

PAPER • OPEN ACCESS

## Many body density of states of a system of non interacting spinless fermions

To cite this article: Rémi Lefèvre *et al* 2023 *New J. Phys.* **25** 063004

View the [article online](#) for updates and enhancements.

You may also like

- [Multiworm algorithm quantum Monte Carlo](#)  
F Lingua, B Capogrosso-Sansone, A Safavi-Naini et al.
- [Quantum point spread function for imaging trapped few-body systems with a quantum gas microscope](#)  
Maxim Pyzh, Sven Krönke, Christof Weitenberg et al.
- [Many body density of states in the edge of the spectrum: non-interacting limit](#)  
Pragya Shukla


**PAPER**

# Many body density of states of a system of non interacting spinless fermions

**OPEN ACCESS****RECEIVED**  
21 August 2022**REVISED**  
11 May 2023**ACCEPTED FOR PUBLICATION**  
25 May 2023**PUBLISHED**  
7 June 2023

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.

Rémi Lefèvre, Krissia Zawadzki<sup>†</sup> and Grégoire Ithier<sup>\* </sup>

Department of Physics, Royal Holloway, University of London, Egham, United Kingdom

<sup>\*</sup> Author to whom any correspondence should be addressed.<sup>†</sup> Present address: School of Physics, Trinity College Dublin, College Green, Dublin 2, Ireland**E-mail:** [gregoire.ithier@rhul.ac.uk](mailto:gregoire.ithier@rhul.ac.uk)**Keywords:** density of states, spinless fermions, quantum many body systems, many body density of states**Abstract**

The modeling of out-of-equilibrium many-body quantum systems requires to go beyond low-energy physics and single or few bodies densities of states. Many-body localization, presence or lack of thermalization and quantum chaos are examples of phenomena in which states at different energy scales, including highly excited ones, contribute to dynamics and therefore affect the system's properties. Quantifying these contributions requires the many-body density of states (MBDoS), a function whose calculation becomes challenging even for non-interacting identical particles due to the difficulty to enumerate accessible states while enforcing the exchange symmetry. In the present work, we introduce a new approach to evaluate the MBDoS in the general case of non-interacting systems of identical quantum particles. The starting point of our method is the principal component analysis of a filling matrix  $F$  describing how  $N$  particles can be distributed into  $L$  single-particle energy levels. We show that the many body spectrum can be expanded as a weighted sum of singular vectors of the filling matrix. The weighting coefficients only involve renormalized energies obtained from the single body spectrum. We illustrate our method in two classes of problems that are mapped into spinless fermions : (i) non-interacting electrons in a homogeneous tight-binding model in 1D and 2D, and (ii) interacting spins in a chain under a transverse field.

**1. Introduction**

The concept of density of states (DoS) is at the heart of statistical physics where it defines both partition function and temperature. In nuclear physics, quantifying the level density is necessary to describe nuclear reactions involving excited states [1]. No less importantly, it is one of the most appealing quantity in condensed matter physics, where one is interested in investigating how electrons and holes populate energy bands to give rise to materials properties [2]. In all these fields, the DoS is crucial to characterize a many-body system and determine which states are accessible at energy scales of interest.

For a long time, the success of mean field theories and the quasi-particle picture to describe a highly degenerate Fermi liquid has made the *single*-body density of states (SBDOS) and its local counterpart the focus of investigations in electronic systems. In the presence of interactions, efforts have been concentrated in the physics at low temperatures, such that the SBDOS around the Fermi level suffices to obtain most of their properties. Nonetheless, the quest for calculating a *many*-body density of states (MBDoS) has become arguably necessary in the context of isolated quantum systems undergoing out-of-equilibrium dynamics [3]. In such systems, contributions from different parts of the spectrum prevent one from relying on a description solely based on low-lying states and the SBDOS. Quantifying these contributions is crucial to shed light on phenomena such as quantum chaos [4], thermalization (or the lack of) [5–8], many-body localization [9], and more generally unconventional stationary states [10, 11] and phase transitions [12].

In this respect, a MBDoS provides useful information, as it allows to quantify how interactions between individual constituents lead to complex many-body dynamics, with coexisting single-particle and collective

effects. Even non-interacting systems pose a challenge due to the combinatorial nature of the problem of how single-body levels can be populated to define a distribution of many-body energy levels. For both non-interacting and interacting systems, a direct numerical approach only allows to compute the full spectrum of energies and eigenstates for system sizes that do not exceed a dozen of particles.

Symmetries can improve this computation by allowing one to split the total Hilbert space into blocks, associated with conserved quantum numbers. This idea has led to exact diagonalization methods [13] and is also implemented in well-established numerical methods such as kernel polynomial methods (KPMs) [14, 15] (see also the review in [16]) which can benefit from tensor network techniques to approximate states [8, 17, 18]. It is also worth noticing that the Fourier transformation of the temporal evolution of the fidelity (or survival probability) is connected to the local counterpart of the MBDoS [19, 20]. On the analytical side, methods to compute the MBDoS started in nuclear physics with a Fermi gas approximation derived by Bethe [21], which inspired approaches such as constant temperature or continuum shell models (see e.g. reviews in [1, 22]). More involved methods using exact combinatorial counting [23, 24], recursive relations [25], or saddle approximations [26] provided some approximate results. Finally, it is also possible to use quantum fields to obtain an analytical expression of the spectral density expressed in terms of 2-point Green's functions [27].

Exact computation of the MBDoS of systems of non-interacting identical particles is the focus of the present paper. We propose a new approach based on the symmetries of a rectangular filling matrix describing how  $N$  particles can be distributed into  $L$  single-particle energy levels to generate many-body states. The core of our method relies on the singular value decomposition (SVD) of the filling matrix, which allows to expand the many-body spectrum as a weighted sum of singular vectors. These 'principal components' *only* depend on the combinatorics of how a given number of identical particles can be dispatched over a given number of single body states. Weighting factors will be given by a discrete Fourier transform of the single body spectrum. This separation is of crucial importance for numerical computations since the expensive combinatorics can be computed once and for all then applied to any system with the same  $L$  and  $N$  values. Moreover, as we will see, we can take advantage of several symmetries which allow the very large set of configurations to be represented in a highly compressed form, suitable for efficient computation. System-dependent quantities only require a small  $O(L)$  inner product to compute.

We organize the paper as follows. In section 2, we introduce the SVD of the filling matrix and explain how it allows to access the non-interacting many body spectrum. In section 3, we develop the method in the case of spinless fermions by introducing the various symmetries which can be exploited before giving a summary of the complete procedure. We then give some numerical values to emphasize on the computational improvements provided. In section 5, we explain how the whole method can be extended to the general case, when additional conserved quantum numbers are present or when dealing with bosons. Applications to tight-binding models and Ising chains are discussed in section 4. Finally, our main findings are summarized in section 6.

## 2. Principal components analysis

We start by considering a system for which we know the single-body spectrum given by the energies  $\{\epsilon_0, \dots, \epsilon_{L-1}\}$ . Our goal is to construct the many-body spectrum, which has energies  $E_p = \sum_{k=0}^{L-1} F_p^k \epsilon_k$ , where  $F_p^k$  is the occupation number of the  $k$ th single body level for the  $p$ th many-body state. Our first step is to rewrite this expression in matrix form by collecting all single-body energies  $\epsilon_k$  into a column vector  $\epsilon$ , all many body energies  $E_p$  into a column vector  $E$ , and finally construct the filling rectangular matrix  $F$  with matrix elements  $F_p^k$ , each row of size  $L$  containing a configuration. We rewrite the relation between many body and single body energies in matrix form :

$$E = F \cdot \epsilon, \quad (1)$$

where  $\cdot$  denotes the matrix product.

The SVD of the filling matrix is defined as  $F = U \cdot \Sigma \cdot V^\dagger$  where  $U, V$  are square unitary matrices and  $\Sigma$  is a rectangular diagonal matrix. The principal component expansion of  $F$  follows:

$$F = \sum_{\ell=0}^{L-1} \sigma_\ell U_\ell \cdot V_\ell^\dagger \quad (2)$$

where  $U_\ell$  is the  $\ell$ th left-singular vector (i.e. the  $\ell$ th column of  $U$ ),  $\sigma_\ell$  are the singular values (i.e. the main diagonal of  $\Sigma$ ), and  $V_\ell$  is the  $\ell$ th right-singular vector (i.e. the  $\ell$ th column of  $V$ ).

To ease notations, in the following we will incorporate the singular value  $\sigma_\ell$  into the definition of  $U_\ell$  such that the left-singular vectors are orthogonal but *not* normalized anymore. With this convention,  $U_\ell$  and  $V_\ell$  are simply related by  $U_\ell = F \cdot V_\ell$ . By combining equations (1) and (2), we obtain :

$$E = \sum_{\ell=0}^{L-1} U_\ell (V_\ell^\dagger \cdot \epsilon) = \sum_{\ell=0}^{L-1} \tilde{\epsilon}_\ell U_\ell \quad (3)$$

where we defined the scalars  $\tilde{\epsilon}_\ell = V_\ell^\dagger \cdot \epsilon$ .

The many body spectrum now appears as a weighted sum of *principal components*, the  $U_\ell$  vectors. These spectral components are weighted by new effective energies  $\tilde{\epsilon}_\ell$  of a *renormalized* single-body spectrum. In particular, each  $V_\ell$  defines which single-body energies will contribute to each component  $U_\ell$  of the many-body spectrum. As we will see later, those  $V_\ell$  take a very simple form and simply contain Fourier modes. Note that  $\tilde{\epsilon}_\ell$  are the only system-dependent quantities here. Depending on the dispersion relation of the system (i.e. the relation between the index  $k$  and the energy  $\epsilon_k$ ), it is possible that some of those will vanish which allows the associated  $U_\ell$  vectors to be discarded.

The problem of computing the many body spectrum is now split in two parts. The first part is to compute the renormalized single body spectrum only depending on the right-singular part of the SVD (i.e. the  $V_\ell$ 's) which only require a small computational cost linear in  $L$ . The second part depends on the left-singular part of the SVD (i.e. the  $U_\ell$ 's) which only contains information about the universal properties of many-body systems encoded in the combinatoric structure of the  $F$  matrix.

Since the number of configurations scales exponentially with the number of levels  $L$  and particles  $N$  considered, so does the height of the  $F$  matrix and thus the size of the  $U_\ell$  vectors. We will use a combinatorial approach to avoid using the  $F$  matrix explicitly and allow computations to be done for large systems. It should be noted that  $F$  is a real matrix, and so are its singular values  $\sigma_\ell$ , however its left and right singular vectors will be chosen in a complex basis. Therefore, renormalized energies  $\tilde{\epsilon}_\ell$  will also be complex.

In the following, we focus on the simplest case of non-interacting fermions with no other quantum number in order to describe our framework to compute the singular decomposition of  $F$ .

### 3. Spinless fermions

In the case of fermions without any additional quantum number, occupation numbers can only be 0 or 1, which are the possible matrix elements of  $F$ . Each row of the filling matrix  $F$  (i.e. some configuration of a many body state) is a binary string. All rows share the same amount of 1's to account for the fixed number of particles  $N$ , there are a total of  $C_L^N$  many body state configurations and the filling matrix  $F$  has dimensions  $L \times C_L^N$ .

#### 3.1. Right part of the SVD

We first focus on the right-singular vectors  $V_\ell$  and the singular values  $\sigma_\ell$ . Both can be computed by considering the eigen decomposition of the square matrix  $F^\dagger \cdot F$  which has dimensions  $(L \times L)$ . Each matrix element  $(F^\dagger \cdot F)_{i,j}$  is the scalar product between the  $i$ th and  $j$ th columns of  $F$ , each encoding the occupation number of the  $i$ th and  $j$ th single body level across all many body states. If  $i = j$ , simple combinatorics considerations indicate there are  $C_{L-1}^{N-1}$  many body levels with the  $i$ th single body level filled with one particle. If  $i \neq j$ , similarly, there are  $C_{L-2}^{N-2}$  many body levels with both the  $i$ th and  $j$ th single body levels filled with a particle. From this reasoning we deduce that the  $F^\dagger F$  matrix has the form of a circulant matrix with only two possible entry values:

$$F^\dagger F = \begin{pmatrix} a & b & \dots & b \\ b & \ddots & & \vdots \\ \vdots & & \ddots & b \\ b & \dots & b & a \end{pmatrix}_{L \times L}$$

where  $a = C_{L-1}^{N-1}$  and  $b = C_{L-2}^{N-2}$ . Circulant matrices are well known and can be diagonalized by using Fourier modes. Let  $\omega_L = e^{i2\pi/L}$ , then eigenvectors of  $F^\dagger \cdot F$ , which are the right-singular vectors  $V_\ell$  of  $F$ , are given by:

$$V_\ell = \frac{1}{\sqrt{L}} \left( 1, \omega_L^\ell, \omega_L^{2\ell}, \dots, \omega_L^{(L-1)\ell} \right)_L \quad (4)$$

The eigenvalues of  $F^\dagger \cdot F$  take the following form:

$$\lambda_\ell = C_{L-1}^{N-1} + C_{L-2}^{N-2} \sum_{k=1}^{L-1} \omega_L^{k\ell},$$

which can only take two distinct values since:

$$\begin{aligned} \sum_{k=1}^{L-1} \omega_L^{k\ell} &= (L-1) & \text{if } \ell = 0, \\ &= -1 & \text{if } \ell \geq 1. \end{aligned}$$

Taking the square root, we obtain the singular values of  $F$ :

$$\begin{aligned} \sigma_0 &= \sqrt{NC_{L-1}^{N-1}} & \text{with multiplicity } 1, \\ \sigma_{\ell \geq 1} &= \sqrt{C_{L-2}^{N-1}} & \text{with multiplicity } L-1. \end{aligned} \quad (5)$$

One can easily check that the  $V_\ell$  vectors form an orthonormal basis associated to these eigenspaces, the  $\ell = 0$  case matching the 1-dimensional eigenspace spanned by  $V_0$  and the  $\ell \geq 1$  case matching the  $L-1$ -dimensional eigenspace spanned by  $V_{\ell \geq 1}$ .

### 3.2. Left part of the SVD

In order to compute the left-singular vectors  $U_\ell$ , we avoid dealing directly with the matrix  $F \cdot F^\dagger$  which is  $C_L^N \times C_L^N$ , and prefer using the identity  $U_\ell = F \cdot V_\ell$  to write the components of  $U_\ell$  as follows:

$$U_\ell^n = \frac{1}{\sqrt{L}} \sum_{k=0}^{L-1} F_n^k \omega_L^{k\ell}.$$

In other words, each component of the  $U_\ell$  vector is the  $\ell$ th Fourier series coefficient of each binary string representing a many body state. The  $\ell = 0$  case simplifies to  $U_0^n = \frac{N}{\sqrt{L}}$  and is connected to the filling factor  $N/L$  which is identical for each many body state. In addition, since the occupation numbers  $F_n^k$  are all real numbers, one has  $U_\ell = U_{L-\ell}^*$ , which provides a first symmetry simplifying the calculation.

For  $\ell \geq 1$ , we will investigate two symmetries which provide insight on the universal properties of  $F$  and allow to reduce the very large set of configurations to a more manageable size for numerical applications.

### 3.3. The $k$ -Symmetry

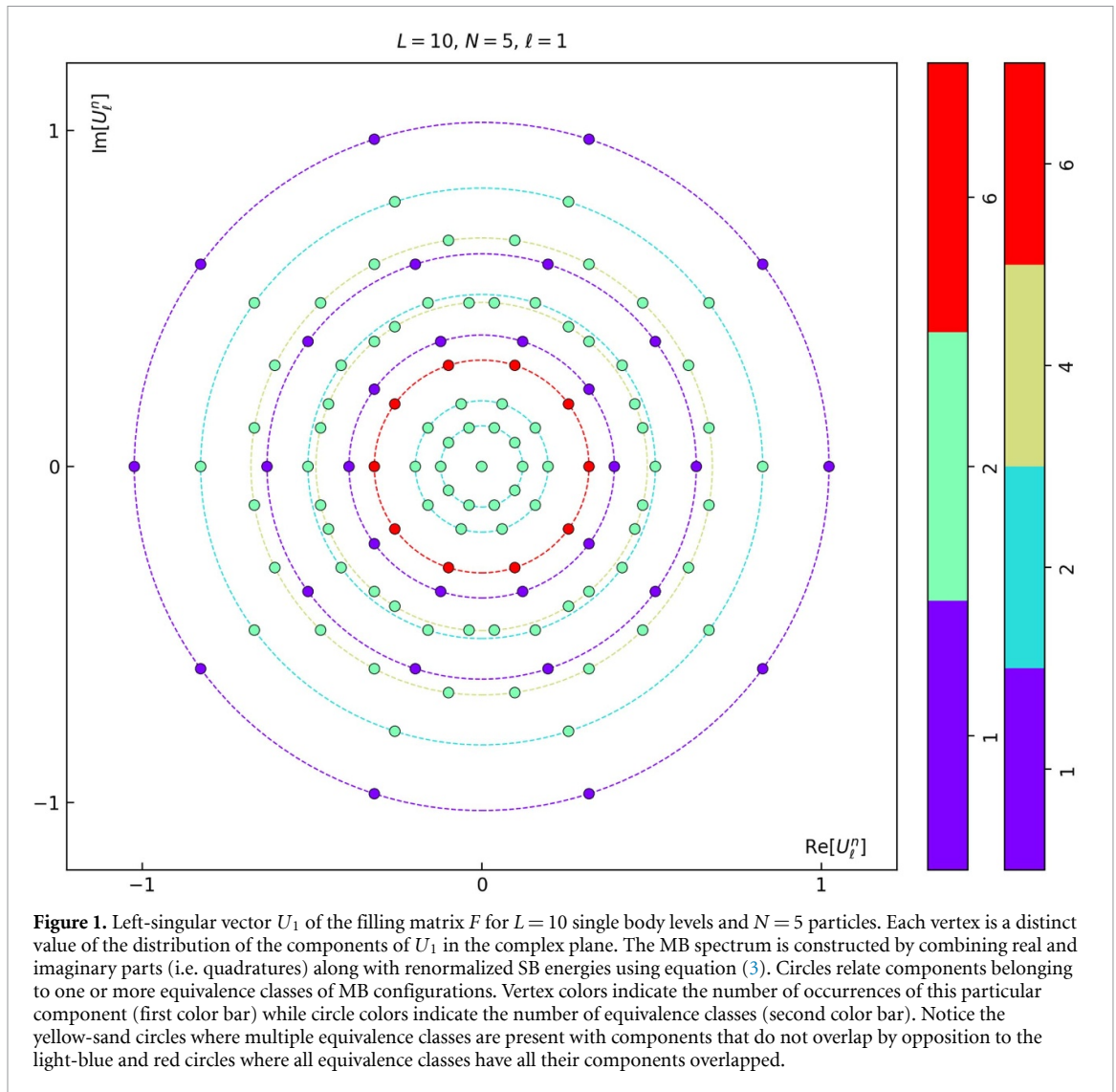
The first symmetry we observe is on the  $V_\ell$  vectors containing Fourier modes based on  $L$ th-roots of unity. Since they contain components of the form  $\omega_L^{k\ell}$ , when computing the inner product of  $V_\ell$  against a configuration represented by a binary string, a circular permutation of those bits by  $r$  positions to the right will multiply the Fourier modes by  $\omega_L^{r\ell}$ , i.e. it will rotate the result  $U_\ell^n$  in the complex plane. We will denote the angle associated to a circular permutation by 1 position  $\theta_{\text{step}} = 2\pi \ell / L$ . Note that this angle does not depend on  $N$  and remains valid for all components of a given  $U_\ell$ .

It follows that all the components  $U_\ell^n$  lie on various circles in the complex plane, and are located at relative angles which are multiples of  $\theta_{\text{step}}$ . We call this structure the  $k$ -symmetry, which is illustrated on figure 1.

The  $k$ -symmetry suggests to group configurations together in equivalence classes defined by the underlying relation of circular permutations. Each class contains configurations that can be transformed into each other by circular permutation, we choose the lowest of them in lexicographic order as class representatives which we will refer to as *seeds*. All members of such an equivalence class, when taking the inner product with  $V_\ell$ , will produce  $U_\ell^n$  components which lie on the same circle in the complex plane, for any fixed given  $\ell$ . Note that two distinct classes can still map to the same circle by having the same radius.

There will be two kinds of such equivalence classes: non-degenerate classes contain exactly  $L$  binary strings where bits do not provide any additional symmetry, while degenerate classes contain less than  $L$  strings as their bits expose a shorter pattern that repeats under circular permutation, for instance  $(0, 0, 1, 1, 0, 0, 1, 1)$  is an 8-bits sequence but loops back early with a circular permutation by 4 positions. This distinction will become important later on.

Next, we notice that equivalence classes are closely related to the factors of  $L$ . In particular, if  $L$  is a prime number, all classes are non-degenerate. Degenerate classes can only have a cardinal which has a common



factor with  $L$ . The prime decomposition of  $L$  is a decisive feature for the symmetries of the filling matrix and its SVD.

A counting argument can be made about those equivalence classes by using the Pólya enumeration theorem [28], from which we can estimate the number of equivalence classes to be of the order of  $C_L^N/L$ .

Generating seeds will be a bottleneck for numerical computations and is a known problem in the theory of necklaces. For binary strings, enumerating 2-colors necklaces can be done using a simple algorithm given in [29] which easily generalizes to  $\ell$ -colors necklaces (required later on). A more involved algorithm is given in [30].

### 3.4. The $\ell$ -Symmetry

On the other hand, we can look at what happens when we consider a different Fourier mode and go from some  $\ell$  to another  $\ell'$ . By definition,  $\omega_L$  is a primitive  $L$ th-root of unity if  $\omega_L^L = 1$  but  $\omega_L^k \neq 1$  for all  $k < L$ , each  $L$ th-root of unity is a distinct power of  $\omega_L$ . A well-known property is that  $\omega_L^\ell$  is a primitive  $r$ th-root of unity for  $r = L/\text{gcd}(L, \ell)$ . Considering another value  $\ell'$ , if the condition

$$\text{gcd}(L, \ell) = \text{gcd}(L, \ell') \tag{6}$$

is verified, then  $\omega_L^{\ell'}$  is also a primitive  $r$ th-root of unity. In other words, each power of  $\omega_L^{\ell'}$  is in a one to one correspondence with a power of  $\omega_L^\ell$ , so the components of  $V_{\ell'}$  are just a permutation of the components of  $V_\ell$ . In this case, since the permutation of  $V_\ell$  components is equivalent to the permutation of configurations in the  $F$  matrix, then vectors  $U_\ell$  and  $U_{\ell'}$  are also the same up to a permutation, i.e. the distributions of their components are identical that is, their pattern in the complex plane is identical. Moreover, since this is a one

to one correspondence, the number of configurations (i.e. degeneracy) landing on a particular vertex in the complex plane are also conserved. We call this property the  $\ell$ -symmetry, verified when (6) is satisfied.

The  $\ell$ -symmetry simplifies the study of  $U_\ell$  vectors, as only a small amount of  $\ell$  values are required to obtain the complete set of  $U_\ell$  distributions in the complex plane, namely the divisors of  $L$ . For any  $\ell$  divisor of  $L$ , we can define the integer  $q = L/\ell$  and rewrite:

$$U_\ell^n = \frac{1}{\sqrt{L}} \sum_{k=0}^{q-1} m_k e^{i2\pi \frac{k}{q}} \quad \text{with} \quad m_k = \sum_{s=0}^{\ell-1} F_n^{sq+k},$$

where a configuration represented by a binary string is *folded* into  $q$  parts. Each of these parts (i.e. each  $m_k$ ) can take any value between 0 and  $\ell$ , and their sum still adds up to  $N$ . This is called a  $\ell$ -restricted weak  $q$ -composition of  $N$  (see [31]). In fact, binary strings themselves are 1-restricted weak  $L$ -compositions of  $N$ . We redefine effective vectors  $V_\ell$  of size  $q$ , which now contain Fourier modes based on the  $q$ th-roots of unity  $\omega_q = e^{i2\pi/q}$ . Its components are of the form  $\omega_q^k$  with  $k \in \{0, q - 1\}$ .

This operation of folding is not injective, many binary configurations are mapped to a given composition. The total number of binary strings mapped to the same composition is given by a winding factor :

$$Q = \prod_{k=0}^{q-1} C_\ell^{m_k} \tag{7}$$

where one can observe that when  $\ell$  increases,  $Q$  increases exponentially, meaning that each component  $U_\ell^n$  in the complex plane will correspond to more and more configurations. Since the total number of configurations does not change, the number of distinct vertices must decrease. This can be observed from figures 2 and 3, where increasing  $\ell$  values lead to a fast decreasing density of  $U_\ell^n$  components in the complex plane, while the degeneracy of each point necessarily goes up : we transition from a spread distribution to a clustered distribution.

In the general case, the number of  $\ell$ -restricted weak  $q$ -compositions is given by the extended polynomial coefficient defined as:

$$[x^N] \left( \sum_{a=0}^{\ell} x^a \right)^q$$

where  $[x^N]$  means extracting the coefficient of the  $x^N$  term in the polynomial. This becomes our effective set of configurations and is much smaller than  $C_L^N$ . There is no simple analytic form for such a number but it can be shown that it decreases very fast when  $\ell$  increases. This will be further divided by  $q$  when using  $k$ -symmetry as we will see below. This is the symmetry that provides the most dramatic gain in performance for numerical computations.

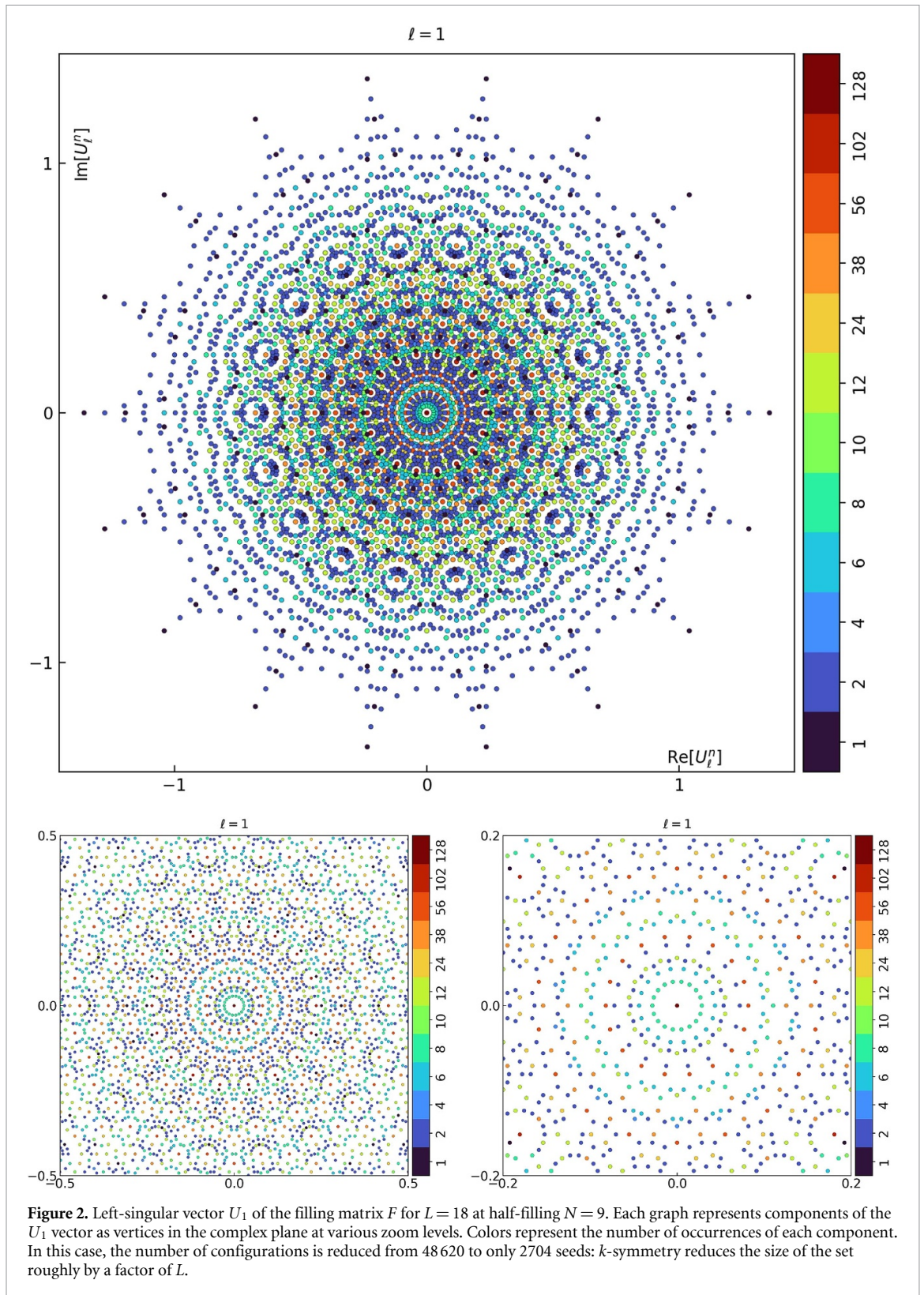
### 3.5. Enumeration

Both symmetries described above simplify the study of the problem. The last ingredient is to apply the  $k$ -symmetry on top of compositions, by again noticing that any circular permutation of a composition is only a multiplication of  $V_\ell$  by a power of  $\omega_q^k$  : we group compositions together in equivalence classes related by circular permutations.

This is where the distinction between degenerate and non-degenerate classes is important. Let's consider a configuration  $C$  of a degenerate class. Then that configuration admits a repeating pattern of size  $N$  which is repeated  $K$  times such that  $L = KN$ . If this is the case, then we can write the  $U_\ell$  component generated by this configuration when the inner product with  $V_\ell$  is performed. For any  $\ell \geq 1$  :

$$\begin{aligned} U_\ell^{(c)} &= \sum_{p=0}^{L-1} C_p e^{i\frac{2\pi\ell}{L} p} \\ &= \sum_{n=0}^{N-1} C_n \sum_{k=0}^{K-1} e^{i\frac{2\pi\ell}{L} (kN+n)} \\ &= \sum_{n=0}^{N-1} C_n e^{i\frac{2\pi\ell}{L} n} \left( \sum_{k=0}^{K-1} e^{i2\pi\ell \frac{k}{K}} \right) = 0 \end{aligned}$$

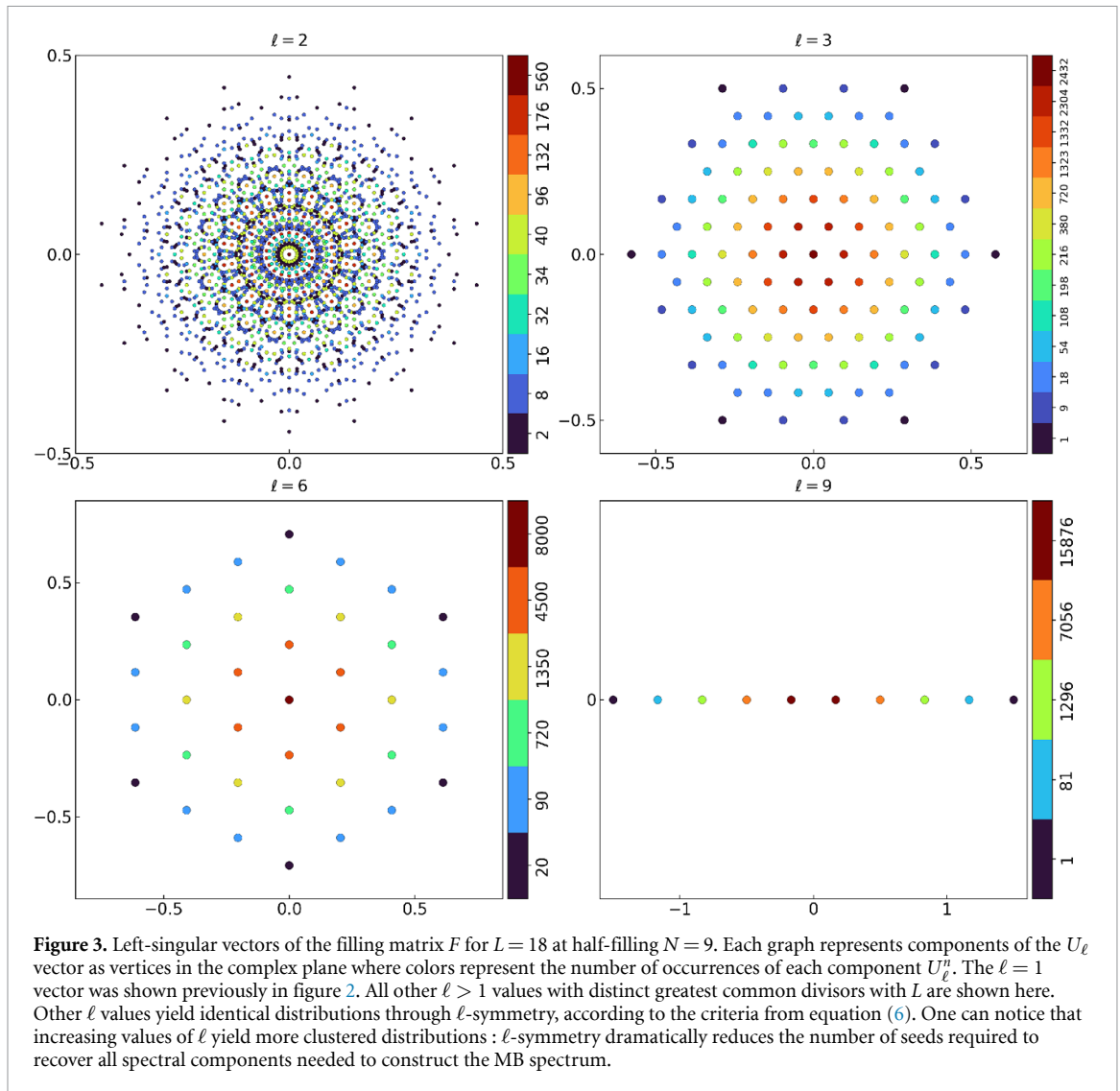
where we can see on the last line that the term in parentheses vanishes. Therefore, any projection on  $V_\ell$  of a member of a degenerate class is null. This means that degenerate classes all land at the origin on the circle of



radius zero. Basically, we only need to count those classes using their winding factors to obtain the occurrence count of zero  $U_\ell^n$  components.

We now have all the tools required for our framework. We can describe a simple recipe to obtain the exact distribution of the components of all  $U_\ell$  vectors. First, we look at the value of  $L$  and construct the list of all its divisors : this gives us the  $\ell$  values we have to deal with. For each of them, we compute  $q = L/\ell$  and proceed by building the list of seeds : all  $\ell$ -restricted weak  $q$ -compositions which are not related by circular permutation. Each seed defines an equivalence class from which we can extract various quantities. We first





**Figure 3.** Left-singular vectors of the filling matrix  $F$  for  $L = 18$  at half-filling  $N = 9$ . Each graph represents components of the  $U_\ell$  vector as vertices in the complex plane where colors represent the number of occurrences of each component  $U_\ell^n$ . The  $\ell = 1$  vector was shown previously in figure 2. All other  $\ell > 1$  values with distinct greatest common divisors with  $L$  are shown here. Other  $\ell$  values yield identical distributions through  $\ell$ -symmetry, according to the criteria from equation (6). One can notice that increasing values of  $\ell$  yield more clustered distributions :  $\ell$ -symmetry dramatically reduces the number of seeds required to recover all spectral components needed to construct the MB spectrum.

look at the size of the equivalence class to check if it is degenerate or not. If it is degenerate, i.e. lower than  $q$ , then the class lies on a 0-radius ‘circle’ i.e. the origin in the complex plane. In this case, we only need to recover the occurrence count which is the winding factor  $Q$  from equation (7) times the degenerate class size. Otherwise, the class has size  $q$  and is non-degenerate : we can extract some key quantities. The radius of the circle in the complex plane where  $U_\ell^n$  values are located, an offset angle where the first vertex is located, the previously defined  $\theta_{step}$  which gives the relative angle between vertices in this class, which allows to reconstruct all  $U_\ell^n$  values and finally their winding factor  $Q$  given by equation (7). The total number of  $U_\ell^n$  components belonging to this class is simply  $q$  times  $Q$ . Note the radius and offset angle are easily computed from the inner product of the seed with  $V_\ell$  where the module needs to be normalized by  $1/\sqrt{\ell}$  to give the radius of this class’s circle. We now know everything about that particular seed and its class. While processing all unique seeds and classes, we can accumulate the ones found to lie on the same circle, i.e. having the same radius, and simply add up occurrence counts accordingly. The end result is a list of all the circles, their respective radiuses, angular positions of all vertices if needed, and occurrence numbers (by point or by circle, whichever is needed). This gives the exact distribution of  $U_\ell^n$  components. The distributions of all other  $U_\ell^n$  vectors for which  $\ell$  was not a divisor of  $L$  are recovered from  $\ell$ -symmetry by the criteria from equation (6).

As an example, let us consider the case  $L = 8$  at half-filling  $N = 4$ . There are a total of 70 possible many body configurations written as binary strings. The divisors of  $L$  are  $\{1, 2, 4\}$ . For  $\ell = 1$  (i.e. binary strings), the  $k$ -symmetry gives 8 non-degenerate classes of size 8 and 2 degenerate classes of sizes 4 and 2, for a total of 10 seeds. For  $\ell = 2$ , applying both symmetries, we have 2-restricted weak 4-compositions, 4 non-degenerate classes of size 4 and 2 degenerate classes of sizes 2 and 1, for a total of 6 seeds. For  $\ell = 4$ , applying both symmetries, we have 4-restricted weak 2-compositions, 2 non-degenerate classes of size 2 and 1 degenerate class of size 1, for a total of 3 seeds. As one can see, the number of configurations to consider is significantly

reduced. For larger systems, e.g.  $L = 20$  at half-filling  $N = 10$ , the number of many body configurations written as binary strings is 184 756. In the worse case,  $\ell = 1$ , this is reduced to 9252 binary seeds. In better cases like  $\ell = 4$ , using  $\ell$ -symmetry leads to a much smaller set of 381 compositions which is further reduced to only 77 seeds.

For even larger systems, i.e.  $L$  of the order of 100, the lower  $\ell$  values still require significant computational time, mostly for  $\ell = 1$ , unless they are truncated or a statistical approach is added to sample the set of seeds. However, it should be noted that those computations are independent of the single body spectrum : they should be performed only *once* for relevant values of  $L$  and  $N$  (i.e. half-filling, quarter-filling, ...) and re-used for many different systems. Once the distribution of  $U_\ell^n$  components is known, the many-body spectrum is obtained from equation (3).

Finally, one should note that our procedure allows to relate any  $U_\ell$  component's index  $n$  to the configuration at the origin of its value. It is possible to provide a two-way mapping between configurations and their indices in lexicographic order, in logarithmic time in the number of configurations, i.e. linear time in  $L$ . Those indices are the indices  $n$  of  $U_\ell^n$  components inside an  $U_\ell$  vector. While requiring more computational time due to the very large number of  $U_\ell^n$  components, this is very useful to partially reconstruct the microscopic information which was lost when we chose to use a distribution of  $U_\ell^n$  components to represent  $U_\ell$  vectors. This allows to obtain a fully exact local information for the MBDoS in a desired energy interval, which becomes crucial in practice when one introduces dynamics and needs to simulate time evolution.

## 4. Applications: spinless fermions

A variety of systems can be described in terms of spinless fermions, including hard core bosons [32], Mott insulators in ladders [33], spin liquids [34], and, very recently, they served as the basis to study topological phases [35] and systems supporting Majorana fermions [36, 37].

To illustrate our method, we will apply it to two classes of solvable Hamiltonians which can be formulated in terms of a single-body spectrum. We first look at tight-binding of electrons in both one-dimensional chain and two-dimensional square lattice then consider the transverse field Ising model which can be mapped using the Jordan Wigner (JW) transformation. In both cases, we consider periodic boundary conditions (PBCs).

### 4.1. Tight-Binding 1D

The Hamiltonian for a chain of  $L$  sites reads:

$$H = -t \sum_{\langle i,j \rangle} (c_i^\dagger c_j + c_i c_j^\dagger), \quad (8)$$

where the sum  $\langle i,j \rangle$  runs over nearest-neighbor sites,  $c_i$  and  $c_i^\dagger$  are the annihilation and creation operators, and  $t$  is the hopping amplitude. In 1D, the single-body energies are:

$$\epsilon_k = -2t \cos(k), \quad (9)$$

where the momentum  $k$  depends on the boundary conditions of the model. For PBC,  $k = \frac{2\pi n}{L}$ , with  $n \in \{0, \dots, L-1\}$ .

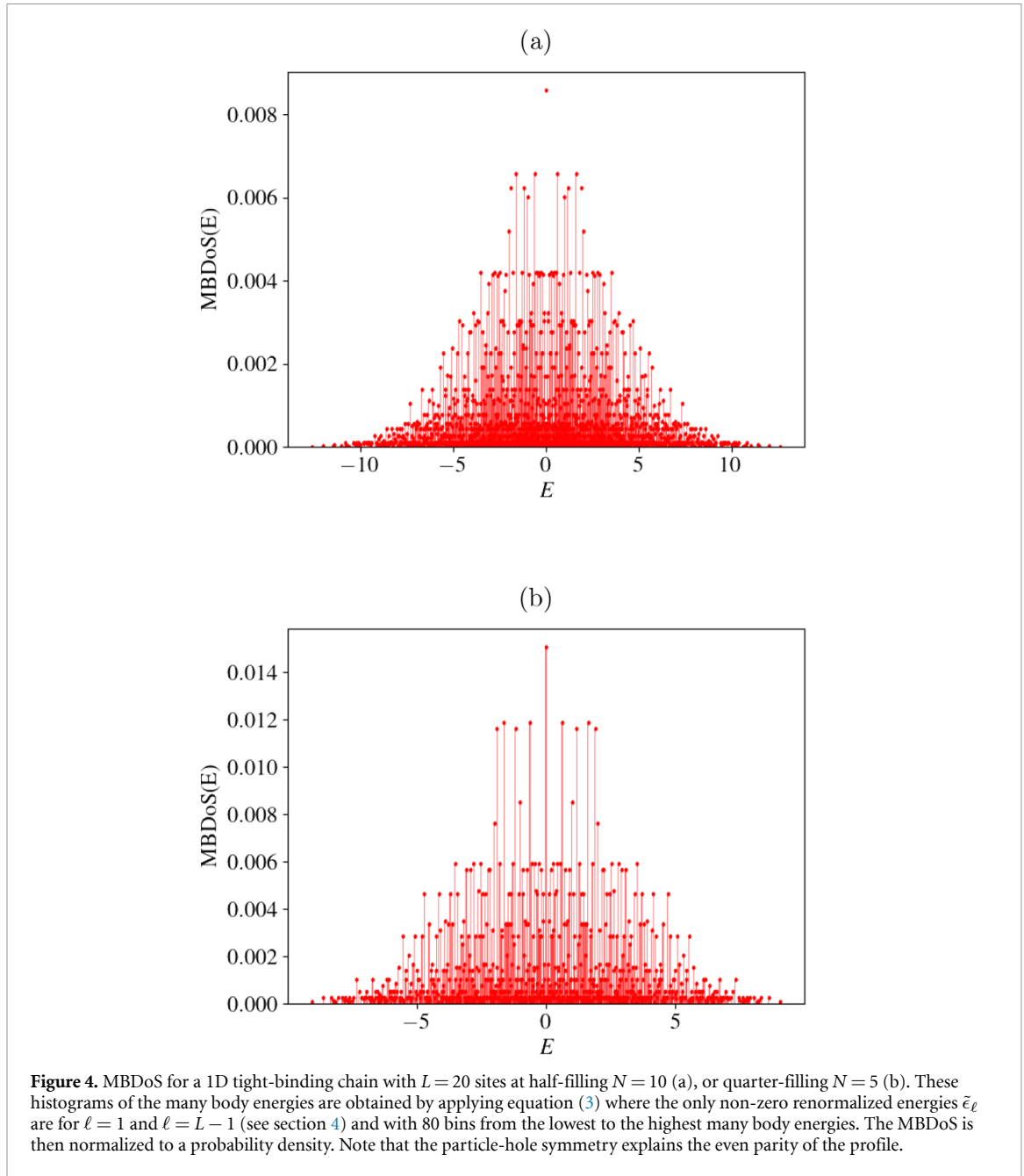
We can write the renormalized energies:

$$\begin{aligned} \tilde{\epsilon}_\ell &= \sum_{n=0}^{L-1} \left( -2t \cos\left(\frac{2\pi n}{L}\right) \right) \frac{1}{\sqrt{L}} \omega_L^{\ell n} \\ &= -\frac{t}{\sqrt{L}} \sum_{n=0}^{L-1} \left( e^{i \frac{2\pi n}{L} (\ell+1)} + e^{i \frac{2\pi n}{L} (\ell-1)} \right) \end{aligned}$$

where one can note that the first term is equal to  $L$  when  $\ell = L-1$  and 0 otherwise. Similarly, the second term is equal to  $L$  when  $\ell = 1$  and 0 otherwise. We have  $\tilde{\epsilon}_1 = \tilde{\epsilon}_{L-1} = -t\sqrt{L}$  and all other  $\tilde{\epsilon}_\ell$  are zero. Using equation (3), we obtain the MB spectrum:

$$E = -t\sqrt{L}(U_1 + U_{L-1}) = -2t\sqrt{L} \operatorname{Re}[U_1]$$

since  $U_1$  and  $U_{L-1}$  are conjugate to each other. The obtained MBDoS is displayed in figure 4.



#### 4.2. Tight-Binding 2D

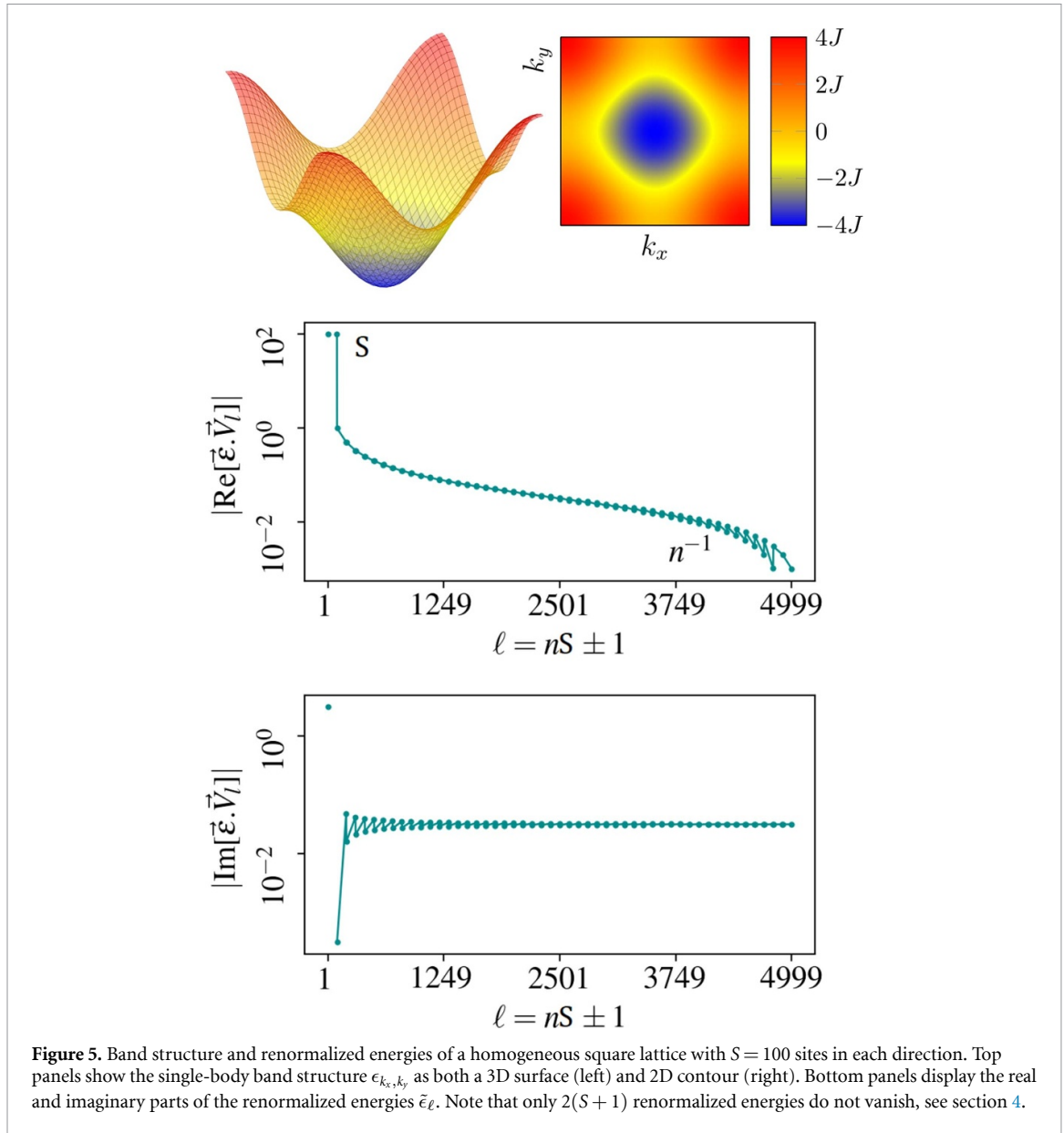
For a 2D square lattice with  $S \times S$  sites, the band structure contains  $L = S^2$  single-body energies given by:

$$\epsilon_{k_x, k_y} = -2t \cos(k_x) - 2t \cos(k_y), \quad (10)$$

where, under PBC, the momenta in each direction are  $k_x = \frac{2\pi n_x}{S}$  and  $k_y = \frac{2\pi n_y}{S}$ , with  $n_x, n_y \in \{0, \dots, S-1\}$ .

We can flatten this 2D array of energies row by row to obtain a 1D array of size  $L$  (indexed by  $n_y S + n_x$ ) and re-use the 1D case method. After some manipulations, we obtain

$$\begin{aligned} \tilde{\epsilon}_{nS \pm 1} &= -t \frac{\sin(\frac{\pi}{S})}{\sin(\frac{\pi}{L}(nS \pm 1))} e^{-i\frac{\pi}{L}((n \mp 1)S \pm 1)}, \\ \tilde{\epsilon}_1 &= \tilde{\epsilon}_{L-1}^* = -t \frac{\sin(\frac{\pi}{S})}{\sin(\frac{\pi}{L})} e^{-i\frac{\pi}{L}(-S+1)}, \\ \tilde{\epsilon}_S &= \tilde{\epsilon}_{L-S} = -tS, \end{aligned}$$



where  $n \in \{1, \dots, S - 1\}$  and all other  $\tilde{\epsilon}_\ell$  are zero. As expected, they come in conjugate pairs. Interestingly, for  $S$  large,  $\tilde{\epsilon}_1 \approx -tS$  and other renormalized energies  $\tilde{\epsilon}_{nS \pm 1} \approx -\frac{t}{n}$  are of order  $S^{-1}$ . In other words, there are two dominant renormalized energies :  $\tilde{\epsilon}_1$  and  $\tilde{\epsilon}_S$  which are almost equal (and their conjugate  $\tilde{\epsilon}_{L-1}$  and  $\tilde{\epsilon}_{L-S}$ ). As a result, the many body energies in the large  $S$  limit are given by :

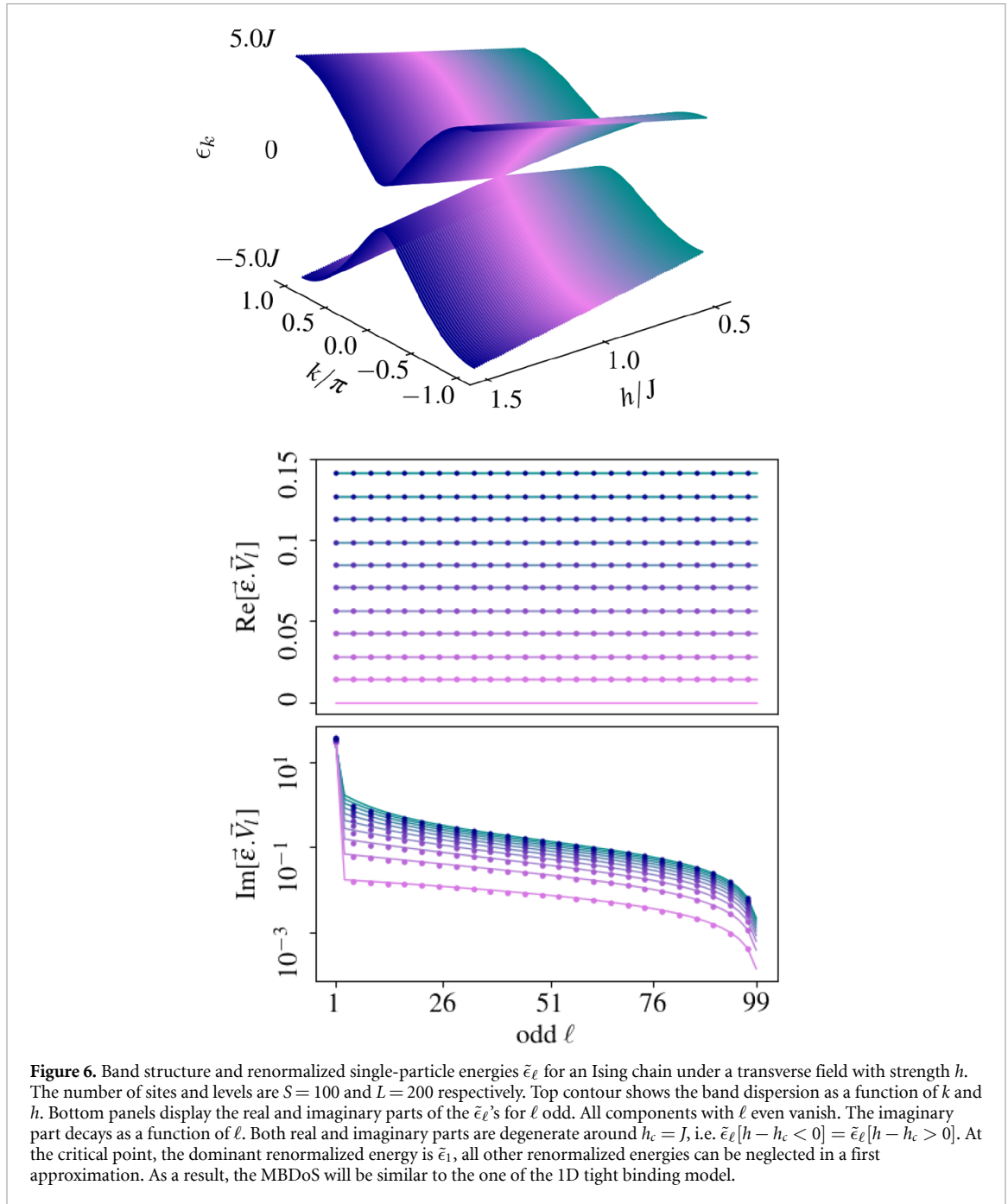
$$E \approx -2tS(\text{Re}[U_1] + \text{Re}[U_S]).$$

Since all vectors  $U_\ell$  are centered ( $\sum_n U_\ell^n = 0$ ) and orthogonal ( $U_\ell \cdot U_{\ell'} = 0$ ), the distributions of their components can be considered as two uncorrelated random processes. In a first approximation, the distribution of the sum will be the convolution of the two distributions. The band structure in equation (10) and the non-vanishing renormalized energies  $\tilde{\epsilon}_\ell$  for a square lattice with  $S = 100$  sites are shown in figure 5.

### 4.3. Transverse Ising

We now consider the Hamiltonian for a 1D chain of spins of size  $S$ :

$$H = -J \sum_{j=0}^{S-1} \sigma_j^x \sigma_{j+1}^x - h \sum_{j=0}^{S-1} \sigma_j^z \quad (11)$$



where  $\sigma^d$  are the Pauli matrices,  $J$  is the nearest neighbor coupling and  $h$  is the transverse field, and we set  $\sigma_S^x = \sigma_0^x$  for convenience in the notation.

After applying the JW transformation, we can obtain the dispersion relation of this model as follows:

$$\epsilon_k^\pm = \pm 2\sqrt{h^2 + J^2 - 2hJ\cos(k)} - 2h \quad (12)$$

where  $k = \frac{2\pi n}{S}$  is the momentum with PBC ( $n \in \{0, \dots, S-1\}$ ). Note that now, two bands contribute to the DoSs with a separating gap closing at the critical point for  $h = J$  (see figure 6.).

In this case, the dispersion relation is bi-valued: we have  $L = 2S$  single body energies which can be arranged into a single vector. For instance, we can concatenate all  $\epsilon_k^+$  followed by all  $\epsilon_k^-$ :

$$\vec{\epsilon} = (\epsilon_0^+, \dots, \epsilon_{S-1}^+, \epsilon_0^-, \dots, \epsilon_{S-1}^-)$$

which gives the following renormalized SB spectrum :

$$\tilde{\epsilon}_\ell = \frac{1}{\sqrt{L}} \sum_{k=0}^{S-1} (\epsilon_k^+ + (-1)^\ell \epsilon_k^-) e^{i \frac{2\pi \ell}{L} k}$$

where  $\tilde{\epsilon}_0 = -2h\sqrt{L}$ , all  $\tilde{\epsilon}_\ell$  vanish for  $\ell$  even and the real part for  $\ell$  odd has a single fixed value  $\text{Re}[\tilde{\epsilon}_\ell] = 4|h - J|/\sqrt{L}$  independent of  $\ell$ . For the imaginary part,  $\text{Im}[\tilde{\epsilon}_\ell]$  is non-zero but decays very fast with faster rate as the system size increases. The factor  $|h - J|$  in the real part introduces a symmetry around the critical point  $h_c = J$  where the only surviving energies are  $\tilde{\epsilon}_1 = -\tilde{\epsilon}_{L-1} = i2J\sqrt{L}$  : we recover the same renormalized single body energies as for the 1D tight binding model, up to a Wick rotation. Those features can be observed in figure 6, where we show the real and imaginary parts of  $\tilde{\epsilon}_\ell$  for  $h/J \in [0.5, 1.5]$ .

Using these previous results, we have all ingredients to compute the MBDoS for the applications aforementioned. In particular, for the 1D tight binding model and the transverse Ising chain at the critical point which have the same renormalized single body energies. As a result, they have the same MBDoS, displayed on figure 4.

## 5. Generalization

Extending the method to additional quantum numbers and to bosons requires one additional step. The only part in our method which is specific to spinless fermions is the very first step where we started from binary strings to represent many body configurations. Suppose the set of conserved quantum numbers now allows for  $K$  particles to occupy each single-body energy level, the representation of many body configurations is simply changed from binary strings to  $K$ -restricted weak  $L$ -compositions of the number of particles  $N$ . We already introduced how to fold binary strings into compositions and the winding factor  $Q$  is known, i.e. how many binary strings were mapped to a given composition. This can also be done in the reverse way, that is, *unfolding* compositions into binary strings. We are effectively duplicating single body energy levels and assigning them new indices in the list of configurations and single body energies. This allows to encode degeneracies in the single body spectrum. Once this additional preliminary step is done, the method can be applied without further modification. The bosonic case is just a particular case where  $K = N$ , as long as the number of particles is kept finite.

On a different aspect, we cannot generalize this approach to open quantum systems, we remain in the context of a closed quantum system without environment and as such, there is *a priori* no concept of temperature. However, one can assume a micro-canonical hypothesis and consider the derivative with respect to energy of the Boltzmann entropy  $S = k \log W$  where  $W$  should be replaced by the MBDoS times an energy interval defining the so called accessible states. Such logarithmic derivative of the MBDoS actually defines an *effective* inverse temperature. The choice of the energy interval is crucial to define temperature and entropy in a way that remains valid for small finite system sizes which are exactly diagonalizable in numerical tests and has been recently investigated in [38].

Finally, regarding interactions between particles, it is worth noticing that the  $U_\ell$  vectors, through the distributions of their components, provide an interesting classification of the many body states into degenerate classes (i.e. states clustered on a single many body energy  $U_\ell^m$ ). At first order in perturbation theory, the effect of interactions should be considered on such states lying in a given class since degeneracies might be partially lifted. This will be investigated in a further publication. In addition, in the case of a random interaction like the Gaussian Unitary or Gaussian Orthogonal ensembles, the calculation of the MBDoS *with interactions* can be done from the non interacting MBDoS, by considering the subordination property between their Stieltjes transforms [11, 39]:  $m_{H_0+W}(z) = m_{H_0}(z + \sigma_w^2 m_H(z))$ , with  $m_H(z) = \text{Tr}((H - z)^{-1})$  where  $\sigma_w^2 m_H(z)$  can be considered as a many body self energy.

## 6. Conclusion

In the present paper, we explored a new approach to compute the many-body density of states of non-interacting systems of identical particles, and we focused on the case of spinless fermions.

We showed that the many-body spectrum can be expanded over the principal components of a filling matrix encoding the allowed many body states. These principal components describe universal spectral properties of quantum many body systems which only depend on the number of one body levels and the number of particles. All system-specific quantities are represented as renormalized energies of the single body spectrum which act by weighting those principal components.

For gapless systems, such as the tight-binding and the critical transverse field Ising chains, we demonstrated that only two renormalized energies are non zero. Even in more general scenari, such as the

square lattice or the Ising chain away from the critical point, many renormalized energies still vanish. In all cases, this will significantly reduce the number of relevant spectral components of the filling matrix involved in the calculation of the many body density of states.

Our framework allows efficient numerical computations that only need to be done once for all systems sharing the same number of single body energies and particle counts, which is achieved by separating costly combinatorial operations from system-dependent quantities. After a full description of the spinless fermionic case, we gave a simple way to extend the method to include additional quantum numbers like spin, and to handle bosonic systems.

## Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

## Acknowledgment

We acknowledge support from the Leverhulme Trust under Grant RPG-2020-094.

## ORCID iD

Grégoire Ithier  <https://orcid.org/0000-0001-5643-7349>

## References

- [1] Zelevinsky V and Horoi M 2019 *Prog. Part. Nucl. Phys.* **105** 180
- [2] Ashcroft N W and Mermin N D 1976 *Solid State Physics* HRW international edn (New York: Holt, Rinehart and Winston)
- [3] Gogolin C and Eisert J 2016 *Rep. Prog. Phys.* **79** 056001
- [4] Santos L F and Rigol M 2010 *Phys. Rev. E* **81** 036206
- [5] Borgonovi F, Izrailev F, Santos L and Zelevinsky V 2016 *Phys. Rep.* **626** 1
- [6] Polkovnikov A, Sengupta K, Silva A and Vengalattore M 2011 *Rev. Mod. Phys.* **83** 863
- [7] Nandkishore R and Huse D A 2015 *Annu. Rev. Condens. Matter Phys.* **6** 15
- [8] Yang Y, Iblisdir S, Cirac J I and Bañuls M C 2020 *Phys. Rev. Lett.* **124** 100602
- [9] Abanin D A, Altman E, Bloch I and Serbyn M 2019 *Rev. Mod. Phys.* **91** 021001
- [10] Ithier G and Benaych-Georges F 2017 *Phys. Rev. A* **96** 012108
- [11] Ithier G, Ascroft S and Benaych-Georges F 2017 *Phys. Rev. E* **96** 060102
- [12] Santos L F, Távora M and Pérez-Bernal F 2016 *Phys. Rev. A* **94** 012113
- [13] Weiße A and Fehske H 2008 *Computational Many-Particle Physics* (Berlin: Springer) pp 529–44
- [14] Silver R N and Röder H 1994 Densities of states of mega-dimensional hamiltonian matrices *Int. J. Mod. Phys. C* **5** 735–53
- [15] Silver R, Roeder H, Voter A and Kress J 1996 *J. Comput. Phys.* **124** 115
- [16] Weiße A, Wellein G, Alvermann A and Fehske H 2006 *Rev. Mod. Phys.* **78** 275
- [17] Schrodri F, Silvi P, Tschirsich F, Fazio R and Montangero S 2017 *Phys. Rev. B* **96** 094303
- [18] Papaefstathiou I, Robaina D, Cirac J I and Bañuls M C 2021 *Phys. Rev. D* **104** 014514
- [19] Torres-Herrera E J and Santos L F 2014 *Phys. Rev. A* **90** 033623
- [20] Torres-Herrera E J, Vyas M and Santos L F 2014 *New J. Phys.* **16** 063010
- [21] Bethe H A 1936 *Phys. Rev.* **50** 332
- [22] Volya A and Zelevinsky V 2006 *Phys. Rev. C* **74** 064314
- [23] Hillman M and Grover J R 1969 *Phys. Rev.* **185** 1303
- [24] Berger J and Martinot M 1974 *Nucl. Phys. A* **226** 391
- [25] Jacquemin C and Kataria S 1986 *Z. Phys. A* **324** 261
- [26] Bohr A N and Mottelson B R 1998 *Nuclear Structure* vol 2 (Singapore: World Scientific)
- [27] Fetter A and Walecka J 2012 *Quantum Theory of Many-Particle Systems (Dover Books on Physics)* (New York: Dover Publications)
- [28] Redfield J H 1927 *Am. J. Math.* **49** 433
- [29] Ruskey F, Savage C and Min Yih Wang T 1992 *J. Algorithms* **13** 414
- [30] Sawada J and Williams A 2013 *Theor. Comput. Sci.* **502** 46–54
- [31] Stanley R P 2012 *Enumerative Combinatorics* 2nd edn 1 Cambridge Studies in Advanced Mathematics (New York)
- [32] Girardeau M 1960 *J. Math. Phys.* **1** 516
- [33] Donohue P, Tsuchiizu M, Giamarchi T and Suzumura Y 2001 *Phys. Rev. B* **63** 045121
- [34] Imambekov A, Schmidt T L and Glazman L I 2012 *Rev. Mod. Phys.* **84** 1253
- [35] Turner A M, Pollmann F and Berg E 2011 *Phys. Rev. B* **83** 075102
- [36] Kitaev A Y 2001 *Phys.-Usp.* **44** 131
- [37] Alicea J 2010 *Phys. Rev. B* **81** 125318
- [38] Burke P C and Haque M 2023 *Phys. Rev. E* **107** 034125
- [39] Ithier G and Ascroft S 2018 *J. Phys. A: Math. Theor.* **51** 48LT01