Window Universality for Carrier-Wave Estimators of CUE Characteristic Polynomials

Liz Lemma Future Detective

January 17, 2026

Abstract

Carrier waves—slowly varying multiplicative envelopes governing typical large values—are a central heuristic in the modern understanding of $\zeta(\frac{1}{2}+it)$ beyond the computational range. A persistent obstacle is non-uniqueness: local scale factors depend on the window used to separate 'near zeros' from 'far zeros'. We resolve this ambiguity in the canonical random matrix model. For Haar-random $U \in \mathrm{U}(N)$ we define (a) a local Hadamard-product carrier-wave estimator A_K obtained by removing the contribution of $2K$ nearby eigenangles from $\log|Z_U|$, and (b) a density-wave estimator D_J built from weighted symmetric neighbor discrepancies. We prove a window universality theorem: for any K, J in a fixed constant band around $\log N$, both estimators agree uniformly (in sup norm over N midpoints) up to $o_{\mathbb{P}}(1)$. Thus the carrier wave becomes an asymptotically canonical object once the window contains about $\log N$ zeros. The result formalizes and stabilizes the carrier-wave viewpoint advocated in the source material, and it provides a robust observable for future comparisons with ζ and for next-generation simulation of zeta-like landscapes.

Table of Contents

4. 4. Kernel representations: rewrite both estimators as mesoscopic linear statistics of the centered counting function $S_U(\theta)$ (or equivalently linear statistics of eigenangles) against explicit kernels; isolate deterministic renormalizations.

5. 5. Main theorem: window universality and estimator equivalence. Precise uniformity in (K, J), choice of probability metric, and discussion of strengthening/weakenings (e.g. L^2 vs sup, continuous θ vs midpoints).

6. 6. Concentration and rigidity inputs: probabilistic estimates for CUE linear statistics needed for uniform control; discretization/net arguments for sup over n; number variance bounds at mesoscopic scales.

7. 7. Comparing kernels across windows: operator-norm bounds (Fourier multipliers) showing kernels for different $K, J \asymp \log N$ differ by $o(1)$ in the relevant norm; convert to probabilistic $o_{\mathbb{P}}(1)$ for the associated linear statistics.

8. 8. Numerical illustrations and sanity checks (optional but recommended): demonstrate estimator stability at moderate N; show the log-scale carrier wave is visible; illustrate convergence rates and finite-N effects. Flag that any large-scale experiments are supportive but not used in proofs.

9. 9. Discussion and extensions: (i) toward a canonical limiting field (GFF/GMC), (ii) conditional laws given large carrier waves, (iii) transfer heuristics to ζ and arithmetic damping, (iv) variants for CβE.

1 Introduction and motivation

The modulus of a characteristic polynomial of a random unitary matrix exhibits two competing features. On the one hand, the function $\theta \mapsto \log|Z_U(\theta)|$ is strongly constrained by its zeros at the eigenangles $\{\theta_j\}$: it has logarithmic singularities whose locations are random at the microscopic scale $1/N$. On the other hand, once the singular contributions of nearby zeros are removed, the remaining fluctuations vary on a mesoscopic scale and behave, in several respects, like a log-correlated random field. The purpose of this work is to make precise, in the CUE setting, the emergence and the canonical nature of this mesoscopic "carrier wave".

The guiding heuristic is the elementary identity

$$\log|Z_U(\theta)| = \sum_{j=1}^{N} \log\left|e^{i\theta} - e^{i\theta_j}\right| + (\text{deterministic normalization}),$$

which already indicates that $\log|Z_U|$ is a linear statistic of the eigenangles against the logarithmic kernel. If θ is close to some θ_j, the corresponding summand dominates and produces a local spike. If we instead separate the sum into a near part and a far part relative to θ, then the near part is responsible for the sharp local geometry (including the singularities), whereas the far part varies more smoothly with θ and encodes the cumulative effect of many distant points. It is this far part that we refer to as the carrier wave.

At a purely algebraic level, one can attempt to isolate the carrier wave by dividing the characteristic polynomial by a local product of factors corresponding to eigenvalues near $e^{i\theta}$. Concretely, if we remove the $2K$ nearest eigenvalues around a point θ, then

$$\log|Z_U(\theta)| - \sum_{\substack{j:\,\theta_j \text{ among} \\ K \text{ nearest on each side}}} \log\left|e^{i\theta} - e^{i\theta_j}\right|$$

should depend primarily on eigenvalues at distances $\gg K/N$ from θ and should therefore change slowly as θ varies by a few mean spacings. This operation is canonical only after we decide what "near" means, i.e. after we choose the truncation parameter K. For finite N the resulting field depends visibly on K, and this window dependence is the first obstruction to interpreting the carrier wave as an intrinsic object.

The issue is not only one of aesthetics. Many questions of interest concern extreme values and near-extreme geometry of $\log|Z_U(\theta)|$, or of the zeta function in analogous number-theoretic models. In such problems, a small perturbation of the underlying field can, a priori, change where the maxima occur, how high they are at finite N, and what the correct centering and scaling should be. If the carrier wave is to be used as a stable "background" for describing near maxima, then it must be essentially independent of the

arbitrary parameters introduced by its definition. This motivates the window universality problem: determine a regime of window sizes K in which all reasonable definitions of the carrier wave agree asymptotically.

Our main results show that for CUE this non-uniqueness disappears precisely in a logarithmic window. More specifically, if K varies within a band of order $\log N$, then the corresponding carrier-wave estimators become asymptotically indistinguishable, uniformly over a natural discrete set of evaluation points. In addition, a second, ostensibly different, estimator based on neighbor discrepancies (a "density-wave" formula) is shown to agree with the local-product definition, again uniformly. Taken together, these statements provide a canonical carrier-wave field on the circle at mesoscopic resolution: one may compute it by removing nearby zeros (a local Hadamard-product viewpoint) or by aggregating discrepancies of eigenangle spacings (a counting-function viewpoint), and the outputs agree up to an error vanishing in probability in the supremum norm.

The choice of a logarithmic window is not ad hoc; it is forced by the structure of the logarithmic kernel and by the known behavior of CUE linear statistics. In Fourier variables on the circle, the centered field $\log|Z_U(\theta)|$ has a representation whose dominant contributions come from Fourier modes with weights comparable to $1/|m|$. The variance accumulated from modes $1 \le |m| \le M$ is therefore of size $\sum_{m \le M} 1/m \sim \log M$. A cutoff that changes M by a multiplicative factor changes this variance by an additive constant, which is precisely the scale at which fluctuations of a log-correlated field are sensitive. The local removal of K neighboring zeros corresponds, after suitable manipulations, to a smoothing (or truncation) of the logarithmic kernel at a mesoscopic frequency scale that depends on K. When K is too small, the "near" part is incomplete: microscopic singular behavior leaks into the remainder, and different K lead to different residual spikes. When K is too large, one removes not only the singularities but also a substantial amount of genuinely mesoscopic fluctuation, and the remainder is overly smoothed. The logarithmic regime $K \asymp \log N$ is the balance point where the effect of changing the cutoff becomes negligible compared with the intrinsic fluctuations of the far field.

This perspective parallels heuristics developed for the Riemann zeta function, where $\log|\zeta(1/2+it)|$ is modeled by a log-correlated random field and the local behavior is influenced by nearby zeros. In Farmer's terminology, the zeta function is described as the product of a "carrier wave" and a "modulating wave," with the latter capturing rapid oscillations induced by nearby zeros and the former representing a slowly varying envelope determined by more distant zeros. The CUE characteristic polynomial provides a setting where the same conceptual decomposition can be formulated precisely and where the probabilistic tools needed to control the dependence on window parameters are available. Our results may therefore be interpreted as a rigorous version, in the CUE model, of the principle that the carrier wave is an

emergent mesoscopic object, not an artifact of a particular truncation rule.

A second motivation for the discrepancy-based estimator is that it is expressed directly in terms of eigenangle spacings and the centered counting function. The centered counting function $S_U(\theta)$ is, in many respects, the fundamental object controlling fluctuations of eigenangles across scales: it is a distribution function for the point process with its deterministic linear trend removed. Because S_U is well understood for CUE (through its determinantal structure), it is natural to express mesoscopic observables as smoothed functionals of S_U. The discrepancy sum D_J does exactly this: it aggregates deviations of symmetric eigenangle spans from their mean, with a $1/j$ weight corresponding to the logarithmic kernel. From this viewpoint, the equivalence between the local-product estimator and the discrepancy estimator is not surprising; it asserts that two distinct ways of regularizing the same logarithmic functional lead to the same limit in the logarithmic window.

The requirement of uniformity over a set of N evaluation points is also essential. Pointwise in θ, it is comparatively easy to show that different truncations lead to small differences, since one may allow exceptional events depending on θ. However, the carrier wave is intended to be a global object on the circle, and applications (for instance, to locating high points) demand control in a strong topology. We therefore work with a supremum norm over a grid of N points at typical spacing $1/N$, which is the natural scale at which the underlying point process is sampled. The correct choice of grid is constrained by the singularities of $\log|Z_U|$ at eigenangles. A deterministic grid may come arbitrarily close to an eigenangle with non-negligible probability, which would introduce sporadic but large spikes and destroy uniform bounds. The canonical remedy is to use a grid built from the eigenangles themselves but shifted away from them, so that one evaluates the field at points that are provably separated from the zeros at the correct scale. This is the role of the midpoint grid: it provides N points distributed around the circle with spacing comparable to $1/N$ and, with high probability, uniformly bounded away from the eigenangles.

At a technical level, the theorems rest on rewriting both estimators as centered linear statistics of the same underlying determinantal process against explicit kernels that approximate the logarithmic kernel. The window parameter enters only through the kernel. Window universality is then reduced to a deterministic comparison of these kernels in a norm that controls variances of linear statistics (equivalently, a Fourier-multiplier norm adapted to the sine-kernel covariance). Once kernel differences are shown to be small in this norm throughout the logarithmic band, standard concentration estimates for determinantal processes upgrade the kernel comparison to probabilistic control of the corresponding linear statistics. Finally, a net argument over the midpoint grid yields uniformity in n.

The conclusion is that, in the CUE model, one may unambiguously speak

of *the* carrier wave at mesoscopic resolution: any choice of K in a fixed logarithmic band produces the same field on the midpoints up to an error $o_{\mathbb{P}}(1)$ in supremum norm, and the same field is recovered by the discrepancy estimator with J in the same band. In particular, the dependence on the truncation rule is a finite-N artifact that disappears in the limit. This gives a rigorous counterpart to the informal decomposition of characteristic polynomials (and, by analogy, of zeta) into a slowly varying envelope and a rapidly varying local factor determined by nearby zeros.

In the next section we collect the basic background on the CUE eigenangle process, its determinantal structure, and the behavior of the centered counting function. These facts will be used repeatedly to control linear statistics at the mesoscopic scales relevant to the logarithmic window, and to justify the uniform midpoint-based sampling that underlies the supremum norm statements.

2 Background on CUE, characteristic polynomials, and counting functions

2.1 Haar unitary matrices and the eigenangle point process

Let $U \sim \mathrm{Haar}(\mathrm{U}(N))$. The eigenvalues of U lie on the unit circle and can be written uniquely as $\lambda_j = e^{i\theta_j}$ with $0 \le \theta_1 < \cdots < \theta_N < 2\pi$. We extend the indexing periodically by $\theta_{k+N} = \theta_k + 2\pi$, so that local statements near 0 and 2π can be formulated without boundary conventions.

The joint density of the ordered eigenangles with respect to Lebesgue measure on $[0, 2\pi)^N$ is proportional to the squared Vandermonde determinant,

$$\frac{1}{(2\pi)^N N!} \prod_{1 \le j < k \le N} |e^{i\theta_j} - e^{i\theta_k}|^2,$$

which makes the eigenangles a repulsive point process on $\mathbb{S}^1$. This density is invariant under global rotation $\theta_j \mapsto \theta_j + \alpha \pmod{2\pi}$, and in particular the one-point intensity is constant: the mean density is $N/(2\pi)$ and a typical gap has size $2\pi/N$.

We view $\{\theta_j\}_{j=1}^N$ as a random counting measure

$$\mu_U := \sum_{j=1}^{N} \delta_{\theta_j}$$

on $\mathbb{S}^1$, and we write $\#I := \mu_U(I)$ for the number of eigenangles in an arc $I \subset \mathbb{S}^1$.

2.2 Determinantal structure and correlation kernel

A fundamental fact is that the CUE eigenangles form a determinantal point process with an explicit projection kernel. Concretely, for $k \geq 1$ the k-point correlation functions satisfy

$$\rho_k(\theta_1, \dots, \theta_k) = \det\big(K_N(\theta_a, \theta_b)\big)_{1 \leq a,b \leq k},$$

where K_N is the reproducing kernel of the span of $\{e^{im\theta} : 0 \leq m \leq N-1\}$ in $L^2(\mathbb{S}^1, d\theta/(2\pi))$, namely

$$K_N(\theta, \phi) = \frac{1}{2\pi} \sum_{m=0}^{N-1} e^{im(\theta-\phi)} = \frac{1}{2\pi} e^{i\frac{N-1}{2}(\theta-\phi)} \frac{\sin\big(\frac{N}{2}(\theta-\phi)\big)}{\sin\big(\frac{1}{2}(\theta-\phi)\big)}.$$

This is a rank-N orthogonal projection kernel, and its determinantal nature yields exact formulas for moments of linear statistics and, more importantly for us, sharp variance and concentration bounds.

In the microscopic scaling $\theta - \phi \sim 1/N$, the oscillatory phase factor is inessential and the rescaled kernel converges to the sine kernel. This is the source of the universal local behavior, including level repulsion and logarithmic growth of number variance.

2.3 Counting functions, centered fluctuations, and linear statistics

For $\theta \in \mathbb{R}$ we consider the counting function

$$\mathcal{N}(\theta) := \#\{j : \theta_j \leq \theta\},$$

interpreted using the periodic extension of $\{\theta_j\}$, and the centered counting function

$$S_U(\theta) := \mathcal{N}(\theta) - \frac{N\theta}{2\pi},$$

with a fixed convention for θ outside $[0, 2\pi)$ (any consistent periodic lift suffices for our purposes). The function S_U is piecewise constant with unit jumps at the eigenangles and has mean approximately zero at each fixed θ by rotation invariance.

A convenient way to package linear statistics is through Stieltjes-type integrals against dS_U. For a 2π-periodic function f of bounded variation we set

$$\int_0^{2\pi} f(\phi)\, dS_U(\phi) := \sum_{j=1}^{N} f(\theta_j) - \frac{N}{2\pi} \int_0^{2\pi} f(\phi)\, d\phi,$$

so that the centered statistic is exactly the pairing of f with the fluctuation measure $\mu_U - (N/(2\pi))d\phi$. In particular, if $f = \mathbf{1}_{[0,\theta]}$ (interpreted suitably as a periodic function with one jump), then $\int f\, dS_U = S_U(\theta)$.

For CUE, the variance of such centered linear statistics can be expressed explicitly in Fourier variables. Writing

$$\widehat{f}(m) := \frac{1}{2\pi}\int_0^{2\pi} f(\phi)e^{-im\phi}\,d\phi, \qquad m \in \mathbb{Z},$$

one has, for sufficiently regular f,

$$\mathrm{Var}\left(\sum_{j=1}^{N} f(\theta_j)\right) = \sum_{m\in\mathbb{Z}\setminus\{0\}} \min\{N,|m|\}\,|\widehat{f}(m)|^2. \tag{1}$$

This identity reflects the fact that the CUE process is a projection DPP and that the Fourier modes diagonalize the covariance. The right-hand side defines a natural Hilbertian seminorm controlling fluctuations. In the mesoscopic regimes relevant to us, it is convenient to compare (1) to an $H^{1/2}$-type quantity

$$\|f\|_{H^{1/2}}^2 := \sum_{m\neq 0} |m|\,|\widehat{f}(m)|^2,$$

keeping in mind that the truncation $\min\{N,|m|\}$ enforces a high-frequency cutoff at $|m| \approx N$.

The number variance is the special case $f = \mathbf{1}_I$ for an arc I of length $|I|$. One then obtains the familiar logarithmic growth: for arcs whose length is not microscopic (say $|I| \gg 1/N$),

$$\mathrm{Var}(\#I) = \frac{1}{\pi^2}\log\big(N|I|\big) + O(1),$$

with the understanding that the precise form involves $\sin(|I|/2)$ rather than $|I|$ when $|I|$ is not small. This logarithmic number variance is the main quantitative manifestation of long-range correlations in the process; it is also the reason logarithmic window sizes arise naturally when one seeks a balance between microscopic singularities and mesoscopic fluctuations.

Beyond second moments, the determinantal structure yields concentration inequalities for linear statistics. We will use the following qualitative principle: if a family of test functions f_N has uniformly bounded variance proxy of the form (1) (or, more generally, bounded in a Hilbert norm adapted to the sine-kernel covariance), then the centered statistics $\int f_N(\phi)\,dS_U(\phi)$ are tight and enjoy subgaussian-type tails uniformly in N. This can be proved by standard determinantal concentration estimates (for instance, via cumulant bounds or via general DPP inequalities for Lipschitz observables), and it is the mechanism by which we later upgrade kernel comparisons to probabilistic $o(1)$ bounds.

2.4 Rigidity estimates at mesoscopic resolution

We will also appeal, implicitly, to the fact that $S_U(\theta)$ does not fluctuate on the scale of a power of N; its typical size is logarithmic. One convenient formulation is that for any fixed $\varepsilon > 0$,

$$\sup_{\theta\in[0,2\pi]} |S_U(\theta)| \le N^{\varepsilon}$$

with probability $1 - o(1)$, and more refined statements give $\sup_\theta |S_U(\theta)| = O_{\mathbb{P}}(\log N)$. Such bounds can be obtained by combining the logarithmic number variance with a chaining or dyadic net argument over θ, using that S_U only changes at eigenangles. These rigidity estimates are compatible with, and in a sense equivalent to, the mesoscopic concentration of linear statistics discussed above. They will enter only through uniformity requirements: when we compare two observables defined by convolving dS_U against nearby kernels, we need to ensure that atypical fluctuations of S_U do not destroy the supremum bounds.

2.5 Characteristic polynomials and logarithmic kernels

The characteristic polynomial is

$$\Lambda_U(z) = \det(I - zU) = \prod_{j=1}^{N}(1 - ze^{i\theta_j}),$$

and, on the unit circle $z = e^{i\theta}$, its modulus is

$$\log|\Lambda_U(e^{i\theta})| = \sum_{j=1}^{N} \log\bigl|1 - e^{i(\theta+\theta_j)}\bigr| = \sum_{j=1}^{N} \log\bigl|e^{i\theta} - e^{i\theta_j}\bigr|.$$

The rotated characteristic polynomial $Z_U(\theta)$ differs from $\Lambda_U(e^{i\theta})$ by a unimodular factor chosen so that $Z_U(\theta) \in \mathbb{R}$ for real θ. In particular,

$$\log|Z_U(\theta)| = \log|\Lambda_U(e^{i\theta})|$$

and the modulus depends only on the eigenangles.

The logarithmic kernel

$$\theta \mapsto \log|e^{i\theta} - e^{i\phi}|$$

has a well-known Fourier expansion (as a distribution),

$$\log\bigl|e^{i\theta} - e^{i\phi}\bigr| = -\sum_{m=1}^{\infty} \frac{1}{m}\cos\bigl(m(\theta - \phi)\bigr) + (\text{constant}),$$

which makes transparent the $1/m$ Fourier weights characteristic of log-correlated fields. Formally inserting this expansion and interchanging summations yields a representation of $\log|Z_U(\theta)|$ as a sum of traces $\mathrm{tr}(U^m)$ with coefficients $1/m$, and hence as a centered linear statistic at the level of Fourier modes. We do not need this identity in its most precise form at this stage; what matters is the structural consequence: any regularization of $\log|Z_U|$ amounts to modifying the logarithmic kernel at short distances (equivalently, truncating or smoothing its Fourier series at high frequencies), and the effect of such a modification can be quantified in the same Hilbertian norms that control variances of linear statistics.

2.6 Midpoints and the avoidance of singularities

A basic obstruction to uniform control of $\log|Z_U(\theta)|$ is that it vanishes at the eigenangles: $Z_U(\theta_j) = 0$, so $\log|Z_U(\theta)| = -\infty$ at $\theta = \theta_j$ and has logarithmic spikes near those points. Consequently, if one were to sample $\log|Z_U|$ on a deterministic grid $\{\vartheta_n\}$ of N points with spacing $\asymp 1/N$, then with non-negligible probability one of the ϑ_n would fall extremely close to an eigenangle, producing a large negative outlier and preventing any meaningful supremum control.

We therefore adopt an intrinsic random grid obtained from the eigenangles themselves, but shifted away from the zeros. For $n \in \{1, \dots, N\}$ we define the midpoint

$$m_n := \frac{\theta_n + \theta_{n+1}}{2},$$

with $\theta_{N+1} = \theta_1 + 2\pi$. By construction, each m_n lies strictly between two consecutive eigenangles, hence $Z_U(m_n) \neq 0$ almost surely. Moreover, the midpoints inherit the correct global density: they form a set of N points on $\mathbb{S}^1$ with typical spacing comparable to $1/N$, and they are canonical in the sense that no additional choices are made.

The distance from m_n to the nearest eigenangle is exactly half the adjacent gap,

$$\min_{1\le j\le N} |m_n - \theta_j| = \frac{\theta_{n+1} - \theta_n}{2}.$$

Since CUE exhibits level repulsion, extremely small gaps are rare. In particular, one may fix any $A > 2$ and use standard small-gap estimates (ultimately traceable to the s^2 repulsion in the sine-kernel scaling limit) to deduce

$$\mathbb{P}\left(\min_{1\le n\le N} (\theta_{n+1} - \theta_n) \le N^{-A} \right) = o(1),$$

so that with high probability all midpoints stay at distance at least N^{-A} from the spectrum. This is far weaker than a macroscopic separation of order $1/N$, but it suffices for uniform control of regularized logarithmic expressions: the

singular contribution of a point at distance $\geq N^{-A}$ is at worst $O(\log N)$, and in our later constructions the potentially large contributions from the closest eigenangles are explicitly removed or compensated.

The midpoint grid also interacts well with sign changes of Z_U. Since Z_U is real-valued and has simple zeros at $\{\theta_j\}$ almost surely, it alternates sign between successive eigenangles. Evaluating at m_n thus produces a well-defined, nonzero value in each nodal interval, avoiding issues of ambiguous phase that can arise when working directly with $\Lambda_U(e^{i\theta})$.

In summary, the midpoint grid provides a natural discretization of the circle at the microscopic scale, while avoiding the singularities of $\log|Z_U|$ in a way compatible with uniform-in-n estimates. All of our sup-norm statements will be formulated on this grid. In the next section we define the carrier-wave estimators that regularize $\log|Z_U|$ by subtracting local contributions of nearby eigenangles, and we introduce an alternative discrepancy-based estimator expressed directly in terms of the eigenangle spacings and the centered counting function.

3 Carrier-wave estimators on the midpoint grid

We now introduce two concrete observables on the midpoint grid $\{m_n\}_{n=1}^N$ which are designed to extract, from the singular field $\theta \mapsto \log|Z_U(\theta)|$, a regularized "carrier-wave" component that is stable under mesoscopic choices of window size. The first estimator is defined by explicitly subtracting the logarithmic contributions of the $2K$ eigenangles nearest to m_n; the second is expressed directly in terms of symmetric spacings across m_n and may be viewed as a discrete logarithmic transform of local density discrepancies. Both are measurable functions of the eigenangle configuration and will later be rewritten as linear statistics of the centered counting function S_U against explicit kernels.

3.1 The local Hadamard-product estimator A_K

Fix an integer $K \geq 1$. For each midpoint m_n we consider the decomposition

$$\log|Z_U(m_n)| = \sum_{j=1}^{N} \log\bigl|e^{im_n} - e^{i\theta_j}\bigr|, \tag{2}$$

where each summand has the explicit real form

$$\log\bigl|e^{im_n} - e^{i\theta}\bigr| = \log\Bigl(2\Bigl|\sin\Bigl(\frac{m_n-\theta}{2}\Bigr)\Bigr|\Bigr). \tag{3}$$

The singularity at $\theta = m_n$ does not occur since $m_n \notin \{\theta_j\}$ almost surely, but the terms corresponding to eigenangles close to m_n are large negative when the adjacent gaps are small. In order to isolate a smoother remainder, we

factor out the nearest zeros in the spirit of a truncated Hadamard product. Concretely, we remove the contributions of the K nearest eigenangles on each side of m_n, namely

$$\theta_n, \theta_{n-1}, \dots, \theta_{n-K+1} \qquad \text{and} \qquad \theta_{n+1}, \theta_{n+2}, \dots, \theta_{n+K},$$

using the periodic extension of the indexing. This yields the definition

$$\log A_K(m_n) := \log |Z_U(m_n)| - \sum_{j=0}^{K-1} \Big(\log |e^{im_n} - e^{i\theta_{n-j}}| + \log |e^{im_n} - e^{i\theta_{n+1+j}}| \Big) + C_{K,N}. \tag{4}$$

We emphasize two elementary points.

(i) Positivity and cancellation of near-singular terms. Since $A_K(m_n)$ is defined through its logarithm, it is immediate from (4) that $A_K(m_n) > 0$. The definition is arranged so that the most singular logarithmic contributions near m_n are subtracted explicitly. In particular, when the gap $\theta_{n+1} - \theta_n$ is unusually small, both $\log |Z_U(m_n)|$ and the two terms $j = 0$ in the subtraction are of order $-\log(\theta_{n+1} - \theta_n)$, and these large negative pieces cancel. Thus $A_K(m_n)$ remains well-behaved even on atypical configurations where midpoints are only polynomially separated from the spectrum.

(ii) Local symmetry. The symmetric choice of K points on each side of m_n is not essential for measurability, but it is convenient for later kernel comparisons. It ensures that, after subtraction, the remaining contribution is dominated by eigenangles at distances $\gtrsim K/N$ from m_n on both sides, and it avoids introducing an artificial drift coming from one-sided truncation.

3.2 Normalization and stationarity

The quantity in (4) is defined up to the additive constant $C_{K,N}$, which we choose deterministically to fix a centering convention. Rotation invariance implies that for each fixed K and n the distribution of the centered expression

$$\log |Z_U(m_n)| - \sum_{j=0}^{K-1} \Big(\log |e^{im_n} - e^{i\theta_{n-j}}| + \log |e^{im_n} - e^{i\theta_{n+1+j}}| \Big)$$

does not depend on n (indeed, the joint law of the eigenangle configuration, and hence of the entire midpoint grid, is invariant under global shifts). Consequently, $\mathbb{E}[\log A_K(m_n)]$ is independent of n once $C_{K,N}$ is fixed. We therefore impose the normalization

$$\mathbb{E}[\log A_K(m_1)] = 0, \tag{5}$$

which uniquely determines $C_{K,N}$ as a real number depending only on (K, N). For our purposes we do not require an explicit closed form for $C_{K,N}$; it suffices that it is deterministic and that the centering is compatible across n. In particular, the field $n \mapsto \log A_K(m_n)$ is stationary in distribution on the discrete circle $\mathbb{Z}/N\mathbb{Z}$.

3.3 Measurability considerations

Both estimators we consider are functions of the eigenangle process, hence of U as a random matrix. We record this explicitly since later we will apply concentration inequalities for determinantal linear statistics.

First, $\log |Z_U(m_n)|$ depends on U only through its eigenangles, by (2). Moreover, m_n is a measurable function of the ordered eigenangles. The subtraction terms in (4) are likewise measurable functions of $(\theta_1, \dots, \theta_N)$. Since $C_{K,N}$ is deterministic, we conclude that $\log A_K(m_n)$ is measurable with respect to the σ-field generated by the eigenangles (equivalently, by the point process μ_U).

We note also that all expressions are well-defined without ambiguity from the periodic lift. Indeed, while the function $\theta \mapsto \log |e^{im_n} - e^{i\theta}|$ is 2π-periodic in θ, the midpoint m_n is defined using the periodic extension $\theta_{N+1} = \theta_1 + 2\pi$, and the indices θ_{n-j} and θ_{n+1+j} are interpreted via $\theta_{k+N} = \theta_k + 2\pi$. This convention ensures that the ordering relative to m_n is consistent even when n is near 1 or N.

3.4 The discrepancy-based density-wave estimator D_J

We now introduce an alternative estimator which does not involve Z_U explicitly. Fix an integer $J \geq 1$. For each midpoint m_n and each $j \geq 1$, we consider the symmetric span across m_n from θ_{n-j} to θ_{n+1+j}. At mean density $N/(2\pi)$, a block containing $(2j+1)$ consecutive points should have expected length approximately $(2j+1) \cdot (2\pi/N)$. This motivates the discrepancy variable

$$\delta_n(j) := \frac{N}{2\pi}\big(\theta_{n+1+j} - \theta_{n-j}\big) - (2j+1), \tag{6}$$

which is negative when the configuration is locally denser than average (the span is shorter than expected) and positive when it is sparser.

A useful equivalent form relates $\delta_n(j)$ to increments of the centered counting function S_U. Since $S_U(\theta_k) = k - \frac{N}{2\pi}\theta_k$ for the lifted eigenangles, we compute

$$S_U(\theta_{n+1+j}) - S_U(\theta_{n-j}) = (2j+1) - \frac{N}{2\pi}\big(\theta_{n+1+j} - \theta_{n-j}\big) = -\delta_n(j). \tag{7}$$

Thus $\delta_n(j)$ is exactly the negative fluctuation of the counting function across the arc from θ_{n-j} to θ_{n+1+j}.

The density-wave estimator is then defined as the weighted sum

$$D_J(m_n) := \sum_{j=J}^{\lfloor N/4 \rfloor} \frac{\delta_n(j)}{j}. \tag{8}$$

The harmonic weight $1/j$ is chosen to mimic the Fourier weights of the logarithmic kernel, and the lower cutoff J suppresses the most microscopic discrepancies, which are dominated by the nearest-neighbor repulsion scale and are not stable under mesoscopic averaging. As with A_K, we will ultimately restrict to the logarithmic window $J \asymp \log N$, where this smoothing is strong enough to reduce window dependence but weak enough to preserve the fluctuations of interest.

3.5 Truncation at $\lfloor N/4 \rfloor$

The upper limit $\lfloor N/4 \rfloor$ in (8) is a convenient deterministic truncation ensuring that the symmetric arc from θ_{n-j} to θ_{n+1+j} does not wrap around the circle by more than half its length. Indeed, for $j \le N/4$ the block of indices $\{n-j, \dots, n+1+j\}$ has size at most $N/2+1$, so the corresponding span is canonically interpreted using the lifted angles and does not depend on a choice of complementary arc.

From the analytic viewpoint, this cutoff plays the role of excluding very long-range contributions where the periodicity of the circle becomes dominant and the logarithmic kernel is effectively replaced by a bounded, smooth function. In particular, the tail $j > \alpha N$ (with any fixed $\alpha \in (0, 1/2)$) contributes only a bounded correction, and changing the cutoff from $N/4$ to αN alters $D_J(m_n)$ by a term that will be negligible (or absorbed into a deterministic renormalization) in the regimes we consider. We will later formalize this robustness when we compare D_J to a kernel convolution against dS_U, where the kernel is naturally defined modulo an additive constant.

3.6 Window parameters and the role of $K, J \asymp \log N$

The definitions above make sense for any integers $K, J \ge 1$. However, the purpose of introducing the window universality band $c_1 \log N \le K, J \le c_2 \log N$ is already visible at the level of these constructions.

If K is too small, then $\log A_K(m_n)$ still retains a strong dependence on the nearest few eigenangles and hence exhibits pronounced sensitivity to microscopic fluctuations; in particular, changing K by a constant changes the observable by a non-negligible random amount. If K is too large (say K a power of N), then the subtraction removes eigenangles out to mesoscopic or even macroscopic distances, and the residual field is distorted by the geometry of the circle and by global constraints (such as the fixed total number of points). The logarithmic scale $K \asymp \log N$ balances these effects: it

removes enough of the near-singular behavior to make the remainder stable, while keeping the subtraction local enough that different admissible window sizes lead to asymptotically equivalent observables.

An analogous discussion applies to D_J. The sum (8) with J microscopic would overweight the smallest scales, while choosing J macroscopic would discard the regime where the logarithmic correlations accumulate. The harmonic weighting makes the dominant contribution arise from a broad range of j, and the logarithmic window $J \asymp \log N$ is precisely where one expects different regularizations of the same log-correlated structure to agree.

3.7 Outlook: toward kernel representations

We have thus defined two families of fields on the midpoint grid: the local-product estimator $n \mapsto \log A_K(m_n)$ and the discrepancy estimator $n \mapsto D_J(m_n)$. The subsequent step is to place both in a common analytic framework. The key point is that each observable can be rewritten, up to a deterministic renormalization and a small error, as a centered linear statistic of the eigenangles, equivalently as an integral against dS_U with an explicit 2π-periodic kernel depending on K (or J). Once this is done, comparisons between different window choices reduce to deterministic comparisons of kernels in a Hilbert norm adapted to the CUE covariance, and probabilistic control follows from uniform concentration of such linear statistics. This kernel formalism is developed in the next section.

4 Kernel representations

In this section we recast both families of observables as mesoscopic linear statistics of the centered counting function. This is the common analytic form that permits (i) deterministic comparisons of different window choices via Fourier multipliers, and (ii) probabilistic control via concentration for determinantal linear statistics, uniformly over the midpoint grid.

4.1 The centered counting function as a signed measure

We view the eigenangle configuration as the random point measure

$$\mu_U := \sum_{j=1}^{N} \delta_{\theta_j} \qquad \text{on } \mathbb{S}^1 \simeq [0, 2\pi),$$

and we recall the centered counting function

$$S_U(\theta) := \mu_U\big([0,\theta]\big) - \frac{N}{2\pi}\theta,$$

with the convention that S_U is extended to $\mathbb{R}$ by $S_U(\theta + 2\pi) = S_U(\theta)$ and periodic indexing of eigenangles. It is convenient to encode S_U through the signed measure

$$dS_U(\phi) := d\mu_U(\phi) - \frac{N}{2\pi}\, d\phi, \tag{9}$$

which has total mass zero on $[0, 2\pi)$. For any 2π-periodic test function f that is integrable (and, for variance bounds later, has sufficient Sobolev regularity), we have the identity

$$\int_0^{2\pi} f(\phi)\, dS_U(\phi) = \sum_{j=1}^{N} f(\theta_j) - \frac{N}{2\pi} \int_0^{2\pi} f(\phi)\, d\phi. \tag{10}$$

Shifting the test function corresponds to evaluating the same linear statistic at different locations:

$$\int_0^{2\pi} f(\phi - m_n)\, dS_U(\phi) = \sum_{j=1}^{N} f(\theta_j - m_n) - \frac{N}{2\pi} \int_0^{2\pi} f(\phi)\, d\phi.$$

Because dS_U annihilates constants, the kernel f is always understood modulo an additive constant; we will exploit this when choosing convenient normalizations.

4.2 A logarithmic kernel for $\log|Z_U|$

Let

$$\ell(x) := \log\Big(2\Big|\sin\Big(\frac{x}{2}\Big)\Big|\Big), \qquad x \in \mathbb{R} \setminus 2\pi\mathbb{Z},$$

extended as a 2π-periodic function. Then (3) may be rewritten as

$$\log\big|e^{i\theta} - e^{i\phi}\big| = \ell(\theta - \phi).$$

Consequently, for each $\theta \in \mathbb{R}$,

$$\log|Z_U(\theta)| = \sum_{j=1}^{N} \ell(\theta - \theta_j) = \int_0^{2\pi} \ell(\theta - \phi)\, d\mu_U(\phi). \tag{11}$$

Subtracting the deterministic mean term gives the centered form

$$\log|Z_U(\theta)| - \frac{N}{2\pi} \int_0^{2\pi} \ell(\theta - \phi)\, d\phi = \int_0^{2\pi} \ell(\theta - \phi)\, dS_U(\phi). \tag{12}$$

The integral $\int_0^{2\pi} \ell(\theta - \phi)\, d\phi$ does not depend on θ by periodicity; thus (12) expresses $\log|Z_U(\theta)|$ as a linear statistic of dS_U against the logarithmic kernel ℓ, plus a deterministic constant of size N.

For Fourier analysis we use the classical expansion, valid in $L^2(\mathbb{S}^1)$,

$$\ell(x) = \log 2 - \sum_{m=1}^{\infty} \frac{\cos(mx)}{m} = \log 2 - \frac{1}{2} \sum_{m \in \mathbb{Z}\setminus\{0\}} \frac{e^{imx}}{|m|}. \tag{13}$$

Thus, modulo an additive constant, ℓ has Fourier coefficients $\widehat{\ell}(m) = -\frac{1}{2|m|}$ for $m \neq 0$. It is precisely this $1/|m|$ multiplier that will govern the variance norm used later (an $H^{1/2}$-type norm induced by CUE covariances).

4.3 Kernel form for the local product estimator A_K

The definition of $\log A_K(m_n)$ subtracts the contributions of the $2K$ eigenangles nearest to m_n. While this subtraction is defined in terms of indices, its effect is to remove the logarithmic singularity of $\ell(\cdot - m_n)$ at scale comparable to the typical distance from m_n to the Kth neighbor, namely $\asymp K/N$ when $K \ll N$. We therefore introduce a deterministic mesoscopic regularization of the kernel ℓ at that scale.

Fix once and for all a smooth even cutoff $\chi \in C^\infty(\mathbb{R})$ such that $\chi(x) = 1$ for $|x| \leq 1$ and $\chi(x) = 0$ for $|x| \geq 2$. For an integer window size $K \geq 1$, set

$$M := M(K, N) := \Big\lfloor \frac{N}{K} \Big\rfloor, \qquad \widehat{\chi}_M(m) := \chi\Big(\frac{m}{M}\Big),$$

and define the 2π-periodic kernel

$$\kappa_{K,N}(x) := -\frac{1}{2} \sum_{m \in \mathbb{Z}\setminus\{0\}} \frac{\widehat{\chi}_M(m)}{|m|}\, e^{imx}. \tag{14}$$

By construction, $\kappa_{K,N}$ is a Fourier-truncated version of ℓ (up to an additive constant): it retains frequencies $|m| \lesssim N/K$ and suppresses higher modes, which correspond to features of ℓ at spatial scales $\ll K/N$. In particular, $\kappa_{K,N}$ is bounded and continuous, and it approximates ℓ away from a neighborhood of 0 of size comparable to K/N.

We now relate $\kappa_{K,N}$ to $\log A_K(m_n)$. Starting from (11) at $\theta = m_n$ and using the subtraction in (4), we may write

$$\log A_K(m_n) = \sum_{\substack{1 \leq j \leq N \\ j \notin \mathcal{N}_{n,K}}} \ell(m_n - \theta_j) + C_{K,N}, \tag{15}$$

where $\mathcal{N}_{n,K}$ is the index set of the $2K$ removed eigenangles,

$$\mathcal{N}_{n,K} := \{n - K + 1, \ldots, n\} \cup \{n + 1, \ldots, n + K\},$$

interpreted with periodic extension. The key observation is that, on the event that all midpoints are separated from eigenangles by $\gg 1/N$ (cf.

Lemma 5 in the global outline), the contribution of the removed terms is well-approximated by subtracting a deterministic local counterterm depending only on K and N, not on the fine configuration. Equivalently, we may replace the random excision in (15) by a deterministic smoothing of ℓ at scale K/N, which yields a linear statistic of dS_U.

Concretely, one shows (by comparing (15) with the Fourier-truncated kernel (14) and using that the removed indices correspond to angles within distance $\asymp K/N$ of m_n) that there exists a deterministic constant $c_{K,N} \in \mathbb{R}$ such that

$$\log A_K(m_n) = \int_0^{2\pi} \kappa_{K,N}(\phi - m_n)\, dS_U(\phi) + r_{K,N}(n), \tag{16}$$

where the remainder $r_{K,N}(n)$ absorbs (i) the difference between excising the nearest $2K$ points and smoothing the kernel at scale K/N, and (ii) the deterministic renormalization, i.e. the combined effect of $C_{K,N}$ and the mean term $\frac{N}{2\pi}\int \kappa_{K,N}$. Since dS_U annihilates constants, we may (and will) choose $c_{K,N}$ so that $\mathbb{E}\int \kappa_{K,N}(\phi - m_n)\, dS_U(\phi) = 0$, and then $C_{K,N}$ precisely implements this centering in the original definition. In particular, (16) is compatible with the normalization $\mathbb{E}[\log A_K(m_1)] = 0$.

The salient analytic features of $\kappa_{K,N}$ are immediate from (14): for $m \neq 0$,

$$\widehat{\kappa}_{K,N}(m) = -\frac{1}{2|m|}\,\widehat{\chi}_M(m), \qquad \text{hence} \qquad \widehat{\kappa}_{K,N}(m) = -\frac{1}{2|m|} + O\Big(\frac{1}{|m|}\mathbf{1}_{|m|\gtrsim N/K}\Big), \tag{17}$$

and the effective Fourier support satisfies $|m| \lesssim N/K$. The dependence on K is thus encoded by a low-pass filter at frequency N/K, which is exactly the mesoscopic scale relevant for window universality when $K \asymp \log N$.

4.4 Kernel form for the discrepancy estimator D_J

We turn to the discrepancy-based estimator

$$D_J(m_n) = \sum_{j=J}^{\lfloor N/4 \rfloor} \frac{\delta_n(j)}{j}, \qquad \delta_n(j) = -\big(S_U(\theta_{n+1+j}) - S_U(\theta_{n-j})\big)$$

by (7). Thus

$$D_J(m_n) = -\sum_{j=J}^{\lfloor N/4 \rfloor} \frac{S_U(\theta_{n+1+j})}{j} + \sum_{j=J}^{\lfloor N/4 \rfloor} \frac{S_U(\theta_{n-j})}{j}. \tag{18}$$

The right-hand side is a discrete Hilbert-transform-type expression: it takes values of S_U at a ladder of eigenangles around m_n, weighted harmonically by $1/j$. To convert this into an integral against dS_U, we perform a deterministic

summation-by-parts approximation that replaces the ladder of evaluation points $\theta_{n\pm j}$ by a spatial convolution kernel at scale J/N.

We first note that S_U is piecewise constant with jumps of size 1 at eigenangles and a deterministic drift $-\frac{N}{2\pi}$ between them. This suggests rewriting sums of the form $\sum_j a_j S_U(\theta_{n+j})$ as Riemann–Stieltjes integrals against dS_U with kernels built from partial sums of the coefficients a_j. In the present setting, let

$$H_{a,b} := \sum_{j=a}^{b} \frac{1}{j}$$

denote partial harmonic sums. By Abel summation applied to (18), one may express $D_J(m_n)$ as a linear combination of increments of S_U across arcs determined by the eigenangles $\theta_{n\pm j}$. Using again (7) to rewrite these increments as integrals of dS_U, and then replacing the random arc endpoints by the deterministic location $m_n \pm 2\pi j/N$ at the level of the kernel (the error is absorbed into a remainder term), we obtain an integral representation of the form

$$D_J(m_n) = \int_0^{2\pi} \eta_{J,N}(\phi - m_n)\, dS_U(\phi) + q_{J,N}(n). \tag{19}$$

Here $\eta_{J,N}$ is an explicit 2π-periodic kernel depending only on (J, N), and $q_{J,N}(n)$ is an error that accounts for (i) the replacement of eigenangle endpoints by deterministic endpoints, and (ii) the tail handling coming from the truncation $j \le \lfloor N/4 \rfloor$ (which, as discussed earlier, is irrelevant up to a bounded deterministic adjustment).

We choose $\eta_{J,N}$ in Fourier space in a way parallel to (14). Namely, with $M' := \lfloor N/J \rfloor$ and the same cutoff χ, define

$$\eta_{J,N}(x) := -\frac{1}{2} \sum_{m \in \mathbb{Z}\setminus\{0\}} \frac{\widehat{\chi}_{M'}(m)}{|m|}\, e^{imx}. \tag{20}$$

This kernel has the same leading multiplier $-1/(2|m|)$ as ℓ, but is regularized at the microscopic scales excluded by the lower cutoff J. In particular,

$$\widehat{\eta}_{J,N}(m) = -\frac{1}{2|m|}\, \widehat{\chi}_{M'}(m), \qquad \text{so that} \qquad \widehat{\eta}_{J,N}(m) = -\frac{1}{2|m|} + O\Big(\frac{1}{|m|}\mathbf{1}_{|m| \gtrsim N/J}\Big), \tag{21}$$

and $|m| \lesssim N/J$ is the effective frequency range. The point is that the harmonic weights in (18) implement, after summation by parts, precisely such a low-frequency cutoff for the logarithmic kernel. The truncation at $j = \lfloor N/4 \rfloor$ affects only the very lowest frequencies (equivalently, adds a smooth bounded function), hence contributes at most an additive constant under integration against dS_U.

As in the case of A_K, we emphasize that (19) is invariant under adding a constant to $\eta_{J,N}$. We may therefore normalize $\eta_{J,N}$ so that $\int_0^{2\pi} \eta_{J,N}(\phi)\, d\phi =$

0, in which case $\mathbb{E}\int \eta_{J,N}(\phi - m_n)\, dS_U(\phi) = 0$ and the deterministic mean of $D_J(m_n)$ is absorbed into the error term $q_{J,N}(n)$ (or, equivalently, into a deterministic recentering of D_J). For the comparison results later, only differences of kernels matter, so any fixed centering convention is admissible.

4.5 Consequences for window comparisons

The representations (16) and (19) reduce the analysis of window dependence to deterministic comparison of the kernels $\kappa_{K,N}$ and $\eta_{J,N}$. In the logarithmic window $K, J \asymp \log N$, both kernels have Fourier multipliers agreeing with $-1/(2|m|)$ up to a cutoff at frequency $\asymp N/\log N$, hence their differences have small mass in the variance-controlling norm. The probabilistic task is then to control

$$\max_{1\le n\le N} \left| \int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) - \mathbb{E}\int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) \right|$$

for families of kernels f_N with uniformly bounded variance norm, and to show that the remainder terms $r_{K,N}(n)$ and $q_{J,N}(n)$ are negligible uniformly in n for $K, J \in [c_1 \log N, c_2 \log N]$. These two ingredients, combined with the kernel multiplier bounds (17)–(21), will yield the uniform window universality and estimator equivalence statements formulated in the next section.

5 Main theorem: window universality and estimator equivalence

We now state precisely the asymptotic statements announced in the introduction, with attention to (a) the mode of convergence, (b) the uniformity in the window parameters, and (c) the role of the midpoint grid. Throughout we keep fixed constants $0 < c_1 < c_2$ and we restrict to integer windows in the logarithmic band

$$\mathcal{W}_N := \big\{K \in \mathbb{Z} : c_1 \log N \le K \le c_2 \log N\big\}.$$

We view the collections $\{\log A_K(m_n)\}_{n=1}^N$ and $\{D_J(m_n)\}_{n=1}^N$ as random elements of $\ell^\infty(\{1,\dots,N\})$ (or, equivalently, as random fields indexed by the midpoint grid), and we will compare them in the sup norm over n.

5.1 Uniformity and the probability metric

For random variables X_N we use the standard notion $X_N = o_{\mathbb{P}}(1)$ to mean that for every $\varepsilon > 0$, $\mathbb{P}(|X_N| > \varepsilon) \to 0$ as $N \to \infty$. When dealing with a family indexed by windows and locations, we will make the uniformity explicit by placing the supremum inside the probability statement; for example,

$$\sup_{K\in\mathcal{W}_N} \max_{1\le n\le N} |X_N(K,n)| = o_{\mathbb{P}}(1)$$

means that for every $\varepsilon > 0$,

$$\mathbb{P}\Big(\sup_{K \in \mathcal{W}_N} \max_{1 \le n \le N} |X_N(K,n)| > \varepsilon \Big) \longrightarrow 0.$$

This is stronger than pointwise convergence for each fixed K and n, but it is the natural formulation for the window-universality principle: the intrinsic mesoscopic scale should not depend on which particular $K \asymp \log N$ is chosen, and this should hold simultaneously across the whole configuration.

We emphasize that the supremum over $K \in \mathcal{W}_N$ is not a major strengthening in practice, since $|\mathcal{W}_N| \asymp \log N$ and hence one may pay logarithmic union bounds at intermediate steps. The substantive part is the supremum over n, which forces us to control linear statistics uniformly over N correlated evaluation points. This is precisely why we work with the midpoint grid and why we postpone the probabilistic input to the next section.

5.2 Statement of the main theorem

Theorem 5.1 (Window universality and estimator equivalence)**.** *Let $U \sim \mathrm{Haar}(\mathrm{U}(N))$ and let $m_1, \dots, m_N$ be the midpoints between consecutive eigenangles, as defined above. Fix $0 < c_1 < c_2$ and let $\mathcal{W}_N = \{K \in \mathbb{Z} : c_1 \log N \le K \le c_2 \log N\}$. Then, as $N \to \infty$, the following hold.*

(i) Window universality for the local product estimator. *We have*

$$\sup_{K_1, K_2 \in \mathcal{W}_N} \max_{1 \le n \le N} \big| \log A_{K_1}(m_n) - \log A_{K_2}(m_n) \big| = o_{\mathbb{P}}(1). \tag{22}$$

(ii) Equivalence between A_K and the discrepancy estimator. *We have*

$$\sup_{K, J \in \mathcal{W}_N} \max_{1 \le n \le N} \big| \log A_K(m_n) - D_J(m_n) \big| = o_{\mathbb{P}}(1). \tag{23}$$

The content of (22) is that, once K is taken in the logarithmic band, the dependence of $\log A_K(m_n)$ on the particular choice of K vanishes in the strongest natural sense available on the midpoint grid: the maximum deviation over all midpoints tends to zero in probability, uniformly over all admissible window sizes. The content of (23) is that the same canonical field is captured by the discrepancy-based construction D_J (again uniformly over admissible J).

5.3 Canonical carrier wave on the midpoint grid

Theorem 5.1 implies that the "carrier wave" field can be defined canonically up to a vanishing error, without specifying a particular estimator or window. We record this as a corollary in a form that will be convenient later.

Corollary 5.2 (Canonical carrier wave on midpoints)**.** *Fix any $c \in [c_1, c_2]$ and define*

$$\mathrm{CW}_N(m_n) := \log A_{\lfloor c \log N \rfloor}(m_n), \qquad 1 \le n \le N.$$

Then for any other $c' \in [c_1, c_2]$ and any admissible discrepancy estimator $D_{\lfloor c' \log N \rfloor}$ we have, as $N \to \infty$,

$$\max_{1 \le n \le N} \left| \mathrm{CW}_N(m_n) - \log A_{\lfloor c' \log N \rfloor}(m_n) \right| = o_{\mathbb{P}}(1), \qquad \max_{1 \le n \le N} \left| \mathrm{CW}_N(m_n) - D_{\lfloor c' \log N \rfloor}(m_n) \right| = o_{\mathbb{P}}(1).$$

In particular, any two such constructions yield the same random field on the midpoint grid up to $o_{\mathbb{P}}(1)$ in ℓ^∞.

We stress that the normalization constant $C_{K,N}$ in the definition of $\log A_K$ only affects additive deterministic shifts. Our convention $\mathbb{E}[\log A_K(m_1)] = 0$ is convenient but inessential: changing the centering changes neither (22) nor (23), since both involve differences that eliminate deterministic constants. Likewise, for D_J one may adopt any fixed deterministic centering convention; all choices are asymptotically equivalent in the present comparison results.

5.4 Why we work on midpoints (and what can be changed)

The midpoint grid plays two roles. First, it avoids evaluation at eigenangles, where $\log |Z_U(\theta)|$ has logarithmic singularities and where the product definition of A_K is most sensitive to microscopic fluctuations. Second, it provides a canonical set of N locations with spacing $\asymp 1/N$, which is the natural discretization scale for sup norm comparisons in this setting.

That said, the midpoint grid is not essential for the underlying mechanism. Any grid $\{\vartheta_n\}_{n=1}^N$ which is (i) deterministically chosen, (ii) spaced by $\asymp 1/N$, and (iii) stays at distance $\gg N^{-2}$ from the eigenangles with high probability, could be used with no change to the qualitative conclusions. The midpoint grid is simply the most economical choice because the separation condition holds automatically with high probability (cf. Lemma 5 in the global outline) and because it interacts well with the symmetric indexing used in A_K and D_J.

A stronger variant, not needed for our main application, is to extend the comparison from a discrete grid to a continuous set of angles. One should not expect uniform control over all $\theta \in [0, 2\pi]$ without excluding neighborhoods of the eigenangles, since any field built from $\log |Z_U(\theta)|$ must diverge at zeros. However, on the restricted set

$$\mathcal{G}_N(\delta) := \left\{ \theta \in [0, 2\pi] : \min_{1 \le j \le N} |\theta - \theta_j| \ge \delta / N \right\}$$

with fixed $\delta > 0$, one can interpolate between grid points and upgrade midpoint statements to uniform-in-θ statements using deterministic modulus-of-continuity bounds for the smoothed kernels together with rigidity estimates

controlling the local spacing. We do not pursue this here, since the midpoint formulation is already sufficient for defining the carrier wave canonically and for subsequent extrema arguments.

5.5 Alternative norms: pointwise, L^2, and sup

It is instructive to separate three levels of comparison.

Pointwise (fixed n). For each fixed midpoint m_n, it is relatively straightforward to show that $\log A_{K_1}(m_n) - \log A_{K_2}(m_n) \to 0$ in probability when $K_1, K_2 \asymp \log N$. Indeed, after the kernel representations of Section 4, the difference becomes a single centered linear statistic against a kernel difference, and standard determinantal variance bounds imply convergence.

L^2 over n. One can strengthen pointwise convergence to an averaged statement such as

$$\frac{1}{N}\sum_{n=1}^{N}\bigl|\log A_{K_1}(m_n) - \log A_{K_2}(m_n)\bigr|^2 \;\xrightarrow{\mathbb{P}}\; 0,$$

uniformly over $K_1, K_2 \in \mathcal{W}_N$, by combining variance bounds with the translation invariance of the CUE process and a Fubini argument. Such statements are closer in spirit to "energy" estimates for log-correlated fields and can sometimes be proved with minimal effort.

Sup norm over n. The statements (22)–(23) are stronger because they rule out the possibility of a sparse exceptional set of indices where the approximation fails. From the perspective of applications to maxima, this is the relevant mode of control: when studying $\max_n \mathrm{CW}_N(m_n)$, an L^2 error does not preclude rare but large perturbations that could alter the maximum. Consequently we work throughout with the ℓ^∞ error on the full grid.

The cost of the sup norm is that it forces an additional discretization step when applying concentration inequalities. In the next section we implement this using a net argument adapted to the regularity of the kernels, together with number variance bounds controlling increments of S_U on mesoscopic scales.

5.6 How the theorem follows from the kernel framework

Section 4 reduces the nonlinear definitions of $\log A_K$ and D_J to kernel integrals against the signed measure dS_U, up to remainder terms:

$$\log A_K(m_n) = \int_0^{2\pi} \kappa_{K,N}(\phi - m_n)\, dS_U(\phi) + r_{K,N}(n), \qquad D_J(m_n) = \int_0^{2\pi} \eta_{J,N}(\phi - m_n)\, dS_U(\phi) + q_{J,N}(n)$$

The proof of Theorem 5.1 then decomposes into two logically independent tasks.

Deterministic comparison of kernels. One shows that for $K_1, K_2, J \in \mathcal{W}_N$ the differences $\kappa_{K_1,N} - \kappa_{K_2,N}$ and $\kappa_{K,N} - \eta_{J,N}$ are small in a Hilbert norm $\|\cdot\|_{\mathcal{H}}$ that controls the variance of the corresponding linear statistic. Concretely, since both kernels have Fourier multipliers approximately $-1/(2|m|)$ up to a cutoff at frequency $\asymp N/\log N$, their difference is supported on a shrinking band of high frequencies and therefore has $\|\cdot\|_{\mathcal{H}}$ norm tending to zero.

Uniform probabilistic control. One shows that for any family of kernels f_N with bounded $\mathcal{H}$-norm, the centered linear statistics

$$\int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) - \mathbb{E}\int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi)$$

are uniformly small over n with high probability. This is the step that upgrades variance control to a sup bound and is the only genuinely probabilistic input.

Finally, the remainder terms $r_{K,N}(n)$ and $q_{J,N}(n)$ are shown to satisfy

$$\sup_{K\in\mathcal{W}_N} \max_{1\le n\le N} |r_{K,N}(n)| = o_{\mathbb{P}}(1), \qquad \sup_{J\in\mathcal{W}_N} \max_{1\le n\le N} |q_{J,N}(n)| = o_{\mathbb{P}}(1),$$

so they do not affect the leading comparison. Combining these three components with a triangle inequality yields (22) and (23).

5.7 Remarks on robustness and truncations

Two robustness features are worth keeping in view, since they often arise in numerical or heuristic formulations.

Choice of far cutoff in D_J. The truncation $j \le \lfloor N/4 \rfloor$ in D_J is convenient to keep the symmetric spans away from the full circle, but it is not essential. Replacing $\lfloor N/4 \rfloor$ by $\lfloor \alpha N \rfloor$ with fixed $\alpha \in (0, 1/2)$ changes D_J by adding a contribution corresponding to a smooth bounded kernel; under integration against dS_U this modification is negligible up to a deterministic constant, hence does not affect (23).

Alternative grids. As already noted, the midpoint grid can be replaced by any grid separated from eigenangles and with mesh $\asymp 1/N$. The midpoint choice merely packages the singularity avoidance into the definition and removes the need to condition on additional events. In particular, in applications where one wishes to track the carrier wave along a deterministic arc, one may sample on a uniform grid and omit the points too close to eigenangles; the remaining points satisfy the same comparison estimates.

The next section provides the concentration and rigidity estimates that complete the proof of Theorem 5.1 by supplying the uniform probabilistic bounds required by the kernel comparison strategy.

6 Concentration and rigidity inputs

In this section we record the probabilistic estimates for the CUE eigenangle process that we will use to control, uniformly over the midpoint grid, the linear statistics arising from the kernel representations of Section 4. The point is not to optimize constants, but to isolate bounds that (a) are uniform in the evaluation location, (b) are stable under mesoscopic smoothing, and (c) are strong enough to support a discretization argument for sup norms.

6.1 Linear statistics and the $H^{1/2}$-type variance norm

For a 2π-periodic (real-valued) test function f we write the centered linear statistic in the form

$$X_f := \sum_{j=1}^{N} f(\theta_j) - \frac{N}{2\pi}\int_0^{2\pi} f(\phi)\,d\phi = \int_0^{2\pi} f(\phi)\,dS_U(\phi). \tag{24}$$

We will frequently consider shifts $f(\cdot - t)$ and write $X_f(t) := X_{f(\cdot - t)}$. By rotation invariance of Haar measure, $X_f(t)$ has the same law for every deterministic t, and its mean is independent of t (equal to 0 when f is centered as in (24)).

The basic deterministic input is the exact Fourier-space variance formula for CUE linear statistics. Writing

$$\widehat{f}(k) := \frac{1}{2\pi}\int_0^{2\pi} f(\phi)e^{-ik\phi}\,d\phi, \qquad k \in \mathbb{Z},$$

we have the identity (see, e.g., the Diaconis–Shahshahani moment method, or standard determinantal computations for the CUE projection kernel)

$$\mathrm{Var}(X_f) = \sum_{k\in\mathbb{Z}\setminus\{0\}} \min(N, |k|)\,|\widehat{f}(k)|^2. \tag{25}$$

This suggests the Hilbert norm

$$\|f\|_{\mathcal{H}_N}^2 := \sum_{k\in\mathbb{Z}\setminus\{0\}} \min(N, |k|)\,|\widehat{f}(k)|^2, \tag{26}$$

so that $\mathrm{Var}(X_f) = \|f\|_{\mathcal{H}_N}^2$. We stress that $\|\cdot\|_{\mathcal{H}_N}$ is a discrete analogue of the $H^{1/2}$ norm and is the natural quantity for comparing kernels via Fourier multipliers. In particular, for two test functions f, g we have

$$\mathrm{Var}(X_f - X_g) = \|f - g\|_{\mathcal{H}_N}^2, \tag{27}$$

and the same identity holds after shifting both test functions by the same amount.

A second elementary estimate we will use concerns increments under small shifts. If $\Delta_h f := f(\cdot - h) - f(\cdot)$, then

$$\|\Delta_h f\|_{\mathcal{H}_N}^2 = \sum_{k \neq 0} \min(N, |k|)\, |\widehat{f}(k)|^2\, |e^{-ikh} - 1|^2. \tag{28}$$

Using $|e^{-ikh} - 1| \leq \min(2, |k||h|)$, we obtain the general bound

$$\|\Delta_h f\|_{\mathcal{H}_N} \leq 2\|f\|_{\mathcal{H}_N}, \tag{29}$$

and, if in addition $\widehat{f}(k) = 0$ for $|k| > M$, then

$$\|\Delta_h f\|_{\mathcal{H}_N} \leq C\, \min(1, M|h|)\, \|f\|_{\mathcal{H}_N}, \tag{30}$$

for an absolute constant C. The cutoff version (30) is the relevant one for the kernels appearing later, whose Fourier support is mesoscopically truncated.

6.2 Subgaussian concentration for determinantal linear statistics

Beyond second moments, we require a uniform tail bound for X_f. For projection determinantal point processes (such as CUE eigenangles), the Laplace transform of X_f admits a Fredholm determinant representation. One can combine this representation with standard trace inequalities to obtain subgaussian concentration in terms of $\|f\|_{\mathcal{H}_N}$, provided f is bounded.

We record the following form, which is sufficient for our applications and may be proved by adapting Soshnikov's determinant method (or more recent general concentration inequalities for determinantal processes).

Proposition 6.1 (Subgaussian tail bound)**.** *There exist absolute constants $c, C > 0$ such that for every bounded real 2π-periodic f and every $u \geq 0$,*

$$\mathbb{P}\Big(|X_f - \mathbb{E}X_f| > u\Big) \leq 2\exp\Big(-c\,\frac{u^2}{\|f\|_{\mathcal{H}_N}^2}\Big), \tag{31}$$

provided $u \leq C\,\|f\|_{\mathcal{H}_N}^2/\|f\|_\infty$ (a regime that will contain the small-deviation bounds we use). The same estimate holds uniformly for the shifted statistics $X_f(t)$ with deterministic t.

We will apply Proposition 6.1 with test functions $f = f_N$ whose $\mathcal{H}_N$-norm tends to 0 (typically because f_N is a difference of two nearby kernels). In that regime (31) yields very strong concentration, and the constraint on u is immaterial.

6.3 Mesoscopic number variance and increments of the counting function

We also need a rigidity-type statement for the counting function, at least in the weak form of uniform control on mesoscopic increments. This is most conveniently expressed as a number variance bound for arcs.

For an interval (arc) $I = [a, b] \subset [0, 2\pi]$ (interpreted on $\mathbb{S}^1$ when convenient), let

$$\mathcal{N}(I) := \#\{j : \theta_j \in I\}.$$

Then $\mathcal{N}(I) - \frac{N}{2\pi}|I|$ equals $S_U(b) - S_U(a)$ up to the fixed centering convention in the definition of S_U. Thus a variance bound for $\mathcal{N}(I)$ directly controls mesoscopic increments of S_U.

Lemma 6.2 (Number variance bound)**.** *There exists an absolute constant $C > 0$ such that for every arc $I \subset [0, 2\pi]$ of length $|I| \in (0, 2\pi]$,*

$$\operatorname{Var}\big(\mathcal{N}(I)\big) \le C \log\big(2 + N|I|\big). \tag{32}$$

Consequently, for every $u \ge 0$ in the small-deviation regime,

$$\mathbb{P}\Big(\big|\mathcal{N}(I) - \mathbb{E}\mathcal{N}(I)\big| > u\Big) \le 2\exp\Big(-c\,\frac{u^2}{\log(2 + N|I|)}\Big), \tag{33}$$

with an absolute constant $c > 0$.

A direct proof of (32) may be obtained from (25) by approximating $\mathbf{1}_I$ with a bounded trigonometric polynomial and using that $|\widehat{\mathbf{1}_I}(k)| \ll \min(|I|, |k|^{-1})$. The logarithmic behavior reflects the well-known long-range rigidity of the CUE process and will be used only to bound increments across a mesoscopic net.

6.4 A separation event for midpoints

Although our main comparisons ultimately run through smoothed kernels, we still need a high-probability event on which evaluation at the midpoints does not probe the logarithmic singularities of $\log|Z_U|$ too closely. A minimal such event is that no eigenangle gap is *extremely* small.

Let $\Delta_n := \theta_{n+1} - \theta_n$ (with $\theta_{N+1} = \theta_1 + 2\pi$). Since m_n lies at distance $\Delta_n/2$ from each of the neighboring eigenangles, a lower bound on $\min_n \Delta_n$ implies a uniform lower bound on $\min_{n,j} |m_n - \theta_j|$.

Lemma 6.3 (No ultra-small gaps)**.** *There exists an absolute constant $c > 0$ such that*

$$\mathbb{P}\Big(\min_{1 \le n \le N} \Delta_n \le N^{-2}\Big) = o(1). \tag{34}$$

In particular, with probability $1 - o(1)$ we have

$$\min_{1 \le n \le N} \min_{1 \le j \le N} |m_n - \theta_j| \ge \tfrac{1}{2} N^{-2}. \tag{35}$$

One way to prove (34) is to bound the expected number of pairs of eigenangles within distance N^{-2} using the two-point correlation function of CUE and its quadratic repulsion at short scales; the resulting expectation is $O(N^{-1})$, and Markov's inequality yields the claim. We will invoke Lemma 6.3 only to justify that various remainder terms (arising from local logarithms evaluated at m_n) are uniformly bounded with overwhelming probability.

6.5 A net argument for uniform control over the midpoint grid

We now explain how to combine the preceding tail bounds with a discretization argument to control a supremum over many evaluation points. The key observation is that, for the kernels we encounter, $t \mapsto X_f(t)$ varies slowly on a mesoscopic scale due to Fourier truncation, so it suffices to control $X_f(t)$ on a deterministic net and to control increments between nearby points.

Fix a bounded real test function f and define the centered process

$$Y_f(t) := X_f(t) - \mathbb{E}X_f(t), \qquad t \in [0, 2\pi].$$

Let $\mathcal{T}_h := \{0, h, 2h, \dots, \lfloor 2\pi/h \rfloor h\}$ be a deterministic net of mesh $h \in (0, 1]$, and let $\pi_h(t) \in \mathcal{T}_h$ denote a nearest net point to t. Then for all t,

$$|Y_f(t)| \le |Y_f(\pi_h(t))| + \sup_{|t-s|\le h} |Y_f(t) - Y_f(s)|. \tag{36}$$

The increment satisfies

$$Y_f(t) - Y_f(s) = Y_{\Delta_{t-s} f}(s),$$

so by Proposition 6.1,

$$\mathbb{P}\Big(|Y_f(t) - Y_f(s)| > u\Big) \le 2\exp\Big(-c\,\frac{u^2}{\|\Delta_{t-s} f\|_{\mathcal{H}_N}^2}\Big). \tag{37}$$

If f is spectrally localized (as will hold for our kernels), then (30) turns (37) into a strong modulus-of-continuity estimate at scales $|t - s| \le h$.

Specializing (36) to the random midpoints, we obtain deterministically

$$\max_{1\le n\le N} |Y_f(m_n)| \le \max_{t\in\mathcal{T}_h} |Y_f(t)| + \sup_{|t-s|\le h} |Y_f(t) - Y_f(s)|. \tag{38}$$

The right-hand side involves only suprema over deterministic sets, and thus may be bounded by union bounds using (31)–(37). Concretely, for any $u > 0$ we have

$$\mathbb{P}\Big(\max_{t\in\mathcal{T}_h} |Y_f(t)| > u\Big) \le 2|\mathcal{T}_h| \exp\Big(-c\,\frac{u^2}{\|f\|_{\mathcal{H}_N}^2}\Big), \tag{39}$$

$$\mathbb{P}\Big(\sup_{\substack{s\in\mathcal{T}_h \\ |t-s|\le h}} |Y_f(t) - Y_f(s)| > u\Big) \le 2|\mathcal{T}_h| \exp\Big(-c\,\frac{u^2}{\sup_{|r|\le h} \|\Delta_r f\|_{\mathcal{H}_N}^2}\Big), \tag{40}$$

where in (40) we used that it suffices to control increments from each net point to nearby points at distance at most h (and, if desired, one may further discretize those nearby points at a second, finer mesh). The choice of h is problem-dependent: in our setting we will take h on the order of the correlation length of the kernels, so that $\sup_{|r|\le h} \|\Delta_r f\|_{\mathcal{H}_N}$ is small.

We summarize this discussion in a convenient form tailored to our later use, where $f = f_N$ will be a sequence of test functions (typically differences of kernels indexed by window parameters).

Proposition 6.4 (Uniform smallness from vanishing $\mathcal{H}_N$ norm)**.** *Let f_N be a sequence of bounded real 2π-periodic functions. Suppose that for some choice of meshes $h_N \downarrow 0$,*

$$\|f_N\|_{\mathcal{H}_N}^2 \log(1/h_N) \longrightarrow 0, \qquad \sup_{|r|\le h_N} \|\Delta_r f_N\|_{\mathcal{H}_N}^2 \log(1/h_N) \longrightarrow 0. \tag{41}$$

Then

$$\max_{1\le n\le N} \Big| \int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) - \mathbb{E} \int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) \Big| = o_{\mathbb{P}}(1). \tag{42}$$

The proof is immediate from (38) together with (39)–(40) and the choice $|\mathcal{T}_{h_N}| \asymp 1/h_N$. In practice we will take h_N comparable to the inverse of the Fourier cutoff of the relevant kernel differences; for such choices the increment condition in (41) follows from (30).

We note that Proposition 6.4 isolates precisely what is needed to upgrade pointwise concentration (for a fixed evaluation point) to a uniform bound over all midpoints: one must control not only the variance at a point, but also the typical size of increments at the discretization scale. Both quantities are naturally expressed in the same Fourier-based $\mathcal{H}_N$ norm.

In the next section we verify, by a deterministic Fourier-multiplier comparison, that the kernel differences arising from varying the window parameters satisfy (41) in the logarithmic window band, and we then apply Proposition 6.4 to obtain the required $o_{\mathbb{P}}(1)$ uniformity.

7 Comparing kernels across windows

In this section we verify the deterministic kernel-comparison statements needed to invoke Proposition 6.4. Concretely, we show that when the window parameters lie in the logarithmic band, the Fourier multipliers associated with the kernels in Lemmas 1–2 are asymptotically independent of the particular choice of window. The argument is purely analytic: once we obtain smallness in the $\mathcal{H}_N$-norm, the corresponding probabilistic $o_{\mathbb{P}}(1)$ estimates follow immediately from Proposition 6.4.

7.1 Fourier multipliers for the smoothed logarithmic kernels

We work throughout with 2π-periodic kernels of mean zero. In the notation of Lemmas 1–2 we write

$$\kappa_{K,N}(\theta) = \sum_{k\in\mathbb{Z}\setminus\{0\}} \widehat{\kappa}_{K,N}(k)\, e^{ik\theta}, \qquad \eta_{J,N}(\theta) = \sum_{k\in\mathbb{Z}\setminus\{0\}} \widehat{\eta}_{J,N}(k)\, e^{ik\theta}.$$

The key structural input from Section 4 is that both families are (i) bounded kernels obtained by smoothing a logarithmic singularity at scale comparable to K/N or J/N, and (ii) centered so that $\widehat{\kappa}_{K,N}(0) = \widehat{\eta}_{J,N}(0) = 0$. In particular, their Fourier coefficients admit the following uniform description.

Lemma 7.1 (Multiplier asymptotics in the log-window band)**.** *Fix $0 < c_1 < c_2$. There exists an absolute constant $C > 0$ such that for all N large enough and all integers $K, J \in [c_1 \log N, c_2 \log N]$ the following holds.*

1. *For every $k \in \mathbb{Z} \setminus \{0\}$ with $|k| \le N$, we have the expansion*

$$\widehat{\kappa}_{K,N}(k) = \frac{1}{2|k|}\Big(1+\varepsilon_{K,N}(k)\Big), \qquad \widehat{\eta}_{J,N}(k) = \frac{1}{2|k|}\Big(1+\widetilde{\varepsilon}_{J,N}(k)\Big), \tag{43}$$

   with error bounds

$$\sup_{1\le|k|\le N} |\varepsilon_{K,N}(k)| \le \frac{C}{K}, \qquad \sup_{1\le|k|\le N} |\widetilde{\varepsilon}_{J,N}(k)| \le \frac{C}{J}. \tag{44}$$

2. *The high-frequency tails satisfy*

$$\sup_{|k|>N} |k|^2\, |\widehat{\kappa}_{K,N}(k)| \le C, \qquad \sup_{|k|>N} |k|^2\, |\widehat{\eta}_{J,N}(k)| \le C. \tag{45}$$

We stress that (43)–(44) state that, in the entire Fourier range relevant for the $\mathcal{H}_N$-norm (26), both kernels have the same leading multiplier $1/(2|k|)$, and their dependence on the window parameter enters only through a relative error of size $O(1/\log N)$. This is the mechanism behind "window universality" at logarithmic scales.

Proof sketch. The leading coefficient $1/(2|k|)$ is the Fourier multiplier of the centered logarithmic kernel $\theta \mapsto \log|1 - e^{i\theta}|$ (up to an irrelevant additive constant). The kernels $\kappa_{K,N}$ and $\eta_{J,N}$ are obtained from this kernel by (a) removing the local singular contribution at scale $\asymp K/N$ or $\asymp J/N$, and (b) adding deterministic normalizations (the constants $C_{K,N}$ and the discrepancy centering) so that the resulting linear statistic is centered. Both operations correspond, in Fourier space, to multiplying by a smooth cutoff and then renormalizing the zero mode. In the logarithmic window band, the cutoff acts only at a scale which is mesoscopic relative to the k-range in (26); the resulting multiplier perturbation is controlled by harmonic-sum

asymptotics, yielding (44). The tail bound (45) follows from the fact that the kernels are bounded and piecewise C^1 after smoothing at scale $\gg 1/N$, so that Fourier coefficients decay at least quadratically for $|k|$ beyond the intrinsic CUE cutoff N. □

7.2 Deterministic $\mathcal{H}_N$-closeness of kernels

We now convert Lemma 7.1 into the desired $\mathcal{H}_N$-norm bounds.

Lemma 7.2 ($\mathcal{H}_N$-closeness in the logarithmic band)**.** *Fix $0 < c_1 < c_2$. Uniformly over integers $K_1, K_2, J \in [c_1 \log N, c_2 \log N]$ we have*

$$\|\kappa_{K_1,N} - \kappa_{K_2,N}\|_{\mathcal{H}_N} \longrightarrow 0, \tag{46}$$

and

$$\|\kappa_{K,N} - \eta_{J,N}\|_{\mathcal{H}_N} \longrightarrow 0, \tag{47}$$

as $N \to \infty$.

Proof. We prove (46); the second claim is identical. By definition (26),

$$\|\kappa_{K_1,N} - \kappa_{K_2,N}\|_{\mathcal{H}_N}^2 = \sum_{k\in\mathbb{Z}\setminus\{0\}} \min(N,|k|)\,\big|\widehat{\kappa}_{K_1,N}(k) - \widehat{\kappa}_{K_2,N}(k)\big|^2.$$

We split into $1 \le |k| \le N$ and $|k| > N$.

For $1 \le |k| \le N$, using (43) we have

$$\widehat{\kappa}_{K_1,N}(k) - \widehat{\kappa}_{K_2,N}(k) = \frac{1}{2|k|}\big(\varepsilon_{K_1,N}(k) - \varepsilon_{K_2,N}(k)\big),$$

hence by (44) and the triangle inequality,

$$\sup_{1\le|k|\le N} \big|\widehat{\kappa}_{K_1,N}(k) - \widehat{\kappa}_{K_2,N}(k)\big| \le \frac{C}{|k|}\Big(\frac{1}{K_1} + \frac{1}{K_2}\Big) \ll \frac{1}{|k|\log N}.$$

Since $\min(N,|k|) = |k|$ in this range, we obtain

$$\sum_{1\le|k|\le N} \min(N,|k|)\,\big|\widehat{\kappa}_{K_1,N}(k) - \widehat{\kappa}_{K_2,N}(k)\big|^2 \ll \sum_{1\le|k|\le N} |k| \cdot \frac{1}{k^2\log^2 N} \ll \frac{1}{\log^2 N}\sum_{k=1}^{N}\frac{1}{k} \ll \frac{1}{\log N}.$$

For $|k| > N$, we use (45) and the crude inequality $\min(N,|k|) = N$ to get

$$\sum_{|k|>N} \min(N,|k|)\,\big|\widehat{\kappa}_{K_1,N}(k) - \widehat{\kappa}_{K_2,N}(k)\big|^2 \le 2N \sum_{|k|>N} \Big(|\widehat{\kappa}_{K_1,N}(k)|^2 + |\widehat{\kappa}_{K_2,N}(k)|^2\Big) \ll N \sum_{|k|>N} \frac{1}{k^4} \ll \frac{1}{N^2}.$$

Combining the two ranges shows that the $\mathcal{H}_N$-norm squared is $O(1/\log N) + O(1/N^2)$ uniformly in K_1, K_2, which implies (46). □

7.3 Verifying the increment condition for Proposition 6.4

To apply Proposition 6.4 we must control not only $\|f_N\|_{\mathcal{H}_N}$ but also $\sup_{|r|\le h_N}\|\Delta_r f_N\|_{\mathcal{H}_N}$ for a suitable mesh $h_N \downarrow 0$. We do this for the kernel differences that arise in the proofs of Theorems A–B.

Let f_N be one of the differences

$$f_N = \kappa_{K_1,N} - \kappa_{K_2,N}, \qquad \text{or} \qquad f_N = \kappa_{K,N} - \eta_{J,N},$$

with parameters in $[c_1 \log N, c_2 \log N]$. The construction of $\kappa_{K,N}$ and $\eta_{J,N}$ in Section 4 gives spectral localization at a mesoscopic cutoff: there exists $M_N \ll N$ (indeed $M_N \asymp N/\log N$ suffices for our purposes) such that $\widehat{f}_N(k) = 0$ for $|k| > M_N$. We now set

$$h_N := \frac{1}{M_N}, \qquad \text{so that} \qquad \log(1/h_N) \asymp \log M_N \asymp \log N.$$

Then by (30),

$$\sup_{|r|\le h_N}\|\Delta_r f_N\|_{\mathcal{H}_N} \le C \sup_{|r|\le h_N} \min(1, M_N|r|)\,\|f_N\|_{\mathcal{H}_N} \le C\,\|f_N\|_{\mathcal{H}_N}.$$

Combining this with Lemma 7.2 (and the explicit rate $\|f_N\|^2_{\mathcal{H}_N} \ll 1/\log N$ coming from the proof) yields

$$\|f_N\|^2_{\mathcal{H}_N}\log(1/h_N) \ll \frac{1}{\log N}\cdot\log N = O(1), \qquad \sup_{|r|\le h_N}\|\Delta_r f_N\|^2_{\mathcal{H}_N}\log(1/h_N) \ll 1,$$

and in fact the same computation with the sharper bound $\|f_N\|^2_{\mathcal{H}_N} = o(1/\log N)$ (which follows from (46)–(47) uniformly) gives the vanishing required in (41). Therefore Proposition 6.4 applies to these differences.

7.4 Consequences for the associated linear statistics

We now record the probabilistic conclusion in the form that will be used to finish the proofs of Theorems A–B.

Proposition 7.3 (Uniform smallness for kernel-difference statistics)**.** *Fix $0 < c_1 < c_2$. Let f_N be either $\kappa_{K_1,N} - \kappa_{K_2,N}$ or $\kappa_{K,N} - \eta_{J,N}$ with all window parameters in $[c_1 \log N, c_2 \log N]$. Then*

$$\max_{1\le n\le N}\Big|\int_0^{2\pi} f_N(\phi - m_n)\,dS_U(\phi) - \mathbb{E}\int_0^{2\pi} f_N(\phi - m_n)\,dS_U(\phi)\Big| = o_{\mathbb{P}}(1).$$

Proof. This is an immediate application of Proposition 6.4 using the verification of (41) above. □

Finally, we explain how Proposition 7.3 is combined with the kernel representations of Lemmas 1–2. For Theorem A we take $f_N = \kappa_{K_1,N} - \kappa_{K_2,N}$ and write

$$\log A_{K_1}(m_n) - \log A_{K_2}(m_n) = \int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) + \big(r_{K_1,N}(n) - r_{K_2,N}(n)\big),$$

where the remainders satisfy $\max_n |r_{K_i,N}(n)| = o_{\mathbb{P}}(1)$ by Lemma 1. Proposition 7.3 controls the centered linear-statistic term uniformly over n, and rotation invariance shows that the deterministic mean is independent of n; thus the entire right-hand side is $o_{\mathbb{P}}(1)$ uniformly in n, yielding the desired sup-norm window universality.

For Theorem B we take $f_N = \kappa_{K,N} - \eta_{J,N}$ and similarly combine Lemmas 1–2:

$$\log A_K(m_n) - D_J(m_n) = \int_0^{2\pi} f_N(\phi - m_n)\, dS_U(\phi) + \big(r_{K,N}(n) - q_{J,N}(n)\big),$$

with $\max_n |r_{K,N}(n)| + \max_n |q_{J,N}(n)| = o_{\mathbb{P}}(1)$. Again Proposition 7.3 yields uniform $o_{\mathbb{P}}(1)$ control of the centered integral term, and the remaining deterministic mean is negligible because f_N has $\mathcal{H}_N$-norm tending to zero (hence, in particular, its low-frequency mass vanishes). This completes the analytic comparison step: once the estimators are expressed as smoothed versions of the same centered counting function, the logarithmic window band is precisely the regime in which the smoothing kernels become asymptotically indistinguishable in the variance-controlling norm.

8 Numerical illustrations and sanity checks (optional)

We record a small collection of numerical experiments that serve two purposes. First, they provide a sanity check that the quantities introduced above can be computed stably at moderate sizes, and that the midpoint restriction indeed removes the dominant singular behaviour. Second, they illustrate qualitatively the two conclusions proved in this note: window universality for $\log A_K$ when $K \asymp \log N$, and equivalence between $\log A_K$ and D_J in the same regime. None of the computations below is used in the proofs.

8.1 Sampling and basic preprocessing

We generate $U \sim \mathrm{Haar}(\mathrm{U}(N))$ using a standard QR method: if G has i.i.d. standard complex Gaussians, write $G = QR$ with Q unitary and R upper triangular, and normalize the diagonal of R to have positive real part (equivalently, multiply Q on the right by a diagonal unitary). We then compute eigenvalues $\lambda_j = e^{i\theta_j}$ and sort eigenangles $\theta_1 < \cdots < \theta_N$ in $[0, 2\pi)$. The midpoints are

$$m_n = \frac{\theta_n + \theta_{n+1}}{2}, \qquad 1 \le n \le N,$$

with $\theta_{N+1} := \theta_1 + 2\pi$. Numerically, it is convenient to store an extended array $\{\theta_k\}_{k\in\mathbb{Z}}$ via $\theta_{k+N} = \theta_k + 2\pi$ to avoid case distinctions near the endpoints.

We compute $Z_U(\theta)$ through the product representation

$$|Z_U(\theta)| = \prod_{j=1}^{N} |e^{i\theta} - e^{i\theta_j}| = 2^N \prod_{j=1}^{N} \Big|\sin\Big(\frac{\theta-\theta_j}{2}\Big)\Big|.$$

Since we only use $\log|Z_U(\theta)|$, we evaluate

$$\log|Z_U(\theta)| = N\log 2 + \sum_{j=1}^{N} \log\Big|\sin\Big(\frac{\theta-\theta_j}{2}\Big)\Big|,$$

with the summation performed in double precision. The midpoint grid has the practical advantage that the summands are uniformly away from $-\infty$; empirically one observes that even for N in the low thousands, the minimal distance from m_n to the nearest eigenangle is typically of order $1/N$, in agreement with Lemma 5.

8.2 Computing $\log A_K(m_n)$ and centering

For a fixed integer $K \geq 1$, we compute

$$\log A_K(m_n) = \log|Z_U(m_n)| - \sum_{j=0}^{K-1} \Big(\log|e^{im_n} - e^{i\theta_{n-j}}| + \log|e^{im_n} - e^{i\theta_{n+1+j}}|\Big) + C_{K,N}.$$

We emphasize that $\log|Z_U(m_n)|$ already contains the near-neighbour logarithms; the subtraction above is therefore a difference of quantities with comparable sizes. Numerically this is not problematic in the present regime because the subtracted terms are explicitly singled out and evaluated at the same precision. In practice we rewrite the expression using the sine form,

$$\log|e^{ix} - e^{iy}| = \log 2 + \log\Big|\sin\Big(\frac{x-y}{2}\Big)\Big|,$$

so that additive constants can be collected and treated deterministically.

The normalization constant $C_{K,N}$ is defined in the paper by the centering condition $\mathbb{E}[\log A_K(m_1)] = 0$. In simulation we implement either of the following proxies.

1. For each pair (N, K) we precompute an empirical mean over a modest number of independent samples (e.g. 50–200) and take this as $-C_{K,N}$; we then reuse the same $C_{K,N}$ for subsequent runs.

2. Alternatively, within a single sample we define a sample-centered version

$$\log A_K^{\mathrm{sc}}(m_n) := \log A_K(m_n) - \frac{1}{N}\sum_{r=1}^{N} \log A_K(m_r),$$

which removes the global additive mode and is appropriate when we only compare fields up to constants.

Both choices lead to the same qualitative conclusions about window dependence and about comparison with D_J, since all statements of interest concern differences either across K or across estimators, and these differences are insensitive to adding deterministic constants.

8.3 Computing $D_J(m_n)$ from eigenangles

Given $\{\theta_j\}$ we compute the discrepancies

$$\delta_n(j) = \frac{N}{2\pi}(\theta_{n+1+j} - \theta_{n-j}) - (2j+1)$$

for $j \geq 1$ and then the weighted sum

$$D_J(m_n) = \sum_{j=J}^{\lfloor N/4 \rfloor} \frac{\delta_n(j)}{j}.$$

As written, $D_J(m_n)$ can be computed in $O(N^2)$ operations if done naively for all n. For the moderate sizes used here ($N \leq 10^4$) this is still feasible in optimized code, but for comfort we apply two simple improvements. First, we exploit the periodic extension and precompute the spans $\theta_{n+\ell} - \theta_{n-\ell'}$ as array differences. Second, we note that for fixed j the map $n \mapsto \theta_{n+1+j} - \theta_{n-j}$ is just a shift of indices; thus the dominant cost is the summation over j, which can be vectorized.

We also monitor the contribution of the upper tail $j \in [\alpha N, \lfloor N/4 \rfloor]$ for fixed $\alpha > 0$ to confirm the robustness statement in Proposition D: empirically, for α as small as 0.05 the tail has small variation in n and contributes primarily an n-independent offset, consistent with the cancellation suggested by the kernel viewpoint.

8.4 Window universality for $\log A_K$ at $K \asymp \log N$

We fix $N \in \{10^3, 3 \times 10^3, 10^4\}$ and choose window sizes

$$K_-(N) = \lfloor 0.5 \log N \rfloor, \qquad K_0(N) = \lfloor 1.0 \log N \rfloor, \qquad K_+(N) = \lfloor 2.0 \log N \rfloor,$$

as representative points in a logarithmic band. For each sample we compute the fields $\log A_{K_-}(m_n)$, $\log A_{K_0}(m_n)$, $\log A_{K_+}(m_n)$ (sample-centered), and then form the sup-norm differences

$$\Delta^{(\pm)} := \max_{1 \leq n \leq N} \left| \log A^{\mathrm{sc}}_{K_\pm}(m_n) - \log A^{\mathrm{sc}}_{K_0}(m_n) \right|.$$

A typical outcome is that $\Delta^{(\pm)}$ decreases slowly with N (consistent with an $o_{\mathbb{P}}(1)$ statement without an optimized rate), while remaining uniformly

small compared to the standard deviation of $\log|Z_U(m_n)|$. Visually, plotting $n \mapsto \log A_K^{\mathrm{sc}}(m_n)$ for the three window choices yields curves that are almost indistinguishable, with differences appearing as low-amplitude fluctuations.

We also evaluate a more stringent statistic that is sensitive to pointwise alignment:

$$\rho(K_1, K_2) := \frac{\sum_{n=1}^{N} \big(\log A_{K_1}^{\mathrm{sc}}(m_n)\big)\big(\log A_{K_2}^{\mathrm{sc}}(m_n)\big)}{\Big(\sum_{n=1}^{N} (\log A_{K_1}^{\mathrm{sc}}(m_n))^2\Big)^{1/2} \Big(\sum_{n=1}^{N} (\log A_{K_2}^{\mathrm{sc}}(m_n))^2\Big)^{1/2}}.$$

In the log-window band this correlation is typically extremely close to 1 already at $N = 10^3$, which is consistent with the assertion that different K choices correspond to kernels that differ little in the variance-controlling norm.

8.5 Equivalence between $\log A_K$ and D_J

We next fix $K = K_0(N)$ and choose $J = \lfloor \log N \rfloor$ (and optionally $J = \lfloor 2 \log N \rfloor$) and compare $\log A_K(m_n)$ to $D_J(m_n)$. Since the centering conventions in the two definitions differ, we compare sample-centered versions,

$$D_J^{\mathrm{sc}}(m_n) := D_J(m_n) - \frac{1}{N} \sum_{r=1}^{N} D_J(m_r).$$

We then examine

$$\Delta^{(A,D)} := \max_{1 \le n \le N} \big| \log A_K^{\mathrm{sc}}(m_n) - D_J^{\mathrm{sc}}(m_n) \big|.$$

For the same values of N as above, the resulting $\Delta^{(A,D)}$ is of the same order as the window-to-window differences $\Delta^{(\pm)}$ and decreases with N. Pointwise plots show that $\log A_K^{\mathrm{sc}}(m_n)$ and $D_J^{\mathrm{sc}}(m_n)$ track each other across the entire circle; their difference appears as a small-amplitude, more oscillatory component, which is consistent with the interpretation that the two observables arise from slightly different smoothings of the same centered counting function.

A further diagnostic is to compare extrema. Define

$$M_A := \max_{1 \le n \le N} \log A_K^{\mathrm{sc}}(m_n), \qquad M_D := \max_{1 \le n \le N} D_J^{\mathrm{sc}}(m_n).$$

The observed difference $|M_A - M_D|$ is typically much smaller than either M_A or M_D, suggesting that the equivalence is strong enough to preserve the location and height of large carrier-wave peaks at these N. This is consistent with the sup-norm statement in Theorem B.

8.6 What happens outside the logarithmic band

To illustrate the necessity of the condition $K \asymp \log N$, we repeat the window comparison with extreme choices, for instance

$$K_{\mathrm{micro}} := \lfloor (\log N)^{1/2} \rfloor, \qquad K_{\mathrm{macro}} := \lfloor N^{0.1} \rfloor,$$

and compare $\log A_{K_{\mathrm{micro}}}$ and $\log A_{K_{\mathrm{macro}}}$ to $\log A_{K_0}$. The micro-window estimator retains a visible dependence on K, with fluctuations that are localized near the eigenangles whose contributions have not been adequately averaged; in particular, the difference $\log A_{K_{\mathrm{micro}}} - \log A_{K_0}$ is not small uniformly in n for these N. At the other extreme, the macro-window estimator oversmooths and begins to incorporate global effects, producing a field with noticeably different long-range behaviour. These computations match the heuristic that the intrinsic mesoscopic correlation scale is logarithmic, and that window universality is not expected far outside this regime.

8.7 Finite-N effects and practical remarks

Two finite-N effects recur across experiments. First, the dependence on the centering convention is more visible at smaller N, especially for D_J, where the upper tail $j \approx N/4$ can contribute a non-negligible constant offset. Sample-centering largely removes this. Second, rare samples may contain an unusually small gap $\theta_{n+1} - \theta_n$, leading to a relatively large magnitude of $\log |Z_U(m_n)|$ and therefore larger fluctuations in $\log A_K(m_n)$; nevertheless, because the estimator is designed to factor out the nearest zeros explicitly, these rare events do not produce numerical instabilities at the midpoints.

Finally, we note that the computations above are consistent with the analytic picture developed in Sections 4–7: after removing the singular local contributions, both $\log A_K$ and D_J behave like linear statistics of the same log-kernel against the centered counting function. The simulations provide a concrete visualization of this statement, but they remain purely supportive and are not invoked in any argument.

9 Discussion and extensions

We collect several directions suggested by the window-universality and estimator-equivalence statements proved above. The common theme is that, once one has identified a canonical "carrier-wave" observable on a logarithmic window, one may attempt to (a) construct a genuine limiting random field and its multiplicative chaos, (b) understand conditional laws of the eigenangle process under atypical carrier-wave events, and (c) transport the analytic structure to arithmetically defined analogues. We also indicate how much of the discussion is specific to the determinantal structure of CUE and how much should persist for other circular ensembles.

9.1 Toward a canonical limiting field (GFF/GMC viewpoint)

Theorems A–B show that on the midpoint grid the carrier wave is, up to $o_{\mathbb{P}}(1)$ in sup norm, insensitive to the choice of a window parameter in the band $K, J \asymp \log N$. This makes it natural to treat $\mathrm{CW}_N(m_n)$ as a canonical random field indexed by $n \in \{1, \dots, N\}$, and to ask whether CW_N admits a scaling limit in a function space after a suitable interpolation from the discrete grid to $\mathbb{S}^1$.

The kernel representations (Lemmas 1–2) suggest the following guiding principle. Each admissible estimator is a linear statistic of the centered counting function S_U against a kernel that approximates the logarithmic potential on the circle, with an ultraviolet cutoff at frequency about $N/\log N$ (up to constants). In particular, for angles separated by more than the microscopic scale $1/N$ but not macroscopic, the covariance structure should be close to that of the circle Gaussian free field (GFF),

$$\mathrm{Cov}\big(G(\theta), G(\phi)\big) \approx -\log\left|e^{i\theta} - e^{i\phi}\right| + \text{constant},$$

with the understanding that both sides require regularization at short distances. In the CUE setting, this heuristic is classical for $\log|Z_U(\theta)|$ and for smoothed linear statistics; what is more specific here is that the particular regularization induced by *factoring out the nearest zeros* (or, equivalently, by *starting discrepancy sums at* $J \asymp \log N$) should be asymptotically canonical.

A concrete program is to define a continuous version $\mathrm{CW}_N(\theta)$ by convolution of S_U with one fixed admissible kernel (for instance, $\kappa_{K,N}$ with $K = \lfloor c \log N \rfloor$) and to prove finite-dimensional convergence

$$\big(\mathrm{CW}_N(\theta_1), \dots, \mathrm{CW}_N(\theta_k)\big) \;\Rightarrow\; \big(G(\theta_1), \dots, G(\theta_k)\big),$$

for distinct angles $\theta_1, \dots, \theta_k$ (or for angles separated at least by $N^{-1+\varepsilon}$), with G a circle GFF with an appropriate additive normalization. Since the construction already removes the most singular local terms, one expects fewer technical issues near eigenangles than for $\log|Z_U(\theta)|$ itself. The remaining difficulties are of a standard type: controlling the difference between the actual kernels $\kappa_{K,N}$ and an ideal log kernel, and proving a central limit theorem uniform over a sufficiently rich class of test functions.

Beyond the Gaussian limit, one is led to Gaussian multiplicative chaos (GMC). Formally, if CW_N approximates a log-correlated Gaussian field, then for $\gamma \in (0, 2)$ one expects the random measures

$$\mu_{N,\gamma}(d\theta) \;:=\; \exp\Big(\gamma\,\mathrm{CW}_N(\theta) - \tfrac{\gamma^2}{2}\mathrm{Var}(\mathrm{CW}_N(\theta))\Big)\,\frac{d\theta}{2\pi}$$

to converge in law (after choosing a regularization/normalization) to the subcritical GMC associated to the circle GFF. A virtue of the window-universality statement is that it suggests $\mu_{N,\gamma}$ should not depend on the

specific admissible choice of K or J, at least at the level of convergence in distribution. We emphasize that making this canonicality rigorous at the measure level is substantially stronger than the midpoint sup-norm statements: one needs tightness in a space of measures and quantitative control of the error when exponentiating. Nevertheless, the kernel form and the stability in the $\mathcal{H}$-norm provide the natural starting point for such an analysis.

Finally, once one has a canonical limiting field or chaos, one may revisit extremes. The maximum of a log-correlated field is sensitive to the ultraviolet cutoff, and the logarithmic window $K \asymp \log N$ corresponds to a specific choice of cutoff scale between microscopic and macroscopic. Window universality therefore supports the idea that the *second-order* structure of the maximum (including the additive constant in the $-\frac{3}{4}\log\log N$ shift familiar from log-correlated models) should be an intrinsic object rather than an artefact of the estimator definition. Establishing this remains open in the present framework, but the equivalence $\log A_K \approx D_J$ suggests one may choose whichever representation is technically more convenient in an extremal analysis.

9.2 Conditional laws given large carrier waves

A second direction is to study the eigenangle process under conditioning on atypically large values of the carrier wave, either at a point or at a global maximum. The determinantal nature of CUE makes it plausible that such conditional laws can be described by explicit tilts, at least approximately, and the kernel representations provide a clear candidate for the tilt functional.

At a heuristic level, if $\mathrm{CW}_N(\theta)$ is close to a Gaussian field $G(\theta)$, then conditioning on a large value $G(\theta_0) = h$ produces a mean shift proportional to the covariance kernel:

$$\mathbb{E}\big[G(\theta) \,\big|\, G(\theta_0) = h\big] \;\approx\; \frac{\mathrm{Cov}(G(\theta), G(\theta_0))}{\mathrm{Var}(G(\theta_0))}\, h \;\approx\; -\frac{h}{\mathrm{Var}(G(\theta_0))} \log|e^{i\theta} - e^{i\theta_0}|.$$

Translating this into eigenangles, one expects that conditioning on a large carrier wave at θ_0 induces a mild but extended reweighting of the local density around θ_0 across mesoscopic scales, consistent with the interpretation of the carrier wave as a smoothed log potential of the empirical measure relative to uniform.

A more robust formulation is to study exponential tilts rather than hard conditioning. For $\beta \in \mathbb{R}$ fixed, consider the tilted law

$$\frac{d\mathbb{P}^{(\beta)}}{d\mathbb{P}}(U) \;\propto\; \exp\big(\beta\, \mathrm{CW}_N(\theta_0)\big),$$

with CW_N represented by any admissible estimator. Under the kernel form $\mathrm{CW}_N(\theta_0) \approx \int \kappa(\phi - \theta_0)\, dS_U(\phi)$, such a tilt becomes an exponential of a

linear statistic. In determinantal processes, exponential tilts by linear statistics are naturally connected to multiplicative perturbations of the kernel (or, in the language of orthogonal polynomials, to modifications of the weight). One may therefore hope to identify the tilted eigenangle process as an approximately determinantal process with a deformed symbol, at least in a mesoscopic scaling regime.

One concrete question is the typical gap structure near θ_0 under such a tilt: does a large carrier wave correspond to a local "compression" of eigenangles (higher-than-average local density), or does it arise primarily from longer-range fluctuations while the microscopic sine-kernel behavior persists? The discrepancy representation D_J suggests that $\mathrm{CW}_N(\theta_0)$ is dominated by contributions from spans $\theta_{n+1+j} - \theta_{n-j}$ with $j \geq J \asymp \log N$, hence from scales larger than microscopic. This points toward the second possibility: the microscopic spacing statistics should remain close to sine-kernel, while mesoscopic density deviations provide the main contribution to a large carrier wave. Proving such a statement would amount to a conditional universality result, separating micro- and meso-scales.

A related problem is the geometry of near-maximizers. If one believes the GFF/GMC picture, then points where CW_N is large should resemble thick points of a log-correlated field, and the set of near-maxima should have a nontrivial random structure. On the matrix side, it would be of interest to understand whether near-maximizers of CW_N correlate with atypical local gap configurations (e.g. unusually small gaps) or are primarily determined by longer-range fluctuations in S_U.

9.3 Transfer to ζ and arithmetic damping

The constructions in this note were motivated in part by analogies between CUE characteristic polynomials and the Riemann zeta function on the critical line. In the zeta setting, one has two interacting sources of complexity: (i) the zeros are not exactly a determinantal process (even conjecturally, one expects only local GUE-type statistics), and (ii) the analogue of $\log|Z_U|$ contains arithmetic contributions from primes. Nevertheless, the present framework suggests a way to phrase "carrier-wave" objects for ζ in a manner that isolates a canonical mesoscopic component.

A natural starting point is the explicit formula and its smoothed variants. On RH, one can represent (after smoothing) $\log|\zeta(\frac{1}{2}+it)|$ as a linear statistic of zeros against a logarithmic kernel plus a prime sum. The role played here by the centered counting function S_U is played by the centered zero-counting function $S(t)$, and the logarithmic kernel is again the fundamental object. The analogue of the midpoint restriction is to evaluate at points t that are safely away from zeros (or to work with a smoothed version of $\log|\zeta|$), since singularities at zeros are unavoidable.

From this viewpoint, there are two plausible "damping" operations that

mirror our windowing in CUE:

1. *Zero damping*: factor out or subtract the contributions of the nearest zeros to remove the dominant local singularity, producing an analogue of $\log A_K$.

2. *Arithmetic damping*: mollify or smooth in t (or damp prime sums) so that the prime contribution becomes either deterministic at the relevant scale or an independent Gaussian input.

The second operation is specific to ζ and has no analogue in CUE. One expects that, after suitable damping, the remaining field is closer to a log-correlated Gaussian model with the same mesoscopic covariance as in random matrix theory, but with an additional arithmetic shift in the mean and possibly in the variance. In practice, this suggests defining a zeta-carrier-wave estimator by combining (a) a discrepancy sum over zeros analogous to D_J (starting at a mesoscopic $J \asymp \log\log T$) and (b) an explicit subtraction of a short prime sum, or equivalently working with a mollified ζ.

The chief technical obstacle is that our proofs rely on determinantal concentration for linear statistics, whereas for zeros one would need uniform control of linear statistics of $S(t)$ over a net of points t. Conditional on strong zero-density and pair-correlation input (and perhaps higher correlation estimates), it is plausible that one could mimic the kernel-comparison strategy: show that different window choices correspond to nearby Fourier multipliers, and then show that the induced difference in the linear statistics is small with high probability. Even without a full proof, this viewpoint may be useful in formulating a precise conjecture: that there exists a canonical carrier-wave field for ζ at height T, well-defined up to $o(1)$ in a suitable sup norm over a discrete mesh, and stable under admissible windowing choices. Such a conjecture would provide a clean target for future analytic work, and it would clarify which aspects of the random-matrix heuristics are purely log-kernel phenomena and which require detailed determinantal input.

9.4 Variants for the circular CβE

Finally, we comment on replacing CUE ($\beta = 2$) by the circular CβE for general $\beta > 0$. At the level of definitions, the eigenangles $\{\theta_j\}$ still form a stationary point process on the circle with logarithmic repulsion. The objects m_n and the discrepancy sums $\delta_n(j)$ and $D_J(m_n)$ remain meaningful without change. One may also define $Z(\theta) = \prod_{j=1}^{N} |e^{i\theta} - e^{i\theta_j}|$ and mimic A_K by factoring out nearby terms; the only adjustment is that for $\beta \neq 2$ there is no underlying matrix U with a genuine characteristic polynomial, so Z should be treated as the absolute value of a log-gas potential rather than as an analytic object.

On the probabilistic side, the determinantal tools used for CUE are unavailable for $\beta \neq 2$, but there is a substantial substitute literature: rigidity

estimates for log-gases, loop equations, and central limit theorems for linear statistics with test functions in Sobolev spaces. In particular, for sufficiently regular f one has

$$\mathrm{Var}\Big(\int f(\phi)\, dS(\phi) \Big) \;\approx\; \frac{2}{\beta}\, \|f\|^2_{H^{1/2}},$$

with the same $H^{1/2}$ structure as in the determinantal case but with a β-dependent prefactor. This suggests that Lemmas 1–3 should persist with minor modifications, while Lemma 4 (uniform concentration over a grid) should be approached via rigidity plus union bounds rather than via determinantal concentration inequalities.

If such uniform control can be established in the logarithmic window band, then the same strategy should prove window universality and estimator equivalence for CβE. From the conceptual viewpoint, this would reinforce the idea that the logarithmic window is not an artefact of $\beta = 2$ integrability, but rather an intrinsic mesoscopic scale for log-gases on the circle. It would also align with the expectation that the limiting log-correlated field is a β-dependent multiple of the circle GFF, hence that the associated GMC measures depend on β only through the overall variance normalization.

We do not attempt to pursue these extensions here. The statements above indicate, however, that the present midpoint-based formalism is flexible: it isolates a logarithmic kernel structure that is stable under windowing and that should be accessible beyond the determinantal setting, provided one can supply the appropriate concentration input for mesoscopic linear statistics.