

Interaction-independent zero modes of the half-filled Hubbard model on arbitrary hopping graphs: exact dimension, explicit basis and complete $\text{SO}(4)$ content

scientific-author-90*

9 September 2026

Abstract

Eigenstates of the Hubbard model whose eigenvectors do not depend on the interaction strength are exact at every coupling, yet they are known only in scattered families, and classification schemes phrased in $\text{SO}(4)$ quantum numbers presuppose that each of them carries a definite pseudospin and spin label. This paper settles them for the no-doublon sector at one electron per site on an arbitrary finite hopping graph—disconnected or non-bipartite included—with arbitrary nonzero hopping amplitudes, in every magnetization sector. The descending chain of invariant subspaces collapses after one step, so the interaction-independent zero modes are exactly the kernel of the hopping operator on the no-doublon space, with no generalized eigenvectors; the dimension of that kernel in a given sector is the number of ways of distributing the up-spins over the connected components subject only to the component-size bounds; and an explicit basis is given by all-plus-one integer vectors indexed by these distributions, which coincide coefficient by coefficient with products of component spin-lowering operators applied to the all-up state. Every basis state is an exact pseudospin singlet that factorizes into maximal-spin states of the components; its minimum number of Slater determinants is a product of binomial coefficients, its up/down entanglement spectrum is flat, and its entropy is the corresponding sum of logarithms of binomial coefficients. The zero-mode space itself is a tensor product of one spin multiplet per connected component: it decomposes into $\text{SO}(4)$ sectors of vanishing pseudospin and definite spin with exact multiplicities, and the total-spin operator acts on it through an explicit integer matrix that involves neither the hopping amplitudes nor fermionic operators. A basis state has definite total spin exactly when the graph is connected or the sector is extremal, so bond dilution destroys the purity of the natural basis while preserving the sector labels of the space, and it converts the unique maximal-spin zero mode of a connected lattice into an extensive residual degeneracy growing between $\log(3)/2$ and $\log 2$ nats per site for the two diluted patterns evaluated here exactly. Every statement is unchanged when the Hubbard interaction is replaced by an arbitrary density-density interaction. The theorems hold for every graph size and are accompanied by exact rational-arithmetic certificates on exhaustive small-graph families and on curated larger ones, and by exact integer combinatorics at large sizes. Nothing here locates these zero modes within the spectrum, takes a thermodynamic limit, or treats sectors containing doublons.

1 Introduction

The Hubbard model [3] is the standard arena for strong correlations in lattice fermions, and its exact eigenstates are correspondingly precious. Among them sits a class that is exact for *every* value of the interaction strength: states whose eigenvector does not depend on the interaction parameter at all. Because such a state diagonalizes the hopping and interaction parts of the

*Corresponding author. GitLab identity `scientific-author-90`; Solution repository `scientific-author-90/hubbard-w-independent-general-bipartite-blocks` (internal), the exact revision of which is bound to the Journal intake.

Hamiltonian separately, no level crossing and no renormalization can move its eigenvector as the coupling is varied; it is an anchor of the spectrum at weak and strong coupling alike. The celebrated examples are the η -pairing towers of Yang [10] and the $\text{SO}(4)$ structure that organizes them [11, 12], and a sustained program has since constructed and classified interaction-independent exact eigenstates of Hubbard-type models, on lattices and on general graphs, with and without bond-charge interactions and in multicomponent generalizations [14, 8, 7, 13]. Such states also supply the exactly solvable building blocks from which quantum many-body scars and Hilbert-space fragmentation are assembled [6, 4, 5], and their entanglement structure is atypical enough to be interesting in its own right [9].

This paper answers the structure question for one natural sector completely. Fix a finite simple graph G and consider the Hubbard Hamiltonian $H = T + WD$ on G with arbitrary nonzero hopping amplitudes (Section 2). In the no-doublon space—states in which no site carries two electrons—at one electron per site, an interaction-independent eigenstate with zero doublon number is a simultaneous eigenvector of T and of D , and the set of all such candidates is the largest T -invariant subspace of $\ker D$, obtained as the stable limit of a descending chain. Earlier work in this program established the connected bipartite case with nonzero amplitudes: the half-filled no-doublon space of interaction-independent states is one-dimensional, spanned by the maximal-spin tower state, with maximal Slater rank [2]. Any lattice with a missing bond—a bond-diluted, vacancy-ordered or decoupled-cluster Hubbard system—lies outside that result, and a classification of interaction-independent eigenstates by $\text{SO}(4)$ quantum numbers [1, 14, 13] presupposes that each of them sits in a definite pseudospin-spin sector, which a disconnected hopping graph a priori need not provide. The questions treated here are therefore:

1. the dimension of the interaction-independent zero-mode space, as a closed formula in the connected-component sizes;
2. an explicit basis, with the fermion sign convention settled by computation rather than asserted;
3. whether the descending chain has genuine generalized eigenvectors;
4. the exact minimum number of Slater determinants, the up/down entanglement spectrum and the von Neumann entropy of every basis state;
5. the exact $\text{SO}(4)$ content—pseudospin j and total spin S —of every basis state, and of the whole zero-mode space;
6. how much of this depends on the interaction being the Hubbard term rather than a general density-density interaction;
7. where the hypothesis that every hopping amplitude is nonzero actually bites: what a zero amplitude does.

All seven are answered by proofs that hold for every graph size, every graph, and every choice of nonzero amplitudes; the answers are summarized here and proved in Sections 3–8. The chain collapses after one step—at one electron per site every hop lands on an occupied site—so the zero modes are exactly the kernel of the hopping operator on the no-doublon space, and every one of them is a genuine eigenvector of T with eigenvalue zero (Lemma 1, Theorem 5). In each sector the dimension is the number of ways of distributing the up-spins over the connected components subject only to the component-size bounds, and an explicit basis consists of all-plus-one integer vectors indexed by these distributions (Theorem 5); the connected case collapses to a single state, recovering the earlier result as the special case $r = 1$. Each basis state is simultaneously a product of component spin-lowering operators applied to the all-up state, which identifies the whole zero-mode space as a tensor product of one maximal-spin multiplet per connected component, with an integer, amplitude-free matrix for the total-spin operator (Theorem 6). From that matrix the

purity criterion is exact: a basis state has definite total spin if and only if the graph is connected or the sector is extremal (Theorem 7), and the abstract spin-coordinate form of the criterion is a two-clause pair condition, complete for arbitrary products of spin representations (Proposition 12). Every zero mode is an exact pseudospin singlet for any phase assignment, so the $\text{SO}(4)$ content of the space is a direct sum of sectors $\{0, S\}$ with exact multiplicities n_S , and a basis of definite- $\text{SO}(4)$ -sector zero modes exists on every graph (Theorems 6 and 8). The Slater rank of a basis state is exactly the product of binomial coefficients over the components, its up/down spectrum is flat, and its entropy is the sum of the corresponding logarithms (Theorem 9); for definite-spin zero modes the Slater rank equals the support size (Corollary 10). Nothing changes for an arbitrary density-density interaction (Corollary 11), a zero amplitude is exactly an edge deletion (Remark 15), and bond dilution converts the single zero mode of a connected lattice into an extensive residual degeneracy growing, for the two patterns bracketing the dilution, between $\log(3)/2$ and $\log 2$ nats per site (Corollary 16).

Every structural claim is accompanied by an independent exact certificate: integer and rational arithmetic with no floating-point input, produced by three implementations that share only the mode-ordering convention and agree state by state where they overlap. The certificates cover every graph at two and four sites, every bipartite graph at six sites, and curated non-bipartite and larger families at six, eight and ten sites, in every magnetization sector and at unit as well as pseudo-random nonzero amplitudes; the purely combinatorial consequences are evaluated exactly up to four hundred sites. They are inventoried as C1–C22 in Section 9, which also states precisely what was *not* computed. The programs, the complete numerical outputs and the underlying proofs are archived with the source of this work [2].

The scope is equally precise: one electron per site, zero doublon number, exact eigenvectors. Nothing here locates these energies within the spectrum, takes a thermodynamic limit, addresses dynamics, or treats sectors containing doublons or blocks with fewer than M electrons; each of these is named as an open direction in Section 10.

2 Model, conventions and the zero-mode problem

2.1 The Hamiltonian and the fermion conventions

Let $G = (V, E)$ be a finite simple graph, $V = \{0, 1, \dots, M-1\}$, with real hopping amplitudes $w_{ij} = w_{ji}$ on the edges. The Hubbard Hamiltonian is

$$H = T + WD, \quad T = - \sum_{(i,j) \in E} \sum_{\sigma \in \{\uparrow, \downarrow\}} w_{ij} (c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma}), \quad D = \sum_{i \in V} n_{i\uparrow} n_{i\downarrow}, \quad (1)$$

where $c_{i\sigma}^\dagger$ creates an electron of spin σ on site i , $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$, $n_i = n_{i\uparrow} + n_{i\downarrow}$, and $W \in \mathbb{R}$ is the interaction strength, left completely arbitrary. D counts doublons: sites carrying both spins. Throughout Sections 3–8 every amplitude on an edge of E is assumed nonzero, $w_{ij} \neq 0$; the boundary case $w_{ij} = 0$ is treated in Remark 15. Neither connectedness nor bipartiteness of G is assumed at any point.

Fermion modes are ordered in the *interleaved* convention: the up mode of site i is mode $2i$ and the down mode is mode $2i + 1$. For a set S of modes, $|S\rangle = \prod_{q \in S}^\nearrow a_q^\dagger |\text{vac}\rangle$ is the basis state with the modes of S occupied, the product taken in increasing mode order, and

$$c_q |S\rangle = (-1)^{p(S,q)} |S \setminus \{q\}\rangle \quad (q \in S), \quad p(S,q) = \#\{q' \in S : q' < q\}, \quad (2)$$

with $c_q |S\rangle = 0$ for $q \notin S$ and the analogous rule for creation. This convention fixes every sign that appears below; the sign question raised by the level-set basis—are its primitive integer coefficient vectors all +1, or a signed pattern?—is settled by an exact parity computation in Lemma 3, not by assertion.

2.2 Blocks, the no-doublon space and interaction independence

The total particle number and both spin populations are conserved by H , so the Fock space splits into blocks $(N_\uparrow, N_\downarrow)$. This paper works in the block of total particle number M , denoted $\mathcal{F}_M$: one electron per site on average. Inside $\mathcal{F}_M$ the *no-doublon space* is $\ker D \cap \mathcal{F}_M$; a state in it has exactly one electron on every site and is therefore determined by its *up-set* $I \subseteq V$, the set of sites carrying the up electron, the down-set being the complement I^c . Write

$$|I\rangle \equiv |I \text{ up}, I^c \text{ down}\rangle, \quad \ker D \cap \mathcal{F}_M = \text{span}\{|I\rangle : I \subseteq V\}, \quad \dim \ker D \cap \mathcal{F}_M = 2^M. \quad (3)$$

The total spin projection grades this space into the blocks $(p, M-p)$ with $p = |I| = N_\uparrow$ and $S^z = p - M/2$, of dimension $\binom{M}{p}$. *Half filling* is the block (n, n) with $M = 2n$ even, $p = n$, $S^z = 0$, of dimension $\binom{M}{n}$. Because every site of a state $|I\rangle$ carries exactly one electron, the no-doublon space at particle number M is literally the M -site spin- $\frac{1}{2}$ space, site i carrying the spin of its electron; this identification is what makes the $\text{SO}(4)$ analysis of Sections 4–7 possible.

An eigenstate of H whose eigenvector does not depend on W is called *interaction-independent* (or *W -independent*). Since $H = T + WD$, a vector v is an eigenvector of H for every value of W if and only if v is a simultaneous eigenvector of T and of D . Restricting to D -eigenvalue 0, the natural home of such states is the largest T -invariant subspace of the no-doublon space, obtained as the stable limit of the descending chain

$$\mathcal{V}_0 \supseteq \mathcal{V}_1 \supseteq \mathcal{V}_2 \supseteq \cdots, \quad \mathcal{V}_0 = \ker D \cap \mathcal{F}_M, \quad \mathcal{V}_{k+1} = T^{-1}(\mathcal{V}_k) \cap \mathcal{V}_k, \quad (4)$$

which stabilizes because the space is finite-dimensional; write $\mathcal{V}_\infty$ for the stable subspace. Every eigenvector of T that lies in $\ker D$ belongs to $\mathcal{V}_\infty$, and if T acts on $\mathcal{V}_\infty$ as a scalar then every vector of $\mathcal{V}_\infty$ is an exact eigenstate of H for every W , with the same eigenvector. The vectors of $\mathcal{V}_\infty$ are the *interaction-independent zero modes* of this paper, for the following reason: Lemma 1 below shows that T annihilates $\mathcal{V}_\infty$, so the chain collapses to the kernel of T , the scalar is 0, and no generalized (Jordan) eigenvectors occur. All operators in (4) preserve the block decomposition, so the chain splits over the sectors; write

$$Z_p = \mathcal{V}_\infty \cap \{S^z = p - M/2\}, \quad Z = \bigoplus_{p=0}^M Z_p, \quad (5)$$

so that at half filling $\mathcal{V}_\infty$ in the block (n, n) is Z_n .

2.3 Components, occupancies and level sets

Let $V_1, \dots, V_r$ be the connected components of G and $b_i = |V_i|$ their sizes, $\sum_i b_i = M$. (Chain subspaces are calligraphic, $\mathcal{V}_k$; components are roman, V_i ; spin representations will be $\mathcal{H}_j$.) Each up-set I has an *occupancy vector*

$$c(I) = (|I \cap V_1|, \dots, |I \cap V_r|) \in \mathbb{Z}^r. \quad (6)$$

A vector c is *admissible* for the block $(p, M-p)$ if $0 \leq c_i \leq b_i$ for all i and $\sum_i c_i = p$. The *level set* of an admissible c is

$$L_c = \{I \subseteq V : c(I) = c\}, \quad |L_c| = \prod_{i=1}^r \binom{b_i}{c_i}, \quad (7)$$

by independent choices inside the components; summing over the admissible c recovers the whole block, $\sum_c |L_c| = \binom{M}{p}$ (Vandermonde's identity). The *all-plus-one indicator* of a level set is the integer vector

$$\mathbf{1}_c = \sum_{I \in L_c} |I\rangle \in \ker D \cap \mathcal{F}_M. \quad (8)$$

Distinct level sets have disjoint supports, so any collection of $\mathbf{1}_c$ is linearly independent. Finally, a *token slide* is a move $I \rightarrow I - \{x\} + \{y\}$ with $(x, y) \in E$, $x \in I$, $y \notin I$: an up-electron is moved along an edge to an adjacent site, whose electron is flipped down in exchange. Sliding preserves $c(I)$, so each L_c is a union of sliding-connected classes; Lemma 4 shows it is a single class.

3 Structure of the zero-mode space

The argument has four steps: the chain collapses (Lemma 1); the hopping amplitudes factor out of every equation (Lemma 2); the fermion signs force the coefficient function to be constant along token slides (Lemma 3); and the sliding classes are exactly the level sets (Lemma 4). Theorem 5 assembles them.

Lemma 1 (No $d = 0$ matrix elements; the chain stops after one step). *For any graph and any amplitudes, T maps $\ker D \cap \mathcal{F}_M$ into the one-doublon subspace of $\mathcal{F}_M$: T has no matrix element between two no-doublon states at particle number M . Consequently $\mathcal{V}_1 = \ker(T|_{\ker D \cap \mathcal{F}_M})$ and*

$$\mathcal{V}_\infty = \mathcal{V}_1 = \mathcal{V}_2 = \dots = \ker(T|_{\ker D \cap \mathcal{F}_M}), \quad (9)$$

sector by sector: $Z_p = \ker(T|_{\ker D \cap \mathcal{F}_M, S^z=p-M/2})$. Every interaction-independent no-doublon state at particle number M is a genuine eigenvector of T with eigenvalue 0, and there is no generalized eigenvector structure.

Proof. A state of $\ker D \cap \mathcal{F}_M$ has every site singly occupied, so there is no empty site. One application of a hopping term moves one electron from a site a to a site b ; the target site b already carries an electron, and it carries the opposite spin, for otherwise the source state would be doubly occupied at b . The image therefore has a doublon at b and a hole at a : it lies in the one-doublon subspace, whose intersection with $\ker D$ is $\{0\}$. Hence every no-doublon column of $T|_{\ker D}$ vanishes, and $v \in \mathcal{V}_1$ means $Tv \in \ker D$, which forces $Tv = 0$: $\mathcal{V}_1 = \ker(T|_{\ker D})$. Since $\ker(T|_{\ker D}) \subseteq \mathcal{V}_\infty \subseteq \mathcal{V}_1$, all terms coincide. An eigenvector of T inside $\ker D$ with eigenvalue E has $E \cdot v$ equal to the no-doublon part of Tv , which is zero, so $E = 0$: every interaction-independent zero mode is a zero-energy eigenstate of H for every W . $\square$

This is the mechanism behind the whole paper: one electron per site leaves nowhere to hop *within* the no-doublon space. In a block with fewer than M electrons empty sites exist, T has nonzero $d = 0$ matrix elements, and the chain need not collapse (Section 10).

Lemma 2 (Two contributions per target; the amplitudes factor out). *Fix an edge $(a, b) \in E$ and an up-set J with $a \notin J$, $b \in J$, and let $\tau(J; a, b)$ be the one-doublon state with up-set J , doublon at b and hole at a . Then $\tau(J; a, b)$ is reached from the no-doublon space by exactly two processes: the down-hop $a \rightarrow b$ out of the source state $|J\rangle$, and the up-hop $a \rightarrow b$ out of the source state $|J + \{a\} - \{b\}\rangle$. Both carry the same prefactor $-w_{ab}$, and the coefficient of $\tau(J; a, b)$ in Tv is*

$$-w_{ab}(s_{\text{up}}\psi_B + s_{\text{dn}}\psi_A), \quad \psi_A = \langle J|v\rangle, \quad \psi_B = \langle J + \{a\} - \{b\}|v\rangle, \quad (10)$$

with fermion signs $s_{\text{up}}, s_{\text{dn}} \in \{-1, +1\}$. For $w_{ab} \neq 0$ the equation “coefficient of τ vanishes” is independent of the value of w_{ab} : everything that follows holds for every choice of nonzero amplitudes, with the same integer solution vectors.

Proof. The hole and the doublon of τ are a and b , so the electron that moved went from a to b , which requires $(a, b) \in E$; the moved electron is either the down electron of a —source $|J\rangle$, in which site a holds the down electron since $a \notin J$ and site b holds the up electron since $b \in J$ —or the up electron of a —source $|J + \{a\} - \{b\}\rangle$, in which site a holds up and site b holds down. No other edge or spin contributes, and both terms come from the summand of T labeled by (a, b) , hence the common prefactor $-w_{ab}$, which divides out of the vanishing equation when $w_{ab} \neq 0$. $\square$

There are $2|E|\binom{M-2}{p-1}$ such target equations in the block $(p, M-p)$ —an ordered edge, then $p-1$ further sites of J among the remaining $M-2$ —and none when $E = \emptyset$.

Lemma 3 (The sign identity). *In the interleaved convention (2), $s_{\text{up}} = -s_{\text{dn}}$ for every edge, direction and target. The equations $Tv = 0$ on $\ker D \cap \mathcal{F}_M$ are therefore exactly*

$$\psi_{J+\{a\}-\{b\}} = \psi_J \quad \text{for every } (a, b) \in E, \ a \notin J, \ b \in J, \quad (11)$$

i.e. the coefficient function $\psi_I = \langle I|v \rangle$ is constant along every token slide. In particular the primitive integer solutions are exactly the indicator vectors of unions of sliding classes, with all coefficients $+1$: the global minus sign of T lives in the operator, never in the eigenvectors.

Proof. Count occupied modes strictly below the mode acted upon. In both source states every site below a contributes exactly one occupied mode, and site a itself contributes the annihilated mode only, so the annihilation signs are $(-1)^a$ in both cases: $p(S_A, 2a+1) = a$ for the down-hop out of $A = |J\rangle$ (site a holds the down electron, mode $2a$ empty) and $p(S_B, 2a) = a$ for the up-hop out of $B = |J+\{a\}-\{b\}\rangle$ (site a holds the up electron). After the annihilation, count again below the created mode. For the down-hop, created at mode $2b+1$: the sites strictly between a and b contribute $b-a-1$ occupied modes, and site b still contributes its up electron at mode $2b < 2b+1$, so $p(S_A \setminus \{2a+1\}, 2b+1) = a + (b-a-1) + 1 = b$ and $s_{\text{dn}} = (-1)^{a+b}$. For the up-hop, created at mode $2b$: site b of S_B holds the down electron at mode $2b+1 \geq 2b$, which is not counted, so $p(S_B \setminus \{2a\}, 2b) = a + (b-a-1) = b-1$ and $s_{\text{up}} = (-1)^{a+b-1} = -s_{\text{dn}}$. Substituting into (10) gives $-w_{ab}s_{\text{dn}}(\psi_A - \psi_B) = 0$, i.e. (11). Every equation of the kernel equates two coefficients, so the integer solutions are spanned by $\{0, 1\}$ -valued indicators of sliding-closed sets, all of whose coefficients are $+1$. $\square$

Lemma 4 (The sliding classes are the level sets). *If G is connected then the token-sliding graph on the p -subsets of V is connected, for every p . Hence for general G the sliding classes are exactly the level sets L_c of (7), and $|L_c| = \prod_i \binom{b_i}{c_i}$.*

Proof. Let G be connected and let $I \neq I'$ be two p -subsets. Define the potential

$$\Phi(I) = \min_{\sigma: I \rightarrow I'} \sum_{x \in I} d(x, \sigma(x)), \quad (12)$$

the minimum over bijections σ of the total graph distance; Φ is a nonnegative integer and $\Phi(I) = 0$ if and only if $I = I'$. Pick $y \in I' \setminus I$, pick a token $x \in I$ minimizing $d(x, y)$, and pick a shortest path $x = v_t, v_{t-1}, \dots, v_1, v_0 = y$, so $t = d(x, y) \geq 1$. Let j be the smallest index ≥ 1 with $v_j \in I$; such a j exists because $v_t = x \in I$. By minimality of $d(x, y)$ among tokens, $j = d(v_j, y) \geq t$; because v_j lies on the chosen shortest path, $j \leq t$. Hence $j = t$ and $v_j = x$. Minimality of j gives $v_{j-1} \notin I$, so sliding the token $x = v_t$ to v_{t-1} is legal; call the result $I_1 = I - \{v_t\} + \{v_{t-1}\}$. Take an optimal bijection σ for I and set $\sigma_1(v_{t-1}) = y$, $\sigma_1(z) = \sigma(z)$ for $z \neq v_t$; this is a bijection $I_1 \rightarrow I'$ of cost $\Phi(I) - d(v_t, y) + d(v_{t-1}, y) = \Phi(I) - 1$, so $\Phi(I_1) \leq \Phi(I) - 1$. Iterating, a strictly decreasing nonnegative integer terminates only at $\Phi = 0$, i.e. at I' . Every slide preserves the occupancy vector, so for general G the classes refine the level sets; the same argument applied inside each component shows the refinement is an equality, and (7) counts $|L_c|$. $\square$

The potential Φ is a minimum-matching cost, and the descent moves one token one step along an optimal matching, strictly decreasing the total matching length. The same statement is also verified combinatorially, with no linear algebra, by breadth-first search on the sliding graph of every certified graph, in the number of classes and in the size of every class (certificate C7, Table 3).

Theorem 5 (Structure and dimension in every sector). *Let G be any simple graph on M sites with connected components $V_1, \dots, V_r$ of sizes $b_1, \dots, b_r$, and let every $w_{ij} \neq 0$. For every $p = 0, \dots, M$, in the block $(p, M - p)$:*

$$\begin{aligned} Z_p &= \text{span}\{\mathbf{1}_c : c \text{ admissible}, \sum_i c_i = p\}, \\ \dim Z_p &= \#\left\{c \in \mathbb{Z}^r : 0 \leq c_i \leq b_i, \sum_i c_i = p\right\} =: w_{p-M/2}. \end{aligned} \quad (13)$$

The whole zero-mode space has

$$\dim Z = \sum_{p=0}^M \dim Z_p = \prod_{i=1}^r (b_i + 1), \quad (14)$$

and at half filling, $M = 2n$ and $p = n$,

$$\mathcal{V}_\infty = Z_n = \ker(T|_{\ker(D-0)}), \quad \dim \mathcal{V}_\infty = \#\left\{c : 0 \leq c_i \leq b_i, \sum_i c_i = n\right\}, \quad (15)$$

with basis $\{\mathbf{1}_c\}$. For connected G this is $\dim \mathcal{V}_\infty = 1$, spanned by the all-plus-one vector $\sum_{|I|=n} |I\rangle$.

Proof. By Lemmas 1–3, $v \in \ker D \cap \mathcal{F}_M$ satisfies $Tv = 0$ if and only if its coefficient function is constant on every sliding class, and by Lemma 4 the classes are the L_c with $\sum_i c_i = p$. The solution space of (11) is therefore spanned by the indicators $\mathbf{1}_c$, which are linearly independent because their supports are disjoint. Lemma 1 identifies this solution space with Z_p and with the half-filled chain limit. Summing $\dim Z_p$ over p evaluates $\prod_i (1 + x + \dots + x^{b_i})$ at $x = 1$, which is (14). For $r = 1$ the box constraints $0 \leq c_1 \leq M$, $c_1 = p$ leave a single admissible vector in every sector. $\square$

Two consistency identities accompany the theorem and are verified exactly on every certified graph. First, the level sets partition the block: $\sum_c |L_c| = \sum_c \prod_i \binom{b_i}{c_i} = \binom{M}{p}$ (Vandermonde), with half-filled per-graph totals 20, 70 and 252 at $M = 6, 8$ and 10, connected or not. Second, when $E = \emptyset$ then $T = 0$ and every no-doublon state is a zero mode: (13) gives $\dim Z_p = \binom{M}{p}$ for the empty graph, as it must. A caution that the data make sharp: the supports of the $\mathbf{1}_c$ partition the 2^M no-doublon configurations— $\sum_c |L_c| = 2^M$ summed over all sectors—but $\dim Z = \prod_i (b_i + 1)$ equals 2^M only for the empty graph; on the certified families the identity $\dim Z = 2^M$ fails on 70 of the 74 graphs. Support size and dimension are different quantities and are never interchanged below.

Table 1 records the dimension (15) for every component-size pattern occurring in the certified families, and the seed example of the paper, the path P_6 on six sites, already shows the mechanism: connected, so $\dim \mathcal{V}_\infty = 1$; cut the middle bond and the components are $(3, 3)$, so $\dim \mathcal{V}_\infty = \#\{c_1 + c_2 = 3 : 0 \leq c_i \leq 3\} = 4$; cut every bond and $\dim \mathcal{V}_\infty = \binom{6}{3} = 20$.

4 The spin construction and the decomposition of the zero-mode space

The level-set basis has a second, representation-theoretic description, and the two descriptions coincide exactly. Let $S_i^- = \sum_{k \in V_i} c_{k\downarrow}^\dagger c_{k\uparrow}$ be the spin-lowering operator of component i , let $S^\pm = \sum_i S_i^\pm$, $S^z = \frac{1}{2} \sum_{k \in V} (n_{k\uparrow} - n_{k\downarrow})$ and $S^2 = S^- S^+ + S^z(S^z + 1)$, and let $|\uparrow\rangle$ denote the all-up state (every site carries its up electron). Denote by $\mathcal{H}_j$ the spin- j irreducible representation of $\text{SU}(2)$, and write

$$j_i = \frac{b_i}{2}, \quad m_i = c_i - j_i, \quad m = p - \frac{M}{2} = \sum_i m_i, \quad \psi_c = \prod_{i=1}^r \frac{(S_i^-)^{b_i - c_i}}{(b_i - c_i)!} |\uparrow\rangle. \quad (16)$$

Table 1: Half-filled zero-mode dimension $\dim \mathcal{V}_\infty$ by connected-component sizes, for every component-size pattern occurring in the certified families of Section 9 (74 graphs). Patterns list the component sizes in non-decreasing order, (1^k) meaning k isolated sites. At $M = 10$ the value is the count formula of Theorem 5, confirmed by the sliding-class certificate C7; the exact nullity was not computed at this size (Section 9).

M	component sizes	$\dim \mathcal{V}_\infty$	graphs	M	component sizes	$\dim \mathcal{V}_\infty$	graphs
6	(6)	1	22	8	(8)	1	5
6	(1, 5)	2	6	8	(2, 6)	3	1
6	(2, 4)	3	4	8	(3, 5)	4	1
6	(3, 3)	4	2	8	(4, 4)	5	1
6	(1, 1, 4)	4	3	8	(2 ⁴)	19	1
6	(1, 2, 3)	6	1	8	(1 ⁸)	70	1
6	(2, 2, 2)	7	1	10	(10)	1	4
6	(1, 1, 1, 3)	8	1	10	(4, 6)	5	2
6	(1, 1, 2, 2)	10	1	10	(5, 5)	6	1
6	(1, 1, 1, 1, 2)	14	1	10	(2 ⁵)	51	1
6	(1 ⁶)	20	1				

Theorem 6 (SU(2) product form and decomposition of Z). *For every graph and every nonzero amplitudes:*

(i) **Product form.** *For every admissible c , $\psi_c = \mathbf{1}_c$, coefficient by coefficient in the convention (2). In particular $T\psi_c = 0$: the zero-mode property of the basis follows from $[T, S^-] = 0$ and the Pauli principle alone.*

(ii) **Intertwiner.** *The linear map*

$$\varphi : \bigotimes_{i=1}^r \mathcal{H}_{j_i} \longrightarrow Z, \quad \varphi \left(\bigotimes_i |j_i, m_i\rangle \right) = \frac{\mathbf{1}_c}{\sqrt{|L_c|}}, \quad (17)$$

is well defined, isometric, bijective and intertwines the SU(2) action; hence, as SU(2) representations, with Z_p the $S^z = m$ weight space,

$$Z \cong \bigotimes_{i=1}^r \mathcal{H}_{b_i/2}. \quad (18)$$

(iii) **Spin multiplicities.** *The multiplicity of total spin S in Z is*

$$n_S = w_S - w_{S+1}, \quad w_m = \dim Z_{M/2+m} = [x^{m+M/2}] \prod_{i=1}^r (1 + x + \cdots + x^{b_i}), \quad (19)$$

and the multiplicity of the eigenvalue $S(S+1)$ of $S^2|_{Z_p}$ is n_S for $S \geq |m|$ and 0 otherwise. Equivalently, by inclusion-exclusion on the box constraints, with $p = m + M/2$ and $\binom{a}{b} = 0$ whenever $a < b$ or $a < 0$,

$$w_m = \sum_{A \subseteq \{1, \dots, r\}} (-1)^{|A|} \binom{p - \sum_{i \in A} (b_i + 1) + r - 1}{r - 1}. \quad (20)$$

(iv) **Closed-form integer matrix for S^2 .** *In the basis $\{\mathbf{1}_c\}$ of Z_p , with rows indexed by c' and columns by c ,*

$$(S^2)_{c',c} = \delta_{c',c} \left[\sum_i (c_i + 1)(b_i - c_i) + m(m+1) \right] + \sum_{i \neq k} \delta_{c', c - e_i + e_k} (c_k + 1)(b_i - c_i + 1), \quad (21)$$

where e_i is the i -th unit vector and the (i, k) term is present only when $c_i \geq 1$ and $c_k \leq b_k - 1$. Every entry is a nonnegative integer and none depends on the amplitudes w_{ij} . In the orthonormal basis $\varphi(\bigotimes_i |j_i, m_i\rangle)$ the same operator is symmetric, hence diagonalizable, and geometric and algebraic multiplicities agree.

(v) **Ladder actions.**

$$S^+ \mathbf{1}_c = \sum_i (c_i + 1) \mathbf{1}_{c+e_i}, \quad S^- \mathbf{1}_c = \sum_i (b_i - c_i + 1) \mathbf{1}_{c-e_i}, \quad (22)$$

with terms dropped when the target is not admissible.

(vi) **SO(4) content of the space.** Every vector of Z is an exact pseudospin singlet (Theorem 8), so

$$Z \cong \bigoplus_S n_S \cdot \{0, S\} \quad (23)$$

as an $\text{SO}(4) \simeq (\text{SU}(2)_\eta \times \text{SU}(2)_S)/\mathbb{Z}_2$ representation: n_S copies of the sector of vanishing pseudospin and spin S , and nothing else.

(vii) **Definite-spin bases.** For every graph, connected or not, every sector p and every nonzero amplitudes, Z_p admits a basis of definite-total-spin interaction-independent zero modes, obtained either by Clebsch–Gordan coupling the j_i in any fixed order or by the exact rational idempotents of the matrix (21).

Proof. (i) Write the site flip $f_k = c_{k\downarrow}^\dagger c_{k\uparrow}$, so $S_i^- = \sum_{k \in V_i} f_k$. Distinct flips commute, $[f_k, f_l] = 0$ for $k \neq l$ (each f is even in the fermion operators), and $f_k^2 = 0$ by the Pauli principle; hence $(S_i^-)^n = n! \sum_{F \subseteq V_i, |F|=n} f_F$ with $f_F = \prod_{k \in F} f_k$, and

$$\psi_c = \prod_i \left(\sum_{F_i \subseteq V_i, |F_i|=b_i-c_i} f_{F_i} \right) |\uparrow\rangle = \sum_{F_1, \dots, F_r} f_{F_1} \cdots f_{F_r} |\uparrow\rangle. \quad (24)$$

Each flip acts on a single site, and a direct evaluation of (2) gives $f_k |S\rangle = +|S - \{2k\} + \{2k+1\}\rangle$ whenever site k is up: the annihilation at mode $2k$ carries the sign $(-1)^k$ (the k sites below contribute one occupied mode each), and the creation at mode $2k+1$ carries $(-1)^k$ again, since flips already performed at other sites do not change the number of occupied modes below $2k+1$. Every term therefore enters with coefficient $+1$, and the surviving up-set is $V \setminus (F_1 \cup \dots \cup F_r)$, which runs over L_c exactly once: $\psi_c = \mathbf{1}_c$. For the second statement, $T = \sum_i T_i$ with T_i acting only on the modes of V_i , and each T_i is a spin scalar, so $[T_i, S_j^-] = 0$ for all i, j ; the Pauli principle gives $T_i |\uparrow\rangle = 0$ (every up mode is occupied, so each hopping term annihilates), hence $T \psi_c = 0$.

(ii) The vectors $\mathbf{1}_c$ have disjoint supports and unit coefficients, so $\langle \mathbf{1}_c, \mathbf{1}_{c'} \rangle = |L_c| \delta_{c,c'}$ and φ maps the orthonormal product weight basis to an orthonormal basis of Z : it is isometric, and bijective because both sides have dimension $\prod_i (b_i + 1)$ by (14). For the intertwining property it suffices to compare S^+ on weight vectors, S^- being its adjoint and S^z being diagonal on both sides. From the product form (16) and the $\text{SU}(2)$ relation $S^+(S^-)^n |j, j\rangle = n(2j - n + 1)(S^-)^{n-1} |j, j\rangle$ applied with $n = b_i - c_i$ and $2j = b_i$,

$$S_i^+ \psi_c = \frac{(b_i - c_i)(c_i + 1)}{b_i - c_i} \psi_{c+e_i} = (c_i + 1) \psi_{c+e_i}, \quad (25)$$

which with (i) is (22). On the tensor side, $S^+ \bigotimes_i |j_i, m_i\rangle = \sum_k A_k \bigotimes_i |j_i, m_i + \delta_{ik}\rangle$ with $A_k = \sqrt{(j_k - m_k)(j_k + m_k + 1)} = \sqrt{(b_k - c_k)(c_k + 1)}$, and

$$A_i \sqrt{\frac{|L_c|}{|L_{c+e_i}|}} = \sqrt{(b_i - c_i)(c_i + 1)} \sqrt{\frac{\binom{b_i}{c_i}}{\binom{b_i}{c_i+1}}} = \sqrt{(b_i - c_i)(c_i + 1)} \sqrt{\frac{c_i + 1}{b_i - c_i}} = c_i + 1, \quad (26)$$

so φ carries the tensor-side action of S^+ to (22): the two sides agree through φ .

(iii) Decomposing $\bigotimes_i \mathcal{H}_{j_i} = \bigoplus_S n_S \mathcal{H}_S$ gives $w_m = \sum_{S \geq |m|} n_S$; the sequence w_m is non-increasing in $|m|$ and vanishes for $|m| > M/2$, so $n_S = w_S - w_{S+1}$, and the $S^z = m$ weight space of $\mathcal{H}_S$ is one-dimensional for $S \geq |m|$ and zero otherwise. The generating-function form of w_m counts the box constraints directly, and the inclusion-exclusion form (20) removes them.

(iv) Using $S^2 = S^- S^+ + S^z(S^z + 1)$, $S^\pm = \sum_i S_i^\pm$ and (22),

$$S^- S^+ \mathbf{1}_c = \sum_k (c_k + 1) S^- \mathbf{1}_{c+e_k} = \sum_k (c_k + 1) \sum_i (b_i - (c + e_k)_i + 1) \mathbf{1}_{c+e_k-e_i}. \quad (27)$$

The diagonal terms $i = k$ contribute $(c_k + 1)(b_k - c_k) \mathbf{1}_c$; the off-diagonal terms $i \neq k$ contribute $(c_k + 1)(b_i - c_i + 1) \mathbf{1}_{c-e_i+e_k}$ and are present exactly when $S_k^+ \mathbf{1}_c \neq 0$ (i.e. $c_k \leq b_k - 1$) and $S_i^- \mathbf{1}_{c+e_k} \neq 0$ (i.e. $c_i \geq 1$). The term $S^z(S^z + 1)$ contributes the scalar $m(m + 1)$. Self-adjointness of S^2 makes the matrix symmetric in the orthonormal basis of (ii).

(v) Proved within (ii) and (iv).

(vi) Proved with Theorem 8(i) below: $\eta^\pm$ annihilate every $\mathbf{1}_c$ and η^z vanishes at one electron per site, so the pseudospin of every vector of Z is $j = 0$, and (18) refines to the sector decomposition with multiplicities n_S .

(vii) By (iv), $S^2|_{Z_p}$ is a diagonalizable integer matrix; its eigenvectors lie in Z_p and are therefore zero modes of definite total spin. The rational idempotent $\prod_{S' \neq S} (X - S'(S' + 1)) / (S(S + 1) - S'(S' + 1))$ of (21) projects onto the spin- S eigenspace exactly over the rationals. $\square$

For a connected graph, (18) is the single multiplet $\mathcal{H}_{M/2}$ and ψ_c is the maximal-spin descendant $(S^-)^{M/2-c_1} |\uparrow\rangle / (M/2 - c_1)!$; at half filling it is $(S^-)^n |\uparrow\rangle / n!$, the maximal-spin tower state of the $\text{SO}(4)$ sector $\{0, M/2\}$ [10, 11]. For a disconnected graph (18) is a genuine Clebsch–Gordan problem, and Sections 5–7 decide exactly which vectors are pure. Two independent constructions of the same basis—the level-set indicator and the $\text{SU}(2)$ product—are compared coefficient by coefficient on all 334 certified half-filled states, with the recorded sign set $\{+1\}$ and nothing else (certificate C3).

5 Definite-spin purity and the exact $\text{SO}(4)$ labels

The matrix (21) answers exactly which basis vectors carry a definite total spin—the question a classification of interaction-independent eigenstates by $\text{SO}(4)$ quantum numbers has to settle before it can be applied to a disconnected hopping graph. The answer is sharp: purity occurs only where the sector is one-dimensional.

Theorem 7 (Purity criterion). *Let $r \geq 1$ and let $\mathbf{1}_c \in Z_p$ be a level-set basis vector of any graph with any nonzero amplitudes. Then $\mathbf{1}_c$ is a total-spin eigenvector if and only if*

$$r = 1 \quad \text{or} \quad p = 0 \quad \text{or} \quad p = M, \quad (28)$$

and in every definite case the spin is maximal,

$$S^2 \mathbf{1}_c = \frac{M}{2} \left(\frac{M}{2} + 1 \right) \mathbf{1}_c, \quad S = \frac{M}{2}. \quad (29)$$

Equivalently, $\mathbf{1}_c$ is definite-spin if and only if $\dim Z_p = 1$. In particular: on a connected graph every zero mode of every sector is a maximal-spin descendant; on every disconnected graph ($r \geq 2$) no level-set basis state of any interior sector $0 < p < M$ —including the whole half-filled block for $n \geq 1$ —is a total-spin eigenstate; and the only definite-spin level-set states of a disconnected graph are the two fully polarized ones, all-up ($p = M$) and all-down ($p = 0$).

Proof. By (21), the vector $S^2 \mathbf{1}_c - [\sum_i (c_i + 1)(b_i - c_i) + m(m + 1)] \mathbf{1}_c$ is a combination of the $\mathbf{1}_{c-e_i+e_k}$, $i \neq k$, in which the term (i, k) is present exactly when $c_i \geq 1$ and $c_k \leq b_k - 1$, with coefficient $(c_k + 1)(b_i - c_i + 1)$, a positive integer. Distinct level sets have disjoint supports, so the $\mathbf{1}_{c'}$ are linearly independent and no cancellation is possible: $\mathbf{1}_c$ is an eigenvector if and only if

$$\text{for all } i \neq k: \quad c_i = 0 \quad \text{or} \quad c_k = b_k. \quad (*)$$

Physically, the off-diagonal terms are the flip-flop moves $S_i^- S_k^+$: purity fails exactly when one component can still lower ($c_i \geq 1$) while another can still raise ($c_k \leq b_k - 1$).

If $r = 1$ there are no pairs and $(*)$ holds vacuously; the diagonal entry evaluates to

$$(p + 1)(M - p) + \left(p - \frac{M}{2}\right) \left(p - \frac{M}{2} + 1\right) = \frac{M}{2} \left(\frac{M}{2} + 1\right), \quad (30)$$

independently of p , and $\mathbf{1}_c$ is the maximal-spin descendant $(S^-)^{M/2-p} |\uparrow\rangle / (M/2 - p)!$ of Theorem 6. If $p = 0$ then $c = 0$ and if $p = M$ then $c_i = b_i$ for all i ; in both cases $(*)$ holds, $|m| = M/2$ forces $S = M/2$, and the diagonal entry is $M + (M^2/4 - M/2)$ respectively $(M/2)(M/2 + 1)$, i.e. $(M/2)(M/2 + 1)$ in both cases.

Conversely, let $r \geq 2$, $0 < p < M$, and suppose $(*)$ holds. The sets $A = \{i : c_i > 0\}$ and $B = \{i : c_i < b_i\}$ are both nonempty. If some $q \in A$ and some $l \in B$ differ, applying $(*)$ to the ordered pair $(i, k) = (q, l)$ demands $c_q = 0$ or $c_l = b_l$, both false; hence every element of A equals every element of B , so $A = B = \{q\}$ with $0 < c_q < b_q$. For any $i \neq q$: $i \notin A$ gives $c_i = 0$ and $i \notin B$ gives $c_i = b_i$, hence $b_i = 0$ —impossible, since every component is nonempty and $r \geq 2$ supplies such an i . So $(*)$ fails.

For the equivalence with $\dim Z_p = 1$: if $r = 1$ or $p \in \{0, M\}$ the admissible set is a single vector, so $\dim Z_p = 1$ by Theorem 5; and if $r \geq 2$ with $0 < p < M$ then $\dim Z_p \geq 2$, because with A, B as above either some $i \in A$ and $k \in B$ differ, making $c - e_i + e_k$ a second admissible vector, or $A = B = \{q\}$ with $0 < c_q < b_q$, and for any $k \neq q$ necessarily $c_k = b_k > 0$, making $c + e_q - e_k$ a second admissible vector. $\square$

The criterion is certified exhaustively two independent ways on every graph, sector and admissible c (certificate C14, Figure 1): as the vanishing of the off-diagonal entries of (21), and fermionically, by requiring $S^2 \mathbf{1}_c$ to be supported on L_c alone. Of the 1544 level-set basis states of the certified families—all $M + 1$ sectors of all 74 graphs—exactly 348 are pure, and they are exactly the predicted set: all 276 states of the 38 connected graphs, plus precisely the two fully polarized states of each of the 36 disconnected graphs. In the half-filled block alone, none of the 296 states of a disconnected graph is a spin eigenstate (certificate C9).

Theorem 8 (Exact pseudospin singlets and the $\text{SO}(4)$ sector basis). *For every finite graph, every nonzero amplitudes and every sector p :*

- (i) **Exact pseudospin singlets.** *For the pair operators $\eta_\Phi^\pm$ of any phase assignment $\Phi = (\Phi_i)_{i \in V}$, $\eta_\Phi^+ = \sum_i e^{i\Phi_i} c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger$, $\eta_\Phi^- = (\eta_\Phi^+)^\dagger$, $\eta_\Phi^z = \frac{1}{2} \sum_i (n_i - 1)$, every vector v of Z satisfies*

$$\eta_\Phi^+ v = 0, \quad \eta_\Phi^- v = 0, \quad \eta_\Phi^z v = 0, \quad \text{hence} \quad \eta_\Phi^2 v = 0, \quad (31)$$

so every interaction-independent zero mode is an exact pseudospin singlet, $j = 0$, for every phase convention at once, on bipartite and non-bipartite graphs alike.

- (ii) **Sector basis.** *Each Z_p admits a basis of exact $\text{SO}(4)$ -sector zero modes: coupling the r spin- j_i representations of Theorem 6 to total spin S by Clebsch–Gordan coefficients (or, equivalently and exactly over the rationals, by the idempotents of (21)) yields vectors with definite labels $\{0, S\}$, $S \geq |m|$, and the multiplicity of S in Z_p is n_S of (19). As an $\text{SO}(4)$ representation,*

$$Z \cong \bigoplus_S n_S \cdot \{0, S\}, \quad (32)$$

and no interaction-independent zero mode of the no-doublon space has $j \neq 0$.

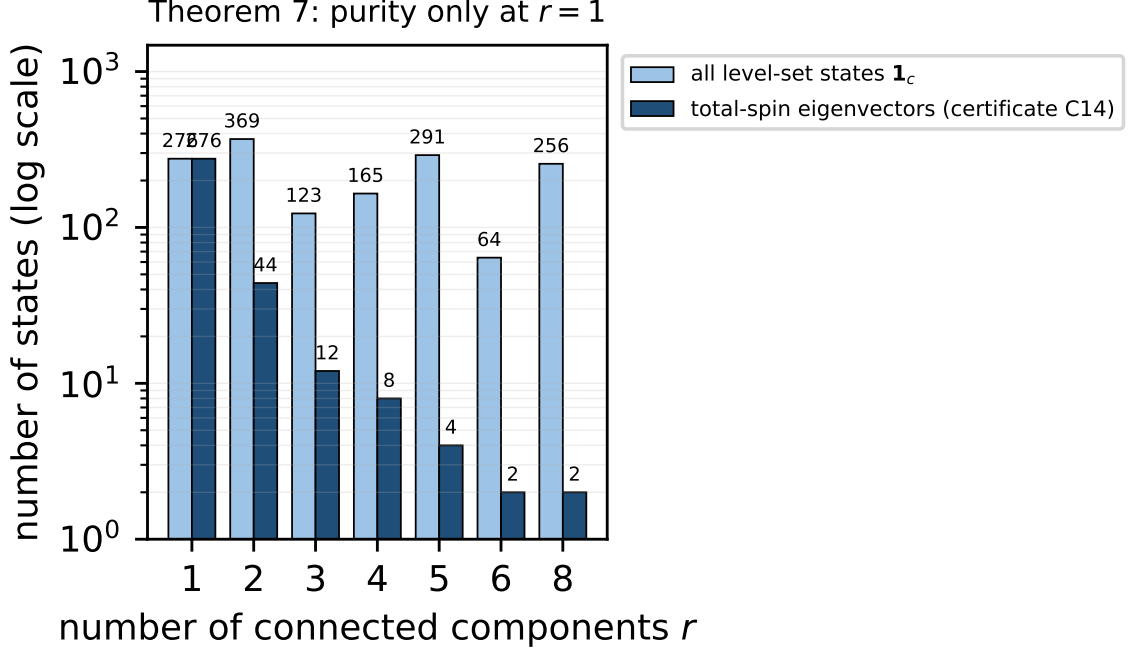


Figure 1: Exact purity census of the level-set basis (certificate C14; the figure title refers to the same theorem number as Theorem 7 here). Bars, by number r of connected components of the certified graphs: all level-set basis states $\mathbf{1}_c$ of all sectors—276, 369, 123, 165, 291, 64, 256 for $r = 1, 2, 3, 4, 5, 6, 8$, in total 1544—against the total-spin eigenvectors among them—276, 44, 12, 8, 4, 2, 2, in total 348. The pure states are exactly those predicted by Theorem 7: every state of every connected graph, and exactly two states ($p = 0$ and $p = M$) per disconnected graph. Purity occurs only at $r = 1$ or at the extremal fillings.

Proof. (i) No state of $\ker D \cap \mathcal{F}_M$ has an empty site, so every term $c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger |I\rangle$ of η_Φ^+ vanishes by the Pauli principle, whatever the phases; no such state has a doublon, so every term of η_Φ^- vanishes likewise; and $n_i = 1$ on the whole space gives $\eta_\Phi^z = 0$ identically. The statement uses single occupancy only—never bipartiteness, never connectedness, never the amplitudes.

(ii) By (i) the pseudospin label of every vector of Z is $j = 0$, so the $\text{SO}(4)$ content of Z is carried entirely by the $\text{SU}(2)$ decomposition (18): the coupled basis of Theorem 6(vii) consists of simultaneous eigenvectors of η^2 (eigenvalue 0) and S^2 (eigenvalue $S(S+1)$), one for each (S, M_S, ν) with multiplicities n_S , and each lies in Z , hence is an interaction-independent zero mode by Lemma 1. $\square$

Two boundary remarks delimit the pseudospin statement. First, (i) is an *annihilation* statement, not a conservation law: that η generates a symmetry of H —the $\text{SO}(4)$ of Yang and Zhang [10, 11, 1]—requires bipartiteness, while $j = 0$ on Z requires only one electron per site and holds on every certified graph, including K_{10} and $2 \times C_5$. Second, the statement is confined to $d = 0$: in the one-doublon subspace, where T maps $\ker D$, an empty site exists, η^+ no longer annihilates, and Lemma 1 fails; whether η -nontrivial interaction-independent zero modes exist there is open (Section 10).

6 Slater rank, entanglement spectrum and entropy

Every state of $\ker D \cap \mathcal{F}_M$ factorizes across the species split into a definite-up part and a definite-down part: with the interleaved convention (2), $|I\rangle = (-1)^{\sigma_I} |I\rangle_\uparrow \otimes |I^c\rangle_\downarrow$, where $|J\rangle_\sigma$ is the Slater

determinant of the spin- σ modes of J in increasing order and σ_I is the parity of the permutation that reorders the mode list. A state $v = \sum_I \psi_I |I\rangle$ therefore has an up/down coefficient array $A[J, K]$, rows indexed by the p -subsets J (the up configuration) and columns by the $(M - p)$ -subsets K (the down configuration), whose entry is $\pm \psi_{J \cup K^c}$ when $K = J^c$ and zero otherwise: *the support of the array is complement-matched*. Its Schmidt form across the species split has singular values $\{|\psi_I| : \psi_I \neq 0\}$ with Schmidt terms the signed factorizations above, and the minimum number of Slater determinants of any representation of v equals the rank of A .

Theorem 9 (Exact Slater rank, flat spectrum, entropy). *Let G have component sizes $b_1, \dots, b_r$ and let c be admissible in the block $(p, M - p)$. Then:*

- (i) *for every $v \in \ker D \cap \mathcal{F}_M$, the minimum number of Slater determinants equals the support size, $N_{\text{Slater}}(v) = \text{rank } A = \#\{I : \psi_I \neq 0\}$;*
- (ii) *for $v = \mathbf{1}_c$ this number is exactly*

$$N_{\text{Slater}}(\mathbf{1}_c) = |L_c| = \prod_{i=1}^r \binom{b_i}{c_i} \quad (33)$$

(an exact value, not a bound); the up/down Schmidt spectrum of $\mathbf{1}_c$ is flat— $|L_c|$ equal singular values, each $1/\sqrt{|L_c|}$ after normalization—and the von Neumann entropy of either reduced state is

$$S_{\text{vN}}(\mathbf{1}_c) = \log |L_c| = \sum_{i=1}^r \log \binom{b_i}{c_i} \text{ nats}, \quad (34)$$

the maximum for any state supported on L_c ;

- (iii) *for connected G at half filling, $N_{\text{Slater}} = \binom{M}{n}$ and $S_{\text{vN}} = \log \binom{M}{n}$, both maximal in the whole block;*
- (iv) *every definite-spin zero mode of Theorem 8(ii) satisfies $N_{\text{Slater}} =$ its support size; among the 699 certified definite-spin eigenvectors at $M \leq 6$ (per amplitude set), 665 have full support and hence maximal Slater number $\binom{M}{p}$.*

Proof. (i) The complement-matched array A is a generalized diagonal matrix—its nonzero entries sit at positions (J, J^c) , one per occupied support configuration—so its rank is its number of nonzeros, which equals the support size of ψ ; the minimum number of product terms of a bipartite vector is its Schmidt rank, which is $\text{rank } A$. The complement-matching constraint is also what removes any Plücker-relation freedom: no decision procedure for decomposability is needed at any M . (ii) For $\mathbf{1}_c$ every ψ_I with $I \in L_c$ equals $+1$ (Lemma 3) and vanishes off L_c , so the singular values are $\{1\}^{|L_c|}$: flat, of multiplicity $|L_c| = \prod_i \binom{b_i}{c_i}$, and the normalized reduced state is maximally mixed on its support, with entropy $\log |L_c|$, maximal for the given support. (iii) For $r = 1$ the only half-filled level set is $c = (n)$ with $|L_c| = \binom{M}{n}$, the dimension of the block. (iv) Immediate from (i); the counts are certificate C20. $\square$

Corollary 10 (Slater rank equals support size for definite-spin zero modes). *For every definite-spin interaction-independent zero mode ψ at one electron per site, $N_{\text{Slater}}(\psi)$ equals its number of nonzero coefficients; for a level-set state $\mathbf{1}_c$ the common value is $\prod_i \binom{b_i}{c_i}$, and where $n_S \geq 2$ the value depends on the eigenvector chosen inside the spin- S eigenspace (Section 10).*

Proof. Immediate from Theorem 9(i)–(ii): the complement-matched array has rank equal to the support size, and for $\mathbf{1}_c$ the support is L_c . Where $n_S \geq 2$ different eigenvectors of the same eigenspace can have different supports: certificate C20 records 665 full-support vectors and 34 others at $M \leq 6$, so both behaviors occur. $\square$

The certified extremes frame the scale: the largest entropies recorded over the whole family are $\log 20 = 2.995732273553991$ at $M = 6$, $\log 70 = 4.248495242049359$ at $M = 8$ and $\log 252 = 5.529429087511423$ at $M = 10$, each attained by the unique zero mode of a connected graph, whose support is the entire half-filled block. Dilution moves these numbers in opposite directions: on the empty graph every level-set state is a single Slater determinant ($|L_c| = 1$, entropy 0), while the definite-spin combinations of them are generically full-support; on a dimer cover the maximal-support level set has $|L_c| = 2^{M/2}$ (all $c_i = 1$). Slater rank here is therefore not a small-number invariant but an extensive one, and the flatness of the spectrum—maximal Schmidt multiplicity at fixed support—is what separates the level-set states from generic fermionic states of the same support size. Every rank statement of certificate C5 is verified exactly: nonzero count = rank = $\prod_i \binom{b_i}{c_i}$ on all 334 half-filled states, all coefficients unit, rows and columns distinct, Gram matrix exactly the identity, cross-checked by dense exact row reduction and by forming AA^T wherever the array fits in 128×128 , and by the signed-permutation structure alone at $M = 10$ (arrays up to 252×252), as the records state.

7 Invariances and the pair condition

7.1 The interaction is irrelevant, and so are its details

Corollary 11 (Arbitrary density-density interactions). *Replace WD in (1) by any density-density interaction $V = \sum_{i \leq j} V_{ij} n_i n_j$ with arbitrary real coefficients, including inhomogeneous on-site terms $U_i = V_{ii}$ and interactions of any range. Every $\mathbf{1}_c$, and every definite-spin combination of Theorem 8(ii), is an exact eigenstate of $T + V$ with eigenvector independent of V and interaction eigenvalue the c -number $\sum_{i \leq j} V_{ij}$. All statements of Sections 3–6 hold verbatim for $T + V$ in place of H .*

Proof. On $\ker D \cap \mathcal{F}_M$ every site is singly occupied, $n_i = 1$, so V acts as the scalar $\sum_{i \leq j} V_{ij}$ on the whole space, and $T + V$ has the eigenvectors of T with energies shifted by that scalar. Certificate C8 confirms it with random inhomogeneous V_{ij} on every certified graph: every support state is singly occupied on every site and the interaction eigenvalue takes exactly one value. $\square$

The zero-mode structure is thus not a Hubbard- U artifact: the whole $\prod_i (b_i + 1)$ -dimensional space is interaction-independent for *any* density interaction, which merely shifts the common energy. What (1) contributes is the grading by D that isolates the no-doublon space in the first place.

7.2 The pair condition: the criterion in abstract spin coordinates

Theorem 7 is phrased in occupancy coordinates tied to the Hubbard realization, where every component is nonempty and $j_i = b_i/2 \geq \frac{1}{2}$. The underlying statement is a fact about arbitrary products of $SU(2)$ representations, complete with the zero-spin factors that no graph can produce; stating it separately makes the criterion quotable outside this setting and exhibits $(*)$ as one clause of a two-sided condition.

Proposition 12 (Spin-coordinate pair condition). *Let $r \geq 1$, let $j = (j_1, \dots, j_r)$ with each j_i a nonnegative integer or half-integer, zero allowed, and let $v = \bigotimes_i |j_i, m_i\rangle$ be a product weight vector of $\bigotimes_i \mathcal{H}_{j_i}$. Then v is an eigenvector of S^2 if and only if for every ordered pair (i, k) , $i \neq k$,*

$$(P) \quad (m_i = j_i \text{ or } m_k = -j_k) \quad \text{and} \quad (m_i = -j_i \text{ or } m_k = j_k), \quad (35)$$

and the eigenvalue is then

$$\lambda = \sum_i j_i(j_i + 1) + 2 \sum_{i < k} m_i m_k. \quad (36)$$

Proof. Expand $S^2 = \sum_i S_i^2 + 2 \sum_{i < k} [S_i^z S_k^z + \frac{1}{2}(S_i^+ S_k^- + S_i^- S_k^+)]$. Acting on v , the diagonal terms give λ , and the only off-diagonal components are the paired cross moves $(m_i, m_k) \rightarrow (m_i + 1, m_k - 1)$ with coefficient $A_i^+ A_k^-$ and $(m_i, m_k) \rightarrow (m_i - 1, m_k + 1)$ with coefficient $A_i^- A_k^+$, where $A_i^\pm = \sqrt{j_i(j_i + 1) - m_i(m_i \pm 1)}$. The product weight basis is orthogonal, so $S^2 v = \lambda v$ if and only if every cross-move coefficient vanishes. Now $m \mapsto m(m + 1)$ is strictly increasing on $m \geq -\frac{1}{2}$, so within $|m_i| \leq j_i$ the equation $j_i(j_i + 1) = m_i(m_i + 1)$ has the single solution $m_i = j_i$ (its other root, $-j_i - 1$, lies outside the weight system), and likewise $j_i(j_i + 1) = m_i(m_i - 1)$ only at $m_i = -j_i$. Hence $A_i^+ = 0$ if and only if $m_i = j_i$, and $A_i^- = 0$ if and only if $m_i = -j_i$, and the vanishing of $A_i^- A_k^+$ respectively $A_i^+ A_k^-$ is the second respectively first clause of (P). No radicals need normalizing, no bound on r enters, and no coupling scheme is chosen. $\square$

Corollary 13 (The aligned-only boundary, exactly). *In the setting of Proposition 12, delete every zero-spin factor of j (a factor with $j_i = 0$ has $m_i = 0 = \pm j_i$, so both clauses of (P) hold automatically for every pair containing it, and zero-spin factors drop out of the constraint entirely). Then:*

- (1) *if two or more nonzero factors remain, every definite-spin product weight vector is extremal in every factor, all signs agree, and the eigenvectors are exactly the two aligned vectors $m = +j$ and $m = -j$, with $S = \sum_i j_i$;*
- (2) *if exactly one nonzero factor j remains, (P) is vacuous and all $2j + 1$ weight vectors are eigenvectors: “aligned-only” fails if and only if $j \geq 1$;*
- (3) *if no nonzero factor remains, there is a single basis vector, aligned by both branches.*

So aligned-only fails exactly when some $j_i = 0$ and the nonzero factors reduce to a single $j \geq 1$, with $r = 1$ and no zero factor the degenerate case of clause (2).

Proof. With two or more nonzero factors, an interior weight $|m_i| < j_i$ would force, through (P) applied to the pair (i, k) , both $m_k = j_k$ and $m_k = -j_k$ for every other nonzero k —impossible since $j_k > 0$; so every nonzero factor is extremal, and two opposite signs violate (P) on that pair. The eigenvalue of an aligned vector is $\sum_i j_i(j_i + 1) + 2 \sum_{i < k} (\pm j_i)(\pm j_k) = (\sum_i j_i)(\sum_i j_i + 1)$ with matching signs. With exactly one nonzero factor there are no pairs to constrain, and S^2 is diagonal on each $\mathcal{H}_j$ weight vector. Clause (3) is $\mathcal{H}_0^{\otimes r}$. $\square$

Corollary 14 (Transfer to the Hubbard realization). *Under the dictionary $j_i = b_i/2$, $m_i = c_i - b_i/2$ of Theorem 6(ii), $A_i^- \neq 0$ if and only if $c_i > 0$ and $A_i^+ \neq 0$ if and only if $c_i < b_i$, so (P) becomes*

$$\text{for all } i \neq k: \quad (c_i = b_i \text{ or } c_k = 0) \quad \text{and} \quad (c_i = 0 \text{ or } c_k = b_k), \quad (37)$$

whose second clause is () and whose first clause is (*) with i and k exchanged. Since (*) is demanded of every ordered pair, the two are one statement: Theorem 7 is (P) restricted to $b_i \geq 1$. The zero-spin factors of Proposition 12 are formal here—a zero-spin factor would be an empty component—so the zero-spin branch of Corollary 13, namely its parenthetical clause and clause (3), is certified only in the abstract; a graph with $r = 1$ realizes the no-zero-factor case of clause (2), which is exactly the $r = 1$ branch of Theorem 7.*

Proof. By the proof of Proposition 12, $A_i^- = 0$ if and only if $m_i = -j_i$ and $A_i^+ = 0$ if and only if $m_i = j_i$; the dictionary turns these into $c_i = 0$ respectively $c_i = b_i$, and substituting them into (P) gives the displayed two-clause condition. Its second clause is (*) for the ordered pair (i, k) and its first clause is (*) for the ordered pair (k, i) ; since (*) ranges over all ordered pairs, the two conditions coincide. With $b_i \geq 1$ there is no zero-spin factor, so Corollary 13 reduces to its clause (1) for $r \geq 2$ —the two aligned vectors, i.e. $p = M$ and $p = 0$ —and to clause (2) for $r = 1$, in which every weight vector is an eigenvector: Theorem 7. $\square$

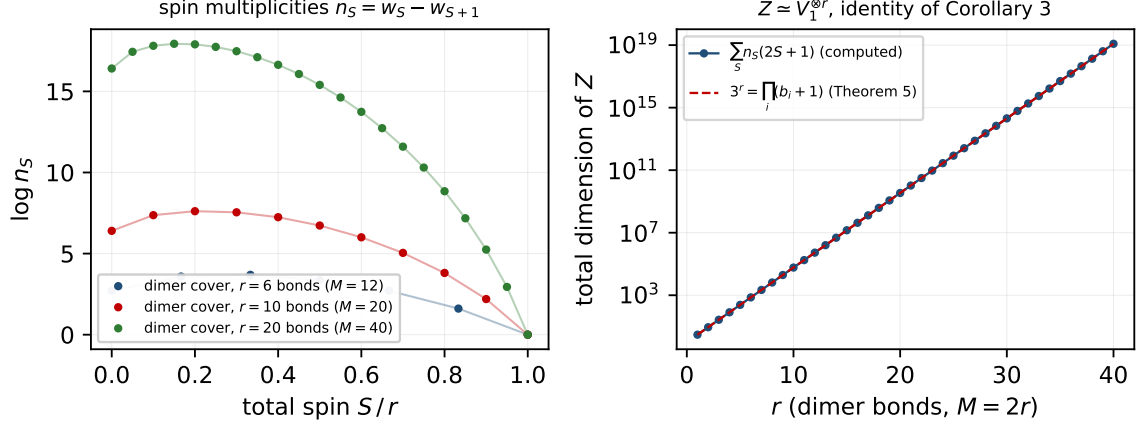


Figure 2: Exact spin content of the diluted patterns (the figure labels refer to the internal numbering of the research notes: their “Theorem 5” is Theorem 5 here, their “Corollary 3” is Corollary 16 below). **Left:** spin multiplicities $n_S = w_S - w_{S+1}$ of the perfect dimer cover with $r = 6, 10$ and 20 bonds ($M = 2r$), against S/r : the trinomial profile of (19) with $b_i = 2$, spreading over all $r + 1$ spins $S = 0, \dots, r$ of $Z \cong \mathcal{H}_1^{\otimes r}$. **Right:** the dimension identity $\sum_S n_S(2S + 1) = 3^r = \prod_i (b_i + 1)$ for the dimer cover, computed integer multiplicities against the closed form, for $r \leq 40$ (to $3^{40} \approx 1.2 \times 10^{19}$); the same identity is checked exactly to $r = 200$, where $\dim Z = 3^{200}$ is a 96-digit integer and every spin $S = 0, \dots, 200$ occurs (certificate C18).

The pair condition is certified by a third, symbolic engine along three routes (certificate C22): exact S^2 in the orthonormal product weight basis, with eigenvectorhood read off the emptiness of the off-diagonal support; exact iterated Clebsch–Gordan expansion, with purity read off a single total S carrying all the squared weight; and the radical-free dictionary $(P) \Leftrightarrow (*) \Leftrightarrow (r = 1 \text{ or } p \in \{0, M\})$ evaluated over the recorded dataset of Section 9. On 33 j -tuples and 270 weight vectors the three routes agree with 0 mismatches, the aligned vectors are eigenvectors in all 33 cases, the clause (1)–(3) prediction is correct in all 33, and the eigenvector counts where aligned-only fails are 3, 4, 6, 3, 3, 5 against aligned counts 2 for $j = (1), (\frac{3}{2}), (\frac{5}{2}), (0, 1), (0, 0, 1), (2)$; on 3088 (state, amplitude-set) instances in 1080 sectors there are 0 mismatches and both the basis dimension and the pure-state count reproduce in all 1080. Figure 2 displays the multiplicities (19) for the dimer cover—the diluted pattern with the richest spin content—together with the dimension identity $\sum_S n_S(2S + 1) = \prod_i (b_i + 1)$ that closes the check.

8 Robustness: zero amplitudes and bond dilution

Remark 15 (Zero amplitudes are edge deletions). The hypothesis $w_{ij} \neq 0$ enters at exactly one place: dividing the two-term equation of Lemma 2 by w_{ab} . Setting one amplitude to zero is therefore exactly deleting that edge, and every theorem of Sections 3–7 applies verbatim to the support graph $G' = (V, \{(i, j) : w_{ij} \neq 0\})$, with the component structure of G' and no other change. Two consequences are certified exhaustively at $M \leq 6$ (all 45 edge-bearing graphs of the exhaustive families, every one of their edges: 201 single deletions, certificate of the zero-amplitude probe): the dimension of every remaining graph equals its own count formula, and all claims C1–C9 hold on every remaining graph. Moreover, a deletion changes $\dim Z_n$ if and only if the deleted edge is a bridge:

$$\begin{array}{ll} \text{deleting a bridge raised } \dim Z_n : & 96 \text{ of } 96; \\ \text{deleting a non-bridge changed it:} & 0 \text{ of } 105. \end{array} \quad (38)$$

The mechanism is Theorem 5: deleting a cycle edge leaves the component sizes, hence the count, untouched; deleting a bridge splits a component b into $(s, b - s)$, and each old occupancy value c_b

is replaced by $\min(s, c_b) - \max(0, c_b - (b - s)) + 1 \geq 1$ new values, with strict increase whenever some admissible vector has $1 \leq c_b \leq b - 1$ —which at interior filling $0 < n < M$ always exists. The seed example: P_6 has $\dim Z_3 = 1$; deleting its middle bond gives components $(3, 3)$ and $\dim Z_3 = \#\{c_1 + c_2 = 3 : 0 \leq c_i \leq 3\} = 4$; deleting every bond gives $\dim Z_3 = \binom{6}{3} = 20$.

Corollary 16 (Extensive residual degeneracy under bond dilution). *Let G be a connected lattice on $M = 2n$ sites and G_t a bond-diluted version with component sizes $b_1(t), \dots, b_{r(t)}(t)$ and arbitrary nonzero amplitudes on the surviving bonds. Then the half-filled zero-mode space jumps from one dimension to*

$$\dim Z_n(G_t) = \#\left\{c : 0 \leq c_i \leq b_i(t), \sum_i c_i = n\right\}, \quad (39)$$

a count that depends only on the component sizes, never on which bonds survive, and grows weakly monotonically under further dilution (Figure 3). For the two patterns that bracket the growth, evaluated as exact integers up to $M = 400$:

- (a) **perfect dimer cover** ($r = M/2$ disjoint bonds, $b = (2^r)$): $Z \cong \mathcal{H}_1^{\otimes r}$, $\dim Z = 3^r$, $w_m = [x^{m+r}](1 + x + x^2)^r$,

$$\dim Z_n = T(r) = [x^r](1 + x + x^2)^r \quad (\text{the central trinomial coefficient}), \quad (40)$$

with $T(r+1)/T(r) \rightarrow 3$ and entropy density $\frac{1}{M} \log \dim Z_n \rightarrow \frac{1}{2} \log 3 \approx 0.549$ nats per site;

- (b) **empty graph** (M isolated sites, $b = (1^M)$): $T = 0$, every no-doublon state is a zero mode, $Z \cong (\mathbb{C}^2)^{\otimes M}$, $w_m = \binom{M}{m+M/2}$,

$$\dim Z_n = \binom{M}{n}, \quad (41)$$

with entropy density $\frac{1}{M} \log \dim Z_n \rightarrow \log 2 \approx 0.693$ nats per site and spin multiplicities $n_S = \binom{M}{M/2-S} - \binom{M}{M/2-S-1}$, the standard multiplicities of M uncoupled spin- $\frac{1}{2}$'s.

In both patterns, for $r \geq 2$, no half-filled level-set state is definite-spin (Theorem 7); the definite-spin zero modes are the Clebsch–Gordan combinations of Theorem 8(ii), with multiplicities $n_S = w_S - w_{S+1}$. These are degeneracies of the eigenstate set, not ground-state degeneracies: nothing here locates these energies within the spectrum.

Proof. The dimension is Theorem 5 applied to G_t ; weak monotonicity holds because every occupancy vector of a coarser component partition lifts to at least one vector of any refinement, and distinct coarse vectors have disjoint lift sets. For (a), $\prod_i (1 + x + x^2) = (1 + x + x^2)^r$ gives w_m and $\dim Z = 3^r$; the central coefficient is $T(r)$, whose consecutive ratios approach 3, recorded as exact fractions up to $r = 200$ (the last recorded ratio is $2.9925\dots$), giving the density $\frac{1}{2} \log 3$. For (b), $\prod_i (1 + x) = (1 + x)^M$ gives $w_m = \binom{M}{m+M/2}$ by Vandermonde and $\dim Z_n = \binom{M}{n}$, whose M -th root tends to 2; the multiplicities n_S are those of the M -fold spin- $\frac{1}{2}$ product, i.e. of $(\mathbb{C}^2)^{\otimes M}$. The spin decompositions are Theorem 6(iii) on the two component patterns. All counts are exact integers; only the reported densities are floating-point. At $M = 8$ and $M = 10$ the pattern values $T(4) = 19$, $T(5) = 51$ and $\binom{8}{4} = 70$ are the same integers the linear-algebra engine found for $4 \times K_2$, $5 \times K_2$ and the 8-site empty graph, so the combinatorial and fermionic routes agree wherever both can run. $\square$

9 Exact verification

Every structural claim of Sections 3–8 is a theorem for every graph, every size and every nonzero amplitudes. The computations below are independent checks, not evidence for the theorems: they

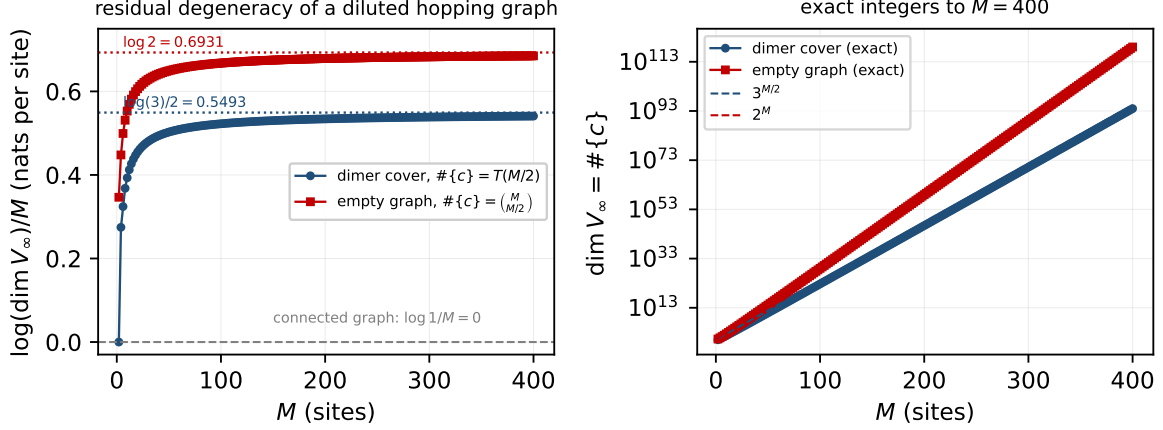


Figure 3: Residual degeneracy of a diluted hopping graph (Corollary 16; exact integer counts, only the plotted densities are floating-point). **Left:** entropy density $\log(\dim Z_n)/M$ against M for connected lattices ($\dim Z_n = 1$, density 0), the perfect dimer cover ($\dim Z_n = T(M/2)$: 0.36805 at $M = 8$, 0.45499 at $M = 20$, 0.54089 at $M = 400$, tending to $\log(3)/2 = 0.54931$) and the empty graph ($\dim Z_n = \binom{M}{M/2}$: 0.53106, 0.60634, 0.68509, tending to $\log 2 = 0.69315$). **Right:** the same counts as exact integers up to $M = 400$ against the all-sector references $3^{M/2}$ and 2^M : $T(200)$ is a 94-digit integer and $\binom{400}{200}$ a 120-digit one. Bond dilution converts the unique maximal-spin zero mode of a connected lattice into an extensive residual degeneracy of exact interaction-independent eigenstates.

are exact rational and integer arithmetic—no floating-point number enters any structural claim, and floats appear only in reported entropies and densities—produced by three engines that share nothing but the interleaved mode-ordering convention (2). The first is a descending-chain linear-algebra engine over exact rationals, which also builds the $SU(2)$ product states; the second is an all-sector spin engine with its own, separately written fermion primitives, compared against the first state by state (C19: 16484 states, 0 mismatches for T , S^+ and S^-); the third is a symbolic checker of the pair condition that builds S^2 in the orthonormal product weight basis rather than the fermionic level-set basis and reads the recorded dataset instead of recomputing it, so its test of Theorem 7 is independent of the second engine’s. Two further exact evaluators produce the large- M dilution integers and the zero-amplitude probe. All fermion eigenvector checks use the *unprojected* operators: certificate C17 records why this matters—the projected hopping operator PTP on $\ker D \cap \mathcal{F}_M$ vanishes identically at one electron per site, so a verification built on it would be passed by all 2^M no-doublon states against the true $\prod_i (b_i + 1)$, and at $N < M$ its kernel is still strictly larger than the invariant subspace (8 vs. 7, 16 vs. 14, 68 vs. 25 on the three recorded probes). The programs, their complete numerical outputs and the underlying proof notes are archived with the source of this work [2]. Tables 2, 3, 4 and 5 inventory the certified families and the certificates C1–C22; Table 6 records four small exact blocks.

The curated samples are listed by name in the records: at $M = 6$ the non-bipartite family is C_5 plus an isolated vertex, $K_3 + K_3$, $K_4 + K_2$, C_6 with one chord, C_6 with a diagonal (two odd cycles), a triangle with a three-site tree attached, $K_{3,3}$ plus one same-sublattice edge, and K_6 minus a perfect matching; at $M = 8$ the family is C_8 , P_8 , $K_{4,4}$, $K_{2,6}$, K_8 , $C_4 + C_4$, $C_6 + K_2$, $4 \times K_2$, eight isolated vertices and $K_3 + K_5$; at $M = 10$ it is C_{10} , P_{10} , $K_{5,5}$, K_{10} , $C_4 + C_6$, $K_4 + K_6$, $2 \times C_5$ and $5 \times K_2$. The totals over the family: 334 half-filled level-set basis states and 1544 basis states over all $M + 1$ sectors of all 74 graphs (per amplitude set; 540 sectors, 1080 with both amplitude sets); 452 exact sector nullities at $M \leq 8$; 25690 pairwise fermion-sign checks with 0 violations and recorded sign set $\{+1\}$. At $M = 10$ the block has $\binom{10}{5}^2 = 63504$ states and the hopping matrix restricted to $d \leq 1$ is 252×6552 ; the exact nullity computations were deliberately not run

at that size, and the records say so—nullity flagged as not computed and the chain dimension reported as absent, rather than a number obtained by a cheaper route. The dimension claim at $M = 10$ rests on Theorem 5 plus the purely combinatorial sliding-class certificate C7, with no linear algebra.

Independence is built into the design, so that no structural claim rests on a single computation: the chain dimension against the nullity against the count formula against the sliding-class breadth-first search (C1, C2, C7, C10); the $SU(2)$ product state against the level-set indicator (C3); the structural Slater rank against dense exact row reduction and explicit Gram matrices (C5); unit against random amplitudes (C6); the fermionic S^2 against the closed-form integer matrix (C12); idempotent ranks against characteristic polynomials against recursive coupling against the highest-weight count (C13, C21); the matrix purity test against the fermionic support test (C14); and the operator route against the Clebsch–Gordan route against the occupancy dictionary for the pair condition (C22). One methodological trap is recorded because it cost a silently wrong first pass elsewhere and would cost one here: the Clebsch–Gordan constructor of the symbolic library takes interleaved arguments, and read in the swapped order it fails *silently*—returning zero whenever the permuted arguments violate $m_1 + m_2 = M$, and otherwise a valid coefficient of a different problem. The symbolic engine therefore asserts two canonical coefficient values before trusting any expansion, and route B of C22 reproducing route A on all 270 weight vectors is what pins the convention down.

What was not computed. The honest boundary of the certificates: exhaustive isomorphism coverage exists only at $M = 2$ and $M = 4$ (all graphs) and $M = 6$ (all bipartite graphs); the non-bipartite graphs at $M = 6$ and the families at $M = 8$ and $M = 10$ are curated samples, listed by name above and in the records; $M \geq 12$ was never touched by any fermionic linear algebra, and the large- M integers of Corollary 16 rest on Theorem 5 and three combinatorial routes, not on nullity computations; the exact nullity at $M = 10$ was deliberately skipped, as the records state. Nothing locates these zero modes within the spectrum; no dynamics, no spectral flow and no thermodynamic limit is computed; sectors with $d \geq 1$ and blocks with $N_\uparrow + N_\downarrow < M$ are untouched except for the C17 probes showing that the half-filling collapse does not extend there. The pseudospin statement is annihilation, $\eta^\pm = 0$ on Z , not η -conservation, which needs bipartiteness. The definite-spin basis of C13 and C20 is deterministic (rational idempotent columns) but not canonical: where $n_S \geq 2$ the individual Slater numbers depend on the eigenvector chosen. The zero-spin factors of Proposition 12 are formal in this setting—no graph has an empty component—so that branch is certified only by the abstract routes of C22, never fermionically. And no floating-point number enters any structural claim anywhere: the only floats in the records are the reported entropies and entropy densities.

10 Discussion and outlook

For the half-filled no-doublon sector of the Hubbard model on an arbitrary finite hopping graph with arbitrary nonzero amplitudes, the picture is now complete. The descending chain of invariant subspaces collapses after one step, because one electron per site leaves no empty target: the interaction-independent zero modes are exactly the kernel of the hopping operator, all at eigenvalue zero, with no generalized eigenvectors (Lemma 1, Theorem 5). Their number in a sector is the number of ways of distributing the up-spins over the connected components subject only to the component-size bounds, and an explicit basis is given by the all-plus-one indicator vectors of the level sets, whose coefficients are all +1 by an exact fermion parity computation (Lemmas 3 and 4). The same basis is a product of component spin-lowering operators applied to the all-up state, which identifies the zero-mode space as a tensor product of one maximal-spin multiplet per component, with an integer, amplitude-free matrix for S^2 from which everything else is computable at any size without fermions (Theorem 6). The purity criterion is exact: a basis state

Table 2: Certified graph families (seed 20260908; Python 3.12.3; exact `Fraction` arithmetic throughout). “Nullity computed” records whether the exact nullity of the full $d \leq 1$ hopping matrix was run at that size. Every graph was run twice: at unit amplitudes and at pseudo-random nonzero integer amplitudes drawn from $\{\pm 1, \pm 2, \pm 3, 5, 7, 11\}$ (C6).

family	graphs	exhaustive?	nullity computed	claims
every graph, $M = 2$	2	yes (all isomorphism classes)	yes	C1–C9
every graph, $M = 4$	11	yes (all isomorphism classes)	yes	C1–C9
every bipartite graph, $M = 6$	35	yes	yes	C1–C9
curated non-bipartite, $M = 6$	8	no (sample)	yes	C1–C9
curated, $M = 8$	10	no (sample)	yes	C1–C9
curated, $M = 10$	8	no (sample)	no	C2–C9
total	74	—	66 of 74	0 failures

Table 3: Certificates C1–C9 (descending-chain engine, half-filled block and its structure). “Exact” means integer or rational arithmetic with no floating-point input.

id	what was computed	outcome
C1	$\dim \mathcal{V}_\infty$ by the exact rational descending chain (4), against the count $\#\{c\}$; combinatorial variant: number of sliding classes against $\#\{c\}$	equal on all 66 nullity graphs ($M \leq 8$); class count equal on all 74
C2	three independent numbers per graph: chain dimension, exact nullity of the full $d \leq 1$ matrix, count formula	equal on every graph where the nullity was computed
C3	$T\mathbf{1}_c = 0$ and $T\psi_c = 0$ exactly; $\psi_c = \mathbf{1}_c$ coefficient by coefficient (Theorem 6(i))	all 334 half-filled states, exact
C4	fermion signs of Lemma 3: both processes per target compared bitwise ($s_{\text{up}} + s_{\text{dn}} = 0$, identical target bitmask)	25690 checks, 0 violations, sign set $\{+1\}$
C5	Slater array of every half-filled state: nonzero count, exact rank, unit coefficients, distinct rows and columns, Gram matrix; dense exact row reduction and AA^T where the array is at most 128×128	$\text{nnz} = \text{rank} = \prod_i \binom{b_i}{c_i}$, Gram = identity, on all 334 states; $M = 10$ ranks from the signed-permutation structure, as the record states
C6	everything structural re-run at pseudo-random nonzero integer amplitudes from $\{\pm 1, \pm 2, \pm 3, 5, 7, 11\}$	dimensions, count formula, the same integer vectors annihilated, Slater ranks and spin weights reproduce exactly
C7	breadth-first search on the token-sliding graph over all $\binom{M}{n}$ up-sets: class count and every class size, no linear algebra	classes = $\#\{c\}$ and sizes = $ L_c $ on every certified graph, including $M = 10$
C8	random inhomogeneous density-density V_{ij} on every graph (Corollary 11)	every support state singly occupied; exactly one interaction eigenvalue, $\sum_{i \leq j} V_{ij}$
C9	exact rational spin weights $w_S = \langle \psi P_S \psi \rangle / \langle \psi \psi \rangle$ through rational idempotents, plus explicit $\eta^\pm$ annihilation	over all 334 half-filled states: weights sum to 1; all η -singlets; connected-graph states have spin exactly $M/2$; 0 of the 296 disconnected-graph states is a spin eigenstate; $M = 4$ cross-check reproduces the textbook Clebsch–Gordan squares of $1 \otimes 1$ (weights $\frac{1}{3}, \frac{1}{2}, \frac{1}{6}$ and $\frac{1}{3}, 0, \frac{2}{3}$)

Table 4: Certificates C10–C16 (all-sector spin engine), on 74 graphs, 540 sectors at unit amplitudes and the same 540 at pseudo-random nonzero integer amplitudes: 1544 level-set basis states each way, 0 failures. Certificates C17–C22 continue in Table 5.

id	what was computed	outcome
C10	$\dim \ker(T _{d=0,p})$ by exact nullity of the full $d = 0 \rightarrow d = 1$ matrix in every sector of every graph, against $\#\{c\}$ and $\text{rank}\{\mathbf{1}_c\}$; w_m by three routes	452 sector nullities, all equal to $\#\{c\}$; $\sum_p \#\{c\} = \prod_i (b_i + 1)$ on every graph
C11	$[T, S^\pm] = 0$ on all of $\mathcal{F}_M$, doublons included, state by state; $T(S^\pm v) = 0$ for basis vectors of every Z_p ; $S^\pm \mathbf{1}_c$ against the integer ladder formulas (22)	81028 commutator checks on 56 graphs, 0 mismatches; 3728 intertwiner checks, all exact
C12	fermionic $S^2 = S^- S^+ + S^z(S^z + 1)$ applied to every $\mathbf{1}_c$, read off class by class (constancy on each L_c itself tested), against the closed-form integer matrix (21)	exact equality on all 1544 states, both amplitude sets; all entries integers; no dependence on the w_{ij}
C13	geometric multiplicities of $S(S+1)$ from exact rational idempotents of (21); characteristic polynomial against $\prod_S (x - S(S+1))^{n_S}$; n_S from recursive irrep coupling; every eigenvector re-verified in the Fock space, $S^2 v = S(S+1)v$ and $Tv = 0$ unprojected	1544 definite-total-spin zero modes at unit amplitudes and 1544 at random amplitudes, 0 failures
C14	purity of each $\mathbf{1}_c$ two ways—off-diagonal vanishing in the C12 matrix and the fermionic support test—against the criterion $r = 1$ or $p \in \{0, M\}$	348 of 1544 pure, exactly the predicted set: all 276 connected-graph states plus $p = 0$ and $p = M$ on each of the 36 disconnected graphs (72)
C15	$\eta^\pm$ annihilation of every $\mathbf{1}_c$ in every sector	all singlets, both amplitude sets
C16	the four small explicit blocks of Table 6 recomputed in the physical normalization	matrices and characteristic polynomials as tabulated; the earlier illustrative tables are exactly $4\times$ these on the $S^z = 0$ blocks; the (scale-invariant) eigenvectors reproduce

Table 5: Certificates C17–C22 (all-sector spin engine, continued from Table 4, and the symbolic pair-condition checker).

id	what was computed	outcome
C17	$\text{rank}(PTP)$ on $\ker D \cap \mathcal{F}_M$ and $\dim \ker(PTP)$ against $\dim Z$; the same comparison at $N < M$	0 nonzero projected matrix elements on every graph, $\dim \ker(PTP) = 2^M$ against $\prod_i (b_i + 1)$; at $N < M$ the projected kernel is strictly larger: 8 vs. 7, 16 vs. 14, 68 vs. 25
C18	dimer cover and empty graph at $M = 4, \dots, 400$ in exact Python integers; every ordered component-size pattern at $M \leq 14$	$\sum_m w_m = \sum_S n_S (2S + 1) = \prod_i (b_i + 1)$ everywhere; trinomial and binomial closed forms confirmed; all $M \leq 14$ patterns agree across polynomial product, inclusion-exclusion (20) and recursive coupling
C19	the two fermion engines compared state by state on T, S^+, S^-	16484 states, 0 mismatches
C20	exact rank of the up/down Slater array of every C13 eigenvector at $M \leq 6$, against its support size	699 eigenvectors per amplitude set: rank = support in every case, 665 of them full support (maximal Slater number $\binom{M}{p}$)
C21	$\dim \ker(S^+ _{Z_p})$ from the fermionic integer ladder matrix, against $n_{ m }$ from recursive coupling	exact agreement in every sector of every graph, both amplitude sets
C22	pair condition, three routes: (A) exact symbolic S^2 in the orthonormal product weight basis (eigenvector iff the off-diagonal support is empty); (B) exact iterated Clebsch–Gordan expansion (pure iff a single total S carries nonzero squared weight); (C) radical-free dictionary $(P) \Leftrightarrow (*) \Leftrightarrow (r = 1 \text{ or } p \in \{0, M\})$ over the recorded C10–C21 dataset	33 j -tuples, 270 weight vectors: $A = B = (P)$, 0 mismatches; aligned $\subseteq$ eigenvectors in all 33; clause predictions correct in all 33; failure counts 3, 4, 6, 3, 3, 5 against aligned 2; and 3088 (state, amplitude-set) instances in 1080 sectors: 0 mismatches, dimension and pure-state count reproduced in all 1080

Table 6: The four small exact blocks of certificate C16: the matrix of S^2 in the unnormalized level-set basis $\{\mathbf{1}_c\}$ (rows c' , columns c , integer entries from (21)), its characteristic polynomial, and the spin reading. The characteristic polynomial factors as $\prod_S (x - S(S+1))^{n_S}$ with the multiplicities of (19). Normalization matters: the doubled spin convention ($2S^z$, $2S^\pm$) gives $S'^2 = 4S^2 - 2S^z$, i.e. exactly $4S^2$ on an $S^z = 0$ block, and tables quoted in that convention are $4\times$ the physical matrices below while their eigenvectors, being scale-invariant, are unchanged.

block	S^2 in the $\mathbf{1}_c$ basis	characteristic polynomial	spins
$2K_1$, $M = 2$, $p = 1$	$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$x^2 - 2x = x(x - 2)$	$S = 0, 1$
$K_2 + K_2$, $M = 4$, $p = 2$	$\begin{pmatrix} 2 & 4 & 0 \\ 1 & 4 & 1 \\ 0 & 4 & 2 \end{pmatrix}$	$x^3 - 8x^2 + 12x = x(x - 2)(x - 6)$	$S = 0, 1, 2$
$K_3 + K_1$, $M = 4$, $p = 2$	$\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$	$x^2 - 8x + 12 = (x - 2)(x - 6)$	$S = 1, 2$
extremal block, $M = 4$, $p \in \{0, 4\}$	$\begin{pmatrix} 6 \end{pmatrix}$	$x - 6$	$S = 2$

has definite total spin precisely when the graph is connected or the sector is extremal—precisely when the sector is one-dimensional—and the spin is then maximal (Theorem 7); in abstract spin coordinates the criterion is the two-clause pair condition, complete for arbitrary products of spin representations including the zero-spin factors no graph can realize (Proposition 12, Corollaries 13 and 14). Every zero mode is an exact pseudospin singlet for any phase convention, so the whole space decomposes into $\text{SO}(4)$ sectors $\{0, S\}$ with exact multiplicities $n_S = w_S - w_{S+1}$, and a basis of definite-sector zero modes exists on every graph (Theorem 8). The minimum number of Slater determinants of any no-doublon state equals its support size; for the level-set basis it is the product of binomial coefficients, the up/down spectrum is flat and the entropy is its logarithm (Theorem 9). None of this depends on the interaction being the Hubbard term—any density-density interaction acts as a c -number here (Corollary 11)—and zero amplitudes are exactly edge deletions, so every theorem transfers to the support graph (Remark 15).

The sharp physical statement is what bond dilution destroys and what it preserves. Cutting the bonds of a connected lattice into r components multiplies the half-filled zero-mode space from one state to $\#\{c\}$, an extensive residual degeneracy of exact interaction-independent eigenstates growing between $\log(3)/2$ and $\log 2$ nats per site for the two bracketing patterns (Corollary 16). What survives is the sector structure: the $\text{SO}(4)$ labels remain well defined on every graph, with exact multiplicities. What is destroyed is the purity of the natural integer basis: on a disconnected graph no half-filled level-set state carries a definite total spin, the labelled vectors are Clebsch–Gordan combinations of level sets rather than the level sets themselves, and n_S states, not one, carry each label. A classification of interaction-independent eigenstates by $\text{SO}(4)$ quantum numbers therefore remains meaningful beyond connected bipartite lattices [14, 13], at the cost of multiplicity. The results also put the earlier connected bipartite theorem—one state, the maximal-spin tower state $(S^-)^n |\uparrow\rangle/n!$, Slater rank $\binom{M}{n}$, maximal species entanglement—in its place as the $r = 1$ corner of Theorems 5, 6 and 9: the sector of the connected half-filled state is $\{0, M/2\}$, the maximal-spin, zero-pseudospin corner saturated by the Yang–Zhang family [10, 11], and “maximal entanglement” becomes “flat spectrum on the support” for general graphs.

The limitations are those of Section 9, and they are of kind, not degree. The theorems hold for every graph and every size; the fermionic certificates are finite—exhaustive at $M = 2, 4$ and for bipartite graphs at $M = 6$, curated up to $M = 10$, combinatorial up to $M = 400$ —and the nullity at $M = 10$ was deliberately not computed. The pseudospin statement is annihilation, not conservation: η remains a symmetry of the Hamiltonian only on bipartite graphs. Where $n_S \geq 2$ the definite-spin basis is deterministic but not canonical, and individual Slater numbers depend on the choice of eigenvector inside the eigenspace. Disconnected hopping graphs model bond-diluted, vacancy-ordered or decoupled-cluster systems, not generic translation-invariant lattices,

and no thermodynamic limit is taken. Nothing locates these energies within the spectrum.

Four directions are open, in increasing distance from the present methods.

1. **The canonical pure-spin basis and its species entanglement.** The isomorphism $Z \cong \bigotimes_i \mathcal{H}_{b_i/2}$ yields a canonical definite-spin basis without further choice: the joint eigenbasis of the intermediate Casimir operators $(\sum_{i \in A_k} \vec{S}_i)^2$ for any fixed nested chain of component sets $\emptyset \subset A_1 \subset \dots \subset A_{r-1} \subset \{1, \dots, r\}$, together with S^2 and S^z ; the joint eigenspaces are one-dimensional. By Theorem 9(i) the exact minimum Slater number of such a vector is its support size, i.e. the number of nonzero Clebsch–Gordan coefficients of that coupling chain in the occupancy basis—a number not computed anywhere here. Three specific questions: is the canonical highest-weight vector of every spin- S copy always full support, hence maximally species-entangled with entropy $\log \binom{M}{p}$? For the perfect dimer cover, where $Z \cong \mathcal{H}_1^{\otimes M/2}$ and the multiplicities n_S grow exponentially, what is the support and entropy distribution over the canonical basis as M grows—does the typical pure-spin zero mode saturate the block, and does a limit shape exist? And is any of this independent of the coupling chain chosen? The certified basis is not canonical—at $M \leq 6$, 665 of its 699 eigenvectors have full support and 34 do not—so both behaviors occur and the canonical count is genuinely open.
2. **The one-doublon sector at half filling.** This paper settles D -eigenvalue 0; the $d = 1$ sector is where T maps $\ker D$, and earlier work on connected graphs records that the below-saturated remainder is not a function of one-particle zero-mode data and that no $M = 4$ remainder state is a single determinant. For disconnected graphs neither the dimension of the largest T -invariant subspace inside $\ker(D - 1)$ nor its Slater structure is known. Lemma 1 fails there—a one-doublon state has an empty site—so T can map $d = 1$ to $d = 1$, and genuine nonzero T -eigenvalues and generalized eigenvectors become possible. What is $\dim \mathcal{V}_\infty$ inside $\ker(D - 1)$ as a function of the component sizes, and are its vectors Slater decomposable?
3. **Observable consequences of the extensive degeneracy.** The states counted here are exact eigenstates at every interaction strength, but nothing locates them in the spectrum or couples them to a thermal or driven state. For a bond-diluted Hubbard model above the percolation threshold—one giant component plus a finite density of small ones—does the residual degeneracy produce an interaction-independent residual entropy, and how does it compete with the giant component’s single zero mode? Answering this needs the spectral position of these states, which is outside the present scope.
4. **What replaces Lemma 1 away from half filling?** In a block with $N_\uparrow + N_\downarrow < M$ empty sites exist, T has nonzero $d = 0$ matrix elements, and $\mathcal{V}_\infty$ can contain genuine eigenvectors with nonzero eigenvalue as well as generalized ones. Is $\dim \mathcal{V}_\infty$ still a function of the component sizes alone, and still independent of the nonzero amplitudes? The descending-chain computation (4) does not use half filling, so the dimension question is directly attackable with the present methods; only the basis construction ψ_c and the all-plus-one sign lemma are specific to one electron per site.

References

- [1] Fabian H. L. Essler, Vladimir E. Korepin, and Kareljan Schoutens. Completeness of the $SO(4)$ extended Bethe ansatz for the one-dimensional Hubbard model. *Nuclear Physics B*, 384:431–458, 1992.
- [2] GeWu Journal Auto Author. Source research repository for this paper. <https://git.gewu-lab.ai/scientific-author-90/hubbard-w-independent-general-bipartite-blocks>, 2026. Solution repository at commit 9febb40c3e96a2e77e53a95549e6197c87513fc3.
- [3] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A*, 276:238–257, 1963.
- [4] Sanjay Moudgalya, B. Andrei Bernevig, and Nicolas Regnault. Quantum many-body scars and Hilbert space fragmentation: a review of exact results. *Reports on Progress in Physics*, 85:086501, 2022.
- [5] Sanjay Moudgalya and Olexei I. Motrunich. Exhaustive characterization of quantum many-body scars using commutant algebras. *Physical Review X*, 14:041069, 2024.
- [6] Sanjay Moudgalya, Nicolas Regnault, and B. Andrei Bernevig. η -pairing in Hubbard models: From spectrum generating algebras to quantum many-body scars. *Physical Review B*, 102:085140, 2020.
- [7] Masaya Nakagawa, Hosho Katsura, and Masahito Ueda. Exact eigenstates of multicomponent Hubbard models: $SU(N)$ magnetic η pairing, weak ergodicity breaking, and partial integrability. *Physical Review Research*, 6:043259, 2024.
- [8] Jun-Hang Ren, Ming-Yong Ye, and Xiu-Min Lin. Slater determinant and exact eigenstates of the two-dimensional Fermi–Hubbard model. *Chinese Physics B*, 27:073102, 2018.
- [9] Oskar Vafek, Nicolas Regnault, and B. Andrei Bernevig. Entanglement of exact excited eigenstates of the Hubbard model in arbitrary dimension. *SciPost Physics*, 3:043, 2017.
- [10] Chen Ning Yang. η pairing and off-diagonal long-range order in a Hubbard model. *Physical Review Letters*, 63:2144–2147, 1989.
- [11] Chen Ning Yang and S. C. Zhang. SO_4 symmetry in a Hubbard model. *Modern Physics Letters B*, 4:759–766, 1990.
- [12] Chen Ning Yang and S. C. Zhang. SO_4 symmetry in a Hubbard model. *International Journal of Modern Physics B*, 5:977–984, 1991. Republication of the 1990 paper.
- [13] Ming-Yong Ye. Exact eigenstates of Hubbard models with and without bond-charge interactions on general graphs. *Communications in Theoretical Physics*, 78:045703, 2026. Preprint arXiv:2506.01553.
- [14] Ming-Yong Ye and Xiu-Min Lin. Exact eigenstates of the Hubbard model independent of the on-site electron interaction. *physica status solidi (b)*, 255:1700321, 2018. Preprint arXiv:1606.08364.