BIOREPS Problem Set #6
Cost and Precision of Brownian Clocks

1 Background

In a previous problem set, we saw how energetically expensive it was to correct protein misfolding. We could imagine that there is some degree of tolerance associated with the final form of the protein; in other words, the final protein structure may not have every disulfide bond, alpha helix, and beta pleated sheet constructed with unlimited precision. We saw that there was a finite energy cost for the precise construction of populations of functional proteins. This relationship between *energy cost* and *precision* is much more ubiquitous to biological processes.

In this problem set, instead of investigating the energy cost of *making* proteins, we will investigate the energy cost (and precision) of *operating* proteins, in particular, as biological Brownian clocks. Biological clocks are different from everyday clocks in that the intervals between 'ticks' varies randomly - this randomness or *stochasticity* is Brownian which means that the variation in ticks falls off as $1/f^2$. Warning: the usage of the word 'clock' here is wrought with peril since the everyday use of the word intuitively implies a high degree of precision - perhaps pacemaker or stochastic pump is more apt - but 'clock' is still appropriate in some sense because, similar to wristwatches, biological clocks signal and coordinate the timing of events like the daily release of hormones or the coupling between the heart and breathing.

Brownian clocks underlie almost all biological rhythms, many of them crucial to biological function (autonomic rhythms governing breathing, eating, etc.) and behavior (oscillations in the nervous system). Populations of Brownian clocks build pacemaker systems that cover at least 15 orders of magnitude (nanoseconds to days). For instance, pacemaker circuits range from the circuits controlling heart rate to the famed superchiasmatic nucleus governing circadian rhythms [1]. Because they are so ubiquitous and numerous, energy cost and precision are fundamental to the nature, design, and understanding of these systems. The operation of Brownian clocks can ultimately be reduced to the protein conformation/phosphorylation cycles. These protein cycles are fueled by ATP and operate with finite precision. One example is the phosphorylation-dephosphorylation cycle of the KaiC protein which governs the circadian rhythm of cyanobacteria.

The energy cost of protein phosphorylation cycles is inextricably tied with their timing precision. Intuitively, higher precision costs more energy and infinite precision requires infinite energy. What if that wasn't the case? As we'll see in the following problem set, unlike everyday clocks, Brownian clocks can achieve arbitrary precision at vanishingly low costs. This is an extremely unintuitive conclusion but could offer profound insight into the role of indeterminacy in physical systems and the (elected?) messiness of biological systems. First we will develop models for two classes of clocks: (1) a clock driven by a constant thermodynamic potential and (2) a Brownian clock, conceptually consistent with the description above. We will evaluate some of the properties of these two clocks, especially the relationship between clocking precision and energetic cost. Second, we will develop some of the theoretical framework on which this conclusion rests, in particular, the central proof that it is always possible to build what is known as a *steady state bipartite Markov process* for a Brownian clock [2].

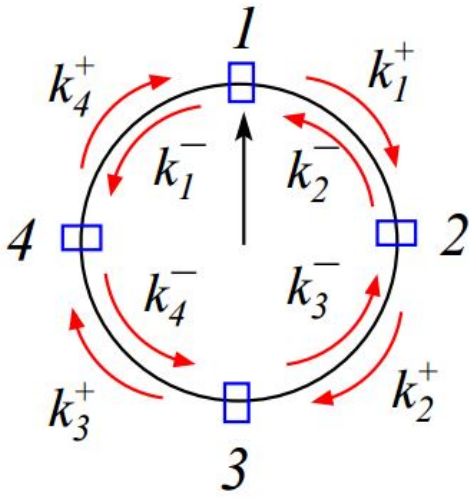

Figure 1: Illustration of a Brownian clock with four states. The clock "hand", while in state 1, can either move to state 2 with rate k_1^+, or to state 4 with rate k_1^-, etc. Figure credit: Ref. [2].

2 Questions

2.1 What is a Brownian Clock?

In a Brownian clock (Figure 1), we cannot measure the time directly by the clockwise (CW) cycle-completion time as we do a traditional clock. This is well-defined only when $k_n^- = 0$, which macroscopically corresponds to the fact that the second hand on an analog clock never ticks counter-clockwise (CCW). Instead, we must consider the probability current J of the system, which is defined as the average of the fluctuating current in the system divided by the total running time of the clock: $J \equiv \langle X \rangle / \mathcal{T}$. The inverse of the probability current (J^{-1}) is the average time for the clock to complete one cycle. Associated with this probability current is an uncertainty that we will define as the following:

$$\epsilon^2 \equiv \frac{\langle X^2 \rangle - \langle X \rangle^2}{\langle X \rangle^2} = \frac{2D}{J^2 \mathcal{T}} \tag{1}$$

a) Why is ϵ^2 an appropriate measure of the relative uncertainty in this system?

b) The diffusion coefficient has been defined as $D \equiv \frac{\langle X^2 \rangle - \langle X \rangle^2}{2\mathcal{T}}$. What is the difference between the definition of D here compared to the way it is usually defined?

c) If it costs us an amount $\mathcal{A}$ of free energy (i.e. chemical potential μ) to traverse the loop once, what would be a good measure of the average energetic cost to run this Brownian clock for a total time $\mathcal{T}$? Define this quantity as $\mathcal{C}$.

2.2 Brownian Clock Driven by a Fixed Thermodynamic Force

a) Referring to Figure 1, we will fix all the forward (CW) transition rates to k_+ and backward (CCW) transition rates to k_-. Write down the probability current, J, for this clock. *Hint:* You can

think of this as a random walk with step size $1/N$ so that the probability of being in any given state is uniform ($P_i = 1/N$ for $i = 1, ..., N$).

b) The diffusion coefficient in this case is given by $(k_+ + k_-)/(2N^2)$. Rationalize this from your knowledge of means, variance, and the specific functional form of J in Eq. (1).

c) Show that $\mathcal{C}\epsilon^2 = 2D\mathcal{A}/J$.

d) Show that $\mathcal{A}/N = \ln\left(\frac{k_+}{k_-}\right)$.

e) Use the results from parts c) and d) as well as the identity $\coth(x) = (e^{2x} + 1)/(e^{2x} - 1)$ to show:

$$\mathcal{C}\epsilon^2 = \frac{\mathcal{A}}{N}\coth\left(\frac{\mathcal{A}}{2N}\right) \tag{2}$$

f) The above result thus implies $\mathcal{C}\epsilon^2 \geq \max(2, \mathcal{A}/N)$. *Hint:* Expand e) to first order about small $\mathcal{A}$ and compare this to $\mathcal{A}/N$ with the knowledge that $|\coth(x)| > 1$. To what regime in $\mathcal{A}$ does this correspond?

g) We can further explore the physical manifestation of this uncertainty relation by considering the following examples to better understand the design, precision, and energetic cost of such a Brownian clock. If the system is far from equilibrium, your answer from f) implies a minimum number of states needed to ensure an uncertainty of no more than ϵ. Find $N_{\min}$ as a function of ϵ and X. We wish to keep our value of $\epsilon \leq 0.1$. Explicitly calculate $N_{\min}$ for $\langle X \rangle = 1$, 10, and 100.

h) Say we wish to measure a time of one hour to a certain precision $\epsilon = 10^{-2}$ with either a "slow" clock that ticks once every minute ($\langle X_{\text{slow}} \rangle = 60$) or a "fast" clock that ticks once every second ($\langle X_{\text{fast}} \rangle = 3600$), where one tick corresponds to one complete cycle of the given system. Find the analytical expression for the affinity $\mathcal{A}_{\min}$ required to drive a given N-state clock with a fixed $\mathcal{C}\epsilon^2$? Calculate this explicitly for both the "fast" and "slow" clocks.

The above examples of the "fast" and "slow" clocks may have different $\mathcal{A}_{\min}$, but their energetic cost is actually very close to the same; they are both bounded by ~ 20000! What does this mean? Our "fast" clock is not necessarily any more precise *or* energetically favorable than our "slow" clock, given that each cycle has a sufficient number of steps, as shown in g). This universal bound of $\mathcal{C}\epsilon^2 \geq 2$ insures that higher precision requires a higher cost in energy. A physical example of such a cost is found in ATP hydrolysis, which yields approximately $20k_BT$. To obtain an uncertainty ϵ, a consumption of $1/(10\epsilon^2)$ is required. We will contrast this with a clock driven by an external protocol, which does not share this same relation with energy consumption and precision.

2.3 Periodic Forcing

Another class of Brownian clocks are those driven by external protocols (Figure 2). These external protocols are not the result of fuel spending in the clock as with the thermodynamic clock,

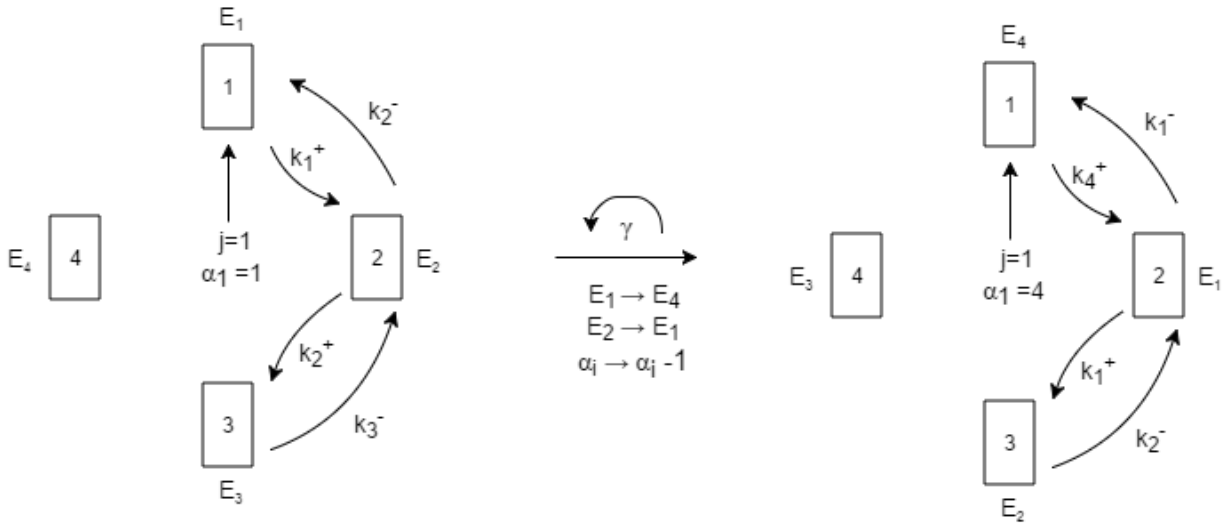

Figure 2: Illustration of a Brownian clock with a stochastic periodic protocol. The states are labeled by $i = 1, 2, \ldots$. The variable α_i undergoes a backwards jump, $\alpha_i \to \alpha_i - 1$, with rate $\gamma = \frac{N}{\tau}$. This corresponds to a jump in energies $E_{\alpha_i} \to E_{\alpha_i - 1}$, and likewise for the rates, $k^+_{\alpha_i}$ and $k^-_{\alpha_{i+1}}$ and barriers, $B_{\alpha_i}(t)$. Half of the network and the barriers, , are not included for simplicity. The pointer indicates the clock is in position $j = 1$ and remains there for the duration of a jump in the α_i since a transition changing in both variables α_i and the position of the clock j is not allowed by definition of a bipartite Markov process. Figure credit: Ref. [2].

but rather the result of diffusion according to the orientation of the site energies and energy barriers of the regime. This orientation changes counterclockwise with period τ and rate $\gamma = \frac{N}{\tau}$, resulting in an effective clockwise shift in the transition rates from i to $i + 1$. To determine these transition rates, we again consider a ring with equidistant sites and N states. With site i having energy $E_i(t)$ and the energy barrier between i and $i + 1$ being $B_i(t)$, we define

$$\epsilon_i \equiv e^{E_i(t)} \qquad \chi_i(t) \equiv e^{-B_i(t)} \tag{3}$$

Then the transition rates are given as

$$k_{i,i+1}(t) = \chi_i(t)\epsilon_i(t) \qquad k_{i,i-1}(t) = \chi_{i-1}\epsilon_i(t). \tag{4}$$

You should notice that this is just a clean way of writing some very familiar equations!

The calculation of useful quantities for externally driven clocks is, in general, extremely difficult when not impossible. Here we'll lay out the process of calculating D, J, and σ, but only calculate them for a couple special cases. Together, the clock and the external protocol driving it form a bipartite Markov process. There are two sets of indices which will only jump to the other set, making it bipartite, and they do so in a manner satisfying the conditions of a Markov process. The matrix describing this process has the following form:

$$\begin{aligned}(\mathbf{L}_n)_{i+1,i} &= \chi_{i-n}\epsilon_{i-n} \\ (\mathbf{L}_n)_{i,i} &= -(\chi_{i-n} + \chi_{i-1-n})\epsilon_{i-n}. \\ (\mathbf{L}_n)_{i-1,i} &= \chi_{i-1-n}\epsilon_{i-n}\end{aligned}$$

Note that jumps from one state to another can only happen between adjacent states, meaning all other elements of the matrix must be zero. It is useful here to relabel by transition rate, instead of position on the ring. We will define the index $\alpha = 1$ to be the index of the state with transition rate $\epsilon_i \chi_i$ to jump to state $\alpha = 2$ and the transition rate $\epsilon_i \chi_N$ to jump to $\alpha = N$. The rate of the change in the protocol is γ. Then the above matrix in conjunction with the modified generator associated with current X then gives the matrix:

$$\begin{aligned}
\mathbf{L}^*(z)_{\alpha+1,\alpha} &= \chi_\alpha \epsilon_\alpha e^{z/N} \\
\mathbf{L}^*(z)_{\alpha,\alpha} &= -(\chi_\alpha + \chi_{\alpha-1})\epsilon_\alpha - \gamma. \\
\mathbf{L}^*(z)_{\alpha-1,\alpha} &= \chi_{\alpha-1}\epsilon_\alpha e^{-z/N} + \gamma
\end{aligned}$$

By calculating the coefficients $C_m(z)$ according to

$$\sum_{m=0}^{N} C_m(z) x^m \equiv det[x\mathbf{I} - \mathbf{L}^*(z)]$$

we can calculate J and D according to

$$J = -C_o'/C_1$$

$$D = -(C_0'' + 2C_1' J + 2C_1 J^2)/2C_1,$$

where the prime indicates a derivative in z and the coefficients are evaluated at $z = 0$. From this, the entropy is calculated by

$$\sigma = \sum_n P^n \left[\sum_{i<j} J_{ij}^n d_{ij} A^n + \gamma_n \sum_i P(i|n)(E_i^{n+1} - E_i^n) \right]$$

where P^n denotes the probability of the external protocol being in state n, $J_{ij}^n = P(i|n)k_{ij}^n - P(j|n)k_{ji}^n$ is the probability current, and E_i^n is the free energy of of the state i with the external protocol in state n. It's not hard to imagine that an actual calculation of any of these quantities would be exceptionally difficult except in special cases. We'll look at two such special cases now. When $C\epsilon^2$ is optimized, $\chi_N = 0, \chi_1 = \chi_2 = \cdots = \chi_{N-1} = \chi$, and $\chi >> \gamma$.

a) What does $\chi_N = 0$ tell you about the system? What about $\chi >> \gamma$?

The equilibrium distribution of this system is given by

$$P_\alpha^\star = e^{-E_\alpha}/Z$$

where Z is the partition function. If we define $k_+^{eff} \equiv \gamma \sum_{\alpha=2}^{N} P_\alpha^\star$ and $k_-^{eff} \equiv \gamma P_1^\star$, the master equation can be written as

$$\frac{d}{dt}P(X,t) = k_+^{eff} P(x - 1/N, t) + k_-^{eff} P(X + 1 - 1/N, t) - (k_-^{eff} + k_+^{eff})P(X,t).$$

This master equation is not very fun to solve in X space, but, as is the case with most ODEs, we can project our very hard problem onto a space where it becomes trivial. Define the Laplace

transform as $\tilde{P}(z,t) \equiv \sum_{X\geq 0} e^{zX} P(X,t)$. $\tilde{P}(z,t)$ is also known as a moment generating function of $P(X,t)$. To see why this is, take the first couple derivatives of $\tilde{P}(z,t)$ with respect to z and evaluate them at $z = 0$.

Multiplying both sides of the master equation by e^{zX} and summing over X yields

$$\frac{d}{dt}\tilde{P}(z,t) = \left[k_+^{eff} e^{z/N} + k_-^{eff} e^{-(N-1)z/N} - (k_+^{eff} + k_-^{eff})\right]\tilde{P}(z,t) \tag{5}$$

b) From the definition of $\tilde{P}(z,t)$ above, show that the natural boundary condition for Eq. (5) is $\tilde{P}(0,t) = 1$. Solve (5), writing your answer in the form $\tilde{P}(z,t) = e^{\psi(z)t}$, for some $\psi(z)$. Show that $J = \frac{d}{dz}\psi(z)|_{z=0}$ is consistent with the J given in the paper for this case: $J = \gamma Z^{-1} N^{-1}[\sum_{\alpha=2}^{N} e^{-E_\alpha} - (N-1)e^{-E_1}]$.

c) Show that $2D = \frac{d^2}{dz^2}\psi(z)|_{z=0}$ is consistent with the D given in the paper for this case, $D = \frac{1}{2}\gamma Z^{-1} N^{-2}[\sum_{\alpha=2}^{N} e^{-E_\alpha} + (N-1)^2 e^{-E_1}]$.

d) Show that σ for this system with the above equation matches the σ given in the paper for this system, $\sigma = \gamma Z^{-1}[\sum_{\alpha=1}^{N} e^{-E_\alpha}(E_{\alpha-1} - E_\alpha)]$. Note that only the second term in the general expression for σ needs to be calculated here.

Now we'll shift our focus to another special case which is simple to solve, and elucidates one of the main results of the paper. Consider a Dissipation-less clock with $E_\alpha = E\delta_{\alpha,1}$ where $\delta_{\alpha,1}$ is the Kronecker delta.

e) Using the equations you verified in b), c), and d) above, use the relation $\mathcal{C}\epsilon^2 = 2D\sigma/J^2$ to calculate $\mathcal{C}\epsilon^2$ for this system. You should find that

$$\mathcal{C}\epsilon^2 = \frac{[1 + e^{-E}(N-1)]E}{(N-1)(1-e^{-E})}. \tag{6}$$

f) Consider what you would expect from the value of $\mathcal{C}\epsilon^2$ if we chose E and N in such a way that $e^E >> N >> E$. Show that, under these conditions, $\mathcal{C}\epsilon^2 \sim E/N$. What does this say about the clock's uncertainty? What does it say about how much energy the system dissipates? This is touted as the "main result" of the paper!

The asymptotic analysis you performed in the previous question illuminates the fundamental difference between a Brownian clock driven with a constant force and one that is pumped periodically, but what can we say about the optimal $\mathcal{C}\epsilon^2$ for any given N? One cannot analytically minimize Eq. (6), so we'll have to do some numerical analysis. Note that for the following questions, the `Mathematica` functions NMinimize, List(LogLinear)Plot, NonlinearModelFit, and Series might be useful. You can read all about them and play around with interactive examples in the MM documentation.

g) For $N = 3$ to $N = 1000$, numerically minimize (6) with respect to E and plot the 'optimal energy profile' vs. N (make one regular plot and one log-linear plot).

h) From your second plot, argue that $E_{opt}(N)$ can be fit to the form $E_{opt}(N) = a + b\ln(N)$ for some constants a and b. Do the fitting and show that $E_{opt}(N) \approx \frac{3}{5}(1 + 2\ln(N))$.

i) We can now plug $E_{opt}(N)$ into (6) to get $\mathcal{C}\epsilon_{opt}^2$, but the answer is kind of messy. Expand around large N to first order to show that in this case $\mathcal{C}\epsilon_{opt}^2 \sim E_{opt}/N$, as we know it must.

This main result is only achievable through the theoretical framework that describes systems driven by a stochastically-varying protocol: the bipartite Markov process. No parts of this process conflict with thermodynamics. Indeed, we could include the (well-defined) entropy generated by: 1) a stochastic protocol with allowed backward jumps, 2) the changes of the protocol, and 3) the generation of the stochastic protocol itself. Barato and Seifert note that for a physically realizable system with a truly external process, the entropy of such protocols–stochastic or deterministic–is irrelevant (e.g. night/day changes). Experiments have yet to be performed to realize the results obtained in this problem set, but possible candidates for a Brownian clock could include single molecules, colloidal particles, and small electronic systems [2].

References

[1] G. Buzsaki, *Rhythms of the Brain* (Oxford University Press, 2006).

[2] A.C. Barato and U. Seifert, "Cost and Precision of Brownian Clocks," (arxiv.org/pdf/1610.07960v1.pdf).