Fraudulent white noise: Flat power spectra belie arbitrarily complex processes

Paul M. Riechers ⬤[*]

Complexity Institute, Nanyang Technological University, 639798 Singapore, Singapore
and School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore, Singapore

James P. Crutchfield ⬤[†]

Department of Physics, Complexity Sciences Center, University of California at Davis, Davis, California 95616, USA

Power spectral densities are a common, convenient, and powerful way to analyze signals, so much so that they are now broadly deployed across the sciences and engineering—from quantum physics to cosmology and from crystallography to neuroscience to speech recognition. The features they reveal not only identify prominent signal frequencies but also hint at mechanisms that generate correlation and lead to resonance. Despite their near-centuries-long run of successes in signal analysis, here we show that flat power spectra can be generated by highly complex processes, effectively hiding all inherent structure in complex signals. Historically, this circumstance has been widely misinterpreted, being taken as the renowned signature of "structureless" *white noise*—the benchmark of randomness. We argue, in contrast, to the extent that most real-world complex systems exhibit correlations beyond pairwise statistics their structures evade power spectra and other pairwise statistical measures. As concrete physical examples, we demonstrate that fraudulent white noise hides the predictable structure of both entangled quantum systems and chaotic crystals. To make these words of warning operational, we present constructive results that explore how this situation comes about and the high toll it takes in understanding complex mechanisms. First, we give the closed-form solution for the power spectrum of a very broad class of structurally complex signal generators. Second, we demonstrate the close relationship between eigenspectra of evolution operators and power spectra. Third, we characterize the minimal generative structure implied by *any* power spectrum. Fourth, we show how to construct arbitrarily complex processes with flat power spectra. Finally, leveraging this diagnosis of the problem, we point the way to developing more incisive tools for discovering structure in complex signals.

I. INTRODUCTION

Innovative science probes the unknown. Success in discovering the mechanisms that underlie the systems we seek to understand, though, requires distinguishing structure from noise. Often, this distinction falls to discretion: Structure is that part of a signal we can predict, while noise stands in as a catch-all for everything else. This conundrum holds especially in the analysis of signals from truly complex systems, as when analyzing data from multielectrode arrays in brain tissue [1] or social experiments [2]. These systems are often said to be "noisy" even though the so-called noise may be entirely functionally relevant but in an unknown way [3]. Such descriptions fall far short of a principled approach that explains all trends and correlational structure, which would claim success only when all that remains unexplained in the signal is structureless white noise. Even this principled approach ultimately begs the central question, though: How do we test whether an apparently random signal is truly white noise?

The challenge of discovering structure in noisy signals is compounded manifold, as we demonstrate in the following, when our chosen observables hide arbitrary amounts of in-principle-predictable structure behind a familiar signature of white noise—the flat power spectrum. Said simply, observables can be completely devoid of pairwise correlation, while still embodying structure in higher-order correlations. More precisely, we will show that structure can be hidden beyond any arbitrarily large order-N correlation—that not appearing in pairwise, three-way, or any n-way statistics, up to some arbitrarily large N. Moreover, the hidden structure can be arbitrarily sophisticated. It can be used, for example, to embed messages while shifting (and so hiding) the messages' content beyond N-way correlation. Here we explore the structures conveyed and hidden by power spectra, revealing a novel perspective on the interplay between structure and noise in Fourier analysis.

Section II discusses temporal structure and provides closed-form expressions for the power spectra from autonomous signal generators. It highlights the intimate

connection between power spectra and eigenspectra of a system's time-evolution generator. Section III then introduces a suite of results on structure that is *hidden* by power spectra. Notably, it introduces a general condition for *fraudulent white noise* processes—structured processes with a flat power spectrum—which applies very broadly, including to input-dependent processes with nonstationary high-order statistics. Section IV demonstrates that fraudulent white noise is observed in important physical systems. We show that fraudulent white noise arises in measurements of entangled quantum systems. We also show that flat diffraction patterns belie the predictable structure of chaotic crystals. Taken together the results emphasize the power spectrum's shortcomings for the task of structure detection. In response, Sec. V considers more sophisticated measures of structure. We give closed-form expressions for polyspectra—which are often advocated as the natural next step for detecting higher-order structure—but show that these too have severe blind spots. This motivates us to introduce the *dependence function* which identifies the presence of novel finite-range dependencies that contribute to total correlation. Section VI concludes the development. Appendices present detailed derivations, as well as several generalizations, of the main results.

II. STRUCTURE IN SPACE AND TIME?

Pairwise correlations are encountered throughout the sciences and engineering, especially in statistical physics. They are assumed, estimated, relied on, designed with, and used for interpretation widely. The following explores several specific examples of pairwise correlation that arise in different fields. These will set the context for our development, particularly for experts in the associated fields. However, our general results should be accessible and relevant across disciplines, as they rely primarily on basic probability theory and linear algebra.

A well-studied lesson from statistical physics is that diverging correlation length heralds the emergence of new types of order. Remarkably, mechanistically distinct physical systems share many universal behaviors near a critical point of emergent order, including the scaling of spatial pairwise correlation length [4]. More broadly, pairwise correlations are indicators of fundamental physical processes. For example, the fluctuation–dissipation theorem says that pairwise temporal correlations in equilibrium determine the friction encountered in transport processes. The Green-Kubo relations [5] make this explicit. Far from equilibrium, say, in computing devices and biological systems composed of excitable media, temporal correlations are signatures of richly coordinated state trajectories.

Pairwise correlations are directly viewed in the frequency domain via power spectral densities. Indeed, power spectra are employed as a basic data analysis tool in many scientific domains and have been key to major scientific discoveries. For example, comparing alternative theoretical predictions for power spectra of incident electromagnetic radiation from locally thermalized bodies, an unexpected discrepancy—the ultraviolet catastrophe—led to the acceptance of Planck's theory of quantized energies and the subsequent birth of quantum theory [6–8]. A contemporary example of the prominent role of power spectra is seen in the exquisitely detailed map of the cosmic microwave background (CMB)—a snapshot of the early universe's spatial correlations. In fact, models of the early universe are now benchmarked against their ability to replicate the CMB power spectrum [9].

In applied mathematics, power spectra played a key role in highlighting the defining features of the strange attractors of dynamical systems theory [10,11]. This led to the discovery of Ruelle-Pollicott resonances, where mixing and the decay of correlations in chaotic systems were related to the point spectrum of the Ruelle-Perron-Frobenius operator [12–14]. Indeed, the power spectra of chaotic systems are still actively used to analyze the behavior of everything from open quantum systems [15,16] to climate models [17].

The famous $1/f$ decay of power spectra found in many complex systems has received considerable attention throughout many decades [18–20]—sometimes being attributed to self-organized criticality [21]; almost always being taken as a signature of truly complex systems. More recently, the value of α in $1/f^\alpha$ noise—and deviations from this mean behavior—are used to interpret particle tracking experiments [22,23]. Related advances have enabled extraction of physical properties from power spectral analyses of nonstationary processes [24–26].

Power spectra are regularly used to discover structure in materials science and biology, too. x-Ray diffraction patterns—used to identify crystalline and molecular organization and central to discovering DNA's double helix [27–30]—are power spectra of scatterer densities, as we explain in Appendix A. Power spectra have been used to identify temporal correlations in single-neuron spike trains, refuting the common Poissonian white-noise assumption common in theoretical and computational neuroscience [31–34]. This allows the possibility that temporal correlations in the spike train—rather than just the firing rate—can play an important role in the neural code [35,36]. On a much larger (mean-field) scale, brain wave activity in different frequency bands gives signatures of normal brain functioning, as well as pathological conditions. Rhythmic brain-wave activity is clinically assessed through real-time power spectra of electroencephalography (EEG) signals [37–39].

From the smallest to the largest scales in the universe, when probing both the inanimate and the animate, power spectra are a central diagnostic tool for structure and validating scientific models. Their use is so important that special-purpose spectrum analyzers are standard laboratory test equipment; they can be readily purchased from dozens of major manufacturers.

Power spectra report pairwise correlations in a signal. But how much of a system's structure is faithfully represented by pairwise correlation? Are there important types of order that evade power spectra completely? To answer these questions, we first consider the problem of hidden structure through the lens of autocorrelation and power spectra. Only then, once the strengths and weakness of power spectra are clear, do we move on to more sophisticated measures of structure. Along the way we trace a path that begins to reveal what one can mean by "statistical dependency," "correlation," and "structure."

A. Correlation and power spectra

To provide a common ground, consider discrete-time processes described by an interdependent sequence $\ldots X_0 X_1 X_2 \ldots$ of random variables X_t that take on values $x \in \mathcal{A}$ within an alphabet assumed (for now) to be a subset of the complex numbers: $\mathcal{A} \subset \mathbb{C}$. (For concreteness here, we interpret t as indexing time $t = t\tau_0$, where τ_0 is the duration of each time step. For other kinds of stochastic process, t may represent spatial or angular coordinates.) An observed process may have a discrete domain, as with a classical discrete-time communication channel or a series of quantum measurements or, otherwise, may be a regularly sampled process evolving in continuous time.

A signal's *power spectrum* or, more properly, its *power spectral density* quantifies how its power is distributed across frequency [40,41]. For a discrete-domain process it is as follows:

$$P(\omega) = \lim_{N \to \infty} \frac{1}{f_0 N} \left\langle \left| \sum_{t=1}^{N} X_t e^{-i\omega t} \right|^2 \right\rangle, \qquad (1)$$

where the angle brackets denote the expected value over the random variable chain $X_1 X_2 X_3 \cdots X_N$, $\omega = 2\pi f / f_0$ is the angular frequency, f is the frequency, and $f_0 = 1/\tau_0$ is the fundamental frequency. We set f_0 to unity in the discrete-time case. In the continuous-time limit where $\tau_0 = dt \to 0$, the power spectrum becomes

$$P(f) = \lim_{L \to \infty} \frac{1}{L} \left\langle \left| \int_0^L X_t e^{-i2\pi f t} \, dt \right|^2 \right\rangle,$$

where we use the fact that $\omega t = 2\pi f t$. In either discrete or continuous time, integrating over any band of frequencies gives the power in that band.

For *wide-sense stationary* stochastic processes the *autocorrelation function*,

$$\gamma(\tau) = \langle \overline{X}_t X_{t+\tau} \rangle, \qquad (2)$$

is independent of the global time shift t and depends only on the relative time separation τ between observables [42]. The bar above X_t denotes its complex conjugate. Equation (2) makes plain the connection between pairwise statistics and the pairwise correlation function. For wide-sense stationary stochastic processes, the power spectrum is also determined by the signal's autocorrelation function $\gamma(\tau)$:

$$P(\omega) = \lim_{N \to \infty} \frac{1}{f_0 N} \sum_{\tau=-N}^{N} (N - |\tau|) \gamma(\tau) e^{-i\omega\tau}. \qquad (3)$$

The windowing function $N - |\tau|$ appearing in Eq. (3) is a direct consequence of Eq. (1); it is not imposed externally, as is common practice in signal analysis. (This factor is important for controlling convergence in our subsequent derivations.)

Equation (3) suggests that the power spectrum is very *nearly* the Fourier transform of the autocorrelation function, except for the $N - |\tau|$ term. In fact, the Wiener-Khinchin theorem proves that the power spectrum is indeed *equal* to the Fourier transform of the autocorrelation function for wide-sense stationary processes [43,44]. Note, too, that the pairwise

correlation function $\gamma(\tau)$ can be obtained via the inverse Fourier transform of the power spectrum $P(\omega)$.

B. Temporal structurelessness

Our goal is to understand temporal structure and to identify it in stochastic processes. To detect structure, even when hidden, we first must establish a baseline reference process that has no temporal structure: genuine white noise.

White noise processes, if we remove their mean value, have zero autocorrelation for all $\tau > 0$. Colloquially, white noise is often taken as a synonym for any completely random process with no statistical dependencies whatsoever. To be precise, we define *genuine white noise* as those processes for which each random variable X_t is statistically independent of all others $X_{t' \neq t}$, and each is identically distributed according to the same probability density function (PDF) over the alphabet. That is, the random variables in the sequence are *independent and identically distributed* (IID).

Familiar examples include a sequence of coin flips or the sequence of sums when rolling a pair of dice. As an example from contemporary physics, consider the (classical) process that results from observing a sequence of Bell-pair quantum states [45]. For each Bell pair, one of the entangled particles is sent to Alice and the other sent to Bob. Alice makes a sequence of measurements (along any measurement axis). The measurement output sequence she observes is pure white noise, with each measurement outcome having equal and independent probability of being up or down along the measurement axis. In fact, more sophisticated deployments of Bell pairs are being developed to provide certifiable random number generation [46]. Experiments now concentrate on increasing the rate of generating white noise [47,48].

The most recognizable feature of all white noise processes is their flat power spectrum. For any IID process, it follows directly from Eq. (2) that $\gamma(0) = \langle |X_t|^2 \rangle$, whereas $\gamma(\tau) = |\langle X_t \rangle|^2$ for $\tau \neq 0$. From Eq. (3), this immediately yields the familiar flat power spectrum of white noise, together with a δ function at zero frequency, corresponding to the signal's constant offset. For real-valued IID processes with zero mean (and $f_0 = 1$), this simplifies further to $\gamma(\tau) = \sigma^2 \delta_{0,\tau}$ and so $P(\omega) = \sigma^2$. In fact, the flat power spectrum has height equal to the variance $\sigma^2 = \langle X_t^2 \rangle - \langle X_t \rangle^2$ of the white noise for any real-valued IID process. The flat power spectrum for IID processes indicates that any temporal structure in the generating source has such short memory that it vanishes within the short sampling time τ_0 between each observation.

Gaussian white noises tend to be the most commonly employed white noise processes and, usually, for good reason. By the central limit theorem, Gaussian white noise arises generically in systems whenever *many* events—with amplitude of finite variance and with rapidly decaying correlation (compared to the timescale between observations)—contribute additively to each individual observation. Suppose, for example, that the expected number of these contributions to each new observation is simply proportional to the time since the last observation. When sampled at interval $dt = \tau_0$, the central limit theorem then tells us that each observation of the accumulated noise is IID and Gaussian distributed with variance $\sigma_\eta^2 \propto dt$. This immediately leads to the familiar standard

(a)

(b)
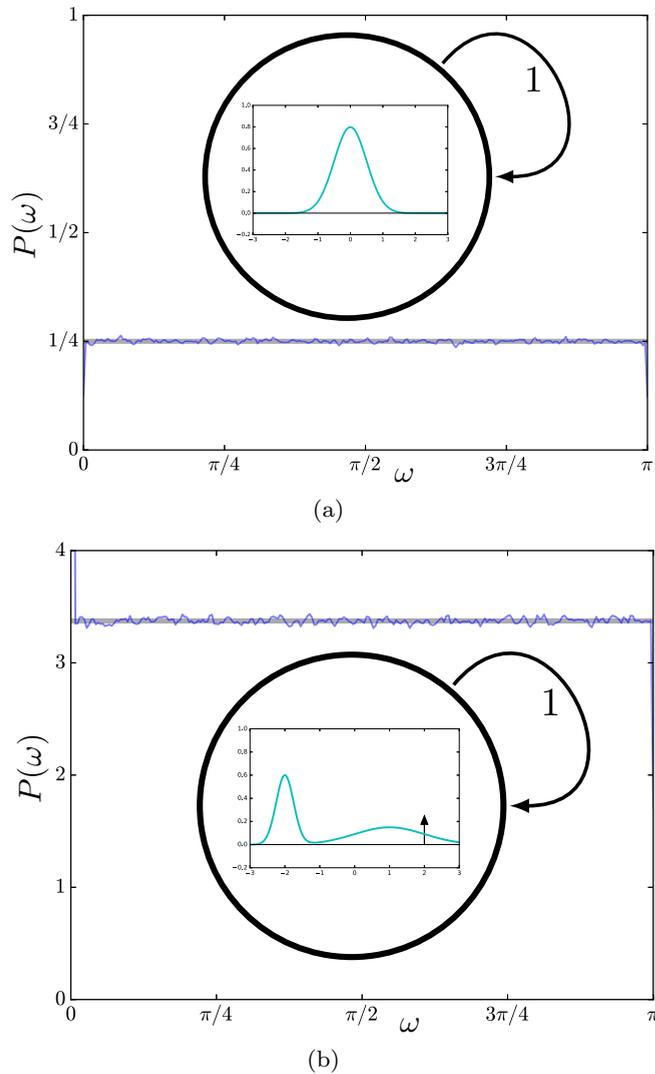
FIG. 1. Genuine white noise processes have no memory: Represented structurally by a state machine with a single state that is repeatedly visited with each observation. The same probability density function, inscribed in the state, is sampled at each time step. (a) Gaussian white noise process (inset) and its flat power spectrum. (b) Non-Gaussian white noise process (inset) and its flat power spectrum. For each (a) and (b), the flat power spectrum is given theoretically (thick gray), with height equal to the variance of the probability density function. We also display the numerically obtained power spectrum (thin blue) for each. The class of all possible (not-necessarily Gaussian) memoryless white noises is identical with the class of processes generated by single-state machines. This class, in turn, is identical to that of all IID processes (spanning all possible probability density functions). These temporally structureless processes constitute all possible varieties of genuine white noise.

deviation $\sigma_\eta \propto \sqrt{dt}$ of the additive noise $\eta(t)$ that appears when numerically integrating stochastic differential equations (e.g., Langevin equations); this, in turn, produces the trajectories of slower random variables [49].

The memoryless nature of repetitive sampling from a distribution is apparent in the state machine shown in Fig. 1(a). The same Gaussian distribution is repeatedly sampled with

probability 1 (as depicted by the self-transition probability there) for each observation, regardless of what happened previously [50].

Other "structureless" white noises are also possible. In fact, any of an uncountably infinite set of different IID processes—Gaussian, Poisson, Bernoulli, or any process that resamples a particular distribution at each time step—all yield the flat power spectrum or white noise. Non-Gaussian noise can emerge from repetitive sampling of a system's (non-Gaussian) stationary distribution when the relaxation timescales are far shorter than the time elapsed between samples. Alternatively, non-Gaussian white noise can arise when only a few physical events contribute to each observation, in which case the non-Gaussianity may reveal features of the physical generative mechanism. Nevertheless, these processes possess no temporal structure on the timescale of observation and, in particular, generate absolutely no correlations in the sequence of observations.

The hallmark of this structural paucity is the single state for the hidden Markov model (HMM) that describes all of these IID processes, as depicted in Fig. 1(b) [51]. The single state means that no influences from the past can affect the next or future samples. These are the genuine white noises.

In sharp contrast, we explore stochastic processes with arbitrarily sophisticated temporal structure on the timescale of observation. The much more general class we next consider allows for a thorough investigation of temporally structured stochastic processes. One surprising feature is that these very structured processes, described by arbitrarily complicated transition dynamics within memoryful collections of internal states, can have the flat power spectrum of white noise. These are the *fraudulent white noise processes*: white noise processes with a flat power spectrum that are nevertheless not *genuine* white noise. Fraudulent white noise contains statistical dependencies—predictable structure completely veiled by common measures of correlation.

C. Models of temporal structure

Structure arises over time from the interdependence among observables. To explicitly address structure in a broad class of temporally structured processes, we use HMMs as our preferred representation for autonomous signal generators [54–60]. Later sections introduce yet more sophisticated models with input dependence.

Despite Markovian state-to-state transitions, HMMs can generate temporally structured non-Markovian stochastic processes—those with infinite history dependence (infinite Markov order). Processes generated by even finite-state HMMs, in fact, *typically* have infinite-range statistical dependencies between observables since simple state-transition motifs guarantee this feature [61]. In addition to this richness and their ability to compactly generate the exact temporal statistics of nonlinear dynamical systems, HMMs are attractive since they are amenable to linear operator techniques [62–68].

Section IV employs HMMs to represent (i) sequential measurements of entangled quantum systems, (ii) scattering factors of disordered materials, and (iii) ion transport through
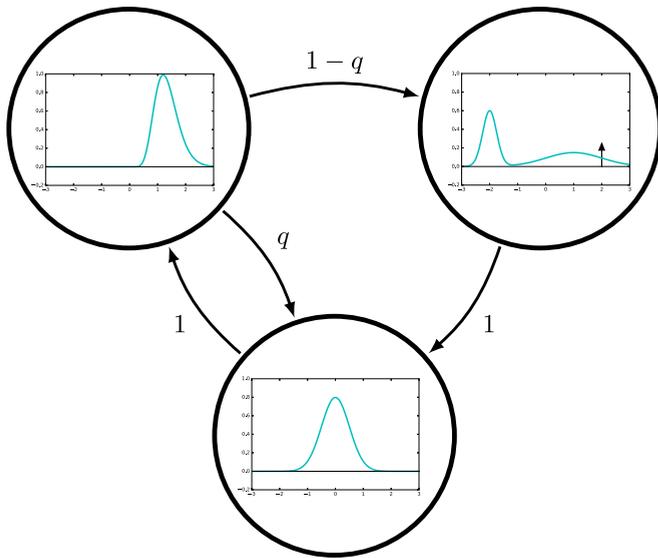
FIG. 2. Simple 3-state HMM that generates a stochastic process according to the state-to-state transition dynamic T and the PDFs $\{p(X|s)\}_{s \in \mathcal{S}}$ associated with each state. Theorem 1 asserts that its power spectrum is the same (modulo constant offset) as the power spectrum generated from an alternative process where each state's PDF is solely concentrated at the average value $\langle X \rangle_{p(X|s)}$ of the original PDF associated with the state.

biomolecular channels. But, to get there, we must first introduce the general properties of HMMs.

Let the 4-tuple $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, T)$ be a discrete-time HMM that generates the stationary stochastic process $\ldots X_{-2}X_{-1}X_0X_1X_2 \ldots$ according to the following. $\mathcal{S}$ is the (finite) set of states of the internal Markov chain and $\mathcal{A} \subseteq \mathbb{C}$ is the observable alphabet. $\mathcal{S}_t$ is the random variable for the hidden state at time t that takes on values $s \in \mathcal{S}$. X_t is the random variable for the observation at time t that takes on values $x \in \mathcal{A}$.

Given the hidden state at time t, the possible observations are distributed according to the conditional probability density functions: $\mathcal{P} = \{p(X_t | \mathcal{S}_t = s)\}_{s \in \mathcal{S}}$. For each $s \in \mathcal{S}$, $p(X_t | \mathcal{S}_t = s)$ may be abbreviated as $p(X|s)$ since the probability density function in each state is assumed to not change over t. Similarly, we will write $p(x|s)$ for $p(X_t = x | \mathcal{S}_t = s)$. Finally, the hidden-state-to-state stochastic transition matrix T has elements $T_{s,s'} = \Pr(\mathcal{S}_{t+1} = s' | \mathcal{S}_t = s)$, which give the probability of transitioning from hidden state s to s' given that the system is in state s, where $s, s' \in \mathcal{S}$. It is important for subsequent developments that $\Pr(\cdot)$ denotes a probability in contrast to $p(\cdot)$ which denotes a probability *density*.

Epitomizing the processes in the class considered, Fig. 2 presents a rather simple HMM with continuous observable alphabet $\mathcal{A} = \mathbb{R}$, whose samples are distributed according to the probability density function shown within each hidden state. As seen in the HMM's top-right state, both continuous probability density functions and discrete output probabilities can be accommodated in this framework: Finite probability of a particular observable is accomplished by an appropriately weighted Dirac δ function in the probability density function. The memoryful structure in Fig. 2 should be contrasted with

the completely memoryless processes of genuine white noise shown in Fig. 1.

The Bayes network in Fig. 3 depicts the structure of conditional independence among the random variables for these memoryful signal generators. For example, for a generic HMM, $p(X_t | X_{t-N} \ldots X_{t-2}X_{t-1} = x_{t-N} \ldots x_{t-2}x_{t-1})$ cannot be simplified since the condition on even arbitrarily distant past observables can influence the probability of the current observable. However, when conditioning on hidden states, the situation can simplify markedly. For example:

$$
\begin{aligned}
p(X_t | X_{t-N} &\ldots X_{t-2}X_{t-1} = x_{t-N} \ldots x_{t-2}x_{t-1}, \\
&\mathcal{S}_{t-N} \ldots \mathcal{S}_{t-2}\mathcal{S}_{t-1} = s_{t-N} \ldots s_{t-2}s_{t-1}) \\
&= p(X_t | \mathcal{S}_{t-1} = s_{t-1}) \\
&= \sum_{s \in \mathcal{S}} T_{s_{t-1},s}\, p(X|s).
\end{aligned}
$$

FIG. 3. Bayesian network for a state-emitting hidden Markov model graphically depicts the structure of conditional independence among random variables for the hidden state $\{\mathcal{S}_n\}_{n \in \mathbb{Z}}$ at each time n and the random variables $\{X_n\}_{n \in \mathbb{Z}}$ for the observation at each time n.

The general properties of HMMs allow one to calculate any statistic about the generated process from the hidden-state-to-state transition matrix T and set $\mathcal{P}$ of conditional probability density functions. For simplicity in the following, assume a finite set of hidden states and a single attracting component. Then every transition matrix T admits a unique stationary distribution $\boldsymbol{\pi}$. This is determined as T's left eigenvector associated with the eigenvalue of unity: $\langle \boldsymbol{\pi} | T = \langle \boldsymbol{\pi} |$. The eigenvector is normalized in probability: $\langle \boldsymbol{\pi} | \mathbf{1} \rangle = 1$, where $| \mathbf{1} \rangle$ is the column vector of all ones. Note also that $| \mathbf{1} \rangle$ is the right eigenvector of T associated with the eigenvalue of unity, $T | \mathbf{1} \rangle = | \mathbf{1} \rangle$. This property conserves state probability in hidden Markov chain evolution.

We can now provide the correlation functions and power spectral density in general and in closed form for the entire class of stochastic process generated by finite-state HMMs. Helpfully, for particular HMMs, the expressions become analytic in the model parameters.

Appendix B shows that the autocorrelation function is given by:

$$
\gamma(\tau) = \begin{cases} \langle \boldsymbol{\pi} | \Omega \, T^{|\tau|} \, \overline{\Omega} | \mathbf{1} \rangle & \text{if } \tau \leqslant 1 \\ \langle |x|^2 \rangle & \text{if } \tau = 0, \\ \langle \boldsymbol{\pi} | \overline{\Omega} \, T^{|\tau|} \, \Omega | \mathbf{1} \rangle & \text{if } \tau \geqslant 1 \end{cases} \tag{4}
$$

where Ω is the $|\mathcal{S}|$-by-$|\mathcal{S}|$ *average-observation matrix* defined by:

$$\Omega = \sum_{s \in \mathcal{S}} \langle X \rangle_{\mathrm{p}(X|s)} |s\rangle\langle s|. \qquad (5)$$

We use the hidden-state basis in which $|s\rangle$ is the column vector of all 0s except for a 1 at the index corresponding to state s. $\langle s|$ is simply its transpose. This yields a natural decomposition of the identity operator: $I = \sum_{s \in \mathcal{S}} |s\rangle\langle s|$. In the hidden-state basis, then, the Ω matrix simply places state-conditioned average outputs along its diagonal.

The power spectrum is calculated starting from Eq. (3) together with Eq. (4), using the spectral decomposition techniques developed for nonnormal and nondiagonalizable operators in Ref. [68]. In the derivation it is important to treat individual eigenspaces separately, as our generalized framework naturally accommodates. Appendix C gives the derivation's full details. Qualitatively, the power spectrum decomposes naturally into a discrete part $P_d(\omega)$ (a weighted sum of Dirac δ functions) and a continuous part $P_c(\omega)$ (a collection of diffuse peaks):

$$P(\omega) = P_c(\omega) + P_d(\omega).$$

For the power spectrum's continuous part the end result is as follows:

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \operatorname{Re}\langle \boldsymbol{\pi} | \overline{\Omega}\, T\, (e^{i\omega} I - T)^{-1} \Omega |\mathbf{1}\rangle, \quad (6)$$

where $\operatorname{Re}(\cdot)$ denotes the real part of its argument.

Remarkably, *all* of the ω dependence is in the apparently simple term $(e^{i\omega} I - T)^{-1}$. This is the *resolvent* of T along the unit circle in the complex plane. However, and central to our main results, this frequency dependence is *filtered* through $\langle \boldsymbol{\pi} | \overline{\Omega}$ and $\Omega |\mathbf{1}\rangle$. Notably, if the average-observation matrix was proportional to the identity, then all frequency dependence would be lost since $\operatorname{Re}\langle \boldsymbol{\pi} |(e^{i\omega} I - T)^{-1}|\mathbf{1}\rangle = -1/2$ is independent of ω [69]. Frequency dependence of the power spectrum thus requires that there are different averages associated with different states. Surprisingly though, none of the structure of the conditional probability density functions $\{\mathrm{p}(X|s)\}_s$ matters for the power spectrum, except for the average value observed in each state. Structure beyond averages is simply not captured.

D. Apparent structure

To fully appreciate the structure that *is* captured by the power spectrum requires a spectral decomposition of the transition matrix. The set Λ_T of T's eigenvalues is calculated as usual. However, since transition matrices are generically nonnormal and often nondiagonalizable, the spectral projection operators associated with T deserve a brief review.

In particular, the spectral projection operator T_λ associated with eigenvalue λ can be defined as the residue of $(zI - T)^{-1}$ as $z \to \lambda$:

$$T_\lambda = \frac{1}{2\pi i} \oint_{C_\lambda} (zI - T)^{-1}\, dz, \qquad (7)$$

where $z \in \mathbb{C}$ and C_λ is a small counterclockwise contour around the eigenvalue λ. Alternatively, the spectral projection

operators can be constructed from all left eigenvectors, generalized left eigenvectors, right eigenvectors, and generalized right eigenvectors associated with λ. The construction is given explicitly in Ref. [68]. In the simple case where the eigenvalue is nondegenerate, the eigenprojector takes on the simple form:

$$T_\lambda = \frac{|\lambda\rangle\langle\lambda|}{\langle\lambda|\lambda\rangle}.$$

However, the left $\langle\lambda|$ and right $|\lambda\rangle$ eigenvectors are *not* simply complex-conjugate transposes of each other, as they would be in the normal-operator case familiar from closed quantum systems and undirected networks. For example, the spectral projection operator associated with stationarity—$T_1 = |\mathbf{1}\rangle\langle\boldsymbol{\pi}|$—can be interpreted as the classical version of a density matrix but, typically, the stationary distribution is not uniform and so $\langle\boldsymbol{\pi}|$ is not proportional to the transpose of $|\mathbf{1}\rangle$.

We will also use the broader class of spectral companion operators:

$$T_{\lambda,m} = T_\lambda (T - \lambda I)^m. \qquad (8)$$

They have the useful property that $T_{\lambda,m} T_{\zeta,n} = \delta_{\lambda,\zeta} T_{\lambda,m+n}$. Clearly, the spectral projection operator is contained in this set, as $T_\lambda = T_{\lambda,0}$. It should be noted that $T_{\lambda,m} = \mathbf{0}$ for $m \geqslant \nu_\lambda$, where ν_λ is the *index* of the eigenvalue λ—i.e., the size of the largest Jordan block associated with λ. One should keep in mind that the transition matrix can be represented as:

$$T = \sum_\lambda (\lambda T_{\lambda,0} + T_{\lambda,1}).$$

While the resolvent has the general spectral decomposition:

$$(zI - T)^{-1} = \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} \frac{1}{(z-\lambda)^{m+1}} T_{\lambda,m}. \qquad (9)$$

The spectral expansion of the resolvent given by Eq. (9) allows us to better interpret the qualitative shape of the power spectrum Eq. (6):

$$P_c(\omega) = \langle |x|^2 \rangle + \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} 2 \operatorname{Re}\frac{\langle \boldsymbol{\pi} | \overline{\Omega}\, T\, T_{\lambda,m}\Omega |\mathbf{1}\rangle}{(e^{i\omega} - \lambda)^{m+1}}. \qquad (10)$$

Note that $\langle \boldsymbol{\pi} | \overline{\Omega}\, T\, T_{\lambda,m}\Omega |\mathbf{1}\rangle$ is a complex-valued scalar and all of the frequency dependence now handily resides in the denominator. When T is diagonalizable, Eq. (10) reduces to:

$$P_c(\omega) = \langle |x|^2 \rangle + \sum_{\lambda \in \Lambda_T} 2 \operatorname{Re}\left(\frac{\lambda \langle \boldsymbol{\pi} | \overline{\Omega}\, T_\lambda \Omega |\mathbf{1}\rangle}{e^{i\omega} - \lambda} \right).$$

The discrete (δ function) portion of the power spectrum is

$$P_d(\omega) = \sum_{k=-\infty}^{\infty} \sum_{\substack{\lambda \in \Lambda_T \\ |\lambda|=1}} 2\pi\, \delta(\omega - \omega_\lambda + 2\pi k) \operatorname{Re}\langle \boldsymbol{\pi} | \overline{\Omega}\, T_\lambda \Omega |\mathbf{1}\rangle, \quad (11)$$

where ω_λ is related to λ by $\lambda = e^{i\omega_\lambda}$. Equation (11) is valid even when T is nondiagonalizable: An extension of the Perron-Frobenius theorem guarantees that T's eigenvalues on the unit circle have index $\nu_\lambda = 1$. With $T_1 = |\mathbf{1}\rangle\langle\boldsymbol{\pi}|$, it is useful to note that $\langle \boldsymbol{\pi} | \overline{\Omega}\, T_1 \Omega |\mathbf{1}\rangle = |\langle x\rangle|^2$, so that the δ function at zero frequency appears whenever the average observation is nonzero.
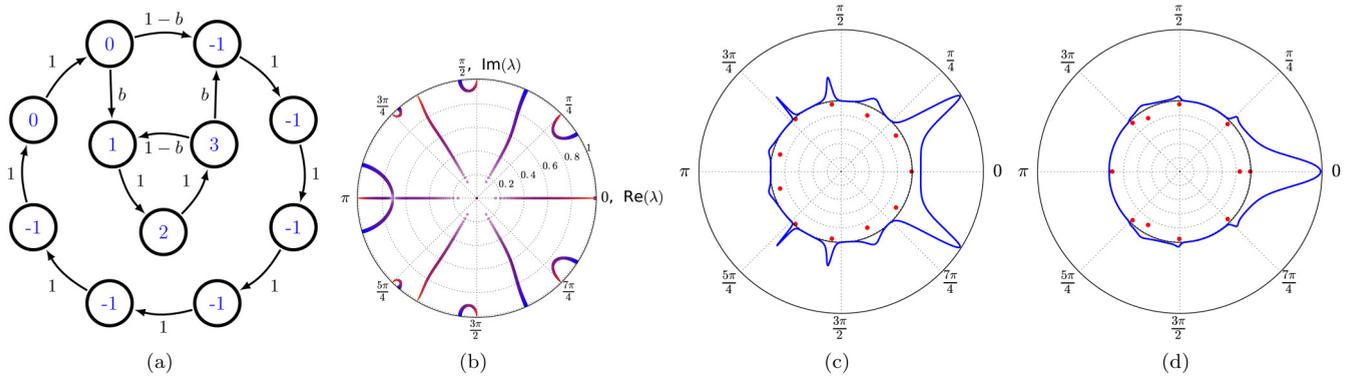
FIG. 4. Parametrized HMM of a stochastic process, its eigenvalue evolution, and two coronal spectrograms showing power spectra emanating from eigenspectra. (a) A b-parametrized HMM with mean values of each state's pdf $\langle x \rangle_{\mathrm{p}(X|s)}$ indicated as the number inside each state. (b) Eigenvalue evolution for all $\lambda \in \Lambda_T$ sweeping transition parameter b from 1 (thick blue) to 0 (thin red). (c) Power spectrum (blue line) and eigenvalues (red dots) at $b = 3/4$. (d) Power spectrum and eigenvalues at $b = 1/4$.

When plotted as a function of the angular frequency ω around the unit circle, the power spectrum suggestively appears to emanate from the eigenvalues $\lambda \in \Lambda_T$ of the hidden linear dynamic. This is illustrated by the *coronal spectrograms* in Figs. 4(c) and 4(d); these are discussed once the general phenomenon is explained.

T's eigenvalues *on* the unit circle yield Dirac δ functions in the power spectrum. T's eigenvalues *within* the unit circle yield more diffuse line profiles, increasingly diffuse as the magnitude of the eigenvalues retreats toward the origin. Moreover, the integrated magnitude of each contribution is determined from the amplitude $\langle \pi | \overline{\Omega} \, T_\lambda \Omega | \mathbf{1} \rangle$. Finally, we note that nondiagonalizable eigenmodes yield qualitatively different line profiles.

The spectral decomposition of the power spectrum offers several insights into the minimal temporal structure required to generate the observed power spectrum. In particular, since (i) each local maxima in the power spectrum emanates from an eigenvalue of the hidden state-to-state transition matrix and (ii) since the number of unique eigenvalues is upper bounded by the number of hidden states (i.e., $|\Lambda_T| \leqslant |\mathcal{S}|$), we have the following result: Counting both diffuse peaks and δ functions, the number of observed peaks in the power spectrum (from $\omega \in (-\pi, \pi]$ in the discrete-time setting) puts a lower bound on the number of hidden states of *any* model capable of generating the observed stochastic process. Note further that all transition matrices must have an eigenvalue of unity and that this eigenvalue can only produce a δ function at $\omega = 0$ with no other way to shape the power spectrum over other frequencies. This gives the immediate consequence that all single-state HMMs (i.e., all IID processes) have a flat power spectrum, as suggested earlier. In such cases, $\Lambda_T = \{1\}$, and there are no other eigenvalues to shape the power spectrum.

Figure 4 shows the power spectrum of a particular parametrized family of stochastic processes. Figure 4(a) displays the HMM's skeleton with state-to-state transition probabilities parametrized by b. The mean values $\langle x \rangle_{\mathrm{p}(X|s)}$ observed from each state are indicated as the blue number inside each state. The process generated depends on the actual PDFs and the transition parameter b. Although, and this is one of our main points, the power spectrum is ignorant to the PDFs' details.

The evolution of the eigenvalues Λ_T of the hidden-state transition dynamic is shown from thick blue to thin red markers in Fig. 4(b), as we sweep the transition parameter b from 1 to 0. A subset of the eigenvalues pass continuously but very quickly through the origin of the complex plane as b passes through $1/2$. The continuity of this is not immediately apparent numerically but can be revealed with a finer increment of b near $b \approx 1/2$. Notice the persistent eigenvalue of $\lambda_T = 1$, which is guaranteed by the Perron-Frobenius theorem.

Using coronal spectrograms, introduced in Refs. [70] and [67], Figs. 4(c) and 4(d) illustrate how the observed power spectrum $P(\omega)$ emanates from the *eigen*spectrum Λ_T of the hidden linear state dynamic. Specifically, in Fig. 4(c) and again, at another parameter setting, in Fig. 4(d), we show the power spectrum $P(\omega)$ (plotted around the unit circle in solid blue) and the eigenspectrum Λ_T (plotted as red dots on and within the unit circle) of the state-to-state transition matrix for the 11-state hidden Markov chain [Fig. 4(a)] that generates it. As anticipated from Eq. (10), the power spectrum has sharper peaks when the eigenvalues are closer to the unit circle. The integrated magnitude of each peak depends on $\langle \pi | \overline{\Omega} | \lambda \rangle \langle \lambda | \Omega | \mathbf{1} \rangle$.

It is easy to verify for this example that the stationary distribution $\langle \pi |$ is uniform for any $b \in (0, 1]$ and that there is no δ function at zero frequency since the average observation is zero. Nevertheless, as $b \to 1$, 10 δ functions (with five different integrated magnitudes) emerge (per 2π band of angular frequency) as the nonunity eigenvalues of the transition matrix approach the points $\{e^{i2n\pi/11}\}_{n=1}^{10}$ on the unit circle. At $b = 1$, the power spectrum is (up to a constant offset) the same as its discrete part: $P(\omega) = P_d(\omega) + \text{const}$, whereas for $b \in (0, 1)$, the power spectrum is diffuse and is the same as its continuous part: $P(\omega) = P_c(\omega)$.

Interestingly, our continuous *power* spectrum is the shadow of the discrete *eigen*spectrum of nonunitary dynamics. (The former is closely related to the continuous eigenspectrum of unitary models of chaotic dynamics.) This suggests that resonances in various physics domains concerned with a continuous spectrum can be modeled as consequences of simpler nonunitary dynamics. Indeed, hints of this already appear in Refs. [71–73].

While we frame our main results in terms of HMMs, in fact, they apply broadly to regularly observed physical systems. Many physical systems have exact representations as finite latent-state models, as in the examples of Sec. IV. However, even when the mapping is not exact, most if not all dynamical systems encountered in physics can be approximated to an arbitrary accuracy by either an autonomous or an input-dependent HMM [66,74]. *The eigendecomposition then serves to re-express the physical system and its power spectrum in its natural state space.*

The interpretation for discrete-state physical systems is obvious. While there are additional mathematical nuances with a continuous state space, the overall picture remains intact [75]. Specifically, Eq. (10) (and our subsequent analysis) applies to most dynamical systems encountered in physics—including quantum systems represented in Liouville space [76]—since these dynamical systems have a countable number of discrete eigenmodes.

E. Continuous-time processes

For both simplicity and generality, we focused on discrete-time dynamics [77]. However, correlation and power spectra are often applied to continuous-time processes. This section makes a more explicit connection to continuous-time processes and points out important features.

First, continuous-time processes are typically observed not continuously but periodically at some sampling frequency f_0. The duration $\tau_0 = 1/f_0$ between observations thus induces a discrete-time transition operator T_{τ_0} between states in that time interval. In such cases, the discrete-time transition matrix is related to the continuous-time generator G of time evolution by $T_{\tau_0} = e^{\tau_0 G}$. Accordingly, the continuous-time generator can be obtained from the discrete-time dynamic via $G = f_0 \ln T_{\tau_0}$ [78]. And the eigenvalues of T_{τ_0} and G are simply related by $\Lambda_{T_{\tau_0}} = \bigcup_{\zeta \in \Lambda_G} \{e^{\tau_0 \zeta}\}$ [79].

1. Autocorrelation and power spectra

Continuous-time representations can be analyzed directly, though. Consider the generic case of a continuous-time dynamic over a hidden state-space, with two types of example in mind:

(1) The system evolves through a continuous state-space. This describes both typical linear and nonlinear systems, including chaotic dynamical systems and Fokker-Planck dynamics. Then G is the generator that induces the finite-time Ruelle-Perron-Frobenius operator.

(2) Or observations are functions of a finite-state space with continuous-time transition rates. An example is current flowing or not, depending on the conformation of a biomolecular ion channel. Then G is the rate matrix of the master equation.

These different settings have the same expression for the autocorrelation and power spectrum. We now give these in closed form.

For real-valued $\tau > 0$, the autocorrelation is

$$\gamma(\tau) = \langle \overline{X}_t X_{t+\tau} \rangle = \langle \boldsymbol{\pi} | \overline{\Omega} \, e^{\tau G} \, \Omega | \mathbf{1} \rangle. \tag{12}$$

From this, we derive the continuous part of the power spectrum $P_c(f)$ with respect to frequency $f \in \mathbb{R}$, with the result

that:

$$P_c(f) = 2 \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \, (2\pi i f I - G)^{-1} \Omega | \mathbf{1} \rangle. \tag{13}$$

Appealing to the resolvent's spectral expansion again allows us to better understand the possible shapes of the power spectrum:

$$P_c(f) = \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_\lambda - 1} 2 \operatorname{Re} \frac{\langle \boldsymbol{\pi} | \overline{\Omega} \, G_{\lambda,m} \Omega | \mathbf{1} \rangle}{(2\pi i f - \lambda)^{m+1}}. \tag{14}$$

Since all of the frequency dependence is isolated in the denominator and since $\langle \boldsymbol{\pi} | \overline{\Omega} \, G_{\lambda,m} \Omega | \mathbf{1} \rangle$ is a frequency-independent complex-valued constant, peaks in $P_c(f)$ arise only via contributions of the form $\operatorname{Re} \frac{c}{(2\pi i f - \lambda)^n}$ for $c \in \mathbb{C}$, $f \in \mathbb{R}$, $\lambda \in \Lambda_G$, and $n \in \mathbb{Z}_+$.

2. Applications

Equation (14) helps explain the shapes of power spectra of chaotic dynamical systems, as appeared some time ago, e.g., in Ref. [11]. In that case, the eigenvalues of the time-evolution operator—whether the Ruelle-Perron-Frobenius transfer operator or the Koopman operator [80]—are known as Ruelle-Pollicott resonances [12–14], and $\langle \boldsymbol{\pi} |$ is the stationary distribution on the attractor. Stochastic differential equations leading to Fokker-Planck dynamics, ubiquitous in statistical physics, also obey Eq. (14). In these cases, the spectral projection operators describe the decay modes of probability densities on the continuous state space.

Even when the exact operator for time evolution is unknown, Eq. (14) can be used for the inverse problem of inferring the hidden linear dynamic from data—since the empirical power spectrum constrains the system's eigenspectrum.

It should be noted, however, that power spectra obtained either experimentally or numerically at finite sampling rate can deviate significantly from Eq. (14) as $f \to f_0/2$. Equation (14) only describes the empirical power spectrum of continuous-time processes for frequencies much less than the sampling frequency such that $f/f_0 \ll 1$, whereas Eq. (10) describes the empirical power spectrum exactly over all frequencies. The empirical power spectrum will approach Eq. (14) over any finite frequency band as the sampling frequency is increased, coinciding in the limit that $f_0/f \to \infty$.

3. Lorentzians and $1/f$ noise

When $c_\lambda \equiv \langle \boldsymbol{\pi} | \overline{\Omega} \, G_{\lambda,0} \Omega | \mathbf{1} \rangle$ is real valued, then the eigenmode's contribution to the power spectrum is c_λ times a Cauchy-Lorentz distribution over frequencies, centered at $f = \operatorname{Im}(\lambda)/2\pi$ with full width at half maximum of $\operatorname{Re}(\lambda)/\pi$. This becomes a delta function in the limit $\operatorname{Re}(\lambda) \to 0$. It is notable that nondiagonalizable eigenmodes contribute qualitatively distinct line profiles to the power spectrum.

Still one may wonder—since Eq. (14) is fully general for continuous-time dynamics—where the commonly encountered feature of $1/f$ noise could possibly originate. Inspired by Bernamont's 1937 insight that superposed Lorentzians can lead to $1/f$ noise [81], we can identify a source of $1/f$ noise in our more general setting.

Definition 1. A continuous-time process has *doubly harmonic diminution* if its:

(i) time-evolution generator G is diagonalizable and has $N + 1$ evenly spaced eigenvalues along the real line $\Lambda_G = \{-na\}_{n=0}^{N}$ for some $a > 0$ and

(ii) spectral intensity fades with increasing frequency according to $c_{-na} = c/n$ for $n \geqslant 1$ and some $c \in \mathbb{R}$.

Appendix D shows that any process with doubly harmonic diminution produces $1/f$ noise over a frequency bandwidth proportional to N, such that:

$$P(f) \sim \begin{cases} \text{constant} & \text{if } f < 3a/2\pi^2 \\ 1/f & \text{if } 3a/2\pi^2 < f \lesssim aN/4\pi^2 \\ 1/f^2 & \text{if } f \gtrsim aN/4\pi^2 \end{cases}.$$

Note that the power spectrum's $1/f$ portion can start at very low frequencies if a is small.

The surprising prevalence of $1/f$ noise in nature can now be reframed in light of our spectral results: Why would doubly harmonic diminution be so common in nature? We suggest that doubly harmonic diminution is a consequence of common motifs of causal dependence in processes. These dependencies impose structural constraints on transition rate matrices that could characteristically shape their spectral properties. Hopefully, this spectral reframing of $1/f$ noise will stimulate further attempts to explain its ubiquity.

F. Transducing structured noise

For certain dynamics, it is profitable to split the generator into deterministic and random components. This is especially useful when a linear time-invariant (LTI) system takes the structured noise as input. Random thermal motion in a harmonic trap is a simple example.

When a LTI system transduces structured noise—taking process X to process Y—the output is generically a simple transformation of the noise, modulated by the square magnitude of the LTI system's transfer function, $H_{X \rightarrow Y}(\omega)$ or $H_{X \rightarrow Y}(f)$ [82]. In discrete time the power spectrum is

$$P_{YY}(\omega) = |H_{X \rightarrow Y}(\omega)|^2 P_{XX}(\omega). \tag{15}$$

This requires modification, however, when the eigenvalues of the noise coincide with the poles and zeros of the LTI system's transfer function.

Consider a LTI system described by polynomials $\mathcal{P}(\mathcal{D})$ and $\mathcal{Q}(\mathcal{D})$ of either the discrete-time delay operator (i.e., $\mathcal{D}Y_t = Y_{t-1}$) or the continuous-time differential operator (i.e., $\mathcal{D}Y_t = \frac{d}{dt}Y_t$) such that:

$$\mathcal{P}(\mathcal{D})Y_t = \mathcal{Q}(\mathcal{D})X_t.$$

Then the square magnitude of the transfer function is given by:

$$|H_{X \rightarrow Y}(\omega)|^2 = \frac{|\mathcal{Q}(e^{i\omega})|^2}{|\mathcal{P}(e^{i\omega})|^2}$$

or

$$|H_{X \rightarrow Y}(f)|^2 = \frac{|\mathcal{Q}(i2\pi f)|^2}{|\mathcal{P}(i2\pi f)|^2}$$

for discrete-time or continuous-time models, respectively. In particular, X_t can be generated from a noise model that can be any HMM type discussed here.

For example, each spatial dimension of a Brownian trajectory simply integrates a white noise X_t according to the finite-difference equation: $Y_t - Y_{t-1} = X_t$. Appendix E shows this leads to the well-known power spectrum of Brownian noise $\sim 1/f^2$ in the limit of $f_0/f \rightarrow \infty$ and gives the correction for finite sampling rates. More generally, Eq. (15) can be used to evaluate the power spectrum from Langevin-type differential equations that transduce arbitrarily sophisticated noise processes.

Notably, any noise structure not revealed by X_t's power spectrum $P_{XX}(\omega)$ remains veiled by $P_{YY}(\omega)$ after passing through any LTI system. This begs the question of what has been hidden.

III. HIDDEN STRUCTURE

In fact, quite a lot is hidden. Remarkably, the power spectrum generated by any hidden-Markov process with continuous random variables for the state observables is the same as that generated by a potentially much simpler process—one that is a function of the same underlying Markov chain that instead emits the *expectation value* of the state observable.

Theorem 1. Let $\mathcal{P} = \{p(X|s)\}_{s \in \mathcal{S}}$ specify any state-paired collection of probability density functions over the domain $\mathcal{A} \subseteq \mathbb{C}$. Let $\mathcal{B} = \{\langle X \rangle_{p(X|s)}\}_{s \in \mathcal{S}}$ and let $\mathcal{Q} = \{\delta(X - \langle X' \rangle_{p(X'|s)})\}_{s \in \mathcal{S}}$. Then the power spectrum generated by any hidden Markov model $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, T)$ differs at most by a constant offset from the power spectrum generated by the hidden Markov model $\mathcal{M}' = (\mathcal{S}, \mathcal{B}, \mathcal{Q}, T)$ that has the same hidden Markov chain but in any state $s \in \mathcal{S}$ emits, with probability 1, the state-conditioned expected value $\langle X \rangle_{p(X|s)}$.

Proof. From Eqs. (6) and (11), we see that $P_c(\omega) + P_d(\omega) - \langle |x|^2 \rangle$ depends only on T and $\{\langle X \rangle_{p(X|s)}\}_{s \in \mathcal{S}}$. Thus, all HMMs sharing the same T and $\{\langle X \rangle_{p(X|s)}\}_{s \in \mathcal{S}}$ have the same power spectrum $P(\omega) = P_c(\omega) + P_d(\omega)$, modulo a constant offset determined by differences in $\langle |x|^2 \rangle$.

Figure 5 demonstrates Theorem 1 for the power spectrum in Fig. 4(c).

One immediate consequence is the following.

Corollary 1. Any hidden Markov chain with any arbitrary state-paired collection of zero-mean distributions, i.e.:

$$\mathcal{P} \in \{\{p(X|s)\}_{s \in \mathcal{S}} : \langle X \rangle_{p(X|s)} = 0 \text{ for all } s \in \mathcal{S}\},$$

generates a flat power spectrum indistinguishable from white noise.

Proof. This follows immediately from Theorem 1 and the fact that the all-zero sequence has a power spectrum that is zero everywhere. Thus, the corresponding power spectrum of the actual process is a flat (nonzero) power spectrum of uniform height $\langle |x|^2 \rangle$.

We can relax the corollary to include cases where the state-conditioned PDFs are all equal to a potentially nonzero constant. Although, a δ function at zero frequency (of integrated magnitude equal to the square magnitude of the constant) will then be observed in addition to the flat power spectrum.
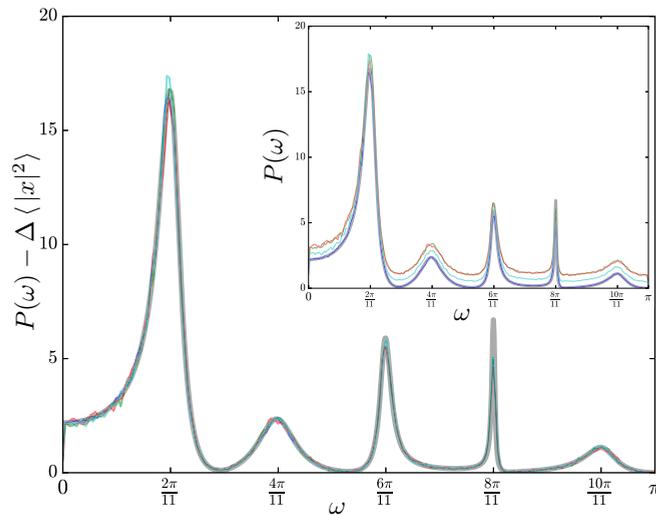
FIG. 5. Demonstrating Theorem 1 for the processes generated by the HMM skeleton of Fig. 4(a), using transition parameter $b = 3/4$ as in Fig. 4(c). Besides an overall constant offset of $\langle |x|^2 \rangle$, the power spectrum is insensitive to all details of the state-conditioned PDFs except for their averages. On top of the theoretical curve (thick gray) given by Eq. (6) we overlay horizontal offsets of the power spectra calculated numerically for stochastically generated time series. The state-conditioned PDFs used to define the different stochastic processes are as follows: (i) single δ functions, (ii) single Gaussians, (iii) two symmetrically spaced δ functions (with no support at the mean), and (iv) weighted δ functions with asymmetric spacing. For each, a time series of length 2^{18} was generated. The Welch method was used to calculate the average power spectrum for each process using FFTs of segments of length 2^9. The inset shows the raw power spectrum for each process without the offset.

The corollary's implications are striking. It is quite surprising, to consider one broad class of examples, that a power spectrum can be completely flat even when a ring of sequential states are visited that emit observables with probability density functions having no overlapping support. Figure 6 gives an example. In such a case, any cogent observer immediately detects the obvious structure in the mismatched supports—observed values are distinct—and forbidden realizations. Yet the power spectrum remains silent about this structure, reporting only the featureless signature of white noise.

In other more challenging settings, structure is not always so obvious without a reliable aid. Indeed, structure becomes increasingly difficult to detect (by any means) when the state-conditioned probability density functions have overlapping support. This is the generic case of non-Markovian processes. The hidden states cannot be detected via casual inspection.

While they give a concrete sense of missing structure, these cases fall far short of telling the full story of how power spectra mask structure. The following sections, culminating in Theorem 2, address additional ways white noise appears without needing to meet the requirements of Corollary 1.

A. Nonlinear pairwise correlation

In a sense, the structure of the stochastic process in Fig. 6 was hidden as shallowly as possible to evade appearing in
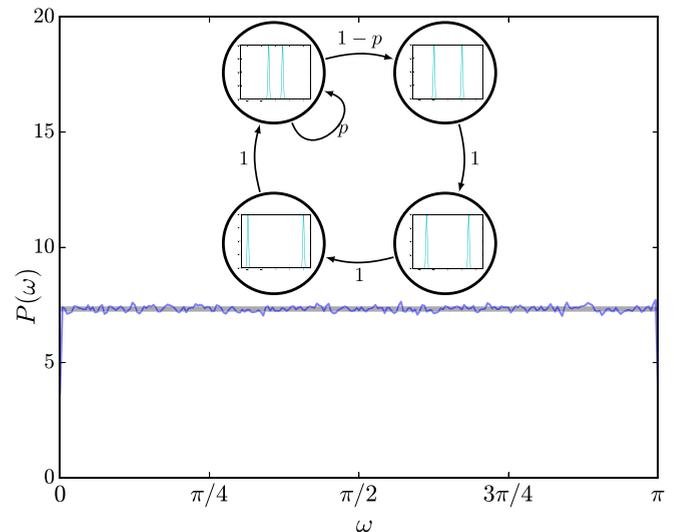
FIG. 6. Demonstrating Corollary 1 on the *noisy phase-slip process*: The overtly structured stochastic process generated by the HMM (inset) has a flat power spectrum for all values of the phase-slip transition parameter $p \in [0, 1]$. The flat power spectrum is shown analytically (thick gray) and numerically (thin blue) for $p = 1/10$. The numerical power spectrum was calculated from a simulated time series of length 2^{20} using the Welch method, performing FFTs on segments of length 2^9.

the power spectrum. As mentioned, the structure should be trivial to detect by other means. Indeed, while the *linear* pairwise correlation $\gamma(\tau)$ vanished for all $\tau > 0$, there is still pairwise dependence between the generated random variables, which is nonlinear. This pairwise dependence can be teased out using the pairwise mutual information $I(X_0; X_\tau)$ between observables at different times [83]. For the process of Fig. 6, if we take the limit of the narrow Gaussians in the state-conditioned PDFs to be pairs of δ functions, then the pairwise mutual information can be calculated exactly as shown in Appendix L. In fact, $I(X_0; X_\tau)$ will be unchanged for any set of four PDFs we could have chosen for the states of the example HMM, as long as the PDFs all have mutually exclusive support for the observable output. (This then makes the hidden state a function of the instantaneous observable.)

A concise summary of the pairwise mutual information is provided via Ref. [67]'s *power-of-pairwise-information* (POPI) spectrum:

$$\mathcal{I}(\omega) = -H(X_0) + \lim_{N \to \infty} \sum_{\tau = -N}^{N} e^{-i\omega\tau} I(X_0; X_\tau),$$

where $H(\cdot)$ is the Shannon entropy of its argument [83]. Examining the pairwise mutual informations and the POPI spectrum for this example (see Figs. 14 and 15), we find the decay of pairwise information to scale intuitively with the phase-slip-parameter p. While Fig. 6's example has no *linear* correlation, nevertheless it *does* have pairwise structure. Thus, the structure of the example process was hidden from power spectra, but not hidden from the POPI spectrum.

The following sections continue investigating temporally structured processes but focus on those with no linear pairwise

correlation (and so a flat power spectrum) and no pairwise mutual information (and so a flat POPI spectrum). These will lead us to introduce a general condition for flat power spectra. And, since power spectra fail so often to detect structure, we turn from criticizing them to being constructive: introducing ways to detect hidden structure.

B. Sophisticated fraudulent white noise

Theorem 1 established that the power spectrum from processes with *continuous* observable random variables is the same as the power spectrum from much simpler corresponding processes with *discrete* observable random variables. Accordingly, Theorem 1 motivates studying the power spectra of processes with discrete observable random variables to determine if there are further ways to achieve a flat power spectrum, beyond Corollary 1's possibilities. For observables that are discrete random variables, it is sufficient to consider their probability *distributions* rather than their probability *density* functions.

We begin this next step of the development by establishing the following simple lemma:

Lemma 1. Any stochastic process (not necessarily stationary) with the *single-condition-independent property* (SCIP):

$$\Pr(X_t | X_{t'} = x) = \Pr(X_t) = \Pr(X_{t'}),$$

for all $x \in \mathcal{A}$ and all $t \neq t'$, generates a flat power spectrum, mimicking white noise.

Proof. See Appendix F.

SCIP processes not only have a flat power spectrum but also a flat POPI spectrum. SCIP implies $I(X_0; X_\tau) = 0$ for all $\tau \neq 0$ which, in turn, implies $\mathcal{I}(\omega) = 0$. These processes completely lack any pairwise correlation, whether linear or nonlinear.

Notably, Lemma 1 is not covered by Corollary 1, nor is Corollary 1 subsumed by Lemma 1. Accordingly, the following develops a single simple condition (culminating in Theorem 2) that covers all of these cases of fraudulent white noise.

Crucially, the class of potentially fraudulent-white-noise processes suggested by Lemma 1 is nontrivial. In addition to IID processes, this class includes non-Markovian processes that hide all of their structure beyond pairwise correlations.

The random-random-XOR process (RRXOR), discussed at length in Ref. [67], is an example. Over blocks of length 3, the first two bits are generated randomly from a uniform distribution and the third bit is then the logical XOR operation of the last two. Explicitly:

$$X_{3n+\phi} = \text{XOR}(X_{3n-2+\phi}, X_{3n-1+\phi}), \quad \text{whereas}$$

$$X_{3n-2+\phi} \sim \left(\tfrac{1}{2}, \tfrac{1}{2}\right) \quad \text{and}$$

$$X_{3n-1+\phi} \sim \left(\tfrac{1}{2}, \tfrac{1}{2}\right),$$

for all $n \in \{1, 2, \dots\}$. As a SCIP process, the RRXOR process has a flat power spectrum although it does not fall under the purview of Corollary 1. Indeed, the RRXOR process has no pairwise correlation at all since $I(X_0; X_\tau) = 0$ for all $\tau > 0$. Accordingly, the POPI spectrum is zero over all frequencies. The structure in this process is strictly *three-way* correlation. In Ref. [67], the phase ϕ itself is a random variable, and

synchronizing to the phase is a surprisingly difficult task [84]. No matter whether the phase ϕ is given, the process has no pairwise correlation—resulting in a flat power spectrum and flat POPI spectrum—and only reveals correlation in its three-way structure.

It is interesting to note that the related RRX*NOR* process, where $X_{3n} = \text{XNOR}(X_{3n-2}, X_{3n-1})$, also has a flat power spectrum. In fact, this suggests a new method to hide structure: embed a correlated message into a sequence of RRXOR and RRXNOR 3-bit sequences that lifts all correlation beyond pairwise. Specifically, the original message is transformed into a sequence of choices about whether to use XOR or XNOR on the previous two random bits. As long as the read frame and the embedding mechanism is known, the message can be easily extracted. But if it is not known that a message is embedded, then it cannot be detected simply by looking for pairwise correlations.

Through similar construction, structure can be shifted up to arbitrarily high orders of correlation. Stochastic processes can be constructed with N-way correlation but no n-way correlation for all $n < N$. Moreover, an arbitrarily correlated message can be embedded within such a process, such that its structure is lifted beyond any desired order of correlation.

C. Content-preserving whitening

Corollary 1 gave a method to construct an arbitrarily complex process with a truly flat power spectrum, so long as all hidden states have the same average output. Here, we introduce an alternate method to construct arbitrarily complex processes with truly flat power spectra. These processes, in addition, are devoid of n-way correlation for all $n < N$.

(1) Choose an embedding block length $N \geqslant 3$.

(2) Choose any stochastic process ("Process A") with a binary output alphabet.

(3) Construct "Process B" as follows:

(a) Whenever Process A would produce a 0, Process B will sample a word uniformly from the set of all words of length N with an even number of 1s.

(b) Whenever Process A would produce a 1, Process B will sample a word uniformly from the set of all words of length N with an odd number of 1s.

Any Process B constructed in this manner has a truly flat power spectrum. Process B will also be devoid of n-way correlation for all $n < N$. Moreover, if A is a stationary process such that its statistical complexity $C_\mu(A)$ is well defined [85,86], then Process B is also a stationary process with $C_\mu(B) \geqslant C_\mu(A)$.

This also works for "infinitely structured" processes, those with divergent statistical complexity. Choose *any* binary Process-A family with $C_\mu \to \infty$. This can be, for example, Ref. [87]'s *heavy-tailed periodic mixture process* that has infinite past-future mutual information: $\mathbf{E} \to \infty$. Then add some structure, via content-preserving whitening, to obtain a binary Process-B family with $C_\mu \to \infty$ and a truly flat power spectrum.

Similar constructions can also be developed for processes with larger alphabets.

Through the lens of pairwise correlation, such structure is simply missed. However, before moving on to consider
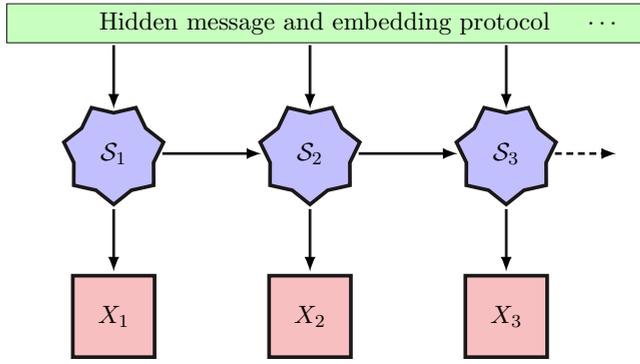
FIG. 7. Bayesian network for memoryful input-dependent generators.

more advanced methods to detect such structure, we finish our investigation of flat power spectra from structured processes. The next section addresses a broad class of possibly-input-dependent process generators and we give a very general condition for when a flat power spectrum results.

D. Input-dependent generators and fraudulent white noise

Probing fraudulent white noise more broadly, consider an arbitrarily correlated message $\vec{m}$ and an input-dependent generator $\mathcal{M}(\vec{m})$ of an observable output process $\{X_t\}_{t \in \mathcal{T}}$. The lengths of the inputs and outputs need not be commensurate, and the input and output alphabets may also be distinct. The generator is fully specified by the tuple $\mathcal{M}(\vec{m}) = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \{T_t(\vec{m})\}_t, \boldsymbol{\mu}_1)$. That is, the internal states $\mathcal{S}$, output alphabet $\mathcal{A}$, and state-dependent PDFs $\mathcal{P}$ are static. However, the hidden-state-to-state transition matrix $T_t(\vec{m})$ at time t is potentially a function of the full input $\vec{m}$. Since stationarity is no longer assumed, the initial distribution $\boldsymbol{\mu}_1$ over hidden states must be specified for the statistics of the output process to be well defined.

Figure 7 shows the relevant Bayes network for this general type of input-dependent generator. Contrast this with Fig. 3, which showed the Bayes network of autonomous HMM generators. Autonomous HMMs can be seen as a special case of these possibly-input-dependent generators when the process $\mathcal{M}(\vec{m}) = \mathcal{M}$ is input independent and the initial distribution $\boldsymbol{\mu}_1 = \boldsymbol{\pi}$ is taken to be the stationary distribution $\langle \boldsymbol{\pi}| = \langle \boldsymbol{\pi}|T$ of the time-independent transition matrix $T_t(\vec{m}) = T$.

The memoryful input-dependent generators we now consider also generalize the memoryful transducers introduced in Ref. [88] to use continuous-variable outputs and allow the lengths of input and output to be incommensurate. Via any of the above models, very general message-embedding schemes can be developed that produce sophisticated fraudulent white noise.

Even with all the generalizations, we can determine autocorrelation and power spectra. Similarly to the derivation for HMMs, we find that *if* the process is wide-sense stationary, then (for $\tau \geqslant 1$):

$$\gamma(\tau) = \langle \boldsymbol{\mu}_t | \overline{\Omega} \, T_{t:t+\tau}(\vec{m}) \, \Omega | \mathbf{1} \rangle, \tag{16}$$

which must overall be t-independent (so long as $t \geqslant 1$). Here $\langle \boldsymbol{\mu}_t | = \langle \boldsymbol{\mu}_1 | T_{1:t}(\vec{m})$ and $T_{a:b}(\vec{m}) = \prod_{t=a}^{b-1} T_t(\vec{m})$, and Ω is again given by Eq. (5). [Notice that $T_{a:a+\tau}(\vec{m}) = T^\tau$ for the special case of autonomous HMMs.]

Thus, autocorrelation for $\tau \geqslant 1$ can be calculated as $\langle \boldsymbol{\mu}_1 | \overline{\Omega} \, T_{1:1+\tau}(\vec{m}) \, \Omega | \mathbf{1} \rangle$, assuming that the pairwise statistics are stationary. This can also be written as:

$$\gamma(\tau) = \left\langle \langle \overline{x} \rangle_{\mathrm{p}(X|\mathcal{S}_t)} \langle x \rangle_{\mathrm{p}(X|\mathcal{S}_{t+\tau})} \right\rangle_{\mathrm{Pr}(\mathcal{S}_t, \mathcal{S}_{t+\tau})}, \tag{17}$$

where we treat $\langle x \rangle_{\mathrm{p}(X|\mathcal{S}_t)}$ as a random variable that depends on $\mathcal{S}_t$ and the whole expression becomes t-independent assuming stationary pairwise statistics. Accordingly, the autocorrelation function is constant and the power spectrum is flat whenever:

$$\mathrm{Pr}\left[\langle x \rangle_{\mathrm{p}(X|\mathcal{S}_{t+\tau})} | \mathcal{S}_t = s \right] = \mathrm{Pr}\left[\langle x \rangle_{\mathrm{p}(X|\mathcal{S}_{t+\tau})} \right] = \mathrm{Pr}\left[\langle x \rangle_{\mathrm{p}(X|\mathcal{S}_t)} \right],$$

for all τ, for all $t \in \mathcal{T}$, and for all $s \in \mathcal{S}$.

However, this requirement is too strict to cover all cases of interest. For example, it does *not* yet imply the flat power spectrum of the RRXOR process. More generally, constant autocorrelation and flat power spectra can be guaranteed by an even weaker condition.

To appreciate this, define the set Ξ of average outputs emitted by the states: $\Xi \equiv \bigcup_{s \in \mathcal{S}} \{ \langle x \rangle_{\mathrm{p}(X|s)} \}$. Furthermore, we define $\mathcal{S}_\xi \subset \mathcal{S}$ as the set of states that all emit the same average output $\xi \in \Xi$. Explicitly, $\mathcal{S}_\xi \equiv \{ s \in \mathcal{S} : \langle x \rangle_{\mathrm{p}(X|s)} = \xi \}$. Using these entities, we can state our result more precisely as the following theorem.

Theorem 2. Let $\{X_t\}_t$ be a stochastic process generated by any of the hidden-state models $\mathcal{M}(\vec{m})$ discussed above, including autonomous HMMs and input-dependent generators, X_t the random variable for the observable at time t, and $\mathcal{S}_t$ the random variables for the hidden state at time t. Such processes have *constant autocorrelation and a flat power spectrum* if:

$$\mathrm{Pr}(\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'} | \mathcal{S}_t \in \mathcal{S}_\xi) = \mathrm{Pr}(\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) = \mathrm{Pr}(\mathcal{S}_t \in \mathcal{S}_{\xi'}),$$

for all separations $\tau > 0$, for all $t \in \mathcal{T}$, and for all $\xi, \xi' \in \Xi$.

Proof. See Appendix G.

Theorem 2 says that a flat power spectrum results whenever the average output of the future hidden state is independent of the average output of the current hidden state.

This generalized condition for flat power spectra covers the special case for HMMs as well as fraudulent white noise from message-embedding schemes with stationary pairwise statistics, but nonstationary high-order statistics. Appendix M shows that a modified version of Theorem 2 also applies to another class of generators that can be more natural for measured quantum systems and systems with computational dependencies. Theorem 2 subsumes Corollary 1 as well as Lemma 1. And, it offers the most general guarantee yet for constant autocorrelation and flat power spectrum.

By way of contrast consider the following. While zero pairwise mutual information is always a sufficient condition for flat power spectrum, it is not a necessary condition. Here, in Theorem 2, we find a very general condition for a flat power spectrum. Appendix N established a related theorem (Theorem 5) that further generalizes the condition for flat power spectra, allowing for time-dependent PDFs associated with each state. Moreover, Theorems 2 and 5 constructively suggest how to design such processes. Notably, these

generalized conditions do not require a stationary dynamic over the hidden states of the observation-generating mechanism, which furthermore allows messages to hitchhike undetected aboard fraudulent white noise.

More broadly, we may ask when two processes generate the same power spectrum, whether or not it is flat.

Theorem 3. Let $\{X_t\}_t$ and $\{Y_t\}_t$ be two stochastic processes generated by any of the hidden-state models $\mathcal{M}(\vec{m})$ discussed above, including autonomous HMMs and input-dependent generators, X_t and Y_t the random variables for the observables at time t, and $\mathcal{S}_t \in \mathcal{S}$ and $\mathcal{R}_t \in \mathcal{R}$ the random variables for the respective hidden states at time t. These processes have *identical power spectra*, up to a constant offset, if:

$$\Pr(\mathcal{S}_t \in \mathcal{S}_\xi, \mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) = \Pr(\mathcal{R}_t \in \mathcal{R}_\xi, \mathcal{R}_{t+\tau} \in \mathcal{R}_{\xi'}),$$

for all separations $\tau > 0$, for all $t \in \mathcal{T}$, and for all $\xi, \xi' \in \Xi$.

Proof. See Appendix H.

Section IV B below leverages Theorem 3 to determine the degeneracy of diffraction patterns from distinct physical structures.

This suite of results emphasizes our main argument's generality: Power spectra are mute when detecting a broad range of observable structure. Whether observing physical, biological, or social systems, we seek structure that reveals mechanism and begets predictability. Through the lens of power spectra, or pairwise correlation more generally, much structure is simply missed. The challenge then is to look for structure beyond pairwise. Section V addresses this challenge shortly. First, though, to motivate the extra effort, we show that fraudulent white noise is indeed a feature of real physical systems.

IV. HIDDEN PHYSICAL STRUCTURE

To ground the theoretical consequences in natural, even familiar phenomena, this section takes on three rather disparate physical systems. It draws out important physical implications of fraudulent white noise and power spectral degeneracy in quantum entanglement, chaotic crystallography, and neural-membrane ion channels.

A. Fraudulent white noise from quantum entanglement

Correlated measurements of entangled quantum systems indelibly confirmed the reality of nonlocal physical states. In particular, Bell tests conclusively showed that no local hidden variable theory is consistent with certain strongly correlated observations [89–91]. Detecting correlation in more general quantum states should similarly yield a deeper appreciation of quantum correlation's important role in everything from thermodynamics [92] to gravity [93,94]. But what if our tools mask correlations?

Entangled many-body systems, as it turns out, easily generate fraudulent white noise when they are measured. The following demonstrates that repeated measurements of even quite simple entangled states leads to fraudulent white noise. As a consequence, one is at risk of inadvertently inferring randomness where there is essential correlation. Recognizing these high-order correlations in fraudulent white noise, in contrast, could reveal the ubiquity of entangle-

ment naturally induced in the time evolution of physical systems.

As a particular example, consider the entangled three-body quantum state:

$$|\Xi\rangle \equiv \tfrac{1}{2}(|000\rangle + |011\rangle + |101\rangle + |110\rangle),$$

where, for example, $|011\rangle = |0\rangle_A \otimes |1\rangle_B \otimes |1\rangle_C$. The quantum circuit diagram:
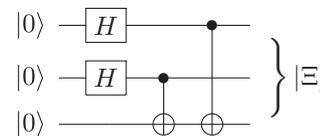
shows that $|\Xi\rangle$ is directly generated by a sequence of two Hadamard gates and two controlled-NOT gates applied to the unentangled state $|000\rangle$. Recall that the Hadamard gate H_A maps $|0\rangle_A$ to $|+\rangle_A \equiv (|0\rangle_A + |1\rangle_A)/\sqrt{2}$. (Reference [45] may be consulted for further explanation of these standard elements of quantum circuit diagrams.)

When measured in the *computational basis* of 0s and 1s, repeated preparation and measurement of $|\Xi\rangle$ states leads exactly to the RRXOR process discussed above, as the reader can directly verify [95]. This quantum preparation and measurement setup is shown explicitly in Fig. 8(d). Certainly, observations contain predictable correlations. A pairwise analysis of the observation sequence, however, gives the statistics of white noise. This holds whether the analysis used either power spectra or POPI spectra or, indeed, any analysis that can be performed in one-on-one meetings among Alice, Bob, and Charlie, who each hold one of the component qubits.

Figure 8 compares additional examples of stochastic processes generated by fixed measurement of unitarily transformed blank quantum inputs [96]. Panel 8(a) reminds us that almost any measurement of a quantum system yields some randomness. The amount of uncertainty, though, depends on how well the measurement basis aligns with the system's quantum state. However, Panel 8(b) reminds us that local properties of a maximally entangled state are maximally unpredictable, regardless of the local measurement basis. The entire structure of a maximally entangled state exists only nonlocally among constituents, yielding correlations when measurements on different parts of the system are compared—as in the Bell process of Fig. 8(c).

When the number of entangled parties is larger than two, correlation becomes much harder to detect. Nevertheless, in each case the physical input, unitary transformation, and measurement protocol together determine the HMM that exactly describes the correlated output process. Figure 8(d) shows how fraudulent white noise in the form of the RRXOR process can arise from measurements of entangled three-bodied

quantum systems. Moreover, adding two swap gates:
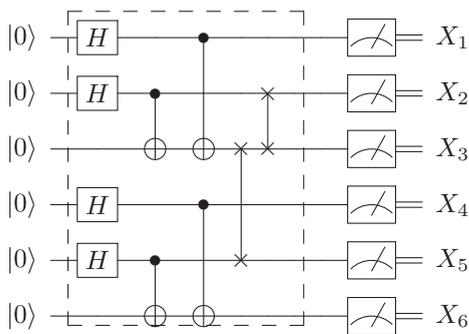
transforms the output into the even more cryptic Interlaced RRXOR process, discussed shortly in Sec. V B.

These constructions demonstrate that *simple sequences of two-body interactions can generate high-order correlations while revealing no low-order correlation whatsoever.*

B. Silent crystals

Many icons of natural structure ensue from the atomic placements encoded in crystals. Semiconductor crystal structures tell electrons and light how to move within them, while the aperiodic crystals of our genetic script instruct our cells how to behave. The typical way to probe crystal structure is x-ray diffraction—the power spectrum of a crystal's electron density. Thus, the preceding results on the degeneracy of power spectra highlight *which features of crystal structure can be inferred from diffraction patterns.*

Close-packed structures, which mimic the dense packing of hard spheres, offer an interesting case study due to their multiplicity and natural abundance [103,104]. All close-packed structures are composed of *modular layers* {A, B, C}, with a material-dependent basis attached to a two-dimensional (2D) hexagonal crystal lattice. Assembling these modular layers, there are two choices for how to nestle the next layer to fill the holes as tightly as possible. For a particular material, differences in diffraction patterns arise from this sequence of stacking choices [70,105–107].

Besides the ABABAB... period-two stacking of close-packed two-dimensional monolayers that leads to *hexagonal close packing* (hcp) and the ABCABC... period-three stacking of these monolayers that leads to *cubic close packing* (ccp), there is an infinite number of ways to stack the monolayers as tightly as possible. The only constraint is the stacking rule that no layer (whether A, B, or C) can appear twice in succession. Nature, it turns out, is fully aware of all the possibilities.

For close-packed materials, the net energy from nearest-neighbor interactions is indifferent to which of the infinitely many close-packed structures is realized. This facilitates great diversity, in both natural and fabricated materials, via *polytypism* and *random stacking* [108]. Prominent examples of polytypic layered structures include SiC, ZnS, stacked graphene, and ice [109]. Different polytypes of the same material can have very different electronic, optical, and mechanical properties [110].

1. Diffraction theory of layered structures

Appendix A reviews the basics of diffraction theory and shows that the diffracted intensity (as a function of the scattering vector $\vec{q}$) can be written as a power spectrum of layer *form factors* $X_n = F^{(n)}(\vec{q}) \in \mathbb{C}$. Each layer form factor is the Fourier transform of the spatially-extended scatterer density (e.g., the electron density) associated with the layer. In particular, the expected diffracted intensity can be written as:

$$\langle I_{\text{diff}}(\vec{q}) \rangle = cNP(\omega) = c \left\langle \left| \sum_{n=1}^{N} F^{(n)}(\vec{q}) e^{-i\omega n} \right|^2 \right\rangle, \quad (18)$$

where $\omega = \tau_0 \vec{q} \cdot \hat{\ell}$ quantifies the change in wave number along the stacking direction $\hat{\ell}$ of N sequential layers of thickness τ_0.

For typical layered structures, there is only a small number of layer types. For close-packed structures, to take one example, each layer realizes one of only three allowed relative offsets in its plane. Yet, in detail, we know that each layer type is subject to both thermal fluctuations and quantum uncertainty of atomic positions. What are the consequences for the diffraction pattern?

Suppose there is a hidden-state model $\mathcal{M}(\vec{m}) = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \{T_t(\vec{m})\}_t, \boldsymbol{\mu}_1)$ that generates the correct statistics of the layer form factors in the material—taking the stochastic stacking process, thermal motion, and quantum uncertainty into account. Theorems 1 and 3 imply that—up to a constant offset—*the diffraction pattern will be the same if we instead consider the much simpler hidden-state model* $\mathcal{M}'(\vec{m}) = (\mathcal{S}, \mathcal{B}, \mathcal{Q}, \{T_t(\vec{m})\}_t, \boldsymbol{\mu}_1)$ *that outputs only the* expected *layer form factor from each hidden state.*

Appendix A 1 shows that this allows us to easily and rigorously produce the surprising results of Debye-Waller theory for exactly periodic lattices in the general setting of randomly stacked structures. Specifically, the state-conditioned expectation value of the form factor directly leads to the Debye-Waller factor that exponentially suppresses the intensity of the diffraction pattern at large magnitude of the scattering vector. Surprisingly, however, thermal and quantum fluctuations do not lead to broadening of the diffraction pattern.

For close-packed structures, there are only three types of layers, differing only via relative displacements of $1/3$ of a lattice-translation vector in the plane of the layer. As a result, if type-A layers have an expected layer form factor of $A = \psi$, then the other layer types are simply related by the third roots of unity:

$$A = \psi, \quad B = \psi e^{i2\pi/3}, \quad \text{and} \quad C = \psi e^{-i2\pi/3}.$$

(See Appendix A 2.) The three possible state-conditioned expectation values for the layer form factors serve as the alphabet $\mathcal{B} = \{A, B, C\}$ for the stacking process $\{X_n\}_n$, where n indexes the layer and adjacent layers are separated by a distance of τ_0. The diffracted intensity from any close-packed structure is then given by the power spectrum of the stacking process.

Information about the stacking process is most directly revealed via $P(\omega)/|\psi|^2$ wherever $|\psi|^2$ is nonzero, which discounts the expected diffraction pattern of a single layer [111].

Traditional crystals are described by periodic patterns. Much more generally, crystal structure can be defined by the stochastic process that generates it. (Traditional crystals,
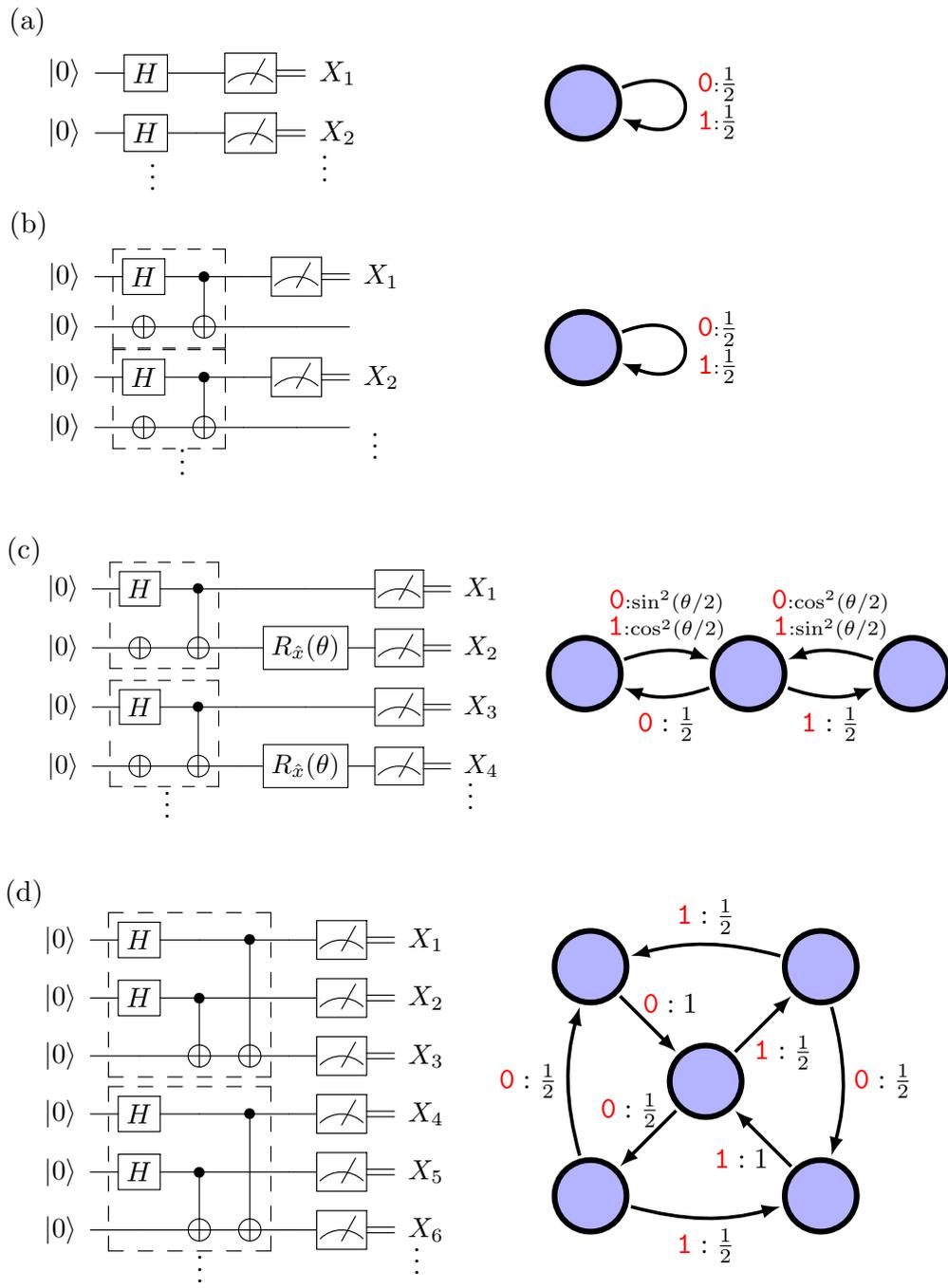
FIG. 8. Stochastic processes generated by fixed measurements of unitarily transformed blank quantum inputs. These include the following: (a) measurement-basis-dependent genuine white noise, (b) measurement-basis-independent uniform white noise, (c) a correlated Bell process, and (d) entanglement-enabled fraudulent white noise. Dashed boxes are drawn around the entangling unitary modules in each case, except (a), where there is no entanglement. The induced Mealy-type HMMs shown on the right are the minimal descriptors of the output process. The edge label "$\mathbf{x} : q$" on the transition from state s to s' indicates the joint probability $\Pr(X_t = \mathbf{x}, \mathcal{S}_{t+1} = s' | \mathcal{S}_t = s) = q$ of observing $\mathbf{x} \in \mathcal{A}$ and transitioning to s', given the current state s. Mealy-type HMMs are a simple case of the more general measurement feedback models discussed in Appendix M.

then, are the special case in which the stochastic stacking process is deterministic and periodic.) For close-packed structures layered according to a stochastic process that can be expressed by a hidden Markov model, our results imply that the diffraction spectrum is intimately related to the HMM's eigenspectrum.

2. Random stacking example

Both hcp and ccp crystals are described by very simple deterministic Markov models. More generally, crystal structure can integrate both features of randomness and features of determinism. Moreover, the randomness need not be simply statistically independent errors (possibly, *faults*) in an
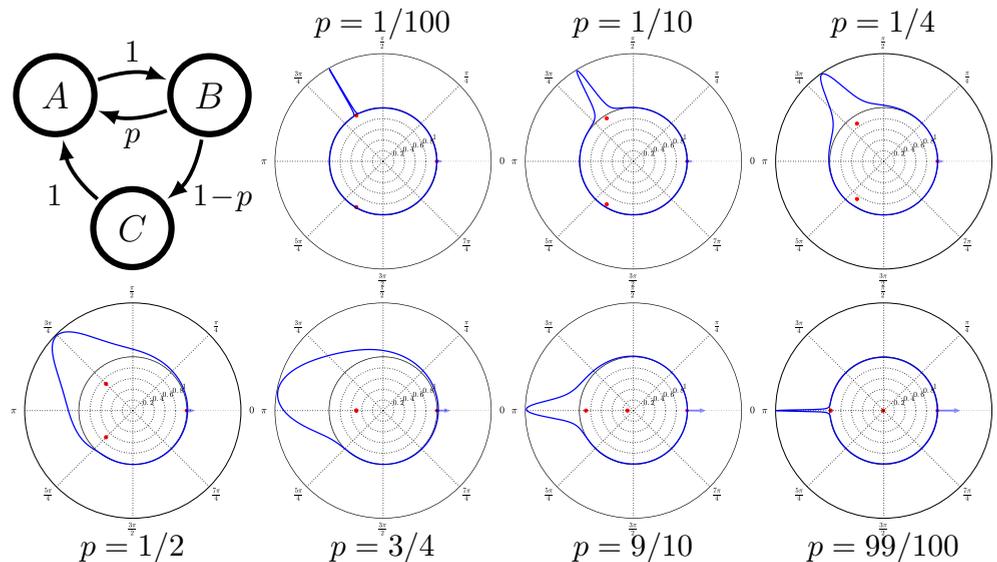
FIG. 9. Parametrized HMM that generates a family of stochastic stacking processes (top left) and the diffracted intensity $P(\omega)/|\psi|^2$ at different values of the faulting parameter p. Plotted as coronal spectrograms, it is clear that the diffraction spectrum emanates from the eigenspectrum of the HMM that generates the crystal.

otherwise periodic parent crystal. Rather, the randomness itself can have a rich causal architecture.

As a first example, consider the p-parametrized family of stochastic stacking processes depicted in Fig. 9. For $p = 1$, we recover the deterministic period-two hcp structure. The period-two nature is reflected in the Bragg reflection at $\omega = \pi$. For $p = 0$, we recover the deterministic period-three ccp structure. The period-three nature is reflected in the Bragg reflection at $\omega = 2\pi/3$. For other values of p, the structure is described by a stochastic stacking process.

For any p, the transition matrix and average-observation matrix are as follows:

$$T = \begin{bmatrix} 0 & 1 & 0 \\ p & 0 & 1-p \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \Omega = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix},$$

respectively. The transition-matrix eigenvalues are $\Lambda_T = \{1, -\frac{1}{2} \pm \sqrt{p - \frac{3}{4}}\}$. The transition matrix is diagonalizable unless $p = 3/4$, where it becomes nondiagonalizable.

For $p \neq 3/4$, each spectral projection operator is given by $T_\lambda = |\lambda\rangle\langle\lambda|$, with $\langle\lambda| = \frac{1}{3\lambda^2 - p}[\lambda \quad 1 \quad \lambda^2 - p]$ and $|\lambda\rangle = [\lambda \quad \lambda^2 \quad 1]^\top$, where $\top$ denotes transposition. Recall that the stationary distribution is the left eigenvector $\langle\boldsymbol{\pi}| = \langle 1| = \frac{1}{3-p}[1 \quad 1 \quad 1-p]$. From these elements, we can calculate $\langle\boldsymbol{\pi}|\overline{\Omega} T_\lambda \Omega|\mathbf{1}\rangle$ and the diffracted intensity analytically as a function of the transition parameter p. Appendix I gives the calculation details.

a. Bragg reflections without periodicity. For $p \in (0, 1)$, the transition matrix T only has a single eigenvalue on the unit circle, so the discrete (Bragg) spectrum has a single contribution from the eigenvalue of unity:

$$P_d(\omega) = \frac{2\pi p^2 |\psi|^2}{(3-p)^2} \sum_{\ell=-\infty}^{\infty} \delta(\omega + 2\pi\ell).$$

It is interesting that this Bragg reflection persists despite the lack of any long-range deterministic periodicities for $p \in (0, 1)$. This rather reflects a different type of long-range order: the persistent imbalance of layer types within each realization of the stochastic stacking process. More generally, Bragg reflections can be attributed to statistical symmetry breaking. Deterministic periodicities are but one special case.

b. Diffuse spectrum. There is a diffuse contribution to the power spectrum for all $p \in (0, 1)$. For $p \in (0, 3/4) \cup (3/4, 1)$, this contribution is as follows:

$$P_c(\omega) = |\psi|^2 \left[1 - \frac{p^2}{(3-p)^2} \right] + \sum_{\lambda \in \Lambda_T \setminus \{1\}} 2\operatorname{Re} \frac{\langle\boldsymbol{\pi}|\overline{\Omega} T_\lambda \Omega|\mathbf{1}\rangle}{e^{i\omega}/\lambda - 1}.$$

However, the expanded expressions are significantly different for $p > 3/4$, where all eigenvalues are real valued and distinct, and for $p < 3/4$, where two of the eigenvalues are complex conjugate pairs.

c. Nondiagonalizable diffraction profiles. At $p = 3/4$, the transition matrix of the stochastic stacking process becomes nondiagonalizable. Curiously, this nondiagonalizability is a generic feature of parametrized transition matrices at the point where real eigenvalues collide and interact to gain complementary imaginary components. That is, nondiagonalizability marks the onset of new behavior. In this case, nondiagonalizability marks the transition from primarily period-2 to primarily period-3 behavior. This critical point of nondiagonalizability is accompanied by a qualitatively distinct diffraction profile—no longer exhibiting the typical Lorentzian line profile. Observing such a line profile experimentally indicates a material at the crossroads of structural transformation.

3. Degenerate diffraction patterns

Our general results on the degeneracy of power spectra directly bear on the degeneracy of diffraction patterns from different crystals. The enhanced understanding of this
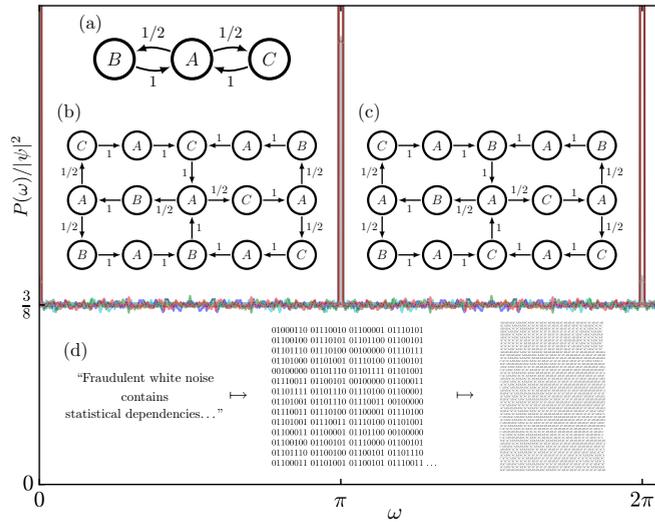
FIG. 10. Diffraction pattern (overall figure) consisting of a white noise background with two Bragg reflections per 2π change of angular frequency along the stacking direction. This pattern will be observed from infinitely-many distinct stochastic processes that generate close-packed structures. The flat diffraction pattern is given analytically (thick gray) by Eq. (19). We verify numerically that this diffraction pattern is observed from a crystal stacked according to the simple stochastic process of panel (a) (thin blue). The same diffraction pattern results from the stochastic processes of panels (b) (thin cyan) and (c) (thin green) that have distinct nontrivial higher-order correlations. And, the same diffraction pattern results also from a crystal that contains the information needed to faithfully reconstruct the entire contents of the present manuscript. (d) To demonstrate this, we extracted an extended excerpt from the manuscript, converted the text to binary ASCII, and then converted each binary character to six layers of the crystal—sampled from process-(b) if 0 and sampled from process-(c) if 1, starting in the central A state each time. The corresponding diffracted intensity is shown in thin red, coinciding with the others.

degeneracy, in turn, sheds new light on the well-known difficulty of the inverse problem of discovering crystal structure from diffraction patterns [112].

Consider a chaotic crystal with a stochastic stacking process described by the simple HMM shown in Fig. 10(a). The transition-matrix eigenvalues are $\Lambda_T = \{0, \pm 1\}$. Appendix I 2 shows that $\langle \boldsymbol{\pi} | \overline{\Omega}\, T_1 \Omega | \mathbf{1} \rangle = \frac{1}{16}|\psi|^2$ and $\langle \boldsymbol{\pi} | \overline{\Omega}\, T_{-1} \Omega | \mathbf{1} \rangle = \frac{9}{16}|\psi|^2$. The resulting diffraction pattern consists of a flat "white noise" background:

$$P_c(\omega) = \frac{3}{8}|\psi|^2,$$

together with two Bragg reflections per 2π of angular-frequency bandwidth:

$$P_d(\omega) = \frac{\pi|\psi|^2}{8} \sum_{\ell=-\infty}^{\infty} [\delta(\omega+2\pi\ell) + 9\delta(\omega-\pi+2\pi\ell)].$$

The exact same diffraction pattern, however, results from an infinite number of distinct and arbitrarily complex stochastic stacking processes. In these cases, the flat diffraction background—a fraudulent white noise—belies the material's sophisticated correlated structure.

For example, the HMMs shown in Figs. 10(b) and 10(c) each contain nontrivial high-order correlation between layer types. However, each produces the same diffracted intensity as before:

$$P(\omega) = \frac{3}{8}|\psi|^2 + \frac{\pi|\psi|^2}{8} \sum_{\ell=-\infty}^{\infty} [\delta(\omega+2\pi\ell)+9\delta(\omega-\pi+2\pi\ell)]. \tag{19}$$

As another example that helps to drive home the point, a binary encoding of the entire contents of this manuscript can be stored in the stacking sequence of a close-packed crystal with exactly the same diffraction pattern as Eq. (19). In fact, *any* sufficiently long binary sequence can be encoded in a crystal with this diffraction pattern.

To construct this crystal, each 0 is mapped to one of the layer sequences in

$$\mathcal{L}_0 = \{ABABAB, ABACAC, ACABAC, ACACAB\}$$

with equal probability, while each 1 is mapped to one of the layer sequences in

$$\mathcal{L}_1 = \{ABABAC, ABACAB, ACABAB, ACACAC\}$$

with equal probability. This is equivalent to applying six iterations of the transition dynamic of Fig. 10(b) for each 0 and then applying six iterations of the transition dynamic of Fig. 10(c) for each 1, starting in the central A state each time. *Theorem 3 guarantees that the diffraction pattern of the resulting crystal is always given by Eq. (19).* This is demonstrated in Fig. 10. Diffracted intensity is completely blind to these correlated binary messages, but the original binary message can nevertheless be recovered by other means [e.g., via scanning tunneling microscopy].

A similar story can be told for our human genome—our DNA is the prototypical "aperiodic crystal" [113]. Its diffraction pattern allowed scientists to uncover its general double-helix structure [27–30]. However, the particular content encoded by the DNA can only be extracted by more refined structure-detection methods—carried out by a team of cooperative enzymes *in vivo*.

To summarize our view of diffraction spectra for chaotic crystals, we showed that (i) state-conditioned expectations of layer form factors simplify diffraction analyses, (ii) Bragg reflections persist in close-packed structures without periodic order, (iii) nondiagonalizability heralds structural transformation and yields qualitatively distinct line profiles, and (iv) an infinite number of arbitrarily complex crystal structures all produce the same flat diffraction pattern (plus two Bragg reflections). These lessons supplement a growing awareness of the diversity of "order" in solid-state physical systems—order beyond what can be described by Patterson autocorrelation functions and diffraction patterns [109,114,115].

C. Which ion-channel features do power spectra capture?

Voltage-gated ion channels embedded in cellular membranes are the engines that propagate signals among cells—coordinating electrical communication in our brains, hearts, and throughout our bodies. Better understanding the dynamics among the macromolecular conformations of these

ion channels allows a better understanding of biological function, malfunction, and possible intervention. However, the ion channel conformations cannot be observed directly. Rather, it is only possible to observe a *function* of the hidden conformational state—whether the instantaneous conformation allows current to flow or not. This non-Markovian observable makes the inverse problem (of inferring the dynamic over hidden conformational states) a difficult task [116,117].

Fortunately, a large body of investigation over many decades elucidated the biology of ion channels [118–121]. Nevertheless, questions remain about how the measured power spectral features, like $1/f$ noise, arise in electrical measurements of ion channels. Does it derive from the conformational switching dynamics? Is it from current fluctuations in a particular conformation? If only power spectra are available, then what can be inferred?

Our results offer insight into which features of the power spectrum can be attributed to the channel's conformational switching dynamics. Most notably, our Theorem 1 says that the conditionally IID distributions associated with each conformational state cannot possibly change the observed power spectrum, so long as the average output from each state is left unchanged. So, for example, state-dependent (conditionally IID) noise cannot be the source of $1/f$ noise since it cannot modulate the power spectrum. Previously, this and related questions could only be explored experimentally and numerically [122, Fig. 3].

To contribute to these issues concretely, let us consider current fluctuations in voltage-gated potassium ion channels. Figure 11 illustrates an important biophysical application of Theorem 1: The power spectrum of current through a voltage-gated K^+ channel is invariant to mean-preserving changes in the ion-current PDFs for each channel conformation. We demonstrate this for a particular physiologically motivated model of gating kinetics; see Appendix J for details. However, it must also hold for any model of potassium ion current, which may include many hidden open conformations and electronic states, so long as the output is conditionally IID in each hidden state.

Figure 11(a) shows the continuous-time model of transition rates between conformations. Each of these conformations has a different number of activation gates blocking the channel: from zero in the open state (leftmost, green) through four. Current only flows in the open state, so the K^+ current dynamics is non-Markovian, as is well known. The average current in each state is either I_0 or 0, depending on whether the channel is in an open or closed conformation, respectively. The Hodgkin-Huxley parameters α_n and β_n are voltage-dependent rates of an individual gate opening or closing. Experiments on ion channels are typically performed at a fixed membrane voltage [118,122,123]. With fixed voltage and sampling rate, the continuous-time model generates a simple discrete-time HMM. (Time-varying voltages produce more complicated HMMs.)

Figures 11(b) and 11(c) correspond to a fixed membrane potential of $v = -40$ mV, with potassium current sampled
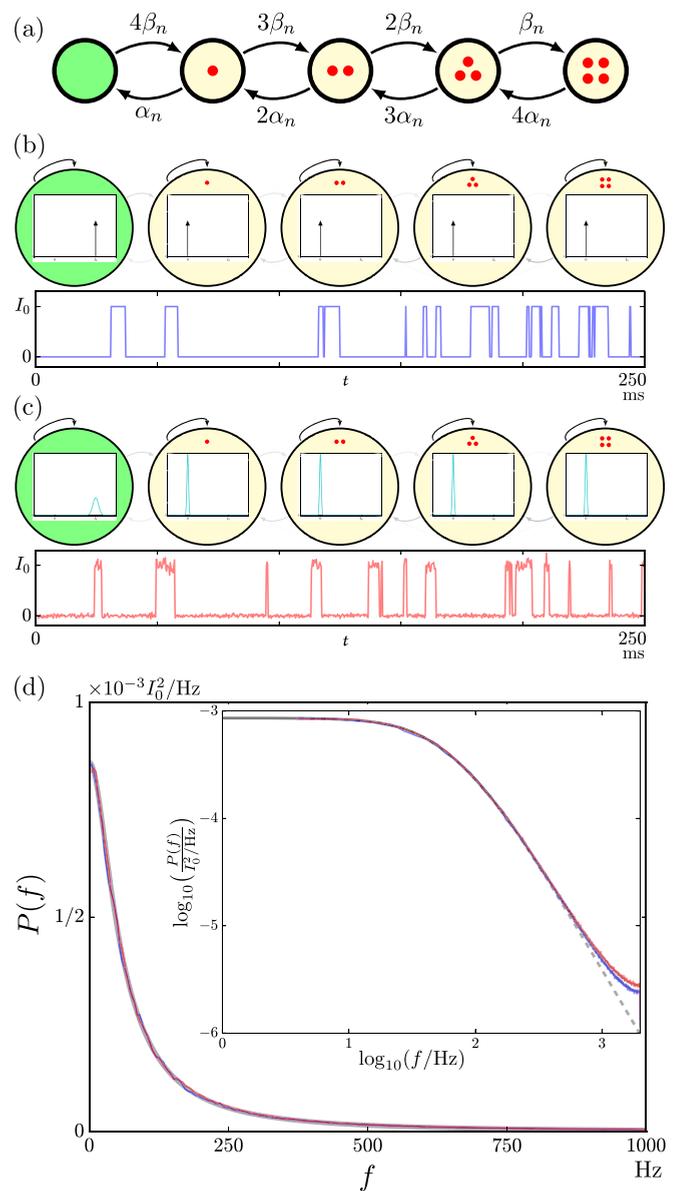
FIG. 11. Biophysical application of Theorem 1: Power spectrum of current through a potassium ion channel does *not* depend on the details of the probabilistic current in each channel conformation. (a) The continuous-time model of transition rates between conformations of the channel. Each state has a different number of activation gates that block the channel (from zero to four). Panels (b) and (c) show HMMs and representative time series of ion current generated from this continuous-time model, at a membrane potential of $v = -40$ mV and sampling rate of $f_0 = 4$ kHz. (b) Binary output. (c) Continuous-valued output, representing both measurement noise and current fluctuations. (d) The power spectrum is shown analytically (thick gray) and numerically [thin blue for binary model (b) and thin red for continuous-valued model (c)]. The numerical power spectra were each calculated from a simulated time series of length 2^{20} using the Welch method, performing FFTs on segments of length 2^{10}. The inset log-log plot shows ~constant behavior at low frequency, $\sim 1/f^2$ behavior at high frequencies, and the effect of finite sampling rate at very high frequencies.

every $\tau_0 = 250\mu$s. Figures 11(b) and 11(c) each show a HMM and a randomly sampled time series. For visualization of the HMMs, the opacity of the directed edges is a simple concave function of the transition probability. (See Appendix J for the exact form of the rate matrix and transition matrix.)

Previous analyses considered the power spectrum from a binary output model similar to Fig. 11(b) [124,125]. Yet with both measurement noise and current fluctuations, a continuous-valued model like Fig. 11(c) better represents the stochastic process observed in experiments. Nevertheless, our Theorem 1 asserts that both of these models produce exactly the same power spectrum, up to a frequency-independent offset. Moreover, for continuous-time processes, this offset vanishes as the sampling rate increases.

We can state this more precisely as a general corollary of Theorem 1.

Corollary 2. For any two HMMs whose transition matrix comes from the same continuous-time generator (via $e^{\tau_0 G}$): If the two models have the same average output in each state, then their power spectra differ only by a frequency-independent offset:

$$P^{(\mathcal{M}')}(\omega) - P^{(\mathcal{M})}(\omega) = (\langle |x|^2 \rangle_{\mathcal{M}'} - \langle |x|^2 \rangle_{\mathcal{M}})/f_0.$$

For a family of such processes with bounded variance of the instantaneous observable, this offset must approach zero as the sampling rate increases.

Proof. This follows immediately from Theorem 1 when we treat f_0 explicitly. (Recall that f_0 was set to unity in the discrete-time case.)

Nevertheless, small constant offsets can be observed between the empirical power spectra whenever a finite sampling frequency is used.

Figure 11(d) shows that the power spectra from model (b) (blue) and model (c) (red) are indeed the same, up to a very small constant offset of $[\boldsymbol{\pi}_{\text{open}}\sigma^2_{\text{open}} + (1 - \boldsymbol{\pi}_{\text{open}})\sigma^2_{\text{closed}}]/f_0 \approx 3.6 \times 10^{-7}I_0^2$/Hz. This small frequency-independent offset predicted between the two models becomes visible at the highest frequencies of the inset log-log plot where power is lowest.

The power spectrum in the continuous-time limit of $f_0 \to \infty$ is derived in Appendix J. The analytic curve (dashed thick gray) is shown in the log-log inset of Fig. 11(d). It is flat at low frequencies and falls off as $1/f^2$ at high frequencies. However, these processes are sampled at a finite rate of $f_0 = 4$ kHz. The analytic curves for the expected *empirical* power spectra [from Figs. 11(b) and 11(c)] are shown in thick gray in the log-log inset. They deviate from the continuous-time model's $1/f^2$ behavior but match the numerical power spectra extremely well up to arbitrarily high frequencies. This whitening of empirical power spectra at high frequencies is predicted by Eq. (6).

Our results suggest that observed $1/f$ noise is likely due to non-IID current fluctuations in the channel's open conformation. This conclusion is at odds with the conclusion of Ref. [122] but is consistent with theoretical [124,125] and experimental [123] observations in much earlier work, where the Lorentzian-like power spectrum of the channel's conductance fluctuations appears to be additive to the $1/f$ flicker noise background.

Despite 70 or more years of ongoing investigation and great advances, potassium ion-channel conduction is still not fully understood [120,126]. Fortunately, the analytic results here can help—they can be applied to evaluate the power spectrum from any proposed model and so aid in bridging theory to experiment. To make genuine progress, these models will necessarily be more complicated, including transitions between distinct electronic conduction states in the channel's open conformation. On the one hand, the results emphasized that power spectra are indifferent to several stochastic features of alternative models. Yet, on the other, the relationship between power spectra and eigenvalues of the rate matrix immediately tells us much about which models can be ruled out based on nontrivial features of observed power spectra.

V. STRUCTURE IN NOISE?

Surely leveraging predictions to exclude alternative mechanisms is a central strategy in physical science, but is there a direct way to discover structure in apparent noise? One approach immediately suggests itself. We first reflect on, and further develop, the theory of higher-order spectra—which maintain much of the familiarity and convenience of power spectra. However, enumerating and interpreting higher-order spectra in general is difficult. Not the least reason for this is that the number of possible spectral descriptions multiplies combinatorially. Or sometimes the motivating questions are more pointed. In these cases, it is often more incisive to develop an information-theoretic probe of statistical interdependencies.

The ultimate goal, though, in using any of these tools is constructing a testable model that generates the observed features of interest. In the deterministic case familiar in classical physics, this is synonymous with learning the equations of motion. In open complex systems with noise and many layers of feedback, this may instead take the form of a hidden-state model—whose input-dependent time-evolution operator generalizes the deterministic equations of motion. By directly expressing mechanisms, developing such models allows thoughtful reflection on assumptions, generalizations, and interventions.

A. Polyspectra

Higher-order spectra—often simply *polyspectra*—are a natural next step to detecting structure beyond the pairwise correlations conveyed by power spectra [127,128]. As we will show, polyspectra are not the ultimate answer to structure detection; however, they certainly are a tool that practitioners should be aware of. The following derives new analytical expressions for polyspectra useful for both experimentalists and theoreticians. In emphasizing properties already implicit in the foregoing, the analysis reveals that polyspectra, too, are blind to predictable structure in processes.

Following Ref. [127], we introduce a general formulation for polyspectra that implicates expectation values—such as $\langle g_0(X_{t_0})g_1(X_{t_1})g_2(X_{t_2}) \rangle$—of time-displaced functions of the observables. As part of the generalization, let $g_k : \mathcal{A} \to \mathbb{C}$ be any function taking observables to complex numbers. If $\mathcal{A}$ is an abstract set—representing, say, observing colors yellow

or red $\mathcal{A} = \{y, r\}$—then the g_k functions allow a polyspectral analysis that is not possible otherwise.

Consider the $(g_0, \ldots, g_K)$ *polyspectrum*:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K) = \lim_{N \to \infty} \frac{1}{N} \left\langle \prod_{k=0}^{K} \widetilde{g}_k^{(N)}(\omega_k) \right\rangle, \quad (20)$$

where $\omega_0 \equiv -\sum_{k=1}^{K} \omega_k$ and

$$\widetilde{g}^{(N)}(\omega) \equiv \sum_{t=1}^{N} g(X_t) e^{-i\omega t}. \quad (21)$$

Although challenging to interpret in full generality, in principle polyspectra provide a window into a process' high-order nonlinear dependencies. Equation (20) says that polyspectra are the expected products of Fourier components—components that, in practice, can be obtained from the FFT. Given the FFT's well-known computational efficiency, polyspectra are an especially appealing probe of higher-order structure.

Many special cases of the $(g_0, \ldots, g_K)$ polyspectrum have been well studied. For example, $S_{\overline{X},X}(\omega) = P(\omega)$ is the power spectrum, $S_{\overline{X},Y}(\omega)$ is the cross-spectrum discussed in Appendix K, $S_{\overline{X},X,X}(\omega_1, \omega_2)$ is the *moment bispectrum*, $S_{\overline{X},X,X,X}(\omega_1, \omega_2, \omega_3)$ is the *moment trispectrum*, $S_{\overline{X-\langle X\rangle},X-\langle X\rangle,X-\langle X\rangle}(\omega_1, \omega_2)$ is the *cumulant bispectrum*, and so on. The following, in contrast, addresses $(g_0, \ldots, g_K)$ polyspectra generally.

Combining Eqs. (20) and (21) yields:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K)$$
$$= \lim_{N \to \infty} \frac{1}{N} \sum_{t_0=1}^{N} \cdots \sum_{t_K=1}^{N} \left\langle \prod_{k=0}^{K} g_k(X_{t_k}) \right\rangle \prod_{k=0}^{K} e^{-i\omega_k t_k}. \quad (22)$$

Thus, the $(g_0, \ldots, g_K)$ polyspectrum is closely related to the expectations $\langle \prod_{k=0}^{K} g_k(X_{t_k}) \rangle$, as suggested. And, crucially, the expectation values can be calculated exactly from any hidden-state model. Unraveling this exact relationship gives new insight into what the polyspectrum conveys about a process.

The time variables $(t_k)_{k=0}^{K}$ in Eq. (22) are not necessarily time ordered by the index k. Moreover, time variables may coincide; i.e., it is possible to have $t_j = t_k$ for $j \neq k$. To remove these complications, one can work with a reduced and time-ordered collection of time variables $(t'_k)_{k=0}^{\kappa}$ such that $t'_k > t'_{k-1}$, where $\kappa + 1 = |\{t_k\}_{k=0}^{K}| \leqslant K + 1$ is the number of distinct values of the time variables. These time-ordered variables are defined recursively via $t'_0 = \min(\{t_k\}_{k=0}^{K})$ and $t'_\ell = \min(\{t_k\}_{k=0}^{K} \setminus \{t'_k\}_{k=0}^{\ell-1})$.

The original time variables $(t_k)_{k=0}^{K}$ induce a function $\alpha : \{0, 1, \ldots K\} \to \{0, 1, \ldots \kappa\}$ that compresses and time-orders the indices, such that $t_k = t'_{\alpha(k)}$. Although α does not generally have a unique inverse, we define $\alpha^{-1}(\ell) = \{k \in \{0, 1, \ldots K\} : \alpha(k) = \ell\}$ to be the set of indices that map to ℓ.

For HMMs, we can then express the expectations in Eq. (22) as:

$$\left\langle \prod_{k=0}^{K} g_k(X_{t_k}) \right\rangle = \left\langle \prod_{\ell=0}^{\kappa} g_{\alpha^{-1}(\ell)}(X_{t'_\ell}) \right\rangle$$
$$= \mathrm{tr}\left[|\mathbf{1}\rangle\langle\boldsymbol{\pi}| \Omega_{g_{\alpha^{-1}(0)}} \prod_{\ell=1}^{\kappa} T^{t'_\ell - t'_{\ell-1}} \Omega_{g_{\alpha^{-1}(\ell)}} \right], \quad (23)$$

where $\mathrm{tr}(\cdot)$ denotes the trace, the product of operators maintains time ordering, we have introduced the new functions $g_{\alpha^{-1}(\ell)}(x) \equiv \prod_{k \in \alpha^{-1}(\ell)} g_k(x)$, and we used the generalized average-observation matrices:

$$\Omega_g \equiv \sum_{s \in \mathcal{S}} \langle g(X) \rangle_{\mathrm{p}(X|s)} |s\rangle\langle s|. \quad (24)$$

Note that the summations over all time variables in Eq. (22) induce all possible functions α that permute and compress the indices. And, within each compressed time ordering, all possible values of the indices consistent with that ordering are summed over. To enumerate all possible compressed time orderings, it is useful to explicitly introduce the set $\mathbb{F}_K^{(\kappa)}$ of all surjective functions mapping $\{0, 1, \ldots K\}$ onto $\{0, 1, \ldots \kappa\}$. For HMMs, we can then express the expectations [129] in Eq. (22) as:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K) = \lim_{N \to \infty} \frac{1}{N} \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \sum_{t'_0=1}^{N-\kappa} \sum_{t'_1=t'_0+1}^{N-\kappa+1} \cdots \sum_{t'_\kappa = t'_{\kappa-1}+1}^{N} \left\langle \prod_{\ell=0}^{\kappa} g_{\alpha^{-1}(\ell)}(X_{t'_\ell}) \right\rangle \prod_{\ell=0}^{\kappa} e^{-i\omega_{\alpha^{-1}(\ell)} t'_\ell}, \quad (25)$$

where $\omega_{\alpha^{-1}(\ell)} \equiv \sum_{k \in \alpha^{-1}(\ell)} \omega_k$.

Leveraging Eq. (23), Appendix O shows that Eq. (25) yields the closed-form expression for the continuous part of the $(g_0, \ldots, g_K)$ polyspectrum:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K) = \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \langle\boldsymbol{\pi}| \Omega_{g_{\alpha^{-1}(0)}} \left\{ \prod_{\ell=1}^{\kappa} T\left[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T \right]^{-1} \Omega_{g_{\alpha^{-1}(\ell)}} \right\} |\mathbf{1}\rangle, \quad (26)$$

where $\mathfrak{z}_{\ell:\kappa}^{(\alpha)} \equiv \prod_{k=\ell}^{\kappa} \mathfrak{z}_k^{(\alpha)} = e^{-i\sum_{k=\ell}^{\kappa} \omega_{\alpha^{-1}(k)}}$.

We see that the $(g_0, \ldots, g_K)$ polyspectrum sandwiches up to K resolvents of the time evolution operator T, with each resolvent separated by average-observation matrices. The resolvents couple the chain of observation matrices, and the

polyspectrum reports the average interaction among these observables over arbitrary displacements.

Using Eq. (9) to express the resolvent $[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T]^{-1}$ in terms of T's eigenvalues and spectral projection operators,

we again see that the eigenspectrum of the time evolution operator directly controls the polyspectrum of the stochastic process. Appendix O 1 discusses this further.

Note, too, that, in addition to the continuous part of the polyspectrum, there are possible discontinuities in the $(g_0, \ldots, g_K)$ polyspectrum wherever $1/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} \in \Lambda_T$. This "discrete" part of the polyspectrum consists of $(K-1)$-dimensional hyperplanes in the K-dimensional $(\omega_k)_{\ell=1}^K$ frequency space where the magnitude of the polyspectrum diverges. Our analytic formulas imply that polyspectra may diverge only where subsets of the frequencies sum to an eigenfrequency ω_λ of an eigenvalue $\lambda = e^{i\omega_\lambda}$ on the unit circle. For example, in a generic bispectrum—which has a two-dimensional frequency space—this leads to coexisting diagonal ($\omega_1 + \omega_2 = \omega_\lambda$), vertical ($\omega_1 = \omega_\lambda$), and horizontal ($\omega_2 = \omega_\lambda$) streaks of high intensity.

It is useful to probe several special cases of the $(g_0, \ldots, g_K)$ polyspectrum. Consider, first, the $(\overline{X}, X)$ polyspectrum, $S_{\overline{X}, X}(\omega_1)$, which is simply the power spectrum $P(\omega_1)$. In this case, $K = 1$. So we must consider the functions contained in $\mathbb{F}_1^{(0)} = \left\{ \boxed{\begin{smallmatrix}1\\0\end{smallmatrix}\!\searrow\!\begin{smallmatrix}1\\0\end{smallmatrix}} \right\}$ and $\mathbb{F}_1^{(1)} = \left\{ \boxed{\begin{smallmatrix}1\rightarrow 1\\0\rightarrow 0\end{smallmatrix}}, \boxed{\begin{smallmatrix}1\\0\end{smallmatrix}\!\times\!\begin{smallmatrix}1\\0\end{smallmatrix}} \right\}$. For the compressive function $\alpha = \boxed{\begin{smallmatrix}1\\0\end{smallmatrix}\!\searrow\!\begin{smallmatrix}1\\0\end{smallmatrix}}$, we obtain $\alpha^{-1}(0) = \{0, 1\}$, yielding:

$$\Omega_{g_{\alpha^{-1}(0)}} = \Omega_{g_{\{0,1\}}} = \Omega_{|X|^2} = \sum_{s \in \mathcal{S}} \langle |X|^2 \rangle_{\mathrm{p}(X|s)} |s\rangle \langle s|.$$

The ($\kappa = 0$) contribution to the power spectrum is thus:

$$\langle \boldsymbol{\pi} | \Omega_{|X|^2} | \mathbf{1} \rangle = \sum_{s \in \mathcal{S}} \langle |X|^2 \rangle_{\mathrm{p}(X|s)} \langle \boldsymbol{\pi} | s \rangle = \langle |x|^2 \rangle,$$

which is indeed the first term in Eq. (6). The ($\kappa = 1$) contribution to the power spectrum is as follows:

$$\sum_{\alpha \in \mathbb{F}_1^{(1)}} \langle \boldsymbol{\pi} | \Omega_{g_{\alpha^{-1}(0)}} T [e^{i\omega_{\alpha^{-1}(1)}} I - T]^{-1} \Omega_{g_{\alpha^{-1}(1)}} | \mathbf{1} \rangle,$$

where it should be recalled that $\omega_0 = -\omega_1$. Plugging in the identity and swap functions of $\mathbb{F}_1^{(1)}$, this becomes

$$2\mathrm{Re} \langle \boldsymbol{\pi} | \Omega_{\overline{X}} T (e^{i\omega_1} I - T)^{-1} \Omega_X | \mathbf{1} \rangle,$$

which is indeed the last term of Eq. (6).

Appendix O 2 gives a similar analysis of the cumulant bispectrum. Analogously to Corollary 1, we find in Theorem 6 of Appendix O 2 that:

The cumulant bispectrum is completely flat for any process generated by a HMM with the same average output $\langle X \rangle_{\mathrm{p}(X|s)} = \langle x \rangle$ from each hidden state.

This serves as a stark warning against overreliance on any particular polyspectrum: Structure and interdependence will be missed and it is challenging to predict for which polyspectra this will happen.

Can polyspectra overcome the shortcomings of power spectra and avoid the inherent pitfalls? Only indirectly. For example, the cumulant bispectrum—often championed as the next-step tool for detecting nonlinearities in a process [128,131–133]—is completely flat for the example process from Fig. 6 for all values of the transition parameter $p \in [0, 1]$. That is, the cumulant bispectrum tells us no more than

the power spectrum. Yet the moment bispectrum should be useful in this case, if one only knew how to interpret it. Alternatively, and more simply, if one is sharp enough to use (in fact, guess) $g(X) = X^2$, then the change in observable reveals the process' structure through the single-frequency $S_{X^2, X^2}(\omega_1)$ polyspectrum.

Such guesswork is inescapable and, more to the point, reveals a fundamental problem: If a process' structure is unknown *a priori*, then there is no guarantee that the structure will be revealed, even after an infinite number of higher-order polyspectra have been inspected. Generically, it is not clear *which* set of polyspectra to use to detect structure. Fortunately, information theory and model reconstruction both provide more principled approaches to extracting a process' statistical dependencies [66,86].

B. Becoming informed

A more systematic and direct method for exploring beyond-pairwise correlations in stationary stochastic processes is through the sequence of *myopic entropy rates* [63,66,67,134–136]:

$$h_L = H(X_L | X_1 X_2 \ldots X_{L-1}),$$

with $h_1 = H(X_1)$. For example, the RRXOR process has $h_1 = h_2 = \log |\mathcal{A}| = 1$ bit/symbol—it appears as random as possible when considering symbols individually or in pairs. Structure is unveiled, though, for $L \geqslant 3$ when $h_L < 1$. That is, progressively longer Markov-order-L approximations of the infinite-Markov-order process reveal progressively more of its hidden structure.

In fact, h_L's convergence reflects how structure is hidden in the stochastic process [136]. As $L \to \infty$, h_L approaches the process' *Shannon entropy rate h*—the irreducible randomness per symbol after all orders of correlation have been taken into account. Notably, the accumulation of the excess myopic entropy $\sum_{L=1}^{\infty} (h_L - h) = \mathbf{E}$—the *excess entropy*—quantifies the total mutual information between the past and future of a process: $\mathbf{E} = I(\ldots, X_{-1}, X_0 ; X_1, X_2, \ldots)$. So while $I(X_0; X_\tau) = 0$ for all $\tau > 0$ for the RRXOR process, the past and future are nevertheless correlated since $\mathbf{E} > 0$. And the convergence to predictability can be viewed in the frequency domain through the excess-entropy spectrum introduced in Ref. [67]. Taken together, this suggests that myopic entropy rates serve well to identify hidden structure beyond pairwise correlation. They show how predictability improves as progressively longer historical context is used.

However, correlations are not always restricted to contiguous blocks. Therefore, there can be pairwise correlations among distant observables while $h_2 = 0$. Moreover, the myopic entropy rates as defined above are restricted to stationary processes. Consequently, despite their utility, myopic entropies are not ideal for direct indication of L-way correlation in the most general setting.

A more direct indicator of L-way correlation is found in the *dependence function D_L*, which quantifies the maximal uniquely-L-way correlation that exists in a process. We say a set $\boldsymbol{\chi}$ of random variables is *fully correlated* if all constituent random variables inform all of the others;

that is, if:

$$
H(X|\boldsymbol{\chi} \setminus \{X, X'\}) - H(X|\boldsymbol{\chi} \setminus \{X\})
$$
$$
= I(X ; X' |\boldsymbol{\chi} \setminus \{X, X'\})
$$
$$
> 0,
$$

for all $X, X' \in \boldsymbol{\chi}$. A process is then L-way correlated if it has a set of L random variables that are fully correlated. One way to quantify this L-way correlation is through the following dependence function:

$$
D_L \equiv \sup_{\{\boldsymbol{\chi} \subset \{X_t\}_t : |\boldsymbol{\chi}| = L\}} \min_{X, X' \in \boldsymbol{\chi}} I(X ; X' |\boldsymbol{\chi} \setminus \{X, X'\}),
$$

defined here only for $L \geqslant 2$. L-way dependence is nonzero if and only if there are novel L-way contributions to a process' total correlation. Note that dependence can be applied to non-stationary processes, processes of finite duration and indeed to any collections of random variables.

Consider, as a simple example of noncontiguous dependencies, the process consisting of two interlaced RRXOR processes with unambiguous phase, which arose from measurement of an entangled quantum system in Sec. IV A. Explicitly:

$$
X_{6n} = \text{XOR}(X_{6n-4}, X_{6n-2}) \quad \text{and}
$$
$$
X_{6n-1} = \text{XOR}(X_{6n-5}, X_{6n-3}),
$$

whereas $X_{6n-5}, X_{6n-4}, X_{6n-3}$, and X_{6n-2} are all generated from a uniform distribution for all $n \in \{1, 2, \dots\}$. Joint probabilities over contiguous variables are completely uncorrelated and as random as possible, up until a block-length of five. Let us treat the example as a stationary process: Calculating probabilities from word frequencies in a single realization, with the implicit assumption of stationarity, effectively inducing random phase. Then we find full randomness in the myopic entropy rates up to block length five: $h_L = \log |\mathcal{A}| = 1$ bit for $1 \leqslant L < 5$. Then, finally, a reduction in apparent entropy occurs at h_5, after which $h_L < h_{L-1}$ for $L \geqslant 5$. Notably, h_3 reflects maximal randomness within its purview, whereas the process actually has three-way but no lower-order dependencies. This yields $D_1 = D_2 = 0$ and $D_3 > 0$. With known phase, we would have $D_3 = 1$ bit.

However, when the process is unknown and only a single realization is available for analysis, probabilities can be inferred only from motifs of random-variable clusters. For example, estimating $\Pr(X_{t-2}, X_t, X_{t+2})$ *as if* the process were stationary leads to finding $0 < \widetilde{D}_3 < 1$, where $\widetilde{D}_L$ denotes approximating the dependence function assuming stationarity and testing a limited set of motifs. Usefully, $\widetilde{D}_L$ sets a lower bound on D_L. So nonzero $\widetilde{D}_L$ implies L-way dependence. Curiously, the assumption of stationarity induces $\widetilde{D}_L > 0$ for all $L \geqslant 3$; reminiscent of how $h_L - h_{L-1} > 0$ for all $L \geqslant 3$ for the RRXOR process with ambiguous phase. In each case, these higher-order correlations correspond to the observer's ability to resolve phase ambiguity.

The dependence function seems to fulfill its desired role of identifying high-order correlations that cannot be explained by lower-order phenomena. Taking a step back, though, we might question the whole endeavor. Can a single model-free signal-analysis method ever reliably detect information processing and thus complex structure in the world around us?

We clearly ousted power spectra for this task. Nevertheless, our arguments here lend support to an affirmative answer, but at the cost of more nuanced and computationally intensive techniques. What is the range of validity of the informational measures discussed above? Can they be entrusted with finding structure in the noise?

First, it should be noted that Shannon entropy is only fully justifiable for alphabets $\mathcal{A}$ of countable cardinality. So apparently continuous observables must be partitioned into measurable sets to apply the informational measures like the myopic entropy rates and the dependencies D_L. Nevertheless, quantum physics suggests that even very large and apparently continuous systems are, in principle, always represented in a countable basis. Practically, too, measurement devices only have a finite precision, so observations are discretized in practice anyway. Therefore, Shannon entropies (like the myopic entropy rates and the dependencies) can be applied in principle.

Second, a likely more-severe challenge arises from limitations built into information theory itself. Specifically, there are more nuanced interpretations of multiway statistical dependencies that are missed by all joint and conditional entropies and all mutual and conditional mutual information [137,138].

Finally, a third and practical challenge arises from limited data: Reliable estimates of probabilities are not always available. Model building offers the strongest response to this challenge. Generative models inferred from low-order statistics sometimes encapsulate predictions of rare events [139]. And, at least, they give a prediction for high-order statistics. Testing these predictions against observation allows refining one's model and discovering new structure.

VI. CONCLUSION

Our investigation began with the modest task of showing how to calculate the correlation function and power spectrum given a signal's generator. To this end, we briefly introduced hidden Markov models as signal generators and then used the linear-operator techniques of Ref. [68] to calculate their autocorrelation and power spectra in closed form.

This led to several lessons. First, we saw that the power spectrum is a direct fingerprint of the resolvent of the model's time-evolution operator, analyzed along the unit circle. Second, spectrally decomposing the not-necessarily-diagonalizable time-evolution operator, we discovered the range of qualitative behaviors that can be exhibited by autocorrelation functions and power spectra. Third, contributions from eigenvalues on the unit circle had to be extracted and dealt with separately. Contributions from eigenvalues on the unit circle correspond to Dirac δ functions—the analog of Bragg reflections in diffraction—whereas eigencontributions from inside the unit circle correspond to diffuse peaks that become sharper for eigenvalues closer to the unit circle. Finally, we found that nondiagonalizable eigenmodes yield qualitatively different line profiles than their diagonalizable counterparts.

These first results incisively answer the challenges raised by Ruelle-Pollicott resonance theory about the possible relationship between complex eigenvalues of time-evolution operators and the correlation and power spectra of observables

[12–14]. In short, we provided the exact relationship between the time-evolution operator and the correlation functions and power spectra, as well as the possible behavior modes of each. The result is a deeper theoretical understanding and constructive calculational methods. These complement early investigations that experimentally delivered meromorphic power spectra from chaotic dynamical systems [10,11].

Accordingly, our findings bear on modern applications of Ruelle-Pollicott resonance theory. These applications are leading, for example, to better understanding of sensitivities in climate models [17] and the dynamics of open quantum systems via their correspondence to classical chaotic dynamical systems [15,16]. Our results provide full analytical correspondence between observed correlation and the spectral properties of nonunitary models. Our approach also bears on Koopman operator theory and its applications, which has received a new wave of attention due to the success of recent data-driven algorithms [140]. However, our results also clarify that resonances discovered via pairwise correlation are generically an insufficient representation of the spectral features of such nonnormal dynamics. This emphasizes that the full spectral representation of the effective nonnormal dynamics [68], generically inaccessible via pairwise correlation, is worth pursuing. Success in this will immediately yield predictions about many complex systems of interest.

The most surprising and more immediate finding, though, is that temporal structure can fully evade detection by power spectra. Arbitrarily sophisticated processes can have exactly flat power spectra and so masquerade as white noise. Accordingly, we called such processes *fraudulent white noise* processes. Theorem 1 and Corollary 1 characterized the many ways that structure can be hidden from power spectra. And, ultimately, Theorem 2 addressed the more general condition for fraudulent white noise, in which the generated time series could be input-dependent and nonstationary.

We showed that fraudulent white noise and the degeneracy of power spectra have important physical implications. We found that fraudulent white noise arises from sequential measurements of entangled quantum systems. Moreover, the generation of high-order structure and the complete absence of pairwise structure occurred despite the fact that these quantum states resulted from a simple sequence of pairwise interactions. Beyond quantum physics, our results on the degeneracy of power spectra have consequences throughout the sciences. We derived new results on the degeneracy of diffraction patterns and showed how the entire contents of the present work can be encoded in a crystal with a flat diffraction pattern. We then leveraged our results to comment on a longstanding debate about $1/f$ noise in biomolecular ion channels.

We started out noting that, on the one hand, divergent correlation length often heralds the emergence of new types of order. And, on the other, that pairwise correlation is generically identified as the structure in random systems. However, we showed that there is often rich structure even in the absence of pairwise correlations. What types of order are we *failing* to predict due to an historical emphasis on pairwise correlations? Complex systems surely exhibit emergent structure beyond the reach of pairwise statistics. There is almost surely more functionally relevant brain activity available in EEGs beyond what is reported in their power spectra. Perhaps, however,

we should consider beyond-pairwise structure for even simple generators of structure. For example, cosmological models could be more thoroughly tested against structure in the CMB beyond what is contained in the two-point angular correlation functions.

Having diagnosed the structures inaccessible via power spectra, we discussed how to detect beyond-pairwise structure. We obtained a closed-form expression for all polyspectra but showed that higher-order spectra are also completely flat in some cases where structure should have been apparent. In response, we introduced the dependence function to detect any L-way correlations for any L. We also stressed the importance of model building whenever possible. In particular, it can help anticipate and perhaps avoid not-yet-encountered catastrophes, which are often a by-product of the high interconnectivity of complex socio-economic systems [141]. Model building, beyond pure signal analysis, is key in this—it allows us to discover new mechanisms in nature.

This all said, nature still keeps us in the dark. We showed that the correlations in a message can be shifted to arbitrarily high orders of correlation. The result is that, for finite length messages, statistical inference can be made effectively impossible regardless of one's sophistication. Nature herself employs this technique whenever we observe an increase in entropy—giving the impression of randomness generated, when it is only ever structure hidden in inaccessibly obscure high-order correlations. Waking up to the true hues of reality—prying open the black box, dispelling apparent white noise—continues to require new theory and new experimentation.

"… it is clearly wise to learn what a procedure really seems to be telling us about" [142, p. 60].

APPENDIX A: DIFFRACTION PATTERNS AS POWER SPECTRA

Diffraction patterns are used extensively to infer material structure from the scattering of, for example, an incident x-ray beam [143–147]. Generally, consider $\vec{r} \in \mathbb{R}^d$ to be a vector in d-dimensional real space. The spatial arrangement of elastic scatterers is given by the scatterers' density $f(\vec{r})$. Ideally, we wish to recover $f(\vec{r})$ from our diffraction experiments, which provide measured intensities. However, far-field patterns of diffracted intensity yield only $I_{\text{diff}}(\vec{q}) = c|F(\vec{q})|^2$, where $F(\vec{q}) = \int_{\mathbb{R}^d} f(\vec{r}) e^{-i\vec{q}\cdot\vec{r}} d^d\vec{r}$ is the d-dimensional Fourier transform of $f(\vec{r})$, c is some constant, and $\vec{q} = 2\pi(\vec{k}_{\text{out}} - \vec{k}_{\text{in}})$ is the scattering vector that quantifies the change in the incident wave vector. In other words, $F(\vec{q})$'s phase information is lost when only intensity is measured. This is known

as the "phase problem" [112]. The x-ray beam's expected diffracted intensity is proportional to $\langle |F(\vec{q})|^2 \rangle$, which is the d-dimensional generalization of a power spectrum. However, it is also interesting to relate the d-dimensional diffraction pattern, along a curve in reciprocal space, to the more familiar one-dimensional power spectrum.

For a given scattering vector $\vec{q}$, decompose $\vec{r} = \vec{r}_\parallel + \vec{r}_\perp$, where $\vec{r}_\parallel \equiv (\vec{r} \cdot \hat{q})\hat{q}$ and $\hat{q} = \vec{q}/|\vec{q}|$. Then let $\mu_\perp(\vec{r}_\parallel)$ be the accumulated density within the entire cross-sectional plane perpendicular to and uniquely identified by $\vec{r}_\parallel$; i.e., $\mu_\perp(\vec{r}_\parallel) \equiv \int_{\mathbb{R}^{d-1}} f(\vec{r}_\parallel + \vec{r}_\perp)\, d^{d-1}\vec{r}_\perp$. We then find that in general:

$$I_{\text{diff}}(\vec{q}) = c \left| \int_{\mathbb{R}} \mu_\perp(\vec{r}_\parallel) e^{-iqr_\parallel}\, dr_\parallel \right|^2. \tag{A1}$$

In particular, we see that the diffraction pattern along any line $\vec{q} = q\hat{q}$ (with varying q but fixed $\hat{q}$) is the power spectrum of the net magnitude of scatterers within sequential cross sections of real space perpendicular to $\hat{q}$.

For molecular or crystalline structures, the net scatterer density may often be well approximated by a superposition of more elementary densities $f(\vec{r}) = \sum_j f_j(\vec{r} - \vec{r}_j)$. If we partition the real space occupied by the material into N layers of thickness τ_0, stacked along a particular direction $\hat{\ell}$, then we obtain the alternative expression:

$$I_{\text{diff}}(\vec{q}) = c \left| \sum_{n=1}^{N} F^{(n)}(\vec{q}) e^{-i\omega n} \right|^2, \tag{A2}$$

where $\omega = \tau_0 \vec{q} \cdot \hat{\ell}$ is (2π times) the change in wave number per layer in the stacking direction. In such cases, the *layer form factors* are as follows:

$$F^{(n)}(\vec{q}) \equiv \sum_{j \in n\text{th layer}} F_j(\vec{q}) e^{-i(\vec{q} \cdot \vec{r}_j - n\omega)},$$

where the "nth layer" is the set of indices $\{j : n\tau_0 \leqslant \langle \vec{r}_j \rangle \cdot \hat{\ell} < (n+1)\tau_0\}$ for the elementary constituents typically contained in the layer. And the *atomic form factor*

$$F_j(\vec{q}) = \int_{\mathbb{R}^d} f_j(\vec{r}) e^{-i\vec{q} \cdot \vec{r}}\, d^d\vec{r}$$

is the d-dimensional Fourier transform of $f_j(\vec{r})$. As a result, we see that the expected diffraction pattern can always be written as the power spectrum of layer form factors:

$$\langle I_{\text{diff}}(\vec{q}) \rangle = cNP(\omega) = c \left\langle \left| \sum_{n=1}^{N} X_n e^{-i\omega n} \right|^2 \right\rangle, \tag{A3}$$

with $X_n = F^{(n)}(\vec{q}) \in \mathbb{C}$ as the layer form factor of the nth layer of the material.

The frequency dependence of $F^{(n)}(\vec{q})$ is often factored out to "correct" the diffraction pattern, so that only the structure of interest—features due to the stacking sequence—remains [148,149].

1. From fraudulent white noise to Debye-Waller theory

It is important to recognize that the elementary positions $\{\vec{r}_j\}_j$ are random variables, since thermal motion—and even quantum uncertainty at zero temperature—can significantly

displace them from their average value. Indeed, the observed diffraction pattern is not consistent with evaluating $\{\vec{r}_j\}_j$ at their average values. This is because the expected value of a structure factor is not the same as the structure factor evaluated at the expected value of elementary positions. Nevertheless, the observed diffraction pattern *is* consistent with $\langle |F(\vec{q})|^2 \rangle$, where the averaging over realizations induces the proper thermal (and quantum-uncertainty) averaging. However, the thermal averaging appears unwieldy in the general case. Fortunately, we can leverage our Theorems 1 and 3 to rigorously recover the simplifications of Debye-Waller theory in our setting of randomly stacked structures.

Suppose there is a hidden-state model $\mathcal{M}(\vec{m}) = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \{T_t(\vec{m})\}_t, \boldsymbol{\mu}_1)$ that generates the correct statistics of the layer form factors in the material—taking the stochastic stacking process, thermal motion, and quantum uncertainty into account. Theorems 1 and 3 imply that the diffraction pattern will be the same (up to a constant offset) if we instead consider the much simpler hidden-state model $\mathcal{M}'(\vec{m}) = (\mathcal{S}, \mathcal{B}, \mathcal{Q}, \{T_t(\vec{m})\}_t, \boldsymbol{\mu}_1)$ that outputs only the *expected* layer form factor from each latent state.

Each of the *expected* layer form factors $b \in \mathcal{B}$ can be expressed as:

$$
\begin{aligned}
b = \langle X \rangle_{p(X|s \in \mathcal{S}_b)} &= \sum_{j \in \text{type-}b \text{ layer}} F_j(\vec{q}) \langle e^{-i\vec{q} \cdot \vec{r}_j} \rangle \\
&= \sum_{j \in \text{type-}b \text{ layer}} F_j(\vec{q}) e^{-i\vec{q} \cdot \langle \vec{r}_j \rangle} D_j(\vec{q}).
\end{aligned}
$$

Notably,

$$D_j(\vec{q}) \equiv \langle e^{-i\vec{q} \cdot (\vec{r}_j - \langle \vec{r}_j \rangle)} \rangle \approx e^{-\frac{1}{6}\sigma_{\vec{r}_j}^2 q^2}$$

is exactly the Debye-Waller factor for an elementary scattering site of type j [106]. The variance $\sigma_{\vec{r}_j}^2$ scales as $k_B T$ at high temperatures (via the equipartition theorem), although it is still nonzero as $T \to 0$ due to zero-point energy.

In the case that the Debye-Waller factors from all scattering sites are the same (i.e., $D_j(\vec{q}) = D(\vec{q})$), the thermal averaging over positions does *not* broaden the diffraction pattern at all. Rather, the Debye-Waller factor only suppresses the diffracted intensity at large scattering magnitudes by an approximately Gaussian envelope (centered at $\vec{q} = 0$).

In contrast, *thermal broadening*—expected of spectral lines throughout the domains of physics—is due to a Doppler effect from the *velocity* of the elementary scatterers (rather than their random positions). This induces a Gaussian convolution on the otherwise Lorentzian line profile. Whereas the Debye-Waller factor is important, thermal broadening is not a significant source of line broadening for x-ray diffraction [150].

2. Close-packed structures

Recall that each layer of a close-packed structure has a two-dimensional hexagonal close-packed lattice. The diffracted intensity will thus only be nonzero at scattering vectors that satisfy the Laue condition for allowed reflections from the two-dimensional crystal:

$$\vec{q} - (\vec{q} \cdot \hat{\ell})\hat{\ell} = \vec{G}, \tag{A4}$$

where $\vec{G}$ is in the set of reciprocal lattice vectors of the 2D hexagonal lattice.

For close-packed structures, there are only three types of layers, differing only via relative displacements of $1/3$ of a lattice translation vector $\vec{t}$ in the plane of the layer [105]. As a result, if type-A layers have an expected layer form factor of

$$A = \sum_{j \in \text{type-}A \text{ layer}} D_j(\vec{q}) F_j(\vec{q}) e^{-i\vec{q}\cdot\langle\vec{r}_j\rangle},$$

then type-B layers will have an expected form factor of

$$B = \sum_{j \in \text{type-}A \text{ layer}} D_j(\vec{q}) F_j(\vec{q}) e^{-i\vec{q}\cdot(\langle\vec{r}_j\rangle - \vec{t}/3)} = e^{i\vec{q}\cdot\vec{t}/3} A,$$

and type-C layers will have an expected form factor of

$$C = \sum_{j \in \text{type-}A \text{ layer}} D_j(\vec{q}) F_j(\vec{q}) e^{-i\vec{q}\cdot(\langle\vec{r}_j\rangle + \vec{t}/3)} = e^{-i\vec{q}\cdot\vec{t}/3} A.$$

However, due to the periodic crystallinity in two dimensions, A is only nonzero when Eq. (A4) is satisfied. By definition of the reciprocal lattice, $\vec{G} \cdot \vec{t} = 2\pi m$ with $m \in \mathbb{Z}$. Hence, for all values of the scattering vector $\vec{q}$ where the expected layer form factors are nonzero, the expected layer form factors are simply related by the third roots of unity. In particular, when we look along a row in reciprocal space satisfying $\vec{q} \cdot \vec{t} = 2\pi m$ with $m \bmod 3 = 1$, the expected layer form factors are related by:

$$B = e^{i2\pi/3} A \quad \text{and} \quad C = e^{-i2\pi/3} A.$$

APPENDIX B: AUTOCORRELATION FOR PROCESSES GENERATED BY AUTONOMOUS HMMs

Let us derive the autocorrelation function in general and in closed form for the class of autonomous HMMs introduced in the main body. Helpfully, for particular models, the expressions become analytic in terms of the model parameters.

Directly calculating, we find that the autocorrelation function, for $\tau > 0$, for any such HMM is

$$
\begin{aligned}
\gamma(\tau) &= \langle \overline{X_t} X_{t+\tau} \rangle \\
&= \int_{x \in \mathcal{A}} \int_{x' \in \mathcal{A}} \bar{x} x' \, \mathrm{p}(X_0 = x, X_\tau = x') \, dx \, dx' \\
&= \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \int_{x \in \mathcal{A}} \int_{x' \in \mathcal{A}} \bar{x} x' \, \mathrm{p}(X_0 = x, X_\tau = x', \mathcal{S}_0 = s, \mathcal{S}_\tau = s') \, dx \, dx' \\
&= \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \int_{x \in \mathcal{A}} \int_{x' \in \mathcal{A}} \bar{x} x' \, \mathrm{Pr}(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') \, \mathrm{p}(X_0 = x | \mathcal{S}_0 = s) \, \mathrm{p}(X_\tau = x' | \mathcal{S}_\tau = s') \, dx \, dx' \\
&= \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \langle \boldsymbol{\pi} | s \rangle \langle s | T^\tau | s' \rangle \langle s' | \mathbf{1} \rangle \left[\int_{x \in \mathcal{A}} \bar{x} \, \mathrm{p}(x|s) \, dx \right] \left[\int_{x' \in \mathcal{A}} x' \, \mathrm{p}(x'|s') \, dx' \right] \\
&= \langle \boldsymbol{\pi} | \left[\sum_{s \in \mathcal{S}} \langle \overline{X} \rangle_{\mathrm{p}(X|s)} | s \rangle \langle s | \right] T^\tau \left[\sum_{s' \in \mathcal{S}} \langle X \rangle_{\mathrm{p}(X|s')} | s' \rangle \langle s' | \right] | \mathbf{1} \rangle,
\end{aligned}
$$

where the integrals are written in a form meant to be easily accessible but should generally be interpreted as Lebesgue integrals. In the above derivation, note that

$$\mathrm{p}(X_0 = x, X_\tau = x', \mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \mathrm{Pr}(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') \mathrm{p}(X_0 = x, X_\tau = x' | \mathcal{S}_0 = s, \mathcal{S}_\tau = s')$$

holds by definition of conditional probability. The decomposition of

$$\mathrm{p}(X_0 = x, X_\tau = x' | \mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \mathrm{p}(X_0 = x | \mathcal{S}_0 = s) \mathrm{p}(X_\tau = x' | \mathcal{S}_\tau = s')$$

for $\tau \neq 0$ follows from the conditional independence in the relevant Bayesian network shown in Fig. 3. Moreover, the equality

$$\mathrm{Pr}(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \langle \boldsymbol{\pi} | s \rangle \langle s | T^\tau | s' \rangle \langle s' | \mathbf{1} \rangle$$

can be derived by marginalizing over all possible intervening state sequences. We can use the hidden-state basis, where $|s\rangle$ is the column vector of all 0s except for a 1 at the index corresponding to state s, while $\langle s|$ is simply its transpose. This yields a natural decomposition of the identity operator: $I = \sum_{s \in \mathcal{S}} |s\rangle\langle s|$.

Since the autocorrelation is a Hermitian function—i.e., $\gamma(-\tau) = \overline{\gamma}(\tau)$—and

$$\gamma(0) = \langle |X|^2 \rangle_{\pi(X)} = \langle \boldsymbol{\pi} | \sum_{s \in \mathcal{S}} \langle |X|^2 \rangle_{\mathrm{p}(X|s)} | s \rangle,$$

we find the full autocorrelation function is given by:

$$\gamma(\tau) = \begin{cases} \langle \boldsymbol{\pi} | \Omega \, T^{|\tau|} \, \overline{\Omega} | \mathbf{1} \rangle & \text{if } \tau \leqslant 1 \\ \langle |x|^2 \rangle & \text{if } \tau = 0, \\ \langle \boldsymbol{\pi} | \overline{\Omega} \, T^{|\tau|} \, \Omega | \mathbf{1} \rangle & \text{if } \tau \geqslant 1 \end{cases} \tag{B1}$$

where Ω is the $|\mathcal{S}|$-by-$|\mathcal{S}|$ matrix defined by:

$$\Omega = \sum_{s \in \mathcal{S}} \langle X \rangle_{\mathrm{p}(X|s)} |s\rangle \langle s|.$$

The Ω matrix simply places state-conditioned average outputs along its diagonal.

To better understand the range of possible behaviors of autocorrelation, we can go a step further. In particular, we employ the general spectral decomposition of T^τ derived in Ref. [68] for nonnormal and potentially nondiagonalizable operators:

$$T^\tau = \left[\sum_{m=0}^{\nu_0 - 1} \delta_{\tau,m} T_{0,m} \right] + \sum_{\lambda \in \Lambda_T \backslash \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{\tau}{m} \lambda^{\tau - m} T_{\lambda,m}, \tag{B2}$$

where $\binom{\tau}{m}$ is the generalized binomial coefficient:

$$\binom{\tau}{m} = \frac{1}{m!} \prod_{n=1}^{m} (\tau - n + 1),$$

with $\binom{\tau}{0} = 1$. As briefly summarized in Sec. II D, Λ_T is the set of T's eigenvalues while T_λ is the spectral projection operator associated with the eigenvalue λ. Recall that ν_λ is the *index* of the eigenvalue λ, i.e., the size of the largest Jordan block associated with λ, and $T_{\lambda,m} = T_\lambda (T - \lambda I)^m$. Substituting Eq. (B2) into Eq. (B1) yields:

$$\gamma(\tau) = \left[\sum_{m=1}^{\nu_0 - 1} \delta_{\tau,m} \langle \boldsymbol{\pi} | \overline{\Omega}\, T_{0,m}\, \Omega | \mathbf{1} \rangle \right] + \sum_{\lambda \in \Lambda_T \backslash \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{\tau}{m} \lambda^{\tau - m} \langle \boldsymbol{\pi} | \overline{\Omega}\, T_{\lambda,m}\, \Omega | \mathbf{1} \rangle,$$

for $\tau > 0$.

It is significant that the zero eigenvalue contributes a qualitatively distinct ephemeral behavior to the autocorrelation while $|\tau| < \nu_0$. All other eigenmodes contribute products of polynomials times decaying exponentials in τ. When T is diagonalizable, the autocorrelation is simply a sum of decaying exponentials.

APPENDIX C: ANALYTICAL POWER SPECTRA

The following derives both the continuous and discrete part of the power spectrum for HMM-generated processes. The development parallels that in Ref. [70], although that derivation was restricted to the special case of diffraction patterns from Mealy (i.e., edge-emitting) HMMs with countable alphabets. In contrast, the following derives analytical expressions for the power spectrum of any stochastic process generated by an HMM. Notably, it also allows uncountably infinite alphabets. Also, it is developed for Moore (i.e., state-emitting) HMMs—although Mealy and Moore HMMs are class equivalent and can be easily transformed from one to the other.

1. Diffuse spectra

Recall Eq. (3):

$$P(\omega) = \lim_{N \to \infty} \frac{1}{N} \sum_{\tau=-N}^{N} (N - |\tau|) \gamma(\tau) e^{-i\omega\tau},$$

and Eq. (4)'s explicit expression for the correlation function:

$$\gamma(\tau) = \begin{cases} \langle \boldsymbol{\pi} | \Omega\, T^{|\tau|}\, \overline{\Omega} | \mathbf{1} \rangle & \text{if } \tau \leqslant 1 \\ \langle |x|^2 \rangle & \text{if } \tau = 0\,. \\ \langle \boldsymbol{\pi} | \overline{\Omega}\, T^{|\tau|}\, \Omega | \mathbf{1} \rangle & \text{if } \tau \geqslant 1 \end{cases}$$

From these we can rewrite the power spectrum directly in terms of the generating HMM's transition matrix:

$$P(\omega) = \langle |x|^2 \rangle + \lim_{N \to \infty} \frac{2}{N} \mathrm{Re} \sum_{\tau=1}^{N} (N - \tau) \langle \boldsymbol{\pi} | \overline{\Omega}\, T^\tau\, \Omega | \mathbf{1} \rangle e^{-i\omega\tau} = \langle |x|^2 \rangle + \lim_{N \to \infty} \frac{2}{N} \mathrm{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \left[\sum_{\tau=1}^{N} (N - \tau) T^\tau e^{-i\omega\tau} \right] \Omega | \mathbf{1} \rangle. \tag{C1}$$

We used the fact that $z + \bar{z} = 2\mathrm{Re}(z)$ for any $z \in \mathbb{C}$. For convenience, we introduce the variable $\mathfrak{z} \equiv e^{-i\omega}$. We then note that the summation splits:

$$\sum_{\tau=1}^{N} (N - \tau) T^\tau e^{-i\omega\tau} = N \sum_{\tau=1}^{N} (\mathfrak{z}T)^\tau - \sum_{\tau=1}^{N} \tau (\mathfrak{z}T)^\tau.$$

For positive integer N, it is always true that:

$$(I - \mathfrak{z}T) \sum_{\tau=1}^{N} (\mathfrak{z}T)^{\tau} = \mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1},$$

and

$$(I - \mathfrak{z}T) \sum_{\tau=1}^{N} \tau(\mathfrak{z}T)^{\tau} = -N\mathfrak{z}^{N+1}T^{N+1} + \sum_{\tau=1}^{N} (\mathfrak{z}T)^{\tau}.$$

Hence, whenever $I - \mathfrak{z}T$ is invertible (i.e., whenever $e^{i\omega} \notin \Lambda_T$), we have:

$$\sum_{\tau=1}^{N} (\mathfrak{z}T)^{\tau} = (I - \mathfrak{z}T)^{-1}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1}),$$

and

$$\sum_{\tau=1}^{N} \tau(\mathfrak{z}T)^{\tau} = (I - \mathfrak{z}T)^{-1}[-N\mathfrak{z}^{N+1}T^{N+1} + (I - \mathfrak{z}T)^{-1}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1})].$$

Together, this yields:

$$\begin{aligned}
\sum_{\tau=1}^{N} (N - \tau)T^{\tau}e^{-i\omega\tau} &= N \sum_{\tau=1}^{N} (\mathfrak{z}T)^{\tau} - \sum_{\tau=1}^{N} \tau(\mathfrak{z}T)^{\tau} \\
&= N(I - \mathfrak{z}T)^{-1}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1} + \mathfrak{z}^{N+1}T^{N+1}) - (I - \mathfrak{z}T)^{-2}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1}) \\
&= NT(\mathfrak{z}^{-1}I - T)^{-1} - (I - \mathfrak{z}T)^{-2}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1}).
\end{aligned}$$

Noting that $(\mathfrak{z}^{-1}I - T)^{-1} = (e^{i\omega}I - T)^{-1}$, this implies that the continuous (i.e., diffuse) part of the power spectrum becomes

$$\begin{aligned}
P_c(\omega) &= \langle |x|^2 \rangle + \lim_{N\to\infty} \frac{2}{N} \operatorname{Re}\langle \boldsymbol{\pi}|\overline{\Omega}\left[\sum_{\tau=1}^{N} (N - \tau)T^{\tau}e^{-i\omega\tau}\right]\Omega|\mathbf{1}\rangle \\
&= \langle |x|^2 \rangle + \lim_{N\to\infty} \frac{2}{N} \operatorname{Re}\langle \boldsymbol{\pi}|\overline{\Omega}[NT(\mathfrak{z}^{-1}I - T)^{-1} - (I - \mathfrak{z}T)^{-2}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1})]\Omega|\mathbf{1}\rangle \\
&= \langle |x|^2 \rangle + 2\operatorname{Re}\langle \boldsymbol{\pi}|\overline{\Omega}\,T(\mathfrak{z}^{-1}I - T)^{-1}\,\Omega|\mathbf{1}\rangle - \lim_{N\to\infty}\frac{2}{N}\operatorname{Re}\langle \boldsymbol{\pi}|\overline{\Omega}\,(I - \mathfrak{z}T)^{-2}(\mathfrak{z}T - \mathfrak{z}^{N+1}T^{N+1})\,\Omega|\mathbf{1}\rangle, \quad\quad (\text{C2}) \\
&= \langle |x|^2 \rangle + 2\operatorname{Re}\langle \boldsymbol{\pi}|\overline{\Omega}\,T(e^{i\omega}I - T)^{-1}\,\Omega|\mathbf{1}\rangle. \quad\quad (\text{C3})
\end{aligned}$$

Equation (C3) is the principle result, yielding the continuous part of the power spectrum in closed form. However, it is also worth noting that Eq. (C2) (without the $N \to \infty$ limit yet being taken) provides the exact result for the expected periodogram from finite length-N samples.

2. Discrete spectra

The transition dynamic's eigenvalues $\Lambda_{\rho(T)} = \{\lambda \in \Lambda_T : |\lambda| = 1\}$ on the unit circle are responsible for a power spectrum's Dirac δ functions. In the physical context of diffraction patterns, these δ functions are the familiar Bragg reflections. For finite length-N samples, eigenvalues on the unit circle give rise to Dirichlet kernels. As $N \to \infty$, the analysis simplifies since the Dirichlet kernels converge to δ functions.

The following derives the exact form of the δ-function contributions, showing how their presence and integrated magnitude can be calculated directly from the stochastic transition dynamic. Recall that the spectral projection operator $T_{\lambda,0}$ associated with the eigenvalue λ can be defined as the residue of $(zI - T)^{-1}$ as $z \to \lambda$:

$$T_{\lambda,0} = \frac{1}{2\pi i}\oint_{C_\lambda}(zI - T)^{-1}\,dz.$$

The spectral companion operators are

$$T_{\lambda,m} = T_{\lambda,0}(T - \lambda I)^m,$$

with the useful property that $T_{\lambda,m}T_{\zeta,n} = \delta_{\lambda,\zeta}T_{\lambda,m+n}$ and $T_{\lambda,m} = \mathbf{0}$ for $m \geqslant \nu_\lambda$. The index ν_λ of the eigenvalue λ is the size of the largest Jordan block associated with λ.

The Perron-Frobenius theorem guarantees that all eigenvalues on the unit circle have an index of one: i.e., $\nu_\lambda = 1$ for all $\lambda \in \Lambda_{\rho(T)}$. This means that the algebraic and geometric multiplicities of these eigenvalues coincide and they are all associated with diagonalizable subspaces.

Taking advantage of the index-one nature of the eigenvalues on the unit circle, and using the shorthand $T_\lambda \equiv T_{\lambda,0}$ for the spectral projection operators, we define:

$$\Theta \equiv \sum_{\lambda \in \Lambda_{\rho(T)}} \lambda T_\lambda$$

and

$$F \equiv T - \Theta.$$

We then consider how the spectral decomposition of T^τ splits into contributions from these two independent components: From Ref. [68], and employing the simplifying notation that $0^{\tau - m} = \delta_{\tau - m, 0}$, we find:

$$T^\tau = \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} \lambda^{\tau - m} \binom{\tau}{m} T_{\lambda, m} = \left(\sum_{\lambda \in \Lambda_{\rho(T)}} \lambda^\tau T_\lambda \right) + \left[\sum_{\lambda \in \Lambda_T \setminus \Lambda_{\rho(T)}} \sum_{m=0}^{\nu_\lambda - 1} \lambda^{\tau - m} \binom{\tau}{m} T_{\lambda, m} \right] = \Theta^\tau + F^\tau,$$

where $\binom{\tau}{m} = \frac{1}{m!} \prod_{n=1}^{m} (\tau - n + 1)$ is the generalized binomial coefficient.

As the sequence length $N \to \infty$, the summation over τ in Eq. (C1) divided by the sequence length becomes

$$\lim_{N \to \infty} \sum_{\tau=1}^{N} \frac{N - \tau}{N} T^\tau e^{-i\omega\tau} = \sum_{\tau=1}^{\infty} T^\tau e^{-i\omega\tau} = \left(\sum_{\tau=1}^{\infty} \Theta^\tau e^{-i\omega\tau} \right) + \left(\sum_{\tau=1}^{\infty} F^\tau e^{-i\omega\tau} \right). \tag{C4}$$

In Eq. (C4), only the summation involving Θ is capable of contributing δ functions. Expanding that sum yields:

$$\sum_{\tau=1}^{\infty} \Theta^\tau e^{-i\omega\tau} = \sum_{\lambda \in \Lambda_{\rho(T)}} T_\lambda \sum_{\tau=1}^{\infty} (\lambda e^{-i\omega})^\tau = \sum_{\lambda \in \Lambda_{\rho(T)}} T_\lambda \left[-1 + \sum_{\tau=0}^{\infty} e^{i(\omega_\lambda - \omega)\tau} \right]$$

$$= \sum_{\lambda \in \Lambda_{\rho(T)}} T_\lambda \left[\frac{-1}{1 - e^{i(\omega - \omega_\lambda)}} + \sum_{k=-\infty}^{\infty} \pi \, \delta(\omega - \omega_\lambda + 2\pi k) \right], \tag{C5}$$

where ω_λ is related to λ by $\lambda = e^{i\omega_\lambda}$. The last line is obtained using well-known properties of the discrete-time Fourier transform [151].

From Eqs. (C1), (C4), and (C5), we find that the potential δ function at ω_λ (and its 2π-periodic offsets) has integrated magnitude:

$$\Delta_\lambda \equiv \lim_{\epsilon \to 0} \int_{\omega_\lambda - \epsilon}^{\omega_\lambda + \epsilon} P(\omega) \, d\omega$$

$$= \lim_{\epsilon \to 0} \int_{\omega_\lambda - \epsilon}^{\omega_\lambda + \epsilon} 2 \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \left(\lim_{N \to \infty} \sum_{\tau=1}^{N} \frac{N - \tau}{N} T^\tau e^{-i\omega\tau} \right) \Omega | \mathbf{1} \rangle \, d\omega$$

$$= \lim_{\epsilon \to 0} \int_{\omega_\lambda - \epsilon}^{\omega_\lambda + \epsilon} 2 \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \left(\sum_{\tau=1}^{\infty} \Theta^\tau e^{-i\omega\tau} \right) \Omega | \mathbf{1} \rangle \, d\omega$$

$$= \lim_{\epsilon \to 0} \int_{\omega_\lambda - \epsilon}^{\omega_\lambda + \epsilon} 2 \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \sum_{\zeta \in \Lambda_{\rho(T)}} T_\zeta \left[\frac{-1}{1 - e^{i(\omega - \omega_\zeta)}} + \sum_{k=-\infty}^{\infty} \pi \, \delta(\omega - \omega_\zeta + 2\pi k) \right] \Omega | \mathbf{1} \rangle \, d\omega$$

$$= 2\pi \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \, T_\lambda \, \Omega | \mathbf{1} \rangle \lim_{\epsilon \to 0} \int_{\omega_\lambda - \epsilon}^{\omega_\lambda + \epsilon} \delta(\omega - \omega_\lambda) \, d\omega$$

$$= 2\pi \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \, T_\lambda \, \Omega | \mathbf{1} \rangle. \tag{C6}$$

Finally, from Eq. (C6) and the 2π-periodicity of the power spectrum, we obtain the full discrete (i.e., δ function) contribution to the power spectrum:

$$P_d(\omega) = \sum_{\lambda \in \Lambda_{\rho(T)}} 2\pi \operatorname{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \, T_\lambda \Omega | \mathbf{1} \rangle \sum_{k=-\infty}^{\infty} \delta(\omega - \omega_\lambda + 2\pi k). \tag{C7}$$

APPENDIX D: A NEW CONDITION FOR $1/f$ NOISE

Here we obtain a sufficient condition for $1/f$ noise.

Equation (14) gave the general formula for power spectra from continuous-time processes:

$$P_c(f) = \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_\lambda-1} 2\operatorname{Re} \frac{\langle \boldsymbol{\pi}|\overline{\Omega}\, G_{\lambda,m}\Omega|\mathbf{1}\rangle}{(i2\pi f - \lambda)^{m+1}}.$$

We now restrict attention to diagonalizable transition rate operators. To simplify notation, we relabel the spectral intensity as $c_\lambda \equiv \langle \boldsymbol{\pi}|\overline{\Omega}\, G_{\lambda,0}\Omega|\mathbf{1}\rangle$. Recall the following.

Definition 1. A continuous-time process has *doubly harmonic diminution* if:

(i) its generator of time evolution G is diagonalizable and has $N+1$ evenly spaced eigenvalues along the real line $\Lambda_G = \{-na\}_{n=0}^N$ for some $a > 0$, and

(ii) its spectral intensity fades with increasing frequency according to $c_{-na} = c/n$ for $n \geqslant 1$ and some $c \in \mathbb{R}$.

We will show that any process with doubly harmonic diminution produces $1/f$ noise over a frequency bandwidth proportional to N.

For a process with doubly harmonic diminution, the power spectrum simplifies considerably to:

$$\begin{aligned}
P_c(f) &= \sum_{\lambda \in \Lambda_G} 2\operatorname{Re} \frac{c_\lambda}{i2\pi f - \lambda} \\
&= \sum_{n=1}^N 2\operatorname{Re} \frac{c/n}{i2\pi f + na} \\
&= \frac{2c}{a} \sum_{n=1}^N \frac{1}{n^2 + \left(\frac{2\pi f}{a}\right)^2}. \quad \text{(D1)}
\end{aligned}$$

By considering various limits, we see that Eq. (D1) leads to nearly perfect $1/f$ noise over a significant bandwidth.

1. Constant spectrum for $f \ll a/2\pi$

If $2\pi f \ll a$, then $1 + \left(\frac{2\pi f}{na}\right)^2 \approx 1$ for all $n \geqslant 1$. Accordingly:

$$\begin{aligned}
P(f) &= \frac{2c}{a} \sum_{n=1}^N \frac{1}{n^2\left[1 + \left(\frac{2\pi f}{na}\right)^2\right]} \\
&\approx \frac{2c}{a} \sum_{n=1}^N \frac{1}{n^2} = \frac{2c}{a} H_{N,2} \\
&\to \frac{c\pi^2}{3a} \quad \text{as } N \to \infty,
\end{aligned}$$

where $H_{N,2} = \sum_{n=1}^N \frac{1}{n^2}$ is a generalized harmonic number. Notably, $H_{N,2} \to \pi^2/6$ as $N \to \infty$.

2. $1/f^2$ spectrum for $f \gg Na/2\pi$

If $2\pi f \gg Na$, then $2\pi f \gg na$ and $1 + \left(\frac{na}{2\pi f}\right)^2 \approx 1$ for all $n \leqslant N$. Accordingly:

$$\begin{aligned}
P(f) &= \frac{2c}{a} \sum_{n=1}^N \frac{1}{\left(\frac{2\pi f}{a}\right)^2\left[1 + \left(\frac{na}{2\pi f}\right)^2\right]} \\
&\approx \frac{ac}{2\pi^2 f^2} \sum_{n=1}^N 1 = \frac{caN}{2\pi^2 f^2}.
\end{aligned}$$

3. $1/f$ spectrum for $\frac{a}{2\pi} \ll f \ll \frac{Na}{2\pi}$

If $2\pi f \ll Na$, then $2\pi f \ll na$ and $1 + \left(\frac{2\pi f}{na}\right)^2 \approx 1$ for any $n \geqslant N$. Then:

$$\begin{aligned}
P(f) &= \frac{2c}{a}\left(\left[\sum_{n=1}^\infty \frac{1}{n^2 + \left(\frac{2\pi f}{a}\right)^2}\right] - \left\{\sum_{n=N+1}^\infty \frac{1}{n^2\left[1 + \left(\frac{2\pi f}{na}\right)^2\right]}\right\}\right) \\
&\approx \frac{2c}{a}\left\{\left[\sum_{n=1}^\infty \frac{1}{n^2 + \left(\frac{2\pi f}{a}\right)^2}\right] - \left(\sum_{n=N+1}^\infty \frac{1}{n^2}\right)\right\} \\
&= \frac{2c}{a}\left\{\left[\sum_{n=1}^\infty \frac{1}{n^2 + \left(\frac{2\pi f}{a}\right)^2}\right] - \left(\frac{\pi^2}{6} - H_{N,2}\right)\right\}.
\end{aligned}$$

With the identity

$$\sum_{n=1}^\infty \frac{1}{n^2 + \left(\frac{2\pi f}{a}\right)^2} = \frac{a\coth(2\pi^2 f/a)}{4f} - \frac{1}{2\left(\frac{2\pi f}{a}\right)^2},$$

this yields:

$$P(f) \approx \frac{c\coth(2\pi^2 f/a)}{2f} - \frac{ac}{(2\pi f)^2} - \frac{2c}{a}\left(\frac{\pi^2}{6} - H_{N,2}\right)$$

for $2\pi f \ll Na$.

For $f > \frac{a}{2\pi^2}$, the hyperbolic cotangent $\coth(2\pi^2 f/a)$ quickly converges to unity. Hence, for $\frac{a}{2\pi^2} \ll f \ll \frac{Na}{2\pi}$, the power spectrum is well approximated by:

$$\begin{aligned}
P(f) &\approx \frac{c}{2f}\left(1 - \frac{a}{2\pi^2 f}\right) - \frac{2c}{a}\left(\frac{\pi^2}{6} - H_{N,2}\right) \\
&\approx \frac{c}{2f} - \frac{2c}{a}\left(\frac{\pi^2}{6} - H_{N,2}\right).
\end{aligned}$$

Moreover, $\frac{\pi^2}{6} - H_{N,2} \to 0$ as $N \to \infty$.

4. Combining the regimes

We showed that any process with doubly harmonic diminution has three distinctive regimes in its power spectrum: nearly constant for very low frequency, $1/f$ decay over a broad bandwidth, and $1/f^2$ decay at very large frequencies.

The transition frequencies between these three behavior regimes is found more specifically by looking for the crossover frequencies—f^* where the constant and $1/f$ approximations meet, and f^{**} where the $1/f$ and $1/f^2$ approximations meet.

The first transition frequency f^*, from constant to $1/f$ behavior, satisfies $\frac{2c}{a}H_{N,2} = \frac{c}{2f^*} - \frac{2c}{a}\left(\frac{\pi^2}{6} - H_{N,2}\right)$. We find that:

$$f^* = \frac{3a}{2\pi^2}. \quad \text{(D2)}$$

The second transition frequency f^{**}, from $1/f$ behavior to $1/f^2$ behavior, satisfies:

$$\frac{c}{2f^{**}} - \frac{2c}{a}\left(\frac{\pi^2}{6} - H_{N,2}\right) = \frac{acN}{2\pi^2(f^{**})^2}.$$

We find that:

$$f^{**} = \frac{a}{8\left(\frac{\pi^2}{6} - H_{N,2}\right)}\left[1 - \sqrt{1 - \frac{8}{3}N\left(1 - \frac{6H_{N,2}}{\pi^2}\right)}\right].$$

This exact expression for f^{**} can be expanded in terms of the small parameter:

$$\epsilon = \frac{8}{3}N\left(1 - \frac{6H_{N,2}}{\pi^2}\right)$$

such that:

$$
\begin{aligned}
f^{**} &= \frac{aN}{2\pi^2 \epsilon}[1 - (1-\epsilon)^{1/2}] \\
&= \frac{aN}{2\pi^2}\sum_{k=1}^{\infty}\binom{1/2}{k}(-\epsilon)^{k-1} \\
&= \frac{aN}{4\pi^2}\left[1 - \frac{1}{4}\epsilon + \frac{1}{8}\epsilon^2 - \mathcal{O}(\epsilon^3)\right],
\end{aligned}
$$

that, to first order, yields the approximation $f^{**} \approx \frac{aN}{4\pi^2}$.

Altogether, this leads to:

$$
P(f) \approx
\begin{cases}
\frac{2c}{a}H_{N,2} & \text{if } f < \frac{3a}{2\pi^2} \\
\frac{c}{2f} - \frac{2c}{a}\left(\frac{\pi^2}{6} - H_{N,2}\right) & \text{if } \frac{3a}{2\pi^2} < f < f^{**} \\
\frac{caN}{2\pi^2 f^2} & \text{if } f > f^{**}
\end{cases}
$$

or, more simply:

$$
P(f) \sim
\begin{cases}
\text{constant} & \text{if } f < \frac{3a}{2\pi^2} \\
1/f & \text{if } \frac{3a}{2\pi^2} < f \lesssim \frac{aN}{4\pi^2} \\
1/f^2 & \text{if } f \gtrsim \frac{aN}{4\pi^2}
\end{cases}.
$$

APPENDIX E: BROWNIAN NOISE

Here we show how to recover the power spectrum of Brownian motion using the tools of Sec. II F. This simple example indicates how to leverage the tools more generally to analyze the power spectra of more sophisticated Langevin-type differential equations that can transduce arbitrarily sophisticated noise models.

Each spatial dimension of a Brownian trajectory behaves independently and simply integrates white noise. In the discrete-time case, the fundamental equation for Brownian noise is

$$Y_t - Y_{t-1} = X_t, \tag{E1}$$

where X_t is a Gaussian white noise of variance $\sigma^2 = 2D\tau_0$, where D is the diffusion coefficient, which implies $P_{XX}(\omega) = \sigma^2/f_0 = 2D/f_0^2$. Equation (E1) corresponds to $\mathcal{P}(\mathcal{D}) = \mathcal{D}^0 - \mathcal{D}$ and $\mathcal{Q}(\mathcal{D}) = \mathcal{D}^0$, which leads to:

$$
\begin{aligned}
|H_{X \to Y}(\omega)|^2 &= \frac{|\mathcal{Q}(e^{i\omega})|^2}{|\mathcal{P}(e^{i\omega})|^2} \\
&= \frac{1}{|1 - e^{i\omega}|^2} = \frac{1}{2[1 - \cos(\omega)]}
\end{aligned} \tag{E2}
$$

and

$$
\begin{aligned}
P_{YY}(\omega) &= \frac{2D/f_0^2}{2[1 - \cos(\omega)]} \tag{E3} \\
&= \frac{D}{2\pi^2 f^2\left\{1 - \frac{\pi^2}{3}\left(\frac{f}{f_0}\right)^2 + \mathcal{O}\left[\left(\frac{f}{f_0}\right)^4\right]\right\}}, \tag{E4} \\
&\to \frac{D}{2\pi^2 f^2} \quad \text{as } \frac{f}{f_0} \to 0. \tag{E5}
\end{aligned}
$$

The last line gives the limiting power spectrum of Brownian noise in the continuous-time case, where it is well known that $P_{YY}(f) \propto 1/f^2$.

It is worth noting that, at finite sampling frequency, the experimentally or numerically obtained power spectrum deviates significantly from the $1/f^2$ spectrum as $f \to f_0/2$, according to Eq. (E4). This could lead to misidentifying $1/f^{\alpha}$ noise.

APPENDIX F: PROOF OF LEMMA 1

Recall Lemma 1:

Any stochastic process (not necessarily stationary) with the SCIP:

$$\Pr(X_t | X_{t'} = x) = \Pr(X_t) = \Pr(X_{t'}),$$

for all $x \in \mathcal{A}$ and all $t \neq t'$, generates a flat power spectrum, mimicking white noise.

Proof. For any such process, $\Pr(X_t)$ is the stationary distribution $\boldsymbol{\mu}_X$ of the instantaneous observable under the stochastic dynamic. Moreover, SCIP means that the joint probability of any two observations decomposes:

$$
\begin{aligned}
\Pr(X_t = x, X_{t+\tau} = x') &= \Pr(X_{t+\tau} = x' | X_t = x)\Pr(X_t = x) \\
&= \Pr(X_{t+\tau} = x')\Pr(X_t = x) \\
&= \boldsymbol{\mu}_X(x')\boldsymbol{\mu}_X(x).
\end{aligned}
$$

Substituting $\boldsymbol{\mu}_X(x')\boldsymbol{\mu}_X(x)$ for $\Pr(X_t = x, X_{t+\tau} = x')$ in the autocorrelation definition of Eq. (2) immediately implies that SCIP processes have τ-independent pairwise correlation $\gamma(\tau) = |\langle x \rangle|^2$ for $\tau \neq 0$. The power spectrum is thus flat over all frequencies, except possibly with a δ function at $\omega = 0$.

APPENDIX G: PROOF OF THEOREM 2

We define the set Ξ of average outputs exhibited by the states: $\Xi \equiv \bigcup_{s \in \mathcal{S}}\{\langle x \rangle_{\text{p}(X|s)}\}$. Furthermore, we define $\boldsymbol{\mathcal{S}}_{\xi} \subset \boldsymbol{\mathcal{S}}$ as the set of states that all exhibit the same average output $\xi \in \Xi$. Explicitly, $\boldsymbol{\mathcal{S}}_{\xi} \equiv \{s \in \boldsymbol{\mathcal{S}} : \langle x \rangle_{\text{p}(X|s)} = \xi\}$.

Recall Theorem 2:

Let $\{X_t\}_t$ be a stochastic process generated by a hidden-state model $\mathcal{M}(\vec{m})$. X_t is the random variable for the observable at time t, and $\mathcal{S}_t$ is the random variable for the hidden state at time t. Such processes have constant autocorrelation and a flat power spectrum *if*:

$$
\begin{aligned}
\Pr(\mathcal{S}_{t+\tau} \in \boldsymbol{\mathcal{S}}_{\xi'} | \mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi}) &= \Pr(\mathcal{S}_{t+\tau} \in \boldsymbol{\mathcal{S}}_{\xi'}) \\
&= \Pr(\mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi'}), \tag{G1}
\end{aligned}
$$

for all separations $\tau > 0$, for all $t \in \mathcal{T}$, and for all $\xi, \xi' \in \Xi$.

Proof. Starting from Eq. (17), we find the autocorrelation for all such processes (for $\tau \geqslant 1$):

$$\gamma(\tau) = \left\langle \langle \overline{x} \rangle_{\text{p}(X|\mathcal{S}_t)} \langle x \rangle_{\text{p}(X|\mathcal{S}_{t+\tau})} \right\rangle_{\text{Pr}(\mathcal{S}_t, \mathcal{S}_{t+\tau})}$$

$$= \sum_{s,s' \in \mathcal{S}} \text{Pr}(\mathcal{S}_t = s, \mathcal{S}_{t+\tau} = s') \langle \overline{x} \rangle_{\text{p}(X|s)} \langle x \rangle_{\text{p}(X|s')}$$

$$= \sum_{\xi, \xi' \in \Xi} \text{Pr}(\mathcal{S}_t \in \mathcal{S}_\xi, \mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) \overline{\xi} \, \xi'$$

$$= \sum_{\xi \in \Xi} \text{Pr}(\mathcal{S}_t \in \mathcal{S}_\xi) \overline{\xi}$$

$$\times \sum_{\xi' \in \Xi} \text{Pr}(\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'} | \mathcal{S}_t \in \mathcal{S}_\xi) \xi'. \qquad \text{(G2)}$$

Combining Eqs. (G1) and (G2), we see that:

$$\gamma(\tau) = \sum_{\xi \in \Xi} \text{Pr}(\mathcal{S}_t \in \mathcal{S}_\xi) \overline{\xi} \sum_{\xi' \in \Xi} \text{Pr}(\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) \xi' = |\langle \xi \rangle|^2,$$

which is a constant. With the same reasoning, we likewise find that $\gamma(\tau) = |\langle \xi \rangle|^2$ for $\tau \leqslant -1$. The autocorrelation function is thus $\gamma(\tau) = |\langle \xi \rangle|^2 + c\delta_{\tau,0}$, where $c \equiv \gamma(0) - |\langle \xi \rangle|^2$ is a constant. Thus, the power spectrum is flat, *if* Eq. (G1) holds.

APPENDIX H: PROOF OF THEOREM 3

Recall Theorem 3: Let $\{X_t\}_t$ and $\{Y_t\}_t$ be two stochastic processes generated by any of the hidden-state models $\mathcal{M}(\vec{m})$ discussed above, including autonomous HMMs and input-dependent generators, X_t and Y_t the random variables for the observables at time t, and $\mathcal{S}_t \in \mathcal{S}$ and $\mathcal{R}_t \in \mathcal{R}$ the random variables for the respective hidden states at time t. These processes have identical power spectra, up to a constant offset, if:

$$\text{Pr}(\mathcal{S}_t \in \mathcal{S}_\xi, \mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) = \text{Pr}(\mathcal{R}_t \in \mathcal{R}_\xi, \mathcal{R}_{t+\tau} \in \mathcal{R}_{\xi'}),$$

for all separations $\tau > 0$, for all $t \in \mathcal{T}$, and for all $\xi, \xi' \in \Xi$, which is the set of average outputs emitted by the states.

Proof. Let $\gamma(\tau)$ be the autocorrelation function for the first process $\{X_t\}_t$, and let $\gamma'(\tau)$ be the autocorrelation function for the second process $\{Y_t\}_t$. Assume:

$$\text{Pr}(\mathcal{S}_t \in \mathcal{S}_\xi, \mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) = \text{Pr}(\mathcal{R}_t \in \mathcal{R}_\xi, \mathcal{R}_{t+\tau} \in \mathcal{R}_{\xi'})$$

for all separations $\tau > 0$, for all $t \in \mathcal{T}$, and for all $\xi, \xi' \in \Xi$. Then, starting from Eq. (17), we find the autocorrelation for the first process (for $\tau \geqslant 1$):

$$\gamma(\tau) = \left\langle \langle \overline{x} \rangle_{\text{p}(X|\mathcal{S}_t)} \langle x \rangle_{\text{p}(X|\mathcal{S}_{t+\tau})} \right\rangle_{\text{Pr}(\mathcal{S}_t, \mathcal{S}_{t+\tau})}$$

$$= \sum_{s,s' \in \mathcal{S}} \text{Pr}(\mathcal{S}_t = s, \mathcal{S}_{t+\tau} = s') \langle \overline{x} \rangle_{\text{p}(X|s)} \langle x \rangle_{\text{p}(X|s')}$$

$$= \sum_{\xi, \xi' \in \Xi} \text{Pr}(\mathcal{S}_t \in \mathcal{S}_\xi, \mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}) \overline{\xi} \, \xi'$$

$$= \sum_{\xi, \xi' \in \Xi} \text{Pr}(\mathcal{R}_t \in \mathcal{R}_\xi, \mathcal{R}_{t+\tau} \in \mathcal{R}_{\xi'}) \overline{\xi} \, \xi'$$

$$= \sum_{r,r' \in \mathcal{R}} \text{Pr}(\mathcal{R}_t = r, \mathcal{R}_{t+\tau} = r') \langle \overline{x} \rangle_{\text{p}(X|r)} \langle x \rangle_{\text{p}(X|r')}$$

$$= \left\langle \langle \overline{x} \rangle_{\text{p}(X|\mathcal{R}_t)} \langle x \rangle_{\text{p}(X|\mathcal{R}_{t+\tau})} \right\rangle_{\text{Pr}(\mathcal{R}_t, \mathcal{R}_{t+\tau})}$$

$$= \gamma'(\tau).$$

With the same reasoning, we find that $\gamma(\tau) = \gamma'(\tau)$ for $\tau \leqslant -1$. Hence, the autocorrelations for the two processes agree everywhere except possibly at $\tau = 0$.

Define the constant $c \equiv \gamma(0) - \gamma'(0)$. The autocorrelation functions for the two processes are then related by $\gamma(\tau) = \gamma'(\tau) + c\delta_{\tau,0}$ for all τ. It then follows that the power spectrum of the processes differ at most by a constant offset.

APPENDIX I: DIFFRACTION PATTERNS OF CHAOTIC CRYSTALS FROM HMMs

Let us analyze two examples of HMM-designed chaotic crystals.

1. Example One

Consider a *p*-parametrized stochastic process for the stacking of layers of a close-packed structure. The stochastic stacking process is described by a HMM, where the transition matrix and average-observation matrix are as follows:

$$T = \begin{bmatrix} 0 & 1 & 0 \\ p & 0 & 1-p \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \Omega = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix},$$

respectively. For $p = 1$, we recover the deterministic period-2 hcp structure. For $p = 0$, we recover the deterministic period-3 ccp structure. For other values of p, the structure is described by a stochastic stacking process.

For any p, the eigenvalues of the transition matrix are $\Lambda_T = \{1, -\frac{1}{2} \pm \sqrt{p - \frac{3}{4}}\}$. The transition matrix is diagonalizable unless $p = 3/4$, where it becomes nondiagonalizable.

We aim to calculate the diffracted intensity for any p in closed form via Eq. (10):

$$P_c(\omega) = \langle |x|^2 \rangle + \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} 2 \, \text{Re} \frac{\langle \boldsymbol{\pi} | \overline{\Omega} \, T \, T_{\lambda,m} \Omega | \mathbf{1} \rangle}{(e^{i\omega} - \lambda)^{m+1}}$$

and Eq. (11):

$$P_d(\omega) = \sum_{k=-\infty}^{\infty} \sum_{\substack{\lambda \in \Lambda_T \\ |\lambda|=1}} 2\pi \, \delta(\omega - \omega_\lambda + 2\pi k) \, \text{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \, T_\lambda \Omega | \mathbf{1} \rangle.$$

We note that $\langle |x|^2 \rangle = \langle |\psi|^2 \rangle = |\psi|^2$.

For $p \neq 3/4$, the continuous spectrum simplifies to:

$$P_c(\omega) = |\psi|^2 + \sum_{\lambda \in \Lambda_T} 2 \, \text{Re} \frac{\lambda \langle \boldsymbol{\pi} | \overline{\Omega} \, T_\lambda \Omega | \mathbf{1} \rangle}{e^{i\omega} - \lambda},$$

and each spectral projection operator is given by $T_\lambda = |\lambda\rangle\langle\lambda|$, with:

$$\langle\lambda| = \frac{1}{3\lambda^2 - p} [\lambda \quad 1 \quad \lambda^2 - p] \quad \text{and} \quad |\lambda\rangle = [\lambda \quad \lambda^2 \quad 1]^\top,$$

where $\top$ denotes transposition. Recall that the stationary distribution is the left eigenvector $\langle \boldsymbol{\pi} | = \langle 1 | = \frac{1}{3-p}[1 \quad 1 \quad 1-p]$.

From these elements, we can calculate the spectral intensity:

$$\langle \boldsymbol{\pi} | \overline{\Omega}\, T_\lambda \Omega | \mathbf{1} \rangle = \frac{|\psi|^2}{(3-p)(3\lambda^2-p)} \begin{bmatrix} 1 & 1 & 1-p \end{bmatrix} \begin{bmatrix} 1 & & \\ & e^{-i2\pi/3} & \\ & & e^{i2\pi/3} \end{bmatrix} \begin{bmatrix} \lambda \\ \lambda^2 \\ 1 \end{bmatrix} \begin{bmatrix} \lambda & 1 & \lambda^2-p \end{bmatrix} \begin{bmatrix} 1 & & \\ & e^{i2\pi/3} & \\ & & e^{-i2\pi/3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \frac{|\psi|^2}{(3-p)(3\lambda^2-p)} [\lambda + \lambda^2 e^{-i2\pi/3} + (1-p)e^{i2\pi/3}][\lambda + e^{i2\pi/3} + (\lambda^2-p)e^{-i2\pi/3}] \tag{I1}$$

for any $\lambda \in \Lambda_T$ and for any $p \neq 3/4$.

For $\lambda = 1$, Eq. (I1) reduces to:

$$\langle \boldsymbol{\pi} | \overline{\Omega}\, T_1 \Omega | \mathbf{1} \rangle = \frac{p^2 |\psi|^2}{(3-p)^2}, \tag{I2}$$

where we have used the identity $1 + e^{i2\pi/3} + e^{-i2\pi/3} = 0$. Equation (I2) is in fact valid for any $p \in [0, 1]$.

For $p \in (0, 1)$, transition matrix T only has one eigenvalue on the unit circle, so the discrete (Bragg) spectrum has a single contribution from the eigenvalue of unity:

$$P_d(\omega) = \frac{2\pi p^2 |\psi|^2}{(3-p)^2} \sum_{k=-\infty}^{\infty} \delta(\omega + 2\pi k). \tag{I3}$$

Although not resulting from a deterministic periodicity, this Bragg reflection can nevertheless be regarded as a result of spatial periodicity in probabilistic behavior.

In fact, for any $p > 0$, the top-left panel of Fig. 9 shows that orientations A and B are more common than orientation C. However, Eq. (I3) survives a cyclic permutation of the alphabet (i.e., $A \mapsto B$, $B \mapsto C$, and $C \mapsto A$). So, this Bragg reflection persists even in multicrystalline materials—where each component chaotic crystal, with its own absolute orientation, is stacked according to either the process in Fig. 9 or one of its cyclic permutations.

There is a diffuse contribution to the power spectrum for all $p \in (0, 1)$. For $p \in (0, 3/4) \cup (3/4, 1)$, this contribution is as follows:

$$P_c(\omega) = |\psi|^2 \left[1 - \frac{p^2}{(3-p)^2} \right] + \sum_{\lambda \in \Lambda_T \setminus \{1\}} 2\,\mathrm{Re}\, \frac{\langle \boldsymbol{\pi} | \overline{\Omega}\, T_\lambda \Omega | \mathbf{1} \rangle}{e^{i\omega}/\lambda - 1}.$$

Expanding this via Eq. (I1) initially appears unwieldy, but the expressions can be simplified as soon as one recognizes that $\lambda^2 = p - 1 - \lambda$ for $\lambda \in \Lambda_T \setminus \{1\}$. Further simplification leverages the properties of the eigenvalues in the distinct regimes of $p > 3/4$ and $p < 3/4$. For $p > 3/4$, all eigenvalues have distinct real values. For $p < 3/4$, the two nonunity eigenvalues are complex conjugate pairs and, accordingly, have the same real part $[\mathrm{Re}(\lambda) = -1/2]$ and the same magnitude ($|\lambda| = \sqrt{1-p}$), with angular frequencies $\omega_\lambda = \pi \pm \arctan(\sqrt{3 - 4p})$.

Figure 12 shows the "corrected" diffraction pattern $P(\omega)/|\psi|^2$ for $p = 1/2$. There is a Bragg reflection at $\omega = 2\pi n$ (for all $n \in \mathbb{Z}$) due to the eigenvalue of unity. The nonunity eigenvalues $\lambda = -1/2 \pm i/2$ appear at angular frequencies $\omega_\lambda \in \{3\pi/4, 5\pi/4\}$. The Lorentzian line profile contributed at $\omega_\lambda = 3\pi/4$ is prominent. There is a local feature around $\omega_\lambda = 5\pi/4$, but it is more nuanced since it is not a peak in the diffracted intensity. Rather, the contribution from the eigenvalue at $\omega_\lambda = 5\pi/4$ primarily *depresses* the diffrac-

tion pattern around it, which allows for the zero at $\omega = 4\pi/3$. Diffracted intensity is forbidden at $4\pi/3$ since the stochastic process does not allow for a full anti-cyclic sequence of layers; i.e., *CBA*, *BAC*, and *ACB* are all forbidden sequences.

For $p = 1$, the continuous spectrum vanishes while the discrete spectrum picks up another Bragg reflection at $\omega = \pi$ with intensity $\langle \boldsymbol{\pi} | \overline{\Omega}\, T_{-1} \Omega | \mathbf{1} \rangle = \frac{3}{4} |\psi|^2$, yielding the diffraction pattern for a 2H hcp crystal:

$$P(\omega) = 2\pi |\psi|^2 \sum_{k=-\infty}^{\infty} \left[\frac{1}{4} \delta(\omega + 2\pi k) + \frac{3}{4} \delta(\omega - \pi + 2\pi k) \right].$$

Similarly, for $p = 0$, the continuous spectrum vanishes while the discrete spectrum picks up a Bragg reflection at $\omega = 2\pi/3$ with intensity $\langle \boldsymbol{\pi} | \overline{\Omega}\, T_{e^{i2\pi/3}} \Omega | \mathbf{1} \rangle = |\psi|^2$, yielding the diffraction pattern for a 3C$^+$ ccp crystal:

$$P(\omega) = 2\pi |\psi|^2 \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi/3 + 2\pi k).$$

Notice from Eq. (I2) that the former Bragg reflection at $\omega = 0$ has vanished at $p = 0$.

For $p = 3/4$, the transition matrix is nondiagonalizable. Since the spectral projection operators always sum to the
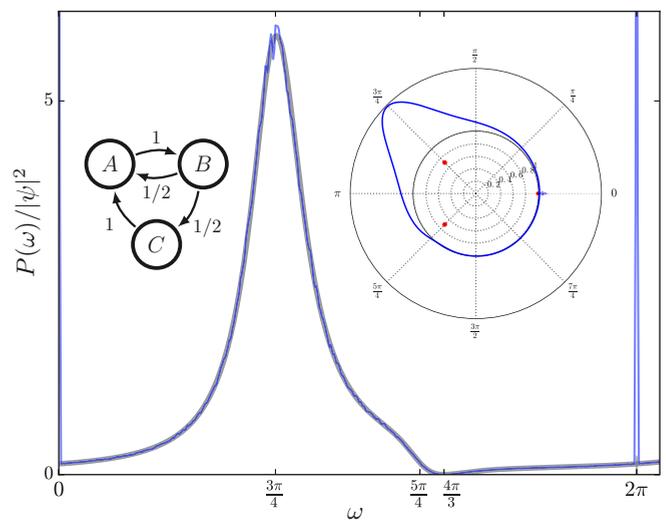
FIG. 12. Example One stochastic stacking process at $p = 1/2$ (left inset) and its diffraction pattern. Main: Numerical diffraction pattern (thin blue line) generated from a sampled stacking sequence of 2^{20} layers, using the Welch method to calculate the power spectrum on subsamples of length 2^9. It closely matches the thick gray line, which is the analytic solution for the diffracted intensity. Right inset: HMM stacking process diffraction pattern and eigenvalues, as a coronal spectrogram.

identity, we can calculate $T_{-1/2}$ easily via $T_{-1/2} = I - |\mathbf{1}\rangle\langle\boldsymbol{\pi}|$, with $\langle\boldsymbol{\pi}| = \frac{1}{9}[4 \quad 4 \quad 1]$, which yields:

$$T_{-1/2} = \frac{1}{9}\begin{bmatrix} 5 & -4 & -1 \\ -4 & 5 & -1 \\ -4 & -4 & 8 \end{bmatrix}.$$

The spectral companion operator $T_{-1/2,1}$ is then found as:

$$T_{-1/2,1} = T_{-1/2}\left(T + \frac{1}{2}I\right)$$

$$= \frac{1}{12}\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}[1 \quad -2 \quad 1].$$

To obtain the diffracted intensity, we calculate:

$$\langle\boldsymbol{\pi}|\overline{\Omega}T_{-1/2}\Omega|\mathbf{1}\rangle = 8|\psi|^2/9$$

and

$$\langle\boldsymbol{\pi}|\overline{\Omega}T_{-1/2,1}\Omega|\mathbf{1}\rangle = \frac{|\psi|^2}{3}e^{i2\pi/3}.$$

We then leverage the fact that $T T_{-1/2,1} = -\frac{1}{2}T_{-1/2,1}$ and $T T_{-1/2} = -\frac{1}{2}T_{-1/2} + T_{-1/2,1}$ to calculate:

$$\langle\boldsymbol{\pi}|\overline{\Omega}T T_{-1/2,1}\Omega|\mathbf{1}\rangle = -\frac{|\psi|^2}{6}e^{i2\pi/3}$$

and

$$\langle\boldsymbol{\pi}|\overline{\Omega}T T_{-1/2}\Omega|\mathbf{1}\rangle = \left(-\frac{4}{9} + \frac{1}{3}e^{i2\pi/3}\right)|\psi|^2.$$

Finally, this yields the nondiagonalizable power spectrum at $p = 3/4$:

$$\frac{P_c(\omega)}{|\psi|^2} = \frac{8}{9} - \frac{1}{3}\text{Re}\frac{e^{i2\pi/3}}{\left(e^{i\omega} + \frac{1}{2}\right)^2} + \frac{2}{3}\text{Re}\frac{e^{i2\pi/3} - \frac{4}{3}}{e^{i\omega} + \frac{1}{2}}.$$

2. Example Two

Here we analyze a generalization of the second chaotic crystal discussed in the main body. For any q, the transition matrix and average-observation matrix are as follows:

$$T = \begin{bmatrix} 0 & 1-q & q \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \Omega = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix},$$

respectively. The transition matrix eigenvalues are $\Lambda_T = \{0, \pm 1\}$, independent of q.

Each spectral projection operator is given by $T_\lambda = |\lambda\rangle\langle\lambda|$, with:

$$\langle\lambda| = \frac{1}{3\lambda - 1}[\lambda \quad \lambda^2 - q \quad q]$$

and

$$|\lambda\rangle = [\lambda \quad 1 \quad (\lambda^2 + q - 1)/q]^\top.$$

Recall that the stationary distribution is the left eigenvector $\langle\boldsymbol{\pi}| = \langle 1| = \frac{1}{2}[1 \quad 1-q \quad q]$. From these elements, we calculate $\langle\boldsymbol{\pi}|\overline{\Omega}\,T_\lambda\Omega|\mathbf{1}\rangle$ and the power spectrum analytically as a function of the transition parameter q. In particular:

$$\langle\boldsymbol{\pi}|\overline{\Omega}\,T_1\Omega|\mathbf{1}\rangle = \frac{1}{4}(3q^2 - 3q + 1)|\psi|^2$$
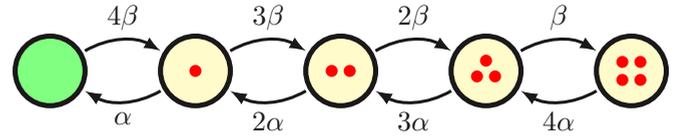
FIG. 13. Voltage-dependent continuous-time Markov chain specifying the transition rates between the conformational states of the K^+ channel. α and β are voltage-dependent transition rates. Only the empty (green) state conducts current. The other states have between one and four activation gates (indicated by the number of red dots) blocking the channel. This model is thus an input-dependent continuous-time HMM for potassium ion current through the channel.

and

$$\langle\boldsymbol{\pi}|\overline{\Omega}\,T_{-1}\Omega|\mathbf{1}\rangle = \frac{3}{4}(q^2 - q + 1)|\psi|^2.$$

The net power spectrum thus consists of a flat "white noise" component:

$$P_c(\omega) = \frac{3}{2}q(1 - q)|\psi|^2$$

in addition to two Bragg reflections per 2π of angular frequency bandwidth

$$P_d(\omega) = \frac{\pi|\psi|^2}{2}\sum_{k=-\infty}^{\infty}[(3q^2 - 3q + 1)\delta(\omega + 2\pi k)$$

$$+ 3(q^2 - q + 1)\delta(\omega - \pi + 2\pi k)].$$

APPENDIX J: POTASSIUM ION CHANNEL—DETAILS

This section lays out the details for the voltage-gated potassium ion channel, as an input-dependent transition rate matrix for partially observable conformational states—i.e., a continuous-time input-dependent HMM. Potassium ion channels are embedded in neural membranes and, together with sodium ion channels, are critical to generating and propagating action potentials that transmit and process information throughout the brain.

What are the dynamics and power spectra of potassium current flowing through the channel? Only one of the five conformational states corresponds to an open channel where current can flow. The other states are distinguished by the number of activation gates closing the channels (from one to four), but observation of the current does not allow for direct observation of these conformational states. Nevertheless, the dynamics among these states influence the statistical properties of the current. In particular, the current is non-Markovian and exhibits a nonexponential distribution of closure durations.

The transition structure between conformational states of the K^+ channel is depicted in Fig. 13. The Hodgkin-Huxley model's voltage-dependent transition rates α and β—often denoted α_n and β_n—describe the probability that an activation gate opens or closes (respectively) at a given voltage:

$$\alpha = \frac{(v + 55)/100 \text{ ms}}{1 - e^{-(v+55)/10}} \quad \text{and} \quad \beta = \frac{1}{8 \text{ ms}}e^{-(v+65)/80},$$

where v is the voltage (in mV) across the membrane [118,121]. The voltage-dependent transition rate matrix can be written explicitly as:

$$G^{(\mathcal{S} \to \mathcal{S}|v)}$$

$$\equiv \begin{bmatrix} -4\beta & 4\beta & 0 & 0 & 0 \\ \alpha & -(\alpha+3\beta) & 3\beta & 0 & 0 \\ 0 & 2\alpha & -(2\alpha+2\beta) & 2\beta & 0 \\ 0 & 0 & 3\alpha & -(3\alpha+\beta) & \beta \\ 0 & 0 & 0 & 4\alpha & -4\alpha \end{bmatrix}.$$

The average current through a single channel is binary—either 0 or I_0. Appreciable current only flows in the open conformation. In the open conformation, $I_0 = g_0(v - V_K)$, where g_0 is the conductance of an open K^+ channel and V_K is the Nernst potential for potassium. Accordingly, the average-observation operator for the ion current is $\Omega = I_0|\text{open}\rangle\langle\text{open}|$, where $\langle\text{open}| = [1 \quad 0 \quad 0 \quad 0 \quad 0]$ denotes the open conformation (i.e., the green state in Fig. 13).

The rate matrix eigenvalues are $\Lambda_G = \{-n(\alpha+\beta)\}_{n=0}^4$.

Applying Eq. (14), the power spectrum at a fixed voltage is

$$P_c(f) = \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_\lambda - 1} 2 \operatorname{Re} \frac{\langle \pi | \overline{\Omega} \, G_{\lambda,m} \Omega | \mathbf{1} \rangle}{(i2\pi f - \lambda)^{m+1}}$$

$$= \sum_{n=0}^4 2 \operatorname{Re} \frac{\langle \pi | \overline{\Omega} \, G_{-n(\alpha+\beta)} \Omega | \mathbf{1} \rangle}{i2\pi f + n(\alpha+\beta)}$$

$$= 2I_0^2 \langle \pi | \text{open} \rangle \sum_{n=0}^4 \operatorname{Re} \frac{\langle \text{open} | G_{-n(\alpha+\beta)} | \text{open} \rangle}{i2\pi f + n(\alpha+\beta)}$$

$$= 2I_0^2 \langle \pi | \text{open} \rangle \sum_{n=1}^4 \frac{\langle \text{open} | G_{-n(\alpha+\beta)} | \text{open} \rangle}{n(\alpha+\beta)\left\{1 + \left[\frac{2\pi f}{n(\alpha+\beta)}\right]^2\right\}}.$$

For convenience, define the opening bias $\psi \equiv \alpha/\beta$ as the ratio between an individual gate's rates of opening versus closing. The spectral projection operators $\{G_{-n(\alpha+\beta)}\}_n$ are simple analytic functions of ψ. We then find the open state's overlap with the spectral projection operators:

$$\langle \text{open} | G_{-n(\alpha+\beta)} | \text{open} \rangle = \binom{4}{n} \frac{\psi^{4-n}}{(1+\psi)^4}.$$

By setting $n = 0$, this expression also yields the stationary probability of the open state: $\langle \pi | \text{open} \rangle = \left(\frac{\psi}{1+\psi}\right)^4$.

The power spectrum for potassium current is thus:

$$P_c(f) = \frac{I_0^2}{\pi}\left(\frac{\psi}{1+\psi}\right)^8 \sum_{n=1}^4 \frac{\binom{4}{n}\psi^{-n}}{nw\left[1+\left(\frac{f}{nw}\right)^2\right]}, \quad (\text{J1})$$

where $w \equiv (\alpha+\beta)/2\pi$. Each nonzero eigenmode contributes a Lorentzian profile to the power spectrum, each with a different crossover frequency $f_c = nw$ depending on n. This spread of crossover frequencies smooths the transition between the flat power spectrum at low frequencies and the $1/f^2$ spectrum at high frequencies. At $v = -40$ mV, the crossover frequency of the net spectrum is $f_c^{\text{net}} \approx 3w$.

In fact, a power spectral signature of this general form has been experimentally observed above the $1/f$ background

noise [123]. That said, the empirically observed crossover frequency suggests that the model is not a complete description of the ion channel dynamics.

Equation (J1), derived from the rate matrix's spectral properties, agrees with the much earlier results calculated via alternative methods in Refs. [124,125]. For ease of comparison with those references, note that the Hodgkin-Huxley parameter n_∞ is related to ψ via $n_\infty = \psi/(1+\psi)$.

In voltage-clamped experiments, a common neurophysiological measurement technique, the voltage is held fixed. Then, the finite-duration transition matrix is simply $T = e^{\tau_0 G}$, where τ_0 is the duration between measurements. Since K^+ current (rather than conformational states) is measured, this gives the transition matrix of a HMM for the observed current. The finite sampling rate associated with such a discrete-time HMM allows exactly predicting the expected *empirical* spectrum. At high frequencies, this deviates from the continuous-time spectrum as the latter implicitly assumes an infinite sampling rate.

APPENDIX K: CROSS-CORRELATION AND SPECTRAL DENSITIES

Cross-correlation and cross-spectral densities are often important in applications [152,153]. These may be especially useful when analyzing input-output processes, to characterize the correlation of input and output, or to characterize the correlation between different aspects of the output. Our results can be easily extended to address these quantities.

Using an HMM that describes the *joint* stochastic process of two observables $(x, y) \in \mathcal{A}$, it is straightforward to generalize our developments to *cross-correlation* $\gamma_{XY}(\tau)$:

$$\gamma_{XY}(\tau) = \langle \overline{X_t} Y_{t+\tau} \rangle$$

(rather than necessarily *auto*correlation $\gamma = \gamma_{XX}$) and the associated *cross-spectral densities* $P_{XY}(\omega)$:

$$P_{XY}(\omega) = \lim_{N \to \infty} \frac{1}{N}\left\langle \left(\sum_{t=1}^N \overline{X_t} e^{i\omega t}\right)\left(\sum_{t=1}^N Y_t e^{-i\omega t}\right)\right\rangle$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{\tau=-N}^N (N - |\tau|)\gamma_{XY}(\tau) e^{-i\omega\tau}$$

of distinct observables $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The individual stochastic processes for each observable by itself can simply be obtained by marginalizing over the other observable.

Explicitly, the expressions take the form:

$$\gamma_{XY}(\tau) = \begin{cases} \langle \pi | \Omega_Y \, T^{|\tau|} \, \overline{\Omega}_X | \mathbf{1} \rangle & \text{if } \tau \leqslant 1 \\ \langle \overline{X_t} Y_t \rangle & \text{if } \tau = 0, \\ \langle \pi | \overline{\Omega}_X \, T^{|\tau|} \, \Omega_Y | \mathbf{1} \rangle & \text{if } \tau \geqslant 1 \end{cases}$$

where:

$$\Omega_Y = \sum_{s \in \mathcal{S}} \langle Y \rangle_{\mathrm{p}(X,Y|s)} |s\rangle\langle s|,$$

and

$$\langle \overline{X_t} Y_t \rangle = \sum_{s \in \mathcal{S}} \langle \pi | s \rangle \langle \overline{X} Y \rangle_{\mathrm{p}(X,Y|s)}.$$

Moreover, the continuous part of the cross-spectral density is given by:

$$P_{XYc}(\omega) = \langle \overline{X_t} Y_t \rangle + \langle \boldsymbol{\pi} | \overline{\Omega}_X \, T \, (e^{i\omega} I - T)^{-1} \Omega_Y | \mathbf{1} \rangle$$
$$+ \langle \boldsymbol{\pi} | \Omega_Y \, T \, (e^{-i\omega} I - T)^{-1} \overline{\Omega}_X | \mathbf{1} \rangle,$$

and so on.

APPENDIX L: PAIRWISE MUTUAL INFORMATION EXAMPLE

For the process generated by the HMM given in Fig. 6, if we take the limit of ever-narrower Gaussians in the state-conditioned PDFs, so that we work with pairs of δ functions, then the process becomes Markovian and the pairwise mutual information can be calculated exactly:

$$I(X_0; X_\tau) = H(X_0) - H(X_\tau | X_0)$$
$$= H(X_0, \mathcal{S}_0) - H(X_\tau, \mathcal{S}_\tau | X_0, \mathcal{S}_0)$$
$$= H(\mathcal{S}_0) + H(X_0 | \mathcal{S}_0) - H(X_\tau, \mathcal{S}_\tau | \mathcal{S}_0)$$
$$= H(\mathcal{S}_0) + H(X_0 | \mathcal{S}_0) - H(\mathcal{S}_\tau | \mathcal{S}_0) - H(X_\tau | \mathcal{S}_\tau)$$
$$= H(\mathcal{S}_0) - H(\mathcal{S}_\tau | \mathcal{S}_0)$$
$$= H(\boldsymbol{\pi}) - \sum_{s \in \mathcal{S}} \boldsymbol{\pi}(s) H(\mathcal{S}_\tau | \mathcal{S}_0 = s)$$
$$= H(\boldsymbol{\pi}) - \sum_{s \in \mathcal{S}} \boldsymbol{\pi}(s) H(\langle s | T^\tau)$$
$$= H(\boldsymbol{\pi}) + \sum_{s,s' \in \mathcal{S}} \boldsymbol{\pi}(s) \langle s | T^\tau | s' \rangle \log \langle s | T^\tau | s' \rangle, \quad \text{(L1)}$$

where $\boldsymbol{\pi} = [1, 1-p, 1-p, 1-p]/(4-3p)$.

Continuing, $\langle s | T^\tau | s' \rangle$ can be calculated via T's spectral decomposition. Since T is diagonalizable and nondegenerate for all values of the transition parameter p, we find:

$$\langle s | T^\tau | s' \rangle = \sum_{\lambda \in \Lambda_T} \lambda^\tau \langle s | T_\lambda | s' \rangle.$$

Moreover:

$$\langle s | T_1 | s' \rangle = \langle s | \mathbf{1} \rangle \langle \boldsymbol{\pi} | s' \rangle$$
$$= \boldsymbol{\pi}(s'),$$

so $\langle s | T^\tau | s' \rangle$ simplifies somewhat to:

$$\langle s | T^\tau | s' \rangle = \boldsymbol{\pi}(s') + \sum_{\lambda \in \Lambda_T \setminus \{1\}} \lambda^\tau \langle s | T_\lambda | s' \rangle.$$

In fact, Eq. (L1) is valid for any set of four PDFs we could have chosen for the example HMM's states, as long as the PDFs all have mutually exclusive support for the observable output, since this then makes the hidden state a function of the instantaneous observable.

Using the linear algebra of Eq. (L1), we calculate the pairwise mutual information and POPI spectrum numerically. The pairwise mutual informations are shown for $p \in \{0.1, 0.5, 0.9\}$ in Fig. 14. Reasonably, the loss of information is monotonic over temporal distance. More surprisingly, the decay of pairwise mutual information is very-nearly exponential as made clear in the inset logarithmic plot.

The POPI spectrum, which can be rewritten for a wide-sense stationary process as:

$$\mathcal{I}(\omega) = \lim_{N \to \infty} 2 \sum_{\tau=1}^{N} \cos(\omega \tau) I(X_0; X_\tau),$$
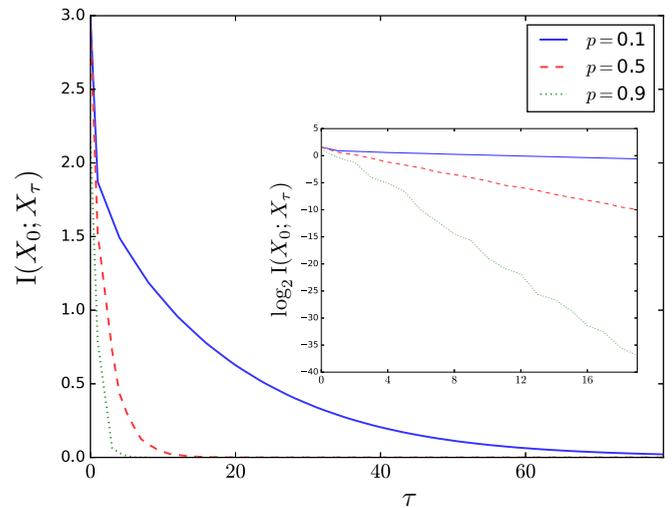
FIG. 14. Nontrivial pairwise mutual information for the process from Fig. 6 with a flat power spectrum.

is shown for these same p values in Fig. 15. The POPI spectrum was approximated by truncating the summation of modulated pairwise mutual informations at a sufficiently large separation of $\tau = 2000$.

APPENDIX M: MEASUREMENT FEEDBACK MODELS

Let us now turn to describe an alternative set of possibly input-dependent models, which may be more convenient for describing certain phenomena. For example, they are more natural for describing measured quantum systems. They also reduce to the canonical models used in *computational mechanics* [86,88] after a number of simplifying assumptions.

After introducing them, we show that Theorem 2 applies to them as well as to the other model types discussed in the main body. In this way, we extend the theory of fraudulent white noise to these models as well.

The models we consider generate observable behavior *during transitions* between states, rather than in the states
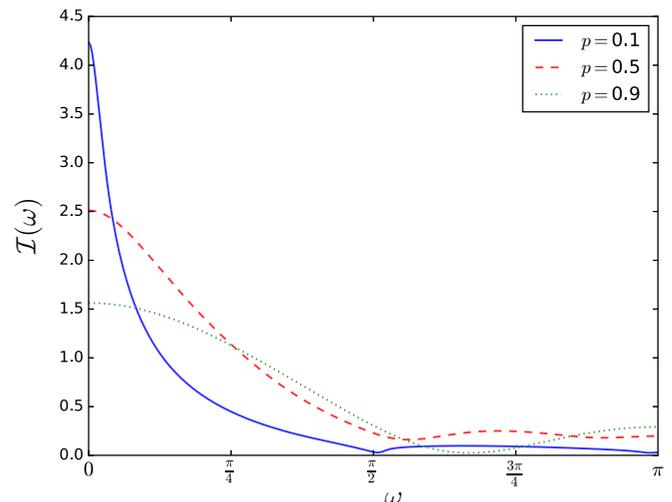
FIG. 15. POPI spectrum for the process from Fig. 6.

themselves. This is a natural approach in the quantum setting since measurement feedback changes the state of the quantum system with dependence on the measurement outcome. For projective measurements, measurement fully defines the new state, but for the much more general class of quantum measurements described by *positive operator valued measures*, the measurement outcome plays a more nuanced role in updating the state. More generally, edge-emitting models can be natural descriptors of complex systems with control and feedback. And, fittingly, edge-emitting models have been used elsewhere as well. For instance, they appear extensively in computer science and computational mechanics—the latter of which spans the study of natural computation in physical systems and the minimal resources required for prediction.

1. Measurement feedback models

Here we introduce measurement feedback models (MFMs) $\mathcal{M}_{\mathrm{MFM}}(\vec{m})$, which are input-dependent generators of an observable output process $\{X_t\}_{t\in\mathcal{T}}$. As before, the lengths and alphabets of the inputs and outputs need not be commensurate. The output is generated via $\mathcal{M}_{\mathrm{MFM}}(\vec{m}) = (\mathcal{S}, \mathcal{A}, \{T_t^{(x)}(\vec{m})\}_{t\in\mathcal{T}, x\in\mathcal{A}}, \boldsymbol{\mu}_1)$, where $\mathcal{S}$ is the countable set of hidden states, $\mathcal{A}$ is the alphabet of observables, and $\boldsymbol{\mu}_1$ is the initial distribution over hidden states. For a given t and x, the matrix elements $\langle s|T_t^{(x)}(\vec{m})|s'\rangle$ provide the probability density of transitioning from state s to s' while emitting the observable x; that is,

$$\langle s|T_t^{(x)}(\vec{m})|s'\rangle = p_{\vec{m}}(X_{t+1} = x, \mathcal{S}_{t+1} = s'|\mathcal{S}_t = s),$$

where $p_{\vec{m}}$ is the probability density (*induced* by $\vec{m}$) of the labeled transition. The symbol-labeled transition matrices $\{T_t^{(x)}(\vec{m})\}_{t\in\mathcal{T}, x\in\mathcal{A}}$ yield the net state-to-state transition probabilities when marginalizing over all possible observations:

$$\int_{x\in\mathcal{A}} T_t^{(x)}(\vec{m}) \, dx = T_t(\vec{m}),$$

where $\langle s|T_t(\vec{m})|s'\rangle = \mathrm{Pr}_{\vec{m}}(\mathcal{S}_{t+1} = s'|\mathcal{S}_t = s)$.

Figure 16 displays two different (but equally valid) Bayesian networks for the decomposition of conditional dependencies among observables and latent states of a MFM. Each decomposition suggests a preferred interpretation. The decomposition of Fig. 16(a) allows identifying a PDF with each directed edge between latent states of a measurement feedback model $\mathcal{M}_{\mathrm{MFM}}(\vec{m})$. Accordingly, Fig. 16(a) suggests that the transited edge determines the probability of the observable, whereas the decomposition of Fig. 16(b) suggests that the observation determines the probability of the latent state transition. The fact that both decompositions are valid insists, perhaps surprisingly, that the interpretations have no physical distinction. The interpretation of causality is ambiguous although each calculus of conditional dependencies is reliable.

The measurement feedback models may initially appear rather restrictive when considering the possibilities of, say, measuring a quantum system in different bases and with different instruments. However, in principle, the different measurement choices are incorporated through the different transformations $T_t(\vec{m})$, both through any predetermined mea-
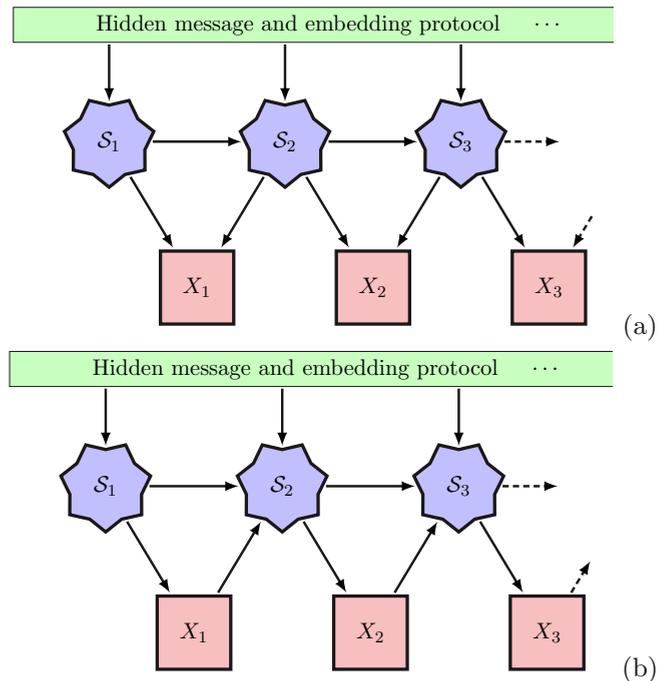
FIG. 16. Alternative Bayesian networks for measurement feedback models.

surement choices in $\vec{m}$ and through dynamic determination via feedback of the measurement outcomes themselves.

Reference [154]'s *process tensors* can also be used to model classical observable processes generated by general quantum dynamics. Although unnecessarily elaborate for most purposes, process tensors are appealing since they rigorously incorporate general quantum measurements. Ultimately though, they, together with a set of "experiments" $\vec{m}$, could be mapped onto the alternative rather-simpler models proposed here if the goal is only to model the observable classical output process.

2. Theorem 2 for measurement feedback

The MFM's average-observation matrices are as follows:

$$\Omega_t = \int_{x\in\mathcal{A}} x \, T_t^{(x)}(\vec{m}) \, dx.$$

Notably, they are no longer diagonal in the hidden-state basis. Rather, they assign to each matrix element the average observation associated with that transition, multiplied by the probability of the edge being traversed when conditioned on occupying the outgoing state. That is,

$$
\begin{aligned}
\langle s|\Omega_t|s'\rangle &= \int_{x\in\mathcal{A}} x \, p_{\vec{m}}(\mathcal{S}_{t+1} = s', X_t = x|\mathcal{S}_t = s) \, dx \\
&= \mathrm{Pr}_{\vec{m}}(\mathcal{S}_{t+1} = s'|\mathcal{S}_t = s) \\
&\quad \times \int_{x\in\mathcal{A}} x \, p_{\vec{m}}(X_t = x|\mathcal{S}_t = s, \mathcal{S}_{t+1} = s') \, dx \\
&= \langle s|T_t(\vec{m})|s'\rangle \langle x\rangle_{p_{\vec{m}}(X_t|\mathcal{S}_t=s, \mathcal{S}_{t+1}=s')}.
\end{aligned}
\tag{M1}
$$

If the process is wide-sense stationary, then for $\tau > 0$:

$$\gamma(\tau) = \langle \boldsymbol{\mu}_t | \overline{\Omega}_t \, T_{t+1:t+\tau}(\vec{m}) \, \Omega_{t+\tau} | \mathbf{1} \rangle, \qquad \text{(M2)}$$

which must be t independent.

For input-independent processes with time-independent transition dynamics—where $T_t^{(x)}(\vec{m}) = T^{(x)}$ and $\boldsymbol{\mu}_1 = \boldsymbol{\pi}$—this simplifies to the autonomous Mealy-type HMMs with continuous PDFs for the observable associated with each hidden-state-to-state transition. The autocorrelation function (for $\tau \geqslant 1$) then reduces to:

$$\gamma(\tau) = \langle \boldsymbol{\pi} | \overline{\Omega} \, T^{\tau-1} \, \Omega | \mathbf{1} \rangle,$$

while the power spectrum's continuous part is

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \, \mathrm{Re} \langle \boldsymbol{\pi} | \overline{\Omega} \, (e^{i\omega} I - T)^{-1} \Omega | \mathbf{1} \rangle. \qquad \text{(M3)}$$

Note that this expression lacks T, the transition dynamic, when compared to Eq. (6). This follows since Ω induces a transition for these Mealy-type HMMs, reducing the number of subsequent transitions by 1.

Let us return to the general setting for autocorrelation given by Eq. (M2) for processes generated by possibly input-dependent models. Developing the analog of Theorem 2 requires recognizing that the average observation *on each edge* matters, rather than previously, where the average observation from each state mattered. For MFMs, constant autocorrelation and flat power spectrum can again be guaranteed by a rather weak condition: The average output of the current edge does not by itself influence the average output of a future edge.

More explicitly, consider the set of all edges:

$$\boldsymbol{\mathcal{E}}^{(t)} \equiv \{(s, s') \in \boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{S}} : \langle s | T_t(\vec{m}) | s' \rangle \neq 0 \},$$

which are transitions between hidden states that can be traversed at time t with positive probability. Since outputs occur during edge transitions, we redefine Ξ as the set of average outputs exhibited by the edges. Equation (M1) indicates that the desired definition is as follows:

$$\Xi \equiv \bigcup_{t \in \mathcal{T}} \bigcup_{(s,s') \in \boldsymbol{\mathcal{E}}^{(t)}} \left\{ \frac{\langle s | \Omega_t | s' \rangle}{\langle s | T_t(\vec{m}) | s' \rangle} \right\}.$$

Furthermore, we define $\mathcal{E}_t$ to be the random variable for the edge traversed at time t; i.e., $\mathcal{E}_t$ is the joint random variable: $\mathcal{E}_t = (\mathcal{S}_t, \mathcal{S}_{t+1})$. And we define $\boldsymbol{\mathcal{E}}_\xi^{(t)} \subset \boldsymbol{\mathcal{E}}^{(t)}$ as the set of edges (at time t) with average output $\xi \in \Xi$:

$$\boldsymbol{\mathcal{E}}_\xi^{(t)} \equiv \left\{ (s, s') \in \boldsymbol{\mathcal{E}}^{(t)} : \frac{\langle s | \Omega_t | s' \rangle}{\langle s | T_t(\vec{m}) | s' \rangle} = \xi \right\}. \qquad \text{(M4)}$$

With these in hand, we can state the theorem analogous to Theorem 2.

Theorem 4. Let $\{X_t\}_t$ be a stochastic process generated by any measurement feedback model $\mathcal{M}_{\mathrm{MFM}}(\vec{m})$, including autonomous Mealy-type HMMs and input-dependent generators. Such processes have *constant autocorrelation and a flat power spectrum* if:

$$\Pr\left[\mathcal{E}_{t+\tau} \in \boldsymbol{\mathcal{E}}_{\xi'}^{(t+\tau)} \big| \mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)} \right] = \Pr\left[\mathcal{E}_{t+\tau} \in \boldsymbol{\mathcal{E}}_{\xi'}^{(t+\tau)} \right]$$

tand there exists a constant $c \in \mathbb{C}$ such that:

$$\sum_{\xi \in \Xi} \xi \, \Pr\left[\mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)} \right] = c,$$

for all separations $\tau > 0$, $t \in \mathcal{T}$, and $\xi, \xi' \in \Xi$.

Proof. Starting from Eq. (M2), we find the autocorrelation for all such processes by calculating:

$$
\begin{aligned}
\gamma(\tau) &= \langle \boldsymbol{\mu}_t | \overline{\Omega}_t \, T_{t+1:t+\tau}(\vec{m}) \, \Omega_{t+\tau} | \mathbf{1} \rangle \\
&= \sum_{s,s',s'',s''' \in \boldsymbol{\mathcal{S}}} \langle \boldsymbol{\mu}_t | s \rangle \langle s | \overline{\Omega}_t | s' \rangle \langle s' | T_{t+1:t+\tau}(\vec{m}) | s'' \rangle \langle s'' | \Omega_{t+\tau} | s''' \rangle \langle s''' | \mathbf{1} \rangle \\
&= \sum_{\substack{(s,s') \in \boldsymbol{\mathcal{E}}^{(t)} \\ (s',s''') \in \boldsymbol{\mathcal{E}}^{(t+\tau)}}} \left[\frac{\langle s | \overline{\Omega}_t | s' \rangle}{\langle s | T_t(\vec{m}) | s' \rangle} \right] \left[\frac{\langle s'' | \Omega_{t+\tau} | s''' \rangle}{\langle s'' | T_{t+\tau}(\vec{m}) | s''' \rangle} \right] \langle \boldsymbol{\mu}_t | s \rangle \langle s | T_t(\vec{m}) | s' \rangle \langle s' | T_{t+1:t+\tau}(\vec{m}) | s'' \rangle \langle s'' | T_{t+\tau}(\vec{m}) | s''' \rangle \langle s''' | \mathbf{1} \rangle \\
&= \sum_{\substack{(s,s') \in \boldsymbol{\mathcal{E}}^{(t)} \\ (s',s''') \in \boldsymbol{\mathcal{E}}^{(t+\tau)}}} \left[\frac{\langle s | \overline{\Omega}_t | s' \rangle}{\langle s | T_t(\vec{m}) | s' \rangle} \right] \left[\frac{\langle s'' | \Omega_{t+\tau} | s''' \rangle}{\langle s'' | T_{t+\tau}(\vec{m}) | s''' \rangle} \right] \Pr(\mathcal{S}_t = s, \mathcal{S}_{t+1} = s', \mathcal{S}_{t+\tau} = s'', \mathcal{S}_{t+\tau+1} = s''') \\
&= \sum_{\substack{(s,s') \in \boldsymbol{\mathcal{E}}^{(t)} \\ (s',s''') \in \boldsymbol{\mathcal{E}}^{(t+\tau)}}} \left[\frac{\langle s | \overline{\Omega}_t | s' \rangle}{\langle s | T_t(\vec{m}) | s' \rangle} \right] \left[\frac{\langle s'' | \Omega_{t+\tau} | s''' \rangle}{\langle s'' | T_{t+\tau}(\vec{m}) | s''' \rangle} \right] \Pr(\mathcal{E}_t = (s, s'), \mathcal{E}_{t+\tau} = (s'', s''')) \\
&= \sum_{\xi, \xi' \in \Xi} \overline{\xi} \xi' \, \Pr\left(\mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)}, \mathcal{E}_{t+\tau} \in \boldsymbol{\mathcal{E}}_{\xi'}^{(t+\tau)} \right) \\
&= \sum_{\xi \in \Xi} \overline{\xi} \, \Pr\left(\mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)} \right) \left[\sum_{\xi' \in \Xi} \xi' \, \Pr\left(\mathcal{E}_{t+\tau} \in \boldsymbol{\mathcal{E}}_{\xi'}^{(t+\tau)} \big| \mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)} \right) \right].
\end{aligned}
$$

Now suppose that:

$$\Pr\left[\mathcal{E}_{t+\tau} \in \boldsymbol{\mathcal{E}}_{\xi'}^{(t+\tau)} \big| \mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)} \right] = \Pr\left[\mathcal{E}_{t+\tau} \in \boldsymbol{\mathcal{E}}_{\xi'}^{(t+\tau)} \right]$$

and there exists some constant $c \in \mathbb{C}$ such that:

$$\sum_{\xi \in \Xi} \xi \, \Pr\left[\mathcal{E}_t \in \boldsymbol{\mathcal{E}}_\xi^{(t)} \right] = c,$$

for all separations $\tau > 0$, $t \in \mathcal{T}$, and $\xi, \xi' \in \Xi$. Then, we find:

$$\gamma(\tau) = \sum_{\xi \in \Xi} \overline{\xi}\, \Pr\left[\mathcal{E}_t \in \mathcal{E}_\xi^{(t)}\right] \left\{ \sum_{\xi' \in \Xi} \xi'\, \Pr\left[\mathcal{E}_{t+\tau} \in \mathcal{E}_{\xi'}^{(t+\tau)}\middle|\mathcal{E}_t \in \mathcal{E}_\xi^{(t)}\right] \right\} = \left\{ \sum_{\xi \in \Xi} \overline{\xi}\, \Pr\left[\mathcal{E}_t \in \mathcal{E}_\xi^{(t)}\right] \right\} \left\{ \sum_{\xi' \in \Xi} \xi'\, \Pr\left[\mathcal{E}_{t+\tau} \in \mathcal{E}_{\xi'}^{(t+\tau)}\right] \right\} = |c|^2,$$

which is a constant for all separations $\tau > 0$, $t \in \mathcal{T}$, and $\xi, \xi' \in \Xi$. Finally, a process with stationary low-order statistics and a flat autocorrelation has a flat power spectrum, as an immediate consequence of Eq. (3). This proves Theorem 4.

For the special case of an autonomous HMM that generates observations during hidden-state-to-state transitions, this condition simplifies significantly. Specifically, $\Omega_t \to \Omega$ and $T_t(\vec{m}) \to T$ become t independent, which furthermore means that $\mathcal{E}_\xi^{(t)} \to \mathcal{E}_\xi$ becomes t independent. For autonomous wide-sense stationary processes, we have $\Pr(\mathcal{E}_t) = \Pr(\mathcal{E}_{t+\tau})$ for all separations $\tau > 0$ and for all $t \in \mathcal{T}$. It then trivially follows that $\sum_{\xi \in \Xi} \xi\, \Pr(\mathcal{E}_t \in \mathcal{E}_\xi)$ is constant for all $t \in \mathcal{T}$. So the only requirement for an autonomous edge-emitting HMM to produce fraudulent white noise is that it satisfies the condition:

$$\Pr(\mathcal{E}_{t+\tau} \in \mathcal{E}_{\xi'}|\mathcal{E}_t \in \mathcal{E}_\xi) = \Pr(\mathcal{E}_{t+\tau} \in \mathcal{E}_{\xi'})$$

for all separations $\tau > 0$, $t \in \mathcal{T}$, and $\xi, \xi' \in \Xi$.

Theorem 4 provides a very general condition for flat power spectra from measurement feedback models.

APPENDIX N: THEOREM 2 FOR TIME-DEPENDENT PDFs

Moreover, Theorem 4 suggests how Theorem 2 generalizes even further to possibly-input-dependent hidden-state models with time-dependent PDFs associated with each state. We will call these *morphing hidden models* (MHMs) $\mathcal{M}_{\mathrm{MHM}}(\vec{m})$. MHMs include, as special cases, all models (Moore-type HMMs and input-dependent generators) considered in the main text. We employ methods similar to those used in Appendix M 2.

A MHM is a possibly-input-dependent generator of an observable output process $\{X_t\}_{t \in \mathcal{T}}$. The output is generated via $\mathcal{M}_{\mathrm{MHM}}(\vec{m}) = (\mathcal{S}, \mathcal{A}, \{\mathcal{P}_t(\vec{m})\}_t, \{T_t(\vec{m})\}_t, \boldsymbol{\mu}_1)$. Here, again, the lengths and alphabets of the inputs and outputs need not be commensurate. That is, the internal states $\mathcal{S}$ and output alphabet $\mathcal{A}$ are static. However, the hidden-state-to-state transition matrix $T_t(\vec{m})$—as well as the state-dependent

PDFs $\mathcal{P}_t(\vec{m})$—are time-dependent such that their values at time t are potentially a function of the full input vector $\vec{m}$. More specifically, $\mathcal{P}_t(\vec{m})$ is the set of hidden-state-dependent probability density functions $p_{\vec{m}}(X_t|s)$ at time t. As before, $\boldsymbol{\mu}_1$ specifies the initial distribution over hidden states: $\mathcal{S}_1 \sim \boldsymbol{\mu}_1$.

For such cases, set:

$$\Omega_t = \sum_{s \in \mathcal{S}} \langle x \rangle_{p_{\vec{m}}(X_t|s)} |s\rangle\langle s|.$$

The Ω_t matrix is time dependent with the state-conditioned expected outputs along its diagonal.

Since the average state output now varies in time, we must generalize Ξ from its more restricted use in the main text. Specifically, redefine Ξ as the set of state-dependent average outputs generated throughout time:

$$\Xi \equiv \bigcup_{t \in \mathcal{T}} \bigcup_{s \in \mathcal{S}} \left\{ \langle x \rangle_{p_{\vec{m}}(X_t|s)} \right\}.$$

Furthermore, we define $\mathcal{S}_\xi^{(t)} \subset \mathcal{S}$ as the set of states (at time t) with average output $\xi \in \Xi$:

$$\mathcal{S}_\xi^{(t)} \equiv \left\{ s \in \mathcal{S} : \langle X_t \rangle_{p_{\vec{m}}(X_t|s)} = \xi \right\}.$$

Using these, we can state the following theorem, which generalizes Theorem 2.

Theorem 5. Let $\{X_t\}_t$ be a stochastic process generated by any morphing hidden model $\mathcal{M}_{\mathrm{MHM}}(\vec{m})$. Such processes have *constant autocorrelation and a flat power spectrum* if:

$$\Pr\left[\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}^{(t+\tau)}\middle|\mathcal{S}_t \in \mathcal{S}_\xi^{(t)}\right] = \Pr\left[\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}^{(t+\tau)}\right]$$

and there exists a constant $c \in \mathbb{C}$ such that:

$$\sum_{\xi \in \Xi} \xi\, \Pr\left[\mathcal{S}_t \in \mathcal{S}_\xi^{(t)}\right] = c,$$

for all separations $\tau > 0$, $t \in \mathcal{T}$, and $\xi, \xi' \in \Xi$.

Proof. For the processes under consideration, we find the linear pairwise correlation (for $\tau \geqslant 1$) to be

$$\begin{aligned}
\langle \overline{X_t} X_{t+\tau} \rangle_{p_{\vec{m}}(X_t, X_{t+\tau})} &= \langle \boldsymbol{\mu}_t | \overline{\Omega}_t\, T_{t:t+\tau}(\vec{m})\, \Omega_{t+\tau} |\mathbf{1}\rangle \\
&= \sum_{s,s' \in \mathcal{S}} \langle \boldsymbol{\mu}_t|s\rangle \langle s|\overline{\Omega}_t|s\rangle \langle s|T_{t:t+\tau}(\vec{m})|s'\rangle \langle s'|\Omega_{t+\tau}|s'\rangle \langle s'|\mathbf{1}\rangle \\
&= \sum_{\xi,\xi' \in \Xi} \overline{\xi}\xi' \sum_{\substack{s \in \mathcal{S}_\xi^{(t)} \\ s' \in \mathcal{S}_{\xi'}^{(t+\tau)}}} \Pr_{\vec{m}}(\mathcal{S}_t = s, \mathcal{S}_{t+\tau} = s') \\
&= \sum_{\xi,\xi' \in \Xi} \overline{\xi}\xi'\, \Pr\left(\mathcal{S}_t \in \mathcal{S}_\xi^{(t)}, \mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}^{(t+\tau)}\right) \\
&= \sum_{\xi \in \Xi} \overline{\xi}\, \Pr\left[\mathcal{S}_t \in \mathcal{S}_\xi^{(t)}\right] \left\{ \sum_{\xi' \in \Xi} \xi'\, \Pr\left[\mathcal{S}_{t+\tau} \in \mathcal{S}_{\xi'}^{(t+\tau)}\middle|\mathcal{S}_t \in \mathcal{S}_\xi^{(t)}\right] \right\}. \quad (\mathrm{N}1)
\end{aligned}$$

Now suppose that:

$$\Pr\left[\mathcal{S}_{t+\tau} \in \boldsymbol{\mathcal{S}}_{\xi'}^{(t+\tau)}\big|\mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi}^{(t)}\right] = \Pr\left[\mathcal{S}_{t+\tau} \in \boldsymbol{\mathcal{S}}_{\xi'}^{(t+\tau)}\right]$$

and there exists some constant $c \in \mathbb{C}$ such that:

$$\sum_{\xi \in \Xi} \xi \, \Pr\left[\mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi}^{(t)}\right] = c$$

for all separations $\tau > 0$, for all $t \in \mathcal{T}$, and for all $\xi, \xi' \in \Xi$. Then, we find:

$$\langle \overline{X_t} X_{t+\tau} \rangle_{p_{\bar{m}}(X_t, X_{t+\tau})} = \sum_{\xi \in \Xi} \overline{\xi} \Pr\left[\mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi}^{(t)}\right] \left\{ \sum_{\xi' \in \Xi} \xi' \Pr\left[\mathcal{S}_{t+\tau} \in \boldsymbol{\mathcal{S}}_{\xi'}^{(t+\tau)}\big|\mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi}^{(t)}\right] \right\}$$

$$= \left\{ \sum_{\xi \in \Xi} \overline{\xi} \Pr\left[\mathcal{S}_t \in \boldsymbol{\mathcal{S}}_{\xi}^{(t)}\right] \right\} \left\{ \sum_{\xi' \in \Xi} \xi' \Pr\left[\mathcal{S}_{t+\tau} \in \boldsymbol{\mathcal{S}}_{\xi'}^{(t+\tau)}\right] \right\} = |c|^2.$$

is constant for all $t \in \mathcal{T}$, and $\forall \xi, \xi' \in \Xi$.

That $\langle \overline{X_t} X_{t+\tau} \rangle_{p_{\bar{m}}(X_t, X_{t+\tau})}$ is constant verifies that the autocorrelation does not depend on the overall time shift of the process, so $\langle \overline{X_t} X_{t+\tau} \rangle_{p_{\bar{m}}(X_t, X_{t+\tau})} = \gamma(\tau)$. Moreover, $\gamma(\tau)$ is constant. Finally, a process with constant autocorrelation has a flat power spectrum, as an immediate consequence of Eq. (3). This proves Theorem 5.

APPENDIX O: ANALYTICAL POLYSPECTRA

This section derives new analytical expressions for polyspectra, revealing their close relationship with the time-evolution operator's eigenspectrum and resolvent.

The $(g_0, \ldots, g_K)$ polyspectrum is defined as:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K) \equiv \lim_{N \to \infty} \frac{1}{N} \left\langle \prod_{k=0}^{K} \widetilde{g}_k^{(N)}(\omega_k) \right\rangle, \tag{O1}$$

where $\omega_0 \equiv -\sum_{k=1}^{K} \omega_k$ and

$$\widetilde{g}^{(N)}(\omega) \equiv \sum_{t=1}^{N} g(X_t) e^{-i\omega t}. \tag{O2}$$

Each $g_k : \mathcal{A} \to \mathbb{C}$ can be any function taking observables to complex numbers.

Combining Eqs. (O1) and (O2) yields:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K) = \lim_{N \to \infty} \frac{1}{N} \sum_{t_0=1}^{N} \cdots \sum_{t_K=1}^{N} \left\langle \prod_{k=0}^{K} g_k(X_{t_k}) \right\rangle \prod_{k=0}^{K} e^{-i\omega_k t_k}. \tag{O3}$$

The original time variables $(t_k)_{k=0}^{K}$ induce a function $\alpha : \{0, 1, \ldots K\} \to \{0, 1, \ldots \kappa\}$ that compresses and time-orders the indices, such that $t_k = t'_{\alpha(k)}$. Since α does not have a unique inverse, we define the preimage $\alpha^{-1}(\ell) = \{k \in \{0, 1, \ldots K\} : \alpha(k) = \ell\}$ to be the set of indices that map to ℓ.

For HMMs, we then express the expectations in Eq. (O3) as:

$$\left\langle \prod_{k=0}^{K} g_k(X_{t_k}) \right\rangle = \left\langle \prod_{\ell=0}^{\kappa} g_{\alpha^{-1}(\ell)}(X_{t'_\ell}) \right\rangle = \text{tr}\left[|\mathbf{1}\rangle\langle\pi| \Omega_{g_{\alpha^{-1}(0)}} \prod_{\ell=1}^{\kappa} T^{t'_\ell - t'_{\ell-1}} \Omega_{g_{\alpha^{-1}(\ell)}} \right], \tag{O4}$$

where $\text{tr}(\cdot)$ denotes the trace, the product of operators on the right maintains time ordering, we define $g_{\alpha^{-1}(\ell)}(x) \equiv \prod_{k \in \alpha^{-1}(\ell)} g_k(x)$, and $\Omega_g \equiv \sum_{s \in \mathcal{S}} \langle g(X) \rangle_{p(X|s)} |s\rangle\langle s|$.

The summations over all time variables in Eq. (O3) induce all possible functions α that permute and compress the indices. And, within each compressed time-ordering, all possible values of the indices consistent with that ordering are summed over. To enumerate all possible compressed time-orderings, it is useful to explicitly introduce the set $\mathbb{F}_K^{(\kappa)}$ where $\mathbb{F}_K^{(\kappa)}$ is the set of all surjective functions mapping $\{0, 1, \ldots K\}$ onto $\{0, 1, \ldots \kappa\}$. Then, we can rewrite Eq. (O3) in terms of the new time-ordered set of variables $(t'_0, t'_1, \ldots t'_\kappa)$ where $t'_k - t'_{k-1} > 0$ for all $k > 0$. Dropping the prime on the t'_k variables, we obtain:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K) = \lim_{N \to \infty} \frac{1}{N} \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \sum_{t_0=1}^{N-\kappa} \sum_{t_1=t_0+1}^{N-\kappa+1} \cdots \sum_{t_\kappa=t_{\kappa-1}+1}^{N} \left\langle \prod_{k=0}^{\kappa} g_{\alpha^{-1}(k)}(X_{t_k}) \right\rangle \prod_{k=0}^{\kappa} e^{-i\omega_{\alpha^{-1}(k)} t_k}, \tag{O5}$$

where $\omega_{\alpha^{-1}(k)} \equiv \sum_{\ell \in \alpha^{-1}(k)} \omega_\ell$.

The manifest time-ordering in Eq. (O5) allows us to use Eq. (O4) for the expectation. It is convenient to rewrite this as:

$$\left\langle \prod_{k=0}^{\kappa} g_{\alpha^{-1}(k)}\big(X_{t_k}\big) \right\rangle = \text{tr}\left\{ |\mathbf{1}\rangle\langle\pi| \Omega_{g_{\alpha^{-1}(0)}} T^{-t_0} \left[\prod_{k=1}^{\kappa-1} T^{t_k} \Omega_{g_{\alpha^{-1}(k)}} T^{-t_k} \right] T^{t_\kappa} \Omega_{g_{\alpha^{-1}(\kappa)}} \right\}. \tag{O6}$$

Technically this assumes that the index of the transition matrix is bounded by $\nu_0(T) \leqslant t_k - t_{k-1}$. This assumption is valid, for example, if T is not singular. Otherwise, a slight modification of the derivation is required, where the zero eigenspace is treated separately. In either case, the final expressions we obtain for polyspectra remain unchanged.

Plugging Eq. (O6) back into Eq. (O5) consolidates and eventually eliminates the t_k dependencies, starting with t_κ. To see this, we introduce $\mathfrak{z}_k^{(\alpha)} \equiv e^{-i\omega_{\alpha^{-1}(k)}}$ and rearrange terms:

$$S_{g_0,\ldots,g_K}(\omega_1, \ldots, \omega_K)$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{\kappa=0}^{K} \sum_{\alpha\in\mathbb{F}_K^{(\kappa)}} \sum_{t_0=1}^{N-\kappa} \sum_{t_1=t_0+1}^{N-\kappa+1} \cdots \sum_{t_\kappa=t_{\kappa-1}+1}^{N} \text{tr}\left\{ |\mathbf{1}\rangle\langle\pi| \Omega_{g_{\alpha^{-1}(0)}} T^{-t_0} \left[\prod_{k=1}^{\kappa-1} T^{t_k} \Omega_{g_{\alpha^{-1}(k)}} T^{-t_k} \right] T^{t_\kappa} \Omega_{g_{\alpha^{-1}(\kappa)}} \right\} \prod_{k=0}^{\kappa} e^{-i\omega_{\alpha^{-1}(k)} t_k}$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{\kappa=0}^{K} \sum_{\alpha\in\mathbb{F}_K^{(\kappa)}} \sum_{t_0=1}^{N-\kappa} \sum_{t_1=t_0+1}^{N-\kappa+1} \cdots \sum_{t_\kappa=t_{\kappa-1}+1}^{N} \langle\pi| \Omega_{g_{\alpha^{-1}(0)}} T^{-t_0} \left\{ \prod_{k=1}^{\kappa-1} \big[\mathfrak{z}_k^{(\alpha)}T\big]^{t_k} \Omega_{g_{\alpha^{-1}(k)}} T^{-t_k} \right\} \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{t_\kappa} \Omega_{g_{\alpha^{-1}(\kappa)}} |\mathbf{1}\rangle \big[\mathfrak{z}_0^{(\alpha)}\big]^{t_0}$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{\kappa=0}^{K} \sum_{\alpha\in\mathbb{F}_K^{(\kappa)}} \sum_{t_0=1}^{N-\kappa} \cdots \sum_{t_\ell=t_{\ell-1}+1}^{N-\kappa+\ell} \cdots \sum_{t_{\kappa-1}=t_{\kappa-2}+1}^{N-1} \langle\pi| \Omega_{g_{\alpha^{-1}(0)}} T^{-t_0} \left\{ \prod_{k=1}^{\kappa-1} \big[\mathfrak{z}_k^{(\alpha)}T\big]^{t_k} \Omega_{g_{\alpha^{-1}(k)}} T^{-t_k} \right\}$$

$$\times \left\{ \sum_{t_\kappa=t_{\kappa-1}+1}^{N} \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{t_\kappa} \right\} \Omega_{g_{\alpha^{-1}(\kappa)}} |\mathbf{1}\rangle \big[\mathfrak{z}_0^{(\alpha)}\big]^{t_0}. \tag{O7}$$

This results in summations of the form $\sum_{t=a}^{b}(zT)^t$. It is always true that $(I - zT)\sum_{t=a}^{b}(zT)^t = (zT)^a - (zT)^{b+1}$. Hence, when $z^{-1} \notin \Lambda_T$, the operator $(I - zT)$ can be inverted to yield:

$$\sum_{t=a}^{b}(zT)^t = (I - zT)^{-1}[(zT)^a - (zT)^{b+1}].$$

The first such summation is as follows:

$$\sum_{t_\kappa=t_{\kappa-1}+1}^{N} \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{t_\kappa} = \big[I - \mathfrak{z}_\kappa^{(\alpha)}T\big]^{-1}\left\{ \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{t_{\kappa-1}+1} - \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{N+1} \right\}$$

$$= \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{t_{\kappa-1}} T\big[I/\mathfrak{z}_\kappa^{(\alpha)} - T\big]^{-1} - \big[I - \mathfrak{z}_\kappa^{(\alpha)}T\big]^{-1}\big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{N+1}. \tag{O8}$$

As $N \to \infty$ the contribution from the rightmost term in Eq. (O8) vanishes. This follows from the fact that

$$(e^{-i\omega}T)^N = e^{-i\omega N}T^N \text{ approaches } \sum_{\lambda\in\Lambda_{\rho(T)}} (\lambda/e^{i\omega})^N T_\lambda \text{ as } N \to \infty.$$

In particular, the contribution from the decaying eigenmodes (with eigenvalue magnitude less than unity) vanishes as $N \to \infty$. However, for eigenvalues $\lambda \in \rho(T)$ on the unit circle, $(\lambda/e^{i\omega})^N$ does not converge for generic ω as $N \to \infty$. Therefore, if the polyspectrum is to be well behaved in the $N \to \infty$ limit, the contribution from these terms also must vanish.

The surviving term, leftmost in Eq. (O8), conveniently has a $T^{t_{\kappa-1}}$ operator on the lefthand side that cancels with the $T^{-t_{\kappa-1}}$ operator in Eq. (O7). In effect, for:

$$\left\{ \prod_{k=1}^{\kappa-1} \big[\mathfrak{z}_k^{(\alpha)}T\big]^{t_k} \Omega_{g_{\alpha^{-1}(k)}} T^{-t_k} \right\}\left\{ \sum_{t_\kappa=t_{\kappa-1}+1}^{N} \big[\mathfrak{z}_\kappa^{(\alpha)}T\big]^{t_\kappa} \right\}$$

we substitute:

$$\left\{ \prod_{k=1}^{\kappa-2} \big[\mathfrak{z}_k^{(\alpha)}T\big]^{t_k} \Omega_{g_{\alpha^{-1}(k)}} T^{-t_k} \right\} \big[\mathfrak{z}_{\kappa-1:\kappa}^{(\alpha)}T\big]^{t_{\kappa-1}} \Omega_{g_{\alpha^{-1}(\kappa-1)}} T\big[I/\mathfrak{z}_\kappa^{(\alpha)} - T\big]^{-1},$$

where $\mathfrak{z}_{\ell:\kappa}^{(\alpha)} \equiv \prod_{k=\ell}^{\kappa} \mathfrak{z}_k^{(\alpha)} = e^{-i\sum_{k=\ell}^{\kappa} \omega_{\alpha^{-1}(k)}}$. The $t_{\kappa-1}$ term can now be summed over in the same fashion as just done for the t_κ term. This summation and annihilation proceeds recursively to yield a surprisingly concise closed-form solution for any polyspectrum.

To carry out the specified recursion, we note that each new summation is of the form:

$$
\sum_{t_\ell=t_{\ell-1}+1}^{N-\kappa+\ell} \left(\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right)^{t_\ell} = \left[I - \mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{-1}\left\{\left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{t_{\ell-1}+1} - \left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{N-\kappa+\ell+1}\right\}
$$

$$
= \left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{t_{\ell-1}}T\left[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T\right]^{-1} - \left[I - \mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{-1}\left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{N-\kappa+\ell+1}
$$

$$
\to_{N\to\infty} \left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{t_{\ell-1}}T\left[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T\right]^{-1}.
$$

This provides the desired annihilation with $T^{-t_{\ell-1}}$, allowing the recursion. (Again, the contribution of the rightmost term vanishes for generic ω in the $N \to \infty$ limit.)

As an intermediate step in this recursive procedure, we obtain:

$$
S_{g_0,\dots,g_K}(\omega_1,\dots,\omega_K) = \lim_{N\to\infty}\frac{1}{N}\sum_{\kappa=0}^{K}\sum_{\alpha\in\mathbb{F}_K^{(\kappa)}}\sum_{t_0=1}^{N-\kappa}\cdots\sum_{t_\ell=t_{\ell-1}+1}^{N-\kappa+\ell}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T^{-t_0}\left\{\prod_{k=1}^{\ell}\left[\mathfrak{z}_k^{(\alpha)}T\right]^{t_k}\Omega_{g_{\alpha^{-1}(k)}}T^{-t_k}\right\}\left[\mathfrak{z}_{\ell+1:\kappa}^{(\alpha)}T\right]^{t_\ell}
$$

$$
\times\left\{\prod_{k=\ell+1}^{\kappa}T\left[I/\mathfrak{z}_{k:\kappa}^{(\alpha)} - T\right]^{-1}\Omega_{g_{\alpha^{-1}(k)}}\right\}|\mathbf{1}\rangle\left[\mathfrak{z}_0^{(\alpha)}\right]^{t_0}
$$

$$
= \lim_{N\to\infty}\frac{1}{N}\sum_{\kappa=0}^{K}\sum_{\alpha\in\mathbb{F}_K^{(\kappa)}}\sum_{t_0=1}^{N-\kappa}\cdots\sum_{t_\ell=t_{\ell-1}+1}^{N-\kappa+\ell}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T^{-t_0}\left\{\prod_{k=1}^{\ell-1}\left[\mathfrak{z}_k^{(\alpha)}T\right]^{t_k}\Omega_{g_{\alpha^{-1}(k)}}T^{-t_k}\right\}\left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{t_\ell}\Omega_{g_{\alpha^{-1}(\ell)}}
$$

$$
\times\left\{\prod_{k=\ell+1}^{\kappa}T\left[I/\mathfrak{z}_{k:\kappa}^{(\alpha)} - T\right]^{-1}\Omega_{g_{\alpha^{-1}(k)}}\right\}|\mathbf{1}\rangle\left[\mathfrak{z}_0^{(\alpha)}\right]^{t_0}
$$

$$
= \lim_{N\to\infty}\frac{1}{N}\sum_{\kappa=0}^{K}\sum_{\alpha\in\mathbb{F}_K^{(\kappa)}}\sum_{t_0=1}^{N-\kappa}\cdots\sum_{t_{\ell-1}=t_{\ell-2}+1}^{N-\kappa+\ell-1}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T^{-t_0}\left\{\prod_{k=1}^{\ell-1}\left[\mathfrak{z}_k^{(\alpha)}T\right]^{t_k}\Omega_{g_{\alpha^{-1}(k)}}T^{-t_k}\right\}\left[\mathfrak{z}_{\ell:\kappa}^{(\alpha)}T\right]^{t_{\ell-1}}
$$

$$
\times\left\{\prod_{k=\ell}^{\kappa}T\left[I/\mathfrak{z}_{k:\kappa}^{(\alpha)} - T\right]^{-1}\Omega_{g_{\alpha^{-1}(k)}}\right\}|\mathbf{1}\rangle\left[\mathfrak{z}_0^{(\alpha)}\right]^{t_0}.
$$

Eventually only the t_0 summation remains:

$$
S_{g_0,\dots,g_K}(\omega_1,\dots,\omega_K) = \lim_{N\to\infty}\frac{1}{N}\sum_{\kappa=0}^{K}\sum_{\alpha\in\mathbb{F}_K^{(\kappa)}}\sum_{t_0=1}^{N-\kappa}\sum_{t_1=t_0+1}^{N-\kappa+1}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T^{-t_0}\left\{\left[\mathfrak{z}_1^{(\alpha)}T\right]^{t_1}\Omega_{g_{\alpha^{-1}(1)}}T^{-t_1}\right\}
$$

$$
\times\left(\mathfrak{z}_{2:\kappa}^{(\alpha)}T\right)^{t_1}\left\{\prod_{k=2}^{\kappa}T\left[I/\mathfrak{z}_{k:\kappa}^{(\alpha)} - T\right]^{-1}\Omega_{g_{\alpha^{-1}(k)}}\right\}|\mathbf{1}\rangle\left[\mathfrak{z}_0^{(\alpha)}\right]^{t_0}
$$

$$
= \lim_{N\to\infty}\frac{1}{N}\sum_{\kappa=0}^{K}\sum_{\alpha\in\mathbb{F}_K^{(\kappa)}}\sum_{t_0=1}^{N-\kappa}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T^{-t_0}\left(\mathfrak{z}_{1:\kappa}^{(\alpha)}T\right)^{t_0}\left\{\prod_{k=1}^{\kappa}T\left[I/\mathfrak{z}_{k:\kappa}^{(\alpha)} - T\right]^{-1}\Omega_{g_{\alpha^{-1}(k)}}\right\}|\mathbf{1}\rangle\left[\mathfrak{z}_0^{(\alpha)}\right]^{t_0}
$$

$$
= \lim_{N\to\infty}\frac{1}{N}\sum_{\kappa=0}^{K}\sum_{\alpha\in\mathbb{F}_K^{(\kappa)}}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}\left\{\prod_{k=1}^{\kappa}T\left[I/\mathfrak{z}_{k:\kappa}^{(\alpha)} - T\right]^{-1}\Omega_{g_{\alpha^{-1}(k)}}\right\}|\mathbf{1}\rangle\sum_{t_0=1}^{N-\kappa}\left(\mathfrak{z}_{0:\kappa}^{(\alpha)}\right)^{t_0}.
$$

It is now crucial to notice that:

$$
\mathfrak{z}_{0:\kappa}^{(\alpha)} = e^{-i\sum_{k=0}^{\kappa}\omega_{\alpha^{-1}(k)}} = e^{-i\sum_{k=0}^{K}\omega_k} = e^{i0} = 1,
$$

since $\omega_0 = -\sum_{k=1}^{K}\omega_k$. Accordingly, the summation over t_0 becomes

$$
\sum_{t_0=1}^{N-\kappa}\left(\mathfrak{z}_{0:\kappa}^{(\alpha)}\right)^{t_0} = \sum_{t_0=1}^{N-\kappa}1^{t_0} = \sum_{t_0=1}^{N-\kappa}1 = N - \kappa.
$$

Thus, for finite K, the continuous part of the $(g_0, \ldots, g_K)$ polyspectrum has the closed-form expression:

$$S_{g_0,\ldots,g_K}(\omega_1,\ldots,\omega_K) = \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \langle \pi | \Omega_{g_{\alpha^{-1}(0)}} \left\{ \prod_{\ell=1}^{\kappa} T \big[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T \big]^{-1} \Omega_{g_{\alpha^{-1}(\ell)}} \right\} | \mathbf{1} \rangle \left(\lim_{N \to \infty} \frac{N - \kappa}{N} \right) \tag{O9}$$

$$= \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \langle \pi | \Omega_{g_{\alpha^{-1}(0)}} \left\{ \prod_{\ell=1}^{\kappa} T \big[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T \big]^{-1} \Omega_{g_{\alpha^{-1}(\ell)}} \right\} | \mathbf{1} \rangle. \tag{O10}$$

In addition to this continuous part, the polyspectrum may also have "discrete" contributions: these are hyperplanes in the K-dimensional $(\omega_k)_{k=1}^K$ frequency space where the magnitude of the polyspectrum diverges. From Eq. (O10), it is evident that divergences in the polyspectrum may only appear where the constituent resolvents $[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T]^{-1}$ diverge. In turn, this would require the scalar $1/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} = e^{i \sum_{n=\ell}^{\kappa} \sum_{m \in \alpha^{-1}(n)} \omega_m}$ to be equal to one of the eigenvalues of the transition matrix T on the unit circle, $\lambda = e^{i\omega_\lambda} \in \Lambda_{\rho(T)}$. Mathematically, this simplifies to the condition that $[\sum_{n=\ell}^{\kappa} \sum_{m \in \alpha^{-1}(n)} \omega_m] \bmod 2\pi = \omega_\lambda \bmod 2\pi$. In words, this means that polyspectra may diverge only where subsets of the frequencies sum to an eigenfrequency ω_λ. For example, it is typical to see diagonal, vertical, and horizontal large-magnitude streaks coexisting in a generic bispectrum, corresponding to $\omega_1 + \omega_2 = \omega_\lambda$, to $\omega_1 = \omega_\lambda$, and to $\omega_2 = \omega_\lambda$, respectively [155].

1. (Eigen)Spectral expansion of polyspectra

Using Eq. (9) to express the resolvent $[I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T]^{-1}$ in terms of the transition-matrix eigenvalues and spectral projection operators:

$$\big(I/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T \big)^{-1} = \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} \frac{1}{\big[1/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - \lambda \big]^{m+1}} T_{\lambda, m}, \tag{O11}$$

we again see that the time-evolution operator T's eigenspectrum directly controls the process' polyspectrum. Furthermore, recall that $T = \sum_{\lambda} (\lambda T_{\lambda,0} + T_{\lambda,1})$ and $T_{\lambda,m} T_{\zeta,n} = \delta_{\lambda,\zeta} T_{\lambda, m+n}$. With this we find:

$$S_{g_0,\ldots,g_K}(\omega_1,\ldots,\omega_K) = \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \sum_{\lambda_1 \in \Lambda_T} \sum_{m_1=0}^{\nu_{\lambda_1}-1} \sum_{\lambda_2 \in \Lambda_T} \sum_{m_2=0}^{\nu_{\lambda_2}-1} \cdots \sum_{\lambda_\kappa \in \Lambda_T} \sum_{m_\kappa=0}^{\nu_{\lambda_\kappa}-1} \frac{\langle \pi | \Omega_{g_{\alpha^{-1}(0)}} \big[\prod_{\ell=1}^{\kappa} T T_{\lambda_j, m_j} \Omega_{g_{\alpha^{-1}(\ell)}} \big] | \mathbf{1} \rangle}{\prod_{\ell=1}^{\kappa} \big[1/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - \lambda_j \big]^{m_j+1}}. \tag{O12}$$

For a diagonalizable transition matrix T, this reduces to:

$$S_{g_0,\ldots,g_K}(\omega_1,\ldots,\omega_K) = \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \sum_{\lambda_1 \in \Lambda_T} \sum_{\lambda_2 \in \Lambda_T} \cdots \sum_{\lambda_\kappa \in \Lambda_T} \frac{\langle \pi | \Omega_{g_{\alpha^{-1}(0)}} \big[\prod_{\ell=1}^{\kappa} T T_{\lambda_j} \Omega_{g_{\alpha^{-1}(\ell)}} \big] | \mathbf{1} \rangle}{\prod_{\ell=1}^{\kappa} \big[1/\mathfrak{z}_{\ell:\kappa}^{(\alpha)} - \lambda_j \big]}. \tag{O13}$$

The relationship between the polyspectrum and the eigenspectrum of the time-evolution operator parallels the lessons already discussed for power spectra. Peaks in the magnitude of general polyspectra likewise emanate from the eigenspectrum—these peaks may occur wherever a subset of the frequencies sum to the angular frequency of an eigenvalue. The peak is sharper when the eigenvalue is closer to the unit circle; the peak is more diffuse for eigenvalues with small magnitude.

2. Polyspectra examples

It is instructive to explore several special cases of the $(g_0, \ldots, g_K)$ polyspectrum. To aid in this, we explicitly construct the surjective function sets $\mathbb{F}_1^{(0)}$, $\mathbb{F}_1^{(1)}$, $\mathbb{F}_2^{(0)}$, $\mathbb{F}_2^{(1)}$, and $\mathbb{F}_2^{(2)}$, shown in Fig. 17

First consider the $(\overline{X}, X)$ polyspectrum, $S_{\overline{X},X}(\omega_1)$, which is simply the power spectrum $P(\omega_1)$. In this case, $K = 1$. So, we consider the functions contained in $\mathbb{F}_1^{(0)} = \left\{ \boxed{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \to 0} \right\}$ and $\mathbb{F}_1^{(1)} = \left\{ \boxed{\begin{smallmatrix} 1 \to 1 \\ 0 \to 0 \end{smallmatrix}}, \boxed{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \times \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}} \right\}$. For the compressive function $\alpha = \boxed{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \to 0}$, we obtain $\alpha^{-1}(0) = \{0, 1\}$, yielding:

$$\Omega_{g_{\alpha^{-1}(0)}} = \Omega_{g_{\{0,1\}}} = \Omega_{|X|^2} = \sum_{s \in \mathcal{S}} \langle |X|^2 \rangle_{\mathrm{p}(X|s)} |s\rangle \langle s|.$$

The ($\kappa = 0$) contribution to the power spectrum is thus:

$$\langle \boldsymbol{\pi} | \Omega_{|X|^2} | \mathbf{1} \rangle = \sum_{s \in \mathcal{S}} \langle |X|^2 \rangle_{\mathrm{p}(X|s)} \langle \boldsymbol{\pi} | s \rangle = \langle |x|^2 \rangle,$$

which is indeed the first term in Eq. (6). The ($\kappa = 1$) contribution to the power spectrum is as follows:

$$\sum_{\alpha \in \mathbb{F}_1^{(1)}} \langle \boldsymbol{\pi} | \Omega_{g_{\alpha^{-1}(0)}} T [e^{i \omega_{\alpha^{-1}(1)}} I - T]^{-1} \Omega_{g_{\alpha^{-1}(1)}} | \mathbf{1} \rangle,$$

FIG. 17. Five examples of $\mathbb{F}_K^{(\kappa)}$: Each is a set of surjective functions, relevant for constructing polyspectra. The two sets $\mathbb{F}_1^{(0)}$ and $\mathbb{F}_1^{(1)}$ are needed to construct general (g_0, g_1) polyspectra $S_{(g_0,g_1)}(\omega_1)$. The three sets $\mathbb{F}_2^{(0)}$, $\mathbb{F}_2^{(1)}$, and $\mathbb{F}_2^{(2)}$ are needed to construct general (g_0, g_1, g_2) polyspectra $S_{(g_0,g_1,g_2)}(\omega_1, \omega_2)$.

where it should be recalled that $\omega_0 = -\omega_1$. Plugging in the identity and swap functions of $\mathbb{F}_1^{(1)}$, this becomes

$$
\langle \boldsymbol{\pi} | \Omega_{g_0} T (e^{i\omega_1} I - T)^{-1} \Omega_{g_1} | \mathbf{1} \rangle + \langle \boldsymbol{\pi} | \Omega_{g_1} T (e^{i\omega_0} I - T)^{-1} \Omega_{g_0} | \mathbf{1} \rangle = \langle \boldsymbol{\pi} | \Omega_{\overline{X}} T (e^{i\omega_1} I - T)^{-1} \Omega_X | \mathbf{1} \rangle + \langle \boldsymbol{\pi} | \Omega_X T (e^{-i\omega_1} I - T)^{-1} \Omega_{\overline{X}} | \mathbf{1} \rangle
$$
$$
= 2 \mathrm{Re} \langle \boldsymbol{\pi} | \Omega_{\overline{X}} T (e^{i\omega_1} I - T)^{-1} \Omega_X | \mathbf{1} \rangle,
$$

which is indeed the last term of Eq. (6).

To see the general structure of other polyspectra, it is helpful to expand the first few κ terms of the general polyspectra analytic expression Eq. (O10). Explicitly expanding the κ terms from 0 to 2 yields:

$$
S_{g_0,...,g_K}(\omega_1, \ldots, \omega_K) = \sum_{\kappa=0}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \langle \pi | \Omega_{g_{\alpha^{-1}(0)}} \left\{ \prod_{\ell=1}^{\kappa} T \left[I / \mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T \right]^{-1} \Omega_{g_{\alpha^{-1}(\ell)}} \right\} | \mathbf{1} \rangle
$$

$$
= \langle \pi | \Omega_{g_{[0,1,\ldots,K]}} | \mathbf{1} \rangle + \left\{ \sum_{\alpha \in \mathbb{F}_K^{(1)}} \langle \pi | \Omega_{g_{\alpha^{-1}(0)}} T \left[e^{i\omega_{\alpha^{-1}(1)}} I - T \right]^{-1} \Omega_{g_{\alpha^{-1}(1)}} | \mathbf{1} \rangle \right\}
$$

$$
+ \left\{ \sum_{\alpha \in \mathbb{F}_K^{(2)}} \langle \pi | \Omega_{g_{\alpha^{-1}(0)}} T \left\{ e^{i[\omega_{\alpha^{-1}(1)} + \omega_{\alpha^{-1}(2)}]} I - T \right\}^{-1} \Omega_{g_{\alpha^{-1}(1)}} T \left[e^{i\omega_{\alpha^{-1}(2)}} I - T \right]^{-1} \Omega_{g_{\alpha^{-1}(2)}} | \mathbf{1} \rangle \right\}
$$

$$
+ \sum_{\kappa=3}^{K} \sum_{\alpha \in \mathbb{F}_K^{(\kappa)}} \langle \pi | \Omega_{g_{\alpha^{-1}(0)}} \left\{ \prod_{\ell=1}^{\kappa} T \left[I / \mathfrak{z}_{\ell:\kappa}^{(\alpha)} - T \right]^{-1} \Omega_{g_{\alpha^{-1}(\ell)}} \right\} | \mathbf{1} \rangle. \tag{O14}
$$

From Eq. (O14), it is now easy to specialize to the $(\overline{X - \langle X \rangle}, X - \langle X \rangle, X - \langle X \rangle)$ polyspectrum denoted $S_{\overline{X - \langle X \rangle}, X - \langle X \rangle, X - \langle X \rangle}(\omega_1, \omega_2)$. This is the *cumulant bispectrum*, $S_{\overline{X}, X, X}^{\mathrm{cumulant}}(\omega_1, \omega_2)$, since the third-order cumulants of the original

time series are the same as the third-order moments of the modified time series with subtracted mean [128,133]. It is as follows:

$$S_{\overline{X},X,X}^{\text{cumulant}}(\omega_1, \omega_2) = S_{\overline{X-\langle X\rangle},X-\langle X\rangle,X-\langle X\rangle}(\omega_1, \omega_2)$$

$$= \langle\pi|\Omega_{|X-\langle x\rangle|^2(X-\langle x\rangle)}|\mathbf{1}\rangle + \left\{\sum_{\alpha\in\mathbb{F}_2^{(1)}}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T[e^{i\omega_{\alpha^{-1}(1)}}I - T]^{-1}\Omega_{g_{\alpha^{-1}(1)}}|\mathbf{1}\rangle\right\}$$

$$+ \sum_{\alpha\in\mathbb{F}_2^{(2)}}\langle\pi|\Omega_{g_{\alpha^{-1}(0)}}T\{e^{i[\omega_{\alpha^{-1}(1)}+\omega_{\alpha^{-1}(2)}]}I - T\}^{-1}\Omega_{g_{\alpha^{-1}(1)}}T[e^{i\omega_{\alpha^{-1}(2)}}I - T]^{-1}\Omega_{g_{\alpha^{-1}(2)}}|\mathbf{1}\rangle. \quad (O15)$$

This leads to a fraudulent white noise theorem for the cumulant bispectrum, reminiscent of Corollary 1.

Theorem 6. Any hidden Markov chain with any arbitrary state-paired collection of equal-mean distributions, i.e.:

$$\mathcal{P} \in \{\{\mathrm{p}(X|s)\}_{s\in\mathcal{S}} : \langle X\rangle_{\mathrm{p}(X|s)} = \langle x\rangle \text{ for all } s \in \mathcal{S}\},$$

generates a flat bispectrum that is constant over all frequencies (ω_1, ω_2).

Proof. Equation (O15) shows that the cumulant bispectrum consists of contributions from $\mathbb{F}_2^{(0)}$, $\mathbb{F}_2^{(1)}$, and $\mathbb{F}_2^{(2)}$. The only $\mathbb{F}_2^{(0)}$ contribution is $\langle\pi|\Omega_{|X-\langle x\rangle|^2(X-\langle x\rangle)}|\mathbf{1}\rangle$, which is a constant independent of frequency. Whereas we show that each contribution from $\mathbb{F}_2^{(1)}$ and $\mathbb{F}_2^{(2)}$ is identically zero if the stochastic process can be generated by a HMM with equal-mean PDFs associated with each state. For such processes, $\langle X\rangle_{\mathrm{p}(X|s)} = \langle x\rangle$, where $\langle x\rangle$ is independent of the latent state s.

With the aid of Fig. 17, it is easy to verify that, for each $\alpha \in \mathbb{F}_2^{(1)}$, either $\Omega_{g_{\alpha^{-1}(0)}}$ or $\Omega_{g_{\alpha^{-1}(1)}}$ equals either $\Omega_{X-\langle x\rangle}$ or $\Omega_{\overline{X-\langle x\rangle}}$. These latter two operators both equal the zero operator $\mathbf{0}$ since:

$$\Omega_{X-\langle x\rangle} = \sum_{s\in\mathcal{S}}\langle X - \langle x\rangle\rangle_{p(X|s)}|s\rangle\langle s| = \sum_{s\in\mathcal{S}}[\langle X\rangle_{p(X|s)} - \langle x\rangle]|s\rangle\langle s| = \mathbf{0}$$

and

$$\Omega_{\overline{X-\langle x\rangle}} = \sum_{s\in\mathcal{S}}\langle\overline{X} - \overline{\langle x\rangle}\rangle_{p(X|s)}|s\rangle\langle s| = \sum_{s\in\mathcal{S}}[\overline{\langle X\rangle_{p(X|s)}} - \overline{\langle x\rangle}]|s\rangle\langle s| = \mathbf{0}.$$

Each potential contribution from $\alpha \in \mathbb{F}_2^{(1)}$ is therefore a product of zero and thus vanishes.

Again with the aid of Fig. 17, it is easy to verify that for each $\alpha \in \mathbb{F}_2^{(2)}$, the operators $\Omega_{g_{\alpha^{-1}(0)}}$, $\Omega_{g_{\alpha^{-1}(1)}}$, and $\Omega_{g_{\alpha^{-1}(2)}}$ are equal to either $\Omega_{X-\langle x\rangle}$ or $\Omega_{\overline{X-\langle x\rangle}}$ which—as we showed—are all zero. Each potential contribution from $\alpha \in \mathbb{F}_2^{(2)}$ is therefore a product of zero and so vanishes.

For such processes, this establishes that the only nonzero contribution to the cumulant bispectrum is independent of frequency. The corresponding cumulant bispectrum is thus flat with uniform height $\langle|x - \langle x\rangle|^2(x - \langle x\rangle)\rangle$.

[1] M. E. J. Obien, K. Deligkaris, T. Bullmann, D. J. Bakkum, and U. Frey, Revealing neuronal function through microelectrode array recordings, Front. Neurosci. **8**, 423 (2015).

[2] S. Kolenikov, D. Steinley, and L. Thombs, *Statistics in the Social Sciences: Current Methodological Developments* (John Wiley & Sons, New York, 2010).

[3] R. B. Stein, E. R. Gossen, and K. E. Jones, Neuronal variability: Noise or part of the signal? Nat. Rev. Neurosci. **6**, 389 (2005).

[4] R. K. Pathria and P. D. Beale, *Statistical Mechanics* (Elsevier Science, Amsterdam, 2011).

[5] R. Kubo, The fluctuation–dissipation theorem, Rep. Prog. Phys. **29**, 255 (1966).

[6] M. Planck, Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum, Verhandl. Dtsc. Phys. Ges. **2**, 237 (1900).

[7] A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, Ann. Phys. **322**, 132 (1905).

[8] M. J. Klein, Max Planck and the beginnings of the quantum theory, Arch. Hist. Exact Sci. **1**, 459 (1961).

[9] N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo *et al.* (Planck Collaboration), Planck 2018 results. I. Overview and the cosmological legacy of Planck, Astron. Astrophys. **641**, A1 (2020).

[10] P. Holmes, 'Strange' phenomena in dynamical systems and their physical implications, Appl. Math. Model. **1**, 362 (1977).

[11] J. D. Farmer, J. P. Crutchfield, H. Froehling, N. H. Packard, and R. S. Shaw, Power spectra and mixing properties of strange attractors, Ann. N. Y. Acad. Sci. **357**, 453 (1980).

[12] M. Pollicott, On the rate of mixing of Axiom A flows, Invent. Math. **81**, 413 (1985).

[13] D. Ruelle, Resonances of Chaotic Dynamical Systems, Phys. Rev. Lett. **56**, 405 (1986).

[14] P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge University Press, Cambridge, UK, 2005), Vol. 9.

[15] K. Pance, W. Lu, and S. Sridhar, Quantum Fingerprints of Classical Ruelle-Pollicott Resonances, Phys. Rev. Lett. **85**, 2737 (2000).

[16] I. García-Mata and M. Saraceno, Spectral properties and classical decays in quantum open systems, Phys. Rev. E **69**, 056211 (2004).

[17] M. D. Chekroun, J. D. Neelin, D. Kondrashov, J. C. McWilliams, and M. Ghil, Rough parameter dependence in

climate models and the role of Ruelle-Pollicott resonances, Proc. Natl. Acad. Sci. USA **111**, 1684 (2014).

[18] W. H. Press, Flicker noises in astronomy and elsewhere, Comments Astrophys. **7**, 103 (1978).

[19] B. B. Mandelbrot, *Multifractals and 1/f Noise: Wild Self-Affinity in Physics (1963–1976)* (Springer, Berlin, 1999).

[20] T. Graves, R. Gramacy, N. Watkins, and C. Franzke, A brief history of long memory: Hurst, mandelbrot and the road to ARFIMA, 1951–1980, Entropy **19**, 437 (2017).

[21] P. Bak, C. Tang, and K. Wiesenfeld, Self-Organized Criticality: An Explanation of the 1/f Noise, Phys. Rev. Lett. **59**, 381 (1987).

[22] D. Krapf, E. Marinari, R. Metzler, G. Oshanin, X. Xu, and A. Squarcini, Power spectral density of a single Brownian trajectory: What one can and cannot learn from it, New J. Phys. **20**, 023029 (2018).

[23] D. Krapf, N. Lukat, E. Marinari, R. Metzler, G. Oshanin, C. Selhuber-Unkel, A. Squarcini, L. Stadler, M. Weiss, and X. Xu, Spectral Content of a Single Non-Brownian Trajectory, Phys. Rev. X **9**, 011019 (2019).

[24] S. Sadegh, E. Barkai, and D. Krapf, 1/f noise for intermittent quantum dots exhibits non-stationarity and critical exponents, New J. Phys. **16**, 113054 (2014).

[25] A. Dechant and E. Lutz, Wiener-Khinchin Theorem for Non-stationary Scale-Invariant Processes, Phys. Rev. Lett. **115**, 080603 (2015).

[26] N. Leibovich and E. Barkai, Aging Wiener-Khinchin Theorem, Phys. Rev. Lett. **115**, 080602 (2015).

[27] J. D. Watson and F. H. C. Crick, A structure for deoxyribose nucleic acid, Nature **171**, 737 (1953).

[28] M. H. F. Wilkins, H. R. Wilson, and A. R. Stokes, Molecular structure of deoxypentose nucleic acids, Nature **171**, 738 (1953).

[29] R. E. Franklin and R. G. Gosling, Molecular configuration in sodium thymonucleate, Nature **171**, 740 (1953).

[30] J. D. Watson and F. H. C. Crick, Genetical implications of the structure of deoxyribonucleic acid, Nature **171**, 964 (1953).

[31] W. Bair, C. Koch, W. Newsome, and K. Britten, Power spectrum analysis of bursting cells in area MT in the behaving monkey, J. Neurosci. **14**, 2870 (1994).

[32] F. Farkhooi, M. F. Strube-Bloss, and M. P. Nawrot, Serial correlation in neural spike trains: Experimental evidence, stochastic modeling, and single neuron variability, Phys. Rev. E **79**, 021905 (2009).

[33] B. Dummer, S. Wieland, and B. Lindner, Self-consistent determination of the spike-train power spectrum in a neural network with sparse connectivity, Front. Comp. Neurosci. **8**, 104 (2014).

[34] R. F. O. Pena, S. Vellmer, D. Bernardi, A. C. Roque, and B. Lindner, Self-consistent scheme for spike-train power spectra in heterogeneous sparse networks, Front. Comp. Neurosci. **12**, 9 (2018).

[35] M. J. Chacron, L. Maler, and J. Bastian, Electroreceptor neuron dynamics shape information transmission, Nat. Neurosci. **8**, 673 (2005).

[36] T. Gollisch and M. Meister, Rapid neural coding in the retina with relative spike latencies, Science **319**, 1108 (2008).

[37] P. Mellado, J. Diedler, and T. Steiner, Continuous EEG monitoring in the ICU, in *Intensive Care Medicine* (Springer, Berlin, 2007), pp. 693–704.

[38] J. Claassen, F. S. Taccone, P. Horn, M. Holtkamp, N. Stocchetti, and M. Oddo, Recommendations on the use of EEG monitoring in critically ill patients: Consensus statement from the neurointensive care section of the ESICM, Intensive Care Med. **39**, 1337 (2013).

[39] W. O. Tatum, Ellen R. Grass lecture: Extraordinary EEG, Neurodiagnostic J. **54**, 3 (2014).

[40] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra—From the Point of View of Communications Engineering* (Dover, New York, 1958).

[41] P. Stoica and R. L. Moses, *Spectral Analysis of Signals* (Pearson Prentice Hall, Upper Saddle River, NJ, 2005).

[42] For wide-sense stationary processes that are also ergodic, the expected value over realizations $\gamma(\tau) = \langle \overline{X_t} X_{t+\tau} \rangle_{\mathrm{p}(X_t, X_{t+\tau})}$ coincides with the time average $\langle \overline{x}_t x_{t+\tau} \rangle_t$ over a single realization. In practice, scientists and engineers typically perform this latter time average to obtain the autocorrelation function $\gamma(\tau) = \langle \overline{X_t} X_{t+\tau} \rangle_{\mathrm{p}(X_t, X_{t+\tau})} = \langle \overline{x}_t x_{t+\tau} \rangle_t$ from experimental data.

[43] N. Wiener, Generalized harmonic analysis, Acta Math. **55**, 117 (1930).

[44] A. Khintchine, Korrelationstheorie der stationären stochastischen prozesse, Math. Ann. **109**, 604 (1934).

[45] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2011).

[46] A. Acín and L. Masanes, Certified randomness in quantum physics, Nature **540**, 213 (2016).

[47] L. Shen, J. Lee, L. P. Thinh, J. D. Bancal, A. Cerè, A. Lamas-Linares, A. Lita, T. Gerrits, S. W. Nam, V. Scarani, and C. Kurtsiefer, Randomness Extraction from Bell Violation with Continuous Parametric Down-Conversion, Phys. Rev. Lett. **121**, 150402 (2018).

[48] Y. Liu, Q. Zhao, M.-H. Li, J.-Y. Guan, Y. Zhang, B. Bai, W. Zhang, W.-Z. Liu, C. Wu, X. Yuan *et al.*, Device-independent quantum random-number generation, Nature **562**, 548 (2018).

[49] The noise standard deviation being proportional to the square root of the time step—$\sigma_\eta \propto \sqrt{dt}$—rather than proportional to the time step itself is familiar in numerically integrating Langevin and other stochastic differential equations. Though sometimes met with confusion, the dependence is a direct consequence of the central limit theorem since many independent noise contributions accumulate within each time step and the variance of the accumulated value is proportional to the number of contributions. The simplest stochastic integral—integrating Gaussian white noise—produces the famous Wiener process W_t which is the canonical model for Brownian motion. Conversely, regularly sampling changes in the Wiener state produces Gaussian white noise: $W_{t+\tau_0} - W_t \sim \mathcal{N}(0, \tau_0)$.

[50] Gaussian genuine white noise is a particular case of a *Gaussian process*. Gaussian processes, often considered in machine learning, are defined as those processes with a multivariate Gaussian distribution over all finite collections of observable random variables $\{X_t\}_{t \in \mathcal{T}}$, where $\mathcal{T}$ is any finite set of times [52, chap. 6.4] [53]. There is only one Gaussian process consistent with each choice of second-order statistics. Therefore, Gaussian genuine white noise is the *only* Gaussian process that is a white noise. (Infinitely many other processes can also have low-order multivariate Gaussian distributions but, if they are

not multivariate Gaussian distributed to all orders, they are not "Gaussian processes" in the technical sense.).

[51] We choose to ignore the power spectrum's δ function at zero frequency when classifying white noise, since it merely corresponds to a constant (dc) offset of the signal. Naturally, we make an exception for *completely* deterministic dc signals (with no noise at all), since these should not be considered among the white noise processes.

[52] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, Berlin, 2006).

[53] J. Görtler, R. Kehlbeck, and O. Deussen, A visual exploration of Gaussian processes, Distill (2019).

[54] A. Paz, *Introduction to Probabilistic Automata* (Academic Press, New York, 1971).

[55] L. R. Rabiner and B. H. Juang, An introduction to hidden Markov models, IEEE ASSP Mag. **3**, 4 (1986).

[56] L. R. Rabiner, A tutorial on hidden Markov models and selected applications, IEEE Proc. **77**, 257 (1989).

[57] R. J. Elliot, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, Vol. 29 of Applications of Mathematics (Springer, New York, 1995).

[58] Y. Ephraim and N. Merhav, Hidden Markov processes, IEEE Trans. Info. Th. **48**, 1518 (2002).

[59] D. Kelly, M. Dillingham, A. Hudson, and K. Wiesner, A new method for inferring hidden Markov models from noisy time sequences, PLoS ONE **7**, e29703 (2012).

[60] J. Bechhoefer, Hidden Markov models for stochastic thermodynamics, New. J. Phys. **17**, 075003 (2015).

[61] R. G. James, J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield, Many roads to synchrony: Natural time scales and their algorithms, Phys. Rev E **89**, 042135 (2014).

[62] B. O. Koopman, Hamiltonian systems and transformation in Hilbert space, Proc. Natl. Acad. Sci. USA **17**, 315 (1931).

[63] J. P. Crutchfield and N. H. Packard, Symbolic dynamics of noisy chaos, Physica D **7**, 201 (1983).

[64] H. Jaeger, Observable operator models for discrete stochastic time series, Neural Comput. **12**, 1371 (2000).

[65] A. M. Fraser, *Hidden Markov Models and Dynamical Systems* (SIAM, Philadelphia, PA, 2008), Vol. 107.

[66] P. M. Riechers and J. P. Crutchfield, Spectral simplicity of apparent complexity, Part I: The nondiagonalizable metadynamics of prediction, Chaos **28**, 033115 (2018).

[67] P. M. Riechers and J. P. Crutchfield, Spectral simplicity of apparent complexity, Part II: Exact complexities and complexity spectra, Chaos **28**, 033116 (2018).

[68] P. M. Riechers and J. P. Crutchfield, Beyond the spectral theorem: Spectrally decomposing arbitrary functions of nondiagonalizable operators, AIP Adv. **8**, 065305 (2018).

[69] An easy explanation for this relies on the spectral decomposition of the resolvent, given in the next section.

[70] P. M. Riechers, D. P. Varn, and J. P. Crutchfield, Diffraction patterns of layered close-packed structures from hidden Markov models, arXiv:1410.5028.

[71] E. Narevicius, P. Serra, and N. Moiseyev, Critical phenomena associated with self-orthogonality in non-Hermitian quantum mechanics, Europhys. Lett. **62**, 789 (2003).

[72] A. V. Sokolov, A. A. Andrianov, and F. Cannata, Non-Hermitian quantum mechanics of non-diagonalizable Hamiltonians: Puzzles with self-orthogonal states, J. Phys. A **39**, 10207 (2006).

[73] A. Mostafazadeh, Spectral Singularities of Complex Scattering Potentials and Infinite Reflection and Transmission Coefficients at Real Energies, Phys. Rev. Lett. **102**, 220402 (2009).

[74] M. Blank, G. Keller, and C. Liverani, Ruelle–Perron–Frobenius spectrum for Anosov maps, Nonlinearity **15**, 1905 (2002).

[75] In general, the spectrum of a time-evolution operator is highly sensitive to the function space it transforms. However, the point spectrum—i.e., the eigenvalues—reflects the operator's simplicity once the appropriate function space has been identified. Fortunately, convergence of eigenvalues and their spectral projections is guaranteed when using a stochastically-smoothed modification to Ulam's method, as in Ref. [74, Theorem 4]. (Note that this can be extended to more general manifolds.) Measurement naturally induces an Ulam-type partitioning, while the impossibility of perfectly isolating a physical system inevitably yields the stochastic smoothing. Ulam-type partitioning of state space, with arbitrary observables on the partitions, turns dynamical systems into HMMs—effectively broadening the scope of our analysis. Extended dynamical mode decomposition (EDMD) can also be used to obtain the spectral projections of the time-evolution operator [80]. The estimated eigenvalues and projectors from either method can then be used to apply our results to generic dynamical systems in physics.

[76] T. Petrosky and I. Prigogine, *The Liouville Space Extension of Quantum Mechanics* (John Wiley & Sons, New York, 1996), chap. 1, pp. 1–120.

[77] Indeed, discrete-time dynamics are, in a sense, more general than continuous-time dynamics. That is, continuous-time dynamics are obtained as the limiting behavior of discrete-time dynamics.

[78] One selects the simplest branch of the complex logarithm since the apparent freedom offered through its different branches corresponds to unobserved cyclic behaviors at inaccessibly small timescales that, if they exist, are aliased by the sampling frequency f_0. In other words, the different branches are a gauge freedom supporting all dynamics that could exist but are suppressed by the mapping $G \rightarrow e^{\tau_0 G}$.

[79] The relationship between discrete and continuous time is the same as that yielding the well-known conformal mapping of the interior of the unit circle in the complex plane to the left-half of the complex plane, which also relates z transforms and Laplace transforms.

[80] S. Klus, P. Koltai, and C. Schütte, On the numerical approximation of the Perron–Frobenius and Koopman operator, J. Compl. Dynam. **3**, 51 (2016).

[81] J. Bernamont, Fluctuations in the resistance of thin films, Proc. Phys. Soc. **49**, 138 (1937).

[82] S. Engelberg, *Random Signals and Noise: A Mathematical Introduction* (CRC Press, Boca Raton, FL, 2018).

[83] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, New York, 2006).

[84] Indeed, the burden of resolving this phase ambiguity renders the RRXOR process infinite Markov order.

[85] J. P. Crutchfield and K. Young, Inferring Statistical Complexity, Phys. Rev. Lett. **63**, 105 (1989).

[86] J. P. Crutchfield, Between order and chaos, Nat. Phys. **8**, 17 (2012).

[87] N. Travers and J. P. Crutchfield, Infinite excess entropy processes with countable-state generators, Entropy **16**, 1396 (2014).

[88] N. Barnett and J. P. Crutchfield, Computational mechanics of input–output processes: Structured transformations and the ϵ-transducer, J. Stat. Phys. **161**, 404 (2015).

[89] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán *et al.*, Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, Nature **526**, 682 (2015).

[90] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J. A. Larsson, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons, Phys. Rev. Lett. **115**, 250401 (2015).

[91] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman *et al.*, Strong Loophole-Free Test of Local Realism, Phys. Rev. Lett. **115**, 250402 (2015).

[92] M. Esposito, K. Lindenberg, and C. Van den Broeck, Entropy production as correlation between system and reservoir, New J. Phys. **12**, 013013 (2010).

[93] M. Van Raamsdonk, Building up spacetime with quantum entanglement, Gen. Relativ. Gravit. **42**, 2323 (2010).

[94] L. Susskind, Entanglement is not enough, Fortschr. Phys. **64**, 49 (2016).

[95] The reader may directly verify the measurement probabilities of the entangled quantum states generated by the provided circuit diagrams, via the standard postulates of quantum mechanics. The associated HMM models necessarily capture the quantum dynamics of repeated measurement, since they were constructed directly from these calculated probabilities.

[96] Time-varying measurement protocols and classically correlated input states can both further enrich the observation sequence. For recent examples of the latter, see Ref. [97]— where the predictable structure of observations becomes exquisitely detailed even without entanglement. Generators of classical stochastic processes may also use persistent quantum memory, in which case they require less memory than otherwise classically necessary [98–102]. It is interesting that the quantum generators (of classical stochastic processes) considered in this section appear to require no supplementary memory at all.

[97] A. E. Venegas-Li, A. M. Jurgens, and J. P. Crutchfield, Measurement-induced randomness and structure in controlled qubit processes, Phys. Rev. E **102**, 040102(R) (2020).

[98] M. Gu, K. Wiesner, E. Rieper, and V. Vedral, Quantum mechanics can reduce the complexity of classical models, Nat. Commun. **3**, 762 (2012).

[99] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield, Occam's quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel, Sci. Rep. **6**, 20495 (2016).

[100] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield, Minimized state complexity of quantum-encoded cryptic processes, Phys. Rev. A **93**, 052317 (2016).

[101] F. C. Binder, J. Thompson, and M. Gu, Practical Unitary Simulator for Non-Markovian Complex Processes, Phys. Rev. Lett. **120**, 240502 (2018).

[102] Q. Liu, T. J. Elliott, F. C. Binder, C. Di Franco, and M. Gu, Optimal stochastic modeling with unitary quantum dynamics, Phys. Rev. A **99**, 062110 (2019).

[103] O. S. Edwards and H. S. Lipson, Imperfections in the structure of cobalt. I. Experimental work and proposed structure, Proc. R. Soc. Lond. A **180**, 268 (1942).

[104] P. M. Riechers, D. P. Varn, and J. P. Crutchfield, Pairwise correlations in layered close-packed structures, Acta Crystallogr. A **71**, 423 (2015).

[105] S. Hendricks and E. Teller, X-ray interference in partially ordered layer lattices, J. Chem. Phys. **10**, 147 (1942).

[106] M. M. J. Treacy, J. M. Newsam, and M. W. Deem, A general recursion method for calculating diffracted intensities from crystals containing planar faults, Proc. R. Soc. Lond. A **433**, 499 (1991).

[107] D. P. Varn, G. S. Canright, and J. P. Crutchfield, ϵ-machine spectral reconstruction theory: A direct method for inferring planar disorder and structure from X-ray diffraction studies, Acta Crystallogr. Sect. A **69**, 197 (2013).

[108] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976), Vol. 8.

[109] D. P. Varn and J. P. Crutchfield, Chaotic crystallography: How the physics of information reveals structural order in materials, Curr. Opin. Chem. Eng. **7**, 47 (2015).

[110] W. Bao, L. Jing, J. Velasco, Y. Lee, G. Liu, D. Tran, B. Standley, M. Aykol, S. B. Cronin, D. Smirnov *et al.*, Stacking-dependent band gap and quantum transport in trilayer graphene, Nat. Phys. **7**, 948 (2011).

[111] Dividing $P(\omega)$ by $|\psi|^2$ is the same as analyzing the "corrected" diffraction pattern (which discounts any ω dependence of ψ) along a row in reciprocal space satisfying the ($h - k = 1 \bmod 3$) restriction of the Miller indices associated with the crystalline modular layers.

[112] G. L. Taylor, Introduction to phasing, Acta Crystallogr. Sect. D **66**, 325 (2010).

[113] E. Schrödinger, *What Is Life? The Physical Aspect of the Living Cell* (Cambridge University Press, Cambridge, UK, 1944).

[114] J. H. E. Cartwright and A. L. Mackay, Beyond crystals: The dialectic of materials and information, Phil. Trans. R. Soc. A. **370**, 2807 (2012).

[115] U. Grimm, Aperiodic crystals and beyond, Acta Crystallogr. B **71**, 258 (2015).

[116] D. Blackwell and L. Koopmans, On the identifiability problem for functions of Markov chains, Ann. Math. Statist. **28**, 1011 (1957).

[117] I. Siekmann, J. Sneyd, and E. J. Crampin, MCMC can detect nonidentifiable models, Biophys. J. **103**, 2275 (2012).

[118] A. L. Hodgkin and A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, J. Physiol. **117**, 500 (1952).

[119] Y. Jiang, A. Lee, J. Chen, V. Ruta, M. Cadene, B. T. Chait, and R. MacKinnon, X-ray structure of a voltage-dependent K$^+$ channel, Nature **423**, 33 (2003).

[120] K. Kasahara, M. Shirota, and K. Kinoshita, Ion concentration and voltage-dependent push and pull mechanisms of potassium channel ion conduction, PLoS ONE **11**, e0150716 (2016).

[121] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, MA, 2005).

[122] Z. Siwy and A. Fuliński, Origin of $1/f^\alpha$ Noise in Membrane Channel Currents, Phys. Rev. Lett. **89**, 158101 (2002).

[123] H. M. Fishman, Relaxation spectra of potassium channel noise from squid axon membranes, Proc. Natl. Acad. Sci. USA **70**, 876 (1973).

[124] T. L. Hill and Y. Chen, On the theory of ion transport across the nerve membrane: IV. Noise from the open-close kinetics of K^+ channels, Biophys. J. **12**, 948 (1972).

[125] C. F. Stevens, Inferences about membrane properties from electrical noise measurements, Biophys. J. **12**, 1028 (1972).

[126] P. Hou, J. Eldstrom, J. Shi, L. Zhong, K. McFarland, Y. Gao, D. Fedida, and J. Cui, Inactivation of KCNQ1 potassium channels reveals dynamic coupling between voltage sensing and pore opening, Nat. Commun. **8**, 1730 (2017).

[127] D. R. Brillinger, An introduction to polyspectra, Ann. Math. Stat. **36**, 1351 (1965).

[128] W. B. Collis, P. R. White, and J. K. Hammond, Higher-order spectra: The bispectrum and trispectrum, Mech. Sys. Sig. Process. **12**, 375 (1998).

[129] Equation (25) is a useful expression for time-ordered correlators of observables. This contrasts with *out-of-time-order correlators* (OTOCs) that recently received renewed interest for investigating temporal correlations. OTOCs are theoretically interesting as they can quantify information scrambling in thermodynamics, chaotic systems, small quantum systems, and black holes [130]. Experimentally obtaining OTOCs requires sufficient experimental finesse to effectively reverse time—and this is simply not an option for most systems. Correlates based on time-ordered observables therefore remain prominent for most systems.

[130] B. Swingle, Unscrambling the physics of out-of-time-order correlators, Nat. Phys. **14**, 988 (2018).

[131] M. Rosenblatt and J. W. Van Ness, Estimation of the bispectrum, Ann. Math. Stat. **36**, 1120 (1965).

[132] C. L. Nikias and M. R. Raghuveer, Bispectrum estimation: A digital signal processing framework, Proc. IEEE **75**, 869 (1987).

[133] A. Petropulu, Higher-order spectral analysis, *Digital Signal Processing Handbook*, edited by V. K. Madisetti and D. B. Williams (Chapman & Hall, London, 1999).

[134] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[135] A. del Junco and M. Rahe, Finitary codings and weak Bernoulli partitions, Proc. Am. Math. Soc. **75**, 259 (1979).

[136] J. P. Crutchfield and D. P. Feldman, Regularities unseen, randomness observed: Levels of entropy convergence, Chaos **13**, 25 (2003).

[137] R. G. James, N. Barnett, and J. P. Crutchfield, Information Flows? A Critique of Transfer Entropies, Phys. Rev. Lett. **116**, 238701 (2016).

[138] R. G. James and J. P. Crutchfield, Multivariate dependence beyond Shannon information, Entropy **19**, 531 (2017).

[139] K. Young and J. P. Crutchfield, Fluctuation spectroscopy, Chaos Solitons Fract. **4**, 5 (1994).

[140] M. Budišić, R. Mohr, and I. Mezić, Applied Koopmanism, Chaos **22**, 047510 (2012).

[141] J. P. Crutchfield, The hidden fragility of complex systems: Consequences of change, changing consequences, in *Cultures of Change—Changing Cultures*, edited by P. Alsina and J. Perello (ACTARBarcelona, Spain, 2009), pp. 98–111.

[142] J. W. Tukey, The future of data analysis, Ann. Math. Stat. **33**, 1 (1962).

[143] A. Guinier, *X-Ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies* (W. H. Freeman & Company, San Francisco, 1963).

[144] B. E. Warren, *X-Ray Diffraction* (Addison-Wesley, New York, 1969).

[145] T. R. Welberry, Diffuse X-ray scattering and models of disorder, Rep. Prog. Phys. **48**, 1543 (1985).

[146] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Sanders College, New York, 1976).

[147] T. R. Welberry, One hundred years of diffuse X-ray scattering, Metallurgical Mater. Trans. A **45**, 75 (2014).

[148] E. Prince, editor, *International Tables for Crystallography Vol. C, Mathematical, Physical and Chemical Tables* (Springer, Berlin, 2004).

[149] M. M. Woolfson, *An Introduction to X-ray Crystallography* (Cambridge University Press, Cambridge, UK, 1997).

[150] D. Balzar, Profile fitting of x-ray diffraction lines and Fourier analysis of broadening, J. Appl. Crystallogr. **25**, 559 (1992).

[151] A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing* (Pearson Education, Upper Saddle River, NJ, 2014).

[152] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis* (Wiley-Interscience, New York, 1980), p. 315.

[153] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures* (John Wiley & Sons, New York, 2011), Vol. 729.

[154] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, Phys. Rev. A **97**, 012127 (2018).

[155] This observation assumes the typical convention for bispectra: that ω_1 is plotted on the horizontal axis while ω_2 is plotted on the vertical axis.