Journal of Physics A: Mathematical and Theoretical

PAPER

Hidden multiscale order in the primes

To cite this article: Salvatore Torquato *et al* 2019 *J. Phys. A: Math. Theor.* **52** 135002

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **52** (2019) 135002 (33pp) https://doi.org/10.1088/1751-8121/ab0588

Hidden multiscale order in the primes

Salvatore Torquato[1,4], **Ge Zhang**[2] **and Matthew De Courcy-Ireland**[3]

[1] Department of Chemistry, Department of Physics, Princeton Institute for the Science and Technology of Materials, and Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, United States of America

[2] Department of Chemistry, Princeton University, Princeton, NJ 08544, United States of America

[3] Department of Mathematics, Princeton University, Princeton, NJ 08544, United States of America

E-mail: torquato@princeton.edu

Received 28 July 2018, revised 12 January 2019
Accepted for publication 8 February 2019
Published 1 March 2019

CrossMark

Abstract

We study the pair correlations between prime numbers in an interval $M \leqslant p \leqslant M+L$ with $M \to \infty$, $L/M \to \beta > 0$. By analyzing the *structure factor*, we prove, conditionally on the Hardy–Littlewood conjecture on prime pairs, that the primes are characterized by unanticipated multiscale order. Specifically, their limiting structure factor is that of a union of an infinite number of periodic systems and is characterized by a dense set of Dirac delta functions (Bragg peaks), similar to but different from the dense Bragg peaks that arise in quasicrystals and standard limit-periodic systems. Primes in dyadic intervals are the first examples of what we call *effectively limit-periodic* point configurations. This behavior implies anomalously suppressed density fluctuations compared to uncorrelated (Poisson) systems at large length scales, which is now known as hyperuniformity. Using a scalar order metric τ calculated from the structure factor, we identify a transition between the order exhibited when L is comparable to M and the uncorrelated behavior when L is only logarithmic in M. Our analysis of the structure factor leads to an algorithm to reconstruct primes in a dyadic interval with high accuracy. The discovery of the hyperuniformity and effective limit-periodic behavior of the primes provide new organizing principles to understand the fundamental nature of patterns in the primes.

Keywords: structure factor, prime numbers, hyperuniformity, point diffraction, order metrics, limit periodicity

[4] Author to whom any correspondence should be addressed.

S Supplementary material for this article is available online

(Some figures may appear in colour only in the online journal)

1. Introduction

While prime numbers are deterministic, by some probabilistic descriptors, they can be regarded as pseudo-random in nature. Indeed, the primes can be difficult to distinguish from a random configuration of the same density. For example, assuming a plausible conjecture, Gallagher proved that the gaps between primes follow a Poisson distribution [1]. Thus the number of $M \leqslant X$ such that there are exactly N primes in the interval $M < p < M + L$ of length $L \sim \lambda \ln X$ is given asymptotically by

$$\#\{M \leqslant X \text{ such that } \pi(M+L) - \pi(M) = N\} \sim X \frac{\mathrm{e}^{-\lambda}\lambda^N}{N!}.$$

Note that Gallagher's interest was in short intervals whereas we analyze those in which the length L is comparable to the lower endpoint M. Our primary observation is that for these longer intervals, the primes are highly correlated and ordered on multiple length scales and hence are drastically different from a Poisson distribution. This is demonstrated by the identification of dense Bragg (Dirac delta function) peaks in the *structure factor* of the primes and by large values of the order parameter τ, both of which we define in section 2. In particular, we use the structure factor to detect a large-scale order known as hyperuniformity [2], very different from the uncorrelated behavior one sees in short intervals.

To study different ranges of primes, it is important to take account of the fact that primes become increasingly sparse in longer intervals. Let $\pi(x)$ denote the *prime counting function*, which gives the number of primes less than x. According to the prime number theorem [3], the prime counting function in the large-x asymptotic limit is given by

$$\pi(x) \sim \frac{x}{\ln(x)} \qquad (x \to \infty). \tag{1}$$

The prime number theorem means that for sufficiently large x, the probability that a randomly selected integer not greater than x is prime is very close to $1/\ln(x)$, which can be viewed as a position-dependent number density $\rho(x)$ (number of primes up to x divided by the interval x). This implies that the primes become sparser as x increases and hence constitute a *statistically inhomogeneous* set of points that are located on a subset of the odd integers (for any prime number greater than 2). This simple observation requires that one carefully choose the interval over which the primes are sampled and characterized in order to obtain meaningful results that in general will depend on the chosen interval. If L is much larger than M, the density $1/\ln(x)$ drops off appreciably as x ranges from M to $M + L$, and then the system is the very opposite of hyperuniform. On the other hand, $\ln(M + L) = \ln(M) + \ln(1 + L/M)$ is asymptotic to $\ln(M)$ as long as L/M is bounded above. In this case, one can treat the primes as homogeneous with constant density $1/\ln(M)$. For this paper, we take $L \sim \beta M$ of the same order as M or sometimes smaller to compare with Gallagher's regime.

The plausible conjecture Gallagher assumed is a version of the Hardy–Littlewood m-tuples conjecture (theorem X1, p 61 of [4]). If $\mathcal{H} = (h_1, \ldots, h_m)$ is a m-tuple of integers, then the conjecture gives the number of $n \leqslant X$ such that all of the shifts $n + h_1, \ldots, n + h_m$ are prime as

$$\#\{n \leqslant X \text{ such that } n + h_j \text{ all prime}\} \sim \mathfrak{S}(\mathcal{H}) \frac{X}{(\ln X)^m}, \tag{2}$$

where

$$\mathfrak{S}(\mathcal{H}) = \prod_{p} \left(1 - \frac{1}{p}\right)^{-m} \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right) \tag{3}$$

$$\nu_{\mathcal{H}}(p) = \#\left\{\text{distinct } h_j \bmod p\right\}. \tag{4}$$

When $m = 1$, say $\mathcal{H} = \{h_1\}$, (2) simply counts primes less than X (or, strictly, less than $X - h_1$). In (3), since every $\nu(p)$ is 1, one has $\mathfrak{S}(\mathcal{H}) = 1$. Thus the case $m = 1$ is the prime number theorem, and it is the only one so far to be proved. Gallagher used all values of m in order to compare the empirical moments of primes in short intervals with the moments of the Poisson distribution.

Our study of the structure factor ultimately leads to an equivalent formulation of the case $m = 2$. The Hardy–Littlewood constant $\mathfrak{S}(\mathcal{H})$ can be understood as a correction to the prediction one would make by imagining that all of the shifts $n + h_j$ are prime independently with probability $1/\ln(X)$; see [5] and, for this and other senses in which the random model fails [6]. To summarize the interpretation, note that, for each p, $(1 - 1/p)^m$ is the naive chance that each of the shifts would be indivisible by p. However, these constraints are not independent, and $\nu_{\mathcal{H}}(p)$ is exactly the number of residue classes modulo p which n must avoid or else one of the numbers $n + h_j$ would have p as a factor. Thus $\mathfrak{S}(\mathcal{H})$ cancels the incorrect guess $(1 - 1/p)^m$ and replaces it with the correct $(p - \nu(\mathcal{H}))/p$. The argument advanced by Hardy–Littlewood is however of an altogether different nature, which we outline in the supplementary material for purposes of comparison for the interested reader (stacks.iop.org/JPhysA/52/135002/mmedia).

Probabilistic methods to treat the primes have yielded fruitful insights about them [7]. Furthermore, there are computationally quick *stochastic* ways to find large primes [8–12]. On the other hand, it is known that primes contain unusual patterns, and hence their distribution is not purely random. Chebyshev observed (circa 1853) that primes congruent to 3 modulo 4 seem to predominate over those congruent to 1. Assuming a generalized Riemann hypothesis, Rubinstein and Sarnak [13] exactly characterized this phenomenon and more general related results. A computational study on the Goldbach conjecture demonstrates a connection based on a modulo 3 geometry between the set of even integers and the set of primes [14]. In 1934, Vinogradov proved that every sufficiently large odd integer is the sum of three primes [15]. This method has been extended to cover many other types of patterns [16–19]. Recently it has been shown that there are infinitely many pairs of primes with some finite gap [20] and that primes with decimal expansion ending in 1 are less likely to be followed by another prime ending in 1 [21]. There is numerical evidence for patterns in the distribution of gaps between primes when these are divided into congruence families [22–24].

The present paper is motivated by certain remarkable global properties of the nontrivial zeros of the Riemann zeta function $\zeta(s)$, which is a function of a complex variable s that is intimately related to the primes. The zeta function has many different representations, one of which is the well-known series formula

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \tag{5}$$

which converges for $\mathrm{Re}(s) > 1$. However, $\zeta(s)$ has a unique analytic continuation to the entire complex plane, excluding the simple pole at $s = 1$. According to the *Riemann hypothesis*, the nontrivial zeros of the zeta function lie along the *critical line* $s = 1/2 + it$ with $t \in \mathbb{R}$ in the complex plane and hence form a one-dimensional point process. The nontrivial zeros tend

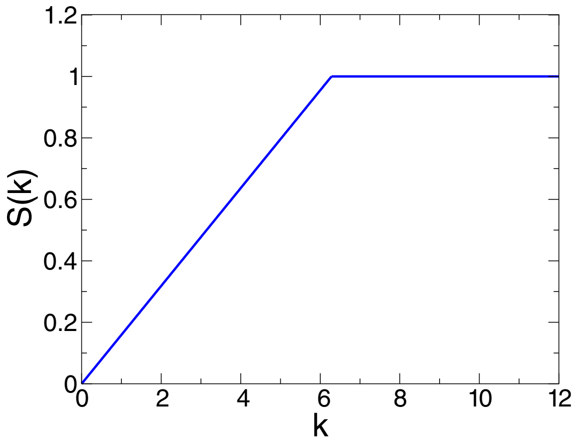

Figure 1. Structure factor of the normalized nontrivial zeros of the Riemann zeta function as a function of the wavenumber (*see* (6)). This falls in a special class of hyperuniform point configurations [30] called class II [43].

to get denser the higher on the critical line. When the spacings of the zeros are appropriately normalized so that they can be treated as a homogeneous point process at unity density in $\mathbb{R}$, the resulting pair correlation function $g_2(\mathbf{r})$ (equal to $g_2(-\mathbf{r})$) takes on the simple form $1 - \sin^2(\pi r)/(\pi r)^2$ [25], where $r \equiv |\mathbf{r}|$. This has consequences for the distribution of the primes in short intervals [26]. The corresponding structure factor $S(\mathbf{k})$ in $\mathbb{R}$ (essentially the Fourier transform of $g_2(\mathbf{r})$) is given by

$$S(\mathbf{k}) = \begin{cases} \frac{k}{2\pi}, & 0 \leqslant k \leqslant 2\pi \\ 1, & k > 2\pi, \end{cases} \tag{6}$$

where $k = |\mathbf{k}|$ is the wavenumber (modulus of the wavevector $\mathbf{k}$) and $S(\mathbf{k}) = S(-\mathbf{k})$. Remarkably, formula (6) exactly matches the structure factor of the eigenvalues of a random matrix in the Gaussian unitary ensemble [27–29]; see figure 1. We see that the structure factor goes to zero linearly in k as the wavenumber goes to zero and is equal to unity for $k > 2\pi$. This implies that the normalized Riemann zeros possess an unusual type of correlated disorder at large length scales known as *hyperuniformity* [2, 30]. A hyperuniform point configuration in d-dimensional Euclidean space $\mathbb{R}^d$ is one in which $S(\mathbf{k})$ tends to zero as the wavenumber k tends to zero [2]. In such systems, density fluctuations are anomalously suppressed at very large length scales, a 'hidden' order that imposes strong global structural constraints. All structurally perfect crystals and quasicrystals are hyperuniform, but typical disordered many-particle systems, including gases, liquids, and glasses, are not. Disordered hyperuniform many-particle systems are exotic states of amorphous matter that have attracted considerable recent attention [30–43].

Because information about the primes can in principle be deduced from information about the nontrivial zeros of the zeta function via explicit formulas [44–46], one might expect the primes to encode hyperuniform correlations seen in the Riemann zeros. For example, von Mangoldt's explicit formula for a weighted counting function $\psi(x) = \sum_{p^n<x} \ln(p)$ is given by

$$\psi(x) = x - \sum_{s} \frac{x^{1/2+\mathrm{i}\gamma}}{\frac{1}{2}+\mathrm{i}\gamma} - \frac{1}{2}\ln(1-x^{-2}) - \ln(2\pi), \tag{7}$$

for $x > 1$ and x not a prime or prime power (where $\psi(x)$ would have a jump discontinuity). Here, $1/2 + \mathrm{i}\gamma$ denotes a nontrivial zero of $\zeta(s)$, meaning that it lies in the critical strip $0 < \mathrm{Re}(s) < 1$. Assuming the Riemann hypothesis, γ is real and the zeros thus form a one-dimensional point process. In any case, allowing for complex γ, the explicit formula applies unconditionally. The trivial zeros $-2, -4, -6, \ldots$ contribute $\ln(1 - x^{-2})$. The explicit formula may be thought of as a Mellin transform of the prime numbers. The structure factor $S(k)$, which is the basis of our investigation, is also a Fourier-type transform of the primes (without any weights) but in a more direct sense:

$$S(k) = \frac{1}{N}\left|\sum_{p} \mathrm{e}^{-\mathrm{i}kp}\right|^2, \tag{8}$$

where p runs over the primes in the interval $[M, M+L]$, the number of which we denote by N. A weighted version of the inner sum, which weights each prime p and also its higher powers p^l by $\ln p$, has been much studied in connection with the circle method (see section 25 of [44], for example). The behavior of $S(k)$ for small values of k reflects the large-scale correlations between primes.

In a very recent numerical study [47], two of us and Martelli adapted a physics-based approach by examining the pair statistics of the primes, especially the structure factor $S(k)$, in an interval $M \leqslant p \leqslant M+L$ with M and L large such that L/M is a positive constant smaller than unity. The simulations strongly suggest that the structure factor exhibits many well-defined Dirac-delta-function (Bragg-like) peaks along with a small 'diffuse' contribution; see figure 2. This means that the primes are characterized by a substantial amount of order on many length scales, especially relative to the uncorrelated lattice gas (i.e. Poisson distribution of points on the integer lattice) that does not have any such peaks; see section 2 for a precise definition.

Motivated by this numerical study, we apply the the classical methods of Hardy and Littlewood [4] as well as Vinogradov [15] to understand rigorously the nature of the primes as a point process by ascertaining the structure factor, scaled by the density, in a distinguished limit, namely, prime numbers in an interval $M \leqslant p \leqslant M+L$ with $M \to \infty$, $L \to \infty$ such that L/M is fixed and positive. We prove that the limiting structure factor is characterized by a dense set of Dirac delta functions (pure point diffraction pattern without any diffuse part); see proposition 1 (equation (44)) and associated corollary (equation (51)). We also show that the corresponding pair correlation function is equivalent to the Hardy–Littlewood m-tuples conjecture with $m = 2$ [4]. Our limiting form for the structure factor vividly elucidates the fact that the primes are characterized by a previously unknown high degree of order across all length scales similar to but different from quasicrystals or limit-periodic systems. Primes in dyadic intervals are the first examples of what we call *effectively limit-periodic* point configurations. This discovery and analysis applied to the local number variance enables us to determine that the primes in such intervals are indeed hyperuniform (see proposition 2 in section 5), which is a new observation. Employing a scalar order metric τ that depends on $S(k)$, we identify a transition between ordered and disordered prime regimes that depends on the intervals studied; see proposition 3 in section 6. While some of the major results were announced in a Letter [48], few mathematical details and derivations were presented there. Here such details are provided and we also report results that are not contained in [48], including an analysis of the number variance and hyperuniformity of the limit-periodic period-doubling chain, demonstration of the hyperuniformity of the primes, analyses of the scaling of the order metric τ with L as well as the value distribution of $S(k)$ of the primes, and a description of an algorithm to reconstruct accurately prime-number configurations.

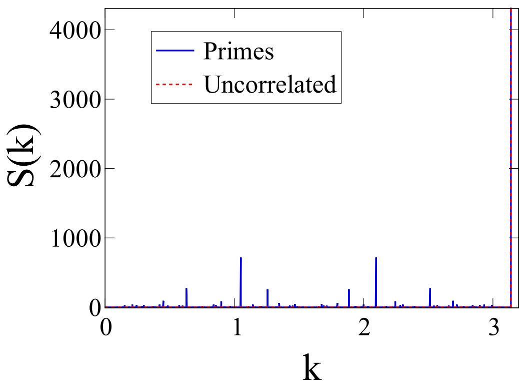

Figure 2. The structure factor $S(k)$ of the prime numbers for $M = 10^{10} + 1$ and $L = 10^5$ obtained in a separate numerical study [47] in which the lattice spacing is assumed to be two (because it is a subset of the odd integer lattice). It is seen that it contains many well-defined Dirac-delta-function like (Bragg-like) peaks of various intensities characterized by a type of self-similarity. Included in the figure is the corresponding structure factor for the uncorrelated lattice (Poisson) gas at the same density on the integer lattice, whose intensities are barely perceptible on the scale of this figure, except for the trivial Bragg peak at $k = 0$ (forward scattering not shown) and associated Bragg peaks that occur at wavenumbers that are multiples of π.

In section 2, we provide relevant definitions. In section 3, we analyze the period-doubling chain, a simple example of a point process with dense Bragg peaks (Dirac delta function) that illustrates the phenomenon of *limit-periodicity*, which we will see applies in a modified form to the primes. In section 4, we show that the structure factor defined in (8) is characterized by sharp peaks at certain rational multiples π, which become progressively denser as M increases, and negligibly small elsewhere. These results are summarized in proposition 1, a corresponding corollary, and a hypothesis. Assuming the Hardy–Littlewood conjecture, we prove that the primes within certain intervals in a distinguished limit are *effectively limit-periodic*. In section 5, we show that, in the infinite-size limit, the primes in a dyadic interval form a hyperuniform point process of class II (see proposition 2). This involves the aforementioned structure factor as well as a cumulative version of it, defined by (23), and the number variance $\sigma^2(R)$ associated with a 'window' of length $2R$. In section 6, we employ a scalar order metric τ, derived from the structure factor, to determine how τ scales with the system size L (see proposition 3) and to identify a transition between large values of τ, when L is comparable to M, and small τ in Gallagher's uncorrelated regime, where L is only logarithmic in M. In section 7, we summarize the classification of the primes as a certain limit-periodic, hyperuniform point process. In section 8, we describe further numerical investigations into the size of the structure factor $S(k)$. In section 9, we discuss the possibility of reconstructing the primes from the limit-periodic form of the inner sum in equation (8). In section 10, we summarize and discuss our conclusions.

2. Definitions

We treat and analyze prime number configurations as point processes and then use a variety of descriptors commonly employed in statistical mechanics to spatially characterize them. In this regard, the ensuing definitions are particularly relevant to the remainder of the paper.

A stochastic point process in a set X is a collection of points with configurational positions $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \ldots$ whose distribution is described by a probability measure on the set of all possible collections. Each configuration in X satisfies two regularity conditions: (i) there are no multiple points ($\mathbf{x}_i \neq \mathbf{x}_j$ if $i \neq j$) and (ii) each bounded subset of X must contain only a finite number of points. Here, we mainly restrict ourselves to one-dimensional point processes. The set X can be one-dimensional Euclidean space $\mathbb{R}$ (continuum systems), discrete systems (e.g. the integer lattice $\mathbb{Z}$), the one-dimensional torus $\mathbb{T}$, or discrete systems on $\mathbb{T}$. The latter two cases constitute *periodic* point processes. A particular configuration (realization) of a point process in X can formally be characterized by the random variable

$$\eta(\mathbf{r}) = \sum_{i=1} \delta(\mathbf{r} - \mathbf{x}_i) \tag{9}$$

called the 'local' density at position $\mathbf{r}$, where $\delta(\mathbf{r})$ is a d-dimensional Dirac delta function. Two particularly important averages are the one-particle and two-particle correlation functions, $\rho_1(\mathbf{r}_1)$ and $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$, respectively. When X is $\mathbb{R}^d$ (continuous systems), they are defined as follows:

$$\rho_1(\mathbf{r}_1) = \langle \eta(\mathbf{r}_1) \rangle, \tag{10}$$

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \langle \eta(\mathbf{r}_1)\eta(\mathbf{r}_2) \rangle - \rho_1(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2), \tag{11}$$

where the angular brackets denote an average with respect to the probability measure. The random setting when X is $\mathbb{R}$ is perfectly general and includes lattices and periodic point processes as special cases. A *lattice* in $\mathbb{R}$ (or in $\mathbb{R}^d$) is a subgroup consisting of the integer linear combinations of vectors that constitute a basis for $\mathbb{R}$ (or $\mathbb{R}^d$). In a lattice in $\mathbb{R}^d$, the space can be geometrically divided into identical regions called *fundamental cells*, each of which contains just one point. In one dimension, there is only one lattice, namely, the integer lattice $\mathbb{Z}$. The dual of the integer lattice with fundamental-cell spacing a is an integer lattice with spacing $2\pi/a$, which we denote by $\mathbb{Z}^*$. A one-dimensional periodic point process (crystal) in $\mathbb{R}$ (points in $\mathbb{T}$) is obtained by placing a fixed configuration of N points (where $N \geqslant 1$) within a fundamental cell $\mathcal{F}$ of the integer lattice, which is then periodically replicated.

In the special case of statistically homogeneous point processes in $\mathbb{R}$, all of the correlation functions are translationally invariant, the first two of which are then simply given by

$$\rho_1(\mathbf{r}_1) = \rho, \tag{12}$$

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \rho^2 g_2(\mathbf{r}_2 - \mathbf{r}_1). \tag{13}$$

Here the constant ρ is the number density (number of points per unit volume) and $g_2(\mathbf{r}) = g_2(-\mathbf{r})$ is the pair correlation function. It is useful to introduce the total correlation function $h(\mathbf{r})$, which is related to the pair correlation function via

$$h(\mathbf{r}) \equiv g_2(\mathbf{r}) - 1 \tag{14}$$

and decays to zero for large $|\mathbf{r}|$ in the absence of long-range order. Note that $h(\mathbf{r}) = 0$ for all $\mathbf{r}$ for the translationally invariant Poisson point process.

The structure factor $S(\mathbf{k})$ in $\mathbb{R}^d$ is defined as follows:

$$S(\mathbf{k}) = 1 + \rho\tilde{h}(\mathbf{k}), \tag{15}$$

where

$$\tilde{f}(\mathbf{k}) = \int_{\mathbb{R}^d} f(\mathbf{r}) \exp\left[-\mathrm{i}(\mathbf{k}\cdot\mathbf{r})\right] \mathrm{d}\mathbf{r} \tag{16}$$

is the Fourier transform of a function $f(\mathbf{r})$ so that

$$f(\mathbf{r}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \tilde{f}(\mathbf{k}) \exp\left[\mathrm{i}(\mathbf{k}\cdot\mathbf{r})\right] \mathrm{d}\mathbf{k}. \tag{17}$$

It follows that the structure factor is always symmetric with respect to the origin, i.e. $S(\mathbf{k}) = S(-\mathbf{k})$. Therefore, in one dimension, one can focus on $S(k)$ (where $k = |\mathbf{k}|$), which conveys complete information about the structure factor. We note that the structure factor $S(\mathbf{k})$ is an important quantity in physics because it can be measured experimentally by x-ray or neutron scattering. Its behavior near $\mathbf{k} = \mathbf{0}$ is directly linked to the hyperuniformity of the point process, as we will detail below. While $S(\mathbf{k})$ is a nontrivial function for spatially correlated point processes, it is identically equal to one for all $\mathbf{k}$ for a translationally invariant Poisson point process.

In general, the structure factor of a statistically homogeneous point process can uniquely be decomposed into three contributions [49]:

$$S(\mathbf{k}) = S(\mathbf{k})_{\mathrm{pp}} + S(\mathbf{k})_{\mathrm{sc}} + S(\mathbf{k})_{\mathrm{ac}}, \tag{18}$$

where $S(\mathbf{k})_{\mathrm{pp}}$ is the 'pure point' (Dirac-delta masses) part, $S(\mathbf{k})_{\mathrm{sc}}$ is the singular-continuous part, and $S(\mathbf{k})_{\mathrm{ac}}$ is the absolutely-continuous part. In the case of the integer lattice, $S(\mathbf{k})$ only consists of the pure-point part. The same is true for a one-dimensional quasicrystal, such as the Fibonacci chain that is characterized by the golden ratio, except here the Dirac-delta functions are dense [50]. The Fibonacci chain is a special case of one-dimensional patterns constructed from substitution rules involving algebraic numbers, and *limit-periodic* chains are closely related patterns but are characterized by rational numbers [49]. One-dimensional point sets generated from substitution rules involving non-Pisot numbers will consist only of singular-continuous contributions [51]. In the case of a Poisson point process, the only contribution to the structure factor is the absolutely continuous part. In stark contrast, we will show that the primes in certain intervals are dominated by a set of dense Bragg peaks.

The following definitions are employed to show the hyperuniformity of the period-doubling chain in section 3 and the primes in section 5. A hyperuniform statistically homogeneous point process in d-dimensional Euclidean space $\mathbb{R}^d$ is one in which the structure factor $S(\mathbf{k})$ tends to zero as the wavenumber $k \equiv |\mathbf{k}|$ tends to zero, i.e.

$$\lim_{|\mathbf{k}|\to 0} S(\mathbf{k}) = 0, \tag{19}$$

implying that single scattering of incident radiation at infinite wavelengths is completely suppressed. This class of point configurations includes perfect crystals, a large class of perfect quasicrystals [33, 52] and special disordered many-particle systems. Observe that the structure-factor definition (15) and the hyperuniformity requirement (19) dictate that the volume integral of $\rho h(\mathbf{r})$ over all space is exactly equal to -1, i.e.

$$\rho \int_{\mathbb{R}^d} h(\mathbf{r}) \mathrm{d}\mathbf{r} = -1, \tag{20}$$

which is a direct-space *sum rule* that a hyperuniform point process must obey. The hyperuniformity property can be stated in terms of the the local number variance $\sigma^2(R)$ associated within an interval (window) of length $2R$ for a one-dimensional homogeneous point process [2]:

$$\sigma^2(R) = \rho 2R\left[1 + \rho \int_{\mathbb{R}} h(\mathbf{r})\alpha_2(r;R)\mathrm{d}\mathbf{r}\right] = \frac{\rho R}{\pi}\int_{\mathbb{R}} S(\mathbf{k})\tilde{\alpha}_2(k;R)\mathrm{d}\mathbf{k}, \tag{21}$$

where $\alpha(r;R) = 1 - r/(2R)$ for $r \leqslant 2R$ and zero otherwise (where $r \equiv |\mathbf{r}|$), and $\tilde{\alpha}(k;R) = 2\sin^2(kR)/(kR)$. A hyperuniform point process is one in which $\sigma^2(R)$ grows more slowly than R in the large-R limit. Three classes of hyperuniformity are to be distinguished: class I, where $\sigma^2(R)$ is bounded; class II, where $\sigma^2(R)$ is logarithmic in R; and class III, where $\sigma^2(R)$ scales as a power $R^{1-\alpha}$ with $0 < \alpha < 1$ (or $d - \alpha$ for a d-dimensional system) [43]. After integrating by parts, the second line of (21) leads to an alternative representation of the number variance [52]:

$$\sigma^2(R) = -\frac{\rho R}{(\pi)}\int_0^\infty Z(k)\frac{\partial\tilde{\alpha}_2(k;R)}{\partial k}\mathrm{d}k, \tag{22}$$

where

$$Z(K) = 2\int_0^K S(k)\mathrm{d}k \tag{23}$$

is the *integrated* or *cumulative* intensity function within a 'sphere' of radius K of the origin in reciprocal space. The quantity $Z(k)$ has advantages over $S(k)$ in the characterization of quasicrystals and other point processes with dense Bragg peaks [52]. If $S(k)$ tends to zero as a power k^α, then its integral $Z(K)$ will tend to zero as a power one higher, $Z(K) \sim K^{\alpha+1}$. Any positive power $\alpha > 0$ yields hyperuniformity and distinguishes the system from a random configuration of Poisson points with the same density.

When X is discrete, such as the integer lattice, it is sometimes convenient to use the same notation as equations (9)–(13) such that $\delta(\mathbf{r} - \mathbf{x}_i)$ is interpreted to be the Kronecker delta $\delta_{\mathbf{r},\mathbf{x}_i}$, which means that $\eta(\mathbf{r})$ takes either the value 0 or 1, depending on whether the site $\mathbf{r} \in X$ is unoccupied (empty) or occupied. In the special case of statistically homogeneous point processes, while the definition (14) remains the same, relations (12) and (13) are modified as follows:

$$\rho_1(\mathbf{r}_1) = f, \tag{24}$$

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = f^2 g_2(\mathbf{r}_2 - \mathbf{r}_1), \tag{25}$$

where f is the occupation fraction (fraction of occupied sites). Similarly, equation (15) for the structure factor becomes in the discrete setting

$$S(\mathbf{k}) = 1 - f + f\tilde{h}(\mathbf{k}), \tag{26}$$

where $\tilde{h}(\mathbf{k})$ is the discrete Fourier transform

$$\tilde{h}(\mathbf{k}) = \sum_{\mathbf{r}\neq\mathbf{0},\mathbf{r}\in X} h(\mathbf{r})\exp\left[-\mathrm{i}(\mathbf{k}\cdot\mathbf{r})\right], \tag{27}$$

where $h(\mathbf{r}) = g_2(\mathbf{r}) - 1$. Note that $1 - f$ is the structure factor of the *uncorrelated lattice gas*, which is a stochastic point process in X in which the occupation of each site is a constant probability f, independent of any other site. While the Fourier-space hyperuniformity condition for discrete X is still given by relation (19), the corresponding direct-space condition is

$$\sum_{\mathbf{r}\neq\mathbf{0},\mathbf{r}\in X} h(\mathbf{r}) = \frac{f-1}{f}. \tag{28}$$

The relation between the local number variance $\sigma^2(R)$ and the structure factor $S(\mathbf{k})$ is unchanged from equation (21). It is noteworthy that the sum rule (28) is a condition for hyperuniformity in the grand-canonical (open-system) ensemble in which the number of particles fluctuates around some average value; see [2] and [43] for details in the continuous-space setting. For a system in which the number of particles is fixed, the sum rule still applies but it is satisfied whether the system is hyperuniform or not.

For a single periodic point configuration of N points within F, specified by its local density $\eta(\mathbf{r})$ (see (9)), it is useful to introduce the *complex collective density variable* $\tilde{\eta}(\mathbf{k})$, which is simply the Fourier transform of $\eta(\mathbf{r})$, i.e.

$$\tilde{\eta}(\mathbf{k}) = \sum_{j=1}^{N} \exp(-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_j). \tag{29}$$

This quantity is directly linked to the *scattering intensity* $\mathcal{S}(\mathbf{k})$ defined as

$$\mathcal{S}(\mathbf{k}) = \frac{|\tilde{\eta}(\mathbf{k})|^2}{N}, \tag{30}$$

which is a nonnegative real function with inversion-symmetry, i.e.

$$\mathcal{S}(\mathbf{k}) = \mathcal{S}(-\mathbf{k}) \tag{31}$$

that obeys the bounds

$$0 \leqslant \mathcal{S}(\mathbf{k}) \leqslant N \qquad (\mathbf{k}\neq\mathbf{0}) \tag{32}$$

with $\mathcal{S}(\mathbf{0}) = N$. For a single periodic configuration with a finite number of N points within a fundamental cell $\mathcal{F}$, the scattering intensity $\mathcal{S}(\mathbf{k})$ is identical to the structure factor $S(\mathbf{k})$ (see (15)), except the latter excludes $\mathbf{k} = \mathbf{0}$ (forward scattering). In general, whether they remain equal in the infinite-system limit depends on the ergodicity of the process, but this issue does not affect our analysis of the primes, and so we will simply take equation (8), which is equivalent to equation (30), to be the definition of the structure factor. Importantly, the definition of hyperuniformity excludes the forward scattering contribution, which is implicit in (19).

To characterize quantitatively order in the primes in various intervals (section 6), we will determine a scalar positive order metric τ that is capable of capturing the degree of translational order across length scales [38]. For a statistically homogeneous point process in $\mathbb{R}^d$ at number density ρ, it is defined by

$$\tau \equiv \frac{1}{D^d}\int_{\mathbb{R}^d} [g_2(\mathbf{r})-1]^2 \mathrm{d}\mathbf{r} \tag{33}$$

$$= \frac{1}{(2\pi)^d D^d}\int_{\mathbb{R}^d} [S(\mathbf{k})-1]^2 \mathrm{d}\mathbf{k}, \tag{34}$$

where D is some characteristic length scale. A convenient choice is $D = \rho^{-1/d}$. For a Poisson point process in $\mathbb{R}^d$, $\tau = 0$ because $g_2(\mathbf{r}) - 1$ is zero for all $\mathbf{r}$. Thus, a deviation of τ from zero measures translational order with respect to the fully uncorrelated case. For example, for the Riemann zeta zeros, $\tau = 2/3$, assuming Montgomery's pair correlation conjecture or, equivalently, the corresponding structure factor (6), which reflects the disordered hyperuniformity of the point process. For any periodic point process in which there are a finite number of points within the fundamental cell $\mathcal{F}$, τ is unbounded because the integrals are carried out over all space. For this reason, one can employ a modified version of τ by carrying out the integral in direct space or reciprocal space over appropriate subsets of $\mathbb{R}^d$, in which case the equality in equation (34) no longer applies.

The discrete-setting counterpart of the order metric τ defined in (34) in which X is a subset of $\mathbb{Z}$ on the torus $\mathbb{T}$ in which the fundamental cell has length L with lattice spacing a is given by

$$\tau \equiv \sum_{j=1}^{Ns-1} f^2[g_2(aj) - 1]^2 \tag{35}$$

$$= \frac{1}{N_s} \sum_{j=1}^{Ns-1} \left(S\left(\frac{2\pi j}{L}\right) - (1-f) \right)^2, \tag{36}$$

where N_s is the number of lattice sites within the fundamental cell (equal to L/a) and N is the number of occupied sites. Strictly speaking, the quantities $g_2(r)$ and $S(k)$ are ensemble averages but, practically, $g_2(r)$ and $S(k)$ can also be interpreted to come from a single but large configuration. In the case of an uncorrelated lattice gas, $S = 1 - f$ in the infinite-system-size limit so that $\tau = 0$. The corresponding expression in the discrete setting for a single configuration in any space dimension was presented and applied in [53]. Analysis of the primes in some fixed interval requires the use of this single-configuration variant of τ, which we will employ. Note that τ for a single configuration of an uncorrelated lattice gas is given by $(1-f)^2$ (not zero) in the infinite-system-size limit.

3. An illustrative example: the period-doubling chain

The spatial distribution of the primes shares some features with *limit-periodic* point sets, and so we analyze an example of such a point configuration, the period-doubling chain, in detail. Consider two types of intervals ('tiles' or 'letters'): a and b. The period-doubling chain is defined by the following iterative substitution rule: $a \to ab$ and $b \to aa$ [49]. Starting from an initial seed a, one obtains a 'one-sided' chain sequence $a, ab, abaa, abaaabab, abaaababababaaabaa, \ldots$. Henceforth, we assume that the chain extends to positive and negative values in $\mathbb{R}$ about the origin (i.e. a 'double-seed' $a|a$ is taken to be the initial condition) in order be consistent with the definitions of the functions presented in section 2 and employed below. In the infinite-size limit, this constitutes a point process on $\mathbb{Z}$ in which a subset of sites are occupied by a's and the remaining sites are occupied by b's. The locations of the b's are given by a superposition of arithmetic progressions $2 + 4j, 8 + 16j, 32 + 64j, \cdots$, with a factor of four from one to the next. Thus the infinite-size limit is a union of periodic systems, which is termed limit-periodic. The limiting densities of a and b sites are 2/3 and 1/3, respectively. The structure factor associated with the a's, as obtained from (30), is given by

$$S(k) = \frac{4\pi}{3} \left(\sum_{m=1}^{\infty} \delta(k - 2\pi m) + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} 2^{-2n} \delta\left(k - \frac{(2m-1)\pi}{2^{n-1}}\right) \right) \tag{37}$$

assuming unit lattice spacing. This is obtained by squaring equation (12) in [49], multiplying it by $2\pi/f$, and then rescaling the function by 2π. The factor 2π accounts for differences in the definition of the Fourier transform. Thus, we have a dense set of Dirac-delta-function peaks, one for each dyadic rational $(2m-1)/2^{n-1}$; see figure 3. These peaks at certain rational

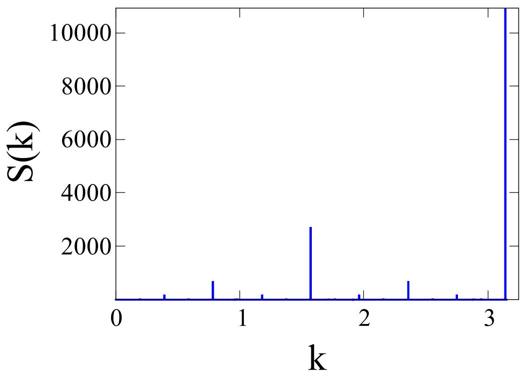

Figure 3. Structure factor of the period-doubling chain as obtained from formula (3) in which the second sum is truncated at $n = 20$. Note the self-similarity in the intensities and locations of the peaks. The height of the peak at $k = \pi$ is larger than shown on the scale of this figure.

numbers arbitrarily close to zero are a feature shared by this example and the prime numbers. Figure 3 depicts the structure factor of the period-doubling chain.

Substitution of this expression (37) for the structure factor into relation (21) yields the local number variance for the period-doubling chain:

$$\sigma^2(R) = \frac{8}{9\pi^2}\left(\sum_{m=1}^{\infty}\frac{\sin^2(2m\pi R)}{m^2} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\sin^2((2m-1)\pi R/2^{n-1})}{(2m-1)^2}\right). \tag{38}$$

The first term in (38) is a periodic function $R(1-2R)/9$ with period 1/2 and the second term is a superposition of periodized 'triangle' functions with heights 1/9 and bases 1, 2, 4, $\cdots$.. Together this results in a number variance that grows logarithmically in R; see figure 4. Therefore, the period-doubling chain is hyperuniform of class II [43].

Having established that the period-doubling chain is hyperuniform, we now want to determine how the structure factor $S(k)$ behaves in the vicinity of the origin. The structure factor $S(k)$ is not a continuous function because there are dense Bragg peaks arbitrarily close to 0, so we do not have $S(k) \to 0$ as $k \to 0$ in the usual sense. We follow the practice of [52] in such instances and pass to a cumulative version of the structure factor, $Z(K)$, defined by (23). This integral simply adds the weights of the δ peaks up to position K. In order to have a peak $(2m-1)\pi/2^{n-1}$ within the range of integration, n must be relatively large:

$$\exists m\ \frac{2m-1}{2^{n-1}}\pi \langle K \iff 2^{n-1}\rangle \pi/K \iff n > \log_2(\pi/K) + 1 = \log_2(2\pi/K), \tag{39}$$

or else there are no integers m in the necessary interval. We have an explicit formula for the tail of a geometric series:

$$\sum_{n>C} b^n = b^{\lceil C \rceil}\frac{1}{1-b} \tag{40}$$

where $\lceil C \rceil$ denotes the least integer greater than C (and, in particular, $C+1$ in case C is already an integer). Using this to sum the series in $Z(K)$ leads to an explicit formula

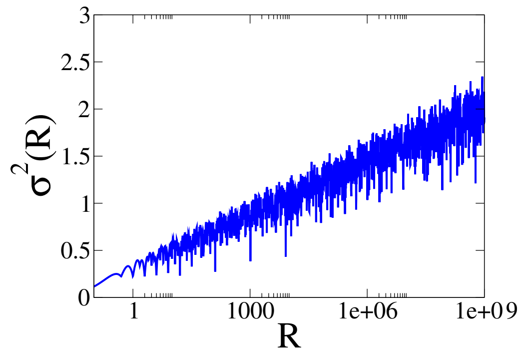

Figure 4. Local number variance of the period-doubling chain as computed from the explicit formula (38) in which the second sum is truncated at $n = 20$.

$$
\begin{aligned}
Z(K) = 2\int_0^K S(k)\mathrm{d}k &= \frac{8\pi}{3}\sum_{n>\log_2(2\pi/K)} 2^{-2n}\left\lfloor \frac{1}{2} + 2^{n-2}\frac{K}{\pi}\right\rfloor \\
&= \frac{8\pi}{3}\sum_{n>\log_2(2\pi/K)} 2^{-2n}\left(\frac{1}{2} + 2^{n-2}\frac{K}{\pi} + \left\{\frac{1}{2} + 2^{n-2}\frac{K}{\pi}\right\}\right) \\
&= \frac{8\pi}{3}\left(\frac{4}{3}2^{-2\lceil\log_2(2\pi/K)\rceil} + \frac{1}{2}\frac{K}{\pi}2^{-\lceil\log_2(2\pi/K)\rceil} + \sum_{n>\log_2(2\pi/K)} 2^{-2n}\left\{\frac{1}{2} + 2^{n-2}\frac{K}{\pi}\right\}\right)
\end{aligned}
\tag{41}
$$

where the braces $\{\cdot\}$ denote fractional part. (Note that the first term in expression (37) for the structure factor makes no contribution to $Z(K)$ for small K.) Taking into account the jump discontinuities when K/π crosses a dyadic rational shows that

$$\frac{1}{6\pi}K^2 \leqslant Z(K) \leqslant \frac{1}{2\pi}K^2. \tag{42}$$

Figure 5 shows the function $Z(k)$ and the aforementioned upper and lower bounds.

Thus $Z(K)$ is bounded between two multiples of K^2, with a self-similar staircase-like behavior in between. Using these bounds and relation (22), we get the following corresponding asymptotic bounds on the number variance $\sigma^2(R)$ in the limit $R \to \infty$:

$$\frac{4}{9\pi^2}\ln(R) \leqslant \sigma^2(R) \leqslant \frac{4}{3\pi^2}\ln(R). \tag{43}$$

This implies that the period-doubling chain falls within class II of hyperuniform systems with a structure factor that effectively behaves as $S(k) \sim k$ as $k \to 0$ [43]. These asymptotic bounds substituted in the number variance expression (22) closely match the upper and lower envelopes (not shown) of the fluctuating number variance function plotted in figure 4.

4. Structure factor of the primes in a distinguished scaling limit

We theoretically study the structure factor $S(k)$ of the primes in an interval $[M, M+L]$ and consider $\beta \equiv L/M$ to be fixed such that $M+L$ is within a constant multiple of M, and thus the prime number theorem implies that the density ρ is effectively $1/\ln(M)$ throughout the interval. We show that the structure factor of the primes in a distinguished limit such that

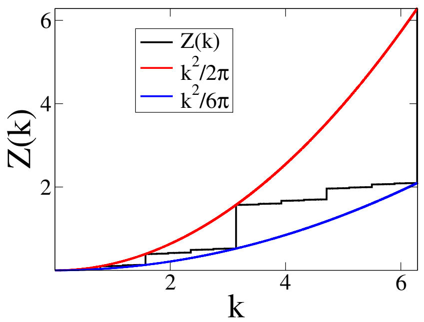

Figure 5. The cumulative intensity function $Z(k)$ of the period-doubling chain as obtained from formula (41) in which the second sum is truncated at $n = 20$. The upper and lower bounds given in (42) are also indicated in the figure.

$M \to \infty$, $\beta > 0$ and after scaling by the density ρ, exhibit peaks that will consist only of Dirac delta functions at rational numbers with odd, square-free denominators, and hence the small diffuse part observed numerically in [47], vanishes in this limit. To do so, we rely on the Hardy–Littlewood circle method [4], but applied to the structure factor, with some modifications. This hypothesis is described by the following proposition:

Proposition 1. *Contingent on the Hardy–Littlewood conjecture, the structure factor of the prime numbers, scaled by the the density* ρ*, in an interval* $M \leqslant p \leqslant M + L$ *in the limit such that* $M \to \infty$*,* $L/M \to \beta > 0$ *is given by*

$$\lim_{M\to\infty} \frac{S(k)}{2\pi\rho} = \sum_{n\in\mathbb{Z}^+}{}^{\flat} \sum_{m\in\mathbb{Z}}{}^{\times} \frac{1}{\phi(n)^2} \delta\left(k - \frac{m\pi}{n}\right), \tag{44}$$

where $\phi(n)$ *is Euler's totient, which counts the numbers up to n with no factor in common with n [45], the symbol* $\flat$ *indicates that the sum over n only involves odd, square-free values of n and the symbol* $\times$ *indicates that m and n have no common factor.*

In what follows, we give a derivation of (44), which is contingent on the Hardy–Littlewood conjecture (2) with $m = 2$. Our approach is different from the original procedure of Hardy and Littlewood and benefits from significant improvements due to Vinogradov [15]. Here we note that Hardy and Littewood begin with a generating function $f(x)$ that is a sum is over all primes, not just those in a finite interval. By contrast, because we want to make contact with the modern concept of hyperuniformity, we begin with the structure factor $S(k)$ for primes in a finite interval $[M, M + L]$. Moreover, we pass to the aforementioned scaling limit, which is not done in usual treatments. For example, in applications like Vinogradov's proof that every large odd number is a sum of three primes, it is important to have a finite version of the peaks with the sum over n in (44) truncated. Passing to the limit in proposition 1 loses some of this information, which is one of the reasons why previous formulations do not state results in terms of Dirac delta functions in the way we do. On the other hand, taking this limit is very important from the point of view of establishing the hyperuniformity of the primes (proposition 2 of section 5). It also leads us to ask different questions and obtain new insights. For example, it enables us to quantify the degree of order in the primes (proposition 3 of section 6) as a function of L and M and allows the first classification of the primes in the aforementioned scaling limit as effectively limit-periodic point process (section 7).

4.1. Derivation of proposition 1

Our starting point is the Fourier-space-based structure factor, which is different from the direct-space generating function $f(x)$ that Hardy and Littlewood employed in their celebrated paper [4]; see also the outline of their approach and commentary in the supplementary material. Our first step is to replace the complex density variable $\tilde{\eta}(k)$ (see (29)) by another sum involving more convenient weights, namely, the von Mangoldt function $\Lambda(n)$, which is given by $\ln(p)$ when n is a power of the prime p, and zero otherwise. To do so, we use the following identity [44]:

$$\sum_{X<p<Y} \mathrm{e}^{\mathrm{i}\alpha p} = \frac{1}{\ln(Y)} \sum_{p<Y} \mathrm{e}^{\mathrm{i}\alpha p} \ln(p) - \frac{1}{\ln(X)} \sum_{p<X} \mathrm{e}^{\mathrm{i}\alpha p} \ln(p) + O\left(\frac{Y-X}{\ln(X)^3}\right) \tag{45}$$

for $X \asymp Y$. We may include powers of primes p^k weighted by $\ln(p)$ at little extra cost, which simplifies certain estimates. This enables us to write the collective density variable in the interval $[M, M+L]$ as

$$\tilde{\eta}(k) \equiv \sum_{M<p\leqslant M+L} \mathrm{e}^{-\mathrm{i}kp} = \frac{1}{\ln(M)} \sum_{n=M+1}^{M+L} \Lambda(n)\mathrm{e}^{-\mathrm{i}\alpha n} + O\left(\frac{L}{\ln(M)^2}\right). \tag{46}$$

The corresponding expression for the structure factor follows immediately from (46) and definition (8), but we scale this result by $2\pi\rho$, arriving at the result

$$\frac{1}{2\pi\rho} S(k) = \frac{1}{2\pi L} \left| \sum_{n=M+1}^{M+L} \Lambda(n)\mathrm{e}^{-\mathrm{i}kn} \right|^2 + O\big(L/\ln(M)^2\big), \tag{47}$$

where, to bound the error, we used the trivial bound $\sum \Lambda(n)\mathrm{e}^{\mathrm{i}\alpha n} \lesssim L$, and remind the reader that $\rho \sim 1/\ln(M)$ and $N \sim \rho L$.

We now apply the circle method of Hardy and Littlewood, but to the sum (47) for the structure factor. We split the interval into major arcs $\mathfrak{M}(q,a) = \{|\alpha - a/q| < \varepsilon\}$ near fractions a/q with small denominator q, and the rest of the interval (minor arcs), denoted by $\mathfrak{m}$. The denominator is restricted to $q \leqslant q_{\max}$ and we choose $q_{\max} = \ln(L)^B$, with a constant B as large as one pleases. The length ε of each major arc is chosen to be $\varepsilon = \ln(L)^B/L$.

Let us first calculate the contribution from the major arcs $\mathfrak{M}(q,a)$ to formula (46) for the collective density variable $\tilde{\eta}(k)$, returning later to the discussion of the minor arcs. On the major arc $\mathfrak{M}(q,a)$, even without the Riemann hypothesis, we still have the following approximation for

$$\ln(M)\tilde{\eta}(k) \approx \frac{\mu(q)}{\phi(q)} \sum_{n=M+1}^{M+L} \mathrm{e}^{-\mathrm{i}(k-2\pi a/q)n} + O\big(L\mathrm{e}^{-c\sqrt{L}}\big). \tag{48}$$

This plays the role of Hardy–Littlewood's approximation to $f(x)$ in lemma 9 of [4]; see also [44] and [54]. They take $M = 1$ but one can of course subtract. We summarize the argument from [44] in appendix for the reader's benefit in order to emphasize the role played by primes in arithmetic progression; see also [48]. Now we employ the elementary identity

$$\left| \sum_{n=M+1}^{M+L} \mathrm{e}^{-\mathrm{i}nt} \right|^2 = \frac{1-\cos(2\pi tL)}{1-\cos(2\pi t)} = \left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2, \tag{49}$$

which immediately follows from summation of the geometric series. Note that $e^{c\sqrt{\ln(L)}}$ is much larger than $\ln(L)^b = e^{b\ln\ln L}$ for any positive values of a and b. Thus the errors from using the prime number theorem are smaller than the error $L/\ln(M)^2$ that we have already exposed ourselves to in (47).

Therefore, the structure factor, which is given by $|\tilde{\eta}(k)|/N$ (see (30)), in the vicinity of a particular peak location k for some a and q and sufficiently small t is given by

$$\frac{1}{2\pi\rho}S(k) = \frac{\mu(q)^2}{\phi(q)^2}\frac{1}{2\pi L}\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2 + O\big(L/\ln(M)^2\big), \tag{50}$$

where we have used (48) and (49). The following corollary immediately follows from (50).

Corollary. *For finite but large N, the structure factor at certain rational values of k/π is given approximately by*

$$S(\pi m/n) \approx N\left(\frac{\mu(2n)}{\phi(2n)}\right)^2, \tag{51}$$

where $\mu(n)$ is the Möbius function defined so that $\mu^2(2n)$ is one whenever $2n$ is square-free and zero otherwise, and note that $\phi(n) = \phi(2n)$ if n is odd. The corresponding representation of the structure factor as a finite sum that accounts for multiple peaks is given by

$$S(k) \approx \sum_{3\leqslant n\leqslant n_{\max}}^{\flat} \sum_{1\leqslant m\leqslant n-1}^{\times} N\left(\frac{\mu(2n)}{\phi(2n)}\right)^2 \delta_{k,m\pi/n}, \tag{52}$$

where the sum over n is truncated at $n = n_{\max}$ and $\delta_{x,y}$ is the Kronecker delta.

Remark 1. Referring to (50), we see that in the limit that t tends to zero faster than L goes to infinity, and $M \to \infty$ such that L/M remains finite, this formula yields Dirac delta functions at rational locations described in the corollary. However, in order to arrive at the limiting form (44) of *proposition 1*, we must show that the structure factor is negligibly small at the irrationals in the infinite-system-size limit, which is a highly nontrivial task. To do so, we rely on the following hypothesis.

Hypothesis. For a general continuous, periodic function $F(x)$, we have

$$\int_0^1 F(\alpha)\frac{1}{L}\left|\sum_{n=M+1}^{M+L} e^{n\alpha}\Lambda(n)\right|^2 d\alpha \approx \sum_q\sum_a \frac{\mu(q)^2}{\phi(q)^2}F(a/q). \tag{53}$$

We now provide strong theoretical arguments, even if far short of rigorous proof, to support this hypothesis. By Fejér's theorem for a periodic function Fx), we have uniform convergence:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F(x-t)\frac{1}{L}\left(\frac{\sin(Lt/2)}{\sin(t/2)}\right)^2 dt \to F(x) \tag{54}$$

as $L \to \infty$. Assuming ε is not too small, the bulk of the integral comes from $|t| < \varepsilon$. Indeed,

$$\int_\varepsilon^{1-\varepsilon} F(a/q+t)\frac{1}{L}\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2 dt \leqslant \frac{2\|F\|_\infty}{L}\int_\varepsilon^{1/2}\frac{dt}{\sin(\pi t)^2} = \frac{2\|f\|_\infty}{L\pi}\cot(\pi\varepsilon) \lesssim \frac{\|F\|_\infty}{L\varepsilon}.$$

This implies that the integral over $\mathfrak{M}(q,a)$ converges to $F(a/q)$ as $L\to\infty$, provided $L\varepsilon\to\infty$. This is the case, for example, with our choice $\varepsilon=\log(L)^B/L$ for any power B. Note that

$$\frac{1}{L}\int_0^1\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2\mathrm{d}t=1.$$

This allows us to write

$$\int_0^1 F(a/q+t)\frac{1}{L}\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2\mathrm{d}t=F(a/q)+\int_0^1\left(F(a/q+t)-F(a/q)\right)\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2\frac{1}{L}\mathrm{d}t$$
$$=F(a/q)+\int_{-\varepsilon}^{\varepsilon}\left(F(a/q+t)-F(a/q)\right)\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2\frac{1}{L}\mathrm{d}t+O\left(\frac{\|F\|_\infty}{L\varepsilon}\right).$$

Suppose that F has modulus of continuity ω, meaning that $|F(x)-F(y)|\leqslant\omega(\varepsilon)$ whenever $|x-y|\leqslant\varepsilon$. For example, if F has continuous derivatives of order up to m, then $\omega(\varepsilon)=\|F\|_{C^m}\varepsilon^m$ is a valid modulus of continuity for F. Using once more the fact that the integral of the Fejér kernel is 1, we have

$$\int_0^1 F(a/q+t)\frac{1}{L}\left(\frac{\sin(\pi Lt)}{\sin(\pi t)}\right)^2\mathrm{d}t=F(a/q)+O\left(\omega(\varepsilon)+\frac{\|F\|_\infty}{L\varepsilon}\right).$$

Dividing the interval $[0,1]$ into the major arcs $\mathfrak{M}(a,q)$ for $q\leqslant q_{\max}$ together with the remaining minor arcs $\mathfrak{m}$, and summing the error made in Fejér's theorem over $q\leqslant q_{\max}$, we get

$$\int_0^1 F(\alpha)\frac{1}{L}\left|\sum_{X<n<Y}\mathrm{e}^{n\alpha}\Lambda(n)\right|^2\mathrm{d}\alpha$$
$$=\sum_q\sideset{}{^\times}\sum_a\frac{\mu(q)^2}{\phi(q)^2}f\left(\frac{a}{q}\right)+O\left(\left(\omega_F(\varepsilon)+\frac{\|F\|_\infty}{L\varepsilon}\right)\ln(q_{\max})\right)+\frac{1}{L}\int_{\mathfrak{m}}F(\alpha)\left|\sum_{X<n<Y}\mathrm{e}^{n\alpha}\Lambda(n)\right|^2\mathrm{d}\alpha.$$

We assume that F is smooth enough to have $\omega_F(\varepsilon)\ln(q_{\max})\to0$, where $\epsilon=\ln(L)^B/L=q_{\max}/L$. For example, this is the case if F is C^1 or even just Hölder continuous with any exponent α, since $\varepsilon^\alpha\ln(q_{\max})\to0$. Likewise, $\ln(q_{\max})/L\varepsilon\to0$. Therefore, assuming the integral over $\mathfrak{m}$ is negligible yields (53) of the hypothesis.

Remark 2. We stress that proving that the minor arcs contribute less than the major arcs is a significant challenge, which we have not solved. Nevertheless, the analysis above is very suggestive.

Now we apply equation (53) of the hypothesis to prime pairs by setting $F(k)=\mathrm{e}^{\mathrm{i}kr}$, yielding

$$\frac{1}{2\pi\rho}\int_0^\pi F(k)S(k)\mathrm{d}k\approx\sideset{}{^\flat}\sum_{n\in\mathbb{Z}^+}\sideset{}{^\times}\sum_{1\leqslant m\leqslant n-1}\frac{1}{\phi(n)^2}F(\pi m/n),\qquad(55)$$

where we have used (47). Expression (55) gives a count of how often p and $p+r$ are both prime for $r\neq0$. Collecting all of the results above and taking the limit of relation (50) so that t tends to zero faster than L goes to infinity, and $M\to\infty$ such that *L*/*M* remains finite yields our hypothesized limiting form (44) of proposition 1.

Remark 3. The interpretation of (44) in our hypothesis of proposition 1 is that $S(k)/(2\pi\rho)$ converges to the sum of peaks $\sum_n\sum_m\phi(n)^{-2}\delta(k-\pi m/n)$ in the sense that their integrals against test functions $F(k)$ are close. Note that the pure point diffraction form of the scaled

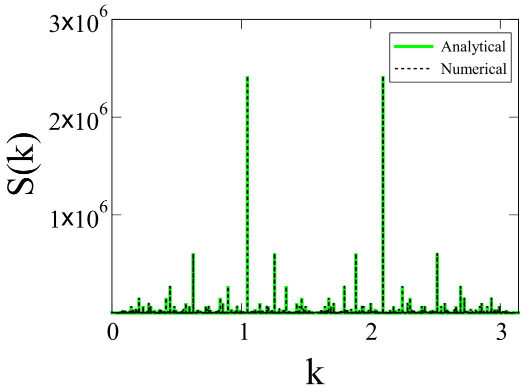

Figure 6. Structure factor for the primes as a function of k (in units of the integer lattice spacing of 2) in the interval $[M, M+L)$, as predicted from formula (51) for $M = 10^{10}+1$, $L = 2.23 \times 10^8$ and $n_{\max} = 100\ln(M)$. This shows many peaks with a type of self-similarity described in the main body of the text and is seen to be in excellent agreement with the corresponding numerically computed structure factor obtained in [47].

structure factor means that the 'diffuse' contribution seen in numerical experiments in finite intervals [47] vanishes in this distinguished limit.

Remark 4. This normalized structure factor of the primes (44) exhibits dense Bragg peaks at certain rational values of k/π, and hence is similar to the structure factor (37) of the limit-periodic period-doubling chain discussed in section 3, which heretofore was not observed. However, there is a fundamental difference between these two systems, to be elaborated in section 7.

Proposition 1 is also supported by our numerical results. Figure 6 compares the prediction of formula (52) for the structure factor of the primes for $M = 10^{10}+1$, $L = 2.23 \times 10^8$ and $n_{\max} = 100\ln(M)$ to the corresponding numerical results reported in [47]. Agreement between the analytical and numerical results is excellent. The structure factor contains many well-defined Bragg-like peaks of various intensities characterized by a type of self-similarity, exact in the limit $M \to \infty$. This self-similar structure arises from the fact that $\phi(n_1 n_2) = \phi(n_1)\phi(n_2)$ for relatively prime n_1 and n_2, so that rescaling preserves the relative heights of the peaks given by equation (44).

4.2. Pair correlation function and equivalence to the Hardy–Littlewood conjecture

We now obtain the pair correlation function $g_2(r)$ of the primes via the limiting form of the structure factor (44) by performing the inverse Fourier transform of $S(k) - 1 \equiv \rho\tilde{h}(k)$ using (17), where $\tilde{h}(k)$ is the Fourier transform of $h(r) \equiv g_2(r) - 1$ for $r \neq 0$:

$$g_2(r) = 1 + \sum_{n \in \mathbb{Z}^+}{}^{\flat} \frac{1}{\phi^2(n)} \sum_{1 \leqslant m \leqslant n-1}^{\times} \exp(rm\pi\mathrm{i}/n). \qquad (56)$$

Recall that the pair correlation function is defined such that $fg_2(r)$ gives the conditional probability that, assuming p is prime, so is $p + r$. Because the density of the primes is $1/\ln(M)$, the occupation fraction is $f = 2/\ln(M)$, and the density of prime pairs with separation r (see (13)) is

$$\begin{aligned}\rho(r) &= \frac{\#\{p \in [M, M+L];\ p,\ p+r \text{ prime}\}}{L} \\ &= \frac{f^2 g_2(r)}{2} \\ &= \frac{2}{(\ln M)^2}\left[1 + \sum_{n\in\mathbb{Z}^+}^{\flat} \frac{1}{\phi^2(n)} \sum_{1\leqslant m\leqslant n-1}^{\times} \exp(rm\pi \mathrm{i}/n)\right].\end{aligned} \tag{57}$$

From (2) and (57), we see that expression (56) for $g_2(r)$ for distinct values of $r = 2, 4, 6, \ldots$ is simply a different representation of the Hardy–Littlewood constant $\mathfrak{S}(\mathcal{H})$, given by (3), for the case $m = 2$ so that $\mathcal{H} = \{0, r\}$. Thus, (44) implies the Hardy–Littlewood conjecture on prime pairs by taking the test function $F(k)$ of (55) to be $\mathrm{e}^{\mathrm{i}kr}$. Conversely, if the Hardy–Littlewood conjecture holds for every r, then one knows (44) for exponential functions $\mathrm{e}^{\mathrm{i}kr}$ and one can deduce it for other functions $F(k)$ by Fourier expansion. The limiting form for $S(k)$ is thus an equivalent formulation of the Hardy–Littlewood conjecture. As such, we certainly do not have a proof of it to offer here.

5. Hyperuniformity

Proposition 2. *Assuming proposition 1, we prove that the primes in the infinite-system-size limit such that $M \to \infty$, $L/M \to \beta > 0$ form a hyperuniform point process of class II.*

Proof. As in the period-doubling chain, to determine whether the primes are hyperuniform, we first determine the cumulative intensity $Z(K)$ defined by (23). Using this relation and assuming (44) from proposition 1, we find that this quantity satisfies

$$\lim_{M\to\infty} \frac{Z(K)}{2\pi\rho} = 2\sum_{n}^{\flat} \sum_{\frac{\pi m}{n}<K}^{\times} \frac{1}{\phi(n)^2}. \tag{58}$$

The largest value of $S(k)$ is N, whereas the weights in (44) lead to values $N/\phi(n)^2 \leqslant N/n^2$. For these to be comparable to the maximum, a cutoff $n \leqslant \sqrt{1/c}$ on the denominator is necessary so that the peak values will be of the same order of magnitude as N. On the other hand, n must be large enough for there to be a peak position $m\pi/n$ less than K:

$$\frac{\pi}{K} \leqslant n \leqslant \sqrt{\frac{1}{c}} = n_{\max}. \tag{59}$$

In particular, the lowest allowable K is $\pi/n_{\max}$.

For the sum over m, note that for any v, and in particular for $v = K/\pi$,

$$\sum_{m<vn}^{\times} 1 = v\phi(n) + O_\epsilon(n^\epsilon). \tag{60}$$

To see this, we use the Möbius inversion property, namely

$$\sum_{d|t} \mu(d) = \begin{cases} 1 \text{ if } t = 1 \\ 0 \text{ if } t > 1 \end{cases} \tag{61}$$

to detect the condition $\gcd(m, n) = 1$. This implies that

$$\sum_{m<vn}^{\times} 1 = \sum_{m<vn} \sum_{d|\gcd(m,n)} \mu(d) = \sum_{d|n} \sum_{b<vn/d} \mu(d) = \sum_{d|n} \mu(d) \lfloor vn/d \rfloor.$$

Now we write $\lfloor vn/d \rfloor = vn/d - \{vn/d\}$ in terms of integer part and fractional part:

$$\sum_{d|n} \mu(d) \lfloor vn/d \rfloor = v \sum_{d|n} \mu(d) \frac{n}{d} - \sum_{d|n} \mu(d) \{vn/d\}.$$

The main term $v\phi(n)$ in (60) thus comes from the identity

$$\phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}.$$

The rest is negligible because $|\mu(d)\{vn/d\}| \leqslant 1$ and the number of divisors of n is asymptotically smaller than any power of n, denoted $O_\epsilon(n^\epsilon)$.

To a good approximation for sufficiently large M, this yields

$$Z(K) \approx \frac{4\,K}{\ln(M)} \sum_{\frac{\pi}{K}<n<n_{\max}}^{\flat} \frac{1}{\phi(n)} + O_\epsilon \left(\frac{1}{\ln(M)} \sum_{\frac{\pi}{K}<n<n_{\max}} \frac{n^\epsilon}{\phi(n)^2} \right) \tag{62}$$

$$\approx \frac{K}{\ln(M)} \left(\ln n_{\max} - \ln \frac{\pi}{K} \right) + O_\epsilon \left(\frac{1}{\ln(M)} \sum_{\frac{\pi}{K}<n<n_{\max}} \frac{n^\epsilon}{\phi(n)^2} \right). \tag{63}$$

Recall that there is a lower bound on K: $K \geqslant \pi/n_{\max}$. Expanding the function $Z(K)$ in (63) around this value, note that

$$\ln n_{\max} - \ln \frac{\pi}{K} \sim \frac{n_{\max}}{\pi} (K - \pi/n_{\max}) \left(1 - \frac{n_{\max}}{\pi} (K - \pi/n_{\max}) \right).$$

If we take $n_{\max}$ to be of order $\ln(M)$ and K close to $\pi/n_{\max}$, then the first term as $K \to 0$ is approximately

$$\frac{K}{\ln(M)} \left(\ln n_{\max} - \ln \frac{\pi}{K} \right) \sim \frac{1}{\ln(M)} K \left(\frac{K n_{\max}}{\pi} - 1 \right).$$

To bound the second term in (63), note that

$$\sum_n \frac{n^\epsilon}{\phi(n)^2} \lesssim \int_{\pi/K}^{n_{\max}} x^{-2+\epsilon} dx \lesssim (\pi/K)^{-2+\epsilon} - n_{\max}^{-2+\epsilon}.$$

The binomial theorem implies that

$$(\pi/K)^{-2+\epsilon} - n_{\max}^{-2+\epsilon} = \left(\frac{K}{\pi} \right)^{2-\epsilon} \left(1 - (K n_{\max}/\pi)^{-2+\epsilon} \right) \lesssim K^{2-\epsilon} (K n_{\max}/\pi - 1).$$

By our choices $n_{\max} \sim \ln(M)$ and $K/\pi \sim n_{\max}^{-1}$, we have $1/\ln(M) \approx K$. Since $K < 1$, $K^{2-\epsilon}$ is larger than $K/\ln(M) \sim K^2$, so it is the second term that dominates. We have

$$Z(K) \lesssim_\epsilon K^{2-\epsilon}\big(Kn_{\max}/\pi - 1\big),$$

where the notation means $Z(K)$ is close to a multiple of the quantity on the right, and this multiple may depend on ϵ. If we choose K/π so close to $n_{\max}^{-1}$ that

$$Kn_{\max}/\pi = 1 + O(K^\epsilon),$$

then we obtain

$$Z(K) \lesssim K^2.$$

This implies hyperuniformity, in the cumulative sense, and subject to this specific calibration of $n_{\max}$ and K. □

Remark. The result $Z(K) \sim K^2$ as $K \to 0$ implies that the primes fall within class II of hyperuniform systems with a structure factor $S(k)$ that effectively is linear in k as $k \to 0$ [43], so that the number variance $\sigma^2(R)$ scales logarithmically with R in the large-R limit.

Figure 7 depicts $Z(K)$ for the primes as determined from the truncated version of (58) with $n_{\max} = 10\ln(M)$. It is seen that for small K, $Z(K)$ is quadratic in K, as predicted.

6. τ order metric

We now study the order metric τ in the discrete setting (see (36)) as a function of the system size L for the primes and the integer lattice.

Proposition 3. *Assuming (51) of the corollary of proposition 1, we prove that the order metric τ of the primes as a function of the system size L in the considered interval $[M, M+L]$ (when divided by ρ^2) obeys the following scaling relation:*

$$\tau/\rho^2 \sim cL, \tag{64}$$

where c is some constant.

This dependence of τ on L can be understood in terms of the peaks in the structure factor, given by (51), via the definition (36). We have

$$\tau = \frac{1}{N_s}\sum_{j=1}^{N_s-1}\left(S(j\pi/N_s) - 1 + f\right)^2. \tag{65}$$

We have seen how to estimate $S(\pi m/n)$ at a rational number in lowest terms: if the denominator is square-free, there will be a peak of height $N/\phi(n)^2$ while, if not, the structure factor will be very small. There could be a common factor between j and N_s, so let

$$n = \frac{N_s}{\gcd(j, N_s)}. \tag{66}$$

This makes $j/N_s = m/n$ in lowest terms with numerator $m = j/\gcd(j, N_s)$. Assume for simplicity that N_s is square-free so that all of the resulting denominators n will be square-free. Then we have

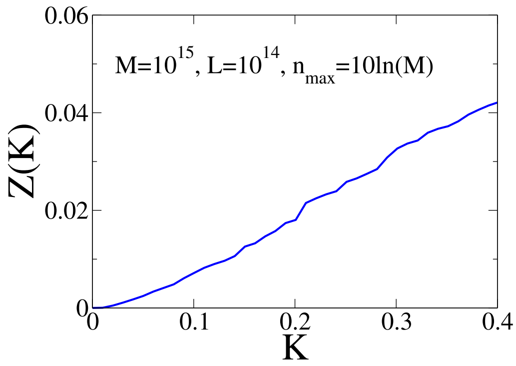

Figure 7. The cumulative intensity function $Z(K)$ of the primes as obtained from the truncated version of (58).

$$\begin{aligned}\tau &\sim \frac{1}{N_s}\sum_{n|N_s}\left(\frac{N}{\phi(n)^2}-1\right)^2 \#\{j \text{ s.t. } N_s/\gcd(j,N_s)=n\}\\ &\sim \frac{N^2}{N_s}\sum_{n|N_s}\frac{1}{\phi(n)^4}\phi(n)\\ &\sim \rho^2 L\sum_{n|N_s}\frac{1}{\phi(n)^3}.\end{aligned}$$

The fact that the last sum is convergent proves proposition 3. Alternatively, one could use expression (35) for τ and substitute the value predicted by Hardy–Littlewood for $g_2(2j)$.

In the case of the integer lattice, τ can be calculated directly from its definition (36). Since of the N_s-1 terms in the sum in (36), $1/f-1$ terms involve $S(k)$ at Bragg peak locations, while the remaining terms involved $S(k)$ values away from Bragg peaks, we have

$$\tau = \frac{1}{N_s}\left[\left(\frac{1}{f}-1\right)(N-1+f)^2+\left(N_s-\frac{1}{f}\right)(0-1+f)^2\right]. \tag{67}$$

For sufficiently large systems, the contribution from the second term is negligible, and N is much larger than $1-f$, yielding

$$\tau \approx \frac{1}{N_s}\left(\frac{1}{f}-1\right)N^2. \tag{68}$$

Plugging in $N_s = L/2$, $f = 2\rho$, and $N = \rho L$, one gets

$$\begin{aligned}\tau &\approx \frac{2}{L}\left(\frac{1}{2\rho}-1\right)(\rho L)^2\\ &= L\rho(1-2\rho),\end{aligned}$$

which is consistent with equation (64) for the primes when the density is fixed.

We have numerically computed τ for the primes and integer lattice by generating such configurations, sampling for their structure factors, and then computing their corresponding values of τ using relation (36). For the primes, we take $M \approx 10^8$, and therefore $f = 2\rho \approx 2/\ln(M) \approx 0.11$. For the integer lattice, we take $f = 0.1$. The constant c is 18.00 for

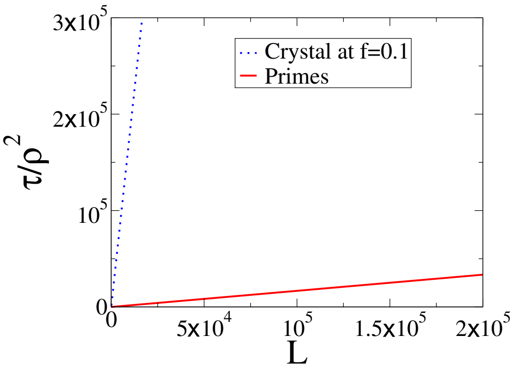

Figure 8. The order metric τ, defined in equation (36), in the discrete setting as a function of the system size L for the primes and integer lattice with filling fraction $f = 0.1$ obtained by direct simulations from equations (36) and (64).

the integer lattice and 0.1674 for the primes. These results are plotted in figure 8. Notice that for a single configuration of an uncorrelated lattice-gas, τ does not grow with L and converges in probability to a constant of order unity (see discussion under equation (36)). This means that the primes in these prescribed intervals are substantially more ordered than the uncorrelated lattice gas and appreciably less ordered than an integer lattice.

Now we study the τ order metric of prime-number configurations with different M and L, obtained by computing their structure factors and evaluating (36). As figure 9 shows, the constant-τ level curves have the form $L \sim \ln^2(M)$, i.e. level curves appear as quadratic curves in the log plot depicted in figure 9. This follows from the fact that τ is proportional to $\rho^2 L$, the density ρ being given by the prime number theorem as $1/\ln(M)$. Thus, $L \sim \ln^2(M)$ is the boundary between regions where primes can be considered to be uncorrelated versus those where they exhibit correlations.

The regime $L \sim \ln(M)^2$ marks another important dividing line. Consider the Cramér model where one replaces the number of primes from M to $M + L$ by a sum of random variables taking the values 0 and 1 with probabilities chosen to match the prime number theorem. A sum of L such random variables will have fluctuations on the order of $L^{1/2}$, by the Central Limit Theorem or by an elementary calculation, while the main term is $L/\ln(M)$. Thus, at least for the random model, $L \sim \ln(M)^2$ is a turning point between short intervals, where $L/\ln(M)^2 \to 0$ and the fluctuations overwhelm the prime number theorem, and longer intervals, where $L/\ln(M)^2 \to \infty$ and the fluctuations can be neglected. Selberg [55], assuming the Riemann hypothesis, proved that for $L/\ln(M)^2 \to \infty$, the number of primes from M to $M + L$ is $L/\ln(M)$ as predicted by the prime number theorem, except possibly for a sparse sequence of exceptional values of M. One might guess that this holds without exception, but Maier proved that there are infinitely many such M [56]. For any power $c > 1$, and setting $L = \ln(M)^c$, Maier proves

$$\limsup_{M\to\infty} \frac{N}{L/\ln(M)} > 1 > \liminf_{M\to\infty} \frac{N}{\ln(M)}. \tag{69}$$

The behavior of primes in intervals of this length is thus a significant departure from the random model. For the uncorrelated regime in which Gallagher's results apply, $L \sim \ln(M)$, and τ will diminish as M increases. As L increases, prime-number configurations move from the uncorrelated regime ($\tau \sim 1$, $L \leqslant \ln^2(M)$) to the effective limit-periodic regime we studied in

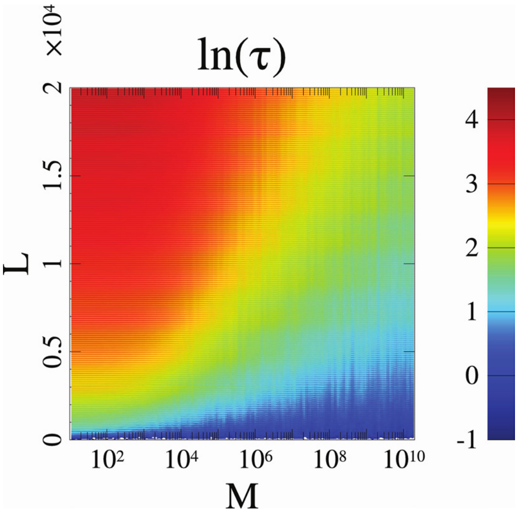

Figure 9. Natural logarithm of the order metric τ, defined in equation (36), of prime numbers for $10 < M < 2 \times 10^{10}$ and $8 < L < 2 \times 10^4$ obtained from numerical simulations.

this paper ($\tau \sim L$, $L \propto M$), and then to the inhomogeneous regime where the density gradient is no longer negligible (e.g. if $L \sim M^{1+\epsilon}$ and $\epsilon > 0$.) In this last phase, the number variance $\sigma^2(R)$ grows faster than the window volume (i.e. faster than R), which is the diametric opposite of hyperuniformity.

7. Classification of the primes

Using equation (44), we have shown that the primes are like a limit-periodic point process, i.e. they are characterized by a structure factor $S(k)$ with dense Bragg peaks at certain rational values of k/π derived from an infinite union of periodic systems in the limit as $M \to \infty$ with $L \sim \beta M$. This is similar to the structure factor of the limit-periodic systems; see section 3. However, the primes show an erratic pattern of occupied and unoccupied sites, very different from the predictable and orderly patterns of standard limit–periodic systems. Hence, the primes represent the first example of a point process that is *effectively* limit-periodic. The reader is referred to [48] for a schematic that illustrates how primes in arithmetic progressions in sufficiently long intervals lead to a superposition of multiple effective periodicities and hence provides a qualitative explanation for the effective limit-periodic behavior of the primes encoded in the structure factor. Each periodicity corresponds to an arithmetic progression, and the primes in a sufficiently long interval compared to the modulus of the progression will show only lower-order deviations from periodicity. The superposition of these approximate periods provides a qualitative explanation for the effective limit-periodic behavior of the primes encoded in the structure factor. It is to be contrasted with the period-doubling chain, which is a superposition of exact periods.

We have also demonstrated that the primes are hyperuniform of class II. Interestingly, this is precisely the same hyperuniformity class to which the (normalized) zeros of the Riemann zeta function belong. However, the primes are substantially more ordered, having dense Bragg peaks instead of a continuous structure factor. As a result, one can only claim that $S(k) \to 0$ in a cumulative sense, unlike the case of the zeros. While the Riemann zeros are disordered and hyperuniform, the primes are effectively limit-periodic and hence are characterized by order on all length scales. In terms of τ, the primes are substantially more ordered than the uncorrelated lattice gas and appreciably less ordered than an integer lattice, but similar in order to the period-doubling chain. It should not go unnoticed that the lack of multiscale order in the Riemann zeros is reflected in the fact that τ is bounded in the large-L limit. Indeed, assuming Montgomery's pair correlation conjecture, it converges to the constant 2/3 (which in the continuous setting should be compared to $\tau = 0$ for the Poisson point process; see discussion under equation (34)). This value 2/3 is what one finds by substituting the conjectured form for the structure factor (see figure 1) in the definition (34) of τ in the continuous setting and integrating $(k/(2\pi) - 1)^2$ over $0 < k < 2\pi$. In a system with multiscale order, such as the primes in dyadic intervals, τ diverges with L.

The condition that $L \sim M$, under which we have shown heretofore unanticipated order in the primes, is to be contrasted with the regime $L \sim \ln(M)$, in which the primes follow Gallagher's uncorrelated behavior. We have also shown that when L grows faster than M, the primes enter the inhomogeneous regime where the density gradient is no longer negligible and hence where the limit-periodicity breaks down.

8. Value distribution of $S(k)$

In this section, we study how frequently the structure factor $S(k)$ exceeds a given threshold t. Here, we apply (8) and let

$$\lambda(t) = \frac{1}{\pi}|\{0 \leqslant k \leqslant \pi;\ S(k) > t\}| \tag{70}$$

be the measure of the set where $S(k) > t$, relative to the total length of the interval $0 < k < \pi$. We think of $\lambda(t)$ as a cumulative distribution function (CDF). The quantity $\lambda(t)$ depends on M, not only t, but we suppress this dependence in the notation. There is also a related quantity which measures how often $S(k) > t$ while excluding forward scattering:

$$\lambda_-(t) = \left|\left\{\frac{2\pi}{L} < k < \pi - \frac{2\pi}{L};\ S(k) > t\right\}\right|. \tag{71}$$

We have

$$\lambda_-(t) \leqslant \lambda(t) \leqslant \lambda_-(t) + 2\pi/L, \tag{72}$$

so the difference between $\lambda(t)$ and $\lambda_-(t)$ is negligible in the limit of large L. By orthogonality of the exponentials $\mathrm{e}^{\mathrm{i}kp}$, we have

$$\int_0^\pi S(k)\mathrm{d}k = \pi. \tag{73}$$

Since $S(k) \geqslant 0$, it follows that

$$\pi = \int_0^\pi S(k)\mathrm{d}k \geqslant \int_{\{S(k)>t\}} S(k)\mathrm{d}k \geqslant t\pi\lambda(t). \tag{74}$$

Thus we have an upper bound

$$\lambda(t) \leqslant \frac{1}{t}. \tag{75}$$

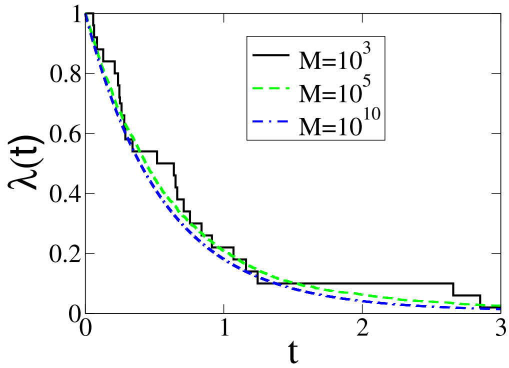

Figure 10. The measure $\lambda(t)$ for relatively small range of values of the threshold t for several values of M. In all cases, $L = 0.1M$. The curves for different values of M converges to an exponential curve well before the largest value of $M = 10^{10}$ is attained. The latter case is well described by the exponential $\exp(-1.779\,89t)$.

For $t > 1$, this is an improvement over the trivial bound $\lambda(t) \leqslant 1$. For small t, figure 10 suggests that $\lambda(t)$ is close to e^{-ct} for some $c > 0$. If one imagines $S(k)$ as a small, noisy contribution on top of the peaks in its limiting form, this suggests that the noise follows an exponential distribution. On the other hand, figure 11 suggests that the upper bound (75) is the correct order of magnitude of $\lambda(t)$ for large t. For t on the order of N, the only way to have $S(k) > t$ is for k to be close to a peak. This leads to the 'heavy-tailed' power-law behavior illustrated in figure 11.

9. Reconstruction of the primes

An interesting question is to what extent we can reconstruct prime configurations from our results in Fourier space? The structure factor is directly related to the modulus of the the complex collective density variable $\tilde{\eta}(k)$ (see (30)) but not its phase. Hence, Fourier inversion of $S(k)$ does not allow us to determine a unique translation of a particle configuration, since a translation of the entire configuration affects the phase of $\tilde{\eta}(k)$ but not its modulus. In a recent work [57], both the modulus and phase of $\tilde{\eta}(k)$ was used to reconstruct one-dimensional point configurations with small N in continuous space. Their algorithm was purposely designed to understand the nature of the inversion procedure for very small particle configurations. We require a different algorithm in order to treat the discrete nature of the odd integers and to reconstruct configurations with many primes.

Using (48), $\tilde{\eta}(k)$ for the primes in a finite interval $[M, M+L)$ is given by the following formula:

$$\tilde{\eta}(k) = \sum_{M \leqslant p < M+L} \exp(-\mathrm{i}kp) \approx \frac{1}{\ln(M)} \frac{\mu(q)}{\phi(q)} \sum_{n=M+1}^{M+L} \mathrm{e}^{-\mathrm{i}(k-2\pi a/q)n}, \tag{76}$$

which we can employ, in principle, to carry out the reconstruction. However, in practice, we do not use the right-most side of (76) for this purpose, since a summation over the entire interval is computationally too time-consuming to calculate for large intervals. Instead, we obtain an approximation to $\tilde{\eta}(k)$ using the following steps:

(i) Calculate $N = (M+L)/\ln(M+L) - M/\ln(M)$. By the Prime Number Theorem, N is the approximate number of primes in this interval.

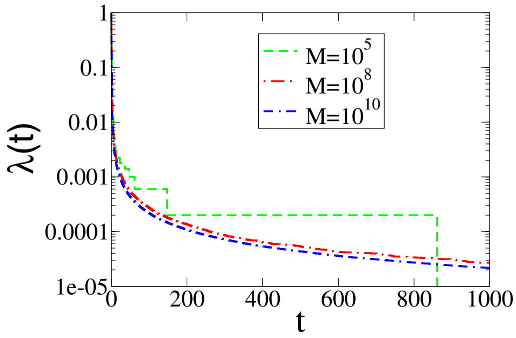

Figure 11. The measure $\lambda(t)$ for relatively large range of values of the threshold t for several values of M. In all cases, $L = 0.1M$. Instead of the exponential curve found in the case shown in figure 10, here $\lambda(t)$ obeys an inverse power law. The curve for $M = 10^{10}$ is well fit by the function $0.021\,74/t$, which we see satisfies the upper bound (75).

(ii) Initialize $\tilde{\eta}(k)$ at $k = (2\pi)/L, 2(2\pi)/L, 3(2\pi)/L, \cdots, (L-1)(2\pi)/L$ to be zero. Set $\tilde{\eta}(0) = N$. We will only need $\tilde{\eta}(k)$ values at these discrete k points for the inverse Fourier transformation.

(iii) As in equation (52) for the the structure factor of a finite system, we account for the peaks only at frequencies $\pi m/n$ with $n \leqslant n_{\max}$ and find all odd square-free numbers for such n.

(iv) For each n, find all integers m such that $0 < m < n$ and m is co-prime with n.

(v) For each n and m, we need to reconstruct a peak at $k = m\pi/n$. For the purpose of reconstructing this peak, the prime-number configuration can be treated as a periodic system of period $2n$. Consider an odd integer $x \in [M, M+L)$, the corresponding site may be occupied only if $\gcd(x, n) = 1$. Thus, if we divide the interval $[M, M+L)$ into $N_c = \lfloor L/(2n) \rfloor$ chunks of length $2n$ (plus a remainder of length $L - 2nN_c$ if L is not divisible by $2n$), each chunk contains $\phi(n)$ such allowed sites. Find the total number of such allowed sites, $N_a = \phi(n)N_c + N_r$, where N_r is the number of allowed sites in the remaining $L/2 - nN_c$ sites. The occupation fraction of allowed sites is thus N/N_a. Assuming equidistribution of primes on allowed sites (which is exact only in the infinite-size limit [48]), calculate the contribution to $\tilde{\eta}(k)$ from the first chunk, i.e. first n sites:

$$C_1 = \frac{N}{N_a} \sum_j^{\times} \exp(-\mathrm{i}kj), \tag{77}$$

where the index j runs over odd integers from M to $M + 2n - 2$ such that $\gcd(j, n) = 1$. The contribution from the jth chunk to $\tilde{\eta}$ is thus $C_j = f^{\,j-1} C_1$, where $f = \exp(-\mathrm{i}k2n)$ is the phase change between the contribution from the first chunk and that from the second chunk. Similarly, also calculate the contribution $\tilde{\eta}(k)$ from the remaining sites:

$$C_r = \frac{N}{N_a} \sum_j^{\times} \exp(-\mathrm{i}kj), \tag{78}$$

where the index j runs over odd numbers from $M + 2nN_c$ to $M + L - 2$ such that $\gcd(j, n) = 1$.

(vi) Whereas the infinite-system limiting form of $\tilde{\eta}$ has perfect Dirac peaks, the actual peaks for finite L have width proportional to $1/L$. The appropriate adjustment for this depends on whether L is divisible by n, since then a whole number of wavelengths pass from M

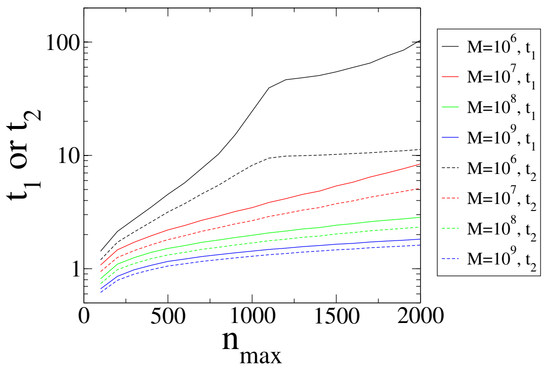

Figure 12. Two measures of the accuracy of the predicted prime numbers, t_1 and t_2, defined by (81) and (82), respectively, versus $n_{\max}$.

to $M + L$. If L is divisible by n, then the peak affects $\tilde{\eta}(k)$ at $k = m\pi/n$, but not nearby k points. Moreover, in this case $f = 1$ and $C_r = 0$. We thus increase $\tilde{\eta}(k)$ at $k = m\pi/n$ by

$$C_1 + C_1 f + C_1 f^2 + \cdots + C_1 f^{N_c - 1} + C_r = N_c C_1. \tag{79}$$

However, if L is not divisible by n, then the peak also affects $\tilde{\eta}(k)$ at all nearby k points. In this case $f \neq 1$, and we increase $\tilde{\eta}(k)$ of affected k points by

$$C_1 + C_1 f + C_1 f^2 + \cdots + C_1 f^{N_c - 1} + C_r = C_1 \left(\frac{1 - f^{N_c}}{1 - f} \right) + C_r. \tag{80}$$

The increment of $\tilde{\eta}(k)$ at nearby k points is truncated when the result from equation (80) is smaller than $\alpha\sqrt{N}$. In practice, we use $\alpha = 1$.

(vii) Perform an inverse Fourier transform of $\tilde{\eta}(k)$ to find $\eta(r)$.

If we had a completely accurate prediction of $\tilde{\eta}(k)$, the resulting local density $\eta(r)$ would be exactly one for each prime number and exactly zero for each composite number. The predicted $\tilde{\eta}(k)$ is not completely accurate. A major reason for the inaccuracy is that there should be infinitely many peaks, but we only consider a finite number of peaks for which $n < n_{\max}$. Another inaccuracy comes from assuming that primes are uniformly distributed in all allowed sites, which is not exact for finite L. Therefore, the resulting $\eta(r)$ is not exactly zero or one. We find N numbers with the highest predicted $\eta(r)$ and predict those numbers to be prime. Thus another source of error is that we always predict N primes, whereas the true number of primes may be more or less than the estimate from the prime number theorem.

We have performed such reconstructions for $L = 510\,510$ and several different M's and $n_{\max}$'s. Let N_c be the number of correctly predicted primes, N_i be the number of incorrect predictions, and N_u be the number of un-predicted primes in this interval. We define the following ratios involving these quantities to measure the accuracy of the reconstruction procedure:

$$t_1 = \frac{N_c}{N_i}, \tag{81}$$

$$t_2 = \frac{N_c}{N_u}. \tag{82}$$

These two ratios versus $n_{\max}$ are plotted in figure 12. We see that for smaller M's and larger $n_{\max}$'s, this reconstruction procedure is highly accurate. When $M = 10^6$ and $n_{\max} = 2000$,

Table 1. Percent accuracy for the prime-number reconstruction process for $L = 510\,510$ for various values of M and $n_{\max}$. Here percent accuracy is defined as $(N_s - N_i - N_u)/N_s$, i.e. the number of correct predictions that a site is either prime or composite divided by the total number of sites.

M	$n_{\max}$	Percent accuracy
10^6	500	0.942
	1000	0.979
	2000	0.987
10^7	500	0.919
	1000	0.940
	2000	0.967
10^8	500	0.912
	1000	0.925
	2000	0.941
10^9	500	0.911
	1000	0.920
	2000	0.931

more than 99% of the predicted prime numbers turn out to be correct. Unfortunately, as M increases, the accuracy declines. For any M, increasing $n_{\max}$ improves the accuracy, but with additional computational cost.

Another measure of the prediction accuracy is the probability of correctly predicting the primality of any site (any odd integer), which is tabulated in table 1 for $L = 510\,510$ and various values of $n_{\max}$ and M. Consistent with the results reported for t_1 and t_2, we see that the accuracy improves for any M as $n_{\max}$ increases, ranging from 93% to about 99% accuracy as M decreases.

We emphasize that an $n_{\max}$ of order $\ln M$ already yields nontrivial results. To find primes between M and $2M$ naively by trial division, one would need to know the primes up to $\sqrt{M}$, much larger than our logarithmic $n_{\max}$. The fact that the first $\ln M$ primes are enough to predict primes in the long and distant interval from M to $2M$, even imperfectly, is an interesting form of long-range order. Bertrand famously postulated that there is always a prime number between M and $2M$, for any value of M. This was proved by Chebyshev using his explicit bounds towards the prime number theorem, but it remains a difficult algorithmic problem to find primes between M and $2M$ quickly when M is large. Our reconstruction procedure does give an algorithm to do so. However, this is not a polynomial-time algorithm. The reason is that the Fast Fourier Transform scales linearithmically in the number of sites L, whereas the goal would be complexity polynomial in $\ln(L)$. Step (iii) involves testing whether each number up to $n_{\max}$ is square-free, which up to logarithmic factors takes $O(n_{\max}^{3/2})$ steps (unless we had a faster way to test square-free than using trial division up to $\sqrt{n}$ to factor n). Step (iv) takes $O(n_{\max}^2)$ steps, again up to logarithmic factors, using the Euclidean algorithm to find greatest common divisors quickly. Step (v) entails computing a sum of length n to find C_1 for each peak. This step takes $O(n_{\max}^2 + N)$ operations. In step (vi), we have to check for the cutoff in updating $\tilde{\eta}(k)$ values for each peak, which involves $O(n_{\max}^2)$ cases. If $n_{\max}$ is negligible compared to N, then (vii) is the slowest step, where we perform the inverse Fourier transform of $\tilde{\eta}(k)$. This step scales as $\mathcal{O}(L \ln L)$. Thus, at a computational cost of $L \ln L$ operations, we reconstruct the primes with imperfect accuracy. One could reconstruct with perfect accuracy by simply testing whether each of the L numbers in the interval is prime. Each primality test takes at most on the order of $\ln(M)^6$ operations for AKS as adapted by Lenstra and Pomerance

[58], or $\ln(M)^4$ assuming the Riemann hypothesis to run a deterministic Miller–Rabin test [8, 9]. Thus our method sacrifices accuracy in exchange for a running time faster by $\ln(L)^3$.

It is interesting to note that when our reconstruction algorithm incorrectly predicts a composite number as prime, the composite number is usually 'almost prime' in the sense that all of its prime factors are large. For example, our algorithm incorrectly predicted $1000\,733 = 809 \times 1237$ and $1001\,423 = 887 \times 1129$ as primes.

10. Conclusions and discussion

The prime numbers display a range of behaviors depending on the interval under consideration. For dyadic intervals $[M, M+L]$ with L comparable to M, we have found order across length scales, very different from the seeming randomness on display for smaller L. However, if L were much larger than M, then one would reach the opposite conclusion of a non-hyperuniform system, purely because of the density gradient without reference to further properties specific to the primes. The substantial order is reflected by the existence of dense Bragg peaks, a consequence of the effective limit-periodicity. The order metric τ gives a quantitative sense in which the primes are substantially more ordered than the uncorrelated lattice gas but less ordered than an integer lattice. When L decreases, τ for the primes in this shorter interval becomes closer to that of an uncorrelated system with a transition visible at $L \sim \ln(M)^2$. But the primes in dyadic intervals are hyperuniform, a fact which would seem almost unbelievable without the effective limit-periodic form for the structure factor; see equation (44). Indeed, the primes fall within the same broad hyperuniformity class as the Riemann zeta zeros (see figure 1), but are substantially more ordered, having dense Bragg peaks instead of a continuous structure factor so that, unlike the Riemann zeros, the order metric τ grows with L. These peaks are located at rational wavenumbers $\pi m/n$ with odd, square-free denominator n, and were discovered numerically in [47]. They are explained in terms of the approximately equal distribution of prime numbers across residue classes modulo $2n$. Assuming the Hardy–Littlewood conjecture on prime pairs, the small 'diffuse part' observed numerically in [47] is negligible in the infinite-system limit. The dense peaks exhibited by the primes are a feature shared with some recently studied quasicrystals, some also in class II and others with even smaller density fluctuations, but these other examples have peaks located at irrational wavenumbers [52]. The primes are distinctive in being a superposition of periodic systems subject to irregularities in the distribution of occupied sites, which we call effective limit-periodicity.

The effective limit-periodic form of the structure factor allows one to predict the Hardy–Littlewood constants for the frequency with which p and $p+r$ are both prime. Of course, it remains an open problem to prove a lower bound establishing that there are infinitely many twin primes. As a more tractable open problem, we ask what further patterns can be found by considering three-particle and higher statistics, beyond what we could discern from the two-particle statistics $S(k)$ and τ. This is related to the full Hardy–Littlewood k-tuples conjecture, going beyond the case of $k = 2$ considered here.

Since our analytical formula for the complex density variable $\tilde{\eta}(k)$, defined by (29), contains phase information, one can employ it to reconstruct a prime-number configuration within an interval $[M, M+L]$ with $L \propto M$ by obtaining the inverse Fourier transform of $\tilde{\eta}(k)$. This leads to an algorithm that enables the reconstruction of the primes in such intervals with high accuracy provided that $n_{\max}$ is sufficiently large and M is not too large.

Acknowledgments

We are grateful to Peter Sarnak, Henry Cohn, Joshua Socolar and Percy Deift for valuable discussions. This work was supported in part by the National Science Foundation under Award No. DMR-1714722 and by the Natural Sciences and Engineering Research Council of Canada.

Appendix. Primes in progressions

The basic idea is to decompose with respect to Dirichlet characters modulo q. These are functions $\chi(n)$ defined for integer values of n and characterized by the properties of periodicity and multiplicativity, namely $\chi(n+q)=\chi(n)$ and $\chi(ab)=\chi(a)\chi(b)$. They are the natural harmonics to use in a situation with period q that preserves multiplicative structure, such as remainders after division by q. One character χ_0 is distinguished, given by $\chi_0(n)=1$ for $\gcd(n,q)=1$ and $\chi_0(n)=0$ in case $\gcd(n,q)>1$. As a rule of thumb, χ_0 provides the main term and we must endeavor to show that the contributions from other characters χ are negligible. The phase of $\mathrm{e}^{an/q}$ can be expanded as a sum over Dirichlet characters χ modulo q. Assuming $\gcd(an,q)=1$, we have

$$\mathrm{e}^{an/q}=\frac{1}{\phi(q)}\sum_{\chi}G(\overline{\chi})\chi(ap). \tag{A.1}$$

The sum is over all Dirichlet characters modulo q. If $\gcd(an,q)>1$, then the sum is 0. We recommend [44] for the theory of Dirichlet characters as well as the Gauss sum $G(\overline{\chi})$. By definition,

$$G(\chi)=\sum_{m=1}^{q}\chi(m)\mathrm{e}^{2\pi\mathrm{i}m/q}, \tag{A.2}$$

and it is important to note that $|G(\chi)|\leqslant q^{1/2}$ while $G(\chi_0)=\mu(q)$ is the Möbius function. Since n is a prime power and a has no factor in common with q, $\gcd(an,q)=1$ does hold unless n is a power of a prime dividing q. Factoring q as $q=p_1p_2\cdots p_w\geqslant 2^w$ shows that there are at most $\ln(q)$ primes dividing q. For each such p, a prime power $n=p^l$ will be less than $M+L$ for $l\lesssim\ln(M)$. In our range, with a value of $q_{\max}$ much less than M, we thus have $\gcd(an,q)=1$ except for at most $\ln(M)^2$ terms n. With an error no worse than a power of $\ln(M)$, we may thus ignore these exceptions, proceeding as if (A.1) held for all n. Writing $\alpha=k/\pi=a/q+t$ and summing over n gives

$$\sum_{n=M+1}^{M+L}\mathrm{e}^{\alpha n}\Lambda(n)=\frac{1}{\phi(q)}\sum_{\chi}\tau(\overline{\chi})\chi(a)\sum_{n=M+1}^{M+L}\mathrm{e}^{nt}\chi(n)\Lambda(n). \tag{A.3}$$

Summing by parts, we have

$$\sum_{n\leqslant Y}\mathrm{e}^{nt}\chi(n)\Lambda(n)=\mathrm{e}^{Yt}\psi(Y,\chi)-2\pi\mathrm{i}t\int_1^Y\mathrm{e}^{ut}\psi(u,\chi)\mathrm{d}u, \tag{A.4}$$

where, for an upper limit u and a character χ to modulus q,

$$\psi(u,\chi)=\sum_{n<u}\Lambda(n)\chi(n). \tag{A.5}$$

By the prime number theorem in progressions (p 125, 132 of [44]), there is a positive $c > 0$ such that for all characters to moduli $q \leqslant \ln(L)^B$, $\psi(u,\chi_0) = u + O\big(\exp(-c\sqrt{\ln(u)})\big)$ while for nontrivial χ, $\psi(u,\chi) = O\big(\exp(-c\sqrt{\ln(u)})\big)$. This leads to the error term stated in (48) (and the Riemann hypothesis would imply an estimate for an even larger range of q).

ORCID iDs

Salvatore Torquato https://orcid.org/0000-0003-4614-335X

References

[1] Gallagher P X 1976 On the distribution of primes in short intervals *Mathematika* **23** 4–9
[2] Torquato S and Stillinger F H 2003 Local density fluctuations, hyperuniform systems, and order metrics *Phys. Rev.* E **68** 041113
[3] Hadamard J 1896 Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques *Bull. Soc. Math. France* **24** 199–220
[4] Hardy G H and Littlewood J E 1923 Some problems of partitio numerorum; III: on the expression of a number as a sum of primes *Acta Math.* **44** 1–70
[5] Cherwell L 1946 On the distribution of the intervals between prime numbers *Q. J. Math.* **17** 46–62
[6] Pintz J 2007 Cramér versus Cramér. On Cramér's probabilistic model for primes *Funct. Approx. Comment. Math.* **37** 361–76
[7] Granville A 1995 Harald Cramér and the distribution of prime numbers *Scand. Actuarial J.* **1995** 12–28
[8] Miller G 1976 Riemann's hypothesis and tests for primality *J. Comput. Syst Sci.* **13** 300–17
[9] Rabin M 1980 Probabilistic algorithm for testing primality *J. Number Theory* **12** 128–38
[10] Pomerance C, Selfridge J L and Wagstaff S S 1980 The pseudoprimes to 25 10? *Math. Comput.* **35** 1003–26
[11] Baillie R and Wagstaff S S 1980 Lucas pseudoprimes *Math. Comput.* **35** 1391–417
[12] Atkin A O L and Morain F 1993 Elliptic curves and primality proving *Math. Comput.* **61** 29–68
[13] Rubinstein M and Sarnak P 1994 Chebyshevs bias *Exp. Math.* **3** 173–97
[14] Martelli F 2013 Dealing with primes I.: on the goldbach conjecture (arXiv:1309.5895 [math.NT])
[15] Vinogradov I M 1947 The method of trigonometrical sums in the theory of numbers *Trudy Mat. Inst. Steklov.* **23** 3–109
[16] Green B and Tao T 2008 The primes contain arbitrarily long arithmetic progressions *Ann. Math.* **167** 481–547
[17] Green B and Tao T 2010 Linear equations in primes *Ann. Math.* **171** 1753–850
[18] Green B and Tao T 2012 The Mobius function is asymptotically orthogonal to nilsequences *Ann. Math.* **175** 541–66
[19] Tao T and Ziegler T 2012 The inverse conjecture for the gowers norm over finite fields in low characteristic *Ann. Comb.* **16** 121–88
[20] Zhang Y 2014 Bounded gaps between primes *Ann. Math.* **3** 1121–74
[21] Lemke Oliver R J and Soundararajan K 2016 Unexpected biases in the distribution of consecutive primes *Proc. Nat. Acad. Sci.* **113** E4446–E4454
[22] Maynard J 2015 Small gaps between primes *Ann. Math.* **181** 383–413
[23] Dahmen S R, Prado S D and Stuermer-Daitx T 2001 Similarity in the statistics of prime numbers *Physica* A **296** 523–8
[24] Wolflf M 1996 Unexpected regularities in the distribution of prime numbers *Proc. 8th Joint EPS—APS Int. Conf. Physics Computing*
[25] Montgomery H L 1973 The pair correlation of zeros of the zeta function *Analytic Number Theory* (*Proc. Sympos. Pure Math., St. Louis, 1972*) vol 22 (Providence, RI: AMS) pp 181–93
[26] Montgomery H L and Soundararajan K 2004 Primes in short intervals *Commun. Math. Phys.* **252** 589–617
[27] Dyson F J 1962 Statistical theory of the energy levels of complex systems. I *J. Math. Phys.* **3** 140–56
[28] Mehta M L 1991 *Random Matrices* (New York: Academic)

[29] Rudnick Z and Sarnak P 1996 Zeros of principal *L*-functions and random matrix theory *Duke Math. J.* **81** 269–322
[30] Torquato S, Scardicchio A and Zachary C E 2008 Point processes in arbitrary dimension from Fermionic gases, random matrix theory, and number theory *J. Stat. Mech.* P11019
[31] Donev A, Stillinger F H and Torquato S 2005 Unexpected density fluctuations in disordered jammed hard-sphere packings *Phys. Rev. Lett.* **95** 090604
[32] Batten R D, Stillinger F H and Torquato S 2008 Classical disordered ground states: super-ideal gases, and stealth and equi-luminous materials *J. Appl. Phys.* **104** 033504
[33] Zachary C E and Torquato S 2009 Hyperuniformity in point patterns and two-phase heterogeneous media *J. Stat. Mech.* P12015
[34] Florescu M, Torquato S and Steinhardt P J 2009 Designer disordered materials with large complete photonic band gaps *Proc. Natl Acad. Sci.* **106** 20658–63
[35] Zachary C E, Jiao Y and Torquato S 2011 Hyperuniform long-range correlations are a signature of disordered jammed hard-particle packings *Phys. Rev. Lett.* **106** 178001
[36] Jiao Y, Lau T, Hatzikirou H, Meyer-Hermann M, Corbo J C and Torquato S 2014 Avian photoreceptor patterns represent a disordered hyperuniform solution to a multiscale packing problem *Phys. Rev.* E **89** 022721
[37] Jack R L, Thompson I R and Sollich P 2015 Hyperuniformity and phase separation in biased ensembles of trajectories for diffusive systems *Phys. Rev. Lett.* **114** 060601
[38] Torquato S, Zhang G and Stillinger F H 2015 Ensemble theory for stealthy hyperuniform disordered ground states *Phys. Rev.* X **5** 021020
[39] Mayer A, Balasubramanian V, Mora T and Walczak A M 2015 How a well-adapted immune system is organized *Proc. Natl Acad. Sci.* **112** 5950–5
[40] Ma T, Guerboukha H, Girard M, Squires A D, Lewis R A and Skorobogatiy M 2016 3d printed hollow-core terahertz optical waveguides with hyperuniform disordered dielectric reflectors *Adv. Opt. Mater.* **4** 2085–94
[41] Goldfriend T, Diamant H and Witten T A 2017 Screening, hyperuniformity, and instability in the sedimentation of irregular objects *Phys. Rev. Lett.* **118** 158005
[42] Hexner D, Chaikin P M and Levine D 2017 Enhanced hyperuniformity from random reorganization *Proc. Natl Acad. Sci.* **114** 4294–9
[43] Torquato S 2018 Hyperuniform states of matter *Phys. Rep.* **745** 1–95
[44] Davenport H 1980 *Multiplicative Number Theory* vol 74 (New York: Springer)
[45] Tenenbaum G 1995 *Introduction to Analytic and Probabilistic Number Theory* vol 46 (Cambridge: Cambridge University Press)
[46] Iwaniec H and Kowalski E 2004 *Analytic Number Theory* vol 53 (Providence, RI: American Mathematical Society)
[47] Zhang G, Martelli F and Torquato S 2018 Structure factor of the primes *J. Phys. A: Math. Theor.* **51** 115001
[48] Torquato S, Zhang G and de Courcy-Ireland M 2018 Uncovering multiscale order in the prime numbers via scattering *J. Stat. Mech.* 093401
[49] Baake M and Grimm U 2011 Diffraction of limit periodic point sets *Phil. Mag.* **91** 2661–70
[50] Levine D and Steinhardt P J 1986 Quasicrystals. I. Definition and structure *Phys. Rev.* B **34** 596
[51] Bombieri E and Taylor J E 1986 Which distributions of matter diffract? An initial investigation *J. Phys. Coll.* **47** C3–19
[52] Oğuz E C, Socolar J E S, Steinhardt P J and Torquato S 2017 Hyperuniformity of quasicrystals *Phys. Rev.* B **95** 054119
[53] DiStasio R A, Zhang G, Stillinger F H and Torquato S 2018 Rational design of stealthy hyperuniform patterns with tunable order *Phys. Rev.* E **97** 023311
[54] Vaughan R C 1997 *The Hardy–Littlewood Method* vol 125 (Cambridge: Cambridge University Press)
[55] Selberg A 1943 On the normal density of primes in small intervals, and the difference between consecutive primes *Arch. Math. Naturvidenskab* B **47** 87–105
[56] Maier H 1985 Primes in short intervals *Michigan Math. J.* **32** 221–5
[57] Kim J, Zhang G, Stillinger F H and Torquato S 2018 Inversion problems for Fourier transforms of particle distributions *J. Stat. Mech.* 113302
[58] Lenstra H W Jr and Pomerance C 2002 Primality testing with Gaussian periods *Foundations of Software Technology and Theoretical Computer Science* p 1