Discrete charging of metallic grains: statistics of addition spectra

2006 J. Phys. A: Math. Gen. 39 4797

(http://iopscience.iop.org/0305-4470/39/18/002)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 4797–4810 doi:10.1088/0305-4470/39/18/002

Discrete charging of metallic grains: statistics of addition spectra

Yshai Avishai[1], **Daniel Berend**[2,4] **and Luba Bromberg**[3]

[1] Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
[2] Departments of Mathematics and Computer Science, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
[3] Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

Received 27 July 2005, in final form 20 January 2006
Published 19 April 2006
Online at stacks.iop.org/JPhysA/39/4797

Abstract

We analyse the statistics of electrostatic energies (and their differences) for a quantum dot system composed of a finite number K of electron islands (metallic grains) with random capacitance–inductance matrix C, for which the total charge is discrete, $Q = Ne$ (where e is the charge of an electron and N is an integer). The analysis is based on a generalized charging model, where the electrons are distributed among the grains such that the electrostatic energy $E(N)$ is minimal. Its second difference (inverse compressibility) $\chi_N = E(N+1) - 2E(N) + E(N-1)$ represents the spacing between adjacent Coulomb-blockade peaks appearing when the conductance of the quantum dot is plotted against gate voltage. The statistics of this quantity for *single* grain quantum dots has been the focus of experimental and theoretical investigations during the last two decades. In the more general case of quantum dots composed of *several* grains, we provide an algorithm for calculating the distribution function corresponding to χ_N and show that this function is piecewise polynomial.

PACS numbers: 05.30.Fk, 05.45.Mt, 41.20.Cv, 73.21.La

1. Introduction

The physics exposed in the addition spectra of quantum dots is rather rich, and hence its investigation is at the focus of both experimental and theoretical studies. After the origin of Coulomb-blockade peaks has been elucidated, investigation is directed towards more subtle questions such as their heights, widths and spacings. The underlying physics is related to the ground-state energy, chemical potential and inverse compressibility of quantum dots composed of a few metallic electron islands coupled capacitively and inductively to each other.

[4] Supported in part by the Israel Science Foundation, Grant 186/01.

The problem of understanding the observed nearest-neighbour Coulomb-blockade peak spacing distribution is of course not new, but its study is focused mainly on *single* grain quantum dots. Theoretical investigations pertaining to diffusive or chaotic quantum dots reveal a marked deviation from Gaussian distribution [1]. Experimentally, it is confirmed that, even for such relatively simple systems, the distribution displays a long tail towards small spacing [2]. Distribution of peak spacings in a magnetic field (which is outside the scope of the present work) has recently been investigated [3].

The present work concentrates on the distribution of spacings between Coulomb-blockade peaks in large semiconductor quantum dots. In particular, we are interested in fluctuations of these quantities with the number N of electrons on the dot. The main problem can be stated as follows: according to the simplest picture (charging model), in which the quantum dot is regarded as a single-electron island whose coupling with the leads is through its capacitance C, the total potential energy of a quantum dot with N electrons and charge $Q = Ne$ is $Q^2/2C - V_g Q$, where V_g is the gate voltage and e is the electron charge. The position of the Nth Coulomb-blockade peak occurs at a gate voltage $V_g = \frac{Ne^2}{C}$. This peak position is then a linear function of N, and therefore the spacing between two adjacent peaks should be a constant e^2/C, independent of N. This is not always confirmed experimentally as we have indicated above [2] even for a single metallic grain. The situation is even more intriguing if the quantum dot is large and might contain more than a single-electron puddle. As indicated in a series of recent experiments [4], the spacing between adjacent Coulomb-blockade peaks occasionally vanishes, namely, Coulomb-blockade peaks tend to bunch. The problem is therefore to explain why the results predicted from a simple charging model deviate substantially from the experimental observation.

In [5], a generalized charging model has been tested, where it is assumed that the large dot used in the experiments [4] could be divided into a set of potential wells (metallic grains) with random capacitances and random mutual inductances. This casts the question of Coulomb-blockade peak spacing distribution into the problem of elucidating the statistics of the addition spectrum of a relatively simple physical system. It consists of K metallic grains (or capacitors), such that the number of electrons on the ith grain is n_i $(i = 1, 2, \ldots, K)$, the total number of electrons being N. The charging model for such a system (at zero temperature) is based on the assumption that the distribution of electrons among the grains is determined by requiring that the electrostatic energy $E(N)$ of a dot containing N electrons is minimal. It is useful at this point to recall the basic facts pertaining to the energy of the electrostatic field of conductors [7, pp 3–7]. The electrostatic energy of the system is a bilinear form in the numbers n_i. This form is given by a $K \times K$ matrix $W = \frac{1}{2}C^{-1}$. Here, C is a positive-definite symmetric matrix (with entries c_{ij}; $i, j = 1, 2, \ldots, K$) of capacitance and inductance coefficients. Physically, the matrix C has positive diagonal entries and *negative* (more precisely, non-positive) non-diagonal entries [7, pp 3–7],

$$c_{ij} = c_{ji}, \qquad c_{ii} > 0, \qquad c_{ij} \leqslant 0 \quad (i \neq j). \tag{1}$$

On the other hand, all the elements of C^{-1} are non-negative. More precisely, the entries $w_{ij}, i, j = 1, 2, \ldots, K$, of the matrix W satisfy

$$w_{ii} > 0, \qquad w_{ij} \geqslant 0 \quad (i \neq j). \tag{2}$$

The off-diagonal entries $c_{ij}, i \neq j$, decay as an inverse power of the distance between the grains, while the diagonal entries c_{ii} are proportional to the geometrical size of the grains. The notion of randomness enters when we recall that, experimentally, the sizes of the grains, as well as the distances between them, are random quantities. This means that the elements of the matrix C are random numbers (subject, of course, to the required symmetries (1)). The

spacing between Coulomb-blockade peaks is equal to the second difference of the ground-state energy. In other words, the distribution of spacing peaks is determined by the statistics of the inverse compressibility,

$$\chi_N \equiv E(N+1) - 2E(N) + E(N-1). \tag{3}$$

When two Coulomb-blockade peaks coincide, the second difference χ_N vanishes. Note that, on the average (and on a large scale), the energy $E(N)$ grows quadratically with N. Therefore, one would expect the second difference to stay finite and N independent. However, there is no simple relation such as $E(N) = aN + bN^2$. The deviation of $E(N)$ from exact quadratic behaviour makes its second difference χ_N non-constant and a fluctuating quantity. It is precisely these fluctuations which we intend to study. As we shall see, the fact that electron charge is quantized makes this task non-trivial.

Having explained the physical motivation, let us recall that the present mathematical problem is also intriguing because it involves the notions of randomness and discreteness. We then pose the mathematical problem: what is the distribution $p(\chi)$ of inverse compressibility for a given system of metallic grains with random capacitive matrix C? As a crude approximation, it was assumed in [6] that the metallic grains are indeed very far apart, and the matrix C is nearly diagonal, its K diagonal entries (capacitances) being random numbers. The energy of the system in the diagonal case is given by

$$E(N) = \min \sum_{i=1}^{K} \frac{1}{2c_i} n_i^2 \qquad \left(\text{subject to } \sum_{i=1}^{K} n_i = N \right). \tag{4}$$

The minimum in (4) is taken over all possible partitions $(n_i)_{i=1}^K$ of N. It was first proved that the minimum problem (4) has the following convenient feature: if $n_1, n_2, \ldots, n_K$ are the argument values bringing $E(N)$ to its minimal value for some N, then the minimum for $N+1$ is obtained by retaining all n_is, except for one which is increased by 1. This allowed an exact determination of the distribution function according to which the sequence χ_N is distributed. For a random set of capacitors ($c_1, c_2, \ldots, c_K$ are the random numbers with probability distribution $P(c_1, c_2, \ldots, c_K)$), the distribution of the inverse compressibility $F(\chi)$ was calculated in [6].

Our next goal is to study this problem for general positive-definite matrices C. The problem turns out to be quite harder. To begin with, it is no longer true that the optimal solution for $N+1$ is obtained in a simple manner from that for N. That is, for each N we need to re-distribute the N electrons between the grains, and it may well happen that, although the n_is grow in general with N, some of them will actually decrease infinitely often as N increases by 1 each time. Namely, there will exist infinitely many values of N for which the optimal value of some n_i decreases as N grows to $N+1$. Thus, the problem entails new behavioural patterns with respect to the diagonal case.

Our main result in this paper is an algorithm for calculating the distribution function corresponding to (χ_N). Moreover, we show that this function is piecewise polynomial. We state the result in section 2. Section 3 is a short digression, discussing a few notions which arise in the proof. The proof of the main theorem is given in section 4. As the algorithm is only implicit in the course of the proof, we summarize it after the end of the proof more explicitly as an algorithm. This is accomplished in section 5.

2. The main results

Mathematically, our problem is as follows. Let $C = (c_{ij})_{i,j=1}^K$ be a positive-definite matrix, with positive diagonal elements and non-positive off-diagonal elements. Assume that the sum

of elements in every row of C is positive and that all entries of $\frac{1}{2}C^{-1} = W = (w_{ij})_{i,j=1}^{K}$ are non-negative. Put

$$E(N) = \min\left\{\sum_{i,j=1}^{K} w_{ij}n_i n_j : n_i \in \mathbf{Z}_+, \sum_{i=1}^{K} n_i = N\right\}, \quad N \in \mathbf{N}, \tag{5}$$

where $\mathbf{N}$ is the set of positive integers and $\mathbf{Z}_+ = \mathbf{N} \cup \{0\}$. We want to understand the statistical behaviour of the sequence $E(N)$, and in particular that of the second difference sequence

$$\chi_N = E(N+1) - 2E(N) + E(N-1). \tag{6}$$

To formulate our main result, we need a few definitions and notation.

Definition 2.1. *Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers and F a distribution function. The sequence (x_n) is asymptotically F-distributed if*

$$\frac{|\{1 \leqslant n \leqslant M : x_n \leqslant x\}|}{M} \underset{M\to\infty}{\longrightarrow} F(x)$$

for every continuity point x of F (where $|S|$ denotes the cardinality of a finite set S).

The definition almost coincides with [8, p 53, definition 7.1], except that there the sequence (x_n) is considered only modulo 1. Note that a sequence need not be asymptotically F-distributed for some F, as the following example shows.

Example 2.1. The sequence of numbers

$$0, \quad \underbrace{1, \ldots, 1}_{10}, \quad \underbrace{0, \ldots, 0}_{10^2}, \quad \underbrace{1, \ldots, 1}_{10^3}, \ldots$$

is not asymptotically F-distributed for any F.

A stronger notion is obtained when we require not only long initial block of the sequence to behave approximately according to F, but rather require any long block to behave so. This leads to the following definition ([8, p 40, definition 5.1] and [8, p 200, definition 3.2]):

Definition 2.2. *In the set-up of definition 2.1, (x_n) is asymptotically well F-distributed if*

$$\frac{|\{L < n \leqslant M : x_n \leqslant x\}|}{M - L} \underset{M-L\to\infty}{\longrightarrow} F(x)$$

for every continuity point x of F.

The following example demonstrates that the property of asymptotic well F-distribution is indeed strictly stronger than that of asymptotic F-distribution.

Example 2.2. The sequence of numbers

$$\underbrace{0}_{1}, \quad \underbrace{1}_{1}, \quad \underbrace{0, 0}_{2}, \quad \underbrace{1, 1}_{2}, \quad \underbrace{0, 0, 0}_{3}, \quad \underbrace{1, 1, 1}_{3}, \ldots$$

is asymptotically F-distributed, where F is the distribution function

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2}, & 0 \leqslant x < 1, \\ 1, & x \geqslant 1, \end{cases}$$

but it is not asymptotically well F-distributed.

Definition 2.3. *A function $g : \mathbf{R} \to \mathbf{R}$ is piecewise polynomial if there exist intervals (finite or infinite) $I_j \subseteq \mathbf{R}$ and polynomials Q_j, $1 \leqslant j \leqslant m$, such that*

$$g(x) = Q_j(x), \qquad x \in I_j, \quad 1 \leqslant j \leqslant m.$$

The degree of g is $\max_{1 \leqslant j \leqslant m} \deg Q_j$.

Returning to our problem, let $b_i = \sum_{j=1}^{K} c_{ij}$, $1 \leqslant i \leqslant K$, be the row sums of the matrix C. Since the matrix C is random, in the generic case the numbers $b_1, b_2, \ldots, b_K$ are linearly independent over the rationals. (That is, considered as vectors in the space $\mathbf{R}$ over the field of rational numbers $\mathbf{Q}$, they are independent.)

Now we can formulate our main result.

Theorem 2.1. *Let C be a positive-definite symmetric matrix, with positive row sums b_1, $b_2, \ldots, b_K$ and let $W = \frac{1}{2}C^{-1}$. Suppose that $b_1, b_2, \ldots, b_K$ are linearly independent over the rationals. Then the sequence $(\chi_N)_{N=1}^{\infty}$ of the second differences, defined via (5) and (6), is asymptotically well F-distributed, where F is a continuous piecewise polynomial function of degree at most $K - 1$, which can be effectively computed.*

As mentioned in the introduction, a phenomenon which occurs in the general case dealt with here, but not in the special case of diagonal matrices C, is that, as we pass from N to $N + 1$, there may be re-distribution of the n_is in the optimal solution. The following example is to that effect.

Example 2.3. Let

$$C = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 3 \end{pmatrix}, \qquad W = \frac{1}{2}C^{-1} = \frac{1}{16}\begin{pmatrix} 5 & 1 & 2 \\ 1 & 5 & 2 \\ 2 & 2 & 4 \end{pmatrix}.$$

Then,

$$E(N) = \tfrac{1}{16} \min_{n_1+n_2+n_3=N, n_i \geqslant 0} \left(5n_1^2 + 2n_1n_2 + 4n_1n_3 + 5n_2^2 + 4n_2n_3 + 4n_3^2\right).$$

Using the techniques in the beginning of section 4, it is easy to verify that the optimal values of n_1, n_2, n_3 are given by

$$(n_1, n_2, n_3) = \begin{cases} \left(\dfrac{N}{3}, \dfrac{N}{3}, \dfrac{N}{3}\right), & N \equiv 0 \,(\mathrm{mod}\, 3), \\ \left(\dfrac{N-1}{3}, \dfrac{N-1}{3}, \dfrac{N+2}{3}\right), & N \equiv 1 \,(\mathrm{mod}\, 3), \\ \left(\dfrac{N+1}{3}, \dfrac{N+1}{3}, \dfrac{N-2}{3}\right), & N \equiv 2 \,(\mathrm{mod}\, 3). \end{cases}$$

Thus, for any non-negative integer K, when passing from $N = 3k+1$ to $N = 3k+2$, the value of n_3 at the optimal point decreases from $k + 1$ to k.

3. Uniform distribution modulo 1

In this section, we briefly discuss the notion of uniform distribution modulo 1 and recall a few related results, which will be needed in the proof of theorem 2.1.

Definition 3.1. *A sequence* $(x_n)_{n=1}^{\infty}$ *of real numbers is uniformly distributed modulo 1 if*

$$\frac{|\{1 \leqslant n \leqslant N : a \leqslant \{x_n\} < b\}|}{N} \underset{N\to\infty}{\longrightarrow} b - a, \quad 0 \leqslant a < b \leqslant 1,$$

where $\{t\}$ *is the fractional part of a real number* t *[8, p 1, definition 1.1].*

In terms of definition 2.1, (x_n) *is uniformly distributed modulo 1 if and only if the sequence* $(\{x_n\})$ *of fractional parts is F-distributed, where F is the distribution function of the uniform distribution on* $[0, 1]$*:*

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leqslant x \leqslant 1, \\ 1, & x > 1. \end{cases} \tag{7}$$

The notion of uniform distribution modulo 1 has a multi-dimensional analogue. A sequence $(\mathbf{x}_n)_{n=1}^{\infty}$ in $\mathbf{R}^s$ is *uniformly distributed modulo 1 in* $\mathbf{R}^s$ if

$$\frac{|\{1 \leqslant n \leqslant N : \mathbf{a} \leqslant \{\mathbf{x}_n\} < \mathbf{b}\}|}{N} \underset{N\to\infty}{\longrightarrow} \prod_{i=1}^{s}(b_i - a_i), \quad \mathbf{0} \leqslant \mathbf{a} < \mathbf{b} \leqslant \mathbf{1},$$

where $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbf{R}^s$, $\mathbf{a} = (a_1, a_2, \ldots, a_s)$, and so forth, and inequalities between vectors in $\mathbf{R}^s$ are to be understood componentwise [8, p 47, definition 6.1].

The notion of uniform distribution modulo 1, both in the one-dimensional and the multi-dimensional cases, has a stronger version, whereby the required property holds not only along initial blocks of the sequence, but along any blocks of larger and larger lengths [8, p 40, definition 5.1]. A sequence satisfying this stronger property is *well distributed modulo 1*. Obviously, well distribution modulo 1 is equivalent in the one-dimensional case to F-distribution of the sequence of fractional parts for the function F given by (7). A basic example of a sequence which is uniformly distributed modulo 1 is $(n\alpha)_{n=1}^{\infty}$, where α is an arbitrary irrational [8, p 8, definition 2.1]. In the multi-dimensional case, the sequence $(n\alpha_1, n\alpha_2, \ldots, n\alpha_s)_{n=1}^{\infty}$ is uniformly distributed modulo 1 in $\mathbf{R}^s$ if and only if the numbers $1, \alpha_1, \alpha_2, \ldots, \alpha_s$ are linearly independent over $\mathbf{Q}$ [8, pp 48–49]. Moreover, for these sequence, well distribution is equivalent to uniform distribution.

Recall that the *density* of a set $A \subseteq \mathbf{N}$ is given by

$$D(A) = \lim_{M\to\infty} \frac{|A \cap [1, M]|}{M}$$

if the limits exists. If, moreover, the limit

$$\mathrm{BD}(A) = \lim_{M-L\to\infty} \frac{|A \cap (L, M]|}{M - L}$$

exists, then it is called the *Banach density* of A.

We can rephrase the definition of uniform distribution modulo 1 using the notion of density of a set. Namely, $(x_n)_{n=1}^{\infty}$ is uniformly distributed modulo 1 if for every interval $I \subseteq [0, 1)$ we have

$$D(\{n : \{x_n\} \in I\}) = |I|, \tag{8}$$

where $|I|$ denotes the length of I. Similarly, $(x_n)_{n=1}^{\infty}$ is well distributed modulo 1 if (8) continues to hold when the density of the left-hand side is replaced by Banach density.

4. Proof of theorem 2.1

To avoid complicated notation, we shall prove in theorem 2.1 only that $(x_N)_{N=1}^{\infty}$ is asymptotically F-distributed and not that it is asymptotically well F-distributed. As will be seen in the proof, our result depends on the fact that the sequence $(\{Nb_1\}, \{Nb_2\}, \ldots, \{Nb_{K-1}\})$ is uniformly distributed modulo 1 in $\mathbf{R}^{K-1}$. Since this sequence is actually well distributed modulo 1, the same proof shows that $(x_N)_{N=1}^{\infty}$ is actually well F-distributed.

Along with the sequence $E(N)$ from (5), it is very useful to consider the sequence $E_1(N)$, defined by

$$E_1(N) = \min\left\{\sum_{i,j=1}^{K} w_{ij}x_ix_j : x_i \in \mathbf{R}, \sum_{i=1}^{K} x_i = N\right\}, \qquad N \in \mathbf{N}. \tag{9}$$

Obviously, $E_1(N) \leqslant E(N)$ for each N. We shall refer to the minimum problems on the right-hand side of (5) and of (9) as the *constrained problem* and the *unconstrained problem*, respectively.

Denote by $\mathbf{e}$ the column K-vector with all entries 1.

Lemma 4.1. *The unique minimum of the unconstrained problem is*

$$\mathbf{x}_0 = \frac{N}{\sum_{i,j=1}^{K} c_{ij}} \cdot C\mathbf{e}.$$

and the corresponding unconstrained minimum is $E_1(N) = \frac{N^2}{2\sum_{i,j=1}^{K} c_{ij}}$.

Due to our assumption regarding the positivity of the row sums of C, all components of $\mathbf{x}_0$ are positive. Multiplying all entries of C by any constant $c > 0$ we obtain an equivalent problem. Taking $c = \left(\sum_{i,j=1}^{K} c_{ij}\right)^{-1}$, *we shall henceforth assume that* $\sum_{i,j=1}^{K} c_{ij} = 1$. In particular, denoting $\mathbf{b} = (b_1, b_2, \ldots b_K)^t$, we have

$$\mathbf{x}_0 = NC\mathbf{e} = N\mathbf{b} \tag{10}$$

and

$$E_1(N) = \frac{N^2}{2}. \tag{11}$$

Proof of lemma 4.1. Let $\mathbf{x} \neq \mathbf{x}_0$ be any feasible solution of the unconstrained problem. Putting $a = \frac{N}{\sum_{i,j=1}^{K} c_{ij}}$ and $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$, we obtain

$$\begin{aligned}\mathbf{x}^tW\mathbf{x} &= (\mathbf{x}_0+\mathbf{y})^tW(\mathbf{x}_0+\mathbf{y}) = \mathbf{x}_0^tW\mathbf{x} + 2\mathbf{x}_0^tW\mathbf{y} + \mathbf{y}^tW\mathbf{y}\\ &= \mathbf{x}_0^tW\mathbf{x}_0 + a\mathbf{1}^tCC^{-1}\mathbf{y} + \mathbf{y}^tW\mathbf{y}\\ &= \mathbf{x}_0^tW\mathbf{x}_0 + a\mathbf{1}^t\mathbf{y} + \mathbf{y}^tW\mathbf{y} = \mathbf{x}_0^tW\mathbf{x}_0 + \mathbf{y}^tW\mathbf{y} > \mathbf{x}_0^tW\mathbf{x}_0.\end{aligned}$$

Consequently,

$$E_1(N) = \frac{N}{\sum_{i,j=1}^{K} c_{ij}}\mathbf{e}^tCW\frac{N}{\sum_{i,j=1}^{K} c_{ij}}C\mathbf{e} = \frac{N^2}{2\sum_{i,j=1}^{K} c_{ij}}.$$

□

Lemma 4.2. $E_1(N) \geqslant E(N) - \sum_{i,j=1}^{K} w_{ij}$ *for every* N.

Proof. Let $\mathbf{x}_0 = (x_{01}, x_{02}, \ldots, x_{0K})$ be the minimum point of the unconstrained problem. Let $r = \sum_{i=1}^{K}\{x_{0i}\}$ be the sum of fractional parts of all coordinates of $\mathbf{x}_0$. Obviously, r is an integer, $0 \leqslant r < K$. Let $i_1, i_2, \ldots, i_K$ be all integers between 1 and K, ordered so that

$\{x_{0,i_1}\} \leqslant \{x_{0,i_2}\} \leqslant \cdots \leqslant \{x_{0,i_K}\}$ (where ties are resolved arbitrarily). Consider the vector $\mathbf{n} = (n_1, n_2, \ldots, n_K)$ defined by

$$n_i = \begin{cases} [x_{0,i}], & i = i_1, i_2, \ldots, i_{K-r}, \\ [x_{0,i}] + 1, & \text{otherwise.} \end{cases}$$

As mentioned in lemma 4.1, all $x_{0,i}$s are positive, and hence $\mathbf{n}$ is a feasible solution of the constrained problem. Set $\mathbf{y} = \mathbf{n} - \mathbf{x}_0$. Since all coordinates of $\mathbf{y}$ lie in the interval $(-1, 1)$, as in the proof of lemma 4.1, we have

$$E(N) \leqslant n^t W \mathbf{n} = x_0^t W \mathbf{x}_0 + y^t W \mathbf{y} \leqslant E_1(N) + \sum_{i,j=1}^{K} w_{ij} \tag{12}$$

which proves the lemma. □

Lemma 4.3. *There exists an effective constant $\Delta = \Delta(C)$ such that, for every N, the distance between the solution of the constrained problem and that of the unconstrained problem does not exceed Δ.*

Proof. Write $W = P^{-1} D P$, where P is orthogonal and D diagonal. Let M be an upper bound on the eigenvalues of C (for example, the L^∞-norm $\max_{1 \leqslant i \leqslant K} \sum_{j=1}^{K} |c_{ij}|$ of C). Then M^{-1} is a lower bound for the eigenvalues of W, namely for the diagonal entries of D. Let F be the diagonal matrix with positive diagonal entries and $F^2 = D$. Obviously, $\|F\mathbf{z}\|_2 \geqslant M^{-1/2} \|\mathbf{z}\|_2$ for every $\mathbf{z} \in \mathbf{R}^K$. Then, for every $\mathbf{y} \in \mathbf{R}^K$ we have

$$\mathbf{y}^t W \mathbf{y} = \mathbf{y}^t P^{-1} F F P \mathbf{y} = \|FP\mathbf{y}\|_2^2 \geqslant \|P\mathbf{y}\|_2^2 / M = \|\mathbf{y}\|_2^2 / M.$$

Now let $\mathbf{x}_0$ and $\mathbf{n} = \mathbf{x}_0 + \mathbf{y}$ be minimum points of the unconstrained problem and of the constrained problem, respectively. Then,

$$E(N) = \mathbf{n}^t W \mathbf{n} = \mathbf{x}_0^t W \mathbf{x}_0 + \mathbf{y}^t W \mathbf{y},$$

which implies by lemma 4.2 that $\mathbf{y}^t W \mathbf{y} \leqslant \sum_{i,j=1}^{K} w_{ij}$. Thus,

$$\|\mathbf{y}\|_2^2 / M \leqslant \sum_{i,j=1}^{K} w_{ij},$$

which yields the conclusion of the lemma with

$$\Delta = \sqrt{M \sum_{i,j=1}^{K} w_{ij}}.$$

□

Proof of theorem 2.1. Lemmas 4.1–4.3 provide a simple algorithm for calculating $E(N)$ for each N in constant time. Namely, we find the point $\mathbf{x}_0$ yielding the optimal $E_1(N)$ according to lemma 4.1, calculate the value of $\mathbf{n}^t W \mathbf{n}$ for all integral points $\mathbf{n}$, with coordinate sum N, within distance Δ from $\mathbf{x}_0$, and take the best of them. If the optimal point turns out to be $\mathbf{n} = \mathbf{x}_0 + \mathbf{y}$, we shall refer to $\mathbf{y}$ as *the correction vector*. We have $\mathbf{y} = \mathbf{n} - \mathbf{x}_0 = \boldsymbol{l} - \{\mathbf{x}_0\}$, where $\{\mathbf{x}_0\}$ denotes the vector of fractional parts of the coordinates of $\mathbf{x}_0$ and $\boldsymbol{l}$ belongs to some finite effective set L of integer vectors. Since the sum of coordinates of the correction vector is always 0, the sum of coordinates of $\boldsymbol{l}$ must equal that of $\{\mathbf{x}_0\}$. Thus, L consists of all integer vectors $\boldsymbol{l}$, for which the vector $\boldsymbol{l} - \{\mathbf{x}_0\}$ is of norm not exceeding the bound in lemma 4.3 and its coordinates sum vanishes. To emphasize the dependence of L on $\mathbf{x}_0$, we shall sometimes write $L(\mathbf{x}_0)$ instead of L. □

Now when choosing the optimal $\boldsymbol{l}$ out of L, we first note that, among any two candidates $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$, the former will be better (or equal) than the latter if and only if

$$(\mathbf{x}_0 + \boldsymbol{l}_1 - \{\mathbf{x}_0\})^t W(\mathbf{x}_0 + \boldsymbol{l}_1 - \{\mathbf{x}_0\}) \leqslant (\mathbf{x}_0 + \boldsymbol{l}_2 - \{\mathbf{x}_0\})^t W(\mathbf{x}_0 + \boldsymbol{l}_2 - \{\mathbf{x}_0\}).$$

This inequality is easily seen to be equivalent to

$$2(\boldsymbol{l}_2 - \boldsymbol{l}_1)^t W\{\mathbf{x}_0\} \leqslant \boldsymbol{l}_2^t W\boldsymbol{l}_2 - \boldsymbol{l}_1^t W\boldsymbol{l}_1.$$

Consequently, $\boldsymbol{l}$ is the optimal choice if and only if

$$2(\boldsymbol{l}' - \boldsymbol{l})^t W\{\mathbf{x}_0\} \leqslant (\boldsymbol{l}')^t W\boldsymbol{l}' - \boldsymbol{l}^t W\boldsymbol{l}, \qquad \boldsymbol{l}' \in L. \tag{13}$$

To study the second differences

$$\chi_N = E(N+1) - 2E(N) + E(N-1),$$

we shall write each term on the right-hand side in the form $E_1(N+j) + d_j$ for an appropriate d_j. In fact, as in (12), denoting by $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$ the correction vectors for $N-1$, N and $N+1$, respectively, we have

$$\chi_N = E_1(N+1) - 2E_1(N) + E_1(N-1) + \mathbf{y}_3^t W\mathbf{y}_3 - 2\mathbf{y}_2^t W\mathbf{y}_2 + \mathbf{y}_1^t W\mathbf{y}_1. \tag{14}$$

By (11),

$$\chi_N = 1 + \mathbf{y}_3^t W 2\mathbf{y}_3 - 2\mathbf{y}_2^t W\mathbf{y}_2 + \mathbf{y}_1^t W\mathbf{y}_1. \tag{15}$$

Let $\mathbf{x}_0, \mathbf{x}_0', \mathbf{x}_0''$ be the points yielding the optimal values of $E_1(N-1)$, $E_1(N)$, $E_1(N+1)$, respectively. In view of (10),

$$\mathbf{x}_0' = \mathbf{x}_0 + \mathbf{b}, \qquad \mathbf{x}_0'' = \mathbf{x}_0 + 2\mathbf{b}. \tag{16}$$

For appropriate integer vectors $\mathbf{p} \in L(\mathbf{x}_0)$, $\mathbf{p}' \in L(\mathbf{x}_0')$, $\mathbf{p}'' \in L(\mathbf{x}_0'')$,

$$\mathbf{y}_1 = \mathbf{p} - \{\mathbf{x}_0\}, \qquad \mathbf{y}_2 = \mathbf{p}' - \{\mathbf{x}_0'\}, \qquad \mathbf{y}_3 = \mathbf{p}'' - \{\mathbf{x}_0''\}. \tag{17}$$

The vectors $\mathbf{p}$, $\mathbf{p}'$, $\mathbf{p}''$ are determined by the system of inequalities:

$$\begin{cases} 2(\boldsymbol{l} - \mathbf{p})^t W\{\mathbf{x}_0\} \leqslant \boldsymbol{l}^t W\boldsymbol{l} - \mathbf{p}^t W\mathbf{p}, & \boldsymbol{l} \in L(\mathbf{x}_0), \\ 2(\boldsymbol{l} - \mathbf{p}')^t W\{\mathbf{x}_0'\} \leqslant \boldsymbol{l}^t W\boldsymbol{l} - (\mathbf{p}')^t W\mathbf{p}', & \boldsymbol{l} \in L(\mathbf{x}_0'), \\ 2(\boldsymbol{l} - \mathbf{p}'')^t W\{\mathbf{x}_0''\} \leqslant \boldsymbol{l}^t W\boldsymbol{l} - (\mathbf{p}'')^t W\mathbf{p}'', & \boldsymbol{l} \in L(\mathbf{x}_0''). \end{cases} \tag{18}$$

Due to (16), it is natural to try to rewrite (18) in terms of $\{\mathbf{x}_0\}$ without referring to $\{\mathbf{x}_0'\}$ and $\{\mathbf{x}_0''\}$. Divide the K-dimensional torus $\mathbf{T}^K$, which we identify with $[0,1)^K$, according to the vector $\mathbf{b}$, as follows.

The ith coordinate $\{x_{0i}'\}$ of $\{x_0'\}$ may be either $\{x_{0i}\} + b_i$ or $\{x_{0i}\} + b_i - 1$, depending on whether $\{x_{0i}'\} + b_i$ is smaller than 1 or not, respectively. Similarly, $\{x_{0i}''\}$ may assume one of the three values $\{x_{0i}\} + 2b_i - c$, where $c = 0, 1, 2$. Divide the circle $\mathbf{T}$ into three disjoint intervals (actually arcs), on each of which both $\{x_{0i}'\}$ and $\{x_{0i}''\}$ assume the same form in terms of $\{x_{0i}\}$. We have to distinguish between two cases:

(1) If $b_i \leqslant \frac{1}{2}$, write

$$\mathbf{T} = [0, 1 - 2b_i) \cup [1 - 2b_i, 1 - b_i) \cup [1 - b_i, 1). \tag{19}$$

If $\{x_{0i}\}$ belongs to the first interval on the right-hand, then

$$\{x_{0i}'\} = \{x_{0i}\} + b_i, \qquad \{x_{0i}''\} = \{x_{0i}\} + 2b_i,$$

if it belongs to the second

$$\{x_{0i}'\} = \{x_{0i}\} + b_i, \qquad \{x_{0i}''\} = \{x_{0i}\} + 2b_i - 1,$$

and if it belongs to the third

$$\{x_{0i}'\} = \{x_{0i}\} + b_i - 1, \qquad \{x_{0i}''\} = \{x_{0i}\} + 2b_i - 1.$$

(2) If $b_i > \frac{1}{2}$, write

$$\mathbf{T} = [0, 1 - b_i) \cup [1 - b_i, 2 - 2b_i) \cup [2 - 2b_i, 1). \tag{20}$$

This time, depending on the interval on the right-hand side containing $\{x_{0i}\}$, we have either

$$\{x'_{0i}\} = \{x_{0i}\} + b_i, \qquad \{x''_{0i}\} = \{x_{0i}\} + 2b_i - 1,$$

or

$$\{x'_{0i}\} = \{x_{0i}\} + b_i - 1, \qquad \{x''_{0i}\} = \{x_{0i}\} + 2b_i - 1,$$

or

$$\{x'_{0i}\} = \{x_{0i}\} + b_i - 1, \qquad \{x''_{0i}\} = \{x_{0i}\} + 2b_i - 2.$$

Let I_{i1}, I_{i2}, I_{i3} be the intervals on the right-hand side of (19) or (20), depending on whether $b_i \leqslant \frac{1}{2}$ or not, respectively. Denote

$$\Omega_{\eta_1\eta_2\ldots\eta_{K-1}} = I_{1\eta_1} \times I_{2\eta_2} \times \cdots \times I_{K-1,\eta_{K-1}}, \qquad \eta_1, \ldots, \eta_{K-1} \in \{1, 2, 3\}. \tag{21}$$

The sets $\Omega_{\eta_1\eta_2\ldots\eta_{K-1}}$ decompose the $(K-1)$-dimensional torus into a union of 3^{K-1} disjoint boxes:

$$\mathbf{T}^{K-1} = \bigcup_{\eta_1=1}^{3} \bigcup_{\eta_2=1}^{3} \cdots \bigcup_{\eta_{K-1}=1}^{3} \Omega_{\eta_1\eta_2\ldots\eta_{K-1}}.$$

The information provided by the vector $\{\mathbf{x}_0\}$ is partly redundant as the fact that $\sum_{i=1}^{K}\{x_{0i}\}$ is an integer determines each component in terms of the others. To avoid this inconvenience, we shall eliminate, say, $\{x_{0K}\}$. Divide $\mathbf{T}^{K-1}$ into K parts as follows:

$$\Omega^s = \{(t_1, \ldots, t_{K-1}) \in \mathbf{T}^{K-1} : s - 1 < \sum_{i=1}^{K-1} t_i \leqslant s\}, \qquad s = 0, \ldots, K-1. \tag{22}$$

(Thus, $\Omega^0 = \{\mathbf{0}\}$, while all other Ω^is have non-empty interior.) Suppose that $(\{x_{01}\}, \ldots, \{x_{0,K-1}\}) \in \Omega^s$. Then,

$$\{x_{0K}\} = s - \{x_{01}\} - \cdots - \{x_{0,K-1}\}. \tag{23}$$

We need a further subdivision to ensure that, in each cell, both $\{x'_{0K}\}$ and $\{x''_{0K}\}$ assume the same form in terms of $\{x_{01}\}, \ldots, \{x_{0,K-1}\}$. To this end, we first split $\mathbf{T}$ into three subintervals, similarly to (19) and (20), depending on whether $b_K \leqslant \frac{1}{2}$ or $b_K > \frac{1}{2}$, namely,

$$\mathbf{T} = [0, 1 - 2b_K) \cup [1 - 2b_K, 1 - b_K) \cup [1 - b_K, 1) \tag{24}$$

or

$$\mathbf{T} = [0, 1 - b_K) \cup [1 - b_K, 2 - 2b_K) \cup [2 - 2b_K, 1). \tag{25}$$

Let I_{K1}, I_{K2}, I_{K3} be the intervals in the splitting. Put

$$\Omega^s_\eta = \left\{(t_1, \ldots, t_{K-1}) \in \Omega^s : s - \sum_{i=1}^{K-1} t_i \in I_{K\eta}\right\}, \qquad \eta = 1, 2, 3. \tag{26}$$

Suppose $b_K \leqslant \frac{1}{2}$. If the point $(\{x_{01}\}, \ldots, \{x_{0,K-1}\})$ belongs to Ω^s_1, then

$$\{x'_{0K}\} = s - \sum_{i=1}^{K-1}\{x_{0i}\} + b_K, \qquad \{x''_{0K}\} = s - \sum_{i=1}^{K-1}\{x_{0i}\} + 2b_K,$$

if it belongs to Ω_2^s, then

$$\{x'_{0K}\} = s - \sum_{i=1}^{K-1}\{x_{0i}\} + b_K, \qquad \{x''_{0K}\} = s - \sum_{i=1}^{K-1}\{x_{0i}\} + 2b_K - 1,$$

and if it belongs to Ω_3^s, then

$$\{x'_{0K}\} = s - \sum_{i=1}^{K-1}\{x_{0i}\} + b_K - 1, \qquad \{x''_{0K}\} = s - \sum_{i=1}^{K-1}\{x_{0i}\} + 2b_K - 1.$$

If $b_K > \frac{1}{2}$, then we similarly find linear expressions for $\{x'_{0K}\}$ and $\{x''_{0K}\}$ in terms of $\{x_{0i}\}$s on each Ω_η^s.

Denote

$$\Omega^s_{\eta_1\eta_2\ldots\eta_K} = \Omega_{\eta_1\eta_2\ldots\eta_{K-1}} \cap \Omega^s_{\eta_K}, \qquad 0 \leqslant s \leqslant K-1, \quad 1 \leqslant \eta_i \leqslant 3. \tag{27}$$

Then,

$$\mathbf{T}^{K-1} = \bigcup_{s=0}^{K-1}\bigcup_{\eta_1=1}^{3}\cdots\bigcup_{\eta_K=1}^{3}\Omega^s_{\eta_1\ldots\eta_K}$$

forms a decomposition of $\mathbf{T}^{K-1}$ into a disjoint union of $K\cdot 3^K$ disjoint polytopes. The important property of this decomposition is that, if $(\{x_{01}\},\ldots,\{x_{0,K-1}\})$ belongs to any cell $\Omega^s_{\eta_1\eta_2\ldots\eta_K}$, the $2K+1$ numbers $\{x_{0K}\}, \{x'_{01}\},\ldots,\{x'_{0K}\}, \{x''_{01}\},\ldots,\{x''_{0K}\}$ depend linearly on the first $K-1$ coordinates $\{x_{01}\},\ldots,\{x_{0,K-1}\}$. That is,

$$\begin{aligned}
\{x'_{0i}\} &= \{x_{0i}\} + b_i - \alpha'_i, \qquad 1 \leqslant i \leqslant K-1,\\
\{x''_{0i}\} &= \{x_{0i}\} + 2b_i - \alpha''_i, \qquad 1 \leqslant i \leqslant K-1,
\end{aligned} \tag{28}$$

$$\begin{aligned}
\{x'_{0K}\} &= s - \sum_{i=1}^{K-1}\{x_{0i}\} + b_K - \alpha'_K,\\
\{x''_{0K}\} &= s - \sum_{i=1}^{K-1}\{x_{0i}\} + 2b_K - \alpha''_K,
\end{aligned} \tag{29}$$

where $\alpha'_i \in \{0,1\}, \alpha''_i \in \{0,1,2\}$, $i = 1,2,\ldots,K$. Altogether, there exists a linear transformation $T : \mathbf{R}^{K-1} \to \mathbf{R}^K$, and for each cell $\Omega^s_{\eta_1\eta_2\ldots\eta_K}$ there exist vectors $\mathbf{v}, \mathbf{v}', \mathbf{v}'' \in \mathbf{R}^K$, such that, denoting $\mathbf{x} = (\{x_{01}\},\ldots,\{x_{0,K-1}\})$, we have

$$\{\mathbf{x}_0\} = T\mathbf{x} + \mathbf{v}, \qquad \{\mathbf{x}'_0\} = T\mathbf{x} + \mathbf{v}', \qquad \{\mathbf{x}''_0\} = T\mathbf{x} + \mathbf{v}''. \tag{30}$$

On each cell $\Omega^s_{\eta_1\eta_2\ldots\eta_K}$ we may now rewrite the system (18), defining the optimal vectors $\mathbf{p}, \mathbf{p}', \mathbf{p}''$ in the form

$$\begin{cases}
2(\boldsymbol{l}-\mathbf{p})^t W T\mathbf{x} \leqslant \boldsymbol{l}^t W \boldsymbol{l} - \mathbf{p}^t W \mathbf{p} - 2(\boldsymbol{l}-\mathbf{p})^t W \mathbf{v}, & \boldsymbol{l} \in L,\\
2(\boldsymbol{l}-\mathbf{p}')^t W T\mathbf{x} \leqslant \boldsymbol{l}^t W \boldsymbol{l} - (\mathbf{p}')^t W \mathbf{p}' - 2(\boldsymbol{l}-\mathbf{p}')^t W \mathbf{v}', & \boldsymbol{l} \in L',\\
2(\boldsymbol{l}-\mathbf{p}'')^t W T\mathbf{x} \leqslant \boldsymbol{l}^t W \boldsymbol{l} - (\mathbf{p}'')^t W \mathbf{p}'' - 2(\boldsymbol{l}-\mathbf{p}'')^t W \mathbf{v}'', & \boldsymbol{l} \in L''.
\end{cases} \tag{31}$$

Note that we have suppressed the dependence of the sets L, L', L'' on $\mathbf{x}_0, \mathbf{x}'_0, \mathbf{x}''_0$. In fact, considering L, for example, it is clear that each candidate $\boldsymbol{l} \in L$ must have sum of coordinates s and, in view of lemma 4.3, its norm is bounded above by $\Delta + \sqrt{K}$. Thus, taking

$$L = \left\{ \boldsymbol{l} \in \mathbf{Z}^K : \|\boldsymbol{l}\| \leqslant \Delta + \sqrt{K}, \sum_{i=1}^{K} l_i = s \right\},$$

$$L' = \left\{ \boldsymbol{l} \in \mathbf{Z}^K : \|\boldsymbol{l}\| \leqslant \Delta + \sqrt{K}, \sum_{i=1}^{K} l_i = s + 1 - \sum_{i=1}^{K} \alpha_i' \right\},$$

$$L'' = \left\{ \boldsymbol{l} \in \mathbf{Z}^K : \|\boldsymbol{l}\| \leqslant \Delta + \sqrt{K}, \sum_{i=1}^{K} l_i = s + 2 - \sum_{i=1}^{K} \alpha_i'' \right\},$$

(where $\boldsymbol{l} = (l_1, l_2, \ldots, l_K)$), we certainly do not miss any potentially optimal vectors $\mathbf{p}, \mathbf{p}', \mathbf{p}''$ by restricting the search to L, L', L'', respectively.

For each choice of $\eta_1, \ldots, \eta_K, s$ and of the vectors $\mathbf{p}, \mathbf{p}', \mathbf{p}''$, let $P^{s\mathbf{p}\mathbf{p}'\mathbf{p}''}_{\eta_1 \ldots \eta_K}$ be the set of all points in $\Omega^s_{\eta_1 \ldots \eta_K}$ satisfying (31). Then,

$$\mathbf{T}^{K-1} = \bigcup_{s=0}^{K-1} \bigcup_{\eta_1=1}^{3} \cdots \bigcup_{\eta_K=1}^{3} \bigcup_{\mathbf{p} \in L} \bigcup_{\mathbf{p}' \in L'} \bigcup_{\mathbf{p}'' \in L''} P^{s\mathbf{p}\mathbf{p}'\mathbf{p}''}_{\eta_1 \ldots \eta_K}, \tag{32}$$

where the sets on the right-hand side are disjoint (up to sets of a smaller dimension).

By (17) and (30), for all points in each sub-polytope $P^{s\mathbf{p}\mathbf{p}'\mathbf{p}''}_{\eta_1 \ldots \eta_K}$ we have optimal correction vectors constructed in the same way, namely,

$$\mathbf{y}_1 = \mathbf{q} - T\mathbf{x}, \qquad \mathbf{y}_2 = \mathbf{q}' - T\mathbf{x}, \qquad \mathbf{y}_3 = \mathbf{q}'' - T\mathbf{x}, \tag{33}$$

where $\mathbf{q} = \mathbf{p} - \mathbf{v}, \mathbf{q}' = \mathbf{p}' - \mathbf{v}', \mathbf{q}'' = \mathbf{p}'' - \mathbf{v}''$. Hence, if the point $\mathbf{x}$ belongs to $P^{s\mathbf{p}\mathbf{p}'\mathbf{p}''}_{\eta_1 \ldots \eta_K}$, then χ_N depends linearly on the coordinates $\{x_{01}\}, \ldots, \{x_{0,K-1}\}$:

$$\begin{aligned} \chi_N &= 1 + (\mathbf{q} - T\mathbf{x})^t W (\mathbf{q} - T\mathbf{x}) - 2(\mathbf{q}' - T\mathbf{x})^t W (\mathbf{q} - T\mathbf{x}) + (\mathbf{q}'' - T\mathbf{x})^t W (\mathbf{q}'' - T\mathbf{x}) \\ &= 2(-\mathbf{q} + 2\mathbf{q}' - \mathbf{q}'')^t W T\mathbf{x} + \text{const}, \end{aligned} \tag{34}$$

where $\text{const} = 1 + \mathbf{q}^t W \mathbf{q} - 2(\mathbf{q}')^t W \mathbf{q}' + (\mathbf{q}'')^t W \mathbf{q}''$. (Note that all coefficients on the right-hand side of (34) depend on $\eta_1, \ldots, \eta_K, s, \mathbf{p}, \mathbf{p}', \mathbf{p}''$.)

We need to find the function F according to which the sequence $(\chi_N)_{N=1}^{\infty}$ is asymptotically F-distributed. To simplify our notation, rewrite (32) in the form

$$\mathbf{T}^{K-1} = \bigcup_{i=1}^{r} P_i, \tag{35}$$

where each P_i is one of the polytopes $P^{s\mathbf{p}\mathbf{p}'\mathbf{p}''}_{\eta_1 \ldots \eta_K}$. Denote

$$\boldsymbol{\theta}_N = (\{Nb_1\}, \ldots, \{Nb_{K-1}\}), \qquad N = 1, 2, \ldots,$$

and

$$A_i = \{N \in \mathbf{N} : \boldsymbol{\theta}_N \in P_i\}, \qquad i = 1, 2, \ldots, r.$$

(A_is may intersect, as P_is may intersect on sets of a smaller dimension. However, this will cause no problem as the intersections are sets of density 0 in $\mathbf{N}$. Alternatively, we may first decide in some arbitrary way where to place 'problematic' integers.)

Let $\boldsymbol{\theta}^{(i)}_N$ be the subsequence of $\boldsymbol{\theta}_N$, consisting of those elements $\boldsymbol{\theta}_N$ with $N \in A_i$, $1 \leqslant i \leqslant r$. That is, $\boldsymbol{\theta}^{(i)}_N$ is the Nth element of $(\boldsymbol{\theta}_N)_{N=1}^{\infty}$ which belongs to A_i. Let $\left(\chi^{(i)}_N\right)_{N=1}^{\infty}$ be the corresponding subsequence of $(\chi_N)_{N=1}^{\infty}$. If the point $\mathbf{x}$ lies in P_i, then it belongs to the subsequence $\left(\boldsymbol{\theta}^{(i)}_N\right)_{N=1}^{\infty}$. Thus, by (34) there exist affine functions $\psi_i : \mathbf{R}^{K-1} \to \mathbf{R}$ such that

$$\chi^{(i)}_N = \psi_i\left(\boldsymbol{\theta}^{(i)}_N\right), \qquad 1 \leqslant i \leqslant r, \quad N = 1, 2, \ldots. \tag{36}$$

Since the numbers $b_1, \ldots, b_K$ are linearly independent over $\mathbf{Q}$, so are the numbers $1, b_1, \ldots, b_{K-1}$, and consequently the sequence $(\boldsymbol{\theta}_N)$ is uniformly distributed modulo 1 in $\mathbf{R}^{K-1}$. Hence, each of the subsequences $\left(\boldsymbol{\theta}^{(i)}_N\right)_{N=1}^{\infty}$, $1 \leqslant i \leqslant r$, is uniformly distributed in

P_i. By (36), the sequence $\left(\chi_N^{(i)}\right)_{N=1}^{\infty}$ is the image of a uniformly distributed sequence in P_i under the mapping ψ_i. Hence, letting $(Y_1, Y_2, \ldots, Y_{K-1})$ be a $(K-1)$-dimensional random variable, uniformly distributed in P_i, we see that $\left(\chi_N^{(i)}\right)_{N=1}^{\infty}$ is F_i-distributed, where F_i is the distribution function of $\psi_i(Y_1, Y_2, \ldots, Y_{K-1})$.

According to [6, theorem 2.3], F_i is a piecewise polynomial function, each polynomial piece being of degree at most $K-1$, and can be effectively computed. Since $\left(\boldsymbol{\theta}_N^{(i)}\right)_{N=1}^{\infty}$ is uniformly distributed modulo 1 in $\mathbf{R}^{K-1}$, the density of each A_i is the measure d_i of the set P_i. As P_i is a polytope, this measure can be effectively computed. By [5, lemma 1], $(\chi_N)_{N=1}^{\infty}$ is asymptotically F-distributed, where $F = \sum_{i=1}^{r} d_i F_i$. This completes the proof.

5. Summary of the algorithm

In this section, we briefly present our method for explicitly finding the distribution function of the sequence $(\chi_N)_{N=1}^{\infty}$ as an algorithm, unlike the implicit form in which it was done in the previous section.

Input. A positive-definite symmetric matrix C, with positive diagonal entries, non-positive off-diagonal entries and positive row sums. The matrix $W = \frac{1}{2}C^{-1}$ has non-negative entries.

Output. The distribution function of the sequence $(\chi_N)_{N=1}^{\infty}$, defined by (6).

Algorithm

(1) Normalize the matrix C as explained after the formulation of lemma 4.1. Denote by $\mathbf{b}$ the vector of row sums of C.
(2) Calculate the constant Δ of lemma 4.3 according to the proof of that lemma.
(3) Split the $(K-1)$-dimensional torus $\mathbf{T}^{K-1}$, which we identify with $[0, 1)^{K-1}$, according to the vector $\mathbf{b}$, as follows:
  (a) For each $1 \leqslant i \leqslant K$, divide the circle $\mathbf{T}$ into three disjoint intervals I_{i1}, I_{i2}, I_{i3} by (19) and (20).
  (b) Write $\mathbf{T}^{K-1}$ as a union of 3^{K-1} disjoint boxes $\Omega_{\eta_1\eta_2\ldots\eta_{K-1}}$ by (21).
  (c) Decompose $\mathbf{T}^{K-1}$ into K parts Ω^s according to (22).
  (d) Divide each Ω^s into three parts Ω^s_η according to (26).
  (e) Form the intersections $\Omega^s_{\eta_1\eta_2\ldots\eta_K}$, defined in (27), and find for each of them the transformation T and the vectors $\mathbf{v}$, $\mathbf{v}'$, $\mathbf{v}''$, defined prior to (30).
  (f) Construct the sub-polytopes $P^{s\mathbf{pp}'\mathbf{p}''}_{\eta_1\ldots\eta_K}$, defined before (32). These sub-polytopes provide the required decomposition of $\mathbf{T}^{K-1}$ (see (31)).
(4) For each sub-polytope $P^{s\mathbf{pp}'\mathbf{p}''}_{\eta_1\ldots\eta_K}$, or P_i according to the new indexing in (35), find the vectors $\mathbf{q}$, $\mathbf{q}'$, $\mathbf{q}''$ of (33), and thus the functions ψ_i of (36).
(5) Calculate each F_i according to the proof of [6, theorem 2.3].
(6) Calculate the volume d_i of each P_i by some algorithm for calculating the volume of a polytope.
(7) Calculate the distribution function of $(\chi_N)_{N=1}^{\infty}$ by the formula $F = \sum_{i=1}^{r} d_i F_i$.

References

[1] Alhassid Y, Weinmüller H A and Wobst A 2004 *Preprint* cond-mat 0406495
Alhassid Y 2000 *Rev. Mod. Phys.* **72** 895
Jallabert R A, Alhassid Y and Stone A D 1992 *Phys. Rev. Lett.* **68** 3468
Alhassid Y and Attias H 1996 *Phys. Rev. Lett.* **76** 1711
Sivan U, Berkovits R, Aloni Y, Prus O, Auerbach A and Ben-Yoseph G 1996 *Phys. Rev. Lett.* **77** 1123

[2] Furlan M, Heinzel T, Jeanneret B, Lotkhov S V and Ensslin K 1999 *Preprint* cond-mat/9909122
[3] Folk J A, Marcus C A, Berkovits R, Kurland I L, Aleiner I L and Altshuler B L 2000 *Preprint* cond-mat/0010441
[4] Ashoori R C, Stormer H L, Weiner J S, Pfeiffer L N, Pearton S J, Baldwin K W and West K W 1992 *Phys. Rev. Lett.* **68** 3088–91
Ashoori R C *et al* 1993 *Physica (Amsterdam)* B **189** 117
Berman D, Zhitenev N B, Ashoori R C and Shayegan M 1999 Observation of quantum fluctuations of charge on a quantum dot *Phys. Rev. Lett.* **82** 161–4
[5] Avishai Y, Berend D and Berkovits R 1998 Statistics of addition spectra of independent quantum systems *J. Phys. A: Math. Gen.* **31** 8063–72
[6] Berend D and Bromberg L Uniform decompositions of polytopes *Preprint*
[7] Landau L D and Livshiz I M 1959 *Electrodynamics of Continuous Media* (Oxford: Pergamon) (See especially the discussion on pp 3–7)
[8] Kuipers L and Niederreiter H 1974 *Uniform Distribution of Sequences* (New York: Wiley)