusoidal signals of periods $0.2667\tau_0, 1.5085\tau_0, 8.5333\tau_0$. For all periods, we set the average value of TNF to be 0.05 and σ to be 0.04.

In Figure 3.6b, NF-κB in the rectified adaptive circuit shows decreasing fraction of cells activated as the period of the input signal increases. Figure 3.6e show how the timescales of the adaptation impacts the activation of NF-κB. Here, the average TNF level and the amplitude of the sine wave remain at 0.05 and 0.04, respectively. The period is varied from $0.1\tau_0$ to $100\tau_0$.

APPENDIX C

TEMPORAL DYNAMICS OF OPTIMIZATION ALGORITHMS

C.1 Text Dataset

Examples of student and teacher texts for both 'UCLA' and 'NEWS' domains			
'UCLA' Domain		'NEWS' Domain	
Student	Teacher	Student	Teacher
..ucofsryhnyyglgp..	..wzuhhuclackshvc..	..clawclucofinesn..	..iftxnnewsrltlvv..
..rgucwslprlaofyw..	..jgckcuclakgixce..	..jdzlqlazrogynps..	..llxjfenewsontzs..
..ukcvoroprwsajbh..	..yoahpprofemyoew..	..ucwspgzmfwjhtyr..	..sngzcutmuprofxr..

Table C.1: Examples of student and teacher texts for both 'UCLA' and 'NEWS' domains with the probability of specialist and generalist features equal to 0.9 and 0.45, respectively. The texts here are the 15-letter samples of the full 400-letter texts in simulations of Fig. 4.2.

We carefully design the text dataset to have only three features words: 'UCLA', 'NEWS', and 'PROF'. Chapter 4 explained that we simply put these features on top of randomly generated texts, but doing so actually create unexpected features. For example, the word 'UCLA' can be detected by a smaller part 'UC' if the chance that 'UC' appears in the random texts is low. We do not want these 2-letter features to exist because we want to study the case with only 3 features of 4-letter words. To get rid of these extra features, we artificially add in the 2-letter and 3-letter features to the random texts: 'UC', 'CL', 'LA', 'UCL', and 'CLA'. We repeated the process for the words: 'NEWS' and 'PROF'. As a result, the network is forced to learn the full 4-letter features 'UCLA'. Similarly, we also add 2-letter and 3-letter words that are parts of 'NEWS' and 'PROF'.

C.2 Finding the intermediate timescale

To estimate the best time for switching between two different data domains, we need to compute the response time, or the amount of epochs the neural network takes to forget an old feature and learn a new one. If the switching time is much smaller than this response time, then the parameters of the model will likely stay at the same minimum, ignoring the new domain of data. If the switching time is much larger than the response time, then the training time is unnecessarily spent. The switching time should be no higher than one order of magnitude from the response time.

We measure the response time by computing how many epochs the model take to forgot an old feature and learn a new one. In Fig. C.1, the model in each of the 49 simulations is initialized with 'NEWS' feature. They are then trained to a dataset from a different domain with 'UCLA' and 'PROF' features. To minimize the loss function, the model tend to overwrite the old feature with a new one. Most 49 simulations in Fig. C.1 take at most 2-3 epochs to reach a plateau in the accuracy. We choose the switching time to be around 20 epochs (or period of 40 epochs), which allow the network to pick up the generalist feature quickly (see Fig. 4.2). The period could have be chosen lower, but the model has shown to pick up the generalist feature almost 100 percents in comparison to one third from the training with mix datasets in Fig. 4.2a.

This response time is sensitive toward many training parameters. For example, doubling the learning rate likely cuts the response time in half. Also, increasing the number of data or decreasing the batch size will lengthen the response time. Surprisingly, even though there are many different local minimum in the loss landscape, this approximation of the response time generalizes well to at least 49 simulations from random initializations.

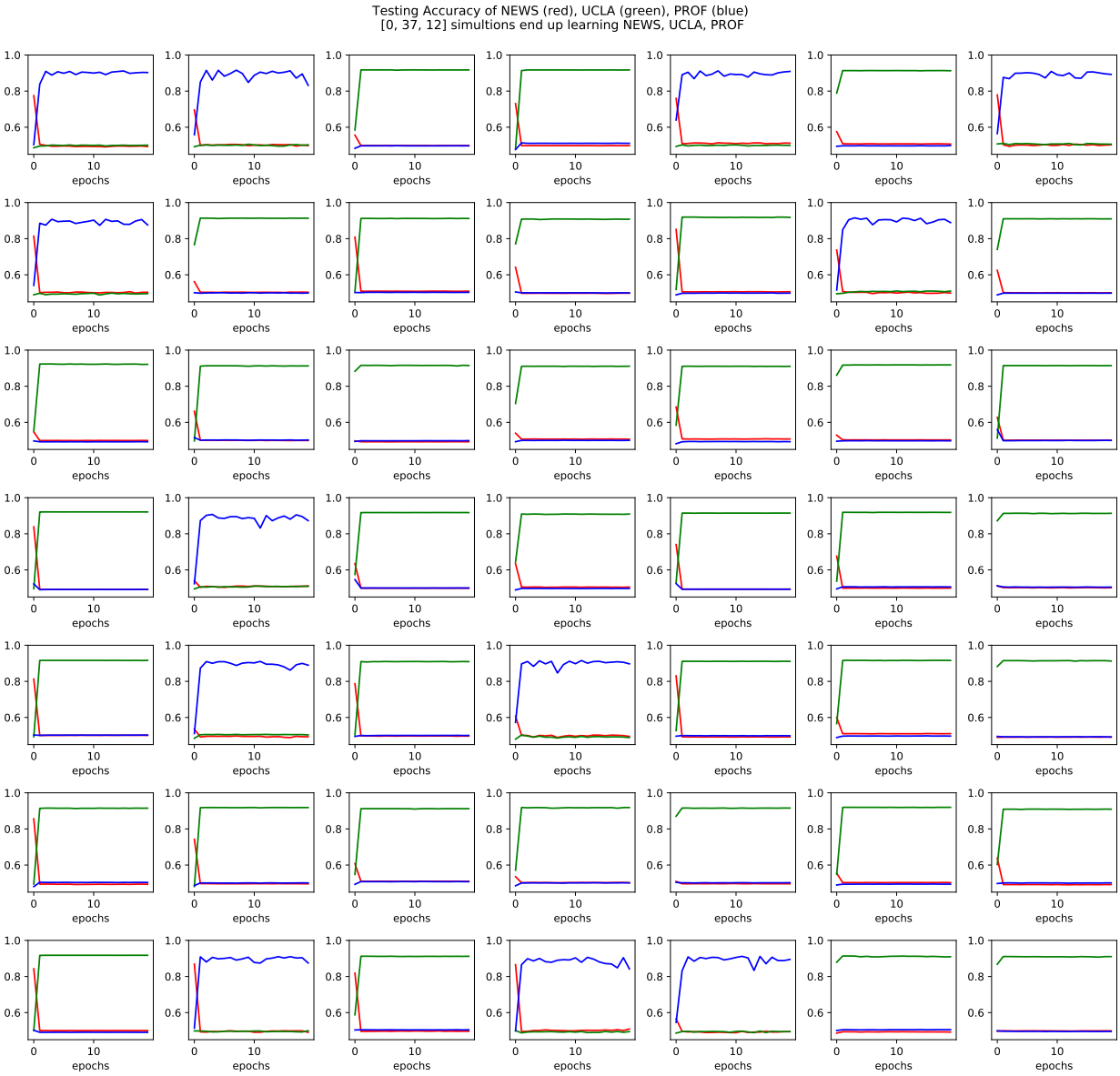

Figure C.1: Each small block show the plot of accuracies over time for 49 different trainings. These are 3 kinds of accuracies toward dataset with features 'NEWS' (red), 'UCLA' (green), and 'PROF' (blue), respectively. Initially, the models were initialized with parameters of high accuracy toward the 'NEWS' feature. Afterwards, the models were trained on dataset with only 'UCLA' and 'PROF' features. As a result, the model needs to forget about the old feature 'NEWS' and learn either of the two new features. Approximately, the time to learn a new feature (or the response time) is around 2 epochs.

C.3 Optimization algorithms

In the simulations of Fig. 4.2, the neural network consists of two components: a 1D convolutional network and a softmax layer. The input texts are converted into a hot vector of size $26L$ where 26 is the number of English alphabets and L is the length of a text. The values of the filter size and the stride for the convolutional network are tuned so that the network convolutes 5 letters at a time. The softmax layer allows the output of the multiple layers to be merged.

Parameters for the training of the neural network			
Parameter	Value	Parameter	Value
length of texts	400	number of data	8000
L2 - weight decay	0.0003	batch size	4
learning rate	0.01	opt. algorithm	torch.optim.Adam
filter size	5 alphabets	number of filters	5
stride	26	padding	26 (on both sides)

Table C.2: Here are the parameters for every simulation in Fig. 4.2.

C.4 Measuring the fractions of learned features

Training a neural network is a heuristic process because the batches of data are shuffled and come in a random order. Thus, we need to repeat the simulations multiple times and find the fraction of models learning the different features. There are multiple ways to extract features out of neural networks. We choose to create three testing datasets, each corresponds to the features 'UCLA', 'NEWS', and 'PROF'. The model with highest testing accuracy on the 'UCLA' testing dataset is interpreting as learning the 'UCLA' feature. The same goes for 'NEWS' and 'PROF'. In the 'NEWS' dataset, the student texts are randomly generated English alphabets of length 400. The teacher texts has the word 'NEWS' put in a random position of the random texts.

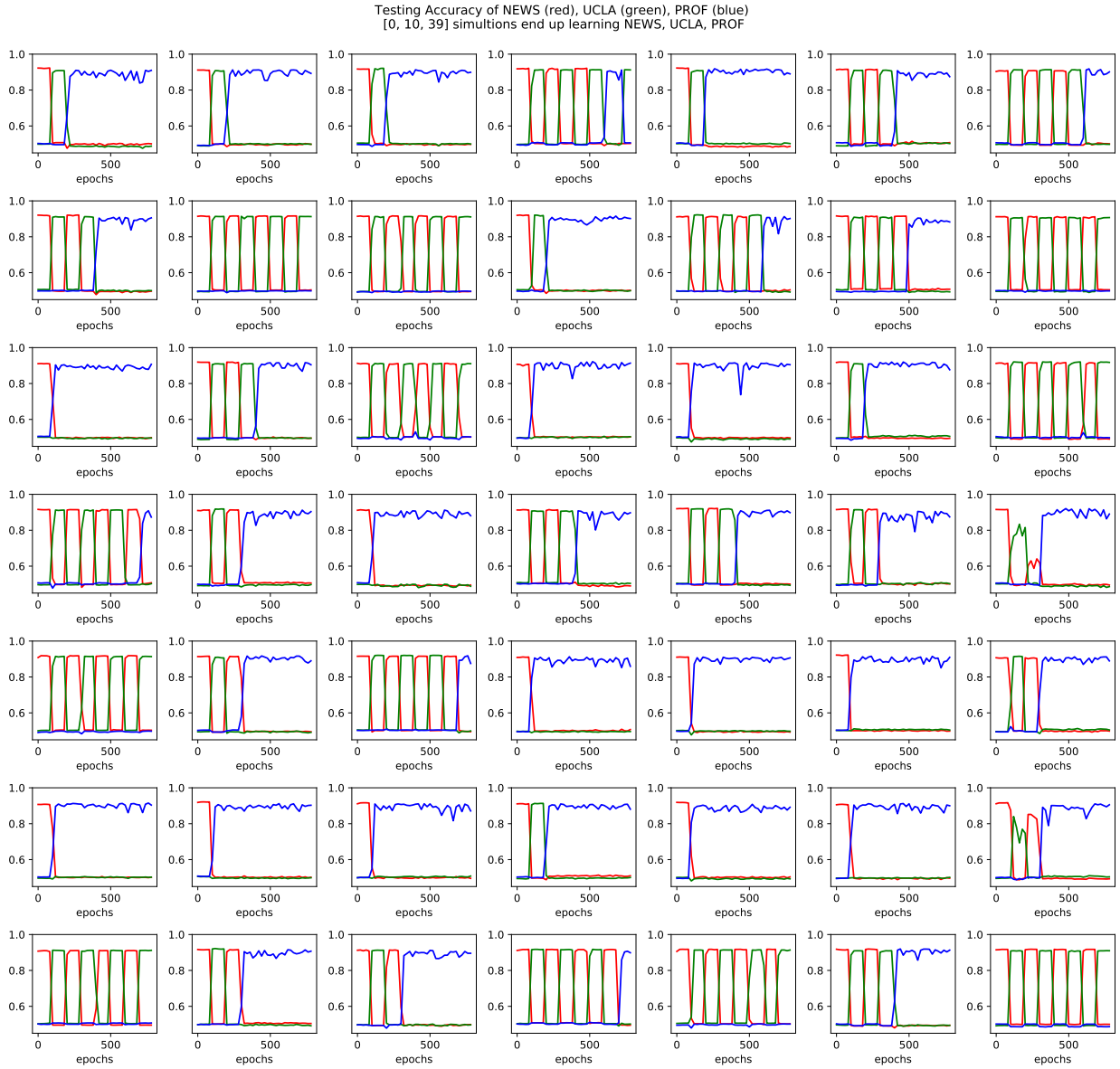

Figure C.2: Each small block show the plot of accuracies over time for 49 different trainings. These are 3 kinds of accuracies toward dataset with features 'NEWS' (red), 'UCLA' (green), and 'PROF' (blue), respectively. Each model is randomly initialized. The training data is switching between 'UCLA' domain and the 'NEWS' domain every 100 epochs. The fraction of learned features of these 49 blocks are summarized in Fig. 4.2(c).

REFERENCES

[1] Jean-Christophe Leloup, Didier Gonze, and Albert Goldbeter. Limit cycle models for circadian rhythms based on transcriptional regulation in drosophila and neurospora. *J. Biol. Rhythms*, 14(6):433–448, 1999.

[2] Christoph Schmal, Jean-Christophe Leloup, and Didier Gonze. Modeling and simulating the arabidopsis thaliana circadian clock using XPP-AUTO. *Methods Mol. Biol.*, 1158:337–358, 2014.

[3] J C W Locke, A J Millar, and M S Turner. Modelling genetic networks with noisy and varied experimental data: the circadian clock in arabidopsis thaliana. *J. Theor. Biol.*, 234(3):383–393, June 2005.

[4] Jean-Christophe Leloup and Albert Goldbeter. Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl. Acad. Sci. U. S. A.*, 100(12):7051–7056, June 2003.

[5] A Goldbeter. A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. *Proceedings of the National Academy of Sciences*, 88(20):9107–9111, October 1991.

[6] B C Goodwin. Oscillatory behavior in enzymatic control processes. *Adv. Enzyme Regul.*, 3:425–438, 1965.

[7] Didier Gonze and Wassim Abou-Jaoudé. The goodwin model: behind the hill function. *PLoS One*, 8(8):e69573, August 2013.

[8] Dilip Kondepudi and Ilya Prigogine. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*. John Wiley & Sons, December 2014.

[9] M B Elowitz and S Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, January 2000.

[10] Olguţa Buşe, Alexey Kuznetsov, and Rodrigo A Pérez. EXISTENCE OF LIMIT CYCLES IN THE REPRESSILATOR EQUATIONS. *Int. J. Bifurcat. Chaos*, 19(12):4097–4106, December 2009.

[11] Eugene Leypunskiy, Jenny Lin, Haneul Yoo, Unjin Lee, Aaron R Dinner, and Michael J Rust. The cyanobacterial circadian clock follows midday in vivo and in vitro. *Elife*, 6, July 2017.

[12] Arthur T Winfree. *The Geometry of Biological Time*. Springer Science & Business Media, June 2001.

[13] D S Saunders. *Insect Clocks, Third Edition*. Elsevier, October 2002.

[14] Lianhong Gu, Jose D Fuentes, Michael Garstang, Julio Tota da Silva, Ryan Heitz, Jeff Sigler, and Herman H Shugart. Cloud modulation of surface solar irradiance at a pasture site in southern brazil. *Agric. For. Meteorol.*, 106(2):117–129, January 2001.

[15] Michael J Rust, Joseph S Markson, William S Lane, Daniel S Fisher, and Erin K O'Shea. Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science*, 318(5851):809–812, November 2007.

[16] Sandeep Krishna, Mogens H Jensen, Kim Sneppen, and Leo P Kadanoff. Minimal model of spiky oscillations in nf-kb signaling. 0.

[17] Mathias Heltberg, Ryan A Kellogg, Sandeep Krishna, Savaş Tay, and Mogens H Jensen. Noise induces hopping between nf-kb entrainment modes. *Cell systems*, 3(6):532–539.e3, 12 2016.

[18] S Mangan and U Alon. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):11980–5, 10 2003.

[19] Steven H. Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC Press, 2019.

[20] Francisco A. Rodrigues, Thomas K.D.M. Peron, Peng Ji, and Jürgen Kurths. The Kuramoto model in complex networks. *Physics Reports*, 610:1–98, 2016.

[21] H. C. Berg and E. M. Purcell. Physics of chemoreception. *Biophysical Journal*, 20(2):193–219, 1977.

[22] J. T. Stuart. On the non-linear mechanics of wave disturbances in stable and unstable parallel flows Part 1. The basic behaviour in plane Poiseuille flow. *Journal of Fluid Mechanics*, 9(3):353–370, 1960.

[23] J Watson. On the non-linear mechanics of wave disturbances. 1960.

[24] W. Pittayakanchit, Z. Lu, J. Chew, M.J. Rust, and A. Murugan. Biophysical clocks face a trade-off between internal and external noise resistance. *eLife*, 7, 2018.

[25] P. Patel, W. Pittayakanchit, M. Son, A. Murugan, and S. Tay. Environmental noise enables sensitive detection and transcriptional decoding of cytokine inputs . Manuscript submitted for publication, 2021.

[26] Clive G Bowsher and Peter S Swain. Environmental sensing, information transfer, and cellular decision-making. *Curr. Opin. Biotechnol.*, 28:149–155, August 2014.

[27] Amir Mitchell, Ping Wei, and Wendell A Lim. Oscillatory stress stimulation uncovers an achilles' heel of the yeast MAPK signaling network. *Science*, 350(6266):1379–1383, December 2015.

[28] Yuhai Tu, Thomas S Shimizu, and Howard C Berg. Modeling the chemotactic response of escherichia coli to time-varying stimuli. *Proc. Natl. Acad. Sci. U. S. A.*, 105(39):14855–14860, September 2008.

[29] Huaqing Cai, Mariko Katoh-Kurasawa, Tetsuya Muramoto, Balaji Santhanam, Yu Long, Lei Li, Masahiro Ueda, Pablo A Iglesias, Gad Shaulsky, and Peter N Devreotes. Nucleocytoplasmic shuttling of a GATA transcription factor functions as a development timer. *Science*, 343(6177):1249531, March 2014.

[30] Eric D Siggia and Massimo Vergassola. Decisions on the fly in cellular sensory systems. *Proceedings of the National Academy of Sciences*, 110(39):E3704–12, September 2013.

[31] Thierry Mora and Ned S Wingreen. Limits of sensing temporal concentration changes by single cells. *Phys. Rev. Lett.*, 104(24):248101, June 2010.

[32] C Troein, Jcw Locke, M S Turner, and A J Millar. Weather and seasons together demand complex biological clocks. *Curr. Biol.*, 19(22):1961–1964, January 2009.

[33] Donald A Bryant. The beauty in small things revealed. *Proc. Natl. Acad. Sci.*, 100(17):9647, August 2003.

[34] Andrian Gutu and Erin K O'Shea. Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. *Mol. Cell*, 50(2):288–294, April 2013.

[35] Julia Holtzendorff, Frédéric Partensky, Daniella Mella, Jean-François Lennon, Wolfgang R Hess, and Laurence Garczarek. Genome streamlining results in loss of robustness of the circadian clock in the marine cyanobacterium prochlorococcus marinus. *J. Biol. Rhythms*, 23(3):187, June 2008.

[36] Alexis Dufresne, Marcel Salanoubat, Frédéric Partensky, François Artiguenave, Ilka M Axmann, Valérie Barbe, Simone Duprat, Michael Y Galperin, Eugene V Koonin, Florence Le Gall, Kira S Makarova, Martin Ostrowski, Sophie Oztas, Catherine Robert, Igor B Rogozin, David J Scanlan, Nicole Tandeau de Marsac, Jean Weissenbach, Patrick Wincker, Yuri I Wolf, and Wolfgang R Hess. Genome sequence of the cyanobacterium prochlorococcus marinus SS120, a nearly minimal oxyphototrophic genome. *Proc. Natl. Acad. Sci.*, 100(17):10020, August 2003.

[37] Yohko Kitayama, Hideo Iwasaki, Taeko Nishiwaki, and Takao Kondo. KaiB functions as an attenuator of KaiC phosphorylation in the cyanobacterial circadian clock system. *EMBO J.*, 22(9):2127–2134, May 2003.

[38] Ioannis Lestas, Glenn Vinnicombe, and Johan Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467(7312):174–178, September 2010.

[39] Lev S Tsimring. Noise in biology. *Rep. Prog. Phys.*, 77(2):026601, January 2014.

[40] Shu-Wen Teng, Shankar Mukherji, Jeffrey R Moffitt, Sophie de Buyl, and Erin K O'Shea. Robust circadian oscillations in growing cyanobacteria require transcriptional feedback. *Science*, 340(6133):737–740, May 2013.

[41] Mark A Woelfle, Yan Ouyang, Kittiporn Phanvijhitsiri, and Carl Hirschie Johnson. The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. *Curr. Biol.*, 14(16):1481–1486, August 2004.

[42] Jennifer A Evans and Alec J Davidson. Health consequences of circadian disruption in humans and animal models. *Prog. Mol. Biol. Transl. Sci.*, 119:283–323, 2013.

[43] Davit A Potoyan and Peter G Wolynes. On the dephasing of genetic oscillators. *Proceedings of the National Academy of Sciences*, 111(6):2391–2396, February 2014.

[44] Laurent Potvin-Trottier, Nathan D Lord, Glenn Vinnicombe, and Johan Paulsson. Synchronous long-term oscillations in a synthetic gene circuit. *Nature*, 538(7626):514–517, October 2016.

[45] S Laughlin. A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch. C*, 36(9-10):910–912, September 1981.

[46] Mirela Domijan and David A Rand. Balance equations can buffer noisy and sustained environmental perturbations of circadian clocks. *Interface Focus*, 1(1):177–186, February 2011.

[47] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. U. S. A.*, 99(20):12795–12800, October 2002.

[48] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58, 2007.

[49] David Zwicker, David K Lubensky, and Pieter Rein ten Wolde. Robust circadian clocks from coupled protein-modification and transcription–translation cycles. *Proc. Natl. Acad. Sci. USA*, 107(52):22540, December 2010.

[50] Gopal K Pattanayak, Connie Phong, and Michael J Rust. Rhythms in energy storage control the ability of the cyanobacterial circadian clock to reset. *Curr. Biol.*, 24(16):1934–1938, August 2014.

[51] Kathleen E Moore, David R Fitzjarrald, Ricardo K Sakai, Michael L Goulden, J William Munger, and Steven C Wofsy. Seasonal variation in radiative and turbulent exchange at a deciduous forest in central massachusetts. *J. Appl. Meteorol.*, 35(1):122–134, January 1996.

[52] Didier Gonze, José Halloy, and Albert Goldbeter. Robustness of circadian rhythms with respect to molecular noise. *Proc. Natl. Acad. Sci. U. S. A.*, 99(2):673–678, January 2002.

[53] Michele Monti, David K Lubensky, and Pieter Rein ten Wolde. Optimal entrainment of circadian clocks in the presence of noise. June 2017.

[54] Irina Mihalcescu, Weihong Hsing, and Stanislas Leibler. Resilient circadian oscillator revealed in individual cyanobacteria. *Nature*, 430(6995):81–85, July 2004.

[55] Yoram Burak and Ila R Fiete. Fundamental limits on persistent activity in networks of noisy neurons. *Proc. Natl. Acad. Sci. U. S. A.*, 109(43):17645–17650, October 2012.

[56] H Sebastian Seung. How the brain keeps the eyes still. *Proceedings of the National Academy of Sciences*, 93(23):13339–13344, 1996.

[57] Raymond Cheong, Alex Rhee, Chiaochun Joanne Wang, Ilya Nemenman, and Andre Levchenko. Information transduction capacity of noisy biochemical signaling networks. *Science*, 334(6054):354–358, October 2011.

[58] Siting Gan and Erin K O'Shea. An unstable singularity underlies stochastic phasing of the circadian clock in individual cyanobacterial cells. *Mol. Cell*, 67(4):659–672.e12, August 2017.

[59] Francis Corson and Eric D Siggia. Gene-free methodology for cell fate dynamics during development. *Elife*, 6, December 2017.

[60] Philip B Kidd, Michael W Young, and Eric D Siggia. Temperature compensation and temperature sensation in the circadian clock. *Proc. Natl. Acad. Sci. U. S. A.*, 112(46):E6284–92, November 2015.

[61] D A Rand, B V Shulgin, D Salazar, and A J Millar. Design principles underlying circadian clocks. *J. R. Soc. Interface*, 1(1):119–130, November 2004.

[62] Alejandra C Ventura, Alan Bush, Gustavo Vasen, Matías A Goldín, Brianne Burkinshaw, Nirveek Bhattacharjee, Albert Folch, Roger Brent, Ariel Chernomoretz, and Alejandro Colman-Lerner. Utilization of extracellular information before ligand-receptor binding reaches equilibrium expands and shifts the input dynamic range. 0.

[63] Ryan Suderman, John A. Bachman, Adam Smith, Peter K. Sorger, and Eric J. Deeds. Fundamental trade-offs between information flow in single cells and cellular populations. *Proceedings of the National Academy of Sciences of the United States of America*, 114(22):5755–5760, 5 2017.

[64] Jeff Hasty, Joel Pradines, Milos Dolnik, and J. J. Collins. Noise-based switches and amplifiers forgene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 97(5):2075–2080, 2 2000.

[65] P. R. Patnaik. Robustness analysis of the e.coli chemosensory system to perturbations in chemoattractant concentrations. *Bioinformatics*, 23(7):875–881, 4 2007.

[66] Bing Liang Xu and Yi Tao. External noise and feedback regulation: Steady-state statistics of auto-regulatory genetic network. *Journal of Theoretical Biology*, 243(2):214–221, 11 2006.

[67] H A Johnson. Thermal noise and biological information. *The Quarterly review of biology*, 62(2):141–52, 6 1987.

[68] Omar P Tabbaa, German Nudelman, Stuart C Sealfon, Fernand Hayot, and Ciriyam Jayaprakash. Noise propagation through extracellular signaling leads to fluctuations in gene expression. *BMC systems biology*, 7:94, 9 2013.

[69] Savaş Tay, Jacob J. Hughey, Timothy K. Lee, Tomasz Lipniacki, Stephen R. Quake, and Markus W. Covert. Single-cell nf-kappab dynamics reveal digital activation and analogue information processing. *Nature*, 466(7303):267–71, 7 2010.

[70] Wenzhe Ma, Ala Trusina, Hana El-Samad, Wendell A Lim, and Chao Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4):760–73, 8 2009.

[71] Toby Lawrence. The nuclear factor nf-kappab pathway in inflammation. *Cold Spring Harbor perspectives in biology*, 1(6):a001651, 12 2009.

[72] Ryan A Kellogg, Chengzhe Tian, Tomasz Lipniacki, Stephen R Quake, and Savaş Tay. Digital signaling decouples activation probability and population heterogeneity. *eLife*, 4:e08931, 10 2015.

[73] Raymond Cheong, Alexander Hoffmann, and Andre Levchenko. Understanding nf-kappab signaling via mathematical modeling. *Molecular systems biology*, 4:192, 2008.

[74] Louise Ashall, Caroline A Horton, David E Nelson, Pawel Paszek, V Claire Harper, Kate Sillitoe, Sheila Ryan, David G Spiller, John F Unitt, David S Broomhead, Douglas B Kell, David A Rand, Violaine Sée, and Michael R H White. Pulsatile stimulation determines timing and specificity of nf-kappab-dependent transcription. *Science (New York, N.Y.)*, 324(5924):242–6, 4 2009.

[75] Ryan A. Kellogg and Savaş Tay. Noise facilitates transcriptional control under dynamic inputs. *Cell*, 160(3):381–392, 1 2015.

[76] Mark D. McDonnell and Derek Abbott. What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology. *PLoS Computational Biology*, 5(5):e1000348, 5 2009.

[77] Mark D. McDonnell and Lawrence M. Ward. The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience*, 12(7):415–426, 6 2011.

[78] J. J. Collins, Carson C. Chow, and Thomas T. Imhoff. Stochastic resonance without tuning. *Nature*, 376(6537):236–238, 7 1995.

[79] Andreas Piehler, Navid Ghorashian, Ce Zhang, and Savaş Tay. Universal signal generator for dynamic cell stimulation. *Lab on a Chip*, 17(13):2218–2224, 6 2017.

[80] Ryan A Kellogg, Rafael Gómez-Sjöberg, Anne A Leyrat, and Savaş Tay. High-throughput microfluidic single-cell analysis pipeline for studies of signaling dynamics. *Nature Protocols*, 9(7):1713–1726, 6 2014.

[81] Alexander Hoffmann, Andre Levchenko, and Martin L Scott. The ikb – nf- kb signaling module : Temporal control and selective gene activation. 298(November):1241–1245, 2002.

[82] T. M. Bui-Nguyen, S. B. Pakala, R. D. Sirigiri, W. Xia, M. C. Hung, S. K. Sarin, V. Kumar, B. L. Slagle, and R. Kumar. Nf-kb signaling mediates the induction of mta1 by hepatitis b virus transactivator protein hbx. *Oncogene*, 29(8):1179–1189, 2 2010.

[83] Bixing Zhao, Yingchao Wang, Xionghong Tan, Kun Ke, Xiaoyuan Zheng, Fei Wang, Shubing Lan, Naishun Liao, Zhixiong Cai, Yingjun Shi, Youshi Zheng, Yongping Lai, Lili Wang, Qin Li, Jingfeng Liu, Aimin Huang, and Xiaolong Liu. Inflammatory microenvironment contributes to stemness properties and metastatic potential of hcc via the nf-kb/mir-497/sall4 axis. *Molecular Therapy - Oncolytics*, 15:79–90, 12 2019.

[84] John M. Ankers, Raheela Awais, Nicholas A. Jones, James Boyd, Sheila Ryan, Antony D. Adamson, V. Claire Harper, Lloyd Bridge, David G. Spiller, Dean A. Jackson, Pawel Paszek, Violaine Sée, and Michael R.H. White. Dynamic nf-kb and e2f interactions control the priority and timing of inflammatory signalling and cell proliferation. *eLife*, 5(MAY2016), 5 2016.

[85] K. Araki, K. Kawauchi, and N. Tanaka. Ikk/nf-kb signaling pathway inhibits cell-cycle progression by a novel rb-independent suppression system for e2f transcription factors. *Oncogene*, 27(43):5696–5705, 9 2008.

[86] Ching Aeng Lim, Fei Yao, Joyce Jing Yi Wong, Joshy George, Han Xu, Kuo Ping Chiu, Wing Kin Sung, Leonard Lipovich, Vinsensius B. Vega, Joanne Chen, Atif Shahab, Xiao Dong Zhao, Martin Hibberd, Chia Lin Wei, Bing Lim, Huck Hui Ng, Yijun Ruan, and Keh Chuang Chin. Genome-wide mapping of rela(p65) binding identifies e2f1 as a transcriptional activator recruited by nf-kb upon tlr4 activation. *Molecular Cell*, 27(4):622–635, 8 2007.

[87] Morgan A. Sammons, Thuy Ai T. Nguyen, Simon S. McDade, and Martin Fischer. Tumor suppressor p53: From engaging dna to target gene regulation. *Nucleic Acids Research*, 48(16):8848–8869, 9 2020.

[88] Marianne Quaas, Gerd A. Müller, and Kurt Engeland. p53 can repress transcription of cell cycle genes through a p21 waf1/cip1-dependent switch from mmb to dream protein complex binding at chr promoter elements. *Cell Cycle*, 11(24):4661–4672, 12 2012.

[89] Iva A. Tchasovnikarova, Richard T. Timms, Christopher H. Douse, Rhys C. Roberts, Gordon Dougan, Robert E. Kingston, Yorgo Modis, and Paul J. Lehner. Hyperactivation of hush complex function by charcot-marie-tooth disease mutation in morc2. *Nature Genetics*, 49(7):1035–1044, 7 2017.

[90] Keiko Kawauchi, Keigo Araki, Kei Tobiume, and Nobuyuki Tanaka. p53 regulates glucose metabolism through an ikk-nf-kb pathway and inhibits cell transformation. *Nature Cell Biology*, 10(5):611–618, 5 2008.

[91] Julie M. Lowe, Daniel Menendez, Pierre R. Bushel, Maria Shatz, Erin L. Kirk, Melissa A. Troester, Stavros Garantziotis, Michael B. Fessler, and Michael A. Resnick. P53 and nf-kb coregulate proinflammatory gene responses in human macrophages. *Cancer Research*, 74(8):2182–2192, 4 2014.

[92] Mari B. Ishak Gabra, Ying Yang, Xazmin H. Lowman, Michael A. Reid, Thai Q. Tran, and Mei Kong. Ikkb activates p53 to promote cancer cell adaptation to glutamine deprivation. *Oncogenesis*, 7(11):1–12, 11 2018.

[93] Matthew S. Hayden and Sankar Ghosh. Signaling to nf-kb. *Genes and Development*, 18(18):2195–2224, 9 2004.

[94] Heike L. Pahl. Activators and target genes of rel/nf-kb transcription factors. *Oncogene*, 18(49):6853–6866, 11 1999.

[95] Qian Zhang, Michael J. Lenardo, and David Baltimore. 30 years of nf-kb: A blossoming of relevance to human pathobiology. *Cell*, 168(1-2):37–57, 1 2017.

[96] Fengyi Wan and Michael J. Lenardo. Specification of dna binding activity of nf-kappab proteins. *Cold Spring Harbor perspectives in biology*, 1(4), 2009.

[97] Ranjan Sen and Stephen T. Smale. Selectivity of the nf-{kappa}b response. *Cold Spring Harbor perspectives in biology*, 2(4), 2010.

[98] Gioacchino Natoli. Control of nf-kappab-dependent transcriptional responses by chromatin organization. *Cold Spring Harbor perspectives in biology*, 1(4), 2009.

[99] Damian Szklarczyk, Annika L. Gable, David Lyon, Alexander Junge, Stefan Wyder, Jaime Huerta-Cepas, Milan Simonovic, Nadezhda T. Doncheva, John H. Morris, Peer Bork, Lars J. Jensen, and Christian Von Mering. String v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1):D607–D613, 1 2019.

[100] Ann Richmond. Nf-kb, chemokine gene transcription and tumour growth. *Nature Reviews Immunology*, 2(9):664–674, 9 2002.

[101] V. Kirti Sawant, Krishna Mohan Sepuru, Emily Lowry, Brigith Penaranda, Charles W. Frevert, Roberto P. Garofalo, and Krishna Rajarathnam. Neutrophil recruitment by

chemokines cxcl1/kc and cxcl2/mip2: Role of cxcr2 activation and glycosaminoglycan interactions. *Journal of Leukocyte Biology*, pages JLB.3A0820–207R, 9 2020.

[102] Yifeng Xia, Roanna C. Padre, Tatiana Hurtado De Mendoza, Virginie Bottero, Vinay B. Tergaonkar, and Inder M. Verma. Phosphorylation of p53 by ikb kinase 2 promotes its degradation by beta-trcp. *Proceedings of the National Academy of Sciences of the United States of America*, 106(8):2629–2634, 2 2009.

[103] Michael Karin. How nf-kb is activated: the role of the ikb kinase (ikk) complex. 0.

[104] Klaus Heger and Vishva M. Dixit. Tbk1 and ikke restrain cell death. *Nature Cell Biology*, 20(12):1330–1331, 12 2018.

[105] Nf-kb activation by a signaling complex containing traf2, tank and tbk1, a novel ikk-related kinase | the embo journal, 0. [Online; accessed 2020-04-13].

[106] Yuhai Tu and Wouter-Jan Rappel. Adaptation in living systems. *Annual Review of Condensed Matter Physics*, 9(1):183–205, 3 2018.

[107] Corentin Briat, Ankit Gupta, and Mustafa Khammash. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Systems*, 2(1):15–26, 1 2016.

[108] Kentaro Inoue, Hisaaki Shinohara, Marcelo Behar, Noriko Yumoto, Gouhei Tanaka, Alexander Hoffmann, Kazuyuki Aihara, and Mariko Okada-Hatakeyama. Oscillation dynamics underlie functional switching of nf-kb for b-cell activation. *npj Systems Biology and Applications*, 2(1):16024, 12 2016.

[109] Lauren M. Workman and Hasem Habelhah. Tnfr1 signaling kinetics: Spatiotemporal control of three phases of ikk activation by posttranslational modification. *Cellular Signalling*, 25(8):1654–1664, 8 2013.

[110] Mogens H. Jensen and Sandeep Krishna. Inducing phase-locking and chaos in cellular oscillators by modulating the driving stimuli. *FEBS Letters*, 586(11):1664–1668, 6 2012.

[111] Martin Bier. Brownian ratchets in physics and biology. *Contemporary Physics*, 38(6):371–379, 11 1997.

[112] Amir Mitchell and Wendell Lim. Cellular perception and misperception: Internal models for decision-making shaped by evolutionary experience. *BioEssays*, 38(9):845–849, 9 2016.

[113] Amir Mitchell, Ping Wei, and Wendell A Lim. Oscillatory stress stimulation uncovers an achilles' heel of the yeast mapk signaling network. *Science (New York, N.Y.)*, 350(6266):1379–83, 12 2015.

[114] Vikas K. Garg, Adam Kalai, Katrina Ligett, and Zhiwei Steven Wu. Learn to Expect the Unexpected: Probably Approximately Correct Domain Generalization. pages 1–17, 2020.

[115] Qi Dou, Daniel C. Castro, Konstantinos Kamnitsas, and Ben Glocker. Domain Generalization via Model-Agnostic Learning of Semantic Features. (NeurIPS), 2019.

[116] Da Li, Yongxin Yang, Yi Zhe Song, and Timothy M Hospedales. Learning to generalize: Meta-learning for domain generalization. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, pages 3490–3497, 2018.

[117] Miguel Ruiz-Garcia, Ge Zhang, Samuel S. Schoenholz, and Andrea J. Liu. Tilting the playing field: Dynamical loss functions for machine learning. pages 1–9, 2021.

[118] Da Li, Yongxin Yang, Yi Zhe Song, and Timothy M. Hospedales. Deeper, Broader and Artier Domain Generalization. *Proceedings of the IEEE International Conference on Computer Vision*, 2017-Octob:5543–5551, 2017.

[119] Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, and Gabriel J. Brostow. Harmonic deep: Networks translation and rotation equivariance. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:7168–7177, 2017.

[120] Ximing Qin, Mark Byrne, Tetsuya Mori, Ping Zou, Dewight R Williams, Hassane McHaourab, and Carl Hirschie Johnson. Intermolecular associations determine the dynamics of the circadian KaiABC oscillator. *Proc. Natl. Acad. Sci. U. S. A.*, 107(33):14805–14810, August 2010.

[121] Joost Snijder, Jan M Schuller, Anika Wiegard, Philip Lössl, Nicolas Schmelling, Ilka M Axmann, Jürgen M Plitzko, Friedrich Förster, and Albert J R Heck. Structures of the cyanobacterial circadian oscillator frozen in a fully assembled state. *Science*, 355(6330):1181–1184, March 2017.

[122] Fumio Hayashi, Hiroki Ito, Masayasu Fujita, Ryo Iwase, Tatsuya Uzumaki, and Masahiro Ishiura. Stoichiometric interactions between cyanobacterial clock proteins KaiA and KaiC. *Biochem. Biophys. Res. Commun.*, 316(1):195–202, March 2004.

[123] Didier Gonze and Albert Goldbeter. Circadian rhythms and molecular noise. *Chaos*, 16(2):026110, June 2006.

[124] Aurore Woller, Didier Gonze, and Thomas Erneux. The goodwin model revisited: Hopf bifurcation, limit-cycle, and periodic entrainment. *Phys. Biol.*, 11(4):045002, August 2014.

[125] Giovanni Russo, Mario di Bernardo, and Eduardo D Sontag. Global entrainment of transcriptional systems to periodic inputs. *PLoS Comput. Biol.*, 6(4):e1000739, April 2010.

[126] Daniel Schultz, Mingyang Lu, Trevor Stavropoulos, José Onuchic, and Eshel Ben-Jacob. Turning oscillations into opportunities: lessons from a bacterial decision gate. *Sci. Rep.*, 3:1668, 2013.

[127] Yuansheng Cao, Hongli Wang, Qi Ouyang, and Yuhai Tu. The free energy cost of accurate biochemical oscillations. *Nat. Phys.*, 11(9):772–778, September 2015.