Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The price of locality in fermionic readout at fixed particle number

Annan Chen

annan.chen24@imperial.ac.uk

Imperial College London

Research Article

Keywords: fermionic correlation functions, single-qubit measurements, classical shadows, particle-number conservation, sample complexity, readout errors, reduced density matrices, Hubbard model

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11820903/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

The price of locality in fermionic readout at fixed particle number

Annan Chen[1, *]

[1]*Department of Physics, Imperial College London, London, United Kingdom*

Random orbital rotations before readout give sample-efficient estimates of fermionic correlation functions, but they require a circuit of entangling gates whose depth grows with the number of modes. We determine the cost of measuring every qubit on its own, with the observable chosen only after the measurements. For N particles in n modes in the Jordan–Wigner encoding and an observable that moves q particles, single-qubit measurements need of order k^{2q}/ϵ^2 copies in the worst case, where $k = \min(N, n-N)$ and ϵ is the additive error, whereas arbitrary single-copy measurements need of order k^q/ϵ^2, which orbital rotations attain. Locality thus doubles the exponent, and with ideal readout neither scaling depends on the number of modes. The lower bound holds for protocols with classical feedback, and random Pauli measurements attain it with an estimator that infers the sign of each Jordan–Wigner string once the measured occupations account for all the other particles. For hopping terms the ratio of the two optimal costs lies between $k/4$ and $3k$, and orbital rotations are within a factor of two of optimal. Independent readout errors of probability e multiply the worst-case cost of a hopping term by a factor of order $(1+ne)^{k-1}$ at fixed k, and a convex program gives the optimal estimator for random Pauli measurements at finite size. In simulations of a 24-site Hubbard chain with 1% readout error, 200 000 copies give the momentum distribution with twice the root-mean-square error of error-free orbital rotations and resolve the depletion of the zero-momentum occupation by the interaction.

I. INTRODUCTION

The reduced density matrices of one and two particles determine the energy of a fermionic system with pairwise interactions, its momentum distribution and many other correlation functions, and measuring them is a central task in quantum simulations of molecules and materials [1, 2]. Classical shadows [3] have changed how this task is approached: the state is measured in randomly chosen bases, and the same data are used afterwards to estimate any observable from a large family. For fermions the natural random bases are those of rotated orbitals [4–7]. With random fermionic Gaussian unitaries, a correlation function of order q on n modes can be estimated from a number of copies proportional to n^q [4, 6, 8], and when the state has a known number N of particles, random orbital rotations reduce this number to N^q, independently of the number of modes [7, 9, 10]. These rotations require a circuit that acts on the state before the qubits are read out. A general orbital rotation of n modes is a network of $n(n-1)/2$ Givens rotations, each acting on two modes, of depth proportional to n on a line of qubits [11, 12], and circuits of this kind have been used to measure one-particle density matrices in experiments [13]. Joint measurements of products of two and four Majorana operators reduce the depth to order $\sqrt{n}$ for qubits on a rectangular lattice [14], but still require entangling gates.

Here we ask what it costs to do without the circuit and to measure every qubit on its own. In the Jordan–Wigner encoding [15] this looks unpromising. A hopping term between two distant sites is a sum of Pauli strings that cover every site between them. The standard estimator for random Pauli measurements, which we call the *string estimator*, can use only the copies in which the whole string was measured in matching bases [3], and its cost grows at least as the square of the distance between the sites. A known particle number changes this. The string contributes only a sign, the parity of the number of particles between the two sites, and a hopping term acts only on configurations in which exactly one of the two sites is occupied. Once the measured occupations account for all the other particles, the sign is known, whichever bases were used at the remaining sites (Fig. 1). The resulting *counting estimator* replaces the dependence on distance by a dependence on the number of particles.

Our results are as follows. Let $k = \min(N, n-N)$, and consider observables $O = gT + \bar{g}T^\dagger$ with $|g| = 1$, in which T moves q particles; for $q = 1$ these are hopping terms.

1. *Locality doubles the exponent.* Every protocol built from single-qubit measurements needs $\Omega(k^{2q}/\epsilon^2)$ copies for the worst-case observable of this kind. The bound holds for arbitrary single-qubit measurements, with classical feedback within a copy and between copies, and for biased estimators; it remains valid when the qubits are measured jointly within each of $2q$ consecutive blocks of nearly equal size but separately across the blocks (Proposition 2). Random Pauli measurements with the counting estimator achieve $O(k^{2q}/\epsilon^2)$ for every such observable (Theorem A). Arbitrary measurements of single copies need $\Theta(k^q/\epsilon^2)$, which orbital rotations attain [9] (Proposition 1). For states without a fixed particle number, n takes the place of k in these expressions (Table I).
2. *For hopping terms the price is between $k/4$ and $3k$.* We measure the cost by the worst-case second moment of an unbiased single-copy estimator, in units of $2|g|^2$ (Sec. II). The optimal cost of a hopping term lies between $k^2/4$ and $3k^2/2$ for single-qubit measurements and between $k/2$ and $k+1$ for arbitrary single-copy measurements (Corollary 1). The price of locality, the ratio of the two, therefore lies between $k/4$ and $3k$, and orbital rotations are within a factor of two of the optimal single-copy measurement. With one particle or one hole, single-qubit readout is itself optimal; from $k = 4$ on, orbital rotations have the lower cost for every number of modes.

* annan.chen24@imperial.ac.uk

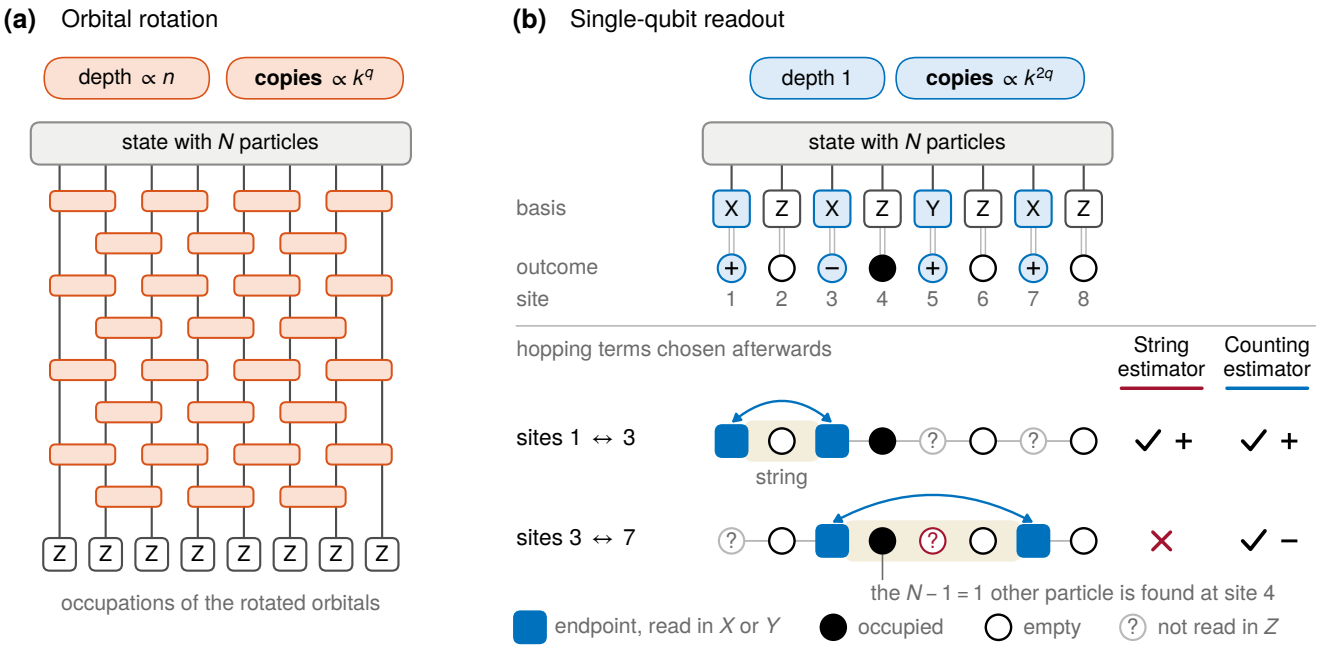


FIG. 1. Two ways of reading out a fermionic state. (a) An orbital rotation on a line of n qubits is a circuit of two-qubit gates of depth proportional to n, followed by a measurement of the occupations of the rotated orbitals. (b) In single-qubit readout every qubit is measured in a randomly chosen Pauli basis, in one layer. For an observable that moves q particles the numbers of copies needed scale as k^q and k^{2q}, with $k = \min(N, n-N)$. The lower part of (b) applies one outcome, for $n = 8$ and $N = 2$, to two hopping terms chosen afterwards. A tick marks an estimator that can use the outcome, and the sign next to it is the sign that the estimator assigns to the Jordan–Wigner string between the endpoints. The string estimator needs every site between the endpoints to have been read in Z; for the term between sites 3 and 7, site 5 was read in Y. The counting estimator needs only that the other particle has been found: the term acts when one of its endpoints is occupied, the other particle is then the one at site 4, and this fixes the sign.

3. *Readout errors enter through the other particles.* When each measured bit is flipped independently with probability e, the cost of the worst-case hopping term grows as $(1+ne)^{k-1}$ at fixed k (Theorem B): each of the $k-1$ particles that the term leaves in place can be confused with a false count, an empty site recorded as occupied, anywhere along the chain. Only errors on empty sites contribute to this factor, so the qubit state that is read more reliably should represent the empty sites (the occupied sites if $N > n/2$).
4. *The optimal estimator is computable.* For random Pauli measurements, the unbiased estimator with the smallest worst-case second moment is the solution of a finite convex program, and a dual solution certifies its optimality (Theorem C).

These results complete the comparison in Table I. The upper bounds for measurements preceded by rotations are known [4, 6–10], as are lower bounds for arbitrary single-copy measurements when all correlation functions of a given order are learned at once [9, 16]. Proposition 1 shows that the same lower bound holds for a single correlation function chosen after the measurements, and Corollary 1 places orbital rotations within a factor of two of the optimum for hopping terms. For single-qubit measurements, the known lower bounds concern general states of the qubits and are governed by the Pauli weight of the observable [3], which for a hopping term in the Jordan–Wigner encoding grows with the distance between its sites; as Koizumi *et al.* point out, such bounds give no guarantee that is independent of the encoding [9]. We show that at fixed particle number the quantity that governs the cost is the particle number, with matching upper and lower bounds.

We test the estimators on simulated measurements. For Hubbard chains of 24 and 16 sites with 1% readout error, 200 000 copies give the momentum distribution with root-mean-square errors of 0.0125 and 0.0231. These are larger by factors of 2.0 and 3.3 than the errors of orbital rotations without gate or readout errors, which we call ideal. On the longer chain the interaction depletes the zero-momentum occupation from 2 to 1.738, and a single data set determines it with a statistical error of 0.016. When the basis set of a water molecule is enlarged from 24 to 115 orbitals per spin, the bound on the cost of every hopping term stays at 61.04 without readout errors and grows from 92.6 to 291.5 with 1% readout error; the string estimator costs between 265 and 6.0×10^3 even without readout errors. In a simple model in which a failed Givens rotation is detected and the copy discarded, single-qubit readout with 1% readout error needs fewer copies than orbital rotations once the failure probability exceeds a threshold, which lies between 0.12% and 0.81% for 24 to 48 modes and two to four particles (Table II and Sec. VIII).

Section II states the setting and the results for ideal readout, and Sec. III explains the counting estimator and the lower bound. Sections IV and V treat readout errors and the optimal

TABLE I. Number of copies needed to estimate an observable that moves q particles and is chosen after the measurements, in units of $|g|^2/\epsilon^2$ and up to factors that depend only on q; $k = \min(N, n-N)$. The lower bounds in the left-hand column follow Refs. [9, 16] (Proposition 1), and the right-hand column is Theorem A.

	Arbitrary single-copy measurements	Single-qubit measurements
Arbitrary states of n modes	$\Theta(n^q)$, attained by Gaussian unitaries [4, 16]	$\Theta(n^{2q})$ (Theorem A)
States of N particles	$\Theta(k^q)$, attained by orbital rotations [9]	$\Theta(k^{2q})$ (Theorem A)

estimator. Section VI contains the numerical tests and applications, Sec. VII the relation to previous work and Sec. VIII a discussion. The proofs are given in the Appendices.

II. SETTING AND MAIN RESULTS

We consider n fermionic modes with annihilation operators $c_1, \dots, c_n$, encoded in n qubits by the Jordan–Wigner transformation along a fixed order of the modes, which we call the chain; the state $|1\rangle$ of qubit i represents an occupied mode. The state ρ has N particles, and $k = \min(N, n-N)$. All statements are symmetric under the exchange of particles and holes.

Observables. A *transfer operator* $T_u = c_{i_1}^\dagger \cdots c_{i_q}^\dagger c_{j_q} \cdots c_{j_1}$ creates particles on q sites and removes them from q other sites; u lists the $2q$ distinct sites, which we call its endpoints. The observables are

$$O = g\,T_u + \bar{g}\,T_u^\dagger, \qquad |g| = a > 0. \tag{1}$$

For $q = 1$ these are the hopping terms $g\,c_i^\dagger c_j + \bar{g}\,c_j^\dagger c_i$. Both the endpoints and the phase of g are chosen after the measurements. The phases $g = a$ and $g = ia$ give $2a\ \mathrm{Re}\langle T_u\rangle$ and $-2a\ \mathrm{Im}\langle T_u\rangle$. The family thus contains every off-diagonal element of the one-particle reduced density matrix and, for larger q, every element of the q-particle reduced density matrix whose $2q$ indices are all distinct. The other off-diagonal elements are products of occupation numbers with transfer operators of lower order, and the estimator of Sec. III A gives them with a second moment no larger than that for the transfer operator alone. The diagonal elements are products of occupation numbers, which are measured directly in the Z basis. We say that the transfer *acts* on a configuration of the particles if T_u or $T_u^\dagger$ does not annihilate it.

Measurements. A measurement of one copy is *local* if each of its effects is separable across the n qubits. This class contains every sequence of single-qubit measurements, with or without local ancillas, in which each measurement may be chosen using the outcomes of the earlier ones [17]. A protocol measures M copies one after another, each an independent preparation of ρ, and may choose the measurement of each copy using the outcomes obtained from earlier copies. The measurements do not depend on the observable.

Accuracy and cost. We require that, for every state of N particles and every observable of the form (1), the estimate of $\langle O\rangle$ lies within ϵ of the true value with probability at least $1-\delta$. This is a guarantee for each observable separately; for L observables simultaneously, δ is replaced by δ/L. For an unbiased estimator f of $\langle O\rangle$ computed from the outcome of a single copy we write $R = \sup_\rho \mathbb{E}_\rho f^2$, the supremum being over the states of N particles, and we call $C = R/(2a^2)$ the normalized cost. The mean of M such estimates has a variance of at most R/M, and the median of the means of $O(\log(1/\delta))$ groups of copies reaches the required accuracy with $M = O(R\,\epsilon^{-2}\log(1/\delta))$ copies. For a transfer operator T_u we denote by R_u the largest value of R over the phases of g, and $C_u = R_u/(2a^2)$. Subscripts on O, Ω and Θ list the parameters on which the implied constants depend, and the upright e is the base of the natural logarithm, as distinct from the error probability e of Sec. IV.

Theorem A (Cost of local measurements)**.** *Let $q \le k \le n/2$.*

(i) Lower bound. *For every local measurement and all unbiased single-copy estimators, the average of R_u over the positions of the endpoints is $\Omega_q(a^2k^{2q})$. Moreover, every protocol that uses local measurements with classical feedback and meets the accuracy requirement for some $0 < \delta \le 1/8$ and $0 < \epsilon \le 2^{-2q-1}a$, with any estimator, needs*

$$M = \Omega_q\Big(k^{2q}a^2\epsilon^{-2}\log(1/\delta)\Big)$$

copies.

(ii) Upper bound. *Independent random Pauli measurements with the counting estimator give $R_u \le 2\mathrm{e}\,a^2k^{2q}$ for every transfer operator, and hence $M = O(k^{2q}a^2\epsilon^{-2}\log(1/\delta))$.*

(iii) No fixed particle number. *For arbitrary states of the n modes, random Pauli measurements with the string estimator give $R_u = O_q(a^2n^{2q})$, and every protocol as in* (i) *that is accurate on all states needs $M = \Omega_q(n^{2q}a^2\epsilon^{-2}\log(1/\delta))$ copies.*

Proposition 1 (Arbitrary single-copy measurements)**.** *Let $q \le k \le n/2$. Every protocol that performs an arbitrary measurement on each copy, chosen using the outcomes from earlier copies, and meets the accuracy requirement needs*

$$M = \Omega_q\Big(k^q a^2\epsilon^{-2}\log(1/\delta)\Big)$$

copies, for ϵ and δ in the range of Theorem A. Orbital rotations attain $M = O_q(k^q a^2\epsilon^{-2}\log(1/\delta))$ [9]. For arbitrary states of the n modes both statements hold with n in place of k, the upper bound being attained by Gaussian unitaries [4].

At half filling $k = n/2$, and the two rows of Table I have the same powers of n. For hopping terms the constants can be bounded as well.

Corollary 1 (Hopping terms)**.** *Let $n \ge 2k \ge 2$, and let one measurement per copy serve every hopping term $g\,c_i^\dagger c_j + \bar{g}\,c_j^\dagger c_i$*

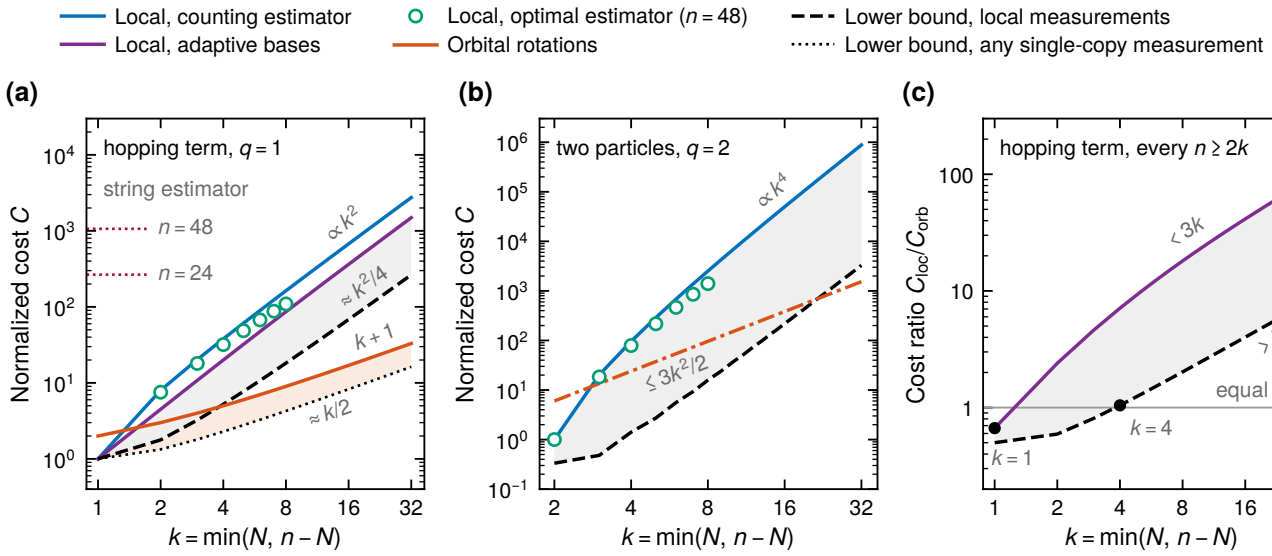


FIG. 2. Cost of local measurements and of arbitrary single-copy measurements without readout errors, in units of $2|g|^2$. (a) Hopping terms. The optimal cost of local measurements lies in the gray band, between the lower bound of Corollary 1 (dashed) and the cost of the adaptive protocol of Sec. III A (purple); the counting estimator with $r = 1/k$ is shown in blue, where r is the probability of measuring a site in X or Y. The optimal cost of arbitrary single-copy measurements lies in the lighter band, between its lower bound (dotted) and the smaller of the cost of the adaptive protocol and the cost $k + 1$ of orbital rotations in the limit of many modes (red). Both bands hold for every $n \geq 2k$ and close at $k = 1$, where both optimal costs equal one. Circles: optimal estimator for random Pauli measurements (Sec. V) at $n = 48$ for the central hopping term, which has equal numbers of sites inside and outside its string, with r chosen optimally. Dotted segments on the left: string estimator for the central hopping term at $n = 24$ and 48, also with the optimal r; its cost is the same for every particle number. (b) Observables that move two particles: counting estimator with $r = 2/k$, lower bound for local measurements (dashed; one endpoint in each of four blocks of $\lfloor k/2 \rfloor$ or $\lceil k/2 \rceil$ modes), optimal estimator with the optimal r at $n = 48$ for endpoints with 22 sites inside and 22 outside the strings (circles), and the upper bound $3k^2/2$ for orbital rotations (dash-dotted). (c) Range of the ratio of the optimal local cost to the cost of orbital rotations for hopping terms, for every $n \geq 2k$. The range lies below one at $k = 1$ and above one from $k = 4$.

with $|g| = a$, the pair of sites and the phase being chosen afterwards. Let C_{loc} and C_{gen} be the smallest values of the largest normalized cost of such a term, for unbiased single-copy estimators, over local measurements and over all single-copy measurements, respectively, and let $C_{\mathrm{orb}}(n, k)$ be the cost of orbital rotations. Then

$$\begin{aligned} \frac{k(n-k)}{n-1} \leq C_{\mathrm{gen}} \leq C_{\mathrm{orb}}(n,k) < k+1, \\ \frac{k^4}{(2k-1)^2} \leq C_{\mathrm{loc}} \leq \frac{3k^2(k+1)}{2(k+2)}. \end{aligned} \tag{2}$$

In particular $k/2 < C_{\mathrm{gen}} < k + 1$ and $k^2/4 < C_{\mathrm{loc}} < 3k^2/2$. Orbital rotations are within a factor $1 + 1/k \leq 2$ of the optimal single-copy measurement, $C_{\mathrm{orb}}(n, k) < (1 + 1/k)\, C_{\mathrm{gen}}$, and the price of locality obeys

$$\frac{k}{4} < \frac{C_{\mathrm{loc}}}{C_{\mathrm{orb}}(n,k)} \leq \frac{C_{\mathrm{loc}}}{C_{\mathrm{gen}}} < 3k. \tag{3}$$

For $k = 1$, $C_{\mathrm{loc}} = C_{\mathrm{gen}} = 1$.

The two lower bounds have a common origin. Take the hopping terms that cross a cut dividing $2k$ half-filled modes into two halves: there are k^2 pairs of sites and two phases for each. A single measurement outcome carries a limited amount of information about these $2k^2$ observables: the squares of their signals, defined in Sec. III B, add up to at most k for an arbitrary measurement and to at most 1 for a local one (Lemma 4). Counting all pairs of modes instead of those across one cut gives the bound on C_{gen} in Eq. (2) (Lemma 10). The cost of orbital rotations is $C_{\mathrm{orb}}(n, k) = (n + 1)(k + 1)(n - k + 1)/[n(n + 2)]$ (Lemma 8). Figure 2 shows these bounds.

Three consequences are worth stating. First, with one particle or one hole, single-qubit readout is optimal among all single-copy measurements: measuring every qubit in X or Y with equal probability has cost 1 for every hopping term, which equals both lower bounds, whereas orbital rotations cost $C_{\mathrm{orb}}(n, 1) = 2(n + 1)/(n + 2)$. Second, for $k \geq 4$ the lower bound on C_{loc} exceeds $k + 1$, so that orbital rotations have the smaller cost for every number of modes. For $k = 3$ the lower bound exceeds $C_{\mathrm{orb}}(n, 3)$ up to $n = 14$; for larger n, and for $k = 2$, the bounds do not decide which is smaller. Third, if the phase of g is fixed in advance for every pair of sites, for instance to real values, each lower bound in Eq. (2) is halved; for $k = 1$ and real g, measuring every qubit in X attains the halved bound $1/2$. Readout errors and the optimal estimator are the subjects of Theorems B and C in Secs. IV and V.

III. WHY LOCALITY DOUBLES THE EXPONENT

A. Counting instead of reading the string

In the Jordan–Wigner encoding, $c_i^\dagger c_j = \sigma_i^+ Z_{i+1}\cdots Z_{j-1}\sigma_j^-$ for $i<j$, where $\sigma^+ = |1\rangle\langle 0|$ and $\sigma^- = |0\rangle\langle 1|$. The string of Z operators between the endpoints contributes the sign $(-1)^m$, where m is the number of particles between the two sites. The string estimator obtains this sign by reading every site of the string in Z. With N known this is more than is needed. A hopping term acts only on configurations in which exactly one of its endpoints is occupied, and the other $k-1$ particles (holes, if $N > n/2$) are then somewhere else. Once Z measurements have found all of them, the number of particles inside the string is known, whichever bases were used at the remaining sites.

Figure 1(b) shows an outcome for $n=8$ and $N=2$ in which site 5 was measured in Y. For the hopping term between sites 3 and 7 the string estimator cannot use this outcome, because the occupation of site 5 is missing. The counting estimator can: one particle is shared between sites 3 and 7 and the other was found at site 4, so the number of particles inside the string is odd, whatever a Z measurement of site 5 would have given.

In general, let each site be measured in Z with probability $1-r$ and in X or Y with probability $r/2$ each. For a transfer operator of order q with creation sites u_+ and annihilation sites u_-, let J be the union of the open intervals between the first and second endpoints, the third and fourth, and so on, in their order along the chain; the strings contribute the parity of the number of particles in J. The counting estimator keeps the outcomes in which all $2q$ endpoints were measured in X or Y and exactly $k-q$ particles were found at the other sites, and returns

$$f = \frac{2\,\mathrm{Re}\big(g\,e^{-i\theta}\big)}{r^{2q}(1-r)^{k-q}}\,(-1)^{y_J}\prod_{i\in u} x_i, \tag{4}$$

and zero for all other outcomes. Here $x_i = \pm 1$ are the outcomes at the endpoints, y_J is the number of particles found in J, and $\theta = \sum_{i\in u_+}\phi_i - \sum_{i\in u_-}\phi_i$ with $\phi_i = 0$ for an endpoint measured in X and $\pi/2$ for one measured in Y. On every configuration on which the transfer acts, the outcome is kept with probability $r^{2q}(1-r)^{k-q}$, and the division by this probability removes the bias of the selection. With equal probabilities for X and Y at the endpoints, the factor $2\,\mathrm{Re}(g\,e^{-i\theta})$ times the product of the endpoint outcomes represents $g\,T_u + \bar g\,T_u^\dagger$ on average, up to the sign of the string. A sign that depends only on u and on the order of the operators in T_u is absorbed into g (Appendix A).

With $r = 1/k$ the other particles are all found with a probability that stays above $1/\mathrm{e}$, while the $2q$ endpoints are measured in X or Y with probability k^{-2q}. This is the origin of the exponent $2q$ in Theorem A, and the normalized cost is at most $\mathrm{e}\,k^{2q}$ (Lemma 1). For a hopping term it is $\max(1,kr)/[r^2(1-r)^{k-1}]$; the factor kr comes from configurations in which both endpoints are empty and one of the k particles elsewhere is not found. An adaptive choice of bases reduces the constant: visiting the sites in random order, measuring Z until $k-1$ particles have been found and then X or Y at all remaining sites costs at most $3k^2(k+1)/[2(k+2)]$ (Lemma 2). Figure 2 compares these protocols with orbital rotations and with the lower bounds.

On a configuration on which the transfer acts, finding all the other particles fixes the occupation of every site away from the endpoints. Multiplying the estimate (4) by the occupations found at some of these sites, l, therefore gives an unbiased estimate of $\langle O\prod_l n_l\rangle$ whose second moment is no larger. These products are the off-diagonal elements of the reduced density matrices of higher order in which the indices l are repeated, so that the bounds on the counting estimator hold for them as well.

B. The information in a local measurement

The lower bound rests on a count of how much a single measurement outcome can reveal about many observables at once. Consider hopping terms, and choose $2k$ modes that hold k particles, all other modes being empty (after an exchange of particles and holes if $N > n/2$). Let τ be the maximally mixed state of k particles on these modes, and divide them into two blocks A and B of k consecutive modes. For a hopping term O between a site of A and a site of B, with either phase, let $H = O/a$. The two states $\tau(1\pm\vartheta H)$ have $\langle O\rangle = \pm\vartheta a\,s_*$ with $s_* = k/(2k-1)$, and an accurate protocol must tell them apart.

An outcome y with effect E_y has probability $\mathrm{Tr}(E_y\tau)[1\pm\vartheta\phi_O(y)]$ in these two states. The *signal* $\phi_O(y) = \mathrm{Tr}(\sigma_y H)$ is the expectation of H in the state $\sigma_y = \tau^{1/2}E_y\tau^{1/2}/\mathrm{Tr}(\tau E_y)$ (Lemma 3). If Γ_y denotes the one-particle density matrix of σ_y, the squared signals of the two phases of the pair (i,j) add up to $4|(\Gamma_y)_{ij}|^2$. The question is therefore how large $\sum_{i\in A,j\in B}|(\Gamma_y)_{ij}|^2$ can be. For an arbitrary measurement the answer is $k/4$, because $\Gamma_y^2 \le \Gamma_y$. For a product of single-qubit effects it is $1/4$, however large k is (Lemma 4). The reason is the string. For a product effect of rank one, $|(\Gamma_y)_{ij}|^2$ is an average over two occupation patterns drawn independently from the distribution that the effect defines: a pair of patterns contributes the product of its two string signs if site j is occupied and site i empty in the first pattern and the reverse holds in the second, and zero otherwise. When an endpoint is moved across the sites at which the two patterns differ, this product alternates in sign, and the alternating sums leave at most one term for each block. A separable effect is a mixture of such effects, so the bound holds for every local measurement. For transfers of q particles the same argument with $2q$ blocks bounds the sum of the squared signals by $2/\binom{2q}{q}$ for local measurements (Lemma 5), against $O_q(k^q)$ for arbitrary ones.

A fixed total shared among many observables leaves little for most of them. There are of order k^{2q} transfer operators and the total is bounded by a constant, so that for some observable the mean squared signal per copy is at most of order k^{-2q}. Telling its two states apart with M copies then requires M to be at least of order k^{2q}/ϑ^2, that is, of order $k^{2q}a^2/\epsilon^2$. Feedback does not change this count: the bound on the total holds for every measurement, and therefore for the measurement of each copy, however it was chosen from the earlier outcomes (Lemma 6). This proves Theorem A(i) and, with the total $O_q(k^q)$ of arbitrary measurements, Proposition 1. For unbiased

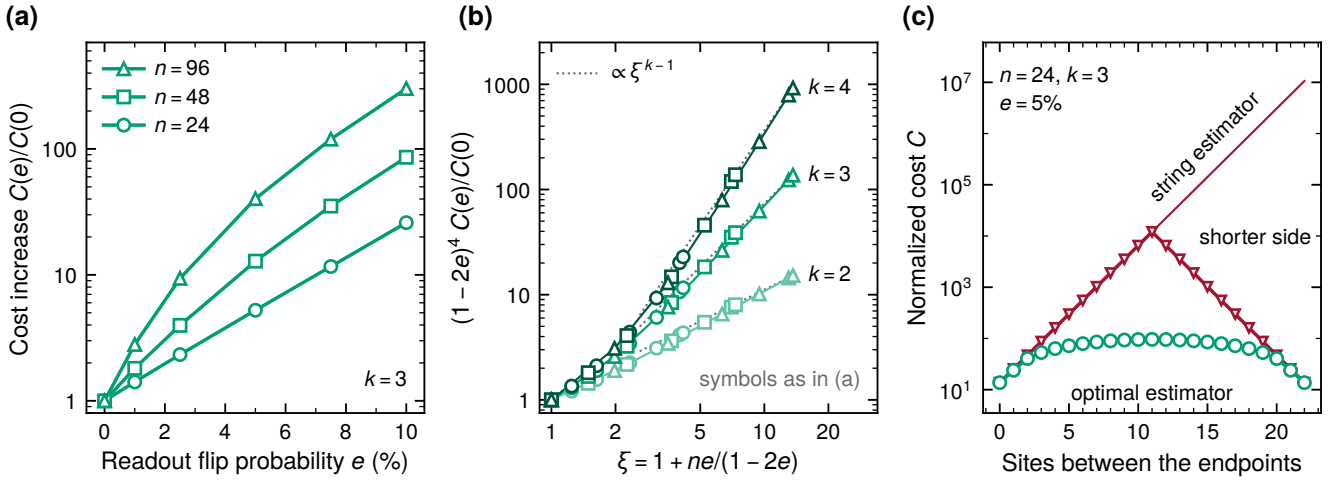


FIG. 3. Readout errors: every measured bit is flipped with probability e, and $r = 1/k$. (a) Cost of the optimal estimator for the central hopping term relative to its value at $e = 0$, for $k = 3$. (b) After division by the factor $(1-2e)^{-4}$ that comes from the endpoints, the costs for the three sizes lie close to one curve for each k as a function of $\xi = 1 + ne/(1-2e)$; thin lines join the points for $n = 96$, and dotted lines have the slope $k-1$ of Theorem B. (c) Cost against the number of sites between the endpoints for $n = 24$, $k = 3$, $e = 5\%$ and $r = 1/3$. Circles: optimal estimator, whose cost depends only on the smaller of the numbers of sites inside and outside the string. Thin line: string estimator, which reads every site between the endpoints. Triangles: string estimator that uses the known particle number to read the smaller of the two regions.

single-copy estimators the same sums give the constants of Corollary 1.

The count also shows what a measurement must do to reach the cost k^q: it must be entangling across cuts of the chain. Divide the $2k$ modes into $2q$ consecutive blocks B_j and take the $p_\times = \prod_j |B_j|$ transfer operators with one endpoint in each block. For measurements that are separable across the blocks, with arbitrary entangling operations inside each block, the squared signals of these observables still add up to at most a constant, and their optimal worst-case cost is $\Theta_q(p_\times)$, which single-qubit measurements attain (Proposition 2); for blocks of nearly equal size $p_\times$ is of order k^{2q}. For $q = 1$ a cut into two halves is crossed by k^2 pairs of sites, and every measurement that is separable across it has a cost at least of order k^2 for one of these pairs, however entangling it is within each half.

IV. READOUT ERRORS

The results above assume ideal readout. With readout errors the number of modes enters the cost.

Theorem B (Readout errors)**.** *Suppose that each single-qubit measurement returns one bit, which is flipped independently with a known probability $e \le e_0 < 1/2$, and that each qubit is measured once per copy. For fixed $k \ge 1$ the normalized cost of the worst-case hopping term is $\Theta_{k,e_0}\big((1+ne)^{k-1}\big)$. Random Pauli measurements attain the upper bound for all hopping terms with the same data. The lower bound $M = \Omega_{k,e_0}\big(a^2(1+ne)^{k-1}/\epsilon^2\big)$ holds for every such protocol with classical feedback and every estimator, at success probability $2/3$ and $0 < \epsilon \le a/4$, even when the hopping term is known before the measurements.*

Readout errors affect the sign of a hopping term through the $k-1$ particles that it leaves in place. At the two endpoints, which are read in X or Y, the errors reduce the expectation of the product of the outcomes by the factor $(1-2e)^2$ and thus multiply the cost by $(1-2e)^{-4}$, whatever the system size. The other particles bring in the size of the system: each lies inside or outside the string, and a false count, an empty site recorded as occupied, can be mistaken for it; the longer the chain, the more empty sites there are on which this can happen.

How the estimator treats false counts decides the cost. Correcting the parity site by site, with the product of unbiased estimates of $1-2n_i$ over the sites of the string, multiplies the second moment by a factor larger than one for every site, which gives a cost exponential in the length of the string. The known particle number avoids this. Since at most $k-1$ particles can occupy the sites of the string on a configuration on which the term acts, the parity is then a polynomial of degree $k-1$ in the occupations, and with this polynomial in place of the full product, false counts enter on at most $k-1$ empty sites at a time. The parity of the sites outside the string differs from that of the inside by the known factor $(-1)^{k-1}$, so the smaller of the two regions can be used. This gives the upper bound of Theorem B (Lemma 11). For the matching lower bound we place each of the $k-1$ particles at an unknown position in one of two sets of sites, one inside and one outside the string, and bound what noisy single-qubit measurements can reveal about which set it is in; the bounds for the $k-1$ particles multiply (Lemmas 12 and 13).

Only errors on empty sites (on occupied sites when $N > n/2$) enter the factor that grows with the system size. If an empty site is recorded as occupied with probability e_{01} and an occupied site as empty with probability e_{10}, the upper bound of Theorem B holds with e_{01} in place of e and the lower bound with $\min(e_{01}, e_{10})$ (Appendix C). The qubit state that

is read more reliably should therefore represent the empty site when $N \le n/2$ and the occupied site otherwise, which can be arranged by flipping every qubit before the readout.

Figure 3 shows the dependence on the system size and on the position of the hopping term, with $r = 1/k$ throughout. Theorem B fixes the power $k-1$. The computed costs show in addition that the single variable $\xi = 1 + ne/(1-2e)$ captures most of the dependence on n and e: after division by the factor $(1-2e)^{-4}$ that comes from the endpoints, the costs for five sizes from $n = 24$ to $n = 96$ differ by at most 7% at $\xi = 2.3$ and at $\xi = 4$.

The position of the hopping term enters through the smaller of the two regions. There are ℓ_{in} sites inside the string and $\ell_{\text{out}} = n - 2 - \ell_{\text{in}}$ outside, and on a configuration on which the term acts the numbers of particles in the two regions add up to $k-1$. The optimal estimator therefore has the same cost when the two lengths are exchanged [Fig. 3(c)]: for $n = 24$, $k = 3$ and $e = 5\%$ it costs 13.7 both for adjacent endpoints and for endpoints at the two ends of the chain, and 95.9 for the central hopping term, for which $\ell_{\text{in}} = \ell_{\text{out}}$. The string estimator, which reads every site between the endpoints, costs $[r^2(1-r)^{\ell_{\text{in}}}(1-2e)^{2\ell_{\text{in}}+4}]^{-1}$ with the same measurement and reaches 1.06×10^7 at 22 sites. A string estimator can also use the known particle number, by reading the smaller of the two regions. Its cost is the same expression with $\min(\ell_{\text{in}}, \ell_{\text{out}})$ in place of ℓ_{in}; it equals the optimal cost at both ends of the range and is 126 times larger for the central hopping term (1.2×10^4 against 95.9).

V. THE OPTIMAL ESTIMATOR

The data contain more information than the counting estimator uses: the numbers of sites that were read in Z, and the outcomes in which the number of particles found differs from $k-q$. At finite size the best use of the data is the solution of a convex program.

Theorem C (Optimal estimator)**.** *Fix $0 < r < 1$ and measure each site independently in Z with probability $1-r$ and in X or Y with probability $r/2$ each, with readout errors as in Theorem B. For a given transfer operator, the smallest worst-case second moment of an unbiased estimator is the value of a finite convex program. An optimal estimator is the product of the X or Y outcomes at the endpoints, a known phase factor, and a function of four numbers: the numbers of sites read in Z and of particles recorded there, inside and outside the string. Every feasible point of the program and every point of its dual bound the optimal value from above and from below. For a hopping term without readout errors, when the numbers of sites inside and outside the string tend to infinity at fixed $k > 1$, the optimal cost, minimized over r, tends to $k^2[k/(k-1)]^{k-1}$, the cost of the counting estimator; this approaches $\mathrm{e}\,k^2$ for large k.*

Two symmetries reduce the search. First, an optimal estimator can be taken to be the product of the X or Y outcomes at the endpoints, the factor $2\,\mathrm{Re}(g\,e^{-i\theta})$ of Eq. (4), and a real function G of the data from the other sites (Lemma 14). Second, G can be taken to depend only on four numbers, $w = (z_{\text{in}}, y_{\text{in}}, z_{\text{out}}, y_{\text{out}})$: the numbers of sites read in Z and of particles recorded there, inside and outside the string (Lemma 15). Correspondingly, the configurations fall into classes $x = (s_{\text{in}}, s_{\text{out}})$ according to the numbers of particles in the two regions, and all configurations of a class give the same distribution $Q_x(w)$ of the data. This distribution is explicit. In a region of ℓ sites of which s are occupied, the probability that z sites are read in Z and that i of them are occupied is $\binom{s}{i}\binom{\ell-s}{z-i}(1-r)^z r^{\ell-z}$, and the number of particles recorded is then the sum of independent $\mathrm{Bin}(i, 1-e)$ and $\mathrm{Bin}(z-i, e)$ variables.

Unbiasedness is a set of linear conditions, one for each class on which the transfer acts: $\sum_w Q_x(w)G(w) = v_x$, where $v_x = \pm 1$ is the sign of the string in that class. The second moment is a maximum over all classes, including those on which the transfer does not act, where the estimate vanishes on average but still fluctuates. In normalized units,

$$C_r = \frac{1}{r^{2q}(1-2e)^{4q}} \min_G \max_x \sum_w Q_x(w)\,G(w)^2, \qquad \sum_w Q_x(w)\,G(w) = v_x \ \text{ if the transfer acts on } x. \tag{5}$$

The prefactor accounts for the probability r^{2q} that all endpoints are read in X or Y and for the attenuation of their outcomes by readout errors. The program is convex, and its dual is simple (Lemma 16): every probability distribution over the classes, together with a multiplier for each condition, gives a lower bound on C_r, and every G that satisfies the conditions gives an upper bound. We solve the program numerically and quote values on which the two bounds agree to the digits shown (Appendix E).

TABLE II. Optimal normalized cost C of the central hopping term for random Pauli measurements, without readout errors and with $e = 1\%$, each with the optimal probability r_{opt} of measuring X or Y; cost C_{orb} of ideal orbital rotations; and failure probability x_* of a Givens rotation at which the model of Sec. VIII gives orbital rotations the cost of single-qubit readout with $e = 1\%$.

		$e = 0$		$e = 1\%$			
n	k	C	r_{opt}	C	r_{opt}	C_{orb}	x_* (%)
24	2	7.14	0.562	9.28	0.578	2.764	0.44
24	3	16.17	0.411	22.73	0.422	3.526	0.67
24	4	27.06	0.334	39.84	0.340	4.207	0.81
48	2	7.54	0.535	11.44	0.574	2.879	0.12
48	3	18.03	0.378	31.71	0.408	3.757	0.19
48	4	31.92	0.299	61.42	0.322	4.594	0.23

Table II lists the optimal costs for the central hopping term. For $n = 24$ and $k = 2$ without readout errors, the optimum is 7.14 at $r = 0.562$, against 7.64 for the optimal estimator at $r = 1/k$ and 8 for the counting estimator; the adaptive protocol of Sec. III A costs at most 4.5, and ideal orbital rotations cost 2.76. Readout errors shift the optimal r upward, and they make the optimization more important: at $e = 1\%$ the optimum is 9.28, whereas the explicit estimator that proves the upper bound of Theorem B (Lemma 11) costs 64 at its optimal r. As the regions inside and outside the string grow at fixed k, the numbers of empty sites read in Z carry less and less information

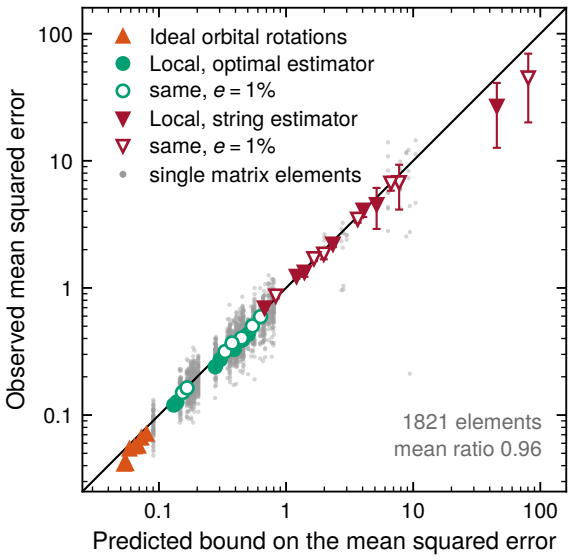


FIG. 4. Simulated measurements: observed mean squared error of a batch of 48 copies against the bound given by the single-copy second moment. Large symbols show, for each of seven combinations (n, k) with $n \leq 16$, the mean over the off-diagonal one-particle elements and three states, for ideal orbital rotations, for the optimal estimator, and for the string estimator applied to the same outcomes; filled symbols are without readout errors, open symbols with 1% readout error, and bars show one standard error. Gray dots are the 1821 individual matrix elements, one-particle and two-particle, for the optimal estimator with 1% readout error. On the line the error equals its bound; each dot is a mean over 24 batches and scatters around its expectation.

about where the particles are. In the limit the data reduce to the numbers of particles found in the two regions, and the optimal cost tends to that of the counting estimator, $k^2[k/(k-1)]^{k-1}$ (Lemma 17).

VI. NUMERICAL TESTS AND APPLICATIONS

A. Single matrix elements

We first compare the errors observed in simulated measurements with the predicted second moments. We take seven combinations (n, k) with $n \leq 16$ and three states for each: a random state of k particles, a random Slater determinant and the ground state of an interacting open chain (Appendix E). On each state we estimate every off-diagonal element $\langle c_i^\dagger c_j \rangle$ with $i < j$, and three transfers of two particles, from 24 independent batches of 48 copies.

For a batch of M copies the single-copy second moment divided by M bounds the mean squared error. Figure 4 compares the observed errors with this bound for the optimal estimator, for the string estimator applied to the same outcomes, and for ideal orbital rotations. The observed errors follow the bounds over three orders of magnitude. For the optimal estimator with 1% readout error, the ratio of the observed error to the bound, averaged over all 1821 matrix elements, is 0.96; for the one-particle elements, with equal weight for the 21 combinations of state and system, it is 0.956 ± 0.008. On the same outcomes the string estimator has mean squared errors between 2 and about 270 times larger. Ideal orbital rotations have mean squared errors smaller than those of the optimal estimator without readout errors by factors of 3 to 6.

B. Hubbard chain

As a physical example we take the momentum distribution

$$n(\kappa) = \frac{1}{L} \sum_{\sigma} \sum_{i,j} e^{-i\kappa(j-i)} \langle c_{i\sigma}^\dagger c_{j\sigma} \rangle$$

of the one-dimensional Hubbard model [18] with on-site repulsion U and hopping amplitude t, on periodic chains of L sites with equal numbers of particles of the two spin species and $U/t = 4$; κ is the lattice momentum, with the lattice spacing set to one. The ground state for these particle numbers is nondegenerate and is obtained by exact diagonalization. The two spin species occupy consecutive blocks of L modes, so that the string of a hopping term of one species involves only particles of that species, and the estimators above apply to each block with k the number of particles per spin. In the simulations each copy is used to measure one species, chosen at random, for single-qubit readout with $r = 1/k$ and for orbital rotations alike.

From 200 000 copies with 1% readout error, the optimal estimator gives the distribution with a root-mean-square error of 0.0125 on 24 sites with two particles per spin, and of 0.0231 on 16 sites with four particles per spin (Fig. 5); without readout errors the values are 0.0113 and 0.0202. Ideal orbital rotations give 0.0062 and 0.0070. On the longer chain the interaction depletes $n(0)$ from 2 to 1.738, and a single data set resolves the depletion: the estimates from the 20 data sets have the mean 1.737 and the standard deviation 0.016.

C. Basis sets of increasing size

Without readout errors the bounds on the cost of single-qubit readout depend on the number of particles and not on the number of modes, which matters when a basis set is enlarged at a fixed number of electrons. The STO-3G, cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets, with spherical functions, give 7, 24, 58 and 115 orbitals per spin for the water molecule [19], with five electrons per spin, so that $k = 2$ for the smallest basis and $k = 5$ for the others. For random Pauli measurements with $r = 1/k$ we compute a bound on the cost that holds for every hopping term and for every state with these numbers of orbitals and electrons, using an estimator that depends only on the numbers of particles recorded inside and outside the string (Appendix E). For the three larger basis sets the bound is 61.04 without readout errors, and 92.6, 146.4 and 291.5 with 1% readout error [Fig. 5(c)]. The string estimator, even when it reads only the smaller of the two regions and r is chosen optimally for its least favorable term, costs 265, 1.6×10^3

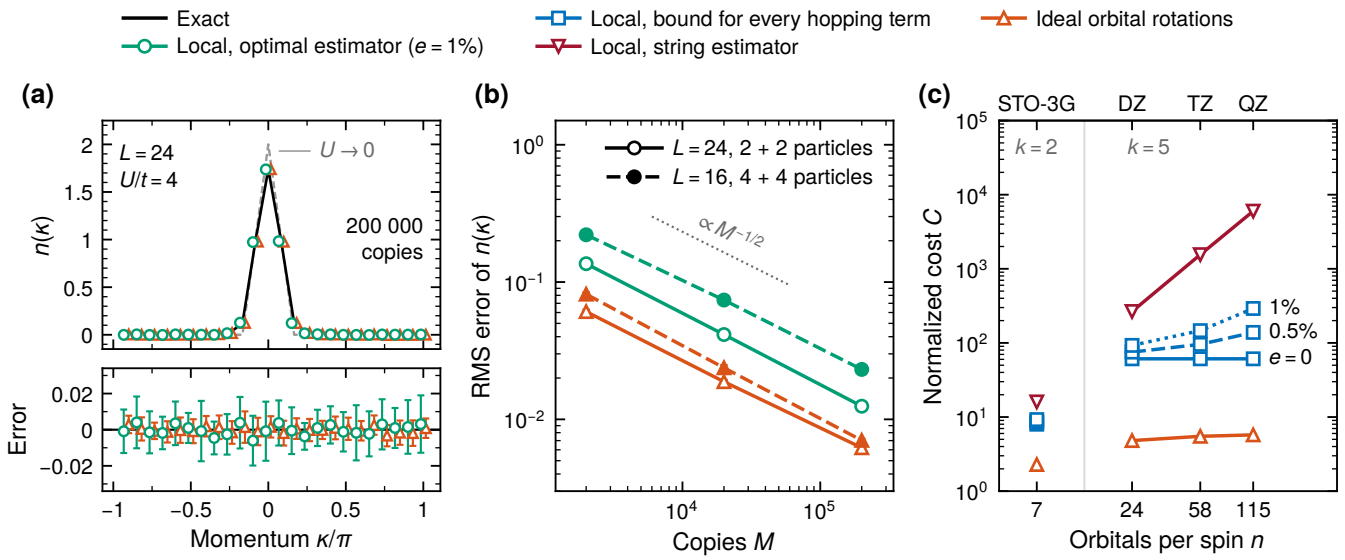


FIG. 5. Applications. (a) Momentum distribution of the Hubbard chain with $L = 24$, two particles per spin and $U/t = 4$, from 200 000 copies: exact values and mean of 20 independent data sets (top; gray, limit $U \to 0$), and deviations from the exact values, with bars giving the standard deviation across the data sets (bottom). (b) Root-mean-square error of the distribution against the number of copies for both chains; mean and standard error over the data sets. The optimal estimator uses $r = 1/k$ and has 1% readout error; orbital rotations are ideal. The dotted line shows the scaling $M^{-1/2}$. (c) Water in the STO-3G, cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets (the last three labeled DZ, TZ and QZ) with five electrons per spin, so that $k = 2$ for $n = 7$ and $k = 5$ otherwise: bound on the cost of every hopping term for $e = 0$, 0.5% and 1% (blue); string estimator that reads the smaller of the two regions, for the central hopping term with r chosen optimally for that term and $e = 0$; and ideal orbital rotations.

and 6.0×10^3 for that term without readout errors, and ideal orbital rotations cost between 4.8 and 5.7. For transfers of two particles the counting estimator with $r = 2/k$ costs 289 for all three basis sets, against at most 37.5 for orbital rotations (Lemma 9). Since these bounds hold for every state, they can be used to plan the number of copies before any data are taken.

VII. RELATION TO PREVIOUS WORK

Measurements preceded by rotations. With random Gaussian unitaries, the estimate of a Majorana operator of degree $2q$ has a second moment of $\binom{2n}{2q}/\binom{n}{q}$, which is of order n^q [4, 6]; a variant adapted to particle-number symmetry halves the depth of the circuit at the cost of more repetitions [4], and bounds of order n^q have recently been extended to arbitrary observables of that degree [8]. Random orbital rotations in a sector of fixed particle number [7] give a second moment of $O_q(k^q)$ for every element of the q-particle reduced density matrix [9]. Lower bounds for arbitrary single-copy measurements were proved for the task of learning all correlation functions of a given order at once [9, 16], and Proposition 1 extends them to a single correlation function chosen after the measurements. The cost of orbital rotations for a hopping term (Lemma 8) follows from Theorem 4 of Koizumi *et al.* [9], and the bound $(q+1)k^q/q!$ for transfers of q particles (Lemma 9) from their Theorem 2; in Refs. [7, 9] the order is denoted by k and the number of particles by η. For a traceless one-body operator h_0, West *et al.* [10] give the bound $(2N+1)\|h_0\|_{\mathrm{F}}^2$ on the variance.

Single-qubit measurements. For general states of qubits, the cost of random single-qubit measurements, and of measurements that entangle a few neighboring qubits, is governed by the Pauli weight of the observable [3, 20]. In the ternary-tree encoding every Majorana operator has a weight of at most $\lceil \log_3(2n+1) \rceil$, and no encoding has an average weight below $\log_3(2n)$ [21]. For the pair of modes with the largest weight, the bound on the second moment of uniformly random Pauli measurements in that encoding corresponds to a normalized cost of 1.1×10^3 at $n = 24$ and 3.3×10^3 at $n = 48$ for a hopping term, whatever the particle number (Appendix E). In the Jordan–Wigner encoding the string estimator costs 265 and 1.1×10^3 for the central hopping term [Fig. 2(a)], and the optimal estimator 7.1 and 7.5 at $k = 2$ (Table II). With 1% readout error the tree encoding gives 1.5×10^3 and 4.5×10^3, against 9.3 and 11.4 for the optimal estimator.

Other approaches. For systems such as ultracold atoms, in which the available operations conserve the particle number and superpositions of different occupations of a single site cannot be measured, Hearth *et al.* construct shadows from one layer of two-site gates, with a number of samples that is polynomial in the number of sites [22]; related protocols use random beam-splitter operations [23] or quenches with free-fermion Hamiltonians [24]. Measuring the two-particle reduced density matrix by groups of compatible operators takes of order n^2 settings [25]. Many works optimize the choice of Pauli bases [26–28] or the estimator for given informationally complete data [29, 30]; Theorem C belongs to the second group and gives, for the present problem, the optimum together with a certificate. Conservation of the particle number is widely used

to detect and mitigate errors [31, 32] and to reduce the number of measurements [33, 34], and readout errors are mitigated by general methods [35] and within shadow estimation [36, 37]; here the particle number is used at an earlier stage, to reconstruct the parity of the string from incomplete occupation data. Finally, our lower bounds belong to the study of state discrimination with separable measurements [17, 38, 39]. A local measurement that is chosen for one hopping term distinguishes the two states $\tau(1 \pm \vartheta H)$ of Sec. III B as well as any measurement; the restriction appears when one measurement has to serve all the terms that cross a cut.

VIII. DISCUSSION

At fixed particle number a measurement circuit lowers the number of copies by a factor of order k^q, and for hopping terms by a factor between $k/4$ and $3k$. Whether this is worth the $n(n-1)/2$ Givens rotations of a general orbital rotation depends on the quality of the gates, and a simple model shows where the balance lies. Suppose that each rotation fails independently with probability x, that a failure is detected, and that the copy is then discarded. The cost of orbital rotations becomes $C_{\mathrm{orb}}/(1-x)^{n(n-1)/2}$, which equals the cost C of single-qubit readout at $x_* = 1 - (C_{\mathrm{orb}}/C)^{2/[n(n-1)]}$. With the costs of the central hopping term at 1% readout error, x_* lies between 0.12% and 0.81% for the sizes of Table II. The model favors orbital rotations in two respects: in practice a failure that is not detected biases the estimate instead of removing a copy, and readout errors are included only for single-qubit readout. For comparison, mean two-qubit gate errors between 0.14% and 0.33%, depending on the gate, and mean measurement errors below 1% have been reported for a recent superconducting processor [40, 41]. Since x_* is approximately $\ln(C/C_{\mathrm{orb}})$ divided by the number of rotations, measurement circuits with fewer gates [14] tolerate larger failure probabilities for a given ratio of the costs.

Several questions remain open. For hopping terms, the optimal cost of local measurements is known to within a factor of six, between $k^2/4$ and $3k^2/2$, and that of arbitrary single-copy measurements to within the factor $1+1/k$. For transfers of more particles the constants of the upper and lower bounds are further apart [Fig. 2(b)], and the dependence of the cost on k in the presence of readout errors remains to be determined. For sums of many hopping terms, such as a kinetic energy, orbital rotations obey a bound in terms of the Frobenius norm [10]; a corresponding bound for single-qubit measurements would be valuable, since the momentum distributions of Sec. VI are sums of this kind. It is also open how the cost decreases with the depth of a measurement circuit between the two limits considered here; by Proposition 2, gates that act only within blocks of the chain do not change its order for the transfers between the blocks. Further directions are measurements on several copies at once [16] and correlated readout errors. Finally, the counting argument relies on the fact that occupations are single-qubit observables in the Jordan–Wigner encoding, and whether other encodings [21, 42, 43] allow single-qubit readout at a cost that is independent of the number of modes is an open question.

DATA AVAILABILITY

The data that support the findings of this article are openly available [44].

Appendix A: Proof of Theorem A

1. Conventions

Exchanging the states $|0\rangle$ and $|1\rangle$ of every qubit exchanges particles and holes. It maps a transfer operator to a transfer operator with the same endpoints in which creation and annihilation are exchanged, up to a sign, and an X or Y measurement to an X or Y measurement. We may therefore assume $N \le n/2$, so that $k = N$, and we speak of particles throughout.

A configuration is a set X of k occupied sites, and $|X\rangle$ is the corresponding basis state. For a transfer operator T_u we write u_+ for the q sites on which it creates particles and u_- for the q sites from which it removes them. Arrange the $2q$ endpoints in increasing order along the chain, and let J be the union of the open intervals between the first and the second endpoint, the third and the fourth, and so on. If X contains u_- and is disjoint from u_+, then

$$T_u|X\rangle = \kappa_u\,(-1)^{|X\cap J|}\,|X^u\rangle, \qquad X^u = (X \setminus u_-) \cup u_+, \quad \text{(A1)}$$

where the sign $\kappa_u = \pm 1$ depends only on u and on the order of the operators in T_u; we absorb it into g. The transfer acts on a configuration X if one of the sets u_+ and u_- is contained in X and the other is disjoint from it. Such a configuration has $K = k - q$ particles away from the endpoints.

A measurement of one copy is described by effects $E_y \ge 0$ with $\sum_y E_y = \mathbf{1}$. We write the proofs for a countable set of outcomes; for a general set of outcomes the sums over y become integrals. A measurement is local if every effect is separable, that is, a sum of products of positive single-qubit operators. States are states of N particles unless stated otherwise.

2. Upper bounds

Lemma 1 (Counting estimator)**.** *Let $q \le k$, $K = k - q$ and $0 < r < 1$, or $r = 1$ if $K = 0$, and measure every site independently in Z with probability $1 - r$ and in X or Y with probability $r/2$ each. The estimator f of Eq. (4) satisfies $\mathbb{E}_\rho f = \mathrm{Tr}(\rho\, O)$ for every state ρ of N particles. On a configuration with $K + j$ particles away from the endpoints, $0 \le j \le q$, its normalized second moment is*

$$\binom{K+j}{j} \frac{r^j}{r^{2q}(1-r)^K}, \quad \text{(A2)}$$

and it vanishes on every other configuration. For $r = 1/k$ the normalized cost is $k^{2q}(1-1/k)^{-(k-q)} \le \mathrm{e}\, k^{2q}$ for every transfer operator of order q. For a hopping term the normalized cost is $\max(1, kr)/[r^2(1-r)^{k-1}]$, which is smallest at $r = 1/k$ when $k > 1$; for $k = 1$, measuring every site in X or Y gives cost one.

Proof. For an endpoint i measured in X ($\phi_i = 0$) or in Y ($\phi_i = \pi/2$), the outcome $x_i = \pm 1$ and the projectors on the corresponding eigenstates satisfy

$$\sum_{x_i=\pm1} x_i \,|x_i\rangle\langle x_i| = e^{i\phi_i}\sigma_i^+ + e^{-i\phi_i}\sigma_i^-.$$

Multiply these operators over the $2q$ endpoints and by $2\,\mathrm{Re}(g\,e^{-i\theta}) = g\,e^{-i\theta} + \bar g\,e^{i\theta}$, and average over the independent choices $\phi_i \in \{0, \pi/2\}$. What remains is $g\prod_{i\in u_+}\sigma_i^+\prod_{j\in u_-}\sigma_j^-$ and its adjoint, because every other term carries a factor $e^{\pm 2i\phi_i}$ for at least one endpoint and the average of $e^{2i\phi_i}$ vanishes. By Eq. (A1) this product differs from T_u by the sign $(-1)^{|X\cap J|}$ of the string. The remaining factor of f depends only on the sites read in Z. If the transfer acts on X, all K particles away from the endpoints are found with probability $(1-r)^K$, and then $y_J = |X \cap J|$; if fewer are found, $f = 0$. Together with the probability r^{2q} that no endpoint is read in Z, this shows that the expectation of f in a state ρ is $\mathrm{Tr}[\rho\,(gT_u + \bar g T_u^\dagger)]$.

The square of f does not depend on the outcomes at the endpoints, and the average of $[2\,\mathrm{Re}(g\,e^{-i\theta})]^2 = 2a^2 + 2\,\mathrm{Re}(g^2e^{-2i\theta})$ over the bases of the endpoints is $2a^2$. The rest of f^2 depends on the choice of bases and on the occupations found in Z, whose distribution is determined by the diagonal of ρ in the occupation basis. It therefore suffices to consider a configuration. If it has $K + j$ particles away from the endpoints, the outcome is kept when all endpoints are measured in X or Y and exactly K of these particles are read in Z, which has probability $r^{2q}\binom{K+j}{K}(1-r)^K r^j$ for $0 \le j \le q$ and cannot happen otherwise. Division by the square of $r^{2q}(1-r)^K$ gives Eq. (A2).

Since $K+j \le k$, $\binom{K+j}{j} \le k^j$, so at $r = 1/k$ the largest value is that for $j = 0$, and $(1-1/k)^{-(k-q)} \le (1-1/k)^{-(k-1)} < \mathrm{e}$. For $q = 1$ the two classes $j = 0$ and $j = 1$ give $\max(1, kr)/[r^2(1-r)^{k-1}]$. The first branch decreases with r for $r \le 1/k$ and the second increases for $r \ge 1/k$, so the minimum is at $r = 1/k$; for $k = 1$ there is no site to be read in Z, and $r = 1$ gives cost one. Finally, the mean of $O(a^2k^{2q}/\epsilon^2)$ independent estimates is within ϵ of $\langle O\rangle$ with probability at least $3/4$ by Chebyshev's inequality, and the median of $O(\log(1/\delta))$ such means is within ϵ with probability at least $1-\delta$. □

The same computation gives the off-diagonal elements with repeated indices. Let D be a set of sites away from the endpoints, and let $\hat n_l = 1$ if site l was found occupied and $\hat n_l = 0$ otherwise. If the transfer acts on X and all K particles away from the endpoints are found, then $\hat n_l$ is the occupation of site l in X. Hence $f\prod_{l\in D}\hat n_l$ has the expectation $\mathrm{Tr}[\rho\,O\prod_{l\in D} n_l]$ in every state of N particles, and its square is at most f^2; the same holds with factors $1-\hat n_l$.

Lemma 2 (Adaptive bases)**.** *Visit the sites in a uniformly random order and measure Z until $K = k - q$ particles have been found; then measure each remaining site in X or Y with equal probability. Let $p_{\mathrm{acc}} = \binom{K+2q}{2q}^{-1}$. The estimator that returns $2\,\mathrm{Re}(g\,e^{-i\theta})(-1)^{y_J}\prod_{i\in u} x_i/p_{\mathrm{acc}}$ when all endpoints are among the sites measured in X or Y, and zero otherwise, is unbiased. Its normalized cost is at most*

$$\binom{K+2q}{2q}^2\binom{k}{K}\Big/\binom{k+2q}{K}, \tag{A3}$$

which equals $3k^2(k+1)/[2(k+2)]$ for a hopping term and is $O_q(k^{2q})$ in general.

Proof. Let the transfer act on a configuration X. No endpoint is read in Z exactly when the K particles of X away from the endpoints precede all $2q$ endpoints in the random order, because an endpoint that is visited earlier is read in Z. The relative order of these $K + 2q$ sites is uniform, so this happens with probability p_{acc}. In that case all particles away from the endpoints have been found and $y_J = |X \cap J|$. The bases of the endpoints are then chosen uniformly and independently of the earlier outcomes, and the outcomes of the other sites that are not read in Z are not used, so the argument of Lemma 1 shows that the estimator is unbiased.

For the second moment consider a configuration with o occupied endpoints, and hence $k - o$ particles away from the endpoints. The outcome is kept when the first K of these $k - o$ particles to be visited precede all $2q$ endpoints, which has probability $\binom{k-o}{K}/\binom{k-o+2q}{K}$; this is the product of the successive probabilities $(k-o-t)/(k-o+2q-t)$, $t = 0, \dots, K-1$, that the next of these sites to be visited is not an endpoint. Each factor increases with $k - o$, so the largest value is that for $o = 0$, and the normalized second moment is at most Eq. (A3). For $q = 1$, $p_{\mathrm{acc}}^{-1} = k(k+1)/2$ and $\binom{k}{k-1}/\binom{k+2}{k-1} = 6/[(k+1)(k+2)]$, which gives $3k^2(k+1)/[2(k+2)]$. For fixed q, Eq. (A3) is $\binom{k+q}{2q}^2\binom{k}{q}/\binom{k+2q}{3q} = O_q(k^{2q})$. □

3. Lower bounds

Lemma 3 (Reference state and signals)**.** *Choose $2k$ modes, which we call active, and let τ be the maximally mixed state of k particles on the active modes, all other modes being empty. Divide the active modes into $2q$ consecutive nonempty blocks $B_1, \dots, B_{2q}$, and let u run over the $p_\times = \prod_j |B_j|$ choices of one endpoint in each block, the blocks that contain the creation sites being the same for all u. For such u and $|g| = a$ let $H = (gT_u + \bar g T_u^\dagger)/a$.*

1. *$\mathrm{Tr}(\tau H^2) = s_* = 2\binom{2k-2q}{k-q}/\binom{2k}{k} \ge 2^{1-2q}$.*
2. *For $0 \le \vartheta \le 1$ the operators $\tau(\mathbf{1} \pm \vartheta H)$ are states of N particles with $\langle O\rangle = \pm\vartheta a s_*$.*
3. *For an effect E_y with $\mathrm{Tr}(\tau E_y) > 0$ let*

$$\sigma_y = \frac{\tau^{1/2}E_y\,\tau^{1/2}}{\mathrm{Tr}(\tau E_y)}, \qquad \phi_O(y) = \mathrm{Tr}(\sigma_y H). \tag{A4}$$

   The probability of the outcome y in the state $\tau(\mathbf{1} \pm \vartheta H)$ is $1 \pm \vartheta\phi_O(y)$ times its probability in τ. Moreover $|\phi_O(y)| \le 1$, and the mean of ϕ_O in the state τ is zero.

4. *The observables with $g = a$ and $g = ia$ have signals whose squares add up to $4|\,\mathrm{Tr}(\sigma_y T_u)|^2$.*

Proof. Since $T_u^2 = 0$, $H^2 = T_u T_u^\dagger + T_u^\dagger T_u$ is the projector on the configurations on which the transfer acts. Among the $\binom{2k}{k}$ configurations of the active modes there are $2\binom{2k-2q}{k-q}$ of them, which gives s_*. Writing $s_*/2$ as the product of the factors $(k-i)^2/[(2k-2i)(2k-2i-1)]$, $i = 0, \dots, q-1$, each of which is at least $1/4$, gives $s_* \geq 2^{1-2q}$. The operator H maps the support of τ to itself and therefore commutes with τ; it has no diagonal elements, so $\mathrm{Tr}(\tau H) = 0$, and $\|H\| \leq 1$ because H^2 is a projector. Hence $\tau(\mathbf{1} \pm \vartheta H)$ is positive with unit trace, and $\langle O \rangle = \pm \vartheta a \,\mathrm{Tr}(\tau H^2)$. For an effect E_y, $\mathrm{Tr}[E_y \tau(\mathbf{1} \pm \vartheta H)] = \mathrm{Tr}(\tau E_y) \pm \vartheta \,\mathrm{Tr}(\tau^{1/2} E_y \tau^{1/2} H)$, which is the third statement; the bound on ϕ_O follows from $\|H\| \leq 1$, and its mean is $\sum_y \mathrm{Tr}(\tau E_y)\phi_O(y) = \mathrm{Tr}(\tau H) = 0$. Finally $\phi_O(y) = 2\,\mathrm{Re}[g\,\mathrm{Tr}(\sigma_y T_u)]/a$, which for $g = a$ and $g = ia$ is twice the real part and minus twice the imaginary part of $\mathrm{Tr}(\sigma_y T_u)$. □

For hopping terms we number the active modes $1, \dots, 2k$ in their order along the chain, which is possible because empty modes contribute no sign to the string, and we take the two blocks $A = \{1, \dots, k\}$ and $B = \{k+1, \dots, 2k\}$. The operator σ_y of Eq. (A4) is a state of k particles on the active modes, and we write $\Gamma_{ij} = \mathrm{Tr}(\sigma_y c_i^\dagger c_j)$ for its one-particle density matrix.

Lemma 4 (Signal carried by one outcome)**.** *Let* $S = \sum_{i\in A}\sum_{j\in B} |\Gamma_{ij}|^2$ *for the state* σ_y *of an effect* E_y *with* $\mathrm{Tr}(\tau E_y) > 0$*. Then*

$$S \leq k/4 \qquad \textit{for every effect,} \tag{A5}$$

$$S \leq 1/4 \qquad \textit{for a separable effect.} \tag{A6}$$

Both bounds are attained.

Proof. Every effect. A one-particle density matrix satisfies $0 \leq \Gamma \leq 1$, hence $\Gamma^2 \leq \Gamma$. The block of this inequality on A reads $\Gamma_{AA}^2 + \Gamma_{AB}\Gamma_{AB}^\dagger \leq \Gamma_{AA}$. Taking the trace,

$$S = \mathrm{Tr}\big(\Gamma_{AB}\Gamma_{AB}^\dagger\big) \leq \mathrm{Tr}\big[\Gamma_{AA}(1-\Gamma_{AA})\big] \leq \frac{k}{4},$$

because Γ_{AA} has k eigenvalues in $[0,1]$. Equality holds for the Slater determinant of the k orbitals $(c_m^\dagger + c_{k+m}^\dagger)/\sqrt{2}$, $m = 1, \dots, k$, for which Γ_{AB} is one half of the identity.

Product effects of rank one. Let E_y be proportional to the projector on $\bigotimes_i(\alpha_i|0\rangle + \beta_i|1\rangle)$. For a configuration X of the active modes write $a_X = \prod_{i\in X}\beta_i \prod_{i\notin X}\alpha_i$ and $w_X = |a_X|^2$. The vector $\sum_X a_X|X\rangle$ is the projection of the product vector on the support of τ, and σ_y is the projector on it divided by $\sum_X w_X$. Let $P(X) = w_X/\sum_{X'} w_{X'}$.

Fix $i \in A$ and $j \in B$. The operator $c_i^\dagger c_j$ maps a configuration X with $j \in X$ and $i \notin X$ to $X^{ij} = (X \setminus \{j\}) \cup \{i\}$ with the sign $(-1)^{|X\cap(i,j)|}$, where (i,j) is the set of modes strictly between i and j, and it annihilates every other configuration. Therefore

$$\Gamma_{ij} = \frac{1}{\sum_{X'} w_{X'}} \sum_{X\ni j,\ X\not\ni i} \overline{a_{X^{ij}}}\, a_X\, (-1)^{|X\cap(i,j)|}.$$

The product $\overline{a_{X^{ij}}} a_X$ equals $\overline{\beta_i\alpha_j}\,\alpha_i\beta_j$ times the nonnegative number $m_X = \prod_{l\in X\setminus\{j\}}|\beta_l|^2 \prod_{l\notin X\cup\{i\}}|\alpha_l|^2$, so its phase is the same for every X, and its modulus is $(w_X w_{X^{ij}})^{1/2}$. Squaring, and using $w_X w_{Y^{ij}} = w_{X^{ij}} w_Y = |\alpha_i\beta_i\alpha_j\beta_j|^2 m_X m_Y$,

$$|\Gamma_{ij}|^2 = \sum_{\substack{X\ni j\\ X\not\ni i}} \sum_{\substack{W\ni i\\ W\not\ni j}} P(X)P(W)\,(-1)^{|X\cap(i,j)|+|W\cap(i,j)|},$$

where $W = Y^{ij}$ and $|Y \cap (i,j)| = |W \cap (i,j)|$. Thus $|\Gamma_{ij}|^2$ is an expectation over two independent configurations X and W drawn from P. With $\Delta = X \triangle W$ their symmetric difference, the sign is $(-1)^{|\Delta\cap(i,j)|}$ and

$$S = \mathbb{E}\left[\sum_{i\in A\cap(W\setminus X)}\ \sum_{j\in B\cap(X\setminus W)} (-1)^{|\Delta\cap(i,j)|}\right]. \tag{A7}$$

Condition on Δ and on $X \cap W$. Since $|X| = |W| = k$, the set Δ has an even number $2s$ of elements, s of them in $X \setminus W$ and s in $W \setminus X$. The weight $w_X w_W = \prod_{l\in X\cap W}|\beta_l|^4 \prod_{l\in\Delta}|\alpha_l\beta_l|^2 \prod_{l\notin X\cup W}|\alpha_l|^4$ depends only on Δ and $X \cap W$, so the s elements of Δ that belong to X form a uniformly random subset. If $s = 0$ the sum is empty. For $s \geq 1$ and two distinct elements i, j of Δ, the conditional probability that $i \in W \setminus X$ and $j \in X \setminus W$ is therefore $\frac{s}{2s} \cdot \frac{s}{2s-1}$.

For $i \in \Delta \cap A$ let $a(i)$ be the number of elements of $\Delta \cap A$ to the right of i, and for $j \in \Delta \cap B$ let $b(j)$ be the number of elements of $\Delta \cap B$ to the left of j. Every mode of A precedes every mode of B, so $|\Delta \cap (i,j)| = a(i) + b(j)$ and the sign factorizes. The conditional expectation of the double sum in Eq. (A7) is

$$\frac{s}{2(2s-1)}\left[\sum_{i\in\Delta\cap A}(-1)^{a(i)}\right]\left[\sum_{j\in\Delta\cap B}(-1)^{b(j)}\right].$$

Each bracket is an alternating sum that starts with +1. It equals one when the number of terms is odd and zero when it is even, and the two numbers of terms add up to $2s$. Since $s/[2(2s-1)] \leq \frac{1}{2}$ for $s \geq 1$,

$$S \leq \tfrac{1}{2}\,\mathrm{Pr}\big[\,|\Delta\cap A| \text{ odd}\,\big].$$

Finally $|\Delta \cap A|$ has the parity of $|X \cap A| + |W \cap A|$. If p is the probability that $|X \cap A|$ is even, independence gives $\mathrm{Pr}[|\Delta \cap A| \text{ odd}] = 2p(1-p) \leq \frac{1}{2}$. This proves Eq. (A6) for product effects of rank one. Equality holds for the product vector with $|\alpha_l| = |\beta_l|$ for one mode of A and one mode of B, $\alpha_l = 0$ for $k-1$ further modes and $\beta_l = 0$ for the remaining ones.

Separable effects. A separable effect is a sum, or an integral, of product effects of rank one, $E_y = \sum_l E_l$. Then $\sigma_y = \sum_l \pi_l \sigma_l$ with $\pi_l = \mathrm{Tr}(\tau E_l)/\mathrm{Tr}(\tau E_y)$, and $\Gamma = \sum_l \pi_l \Gamma^{(l)}$. The quantity S is a convex function of Γ, so the bound for each σ_l gives the bound for σ_y. □

By Lemma 3(4), the squared signals of the two phases of the pair (i,j) add up to $4|\Gamma_{ij}|^2$. For the $2k^2$ hopping terms that cross the cut, with $i \in A$, $j \in B$ and $g \in \{a, ia\}$, Lemma 4 therefore gives

$$\begin{aligned} \sum_O \phi_O(y)^2 &\leq k \quad \text{for every measurement,} \\ \sum_O \phi_O(y)^2 &\leq 1 \quad \text{for a local measurement,} \end{aligned} \tag{A8}$$

for every outcome y.

Lemma 5 (Signal carried by one outcome: transfers of q particles)**.** *In the setting of Lemma 3, for every separable effect E_y with $\operatorname{Tr}(\tau E_y) > 0$,*

$$\sum_u |\operatorname{Tr}(\sigma_y T_u)|^2 \le \frac{1}{2\binom{2q}{q}}.$$

Consequently the squared signals of the $2p_\times$ observables $gT_u + \bar{g}T_u^\dagger$ with $g \in \{a, ia\}$ add up to at most $\Sigma_q = 2/\binom{2q}{q}$.

Proof. For $q = 1$ this is Lemma 4, and the general case follows the same steps. By convexity it suffices to consider an effect proportional to the projector on a product vector. With the notation of the proof of Lemma 4, let X and W be independent configurations with distribution P, and let $\Delta = X \triangle W$ and $|\Delta| = 2s$. Exactly as there,

$$|\operatorname{Tr}(\sigma_y T_u)|^2 = \mathbb{E}\big[\mathbf{1}\{u_- \subset X \setminus W,\ u_+ \subset W \setminus X\}\,(-1)^{|\Delta\cap J|}\big],$$

where J is the set defined in the conventions for the endpoints u. Condition on Δ and on $X \cap W$. The s elements of Δ that belong to X form a uniformly random subset of Δ, so $2q$ given elements of Δ have the required membership with probability $\binom{2s-2q}{s-q}/\binom{2s}{s}$, which is zero for $s < q$. Because the blocks are consecutive, the jth endpoint in the order of the chain lies in the block B_j, and $|\Delta \cap J|$ is the sum over the blocks of the number of elements of $\Delta \cap B_j$ to the right of the endpoint, for odd j, or to its left, for even j. The sign thus factorizes over the blocks, and the sum over the position of the endpoint in B_j is an alternating sum that starts with +1: it equals one if $|\Delta \cap B_j|$ is odd and zero if it is even. Hence

$$\sum_u |\operatorname{Tr}(\sigma_y T_u)|^2 = \mathbb{E}\left[\mathbf{1}_{\mathcal{A}} \binom{2s-2q}{s-q} \Big/ \binom{2s}{s}\right], \tag{A9}$$

where $\mathcal{A}$ is the event that $|\Delta \cap B_j|$ is odd for every j. The ratio of binomial coefficients decreases with s and equals $1/\binom{2q}{q}$ at $s = q$. The event $\mathcal{A}$ requires $|\Delta \cap B_1|$ to be odd, which has probability at most $1/2$ by the argument at the end of the proof of Lemma 4. This gives the bound, and the statement for the signals follows from Lemma 3(4). □

For $q = 1$ the bound is attained. For $q \ge 2$ we do not know whether it is: product vectors with $|\alpha_l| = |\beta_l|$ at one site of each block, $\alpha_l = 0$ at $k - q$ further sites and $\beta_l = 0$ at the remaining ones give $\binom{2q}{q}^{-2}$ for the sum, a factor $\binom{2q}{q}/2$ below the bound.

Lemma 6 (From signals to copies)**.** *In the setting of Lemma 3, let Σ be a number such that, for every outcome of every measurement in the class considered, the squared signals of the $2p_\times$ observables $O = gT_u + \bar{g}T_u^\dagger$ with $g \in \{a, ia\}$ add up to at most Σ.*

1. *For unbiased single-copy estimators, the average of the normalized cost over these $2p_\times$ observables is at least $s_*^2\, p_\times/\Sigma$.*
2. *Every protocol with classical feedback that estimates each of these $\langle O\rangle$ to within ϵ with probability at least $1 - \delta$ on every state of N particles, with $0 < \epsilon \le as_*/4$, $0 < \delta \le 1/8$ and any estimator, uses*

$$M \ge \frac{p_\times a^2 s_*^2}{4\,\Sigma\,\epsilon^2} \log \frac{1}{4\delta(1-\delta)}$$

copies.

3. *For local measurements and blocks of nearly equal sizes, statement 2 holds with $\Sigma = \Sigma_q$ and gives $M = \Omega_q(k^{2q} a^2 \epsilon^{-2} \log(1/\delta))$, and for unbiased single-copy estimators the average of R_u over all positions of the endpoints, for a fixed order of the creation and annihilation sites along the chain, is $\Omega_q(a^2 k^{2q})$.*

Proof. Statement 1. Fix one of the observables and an unbiased estimator f. Its expectation is $\vartheta a s_*$ in the state $\tau(\mathbf{1} + \vartheta H)$ and zero in τ, so by Lemma 3(3) $\mathbb{E}_\tau[f\phi_O] = as_*$. The Cauchy–Schwarz inequality gives $R \ge \mathbb{E}_\tau f^2 \ge a^2 s_*^2/I_O$ with $I_O = \mathbb{E}_\tau \phi_O^2$. The I_O of the $2p_\times$ observables add up to at most Σ, and their harmonic mean is at most their arithmetic mean, so the average of R is at least $2p_\times a^2 s_*^2/\Sigma$.

Statement 2. Let $\mathbb{P}_0$ be the distribution of all outcomes of the protocol when the state is τ, and $\mathbb{P}_{O,\pm}$ the distribution when it is $\tau(\mathbf{1} \pm \vartheta H)$, with $\vartheta = 2\epsilon/(as_*) \le 1/2$. Given the earlier outcomes, the measurement of the next copy is determined, and by Lemma 3(3) the probability of its outcome in the state $\tau(\mathbf{1} \pm \vartheta H)$ is $1 \pm \vartheta\phi_O$ times that in τ, where $|\vartheta\phi_O| \le 1/2$ and ϕ_O has zero mean in τ. The inequality $-\log(1+x) \le -x + x^2$ for $x \ge -1/2$ and the chain rule for the relative entropy give

$$D(\mathbb{P}_0 \| \mathbb{P}_{O,\pm}) \le \vartheta^2 \sum_{m=1}^{M} \mathbb{E}_0\, I_O(\mathcal{M}_m), \tag{A10}$$

where $\mathcal{M}_m$ is the measurement applied to copy m and $I_O(\mathcal{M}) = \mathbb{E}_\tau \phi_O^2$ for the measurement $\mathcal{M}$. For every measurement the I_O add up to at most Σ, so the right-hand sides of Eq. (A10) add up to at most $\vartheta^2 M\Sigma$, and for some observable $D(\mathbb{P}_0\|\mathbb{P}_{O,\pm}) \le \vartheta^2 M\Sigma/(2p_\times)$ for both signs. For this observable, $\langle O\rangle = \pm 2\epsilon$ in the two states, so an estimate within ϵ of the true value has the correct sign. Let $\mathcal{S}$ be the event that the estimate is positive. Then $\mathbb{P}_{O,+}(\mathcal{S}) \ge 1 - \delta$ and $\mathbb{P}_{O,-}(\mathcal{S}) \le \delta$. If $\mathbb{P}_0(\mathcal{S}) \le 1/2$, the monotonicity of the relative entropy gives $D(\mathbb{P}_0\|\mathbb{P}_{O,+}) \ge \frac{1}{2}\log\{1/[4\delta(1-\delta)]\}$, the relative entropy between a fair coin and one with bias $1 - \delta$; otherwise the same holds for $\mathbb{P}_{O,-}$. Hence $\vartheta^2 M\Sigma/(2p_\times) \ge \frac{1}{2}\log\{1/[4\delta(1-\delta)]\}$, which is the claim.

Statement 3. With blocks of sizes $\lfloor k/q\rfloor$ and $\lceil k/q\rceil$, $p_\times \ge (k/2q)^{2q}$, and for $\delta \le 1/8$ the logarithm is at least $0.39\log(1/\delta)$; with Lemma 5 this gives the bound on M. For the average over positions, choose the $2k$ active modes uniformly at random among the n modes. For each choice, statement 1 bounds the average of R over the $p_\times$ choices of u and the two phases, and R_u is at least the average over the two phases. A given set of $2q$ sites can be the set of endpoints only if it is contained in the active modes, which has probability $\binom{2k}{2q}/\binom{n}{2q}$. Under the random choice a set of $2q$ sites is therefore selected with a probability at most $\binom{2k}{2q}/p_\times$ times its probability under the uniform distribution, and the average of R_u over all

positions of the endpoints is at least $2a^2 s_*^2 p_\times^2/[\Sigma_q \binom{2k}{2q}]$, which is $\Omega_q(a^2 k^{2q})$. □

Theorem A(i) is statement 3 of Lemma 6, and part (ii) is Lemma 1. Part (iii) is Lemma 7 below.

4. Measurements that are separable across blocks

Proposition 2 (Block-separable measurements)**.** *In the setting of Lemma 3, call a measurement block-separable if each of its effects is a sum of products of positive operators on the blocks $B_1, \dots, B_{2q}$. For block-separable measurements the conclusions of Lemma 6 hold with $\Sigma = 4$; in particular, for unbiased single-copy estimators the largest normalized cost among the $2p_\times$ observables is at least $s_*^2 p_\times/4 \geq 4^{-2q} p_\times$. Conversely, there is a measurement with single-qubit effects for which each of these observables has normalized cost $p_\times$.*

Proof. Lower bound. By convexity and Lemma 3(4) it suffices to show that $\sum_u |\operatorname{Tr}(\sigma_y T_u)|^2 \leq 1$ when E_y is proportional to the projector on a product $\bigotimes_j \psi_j$ of vectors of the blocks. Decompose each ψ_j into its components $\psi_j^{(m)}$ with m particles in the block. The projection of the product on the support of τ is the sum of the products $\bigotimes_j \psi_j^{(m_j)}$ over the vectors $\boldsymbol{m} = (m_1, \dots, m_{2q})$ with $\sum_j m_j = k$. Let $\pi_{\boldsymbol{m}}$ be the squared norm of such a product divided by the squared norm of the projection, so that $\sum_{\boldsymbol{m}} \pi_{\boldsymbol{m}} = 1$. By Eq. (A1), T_u is a product over the blocks of operators $t_j^{(i)}$, where i is the endpoint in B_j: a creation or annihilation operator at site i multiplied by the parity of the sites of B_j to the right of i, for odd j, or to its left, for even j. It changes the number of particles in B_j by $h_j = \pm 1$, the same for all u. Therefore

$$\operatorname{Tr}(\sigma_y T_u) = \sum_{\boldsymbol{m}} \sqrt{\pi_{\boldsymbol{m}} \pi_{\boldsymbol{m}+\boldsymbol{h}}} \prod_j A_j^{\boldsymbol{m}}(i_j),$$

where $A_j^{\boldsymbol{m}}(i)$ is the matrix element of $t_j^{(i)}$ between the normalized vectors $\psi_j^{(m_j)}$ and $\psi_j^{(m_j+h_j)}$. For fixed $\boldsymbol{m}$ and j, the vector with the components $A_j^{\boldsymbol{m}}(i)$, $i \in B_j$, has Euclidean norm at most one: for every unit vector (x_i) the operator $\sum_i x_i t_j^{(i)}$ is the creation or annihilation operator of a normalized mode and has norm one. The products $\prod_j A_j^{\boldsymbol{m}}(i_j)$, as u runs over its $p_\times$ values, therefore also form a vector of norm at most one, and the triangle and Cauchy–Schwarz inequalities give

$$\Big[\sum_u |\operatorname{Tr}(\sigma_y T_u)|^2\Big]^{1/2} \leq \sum_{\boldsymbol{m}} \sqrt{\pi_{\boldsymbol{m}} \pi_{\boldsymbol{m}+\boldsymbol{h}}} \leq 1.$$

Upper bound. For each copy choose one of the $p_\times$ values of u at random, measure its $2q$ endpoints in X or Y with equal probability and every other site in Z. For the chosen u the strings are read completely, and $p_\times$ times $2\operatorname{Re}(g\, e^{-i\theta})(-1)^{y_J} \prod_{i\in u} x_i$ is an unbiased estimate of $\langle O\rangle$ for these endpoints; for all other endpoints the estimate is zero. Its normalized second moment is $p_\times^2/p_\times = p_\times$. □

5. States without a fixed particle number

Lemma 7 (Arbitrary states of the n modes)**.** *Measure every site in Z with probability $1-r$ and in X or Y with probability $r/2$ each, with $r = 2q/n$. The string estimator, which returns $2\operatorname{Re}(g\, e^{-i\theta})(-1)^{y_J} \prod_{i\in u} x_i/[r^{2q}(1-r)^{|J|}]$ when all endpoints were measured in X or Y and all sites of J in Z, and zero otherwise, is unbiased for every state of the n modes, and its normalized cost is $[r^{2q}(1-r)^{|J|}]^{-1} \leq (\mathrm{e}\, n/2q)^{2q}$. Every protocol with local measurements and classical feedback that meets the accuracy requirement on all states, for ϵ and δ as in Theorem A, uses $\Omega_q(n^{2q} a^2 \epsilon^{-2} \log(1/\delta))$ copies.*

Proof. As in the proof of Lemma 1, the average over X and Y at the endpoints leaves $gT_u + \bar{g}T_u^\dagger$, now on the whole Fock space, because the terms that change the particle number also carry a factor $e^{\pm 2i\phi_i}$; here the sign of the string is read directly from the sites of J. The outcome is used with probability $r^{2q}(1-r)^{|J|}$, and the square of the estimate averages to $2a^2$ divided by this probability. Since $|J| \leq n - 2q$, the choice $r = 2q/n$ gives $(1-r)^{|J|} \geq (1-2q/n)^{n-2q} \geq \mathrm{e}^{-2q}$. For the lower bound, a protocol that is accurate on all states is accurate on the states of $\lfloor n/2 \rfloor$ particles on $2\lfloor n/2 \rfloor$ of the modes, and Lemma 6 applies with $k = \lfloor n/2 \rfloor$. □

Appendix B: Arbitrary single-copy measurements and orbital rotations

Lemma 8 (Orbital rotations: hopping terms)**.** *For $1 \leq N \leq n-1$ and $k = \min(N, n-N)$, the estimator of orbital-rotation shadows for a hopping term has the normalized cost*

$$C_{\mathrm{orb}}(n,k) = \frac{(n+1)(k+1)(n-k+1)}{n(n+2)} \tag{B1}$$

for every phase of g. At fixed k it increases with n, from $(2k+1)(k+1)/(4k)$ at $n = 2k$ to the limit $k+1$.

Proof. Orbital-rotation shadows apply a Haar-random rotation of the n orbitals and measure the occupations. For $i \neq j$ the estimate of $\Gamma_{ij} = \langle c_i^\dagger c_j\rangle$ is $\widehat{\Gamma}_{ij} = (n+1)(\Pi_{\mathrm{occ}})_{ji}$, where Π_{occ} is the projector on the orbitals found occupied, expressed in the original basis [9]. Theorem 4 of Ref. [9] gives the variance of $\widehat{\Gamma}_{ij}$; adding $|\Gamma_{ij}|^2$ to it,

$$\begin{aligned} \mathbb{E}_\rho |\widehat{\Gamma}_{ij}|^2 &= \alpha\big(N + \langle n_i + n_j\rangle_\rho\big) - \beta\langle n_i n_j\rangle_\rho, \\ \alpha &= \frac{(n+1)(n-N+1)}{n(n+2)}, \qquad \beta = \frac{n+1}{n}. \end{aligned} \tag{B2}$$

The estimate of $\langle O\rangle$ is $f = g\,\widehat{\Gamma}_{ij} + \bar{g}\,\overline{\widehat{\Gamma}_{ij}}$, so that $\mathbb{E}_\rho f^2 = 2a^2\, \mathbb{E}_\rho |\widehat{\Gamma}_{ij}|^2 + 2\operatorname{Re}(g^2\, \mathbb{E}_\rho \widehat{\Gamma}_{ij}^2)$. The last term vanishes. Indeed, $\mathbb{E}_\rho \widehat{\Gamma}_{ij}^2$ is the expectation in ρ of an operator W, and because the Haar measure is invariant under changes of the phases of the orbitals, W is multiplied by $e^{2i(\chi_i - \chi_j)}$ when the phase of every mode l is rotated by χ_l. An operator with this property changes the occupation of mode i by two, and is therefore zero. Hence the normalized second moment of f is $\mathbb{E}_\rho |\widehat{\Gamma}_{ij}|^2$, whatever the phase of g.

The right-hand side of Eq. (B2) is the expectation of an operator that is diagonal in the occupation basis, with the values αN, $\alpha(N+1)$ and $\alpha(N+2)-\beta$ on configurations with zero, one and two of the sites i and j occupied. Since $\alpha < \beta$, the largest is $\alpha(N+1)$, and a configuration with exactly one of the two sites occupied exists for $1 \le N \le n-1$. This gives Eq. (B1) with N in place of k, and the expression is unchanged when N is replaced by $n-N$. For the dependence on n write

$$\frac{C_{\mathrm{orb}}(n,k)}{k+1} = 1 - \frac{kn+k-1}{n(n+2)}.$$

The derivative of the subtracted fraction with respect to n has the numerator $-kn^2-(2n+2)(k-1) < 0$. □

Lemma 9 (Orbital rotations: transfers of q particles)**.** *For a transfer of q particles, the estimator of orbital-rotation shadows has a normalized cost of at most* $(q+1)k^q/q!$.

Proof. Let $\widehat{T}_u$ be the estimate of $\langle T_u\rangle$. As in the proof of Lemma 8, $\mathbb{E}_\rho \widehat{T}_u^2 = 0$, and the operator whose expectation is $\mathbb{E}_\rho|\widehat{T}_u|^2$ is invariant under changes of the phases of the orbitals, hence diagonal in the occupation basis. Its largest expectation is therefore attained on a configuration, where $\langle T_u\rangle = 0$ and the second moment equals the variance. Theorem 2 of Ref. [9] bounds this variance, which in our normalization gives $(q+1)N^q/q!$. For $N > n/2$ the same protocol applied to the holes gives the bound with k. □

Proof of Proposition 1. Koizumi *et al.* prove the lower bound for the task of estimating all entries of the q-particle reduced density matrix at once [9]. Their argument applies to a single entry chosen after the measurements, as follows. Take the reference state τ and the $2p_\times$ observables of Lemma 6, with $p_\times \ge (k/2q)^{2q}$. For an effect E_y of positive probability, the numbers $\mathrm{Tr}(\sigma_y T_u)$ are, up to signs, distinct entries of the q-particle reduced density matrix $\Gamma_y^{(q)}$ of the state σ_y in the basis of Slater determinants, normalized to the trace $\binom{k}{q}$. Its Hilbert–Schmidt norm is at most $c_q k^{q/2}$ for every state of k fermions [45]. For $q=1$ this is $\mathrm{Tr}\,\Gamma^2 \le \mathrm{Tr}\,\Gamma = k$, and for $q=2$ it is due to Christiansen [46], whose two-particle matrix has the trace $k(k-1)$; dividing by two gives our normalization and the bound $\sqrt{5}\,k/2$. The bounds extend from pure to mixed states because the norm is convex. By Lemma 3(4),

$$\sum_O \phi_O(y)^2 = 4\sum_u |\,\mathrm{Tr}(\sigma_y T_u)|^2 \le 4\left\|\Gamma_y^{(q)}\right\|_{\mathrm{HS}}^2 \le 4c_q^2 k^q$$

for every outcome of every measurement. Lemma 6 with $\Sigma = 4c_q^2 k^q$ then gives

$$M \ge \frac{p_\times a^2 s_*^2}{16\,c_q^2 k^q \epsilon^2}\log\frac{1}{4\delta(1-\delta)} = \Omega_q\Big(k^q a^2 \epsilon^{-2}\log(1/\delta)\Big).$$

The upper bound is the second moment $O_q(a^2k^q)$ of the estimator of orbital-rotation shadows (Lemma 9) with the median of means. For arbitrary states of the n modes, the states of $\lfloor n/2\rfloor$ particles on $2\lfloor n/2\rfloor$ of the modes give the lower bound with n in place of k, and a transfer operator is a sum of at most 4^q Majorana operators of degree $2q$, each of which is estimated with a second moment of order n^q by Gaussian unitaries [4]. □

Lemma 10 (All pairs of modes)**.** *Let* $1 \le N \le n-1$ *and* $k = \min(N, n-N)$*, and consider the* $n(n-1)$ *hopping terms* $O = g\,c_i^\dagger c_j + \bar g\, c_j^\dagger c_i$ *with* $i<j$ *and* $g \in \{a, ia\}$*. For every single-copy measurement and all unbiased estimators of these observables on the states of N particles, the largest normalized cost is at least the average, and*

$$\frac{1}{n(n-1)}\sum_O C_O \;\ge\; \frac{k(n-k)}{n-1}. \tag{B3}$$

Proof. Let τ_N be the maximally mixed state of N particles in the n modes, and $H = O/a$. The operator $H^2 = n_i(1-n_j)+n_j(1-n_i)$ is the projector on the configurations in which exactly one of the two modes is occupied, so

$$s_N = \mathrm{Tr}(\tau_N H^2) = \frac{2N(n-N)}{n(n-1)}.$$

The statements of Lemma 3 hold with τ_N and s_N in place of τ and s_*, with the same proof. For an effect E_y let Γ_y be the one-particle density matrix of the state $\sigma_y = \tau_N^{1/2} E_y \tau_N^{1/2}/\mathrm{Tr}(\tau_N E_y)$. The squared signals of the two phases of the pair (i,j) add up to $4|(\Gamma_y)_{ij}|^2$, and therefore

$$\sum_O \phi_O(y)^2 = 4\sum_{i<j}\big|(\Gamma_y)_{ij}\big|^2 = 2\Big[\,\mathrm{Tr}\,\Gamma_y^2 - \sum_i (\Gamma_y)_{ii}^2\Big]$$
$$\le 2\Big(N - \frac{N^2}{n}\Big),$$

because $0 \le \Gamma_y \le 1$ gives $\mathrm{Tr}\,\Gamma_y^2 \le \mathrm{Tr}\,\Gamma_y = N$, and the n diagonal entries add up to N. As in the proof of Lemma 6, an unbiased estimator has $R_O \ge a^2 s_N^2/I_O$ with $I_O = \mathbb{E}_{\tau_N}\phi_O^2$, and the harmonic mean of the I_O is at most their arithmetic mean. With $\sum_O I_O \le 2N(n-N)/n$ this gives

$$\frac{1}{n(n-1)}\sum_O \frac{R_O}{2a^2} \;\ge\; \frac{s_N^2\, n(n-1)}{2\sum_O I_O} \;\ge\; \frac{N(n-N)}{n-1},$$

and $N(n-N) = k(n-k)$. □

Proof of Corollary 1. Lower bounds. For C_{loc} use the $2k^2$ hopping terms between the blocks A and B of $2k$ active modes, which exist because $n \ge 2k$. Lemma 6(1) with $p_\times = k^2$, $s_* = k/(2k-1)$ and $\Sigma = 1$, from Eq. (A8), shows that the average normalized cost of these observables, and hence the largest, is at least $k^4/(2k-1)^2$. The bound on C_{gen} is Lemma 10.

Upper bounds. Orbital rotations are a single-copy measurement with an unbiased estimator of cost $C_{\mathrm{orb}}(n,k) < k+1$ for every hopping term (Lemma 8), and the adaptive protocol of Lemma 2 is local with a cost of at most $3k^2(k+1)/[2(k+2)]$.

Simplified bounds. For $n \ge 2k$ the ratio $(n-k)/(n-1)$ is at least $k/(2k-1)$, so $C_{\mathrm{gen}} \ge k^2/(2k-1) > k/2$; the bounds on C_{loc} follow from $4k^2 > (2k-1)^2$ and $k+1 < k+2$. By Lemma 8,

$$C_{\mathrm{orb}}(n,k)\Big/\frac{k(n-k)}{n-1} = \Big(1+\frac{1}{k}\Big)\frac{(n^2-1)(n-k+1)}{n(n+2)(n-k)} < 1+\frac{1}{k},$$

because $n(n+2)(n-k) - (n^2-1)(n-k+1) = n(n-2k) + n - k + 1 > 0$.

Ratios. From $C_{\rm orb}(n,k) < k+1$ and $4k^3-(2k-1)^2(k+1) = 3k-1 > 0$ we obtain $C_{\rm loc}/C_{\rm orb}(n,k) > k/4$, and $C_{\rm gen} \le C_{\rm orb}(n,k)$ gives the same lower bound for $C_{\rm loc}/C_{\rm gen}$. Further, $C_{\rm loc}/C_{\rm gen} \le 3(k+1)(2k-1)/[2(k+2)] < 3k$, because $(k+1)(2k-1) < 2k(k+2)$.

Special cases. For $k = 1$ the adaptive protocol measures every qubit in X or Y and has cost one, which equals both lower bounds; hence $C_{\rm loc} = C_{\rm gen} = 1$, whereas $C_{\rm orb}(n,1) = 2(n+1)/(n+2) \ge 3/2$. For $k \ge 4$, $k^4-(k+1)(2k-1)^2 = k^3(k-4)+3k-1 > 0$, so that $C_{\rm loc} > k+1 > C_{\rm orb}(n,k)$. If the phase of g is fixed in advance for every pair of sites, the sums of squared signals contain one of the two terms for each pair and obey the same bounds, while the number of observables is halved; both lower bounds in Eq. (2) are then halved. For $k = 1$ and real g the halved bounds are attained: on the states of one particle, or of one hole, the product $x_i x_j$ of the outcomes of X measurements of two sites has the expectation of $c_i^\dagger c_j + c_j^\dagger c_i$ up to a known sign, so that $\pm a\, x_i x_j$ is an unbiased estimate with second moment a^2. □

Appendix C: Proof of Theorem B

Throughout this appendix $\lambda = 1-2e$. A site read in Z yields a recorded occupation $y_i \in \{0,1\}$, which differs from the true occupation with probability e. An endpoint read in X or Y yields a recorded outcome $x_i = \pm 1$, whose expectation is λ times that of the ideal outcome.

1. Upper bound

Lemma 11 (The parity as a polynomial)**.** *Let $0 < r < 1$, $0 \le e \le e_0 < 1/2$ and $K = k - q$. Let S be the smaller of the set J and the set of sites that belong neither to J nor to the endpoints, and let $b = |S|$. For $i \in S$ let $\hat n_i = (y_i - e)/[(1-r)\lambda]$ if site i is read in Z and $\hat n_i = 0$ otherwise, and let*

$$F = \sum_{\nu=0}^{K} (-2)^\nu\, \mathcal{E}_\nu(\hat n_S), \tag{C1}$$

where $\mathcal{E}_\nu$ is the elementary symmetric polynomial of degree ν in the variables $\hat n_i$, $i \in S$. Let $\mathbf{1}_\perp$ be the indicator that every endpoint is read in X or Y, and let $\varepsilon_S = 1$ if $S = J$ and $\varepsilon_S = (-1)^K$ otherwise. Then

$$f = \mathbf{1}_\perp \frac{2\,\mathrm{Re}(g\, e^{-i\theta})}{(r\lambda)^{2q}}\, \varepsilon_S\, F \prod_{i\in u} x_i$$

is an unbiased estimator of $\langle O\rangle$, and its normalized cost is $O_{k,e_0,r}\big((1+be)^K\big)$.

Proof. The estimate $\hat n_i$ is unbiased for the occupation n_i: a site is read in Z with probability $1-r$, and the recorded bit then has the expectation $e + \lambda n_i$. The estimates at different sites are independent given the configuration, so the expectation of $\mathcal{E}_\nu(\hat n_S)$ is $\mathcal{E}_\nu(n_S)$. On a configuration on which the transfer acts, at most K sites of S are occupied, so $\mathcal{E}_\nu(n_S) = 0$ for $\nu > K$ and the expectation of F is $\sum_\nu (-2)^\nu \mathcal{E}_\nu(n_S) = \prod_{i\in S}(1-2n_i)$, the parity of S. The K particles away from the endpoints are divided between J and its complement, so the parity of J is ε_S times that of S. The factor that depends on the endpoints is treated as in Lemma 1, with the additional factor λ^{2q} for the attenuation of the $2q$ recorded outcomes. Hence f is unbiased.

For the second moment, the average of f^2 over the bases and outcomes of the endpoints is $2a^2 r^{2q}/(r\lambda)^{4q}$ times F^2, and it suffices to bound $\mathbb{E}F^2$ on a configuration with s occupied sites in S; here s may be as large as k. Write $\hat n_i = n_i + \delta_i$. The δ_i are independent with zero mean and with the variances

$$v_0 = \frac{e(1-e)}{(1-r)\lambda^2}, \qquad v_1 = \frac{e(1-e)+r\lambda^2}{(1-r)\lambda^2}$$

on empty and on occupied sites. Expanding F in the products $\prod_{i\in T}\delta_i$ over subsets $T \subset S$, which are orthogonal, gives

$$\begin{aligned} \mathbb{E}F^2 &= \sum_{i+j\le K} \binom{s}{i}\binom{b-s}{j} (4v_1)^i (4v_0)^j A_{ij}^2, \\ A_{ij} &= \sum_{m=0}^{K-i-j} \binom{s-i}{m} (-2)^m, \end{aligned} \tag{C2}$$

where i and j are the numbers of occupied and of empty sites in T. For fixed k, e_0 and r, the factors $\binom{s}{i}(4v_1)^i A_{ij}^2$ are bounded, and $\binom{b-s}{j} v_0^j = O((be)^j)$ with $j \le K$. Hence $\mathbb{E}F^2 = O_{k,e_0,r}((1+be)^K)$. □

For a hopping term $K = k-1$ and $b \le n/2$, so the estimator of Lemma 11 with, for instance, $r = 1/2$ has a normalized cost of $O_{k,e_0}((1+ne)^{k-1})$ for all hopping terms with the same data. This is the upper bound of Theorem B. If the full product $\prod_{i\in S}(1-2\hat n_i)$ is used in place of F, the second moment on a configuration with s occupied sites in S is $(1+4v_0)^{b-s}(1+4v_1)^s$, which is exponential in b.

2. Lower bound

Lemma 12 (A particle at an unknown position)**.** *Let two disjoint sets of $s_0 \ge 1$ sites be given, and let a sign $S_0 = \pm 1$ be chosen with equal probabilities. Place one particle at a uniformly random site of the first set if $S_0 = +1$ and of the second set if $S_0 = -1$, and let Y be the recorded occupations when all $2s_0$ sites are read in Z, each bit being flipped independently with probability $e < 1/2$. Then $\eta = \mathbb{E}[\mathbb{E}(S_0 \mid Y)^2]$ satisfies*

$$\eta^{-1} \ge 1 + \frac{s_0\, e}{1-2e}. \tag{C3}$$

Proof. For $e = 0$ the statement is trivial; let $e > 0$ and $\theta_e = e/(1-e)$. Form a list that contains the site of the particle and, independently with probability θ_e, each of the other sites. If every site of the list is then recorded as occupied with probability $1-e$, independently, and no other site is, the result has the distribution of Y: the site of the particle is recorded with probability $1-e$ and every other site with probability $\theta_e(1-e) = e$. The list is therefore at least as informative as Y, and $\eta \le \mathbb{E}[\mathbb{E}(S_0 \mid \text{list})^2]$.

Given the list, the particle is equally likely to be at any of its sites, and the number h of sites in the list is distributed

as one plus a binomial variable with $m' = 2s_0 - 1$ trials and probability θ_e. Given h, the list is a uniformly random set of h of the $2s_0$ sites, and the posterior mean of S_0 is the mean of the signs of its sites, where the sites of the first set have the sign +1 and those of the second set −1. For a sample of h out of $2s_0$ signs with zero total, drawn without replacement, the mean has the variance $(2s_0 - h)/(m'h)$. Averaging over h, with $\mathbb{E}(1/h) = [1 - (1 - \theta_e)^{m'+1}]/[(m' + 1)\theta_e]$,

$$\eta \le \bar\eta = \frac{2s_0\,\mathbb{E}(1/h) - 1}{m'} = \frac{1}{m'}\sum_{j=1}^{m'} c^{\,j}, \qquad c = 1 - \theta_e.$$

To bound $\bar\eta$, note that $1 - \bar\eta = (\theta_e/m')\sum_{j=0}^{m'-1}(m' - j)\,c^{\,j}$. Both $m' - j$ and $c^{\,j}$ decrease with j, so the average of their product is at least the product of their averages, which are $(m' + 1)/2 = s_0$ and $\bar\eta/c$. Hence $1 - \bar\eta \ge \theta_e s_0 \bar\eta/c$, and since $\theta_e/c = e/(1 - 2e)$ this is Eq. (C3). □

Lemma 13 (Lower bound for a hopping term)**.** *In the setting of Theorem B, consider a hopping term with $\ell_{\rm in}$ sites inside the string and $\ell_{\rm out}$ outside, and let $b = \min(\ell_{\rm in}, \ell_{\rm out})$, $m_0 = \min(k - 1, b)$ and, if $m_0 \ge 1$, $s_0 = \lfloor b/m_0 \rfloor$. With η as in Lemma 12 for this s_0, every unbiased single-copy estimator has*

$$C \ge \frac{1}{2\lambda^4 \eta^{m_0}},$$

and every protocol with classical feedback that estimates $\langle O\rangle$ to within $\epsilon \le a/4$ with probability at least $2/3$ on every state of N particles uses $M = \Omega(a^2/[\epsilon^2\lambda^4\eta^{m_0}])$ copies, even if the hopping term is known in advance.

Proof. The states. Choose m_0 pairs of sets of sites, the sets being disjoint and each pair consisting of s_0 sites inside the string and s_0 sites outside. For each pair choose a sign $S_m = \pm 1$ with equal probabilities and place one particle at a uniformly random site of the inside set if $S_m = -1$ and of the outside set if $S_m = +1$. Place the remaining $k - 1 - m_0$ particles at fixed sites, and put one particle in the superposition of the two endpoints for which $\langle O\rangle = aV\prod_m S_m$ up to a fixed sign, where $V = \pm 1$ is a further sign chosen with equal probabilities; changing V changes the relative phase of the superposition by π. Let $L = V\prod_m S_m$, and let $\rho_\pm$ be the mixture of these states over all choices with $L = \pm 1$, so that $\langle O\rangle = \pm a$ in $\rho_\pm$ after absorbing the fixed sign into g.

Reduction to recorded occupations. We show that for every measurement allowed in Theorem B, the posterior mean W of L given the outcome, in the mixture $(\rho_+ + \rho_-)/2$, satisfies

$$\mathbb{E}W^2 \le \lambda^4 \eta^{m_0}. \tag{C4}$$

Away from the endpoints the states are diagonal in the occupation basis. A single-qubit measurement that returns one bit, flipped with probability e, records the value 1 with probabilities p_0 and p_1 on an empty and on an occupied site, and both lie in $[e, 1 - e]$. Every such pair is obtained from the recorded occupation of the site by replacing a recorded 0 by 1 with probability $[(1 - e)p_0 - ep_1]/\lambda$ and keeping a recorded 1 with probability $[(1 - e)p_1 - ep_0]/\lambda$; these numbers lie in $[0, 1]$, as one checks at the four corners of the square $[e, 1 - e]^2$. An observer who is given the recorded occupations Y_m of all sites of each pair of sets, and who measures the endpoints as the protocol prescribes, can therefore reproduce the outcome of the protocol site by site, whatever the order of the measurements and their dependence on earlier outcomes. The posterior mean of L given the data of this observer has a mean square at least as large as $\mathbb{E}W^2$.

The observer. The signs $V, S_1, \ldots, S_{m_0}$ are independent, Y_m depends only on S_m and on the position of the particle of the mth pair, and the outcomes at the endpoints depend only on V and on the bases, which may depend on the Y_m. The posterior mean of L is therefore the product of the posterior means of the S_m given Y_m and of that of V given the endpoint outcomes. By Lemma 12 each S_m contributes a factor η to the mean square, and it remains to show that the endpoints contribute at most λ^4, for every choice of their bases.

The endpoints. The two endpoints are in a maximally entangled state of two qubits, and the two values of V give orthogonal states. A measurement of one of them with two outcomes is a randomized function of a projective measurement along some axis, by the argument used above for the other sites, so we may assume that each endpoint is measured projectively and its outcome is flipped with probability e. The first recorded outcome is ±1 with equal probabilities for both values of V. Let the two axes make the angles with cosines c_1 and c_2 with the Z axis; the second may depend on the first outcome. Given the first recorded outcome, the mean of the second is $c_0 + Va_0$ with $|c_0| = \lambda^2|c_1c_2|$ and $|a_0| \le \lambda^2[(1 - c_1^2)(1 - c_2^2)]^{1/2}$. The posterior mean of V then has the mean square $a_0^2/(1 - c_0^2) \le \lambda^4$, because $(1 - c_1^2)(1 - c_2^2) \le 1 - c_1^2c_2^2 \le 1 - \lambda^4 c_1^2 c_2^2$. This proves Eq. (C4).

Unbiased estimators. An unbiased estimator has the expectations $\pm a$ in $\rho_\pm$, so in the mixture $(\rho_+ + \rho_-)/2$ we have $\mathbb{E}[fW] = a$, and the Cauchy–Schwarz inequality gives $R \ge \mathbb{E}f^2 \ge a^2/\mathbb{E}W^2 \ge a^2/(\lambda^4\eta^{m_0})$.

Copies. For $0 < \vartheta \le 1/2$ the states $\frac{1}{2}(1 \pm \vartheta)\rho_+ + \frac{1}{2}(1 \mp \vartheta)\rho_-$ have $\langle O\rangle = \pm\vartheta a$, and the probabilities of an outcome in them are $1 \pm \vartheta W$ times its probability in the mixture. The relative entropy between the two outcome distributions of one copy is at most their chi-squared divergence, which is at most $8\vartheta^2\mathbb{E}W^2$, and by the chain rule the relative entropy between the distributions of all outcomes of a protocol with feedback is at most $8M\vartheta^2\lambda^4\eta^{m_0}$. With $\vartheta = 2\epsilon/a$, an estimate within ϵ of the true value determines the sign of $\langle O\rangle$, and doing so with probability $2/3$ requires a relative entropy bounded below by a constant. This gives the bound on M. □

For the central hopping term and $n \ge 2k$, $b = \lfloor (n-2)/2 \rfloor \ge k - 1$, so that $m_0 = k - 1$ and, for $k > 1$, $s_0 = \lfloor b/(k - 1) \rfloor$ is at least a constant times n/k. With Eq. (C3), Lemma 13 then gives $C = \Omega_{k,e_0}((1 + ne)^{k-1})$ and the corresponding bound on M, which completes the proof of Theorem B.

Remark (Unequal error probabilities). Suppose that an empty site is recorded as occupied with probability e_{01} and an occupied site as empty with probability e_{10}, both at most $e_0 < 1/2$, and that the two outcomes of an X or Y measurement are subject to the same pair of errors. The upper bound of Theorem B then holds with e_{01} in place of e. Indeed, in Lemma 11 the estimate $\hat n_i = (y_i - e_{01})/[(1 - r)(1 - e_{01} - e_{10})]$ is unbiased,

its variance on an empty site is $e_{01}(1-e_{01})/[(1-r)(1-e_{01}-e_{10})^2]$, which takes the place of v_0, and its variance on an occupied site and the second moments of the corrected endpoint outcomes are bounded in terms of e_0 and r. Exchanging the two computational states of every qubit before the readout exchanges e_{01} and e_{10}, so the bound holds with $e_* = \min(e_{01}, e_{10})$. Conversely, the lower bound of Theorem B holds with e_* in place of e. A symmetric channel with flip probability e_*, followed by recording 1 with probability α after a 0 and with probability β after a 1, where

$$\alpha = \frac{(1-e_*)e_{01} - e_*(1-e_{10})}{1-2e_*},$$
$$\beta = \frac{(1-e_*)(1-e_{10}) - e_* e_{01}}{1-2e_*},$$

reproduces the channel with the errors (e_{01}, e_{10}), and $0 \le \alpha \le \beta \le 1$. Every protocol for the unequal channel is therefore a protocol for the symmetric one. With a free choice of the state that represents an empty site, the normalized cost of the worst-case hopping term is thus $\Theta_{k,e_0}((1+ne_*)^{k-1})$.

Appendix D: Proof of Theorem C

We call the list of the bases chosen and the outcomes recorded for one copy its record. The measurement is that of Theorem C, and $\lambda = 1-2e$.

Lemma 14 (Estimators of product form)**.** *Let f be an unbiased estimator of $\langle O\rangle$ on the states of N particles, and let $\mathbf{1}_\perp$ be the indicator of the event that every endpoint is read in X or Y. There is an unbiased estimator of the form*

$$\tilde f = \mathbf{1}_\perp \, 2\,\mathrm{Re}\big(g\, e^{-i\theta}\big) \prod_{i\in u} x_i \; \tilde G(\omega) \tag{D1}$$

with $\sup_\rho \mathbb{E}_\rho \tilde f^2 \le \sup_\rho \mathbb{E}_\rho f^2$. Here θ and the outcomes x_i are those of Eq. (4)*, and $\tilde G$ is a real function of the record ω of the other sites that depends only on which of them were read in Z and on the occupations recorded there.*

Proof. For a site s let $S_s = \exp(i\pi n_s/2)$. It is a unitary on qubit s that preserves the particle number, and the Pauli operators of that qubit satisfy $S_s^\dagger X S_s = -Y$, $S_s^\dagger Y S_s = X$ and $S_s^\dagger Z S_s = Z$. Let Λ_s be the map on records that replaces the entry (X,x) of site s by (Y,x) and the entry (Y,y) by $(X,-y)$ and changes nothing else. Because X and Y are chosen with equal probability and a flip of the recorded bit commutes with a change of sign of the outcome, the record of the state $S_s\rho S_s^\dagger$ has the distribution of Λ_s applied to the record of ρ. Hence $\mathbb{E}_\rho[h\circ\Lambda_s] = \mathbb{E}_{S_s\rho S_s^\dagger}[h]$ for every function h of the record.

The transfer operator satisfies $S_s^\dagger T_u S_s = T_u$ if s is not an endpoint, $-iT_u$ if T_u creates a particle at s, and iT_u if it removes one. Let $\mathcal{G}$ be the group of all products $\gamma = \prod_s \Lambda_s^{j_s}$ for which the unitary $\prod_s S_s^{j_s}$ multiplies T_u by a sign $\chi(\gamma) = \pm 1$. For $\gamma \in \mathcal{G}$ the estimator $\chi(\gamma)\, f\circ\gamma$ is unbiased, and its second moment in the state ρ is that of f in the rotated state, so the two have the same worst-case second moment. The average $\bar f = |\mathcal{G}|^{-1}\sum_\gamma \chi(\gamma)\, f\circ\gamma$ is unbiased, and since $f \mapsto \sup_\rho \mathbb{E}_\rho f^2$ is convex its worst-case second moment is at most that of f. By construction $\bar f \circ \gamma = \chi(\gamma)\bar f$ for every $\gamma \in \mathcal{G}$.

The group contains Λ_s for every site that is not an endpoint, with $\chi = 1$. Invariance under it means that $\bar f$ depends neither on the basis nor on the outcome when such a site is read in X or Y. For an endpoint s the group contains Λ_s^2 with $\chi = -1$; this map changes the sign of an X or Y outcome and leaves a Z entry unchanged. Hence $\bar f$ vanishes unless every endpoint is read in X or Y, and it is odd in each endpoint outcome. Describe the entry of an endpoint s by the angle $\vartheta_s = 0, \pi/2, \pi, 3\pi/2$ for $(X,+), (Y,+), (X,-), (Y,-)$, so that Λ_s adds $\pi/2$ to ϑ_s. A function that changes sign under $\vartheta_s \to \vartheta_s + \pi$ for every endpoint is a combination of the products $\prod_s e^{im_s\vartheta_s}$ with $m_s = \pm 1$. The group also contains $\Lambda_s\Lambda_t$ for a creation site s and an annihilation site t, and $\Lambda_s\Lambda_t^{-1}$ for two sites of the same kind, all with $\chi = 1$; invariance under them leaves only the product with $m_s = 1$ on the creation sites and $m_s = -1$ on the annihilation sites, and its complex conjugate. With $\phi_s = 0$ for X and $\pi/2$ for Y one has $e^{i\vartheta_s} = x_s e^{i\phi_s}$, and $\theta = \sum_{i\in u_+}\phi_i - \sum_{i\in u_-}\phi_i$. Therefore

$$\bar f = \mathbf{1}_\perp \, 2\,\mathrm{Re}\big(c(\omega)\, e^{-i\theta}\big) \prod_{i\in u} x_i$$

with a complex function c of the record ω of the other sites.

Write $c = g\,(\tilde G + i\tilde G')$ with real $\tilde G$ and $\tilde G'$. As in the proof of Lemma 1, averaging over the bases of the endpoints turns the product of their outcomes into T_u times the sign of the string, so the expectation of $\bar f$ in a state ρ is that of $g\,T_u D + \bar g\, D^\dagger T_u^\dagger$, where D is diagonal in the occupations of the other sites and depends linearly on $\tilde G + i\tilde G'$ with real coefficients. Unbiasedness on all states of N particles requires $D = 1$ on every configuration on which T_u acts, which is one real linear condition on $\tilde G$ and the same condition with right-hand side zero on $\tilde G'$. Since $\mathbb{E}e^{2i\theta} = 0$, and since the distribution of ω does not depend on the bases chosen at the endpoints once their outcomes are summed over, the second moment of $\bar f$ given ω and given that every endpoint is read in X or Y is $2|c(\omega)|^2 = 2a^2[\tilde G(\omega)^2 + \tilde G'(\omega)^2]$. Setting $\tilde G' = 0$ therefore keeps the estimator unbiased and does not increase its second moment in any state. □

Lemma 15 (The four numbers)**.** *Among the unbiased estimators of the form* (D1)*, the smallest worst-case second moment is attained with $\tilde G(\omega) = G(w)/(r\lambda)^{2q}$, where G depends on the record only through $w = (z_{\rm in}, y_{\rm in}, z_{\rm out}, y_{\rm out})$, the numbers of sites read in Z and of particles recorded there in J and in the set of the other sites that are not endpoints. The optimization over G is the program* (5)*.*

Proof. With $G = (r\lambda)^{2q}\tilde G$, the argument of Lemma 1 shows that the estimator (D1) is unbiased if and only if, for every configuration X on which the transfer acts, the expectation of G given X is the sign $(-1)^{|X\cap J|}$ of the string; the factor $(r\lambda)^{2q}$ is the probability that no endpoint is read in Z times the attenuation of the recorded endpoint outcomes. The second moment of the estimator on a configuration X, which need not be one on which the transfer acts, is $2a^2r^{2q}$ times the expectation of $\tilde G^2$ given X, that is, $2a^2/(r^{2q}\lambda^{4q})$ times the expectation of G^2.

Let π be a permutation of the sites that maps J to itself and leaves every endpoint in place. If G satisfies the conditions, so does $G \circ \pi$, because the condition for a configuration X becomes the condition for $\pi(X)$ and the sign of the string depends only on the number of particles in J. The largest second moment over the configurations is a convex function of G that is unchanged by π. The average of $G \circ \pi$ over all such permutations therefore satisfies the conditions and has a worst-case second moment no larger than that of G, and it depends on the record only through w.

For a function of w, the expectations given X depend only on the class $x = (s_{\text{in}}, s_{\text{out}})$ of X, the numbers of particles in the two regions. In a region with ℓ sites of which s are occupied, the probability that z sites are read in Z and y particles are recorded is

$$\sum_i \binom{s}{i}\binom{\ell - s}{z - i}(1-r)^z r^{\ell - z} \Pr[Y_i = y], \tag{D2}$$

where i is the number of occupied sites read in Z and Y_i is the sum of independent $\text{Bin}(i, 1-e)$ and $\text{Bin}(z-i, e)$ variables. The product of these probabilities for the two regions is $Q_x(w)$. The conditions are $\sum_w Q_x(w)G(w) = v_x$ for the classes with $s_{\text{in}} + s_{\text{out}} = k - q$, with $v_x = (-1)^{s_{\text{in}}}$, and the normalized second moment is the largest value of $\sum_w Q_x(w)G(w)^2/(r^{2q}\lambda^{4q})$ over all classes. This is the program (5). □

Lemma 16 (Upper and lower bounds)**.** *Let $\gamma = r^{-2q}\lambda^{-4q}$, let $\mathcal{X}_0$ be the set of classes on which the transfer acts, and for a probability distribution μ on the classes let $Q_\mu = \sum_x \mu_x Q_x$. For $0 < r < 1$ the conditions of the program* (5) *can be satisfied, the minimum is attained, and*

$$C_r = \gamma \max_\mu \sup_\nu \Phi(\mu, \nu),$$

$$\Phi(\mu, \nu) = 2\sum_{x \in \mathcal{X}_0} \nu_x v_x - \sum_w \frac{\left[\sum_{x \in \mathcal{X}_0} \nu_x Q_x(w)\right]^2}{Q_\mu(w)}, \tag{D3}$$

where ν ranges over the real vectors indexed by $\mathcal{X}_0$; a term with vanishing numerator and denominator is zero, and a term with a nonzero numerator and a vanishing denominator is infinite. In particular, every G that satisfies the conditions gives the upper bound $\gamma \max_x \sum_w Q_x G^2$ on C_r, and every pair (μ, ν) gives the lower bound $\gamma\,\Phi(\mu, \nu)$.

Proof. For $0 < r < 1$ the estimator of Lemma 11 has the form (D1), so some G satisfies the conditions (Lemma 15). Let G be such a function and let (μ, ν) be given. The largest second moment is at least the average with the weights μ, and

$$\sum_w Q_\mu G^2 = 2\sum_x \nu_x v_x + \sum_w \left[Q_\mu G^2 - 2G\sum_x \nu_x Q_x\right]$$

by the conditions. Minimizing each term of the last sum over the value of $G(w)$ gives $\Phi(\mu, \nu)$, which is therefore a lower bound on C_r/γ. For the equality, the function $\sum_w Q_\mu(w)G(w)^2$ is convex in G and linear in μ, the conditions define a convex set and the distributions μ form a compact convex set, so the minimum over G and the maximum over μ can be exchanged; for fixed μ the minimum over G is a least-squares problem with linear conditions, whose value is the supremum of $\Phi(\mu, \nu)$ over ν. The minimum over G is attained because the largest second moment is a continuous convex function that tends to infinity with every value $G(w)$ that has positive probability in some class. When Q_μ is positive, the optimal G for a given μ is $G(w) = \sum_{x \in \mathcal{X}_0} \nu_x Q_x(w)/Q_\mu(w)$, with ν determined by the conditions. □

For $r = 1$ no site is read in Z, and the conditions can be satisfied only if the sign v_x is the same for all classes on which the transfer acts, as for $k = q$ [the point $k = 2$ in Fig. 2(b)]; the optimal cost is then λ^{-4q}. Otherwise no unbiased estimator exists.

Lemma 17 (Large systems)**.** *For a hopping term without readout errors and fixed $k > 1$, the minimum of C_r over r tends to $k^2[k/(k-1)]^{k-1}$ when the numbers of sites inside and outside the string tend to infinity. This is the cost of the counting estimator with $r = 1/k$.*

Proof. Upper bound. The counting estimator satisfies the conditions of the program for every size, and its cost $\max(1, kr)/[r^2(1-r)^{k-1}]$ equals $k^2[k/(k-1)]^{k-1}$ at $r = 1/k$ (Lemma 1).

A family of lower bounds. Let μ_1 be the distribution on the classes with $s_{\text{in}} + s_{\text{out}} = k - 1$ that results from placing each of $k-1$ particles inside or outside the string with probability $1/2$, and μ_2 the corresponding distribution for k particles; the classes in the support of μ_2 are those of configurations with both endpoints empty. For $0 < \kappa \le 1$ take $\mu = \kappa\mu_1 + (1-\kappa)\mu_2$ and $\nu_x = t\,\mu_1(x)v_x$ in Eq. (D3), and maximize over t. This gives

$$C_r \ge \frac{1}{r^2\,\mathcal{S}_\kappa}, \qquad \mathcal{S}_\kappa = \sum_w \frac{\left[\sum_x \mu_1(x)v_x Q_x(w)\right]^2}{Q_\mu(w)}. \tag{D4}$$

Large systems. In a region with ℓ sites and $s \le k$ particles, the number of sites read in Z is the number of particles found plus the number of empty sites read in Z, which is a $\text{Bin}(\ell - s, 1-r)$ variable. Its distribution differs from $\text{Bin}(\ell, 1-r)$ by at most $O(k[\ell r(1-r)]^{-1/2})$ in total variation, because adding one trial to a binomial variable changes its distribution by at most its largest probability. Up to an error of this order, uniformly for r in a compact subinterval of $(0,1)$, $Q_x(w)$ is therefore the product of a distribution of the numbers of empty sites read in Z that is the same for all classes and of the distribution of the numbers $(y_{\text{in}}, y_{\text{out}})$ of particles found, each particle being found independently with probability $1-r$. Since $|\sum_x \mu_1(x)v_x Q_x| \le Q_{\mu_1} \le Q_\mu/\kappa$, the sum $\mathcal{S}_\kappa$ changes by an amount of the same order when the Q_x are replaced by these products, and we evaluate it for the products. If fewer than $k-1$ particles are found under μ_1, a particle that was not found is equally likely to be inside or outside, so the terms of $\sum_x \mu_1(x)v_x Q_x(w)$ cancel. If all are found, which has probability $(1-r)^{k-1}$, the sign is $(-1)^{y_{\text{in}}}$ and y_{in} is a $\text{Bin}(k-1, 1/2)$ variable. Under μ_2, exactly $k-1$ of the k particles are found with probability $kr(1-r)^{k-1}$, and y_{in} then has the same distribution. Hence, in the limit,

$$\mathcal{S}_\kappa = \frac{(1-r)^{k-1}}{\kappa + (1-\kappa)kr},$$

and letting κ tend to 0 or setting $\kappa = 1$ gives $\liminf C_r \geq \max(1, kr)/[r^2(1-r)^{k-1}]$, uniformly on compact subintervals of $(0, 1)$. The right-hand side is smallest at $r = 1/k$, where it equals $k^2[k/(k-1)]^{k-1}$.

Values of r near 0 *and* 1. It remains to exclude that the minimum over r is approached at the ends of the interval, where the convergence is not uniform. For every size $C_r \geq r^{-2}$, because the conditions imply $\sum_w Q_x G^2 \geq 1$; this diverges as $r \to 0$. For r near 1 we use the states of Lemma 13 with one particle at a random site of one of two sets of s_0 sites, inside and outside the string, and the other $k-2$ particles at fixed sites. Without readout errors a site is read in Z with probability $1-r$, and if the site of the particle is read its set is known. Otherwise the particle is equally likely to be at any of the sites that were not read in Z, and the posterior mean of the sign is the mean of the signs of these sites. As in the proof of Lemma 12, the mean square of the posterior mean is

$$\frac{2s_0(1-r) - (1-r)^{2s_0}}{2s_0 - 1} \leq 2(1-r), \tag{D5}$$

and the argument of Lemma 13, with $m_0 = 1$, $\lambda = 1$ and this bound in place of η, gives $C_r \geq 1/[4(1-r)]$ for every size with $s_0 \geq 1$. The minimum over r may therefore be restricted to a compact subinterval of $(0, 1)$, which completes the proof. □

Appendix E: Numerical details

Convex program. For each transfer operator, each position of its endpoints and each value of r, we compute numerically a function G that satisfies the conditions of the program (5) and a pair (μ, ν) for Eq. (D3), and we evaluate the resulting upper and lower bounds on C_r in double-precision arithmetic with explicit bounds on rounding errors. The two bounds agree to the digits given in Table II and in the figures. The cost quoted for the explicit estimator of Lemma 11 is the largest value of Eq. (C2) over the number s of occupied sites, divided by $r^2(1-2e)^4$ and minimized over r.

The minima over r in Table II are confirmed as follows. Every term of $Q_x(w)$ with $z = z_{\text{in}} + z_{\text{out}}$ sites read in Z contains the factor $r^{n-2q-z}(1-r)^z$. With $\bar{Q}_x$ denoting Q_x without this factor, a pair (μ, ν) that does not depend on r gives, after the scale of ν has been optimized,

$$\frac{1}{C_r} \leq \frac{\lambda^{4q}}{\left(\sum_x \nu_x \nu_x\right)^2} \sum_z r^{n-z}(1-r)^z \, \bar{\mathcal{S}}_z,$$
$$\bar{\mathcal{S}}_z = \sum_{w:\, z(w)=z} \frac{\left[\sum_x \nu_x \bar{Q}_x(w)\right]^2}{\bar{Q}_\mu(w)} \tag{E1}$$

for all r, with $\lambda = 1 - 2e$. The right-hand side is a polynomial in r, and finitely many such pairs, each used on an interval of values of r, bound C_r from below on the whole range $0 < r < 1$.

Simulated matrix elements. The systems are $(n, k) = (8, 4)$ and $n = 12$ and 16 with $k = 2$, 3 and 4. The three states are a normalized vector of the k-particle sector with independent complex Gaussian amplitudes; a Slater determinant of k orbitals drawn from the Haar measure; and the ground state of an open chain with nearest-neighbor hopping of amplitude -1, nearest-neighbor density interaction 1.3 and on-site potential $0.27\cos(1.37\,i)$ at the sites $i = 1, \ldots, n$, an arbitrary choice without spatial symmetry. The transfers of two particles have their endpoints at the sites $(1, 2, n-1, n)$, $(1, \lfloor n/3 \rfloor + 1, \lfloor 2n/3 \rfloor + 1, n)$ and $(2, 4, n-3, n-1)$, with particles created at the first and third of them. We draw 24 independent batches of 48 copies from the exact outcome distributions. The optimal and the string estimator use the same outcomes, taken with $r = 1/k$ at each error rate, and orbital rotations are ideal. In Fig. 4 the squared error of each matrix element is averaged over the batches, the standard error is taken across the batches, and the three states have equal weights.

Hubbard chains. The Hubbard Hamiltonian with periodic boundary conditions is diagonalized at $U/t = 4$ in the sector with two particles per spin for $L = 24$ and four per spin for $L = 16$; in both cases the ground state is nondegenerate and has zero total momentum. Each of 20 independent data sets consists of 200 000 copies, and its first 2 000 and 20 000 copies are used for the smaller numbers of copies. With $\kappa_j = 2\pi j/L$, the error of a data set is $[L^{-1}\sum_j |\widehat{n}(\kappa_j) - n(\kappa_j)|^2]^{1/2}$. Each copy serves one spin species, chosen with equal probabilities, and for single-qubit readout the terms with $i = j$ in $n(\kappa)$ are replaced by their known sum, the number of particles.

Basis sets. The estimator is a function of the numbers of particles recorded inside and outside the string, with $r = 1/k$; numbers above 16 are treated as 16, which restricts the choice of estimators and leaves the bound valid. It satisfies the conditions of the program (5), so that its largest second moment over the classes is an upper bound on the cost for every state. The bound quoted is the largest value over all positions of the endpoints.

Ternary-tree encoding. The n qubits are the nodes of a complete ternary tree filled level by level. Each of the $2n+1$ paths from the root to a vacant child position defines a Pauli string, which acts on every qubit of the path with X, Y or Z according to the branch taken there [21]; from left to right, the first $2n$ strings are the Majorana operators, and consecutive pairs form the modes. A hopping term with real or imaginary g is $a/2$ times the sum of two Pauli strings, of weights w_1 and w_2. With uniformly random Pauli bases a string of weight w is read with probability 3^{-w}, and readout errors attenuate it by $(1-2e)^w$, so that the normalized second moment of the estimate is at most $(s^{w_1} + s^{w_2})^2/8$ with $s = \sqrt{3}/(1-2e)$. The values quoted are the largest over all pairs of modes and both phases, for which $w_1 = w_2 = 7$ at $n = 24$ and $w_1 = w_2 = 8$ at $n = 48$.

[1] N. C. Rubin, R. Babbush, and J. McClean, Application of fermionic marginal constraints to hybrid quantum algorithms, New J. Phys. **20**, 053020 (2018).

[2] W. J. Huggins, J. R. McClean, N. C. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush, Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers, npj Quantum Inf. **7**, 23 (2021).

[3] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[4] A. Zhao, N. C. Rubin, and A. Miyake, Fermionic Partial Tomography via Classical Shadows, Phys. Rev. Lett. **127**, 110504 (2021).

[5] B. O'Gorman, Fermionic tomography and learning (2022), arXiv:2207.14787 [quant-ph].

[6] K. Wan, W. J. Huggins, J. Lee, and R. Babbush, Matchgate Shadows for Fermionic Quantum Simulation, Commun. Math. Phys. **404**, 629 (2023).

[7] G. H. Low, Classical shadows of fermions with particle number symmetry (2022), arXiv:2208.08964 [quant-ph].

[8] M. West, S. Y. Chang, L. Coffman, M. Larocca, and M. Cerezo, Practical fermionic shadows enabled by improved sample-complexity bounds (2026), arXiv:2609.40173 [quant-ph].

[9] Y. Koizumi, K. Wada, T. P. Takama, and N. Yoshioka, Provably Efficient Learning of Fermionic Correlations under Particle-Number Symmetry (2026), arXiv:2606.30601v2 [quant-ph].

[10] M. West, M. Cerezo, and M. Larocca, Particle-preserving fermionic shadows with mode-independent sample complexity (2026), arXiv:2606.27254 [quant-ph].

[11] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, Phys. Rev. Lett. **120**, 110501 (2018).

[12] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions, Phys. Rev. Applied **9**, 044036 (2018).

[13] Google AI Quantum and Collaborators, Hartree-Fock on a superconducting qubit quantum computer, Science **369**, 1084 (2020).

[14] J. Majsak, D. McNulty, and M. Oszmaniec, A simple and efficient joint measurement strategy for estimating fermionic observables and Hamiltonians, npj Quantum Inf. **11**, 61 (2025).

[15] P. Jordan and E. Wigner, Über das Paulische Äquivalenzverbot, Z. Phys. **47**, 631 (1928).

[16] R. King, D. Gosset, R. Kothari, and R. Babbush, Triply Efficient Shadow Tomography, PRX Quantum **6**, 010336 (2025).

[17] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask), Commun. Math. Phys. **328**, 303 (2014).

[18] E. H. Lieb and F. Y. Wu, Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension, Phys. Rev. Lett. **20**, 1445 (1968).

[19] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus, New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community, J. Chem. Inf. Model. **59**, 4814 (2019).

[20] M. Ippoliti, Classical shadows based on locally-entangled measurements, Quantum **8**, 1293 (2024).

[21] Z. Jiang, A. Kalev, W. Mruczkiewicz, and H. Neven, Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning, Quantum **4**, 276 (2020).

[22] S. N. Hearth, M. O. Flynn, A. Chandran, and C. R. Laumann, Efficient Local Classical Shadow Tomography with Number Conservation, Phys. Rev. Lett. **133**, 060802 (2024).

[23] P. Naldesi, A. Elben, A. Minguzzi, D. Clément, P. Zoller, and B. Vermersch, Fermionic Correlation Functions from Randomized Measurements in Programmable Atomic Quantum Devices, Phys. Rev. Lett. **131**, 060601 (2023).

[24] J. Denzler, A. A. Mele, E. Derbyshire, T. Guaita, and J. Eisert, Learning Fermionic Correlations by Evolving with Random Translationally Invariant Hamiltonians, Phys. Rev. Lett. **133**, 240604 (2024).

[25] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, Nearly Optimal Measurement Scheduling for Partial Tomography of Quantum States, Phys. Rev. X **10**, 031064 (2020).

[26] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo, Measurements of Quantum Hamiltonians with Locally-Biased Classical Shadows, Commun. Math. Phys. **391**, 951 (2022).

[27] H.-Y. Huang, R. Kueng, and J. Preskill, Efficient Estimation of Pauli Observables by Derandomization, Phys. Rev. Lett. **127**, 030503 (2021).

[28] B. Wu, J. Sun, Q. Huang, and X. Yuan, Overlapped grouping measurement: A unified framework for measuring quantum states, Quantum **7**, 896 (2023).

[29] L. E. Fischer, T. Dao, I. Tavernelli, and F. Tacchino, Dual-frame optimization for informationally complete quantum measurements, Phys. Rev. A **109**, 062415 (2024).

[30] J. Malmi, K. Korhonen, D. Cavalcanti, and G. García-Pérez, Enhanced observable estimation through classical optimization of informationally overcomplete measurement data: Beyond classical shadows, Phys. Rev. A **109**, 062412 (2024).

[31] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).

[32] S. McArdle, X. Yuan, and S. Benjamin, Error-Mitigated Digital Quantum Simulation, Phys. Rev. Lett. **122**, 180501 (2019).

[33] S. E. Smart and D. A. Mazziotti, Lowering tomography costs in quantum simulation with a symmetry projected operator basis, Phys. Rev. A **103**, 012420 (2021).

[34] G. Davis and J. K. Freericks, Hidden Rotation Symmetry of the Jordan–Wigner Transformation and Its Application to Measurement in Quantum Computation, Symmetry **18**, 251 (2026).

[35] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, Phys. Rev. A **103**, 042605 (2021).

[36] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, Robust Shadow Estimation, PRX Quantum **2**, 030348 (2021).

[37] A. Zhao and A. Miyake, Group-theoretic error mitigation enabled by classical shadows and symmetries, npj Quantum Inf. **10**, 57 (2024).

[38] W. Matthews, S. Wehner, and A. Winter, Distinguishability of Quantum States Under Restricted Families of Measurements with an Application to Quantum Data Hiding, Commun. Math. Phys. **291**, 813 (2009).

[39] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Hiding Bits in Bell States, Phys. Rev. Lett. **86**, 5807 (2001).

[40] Google Quantum AI and Collaborators, Quantum error correction below the surface code threshold, Nature **638**, 920 (2025).

[41] Google Quantum AI, Willow Spec Sheet, `https://quantumai.google/static/site-assets/downloads/willow-spec-sheet.pdf` (2024), published 9 December 2024.

[42] S. B. Bravyi and A. Y. Kitaev, Fermionic Quantum Computation, Ann. Phys. **298**, 210 (2002).
[43] G. González-García, F. M. Gambetta, and R. A. Santos, On the stability to noise of fermion-to-qubit mappings (2026), arXiv:2603.22141 [quant-ph].
[44] A. Chen, Data for "The price of locality in fermionic readout at fixed particle number", Zenodo, version 1.0.0, doi:10.5281/zenodo.24371203 (2026).
[45] F. L. A. Visconti, Hilbert–Schmidt Norm Estimates for Fermionic Reduced Density Matrices, Ann. Henri Poincaré 10.1007/s00023-026-01686-z (2026).
[46] M. R. Christiansen, Hilbert–Schmidt Estimates for Fermionic 2-Body Operators, Commun. Math. Phys. **405**, 18 (2024).