Supplemental Material for "Density-Matrix Renormalization Group for Continuous Quantum Systems"

Shovan Dutta,[1,2,3] Anton Buyskikh,[4] Andrew J. Daley,[4] and Erich J. Mueller[2]

[1] *T.C.M. Group, Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge CB3 0HE, United Kingdom*
[2] *Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853, USA*
[3] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[4] *Department of Physics and SUPA, University of Strathclyde, Glasgow G4 0NG, United Kingdom*

(Dated: May 8, 2022)

Contents

I. Derivation of the partitioned Hamiltonian

We consider the Hamiltonian in Eq. (1) of the main text, $\hat{H} = \int_0^L \mathrm{d}x \left[\hat{\mathcal{K}}(x) + \hat{\mathcal{U}}(x) + \hat{\mathcal{P}}(x)\right]$, where

$$\hat{\mathcal{K}}(x) = \frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}^\dagger(x)\right)\left(\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}(x)\right), \tag{S1}$$

$$\hat{\mathcal{U}}(x) = \frac{g}{2}\,\hat{\psi}^\dagger(x)\,\hat{\psi}^\dagger(x)\,\hat{\psi}(x)\,\hat{\psi}(x), \tag{S2}$$

$$\hat{\mathcal{P}}(x) = V(x)\,\hat{\psi}^\dagger(x)\,\hat{\psi}(x) \tag{S3}$$

denote the kinetic, interaction, and potential energy densities, respectively. We divide the x axis into M continuous segments with boundaries at $\{X_j\}$. Since the wave function is everywhere finite, the interaction and potential energies reduce to sums over the individual segments, $\hat{U} = \sum_j \hat{U}_j$ and $\hat{P} = \sum_j \hat{P}_j$, where

$$\hat{U}_j = \int_{X_{j-1}}^{X_j} \mathrm{d}x\, \hat{\mathcal{U}}(x), \quad \text{and} \quad \hat{P}_j = \int_{X_{j-1}}^{X_j} \mathrm{d}x\, \hat{\mathcal{P}}(x). \tag{S4}$$

However, the kinetic energy $\hat{K}$ will diverge if $\hat{\psi}(x)$ has a discontinuity across any of the boundaries, which gives rise to additional terms in the Hamiltonian. To see this, we write $\hat{\psi}(x)$ close to the boundary at X_j as

$$\hat{\psi}(x) = \big[1 - \theta(x - X_j)\big]\, \hat{\psi}_<(x) + \theta(x - X_j)\, \hat{\psi}_>(x) \tag{S5}$$

where θ is the unit step function and $\hat{\psi}_{<(>)}(x)$ denotes the part of $\hat{\psi}$ to the left (right) of X_j. Thus, one obtains

$$\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}(x) = \delta(x - X_j)\left[\hat{\psi}_>(x) - \hat{\psi}_<(x)\right] + \big[1 - \theta(x - X_j)\big]\left(\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}_<(x)\right) + \theta(x - X_j)\left(\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}_>(x)\right) , \tag{S6}$$

where δ is the Dirac delta function. Substituting this expression into Eq. (S1) and using the sifting property of the delta function yield the following contribution to $\hat{K}$ from the boundary,

$$\begin{aligned}\int_{X_j^-}^{X_j^+} \mathrm{d}x\, \hat{\mathcal{K}} = {} & \Lambda\left[\hat{\psi}(X_j^+) - \hat{\psi}(X_j^-)\right]^\dagger\left[\hat{\psi}(X_j^+) - \hat{\psi}(X_j^-)\right] \\ & + \frac{1}{2}\left[\hat{\psi}(X_j^+) - \hat{\psi}(X_j^-)\right]^\dagger\left[(1-\theta_0)\left(\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}(X_j^-)\right) + \theta_0\left(\frac{\mathrm{d}}{\mathrm{d}x}\hat{\psi}(X_j^+)\right)\right] + \text{h.c.} ,\end{aligned} \tag{S7}$$

where $\hat{\psi}(X_j^+) \equiv \hat{\psi}_>(X_j)$, $\hat{\psi}(X_j^-) \equiv \hat{\psi}_<(X_j)$, $\theta_0 \coloneqq \theta(0)$, and $\Lambda = \delta(0)/2$ is a formally divergent energy penalty that ensures the wave function is continuous, i.e., $\hat{\psi}(X_j^+) = \hat{\psi}(X_j^-)$. For the numerics, we treat Λ as a phenomenological parameter. As Λ is increased, the discontinuity in the wave function falls off as $1/\Lambda$, and so does the second line of Eq. (S7). In fact, these other terms do not impose any constraint on the wave function, and we find they also do not alter the numerical convergence to the ground state. Thus, one can drop these without affecting physical predictions, which gives $\hat{K} = \sum_j \hat{K}_j + \Lambda\hat{\Upsilon}_{j,j+1}$, where

$$\hat{K}_j = \int_{X_{j-1}}^{X_j} \mathrm{d}x\, \hat{\mathcal{K}}(x) , \quad \text{and} \quad \hat{\Upsilon}_{j,j+1} \coloneqq \left[\hat{\psi}(X_j^+) - \hat{\psi}(X_j^-)\right]^\dagger\left[\hat{\psi}(X_j^+) - \hat{\psi}(X_j^-)\right] . \tag{S8}$$

Thus, we arrive at the partitioned Hamiltonian $\hat{H} = \sum_j \hat{K}_j + \hat{U}_j + \hat{P}_j + \Lambda\hat{\Upsilon}_{j,j+1}$.

II. Basis functions and local operators

As explained in the main text, each segment j is spanned by multiple n-body basis functions $\phi_{n,k}^{(j)}(\vec{r})$, with $\vec{r} \equiv \{x_1, x_2, \ldots, x_n\}$, which leads to a matrix representation of the local operators. Here we show how to compute these matrix elements in terms of monomial integrals. We also show how to include the cusp constraint which arises from contact interactions.

A. Characterization in terms of monomials

Given an n-particle basis function $\phi_{n,k}^{(j)}(\vec{r})$, the many-body state may be expressed as

$$\big|\phi_{n,k}^{(j)}\big\rangle = \int_{X_{j-1}}^{X_j} \mathrm{d}^n r\, \phi_{n,k}^{(j)}(\vec{r})\, \frac{\hat{\psi}^\dagger(x_1)\hat{\psi}^\dagger(x_2)\ldots\hat{\psi}^\dagger(x_n)}{\sqrt{n!}}\,|0\rangle , \tag{S9}$$

where $|0\rangle$ is the vacuum, and the integration limits denote that all coordinates lie in the interval $[X_{j-1}, X_j]$. As we are dealing with bosons, the wave function $\phi_{n,k}^{(j)}(\vec{r})$ is symmetric under exchange of any two coordinates, and the field operators satisfy the commutation $[\hat{\psi}(x), \hat{\psi}^\dagger(x')] = \delta(x - x')$. The inner product vanishes between states with unequal

numbers of particles, while

$$\langle \phi^{(j)}_{n,k} | \phi^{(j)}_{n,k'} \rangle = \int_{X_{j-1}}^{X_j} \mathrm{d}^n r \, \phi^{(j)*}_{n,k}(\vec{r}) \, \phi^{(j)}_{n,k'}(\vec{r}) \, . \tag{S10}$$

For the basis to be orthogonal, (S10) must vanish for $k \neq k'$. It is convenient to rescale the coordinates so that the interval runs from 0 to 1,

$$\phi^{(j)}_{n,k}(\vec{r}) \coloneqq \frac{1}{w_j^{n/2}} \chi^{(j)}_{n,k}\left(\frac{\vec{r} - \vec{R}_{j-1}}{w_j} \right) . \tag{S11}$$

where $w_j \coloneqq X_j - X_{j-1}$ and $\vec{R}_{j-1}$ is the coordinate $x_1 = x_2 = \cdots = x_n = X_{j-1}$. Then Eq. (S10) becomes

$$\langle \phi^{(j)}_{n,k} | \phi^{(j)}_{n,k'} \rangle = \int_0^1 \mathrm{d}^n r \, \chi^{(j)*}_{n,k}(\vec{r}) \, \chi^{(j)}_{n,k'}(\vec{r}) \, , \tag{S12}$$

so the basis will be orthonormal provided $\chi^{(j)}_{n,k}$ are orthonormal over the interval $[0,1]^n$.

One can always write $\chi^{(j)}_{n,k}(\vec{r})$ as a sum over symmetrized monomials $\mathbf{p}(\vec{r})$ whose variation in the "canonical" sector

$$\mathcal{S}_n : \quad 0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq 1 \tag{S13}$$

is given by $\mathbf{p}(\vec{r}) = x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n}$, where p_i are nonnegative integers. In this sector,

$$\chi^{(j)}_{n,k}(\vec{r}) = \sum\nolimits_{\mathbf{p}} A^{(j)}_{n,k,\mathbf{p}} \, \mathbf{p}(\vec{r}) \, , \tag{S14}$$

with expansion coefficients $A^{(j)}_{n,k,\mathbf{p}}$, which characterize the basis. In numerical calculations, one has to restrict to a finite basis by constraining $\mathbf{p} \coloneqq \{p_1, p_2, \ldots, p_n\}$. For example, the number of monomials of degree $\sum_{i=1}^n p_i = d$ grows as $\binom{d+n-1}{d}$, and we find it convenient to only retain states with $d \leq d_{\max}$. The monomials are not orthogonal, but their inner product $\langle \mathbf{p} | \mathbf{q} \rangle = n! \, \mathcal{I}(\mathbf{p} + \mathbf{q})$ has a simple form

$$\mathcal{I}(\mathbf{p}) \coloneqq \int_0^1 \mathrm{d}x_n \int_0^{x_n} \mathrm{d}x_{n-1} \cdots \int_0^{x_2} \mathrm{d}x_1 \, x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n} = \prod_{n'=1}^{n} \frac{1}{n' + \sum_{i=1}^{n'} p_i} \, . \tag{S15}$$

As explained in Sec. II C, a convenient way to construct our orthonormal basis is to take the states to be eigenvectors of a Hermitian operator. When expressed in terms of the non-orthogonal monomials, this involves solving a generalized eignevalue problem. Physical constraints, such as the cusp condition arising from short-range interactions, are incorporated by correctly choosing the Hermitian operator.

B. Matrix elements of local operators

1. *Field operator*

The action of the field operator $\hat{\psi}(x)$ on the basis states in Eq. (S9) can be found by using Bose commutation and the symmetry of $\phi^{(j)}_{n,k}(\vec{r})$ under particle exchange, which gives, for $X_{j-1} \leq x \leq X_j$,

$$\hat{\psi}(x) \, | \phi^{(j)}_{n,k} \rangle = \sqrt{n} \int_{X_{j-1}}^{X_j} \mathrm{d}^{n-1} r \, \phi^{(j)}_{n,k}(x, \vec{r}) \, \frac{\hat{\psi}^\dagger(x_1) \hat{\psi}^\dagger(x_2) \ldots \hat{\psi}^\dagger(x_{n-1})}{\sqrt{(n-1)!}} \, |0\rangle \, . \tag{S16}$$

The right-hand side describes a state of $n-1$ particles with the (unnormalized) wave function $\sqrt{n}\,\phi^{(j)}_{n,k}(x,\vec{r})$ – which is simply the original wave function with one of the positions set to x. Because of the Bose symmetry, it does not matter which particle is selected.

The nonzero matrix elements of $\hat{\psi}(x)$ are given by [using Eqs. (S10) and (S11)]

$$\langle \phi^{(j)}_{n-1,k} |\, \hat{\psi}(x)\, | \phi^{(j)}_{n,k'} \rangle = \sqrt{n} \int_{X_{j-1}}^{X_j} \mathrm{d}^{n-1}r\; \phi^{(j)*}_{n-1,k}(\vec{r})\, \phi^{(j)}_{n,k'}(x,\vec{r}) = \sqrt{\frac{n}{w_j}} \int_0^1 \mathrm{d}^{n-1}r\; \chi^{(j)*}_{n-1,k}(\vec{r})\, \chi^{(j)}_{n,k'}(\tilde{x},\vec{r})\;, \tag{S17}$$

where $\tilde{x} \coloneqq (x - X_{j-1})/w_j$. As $\chi^{(j)}_{n,k}$ are linear combinations of symmetrized monomials [Eq. (S14)], it suffices to evaluate this integral for any two such monomials, i.e.,

$$\psi_{\mathbf{p},\mathbf{q}}(\tilde{x}) \coloneqq \sqrt{n} \int_0^1 \mathrm{d}^{n-1}r\; \mathbf{p}(\vec{r})\; \mathbf{q}(\tilde{x},\vec{r})\;, \tag{S18}$$

where $\mathbf{p} = \{p_1, p_2, \ldots, p_{n-1}\}$, $\mathbf{q} = \{q_1, q_2, \ldots, q_n\}$. Using exchange symmetry, one can write $\psi_{\mathbf{p},\mathbf{q}}(\tilde{x}) = \sqrt{n}(n-1)! \int_{\mathcal{S}_{n-1}} \mathrm{d}^{n-1}r\; \mathbf{p}(\vec{r})\; \mathbf{q}(\tilde{x},\vec{r})$ over the canonical sector $\mathcal{S}_{n-1}$ [Eq. (S13)]. In this sector, $\mathbf{p}(\vec{r}) = x_1^{p_1} x_2^{p_2} \ldots x_{n-1}^{p_{n-1}}$, but the expression for $\mathbf{q}(\tilde{x},\vec{r})$ depends on the ordering of $\tilde{x}$ relative to the other coordinates. For $\tilde{x} \to 0^+$ (i.e., $x \to X_{j-1}^+$), $q(\tilde{x},\vec{r}) = \tilde{x}^{q_1} x_1^{q_2} x_2^{q_3} \ldots x_{n-1}^{q_n}$, whereas for $\tilde{x} \to 1^-$ ($x \to X_j^-$), $q(\tilde{x},\vec{r}) = x_1^{q_1} x_2^{q_2} \ldots x_{n-1}^{q_{n-1}} \tilde{x}^{q_n}$. Thus, using Eq. (S15),

$$\psi_{\mathbf{p},\mathbf{q}}(0) = \delta_{q_1,0}\, \sqrt{n}\, (n-1)!\, \mathcal{I}\big(\{p_1+q_2, p_2+q_3, \ldots, p_{n-1}+q_n\}\big)\;, \tag{S19}$$

$$\psi_{\mathbf{p},\mathbf{q}}(1) = \sqrt{n}\, (n-1)!\, \mathcal{I}\big(\{p_1+q_1, p_2+q_2, \ldots, p_{n-1}+q_{n-1}\}\big)\;. \tag{S20}$$

For intermediate values of $\tilde{x}$, the integral can be split into domains where $x_1 \le \cdots \le x_{n'-1} \le \tilde{x} \le x_{n'+1} \le \cdots \le x_{n-1}$ for $n' = 1, 2, \ldots, n$, which gives

$$\psi_{\mathbf{p},\mathbf{q}}(\tilde{x}) = \sqrt{n}\, (n-1)! \sum_{n'=1}^{n} \tilde{x}^{q_{n'}} \mathcal{I}_l\big(\{p_1+q_1, \ldots, p_{n'-1}+q_{n'-1}\}, \tilde{x}\big)\, \mathcal{I}_r\big(\tilde{x}, \{p_{n'}+q_{n'+1}, \ldots, p_{n-1}+q_n\}\big)\;, \tag{S21}$$

where

$$\mathcal{I}_l\big(\{p_1, \ldots, p_n\}, \tilde{x}\big) \coloneqq \int_0^{\tilde{x}} \mathrm{d}x_n \int_0^{x_n} \mathrm{d}x_{n-1} \cdots \int_0^{x_2} \mathrm{d}x_1\; x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n} = \tilde{x}^{n+\sum_{i=1}^n p_i}\, \mathcal{I}\big(\{p_1, \ldots, p_n\}\big)\;, \tag{S22}$$

$$\begin{aligned} \mathcal{I}_r\big(\tilde{x}, \{p_1, \ldots, p_n\}\big) &\coloneqq \int_{\tilde{x}}^1 \mathrm{d}x_n \int_{\tilde{x}}^{x_n} \mathrm{d}x_{n-1} \cdots \int_{\tilde{x}}^{x_2} \mathrm{d}x_1\; x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n} \\ &= \sum_{i=0}^{n} (-1)^i\, \mathcal{I}_l\big(\{p_i, p_{i-1} \ldots, p_1\}, \tilde{x}\big)\, \mathcal{I}\big(\{p_{i+1}, p_{i+2}, \ldots, p_n\}\big)\;, \end{aligned} \tag{S23}$$

and $\mathcal{I}_l(\{\}, \tilde{x}) \coloneqq 1$. Note that $\mathcal{I}_l(\mathbf{p}, 1) = \mathcal{I}_r(0, \mathbf{p}) = \mathcal{I}(\mathbf{p})$. Equation (S21) gives $\psi_{\mathbf{p},\mathbf{q}}(\tilde{x})$ as a polynomial in $\tilde{x}$ of degree $n - 1 + q_n + \sum_{i=1}^{n-1} p_i + q_i$. In numerical simulations, we store these polynomial coefficients.

2. *Density and potential energy*

A similar construction applies for the density $\hat{\rho}(x) = \hat{\psi}^\dagger(x)\hat{\psi}(x)$. Using Eq. (S16), one finds the matrix elements

$$\langle \phi^{(j)}_{n,k} |\, \hat{\rho}(x)\, | \phi^{(j)}_{n,k'} \rangle = n \int_{X_{j-1}}^{X_j} \mathrm{d}^{n-1}r\; \phi^{(j)*}_{n,k}(x,\vec{r})\, \phi^{(j)}_{n,k'}(x,\vec{r}) = \frac{n}{w_j} \int_0^1 \mathrm{d}^{n-1}r\; \chi^{(j)*}_{n,k}(\tilde{x},\vec{r})\, \chi^{(j)}_{n,k'}(\tilde{x},\vec{r}) \tag{S24}$$

for $X_{j-1} \le x \le X_j$. As before, the integral needs to be calculated only for symmetrized monomials,

$$\rho_{\mathbf{p},\mathbf{q}}(\tilde{x}) \coloneqq n \int_0^1 \mathrm{d}^{n-1} r \; \mathbf{p}(\tilde{x},\vec{r}) \; \mathbf{q}(\tilde{x},\vec{r}) \; , \tag{S25}$$

where $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$ and $\mathbf{q} = \{q_1, q_2, \ldots, q_n\}$. Using exchange symmetry, $\rho_{\mathbf{p},\mathbf{q}}(\tilde{x}) = n! \int_{\mathcal{S}_{n-1}} \mathrm{d}^{n-1} r \; \mathbf{p}(\tilde{x},\vec{r}) \; \mathbf{q}(\tilde{x},\vec{r})$ where $\mathcal{S}_n$ denotes the canonical ordering [Eq. (S13)]. At the boundaries $\tilde{x} \to 0^+$ and $\tilde{x} \to 1^-$, both $\mathbf{p}(\tilde{x},\vec{r})$ and $\mathbf{q}(\tilde{x},\vec{r})$ reduce to single monomials, yielding [as in Eqs. (S19) and (S20)]

$$\rho_{\mathbf{p},\mathbf{q}}(0) = \delta_{p_1+q_1,0} \, n! \, \mathcal{I}\big(\{p_2+q_2, p_3+q_3, \ldots, p_n+q_n\}\big) \; , \tag{S26}$$

$$\rho_{\mathbf{p},\mathbf{q}}(1) = n! \, \mathcal{I}\big(\{p_1+q_1, p_2+q_2, \ldots, p_{n-1}+q_{n-1}\}\big) \; . \tag{S27}$$

For intermediate $\tilde{x}$, we split the integral into subdomains $x_1 \le \cdots \le x_{n'-1} \le \tilde{x} \le x_{n'+1} \le \cdots \le x_{n-1}$, obtaining

$$\rho_{\mathbf{p},\mathbf{q}}(\tilde{x}) = n! \sum_{n'=1}^{n} \tilde{x}^{p_{n'}+q_{n'}} \mathcal{I}_l\big(\{p_1+q_1, \ldots, p_{n'-1}+q_{n'-1}\}, \tilde{x}\big) \; \mathcal{I}_r\big(\tilde{x}, \{p_{n'+1}+q_{n'+1}, \ldots, p_n+q_n\}\big) \; , \tag{S28}$$

where $\mathcal{I}_l$ and $\mathcal{I}_r$ are given by Eqs. (S22) and (S23). Again, $\rho_{\mathbf{p},\mathbf{q}}(\tilde{x})$ is a polynomial in $\tilde{x}$ of degree $n - 1 + \sum_{i=1}^{n} p_i + q_i$, and we store the coefficients.

The matrix elements of the potential energy $\hat{P}_j = \int_{X_{j-1}}^{X_j} \mathrm{d}x \, V(x) \, \hat{\rho}(x)$ can be obtained from those of the density in Eq. (S24). In particular, for two symmetrized monomials, one calculates

$$V^{(j)}_{\mathbf{p},\mathbf{q}} \coloneqq \int_0^1 \mathrm{d}\tilde{x} \; V(X_{j-1} + w_j \tilde{x}) \; \rho_{\mathbf{p},\mathbf{q}}(\tilde{x}) \; , \tag{S29}$$

which reduces to computing moments of $V(x)$, since $\rho_{\mathbf{p},\mathbf{q}}(\tilde{x})$ is a polynomial in $\tilde{x}$. For a sinusoidal potential $V(x) = V_0 \cos^2 \kappa x$, these moments can be found exactly using

$$\int_0^1 \mathrm{d}\tilde{x} \, \tilde{x}^{p-1} e^{\mathrm{i}\tilde{\kappa}\tilde{x}} = (\mathrm{i}/\tilde{\kappa})^p \, \Gamma_l(p, -\mathrm{i}\tilde{\kappa}) \tag{S30}$$

$\forall p \ge 1$, where $\Gamma_l(p,z)$ is the lower incomplete gamma function [S1].

3. *Kinetic energy*

The kinetic energy within the j-th segment is given by $\hat{K}_j = (1/2) \int_{X_{j-1}}^{X_j} \mathrm{d}x \, [\partial_x \hat{\psi}^\dagger(x)] \, [\partial_x \hat{\psi}(x)]$. To find its matrix elements, we use Eq. (S16) and the exchange symmetry of the basis functions, obtaining

$$\big\langle \phi^{(j)}_{n,k} \big| \, \hat{K}_j \, \big| \phi^{(j)}_{n,k'} \big\rangle = \frac{1}{2} \int_{X_{j-1}}^{X_j} \mathrm{d}^n r \; \vec{\nabla} \phi^{(j)*}_{n,k}(\vec{r}) . \vec{\nabla} \phi^{(j)}_{n,k'}(\vec{r}) = \frac{1}{2w_j^2} \int_0^1 \mathrm{d}^n r \; \vec{\nabla} \chi^{(j)*}_{n,k}(\vec{r}) . \vec{\nabla} \chi^{(j)}_{n,k'}(\vec{r}) \; . \tag{S31}$$

Replacing $\chi^{(j)}_{n,k}(\vec{r})$ with a symmetrized monomial [Eq. (S14)], we only need to evaluate

$$K_{\mathbf{p},\mathbf{q}} \coloneqq \frac{1}{2} \int_0^1 \mathrm{d}^n r \; \vec{\nabla} \mathbf{p}(\vec{r}) . \vec{\nabla} \mathbf{q}(\vec{r}) = \frac{n!}{2} \int_{\mathcal{S}_n} \mathrm{d}^n r \; \vec{\nabla} \mathbf{p}(\vec{r}) . \vec{\nabla} \mathbf{q}(\vec{r}) \; , \tag{S32}$$

where $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$ and $\mathbf{q} = \{q_1, q_2, \ldots, q_n\}$. In sector $\mathcal{S}_n$, $\partial_{x_i} \mathbf{p}(\vec{r}) = p_i x_i^{p_i - 1} \prod_{i' \neq i} x_{i'}^{p_{i'}}$. Thus, Eq. (S15) yields

$$K_{\mathbf{p},\mathbf{q}} = \frac{n!}{2} \sum_{i=1}^{n} p_i q_i \, \mathcal{I}\big(\texttt{incr}(\mathbf{p}+\mathbf{q}, i, -2)\big) \,, \tag{S33}$$

where $\texttt{incr}(\mathbf{p}, i, s) \coloneqq \{p_1, \ldots, p_{i-1}, p_i + s, p_{i+1}, \ldots, p_n\}$, i.e., it increments the i-th element by s. Note that one can also extract matrix elements of the kinetic energy density $\hat{\mathcal{K}}(x)$ in Eq. (S1) using the procedure in Sec. II B 2.

4. *Interaction energy*

For the interaction energy $\hat{U}_j$ in Eq. (S4), we again use the action of the field operator in Eq. (S16) to obtain

$$\langle \phi^{(j)}_{n,k} | \, \hat{U}_j \, | \phi^{(j)}_{n,k'} \rangle = g \binom{n}{2} \int_{X_{j-1}}^{X_j} \mathrm{d}x \int_{X_{j-1}}^{X_j} \mathrm{d}^{n-2} r \; \phi^{(j)*}_{n,k}(x, x, \vec{r}) \; \phi^{(j)}_{n,k'}(x, x, \vec{r}) \tag{S34}$$

$$= g \int_{X_{j-1}}^{X_j} \mathrm{d}^n r \sum_{i<i'} \delta(x_i - x_{i'}) \; \phi^{(j)*}_{n,k}(\vec{r}) \; \phi^{(j)}_{n,k'}(\vec{r}) = \frac{g}{w_j} \int_0^1 \mathrm{d}^n r \sum_{i<i'} \delta(x_i - x_{i'}) \; \chi^{(j)*}_{n,k}(\vec{r}) \; \chi^{(j)}_{n,k'}(\vec{r}) \,. \tag{S35}$$

Therefore, constructing the matrix for $\hat{U}_j$ boils down to evaluating

$$U_{\mathbf{p},\mathbf{q}} \coloneqq \int_0^1 \mathrm{d}^n r \sum_{i<i'} \delta(x_i - x_{i'}) \; \mathbf{p}(\vec{r}) \; \mathbf{q}(\vec{r}) = \frac{n!}{2} \int_{\mathcal{S}_n} \mathrm{d}^n r \sum_{i=1}^{n-1} \delta(x_i - x_{i+1}) \; \mathbf{p}(\vec{r}) \; \mathbf{q}(\vec{r}) \tag{S36}$$

for symmetrized monomials $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$ and $\mathbf{q} = \{q_1, q_2, \ldots, q_n\}$. Note, in sector $\mathcal{S}_n$, $x_1 \leq x_2 \leq \cdots \leq x_n$, so we have the delta function only between neighboring coordinates x_i and x_{i+1}. Substituting the monomial expressions for $\mathbf{p}(\vec{r})$ and $\mathbf{q}(\vec{r})$, and using Eq. (S15), we find

$$U_{\mathbf{p},\mathbf{q}} = \frac{n!}{2} \sum_{i=1}^{n-1} \mathcal{I}\big(\texttt{merge}(\mathbf{p}+\mathbf{q}, i)\big) \,, \tag{S37}$$

where $\texttt{merge}(\mathbf{p}, i) \coloneqq \{p_1, \ldots, p_{i-1}, p_i + p_{i+1}, p_{i+2}, \ldots, p_n\}$, i.e., it merges the i- and $i+1$-th elements. As with $\hat{\mathcal{K}}(x)$, the matrix elements of the interaction energy density $\hat{\mathcal{U}}(x)$ in Eq. (S2) can be obtained using the method in Sec. II B 2.

C. Basis construction for contact interactions

The ideal choice of basis functions, $\chi^{(j)}_{n,k}(\vec{r})$, would have three properties: (i) A small number of these states should accurately approximate the low-energy eigenstates of the Hamiltonian, (ii) these states should smoothly connect to the wave functions in neighboring sectors, and (iii) they should be orthogonal to one another. The latter can be ensured by taking them to be eigenstates of a Hermitian operator – and choosing the basis is equivalent to choosing the operator.

Insight into the choice of operator comes from the one-particle sector, where selecting the basis functions is related to deciding on a functional form for splines which will be used to piecewise describe the ground state of the Schrödinger equation. In that case, one might naively choose the single-particle basis functions on $[0, 1]$ to be solutions to Laplace's equation, $\partial_{x_1}^2 \chi = \lambda \chi$. Depending on boundary conditions, $\chi = \cos(\pi k x_1)$ with $k = 0, 1, \ldots$ or $\chi = \sin(\pi k x_1)$ with $k = 1, 2, \ldots$. Neither set of states is amenable to continuously connecting across segments – as either the basis function or its derivative vanishes at the boundaries. The solution is to modify the operator so that $x_1 = 0$ and $x_1 = 1$ are regular singular points; for example choosing them to be solutions to Legendre's equation: $\partial_{x_1}[x_1(1 - x_1)\partial_{x_1}\chi] = \lambda\chi$.

The solutions are Legendre polynomials $P_k(2x_1-1)$ [S1], and the resulting wave function expansion amounts to using polynomial splines. Colloquially, one imagines that the factor $x_1(1-x_1)$ "absorbs" the boundary conditions, allowing basis functions to have both nonzero amplitude and slope at the segment boundaries. An equivalent construction of the Legendre polynomials is to perform a Gram-Schmidt orthogonalization on the monomials $\{1, x_1, x_1^2, \ldots, x_1^k\}$.

In our problem the many-body wave function has a kink whenever two particles coincide. Thus, the expansion will perform better if the basis functions also have this kink, motivating the modified Legendre equation,

$$-\frac{1}{2}\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[x_i(1-x_i)\,\frac{\partial}{\partial x_i}\,\chi_{n,k}^{(j)}(\vec{r})\right]+c_j\sum_{i<i'}x_i(1-x_i)\,\delta(x_i-x_{i'})\,\chi_{n,k}^{(j)}(\vec{r})=\mathcal{E}_{n,k}^{(j)}\,\chi_{n,k}^{(j)}(\vec{r})\,, \tag{S38}$$

where c_j gives the slope discontinuity, $\partial_{x_i}\chi_{n,k}^{(j)}(x_i\to x_{i'}^{+})-\partial_{x_i}\chi_{n,k}^{(j)}(x_i\to x_{i'}^{-})=c_j\,\chi_{n,k}^{(j)}(x_i=x_{i'})$. As the coordinates of χ are stretched by a factor of w_j relative to the physical coordinates [Eq. (S11)], we require $c_j=w_j g$. For uniform segments, $w_j=$ constant, thus c_j and $\chi_{n,k}^{(j)}$ do not depend on j.

We expand Eq. (S38) on the (non-orthogonal) symmetrized monomials. Defining

$$\hat{K}_{\mathcal{L}}[\chi(\vec{r})]\coloneqq-\frac{1}{2}\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[x_i(1-x_i)\,\frac{\partial}{\partial x_i}\,\chi(\vec{r})\right] \tag{S39}$$

$$\hat{U}_{\mathcal{L}}[\chi(\vec{r})]\coloneqq\sum_{i<i'}x_i(1-x_i)\,\delta(x_i-x_{i'})\,\chi(\vec{r}) \tag{S40}$$

and their sum $\hat{H}_{\mathcal{L}}^{(j)}=\hat{K}_{\mathcal{L}}+c_j\hat{U}_{\mathcal{L}}$, the matrix elements are

$$\langle\mathbf{p}|\hat{K}_{\mathcal{L}}|\mathbf{q}\rangle=\frac{n!}{2}\int_{\mathcal{S}_n}\mathrm{d}^n r\sum_{i=1}^{n}x_i(1-x_i)\,[\partial_{x_i}\mathbf{p}(\vec{r})]\,[\partial_{x_i}\mathbf{q}(\vec{r})]\,, \tag{S41}$$

$$\langle\mathbf{p}|\hat{U}_{\mathcal{L}}|\mathbf{q}\rangle=\frac{n!}{2}\int_{\mathcal{S}_n}\mathrm{d}^n r\sum_{i=1}^{n-1}x_i(1-x_i)\,\delta(x_i-x_{i+1})\,\mathbf{p}(\vec{r})\,\mathbf{q}(\vec{r})\,. \tag{S42}$$

As before, $\mathbf{p}=\{p_1,p_2,\ldots,p_n\}$, $\mathbf{q}=\{q_1,q_2,\ldots,q_n\}$, and $\mathcal{S}_n$ stands for the canonical ordering in Eq. (S13). Substituting $\mathbf{p}(\vec{r})=x_1^{p_1}x_2^{p_2}\ldots x_n^{p_n}$ and $\mathbf{q}(\vec{r})=x_1^{q_1}x_2^{q_2}\ldots x_n^{q_n}$, we find, similar to Secs. II B 3 and II B 4,

$$\langle\mathbf{p}|\hat{K}_{\mathcal{L}}|\mathbf{q}\rangle=\frac{n!}{2}\sum_{i=1}^{n}p_iq_i\Big[\mathcal{I}\big(\texttt{incr}(\mathbf{p}+\mathbf{q},i,-1)\big)-\mathcal{I}(\mathbf{p}+\mathbf{q})\Big]\,, \tag{S43}$$

$$\langle\mathbf{p}|\hat{U}_{\mathcal{L}}|\mathbf{q}\rangle=\frac{n!}{2}\sum_{i=1}^{n-1}\Big[\mathcal{I}\Big(\texttt{incr}\big(\texttt{merge}(\mathbf{p}+\mathbf{q},i),i,1\big)\Big)-\mathcal{I}\Big(\texttt{incr}\big(\texttt{merge}(\mathbf{p}+\mathbf{q},i),i,2\big)\Big)\Big]\,, \tag{S44}$$

where $\mathcal{I}(\mathbf{p})$ is given by Eq. (S15). These expressions reduce Eq. (S38) to a generalized eigenvalue problem $H_{\mathcal{L}}^{(j)}A_{n,k}^{(j)}=\mathcal{E}_{n,k}^{(j)}O_nA_{n,k}^{(j)}$, where $A_{n,k}^{(j)}$ are the expansion coefficients of $\chi_{n,k}^{(j)}$ in terms of the monomials [Eq. (S14)] and O_n is the matrix of inner products between the monomials, given by $\langle\mathbf{p}|\mathbf{q}\rangle=n!\,\mathcal{I}(\mathbf{p}+\mathbf{q})$. The energy spectrum $\mathcal{E}_{n,k}^{(j)}$ provides a natural ordering of the basis states, which can be truncated at high energies. Although one is solving a many-body problem in generating the basis, the complexity is greatly reduced compared to the original problem, as the number of monomials is limited if the number of particles in a segment, n, is sufficiently small.

III. Single-particle correlations

Here we illustrate the extraction of observables by showing how one can compute the single-particle correlations, how these converge to the continuum limit, and how one recovers expected power laws in the momentum distribution.

A. Extraction as piecewise polynomial

Following Secs. II A and II B 1 one can express the matrix elements of the field operator in the i-th segment as

$$\langle \phi^{(i)}_{n,k} | \, \hat{\psi}(x) \, | \phi^{(i)}_{n',k'} \rangle = \sum_{p=0}^{p_{\max}} \Psi^{i,p}_{nk,n'k'} \left(\frac{x - X_{i-1}}{w_i} \right)^p , \tag{S45}$$

where $p_{\max} = 2d_{\max} + n_{\max} - 1$. Thus, $\hat{\psi}(x)$ can be represented by operators $\hat{\Psi}^{i,p}$ with matrix elements $\Psi^{i,p}_{nk,n'k'}$. We compute the correlator $C^{i,p}_{j,q} := \langle \hat{\Psi}^{i,p\dagger} \hat{\Psi}^{j,q} \rangle$ using efficient tensor contractions to obtain the single-particle correlations

$$\langle \hat{\psi}^\dagger(x) \hat{\psi}(x') \rangle = \sum_{p,q=0}^{p_{\max}} C^{i,p}_{j,q} \left(\frac{x - X_{i-1}}{w_i} \right)^p \left(\frac{x' - X_{j-1}}{w_j} \right)^q , \tag{S46}$$

when x and x' belong to the i-th and j-th segments, respectively. Figure S1 shows how the diagonal and anti-diagonal correlations differ between Mott and superfluid ground states in a shallow potential.

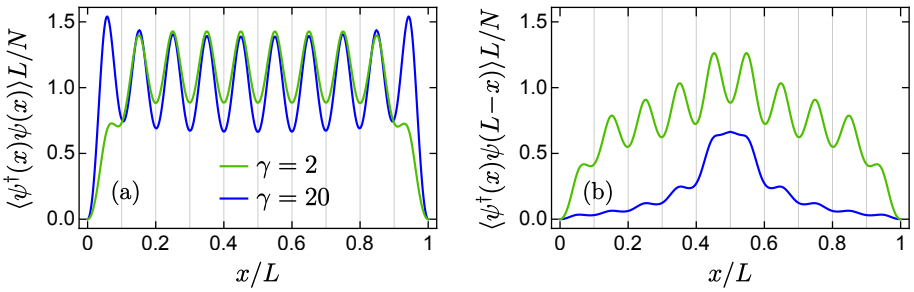

FIG. S1. (a) Density and (b) correlations in Mott-insulating ($\gamma = 20$) and superfluid ($\gamma = 2$) ground states of 10 bosons in 10 wells with $V_0/E_r = 1$, modeled with 20 segments and quartic basis functions, corresponding to Figs. 4(c,d) in the main text.

B. Scaling of error

In Fig. 3(a) of the main text we showed that the error in the ground-state energy falls off as $M^{-2d_{\max}}$ for cDMRG, as opposed to M^{-2} for grid-based discretization. Since the DMRG sweeps have the same dependence on M in both methods, it follows that cDMRG is more efficient for more accurate calculations [Fig. 3(b)]. Here we show that this conclusion also holds for the single-particle correlations. We compare ground states for the Tonks gas ($\gamma \to \infty$) in a box trap ($V_0 = 0$), for which the correlations are known exactly [S2].

In this limit the cusp in the cDMRG basis functions is replaced by the condition that they must vanish whenever two particles coincide. Thus, we solve the modified Legendre equation $\hat{K}_{\mathcal{L}} \, \chi^{(j)}_{n,k} = \mathcal{E}^{(j)}_{n,k} \, \chi^{(j)}_{n,k}$ projected onto the nullspace (or lowest-energy subspace) of $\hat{U}_{\mathcal{L}}$ [Eqs. (S41) and (S42)], effectively generating free-fermion basis functions. Conversely, for discretization one deals with hard-core bosons, limiting the on-site occupations to $\{0, 1\}$.

We evaluate the accuracy of the off-diagonal correlations $\langle \hat{\psi}^\dagger(x) \hat{\psi}(L-x) \rangle$, shown in Fig. 2(d) and Fig. S1(b). We consider an integral measure of the error $\varepsilon(x) := \langle \hat{\psi}^\dagger(x) \hat{\psi}(L-x) \rangle - \langle \hat{\psi}^\dagger(x) \hat{\psi}(L-x) \rangle_{\text{exact}}$, given by

$$\varepsilon_{\text{corr}} := \sqrt{\frac{L}{N} \int_0^L dx \, |\varepsilon(x)|^2} \, . \tag{S47}$$

For discretization we replace the integral by a weighted sum over the grid points. Figure S2(a) shows that $\varepsilon_{\text{corr}}$ falls off as $M^{-(d_{\max}+1)}$ for cDMRG: This is expected since we are able to approximate spatial variations of order $d_{\max}$ in each segment. For discretization $\varepsilon_{\text{corr}}$ scales as M^{-2} since $d_{\max}$ is effectively 1. Thus, one can accelerate convergence

by increasing $d_{\max}$, although $\varepsilon_{\rm corr}$ decays slower than the energy error for $d_{\max} > 1$. In addition, the curves for $N = 5$ and $N = 10$ in Fig. S2(a) roughly coincide if M is rescaled by N, which suggests the scaling $\varepsilon_{\rm corr} \sim (N/M)^{d_{\max}+1}$.

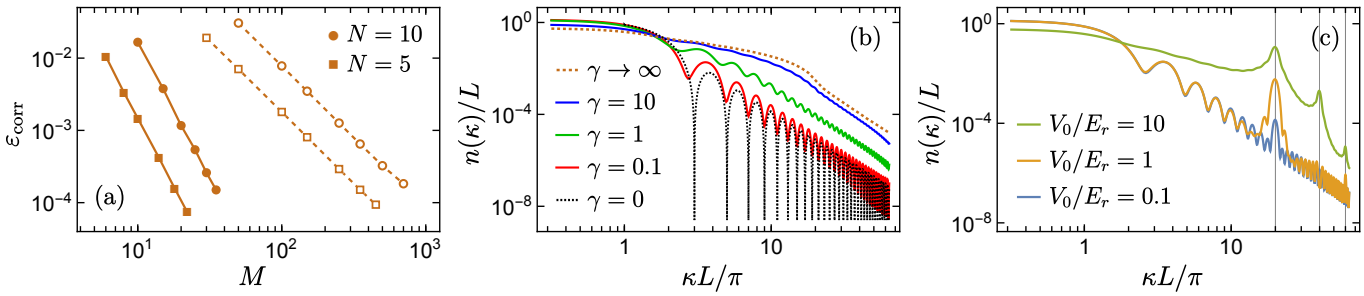

FIG. S2. (a) Correlation error, defined by Eq. (S47), in the ground state of a Tonks gas ($\gamma \to \infty$) with $V_0 = 0$ decaying roughly as $(N/M)^{(d_{\max}+1)}$ for cDMRG with $d_{\max} = 3$ (solid lines) and as $(N/M)^2$ for grid-based discretization (dotted lines). The basis states and DMRG parameters used are listed in Tables S6–S8. (b) Momentum distributions for $V_0 = 0$ and $N = 10$ at different interaction strengths: $\gamma = 0$ and $\gamma \to \infty$ are exact results, $\gamma = 0.1$ and 10 are from cDMRG with $M = 10$ and 20, respectively, corresponding to Fig. 3 in the main text, and $\gamma = 1$ is from cDMRG with $M = 20$, basis A in Table S3, and DMRG parameters in Table S5. All curves fall off asymptotically as $1/\kappa^4$. (c) $n(\kappa)$ for $\gamma = 0.2$, $N = 10$, and $V(x) = V_0 \cos^2(N\pi x/L)$ from cDMRG with the same parameters as in Fig. 4, detailed in Sec. IV. Vertical lines at multiples of $2N$ show locations of satellite peaks.

C. Momentum distribution

The momentum distribution is given by

$$n(\kappa) = \frac{1}{2\pi} \int_0^L dx \int_0^L dx' e^{\mathrm{i}\kappa(x-x')} \langle \hat{\psi}^\dagger(x) \hat{\psi}(x') \rangle \, . \tag{S48}$$

Using the piecewise expression from Eq. (S46) we can compute the distribution as

$$\begin{aligned} n(\kappa) &= \frac{1}{2\pi} \sum_{i,j=1}^{M} \sum_{p,q=0}^{p_{\max}} C_{j,q}^{i,p} \int_{X_{i-1}}^{X_i} dx \, e^{\mathrm{i}\kappa x} \left(\frac{x - X_{i-1}}{w_i} \right)^p \int_{X_{j-1}}^{X_j} dx' \, e^{-\mathrm{i}\kappa x'} \left(\frac{x' - X_{j-1}}{w_j} \right)^q \\ &= \sum_{i,j=1}^{M} \sum_{p,q=0}^{p_{\max}} C_{j,q}^{i,p} \, \frac{e^{\mathrm{i}\kappa(X_{i-1} - X_{j-1})}}{2\pi\kappa^2} \, \frac{\Gamma_l(p+1, -\mathrm{i}\kappa w_i)}{(-\mathrm{i}\kappa w_i)^p} \, \frac{\Gamma_l(q+1, \mathrm{i}\kappa w_j)}{(\mathrm{i}\kappa w_j)^q} \, , \end{aligned} \tag{S49}$$

where Γ_l is the lower incomplete gamma function. For contact interactions one expects $n(\kappa)$ to fall off as κ^{-4} at large momenta [S3]. Figure S2(b) shows that cDMRG faithfully reproduces this universal power law over several orders of magnitude, interpolating between exact results for $\gamma = 0$ and $\gamma \to \infty$ for $V_0 = 0$. The strong modulations at weak interactions can be understood by noting that, for $\gamma = 0$, $\langle \hat{\psi}^\dagger(x) \hat{\psi}(x') \rangle = (2N/L) \sin(\pi x/L) \sin(\pi x'/L)$, such that $n(\kappa) = 2NL\pi(1 + \cos \kappa L)/(\kappa^2 L^2 - \pi^2)^2$, which vanishes for $\kappa L \in \{3\pi, 5\pi, 7\pi, \dots\}$. Figure S2(c) shows how turning on a lattice potential $V(x) = V_0 \cos^2(N\pi x/L)$ produces satellite peaks at $\kappa L = 2N\pi\nu$ for integer ν, as the lattice couples momenta separated by multiples of $2N\pi/L$. For Mott-insulating states in a deep lattice, the correlation length becomes of the order of the lattice spacing, so the power-law tail shifts to higher momenta (large ν).

IV. Parameters used in simulations

As outlined in Sec. II C, we construct the basis by solving an eigenvalue problem, Eq. (S38), in the space of symmetrized monomials of maximum degree $d_{\max}$. The number of monomials grows as $N_{\rm mon} = \binom{n+d_{\max}}{n}$. Provided

the segments are sufficiently narrow, or the repulsive interactions are sufficiently strong ($\gamma \gg 1$), we retain only a few or no basis states for larger n. Table S1 enumerates the total number of monomials and the number of basis states we keep in a typical calculation with strong interactions. Figure S3 shows the average weight, $\bar{N}_{n,k}$, of each basis function in the ground state. These are calculated by averaging over the reduced density matrices describing individual segments, and $\sum_{n,k} \bar{N}_{n,k} = 1$. The basis is ordered so that larger k corresponds to larger $\mathcal{E}_{n,k}$. The weights fall off strongly with both n and k, justifying our truncation. The k dependence is well approximated by a power law, and roughly the same power law is seen for each n.

TABLE S1. Number of monomials and basis states we keep in each segment for the calculation in Fig. 2 of the main text. Here, $d_{\max} = 4$, $N = 5$, $M = 8$, and $\gamma = 50$.

n	N_{mon}	N_{basis}
0	1	1
1	5	5
2	15	15
3	35	35
4	70	10
5	126	2
Total	252	68

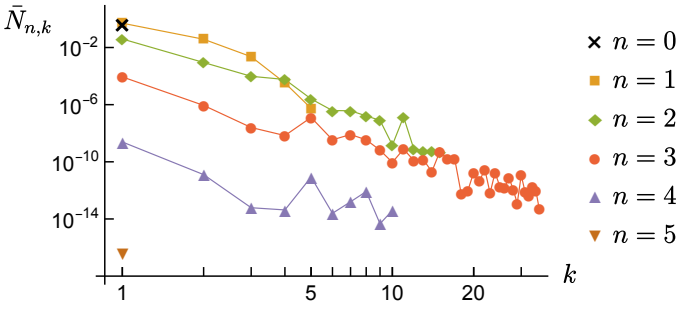

FIG. S3. Average occupation of the basis states enumerated in Table S1 in the ground state shown in Fig. 2. The basis states are indexed by their eigenvalue $\mathcal{E}_{n,k}$ in Eq. (S38).

Once the basis is generated, one can represent the local operators as matrices following Sec. II B and proceed to the DMRG sweeps, for which we used version 3.0.0 of the ITensor library in C++ [S4]. As discussed in the main text, we run multiple DMRG cycles where the energy penalty Λ is successively increased in powers of 10. Here, the practical objective is to produce a good initial state for the final cycle as quickly as possible. Thus, we start from a small maximum bond dimension $D_{\max}$ and gradually increase it with Λ to speed up the initial cycles, as shown in Table S2. Throughout, we discard singular values below a truncation cutoff $\epsilon_{\mathrm{trunc}} = 10^{-14}$. For the final few cycles, $D_{\max}$ is sufficiently large that this threshold is exceeded on all bonds. For each value of Λ, we sweep back and forth until the relative change in the total energy (including the discontinuity penalty) between consecutive sweeps is less than a convergence threshold ϵ_{conv}, which we lower with increasing Λ (as in Table S2). We use a Davidson eigensolver with up to a few tens of maximum iterations $\nu_{\max}$, for which our thresholds are typically met after a small number of sweeps. We terminate the program if, at the end of a cycle, the discontinuity $\sum_j \langle \hat{\Upsilon}_{j,j+1} \rangle / L$ has dropped below a target $T_{\mathrm{disc}} = 10^{-12}$. The CPU- and wall times are measured in seconds for each cycle and for the entire DMRG program using the `clock()` and `chrono::high_resolution_clock::now()` functions, respectively, in C++ on Linux desktops. Since we used quad-core CPUs, wall times are about a quarter of the CPU times.

TABLE S2. DMRG parameters for successive cycles with increasing penalty Λ for the system in Table S1 and Fig. 2: $\nu_{\max}$ is the maximum eigensolver iteration, $\epsilon_{\mathrm{trunc}}$ is the singular-value cutoff, $D_{\max}$ is the maximum bond dimension, ϵ_{conv} is the convergence threshold, and D, N_{sw}, Δt_{CPU}, Δt_{wall} are the resulting bond dimension, number of sweeps, CPU- and wall times.

ΛL	$\nu_{\max}$	$\epsilon_{\mathrm{trunc}}$	$D_{\max}$	ϵ_{conv}	D	N_{sw}	$\Delta t_{\mathrm{CPU}}(s)$	$\Delta t_{\mathrm{wall}}(s)$
10^1	30	10^{-14}	20	10^{-3}	20	3	45	12
10^2	40	10^{-14}	30	10^{-4}	30	3	109	28
10^3	40	10^{-14}	40	10^{-5}	40	4	203	51
10^4	40	10^{-14}	50	10^{-6}	50	4	257	64
10^5	30	10^{-14}	100	10^{-7}	58	5	255	64
10^6	20	10^{-14}	200	10^{-8}	48	5	112	28

Table S3 shows what basis states were used for benchmarking against discretization in Fig. 3 of the main text. The corresponding DMRG parameters are listed in Tables S4 and S5. For the discretized model in Eq. (2), we employed a standard DMRG cycle with $\epsilon_{\rm trunc} = 10^{-14}$ and $\epsilon_{\rm conv} = 10^{-8}$, same as in the final cycle of cDMRG; we used $n_{\rm max} = N$ and $\nu_{\rm max} = 3$ which produced good convergence, and although $D_{\rm max}$ was set to 1000, the actual bond dimensions were comparable to those found using cDMRG, as shown in Fig. S4(a). We varied the number of segments and grid points to control the error ε in the ground-state energy. Figure S4(b) shows that the total number of sweeps is relatively independent of ε for cDMRG, but scales roughly as $\varepsilon^{-1/2}$ for discretization. The wall time for the entire DMRG algorithm, plotted in Fig. S4(c), exhibits a similar scaling as the CPU time in Fig. 3(b) of the main text.

TABLE S3. Number of retained basis states, $N_{\rm basis}$, for $N = 10$ in three cases: A and B were used in benchmarking for $\gamma = 10$ and $\gamma = 0.1$, respectively, in Fig. 3. C was used in Figs. 4 and 5 to explore ground states in a sinusoidal potential with $M = 20$.

Label	$d_{\rm max}$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$	Total
A	3	4	10	20	35	50	30	15	5	2	0	172
B	3	4	10	20	35	56	70	60	50	35	20	361
C	4	5	15	35	70	90	50	25	10	5	0	306

TABLE S4. DMRG parameters used in benchmarking for $N = 10$ and $\gamma = 10$ in Fig. 3. See Table S2 for notation.

ΛL	$\nu_{\rm max}$	$\epsilon_{\rm trunc}$	$D_{\rm max}$	$\epsilon_{\rm conv}$
10^1	20	10^{-14}	20	10^{-4}
10^2	30	10^{-14}	30	10^{-5}
10^3	30	10^{-14}	40	10^{-6}
10^4	30	10^{-14}	50	10^{-7}
10^5	20	10^{-14}	70	10^{-8}
10^6	10	10^{-14}	120	10^{-8}

TABLE S5. DMRG parameters used for $N = 10$ and $\gamma = 0.1$ in Fig. 3 and in the presence of a potential in Figs. 4 and 5.

ΛL	$\nu_{\rm max}$	$\epsilon_{\rm trunc}$	$D_{\rm max}$	$\epsilon_{\rm conv}$
10^1	30	10^{-14}	10	10^{-4}
10^2	40	10^{-14}	20	10^{-5}
10^3	40	10^{-14}	30	10^{-6}
10^4	40	10^{-14}	40	10^{-7}
10^5	30	10^{-14}	60	10^{-7}
10^6	20	10^{-14}	100	10^{-8}

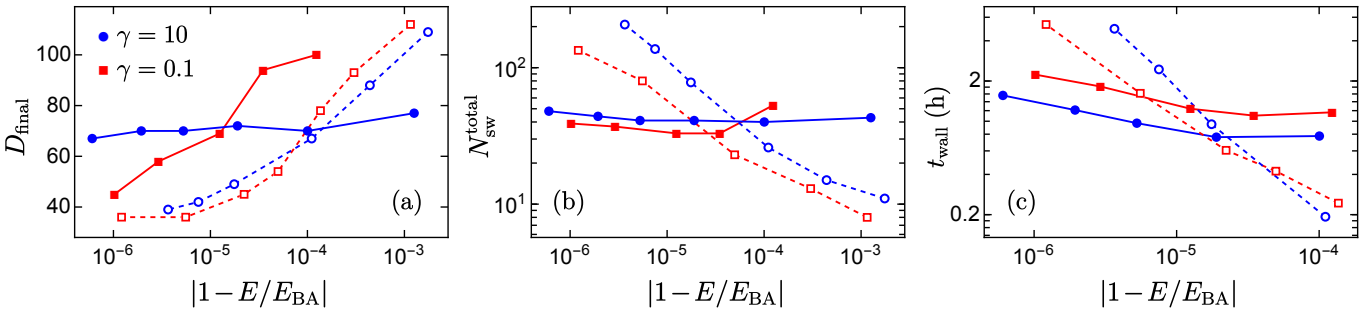

FIG. S4. (a) Final bond dimension, (b) total number of sweeps, and (c) total wall time as a function of the relative error in the ground-state energy for $N = 10$ using cDMRG (solid lines) and the discretized model (dashed lines), corresponding to Fig. 3. The basis and DMRG parameters are given in Tables S3–S5. Dashed lines in (b) approximately follow $N_{\rm sw}^{\rm total} \sim |1-E/E_{\rm BA}|^{-1/2}$.

The ground states in the presence of a sinusoidal potential were obtained using basis C in Table S3 and the sweep parameters in Table S5. The resulting bond dimensions, sweep numbers, and CPU times are shown in Fig. S5. For weak interactions, the first two exhibit peaks where the ground state changes from a Mott insulator to a superfluid. As expected, the CPU time is maximum at weak interactions and weak potentials where the ground state is the most delocalized, requiring a large number of sweeps to populate all basis states.

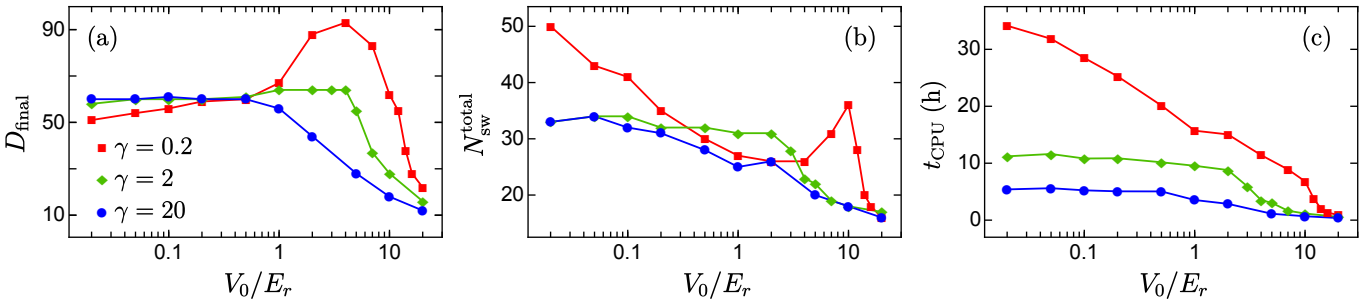

FIG. S5. (a) Final bond dimension, (b) total number of sweeps, and (c) total CPU time as a function of the potential depth V_0, corresponding to Figs. 4 and 5, with $N = 10$, $M = 20$, basis C in Table S3 and DMRG parameters in Table S5.

The basis states and DMRG parameters used in benchmarking the Tonks gas ($\gamma \to \infty$) with $V_0 = 0$ in Fig. S2(a) are given in Tables S6–S8. The small basis dimensions are due to projection onto the lowest-energy subspace of the interaction energy (see Sec. III B). For discretization we use the same settings as above, except with $n_{\rm max} = 1$ and $\epsilon_{\rm conv}$ set the same as in the final cycle of cDMRG.

TABLE S6. Number of retained basis states used in benchmarking the Tonks gas in Fig. S2(a) for $N = 5$ (D) and $N = 10$ (E).

Label	$d_{\rm max}$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$	Total
D	3	4	6	4	1	0	0	0	0	0	0	16
E	3	4	6	4	1	1	1	1	1	1	0	21

TABLE S7. DMRG parameters for $N = 5$ in Fig. S2(a).

ΛL	$\nu_{\rm max}$	$\epsilon_{\rm trunc}$	$D_{\rm max}$	$\epsilon_{\rm conv}$
10^1	40	10^{-14}	20	10^{-5}
10^2	40	10^{-14}	30	10^{-6}
10^3	40	10^{-14}	40	10^{-7}
10^4	40	10^{-14}	50	10^{-8}
10^5	40	10^{-14}	100	10^{-9}
10^6	40	10^{-14}	200	10^{-9}

TABLE S8. DMRG parameters for $N = 10$ in Fig. S2(a).

ΛL	$\nu_{\rm max}$	$\epsilon_{\rm trunc}$	$D_{\rm max}$	$\epsilon_{\rm conv}$
10^1	20	10^{-14}	20	10^{-5}
10^2	30	10^{-14}	30	10^{-6}
10^3	30	10^{-14}	40	10^{-7}
10^4	30	10^{-14}	50	10^{-8}
10^5	20	10^{-14}	70	10^{-9}
10^6	10	10^{-14}	120	10^{-10}

V. Convergence parameters in the algorithm

There are several numerical parameters in our algorithm that control the approach to the continuum limit: the energy penalty for discontinuities (Λ), the maximum number of particles allowed in each segment ($n_{\rm max}$), the maximum polynomial degree in the basis functions ($d_{\rm max}$), the number of segments (M), and the number of basis states retained ($k_{\rm max}$) for $n \in \{0, 1, \ldots, n_{\rm max}\}$. Below we discuss how convergence works with respect to these parameters. There are other parameters intrinsic to the DMRG sweeps, such as the maximum bond dimension (D); however, these behave the same way as in traditional DMRG calculations.

The role of the energy penalty Λ is to impose continuity across neighboring segments. One can rescale the Hamiltonian as $\hat{H}/\Lambda = (1/\Lambda)\hat{H}_s + \hat{\Upsilon}$, where $\hat{H}_s$ is the sum over energies in all segments and $\hat{\Upsilon}$ is the discontinuity operator [see Eq. (4)]. For $\Lambda \to \infty$ the system is projected onto the degenerate nullspace of $\hat{\Upsilon}$, and the ground-state energy E^*

is found by minimizing $\hat{H}_s$ in this space of continuous wave functions. For finite Λ the energy and the wave function of the ground state both gain corrections of order $1/\Lambda$, following degenerate perturbation theory. Using this scaling, one can numerically extrapolate observables to the continuum limit, as shown in Fig. 1(a) inset for the energy. Note this is obtained from a single run of the algorithm as we increase Λ in consecutive DMRG cycles.

The role of $n_{\max}$ is to truncate basis states that are not important at low energies. Thus, it should be large enough that the likelihood of having more particles in a given segment is small. In addition, one requires $n_{\max} \geq 2$ to model the physics of two-body interactions. This is similar to lattice DMRG where, for unity filling, it is usually sufficient, and indeed common practice, to limit the on-site occupations to $n_{\max} \sim 5$ [S5]. For generic continuous systems with equal segments, the particle-number distribution is peaked at $n = N/M$ and falls off exponentially at large n, so it sufficient to have $n_{\max} \gg N/M$. It is generally a good strategy to choose M such that $N/M \lesssim 1$ and set $n_{\max} \gg 1$.

The maximum polynomial degree $d_{\max}$ limits how accurately one can describe spatial variations within a segment. Thus, it is complementary to the number of segments M. As we showed before, the error in the ground-state energy and the single-particle correlations fall off as $M^{-2d_{\max}}$ [Fig. 3(a)] and $M^{-(d_{\max}+1)}$ [Fig. S2(a)], respectively. One can extrapolate such power-law scaling to the continuum limit, e.g. by increasing M for a given $d_{\max}$. Larger values of $d_{\max}$ produces a faster decay in error at the cost of a larger basis size N_{local}, which grows as $N_{\text{local}} = \binom{n_{\max}+d_{\max}+1}{n_{\max}} \sim n_{\max}^{d_{\max}+1}/(d_{\max}+1)!$ if one retains all basis states. The computational cost for a DMRG sweep is $\mathcal{O}(D^3 D_H^2 N_{\text{local}}^2 M)$ for large D [S6], where D_H is the bond dimension of the Hamiltonian MPO. Therefore, if the target error scales as $\varepsilon \sim M^{-\zeta d_{\max}}$, the optimal choice for $d_{\max}$ will minimize $N_{\text{local}}^2 M \sim \left[n_{\max}^{d_{\max}+1}/(d_{\max}+1)!\right]^2 \varepsilon^{-1/(\zeta d_{\max})}$. In particular, this optimal $d_{\max}$ increases for smaller desired error ϵ, smaller ζ, and smaller $n_{\max}$.

Since we have multiple basis states for each value of n for a given $d_{\max}$, it is useful, though not essential, to further truncate the unimportant states and reduce N_{local}. We generate the basis by diagonalizing a local energy functional (see Sec. II C), which gives a natural ordering of the states $\phi_{n,k}^{(j)}$. As shown in Fig. S3, their weights in the ground state fall off exponentially with k, which provides a consistency check for truncating at $k_{\max}$. Generally, these cutoffs $k_{\max}(n)$ should be guided by the local physics. For example, as one expects a much smaller contribution from n close to $n_{\max}$, one may retain fewer states as in Table S3. This is especially true for strong repulsive interactions (Table S6). Typically, $k_{\max}$ is peaked at an intermediate value of n, and there is some trial-and-error involved in checking that the results have indeed converged to the desired precision.

VI. Splitting a basis into finer segments

As discussed in the main text, one needs to split a basis into finer segments for multiscale DMRG approaches [S7] and for obtaining the spatial entanglement at any point other than the segment boundaries. It suffices to consider a single segment with rescaled basis functions $\chi_{n,k}(\vec{r})$ defined over $[0,1]^n$ [see Eq. (S11)]. For notational simplicity, we drop the segment label j in this section. To split $\chi_{n,k}$ at at intermediate point $\tilde{x}$, we first construct basis functions $\chi_{n,k}^{\pm}(\vec{r})$ over the left and right zones, $[0,\tilde{x}]^n$ and $[\tilde{x},1]^n$, in terms of symmetrized monomials, as described below. Then the task is to decompose $\chi_{n,k}(\vec{r})$ in the tensor product basis $\chi^+ \otimes \chi^-$, i.e.,

$$\chi_{n,k}(\vec{r}) = \sum_{n'=0}^{n} \sum_{k^+,k^-} \mathcal{C}_{n',k^+,k^-}^{n,k} \; \mathcal{S}\left[\chi_{n',k^+}^{+}(x_1,x_2,\ldots,x_{n'})\, \chi_{n-n',k^-}^{-}(x_{n'+1},x_{n'+2},\ldots,x_n)\right] , \tag{S50}$$

where $\mathcal{S}$ symmetrizes all n coordinates and the transformation coefficients $\mathcal{C}^{n,k}_{n',k^+,k^-}$ characterize the splitting. Using orthonormality and exchange symmetry of the basis functions, one finds

$$\mathcal{C}^{n,k}_{n',k^+,k^-} = \int_0^1 \mathrm{d}^n r\ \chi_{n,k}(\vec{r})\ \mathcal{S}\big[\chi^{+*}_{n',k^+}(x_1,x_2,\ldots,x_{n'})\ \chi^{-*}_{n-n',k^-}(x_{n'+1},x_{n'+2},\ldots,x_n)\big] \tag{S51}$$

$$= \binom{n}{n'}^{1/2} \int_0^{\tilde{x}} \mathrm{d}x_1 \ldots \int_0^{\tilde{x}} \mathrm{d}x_{n'} \int_{\tilde{x}}^1 \mathrm{d}x_{n'+1}\ \ldots \int_{\tilde{x}}^1 \mathrm{d}x_n\ \chi_{n,k}(\vec{r})\ \chi^{+*}_{n',k^+}(x_1,\ldots,x_{n'})\ \chi^{-*}_{n-n',k^-}(x_{n'+1},\ldots,x_n)\ . \tag{S52}$$

As the basis functions are given in terms of symmetrized monomials [see Eq. (S14)], it is sufficient to compute this integral for $\chi_{n,k}(\vec{r}) = \mathbf{p}(\vec{r})$, $\chi^+_{n',k^+}(x_1,\ldots,x_{n'}) = \mathbf{q}^+(x_1,\ldots,x_{n'})$, and $\chi^-_{n-n',k^-}(x_{n'+1},\ldots,x_n) = \mathbf{q}^-(x_{n'+1},\ldots,x_n)$, with the monomial exponents $\mathbf{p} = \{p_1,\ldots,p_n\}$, $\mathbf{q}^+ = \{q^+_1,\ldots,q^+_{n'}\}$, and $\mathbf{q}^- = \{q^-_{n'+1},\ldots,q^-_n\}$, which yields

$$\mathcal{C}^{\mathbf{p}}_{\mathbf{q}^+,\mathbf{q}^-} = \sqrt{n!\,n'!\,(n-n')!}\ \mathcal{I}_l\big(\{p_1+q^+_1,\ldots,p_{n'}+q^+_{n'}\},\tilde{x}\big)\ \mathcal{I}_r\big(\tilde{x},\{p_{n'+1}+q^-_{n'+1},\ldots,p_n+q^-_n\}\big)\ , \tag{S53}$$

where $\mathcal{I}_l$ and $\mathcal{I}_r$ have closed-form expressions given in Eqs. (S22) and (S23).

One can construct $\chi^\pm_{n,k}$ in terms of monomials using the same procedure as in Sec. II C. For contact interactions, they can be generated from the eigenvalue equations $\big(\hat{K}^\pm_{\mathcal{L}} + c\hat{U}^\pm_{\mathcal{L}}\big)|\chi^\pm_{n,k}\rangle = \mathcal{E}^\pm_{n,k}|\chi^\pm_{n,k}\rangle$ where, in the position basis,

$$\hat{K}^+_{\mathcal{L}} = -\frac{1}{2}\sum_i \frac{\partial}{\partial x_i}\left[x_i(\tilde{x}-x_i)\,\frac{\partial}{\partial x_i}\right], \qquad \hat{U}^+_{\mathcal{L}} = \sum_{i<i'} x_i(\tilde{x}-x_i)\,\delta(x_i-x_{i'})\ , \tag{S54}$$

$$\hat{K}^-_{\mathcal{L}} = -\frac{1}{2}\sum_i \frac{\partial}{\partial x_i}\left[(x_i-\tilde{x})(1-x_i)\,\frac{\partial}{\partial x_i}\right], \qquad \hat{U}^-_{\mathcal{L}} = \sum_{i<i'} (x_i-\tilde{x})(1-x_i)\,\delta(x_i-x_{i'})\ . \tag{S55}$$

Note we have adapted the factors $x_i(1-x_i)$ in Eq. (S38) for the intervals $[0,\tilde{x}]$ and $[\tilde{x},1]$. Following the derivation in Eqs. (S41)–(S44), we find the matrix elements, for $\mathbf{p} = \{p_1,\ldots,p_n\}$ and $\mathbf{q} = \{q_1,\ldots,q_n\}$,

$$\big\langle \mathbf{p}\big|\hat{K}^+_{\mathcal{L}}\big|\mathbf{q}\big\rangle = \frac{n!}{2}\sum_{i=1}^n p_i q_i\Big[\tilde{x}\,\mathcal{I}_l\big(\mathtt{incr}(\mathbf{p}+\mathbf{q},i,-1),\tilde{x}\big) - \mathcal{I}_l(\mathbf{p}+\mathbf{q},\tilde{x})\Big]\ , \tag{S56}$$

$$\big\langle \mathbf{p}\big|\hat{U}^+_{\mathcal{L}}\big|\mathbf{q}\big\rangle = \frac{n!}{2}\sum_{i=1}^{n-1}\Big[\tilde{x}\,\mathcal{I}_l\Big(\mathtt{incr}\big(\mathtt{merge}(\mathbf{p}+\mathbf{q},i),i,1\big),\tilde{x}\Big) - \mathcal{I}_l\Big(\mathtt{incr}\big(\mathtt{merge}(\mathbf{p}+\mathbf{q},i),i,2\big),\tilde{x}\Big)\Big]\ , \tag{S57}$$

$$\big\langle \mathbf{p}\big|\hat{K}^-_{\mathcal{L}}\big|\mathbf{q}\big\rangle = \frac{n!}{2}\sum_{i=1}^n p_i q_i\Big[(1+\tilde{x})\,\mathcal{I}_r\big(\tilde{x},\mathtt{incr}(\mathbf{p}+\mathbf{q},i,-1)\big) - \mathcal{I}_r(\tilde{x},\mathbf{p}+\mathbf{q}) - \tilde{x}\,\mathcal{I}_r\big(\tilde{x},\mathtt{incr}(\mathbf{p}+\mathbf{q},i,-2)\big)\Big]\ , \tag{S58}$$

$$\big\langle \mathbf{p}\big|\hat{U}^-_{\mathcal{L}}\big|\mathbf{q}\big\rangle = \frac{n!}{2}\sum_{i=1}^{n-1}\Big[(1+\tilde{x})\,\mathcal{I}_r\Big(\tilde{x},\mathtt{incr}\big(\mathtt{merge}(\mathbf{p}+\mathbf{q},i),i,1\big)\Big) - \mathcal{I}_r\Big(\tilde{x},\mathtt{incr}\big(\mathtt{merge}(\mathbf{p}+\mathbf{q},i),i,2\big)\Big)$$
$$- \tilde{x}\,\mathcal{I}_r\big(\tilde{x},\mathtt{merge}(\mathbf{p}+\mathbf{q},i)\big)\Big]\ . \tag{S59}$$

The inner product of the symmetrized monomials over $[0,\tilde{x}]^n$ and $[\tilde{x},1]^n$ are simply given by $\langle\mathbf{p}|\mathbf{q}\rangle^+ = n!\,\mathcal{I}_l(\mathbf{p}+\mathbf{q},\tilde{x})$ and $\langle\mathbf{p}|\mathbf{q}\rangle^- = n!\,\mathcal{I}_r(\tilde{x},\mathbf{p}+\mathbf{q})$. Using these results, the construction of $\chi^\pm_{n,k}$ reduces to a generalized eigenvalue problem. In Fig. 5 of the main text, we truncate $\chi^\pm_{n,k}$ the same way as $\chi_{n,k}$ (as detailed in Sec. IV).

VII. Tight-binding approximation with hard walls

In the presence of a sufficiently deep external potential $V(x) = V_0\cos^2(N_w\pi x/L)$, one can approximate the continuum problem by N_w localized Wannier orbitals at the potential minima. To derive this tight-binding model, we consider the single-particle Hamiltonian $\hat{H}_{\mathrm{sp}} = \hat{K} + \hat{P}$, where $\hat{K} = -\partial_x^2/2$ and $\hat{P} = V(x)$ in the position basis. Since

we have hard-wall boundaries at $x = 0$ and $x = L$, the Hilbert space is spanned by the particle-in-a-box wave functions $\alpha_m(x) = \sqrt{2/L}\,\sin(m\pi x/L)$, $m = 1, 2, 3, \ldots$, such that $\langle\alpha_m|\hat{K}|\alpha_{m'}\rangle = \delta_{m,m'}(m/N_w)^2 E_r$, where $E_r = N_w^2\pi^2/(2L^2)$ is the recoil energy. The potential $V(x)$ couples these states with the amplitudes

$$\langle\alpha_m|\hat{P}|\alpha_{m'}\rangle = (V_0/2)\,\delta_{m,m'} + (V_0/4)\big(\delta_{m,m'+2N_w} + \delta_{m,m'-2N_w} - \delta_{m,-m'+2N_w}\big)\,, \tag{S60}$$

which vanish for $m \neq m'$ unless m and m' are separated by or add up to $2N_w$. This selection rule partitions the wave functions into N_w groups characterized by $q = 1, 2, \ldots, N_w$, where $\mathrm{mod}(m \pm q, 2N_w) = 0$. Here, q plays the role of quasimomentum and the lowest-energy eigenstates of $\hat{H}_{\mathrm{sp}}$ for each q constitute the lowest band, $\hat{H}_{\mathrm{sp}}|\psi_q\rangle = E_q|\psi_q\rangle$. We find the Wannier functions as eigenstates of $\hat{X}_{\mathrm{proj}} = \hat{\Pi}\hat{X}\hat{\Pi}$, where $\hat{\Pi}$ is the projector onto the lowest band and $\hat{X}$ is the position operator with matrix elements

$$\langle\alpha_m|\hat{X}|\alpha_{m'}\rangle = [(-1)^{m+m'} - 1]\,\frac{4Lmm'}{\pi^2(m^2 - m'^2)^2}\,. \tag{S61}$$

Figure S6(a) shows the Wannier functions $W_j(x)$ for $N_w = 10$ and $V_0/E_r = 2$, centered at different potential minima j, which become more localized with increasing V_0/E_r. In this Wannier basis, one can calculate the nearest-neighbor tunneling $J_{j,j+1} = -\langle W_j|\hat{H}_{\mathrm{sp}}|W_{j+1}\rangle$ and local energy shifts $\varepsilon_j = \langle W_j|\hat{H}_{\mathrm{sp}}|W_j\rangle$. Contact interactions give rise to the "on-site" interaction energies $U_j = g\int_0^L \mathrm{d}x\,|W_j(x)|^4$. As shown in Figs. S6(b-c), these effective Hubbard parameters are slightly larger close to the edges. In Fig. 4 of the main text, we simulate such nonuniform Hubbard models using a standard DMRG routine in Mathematica version 12.3.0 with singular-value cutoff $\epsilon_{\mathrm{trunc}} = 10^{-12}$, convergence threshold $\epsilon_{\mathrm{conv}} = 10^{-8}$, and maximum bond dimension $D_{\max} = 500$ (cf. Sec. IV).

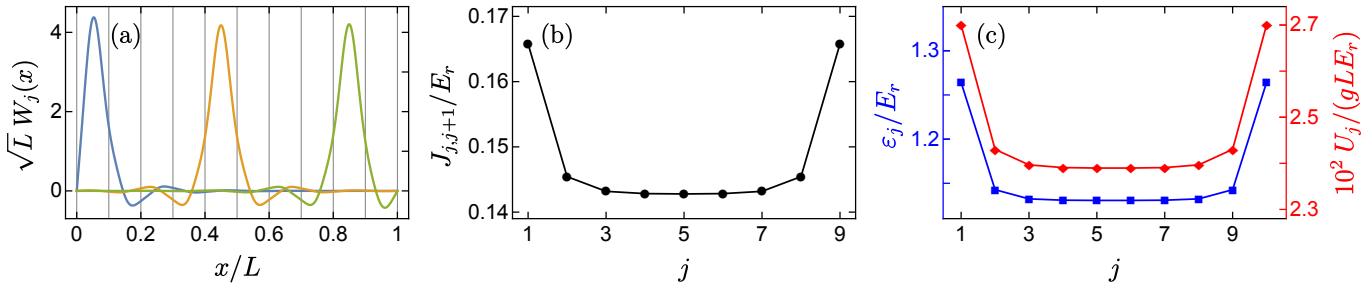

FIG. S6. (a) Wannier functions at the 1st, 5th, and 9th potential minima for $V_0/E_r = 2$ and $N_w = 10$. (b) Tunneling between nearest-neighbor minima and (c) local energy shifts ε_i and interaction energies U_i for the system in (a).

VIII. Condensate fraction in the Tonks gas

In the limit of infinitely strong repulsive interactions, $\gamma \to \infty$, the 1D Tonks gas of impenetrable bosons maps onto free fermions, such that their ground-state wave function $\Psi(\vec{r})$ is given by the absolute value of that of the fermions, $\Psi(\vec{r}) = |\Psi_F(\vec{r})|$ [S8]. For a sinusoidal potential $V(x) = V_0\cos^2(N_w\pi x/L)$ with unit filling, $N_w = N$, $\Psi_F(\vec{r})$ is obtained by populating each of the single-particle eigenstates $\psi_q(x)$ in the lowest band (see Sec. VII) with a fermion and taking the Slater determinant. Therefore,

$$\Psi(\vec{r}) = \frac{1}{\sqrt{N!}}\Big|\det\big[\psi_q(x_i)\big]_{q,i=1,2,\ldots,N}\Big|\,. \tag{S62}$$

The boson occupation of the single-particle modes are contained in one-body density operator $\hat{\rho}_1$, where

$$\rho_1(x,x') = N\int_0^L \mathrm{d}^{N-1}r\ \Psi(x,\vec{r})\,\Psi(x',\vec{r})\ . \tag{S63}$$

In particular, the condensate fraction is given by $f_0 \coloneqq N_0/N$, where N_0 is the occupation of the single-particle ground state (corresponding to $q=1$), $N_0 = \langle\psi_1|\hat{\rho}_1|\psi_1\rangle$. Thus,

$$f_0 = \frac{1}{N}\int_0^L \mathrm{d}x \int_0^L \mathrm{d}x'\ \psi_1(x)\,\rho_1(x,x')\,\psi_1(x')\ . \tag{S64}$$

For $V_0 = 0$, $\psi_q(x) = \sqrt{2/L}\,\sin(q\pi x/L)$ and one can simplify $\rho_1(x,x')$ to a determinant [S2], reducing the calculation of f_0 to a 2D numerical integral. Figure S7(a) shows that, in this case, $f_0 \sim N^{-0.45}$. The condensate fraction f_0 vanishes in the thermodynamic limit, but is finite for fixed N. For $V_0 > 0$, we find the single-particle states $|\psi_q\rangle$ by exact diagonalization and then compute f_0 by an $(N+1)$-dimensional Monte Carlo integration with up to 10^8 points in Mathematica version 12.3.0. The results for $N = 10$ are shown in Fig. 4(b) of the main text.

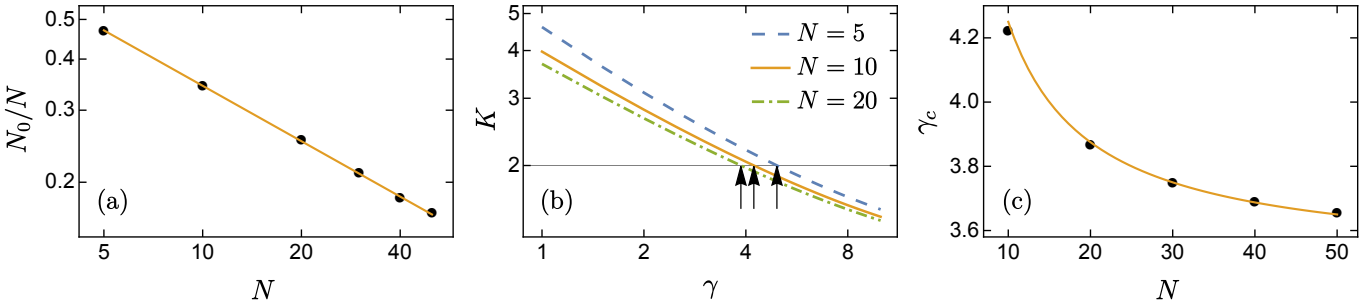

FIG. S7. Ground-state parameters for $V_0 \to 0$: (a) Condensate fraction vs particle number for $\gamma \to \infty$; solid line shows the fit $0.97\,N^{-0.45}$. (b) Luttinger parameter vs interaction strength; arrows show the onset of pinning instability below $K = 2$. (c) Critical interaction strength γ_c, corresponding to $K = 2$, vs particle number; solid line shows the fit $3.5 + 7.5/N$.

IX. Luttinger parameter and pinning instability

The low-energy excitations of our model for the interacting 1D Bose gas are described by a Luttinger liquid [S9], characterized by the parameter $K = v_s/v_N$, where v_s and v_N are the speed of sound and density stiffness, respectively. These can be obtained from the ground-state energy E as

$$v_s = \sqrt{\frac{L^2}{N}\frac{\partial^2 E}{\partial L^2}} \quad \text{and} \quad v_N = \frac{L}{\pi}\frac{\partial^2 E}{\partial N^2}\ , \tag{S65}$$

with $\hbar = m = 1$. In the absence of any external potential $V(x)$, E can be calculated exactly using Bethe Ansatz [S10], thus one can find K as a function of the interaction strength γ for a given particle number N, as shown in Fig. S7(b). Crucially, perturbative calculations have shown [S11] that for $K < 2$, a Luttinger liquid has an instability whereby it is pinned to an insulating state by an arbtrarily weak commensurate potential, $V(x) = V_0\cos^2(N_w\pi x/L)$ with N/N_w = integer. Thus, $K = 2$ marks the transition from a superfluid to a Mott insulator for $V_0 \to 0$. In Fig. S7(c), we plot the corresponding interaction strength γ_c, which is well fitted by $\gamma_c \approx 3.5 + 7.5/N$.

[S1] S. Zhang and J. Jin, *Computation of Special Functions* (Wiley, New York, 1996).

[S2] P. J. Forrester, N. E. Frankel, T. M. Garoni, and N. S. Witte, Painlevé transcendent evaluations of finite system density matrices for 1d impenetrable bosons, Commun. Math. Phys. **238**, 257 (2003).

[S3] M. Olshanii and V. Dunjko, Short-distance correlation properties of the Lieb-Liniger system and momentum distributions of trapped one-dimensional atomic gases, Phys. Rev. Lett. **91**, 090401 (2003).

[S4] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor software library for tensor network calculations (2020), arXiv:2007.14822.

[S5] C. Kollath, U. Schollwöck, J. von Delft, and W. Zwerger, One-dimensional density waves of ultracold bosons in an optical lattice, Phys. Rev. A **71**, 053606 (2005).

[S6] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[S7] M. Dolfi, B. Bauer, M. Troyer, and Z. Ristivojevic, Multigrid algorithms for tensor network states, Phys. Rev. Lett. **109**, 020604 (2012).

[S8] M. Girardeau, Relationship between systems of impenetrable bosons and fermions in one dimension, J. Math. Phys. **1**, 516 (1960).

[S9] M. A. Cazalilla, Bosonizing one-dimensional cold atomic gases, J. Phys. B **37**, S1 (2004).

[S10] M. T. Batchelor, X.-W. Guan, N. Oelkers, and C. Lee, The 1D interacting Bose gas in a hard wall box, J. Phys. A **38**, 7787 (2005).

[S11] H. P. Büchler, G. Blatter, and W. Zwerger, Commensurate-incommensurate transition of cold atoms in an optical lattice, Phys. Rev. Lett. **90**, 130401 (2003).