C. Gross, and I. Bloch, Microscopic evolution of doped Mott insulators from polaronic metal to Fermi liquid, Science **374**, 82 (2021).

[44] M. Lebrat, M. Xu, L. H. Kendrick, A. Kale, Y. Gang, P. Seetharaman, I. Morera, E. Khatami, E. Demler, and M. Greiner, Observation of Nagaoka polarons in a Fermi–Hubbard quantum simulator, Nature **629**, 317 (2024).

[45] M. L. Prichard, B. M. Spar, I. Morera, E. Demler, Z. Z. Yan, and W. S. Bakr, Directly imaging spin polarons in a kinetically frustrated Hubbard system, Nature **629**, 323 (2024).

[46] W. Xu, W. R. McGehee, W. N. Morong, and B. DeMarco, Bad-metal relaxation dynamics in a Fermi lattice gas, Nature Communications **10**, 1588 (2019).

[47] B. Spivak, S. V. Kravchenko, S. A. Kivelson, and X. P. A. Gao, Colloquium: Transport in strongly correlated two dimensional electron fluids, Rev. Mod. Phys. **82**, 1743 (2010).

[48] A. F. Hebard, T. T. M. Palstra, R. C. Haddon, and R. M. Fleming, Absence of saturation in the normal-state resistivity of thin films of K_3C_{60} and Rb_3C_{60}, Phys. Rev. B **48**, 9945 (1993).

[49] O. Gunnarsson, Superconductivity in fullerides, Rev. Mod. Phys. **69**, 575 (1997).

[50] NVIDIA Corporation, *cuSOLVERMp: A High-Performance CUDA Library for Distributed Dense Linear Algebra*, NVIDIA Corporation, version 0.8.0 (CUDA 12), accessed August 1, 2026.

[51] W. Kohn, Theory of the insulating state, Physical Review **133**, A171 (1964).

[52] D. Poilblanc, Twisted boundary conditions in cluster calculations of the optical conductivity in two-dimensional lattice models, Physical Review B **44**, 9562 (1991).

[53] X. Zotos, F. Naef, and P. Prelovšek, Transport and conservation laws, Phys. Rev. B **55**, 11029 (1997).

[54] R. E. Prange and L. P. Kadanoff, Transport Theory for Electron-Phonon Interactions in Metals, Physical Review **134**, A566 (1964).

[55] N. F. Mott, Conduction in non-crystalline systems IX. the minimum metallic conductivity, Philosophical Magazine **26**, 1015 (1972).

[56] P. T. Brown, E. Guardado-Sanchez, B. M. Spar, E. W. Huang, T. P. Devereaux, and W. S. Bakr, Angle-resolved photoemission spectroscopy of a Fermi–Hubbard system, Nature Physics **16**, 26 (2020).

Appendix A: Model, Hilbert space, symmetries, and boundary conditions

Here we briefly provide several basic definitions of the model and discussed observables and outline the structure of the code.

1. Model and Hilbert space

We study the infinite-U Hubbard model on finite clusters of the square lattice,

$$H \;=\; -t \sum_{\langle ij\rangle,\sigma} \left(\tilde{c}^{\dagger}_{i\sigma} \tilde{c}_{j\sigma} + \text{h.c.} \right), \qquad \tilde{c}_{i\sigma} = c_{i\sigma}\,(1 - n_{i\bar{\sigma}}), \tag{A1}$$

where the Gutzwiller-projected operators $\tilde{c}_{i\sigma}$ enforce the no-double-occupancy constraint, and the hole doping is $x = N_h/N$ with N_h the number of holes and N the number of sites. For fixed electron number $N_e = N - N_h$, the dimension of the constrained Hilbert space is $\binom{N}{N_h} 2^{N_e}$, i.e., the number of hole placements times the number of spin configurations of the remaining singly occupied sites.

The symmetries used in the calculation are: (i) $U(1)$ charge conservation, $[H, \hat{N}_e] = 0$; (ii) $SU(2)$ spin symmetry, $[H, \mathbf{S}] = 0$, of which we exploit S_z conservation together with the spin-flip degeneracy between $\pm S_z$ sectors, halving the number of blocks to be diagonalized; (iii) lattice translations, whose quantum numbers depend on the boundary conditions as specified below. Twists and staggering generically reduce the C_4 point group of the untwisted square clusters to C_2 or break it entirely; we therefore do not use point-group symmetries anywhere. The Hamiltonian is block diagonal in $(N_e, S_z, \mathbf{k})$, and each symmetry block is fully diagonalized with dense eigensolvers, retaining the complete many-body spectrum and all current matrix elements. The block structure and the largest dense-block dimensions are collected in Table I.

2. Cluster geometries

Three cluster geometries are used for the transport and thermodynamics results (Fig. 6): the 4×4 square torus, spanned by $\mathbf{L}_1 = (4, 0)$ and $\mathbf{L}_2 = (0, 4)$; the *staggered* 4×4 torus ($s = 1$), in which the y boundary is reconnected with a unit shift, $(x, N_y) \sim (x + s,\, 0)$, corresponding to $\mathbf{L}_2 = (-s, 4)$; and the 18-site tilted (Betts) cluster [37], spanned by $\mathbf{L}_1 = (3, 3)$ and $\mathbf{L}_2 = (3, -3)$. The 4×4 torus is well known to be a pathological geometry at the non-interacting level: its single-particle spectrum has a high degree of fine-tuned degeneracy, which collapses many-body levels onto a few distinct energies and enhances finite-size effects. Tilting (in the spirit of the Betts construction) and staggering modify the set of allowed momenta and lift these fine-tuned degeneracies; comparing the $s = 0$ and $s = 1$ sectors of the 4×4 cluster, and the 4×4 against the 18-site cluster, therefore provides independent finite-size diagnostics (App. C 3). The variational caging analysis of App. H additionally uses a 2×8 ladder with one hole.

3. Twisted and staggered boundary conditions

Twisted boundary conditions $\boldsymbol{\theta} = (\theta_x, \theta_y)$ are implemented by attaching phases to the hopping matrix elements, $t \to t\, e^{-i\boldsymbol{\theta}\cdot\boldsymbol{\delta}}$, with $\boldsymbol{\delta}$ the bond vector. For the simple torus the allowed momenta are

$$\mathbf{k}(\boldsymbol{\theta}) = \left(\frac{2\pi}{L_x} n_x + \theta_x,\ \frac{2\pi}{L_y} n_y + \theta_y\right), \tag{A2}$$

while on the staggered torus, $(x, N_y) \sim (x + s, 0)$,

$$\mathbf{k}(\boldsymbol{\theta}, s) = \left(\frac{2\pi}{L_x} n_x + \theta_x,\ \frac{2\pi}{L_y}\left(n_y + \frac{n_x s}{L_x}\right) + \theta_y\right). \tag{A3}$$

Equation (A3) satisfies the boundary conditions of the torus spanned by $\mathbf{L}_1 = (L_x, 0)$, $\mathbf{L}_2 = (-s, L_y)$: $\mathcal{T}_x^{L_x} = 1$ and $\mathcal{T}_y^{L_y}\mathcal{T}_x^{-s} = 1$. Because the twist enters only through the hopping phases, the eigenstates are strictly periodic on the torus, and the translation eigenvalues involve only the combination $\mathbf{k} - \boldsymbol{\theta}$. The closure conditions therefore require $(\mathbf{k} - \boldsymbol{\theta}) \cdot \mathbf{L}_{1,2} \in 2\pi\mathbb{Z}$, which Eq. (A3) satisfies for every twist:

$$(k_x - \theta_x)\, L_x = 2\pi n_x, \qquad (k_y - \theta_y)\, L_y - (k_x - \theta_x)\, s = 2\pi\left(n_y + \frac{n_x s}{L_x}\right) - \frac{2\pi n_x s}{L_x} = 2\pi n_y.$$

Staggering thus reshuffles the allowed momenta at fixed twist, so the $s = 0$ and $s = 1$ sectors provide independent finite-size samples.

For twist averaging we use a 5×5 "golden-ratio" grid,

$$\theta_x \in \left\{\frac{2\pi(a + \varphi)}{5}\right\}_{a=0}^{4}, \qquad \theta_y \in \left\{\frac{2\pi(b + \varphi)}{5}\right\}_{b=0}^{4}, \qquad \varphi = \frac{1 + \sqrt{5}}{2}, \tag{A4}$$

i.e., 25 twists per boundary-condition sector. The irrational offset φ keeps every twist away from high-symmetry commensurate points, avoiding the fine-tuned degeneracies discussed above. The same grid is used for the 4×4 cluster in both boundary-condition sectors ($s = 0$ and $s = 1$) and, in full, for the 18-site cluster (25 twists); the $x = 3/16$ data shown in Fig. 4 are for a single irrational twist.

4. Numerical implementation

We use a custom Fortran exact-diagonalization code developed for this work. The code diagonalizes each dense symmetry block, and its structure is organized to make this full diagonalization and the subsequent spectral sums feasible at the sizes of Table I.

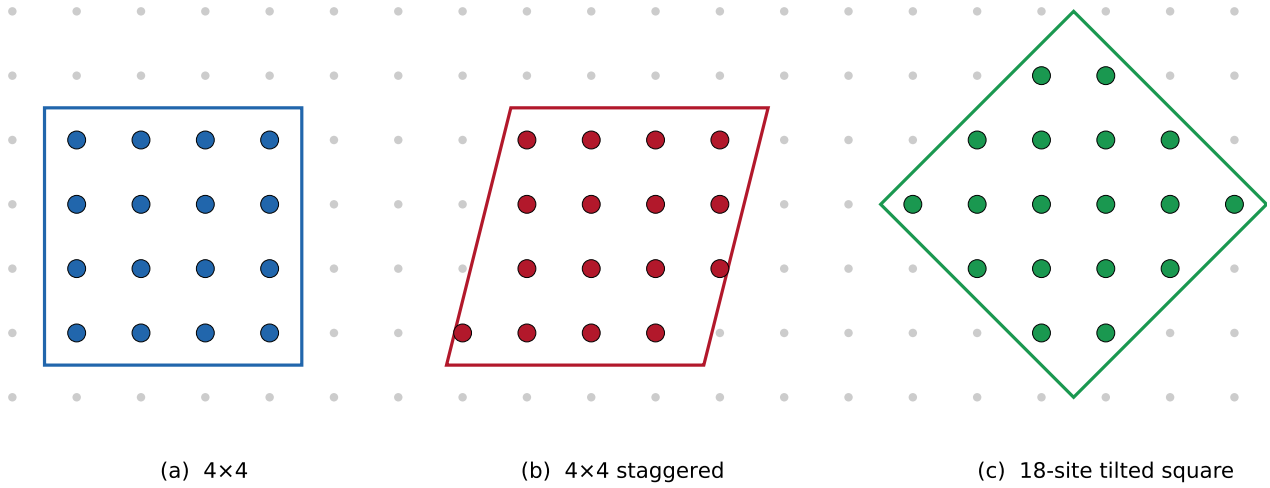


Figure 6. **Cluster geometries.** Representative finite clusters used in the exact-diagonalization calculations: (**a**) the 4×4 square torus, spanned by $\mathbf{L}_1 = (4, 0)$, $\mathbf{L}_2 = (0, 4)$; (**b**) the staggered 4×4 torus ($s = 1$), in which the y boundary is reconnected with a unit shift, $\mathbf{L}_2 = (-s, 4)$, modifying the allowed momenta according to Eq. (A3); (**c**) the 18-site tilted square Betts cluster, $\mathbf{L}_1 = (3, 3)$, $\mathbf{L}_2 = (3, -3)$.

Each (N_e, S_z) sector is generated directly in the constrained (no-double-occupancy) Fock basis, with the Hamiltonian and current operators stored and applied sparsely. When an operator maps one basis state to another, the target index is found by binary search on a lexicographically sorted key map, e.g. a standard sorted-basis lookup, which keeps operator assembly near-linear in the sector dimension. Translation symmetry, carrying the attendant fermionic signs, then blocks each sector by lattice momentum $\mathbf{k}(\boldsymbol{\theta}, s)$; together with particle number and S_z (with the $\pm S_z$ degeneracy) this leaves a set of independent dense blocks whose largest dimension controls the cost and feasibility of the computation.

Every dense block is diagonalized in full on GPUs with the distributed multi-GPU eigensolver NVIDIA `cuSOLVERMp` [50]. The eigenpairs are then consumed in a single streaming pass, i.e., for each batch of eigenstates the current and kinetic-energy matrix elements are formed by sparse scatters and dense rotations, and their contributions to the conductivity and to the thermodynamic traces are accumulated on the fly, simultaneously for all temperatures, frequencies, and broadenings. The full spectral matrices are therefore never stored, which is what keeps the memory footprint manageable at the largest block sizes.

Cluster	N_h	x	$\dim \mathcal{H}_{N_e}$	largest $(S_z, \mathbf{k})$ block	twists / role
4×4 ($s = 0$ and $s = 1$)	0	0	2^{16}	analytic, Eq. (B2)	χ_c reference
4×4 ($s = 0$ and $s = 1$)	1	1/16	524,288	6,435	25 twists
4×4 ($s = 0$ and $s = 1$)	2	2/16	1,966,080	25,740	25 twists
4×4	3	3/16	4,587,520	60,060	1 irrational twist; thermodynamics only
18-site Betts	0	0	2^{18}	analytic	χ_c reference
18-site Betts	1	1/18	2,359,296	24,310	25 twists; untwisted for spin-bond correlator
18-site Betts	2	2/18	10,027,008	109,430	1 irrational twist; thermodynamics only
4×5	1	1/20	10,485,760	92,378	25 twists
2×8 ladder	1	1/16	524,288	6,435	untwisted; caging analysis (App. H)

Table I. **Clusters, dopings, and Hilbert-space dimensions.** $\dim \mathcal{H}_{N_e} = \binom{N}{N_h} 2^{N_e}$ is the dimension of the constrained Hilbert space at fixed hole number; the largest dense block is the largest $(S_z, \mathbf{k})$ sector actually diagonalized (spin-flip degeneracy halves the number of S_z blocks). Common numerical parameters: temperature grid of 80 points, linear on $T \in [0.01, 2]\, t$; frequency grid of 201 points up to $\omega = 10\, t$, densified at low frequency; broadenings $\gamma/t \in \{0.005, 0.01, 0.02, 0.05, 0.1\}$.

Appendix B: Thermodynamic observables

All thermodynamic quantities are computed from the full spectra of the particle-number sectors. For each sector N_e and twist $\boldsymbol{\theta}$ the canonical partition function and free energy are

$$Z_{N_e}(T, \boldsymbol{\theta}) = \sum_n e^{-E_n(\boldsymbol{\theta})/T}, \qquad F_{N_e}(T, \boldsymbol{\theta}) = -T \ln Z_{N_e}(T, \boldsymbol{\theta}), \tag{B1}$$

with the sums running over all eigenstates of the sector. Unless stated otherwise, sector free energies and observables are averaged over the twist grid, Eq. (A4). The entropy and specific heat per site are obtained by numerical differentiation of the exact $F(T)$ and $E(T) = \langle H \rangle_T$ on the temperature grid, $S = -\partial F/\partial T/N$ and $C_V = \partial E/\partial T/N$. The uniform spin susceptibility is computed as in the main text, $\chi_s = (\langle S_z^2 \rangle - \langle S_z \rangle^2)/(NT)$.

The charge compressibility is obtained from the discrete second difference of the sector free energies with respect to the electron number,

$$\chi_c^{-1}(T) \;=\; N\Big[F_{N_e+1}(T) - 2F_{N_e}(T) + F_{N_e-1}(T)\Big], \tag{B2}$$

evaluated at the electron number N_e corresponding to the doping of interest. For the single-hole dopings the half-filled sector enters as one of the neighboring sectors; at $U = \infty$ it has no charge dynamics, so its free energy is exact, $F_N = -NT \ln 2$ (free spins). Concretely, $\chi_c(x{=}1/18)$ uses the

18-site sectors $N_e = 18, 17, 16$, $\chi_c(x{=}1/16)$ uses the 4×4 sectors $N_e = 16, 15, 14$, and $\chi_c(x{=}2/16)$ uses $N_e = 15, 14, 13$. Each sector free energy is averaged over its available twists (Table I): all 4×4 sectors with $N_h \leq 2$ and the 18-site one-hole sector use the full 25-twist grid, while the outermost sectors of the $x = 1/18$ and $x = 2/16$ differences ($N_e = 16$ on the 18-site cluster and $N_e = 13$ on the 4×4 cluster) are available for a single twist.

Appendix C: Optical conductivity and the dc limit

1. Kubo formula and broadening

The longitudinal optical conductivity is computed from the Lehmann representation,

$$\sigma_\mu(\omega) \;=\; \pi\,\frac{1-e^{-\omega/T}}{\omega}\sum_{n,m}\frac{e^{-E_n/T}}{Z}\;|\langle m|J_\mu|n\rangle|^2\,\delta\big(\omega-(E_m-E_n)\big), \qquad \mu = x, y, \tag{C1}$$

with J_μ the (projected) current operator, using every eigenstate and every current matrix element of each symmetry block. Here and throughout the appendices, $\sigma(\omega)$ denotes the real (dissipative) part of the conductivity, and the reported conductivity is the average of the two diagonal components, $\sigma = (\sigma_x + \sigma_y)/2$ (the anisotropy is at most a few percent in the staggered-BC case).

The delta functions are broadened into Lorentzians, $\delta(\omega-\Delta E) \to \frac{1}{\pi}\,\gamma/[\gamma^2+(\omega-\Delta E)^2]$, and the calculation is repeated for $\gamma/t \in \{0.005, 0.01, 0.02, 0.05, 0.1\}$. The physically meaningful broadening window is bounded below by the finite-size level spacing and above by the temperature, $\Delta_{\rm fs} \lesssim \gamma \ll T$: for smaller γ the spectrum resolves into discrete peaks, while larger γ artificially redistributes spectral weight. In practice, to minimize the systematic loss of spectral weight caused by introducing broadening, we take the $\gamma \to 0$ limit by linear extrapolation in γ from the two smallest values, $\sigma(\omega,\gamma) = \sigma_0(\omega)+a_\omega\gamma+\mathcal{O}(\gamma^2)$. The results obtained are robust to other extrapolation schemes, e.g., using more broadening values or adding higher order terms in γ. Representative extrapolations for both clusters are shown in Fig. 7(a,b), and Fig. 7(c,d) shows the resulting dc resistivity obtained from the fixed-broadening spectra alongside the $\gamma \to 0$ extrapolation. In all cases we find that the γ dependence is regular and enables a systematic extrapolation.

We also note that the residual resistivity obtained from extrapolating the T-linear resistivity in the semi-quantum regime to $T = 0$ is reduced for larger systems (see Fig. 2), which suggests it is a finite-size artifact.

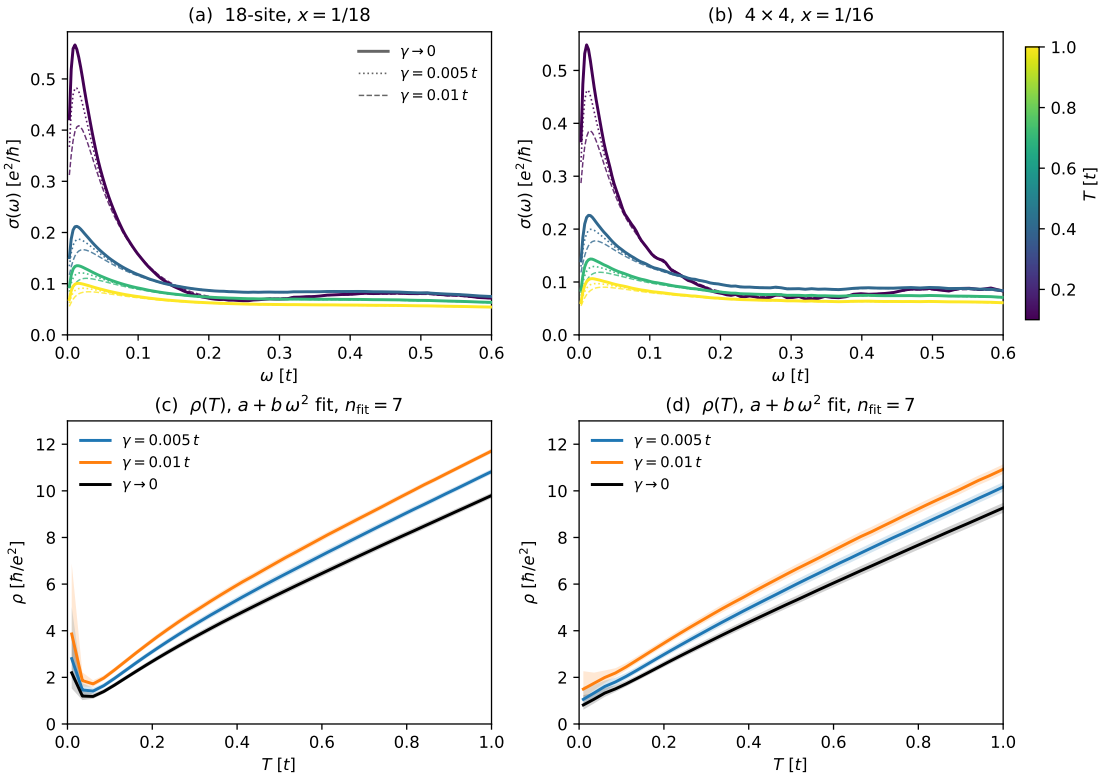


Figure 7. $\gamma \to 0$ **extrapolation of the optical conductivity and the dc resistivity.** (**a**) and (**b**) Optical conductivity at representative temperatures for the sampled broadenings $\gamma/t = 0.005, 0.01$, together with the $\gamma \to 0$ extrapolation from the two smallest values, for the 18-sites and 4×4 clusters, respectively. (**c**) and (**d**) dc resistivity obtained by applying the extraction of App. C 2 to the fixed-broadening spectra with $\gamma = 0.005\,t$ and $0.01\,t$ and to the $\gamma \to 0$-extrapolated spectra. Twist error becomes significant below $T \lesssim 0.08t$, indicating increased sensitivity to finite size effects, similarly to the non-staggered/staggered comparison in the main text.

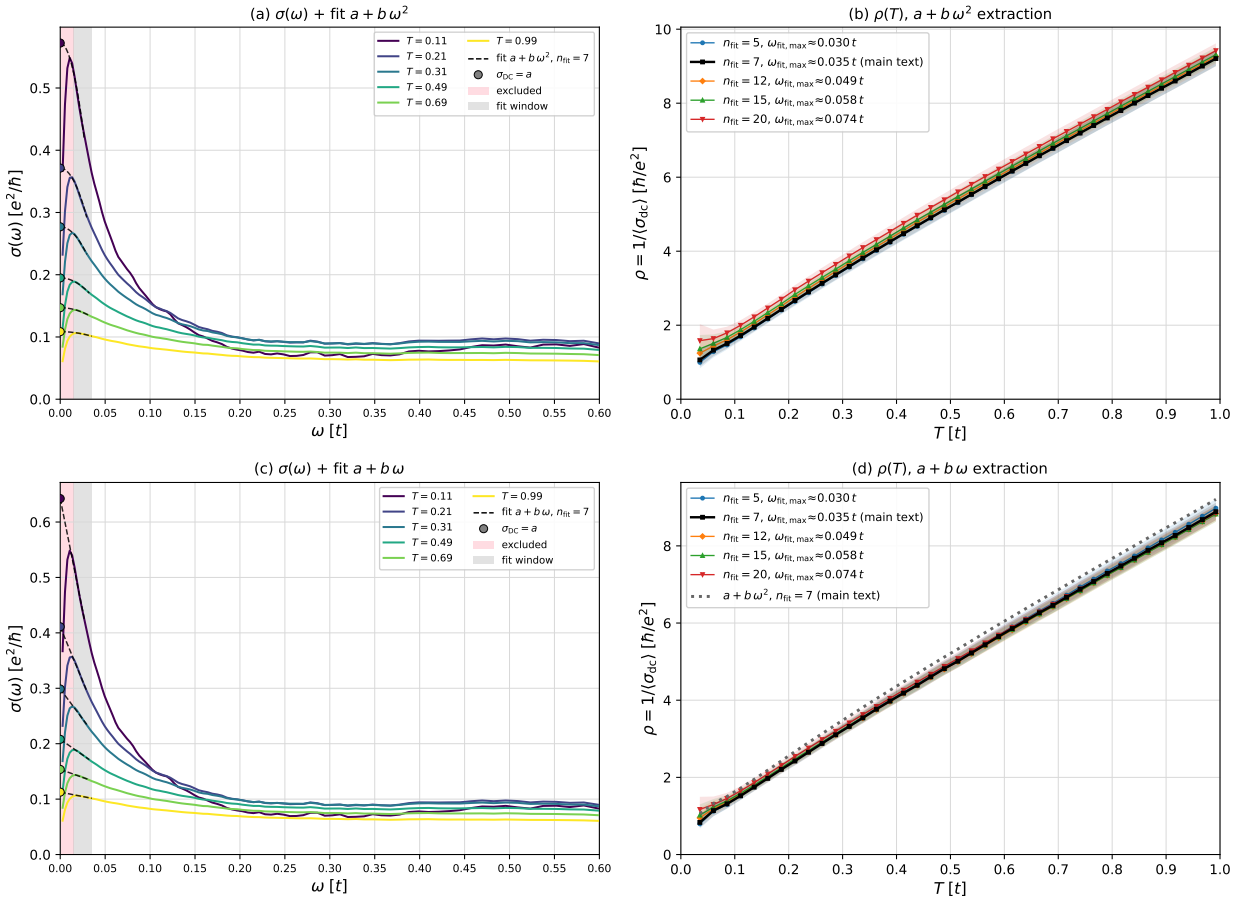


Figure 8. **Extraction of the dc conductivity and its robustness in representative case** 4×4**,** $x = 1/16$**.** (**a**) Twist-averaged, $\gamma \to 0$-extrapolated $\sigma(\omega)$ at representative temperatures. The pink band marks the level-repulsion downturn at $\omega < \omega_{\mathrm{peak}}$, excluded from the fit; the gray band marks the fit window $[\omega_{\mathrm{peak}}, \omega_{\mathrm{fit,max}}]$ used for the quadratic extrapolation $\sigma(\omega) = \sigma_{\mathrm{dc}} + b\,\omega^2$ (dashed lines; circles mark the extracted $\sigma_{\mathrm{dc}} = a$ at $\omega = 0$). (**b**) Twist-averaged resistivity $\rho(T) = 1/\langle\sigma_{\mathrm{dc}}\rangle_{\boldsymbol{\theta}}$ from the quadratic fit for several fit windows, labeled by their upper frequency $\omega_{\mathrm{fit,max}} \approx 0.03$–$0.07\,t$; the main-text choice is $\omega_{\mathrm{fit,max}} \approx 0.05\,t$. Bands show $1/(\langle\sigma_{\mathrm{dc}}\rangle \mp \mathrm{SEM})$ over the twists. (**c**,**d**) The same, using a *linear* extrapolation $\sigma(\omega) = \sigma_{\mathrm{dc}} + b\,\omega$; panel (**d**) overlays the quadratic main-text result (dotted). The two fit forms differ by a modest offset (~ 10, largest near $0.08t$, below which twist variance grows rapidly and finite-size effects dominate) but give the same T-linear slope, and within each form the fit window changes $\rho(T)$ by no more than about the twist standard error for $T > T_{\mathrm{fs}}$.

2. Extraction of the dc limit

At the lowest frequencies the finite cluster deviates from the expected regular $\omega \to 0$ behavior of the thermodynamic limit as finite cluster artifacts depletes the spectral weight in a narrow window

near $\omega = 0$, producing a downturn of $\sigma(\omega)$ below a temperature-dependent scale $\omega_{\rm peak}$ [31]. We therefore extract $\sigma_{\rm dc}$ by excluding the downturn window and fitting

$$\sigma(\omega) \;=\; \sigma_{\rm dc} + b\,\omega^2, \qquad \omega \in [\omega_{\rm peak},\, \omega_{\rm fit,max}], \tag{C2}$$

where $\omega_{\rm peak}$ is the local maximum of the per-twist, $\gamma \to 0$-extrapolated $\sigma(\omega)$ near $\omega = 0$, and the fit window extends to $\omega_{\rm fit,max} \approx 0.05\,t$ in the main text. The extraction is performed independently for each twist, and the resistivity is obtained from the twist-averaged conductivity, $\rho(T) = 1/\langle\sigma_{\rm dc}(\boldsymbol{\theta})\rangle_{\boldsymbol{\theta}}$, with the uncertainty band propagated from the standard error of $\langle\sigma_{\rm dc}\rangle$, i.e. $\rho_\pm = 1/(\langle\sigma_{\rm dc}\rangle \mp \mathrm{SEM})$. The extraction can be done when the excluded window is small compared with the physical scales, $\max(\omega_{\rm peak}, \gamma) \ll \min(T, t)$, which is satisfied throughout the reported temperature range.

Figure 8 tests the stability of this procedure. Varying the fit window over $\omega_{\rm fit,max} \approx 0.03$–$0.07\,t$ [Fig. 8(b)] changes $\rho(T)$ by no more than about the twist standard error for $T > T_{\rm fs}$. Replacing the quadratic fit by a linear one, $\sigma(\omega) = \sigma_{\rm dc} + b\,\omega$ [Fig. 8(c,d)], shifts $\rho(T)$ downward by a small constant amount but leaves the T-linear slope essentially unchanged. The characteristics of the semi-quantum regime discussed in the main text, i.e., the magnitude of ρ, its approximate T-linearity, and its slope, are thus insensitive both to the fit window and to the fit form.

3. More details on finite size effects

We further assess the finite-size dependence of the transport results in Fig. 9. Panel (a) shows the residual resistivity ρ_0, obtained from a linear fit $\rho(T) = \rho_0 + AT$ over the temperature window indicated in the caption; the results are insensitive to moderate changes of this window. In the one-hole sector, ρ_0 decreases systematically upon increasing the system size from $N = 16$ to 18 and 20. Although the doping simultaneously decreases along this sequence, the systematic suppression with system size suggests that the nonzero intercept is predominantly a finite-size effect, consistent with its disappearance toward the thermodynamic limit. The two-hole $N = 16$ result has a substantially larger intercept, indicating stronger finite-size corrections in this sector.

In contrast, the slope of the T-linear resistivity is considerably more stable. At low doping, where $\rho \sim T/x$ is expected, panel (b) shows xA. The one-hole results are nearly independent of system size, even as the doping changes from $x = 1/16$ to $1/20$, as expected for $\rho \sim T/x$. Thus, while the residual resistivity is progressively suppressed with increasing size, the T-linear slope retains the expected $1/x$ scaling. This stability suggests that the slope is an intrinsic property of the semi-quantum regime rather than a finite-size artifact. The two-hole result exhibits a larger

correction, consistent with the stronger size sensitivity seen in panel (a), but remains qualitatively consistent with the dilute one-hole results.

A separate finite-size feature appears in the low-frequency optical conductivity. We characterize the small downturn as $\omega \to 0$ by the position of the low-frequency maximum of $\sigma(\omega)$, denoted ω^*. As shown in panel (c), $\omega^*/T \lesssim 1$ for $T \gtrsim 0.08t$, encompassing most of the semi-quantum regime. The downturn is therefore confined to a frequency scale well below the thermal scale over the regime of interest. This separation of scales, together with the agreement between the different extrapolation procedures discussed above, shows that the extracted dc conductivity is insensitive to this narrow low-frequency feature. Again, the two-hole sector shows larger low-temperature corrections while following the same overall trend.

Finally, panel (d) shows the fraction of spectral weight not contained in the regular part of the optical conductivity,

$$f_D = 1 - \frac{\int_{0^+}^{\infty} d\omega\, \sigma(\omega)}{(\pi/2)|K|}, \tag{C3}$$

i.e., the Drude-weight fraction, which in the present context is a finite-size contribution that decreases systematically with increasing system size, remaining below 20% over the regime of interest, consistent with its expected vanishing in the thermodynamic limit. Here $(\pi/2)|K|$ is the total optical weight on the half-axis $\omega \geq 0$, with K the thermodynamic kinetic energy per site. A finite zero-frequency contribution is a well-known artifact of optical-conductivity calculations on finite clusters with twisted boundary conditions [51–53]: the many-body spectrum remains sensitive to flux threaded through the finite torus, giving a finite charge stiffness and hence a $\delta(\omega)$ peak, physically corresponding to persistent currents. Note that this zero-frequency contribution does not enter our determination of σ_{dc}, which is obtained by extrapolating the regular conductivity using data at $\omega \gtrsim \omega^*$, thereby excluding both the spurious low-frequency downturn and the $\omega = 0$ Drude contribution. The missing weight in panel (d) therefore does not correspond to unresolved finite-frequency transitions contaminating the dc limit; moreover, its systematic decrease with system size indicates that this weight redistributes, in the thermodynamic limit, to frequencies captured by the regular part. Finite-size effects relevant to σ_{dc} are instead assessed in the main text through system-size dependence, boundary-condition comparisons, and the variance over twists.

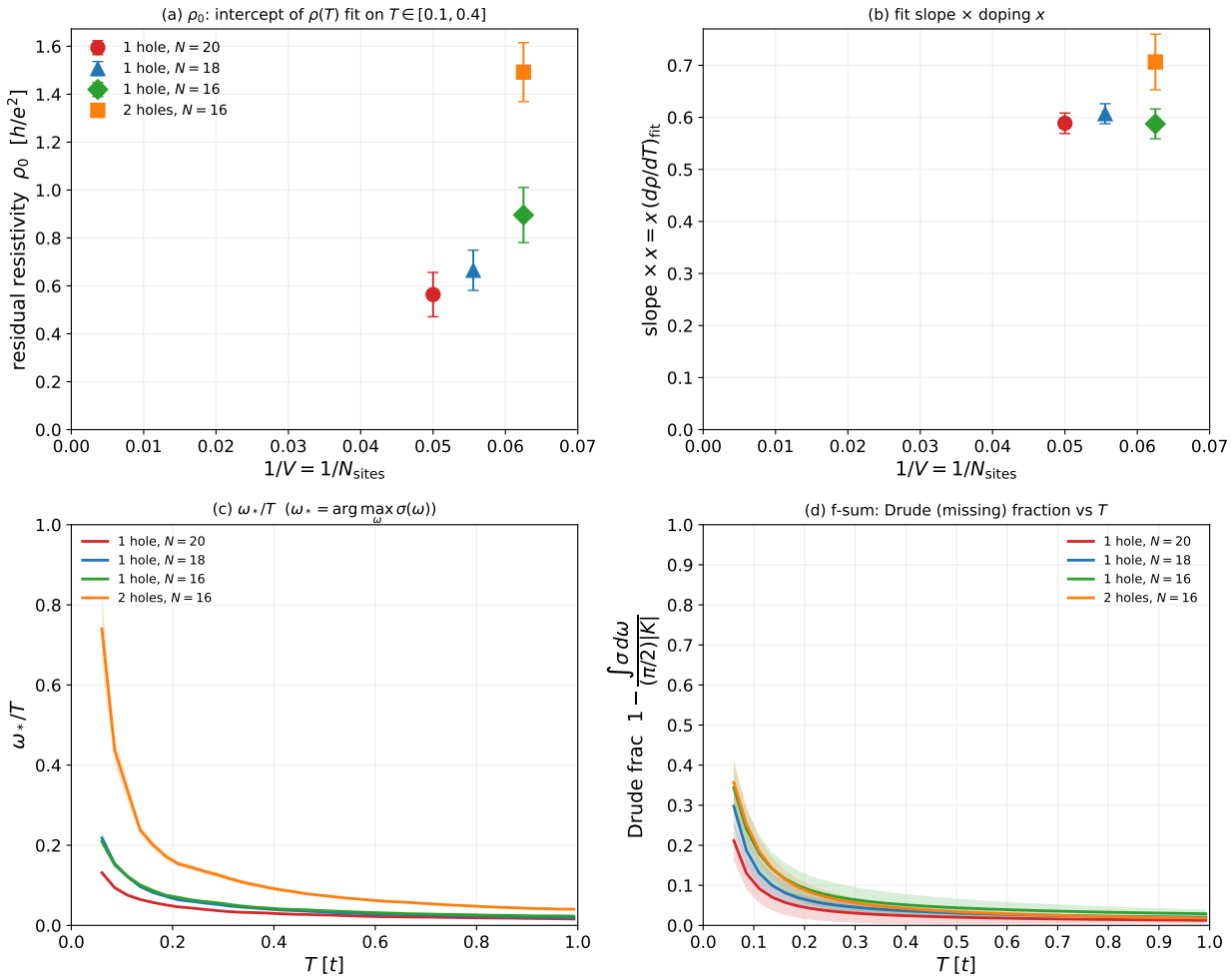


Figure 9. **Finite-size diagnostics of charge transport.** (A) Residual resistivity ρ_0 obtained from fits $\rho(T) = \rho_0 + AT$ over $0.1 \leq T/t \leq 0.4$, shown versus inverse system size. (B) Corresponding slope A multiplied by the doping x, testing the expected low-doping scaling $\rho \sim T/x$. (C) Position ω^* of the low-frequency maximum of $\sigma(\omega)$, normalized by T. (D) Drude-weight fraction, defined as the spectral weight missing from the regular optical conductivity. Results are shown for one hole on $N = 16, 18, 20$ clusters and two holes on $N = 16$.

4. Optical lineshape diagnostics

To characterize the narrowing of the low-frequency feature without assuming a specific lineshape, we define $\Delta\sigma(\omega, T) = \sigma(\omega, T) - \sigma(\omega, T_{\mathrm{ref}})$. We characterize its frequency extent by the high-frequency half-maximum edge $\Delta\omega$, defined by

$$\Delta\sigma(\Delta\omega, T) = \tfrac{1}{2}\Delta\sigma_{\mathrm{dc}}(T), \qquad \Delta\omega > \omega_{\mathrm{max}},$$

where $\omega_{\max}$ is the position of the maximum of $\Delta\sigma(\omega, T)$ [Fig. 10A]. In the main text we use $T_{\mathrm{ref}} = t$; the resulting $\Delta\omega(T)$ is shown in Fig. 3F.

We test the dependence of this construction on the choice of reference temperature in Fig. 10B. Solid and dashed curves show the raw $\Delta\omega(T)$ obtained using $T_{\mathrm{ref}} = t$ and $2t$, respectively. Changing T_{ref} shifts the absolute values somewhat while leaving the qualitative temperature dependence essentially unchanged. Unlike in Fig. 3F, the $x = 2/16$ values are shown here without the factor of $1/2$, so that the absolute frequency scale can be compared directly. The one-hole clusters exhibit an approximately linear decrease of $\Delta\omega$ upon cooling, whereas the two-hole result has a larger characteristic frequency scale and a qualitatively different temperature dependence, as discussed in the main text.

As a further check that $\Delta\omega$ meaningfully characterizes the low-frequency lineshape, Figs. 10C–E show $\Delta\sigma(\omega, T)/\Delta\sigma_{\mathrm{dc}}(T)$ versus $\omega/\Delta\omega(T)$ at representative temperatures. The peak height and high-frequency half-maximum point coincide by construction. Away from these normalization points, the one-hole curves exhibit an approximate lineshape collapse over a broader frequency range, while the two-hole lineshape shows larger deviations, consistent with its qualitatively distinct behavior.

5. Ioffe–Regel convention

For reference we state the convention used for the Ioffe–Regel line in the main text. For a two-dimensional Drude metal,

$$\sigma = \frac{e^2 n \tau}{m} = \frac{e^2}{2\pi\hbar}\, k_F \ell. \tag{C4}$$

The validity of the Boltzmann transport equation (for Fermi liquids or non-Fermi liquids [54]) requires the mean free path to exceed the Fermi wavelength, $\ell \gtrsim \lambda_F = 2\pi/k_F$, i.e., $k_F\ell/2\pi \gtrsim 1$, which translates into

$$\sigma_{\mathrm{IR}} = \frac{e^2}{\hbar} \qquad \Longleftrightarrow \qquad \rho_{\mathrm{IR}} = \frac{\hbar}{e^2}. \tag{C5}$$

Resistivities $\rho \gtrsim \hbar/e^2$ therefore signal transport beyond the Boltzmann transport regime and in particular do not host electronic quasiparticles. Note that this criterion is distinct from the Mott criterion $\ell > a$ [55], which is violated only at correspondingly larger resistivities.

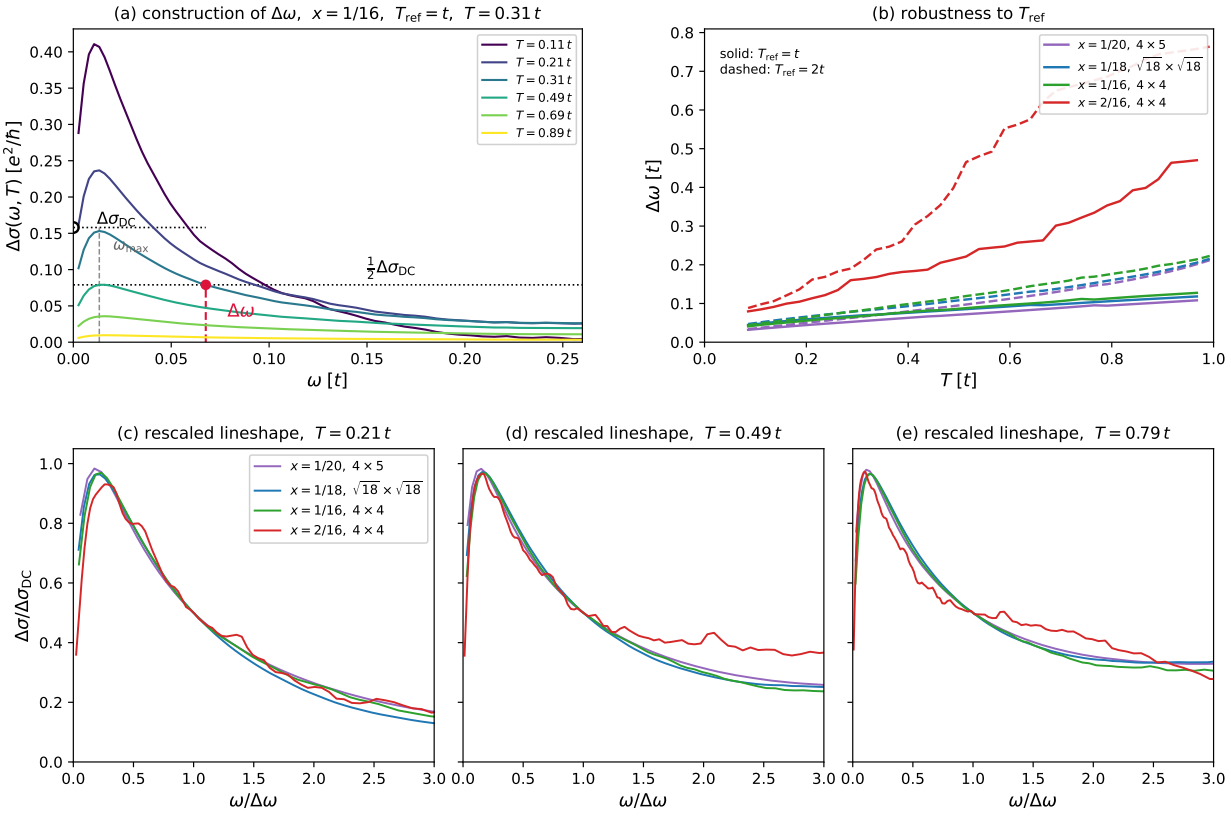


Figure 10. **Characterization of the low-frequency optical feature.** (A) Construction of the high-frequency half-maximum edge $\Delta\omega$ from $\Delta\sigma(\omega,T) = \sigma(\omega,T) - \sigma(\omega,T_{\rm ref})$ for $x = 1/16$ and $T_{\rm ref} = t$. The marked points indicate $\omega_{\rm max}$ and the high-frequency half-maximum crossing $\Delta\omega$. (B) Dependence on the reference temperature. Solid and dashed curves show the raw $\Delta\omega(T)$ obtained using $T_{\rm ref} = t$ and $2t$, respectively, for the four systems shown in Fig. 3. The $x = 2/16$ values are shown without the factor of 1/2 used for visual comparison in Fig. 3F. (C–E) Rescaled low-frequency lineshapes, $\Delta\sigma(\omega,T)/\Delta\sigma_{\rm dc}(T)$ versus $\omega/\Delta\omega(T)$, at the indicated representative temperatures.

Appendix D: Resistivity in a bubble approximation

Here we provide a limited examination of the importance of vertex corrections to the resistivity by comparing the full Kubo result with the "bubble" conductivity constructed from the single-particle spectral function. We restrict this comparison to the 4×4 cluster with one hole, $x = 1/16$. We compute the exact canonical finite-T *spin-summed* spectral function of the projected electrons

$$\tilde{c}_{\boldsymbol{k}\sigma} = \frac{1}{\sqrt{N}} \sum_j e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_j} c_{j\sigma}(1 - n_{j\bar{\sigma}}) \tag{D1}$$

by full diagonalization of the N_e-1, N_e, and N_e+1 sectors,

$$A(\boldsymbol{k},\omega,T)=\sum_{\sigma,n,m}\frac{e^{-E_n^{(N_e)}/T}}{Z_{N_e}}\Big[|\langle m^{(N_e-1)}|\tilde{c}_{\boldsymbol{k}\sigma}|n^{(N_e)}\rangle|^2\delta_\gamma\Big[\omega-(E_n^{(N_e)}-E_m^{(N_e-1)})\Big] \\ +|\langle m^{(N_e+1)}|\tilde{c}^\dagger_{\boldsymbol{k}\sigma}|n^{(N_e)}\rangle|^2\delta_\gamma\Big[\omega-(E_m^{(N_e+1)}-E_n^{(N_e)})\Big]\Big]. \tag{D2}$$

Here δ_γ is a Lorentzian broadening. Frequencies are not shifted by the chemical potential, so that the Fermi level occurs at $\omega=\mu(T)$. The projected spectral function obeys $\int d\omega\, A(\boldsymbol{k},\omega,T)=2-n=1+x$, giving 17/16 for the cluster considered here.

From $A(\boldsymbol{k},\omega,T)$ we evaluate the dc bubble conductivity [23],

$$\sigma_x^{\rm bub}(T)=\frac{\pi}{N}\sum_{\boldsymbol{k}}(v_{\boldsymbol{k}}^x)^2\int d\omega\,\left(-\frac{\partial f(\omega-\mu)}{\partial\omega}\right)\frac{A(\boldsymbol{k},\omega,T)^2}{2}, \tag{D3}$$

where $v_{\boldsymbol{k}}^x=2t\sin\tilde{k}_x$. The factor 1/2 follows from $\sum_\sigma A_\sigma^2=A^2/2$ for the spin-summed spectral function in the spin-symmetric state.

Because this comparison is not the primary focus of the present work and carries a substantially higher computational cost, we restrict the calculation to four boundary twists and a coarser frequency grid. This provides meaningful results for $T\gtrsim 0.2t$, sufficient for the limited purpose here of assessing the role of vertex corrections. Since $A(\boldsymbol{k},\omega,T)$ is evaluated canonically whereas Eq. (D3) contains a grand-canonical Fermi window, for each twist we determine μ from

$$\sum_{\boldsymbol{k}}\int d\omega\, f(\omega-\mu)A(\boldsymbol{k},\omega,T)=N_e \tag{D4}$$

with f the Fermi-Dirac distribution function. We restrict the comparison to temperatures for which the thermally sampled window of width $\sim T$ around μ is sufficiently broad that finite-size ambiguities in the precise location of the chemical potential are unimportant.

Figure 11 shows the resulting bubble resistivity together with the full Kubo result, alongside representative single-particle spectral functions entering Eq. (D3). The bubble resistivity, $\rho^{\rm bub}=1/\langle\sigma^{\rm bub}\rangle_{\boldsymbol{\theta}}$, differs quantitatively from the full Kubo resistivity of the finite cluster, while exhibiting a similar temperature dependence. Over $0.3\lesssim T/t\leq 2$,

$$\frac{\rho^{\rm bub}(T)}{\rho^{\rm Kubo}(T)}\approx 1.6, \tag{D5}$$

with little detectable temperature dependence. Thus we find that vertex corrections are important for quantitatively capturing the finite-cluster resistivity, although the bubble shows similar T-scaling over this range. The importance of such corrections is consistent with the quantitative discrepancy between 16-site FTLM and single-site DMFT resistivities reported for the doped Hubbard

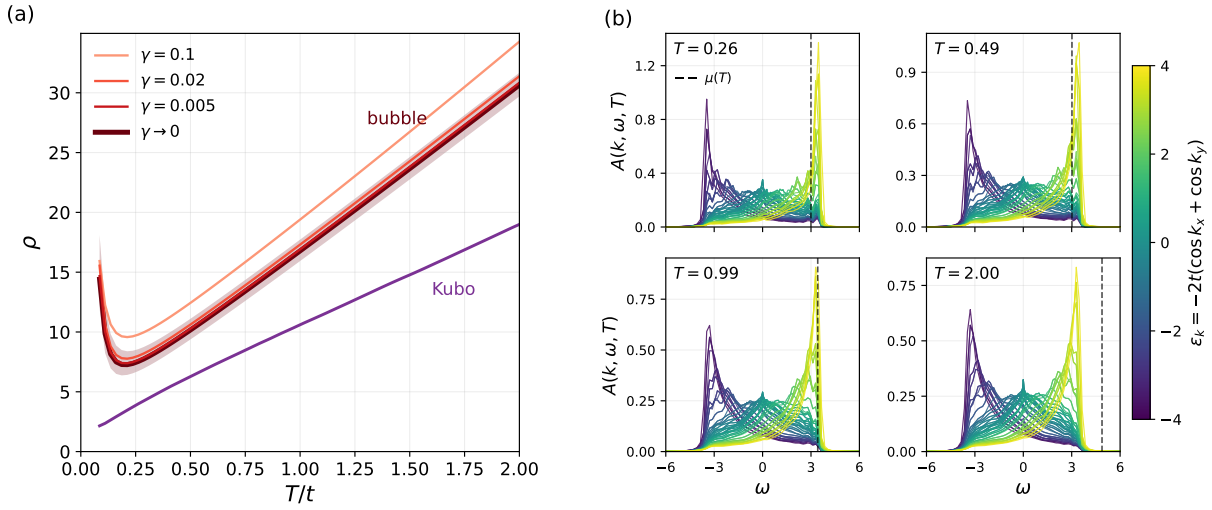


Figure 11. **Single-particle spectral function and bubble resistivity.** (**a**) Full Kubo resistivity and the vertex-free bubble resistivity ρ^{bub} obtained from Eq. (D3) for the 4×4 one-hole cluster. (**b**) Spin-summed spectral function $A(\boldsymbol{k}, \omega, T)$ entering Eq. (D3), shown for $\gamma = 0.02t$ at four representative temperatures. Curves correspond to the 64 twist-shifted momenta obtained by pooling the 16 momenta from each of the four boundary twists used in the bubble calculation; the curves are not averaged over twists. Lines are colored by the bare-band energy $\varepsilon_{\boldsymbol{k}} = -2t(\cos k_x + \cos k_y)$ of the corresponding twist-shifted momentum. Frequencies are not shifted by the chemical potential, and dashed vertical lines indicate $\mu(T)$ determined from Eq. (D4).

model in Ref. [12]. We defer a systematic study of these effects and a more detailed investigation of the single-particle spectral function, which is directly accessible in cold-atom Fermi–Hubbard systems through ARPES-type measurements [56], to future work.

Appendix E: Details on the hole-spin-bond correlator

1. Definition and normalization

Recall the connected hole-spin-bond correlator defined in Eq. (2) is

$$C_h(\mathbf{R}, \mathbf{e}; T) = \left\langle \rho_h(\mathbf{0})\, \mathbf{S}_{\mathbf{R}-\mathbf{e}/2} \cdot \mathbf{S}_{\mathbf{R}+\mathbf{e}/2} \right\rangle - \langle \rho_h(\mathbf{0}) \rangle \left\langle \mathbf{S}_{\mathbf{R}-\mathbf{e}/2} \cdot \mathbf{S}_{\mathbf{R}+\mathbf{e}/2} \right\rangle, \qquad \langle \rho_h(\mathbf{0}) \rangle = x = \frac{1}{N}, \quad \text{(E1)}$$

where $\mathbf{R}$ is the shortest displacement, accounting for periodic boundary conditions, from the origin to the bond center, and $\mathbf{e} = \hat{\mathbf{x}}, \hat{\mathbf{y}}$ specifies the bond orientation. We denote by $C_h(R, T)$ the value of C_h at $R = |\mathbf{R}|$. We normalize throughout by

$$C_0 = \frac{x^2}{2}, \quad \text{(E2)}$$

the $T \to 0$ value of (E1) on any bond away from the hole, derived below. The four $R = 1/2$ bonds incident on the conditioned hole site are called *incident*.

2. A sum rule

For brevity we denote $b = (i, i + \mathbf{e})$ the nearest-neighbor bond and $B = \sum_b \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}}$ the total nearest-neighbor bond operator, with the sum running over all $2N$ bonds. With exactly one hole, $\sum_{\mathbf{r}} \rho_h(\mathbf{r}) = 1$ as an operator identity, while both B and the thermal state are translation invariant. Hence

$$\langle \rho_h(\mathbf{0}) B \rangle = \frac{1}{N} \sum_{\mathbf{r}} \langle \rho_h(\mathbf{r}) B \rangle = \frac{1}{N} \langle B \rangle = x \langle B \rangle. \tag{E3}$$

Summing (E1) over all bonds therefore gives

$$\sum_b C_h(b; T) = \langle \rho_h(\mathbf{0}) B \rangle - x \langle B \rangle = 0, \tag{E4}$$

where $C_h(b; T)$ is shorthand for $C_h(\mathbf{R}_b, \mathbf{e}_b; T)$.

3. The Nagaoka ground state

As $T \to 0$, the $U = \infty$ one-hole ground state is the Nagaoka ferromagnet [33], with maximal total spin $S = (N - 1)/2$ and a uniformly delocalized hole, so that $\langle \rho_h(\mathbf{r}) \rangle = x$ for every $\mathbf{r}$. Two elementary observations then fix (E1) completely.

First, conditioned on the hole sitting at the origin, all remaining spins belong to the maximally polarized $S = (N - 1)/2$ multiplet. Hence, on any bond that avoids the origin,

$$\langle \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle_h = \frac{1}{4}, \tag{E5}$$

i.e.

$$\langle \rho_h(\mathbf{0})\, \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle = \frac{x}{4}. \tag{E6}$$

On an incident bond one endpoint is the conditioned hole site and therefore carries no spin, so the joint expectation vanishes identically as an operator identity.

Second, without conditioning on the hole position, either endpoint of a given bond is empty with probability x, so the probability that both endpoints are occupied is $1 - 2x$. In the fully polarized state this gives

$$\langle \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle = \frac{1}{4}(1 - 2x). \tag{E7}$$

Combining the two, the $T \to 0$ values are

$$C_h^{\text{non-inc}} = \frac{x}{4} - x\,\frac{1-2x}{4} = \frac{x^2}{2} = C_0, \qquad \frac{C_h(R=1/2, T\to 0)}{C_0} = \frac{-x(1-2x)/4}{x^2/2} = -\frac{1-2x}{2x}, \tag{E8}$$

which is -8 for $N = 18$. The first of these is Eq. (E2): in the fully polarized state the normalized correlator on every non-incident bond is exactly $+1$, independently of N and of its distance from the hole. The two limits in (E8) are not independent: with 4 incident and $2N-4$ non-incident bonds, the sum rule (E4) requires

$$4 \times \left[-\frac{N-2}{2}\right] + (2N-4) \times 1 = 0, \tag{E9}$$

which holds identically.

The incident shell is special because it contains no conditional local bond information. Since the conditional part vanishes there identically, (E1) reduces to minus the disconnected piece alone. Defining the orientation-averaged nearest-neighbor spin correlation by

$$\langle \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle_{\text{nn}} \equiv \frac{1}{2} \sum_{\mathbf{e}=\hat{\mathbf{x}},\hat{\mathbf{y}}} \langle \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle\,, \tag{E10}$$

the shell-averaged incident correlator obeys the exact identity

$$\frac{C_h(R=1/2, T)}{C_0} = -\frac{2}{x}\,\langle \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle_{\text{nn}} \qquad \text{at all } T. \tag{E11}$$

The deep blue column at $R = 1/2$ in Fig. 5A is therefore not a statement about the local bonds surrounding a conditioned hole; it directly tracks the global nearest-neighbor spin correlation, up to rescaling.

4. Formation of the ferromagnetic texture

Fig. 12A shows the five non-incident shells. All approach the common Nagaoka value $+1$ as $T \to 0$, but their evolution is quite different. The $R = \sqrt{5}/2$ shell, consisting of the eight bonds forming the 3×3 "spin cage" around the hole, is the only one that overshoots this value, peaking at intermediate T inside the semi-quantum regime. The more distant shells instead become negative at intermediate temperature. By the sum rule (E4), any enhancement of the connected correlation near the hole must be compensated elsewhere.

Since C_h also contains a temperature-dependent disconnected contribution, the overshoot above $+1$ does not mean that the local bond correlation exceeds that of the fully polarized state. To characterize the local texture more directly, we consider the conditional bond correlation

$$\langle \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle_h \equiv \frac{\langle \rho_h(\mathbf{0})\, \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{e}} \rangle}{x}, \tag{E12}$$

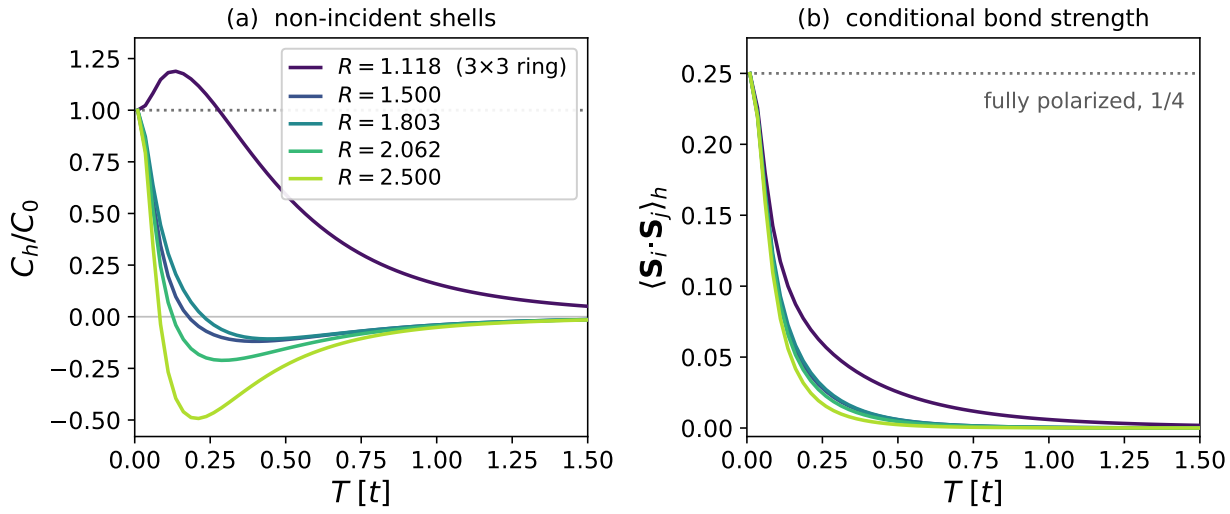


Figure 12. **Shell-resolved hole-spin-bond correlator, 18-site cluster, $x = 1/18$, untwisted.** Bonds are grouped by the distance R from the hole to the bond center; the incident shell $R = 1/2$, fixed exactly by Eq. (E11), is not shown. (**a**) C_h/C_0, Eq. (E1) normalized by its Nagaoka value $C_0 = x^2/2$ (dotted). (**b**) The conditional bond strength $\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_h = \langle \rho_h(0)\, \mathbf{S}_i \cdot \mathbf{S}_j \rangle / x$, same shells and colors, approaching the fully polarized 1/4 (dotted).

shown in Fig. 12B. It vanishes identically on the incident shell, approaches the fully polarized value 1/4 on all other shells as $T \to 0$, and is largest on the $R = \sqrt{5}/2$ ring. Its magnitude, however, remains unsaturated down to very low temperatures. In other words, below T_{quant}, the spins near the hole are clearly correlated ferromagnetically, but the local texture remains fluctuating and rather than fully polarized.

Appendix F: High-temperature expansions

The dashed high-temperature curves in the main-text figures (Figs. 2B and 4) follow from the leading high-temperature expansion (HTE) of the free energy, which we summarize here. At $T \gg t$ the canonical partition function of a cluster of V sites with N_e electrons is expanded as

$$Z_{N_e} = \Omega_{N_e} \left(1 - \beta \langle H \rangle_0 + \tfrac{\beta^2}{2} \langle H^2 \rangle_0 + \dots \right), \qquad \Omega_{N_e} = \binom{V}{N_e} 2^{N_e}, \tag{F1}$$

where $\langle \cdot \rangle_0$ denotes the infinite-temperature average and Ω_{N_e} counts the constrained configurations. Since H is purely off-diagonal in the occupation basis, $\langle H \rangle_0 = 0$; moreover $\langle H^3 \rangle_0 = 0$ on the bipartite clusters (the 4×4 torus and the 18-site Betts cluster), because their shortest closed loops are the four-bond plaquettes. The one caveat concerns the staggered ($s = 1$) torus and the 4×5

torus are not bipartite. Their shortest odd loops wind the boundary and have length $L_y + s = 5$ and $L_y = 5$, respectively. Odd moments therefore first enter at order β^5, well beyond the order kept below, and on the bipartite clusters they vanish identically. The second moment counts the allowed directed nearest-neighbor hops: a hop $i \to j$ requires site i occupied and site j empty, which at fixed N_e occurs with probability $N_e(V - N_e)/[V(V-1)]$; with $4V$ directed nearest-neighbor links this gives

$$a_2 \equiv \langle H^2 \rangle_0 = 4t^2 \, \frac{N_e \, (V - N_e)}{V - 1} \quad \xrightarrow{V \to \infty} \quad 4t^2 n(1-n) \, V, \tag{F2}$$

with $n = N_e/V = 1 - x$. From $F = -T \ln Z = -T \ln \Omega_{N_e} - a_2/(2T) + \dots$ one obtains the leading high-temperature forms, per site,

$$s(T) = s_\infty - \frac{2t^2}{T^2} \, n(1-n), \qquad s_\infty = -n \ln(n/2) - (1-n) \ln(1-n), \tag{F3}$$

$$c_V(T) = \frac{4t^2}{T^2} \, n(1-n), \tag{F4}$$

$$\chi_c^{-1}(T) = \frac{T}{n(1-n)} + \frac{4t^2}{T}. \tag{F5}$$

For comparison with the finite clusters at fixed particle number we use the finite-V expressions, obtained from Eq. (F1) without taking the thermodynamic limit:

$$s = \frac{1}{V} \ln \Omega_{N_e} - \frac{2t^2}{T^2} \, \frac{N_e(V - N_e)}{V(V-1)}, \tag{F6}$$

$$c_V = \frac{4t^2}{T^2} \, \frac{N_e(V - N_e)}{V(V-1)}, \tag{F7}$$

$$\chi_c^{-1} = VT \, \ln\!\left[\frac{(N_e + 1)(V - N_e + 1)}{N_e \, (V - N_e)} \right] + \frac{4Vt^2}{(V-1) \, T}, \tag{F8}$$

where the first term of Eq. (F8) is obtained by inserting $F_{N_e} = -T \ln \Omega_{N_e}$ into the definition Eq. (B2). These finite-cluster forms are the dashed curves shown in the main text.

For the spin sector, independent spin-1/2 moments give the Curie law with the main-text normalization, $T\chi_s = N_e/(4N) = (1 - x)/4$. In the opposite (Nagaoka) limit the ground-state multiplet has total spin $S_{\rm tot} = N_e/2$; averaging S_z^2 over the degenerate multiplet gives $\langle S_z^2 \rangle = S_{\rm tot}(S_{\rm tot} + 1)/3$, hence

$$T\chi_s \;=\; \frac{N_e(N_e + 2)}{12 \, N}. \tag{F9}$$

The departure of each observable from its high-temperature form defines the crossover scale $T_{\rm quant}$ operationally; as expected for a crossover (rather than a transition), the resulting scale varies somewhat between observables, $T_{\rm quant} \approx 0.5$–$0.8 \, t$.

Appendix G: Nagaoka ferromagnetism and spin stiffness

The lower boundary of the semi-quantum regime is set by the coherence scale $T_{\rm coh}$ associated with the ferromagnetic correlations of the low-doping ground state [33, 34]. Here we estimate the associated spin stiffness for a single hole and recall how it sets $T_{\rm coh} \sim xt$.

For one hole on a cluster of N sites, the ground state is the fully polarized Nagaoka ferromagnet, and the hole occupies the top of the single-particle band at $\mathbf{k}_0 = (\pi, \pi)$: with all $\mathbf{k}$ states but $\mathbf{k}_0$ filled by spin-aligned electrons,

$$E_0 = \sum_{\mathbf{k} \neq \mathbf{k}_0} \epsilon(\mathbf{k}) = -\epsilon(\mathbf{k}_0) = -4t, \qquad \epsilon(\mathbf{k}) = -2t\,(\cos k_x + \cos k_y). \tag{G1}$$

To extract the stiffness of the ferromagnet we impose a uniform spin twist of total angle ϕ across the x direction, i.e., a local spin rotation by $n_x\alpha$ with $\alpha = \phi/L$. The local "up" direction then rotates from site to site, and the overlap of neighboring spinors reduces the effective hopping of the polarized band,

$$t \;\to\; \langle \uparrow_{\mathbf{r}} \,|\, \uparrow_{\mathbf{r}+\hat{x}} \rangle\, t = \cos\Big(\frac{\phi}{2L}\Big)\, t, \tag{G2}$$

for x bonds only. Keeping the hole at $\mathbf{k}_0$, the ground-state energy shift is

$$\Delta E(\phi) = 2t \left[1 - \cos\Big(\frac{\phi}{2L}\Big)\right] \approx \frac{t\,\phi^2}{4L^2}. \tag{G3}$$

Matching to the continuum stiffness definition in $d = 2$, $\Delta E/L^d = (\rho_s/2)\,(\phi/L)^2$, gives

$$\rho_s = \frac{x\,t}{2}, \qquad x = \frac{1}{N}, \tag{G4}$$

i.e., a stiffness of order xt, carried by the mobile holes. At any finite temperature the two-dimensional ferromagnet is disordered, but with an exponentially growing correlation length,

$$\xi_{\rm FM}(T) \sim \exp\left(\frac{2\pi\rho_s}{T}\right), \tag{G5}$$

so that for $T \lesssim T_{\rm coh} \sim xt$ the holes effectively move through an ordered spin background as coherent spinless fermions.

Appendix H: Variational bound for the caging free energy

The caging construction of the main text is controlled by the thermodynamic variational principle. Let $H_{\rm tr}$ be obtained from H by cutting all hopping bonds that cross the boundary of a chosen

cage region, with the hole inside the cage and the exterior half filled. The Gibbs–Bogoliubov inequality, $F \le F_{\rm tr} + \langle H - H_{\rm tr}\rangle_{\rm tr}$, immediately gives $F \le F_{\rm tr}$: $H_{\rm tr}$ conserves the particle numbers of the cage and of the exterior separately, every term of $H - H_{\rm tr}$ transfers a particle across the boundary, and hence $\langle H - H_{\rm tr}\rangle_{\rm tr} = 0$. Moreover, at $U = \infty$ the half-filled exterior has no charge dynamics, so the trial free energy is additive,

$$F_{\rm tr}(N_{\rm cage}, T) = F_{\rm cage}(N_{\rm cage}, T) - (N - N_{\rm cage})\, T \ln 2, \tag{H1}$$

with $F_{\rm cage}$ the free energy of the isolated cage (an open-boundary cluster of $N_{\rm cage}$ sites containing one hole), obtained by exact diagonalization of the cage alone. Throughout this work $\Delta F_{\rm cage} \equiv F_{\rm tr} - F \ge 0$ denotes the caging free-energy cost. At high temperatures it approaches the positional-entropy cost of confining the hole, $\Delta F_{\rm cage} \to T \ln(N/N_{\rm cage})$, which follows from the exact $T \to \infty$ state counts $N\, 2^{N-1}$ and $N_{\rm cage}\, 2^{N-1}$ of the full and trial systems, respectively. We therefore consider the residual cost

$$\Delta\widetilde{F}_{\rm cage}(N_{\rm cage}, T) \;\equiv\; \Delta F_{\rm cage}(N_{\rm cage}, T) - T \ln(N/N_{\rm cage}), \tag{H2}$$

which measures the confinement cost beyond positional entropy. The figures show the dimensionless ratio $\Delta\widetilde{F}_{\rm cage}/T$: it vanishes at high temperature, where the free energy is local; is of order unity through much of the semi-quantum window, so that confinement on the cage scale costs of order the thermal scale; and grows rapidly on cooling toward $T_{\rm coh}$, as discussed in the main text.

1. Cage geometries and effective cage length

Figure 13(a) shows the residual confinement cost of the 2×8 ladder with a single hole ($x = 1/16$), for cages $2 \times n$ with $n = 1, \ldots, 7$. This quasi-one-dimensional geometry complements the two-dimensional 18-site probe of the main text and, admitting a longer series of cage sizes, resolves the cage-size dependence. Its confinement cost behaves as on the 18-site cluster, of order of the thermal scales across the semi-quantum window and rising on cooling toward $T_{\rm coh}$, showing that effective caging, with a comparable cage length, is not an artifact of the specific two-dimensional cluster.

Over the accessible range of cage sizes, the dependence on cage size is approximately exponential,

$$\Delta\widetilde{F}_{\rm cage}(T, \ell) \;=\; A(T)\, e^{-\ell/\xi_{\rm cage}(T)}, \qquad \ell = \begin{cases} \sqrt{N_{\rm cage}} & \text{(2D cages)}, \\ n & \text{($2 \times n$ ladder cages)}, \end{cases} \tag{H3}$$

which defines an effective cage length $\xi_{\rm cage}(T)$ [Fig. 13(b,c,d)]. The fitted length is short and weakly temperature dependent across the semi-quantum window, $\xi_{\rm cage} \approx 1.5$ lattice spacings on

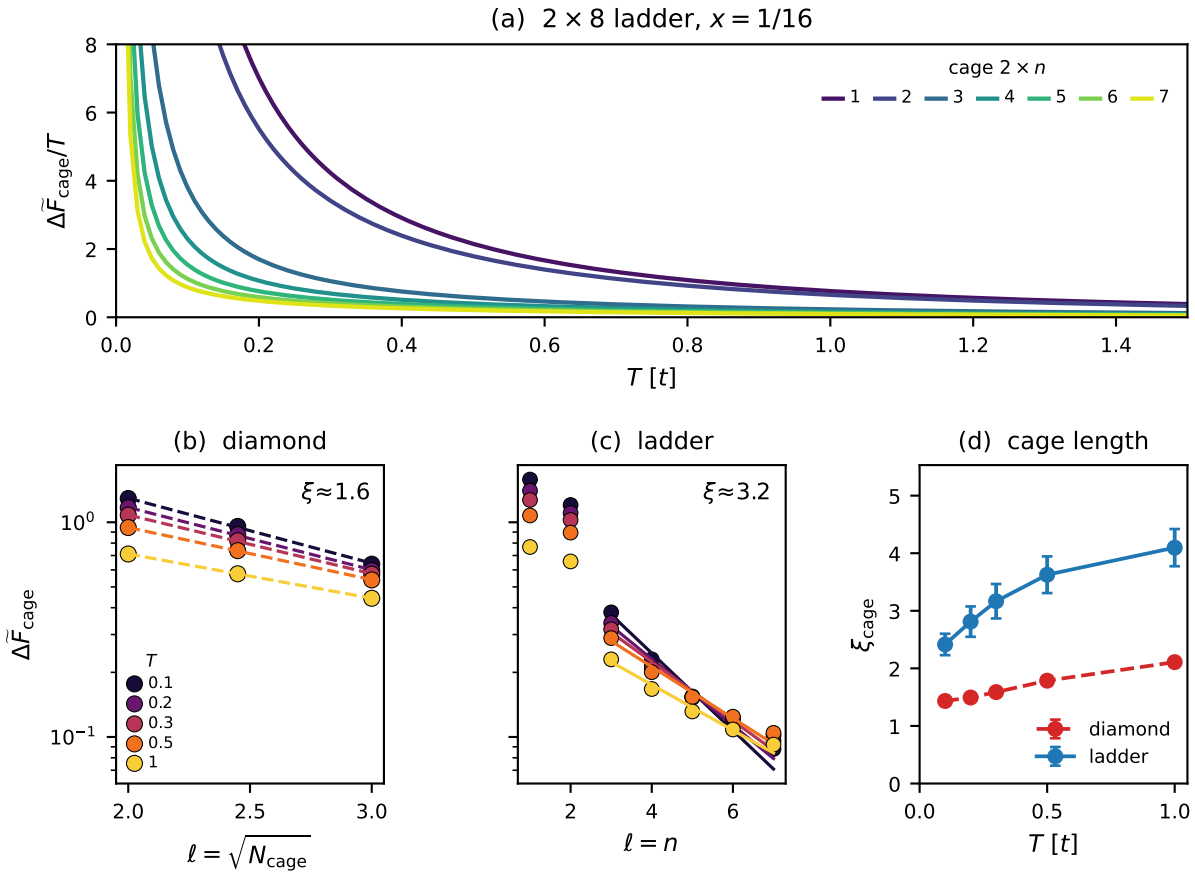


Figure 13. **Cage confinement cost, cage-size scaling, and cage length.** (**a**) Residual confinement cost $\Delta\widetilde{F}_{\rm cage}/T$ [Eq. (H2)] versus temperature for the $2\times n$ cages ($n=1,\ldots,7$) on the 2×8 ladder ($x=1/16$). Bonds crossing the cage boundary are cut; the hole is confined to the cage and the exterior remains half filled. (The corresponding cost for the 18-site diamond is shown in Fig. 5C.) Throughout the semi-quantum window the cost remains order the thermal scale, $\Delta\widetilde{F}_{\rm cage}/T\sim 1$, growing rapidly only on approaching $T_{\rm coh}$. (**b**,**c**) Cage-size dependence of the confinement cost: $\Delta\widetilde{F}_{\rm cage}$ versus cage linear size ℓ on a logarithmic scale, with the exponential fits of Eq. (H3) [$\ell=\sqrt{N_{\rm cage}}$ for the 18-site diamond cages 2×2, 3×2, 3×3; $\ell=n$ for the $2\times n$ ladder cages], at temperatures $T=0.1$–$1.5\,t$. (**d**) Fitted cage length $\xi_{\rm cage}(T)$ for both geometries.

the 18-site cluster and ≈ 2.8 on the ladder. A short, nearly T-independent $\xi_{\rm cage}$ supports the interpretation of charge motion in the semi-quantum regime as quasi-local and controlled by a finite spin environment around the hole rather than by delocalization across the system. On cooling toward $T_{\rm coh}$ the dimensionless confinement cost $\Delta\widetilde{F}_{\rm cage}/T$ rises sharply [Fig. 13(a)], signaling the breakdown of the caging description as ferromagnetic coherence sets in and the hole delocalizes.