Supplemental Material for: "Realizing Exactly Solvable SU(N) Magnets with Thermal Atoms"

Michael E. Beverland,[1] Gorjan Alagic,[2] Michael J. Martin,[1]
Andrew P. Koller,[3] Ana M. Rey,[3] and Alexey V. Gorshkov[4]

[1]*Institute for Quantum Information & Matter, California Institute of Technology, Pasadena, CA 91125, USA*
[2]*Department of Mathematical Sciences, University of Copenhagen*
[3]*JILA, NIST, and Department of Physics, University of Colorado Boulder, CO 80309*
[4]*Joint Quantum Institute and Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, MD 20742*

(Dated: March 28, 2016)

S1. EIGENSTATES AND ENERGIES OF THE HAMILTONIAN

In this Section, we present the details behind the derivation of the eigenstates and the energies of the Hamiltonian given in Eq. (1) of the main text. In particular, we compute the degeneracy of the ground state for $U > 0$ and $U < 0$. As in the main text, we use n and N to mean the number of atoms, and number of nuclear spin states per atom respectively.

Define $\hat{U}(\hat{V}, \sigma)$ which permutes occupied orbitals by $\sigma \in S_n$ and implements the spin rotation $\hat{V} \in SU(N)$:

$$\hat{U}(\hat{V}, \sigma)\ |p_1\rangle|p_2\rangle...|p_n\rangle \quad \equiv \quad \hat{V}|p_{\sigma^{-1}(1)}\rangle\hat{V}|p_{\sigma^{-1}(2)}\rangle...\hat{V}|p_{\sigma^{-1}(n)}\rangle. \tag{S1}$$

These unitaries (for all $\hat{V} \in SU(N)$ and $\sigma \in S_n$) form a well-understood representation of the group $G = S_n \times SU(N)$. Each such unitary commutes with $\hat{H} = -U \sum_{j \neq k} \hat{s}_{jk}$, where for clarity we dropped all constants from Eq. (1). Irreps of $SU(N)$ and S_n are uniquely labeled by Young diagrams $\vec{\mu}$ and $\vec{\nu}$, respectively, which satisfy different conditions: $\vec{\mu} = (\mu_1, \mu_2, ..\mu_N)$, whereas $\sum_i \nu_i = n$. Each irrep of the product group $G = S_n \times SU(N)$ is the tensor product of an irrep of $SU(N)$ and an irrep of S_n and is therefore uniquely labeled by a pair $(\vec{\mu}, \vec{\nu})$. A consequence of Schur-Weyl duality is that representation (S1) block-diagonalizes into exactly one copy of each irrep of G satisfying $\vec{\mu} = \vec{\nu}$, and no other irreps [S1, S2]. Therefore for each Young diagram $\vec{\lambda} = (\lambda_1, \lambda_2, .., \lambda_N)$ such that $\sum_i \lambda_i = n$, there is a subspace of constant energy $E(\vec{\lambda})$. One can form an unnormalized projection operator $\hat{\Pi}_{L(\vec{\lambda})}$ into the $\vec{\lambda}$ subspace [S2]:

$$\hat{\Pi}_{L(\vec{\lambda})} = \sum_{\substack{c \in \mathrm{col}(T) \\ r \in \mathrm{row}(T)}} \mathrm{sgn}(c)\ \ \hat{U}(\hat{I}, c)\ \hat{U}(\hat{I}, r). \tag{S2}$$

Here, $L(\vec{\lambda})$ is the labeling of boxes in the Young diagram $\vec{\lambda}$ from 1 to n as shown in Fig. 2(b) in the main text, and $\mathrm{row}(L)$ ($\mathrm{col}(L)$) is the group of all permutations of the numbers 1 to n that preserve the contents of rows (columns) of $L(\vec{\lambda})$. Applying $\hat{\Pi}_{L(\vec{\lambda})}$ to any state that it does not annihilate returns an eigenstate of energy $E(\vec{\lambda})$. For concreteness we use $|T\rangle \equiv |1, 2, ..., \gamma_1\rangle\ |1, 2, ..., \gamma_2\rangle...\ |1, 2, ..., \gamma_{\lambda_1}\rangle$, where we also define $\vec{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_{\lambda_1})$ as the column heights of the Young diagram $\vec{\lambda}$. For each $\vec{\lambda}$ we obtain an explicit eigenstate: $|\vec{\lambda}\rangle = \hat{\Pi}_{L(\vec{\lambda})}|T\rangle$ as in Eq. (2) of the main text. Now we describe how to obtain the eigenvalue $E(\vec{\lambda})$ such that:

$$\hat{H}|\vec{\lambda}\rangle = E(\vec{\lambda})|\vec{\lambda}\rangle. \tag{S3}$$

Premultiplying by $\langle T|$ we obtain: $E(\vec{\lambda}) = \langle T|\hat{H}|\vec{\lambda}\rangle = -U \sum_{j \neq k} \langle T|\hat{s}_{jk}|\vec{\lambda}\rangle$, noting that $\langle T|\vec{\lambda}\rangle = 1$. For j, k in the same column of the labeled Young diagram $L(\vec{\lambda})$, we know that $\hat{s}_{jk}|\vec{\lambda}\rangle = -|\vec{\lambda}\rangle$. Similarly for j, k in the same row of $L(\vec{\lambda})$ we have $\langle T|\hat{s}_{jk} = \langle T|$. Thus pairs (j, k) in columns contribute -1 to $E(\vec{\lambda})$ and pairs (j, k) in rows contribute $+1$. The number of such pairs can be counted, hence:

$$E(\vec{\lambda})/(-U) = \sum_{i=1}^{N} \binom{\lambda_i}{2} - \sum_{j=1}^{\lambda_1} \binom{\gamma_j}{2} + \sum_{\{j \neq k\}_{\mathrm{diagonal}}} \langle T|\hat{s}_{jk}|\vec{\lambda}\rangle, \tag{S4}$$

The swap $\hat{s}_{jk}$, where j and k are neither in same column nor in same row in $L(\vec{\lambda})$, can always be written as $\hat{s}_{jk} = \hat{s}_{jm}\hat{s}_{km}\hat{s}_{jm} = \hat{s}_{km}\hat{s}_{jm}\hat{s}_{km}$, where m is chosen such that (j, m) and (k, m) lie in a row and a column of $L(\vec{\lambda})$, respectively

(it suffices to consider the case $j > k$). Therefore, $\langle T|\hat{s}_{jk}|\vec{\lambda}\rangle = \langle T|\hat{s}_{km}\hat{s}_{jm}|\vec{\lambda}\rangle = -\langle T|\hat{s}_{km}\hat{s}_{jm}|\vec{\lambda}\rangle = 0$, implying $E(\vec{\lambda})/(-U) = \sum_{i=1}^{N}\binom{\lambda_i}{2} - \sum_{j=1}^{\lambda_1}\binom{\gamma_j}{2}$.

The dimensions of each block can be calculated using the standard hook-length formulae [S3] for any given Young diagram $\vec{\lambda}$. In particular, the ground-state spaces for $U > 0$ (ferromagnetic interaction) and $U < 0$ (antiferromagnetic interaction) are $\vec{\lambda}_F = (n, 0, 0, ..., 0)$ and $\vec{\lambda}_{AF} = (n/N, n/N, ..., n/N)$ and have dimensions D_F and D_{AF}, respectively:

$$D_F = \frac{(n+N-1)!}{n!\ (N-1)!}, \qquad D_{AF} = \frac{n!}{[(n/N)!]^N}\prod_{i=1}^{N-1}\frac{i!}{[n/N+i]}. \tag{S5}$$

S2. DERIVATION OF SPIN-DIFFUSION DYNAMICS

In this Section, we present the derivation of the spin-diffusion dynamics, first for $N = 2$ [i.e. Eq. (4) in the main text] and then for general N.

We are concerned with observable $\hat{Q} = \sum_{j=1}^{m_1}|1\rangle_j\langle 1|_j$. In this section, we use the notation that for any operator $\hat{A}$, $A(t) \equiv \langle\psi(0)|e^{i\hat{H}t}\hat{A}e^{-i\hat{H}t}|\psi(0)\rangle$, where $|\psi(0)\rangle = |1\rangle^{\otimes m_1}|2\rangle^{\otimes m_2}...|N\rangle^{\otimes m_N}$. As most readers are assumed to be familiar with spin-1/2 systems, we outline the $N = 2$ case first before covering the general case more abstractly.

For $N = 2$, we can choose the angular momentum (Dicke) basis to span the Hilbert space: $|S, S_z, k\rangle$, which diagonalizes the Hamiltonian: $\hat{H}|S, S_z, k\rangle = -US(S+1)|S, S_z, k\rangle$ (dropping a constant energy). The initial state is $|\psi(0)\rangle = |\uparrow\rangle^{\otimes m}|\downarrow\rangle^{\otimes n-m}$ where we used $|\uparrow\rangle$ and $|\downarrow\rangle$ in place of $|1\rangle$ and $|2\rangle$. This state can be understood as a tensor product of two Dicke states on subsets of spins: $|\psi(0)\rangle = |m/2, m/2\rangle \otimes |(n-m)/2, -(n-m)/2\rangle$, where there is no need for a k quantum number since states with $|S_z| = S$ have no additional degeneracy. The tensor product of two angular momentum states can be written as a sum of "total" angular momentum states: $|\psi(0)\rangle = \sum_S C(S)|S, S_z = m - n/2, \alpha(S)\rangle$, where $C(S)$ is a Clebsch-Gordan coefficient, and $\alpha(S)$ represents the fact that $|S, S_z = m - n/2, \alpha(S)\rangle$ is some specific linear combination of Dicke states with the same S and S_z, but different k's. Hence, $Q(t) = \sum_{S,S'} C(S')^* C(S) e^{iUt\left[S(S+1)-S'(S'+1)\right]}\langle S', S_z, \alpha(S')|\hat{Q}|S, S_z, \alpha(S)\rangle$. Note that $\hat{Q} = m\hat{I} + \hat{S}^z_m$ with $\vec{S}_m = \sum_{j=1}^{m}\vec{S}_j$, and $\hat{S}^z_m$ is the 0-component of the $(S = 1)$-spherical tensor $\hat{\mathbb{T}} \equiv \{\hat{S}^{-1}_m, \hat{S}^z_m, \hat{S}^{+1}_m\}$, with $\hat{S}^{\pm 1}_m = \mp(\hat{S}^x_m \pm i\hat{S}^y_m)/\sqrt{2}$. We first apply the Wigner-Eckart theorem to write the matrix element in terms of the reduced matrix element and a Clebsch-Gordan coefficient. Then, since $\hat{\mathbb{T}} \equiv \hat{\mathbb{T}}_m \otimes \hat{I}$ acts only on the first m spins, we rewrite [S4, S5] the reduced matrix element on the full system in terms of one on the first m spins and a recoupling coefficient:

$$\langle S', S'_z, \alpha(S')|\hat{Q}|S, S_z, \alpha(S)\rangle = m\delta_{S,S'} + \langle m/2||\hat{\mathbb{T}}_L||m/2\rangle \begin{Bmatrix} 1 & m/2 & m/2 \\ (n-m)/2 & S' & S \end{Bmatrix} (\langle 1, 0| \otimes \langle S, S_z|)\ \ |S', S'_z\rangle, \tag{S6}$$

where $(\langle 1, 0| \otimes \langle S, S_z|)\ \ |S', S'_z\rangle$ is a Clebsch-Gordan coefficient and $\langle m/2||\hat{\mathbb{T}}_L||m/2\rangle$ is the reduced matrix element of $\hat{\mathbb{T}}_L$ on the $S = m/2$ state of the first m spins. The recoupling coefficient $\begin{Bmatrix} S_A & S_B & S_{AB} \\ S_C & S & S_{BC} \end{Bmatrix} \equiv \langle S, S_z, (S_{AB}, S_C)|S, S_z, (S_A, S_{BC})\rangle$ is the overlap between two states of given S and S_z formed from the tensor product of three subsystems with S_A, S_B and S_C in two different ways: by combining A and B to form S_{AB} first, and by combining B and C to form S_{BC} first. Substitution of the Clebsch-Gordan and recoupling coefficients into the matrix element gives Eq. (4) in the main text.

Now we proceed with the calculation for arbitrary N, simplifying our notation by dropping hats and vectors. The initial state [see Fig. S1(a)] can be written as a direct product of spin-symmetric states $|\psi(0)\rangle = \otimes_{j=1}^{m_1}|1\rangle \otimes_{j=1}^{m_2}|2\rangle... \otimes_{j=1}^{m_N}|N\rangle = |\kappa_1, a_1\rangle|\kappa_2, a_2\rangle...|\kappa_N, a_N\rangle$, where a_i labels the particular state in the $\kappa_i \equiv (m_i, 0, ...0)$ irrep which corresponds to $|i\rangle^{\otimes m_i}$. The product of $\kappa = (m, 0, ..., 0)$ with any irrep λ' has no multiplicity [S1]: $|\kappa, a\rangle|\lambda', a'\rangle = \sum_{\lambda'', a''} C(\lambda'', a'')|\lambda'', a''\rangle$, where each irrep λ'' appears at most once and $C(\lambda'', a'') \equiv \langle\lambda'', a''|\left(|\kappa, a\rangle|\kappa', a'\rangle\right)$ is a Clebsch-Gordan coefficient. Applying this iteratively, starting from the right, $|\psi(0)\rangle = \sum_{\lambda,a,\alpha} C(\lambda, a, \alpha)|\lambda, a, \alpha\rangle$, where α labels the set of intermediate irreps, $C(\lambda, a, \alpha)$ can be expressed in terms of Clebsch-Gordan coefficients, and $|\lambda, a, \alpha\rangle$ are orthogonal eigenstates: $H|\lambda, a, \alpha\rangle = E(\lambda)|\lambda, a, \alpha\rangle$. Note: $a \in 1, 2, ..., \dim[\lambda_{SU(N)}]$ labels a basis state within the λ-irrep of $SU(N)$, and each α labels one distinct copy (out of $\dim[\lambda_{S_n}]$ copies) of the λ-irrep of $SU(N)$ in the Hilbert space $\mathcal{H} = (\mathbb{C}_N)^{\otimes n}$ (all copies of irrep λ of $SU(N)$ in $\mathcal{H}$ sit inside a single copy of irrep λ of $S_n \times SU(N)$). Therefore: $Q(t) = \sum_{\lambda,\lambda',a,a',\alpha} C^*(\lambda', a', \alpha)C(\lambda, a, \alpha)e^{i[E(\lambda')-E(\lambda)]t}\langle\lambda', a', \alpha|Q|\lambda, a, \alpha\rangle$, where we set $\alpha' = \alpha$ since Q has support only on the first m_1 spins. We now outline tools to determine the matrix element $\langle\lambda', a', \alpha|Q|\lambda, a, \alpha\rangle$.

The states $|\lambda, a, \alpha\rangle$ transform according to matrix irrep D^λ of $SU(N)$: $V^{\otimes n}|\lambda, a, \alpha\rangle = \sum_{a'} D^\lambda_{a'a}(V)|\lambda, a', \alpha\rangle$. For each N, there is a set of single-spin operators which generate $SU(N)$: $\tau^{\text{adj}} \equiv \{t_1, t_2, ..., t_{N^2-1}\}$ which transform

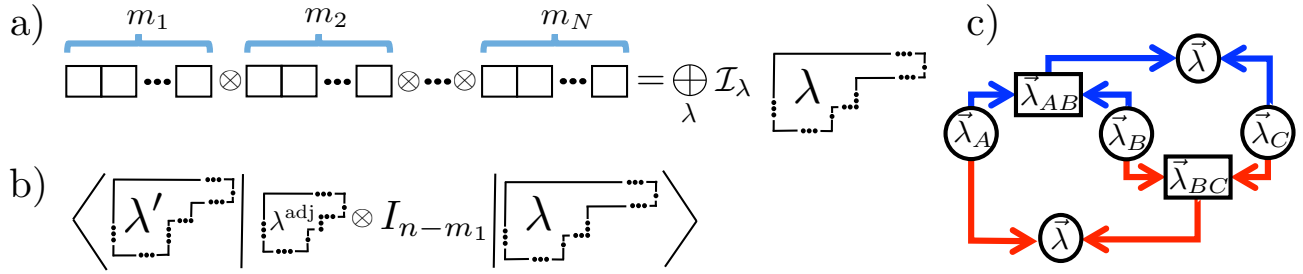

FIG. S1: (a) Initial state $|\psi(0)\rangle$ can be written in terms of energy eigenstates: $|\psi(0)\rangle = |11...1\rangle|22...2\rangle...|NN...N\rangle = \sum_{\lambda,a,\alpha} C(\lambda,a,\alpha)|\lambda,a,\alpha\rangle$. (b) Key simplifications arise in the matrix element $\langle\lambda',a',\alpha|Q|\lambda,a,\alpha\rangle$ (which is used to calculate $Q(t)$) since: $\hat{Q}$ is a component of a "spherical tensor" for $SU(N)$ (allowing us to make use of the Wigner-Eckart theorem) and has support only on the first m_1 sites. (c) The recoupling coefficient is defined by taking the direct product of three irreps A, B and C, and finding the overlap between two copies of the same irrep found in two ways: by combining A and B first (top), and by combining B and C first (bottom).

according to $D^{\rm adj}$ (the *adjoint irrep* $\lambda_{\rm adj}$): $V^{\otimes n} t_a V^{\dagger\otimes n} = \sum_{a'} D^{\rm adj}_{a'a}(V) t_{a'}$. The set $\{t_1, t_2, ..., t_{N^2-1}, \hat{I}\}$ forms a basis for $N \times N$ Hermitian matrices: therefore, any single-atom spin observable can be written as $\hat{q} = c_0\hat{I} + \sum_a c_a t_a$ for some real constants c_a. Therefore $\langle\lambda',a',\alpha|Q|\lambda,a,\alpha\rangle = c_0 + \sum_{a''} c_{a''}\langle\lambda',a',\alpha|T^{\rm adj}_{a''}|\lambda,a,\alpha\rangle$, where $T^{\rm adj}_a = \sum_{j=1}^{m_1} t_{a\,j}$ and $Q = \sum_{j=1}^{m_1} |1\rangle_j\langle 1|_j \equiv \sum_{j=1}^{m_1} \hat{q}_j$. We now prove a generalization of Eq. (S6) to determine the matrix element $\langle\lambda',a',\alpha'|T^{\rm adj}_{a''}|\lambda,a,\alpha\rangle$ [see Fig. S1(b)]. We will need the Wigner-Eckart theorem and recoupling coefficients for $SU(N)$:

$$\langle\lambda',a',\alpha'|T^{\lambda''}_{a''}|\lambda,a,\alpha\rangle = \sum_{\mathcal{I}} \left(\langle\lambda',a',\mathcal{I}||\lambda'',a''\rangle|\lambda,a\rangle\right) \langle\lambda',\alpha'||T^{\lambda''}||\lambda,\alpha\rangle_{\mathcal{I}}, \tag{S7}$$

$$\left\{\begin{array}{ccc}\lambda_A & \lambda_B & \lambda_{AB}\\ \lambda_C & \lambda & \lambda_{BC}\end{array}\right\}_{\mathcal{I}_{AB},\mathcal{I}_C;\mathcal{I}_{BC},\mathcal{I}_A} \equiv \langle\lambda,a,(\lambda_{AB},\mathcal{I}_{AB},\mathcal{I}_C)|\lambda,a,(\lambda_{BC},\mathcal{I}_{BC},\mathcal{I}_A)\rangle. \tag{S8}$$

Note that multiplicity $\mathcal{I}$ appears in the Wigner Eckart theorem for $N > 2$ [Eq. (S7)], since the tensor product of irreps can include multiple appearances of the same irrep. The recoupling coefficient defined in Eq. (S8) relates two copies of the same irrep λ formed from the tensor product of three irreps: λ_A, λ_B, and λ_C, but combined in different orders [see Fig. S1(c)]. To define notation: λ_A and λ_B are combined to make λ_{AB}, whose different copies are labeled by $\mathcal{I}_{AB}$, while $\mathcal{I}_C$ labels different copies of λ when λ_{AB} is combined with λ_C.

One can decompose $|\lambda,a,\alpha\rangle = \sum_{a_1,a_2} D(a_1,a_2)|\kappa_1,a_1\rangle|\lambda_2,a_2\rangle$, where λ_2 is specified by α, and $D \equiv (\langle\kappa_1,a_1|\langle\lambda_2,a_2|)|\lambda,a,\alpha\rangle$. Substituting into $\langle\lambda',a',\alpha|T^{\rm adj}_{a''}|\lambda,a,\alpha\rangle$ and applying Eq. (S7) to the first m_1 spins:

$$\begin{aligned}\langle\lambda',a',\alpha|T^{\rm adj}_{a''}|\lambda,a,\alpha\rangle &= \langle\kappa_1||\mathbb{T}^{\rm adj}||\kappa_1\rangle \sum_{a_1,a_1',a_2} \left[(\langle\kappa_1,a_1'|\langle\lambda_2,a_2|)|\lambda',a'\rangle\right]^* \left[(\langle\kappa_1,a_1|\langle\lambda_2,a_2|)|\lambda,a\rangle\right] \left[\langle\kappa,a_1'|(|\lambda^{\rm adj},a''\rangle|\kappa_1,a_1\rangle)\right]\\ &= \langle\kappa_1||\mathbb{T}^{\rm adj}||\kappa_1\rangle \sum_{\mathcal{I}_1} \left\{\begin{array}{ccc}\lambda^{\rm adj} & \kappa_1 & \kappa_1\\ \lambda_2 & \lambda' & \lambda\end{array}\right\}^*_{\mathcal{I}_1} \left[(\langle\lambda^{\rm adj},a''|\langle\lambda,a|)|\lambda',a',\mathcal{I}_1\rangle\right].\end{aligned} \tag{S9}$$

The second line represents the generalization of Eq. (S6). To derive Eq. (S9), we return to the abstract scenario of three irreps λ_A, λ_B and λ_C used to define recoupling coefficients in Eq. (S8). First write $|\lambda,a,(\lambda_{AB})\rangle$ as a linear combination of $|\lambda,a,(\lambda_{BC},\mathcal{I}_A)\rangle$ with Eq. (S8) as coefficients in the special case where $\lambda_B = \lambda_{AB} = \kappa$ (allowing us to drop $\mathcal{I}_{AB}$, $\mathcal{I}_C$ and $\mathcal{I}_{BC}$). Rewriting states on both sides as the direct product of states in each of the three subsystems, multiplying by $[\langle\lambda'_{BC},a_{BC}|\left(|\lambda_B,a_B\rangle|\lambda_C,a_C\rangle\right)]$, summing over λ'_{BC}, and using orthogonality gives:

$$\begin{aligned}\sum_{a_{AB},a_B,a_C} &\left[(\langle\lambda_{AB},a_{AB}|\langle\lambda_C,a_C|)\,|\lambda,a\rangle\right] \left[(\langle\lambda_A,a_A|\langle\lambda_B,a_B|)\,|\lambda_{AB},a_{1AB}\rangle\right] \left[\langle\lambda_{BC},a_{BC}|\left(|\lambda_B,a_B\rangle|\lambda_C,a_C\rangle\right)\right] =\\ &\sum_{\mathcal{I}_A} \left\{\begin{array}{ccc}\lambda_A & \kappa & \kappa\\ \lambda_C & \lambda & \lambda_{BC}\end{array}\right\}^*_{\mathcal{I}_A} \left[(\langle\lambda_A,a_A|\langle\lambda_{BC},a_{BC}|)\,|\lambda,a,\mathcal{I}_A\rangle\right].\end{aligned} \tag{S10}$$

Using Eq. (S9), the time evolution $T_a(t) \equiv \langle\psi(0)|\exp(iHt)T_a\exp(-iHt)|\psi(0)\rangle$, and therefore $Q(t)$, is written as

an efficiently computable sum (containing $poly(n)$ terms [S6], each calculated in $poly(n)$ operations):

$$T_a(t) = \langle\kappa_1||\mathbb{T}^{\rm adj}||\kappa_1\rangle \sum_{\lambda_1',a_1',\lambda_1,a_1;\alpha} C^*(\lambda_1',a_1',\alpha)\ C(\lambda_1,a_1,\alpha) e^{(i[E(\lambda_1')-E(\lambda_1)]t)} \times \sum_{\mathcal{I}_1} \left\{ \begin{array}{ccc} \lambda^{\rm adj} & \kappa_1 & \kappa_1 \\ \lambda_2 & \lambda_1' & \lambda_1 \end{array} \right\}_{\mathcal{I}_1} \left[\langle\lambda_1',a_1',\mathcal{I}_1|\left(|\lambda^{\rm adj},j\rangle|\lambda_1,a_1\rangle\right)\right]. \tag{S11}$$

The group-theoretic method presented in this Section was crucial for obtaining the analytical result for $SU(2)$ [Eq. (4) in the main text]. It is also crucial for doing numerics for $SU(N>2)$ for large n. However, for sufficiently small n, such as the one shown in Fig. 3, one can do the $SU(N>2)$ numerics using the following simpler method. One first constructs a complete basis of fully symmetric states for the first m_1 spins, for the next m_2 spins, for the next m_3 spins, etc... Then one combines them into a basis for the full system and keeps only those states that have m_1 1's, m_2 2's, m_3 3's, etc... It is straightforward to evaluate the Hamiltonian in this reduced basis and then numerically exponentiate it to calculate time evolution.

S3. HAMILTONIAN DERIVATION: ATOMS WITH CONTACT INTERACTIONS

In this Section, we derive the Hamiltonian describing identical (bosonic or fermionic) multi-component particles in an infinite square well interacting via s-wave interactions. We then specialize to the case of fermionic alkaline-earth atoms and derive Eq. (5) in the main text.

Contact interactions between two identical multi-component fermionic (bosonic) atoms are described by the Hamiltonian

$$\hat{H}_{int}^{12} = 4\pi\hbar\omega_\perp \delta(x_1-x_2)\otimes\hat{A}, \tag{S12}$$

where the operator $\hat{A}$ only has a physical effect on exchange antisymmetric (symmetric) two-particle internal states because exchange symmetric (antisymmetric) spatial states do not interact. In second quantized form, where $\hat{c}_{jr}^\dagger$ creates an atom in internal state r and orbital $\phi_j(x)$ with non-interacting energy E_j, and $W_{k'j'jk} = (4\pi\hbar\omega_\perp)\int_0^L dx\ \ \phi_{k'}(x)\phi_{j'}(x)\phi_j(x)\phi_k(x)$. The interaction becomes: $\hat{H}_{int} = \sum_{j',k',j,k} W_{k'j'jk} \sum_{r',s',r,s} \langle s',r'|\hat{A}|r,s\rangle\ \hat{c}_{j'r'}^\dagger \hat{c}_{k's'}^\dagger \hat{c}_{jr}\hat{c}_{ks}$. Specializing to the infinite square well of width L, to first order in the interaction, only terms satisfying $(j',k')=(j,k)$ or $(j',k')=(k,j)$ survive. Additionally assuming no multiple occupancies, we obtain $W_{kjjk}=W_{jkjk}=W\equiv(4\pi\hbar\omega_\perp)/L$ for $j\neq k$, and the Hamiltonian becomes:

$$\hat{H} = \sum_{j,r} E_j \hat{c}_{jr}^\dagger \hat{c}_{jr} + W\sum_{j,k}\sum_{r',s',r,s} \langle s',r'|\hat{A}|r,s\rangle \left(\hat{c}_{jr'}^\dagger \hat{c}_{ks'}^\dagger \hat{c}_{jr}\hat{c}_{ks} + \hat{c}_{kr'}^\dagger \hat{c}_{js'}^\dagger \hat{c}_{jr}\hat{c}_{ks}\right). \tag{S13}$$

Now we specialize to the case focused on in this paper. For fermionic alkaline-earth atoms, $\hat{A}$ cannot depend on nuclear spin; therefore $\hat{A} = \left(a_{ee}|e,e\rangle\langle e,e| + a_{gg}|g,g\rangle\langle g,g| + a_{eg}^+|e,g\rangle_+\langle e,g|_+ + a_{eg}^-|e,g\rangle_-\langle e,g|_-\right)\otimes\hat{I}_{Nuclear}$, where $|e,g\rangle_\pm = (|e,g\rangle \pm |g,e\rangle)/\sqrt{2}$ [S7]. Under these conditions, and applying a strong magnetic field (which to first order in perturbation theory prevents exchanges $|ep,gq\rangle \leftrightarrow |eq,gp\rangle$ for $p\neq q$), we obtain Eq. (5) with $U_{1g2g}=U_{2g1g}=U_{gg}\equiv 4\pi\omega_\perp a_{gg}/L$, $U_{1e2e}=U_{2e1e}=U_{ee}\equiv 4\pi\omega_\perp a_{ee}/L$, $U_{1g1e}=U_{2g2e}=4\pi\omega_\perp a_{eg}^-/M$, $U_{1g2e}=U_{2g1e}=2\pi\omega_\perp(a_{eg}^+ + a_{eg}^-)/M$. Recently discovered orbital Feshbach resonances may be used to further tune the values of U_{1g2e} and U_{2g1e} [S8, S9, S10].

S4. EXPERIMENTAL DETAILS

Here we expand upon the experimental considerations section in the main text. The bow tie configuration build-up cavity of attractive magic-wavelength (λ =813 nm) beams shown in Fig. 5 in the main text results in orthogonal standing waves in the x-y plane, whose intensity maxima are spaced by $\simeq 3$ μm, with beam waist of 100 μm at the intersection of the two beams. The build-up cavity will increase the beams' intensity by a factor of ~ 100 with a circulating power of 25 W. The resulting 1D trap sites have $\omega_\perp \simeq 2\pi\times 88$ kHz for the initial loading and cooling phase of the experiment. The (much weaker) longitudinal trapping frequency that results is $\omega_z \simeq 2\pi\times 880$ Hz.

As described in the main text, an additional blue-detuned optical potential at 394 nm, the Sr blue magic wavelength, creates sharp caps on the resulting tubes. This potential is formed by a projected image of a Gaussian beam with waist 30 μm and total power 400 mW screened in the center by a rectangular mask of width $L=10$ μm.

The large $\omega_\perp$ enforces a pseudo one-dimensional system as only the lowest radial energy level will be populated. However, the desired condition that $U = (4\pi a_{gg}\hbar\omega_\perp)/L < 3\hbar^2(\pi/L)^2/M$ is not satisfied with this large $\omega_\perp$. After loading into the hybrid red- and blue-detuned optical potential, we propose to ramp the red-detuned optical lattice potentials adiabatically from the 25 W circulating power to 300 mW, resulting in $\omega_\perp \simeq 2\pi \times 10$ kHz and $\omega_z \simeq 2\pi \times 100$ Hz. The adiabatic nature of the ramp ensures that the x and y degrees of freedom remain frozen.

Imperfections on the mask that creates the flat potential and imperfect edges of the trap from the blue-detuned potential contribute to ΔU. In the following section (Sec. S5), we give an analytic bound that a harmonic perturbation of frequency ω_z small enough that $M\omega_z^2 L^2 < \frac{\hbar^2\pi^2}{ML^2}$ leads to $\Delta U/U < 10^{-2}$. Exact diagonalization of the 1D potential confirms that $\Delta U/U$ is even less sensitive to ω_z: our parameters correspond to $M\omega_z^2 L^2 \approx 750\frac{\hbar^2\pi^2}{2ML^2}$, yet $\Delta U/U$ remains below one percent. The imaging system used to form the potential contributes much more significantly to ΔU. With an imaging point spread function of full width at half maximum (FWHM) of 1 μm with atoms at 1 μK, exact diagonalization results in $\Delta U/U \lesssim 5\%$.

Therefore with these parameters, one obtains $U/\hbar = (4\pi a_{gg}\omega_\perp)/L \approx 2\pi \times 10$ Hz, and should be able to meet all three of the key requirements stated in the main text with $\lesssim 20$ atoms in a single tube. In addition, as the pulses in the GHZ protocol should resolve U, they should have a sufficiently long duration $\gg 0.1$ s. With additional effort, it should be possible to reach a regime of higher U and n while satisfying these requirements. By shaking the trap during preparation with frequencies low enough to depopulate the lowest m energy orbitals, the restrictions on L and $\omega_\perp$ from the requirement that $U = (4\pi a_{gg}\hbar\omega_\perp)/L < 3\hbar^2(\pi/L)^2/M$ is relaxed to $(4\pi a_{gg}\hbar\omega_\perp)/L < [(m+2)^2 - (m+1)^2]\hbar^2(\pi/L)^2/M$. Decreasing the ratio between the spatial imperfections of the potential and L will reduce $\Delta U/U$. For example, reducing the FWHM of the point spread function in our numerical calculations described above from 1 μm to 0.5 μm yields $\Delta U/U < 2\%$. Approaches for creating subwavelength potentials can also be envisioned [S11].

Beyond the three key requirements given in the main text, there are a number of other considerations which we now address. Taking a typical recombination rate constant $K_3 \approx 10^{-28}$ cm^6/s for $n = 20$ particles, it should take approximately 1 second before a single particle is lost. This loss time is 10 times longer than the coherent interaction time $2\pi\hbar/U$, a ratio that is comparable (or even superior) to the ratio of the decoherence time to the spin-spin interaction time in superexchange-based systems [S12, S13]. Tunneling between the tubes is negligible due to the large 3 μm spacing between tubes. The approximate magnitude of p-wave terms involving occupied orbitals j and k is $\pi^2(j^2+k^2)(b_{gg}/L)^2(b_{gg}/a_{gg})\,U$, where b_{gg}^3 is the scattering volume for p-wave interactions. This remains small for $j,k < 300$, taking $b_{gg} \approx 3.9$ nm [S14] for ^{87}Sr. Vector and tensor light shifts [S15] in principle break $SU(N)$ symmetry, but tensor polarizability in our system is negligible, while vector shifts can be avoided with the use of linear polarization [S16]. Specifically, to ensure any breaking of the $SU(N)$ symmetry is far below a level which could affect our proposal, beam circularity of below a few percent should be sufficient. An appropriate choice of linear polarization of the blue-detuned beam will ensure minimal longitudinal field components (and hence minimal circularity) induced by imaging the mask.

S5. ROBUSTNESS TO IMPERFECTIONS

In this Section, we consider deviation from a perfect infinite square-well potential $V(x)$. For simplicity, we consider the case in which all atoms are in the ground electronic state. The interaction Hamiltonian Eq. (1) in the main text becomes: $\hat{H}' = -\sum_{j<k} U_{jk}\hat{s}_{jk}$, where $U_{jk} = (UL/2)\int \phi_j^2(x)\phi_k^2(x)dx$, and $\phi_j(x)$ is a single-particle orbital, which is a sine function in the ideal case. As $\hat{H}'$ is a weighted sum of terms $\hat{s}_{jk}$ and therefore has $SU(N)$ symmetry, it cannot mix states in different $\vec{\lambda}$-subspaces. However as $\hat{H}'$ does not exhibit S_n symmetry, the $\vec{\lambda}$ subspace does not have a single energy - but breaks into $D(\vec{\lambda})$ energy subspaces, $D(\vec{\lambda})$ is the dimension of the $\vec{\lambda}$ irrep of S_n. We write the eigenenergies of $\hat{H}'$ as $E'(\vec{\lambda}, b)$, with b labeling distinct energies.

Provided that the inhomogeneity in U_{jk} is small, i.e. that $|U_{jk} - U| \ll U$, the energy splittings $E'(\vec{\lambda}, b)$ within each $\vec{\lambda}$ subspace will be small compared to energy separations between different $\vec{\lambda}$ subspaces. Exact determination of $E(\vec{\lambda}, b)$ can be carried out by projecting $\hat{H}'$ onto the $\vec{\lambda}$ subspace and solving the resulting matrix equation, which is computationally difficult as the matrices have dimension $O(\exp(n))$. Here we are satisfied with an indication of the magnitude of deviation from the ideal energy eigenvalues. We seek the offset: $\Delta E(\vec{\lambda}) \equiv \frac{1}{D(\vec{\lambda})}\sum_{b=1}^{D(\vec{\lambda})}\left[E'(\vec{\lambda}, b) - E(\vec{\lambda})\right]$ and the variance: $\sigma^2(\vec{\lambda}) \equiv \frac{1}{D(\vec{\lambda})}\sum_{b=1}^{D(\vec{\lambda})}\left[\Delta E(\vec{\lambda}, b) - \Delta E(\vec{\lambda})\right]^2$. Defining $E(\vec{\lambda}_0) = -Un(n-1)/2$, where $\vec{\lambda}_0 = (n, 0, 0, .., 0)$, one

can show that

$$\Delta E(\vec{\lambda}) = -\left(\frac{E(\vec{\lambda})}{E(\vec{\lambda}_0)}\right)\sum_{j<k}(U_{jk}-U). \tag{S14}$$

Note that $\left|\frac{E(\vec{\lambda})}{E(\vec{\lambda}_0)}\right| \leq 1$ for all $\vec{\lambda}$. The main technical lemma used to prove this is that for any operator $\hat{O}$,

$$\sum_{b=1}^{D(\vec{\lambda})}\langle\vec{\lambda},b|\hat{O}|\vec{\lambda},b\rangle = \frac{D(\vec{\lambda})}{n!}\sum_{\sigma\in S_n}\langle\vec{\lambda},b'|\sigma^{-1}\hat{O}\sigma|\vec{\lambda},b'\rangle, \tag{S15}$$

where the latter sum is over all permutations σ in the symmetric group S_n. Modeling U_{jk} as a set of $n(n-1)/2$ independent random variables with mean U, one can similarly show that

$$\sigma^2(\vec{\lambda}) = \left[1-\left(\frac{E(\vec{\lambda})}{E(\vec{\lambda}_0)}\right)^2\right]\sum_{j<k}\langle(U_{jk}-U)^2\rangle, \tag{S16}$$

where $\langle\rangle$ indicates that we have taken the ensemble average over realizations [S17] of ΔU_{jk}, which simply allows us to set $\langle\Delta U_{jk}\Delta U_{j'k'}\rangle = 0$ where $j,k \neq j',k'$. These results indicate that the deviations in energy levels from those for the exact case caused by inhomogeneity in U_{jk} generically behave as $\sim n\Delta U$. This is because, to estimate $\Delta E(\vec{\lambda})$, we assume that $\sum_{j<k}(U_{jk}-U)$ is the sum of $n(n-1)/2$ uncorrelated positive and negative terms each of magnitude $\sim \Delta U$, and similarly for the variance $\sigma^2(\vec{\lambda})$, except all terms are positive. We therefore expect that, in order to see p revivals of the kind shown in Fig. 3 of the main text, we need to pick up small phase errors $n\Delta U t \lesssim 1$ over time $t \sim p/U$, which corresponds to $\Delta U/U \lesssim 1/(np)$.

However, note that most symmetric $\vec{\lambda}$ subspaces (which have $E(\vec{\lambda})/E(\vec{\lambda}_0)$ close to unity), experience less splitting due to inhomogeneity in U_{jk}, although they do experience an overall shift. For the GHZ protocol described in the main text, the $\vec{\lambda}$ subspaces involved are $(n,0)$, $(n-1,1)$ and $(n-2,2)$, which will shift relative to one another under inhomogeneity in U_{jk} by an amount independent of n for large n.

To obtain some concrete estimates of the effects of an imperfect square-well potential, we consider the following example: a perfect square well, plus an additional harmonic perturbing potential $V_1(x) = \alpha x^2$ (which in effect "rounds off" the boundary of the well somewhat). With first-order corrections, the single-particle wave functions $\phi_j(x)$ are

$$\phi_j(x) \;=\; \sqrt{\frac{2}{L}}\sin(j\pi x/L) + \frac{8}{\pi^2}\left(\alpha L^2/\frac{\hbar^2\pi^2}{2ML^2}\right)\sum_{k\ \ k\neq j}\frac{jk(-1)^{j+k}}{(j^2-k^2)^3}\ \sqrt{\frac{2}{L}}\sin(k\pi x/L). \tag{S17}$$

Substitution into $U_{jk} = UL\int\phi_j^2(x)\phi_k^2(x)dx$ yields exact expressions for the first order corrections to U, which (for all j and k) satisfy: $|U_{jk}-U| < 10^{-2}\left(\alpha L^2/\frac{\hbar^2\pi^2}{2ML^2}\right)U + O(\alpha^2)$. The inhomogeneity is therefore strictly less than one percent if the magnitude of the perturbation is approximately of the same order as the characteristic energy of the square well. The size of the deviations fall off at the fourth power of j,k, such that for ensembles of atoms, ΔU is typically much better than this bound suggests.

S6. GHZ STATE PREPARATION

In this Section, we present the details behind the GHZ state preparation protocol and explain how m GHZ states can be prepared when $N \geq 2^m$.

The state $|A\rangle = |1g\,1g...1g\rangle$ has energy $E_A = nB_g$. The state $|B\rangle = |\{1e\,1g...1g\}\rangle$ lies in the same energy manifold as the state $(|1g\,1e\rangle - |1e\,1g\rangle)|1g...1g\rangle$, which has energy $E_B = \omega_{eg} + (n-1)B_g + B_e + [(n-1)-(-1)]U_{1g\,1e}$. Similarly, $|C\rangle = |\{1e\,1e...1g\}\rangle$ has the same energy as $(|1g\,1e\rangle - |1e\,1g\rangle)(|1g\,1e\rangle - |1e\,1g\rangle)|1g...1g\rangle$, with energy $E_C = 2\omega_{eg} + (n-2)B_g + 2B_e + [2(n-2)-(-2)]U_{1g\,1e}$. Driving with frequency $(E_B - E_A)$ forms an effective two-level system: $\{|A\rangle \leftrightarrow |B\rangle \not\leftrightarrow |C\rangle\}$ since $(E_B - E_A) = \omega_{eg} - B_g + B_e + nU_{1g\,1e} \neq (E_C - E_B) = \omega_{eg} - B_g + B_e + (n-2)U_{1g\,1e}$.

Now we explain why transition $|A\rangle \to |D\rangle \equiv |2g\,2g...2g\rangle$ occurs, while the transition $|B\rangle \not\to |x\rangle$ is blocked for any energy eigenstate $|x\rangle$. First note that the transition $|A\rangle \to |D\rangle$ actually passes through a ladder of intermediate energy eigenstates: $|A\rangle \equiv |1g\,1g...1g\rangle \to |\mathcal{S}\{2g\,1g...1g\}\rangle \to |\mathcal{S}\{2g\,2g...1g\}\rangle \to ... \to |2g\,2g...2g\rangle \equiv |D\rangle$, where $\mathcal{S}$ symmetrizes

its argument. Each state in the ladder has energy $2B_g$ more than the last, and is connected to the previous through the operator $\hat{P} = \sum_j (|2g\rangle_j\langle 1g|_j + |2e\rangle_j\langle 1e|_j)$, which is applied as a pulse with frequency $2B_g$. To show that $|B\rangle$ does not transition to any other state under the action of this pulse, we must prove that *there exists no state* $|x\rangle$ *such that* $\hat{H}|x\rangle = (E_B + 2B_g)|x\rangle$ *and* $\langle x|\hat{P}|B\rangle \neq 0$. We will assume that $n > 2$, $B_e \neq B_g$ and either $|U_{gg}| > 0$ or $|U_{eg}| > 0$.

Our proof has the following structure: we find four orthonormal states such that $\hat{P}|B\rangle \in \text{span}\{|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, |\phi_4\rangle\} \equiv \mathcal{H}_0$, where subspace $\mathcal{H}_0$ is closed under the action of $\hat{H}$ (i.e. for all $|\psi\rangle \in \mathcal{H}_0$, $\hat{H}|\psi\rangle \in \mathcal{H}_0$). Any eigenstate $|x\rangle$ of $\hat{H}$ coupled to $|B\rangle$ through $\hat{P}$ must be in $\mathcal{H}_0$, but we show the four eigenvalues E_i of $\hat{H}$ in $\mathcal{H}_0$ satisfy $E_i \neq (E_B - 2B_g)$.

To complete the proof, we must present $\{|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, |\phi_4\rangle\}$ explicitly, and show that $E_i \neq (E_B - 2B^g)$ for all four eigenstates ($i = 1, 2, 3, 4$). Without loss of generality, take $|B\rangle = (|1g\,1e\rangle - |1e\,1g\rangle)|1g...1g\rangle$, thus $\hat{P}|B\rangle = \sqrt{2(n-2)}|\phi_1\rangle + \sqrt{2}|\phi_3\rangle + \sqrt{2}|\phi_4\rangle$, where $|\phi_1\rangle \equiv \frac{1}{\sqrt{2(n-2)}}(|1g\,1e\rangle - |1e\,1g\rangle)|\mathcal{S}\{1g2g...1g\}\rangle$, $|\phi_2\rangle \equiv \frac{1}{\sqrt{2}}(|2g\,1e\rangle - |1e\,2g\rangle)|1g1g...1g\rangle$, $|\phi_3\rangle \equiv \frac{1}{\sqrt{2(n-2)}}(|1g\,2g\rangle - |2g\,1g\rangle)|\mathcal{S}\{1g1e...1g\}\rangle$, and $|\phi_4\rangle \equiv \frac{1}{\sqrt{2}}(|1g\,2e\rangle - |2e\,1g\rangle)|1g1g...1g\rangle$ (note that $|\phi_4\rangle$ is an energy eigenstate). $\hat{H}$ is closed on subspace $\mathcal{H}_0$ and takes the form:

$$\hat{H} = (E_B - 2B_g) + \begin{pmatrix} 0 & -\sqrt{n-2}U_{gg} & -U_{ge} & 0 \\ -\sqrt{n-2}U_{gg} & (n-2)U_{gg} & \sqrt{n-2}U_{ge} & 0 \\ -U_{ge} & \sqrt{n-2}U_{ge} & (n-1)U_{gg} - U_{ge} & 0 \\ 0 & 0 & 0 & 2(B_g - B_e) \end{pmatrix}. \tag{S18}$$

The matrix written explicitly in Eq. (S18) can be shown to have non-zero determinant (and therefore no vanishing eigenvalues) provided $n > 2$, $B_e \neq B_g$ and either $|U_{gg}| > 0$ or $|U_{eg}| > 0$, which completes our proof.

In the main text, we note that for integer m such that $N \geq 2^m$, it is possible to create m GHZ states. We describe the procedure here in more detail for $m = 2$. First create a regular GHZ state as described in the main text ($|1g1g..1g\rangle + |2g2g..2g\rangle$) from initial state $|1g...1g\rangle$. Then, apply pulse 1 of two different frequencies to $|1g1g..1g\rangle$ and to $|2g2g..2g\rangle$, resulting in ($|1e1g..1g\rangle + |1g1g..1g\rangle + |2e2g..2g\rangle + |2g2g..2g\rangle$). Now, instead of applying pulse 2, apply a pulse which implements $|p\rangle \mapsto |p+2\rangle$ (for $p = 1, 2$), but only to atoms in a many-body state containing no e atoms. The resulting state is ($|1e1g..1g\rangle + |3g3g..3g\rangle + |2e2g..2g\rangle + |4g4g..4g\rangle$). Finally, apply pulse 3 of two different frequencies to yield ($|1g1g..1g\rangle + |2g2g..2g\rangle + |3g3g..3g\rangle + |4g4g..4g\rangle$). This is precisely equivalent to two GHZ states, which can be seen by defining the basis $\{|\Downarrow\Downarrow\rangle, |\Downarrow\Uparrow\rangle, |\Uparrow\Downarrow\rangle, |\Uparrow\Uparrow\rangle \equiv \{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}\}$. Then ($|11..1\rangle + |22..2\rangle + |33..3\rangle + |44..4\rangle$) = ($|\Downarrow\Downarrow .. \Downarrow\rangle + |\Uparrow\Uparrow .. \Uparrow\rangle$)($|\Downarrow\Downarrow .. \Downarrow\rangle + |\Uparrow\Uparrow .. \Uparrow\rangle$). The process could be continued, where in the ith iteration, the second pulse involves $|p\rangle \mapsto |p + 2^i\rangle$ (for $p = 1, 2, 3...2^i$).

[S1] D. Bacon, I. L. Chuang, and A. W. Harrow, Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms p. 1235 (2007).

[S2] W. Fulton and J. Harris, *Representation Theory: A First Course (Graduate Texts in Mathematics)* (Springer, New York, 1991).

[S3] B. E. Sagan, *The Symmetric Group* (Springer, New York, 2000).

[S4] M. E. Rose, *Elementary Theory of Angular Momentum* (Dover Publications Inc, New York, 1957).

[S5] J. Brown and A. Carrington, *Rotational Spectroscopy of Diatomic Molecules*, Cambridge Molecular Science (Cambridge University Press, 2003).

[S6] A. Alex, M. Kalus, A. Huckleberry, and J. von Delft, J. Math. Phys. **52**, 023507 (2011).

[S7] A. V. Gorshkov, M. Hermele, V. Gurarie, C. Xu, P. S. Julienne, J. Ye, P. Zoller, E. Demler, M. D. Lukin, and A. M. Rey, Nature Phys. **6**, 289 (2010).

[S8] R. Zhang, Y. Cheng, H. Zhai, and P. Zhang, Phys. Rev. Lett. **115**, 135301 (2015).

[S9] G. Pagano, M. Mancini, G. Cappellini, L. Livi, C. Sias, J. Catani, M. Inguscio, and L. Fallani, Phys. Rev. Lett. **115**, 265301 (2015).

[S10] M. Höfer, L. Riegger, F. Scazza, C. Hofrichter, D. R. Fernandes, M. M. Parish, J. Levinsen, I. Bloch, and S. Fölling, Phys. Rev. Lett. **115**, 265302 (2015).

[S11] F. Jendrzejewski et al., in preparation (2014).

[S12] S. Trotzky, P. Cheinet, S. Folling, M. Feld, U. Schnorrberger, A. M. Rey, A. Polkovnikov, E. A. Demler, M. D. Lukin, and I. Bloch, Science **319**, 295 (2008).

[S13] R. C. Brown, R. Wyllie, S. B. Koller, E. A. Goldschmidt, M. Foss-Feig, and J. V. Porto, Science **348**, 540 (2015).

[S14] X. Zhang, M. Bishof, S. L. Bromley, C. V. Kraus, M. S. Safronova, P. Zoller, A. M. Rey, and J. Ye, Science **345**, 1467 (2014).

[S15] M. M. Boyd, T. Zelevinsky, A. D. Ludlow, S. Blatt, T. Zanon-Willette, S. M. Foreman, and J. Ye, Phys. Rev. A **76**, 022510 (2007).

[S16] Small deviations from linear polarization will play a more significant role for the $^3P_0$ state than for the $^1S_0$ state because of the larger vector polarizability of the former. However, the $^3P_0$ state is only used in the GHZ protocol where a vector shift is indistinguishable from a slight change in the value of the applied magnetic field.

[S17] It is not necessary to do this – one can calculate the exact expression without taking an ensemble average, but it is quite complicated, and all we seek is an approximate indication of how much spreading to expect for each subspace.