Renormalization-group approach to ordered phases in music

Ryan Buechele [*]
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Jesse Berezovsky [†]
Department of Physics, Case Western Reserve University, Cleveland, Ohio 44106, USA

(Received 12 April 2024; accepted 9 July 2024; published 29 July 2024)

The organization of disordered sounds into the ordered structures of music can be understood through an analogy to the emergent ordering of physical systems undergoing phase transitions. This work builds off a prior mean-field model for pitch in music [J. Berezovsky, Sci. Adv. **5**, eaav8490 (2019)] by using renormalization-group techniques to study the effects of dimensionality and local correlations. We corroborate the results of the mean-field model by showing convergence of the phase diagram as lattice dimension is increased, while also uncovering new phases which the mean-field model does not reveal. We also compute the nearest-neighbor correlations and provide comparisons to the mean-field model, as well as historical tuning systems used by different groups of musicians. The new phases and resulting correlations revealed in this work suggest a number of possible avenues for further exploration, including generating new music using the pitch distributions suggested by the model.

DOI: 10.1103/PhysRevE.110.014145

I. INTRODUCTION

At a fundamental level, music can be viewed as the organization of sounds into an ordered state. The ordered state of music exhibits symmetry breaking in at least two dimensions: the continuous symmetry of time is broken into discrete beats, and the continuous spectrum of sound frequency is broken into a discrete symmetry of musical pitches. Western music is composed from a set of 12 pitch classes that repeat periodically in every octave (doubling of frequency). In previous work, Berezovsky has introduced a model for quantitatively treating the ordering of pitches into musical harmony via an analogy to statistical mechanics [1]. This model has been explored in the mean-field approximation and in numerical simulation [2]. The results reproduce aspects of Western and non-Western systems of harmony, including particular n-fold octave divisions, chords, and chord progressions.

Historically, the question of how to tune the 12 pitch classes of Western music to specific frequencies has been answered in different ways. The problem lies in the fact that the small-integer frequency ratios that are perceived as consonant are not commensurate with each other. For example, pitches with frequencies in ratios of 2:1, 3:2, or 5:4 are perceived as consonant and referred to as an octave, a perfect fifth, and a major third, respectively, but no number of perfect fifths will equal a whole number of octaves [$(\frac{3}{2})^n \neq 2^m$ for any positive integers n, m]. Moreover, this incommensurability holds for any combination of these three intervals due to the differing prime numbers in each. The task of selecting a set of pitches thus requires a tradeoff between how many pitch intervals closely approximate consonant ratios and how well they approximate those ratios. The predominant system of tuning in Western music today is so-called 12-tone equal temperament, in which the 12 pitches are spaced equally on a logarithmic scale within each octave. In this system only the octaves are perfectly tuned to a 2:1 ratio, but many other combinations of pitches are reasonably close to other ratios. The other extreme is so-called "just intonation," in which certain pitches are tuned to precise ratios with each other. This is typically not used on fixed-pitch instruments such as the piano, as it severely limits the number of pitch intervals that are close to consonance. Ensembles of variable-tuning instruments, such as unfretted string instruments or the human voice, may locally adjust the tuning in real time to bring specific intervals into consonance.

Here, we present a treatment of a model of musical harmony using a renormalization-group (RG) approach. These results allow exploration of the effects of dimensionality and local correlations that are obscured in the mean-field limit. The mean-field model used in previous work relies on the unrealistic assumption that every tone interacts with all others and cannot capture behavior where some pairs of pitches are more likely to interact than others. We find that the RG model at dimension $d > 2$ recovers the same types of phases seen in the mean-field model, with increasing quantitative agreement as d increases. The RG model, however, reveals nearest-neighbor correlations that reflect just intonation of intervals, even within a phase that is equally tempered over long ranges. Moreover, we uncover new phases suggestive of novel systems of intonation, signified by unusual correlations that the mean-field model does not capture. These systems of tuning may contribute to the area of microtonal music, where the octave is divided into more than 12 pitches [3].

[*]Contact author: buechele.5@osu.edu
[†]Contact author: jab298@case.edu

The paper is organized as follows. First, we review the model developed in Ref. [1] and the mean-field results described there. Next, we describe the renormalization group method applied here. We then present results of phase diagrams and local correlations. Finally, we discuss the results in comparison to the mean-field case.

II. STATISTICAL MODEL OF MUSICAL HARMONY

Two core assumptions about human perception and psychology form the basis of the model: a system of musical harmony should simultaneously seek to minimize the perceived dissonance between pitches and maximize the variety of possible compositions. Absolutely minimizing dissonance can be achieved by using only one pitch, but this music would be rather boring for the performer or listener, which drives the second assumption that music should maximize the variety of pitches to create interest and flexibility to allow a range of artistic expression. This tension between minimizing dissonance and maximizing pitch variety mirrors the familiar behavior in physical systems that minimizes their internal energy while maximizing their entropy. Therefore, our model of a musical system of harmony will seek to minimize a "musical free energy" $F = D - TS$, where D is the dissonance between tones, S is an entropy obtained from the variety of pitches, and T is a culturally determined "temperature" parameter which regulates the tradeoff [1]. From this point, we can turn to the usual techniques of statistical mechanics that rely on minimizing free energy to model pitches in a piece of music.

A quantitative description of dissonance is needed to obtain the dissonance D of two tones, in analogy to the energy of two interacting particles. We follow an approach originating with Helmholtz [4], further developed by Plomp and Levelt in the 20th century [5], and later formalized by Sethares [6]. Recent empirical work has verified that perceived dissonance broadly agrees with the prediction of the model we use here, and similar models [7]. Tones are taken to be characterized by an amplitude A and a fundamental frequency Φ. The spectrum of the tone is described by a set of frequencies $\phi_n \propto \Phi$ $(n = 1, 2, ...)$ with associated amplitudes $\alpha_n \propto A$. These frequencies, known as partials, are integer multiples of Φ in the common case of a harmonic spectrum.

Empirical studies have mapped out how perceived dissonance depends on the frequency ratio between two single-frequency (pure) tones [5]. These empirical results implicitly depend on the innate or enculturated preferences of the subjects who contributed data. The two-pure-tone dissonance $\bar{D}$ between two frequencies ϕ_i and ϕ_j can be parameterized as

$$\bar{D}(\phi_i, \phi_j) = \frac{1}{\omega_c} \exp\left[-\ln \left| \frac{\theta_{ij}/2\pi}{\omega_c} \right|^2 \right],$$
$$\text{with } \theta_{ij} = 2\pi \log_2\left(\frac{\phi_i}{\phi_j}\right). \tag{1}$$

$\bar{D}$ depends on the *critical bandwidth* ω_c, an empirically determined parameter describing the fractional frequency difference at which perceived dissonance between two frequencies is a maximum. Examples of $\bar{D}$ at several values of ω_c are shown in Fig. 1(a). Dissonance between two tones

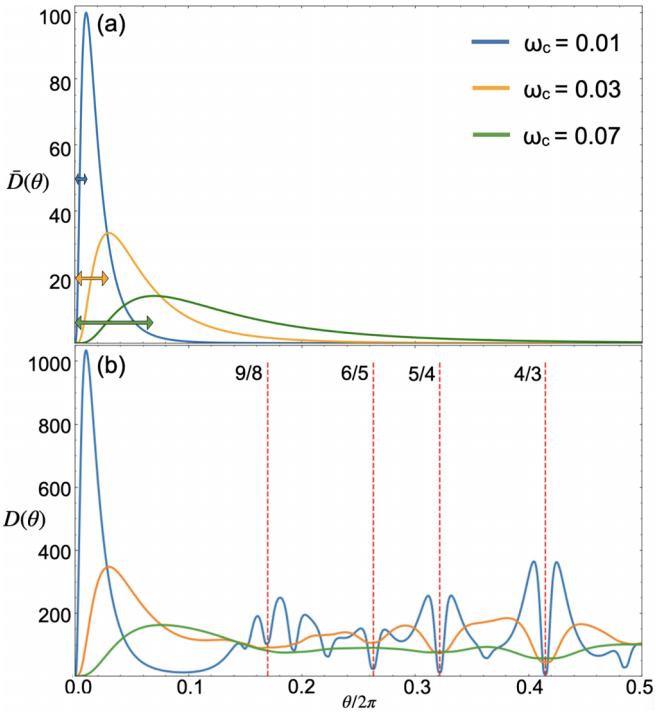

FIG. 1. (a) Two-pure-tone dissonance function $\bar{D}(\theta)$ plotted over one half octave for critical bandwidth $\omega_c = 0.01, 0.03$, and 0.07; the bandwidths are indicated with the double-headed arrows. (b) Full dissonance function $D(\theta)$ over one half octave with $\omega_c = 0.01, 0.03$, and 0.07; the red dashed lines indicate the common musical intervals of the perfect fourth ($\Phi/\Phi_0 = 4/3$), major third ($\Phi/\Phi_0 = 5/4$), minor third ($\Phi/\Phi_0 = 6/5$), and major second ($\Phi/\Phi_0 = 9/8$) which coincide with local minima in the Hamiltonian.

then arises from closely spaced, but not identical, frequency components in the spectra of the tones.

Though ω_c is found to be a function of frequency, typically ranging from the order of 0.01 to 0.1 [8,9], here we take it as a fixed parameter and study the behavior as ω_c is varied. The total dissonance between two tones with fundamental frequencies Φ_1 and Φ_2 is then obtained by summing the two-tone dissonance between all combinations of harmonics of Φ_1 and Φ_2:

$$D_0(\Phi_1, \Phi_2) = \sum_{m,n=1,2,...} \min(\alpha_{1,m},\ \alpha_{2,n})^{0.606} \bar{D}(\phi_{1,m},\ \phi_{2,n},\ \omega_c), \tag{2}$$

where the term $\min(\alpha_{1,m},\ \alpha_{2,n})^{0.606}$ approximates the perceived volume of the given pair of harmonics, so that quieter harmonics contribute less to the dissonance [10].

Throughout, we consider the simple case of a sawtooth wave with fundamental Φ, which well approximates the sound of a bowed vibrating string and has harmonics given by $\phi_n = n\Phi$ and $\alpha_n = A/n$ for $n \in \{1, 2, ..., n_{\max}\}$. We also assume octave periodicity; that is, frequencies related by a factor of 2 are considered equivalent, so that all frequencies live in the half-open interval $[\Phi_0, 2\Phi_0)$ for some arbitrary fundamental frequency Φ_0. In this way, the model can be viewed as similar to an XY model, with the angle θ from 0 to 2π being mapped onto an interval from the unison Φ_0 to the octave $2\Phi_0$ as

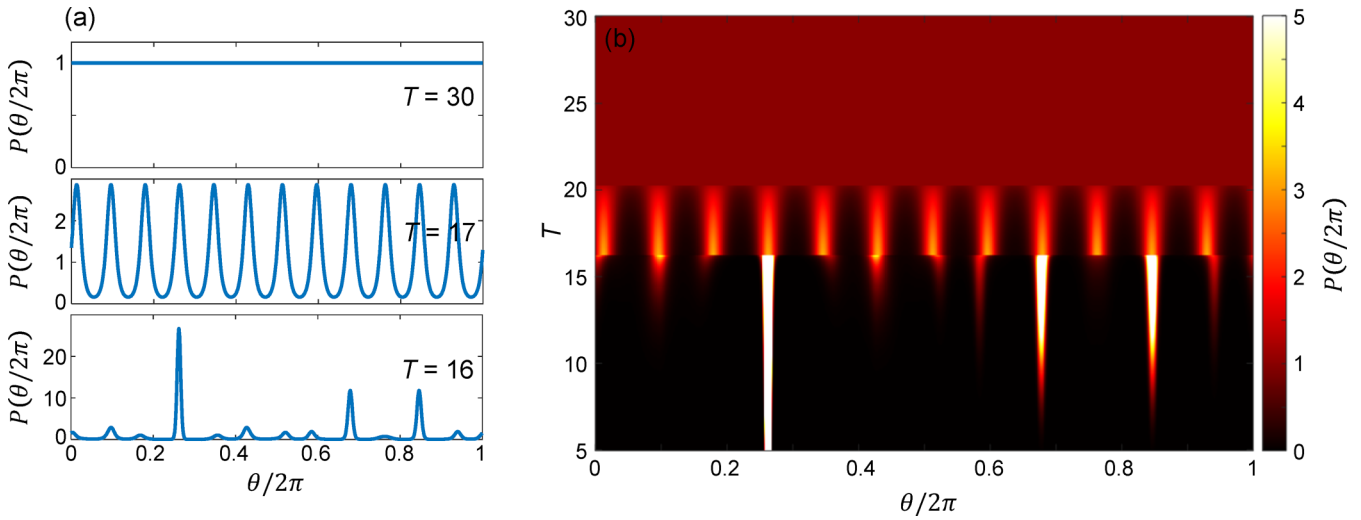

FIG. 2. (a) Pitch probability distributions $P(\theta)$ from the mean-field model in Ref. [1] at three temperatures T and $\omega_c = 0.03$. (b) Solutions $P(\theta)$ from the same model as (a), but plotted vs a range of T.

$\theta(\Phi) = 2\pi \log_2(\Phi/\Phi_0)$. This is implemented by summing each octave of D_0,

$$D(\theta_{ij}) = \sum_{n=-\infty}^{\infty} D_0(\theta_{ij} + 2\pi n), \tag{3}$$

so that $D(\Phi_0, \Phi_1) = D(2^x\Phi_0, 2^y\Phi_1)$ for $x, y \in \mathbb{Z}$.

A plot of the dissonance function $D(\theta)$ used here is shown in Fig. 1 for three different values of ω_c, with the sum over partials truncated at $n_{\max} = 10$. $D(\theta)$ is plotted over half of the octave—the other half is symmetric about $\theta = \pi$. As has been noted elsewhere, $D(\theta)$ has minima at commonly used musical intervals occurring at small-integer ratios of frequency and maxima at small deviations from these ratios. The critical bandwidth sets the width of these features, with smaller ω_c creating sharper peaks and valleys. Note that integer ratios of frequency Φ appear at irrational values of $\theta/2\pi$.

Treating the dissonance function $D(\theta)$ as an interaction energy between tones, previous work used a mean-field approximation to determine a pitch probability distribution $P(\theta)$ that minimizes free energy $F = D - TS$ [1]. At each fixed T and ω_c, D is calculated as the expectation value of $D(\theta)$:

$$D = \frac{1}{2}\int_0^{2\pi} d\theta \int_0^{2\pi} d\theta' P(\theta) D(\theta - \theta') P(\theta') \tag{4}$$

and S is obtained from the integral:

$$S = -\int_0^{2\pi} d\theta P(\theta) \ln P(\theta) \tag{5}$$

[1]; a variational approach is used to find $P(\theta)$ which minimizes F. As T is varied at a fixed ω_c, two phase transitions are observed, in general. For example, Fig. 2(a) shows $P(\theta)$ calculated at three values of T, with $\omega_c = 0.03$. At the highest T, $P(\theta)$ is a constant, representing disordered sound. As T is lowered, a transition occurs to a phase with 12 equally spaced peaks, representing the 12 pitches used in Western music. As T is lowered further, another transition occurs and $P(\theta)$ now displays 12 peaks with unequal height and spacing. The transitions between these three phases can be seen in Fig. 2(b), where $P(\theta)$ is plotted vs T.

The phase transitions seen in the mean-field model can be understood via the order parameters, which are the Fourier coefficients of $P(\theta)$ with $k = 1, 2, ...,$

$$M_k = \langle e^{ik\theta} \rangle = \int_0^{2\pi} P(\theta) e^{ik\theta} d\theta. \tag{6}$$

Figure 3(a) shows $|M_k|$ for $k = 1$–20. When $P(\theta)$ is a constant, then all $M_k = 0$. As temperature is lowered, the first phase transition is seen when M_{12} becomes nonzero. The next transition occurs when all other M_k become nonzero. This behavior is discussed in terms of Landau theory in Ref. [1], where the Landau free energy is calculated by writing $F = D - TS$ in terms of M_k and the Fourier coefficients g_k of the interaction $D(\theta)$:

$$g_k = \int_0^{2\pi} d\theta D(\theta) \cos(k\theta). \tag{7}$$

Near the higher temperature transition, all M_k are small, so inserting the Fourier representations into F and expanding to second order in M_k yields

$$F \approx \frac{1}{2} g_0 + \frac{1}{4} \sum_{k=1}^{\infty} \left(\frac{g_k}{2} + T \right) |M_k|^2. \tag{8}$$

A second-order phase transition can occur as T is lowered when one of the quadratic terms in the sum in Eq. (8) changes sign at $T = -g_k/2$, assuming higher-order terms are positive. The first term to do so is the mode labeled by k_0 such that g_{k_0} is the most negative g_k, and at this temperature M_{k_0} becomes nonzero. Since only M_{k_0} is nonzero, F can be written with a single-mode approximation as

$$F \approx \frac{1}{2} g_0 + \frac{1}{4} \left(\frac{g_{k_0}}{2} + T \right) |M_{k_0}|^2 + \cdots, \tag{9}$$

where all remaining terms are positive. From here, we clearly see that the high-temperature transition occurs at $T_c = -g_{k_0}/2$. Values of g_k are shown in Fig. 3(b) at $\omega_c = 0.03$,

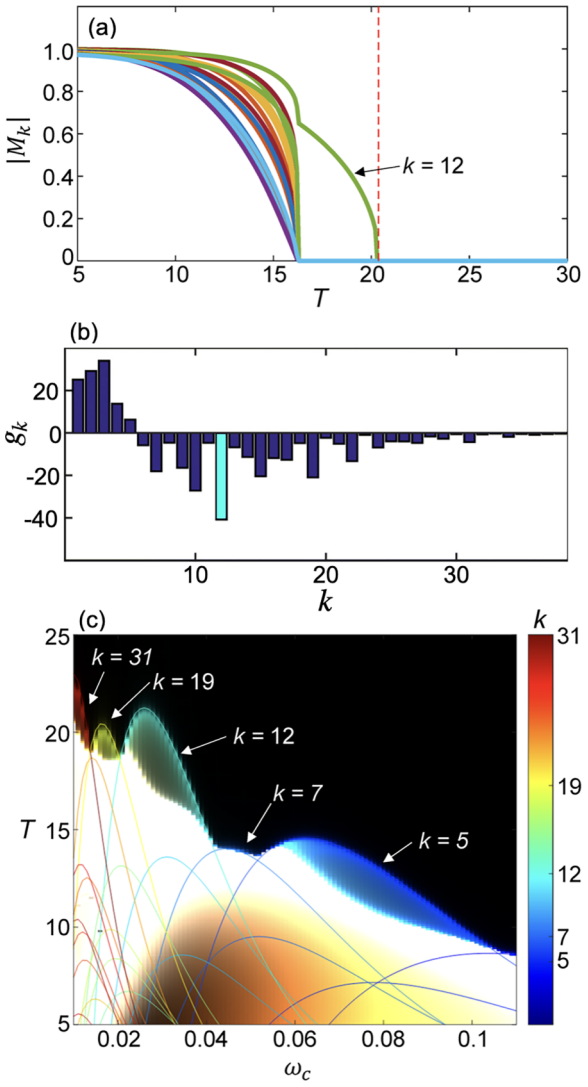

FIG. 3. (a) Order parameters $|M_k|$ from the mean-field calculation in Ref. [1], for $k = 1$–20 and $\omega_c = 0.03$. Red dashed line shows the predicted $T_c = -g_{12}/2$. (b) Fourier components g_k of $D(\theta)$ with $\omega_c = 0.03$; the most negative component (in this case $k_0 = 12$, marked in cyan) defines $T_c = -g_{k_0}/2$. (c) Mean-field phase diagram. Colors are obtained by combining RGB values for each $d|M_k|/dT$. Lines show $-g_k/2$ with k indicated by the color scale.

with g_{k_0} highlighted at $k_0 = 12$. The predicted transition temperature T_c is indicated by a dashed red line in Fig. 3(a). The lower-temperature transition is less straightforward to analyze, because there exist M_k that are not close to zero near this transition, as seen in Fig. 3(a).

When ω_c is varied, qualitatively similar behavior is obtained but with different transition temperatures, and with $P(\theta)$ exhibiting different numbers of peaks, in accordance with the prediction of Landau theory. Figure 3(c) shows this dependence as a phase diagram, where k is represented by a color scale and the plot is colored by mixing those colors according to values of $d|M_k|/dT$. The color is black when all $M_k = 0$ are constant, displays a distinct hue when one M_k becomes nonzero, and tends towards white when many M_k are becoming nonzero. Taking the derivative of M_k highlights the phase boundaries. The lines plotted in Fig. 3(c) show $-g_k/2$ vs ω_c, with the color indicating k. This shows how the higher transition temperature tracks the predicted $T_c = -g_{k_0}/2$.

III. METHODS

The mean-field approach used in Ref. [1] averages over the dissonance of every tone, which inherently assumes that each tone is equally likely to interact with every other tone in the system. In typical music, however, each tone only interacts via dissonance with at most several other tones sounded at the same time. Even in music that uses all 12 pitch classes equally, it is often the case that particular consonant combinations of pitches (i.e., chords) are more likely to be sounded together. This unrealistic mean-field assumption can be removed by using renormalization-group theory, which exploits the fact that a lattice structure is scale invariant to solve for the configuration of pitches while preserving local correlations.

The model used in this work is adapted from a treatment of superfluidity in helium films [11] that utilized a hierarchical lattice to enable the RG method. This technique has proven useful in studying the effects of lattice topology on critical phenomena [12–14], as well as critical phenomena in complex networks [15,16] and spin glasses [17,18]. While Ref. [11] focused on a two-dimensional system, the renormalization procedure can be generalized to any dimension, including noninteger fractal dimensions [19,20]. While physical systems rarely, if ever, organize into hierarchical lattices, the results using this method still shed light on the critical behavior of real systems. In the case of musical phase transitions studied here, the system has no particular *a priori* physical lattice structure, so the hierarchical lattice is a convenient choice. Moreover, a hierarchical network is arguably a reasonable representation of the hierarchical structures in tonal music, which are often understood through a reduction process akin to coarse graining, where related musical elements are combined or eliminated, leaving only the most important structural elements of the music [21,22]. The particular choices of lattices are further justified by the coordination number distribution of those lattices as compared to the typical number of notes combined together in music. (The calculation of the coordination number distribution for a given lattice is shown in Appendix A.) From the formula derived there for the fraction of sites r_j having coordination number z_j, we find, for example, that the three-dimensional (3D) lattice shown in Fig. 6(c) is a reasonable choice, with $r_0 = 87.5\%$ of sites having $z_0 = 2$ neighbors, $r_1 = 10.9\%$ of sites having $z_1 = 8$ neighbors, $r_2 = 1.4\%$ of sites having $z_2 = 32$ neighbors, and so on, with increasingly small fractions for larger coordination numbers.

This work studies the behavior of pitches on a hierarchical lattice which is constructed recursively by replacing all single bonds between sites with a particular arrangement of bonds and sites. For example, consider an initial system of just two sites connected by a bond (the bond indicates that the two sites interact). That single bond is replaced by a motif consisting of additional sites connected to each other and to the original two sites by additional bonds. An example of such a motif is shown in the top of Fig. 4, where the original sites are at the left and right, and the bond between them has been

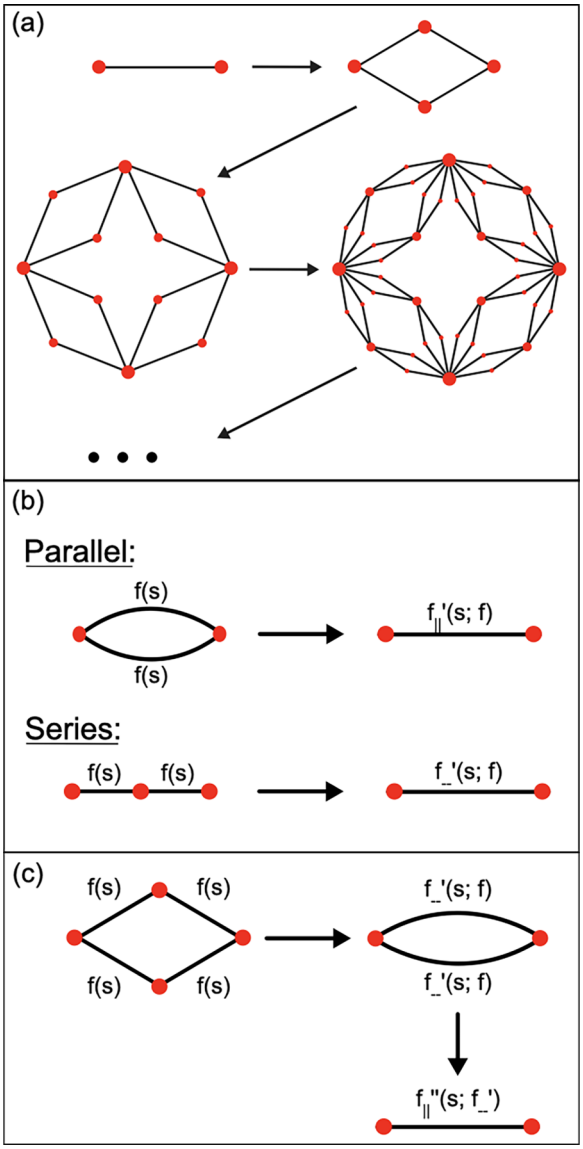

FIG. 4. (a) Iterative process which generates a hierarchical lattice by replacing each bond with a particular motif of bonds and sites. The pictured lattice has $b = q = 2$, $d = 2$. (b) Elementary decimation procedures for sets of parallel bonds and series bonds. (c) Example of how elementary decimations can be combined to deduce the RG relations for a given lattice, shown here for the $d = 2$ lattice from (a).

replaced by two new sites each connected to the original sites by a single bond. This process is then repeated where each bond in the new structure is replaced by the same motif. In general, a motif can contain q parallel branches and b bonds in series in each branch; the example in Fig. 4 has $q = b = 2$. It is permitted to connect the same two sites with multiple bonds within a branch; this increases the interaction strength and also affects the iterative construction process, as each of those bonds will be replaced by a new motif. Iterating this process to infinity creates a structure which lacks a fixed number of nearest neighbors and exhibits scale invariance, meaning that the RG transformations for the system are exact [14,20,23]. The hierarchical construction also guarantees that the sites "decimated" in each RG iteration are in local clusters with a finite number of interacting neighbors, meaning that pitches do not influence far away pitches, as in the mean-field model [12].

The dimension d of a particular lattice is determined by the formula

$$d = 1 + \frac{\ln q}{\ln b}, \tag{10}$$

where q and b represent the number of parallel and series bonds, respectively, added in each construction step; see Appendix A for a derivation of this formula. Once a lattice structure has been established, we can write down RG relations for the system.

The procedure for renormalizing the hierarchical lattice begins by writing the partition function $\mathcal{Z}$ for one step of the lattice construction in terms of its Fourier coefficients $f(s)$, where s are integers. The goal is to determine a recursive relationship for the Fourier coefficients $f'(s; f)$ of $\mathcal{Z}'$, the renormalized partition function, in terms of the coefficients $f(s)$ of $\mathcal{Z}$ [11,20,24,25]. This is accomplished by effectively undoing the lattice construction, removing lattice sites and moving bonds to reduce the construction step back to a single bond between two sites, as illustrated in Fig. 4. There are two elementary steps illustrated in Fig. 4(b), called *decimations*, which can be combined in suitable ways to determine the renormalization for any hierarchical lattice shape [25]; see Appendix B for derivation of the decimation formulas. For a *series decimation*, two bonds in series are reduced to one bond, and the Fourier coefficients $f'(s; f)$ of $\mathcal{Z}'$ for the new, single bond are given by

$$f'_{\_}(s; f) = \frac{f^2(s)}{\sum_t f^2(t)}. \tag{11}$$

For a *parallel decimation*, two bonds in parallel are combined into one bond, and the Fourier coefficients are given by the convolution sum:

$$f'_{\|}(s; f) = \frac{\sum_t f(t) f(s-t)}{\sum_u \left[\sum_v f(v) f(u-v)\right]}. \tag{12}$$

As an example, the bottom of Fig. 4 shows the decimation of a $d = 2$ motif with initial Fourier coefficients $f(s)$. First, the series bonds in each branch are decimated, resulting in renormalized coefficients $f'_{\_}(s; f)$, and an intermediate structure with the two outer sites connected by a double bond. The two parallel bonds are then decimated by applying the parallel RG formula to $f'_{\_}(s; f)$, resulting in Fourier coefficients $f''_{\|}(s; f'_{\_})$.

After determining the RG recursion relations for the lattice, the RG fixed points for the dissonance Hamiltonian with given values of the critical bandwidth ω_c and the temperature $T \equiv 1/\beta$ can be determined numerically in the following manner:

(1) Compute tabulated values of the dissonance function D at values of θ_{ij} spanning half an octave, with sufficiently small spacing. The full octave is not needed due to the symmetry of the dissonance function.

(2) Compute the Boltzmann factor appearing in the partition function $\mathcal{Z}$ for the given temperature T as $\exp^{-D(\theta_{ij})/T}$, and take a discrete Fourier transform of $\mathcal{Z}$ to compute tabulated values of the $f(s)$. The $f(s)$ must be truncated at some sufficiently large s.

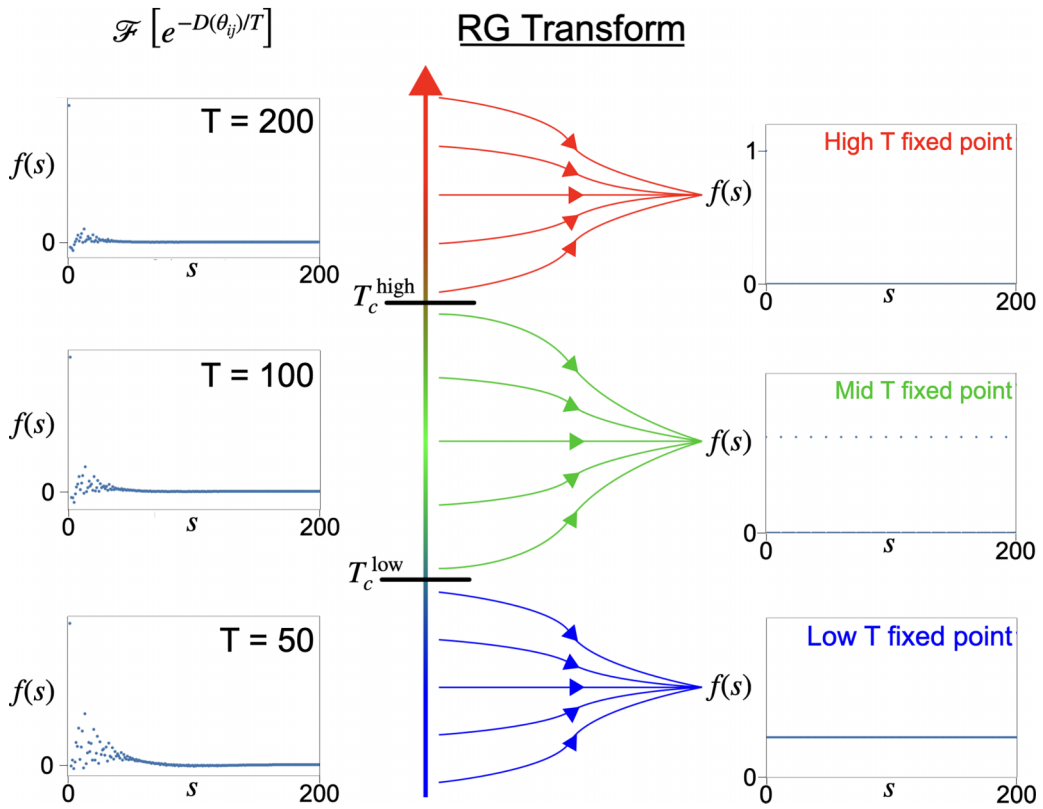

FIG. 5. Schematic overview of the method to determine fixed points with the RG procedure. Results are shown for the case of $\omega_c = 0.03$. There are generally three classes of fixed points (low, mid, and high temperature) identified for an arbitrary lattice dimension and critical bandwidth; phase transitions are identified as the temperatures at which RG flows switch from one fixed point to another.

(3) Iterate the RG recursion relations until the $f'(s;f)$ converge for all s to a fixed point in Fourier space.

In this procedure the ideally exact RG transformations become approximate due to the truncation of s at $\pm s_{\max}$. (We choose the spacing of θ_{ij} values to be small enough that the truncation of s is the limiting factor.) Since the initial values of $f(s)$ asymptotically approach zero with a scale set by $1/\omega_c$ (see left column of Fig. 5), we must choose $s_{\max} \gg 1/\omega_c$. The results shown here use $s_{\max} = 2000$. We have checked the results vs $s_{\max}$ and found little change in critical temperatures beyond $s_{\max} = 1000$. While the original $f(s)$ go to zero rapidly, some of the fixed points do not approach zero at large $|s|$, as shown below. In these cases, the truncation of s affects the computation of the convolution sums in $f'_{\|}(s;f)$. Specifically, the convolution falls off at large $|s|$ due to the finite range of the values. To correct for the reduced number of terms in the sum, we multiply the correlation sums by a correction factor $C = 2s_{\max}/(2s_{\max} - |s|)$. This factor ensures, for example, that the convolution sum of constant $f(s)$ is also a constant.

A schematic of the RG procedure is presented in Fig. 5. At the left, the initial Fourier coefficients $f(s) = \mathcal{F}[\exp(-D(\theta_{ij})/T)]$ are shown at three values of T. The RG transform is then iteratively applied to these $f(s)$. The resulting RG flow is illustrated schematically by the curved lines, where the transformed $f'(s;f)$ converge to one of several fixed points. At the right, the three typical fixed points are shown. With this method implemented, a phase diagram of the system can be constructed by determining regions in (ω_c, T) phase space corresponding to the same fixed point. In particular, a binary search algorithm across temperature on lines of constant ω_c is used to determine the temperatures of the transitions from one RG fixed point to another.

Within the Landau paradigm, phase transitions are indicated by singularities in the free energy; therefore, we can demonstrate that a phase transition occurs at a temperature where the RG flow transitions from one fixed point to another by observing singular behavior in the free energy at that temperature. The free energy F is defined in terms of the partition function $\mathcal{Z}$ as

$$F = -\frac{1}{\beta}\ln \mathcal{Z}; \tag{13}$$

we define a dimensionless free energy per bond $\mathcal{F}$ as

$$\mathcal{F} \equiv -\frac{\beta F}{B} = \frac{\ln \mathcal{Z}}{B}, \tag{14}$$

for a lattice with B bonds. To calculate the free energy, we introduce an external field term BG to the Hamiltonian $-\beta\mathcal{H}$, where G is the field coupling per bond:

$$-\beta\mathcal{H} \to -\beta\tilde{\mathcal{H}} \equiv -\beta\mathcal{H} + BG. \tag{15}$$

The parameter G is initially 0 but becomes nonzero when subjected to its RG recursion relations; for series and parallel

decimations, these are given by the following:

$$\text{Series: } G'_{\_} = 2G + \ln\left(2\pi \sum_s f^2(s)\right), \tag{16}$$

$$\text{Parallel: } G'_{\|} = 2G + \ln\left(\sum_s \left[\sum_t f(t) f(s-t)\right]\right). \tag{17}$$

Carrying out the transformation in Eq. (15), the partition function becomes

$$\tilde{\mathcal{Z}} = \mathrm{Tr}[e^{-\beta\tilde{\mathcal{H}}}] \;= e^{BG}\mathrm{Tr}[e^{-\beta\mathcal{H}}] \;= e^{BG}\mathcal{Z}; \tag{18}$$

therefore, the dimensionless free energy becomes

$$\tilde{\mathcal{F}} = \frac{\ln\tilde{\mathcal{Z}}}{B} = \frac{\ln\mathcal{Z}}{B} + \frac{BG}{B} = \mathcal{F} + G. \tag{19}$$

After one RG step, the number of bonds B' is given by B/qb, and the partition function $\tilde{\mathcal{Z}}$ is unchanged by the RG procedure (by construction); with this we can relate the initial free energy $\mathcal{F}$ to the evolution of G,

$$\mathcal{F} + G = \frac{1}{qb}(\mathcal{F}' + G'), \tag{20}$$

and after N steps,

$$\mathcal{F}^{(0)} + G^{(0)} = \frac{1}{(qb)^N}(\mathcal{F}^{(N)} + G^{(N)}). \tag{21}$$

By construction, $G^{(0)} = 0$; moreover, $\mathcal{F}^{(N)}$ remains bounded under RG transformations, while $G^{(N)}$ grows without bound as N grows. Thus, for N sufficiently large,

$$\mathcal{F} \approx \frac{G^{(N)}}{(qb)^N}, \tag{22}$$

and we can recover the true thermodynamic free energy per bond as

$$\frac{F}{B} = -\frac{\mathcal{F}}{\beta} \approx -\frac{G^{(N)}}{\beta(qb)^N}, \tag{23}$$

and calculate other thermodynamic quantities (such as entropy, heat capacity, etc.) from this.

At exactly $d = 2$, the procedure described above for identifying phases by RG fixed points fails due to the fact that the fixed points never fully converge. This is similar to the behavior of the two-dimensional (2D) XY model, which has a Berezinskii-Kosterlitz-Thouless (BKT) transition [26,27] of infinite order; in this model the renormalization flows at all temperatures drift towards the same high-temperature fixed point, but the number of iterations required to do so changes dramatically at some transition temperature [11]. To clearly identify the crossover between low- and high-temperature behavior, we rely on the free energy calculation and identify the transition temperatures as points at which we observe a peak in the heat capacity, related to the second derivative of the free energy. One could also implement a cutoff point to distinguish low- and high-temperature phases by insisting that the high-temperature phase converges to the appropriate behavior within some fixed number of iterations, and any temperature that fails to converge within that many iterations is in the low-temperature phase [11].

IV. RESULTS

We generally find three categories of fixed points in Fourier coefficient space (see Fig. 5). At low temperatures, $f(s)$ is the same constant, nonzero value for all s; at high temperatures, only $f(0)$ is nonzero. There can also be a range of intermediate temperatures with a periodic fixed point where every nth coefficient has an equal, nonzero value; the integer n changes depending on the lattice dimension and critical bandwidth ω_c.

We can map domains corresponding to the same fixed point in a phase space of T vs ω_c to construct phase diagrams for a given hierarchical lattice, shown in Fig. 6 for a variety of lattice dimensions, with the lattice motif shown in the inset. Each diagram shares the same general structure of a high-temperature phase denoted in red with dots, low-temperature phase marked in blue, and the intermediate phases appearing along the boundary between low and high temperature marked in green with diagonals. The intermediate phases are labeled with the periodicity of the fixed point. The only exception is the phase diagram for $d = 2$, which only has the low- and high-temperature fixed points. The temperature axes are scaled by the mean degree $\bar{z}$ and an additional integer factor in some cases as indicated (see Appendix B for derivation of $\bar{z}$.) These scalings are observed to bring the critical temperatures in the mean-field result into close quantitative agreement with the RG result. It is reasonable that T_c would scale with $\bar{z}$. The fact that in some cases T_c also appears to be scaled by an additional small-integer factor suggests that the scaling of T_c depends on the particular degree distribution, as is seen in other irregular networks [28,29].

As lattice dimension increases beyond $d = 2$, we see the emergence of intermediate fixed points with various periodicities in Fourier space. At $d = 3$ we see periodicities of 5 and 12. The 12-fold periodicity, which appears in an empirically realistic range of ω_c [8,9], hints towards the common scheme of dividing an octave into 12 pitches. Fivefold octave divisions are seen in the Gamelan music of Indonesia [30]. At $d = 5$ we also see small regions with periodicities of 3 and 7, suggesting other more unconventional sets of pitches. We note that many different lattices may be constructed with the same dimension. We have studied several lattices each at $d = 2$, 3, 4, and 5, and found similar phase diagrams for each dimension after scaling by $\bar{z}$.

The purple lines in Fig. 6 show the values of $-g_k/2$ vs ω_c, identical to the colored lines shown in Fig. 3(c). The greatest of these curves at a given ω_c yields the mean-field prediction for T_c and the resulting periodicity of the ordered phase. The quantitative agreement between the RG result and the mean-field result improves with increasing dimension; in particular, the values of $k = 5$, 7, and 12 become quite close at $d = 5$. However, at values $\omega_c < 0.02$ the mean-field result yields higher periodicities of 19 and 31, where the RG result instead shows lower periodicities of 3 and 5. This suggests that there are ordered phases not captured by the mean-field approximation, which will be discussed further below.

A. Free energy

In order to verify that the switching from one RG fixed point to another indicates a true phase transition, we can

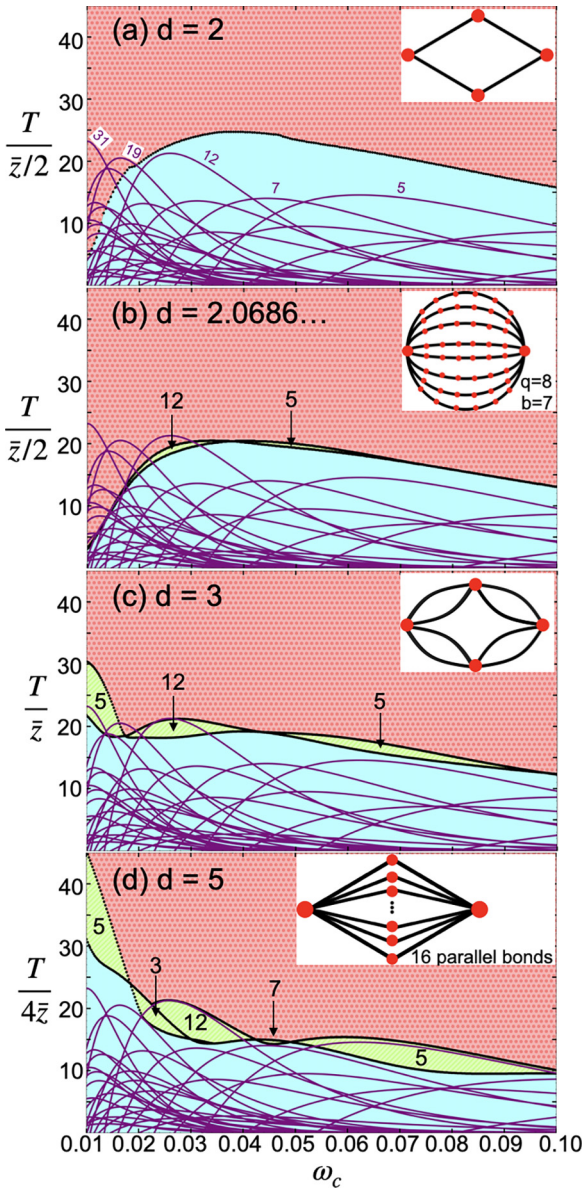

FIG. 6. Phase diagrams of the dissonance Hamiltonian on a variety of hierarchical lattice dimensions (lattice motifs inset top right in each plot). The numbers labeling the intermediate (green) phases indicate the periodicity of the Fourier fixed point in that particular phase. The temperatures here have been scaled by the mean degree of the lattice $\bar{z}$ and an additional power of 2, which normalizes the transitions for all dimensions to around the same scale. The purple lines show the first 31 Fourier coefficients of $D(\theta)$ plotted vs critical bandwidth, showing convergence towards the mean-field result as dimension increases. In (a), the most significant of these mean-field parameters are labeled as in Fig. 3(c).

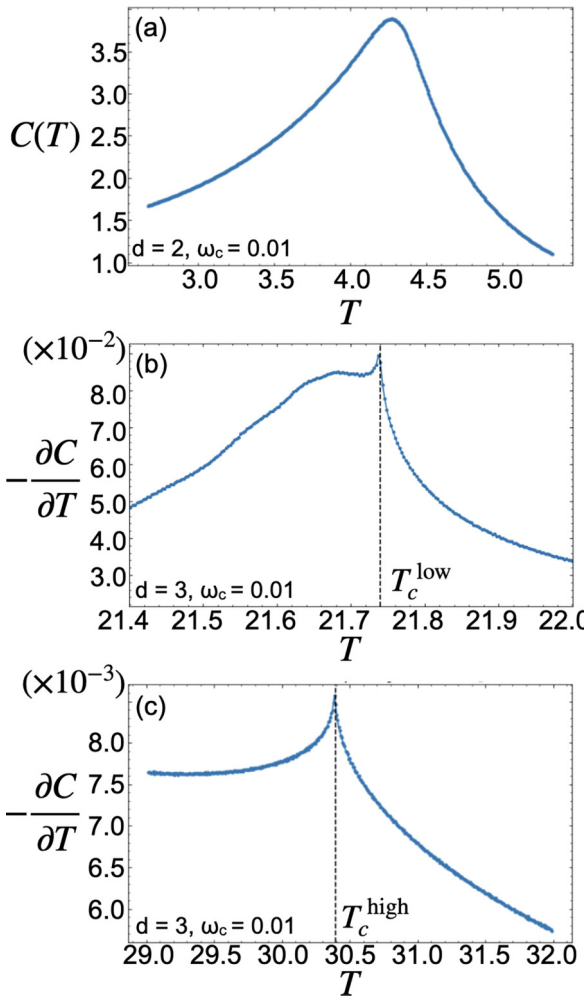

FIG. 7. (a) Heat capacity of a 2D lattice near its transition at $\omega_c = 0.01$. The 2D lattice has no singularity in its free energy, similar to the BKT transition; instead, we identify the transition as the temperature where $C(T)$ is maximal. (b) Temperature derivative of the heat capacity of a 3D lattice near its low-temperature transition at $\omega_c = 0.01$. The dashed line indicates the transition temperature calculated by the RG fixed points, which agrees well with the position of singularities in this quantity. (c) Temperature derivative of the heat capacity of the same 3D lattice near its high-temperature transition at $\omega_c = 0.01$.

compute the free energy as a function of temperature $F(T)$ of a particular hierarchical lattice with a fixed critical bandwidth. Using the free energy, we can take numerical derivatives to determine the entropy $S(T) = -\frac{\partial F}{\partial T}$, heat capacity $C(T) = T\frac{\partial S}{\partial T}$, and its first derivative $\frac{\partial C}{\partial T}$ to verify the existence of a singularity in the free energy, which indicates a phase transition. Plots of the heat capacity for a 2D lattice and derivative $\frac{\partial C}{\partial T}$ for a 3D lattice, both with $\omega_c = 0.01$, are shown in Fig. 7. For $d = 3$ the transition temperatures from the RG phase diagram are denoted with dashed lines, and we notice kinks in $\frac{\partial C}{\partial T}$ that line up with these predicted transition temperatures.

The free energy behavior for $d = 2$ is quite distinct compared to higher dimensions; rather than the sharp kinks seen in the derivative of heat capacity in Fig. 7, there is instead a rounded peak in the heat capacity, shown in the top of Fig. 7, which appears as approaching a jump discontinuity in the derivative. This heat capacity behavior is similar to the BKT transition of the 2D XY model [31,32], which comes as no surprise given that $D(\theta)$ with octave periodicity resembles an XY model with a more complicated periodic structure. The phase boundary constructed in Fig. 6(a) was determined by identifying the peak in heat capacity for each value of the critical bandwidth ω_c.

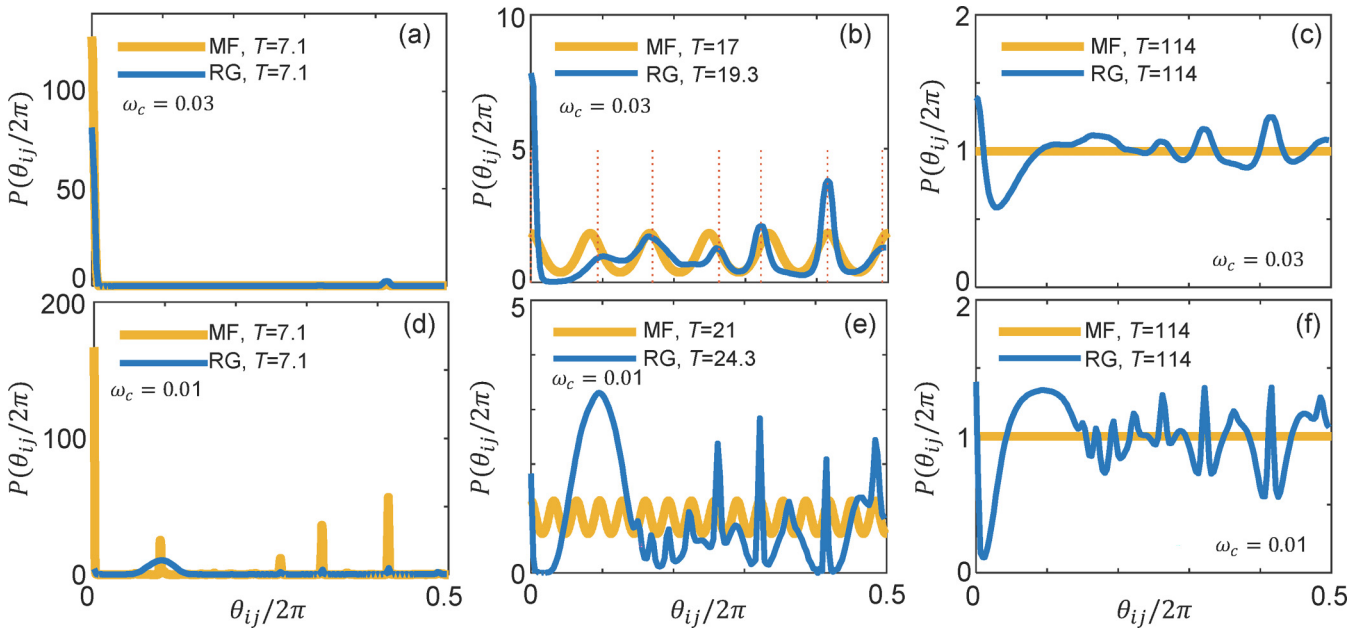

FIG. 8. Nearest-neighbor pitch probability distributions $P(\theta_{ij})$, for $n_{\max} = 10$. (a)–(c) $P(\theta_{ij})$ calculated using the RG (blue) and mean-field (MF - yellow) methods, at increasing temperature in the three different phases at $\omega_c = 0.03$. (d)–(f) The same as (a)–(c), but at $\omega_c = 0.01$. Dotted red lines in (b) show pitches in five-limit just intonation.

V. CORRELATIONS

Observable quantities can be obtained by differentiating the free energy with respect to quantities in the Hamiltonian. Here, we will study the probability distribution of nearest-neighbor pitch differences $P(\theta_{ij})$. We first express the interaction potential $V(\theta_{ij}) = -\beta D(\theta_i - \theta_j)$ in Fourier space $V(\theta_{ij}) = \sum_s g_s e^{is\theta_{ij}}$. The derivative

$$\frac{\partial F}{\partial g_k} = -\frac{1}{\beta \mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial g_k} = -\frac{1}{\beta} \sum_{\{ij\}} \langle e^{ik\theta_{ij}} \rangle \tag{24}$$

brings down a factor of $e^{ik\theta_{ij}}$ from the exponential in $\mathcal{Z}$ and yields the expectation value of that factor, summed over all nearest neighbors $\{ij\}$. In terms of the free energy per bond $\mathcal{F}$,

$$\langle e^{ik\theta_{ij}} \rangle_{\{ij\}} = \frac{\partial \mathcal{F}}{\partial g_k}, \tag{25}$$

where the subscript $\{ij\}$ indicates an average over all bonds.

We can obtain the same quantity,

$$\langle e^{ik\theta_{ij}} \rangle_{\{ij\}} = \int_0^{2\pi} P(\theta_{ij}) e^{ik\theta_{ij}} d\theta_{ij}, \tag{26}$$

from the probability distribution $P(\theta_{ij})$ and hence can obtain that probability distribution using an inverse Fourier series:

$$P(\theta_{ij}) = \sum_k \frac{\partial \mathcal{F}}{\partial g_k} e^{-ik\theta_{ij}}. \tag{27}$$

In practice, we approximate $\frac{\partial \mathcal{F}}{\partial g_k}$ numerically by modifying one Fourier coefficient of $D(\theta_{ij})$ by a small quantity $\delta g_k \ll |g_k|$, calculating $\mathcal{F}$ as above, and taking a finite difference:

$$\frac{\partial \mathcal{F}}{\partial g_k} \approx \frac{\mathcal{F}(g_k + \delta g_k) - \mathcal{F}(g_k - \delta g_k)}{2\delta g_k}. \tag{28}$$

Results of $P(\theta_{ij})$ are shown in Fig. 8 for a 3D lattice at selected values of ω_c in the three different phase regions. [The lattice is the same as shown in Fig. 6(c).] The two rows of Fig. 8 show results for $\omega_c = 0.03$ and 0.01, respectively. The first and third column of plots in Fig. 8 show results at $T = 7.1\bar{z}$ and $T = 114\bar{z}$, respectively, with the temperature in the middle column chosen to be in the intermediate phase at that value of ω_c. In each case, $P(\theta_{ij})$ is shown in blue over the half-octave. For comparison, we also plot in yellow the same quantity as calculated in the mean-field theory of Ref. [1]:

$$P(\Delta\theta) = \int P(\theta) P(\theta - \Delta\theta) d\theta. \tag{29}$$

The differences between the RG and mean-field results highlight the effects of local correlations.

In Figs. 8(a)–8(c) we see the evolution of $P(\theta_{ij})$ from the low-temperature phase, to the 12-fold intermediate phase, to the high-temperature disordered phase. At $T = 7.1\bar{z}$, $P(\theta_{ij})$ for both the RG and mean-field result is strongly peaked at around $\theta_{ij} = 0$ as we approach the expected ground state of uniform pitch. In the intermediate phase [Fig. 8(b)], there is a more pronounced difference between RG and the mean field. The mean-field (MF) result shows regular oscillations with a period of $2\pi/12$, arising from the 12 equal peaks per octave in $P(\theta)$. In the RG result, we can still identify 12 peaks per octave but with unequal height and spacing. This difference is due to local correlations. While on average, pitches may be distributed into 12 equal peaks, nearest neighbors are more likely to favor the more consonant intervals such as unison, octave, and perfect fourth or fifth (frequency ratios of 1, 2, 4/3 or 3/2, respectively). The three peaks with enhanced amplitude relative to the MF result are the unison, perfect fourth, and major third, in decreasing order. (By symmetry, the other half of the octave would see enhanced peaks at the octave, perfect fifth, and minor sixth.) This is seen in music,

which while potentially composed of all 12 pitch classes, favors these consonant intervals for simultaneously sounded tones.

The unequal spacing of peaks in $P(\theta_{ij})$ in Fig. 8(b) reflects the just intonation of chords, where pitches are tuned slightly away from the equally divided octave so as to maximize consonance of pairs of pitches. The dotted red lines indicate the frequency ratios 16/15, 9/8, 6/5, 5/4, 4/3, and 45/32, which correspond to a common just intonation system (so-called five-limit just intonation). The RG peaks are shifted towards these just intonation values relative to the MF result. Just intonation may be used by ensembles of variable-tuning instruments, such as unfretted string instruments or the human voice, where pitches can be adjusted locally to maximize consonance even if the average distribution of pitches is uniformly spaced.

Despite the unequal peaks seen in the nearest-neighbor distribution $P(\theta_{ij})$ in Fig. 8(b), it is plausible that the long-range order still displays 12 equal peaks in the intermediate phase. This can occur because the largest peak at nonzero $\theta_{ij}/2\pi$ occurs at approximately 5/12 (and by symmetry 7/12). Since 5 and 12 are relatively prime, repeated neighbors at this separation yield pitches (mod 1) at all integer multiples of 1/12.

Finally, in the high-temperature region at $T = 114\bar{z}$ [Fig. 8(c)], we see that θ_{ij} is approaching the expected disordered limit of uniformly random $P(\theta_{ij})$.

The results shown in Figs. 8(d)–8(f) at $\omega_c = 0.01$ are qualitatively different than those at $\omega_c = 0.03$. The high-temperature case [Fig. 8(f)] still appears to be converging to the uniform state, but the intermediate phase fixed point showed fivefold symmetry, in conflict with the mean-field prediction of a 31-fold phase. In the intermediate phase [Fig. 8(e)], we see a large broad peak in $P(\theta_{ij})$ emerging at around $\theta_{ij}/2\pi = 0.1$, in addition to narrower peaks corresponding to the expected small-integer frequency ratios. This bears little resemblance to the mean-field result of regular oscillations with a period of $2\pi/31$. Some but not all of the sharp peaks in the RG result line up roughly with maxima of the mean-field result, and the large broad feature is not seen in the mean-field curve. As the temperature is lowered further [Fig. 8(d)], the broad peak in the RG result dominates over the others, with no large spike at zero, suggesting that the ground state will not be the single-pitch state as seen in the mean-field case. Instead, Fig. 8(d) suggests that the ground state will be more like a canted state. This state, which one would not expect to see in the mean-field model, may be an interesting avenue for future study.

The unexpected low-temperature state at $\omega_c = 0.01$ shown in Fig. 8(d) occurs because of a relatively wide dip in $D(\theta_{ij})$ at around $\theta/2\pi = 0.1$, in which no frequency ratios are centered and the surrounding features are narrow enough not to raise the value of D very much (see Fig. 1). If we calculate $D(\theta_{ij})$ using $n_{\max} = 15$ partials (instead of $n_{\max} = 10$ as above), then additional features appear in $D(\theta_{ij})$ in the range around $\theta/2\pi = 0.1$ [see Fig. 9(a)]. The phase diagram for the 3D lattice with $n_{\max} = 15$ is shown in Fig. 9(b). Now a 31-fold fixed point does appear in the approximate range of ω_c predicted by the mean-field model. When the dimension is increased to $d = 5$ [Fig. 9(c)], we again see increased quantitative agreement between the RG and mean-field results, including the

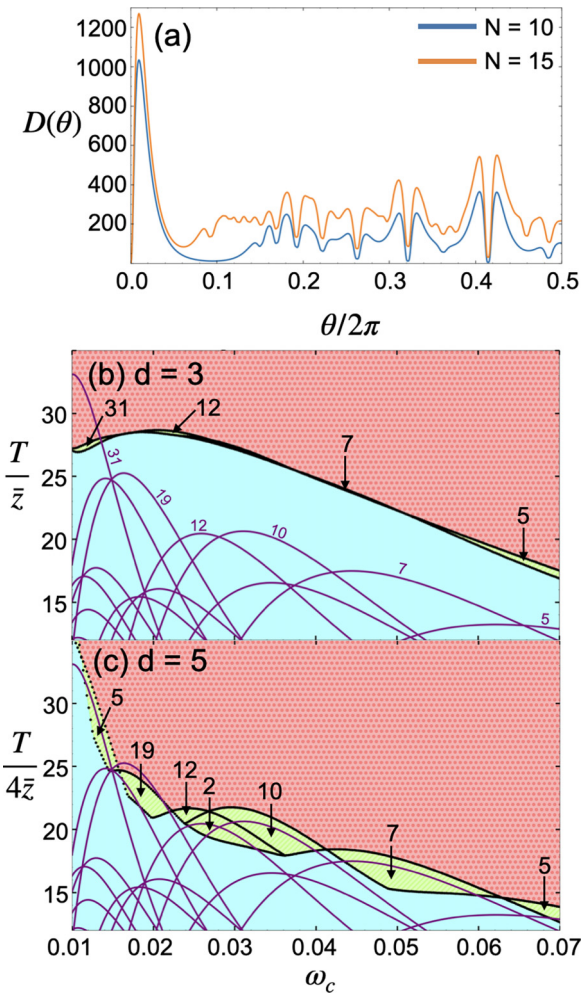

FIG. 9. (a) Dissonance Hamiltonian plotted with $n_{\max} = 10$ and $n_{\max} = 15$. (b), (c) Phase diagrams calculated with $n_{\max} = 15$ for lattices with $d = 3$ (b) and $d = 5$ (c), accompanied by the mean-field order parameters with $n_{\max} = 15$ plotted and labeled in purple.

predicted 19-fold region, although the predicted 31-fold region has now reverted back to the fivefold fixed point.

In Figs. 10(a)–10(c) we show the resulting $P(\theta_{ij})$ as before at $T = 7.1\bar{z}$ and $T = 114\bar{z}$, and within the intermediate phase, at $\omega_c = 0.01$, but now with $n_{\max} = 15$. At the two higher temperatures (b) and (c), we see that the wide peak near $\theta/2\pi = 0.1$ is suppressed as compared to Figs. 8(e) and 8(f). And at low temperature [Fig. 10(a)] we see that the $P(\theta_{ij})$ appears to be approaching the single-pitch ground state, in agreement with the mean-field model. The RG nearest neighbor $P(\theta_{ij})$ in the intermediate phase [Fig. 10(b)] may plausibly lead to the mean-field prediction of long-range order with 31 equally spaced pitches, as the largest nonzero peak is close to 13/31 (and other prominent peaks are close to 8/31 and 10/31). Since 31 is prime, all of these pitch differences will yield pitches (mod 1) at all integer multiples of 1/31.

VI. CONCLUSIONS

By analogy to the free energy of a physical system, ordered phases of sound emerge through a process which minimizes

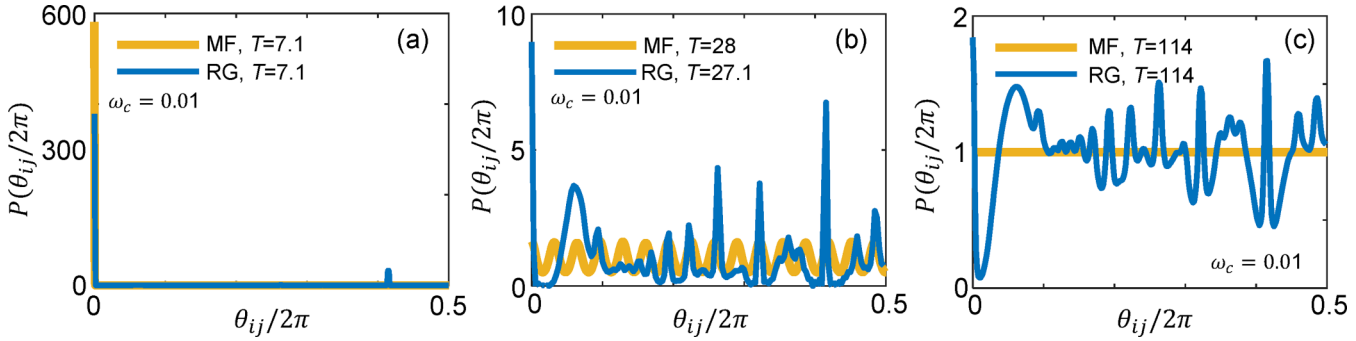

FIG. 10. Nearest-neighbor pitch probability distributions $P(\theta_{ij})$, for $n_{\max} = 15$. (a)–(c) $P(\theta_{ij})$ calculated using the RG (blue) and mean-field (MF - yellow) methods, at increasing temperature in the three different phases at $\omega_c = 0.01$.

the dissonance between tones and maximizes the variety of pitches [1]. Here, we have explored this harmonic ordering on a variety of hierarchical lattice structures using RG theory. By treating dissonance between sounds as an XY-type interaction Hamiltonian, the RG procedure gives rise to three classes of fixed points, corresponding to three different types of phases, depending on the temperature and critical bandwidth of the dissonance Hamiltonian; these three fixed points seem to directly correspond to three distinct phases also seen in the mean-field model, and thus to three systems of intonation.

The periodicities of these fixed points reflect the ways that musicians throughout history and in different cultures have chosen to divide the octave, and clear connections can be made between historical tuning systems and the nearest-neighbor correlations computed here. The 12-fold division of the octave is familiar from Western music but also is used at least approximately in other cultures' music as well. Note that a 12-fold octave division does not imply that all 12 pitches are used equally in a given piece of music. Much of Western music is composed using a seven-note (diatonic) scale, comprised of a subset of the 12 pitches. Or five-note (pentatonic) scales used in, for example, Japanese, Chinese, and Indian music, as well as in the blues, can also be described as a subset of 12 pitch classes per octave. However, tuning systems based (at least approximately) on other octave divisions do exist. Some gamelan music uses a roughly equal fivefold division of the octave [30], whereas some Thai classical music uses a roughly equal sevenfold octave division [33]. Interestingly, both of those musical systems commonly utilize instruments with nonharmonic spectra (xylophones, metallophones, gongs, etc.). As such, the specific RG results shown here may not apply, because they were obtained using an assumption of harmonic partials. But, as is discussed in the Supplemental Material to Ref. [1], such nonharmonic spectra do lead to seven- and fivefold octave divisions in the mean-field approximation (and notably, do not lead to any 12-fold divisions). Studying these spectra using the RG method would be an interesting avenue for future work.

Here we have shown that the 12-fold ordered phase seen in mean-field theory appears in the RG results for any dimension $d > 2$, and the two-phase regions begin to qualitatively agree at around $d = 3$. This demonstrates that this ordered state persists even when the interactions are limited to a relatively small subset of nearest neighbors, as would be the case in music where dissonance is only perceived among tones sounded simultaneously.

The discrepancy between the mean-field and RG results at small values of ω_c may suggest new systems of harmony. Previous experimental microtonal music has explored octave divisions such as 19- and 31-fold seen in the mean-field results at small values of ω_c [3]. The RG results with ten harmonic partials have shown that other orderings can emerge in this range of ω_c that cannot be captured by the mean-field approximation. Moreover, the orderings that emerge depend on the spectrum of partials, which set the timbre, or tone quality, of the sound from a particular instrument. Further studies of these ordered phases of sound using numerical simulation allow for exploration of new systems for making music.

ACKNOWLEDGMENTS

We would like to greatly thank Dr. M. Hinczewski in the CWRU Department of Physics for his assistance with preliminary calculations of the RG model and additional consultation in implementing the methods used.

APPENDIX A: LATTICE DIMENSION FORMULA

Consider the construction of an arbitrary hierarchical lattice wherein a singular bond between two lattice sites is repeatedly replaced by some bond structure containing q parallel branches, each branch with b bonds in series, adding a total of q' new parallel sites. (If multiple parallel bonds connect the same sites, then $q' < q$.) Starting from a singular bond, after n construction steps, the total number of bonds is $B_n = (qb)^n$, and the nth step adds $q'(b-1)B_{n-1}$ new lattice sites. The total number of lattice sites after n steps is given by

$$\begin{aligned} N_n &= 2 + \sum_{k=1}^{n} q'(b-1)B_{n-1} \\ &= 2 + \frac{q'(b-1)((qb)^n - 1)}{qb-1} \\ &\overset{n\to\infty}{\approx} \frac{q'(b-1)}{qb-1}(qb)^n. \end{aligned} \tag{A1}$$

The total number of lattice sites depends on a characteristic length, called the lattice diameter $l_n = b^n$, which is the shortest distance between the two original lattice sites; for a

d-dimensional lattice, the total number of sites at iteration n of the construction should be $N_n \propto l_n^d$. Thus,

$$N_n \propto (qb)^n = q^n b^n = b^{n+n\log_b q} = b^{n(1+\frac{\ln q}{\ln b})} = l_n^{1+\frac{\ln q}{\ln b}} \tag{A2}$$

$$\Rightarrow d = 1 + \frac{\ln q}{\ln b}. \tag{A3}$$

A given lattice can be characterized by its degree and coordination number distribution. We draw a distinction between these two quantities because the lattices we study can have multiple bonds between the same two sites. The degree of a site is defined as the number of bonds attached to that site. The coordination number of a site is defined as the number of other sites to which that site is bonded.

From B_n and N_n, we can obtain the mean degree $\bar{z} = 2B_n/N_n \approx 2\frac{qb-1}{q'(b-1)}$.

For the coordination number distribution, we know that the nth step of lattice construction adds $\nu_0 = q'(b-1)(qb)^{n-1}$ new lattice sites with coordination number $z_0 = 2$. After the nth step, the $\nu_1 = q'(b-1)(qb)^{n-2}$ sites added at the previous step gain q new neighbors for each of their two previous neighbors, for a total $z_1 = 2q$. Iterating, we obtain $\nu_j = q'(b-1)(qb)^{n-j-1}$ sites with coordination number $z_j = 2q^j$. The fraction of sites with coordination number z_j is then $r_j = \nu_j/N_n = (qb-1)/(qb)^{j+1}$.

APPENDIX B: RG DECIMATION FORMULAS

Pitch in this model is treated as periodic, with a period of a one octave pitch interval; this periodicity means the dissonance function behaves like a (more complicated) XY-type Hamiltonian of the form $-\beta\mathcal{H} = V(\theta_i - \theta_j)$, where $\theta_i = 2\pi \log_2(\Phi_i/\Phi_0)$ for an arbitrary fundamental frequency Φ_0. The goal of the RG procedure is to eliminate bonds and lattice sites in a way that preserves the partition function $\mathcal{Z}$; on a hierarchical lattice, this can be done exactly without introducing additional terms to the Hamiltonian [20,23].

1. Series bonds

For a set of three lattice sites, connected by two bonds in a line, the goal is to eliminate the center site θ_2 by integrating over it, leaving in the partition function just a trace over the configurations of θ_1 and θ_3. The partition function of the three sites is given by

$$\mathcal{Z} = \int d\theta_1 d\theta_3 \left[\int d\theta_2 e^{V(\theta_1-\theta_2)+V(\theta_2-\theta_3)} \right]. \tag{B1}$$

It is more convenient to work with Fourier coefficients $f(s)$ by way of a discrete Fourier transform:

$$f(s) = \int \frac{d\theta}{2\pi} e^{is\theta} e^{V(\theta)} \iff e^{V(\theta)} = \sum_s f(s) e^{-is\theta}. \tag{B2}$$

Inserting the Fourier transform of the integrand of $\mathcal{Z}$ gives

$$\mathcal{Z} = \int d\theta_1 d\theta_3 \left[\sum_{s,t} f(s) e^{-is\theta_1} f(t) e^{it\theta_3} \int d\theta_2 e^{i(s-t)\theta_2} \right]. \tag{B3}$$

The integral over θ_2 evaluates to $2\pi\delta_{s,t}$, so this simplifies to

$$\mathcal{Z} = \int d\theta_1 d\theta_3 \left[\sum_s 2\pi f^2(s) e^{-is(\theta_1-\theta_3)} \right]. \tag{B4}$$

The renormalized lattice has just one bond connecting θ_1 and θ_3 with a potential $V'(\theta_1 - \theta_3)$, so this has the partition function

$$\mathcal{Z}' = \int d\theta_1 d\theta_3 [e^{V'(\theta_1-\theta_3)}] = \int d\theta_1 d\theta_3 \left[\sum_s f'(s) e^{-is(\theta_1-\theta_3)} \right]. \tag{B5}$$

Comparing $\mathcal{Z}$ and $\mathcal{Z}'$, we can see that for each term in the sums,

$$f'(s; f) = 2\pi f^2(s). \tag{B6}$$

It is at this point that we implement the constraint that $V'(0) = 0$, so that for each RG iteration, the dissonance for two unison pitches is always zero. This effectively creates a normalization condition for the renormalized coefficients:

$$e^{V'(0)} = \sum_s f'(s; f) \Rightarrow 1 = \sum_s 2\pi f^2(s). \tag{B7}$$

This gives us the final recursion relation for series bonds in Eq. (11):

$$f'(s; f) = \frac{f^2(s)}{\sum_t f^2(t)}. \tag{B8}$$

To calculate the free energy, we need to alter the partition function $\mathcal{Z} \to \tilde{\mathcal{Z}}$ according to Eq. (15); this gives

$$\tilde{\mathcal{Z}} = \int d\theta_1 d\theta_3 \left[\int d\theta_2 e^{V(\theta_1-\theta_2)+V(\theta_2-\theta_3)+2G} \right]. \tag{B9}$$

Following the same steps as before yields the following for $\tilde{\mathcal{Z}}$ before the RG transformation:

$$\tilde{\mathcal{Z}} = e^{2G} \int d\theta_1 d\theta_3 \left[\sum_s 2\pi f^2(s) e^{-is(\theta_1-\theta_3)} \right], \tag{B10}$$

and after the RG transformation:

$$\tilde{\mathcal{Z}}' = e^{G'} \int d\theta_1 d\theta_3 \left[\sum_s f'(s) e^{-is(\theta_1-\theta_3)} \right]. \tag{B11}$$

Equating $\tilde{\mathcal{Z}}$ and $\tilde{\mathcal{Z}}'$ and again enforcing the condition $V'(0) = 0$ gives the recursion relation

$$e^{G'} = 2\pi e^{2G} \sum_s f^2(s), \tag{B12}$$

which yields Eq. (16) after some manipulation.

2. Parallel bonds

For a set of two parallel bonds connecting two lattice sites, the partition function is

$$\mathcal{Z} = \int d\theta_1 d\theta_2 [e^{2V(\theta_1-\theta_2)}] = \int d\theta_1 d\theta_2 [(e^{V(\theta_1-\theta_2)})^2]. \tag{B13}$$

Inserting the Fourier transform of the integrand, the product in the integral becomes a convolution sum:

$$\mathcal{Z} = \int d\theta_1 d\theta_2 \left[\sum_s \left[\sum_t f(t) f(s-t) \right] e^{-is(\theta_1-\theta_2)} \right]. \quad \text{(B14)}$$

After renormalization, the lattice will have just one bond connecting the sites with a potential $V'(\theta_1 - \theta_2)$, giving the renormalized partition function

$$\begin{aligned} \mathcal{Z}' &= \int d\theta_1 d\theta_2 [e^{V'(\theta_1-\theta_2)}] \\ &= \int d\theta_1 d\theta_2 \left[\sum_s f'(s) e^{-is(\theta_1-\theta_2)} \right]. \end{aligned} \quad \text{(B15)}$$

Setting both $\mathcal{Z}$ and $\mathcal{Z}'$ equal, we can see that for each term within the sum over s,

$$f'(s; f) = \sum_t f(t) f(s-t). \quad \text{(B16)}$$

Insisting that $V'(0) = 0$ gives the final relation of Eq. (12).

Introducing the field G for the free energy calculation yields

$$\begin{aligned} \tilde{\mathcal{Z}} &= \int d\theta_1 d\theta_2 [e^{2V(\theta_1-\theta_2)+2G}] \\ &= e^{2G} \int d\theta_1 d\theta_2 [(e^{V(\theta_1-\theta_2)})^2]; \end{aligned} \quad \text{(B17)}$$

$$\tilde{\mathcal{Z}} = e^{2G} \int d\theta_1 d\theta_2 \left[\sum_s \left[\sum_t f(t) f(s-t) \right] e^{-is(\theta_1-\theta_2)} \right]. \quad \text{(B18)}$$

In the renormalized partition function, we have

$$\tilde{\mathcal{Z}}' = e^{G'} \int d\theta_1 d\theta_2 \left[\sum_s f'(s) e^{-is(\theta_1-\theta_2)} \right]. \quad \text{(B19)}$$

Comparison of $\tilde{\mathcal{Z}}$ and $\tilde{\mathcal{Z}}'$ with $V'(0) = 0$ gives the relation

$$e^{G'} = e^{2G} \sum_s \left[\sum_t f(t) f(s-t) \right], \quad \text{(B20)}$$

which simplifies to give Eq. (17). Analytically, the double sum is equal to 1 when the sum is taken over an infinite number of s, t values, but it is necessary to include this factor when performing numeric computations, since the sums must be truncated at some finite value of s, t.

[1] J. Berezovsky, Sci. Adv. **5**, eaav8490 (2019).
[2] H. Din and J. Berezovsky, PLoS ONE **18**, e0280227 (2023).
[3] D. Keislar, Comput. Music J. **11**, 18 (1987).
[4] H. v. Helmholtz, *On the Sensations of Tone* (Courier Corporation, North Chelmsford, MA, 2013).
[5] R. Plomp and W. J. M. Levelt, J. Acoust. Soc. Am. **38**, 548 (1965).
[6] W. A. Sethares, J. Acoust. Soc. Am. **94**, 1218 (1993).
[7] R. Marjieh, P. M. Harrison, H. Lee, F. Deligiannaki, and N. Jacoby, Nat. Commun. **15**, 1482 (2024).
[8] J. V. Tobias, *Foundations of Modern Auditory Theory* (Academic Press, New York, 1970), Vol. 1, pp. 157–202.
[9] E. Zwicker and E. Terhardt, J. Acoust. Soc. Am. **68**, 1523 (1980).
[10] W. A. Sethares, *Tuning, Timbre, Spectrum, Scale* (Springer, New York, 2005).
[11] A. N. Berker and D. R. Nelson, Phys. Rev. B **19**, 2488 (1979).
[12] M. Kaufman and R. B. Griffiths, Phys. Rev. B **24**, 496 (1981).
[13] M. Kaufman and R. B. Griffiths, Phys. Rev. B **30**, 244 (1984).
[14] R. B. Griffiths and M. Kaufman, Phys. Rev. B **26**, 5022 (1982).
[15] M. Hinczewski and A. N. Berker, Phys. Rev. E **73**, 066126 (2006).
[16] M. Hinczewski, Phys. Rev. E **75**, 061104 (2007).
[17] S. R. McKay, A. N. Berker, and S. Kirkpatrick, Phys. Rev. Lett. **48**, 767 (1982).
[18] A. N. Berker and S. R. McKay, J. Stat. Phys. **36**, 787 (1984).
[19] L. P. Kadanoff, Ann. Phys. **100**, 359 (1976).
[20] T. W. Burkhardt, in *Real-Space Renormalization*, edited by T. W. Burkhardt and J. M. J. van Leeuwen (Springer, Berlin, Heidelberg, 1982), pp. 33–56.
[21] F. Lerdahl and R. Jackendoff, Music Percept.: Interdiscip. J. **1**, 229 (1983).
[22] A. Forte and S. E. Gilbert, *Introduction to Schenkerian Analysis* (Norton, New York, 1982).
[23] A. N. Berker and S. Ostlund, J. Phys. C **12**, 4961 (1979).
[24] M. E. Fisher, Rev. Mod. Phys. **70**, 653 (1998).
[25] A. Erbaş, A. Tuncer, B. Yücesoy, and A. N. Berker, Phys. Rev. E **72**, 026129 (2005).
[26] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
[27] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).
[28] Y. Gefen, B. B. Mandelbrot, and A. Aharony, Phys. Rev. Lett. **45**, 855 (1980).
[29] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).
[30] E. C. Carterette and R. A. Kendall, Leonardo Music J. **4**, 59 (1994).
[31] J. E. Van Himbergen and S. Chakravarty, Phys. Rev. B **23**, 359 (1981).
[32] P. H. Nguyen and M. Boninsegni, Appl. Sci. **11**, 4931 (2021).
[33] J. Rahn, Music Sci. **2** (2019).