This is the accepted manuscript made available via CHORUS. The article has been published as:

Compact unary coding for bosonic states as efficient as conventional binary encoding for fermionic states

Hatem Barghathi, Caleb Usadi, Micah Beck, and Adrian Del Maestro

Phys. Rev. B **105**, L121116 — Published 29 March 2022

DOI: 10.1103/PhysRevB.105.L121116

Balls and Walls: A Compact Unary Coding for Bosonic States

Hatem Barghathi,[1] Caleb Usadi,[2] Micah Beck,[3] and Adrian Del Maestro[1,3]

[1] *Department of Physics and Astronomy, University of Tennessee, Knoxville, TN 37996, USA*
[2] *Department of Physics, University of Vermont, Burlington, VT 05405, USA*
[3] *Min H. Kao Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA*

We introduce a unary coding of bosonic occupation states based on the famous "balls and walls" counting for the number of configurations of N indistinguishable particles on L distinguishable sites. Each state is represented by an integer with a human readable bit string that has a compositional structure allowing for the efficient application of operators that locally modify the number of bosons. By exploiting translational and inversion symmetries, we identify a speedup factor of order L over current methods when generating the basis states of bosonic lattice models. The unary coding is applied to a one-dimensional Bose-Hubbard Hamiltonian with up to $L = N = 20$, and the time needed to generate the ground state block is reduced to a fraction of the diagonalization time. For the ground state symmetry resolved entanglement, we demonstrate that variational approaches restricting the local bosonic Hilbert space could result in an error that scales with system size.

Finite lattice models represent an essential simplification in quantum many-body physics, where a set of discrete low energy degrees of freedom are sufficient to capture the relevant phases and phase transitions of a more complete high energy description. The states of these quantum systems form a finite dimensional Hilbert space $\mathcal{H}$ and are thus amenable to a numerical representation through the exact diagonalization (ED) of a lattice Hamiltonian [1–14]. This provides a complete description of the eigenstates, however, it is limited by the exponentially increasing cardinality $|\mathcal{H}| \sim \mathrm{e}^L$ of $\mathcal{H}$ for L lattice sites. This problem has motivated stochastic and variational approaches, such as quantum Monte Carlo (QMC) [15–18], and the density matrix renormalization group (DMRG) [19–22]. These methods are essential to our understanding of quantum many-body phenomena [23–29], however, they are stochastic, or require an approximate truncation of the local Hilbert space, and thus ED still plays a crucial role in benchmarking novel methods, as well as in gaining access to the density matrix. While extremely efficient approaches exist for the Heisenberg model [1] and its SU(N) generalizations [30] with applications to fermionic lattice models of up to ≈ 40 sites, the expanded Hilbert space of bosonic systems, allowing multiple occupancy of a single spatial mode, presents a formidable challenge [9, 10, 31, 32].

In this Letter we introduce a unary basis (UB) coding of bosonic occupation (Fock) states using a highly compact and compositional, yet still human-readable bitstring exploiting the famous *balls and walls* illustration of particles and sites used in Bose-Einstein counting. For example, for $N = L = 11$:

$$|2,0,1,0,3,0,1,0,0,0,4\rangle \equiv |\bullet\bullet|\ |\bullet|\ |\bullet\ \bullet\ \bullet|\ |\bullet|\ |\ |\ |\bullet\ \bullet\ \bullet\ \bullet$$
$$I_{UB} = 2541296 \equiv 10011011000110111110000 \tag{1}$$

where $|n_1, n_2, \ldots, n_L\rangle$ is a characteristic basis state with n_j the occupancy of site j. In the UB coding, 1's denote sites with the following number of 0's corresponding to that site's occupation, and each basis state can be represented as an integer in base 10. In addition to reducing the memory required to store the basis by a factor L, this approach provides rapid access to n_j and significantly accelerates the action of local operators via bitwise operations on I_{UB}. An implementation in the presence of lattice symmetries yields a reduction of computational complexity by a factor of L reducing the time needed to generate the basis for a one dimensional (1D) Bose-Hubbard (BH) model to a fraction of that needed to obtain the ground state via iteration. This yields a practical speedup of over 20× compared to current methods based on unique integer ordering of permanents [9, 10] for system sizes up to $L = N = 20$.

The utility of the unary basis is validated by studying entanglement at the critical point of the 1D BH model for both spatial mode and particle bipartitions as well as in the presence of a U(1) symmetry fixing the total number of particles. In the latter case, we show that the common practice of restricting the number of bosons per site to a small integer $n_{\max}$ ("soft-spin" approach) can lead to large relative errors that scale with the system size. This finding serves to challenge the conventional wisdom that an investigation of the $n_{\max}$ dependence of the energy alone is enough to justify this approximation. To facilitate adoption of the unary basis in future work, a complete software implementation is included in Ref. [33].

Background – For both fermionic and bosonic lattice models, it is natural to diagonalize any local N-particle Hamiltonian in a basis of spatial modes $|n_1, n_2, \ldots, n_L\rangle$. For fermions, the Pauli exclusion principle restricts $n_j = 0, 1$ and each of the $|\mathcal{H}_f| = \binom{L}{N}$ basis states can be encoded as a binary word of length L that can be stored as a $2^{\lceil \log_2 L \rceil}$-bit integer as is implemented in commonly used ED software [11–13, 34]. For bosons, the possibility of multiple occupancy, $n_j = 0, \ldots, N$ on any site enlarges the Hilbert space to $|\mathcal{H}| = \binom{N+L-1}{N}$ with basis vectors for

the occupation states naturally parameterized by $L|\mathcal{H}|$ integers. The memory required to store the Hamiltonian has an upper bound $\sim |\mathcal{H}|^2$ and thus dwarfs that needed to store the basis states. However, in bosonic lattice models with limited range hopping, the Hamiltonian is sparse, reducing its storage cost to $\sim L|\mathcal{H}|$ and encoding the basis as arrays of site occupations now composes a leading share of the required memory. To address this problem, permanent ordering (PO) schemes [9, 10, 31, 32] have been introduced which assign a unique contiguous integer label $I_{PO}(n_1, \ldots, n_L)$ to each occupation state via an iterative procedure of complexity $\mathrm{O}(L)$. To further avoid the memory impact of storing $\{I_{PO}\}$, a lookup table combined with an $\mathrm{O}(L)$ inverse function I_{PO}^{-1} [10], can be implemented to gain access to site occupations numbers n_j. While this method has proven to be effective, e.g., in studying many-body localization [35, 36] it has yet to be extended to exploit lattice symmetries (e.g. translation, inversion) where the maximum number of non-zero Hamiltonian matrix elements is reduced to $\sim |\mathcal{H}|$. Here it is crucially important to suppress the now leading order memory share ($\sim L|\mathcal{H}|$) of the basis.

Unary Basis – To study the largest possible systems, we introduce an integer labelling scheme motivated by the formal equivalence between occupation states of N bosons on L sites, and those of N fermions on $N + L$ lattice sites [32]. By reinterpreting bits corresponding to the presence (`1`) or absence (`0`) of a fermion with a site boundary and boson respectively, the resulting set of non-contiguous $|\mathcal{H}|$ integers $\{I_{UB}\}$ can encode information about the occupation state directly in their bit string (see Eq. (1)) while remaining as efficient as a binary encoding for fermions. This can be quantified by defining the efficiency of the representation as $\eta = \log_2 |\mathcal{H}| / \log_2 2^{L+N}$, providing a maximum efficiency at unit filling ($\nu = N/L = 1$, with N and $L \gg 1$) of $\eta = 1 - (1/4L)\log_2 L + \mathrm{O}(L^{-1})$, where for $L = 20$, $\eta \simeq 0.925$. Comparing with an alternative approach where N bosons are mapped to a spin model with $S = N/2$ [12, 13], that uses a sequence of $L\lceil \log_2(N+1) \rceil$ bits coding $n_j \leq N$ on each site, the corresponding efficiency at unit filling is $\eta \approx 2/\lceil \log_2(L) \rceil - 1/(2L) + \mathrm{O}((L \log_2 L)^{-1})$ which is smaller by a factor of $2/\lceil \log_2(L) \rceil$ compared to the unary coding, and for $L = 20$, $\eta \approx 0.37$.

In addition to memory compactness, the UB coding can accelerate the inverse operation I_{UB}^{-1} to obtain access to occupation numbers. Given a 64-bit integer I_{UB} representing $|n_1, \ldots, n_L\rangle$, n_1 can be found by counting the trailing zeros in the bit string of I_{UB} (a "compiler builtin"). Shifting the bits of I_{UB} by $n_1 + 1$ allows for reading the number of particles on the next site, and by repeating these bitwise operations L times, the corresponding occupation vector (OV) $|n_1, \ldots, n_L\rangle$ can be constructed. This procedure can be further sped up by viewing the 64 bits of I_{UB} as a sequence of four 16-bit integers from which the corresponding number of sites and occupation numbers can be obtained by direct lookup of the 2^{16} possibilities then recomposed to generate the corresponding OV. We find that generating OVs for $\{I_{UB}\}$ is more than $4\times$ faster than obtaining them on the fly using permanent ordering, at a cost of additional memory usage corresponding to only $1/L$ of that needed to store the system Hamiltonian in the absence of translational symmetry.

Lattice Symmetries – The utility of the unary coding can be extended to treat systems that preserve Hamiltonian symmetries such as translation and inversion (see Supplemental Material [37]). If the Hamiltonian commutes with the translation operator $\widehat{T}$, it has a block diagonal structure where each of the L blocks has a quasi-momentum index q and contains a maximum number of non-zero elements that scales as $\sim |\mathcal{H}|$ (for short range hopping). The resulting number of translationally symmetrized basis states (the q^{th} degenerate set of the eigenstates of $\widehat{T}$) scales as $\sim |\mathcal{H}|/L$. For a given q and OV, an eigenvector $|\phi_{\alpha,q}\rangle$ of $\widehat{T}$ can be generated: $|\phi_{\alpha,q}\rangle = L_\alpha^{-1/2} \sum_{j=0}^{L_\alpha - 1} \exp\left[-2\imath\pi jq/L_\alpha\right] \widehat{T}^j |n_1, n_2, \ldots, n_L\rangle$, where α is the index of a *cycle* with length $L_\alpha \leq L$ [38] generated by repeatedly acting on a given $|n_1, n_2, \ldots, n_L\rangle$ with $\widehat{T}$. Each cycle can be mapped to a set of eigenvectors of $\widehat{T}$ with the same number of elements.

Calculating matrix elements of local operators requires access to all OVs in $|\phi_{q,\alpha}\rangle$, as well as the ability to reconstruct the state $|\phi_{q,\alpha}\rangle$ where a given OV appears. The naive storage cost for direct lookup of this information is at least double that needed for $\{I_{UB}\}$. Alternatively, one can trade memory with computational complexity by storing L_α for each cycle, as well as an extremal integer I_α that represents one OV characteristic of the cycle, (i.e. I_{PO} or its unary coding I_{UB}). Thus, when inverting I_α to an OV, the rest of the cycle can be obtained by L_α translations. If $I_\alpha \in \{I_{UB}\}$ the unary coding of the OVs of the cycle can be obtained by bit shifting I_α. As a single 64-bit instruction can shift a vector of 8 bytes (using parallel intra-processor bit paths) faster than moving each byte sequentially (using multiple memory operations), the UB coding provides a reliable constant speedup.

The inverse process of finding the cycle α where a given $|n_1, \ldots, n_L\rangle$ appears can be done in four steps: (i) cyclic shifting L times; (ii) converting each of the L OVs to their integer representation; (iii) finding the characteristic integer; and (iv) performing a fast search for α in the ordered list of integers I_α. While steps (i) and (iii)-(iv) require at most $\mathrm{O}(L)$ operations, the time complexity of step (ii) is $\mathrm{O}(L^2)$, as it consists of repeating the operation of converting an OV to an integer L times. However, by exploiting the unary coding, the asymptotically slowest step (ii) is not required, as the shifting process (i) can be directly applied to I_{UB} to obtain the targeted list of integers. This represents a significant opportunity for a reduction in complexity and speedup over PO methods.

Benchmarking – As a test case, we consider the unary coding for N bosons on a ring of $L = N$ sites ($\nu = 1$) governed by the Bose-Hubbard Hamiltonian:

$$\widehat{H} = -\sum_{i=1}^{L}\left(b_{i+1}^{\dagger}b_i + \text{h.c.}\right) + \frac{U}{2}\sum_{i=1}^{L} n_i(n_i-1), \qquad (2)$$

where $n_i = b_i^{\dagger}b_i$, $[b_i, b_j^{\dagger}] = \delta_{i,j}$ and $b_{L+1} = b_1$. The on site interaction $U > 0$ is measured in units of the hopping. Eq. (2) exhibits a continuous phase transition from a superfluid ($U \ll 1$) to an insulator ($U \gg 1$) at $U \approx 3.3$ [23, 39, 40]. Using the unary coding of the basis, as well as translational ($q = 0, \ldots, L-1$) and inversion ($r = \pm 1$) symmetries on a ring, we have obtained the ground state $|\Psi_0\rangle$ of Eq. (2) via exact diagonalization for up to $N = 20$ bosons on $L = 20$ sites.

To compare the efficiency of the PO and UB methods, we measure the times $t_{\rm PO}$ and $t_{\rm UB}$ spent in constructing the ground state block $\mathcal{M}_{q=0,r=1}$ of the sparse matrix representing $\widehat{H}$. We find that the ratio $t_{\rm PO}/t_{\rm UB}$ (Fig. 1 left), scales as $\mathrm{O}(L)$, as expected from steps (i)–(iv) above. For $L = N = 20$, we find a speedup of $t_{\rm PO}/t_{\rm UB} \gtrapprox 20$. The practicality of this linear in L speedup becomes evident when measuring in units of t_D, the time required to iteratively obtain the ground state $|\Psi_0\rangle$ from $\mathcal{M}_{q=0,r=1}$. Again, for $L = N = 20$, we find $t_{BU}/t_D \approx 0.19$ compared to $t_{PO}/t_D \approx 3.96$, as illustrated in the right panel of Fig. 1. Thus, employing the unary coding can practically reduce computation time by a factor of $(t_{PO} + t_D)/(t_{UB} + t_D) \gtrsim 4$. A larger larger computational speed up ($\gtrapprox 20$ for $L = N = 18$) is identified when calculating the elements of the one-body density matrix (OBDM) $\langle b_i^{\dagger}b_j\rangle$ (right panel of Fig. 1).

Application – As a further illustration of the utility of the unary coding allowing the use of an unrestricted local bosonic Hilbert space, we investigate several measures of entanglement in $|\Psi_0\rangle$. In all calculations, the required memory was less than 1.5 terabyte (TB).

By partitioning the spatial modes of the pure state $\rho = |\Psi_0\rangle\langle\Psi_0|$ into ℓ consecutive sites A_ℓ and its complement $\bar{A}_\ell$, corresponding to the remaining $L-\ell$ modes, we obtain the reduced density matrix $\rho_{A_\ell} = \mathrm{Tr}_{\bar{A}_\ell}\,\rho$ by tracing out all degrees of freedom in $\bar{A}_\ell$. The von Neumann entanglement entropy $S(\ell) = -\mathrm{Tr}\,\rho_{A_\ell}\ln\rho_{A_\ell}$ quantifying the amount of entanglement that exists between A_ℓ and $\bar{A}_\ell$ [41] is shown as square symbols in Fig. 2. The results demonstrate the reduction of spatial mode entanglement as the on-site repulsion U is increased and the system transitions from a superfluid to localized phase.

With access to an exact representation of the ground state $|\Psi_0\rangle$, we can also study particle entanglement for a partition A_n of n particles [42–51]. This is most easily performed using the first quantized basis $|i_1, \ldots, i_N\rangle$, where $i_k \in \{1, \ldots, L\}$ is the lattice coordinate of the particle labeled $k \in \{1, \ldots, N\}$. The familiar n-particle reduced density matrix of partition A_n is then:

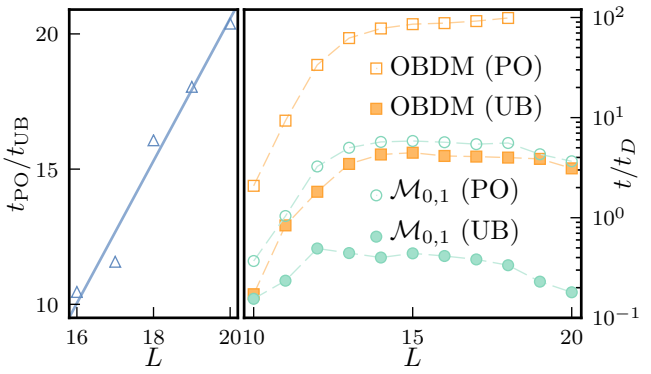


FIG. 1. Computational efficiency of the unary basis (UB) and permanent ordering (PO) encodings. Left: The ratio $t_{\rm PO}/t_{\rm UB}$ of wall clock times needed to construct the ground state block using the permanent ordering (PO) and unary basis (UB) codings as a function of system size L. The line is a linear fit demonstrating the complexity reduction discussed in the text. Right: The wall clock time ratio t/t_D vs. system size L, where t_D is the time consumed in diagonalizing the ground state block $\mathcal{M}_{q=0,r=1}$ of the Bose-Hubbard Hamiltonian and t represents the time spent in constructing $\mathcal{M}_{q=0,r=1}$ (circles) or calculating the elements of the one body density matrix (OBDM) in the ground state (squares).

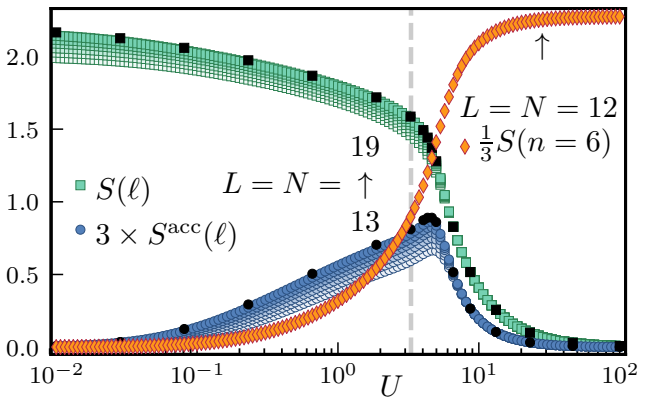


FIG. 2. Mode entanglement $S(\ell)$ and $S^{\rm acc}(\ell)$) and particle entanglement entropy $S(n)$ in the ground state of the BH model as a function of interaction strength U at unit filling ($L = N$). For mode entanglement, the partition size is fixed at $\ell = \lceil L/2\rceil$, where $f(x) = \lceil x\rceil$ is the least integer function

$$\rho_{A_n}^{\{i_k\}_1^n,\{j_k\}_1^n} = \sum_{\{i_k\}_{n+1}^N} \Psi^{\dagger}(\{i_k\}_1^n, \{i_k\}_{n+1}^N)\Psi(\{j_k\}_1^n, \{i_k\}_{n+1}^N),$$

where $\{i_k\}_a^b \equiv \{i_a, i_{a+1}, \ldots, i_{b-1}, i_b\}$. The n-particle entanglement entropy, $S(n) = -\mathrm{Tr}\,\rho_{A_n}\ln\rho_{A_n}$, is shown as diamonds in Fig. 2. In contrast to $S(\ell)$, $S(n)$ is large in the insulating phase, vanishes as $U \to 0$ and is extremely difficult to calculate via DMRG for $n > 2$ [52].

The interplay between fixed particle number and fluctuations between spatial modes, as reflected in symmetry resolved entanglement, is a subject of growing interest [53–80] and can be potentially measured in ultra-cold lattice gases [81, 82]. Here, the presence of conservation laws (e.g. fixed N), reduces the amount of entanglement that is accessible for quantum information processing which can be quantified using the accessible entanglement entropy $S^{\rm acc}(\ell)$ (see Supplemental Material). [83–90]. Figure 2 demonstrates that the accessible entanglement entropy vanishes at the extremes of the two competing phases, and peaks near the critical point.

Restricting the local Hilbert space by enforcing an upper limit $n_{\rm max}$ on the occupation number per site is a widely used approximation when performing the ED of a bosonic Hamiltonian, or the DMRG representation of its ground state [23–25, 39, 91]. The validity of this approximation is judged based on the convergence of observables, (e.g. the ground state energy, the average occupation per site, or its fluctuations [9]) with $n_{\rm max}$. While formally, the approximation is justified only in the insulating phase of the BH model, where the particle number fluctuations are suppressed by repulsive on-site interactions, the resulting error in the ground state energy remains small and finite, even in the superfluid phase down to the non-interacting limit [9]. This empirical finding has yielded the conventional wisdom of utilizing a small $n_{\rm max}$, even deep in the superfluid phase, in extended BH systems [92], and for fermions coupled to bosonic Hilbert spaces (e.g. the Hubbard-Holstein model [93]).

However, the truncated degrees of freedom *could* contribute significantly to other observables that crucially depend on particle number fluctuations, specifically the spectrum of the n-particle reduced density matrix and entanglement in the presence of physical conservation laws. To illustrate, we calculate the error arising from restricting $n_{\rm max} = 4$, in both the ground state energy E and $S^{\rm acc}(\ell)$. Figure 3 shows that while the relative error in E is less than 1%, the relative error in $S^{\rm acc}(\ell)$ could be three orders of magnitude larger in the superfluid phase. It is expected that such large errors would also be present in the density-wave phase of the extended BH model at $U/t \ll 1$ [92]. Moreover, the relative error *increases* with L, (inset Fig. 3), even for $n_{\rm max} = 6$, indicating that large-scale DMRG studies that employ $n_{\rm max}$ may contain logarithmically growing errors. To intuitively understand the origin of such scaling with L, we consider the simpler case of the bosonic density reduced to half-filling, where $S^{\rm acc}$ still vanishes as $U \to 0$. However, enforcing the extreme hard-core constraint of $n_{\rm max} = 1$, which for an appropriate parity of N possesses the same $S^{\rm acc}(\ell = L/2)$ for a partition of consecutive sites as non-interacting fermions. Here, it is known that $S^{\rm acc} \sim \ln L$, and thus, the error will have a similar scaling [78, 90].

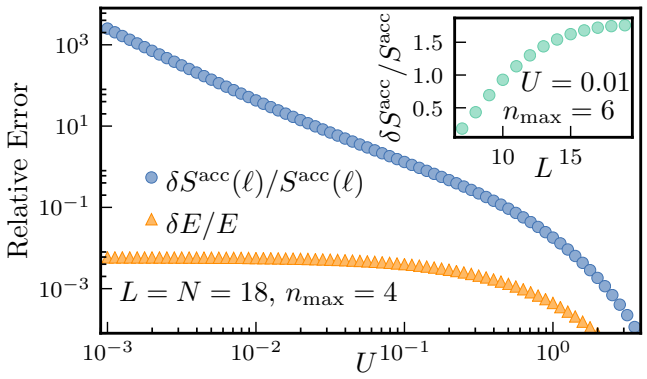


FIG. 3. The relative error in the accessible entanglement $S^{\rm acc}(\ell = 9)$ and the energy E of the BH ground state, arising from fixing the maximum occupancy at all sites to $n_{\rm max} = 4$, as a function of interaction strength U/t at $L = N = 18$. Inset: Scaling of the relative error $\delta S^{\rm acc}(\ell)/S^{\rm acc}(\ell)$ with the system size L at $U = 0.01$ and $n_{\rm max} = 6$, where $\ell = \lceil L/2 \rceil$.

Discussion – In this paper we have introduced a compact unary coding for bosonic basis states that is both human and machine readable by enumerating configurations of indistinguishable particles on distinguishable sites through combinatorial counting of *balls and walls*. Having access to an efficient and compact exact representation of the ground state allows for the study of large system sizes (achieving equivalence for bosonic lattice Hamiltonians with state of the art exact diagonalization methods for fermions) and information quantities such as the particle partition or symmetry resolved entanglement. These are not amenable to measurement via variational (due to the difficulty in computing n-particle density matrices, or errors imposed by restricting the local Hilbert space), or stochastic (limited to Rényi entropies [94]) algorithms. Further constant speedups could be achieved using union data types that allow for the simultaneous view of an occupation vector as an array of smaller integers. This could be useful when studying dynamics in bosonic systems, where access to asymptotic steady states requires considerable computational resources. While we have focused on a one-dimensional Bose-Hubbard Hamiltonian, the unary coding can be applied to any D-dimensional bosonic lattice model utilizing standard large-memory resources. By flattening a D-dimensional configuration, it can be represented by a single integer where translations can be achieved using bitwise operations on I_{UB}. Thus the reduction of time complexity in the presence of translational symmetry is expected to hold even for $D > 1$. We hope that the balls and walls coding will be rapidly incorporated into exact diagonalization software, contributing to the acceleration of future studies of bosonic systems.

This work was supported in part by the NSF under Grant No. DMR-2041995.

[1] H. Q. Lin, "Exact diagonalization of quantum-spin models," Phys. Rev. B **42**, 6561 (1990).

[2] E. Dagotto, A. Moreo, F. Ortolani, D. Poilblanc, and J. Riera, "Static and dynamical properties of doped Hubbard clusters," Phys. Rev. B **45**, 10741 (1992).

[3] H.Q. Lin, J.E. Gubernatis, Harvey Gould, and Jan Tobochnik, "Exact diagonalization methods for quantum systems," Computers in Physics **7**, 400 (1993).

[4] B. Damski, H.-U. Everts, A. Honecker, H. Fehrmann, L. Santos, and M. Lewenstein, "Atomic fermi gas in the trimerized kagomé lattice at 2/3 filling," Phys. Rev. Lett. **95**, 060403 (2005).

[5] Alexander Weiße and Holger Fehske, "Exact diagonalization techniques," in *Computational Many-Particle Physics* (Springer, Berlin, Germany, 2008) p. 529.

[6] J M Zhang and R X Dong, "Exact diagonalization: the Bose–Hubbard model as an example," Eur. J. Phys. **31**, 591 (2010).

[7] Hui Hu, B. Ramachandhran, Han Pu, and Xia-Ji Liu, "Spin-Orbit Coupled Weakly Interacting Bose-Einstein Condensates in Harmonic Traps," Phys. Rev. Lett. **108**, 010402 (2012).

[8] Anders W. Sandvik, "Computational studies of quantum spin systems," AIP Conference Proceedings **1297**, 135 (2010).

[9] Ágnes Szabados, Péter Jeszenszki, and Péter R. Surján, "Efficient iterative diagonalization of the Bose-Hubbard model for ultracold bosons in a periodic optical trap," Chem Phys **401**, 208 (2012).

[10] David Raventós, Tobias Graß, Maciej Lewenstein, and Bruno Juliá-Díaz, "Cold bosons in optical lattices: a tutorial for exact diagonalization," J. Phys. B: At. Mol. Opt. **50**, 113001 (2017).

[11] Mitsuaki Kawamura, Kazuyoshi Yoshimi, Takahiro Misawa, Youhei Yamaji, Synge Todo, and Naoki Kawashima, "Quantum lattice model solver $H\Phi$," Comput. Phys. Commun. **217**, 180 (2017).

[12] Phillip Weinberg and Marin Bukov, "Quspin: a python package for dynamics and exact diagonalisation of quantum many body systems part i: spin chains," SciPost Phys. **2**, 003 (2017).

[13] Phillip Weinberg and Marin Bukov, "Quspin: a python package for dynamics and exact diagonalisation of quantum many body systems. part ii: bosons, fermions and higher spins," SciPost Phys. **7**, 20 (2019).

[14] Tom Westerhout, "lattice-symmetries: A package for working with quantum many-body bases," (2021), arXiv:2104.04011 [cond-mat.str-el].

[15] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, "Monte Carlo calculations of coupled boson-fermion systems. I," Phys. Rev. D **24**, 2278 (1981).

[16] Anders W. Sandvik and Juhani Kurkijärvi, "Quantum monte carlo simulation method for spin systems," Phys. Rev. B **43**, 5950–5961 (1991).

[17] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, "Exact, complete, and universal continuous-time worldline monte carlo approach to the statistics of discrete quantum systems," J. Exp. Theor. Phys. **87**, 310 (1998).

[18] E. Kozik, K. Van Houcke, E. Gull, L. Pollet, N. Prokof'ev, B. Svistunov, and M. Troyer, "Diagrammatic Monte Carlo for correlated fermions," Europhys. Lett. **90**, 10004 (2010).

[19] Steven R. White, "Density matrix formulation for quantum renormalization groups," Phys. Rev. Lett. **69**, 2863 (1992).

[20] Steven R. White, "Density-matrix algorithms for quantum renormalization groups," Phys. Rev. B **48**, 10345 (1993).

[21] U. Schollwöck, "The density-matrix renormalization group," Rev. Mod. Phys. **77**, 259 (2005).

[22] Ulrich Schollwöck, "The density-matrix renormalization group in the age of matrix product states," Annals of Physics **326**, 96 (2011), january 2011 Special Issue.

[23] T. D. Kühner and H. Monien, "Phases of the one-dimensional Bose-Hubbard model," Phys. Rev. B **58**, R14741 (1998).

[24] S. Rapsch, U. Schollwöck, and W. Zwerger, "Density matrix renormalization group for disordered bosons in one dimension," Europhys. Lett. **46**, 559 (1999).

[25] Till D. Kühner, Steven R. White, and H. Monien, "One-dimensional Bose-Hubbard model with nearest-neighbor interaction," Phys. Rev. B **61**, 12474 (2000).

[26] B. Capogrosso-Sansone, N. V. Prokof'ev, and B. V. Svistunov, "Phase diagram and thermodynamics of the three-dimensional Bose-Hubbard model," Phys. Rev. B **75**, 134302 (2007).

[27] J. P. F. LeBlanc, Andrey E. Antipov, Federico Becca, Ireneusz W. Bulik, Garnet Kin-Lic Chan, Chia-Min Chung, Youjin Deng, Michel Ferrero, Thomas M. Henderson, Carlos A. Jiménez-Hoyos, E. Kozik, Xuan-Wen Liu, Andrew J. Millis, N. V. Prokof'ev, Mingpu Qin, Gustavo E. Scuseria, Hao Shi, B. V. Svistunov, Luca F. Tocchio, I. S. Tupitsyn, Steven R. White, Shiwei Zhang, Bo-Xiao Zheng, Zhenyue Zhu, and Emanuel Gull and, "Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms," Phys. Rev. X **5**, 041041 (2015).

[28] Thomas Schäfer, Nils Wentzell, Fedor Šimkovic, Yuan-Yao He, Cornelia Hille, Marcel Klett, Christian J. Eckhardt, Behnam Arzhang, Viktor Harkov, François-Marie Le Régent, Alfred Kirsch, Yan Wang, Aaram J. Kim, Evgeny Kozik, Evgeny A. Stepanov, Anna Kauch, Sabine Andergassen, Philipp Hansmann, Daniel Rohe, Yuri M. Vilk, James P. F. LeBlanc, Shiwei Zhang, A.-M. S. Tremblay, Michel Ferrero, Olivier Parcollet, and Antoine Georges, "Tracking the footprints of spin fluctuations: A multimethod, multimessenger study of the two-dimensional hubbard model," Phys. Rev. X **11**, 011058 (2021).

[29] Mingpu Qin, Thomas Schäfer, Sabine Andergassen, Philippe Corboz, and Emanuel Gull, "The Hubbard model: A computational perspective," arXiv:2104.00064 (2021).

[30] Pierre Nataf and Frédéric Mila, "Exact Diagonalization of Heisenberg SU(N) models," Phys. Rev. Lett. **113**, 127204 (2014).

[31] D. Sundholm and T. Vänskä, "A configuration interaction approach to bosonic systems," J. Phys. B: At. Mol. Opt. **37**, 2933 (2004).

[32] Alexej I. Streltsov, Ofir E. Alon, and Lorenz S. Cederbaum, "General mapping for bosonic and fermionic operators in Fock space," Phys. Rev. A **81**, 022124 (2010).

[33] (2021), All code, scripts and data used in this work are included in a GitHub repository: `https://github.com/DelMaestroGroup/papers-code-UnaryBosonicBasis`

doi:10.5281/zenodo.5509930.

[34] B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler, A. Hehn, R. Igarashi, S. V. Isakov, D. Koop, P. N. Ma, P. Mates, H. Matsuo, O. Parcollet, G. Pawłowski, J. D. Picon, L. Pollet, E. Santos, V. W. Scarola, U. Schollwöck, C. Silva, B. Surer, S. Todo, S. Trebst, M. Troyer, M. L. Wall, P. Werner, and S. Wessel, "The ALPS project release 2.0: open source software for strongly correlated systems," J. Stat. Mech.: Theor. Exp. **2011**, 05001 (2011).

[35] Piotr Sierant, Dominique Delande, and Jakub Zakrzewski, "Many-body localization due to random interactions," Phys. Rev. A **95**, 021601 (2017).

[36] Miroslav Hopjan and Fabian Heidrich-Meisner, "Many-body localization from a one-particle perspective in the disordered one-dimensional Bose-Hubbard model," Phys. Rev. A **101**, 063617 (2020).

[37] (2020), See Supplemental Material for information on the hardware used for benchmarking and more details on the application of lattice symmetries.

[38] L_α is the minimum number of cyclic shifts which maps a given occupation vector onto itself. Except for an exponentially small fraction of states, $L_\alpha = L$ (see supplemental material [37]).

[39] Juan Carrasquilla, Salvatore R. Manmana, and Marcos Rigol, "Scaling of the gap, fidelity susceptibility, and Bloch oscillations across the superfluid-to-Mott-insulator transition in the one-dimensional Bose-Hubbard model," Phys. Rev. A **87**, 043606 (2013).

[40] G. Boéris, L. Gori, M. D. Hoogerland, A. Kumar, E. Lucioni, L. Tanzi, M. Inguscio, T. Giamarchi, C. D'Errico, G. Carleo, G. Modugno, and L. Sanchez-Palencia, "Mott transition for strongly interacting one-dimensional bosons in a shallow periodic potential," Phys. Rev. A **93**, 011601 (2016).

[41] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, "Quantum entanglement," Rev. Mod. Phys. **81**, 865 (2009).

[42] Masudul Haque, Oleksandr Zozulya, and Kareljan Schoutens, "Entanglement Entropy in Fermionic Laughlin States," Phys. Rev. Lett. **98**, 060401 (2007).

[43] O S Zozulya, M Haque, K Schoutens, and E H Rezayi, "Bipartite entanglement entropy in fractional quantum Hall states," Phys. Rev. B **76**, 125310 (2007).

[44] O. Zozulya, Masudul Haque, and K. Schoutens, "Particle partitioning entanglement in itinerant many-particle systems," Phys. Rev. A **78**, 042326 (2008).

[45] Masudul Haque, O. S. Zozulya, and K. Schoutens, "Entanglement between particle partitions in itinerant many-particle states," J. Phys. A: Math. Theor. **42**, 504012 (2009).

[46] Zhao Liu and Heng Fan, "Particle entanglement in rotating gases," Phys. Rev. A **81**, 062302 (2010).

[47] C M Herdman, P N Roy, R G Melko, and A Del Maestro, "Particle entanglement in continuum many-body systems via quantum Monte Carlo," Phys. Rev. B **89**, 140501 (2014).

[48] C. M. Herdman, Stephen Inglis, P. N. Roy, R. G. Melko, and A. Del Maestro, "Path-integral Monte Carlo method for Rényi entanglement entropies," Phys. Rev. E **90**, 013308 (2014).

[49] C. M. Herdman and A. Del Maestro, "Particle partition entanglement of bosonic Luttinger liquids," Phys. Rev. B **91**, 184507 (2015).

[50] Lukas Rammelmüller, William J. Porter, Jens Braun, and Joaqun E. Drut, "Evolution from few- to many-body physics in one-dimensional Fermi systems: One- and two-body density matrices and particle-partition entanglement," Phys. Rev. A **96**, 033635 (2017).

[51] Hatem Barghathi, Emanuel Casiano-Diaz, and Adrian Del Maestro, "Particle partition entanglement of one dimensional spinless fermions," J. Stat. Mech.: Theor. Exp. **2017**, 083108 (2017).

[52] Yuki Kurashige, Jakub Chalupský, Tran Nguyen Lan, and Takeshi Yanai, "Complete active space second-order perturbation theory with cumulant approximation for extended active-space wavefunction from density matrix renormalization group," J. Chem. Phys. **141**, 174111 (2014).

[53] Sara Murciano, Paola Ruggiero, and Pasquale Calabrese, "Symmetry resolved entanglement in two-dimensional systems via dimensional reduction," J. Stat. Mech.: Theor. Exp. **2020**, 083102 (2020).

[54] Mao Tian Tan and Shinsei Ryu, "Particle number fluctuations, Rényi entropy, and symmetry-resolved entanglement entropy in a two-dimensional Fermi gas from multidimensional bosonization," Phys. Rev. B **101**, 235169 (2020).

[55] Luca Capizzi, Paola Ruggiero, and Pasquale Calabrese, "Symmetry resolved entanglement entropy of excited states in a CFT," J. Stat. Mech.: Theor. Exp. **2020**, 073101 (2020).

[56] Shachar Fraenkel and Moshe Goldstein, "Symmetry resolved entanglement: exact results in 1d and beyond," J. Stat. Mech.: Theor. Exp **2020**, 033106 (2020).

[57] Noa Feldman and Moshe Goldstein, "Dynamics of charge-resolved entanglement after a local quench," Phys. Rev. B **100**, 235146 (2019).

[58] Hatem Barghathi, Emanuel Casiano-Diaz, and Adrian Del Maestro, "Operationally accessible entanglement of one-dimensional spinless fermions," Phys. Rev. A **100**, 022324 (2019).

[59] Riccarda Bonsignori, Paola Ruggiero, and Pasquale Calabrese, "Symmetry resolved entanglement in free fermionic systems," J. Phys. A: Math. Theor. **52**, 475302 (2019).

[60] Moshe Goldstein and Eran Sela, "Symmetry-resolved entanglement in many-body systems," Phys. Rev. Lett. **120**, 200602 (2018).

[61] M. Kiefer-Emmanouilidis, R. Unanyan, J. Sirker, and M. Fleischhauer, "Bounds on the entanglement entropy by the number entropy in non-interacting fermionic systems," SciPost Phys. **8**, 83 (2020).

[62] Sara Murciano, Giuseppe Di Giulio, and Pasquale Calabrese, "Symmetry resolved entanglement in gapped integrable systems: a corner transfer matrix approach," SciPost Phys. **8**, 46 (2020).

[63] Sara Murciano, Giuseppe Di Giulio, and Pasquale Calabrese, "Entanglement and symmetry resolution in two dimensional free quantum field theories," J. High Energy Phys. **2020**, 73 (2020).

[64] F. Benatti, R. Floreanini, F. Franchini, and U. Marzolino, "Entanglement in indistinguishable particle systems," Physics Reports **878**, 1–27 (2020), entanglement in indistinguishable particle systems.

[65] Xhek Turkeshi, Paola Ruggiero, Vincenzo Alba, and Pasquale Calabrese, "Entanglement equipartition in crit-

ical random spin chains," Phys. Rev. B **102**, 014455 (2020).
[66] Dana Faiez and Dominik Šafránek, "How much entanglement can be created in a closed system," Phys. Rev. B **101**, 060401 (2020).
[67] Dávid X. Horváth and Pasquale Calabrese, "Symmetry resolved entanglement in integrable field theories via form factor bootstrap," J. High Energy Phys. **2020**, 131 (2020).
[68] Riccarda Bonsignori and Pasquale Calabrese, "Boundary effects on symmetry resolved entanglement," J. Phys. A: Math. Theor. **54**, 015005 (2020).
[69] Caroline de Groot, David T Stephen, Andras Molnar, and Norbert Schuch, "Inaccessible entanglement in symmetry protected topological phases," J Phys. A: Math. Theor. **53**, 335302 (2020).
[70] Suting Zhao, Christian Northe, and René Meyer, "Symmetry-resolved entanglement in AdS_3/CFT_2 coupled to $U(1)$ Chern-Simons Theory," (2021), arXiv:2012.11274 [hep-th].
[71] Shachar Fraenkel and Moshe Goldstein, "Entanglement measures in a nonequilibrium steady state: Exact results in one dimension," (2021), arXiv:2105.00740 [quant-ph].
[72] Dávid X. Horváth, Luca Capizzi, and Pasquale Calabrese, "U(1) symmetry resolved entanglement in free 1+1 dimensional field theories via form factor bootstrap," J. High Energy Phys. **2021**, 197 (2021).
[73] Gilles Parez, Riccarda Bonsignori, and Pasquale Calabrese, "Quasiparticle dynamics of symmetry-resolved entanglement after a quench: Examples of conformal field theories and free fermions," Phys. Rev. B **103**, L041104 (2021).
[74] Luca Capizzi and Pasquale Calabrese, "Symmetry resolved relative entropies and distances in conformal field theory," (2021), arXiv:2105.08596 [hep-th].
[75] Sara Murciano, Riccarda Bonsignori, and Pasquale Calabrese, "Symmetry decomposition of negativity of massless free fermions," SciPost Phys. **10**, 111 (2021).
[76] Benoit Estienne, Yacine Ikhlef, and Alexi Morin-Duchesne, "Finite-size corrections in critical symmetry-resolved entanglement," SciPost Phys. **10**, 54 (2021).
[77] David X. Horvath, Pasquale Calabrese, and Olalla A. Castro-Alvaredo, "Branch point twist field form factors in the sine-Gordon Model II: Composite twist fields and symmetry resolved entanglement," (2021), arXiv:2105.13982 [hep-th].
[78] Hatem Barghathi, C. M. Herdman, and Adrian Del Maestro, "Rényi generalization of the accessible entanglement entropy," Phys. Rev. Lett. **121**, 150501 (2018).
[79] R. G. Melko, C. M. Herdman, D. Iouchtchenko, P. N. Roy, and A. Del Maestro, "Entangling qubit registers via many-body states of ultracold atoms," Phys. Rev. A **93**, 042336 (2016).
[80] Hui-Huang Chen, "Symmetry decomposition of relative entropies in conformal field theory," J. High Energy Phys. **2021**, 84 (2021).
[81] Rajibul Islam, Ruichao Ma, Philipp M. Preiss, M. Eric Tai, Alexander Lukin, Matthew Rispoli, and Markus Greiner, "Measuring entanglement entropy in a quantum many-body system," Nature **528**, 77 (2015).
[82] Alexander Lukin, Matthew Rispoli, Robert Schittko, M. Eric Tai, Adam M. Kaufman, Soonwon Choi, Vedika Khemani, Julian Léonard, and Markus Greiner, "Probing entanglement in a many-body-localized system," Science **364**, 256 (2019).
[83] Stephen D. Bartlett and H. M. Wiseman, "Entanglement Constrained by Superselection Rules," Phys. Rev. Lett. **91**, 097903 (2003).
[84] H. M. Wiseman and John A. Vaccaro, "Entanglement of indistinguishable particles shared between two parties," Phys. Rev. Lett. **91**, 097902 (2003).
[85] Howard M. Wiseman, Stephen D. Bartlett, and John A. Vaccaro, "Ferreting out the fluffy bunnies: Entanglement constrained by generalized superselection rules," in *Laser Spectroscopy* (World Scientific, Singapore, 2004) pp. 307–314.
[86] John A. Vaccaro, Fabio Anselmi, and Howard M. Wiseman, "Entanglement of identical particles and reference phase uncertainty," Int. J. Quantum. Inform. **01**, 427–441 (2003).
[87] N. Schuch, F. Verstraete, and J. I. Cirac, "Nonlocal resources in the presence of superselection rules," Phys. Rev. Lett. **92**, 087904 (2004).
[88] Jacob Dunningham, Alexander Rau, and Keith Burnett, "From pedigree cats to fluffy-bunnies," Science **307**, 872–875 (2005).
[89] M. Cramer, M. B. Plenio, and H. Wunderlich, "Measuring entanglement in condensed matter systems," Phys. Rev. Lett. **106**, 020401 (2011).
[90] I. Klich and L. S. Levitov, "Scaling of entanglement entropy and superselection rules," (2008), arXiv:0812.0006.
[91] Andreas M. Läuchli, "Operator content of real-space entanglement spectra at conformal critical points," (2013), arXiv:1303.0741 [cond-mat.stat-mech].
[92] Davide Rossini and Rosario Fazio, "Phase diagram of the extended Bose–Hubbard model," New J. Phys. **14**, 065012 (2012).
[93] Masaki Tezuka, Ryotaro Arita, and Hideo Aoki, "Phase diagram for the one-dimensional hubbard-holstein model: A density-matrix renormalization group study," Phys. Rev. B **76**, 155114 (2007).
[94] Matthew B. Hastings, Iván González, Ann B. Kallin, and Roger G. Melko, "Measuring Renyi Entanglement Entropy in Quantum Monte Carlo Simulations," Phys. Rev. Lett. **104**, 157201 (2010).