

Dressing Composite Fermions with Artificial Intelligence

Mytraya Gattu *

Department of Physics, 104 Davey Lab, [Pennsylvania State University](#), University Park, Pennsylvania 16802, USA



(Received 30 November 2025; accepted 15 September 2026; published 28 October 2026)

Recent variational studies have demonstrated that the strongly correlated ground states of the fractional quantum Hall (FQH) effect can be captured using machine learning approaches starting from no prior knowledge of the underlying physics. We introduce a complementary framework that instead starts from Jain's composite-fermion (CF) wave functions, which accurately describe FQH states as weakly interacting states of CFs at fillings $\nu = n/(2pn + 1)$ in an idealized limit. As we move away from this idealized limit to one more in line with experimental reality, we expect CFs to become dressed much like the electrons of a noninteracting system, which are dressed by neutral excitations as the interaction is turned on adiabatically, as in Landau's Fermi-liquid theory. We model this dressing using a Feynman-Cohen-style backflow approach, implemented through symmetry-preserving neural networks, which we refer to as CF-Flow. CF-Flow achieves competitive accuracy with substantially greater computational efficiency and scales to systems of $\gtrsim 26$ electrons. At fillings $\nu = 1/3$ and $2/5$, CF-Flow produces ground-state energies as functions of Landau-level mixing strength that are nearly indistinguishable from those obtained using the fixed-phase diffusion Monte Carlo (fp-DMC) method, with low local-energy standard deviations, even though the latter constrains the wave function phase to that of the lowest Landau level, thereby providing insight into why fp-DMC has been successful in giving an accurate quantitative account of several experiments. Finally, the symmetry-preserving architecture of CF-Flow enables access to excited states and computation of the transport gap at $\nu = 1/3$ and $\nu = 2/5$. Within the FQH sectors studied, the extrapolated gap is well described by an exponential decrease with Landau-level mixing and approaches a finite value at large mixing strength.

DOI: [10.1103/zq76-147c](#)

I. INTRODUCTION

That any progress has been achieved in understanding the behavior of interacting systems of quantum mechanical particles (the subject of much of modern condensed matter physics) is extremely remarkable. At the outset, the problem seems impossible: Given that even the problem of three interacting electrons eludes an exact solution, how are we to understand the behavior of 10^{21} electrons?

For metals, Landau's Fermi-liquid theory offers valuable insight [1]. The theory begins with noninteracting electrons, for which we know the full solution: The ground state is a Fermi sea, and the excitations are electron-hole pairs. One then assumes that as the interaction is slowly turned on, no phase transition occurs, and the electrons of the noninteracting system evolve into Landau quasiparticles, i.e., entities that retain the same quantum numbers as electrons but are "dressed" by a screening cloud of particle-hole excitations emerging from the Fermi sea. The properties of these quasiparticles near the Fermi energy can then be estimated through perturbative techniques.

In this paper, we consider a strongly interacting system, namely, electrons in two dimensions exposed to a strong perpendicular magnetic field—the setup of the fractional quantum Hall effect (FQHE) [2,3]. This is actually a much more difficult problem than that in the preceding paragraph. The reason is that when electrons occupy the lowest quantum of the kinetic energy [namely, the lowest Landau level (LLL)], their behavior is determined entirely by interactions, and it is not adiabatically connected to noninteracting electrons (which form a macroscopically degenerate system). We do not have a natural starting point akin to the Fermi sea of the Landau Fermi-liquid theory. How, then, are we to tackle this problem?

Several recent studies [4,5] have applied neural quantum states (NQS) [6–8] to this problem, leveraging the universal expressivity of neural networks to represent quantum many-body wave functions within variational Monte Carlo (VMC). These works employ the Psiformer architecture [9], a transformer-based NQS that uses attention mechanisms [10] to model electron correlations. Remarkably, they demonstrate that the dramatic reorganization from a macroscopically degenerate noninteracting system to gapped, incompressible fractional quantum Hall (FQH) states can be captured by starting from bare electrons and dressing them using neural networks. As noted in Ref. [5], however, learning all correlations *ab initio* can become computationally demanding and currently limits accessible system sizes, though recent pretraining strategies [11] offer promising avenues for scaling Psiformer-like ansätze to larger systems.

*Contact author: mytraya@mit.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](#) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

We take an alternative approach, building on the fact that we do know the solution with a high degree of accuracy for an idealized model: electrons in a strictly two-dimensional space with pure Coulomb interactions and no mixing between different Landau levels (LLs). The accuracy of the Jain wave functions [12,13] has been confirmed by comparisons with exact solutions known for small systems from computer calculations [13–17]. Remarkably, this solution is formulated in terms of weakly interacting (or, in the idealized limit, noninteracting) composite fermions (CFs), related to the solution of noninteracting electrons at an effective magnetic field. As the system deviates from this ideal limit, whether through finite quantum well width [18–24], LL mixing [25,26], or considering CF states in higher LLs such as the second LL [27], the interactions between CFs increase and they become dressed by CF particle-hole excitations. When the system is close to the ideal limit, we expect only quantitative corrections, but eventually these effects will cause a transition into a different state. We model this dressing through a backflow transformation that captures how each CF’s motion influences others, an approach pioneered by Feynman and Cohen for liquid helium [28]. Our framework, which we call CF-Flow, combines Jain’s CF wave functions with symmetry-preserving neural networks [29–32] to implement this backflow [28,33–35]. Importantly, CF-Flow is a general approach applicable whenever interactions between CFs go beyond the perturbative regime where analytical techniques remain tractable [36–44].

In this work, we focus on LL mixing as a first application of CF-Flow and, for simplicity, do not include the effect of finite quantum well width. LL mixing is quantified by the dimensionless parameter κ ,

$$\kappa = \frac{e^2}{\epsilon \ell_B \hbar \omega_c}, \quad (1)$$

where $\ell_B = \sqrt{\hbar c/(eB)}$ is the magnetic length, $\omega_c = eB/m^*$ is the cyclotron frequency, and m^* is the electron effective mass. The “LLL limit” corresponds to $\kappa \rightarrow 0$, but experiments operate far from this idealized regime. Even at the highest accessible magnetic fields, electron-type GaAs samples (which have very small m^*) reach only $\kappa \gtrsim 0.4$, while hole systems reach $\kappa \gtrsim 2.5$; typical experiments operate in the range $\kappa \sim 0.50$ –10 [13]. It has long been thought that this mismatch between κ values in theoretical calculations and experimental conditions contributes to differences in calculated and measured energy gaps [38,39,45].

Moreover, LL mixing is known to induce phase transitions, such as the transition at filling factor $\nu = 1/4$ from a CF metal [46–49] to a paired (superconducting) state of CFs (manifesting as an even-denominator FQH state) [26,50], and transitions at low fillings between liquid FQH states and crystalline phases [25,51,52]. Perhaps most notably, LL mixing is thought to underlie the long-standing discrepancy between theory and experiment regarding the nature of the $\nu = 5/2$ ground state [41,53–61], a proposed platform for topological quantum computing [62]. However, these transitions and typical experimental conditions occur at values of κ beyond the perturbative regime for which analytical techniques have been developed, motivating the use of nonperturbative

variational approaches, such as CF-Flow, to access this physics. (The filling factor ν is defined as $\rho\phi_0/B$, where ρ is the transverse electron density and $\phi_0 = hc/e$ is the magnetic flux quantum.)

Previous approaches have employed the fixed-phase diffusion Monte Carlo (fp-DMC) method [25,26,63–67], which gives the lowest energy in a given phase sector. While this method does give a strict upper bound on the ground-state energy, its accuracy depends on the choice of the phase of the wave function (with any choice being somewhat *ad hoc*). CF-Flow provides an independent approach to this problem, using an order of magnitude fewer parameters than other machine learning (ML) architectures while training substantially faster and scaling to 26 or more electrons. At the Jain filling fractions, CF-Flow yields accurate ground-state energies as a function of the strength of LL mixing, ascertained from the low local-energy variance.

While our primary objective in this article is to develop CF-Flow, the framework already yields useful insight into LL mixing. A notable observation arises from comparison with fp-DMC [25]: Despite fp-DMC constraining the wave function phase to the LLL, the two approaches produce essentially identical ground-state energies at the Jain fillings. This agreement supports the central assumption underlying fp-DMC and helps explain its success in reproducing experimental measurements, from spin-polarization boundaries [67] to transitions between CF metals, CF superconductors [26], and liquid and crystal phases [25,51,68], even though real systems operate far from the LLL limit.

CF-Flow’s symmetry-preserving architecture also enables access to excited states by initializing from the appropriate CF wave function in a chosen angular-momentum sector. We illustrate this by computing the transport gap at $\nu = 1/3$ and at $\nu = 2/5$ across a wide range of LL-mixing strengths. The gap decreases exponentially with increasing κ and approaches a finite value in the $\kappa \rightarrow \infty$ limit. This suggests that, within the FQH phase, the transport gap does not close as a function of LL mixing. Our prediction for the gap as a function of LL mixing can be tested experimentally through activated transport measurements as LL mixing is varied.

The rest of the paper is organized as follows. In Sec. II, we review Haldane’s spherical geometry [69], and in Sec. III, we summarize CF theory and the construction of CF wave functions for ground and excited states at the Jain fillings. Section IV adapts the Feynman-Cohen backflow approach to the sphere and defines the CF-Flow ansatz. Section V presents our training methodology and results: ground-state energies at $\nu = 1/3$ and $2/5$ as functions of κ , and the transport gap at $\nu = 1/3$ and $\nu = 2/5$. Finally, Sec. VI offers discussion and conclusion.

II. BACKGROUND

In this section, we introduce the spherical geometry [69] used to study the bulk properties of FQH liquids and describe the corresponding single-particle wave functions. We then define the many-body problem and briefly discuss the commonly employed LLL limit.

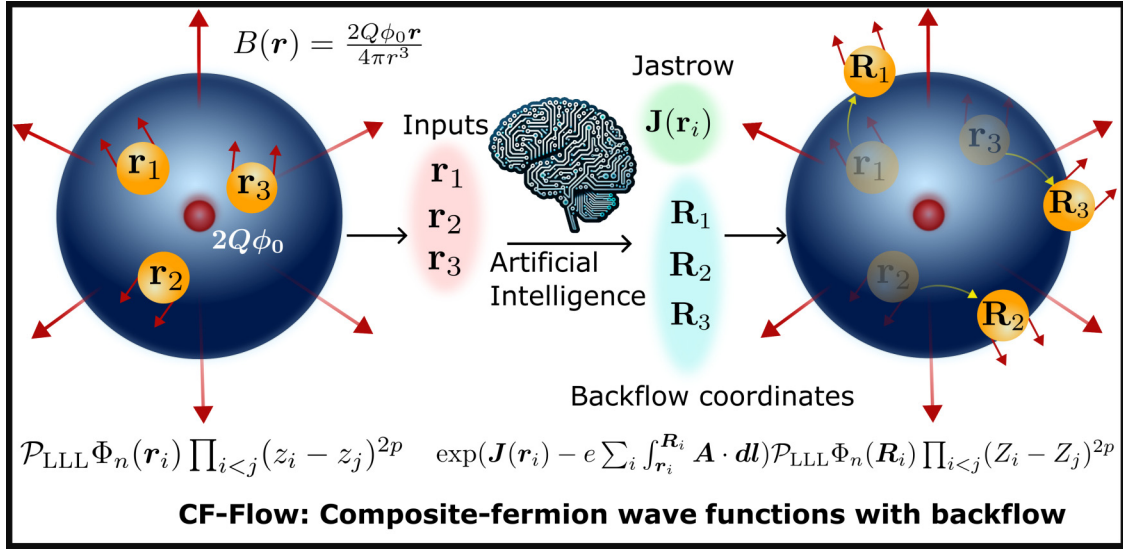


FIG. 1. Schematic illustration of CF-Flow. In the Feynman-Cohen backflow picture [28], turning on interactions causes particles to shift from their noninteracting positions $\mathbf{r}_i$ to backflow positions $\mathbf{R}_i$. CF-Flow implements this idea by combining the backflow approach with CF wave functions defined on the Haldane sphere using symmetry-preserving machine-learning methods, enabling us to study regimes in which interactions between CFs become significant.

A. Single-particle physics on the Haldane sphere

Consider the setup of the FQHE: interacting electrons confined to two dimensions in the presence of a strong perpendicular magnetic field. To study such a system, we employ Haldane's spherical geometry [69], where electrons lie on the surface of a sphere of radius R with a magnetic monopole of charge $Q \in \mathbb{Z}/2$ at its center (Fig. 1).

For bulk studies, this setup offers several advantages. The compact, edge-free surface suppresses finite-size corrections from bulk-edge coupling (in contrast to the disk or planar geometry). Its genus-zero topology [70] ensures a unique gapped ground state for these liquids, unlike the torus, where states at filling factor $\nu = p/q$ form a q -fold degenerate ground-state manifold. Moreover, distinct topological sectors at a given filling ν are labeled by the Wen-Zee shift [71] S , with $N_\phi = N/\nu - S$ for N electrons. We describe the single-particle problem next.

The magnetic monopole generates a uniform radial magnetic field of strength $B = 2Q\phi_0/4\pi R^2$, where $\phi_0 = hc/e$ is the magnetic flux quantum. In the commonly used gauge $\mathbf{A} = -(\hbar c Q \cot \theta / eR) \hat{\phi}$ [13,69], the kinetic Hamiltonian $\mathcal{K}$ is

$$\mathcal{K} = \frac{\hbar\omega_c}{2Q} \left[-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \left(Q \cot \theta + \frac{i}{\sin \theta} \frac{\partial}{\partial \phi} \right)^2 \right], \quad (2)$$

where $\omega_c = eB/m^*c$ is the cyclotron frequency.

The eigenstates of $\mathcal{K}$ are the monopole harmonics $Y_{Q,l,m}(\theta, \phi)$ [13,69,72], labeled by $l = |Q|, |Q| + 1, \dots$, and $m = -l, -l + 1, \dots, l$, with kinetic energy

$$\mathcal{K}Y_{Q,l,m}(\theta, \phi) = \hbar\omega_c \frac{l(l+1) - Q^2}{2Q} Y_{Q,l,m}(\theta, \phi). \quad (3)$$

States with the same l form a $(2l+1)$ -fold degenerate shell (corresponding to the $2l+1$ possible values of m) and constitute the $(l - |Q|)$ th LL on the sphere.

In the LLL, the monopole harmonics take a particularly simple form in terms of the spinor variables $u = \cos(\theta/2)e^{i\phi/2}$ and $v = \sin(\theta/2)e^{-i\phi/2}$:

$$Y_{Q,Q,m}(\theta, \phi) = (-1)^{Q-m} \sqrt{\frac{2Q+1}{4\pi}} \binom{2Q}{Q-m} u^{Q+m} v^{Q-m}. \quad (4)$$

In higher LLs, they are multiterm polynomials involving u, v, u^*, v^* [13]. (Here, for $z = x + iy$, z^* denotes its complex conjugate, $x - iy$.)

An important subtlety distinguishes LL wave functions on the sphere from their planar counterparts. In the plane, the vector potential $\mathbf{A}$ can be chosen to be regular everywhere, and there is no energetic obstruction to mixing orbitals at different magnetic field strengths. On the sphere, $\mathbf{A}$ is singular at the poles (Dirac strings): Any wave function that contains monopole harmonics at charge $Q' \neq Q$ (equivalently $B' \neq B$) generically acquires infinite kinetic energy (see Appendix A and Ref. [5] for Psiformer-like ansätze).

This constrains variational states: They must lie entirely in the correct monopole sector to avoid formally infinite kinetic energy. Energy minimization may eventually favor that sector, but intermediate configurations during optimization can spike to anomalously large energies. For CF-Flow, built from Jain CF wave functions (which are built entirely from monopole harmonics at the correct monopole charge), this requirement is satisfied via the inclusion of an Aharonov-Bohm phase factor (see Sec. IV).

Having established the single-particle physics, we now turn to the many-body problem.

B. The many-body problem

In this work, our focus is on understanding how the Jain FQH states at fillings $\nu = n/(2pn + 1)$ (which have a Wen-Zee shift $S = n + 2p$) evolve as we move away from the zero LL mixing limit. We restrict our attention to the ideal

two-dimensional limit, in which electrons interact via the Coulomb interaction.

Formally, the problem is to identify the ground (and excited) state(s) of the Hamiltonian (in units of $e^2/\epsilon\ell_B$) for N electrons facing a monopole of charge $2Q = N/\nu - (n + 2p)$:

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathcal{K}_i}{\kappa} + \sum_{i<j=1}^N \frac{1}{\sqrt{Q}|\Omega_i - \Omega_j|}. \quad (5)$$

Here, κ is the LL mixing parameter, defined as the ratio of the Coulomb energy scale $e^2/\epsilon\ell_B$ to the cyclotron energy $\hbar\omega_c$. We use the shorthand notation $\Omega_i \equiv (\theta_i, \phi_i)$ to refer to the i th electron's position. The Hamiltonian is rotationally invariant (up to a redefinition of the gauge defining the vector potential). Consequently, the many-body eigenstates are labeled by total angular momentum L and azimuthal quantum number M . For example, the gapped FQH ground states occur at $L = M = 0$.

While experiments are typically conducted in the $\kappa \sim 0.50$ – 10 regime, most of our theoretical understanding of the observed phenomenology of FQH liquids has been achieved by working in the “LLL limit” corresponding to $\kappa \rightarrow 0$. In this limit, the kinetic energy is fully quenched (i.e., all electrons inhabit only the LLL) and the system is governed solely by Coulomb interactions projected into the LLL:

$$\mathcal{H}_{\text{LLL}} = P_{\text{LLL}} \left(\sum_{i<j=1}^N \frac{1}{\sqrt{Q}|\Omega_i - \Omega_j|} \right) P_{\text{LLL}}, \quad (6)$$

$$\Psi_{\frac{n}{2pn+1}}(\Omega_1, \dots, \Omega_N) = P_{\text{LLL}} \underbrace{\begin{vmatrix} Y_{Q^*, Q^*, -Q^*}(\Omega_1) & \dots & Y_{Q^*, Q^*, -Q^*}(\Omega_N) \\ \vdots & & \vdots \\ Y_{Q^*, Q^*, Q^*}(\Omega_1) & \dots & Y_{Q^*, Q^*, Q^*}(\Omega_N) \\ Y_{Q^*, Q^*+1, -(Q^*+1)}(\Omega_1) & \dots & Y_{Q^*, Q^*+1, -(Q^*+1)}(\Omega_N) \\ \vdots & & \vdots \\ Y_{Q^*, Q^*+n-1, (Q^*+n-1)}(\Omega_1) & \dots & Y_{Q^*, Q^*+n-1, (Q^*+n-1)}(\Omega_N) \end{vmatrix}}_{\phi_n(\Omega_1, \dots, \Omega_N; Q^*)} \underbrace{\prod_{i<j=1}^N (u_i v_j - u_j v_i)^{2p}}_{\phi_1^{2p}}. \quad (7)$$

Here, ϕ_n is a Slater determinant of N electrons facing a monopole of charge $Q^* = N/2n - n/2$ and filling n LLs, and $\phi_1^{2p} = \prod_{i<j} (u_i v_j - u_j v_i)^{2p}$ is the Jastrow factor that attaches $2p$ vortices to each electron and turns them into CFs. (Alternatively, a CF state can be thought of as a special case of the more general parton states [73,74] and a product of one $\nu = n$ parton and of $2p$ $\nu = 1$ partons.) $u_i = \cos(\theta_i/2)e^{i\phi_i/2}$ and $v_i = \sin(\theta_i/2)e^{-i\phi_i/2}$ are the spinor representations of Ω_i . P_{LLL} is the projection operator into the LLL, which is essential to the quantitative accuracy of CF wave functions [13]. In practice, P_{LLL} is implemented through the Jain-Kamilla (JK) projection [75,76], which yields a computationally tractable (but slightly modified) form and enables studies of systems with $\gtrsim 900$ CFs [77]. Surprisingly, the JK-projected CF wave function bears a striking resemblance to the Psiformer-like ansatz [5,9] (introduced more than a quarter century later).

where P_{LLL} is the operator that projects each electron's state into the LLL. The zero-point energy has been absorbed into the Hamiltonian definition (we follow this convention henceforth). In this limit, Jain's CF wave functions are known to be extremely accurate (overlaps typically exceeding 99% with exact ground states [13–17]) and furnish the reference wave functions that CF-Flow extends to $\kappa \neq 0$.

III. COMPOSITE-FERMION THEORY

In this section, we provide a brief overview of CF theory and Jain's CF wave functions [12,13].

The CF theory [12] posits that electrons constrained to the LLL capture an even number of vortices ($2p$, where $p \in \mathbb{Z}^+$) to transform into *weakly interacting* composite fermions. This vortex attachment screens the external magnetic field, causing the CFs to experience a reduced monopole charge $Q^* = Q - p(N - 1)$ and inhabit their own LL-like structures, known as Λ levels. These Λ levels are separated by the CF cyclotron energy, $\hbar\omega_c^*$, which arises entirely from interelectron interactions ($\propto e^2/\epsilon\ell_B$) and scales as $\sqrt{Q^*}$ [13].

The central principle behind the theory is that the integer quantum Hall (IQH) effect of CFs inhabiting Λ levels manifests as the FQHE of electrons. In particular, an integer filling ($\nu^* = n$) of Λ levels creates an incompressible state of CFs, which manifests as the FQHE of electrons at the Jain fillings $\nu = n/(2pn + 1)$. The resulting many-body state $\Psi_{\frac{n}{2pn+1}}$ is written as

Both contain multiparticle orbitals $\phi_i(z_i; \{z_{j \neq i}\})$ that depend on the positions of all other electrons; the interested reader can refer to Appendix B for more details.

Beyond ground states, CF theory also provides an accurate description of low-lying excitations. (Higher excited states such as electron/holelike excitations [78–80] can be accessed through the method of CF diagonalization [81,82].) The lowest-energy neutral excitations at filling $\nu = n/(2pn + 1)$ are CF particle-hole excitations (CF excitons), created by promoting a CF from the $(n - 1)$ th Λ level to the n th Λ level [13,83]. We construct an excited state with total angular momentum (L, M) by superposing quasihole-quasiparticle configurations:

$$\Psi_{\frac{n}{2pn+1}}^{L,M} = P_{\text{LLL}} \sum_{m_{\text{QH}}, m_{\text{QP}}} c_{L_{\text{QH}}, -m_{\text{QH}}; L_{\text{QP}}, m_{\text{QP}}}^{L,M} \phi_n^{m_{\text{QH}}, m_{\text{QP}}} \phi_1^{2p}. \quad (8)$$

Here, $\phi_n^{m_{\text{QH}}, m_{\text{QP}}}$ is the Slater determinant of N electrons filling n LLs with an electron missing in the $L_z = m_{\text{QH}}$ orbital of the $(n-1)$ th LL (contributing angular momentum $L_{\text{QH}} = Q^* + n - 1$) and an additional electron in the $L_z = m_{\text{QP}}$ orbital of the n th LL (contributing $L_{\text{QP}} = Q^* + n$). A quasihole corresponds to a missing CF from the $(n-1)$ th Λ level and a quasiparticle corresponds to an additional CF in the n th Λ level. The coefficients $C_{L_{\text{QH}}, -m_{\text{QH}}; L_{\text{QP}}, m_{\text{QP}}}^{L, M}$ are Clebsch-Gordan coefficients [84].

The state with total angular momentum $L = L_{\text{QH}} + L_{\text{QP}}$ and $L_z = L$ contains only a single Fock-space configuration $(m_{\text{QH}}, m_{\text{QP}}) = (-L_{\text{QH}}, L_{\text{QP}})$. In the thermodynamic limit, this state corresponds to an infinitely separated quasihole-quasiparticle pair. The transport gap Δ is defined as the energy difference between this exciton state and the uniform ground state $\Psi_{\frac{n}{2pm+1}}$. This gap governs the activated temperature dependence of the longitudinal resistance near the $\nu = n/(2pn+1)$ plateau, $R_{xx} \sim \exp[-\Delta/(2k_B T)]$ [85]. In Sec. VC, we use this state as the starting point in the CF-Flow framework to evaluate the κ dependence of Δ at the Jain fillings studied below.

Having established the CF framework for the idealized LLL limit, we now describe how to combine the backflow approach with CF wave functions [Eqs. (7) and (8)] to obtain variational ansätze for the lowest-energy states within a given angular-momentum sector.

IV. THE BACKFLOW ANSATZ

In this section, we formulate the backflow approach for the spherical geometry, following the formal construction of Ref. [35].

Let us imagine that we would like to obtain the ground-state wave function $\Psi_{\text{GS}}^{L, M}(\Omega_1, \dots, \Omega_N)$ in the angular-momentum sector (L, M) [recall that the eigenstates of the many-body Hamiltonian $\mathcal{H}$ from Eq. (5) are labeled by (L, M)]. Formally, if we know that a state $\Psi_0^{L, M}(\Omega_1, \dots, \Omega_N)$, for example, a state of CFs such as the one in Eq. (7) or Eq. (8), has a nonzero overlap (however small) with $\Psi_{\text{GS}}^{L, M}$, we can obtain the latter from the former through imaginary time evolution. Focusing on the Hamiltonian from Eq. (5) and dropping the (L, M) indices for visual clarity:

$$\Psi_{\text{GS}}(\Omega_1, \dots, \Omega_N) = \lim_{\tau \rightarrow \infty} e^{-\tau \mathcal{H}} \Psi_0(\Omega_1, \dots, \Omega_N). \quad (9)$$

In terms of the infinite-time real-space propagator $\mathcal{G}(\{\Omega\}; \{\Omega'\})$, which is the infinite-time transition amplitude between a state of N electrons localized at $\Omega'_1, \dots, \Omega'_N$ and one at $\Omega_1, \dots, \Omega_N$,

$$\mathcal{G}(\{\Omega\}; \{\Omega'\}) = \langle \{\Omega\} | \lim_{\tau \rightarrow \infty} e^{-\tau \mathcal{H}} | \{\Omega'\} \rangle, \quad (10)$$

we can rewrite the above equation as

$$\Psi_{\text{GS}}(\{\Omega\}) \propto \int \prod_i d\Omega'_i \mathcal{G}(\{\Omega\}; \{\Omega'\}) \Psi_0(\{\Omega'\}). \quad (11)$$

Formally, the ground state can thus be obtained by evaluating $\mathcal{G}$. An analytic evaluation is of course impossible, but let us see if we can constrain it in some way.

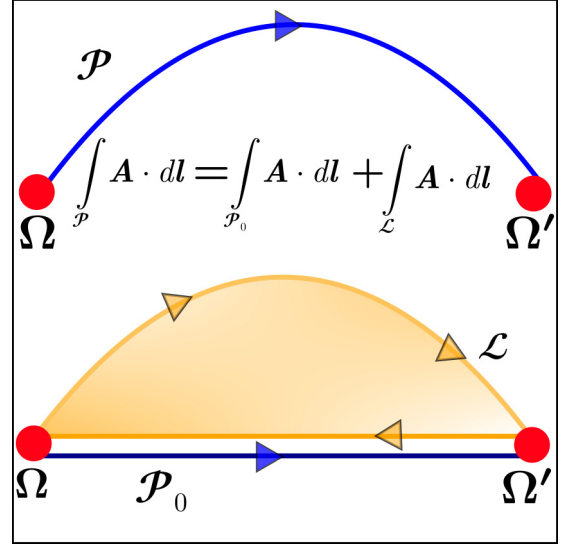


FIG. 2. Decomposition of the Aharonov-Bohm phase. The line integral $\int \mathbf{A} \cdot d\mathbf{l}$ along the path $\mathcal{P}$ between Ω and Ω' (top) can always be decomposed into one along a canonical geodesic path $\mathcal{P}_0$ and along a closed loop path $\mathcal{L}$ (bottom), where the loop contribution depends only on the enclosed flux, a gauge invariant quantity, which consequently does not depend on the location of $\mathcal{L}$ (since the magnetic field B is uniform everywhere).

Let us first note that if the i th electron moves along a path $\mathcal{P}$ between points Ω_i and Ω'_i , then the wave function must acquire a phase:

$$\exp\left(-i \frac{e}{\hbar c} \int_{\mathcal{P}} \mathbf{A} \cdot d\mathbf{l}\right), \quad (12)$$

with $d\mathbf{l}$ being the line element along $\mathcal{P}$. Imagine, however, that we know only the end points of $\mathcal{P}$ [as in Eq. (10)]. In this scenario, let us define a canonical path $\mathcal{P}_0$ (such as the geodesic connecting Ω_i to Ω'_i). We can then always decompose the phase acquired by the wave function as

$$\exp\left(-i \frac{e}{\hbar c} \int_{\mathcal{P}_0} \mathbf{A} \cdot d\mathbf{l}\right) \underbrace{\exp\left(-i \frac{e}{\hbar c} \oint_{\mathcal{L}} \mathbf{A} \cdot d\mathbf{l}\right)}_{\mathcal{W}_{\mathcal{L}}}. \quad (13)$$

Here, $\mathcal{W}_{\mathcal{L}}$ is the phase along some loop (or sum of loops) $\mathcal{L}$; the line integral along any path $\mathcal{P}$ can always be decomposed into the phase acquired along $\mathcal{P}_0$ and $\mathcal{W}_{\mathcal{L}}$. The key insight is that $\mathcal{W}_{\mathcal{L}}$ is gauge invariant and equals the flux through $\mathcal{L}$. This flux depends only on the shape of $\mathcal{L}$, not on its location on the sphere as the magnetic field is uniform everywhere. This decomposition is sketched in Fig. 2.

It therefore follows that we should be able to write $\mathcal{G}$ as

$$\mathcal{G}(\{\Omega\}; \{\Omega'\}) = \exp\left[\iota \sum_i \phi_{\text{AB}}(\Omega'_i, \Omega_i) + \mathcal{G}_{\text{re}}(\{\Omega\}; \{\Omega'\}) + \iota \mathcal{G}_{\text{im}}(\{\Omega\}; \{\Omega'\})\right], \quad (14)$$

where $\phi_{\text{AB}}(\Omega'_i, \Omega_i) = 2Q \arg[\bar{u}'_i u_i + \bar{v}'_i v_i]$ is the Aharonov-Bohm phase accumulated along the geodesic from Ω'_i to Ω_i , i.e., $\mathcal{P}_0$. Here, $\mathcal{G}_{\text{re}}$ is an $\text{SO}(3)$ -invariant function that is even

under inversion, while $\mathcal{G}_{\text{im}}$ is SO(3) invariant but odd under inversion. The parities under inversion follow from noting that inversion is equivalent to flipping the sign of the magnetic field in the kinetic Hamiltonian [see Eq. (2)], i.e., $Q \rightarrow -Q$. Since the wave functions must be related to those at the opposite magnetic field via complex conjugation, the parities of $\mathcal{G}_{\text{re}}$ and $\mathcal{G}_{\text{im}}$ follow. Both functions are equivariant under permutations of electron coordinates.

We now use this formal representation to motivate the backflow form used below. Let us note that the integral in Eq. (11) can be written as

$$\begin{aligned} \Psi_{\text{GS}}(\{\Omega\}) &\propto \int \prod_i d\Omega'_i \text{Re}[\mathcal{G}(\{\Omega\}; \{\Omega'\})\Psi_0(\{\Omega'\})] \\ &+ i \int \prod_i d\Omega'_i \text{Im}[\mathcal{G}(\{\Omega\}; \{\Omega'\})\Psi_0(\{\Omega'\})]. \end{aligned} \quad (15)$$

To the real and imaginary parts separately, we may then apply the mean value theorem: Each scalar integral can be represented by the corresponding integrand evaluated at a representative configuration in the integration domain, multiplied by an overall volume factor. More generally, each scalar integral may be approximated by a K -point quadrature. Since the real and imaginary parts need not share the same representative configurations, a generic complex quadrature of this type involves two independent sets of K points (for each specific point $\{\Omega\}$ in the configuration space).

In this work, we adopt the simplest approximation: a single set of backflow coordinates $\Omega_i^{\text{BF}} \equiv \Omega_i^{\text{BF}}(\{\Omega\})$ for both real and imaginary parts. This leads to the variational form

$$\Psi_{\text{GS}}(\{\Omega\}) \propto \mathcal{G}(\{\Omega\}; \{\Omega^{\text{BF}}\})\Psi_0(\{\Omega^{\text{BF}}\}). \quad (16)$$

The symmetries of the system constrain Ω_i^{BF} to be (1) permutation equivariant, (2) SO(3) equivariant, and (3) odd under inversion [i.e., fully O(3) equivariant]. With these constraints, we can replace $\mathcal{G}_{\text{re}}(\{\Omega\}; \{\Omega^{\text{BF}}(\{\Omega\})\})$ and $\mathcal{G}_{\text{im}}(\{\Omega\}; \{\Omega^{\text{BF}}(\{\Omega\})\})$ by simplified functions $\mathcal{J}_{\text{re}}(\{\Omega\})$ and $\mathcal{J}_{\text{im}}(\{\Omega\})$, which are permutation invariant and SO(3) invariant, with $\mathcal{J}_{\text{re}}$ even and $\mathcal{J}_{\text{im}}$ odd under inversion.

Before proceeding, we note that $\mathcal{J}_{\text{re}}$ faces additional constraints due to the singular Coulomb interaction in Eq. (5). Following Kato [86], if $\Psi_0 \sim d_{ij}^{\mathcal{M}} e^{i\mathcal{M}\phi_{ij}}$ as the distance between particles i and j , $d_{ij} = |\Omega_i - \Omega_j|$, approaches zero, where ϕ_{ij} is the azimuthal orientation of the tangent vector connecting Ω_i and Ω_j and $\mathcal{M}$ is an odd integer (required by fermionic statistics), then Ψ_{GS} must satisfy the cusp condition:

$$\lim_{d_{ij} \rightarrow 0} \Psi_{\text{GS}} \sim e^{i\mathcal{M}\phi_{ij}} d_{ij}^{\mathcal{M}} \left(1 + \frac{\kappa d_{ij}}{(2\mathcal{M} + 1)\ell_B} \right). \quad (17)$$

Here, we have assumed that the exponent $\mathcal{M}$ as we go from Ψ_0 to Ψ_{GS} remains unchanged. (This is true in general for any state of CFs except when CFs are restricted entirely to the lowest Λ level.) Within CF-Flow, this exponent is fixed by construction: The neural networks are analytic, so a zero of order $\mathcal{M}$ at coincidence is carried to a zero of the same order. At $\nu = 1/3$, this is a limitation, since the exact Coulomb ground state, already within the LLL, has $\mathcal{M} = 1$ rather than

the Laughlin value $\mathcal{M} = 3$ [87] (see Appendix E). In this paper, we focus on the simplest case, backflow on top of a single reference wave function. In future work, we will consider a sum of several backflow copies, or backflow on top of a linear combination of CF Slater determinants, for instance, the $\nu^* = 1$ IQH state of CFs together with CF-exciton states, which would allow the zeros to split.

Taking the cusp condition into account, we decompose $\mathcal{J}_{\text{re}}$ as $\tilde{\mathcal{J}}_{\text{re}} - \kappa/(2\mathcal{M} + 1) \sum_{i < j} \mu^2/(\mu + d_{ij})$, where $\tilde{\mathcal{J}}_{\text{re}}$ is a function with vanishing first derivative as $d_{ij} \rightarrow 0$ and $\mu > 0$ is a gating parameter that preserves $d_{ij} \rightarrow 0$ behavior and ensures decay as $d_{ij} \rightarrow \infty$. Combining all these elements, the ground-state wave function takes the form

$$\begin{aligned} \Psi_{\text{GS}}(\{\Omega\}) &= \exp \left[i \sum_i \phi_{\text{AB}}(\Omega_i^{\text{BF}}, \Omega_i) + \tilde{\mathcal{J}}_{\text{re}}(\{\Omega\}) \right. \\ &\quad \left. - \frac{\kappa}{2\mathcal{M} + 1} \sum_{i < j} \frac{\mu^2}{\mu + d_{ij}} + i \mathcal{J}_{\text{im}}(\{\Omega\}) \right] \\ &\times \Psi_0(\{\Omega^{\text{BF}}\}). \end{aligned} \quad (18)$$

This formulation provides the foundation for CF-Flow. We choose Ψ_0 to be a CF wave function [Eq. (7) or (8)], and, because evaluating $\tilde{\mathcal{J}}_{\text{re}}$, $\mathcal{J}_{\text{im}}$, μ , and Ω_i^{BF} exactly is infeasible, approximate the backflow coordinates Ω_i^{BF} , Jastrow factors $\mathcal{J}_{\text{re}}$ and $\mathcal{J}_{\text{im}}$ using symmetry-preserving neural networks whose parameters are optimized via VMC to minimize the energy. These networks constitute the learnable components of our ansatz. The explicit neural-network parametrization and a flowchart for the complete architecture are given in Appendix C.

Equation (18) admits an intuitive physical interpretation closely related to the original backflow picture of Feynman and Cohen [28]. Consider a reference Hamiltonian $\mathcal{H}_0$ for which Ψ_0 represents a state of noninteracting CFs. Now imagine adiabatically turning on interactions between CFs, specifically the additional interactions beyond those already captured by $\mathcal{H}_0$. In the semiclassical picture, these interactions induce two effects: First, they cause each CF to experience a displacement from its original position Ω_i to a new effective position Ω_i^{BF} , with this displacement depending on the configuration of all other CFs. Second, as the i th CF moves from Ω_i to Ω_i^{BF} , it accumulates an Aharonov-Bohm phase $\phi_{\text{AB}}(\Omega_i^{\text{BF}}, \Omega_i)$ due to the background magnetic field. The Jastrow factors $\tilde{\mathcal{J}}_{\text{re}}$ and $\mathcal{J}_{\text{im}}$ capture additional correlations induced by the interactions: $\tilde{\mathcal{J}}_{\text{re}}$ modifies the wave function amplitude (together with the Kato cusp term), while $\mathcal{J}_{\text{im}}$ introduces collective phases arising from the many-body dynamics.

We conclude with an important technical point concerning the imaginary terms in Eq. (18). As discussed in Sec. II A, it is important to ensure that any variational wave function is confined entirely to the physical monopole sector. For the components $\tilde{\mathcal{J}}_{\text{re}}$, $\mathcal{J}_{\text{im}}$, and Ω_i^{BF} , we design architectures that explicitly enforce the symmetry constraints discussed above: The scalar functions are permutation invariant and SO(3) invariant, with $\tilde{\mathcal{J}}_{\text{re}}$ even under inversion and $\mathcal{J}_{\text{im}}$ odd under inversion, while the backflow coordinates are permutation equivariant and O(3) equivariant

(see Appendix C for more details). Together with the Aharonov-Bohm phase factor $\exp[i \sum_i \phi_{AB}(\Omega_i^{\text{BF}}, \Omega_i)]$, these constraints ensure that the full wave function remains in the correct monopole sector and has the required transformation properties under rotations, inversion, and exchange of electron coordinates.

V. RESULTS

Having defined the CF-Flow ansatz, we now describe its optimization and results. In this section, we first describe the training protocol and hyperparameter choices used to implement CF-Flow. We then present the ground-state energies at $\nu = 1/3$ and $\nu = 2/5$ in the $L = 0$ sector (i.e., states with uniform density) as functions of the LL mixing strength κ and compare them with earlier results. Finally, we evaluate the transport gap at $\nu = 1/3$ and $\nu = 2/5$ as a function of κ .

A. Training

We optimize the variational parameters $\{\lambda_\alpha\}$ appearing in the learnable components of CF-Flow using the stochastic reconfiguration (SR) method [88,89]. Following the foundational model approach of Refs. [90,91], we embed κ into every learned component and train a single variational wave function that accepts the LL mixing strength κ as an input parameter, thereby serving as an ansatz for the entire family of Hamiltonians $\mathcal{H}(\kappa)$ defined in Eq. (5).

We evaluate the expectation values through importance sampling with respect to $|\psi|^2$, using the Metropolis-adjusted Langevin algorithm (MALA), adapted to spherical geometry [92]. We employ a staged training protocol in which learnable components are activated sequentially: (1) Kato's cusp term, (2) amplitude Jastrow, (3) phase Jastrow, (4) backflow coordinates, and (5) fine-tuning with reduced learning rate. Each stage initializes from the best parameters of the previous one (determined by the lowest average energy across κ during training). Figure 3 shows a representative training curve. The order reflects the increasing complexity of the modules; we have verified that alternative orderings yield equivalent final results. Further details regarding training are provided in Appendix D.

This staged approach reveals the relative importance of each component to the final energy. The Kato cusp term, which decays at large interparticle separations d_{ij} and connects smoothly to LLL wave functions in the $\kappa \rightarrow 0$ limit, provides modest improvement, with larger relative gains at smaller N . The amplitude Jastrow and backflow coordinates yield the largest energy gains: The amplitude Jastrow alone should, in principle, reproduce fp-DMC results (which constrains only the phase), while backflow also enables optimization of the phase of the wave function [34]. In contrast, the phase Jastrow term provides negligible improvement.

B. Ground-state energies

Figure 4 shows the per-particle energies E/N obtained from CF-Flow at $\nu = 1/3$ and $\nu = 2/5$ starting from the uniform wave functions of Eq. (7), i.e., the $\nu^* = 1$ and $\nu^* = 2$ IQH states of CFs. The results correspond to the parameters yielding the lowest average energy across all sampled κ

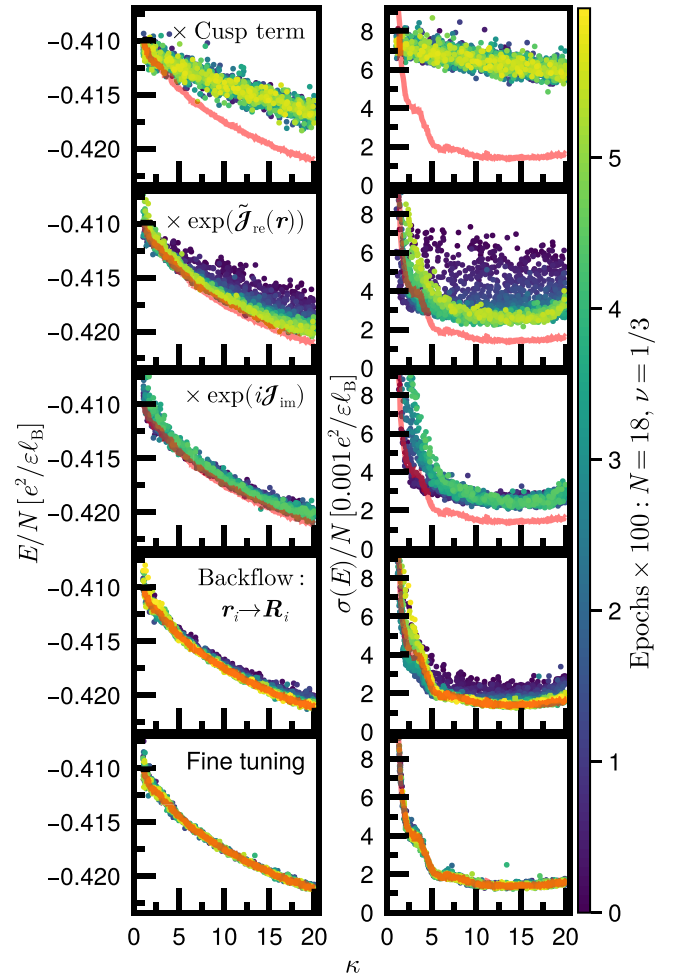


FIG. 3. Per-particle energy E/N (left) and local-energy standard deviation $\sigma(E)/N$ (right) as functions of the LL mixing strength κ during optimization of CF-Flow for $N = 18$ electrons at $\nu = 1/3$. Color indicates the training epoch within each stage. CF-Flow begins from Laughlin's wave function (the uniform ($L = M = 0$) state of CFs filling the $n = 0$ Λ level [see Eq. (7)]) and progressively learns backflow corrections through neural networks (see Fig. 8). Training proceeds in five stages: (1) enabling Kato's cusp term, (2) adding the amplitude Jastrow factor $\exp[\tilde{\mathcal{J}}_{\text{re}}(\mathbf{r})]$, (3) adding the phase Jastrow factor $\exp[i\tilde{\mathcal{J}}_{\text{im}}(\mathbf{r})]$, (4) enabling backflow, and (5) fine-tuning with a reduced learning rate. The red curve marks the best model (lowest mean energy across κ over all training runs). The total energy E includes contributions from the uniform background charge and is corrected for the "density shift" inherent to spherical geometry [93].

values. In all cases, the average (over κ) per-particle local-energy standard deviation lies in the range 0.002–0.004 $e^2/\varepsilon\ell_B$ (see, e.g., Fig. 3). The small local-energy standard deviation $\sigma(E)/N$ indicates that CF-Flow produces states very close to the true ground state. In the limit $\kappa \rightarrow 0$, this deviation grows because even a small component of the wave function outside the LLL leads to a large energy change.

For comparison with fp-DMC and DeepHall [5], which are evaluated at fixed κ , we interpolate the CF-Flow energies over $\kappa \in [1, 20]$ using Gaussian process regression [94] (allowing us to interpolate without enforcing an explicit functional form and to capture somewhat nonsmooth behavior).

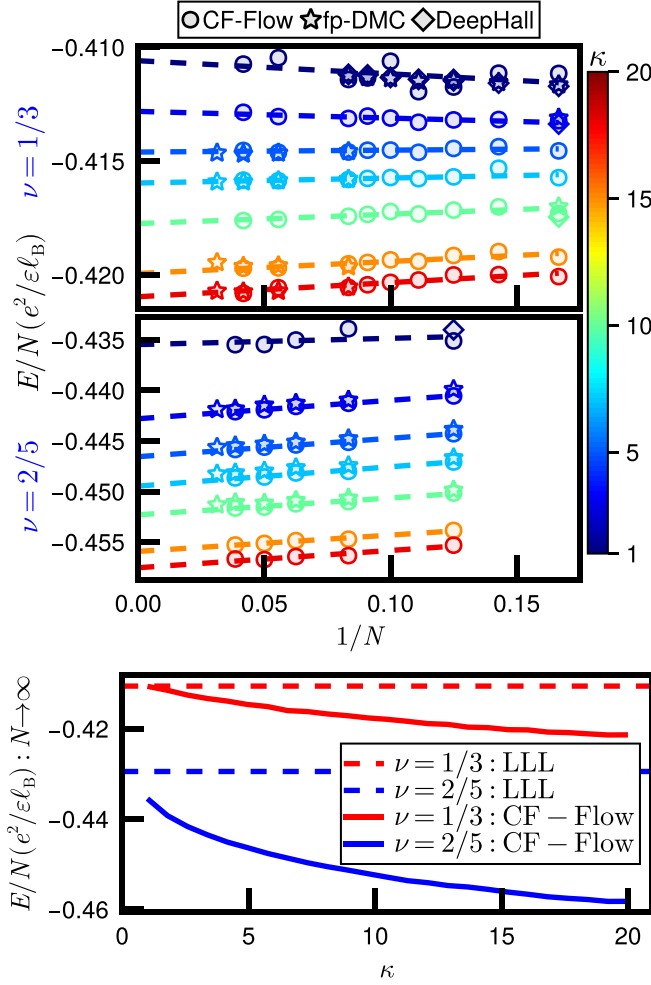


FIG. 4. Per-particle energy E/N for uniform (i.e., FQH ground) states at fillings $\nu = 1/3$ and $\nu = 2/5$ as functions of $1/N$ and the LL mixing strength κ . Top: CF-Flow results compared with DeepHall [5], a recent Psiformer-based neural quantum state approach, and fp-DMC [5,25], which constrains the wave function phase to match Eq. (7) and provides the ground-state energy within this phase sector. Dashed lines denote linear fits in $1/N$ at fixed κ . Bottom: Thermodynamic-limit energies E/N ($N \rightarrow \infty$) from CF-Flow obtained via linear $1/N$ extrapolation, compared with LLL ($\kappa = 0$) values from CF theory [13]. Error bars are smaller than the marker size.

As shown, CF-Flow yields per-particle energies comparable to both DeepHall and fp-DMC at $\nu = 1/3$, and slightly lower (better) energies at $\nu = 2/5$.

Finally, we comment on the role of the initial wave function Ψ_0 by comparing CF-Flow starting from the unprojected (i.e., without LLL projection) and projected CF wave functions at $\nu = 2/5$ [see Eq. (7)]. For $N = 16$ (Fig. 5), the unprojected wave function, despite lying in the same topological phase and symmetry sector (the $L = 0$ angular-momentum sector), gives a higher energy after optimizing the backflow deformations in the small- κ regime, where the initial kinetic-energy cost of the unprojected wave function is large. Thus, the choice of initial wave function can significantly affect the energy reached by CF-Flow. We leave it for future work to determine whether a linear combination of multiple unprojected wave functions,

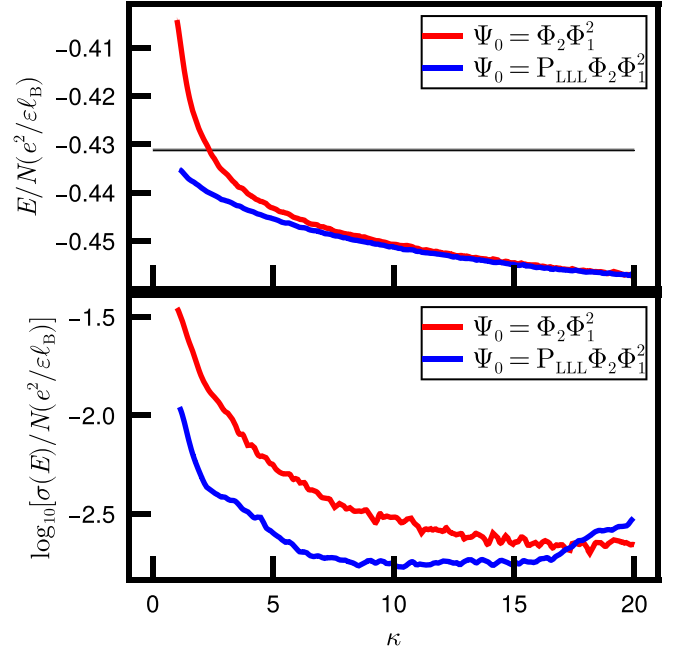


FIG. 5. Top: per-particle energy E/N and bottom: per-particle local-energy standard deviation $\sigma(E)/N$ as functions of LL mixing strength κ for $N = 16$ electrons at filling $\nu = 2/5$, obtained using CF-Flow initialized from the unprojected Jain wave function (red curve) and the projected Jain wave function (blue curve) [see Eq. (7)]. The horizontal line marks the energy of the wave function from Eq. (7). In the small- κ regime, the kinetic-energy cost of the unprojected wave function makes it a worse starting point than the projected wave function: With the same hyperparameters, CF-Flow reaches a higher energy after training.

each with independent backflow deformations, can reach energies similar to those obtained from the projected wave function (see Sec. IV). At the level of the formal quadrature argument, at least two independent copies would be needed, and more may be required in practice.

C. Transport gap

To evaluate the transport gap Δ , we initialize CF-Flow from the largest- L CF exciton wave function [Eq. (8)], corresponding to $L = N$ for a system of N electrons at $\nu = 1/3$ and $L = N/2 + 1$ at $\nu = 2/5$. This state represents a maximally separated quasihole-quasiparticle pair. The symmetry-preserving design of CF-Flow enables such calculations: The ansatz is, by construction, restricted to a single total-angular-momentum sector. The results are shown in Fig. 6. The Coulomb interaction between the quasihole and quasiparticle, treated as point charges, is subtracted from this energy difference. These transport-gap calculations are more demanding than the ground-state calculations because the gap is extracted from the difference of separately optimized extensive energies. For the largest systems, residual optimization uncertainty is therefore larger than for the ground-state benchmarks.

Consequently, the thermodynamic extrapolations should be interpreted with caution; they nevertheless provide useful insight into the trend in the thermodynamic limit and are shown in Fig. 6, bottom panel. We find that the extrapolated gap $\Delta(\kappa)$

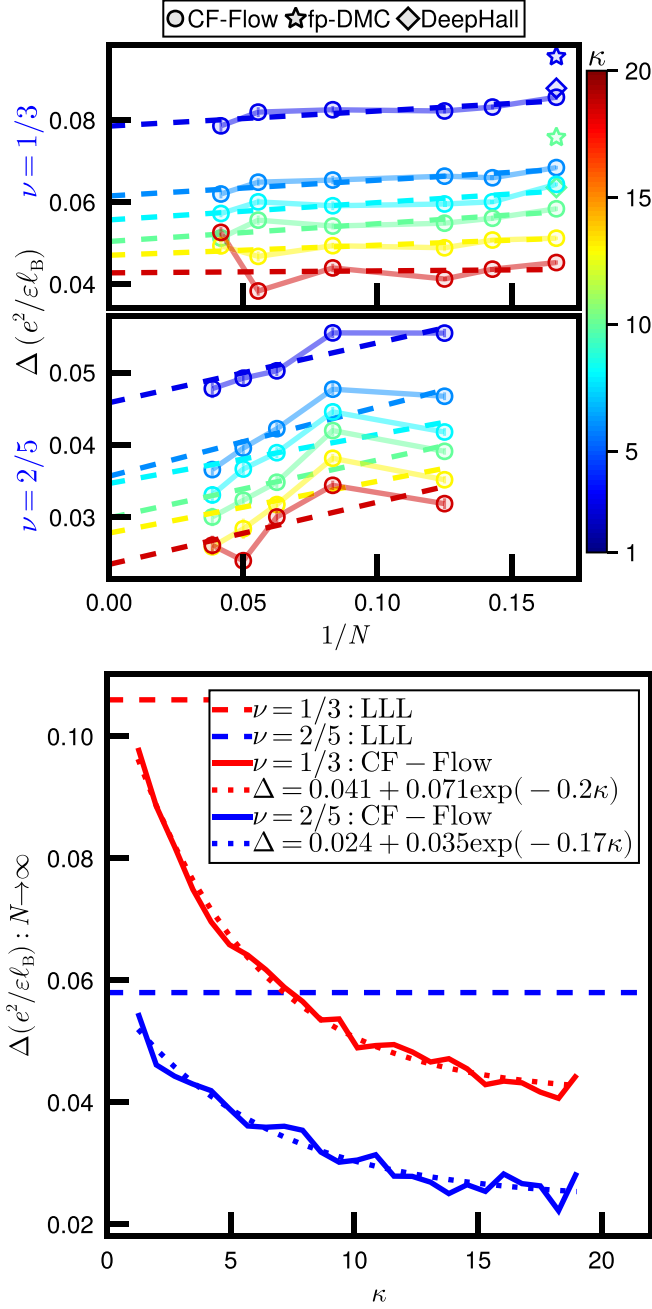


FIG. 6. Top: Transport gap Δ at $\nu = 1/3$ and $\nu = 2/5$ as a function of κ for various system sizes N . The symmetry-preserving architecture of CF-Flow enables construction of variational ansätze for ground states or lowest-energy states within each angular-momentum sector. We compute Δ as the energy difference between the uniform $L = 0$ ground state and the largest- L CF exciton state (corresponding to a maximally separated quasihole-quasiparticle pair in the LLL limit, with interactions between separated particles treated as point charges). CF-Flow results are compared with DeepHall [5], fp-DMC [25], and exact diagonalization in the LLL where available. Bottom: Thermodynamic-limit gap Δ obtained via $1/N$ extrapolation. The κ dependence is well described by exponential fits, with Δ decaying to a finite value in the $\kappa \rightarrow \infty$ limit. [See Eq. (19).]

is well described by exponential fits,

$$\begin{aligned} \nu = \frac{1}{3} : \quad \Delta(\kappa) &= [0.041 + 0.071 \exp(-0.2\kappa)] \frac{e^2}{\varepsilon \ell_B}, \\ \nu = \frac{2}{5} : \quad \Delta(\kappa) &= [0.024 + 0.035 \exp(-0.17\kappa)] \frac{e^2}{\varepsilon \ell_B}. \end{aligned} \quad (19)$$

Importantly, these fits suggest that as $\kappa \rightarrow \infty$, the transport gap remains finite, approaching $\sim 0.041 e^2/\varepsilon \ell_B$ at $\nu = 1/3$ and $\sim 0.024 e^2/\varepsilon \ell_B$ at $\nu = 2/5$. We would like to note that the fits are based on the data in the range $1 \leq \kappa \leq 20$ studied here; the $\kappa \rightarrow \infty$ values are extrapolations that presume the system remains in the FQH phase. We therefore conclude that, within the FQH phases at $\nu = 1/3$ and $\nu = 2/5$, the transport gap does not close as a function of LL mixing. This does not rule out an LL-mixing-induced transition through another mechanism, for example, through the softening of a CF exciton of finite momentum. We leave a comprehensive study of LL-mixing-induced transitions out of the FQH phase to future work.

VI. DISCUSSION AND CONCLUSION

Developing a nonperturbative and scalable method to study LL mixing has been a long-standing challenge in the theory of the FQHE. The LLL limit is central to much of our current understanding, most notably through CF theory, but real experimental systems typically operate at $\kappa \sim 0.50$ – 10 [36]. In this nonperturbative regime, LL mixing can drive phase transitions and substantially modify quantitative properties such as the transport gap.

We have introduced CF-Flow, a variational framework that combines CF wave functions with backflow transformations implemented through symmetry-preserving neural networks. CF-Flow naturally allows us to address LL-mixing physics and, more generally, any regime where interactions between CFs become significant. By starting from CF wave functions, which already encode strong electron correlations, and learning backflow corrections through neural networks, CF-Flow achieves accuracy comparable to or better than recent NQS approaches [5] while using an order of magnitude fewer parameters (5×10^4 versus 10^6) and converging faster (2×10^3 epochs versus 2×10^5): Each calculation for a given system size completes in under one day on a single NVIDIA A100 GPU. This efficiency enables exploration over a substantially larger parameter space in system size, filling factor, and κ . The symmetry-preserving architecture also gives direct access to excited states by initializing from the appropriate CF wave function in any chosen angular-momentum sector. The central idea is that, instead of learning from scratch how electrons reorganize into dressed CFs, we start from CFs and learn only their dressing.

Beyond its computational advantages, CF-Flow yields insight into the structure of many-body ground states at Jain fillings. At $\nu = 1/3$ and $\nu = 2/5$, the close agreement between CF-Flow and fp-DMC suggests that the phase of the IQH wave function of CFs [Eq. (7)], which describes the LLL

limit ($\kappa \rightarrow 0$), remains close to that of the true ground state even up to $\kappa \sim 20$. This provides an independent validation of fp-DMC when constrained to this phase and helps explain its quantitative success in predicting experimental properties. However, this phase robustness is not universal: The difference between CF-Flow and fp-DMC transport gaps shows that the phase of CF exciton wave functions is more sensitive to corrections due to LL mixing. Making the notion of dressing quantitative requires a regime in which CF wave functions are still a good basis: For a perturbation close to the LLL limit, a finite quantum-well width, for instance, one may expand the optimized CF-Flow state in the CF basis and ask how far it has moved from the state of noninteracting CFs, and how many CF exciton states are needed to reproduce it. We leave such an analysis, which would make the notion of CF dressing more concrete, to future work.

The dependence of the optimized energy on the starting point, projected versus unprojected at $\nu = 2/5$ (Fig. 5), raises the question of how much the initial wave function matters, especially when the target phase differs from the starting phase. In future work, it will be important to study the origin of this discrepancy systematically, and to see how the results change when CF-Flow is initialized from non-CF starting points such as Wigner-crystal wave functions at large κ , where neural-network approaches have recently reported crystallization [95]. We also point out that we have made several simplifying assumptions in this work, namely, simple multilayer perceptrons (MLPs) for the backflow and the Jastrow factors and, most importantly, a single Slater determinant ansatz (see Sec. IV); relaxing these is the next step.

Several directions warrant further exploration.

(1) In this work, we have focused on two angular-momentum sectors: the $L = 0$ sector, whose lowest-energy state corresponds in the LLL limit to the uniform FQH liquid, and the sector containing the largest- L CF exciton, corresponding in the LLL limit to a maximally separated quasihole-quasiparticle pair. Among the sectors studied here, the lowest-energy state remains in the $L = 0$ sector. A more comprehensive search for competing phases would require optimizing CF-Flow in additional angular-momentum sectors, which we leave for future work. At sufficiently large κ , for example, the system may transition to a symmetry-broken state such as a charge-density wave or Wigner crystal. The situation is subtle: Such phases may appear as ground states in higher-angular-momentum sectors if rotational symmetry is spontaneously broken, or they may appear in the $L = 0$ sector as symmetry-averaged states, i.e., linear superpositions of different symmetry-broken configurations. Figure 7 shows the pair-correlation function $g(r)$ for $N = 24$ electrons at $\nu = 1/3$ and $N = 26$ electrons at $\nu = 2/5$. With increasing LL mixing, the state develops more crystalline correlations. However, from the system sizes studied in this work, we cannot determine whether this reflects a transition into a crystalline phase in the thermodynamic limit. The peaks in $g(r)$ decay with separation, with the nature of the decay indicating whether or not such a transition has taken place. For instance, exponential decay would indicate a finite correlation length, consistent with a liquid phase, whereas sufficiently slow decay would provide evidence for crystalline order. The limited range of interparticle separations accessible in the

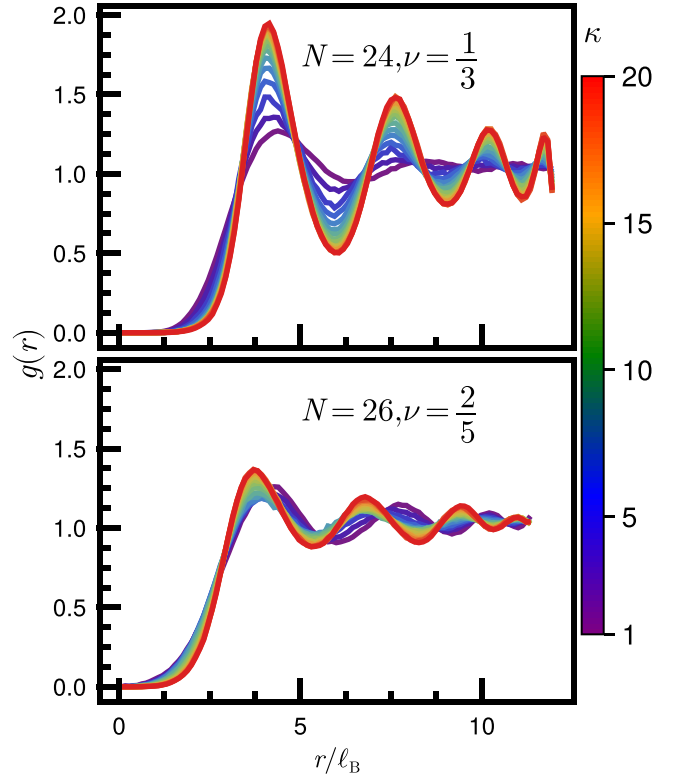


FIG. 7. Pair-correlation function $g(r)$ as a function of the electron separation r for FQH liquids at $\nu = 1/3$ (top, $N = 24$) and $\nu = 2/5$ (bottom, $N = 26$), obtained from CF-Flow. The color indicates the LL mixing strength κ . With increasing LL mixing, the states develop enhanced crystalline correlations. For the finite systems shown, however, the accessible range of separations is insufficient to determine whether these correlations indicate long-range crystalline order in the thermodynamic limit.

largest systems studied here makes it difficult to determine the nature of this decay. In future work, it would be important to study larger systems and additional angular-momentum sectors not considered here. Other properties of FQH liquids, such as quasiparticle and quasihole density profiles, the static structure factor, and the entanglement spectrum [96], may also provide sharper diagnostics of how FQH phases evolve as a function of LL mixing. We leave the study of these properties to future work as well.

(2) Real experimental systems contain disorder, which can strongly modify transport properties and excitation gaps. Extending CF-Flow to incorporate disorder, which the foundational approach can accommodate naturally [90,91], and understanding how disorder and LL mixing together shape observed transport gaps represent an important future direction.

(3) Phases in the second LL, such as the $\nu = 5/2$ FQH state, where LL mixing breaks particle-hole symmetry and controls the competition between the Pfaffian and anti-Pfaffian states, are at present out of reach. Applying CF-Flow there requires a reference state in the second LL. Constructing one in real space is a hard task: One must take the analytic expression for the corresponding LLL state, promote each electron's single-particle state to the second LL, and antisymmetrize the result against the filled LLL. The combinatorial

cost of this antisymmetrization is severe and presumably limits such a construction to on the order of 10 electrons. There is at present no viable route to starting CF-Flow directly from a CF state in the second LL. The backflow construction, however, is not tied to CF reference states: Any Ψ_0 that can be evaluated together with its derivatives, a crystal wave function at large κ , for instance, can serve as the starting point, while the ansatz remains strictly within the physical monopole sector by construction (Appendix A).

(4) The recent observation of the fractional quantum anomalous Hall effect, FQHE in the absence of a magnetic field in Chern bands [97–102], is an exciting area for applying CF-Flow. The prominent FQH states again occur at the Jain fillings $\nu = n/(2n \pm 1)$ and follow a qualitative trend similar to LLs, with smaller n producing more robust states (as at these fillings CFs feel a larger effective field). However, it is *a priori* unclear whether CFs remain weakly interacting in these systems [103]. Extending CF-Flow to this setting, and understanding whether and how CFs are dressed in Chern bands, would be an exciting future direction.

ACKNOWLEDGMENTS

M.G. thanks J. K. Jain for many thoughtful discussions throughout the course of this project, for valuable comments on the manuscript, and for help in writing the introduction. M.G. also thanks D. Jin, P. N. Chakravartula, and G. J. Sreejith for insightful discussions and suggestions. M.G. acknowledges support from the National Science Foundation under Grant No. DMR-2404619, awarded to J. K. Jain. M.G. was partially supported by a 2025/2026 Rising Researcher Grant from Penn State’s Institute for Computational and Data Sciences (RRID: SCR_025154). Computational support was provided by the Penn State Institute for Computational and Data Sciences Roar Core Facility (RRID: SCR_026424). M.G. thanks the Lodha Theoretical Physics Institute, Mumbai, for its hospitality during part of this work. Plots were made using `Makie.jl` [104]. All variational Monte Carlo computations in this work were implemented using JAX [105]. During the course of this project, Anthropic’s Claude models (Fable and Opus) and OpenAI’s Codex were used to help understand JAX [105], to optimize code for GPUs, and to orchestrate the project workflow. They were also used to review drafts of the manuscript and to help write the plotting scripts for the figures, under the direction of M.G.

DATA AVAILABILITY

The data that support the findings of this article are openly available [106]. The code for CF-Flow is available from the author upon request.

APPENDIX A: MONOPOLE-SECTOR CONSTRAINT IN SPHERICAL GEOMETRY

In this Appendix, we expand on the discussion in Sec. II A: Wave functions that mix monopole harmonics from a monopole sector $Q' \neq Q$ generically acquire infinite kinetic energy when evaluated with the kinetic operator at monopole

strength Q [Eqs. (2) and (3)]. The infinite kinetic energy due to monopole mismatch is a direct consequence of the singular nature of the vector potential $\mathbf{A}$ at the poles, i.e., the presence of Dirac strings.

The monopole harmonics can be expressed as sections of the Wigner D matrices [72,76,107,108]:

$$Y_{Q,l,m}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} D_{-m,-Q}^l(\phi, \theta, 0),$$

$$D_{m_1,m_2}^l(\alpha, \beta, \gamma) = \langle l, m_1 | e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z} | l, m_2 \rangle, \quad (\text{A1})$$

where J_x , J_y , and J_z are the generators of rotations satisfying $[J_i, J_j] = i\epsilon_{ijk}J_k$, and $|l, m\rangle$ are eigenstates with angular momentum l and azimuthal quantum number $m = -l, \dots, l$ [84].

To see the constraint explicitly, consider evaluating the kinetic energy of $Y_{Q',l',m'}(\theta, \phi)$ with respect to the kinetic operator at monopole strength Q [Eq. (2)]. For clarity, we adjust the sphere radius such that electrons always experience the same magnetic field strength B and use primes to denote operators and wave functions at monopole strength Q' . From Eqs. (2) and (3), we obtain

$$\mathcal{K}Y_{Q',l',m'} = \frac{\hbar\omega_c}{2Q} \left[-\nabla^2 + Q^2 \cot^2 \theta - \frac{2Qm' \cot \theta}{\sin \theta} \right] Y_{Q',l',m'}.$$

Since $\mathcal{K}'Y_{Q',l',m'} = \frac{\hbar\omega_c[l'(l'+1)-Q'^2]}{2Q'} Y_{Q',l',m'}$, where

$$\mathcal{K}' = \frac{\hbar\omega_c}{2Q'} \left[-\nabla^2 + Q'^2 \cot^2 \theta - \frac{2Q'm' \cot \theta}{\sin \theta} \right],$$

we can rearrange to find

$$\begin{aligned} \mathcal{K}Y_{Q',l',m'} &= \frac{\hbar\omega_c[l'(l'+1)-Q'^2]}{2Q} Y_{Q',l',m'} \\ &\quad + \frac{\hbar\omega_c(Q-Q') \cos \theta}{2Q \sin^2 \theta} \\ &\quad \times [(Q+Q') \cos \theta - 2m'] Y_{Q',l',m'}. \end{aligned}$$

Using the relation in Eq. (A1), the expectation value of the kinetic energy becomes

$$\begin{aligned} \langle Y_{Q',l',m'} | \mathcal{K} | Y_{Q',l',m'} \rangle &= \frac{\hbar\omega_c[l'(l'+1)-Q'^2]}{2Q} + \frac{\pi \hbar\omega_c(Q-Q')}{Q} \\ &\quad \times \int_0^\pi d\theta \frac{\cos \theta}{\sin \theta} [(Q+Q') \cos \theta - 2m'] [d_{-m',-Q'}^{l'}(\theta)]^2. \end{aligned}$$

When $|m'| = |Q'|$ and $Q' \neq Q$, the integral diverges due to the $\cot \theta$ singularity at one of the poles. Since $Y_{Q',l',\pm Q'}$ (proportional to $d_{-m',-Q'}^{l'}(\theta)$) are the only monopole harmonics nonzero at the poles, any function $f(\theta, \phi)$ with support at the poles generically exhibits divergent kinetic energy unless it lies entirely within the correct monopole sector Q .

APPENDIX B: JAIN-KAMILLA PROJECTION

In this Appendix, we briefly present the JK approach to LLL projection of composite-fermion wave functions [75] and

show how the JK-projected wave functions are built from multiparticle orbitals that resemble those used in Psiformer-like ansätze [5,9].

The key insight behind the JK projection approach is to restructure the projection to avoid a many-body calculation (which scales as N^N). We begin by decomposing the Jastrow factor as [see Eq. (7)]

$$\prod_{i<j=1}^N (u_i v_j - u_j v_i)^{2p} = \prod_{i<j=1}^N (u_i v_j - u_j v_i)^{2p-2} \prod_{i=1}^N \underbrace{\left(\prod_{j \neq i} (u_i v_j - u_j v_i) \right)}_{J_i}.$$

Absorbing the “individual” Jastrow factors J_i into each column of the Slater determinant, we can rearrange the CF wave function from Eq. (7) as

$$\Psi_{\frac{n}{2pn+1}}(\Omega_1, \dots, \Omega_N) = P_{\text{LLL}} \begin{vmatrix} Y_{Q^*, Q^*, -Q^*}(\Omega_1) J_1 & \dots & Y_{Q^*, Q^*, -Q^*}(\Omega_N) J_N \\ \vdots & \vdots & \vdots \\ Y_{Q^*, Q^*, Q^*}(\Omega_1) J_1 & \dots & Y_{Q^*, Q^*, Q^*}(\Omega_N) J_N \\ Y_{Q^*, Q^*+1, -(Q^*+1)}(\Omega_1) J_1 & \dots & Y_{Q^*, Q^*+1, -(Q^*+1)}(\Omega_N) J_N \\ \vdots & \vdots & \vdots \\ Y_{Q^*, Q^*+n-1, (Q^*+n-1)}(\Omega_1) J_1 & \dots & Y_{Q^*, Q^*+n-1, (Q^*+n-1)}(\Omega_N) J_N \end{vmatrix} \phi_1^{2p-2}. \quad (\text{B1})$$

This rearrangement has a crucial property: In the i th column, only the state of the i th electron lies outside the LLL; all LLL wave functions are polynomials in only u_i, v_i and not their complex conjugates [Eq. (4)]. Projecting a single electron’s state into the LLL is exponentially simpler than projecting the full many-body state.

Jain and Kamilla therefore proposed moving the projection inside the determinant, projecting each column independently:

$$\Psi_{\frac{n}{2pn+1}}(\Omega_1, \dots, \Omega_N) = \underbrace{\begin{vmatrix} P_{\text{LLL}} Y_{Q^*, Q^*, -Q^*}(\Omega_1) J_1 & \dots & P_{\text{LLL}} Y_{Q^*, Q^*, -Q^*}(\Omega_N) J_N \\ \vdots & \vdots & \vdots \\ P_{\text{LLL}} Y_{Q^*, Q^*, Q^*}(\Omega_1) J_1 & \dots & P_{\text{LLL}} Y_{Q^*, Q^*, Q^*}(\Omega_N) J_N \\ P_{\text{LLL}} Y_{Q^*, Q^*+1, -(Q^*+1)}(\Omega_1) J_1 & \dots & P_{\text{LLL}} Y_{Q^*, Q^*+1, -(Q^*+1)}(\Omega_N) J_N \\ \vdots & \vdots & \vdots \\ P_{\text{LLL}} Y_{Q^*, Q^*+n-1, (Q^*+n-1)}(\Omega_1) J_1 & \dots & P_{\text{LLL}} Y_{Q^*, Q^*+n-1, (Q^*+n-1)}(\Omega_N) J_N \end{vmatrix}}_{\phi_n^{\text{JK}}(\Omega_1, \dots, \Omega_N; Q^* + (N-1))} \phi_1^{2p-2}. \quad (\text{B2})$$

This form ensures the wave function remains entirely within the LLL: By Eq. (4), the product of any two LLL wave functions yields another LLL wave function (albeit at different monopole strength). Remarkably, multiple studies have shown that the JK wave function is nearly indistinguishable from Eq. (7), despite being formally a different trial wave function class [13,75,109]. The determinant ϕ_n^{JK} is often called the CF Slater determinant, with its elements termed CF orbitals.

While we do not discuss the detailed implementation of JK projection, it is instructive to consider the form of a typical JK-projected wave function in planar geometry [75] (after some rearrangement),

$$\Psi(z_1, \dots, z_N) = \underbrace{\begin{vmatrix} 1 & \dots & 1 \\ z_1 & \dots & z_N \\ \vdots & \vdots & \vdots \\ e_1\left(\left\{\frac{1}{z_1 - z_j}\right\}_{j \neq 1}\right) & \dots & e_1\left(\left\{\frac{1}{z_N - z_j}\right\}_{j \neq N}\right) \\ \vdots & \vdots & \vdots \\ z_1^{m_a+n_a} e_{n_a}\left(\left\{\frac{1}{z_1 - z_j}\right\}_{j \neq 1}\right) & \dots & z_N^{m_a+n_a} e_{n_a}\left(\left\{\frac{1}{z_N - z_j}\right\}_{j \neq N}\right) \\ \vdots & \vdots & \vdots \end{vmatrix}}_{\phi_{\text{JK}}(z_1, \dots, z_N)} \prod_{i<j=1}^N (z_i - z_j)^{2p} \exp\left[-\sum_{i=1}^N \frac{|z_i|^2}{4\ell_B^2}\right], \quad (\text{B3})$$

where $z_i = x_i + iy_i$ are the complex coordinates corresponding to the i th electron, e_k are the elementary symmetric polynomials, and (n_a, m_a) are the quantum numbers corresponding to the occupied Λ levels.

Structurally, as seen in Eq. (B3), the JK-projected CF wave functions (introduced in 1997) bear an interesting resemblance to Psiformer-like ansätze [5,9] (introduced in 2023, more than a quarter century later): In both cases, the total wave

function is written as a determinant of coordinate-dependent multiparticle orbitals $\phi_i(z_i; \{z_{j \neq i}\})$ multiplied by a Jastrow-type correlation factor.

APPENDIX C: NEURAL-NETWORK IMPLEMENTATION OF CF-FLOW

This Appendix describes the neural-network architecture used in CF-Flow. We begin with a brief overview of MLPs and the DeepSets architecture [29], which form the basic building blocks of CF-Flow. We then discuss results on O(3)-equivariant functions that are useful for constructing our symmetry-preserving ansatz, before presenting the full CF-Flow architecture.

1. Multilayer perceptrons and DeepSets

MLPs are universal function approximators [7] inspired by neurons and widely used in ML. They consist of a sequence of perceptron layers defined as follows.

A perceptron layer maps an input vector $\mathbf{x} \in \mathbb{R}^{d_{\text{in}}}$ to an output $\mathbf{y} \in \mathbb{R}^{d_{\text{out}}}$ via

$$\mathbf{y} = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b}), \quad (\text{C1})$$

where $\mathbf{W} \in \mathbb{R}^{d_{\text{out}} \times d_{\text{in}}}$ is the weight matrix, $\mathbf{b} \in \mathbb{R}^{d_{\text{out}}}$ is the bias vector, and σ is a nonlinear activation function applied elementwise. The linear transformation $\mathbf{W}\mathbf{x} + \mathbf{b}$ projects the input into a (typically higher-dimensional) feature space, while the nonlinearity σ enables the network to capture interactions between features.

An MLP is formed by composing multiple such layers,

$$\mathbf{y}^{(l+1)} = \sigma_l(\mathbf{W}_l \mathbf{y}^{(l)} + \mathbf{b}_l), \quad (\text{C2})$$

with parameters $\{\mathbf{W}_l, \mathbf{b}_l\}$ optimized during training. A sufficiently expressive MLP can approximate any continuous function on a compact domain, a property we will exploit in constructing our ansatz.

However, an MLP's ability to learn a given function depends crucially on the training data. For instance, if the target function is invariant under exchange of its inputs but the data do not exhibit this symmetry, the trained MLP will likewise fail to respect it. In the context of VMC optimization [8], where only a finite number of samples can be drawn within a finite simulation time, this becomes a significant limitation: Unless symmetries are explicitly enforced, the resulting variational ansatz will not preserve them.

To overcome this, we employ architectures that incorporate the relevant symmetries by construction: the DeepSets framework, which uses MLPs to represent functions over sets (thereby ensuring invariance under permutations of particle indices), and the O(3)-equivariant models introduced in Ref. [31], which we adapt to our system of indistinguishable electrons (for a broader overview of symmetry-preserving neural-network design, see Refs. [30,32]).

The DeepSets framework [29] allows for the approximation of functions invariant under the exchange of inputs. Specifically, consider a function f that satisfies $f(\mathbf{x}_1, \dots, \mathbf{x}_N) = f(\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(N)})$ for any permutation π of

its inputs. Then, f can always be written as

$$f(\mathbf{x}_1, \dots, \mathbf{x}_N) = g\left(\sum_i \mathbf{h}(\mathbf{x}_i)\right), \quad (\text{C3})$$

where $\mathbf{h}$ and g are suitable functions and the sum runs over the input index i (not over the components of $\mathbf{h}$). In the DeepSets framework, both $\mathbf{h}$ and g are represented using MLPs [see Eq. (C2)]. As long as these MLPs are sufficiently wide, this construction can approximate any function f that is invariant under permutations of its inputs [29,110]. We will make use of this property below.

2. O(3)-equivariant functions

CF-Flow provides an ML approximation to the backflow representation of the ground state Ψ_{GS} in Eq. (18). Because evaluating $\tilde{\mathcal{J}}_{\text{re}}$, $\mathcal{J}_{\text{im}}$, μ , and Ω^{BF} exactly is infeasible, we approximate each of them using neural networks whose parameters are optimized via VMC to minimize the energy. These networks constitute the learnable components of our ansatz.

The gating parameter μ is represented by an MLP that maps the physical inputs κ , N , and Q to a positive scalar (see below for details). For the remaining components ($\tilde{\mathcal{J}}_{\text{re}}$, $\mathcal{J}_{\text{im}}$, and Ω^{BF}), we design architectures that explicitly enforce the symmetry constraints discussed in Sec. IV.

To this end, we invoke the theory of invariants of Lie groups [111], following Ref. [31] and adapting it to the case of indistinguishable electrons. The key results [31,111] can be summarized as follows:

Theorem C1. If $\mathbf{v}_1, \dots, \mathbf{v}_N \in \mathbb{R}^3$, any SO(3)-invariant scalar that is even under inversion [i.e., O(3) invariant] can be expressed as a function of dot products $\mathbf{v}_i \cdot \mathbf{v}_j$.

Theorem C2. If $\mathbf{v}_1, \dots, \mathbf{v}_N \in \mathbb{R}^3$, any SO(3)-invariant scalar that is odd under inversion can be written as a sum over scalar triple products $\chi_{ijk} = \mathbf{v}_i \cdot (\mathbf{v}_j \times \mathbf{v}_k)$, multiplied by O(3)-invariant functions of dot products $\mathbf{v}_a \cdot \mathbf{v}_b$.

Theorem C3. If $\mathbf{v}_1, \dots, \mathbf{v}_N \in \mathbb{R}^3$, any O(3)-equivariant vector-valued function can be expressed as a sum over $\mathbf{v}_i$ with O(3)-invariant coefficients built from functions of the dot products $\mathbf{v}_a \cdot \mathbf{v}_b$.

Combined with the DeepSets framework [29], which can approximate any permutation-invariant function using sufficiently expressive MLPs, these theorems provide the foundation for constructing symmetry-preserving networks for each CF-Flow module: $\tilde{\mathcal{J}}_{\text{re}}$ (amplitude Jastrow), $\mathcal{J}_{\text{im}}$ (phase Jastrow), and Ω^{BF} (backflow).

3. Architecture of CF-Flow

In designing CF-Flow, we embed the physical parameters κ , N , and Q into every learned component, following the “foundational” approach introduced in Refs. [90,91]. This enables a single variational ansatz for the ground state within the same angular-momentum sector as the reference CF wave function to describe systems across a continuous range of LL mixing strengths [we have focused on $\kappa \in (1, 20)$] and system sizes N . Embedding N and Q explicitly also opens the possibility of “transfer learning”: training at smaller system sizes, where optimization is computationally cheaper, and using the

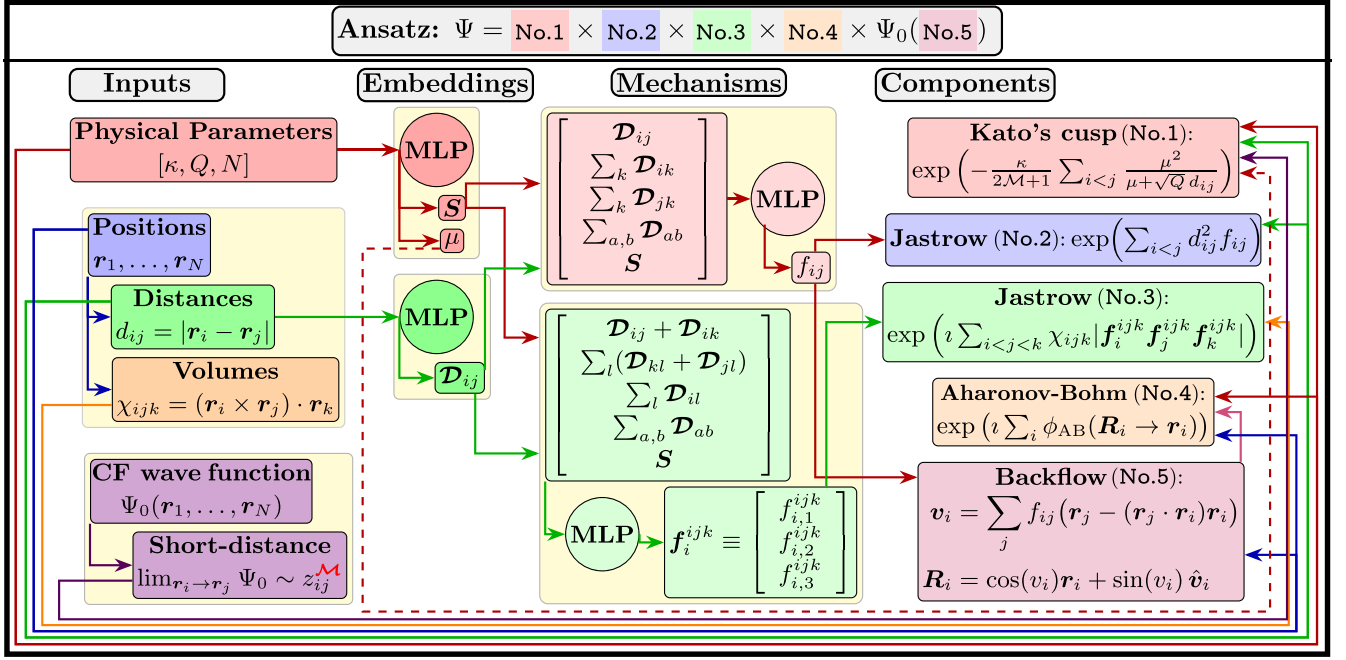


FIG. 8. Schematic illustration of the CF-Flow architecture. The ansatz (top) consists of a CF wave function Ψ_0 evaluated at backflow coordinates $\mathbf{R}_i$ and multiplied by four physically motivated factors. The model takes as input the physical parameters κ , Q , and N (LL mixing strength, monopole strength, and particle number) and constructs geometric features from electron coordinates $\mathbf{r}_1, \dots, \mathbf{r}_N$, pairwise distances d_{ij} , and signed volumes χ_{ijk} . Using the DeepSets architecture [29], MLPs learn O(3)-invariant scalars parametrizing each component: (1) “two-body” scalars f_{ij} (many-body functions with pair indices) for an amplitude-modulating Jastrow factor (No. 2) and backflow coordinates (No. 5), and (2) “three-body” scalars f_a^{ijk} ($a = i, j, k$; many-body functions with triplet indices) for a phase-modulating Jastrow factor (No. 3). The architecture also includes physically enforced terms: Kato’s cusp (No. 1) with short-distance coefficient set by $\mathcal{M}$ (the relative angular momentum at small d_{ij} in Ψ_0) and κ , gated at large distances; and the Aharonov-Bohm phase (No. 4) enforcing our gauge choice (in spinor representation, $\phi_{AB}(U, V; u, v) = 2Q \arg[\bar{U}u + \bar{V}v]$). By construction, the combined Jastrow term is SO(3) invariant and transforms to its complex conjugate under inversion, while the full wave function is manifestly permutation equivariant.

resulting model as initialization for larger systems. Throughout this section, we denote by $\mathbf{r}_i$ the Cartesian representation of Ω_i , i.e., the coordinates of the i th electron on the sphere, and by $\mathbf{R}_i$ the Cartesian representation of Ω_i^{BF} , the corresponding backflow coordinate.

a. The gating parameter

As mentioned previously, we approximate μ using an MLP acting on the inputs (κ, N, Q) . To enforce positivity, the output layer employs the softplus activation function,

$$\text{softplus}(x) = \ln(1 + e^x), \quad (\text{C4})$$

so that

$$\mu = \text{softplus}(\text{MLP}([\kappa, N, Q])). \quad (\text{C5})$$

b. Amplitude Jastrow

The amplitude term $\tilde{\mathcal{J}}_{\text{re}}$ is an O(3)-invariant function of $\mathbf{r}_1, \dots, \mathbf{r}_N$ and can therefore be expressed solely in terms of the pairwise distances $d_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, which are related to the dot products $\mathbf{r}_i \cdot \mathbf{r}_j$ through an invertible mapping. To satisfy the cusp constraint $\partial_{d_{ij}} \tilde{\mathcal{J}}_{\text{re}}|_{d_{ij}=0} = 0$, we write

$$\tilde{\mathcal{J}}_{\text{re}}(\{\mathbf{r}\}) = \sum_{i<j} f_{ij}(\{d_{ab}\}) d_{ij}^2, \quad (\text{C6})$$

which is always possible for analytic $\tilde{\mathcal{J}}_{\text{re}}$. Furthermore, permutation invariance requires f_{ij} to be symmetric under permutations of $\{\mathbf{r}_a : a \neq i, j\}$, which we enforce via the DeepSets architecture (see Fig. 8).

c. Phase Jastrow

The phase term $\mathcal{J}_{\text{im}}$ is SO(3) invariant and odd under inversion, and can thus be expressed as a sum over scalar triple products $\chi_{ijk} = \mathbf{r}_i \cdot (\mathbf{r}_j \times \mathbf{r}_k)$ with coefficient functions $f_{ijk}(\{d_{ab}\})$:

$$\mathcal{J}_{\text{im}}(\{\mathbf{r}\}) = \sum_{i<j<k} f_{ijk}(\{d_{ab}\}) \chi_{ijk}. \quad (\text{C7})$$

Since χ_{ijk} is totally antisymmetric in i, j, k , permutation invariance requires f_{ijk} to also be totally antisymmetric. We construct such an f_{ijk} as a determinant over three O(3)-invariant three-dimensional vectors $\mathbf{f}_a^{ijk}$ ($a = i, j, k$), as shown in Fig. 8.

d. Backflow coordinates

The backflow coordinates $\mathbf{R}_i(\{\mathbf{r}\})$ are O(3) equivariant. To construct them, we express them as a sum over $\mathbf{r}_j$ with coefficient functions $f_{ij}(\{d_{ab}\})$ that are symmetric under

permutations of $\{\mathbf{r}_a : a \neq i, j\}$:

$$\mathbf{R}_i = \sum_j f_{ij}(\{d_{ab}\}) \mathbf{r}_j. \quad (\text{C8})$$

To ensure $\mathbf{R}_i$ lies on the sphere, we first construct a tangent vector

$$\mathbf{v}_i = \sum_j f_{ij} \mathbf{t}_{ij}, \quad \mathbf{t}_{ij} = \mathbf{r}_j - (\mathbf{r}_j \cdot \mathbf{r}_i) \mathbf{r}_i, \quad (\text{C9})$$

where $\mathbf{t}_{ij}$ projects $\mathbf{r}_j$ onto the tangent plane at $\mathbf{r}_i$, and then apply the exponential map on the sphere [112]:

$$\mathbf{R}_i = \exp_{\mathbf{r}_i}(\mathbf{v}_i) = \cos(v_i) \mathbf{r}_i + \sin(v_i) \hat{\mathbf{v}}_i. \quad (\text{C10})$$

The coefficients f_{ij} are constructed as before using DeepSets. Figure 8 summarizes the complete CF-Flow architecture.

APPENDIX D: OPTIMIZATION DETAILS

In this Appendix, we give the optimization details summarized in Sec. V A.

Within the foundational-model framework, parameter updates are determined by ensemble-averaged quantities: the quantum geometric tensor (QGT) $\mathcal{S}_{\alpha\beta}$ and the variational force $\mathcal{F}_\alpha$, defined as [90]

$$\begin{aligned} \mathcal{F}_\alpha &= \frac{1}{\kappa_{\max} - \kappa_{\min}} \int_{\kappa_{\min}}^{\kappa_{\max}} d\kappa F_\alpha(\kappa), \\ \mathcal{S}_{\alpha\beta} &= \frac{1}{\kappa_{\max} - \kappa_{\min}} \int_{\kappa_{\min}}^{\kappa_{\max}} d\kappa S_{\alpha\beta}(\kappa), \end{aligned} \quad (\text{D1})$$

where we take $\kappa_{\min} = 1$ and $\kappa_{\max} = 20$. At each κ , the force $\mathcal{F}_\alpha$ and QGT $\mathcal{S}_{\alpha\beta}$ are computed as [113]

$$\begin{aligned} F_\alpha &= -2 \Re[\langle (\mathcal{H} - \langle \mathcal{H} \rangle)(\mathcal{O}_\alpha - \langle \mathcal{O}_\alpha \rangle)^* \rangle], \\ S_{\alpha\beta} &= \Re[\langle (\mathcal{O}_\alpha - \langle \mathcal{O}_\alpha \rangle)^* (\mathcal{O}_\beta - \langle \mathcal{O}_\beta \rangle) \rangle], \\ \mathcal{O}_\alpha &= \partial_{\lambda_\alpha} \ln \psi(\mathbf{r}_1, \dots, \mathbf{r}_N; \{\lambda_\gamma\}), \end{aligned} \quad (\text{D2})$$

where expectation values are taken with respect to the CF-Flow wave function $\psi(\mathbf{r}_1, \dots, \mathbf{r}_N; \{\lambda_\gamma\})$.

Compared with the standard random-walk Metropolis-Hastings scheme, MALA achieves equivalent statistical accuracy with significantly fewer samples. At each training iteration (epoch), we draw four κ values uniformly from (1, 20) to construct the ensemble averages in Eq. (D1). For each κ , we run eight independent Markov chains, generating 400 samples per chain after discarding 512 initial samples for equilibration. To reduce memory overhead, we employ the “mini-SR” technique of Ref. [113].

The variational parameters are updated via

$$\lambda_\alpha(t+1) = \lambda_\alpha(t) + \eta(t) [\mathcal{S}_{\alpha\beta}(t) + \epsilon(t) \delta_{\alpha\beta}]^{-1} \mathcal{F}_\beta(t), \quad (\text{D3})$$

where $\eta(t)$ is the learning rate and $\epsilon(t)$ is a regularization parameter. Both are decayed during training: $\eta(t)$ is halved whenever the (ensemble-averaged) energy plateaus (i.e., ceases to decrease), and optimization terminates after 250 epochs without improvement.

This staged approach also guides the choice of hyperparameters for the individual learnable modules of CF-Flow. We set $\mathcal{M} = 3$ for $\nu = 1/3$ (Laughlin) and $\mathcal{M} = 1$ for all other states (see Appendix E). We employ a small MLP for μ , while

the backflow and amplitude-Jastrow components use larger MLPs with one or two perceptron layers of width 48–96. The phase-Jastrow network is kept roughly 6 times smaller, as its contribution to the total energy is minor (we have checked this explicitly by considering larger networks). In total, the model contains about 5×10^4 trainable parameters. Networks with 2–4 times fewer parameters yield statistically similar results, but we adopt the larger configuration to ensure sufficient representational capacity across system sizes and filling factors.

We do not believe the energies and local-energy standard deviations reported here represent fundamental limits of CF-Flow. Further improvements could be achieved by reducing stochastic noise through longer sampling chains, though we did not pursue this further due to computational cost.

APPENDIX E: SHORT-DISTANCE CONSTRAINT ON CF-FLOW

In this Appendix, we discuss a formal constraint on CF-Flow arising from our choice of activation function and its implications for the short-distance behavior of the wave function.

1. Constraint from analyticity

The SR method requires second-order spatial derivatives of the wave function [see Eq. (2)] to evaluate the local kinetic energy. We therefore use the smooth GELU activation function in all MLPs (see Fig. 8),

$$\text{GELU}(x) = x \frac{1 + \text{erf}(x/\sqrt{2})}{2}, \quad (\text{E1})$$

which renders the overall ansatz analytic in the particle coordinates $\mathbf{r}_i$.

Because the wave function is analytic, CF-Flow cannot modify the order of zeros when two particles coincide. Specifically, it cannot reduce the relative-angular-momentum exponent $\mathcal{M}$ that governs the short-distance behavior $\psi(\mathbf{r}) \sim |\mathbf{r}_i - \mathbf{r}_j|^\mathcal{M}$ as $\mathbf{r}_i \rightarrow \mathbf{r}_j$. We therefore set $\mathcal{M}$ to match the underlying CF state: $\mathcal{M} = 3$ for the $\nu = 1/3$ Laughlin state [Eq. (7)] and $\mathcal{M} = 1$ for all other CF states considered here. (More generally, for two-vortex CF states, as long as at least one CF resides in an $n \geq 1$ Λ level after LLL projection, we have $\mathcal{M} = 1$.) Higher-vortex ($p \geq 2$) CF states would require different choices, though we have not explored them in this work.

2. Practical implications

The practical implications of this constraint are not fully understood. Even within the LLL, the Laughlin wave function, whose short-distance exponent is fixed at $\mathcal{M} = 3$, is known to approximate the Coulomb ground state extremely well: It reproduces the per-particle energy to within $\sim 10^{-4} e^2/\epsilon \ell_B$ [13, 114]. At the same time, the exact ground state contains a small admixture of the $\nu^* = 1$ IQH state of CFs and CF excitons, which lowers the short-distance exponent to $\mathcal{M} = 1$ [13, 87]. CF-Flow, by construction, retains the exponent of the underlying CF state.

Whether enforcing $\mathcal{M} = 3$ at $\nu = 1/3$ affects the *reachable* accuracy of CF-Flow is unclear. In practice, we believe the precision is limited by stochastic noise in Monte Carlo sampling and finite-precision effects in gradient-based optimization. Within these limits, any consequence of the fixed exponent would be extremely difficult to resolve.

As shown in Fig. 4, CF-Flow nevertheless attains the same level of energy accuracy as existing methods. Determining whether altering the short-distance exponent could yield improvements below the current numerical noise floor remains an open question.

-
- [1] G. Baym and C. Pethick, *Landau Fermi-liquid Theory: Concepts and Applications* (Wiley, Hoboken, NJ, 1991).
 - [2] D. C. Tsui, H. L. Stormer, and A. C. Gossard, Two-dimensional magnetotransport in the extreme quantum limit, *Phys. Rev. Lett.* **48**, 1559 (1982).
 - [3] H. L. Stormer, A. Chang, D. C. Tsui, J. C. M. Hwang, A. C. Gossard, and W. Wiegmann, Fractional quantization of the Hall effect, *Phys. Rev. Lett.* **50**, 1953 (1983).
 - [4] Y. Teng, D. D. Dai, and L. Fu, Solving the fractional quantum Hall problem with self-attention neural network, *Phys. Rev. B* **111**, 205117 (2025).
 - [5] Y. Qian, T. Zhao, J. Zhang, T. Xiang, X. Li, and J. Chen, Describing Landau level mixing in fractional quantum Hall states with deep learning, *Phys. Rev. Lett.* **134**, 176503 (2025).
 - [6] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
 - [7] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
 - [8] H. Lange, A. Van de Walle, A. Abedinnia, and A. Bohrdt, From architectures to applications: A review of neural quantum states, *Quantum Sci. Technol.* **9**, 040501 (2024).
 - [9] I. von Glehn, J. S. Spencer, and D. Pfau, A self-attention ansatz for *ab-initio* quantum chemistry, in *Proceedings of the Eleventh International Conference on Learning Representations (ICLR 2023)*, Kigali, Rwanda (OpenReview.net, 2023).
 - [10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, Attention is all you need, in *Advances in Neural Information Processing Systems (NIPS 2017)* (Curran Associates, Red Hook, NY, 2017), Vol. 30, pp. 5998–6008.
 - [11] K. Nazaryan, F. Gaggioli, Y. Teng, and L. Fu, Artificial intelligence for quantum matter: Finding a needle in a haystack, [arXiv:2507.13322](https://arxiv.org/abs/2507.13322).
 - [12] J. K. Jain, Composite-fermion approach for the fractional quantum Hall effect, *Phys. Rev. Lett.* **63**, 199 (1989).
 - [13] J. K. Jain, *Composite Fermions* (Cambridge University Press, New York, 2007).
 - [14] G. Fano, F. Ortolani, and E. Colombo, Configuration-interaction calculations on the fractional quantum Hall effect, *Phys. Rev. B* **34**, 2670 (1986).
 - [15] G. Dev and J. K. Jain, Band structure of the fractional quantum Hall effect, *Phys. Rev. Lett.* **69**, 2843 (1992).
 - [16] X. G. Wu, G. Dev, and J. K. Jain, Mixed-spin incompressible states in the fractional quantum Hall effect, *Phys. Rev. Lett.* **71**, 153 (1993).
 - [17] A. C. Balram, G. J. Sreejith, and J. K. Jain, Splitting of the Girvin-MacDonald-Platzman density wave and the nature of chiral gravitons in the fractional quantum Hall effect, *Phys. Rev. Lett.* **133**, 246605 (2024).
 - [18] Y. W. Suen, L. W. Engel, M. B. Santos, M. Shayegan, and D. C. Tsui, Observation of a $\nu = 1/2$ fractional quantum Hall state in a double-layer electron system, *Phys. Rev. Lett.* **68**, 1379 (1992).
 - [19] Y. W. Suen, H. C. Manoharan, X. Ying, M. B. Santos, and M. Shayegan, One-component to two-component transitions of fractional quantum Hall states in a wide quantum well, *Surf. Sci.* **305**, 13 (1994).
 - [20] D. R. Luhman, W. Pan, D. C. Tsui, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, Observation of a fractional quantum Hall state at $\nu = 1/4$ in a wide GaAs quantum well, *Phys. Rev. Lett.* **101**, 266804 (2008).
 - [21] J. Shabani, Y. Liu, M. Shayegan, L. N. Pfeiffer, K. W. West, and K. W. Baldwin, Phase diagrams for the stability of the $\nu = \frac{1}{2}$ fractional quantum Hall effect in electron systems confined to symmetric, wide GaAs quantum wells, *Phys. Rev. B* **88**, 245413 (2013).
 - [22] J. Shabani, T. Gokmen, and M. Shayegan, Correlated states of electrons in wide quantum wells at low fillings: The role of charge distribution symmetry, *Phys. Rev. Lett.* **103**, 046805 (2009).
 - [23] J. Shabani, T. Gokmen, Y. T. Chiu, and M. Shayegan, Evidence for developing fractional quantum Hall states at even denominator $1/2$ and $1/4$ fillings in asymmetric wide quantum wells, *Phys. Rev. Lett.* **103**, 256802 (2009).
 - [24] A. Sharma, A. C. Balram, and J. K. Jain, Composite-fermion pairing at half-filled and quarter-filled lowest Landau level, *Phys. Rev. B* **109**, 035306 (2024).
 - [25] J. Zhao, Y. Zhang, and J. K. Jain, Crystallization in the fractional quantum Hall regime induced by Landau-level mixing, *Phys. Rev. Lett.* **121**, 116802 (2018).
 - [26] T. Zhao, A. C. Balram, and J. K. Jain, Composite fermion pairing induced by Landau level mixing, *Phys. Rev. Lett.* **130**, 186302 (2023).
 - [27] C. Töke, M. R. Peterson, G. S. Jeon, and J. K. Jain, Fractional quantum Hall effect in the second Landau level: The importance of inter-composite-fermion interaction, *Phys. Rev. B* **72**, 125315 (2005).
 - [28] R. P. Feynman and M. Cohen, Energy spectrum of the excitations in liquid helium, *Phys. Rev.* **102**, 1189 (1956).
 - [29] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Póczos, R. Salakhutdinov, and A. J. Smola, Deep sets, in *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17* (Curran Associates, Inc., Red Hook, NY, 2017), pp. 3394–3404.
 - [30] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, Geometric deep learning: Going beyond Euclidean data, *IEEE Signal Process. Mag.* **34**, 18 (2017).
 - [31] S. Villar, D. W. Hogg, K. Storey-Fisher, W. Yao, and B. Blum-Smith, Scalars are universal: Equivariant machine learning,

- structured like classical physics, *Neural Inf. Process. Syst.* **34**, 28848 (2021).
- [32] P. G. Debenedetti, J. J. de Pablo, and G. C. Schatz, Machine learning in chemistry, *Proc. Natl. Acad. Sci. USA* **122**, e2517062122 (2025).
- [33] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Inhomogeneous backflow transformations in quantum Monte Carlo calculations, *Phys. Rev. E* **74**, 066701 (2006).
- [34] D. Luo and B. K. Clark, Backflow transformations via neural networks for quantum many-body wave functions, *Phys. Rev. Lett.* **122**, 226401 (2019).
- [35] G. Pescia, J. Nys, J. Kim, A. Lovato, and G. Carleo, Message-passing neural quantum states for the homogeneous electron gas, *Phys. Rev. B* **110**, 035108 (2024).
- [36] I. Sodemann and A. H. MacDonald, Landau level mixing and the fractional quantum Hall effect, *Phys. Rev. B* **87**, 245425 (2013).
- [37] A. H. MacDonald, Influence of Landau-level mixing on the charge-density-wave state of a two-dimensional electron gas in a strong magnetic field, *Phys. Rev. B* **30**, 4392 (1984).
- [38] V. Melik-Alaverdian and N. E. Bonesteel, Composite fermions and Landau-level mixing in the fractional quantum Hall effect, *Phys. Rev. B* **52**, R17032 (1995).
- [39] G. Murthy and R. Shankar, Hamiltonian theory of the fractional quantum Hall effect: Effect of Landau level mixing, *Phys. Rev. B* **65**, 245309 (2002).
- [40] W. Bishara and C. Nayak, Effect of Landau level mixing on the effective interaction between electrons in the fractional quantum Hall regime, *Phys. Rev. B* **80**, 121302 (2009).
- [41] A. Wójs, C. Tóke, and J. K. Jain, Landau-level mixing and the emergence of Pfaffian excitations for the $\nu = 5/2$ fractional quantum Hall effect, *Phys. Rev. Lett.* **105**, 096802 (2010).
- [42] S. H. Simon and E. H. Rezayi, Landau level mixing in the perturbative limit, *Phys. Rev. B* **87**, 155426 (2013).
- [43] M. R. Peterson and C. Nayak, More realistic Hamiltonians for the fractional quantum Hall regime in GaAs and graphene, *Phys. Rev. B* **87**, 245129 (2013).
- [44] M. R. Peterson and C. Nayak, Effects of Landau level mixing on the fractional quantum Hall effect in monolayer graphene, *Phys. Rev. Lett.* **113**, 086401 (2014).
- [45] R. H. Morf, N. d'Ambrumenil, and S. Das Sarma, Excitation gaps in fractional quantum Hall states: An exact diagonalization study, *Phys. Rev. B* **66**, 075408 (2002).
- [46] B. I. Halperin, P. A. Lee, and N. Read, Theory of the half-filled Landau level, *Phys. Rev. B* **47**, 7312 (1993).
- [47] R. L. Willett, R. R. Ruel, K. W. West, and L. N. Pfeiffer, Experimental demonstration of a Fermi surface at one-half filling of the lowest Landau level, *Phys. Rev. Lett.* **71**, 3846 (1993).
- [48] M. S. Hossain, M. K. Ma, M. A. Mueed, D. Kamburov, L. N. Pfeiffer, K. W. West, K. W. Baldwin, R. Winkler, and M. Shayegan, Geometric resonance of four-flux composite fermions, *Phys. Rev. B* **100**, 041112 (2019).
- [49] *Fractional Quantum Hall Effects: New Developments*, edited by B. I. Halperin and J. K. Jain (World Scientific, Singapore, 2020).
- [50] C. Wang, A. Gupta, S. K. Singh, P. T. Madathil, Y. J. Chung, L. N. Pfeiffer, K. W. Baldwin, R. Winkler, and M. Shayegan, Fractional quantum Hall state at filling factor $\nu = 1/4$ in ultra-high-quality GaAs two-dimensional hole systems, *Phys. Rev. Lett.* **131**, 266502 (2023).
- [51] K. A. Villegas Rosales, S. K. Singh, M. K. Ma, M. S. Hossain, Y. J. Chung, L. N. Pfeiffer, K. W. West, K. W. Baldwin, and M. Shayegan, Competition between fractional quantum Hall liquid and Wigner solid at small fillings: Role of layer thickness and Landau level mixing, *Phys. Rev. Res.* **3**, 013181 (2021).
- [52] M. B. Santos, Y. W. Suen, M. Shayegan, Y. P. Li, L. W. Engel, and D. C. Tsui, Observation of a reentrant insulating phase near the $1/3$ fractional quantum Hall liquid in a two-dimensional hole system, *Phys. Rev. Lett.* **68**, 1188 (1992).
- [53] R. Willett, J. P. Eisenstein, H. L. Störmer, D. C. Tsui, A. C. Gossard, and J. H. English, Observation of an even-denominator quantum number in the fractional quantum Hall effect, *Phys. Rev. Lett.* **59**, 1776 (1987).
- [54] W. Pan, J.-S. Xia, V. Shvarts, D. E. Adams, H. L. Stormer, D. C. Tsui, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, Exact quantization of the even-denominator fractional quantum Hall state at $\nu = 5/2$ Landau level filling factor, *Phys. Rev. Lett.* **83**, 3530 (1999).
- [55] R. H. Morf, Transition from quantum Hall to compressible states in the second Landau level: New light on the $\nu = 5/2$ enigma, *Phys. Rev. Lett.* **80**, 1505 (1998).
- [56] G. Moore and N. Read, Nonabelions in the fractional quantum Hall effect, *Nucl. Phys. B* **360**, 362 (1991).
- [57] G. Möller and S. H. Simon, Paired composite-fermion wave functions, *Phys. Rev. B* **77**, 075319 (2008).
- [58] M. R. Peterson, K. Park, and S. Das Sarma, Spontaneous particle-hole symmetry breaking in the $\nu = 5/2$ fractional quantum Hall effect, *Phys. Rev. Lett.* **101**, 156803 (2008).
- [59] E. H. Rezayi and S. H. Simon, Breaking of particle-hole symmetry by Landau level mixing in the $\nu = 5/2$ quantized Hall state, *Phys. Rev. Lett.* **106**, 116801 (2011).
- [60] K. Pakrouski, M. R. Peterson, T. Jolicoeur, V. W. Scarola, C. Nayak, and M. Troyer, Phase diagram of the $\nu = 5/2$ fractional quantum Hall effect: Effects of Landau-level mixing and nonzero width, *Phys. Rev. X* **5**, 021004 (2015).
- [61] E. H. Rezayi, Landau level mixing and the ground state of the $\nu = 5/2$ quantum Hall effect, *Phys. Rev. Lett.* **119**, 026801 (2017).
- [62] S. Das Sarma, M. Freedman, and C. Nayak, Topologically protected qubits from a possible non-Abelian fractional quantum Hall state, *Phys. Rev. Lett.* **94**, 166802 (2005).
- [63] G. Ortiz, D. M. Ceperley, and R. M. Martin, New stochastic method for systems with broken time-reversal symmetry: 2D fermions in a magnetic field, *Phys. Rev. Lett.* **71**, 2777 (1993).
- [64] J. Kolorenč and L. Mitas, Applications of quantum Monte Carlo methods in condensed systems, *Rep. Prog. Phys.* **74**, 026502 (2011).
- [65] V. Melik-Alaverdian, G. Ortiz, and N. E. Bonesteel, Quantum projector method on curved manifolds, *J. Stat. Phys.* **104**, 449 (2001).
- [66] V. Melik-Alaverdian, N. E. Bonesteel, and G. Ortiz, Quantum Hall fluids on the Haldane sphere: A diffusion Monte Carlo study, *Phys. Rev. Lett.* **79**, 5286 (1997).
- [67] Y. Zhang, A. Wójs, and J. K. Jain, Landau-level mixing and particle-hole symmetry breaking for spin transitions in the fractional quantum Hall effect, *Phys. Rev. Lett.* **117**, 116803 (2016).

- [68] M. K. Ma, K. A. Villegas Rosales, H. Deng, Y. J. Chung, L. N. Pfeiffer, K. W. West, K. W. Baldwin, R. Winkler, and M. Shayegan, Thermal and quantum melting phase diagrams for a magnetic-field-induced Wigner solid, *Phys. Rev. Lett.* **125**, 036601 (2020).
- [69] F. D. M. Haldane, Fractional quantization of the Hall effect: A hierarchy of incompressible quantum fluid states, *Phys. Rev. Lett.* **51**, 605 (1983).
- [70] X.-G. Wen and A. Zee, Topological degeneracy of quantum Hall fluids, *Phys. Rev. B* **58**, 15717 (1998).
- [71] X. G. Wen and A. Zee, Shift and spin vector: New topological quantum numbers for the Hall fluids, *Phys. Rev. Lett.* **69**, 953 (1992).
- [72] T. T. Wu and C. N. Yang, Dirac monopole without strings: Monopole harmonics, *Nucl. Phys. B* **107**, 365 (1976).
- [73] J. K. Jain, Incompressible quantum Hall states, *Phys. Rev. B* **40**, 8079 (1989).
- [74] A. C. Balram, M. Barkeshli, and M. S. Rudner, Parton construction of a wave function in the anti-Pfaffian phase, *Phys. Rev. B* **98**, 035127 (2018).
- [75] J. K. Jain and R. K. Kamilla, Composite fermions in the Hilbert space of the lowest electronic Landau level, *Int. J. Mod. Phys. B* **11**, 2621 (1997).
- [76] M. Gattu and J. K. Jain, Unlocking new regimes in fractional quantum Hall effect with quaternions, *Phys. Rev. Lett.* **134**, 156501 (2025).
- [77] A. Anakru, M. Gattu, A. C. Balram, X.-C. Wu, P. Kumar, Z. Bi, and J. K. Jain, Emergent gauge field in composite-fermion metals: A large-scale microscopic study, *Phys. Rev. Lett.* **135**, 246503 (2025).
- [78] J. K. Jain and M. R. Peterson, Reconstructing the electron in a fractionalized quantum fluid, *Phys. Rev. Lett.* **94**, 186808 (2005).
- [79] M. Gattu, G. J. Sreejith, and J. K. Jain, Scanning tunneling microscopy of fractional quantum Hall states: Spectroscopy of composite-fermion bound states, *Phys. Rev. B* **109**, L201123 (2024).
- [80] S. Pu, A. C. Balram, Y. Hu, Y.-C. Tsui, M. He, N. Regnault, M. P. Zaletel, A. Yazdani, and Z. Papić, Fingerprints of composite fermion Lambda levels in scanning tunneling microscopy, *Phys. Rev. B* **110**, L081107 (2024).
- [81] S. S. Mandal and J. K. Jain, Theoretical search for the nested quantum Hall effect of composite fermions, *Phys. Rev. B* **66**, 155302 (2002).
- [82] A. C. Balram, A. Wójs, and J. K. Jain, State counting for excited bands of the fractional quantum Hall effect: Exclusion rules for bound excitons, *Phys. Rev. B* **88**, 205312 (2013).
- [83] R. K. Kamilla, X. G. Wu, and J. K. Jain, Excitons of composite fermions, *Phys. Rev. B* **54**, 4873 (1996).
- [84] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics* (Cambridge University Press, Cambridge, 2020).
- [85] S. M. Girvin, A. H. MacDonald, and P. M. Platzman, Collective-excitation gap in the fractional quantum Hall effect, *Phys. Rev. Lett.* **54**, 581 (1985).
- [86] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [87] M. R. Peterson and J. K. Jain, Possible persistence of fractional quantum Hall effect down to ultralow fillings, *Phys. Rev. B* **68**, 195310 (2003).
- [88] S. Sorella, Green function Monte Carlo with stochastic reconfiguration, *Phys. Rev. Lett.* **80**, 4558 (1998).
- [89] S. Sorella, Wave function optimization in the variational Monte Carlo method, *Phys. Rev. B* **71**, 241103 (2005).
- [90] R. Rende, L. L. Viteritti, F. Becca, A. Scardicchio, A. Laio, and G. Carleo, Foundation neural-networks quantum states as a unified ansatz for multiple Hamiltonians, *Nat. Commun.* **16**, 7213 (2025).
- [91] L. L. Viteritti, R. Rende, G. Bracci-Testasecca, J. Niedda, R. Moessner, G. Carleo, and A. Scardicchio, Quantum spin glass in the two-dimensional disordered Heisenberg model via foundation neural-network quantum states, [arXiv:2507.05073](https://arxiv.org/abs/2507.05073).
- [92] M. A. Girolami and B. Calderhead, Riemann manifold Langevin and Hamiltonian Monte Carlo methods, *J. R. Stat. Soc.: Ser. B* **73**, 123 (2011).
- [93] R. Morf, N. d'Ambrumenil, and B. I. Halperin, Microscopic wave functions for the fractional quantized Hall states at $\nu = \frac{2}{5}$ and $\frac{2}{7}$, *Phys. Rev. B* **34**, 3037 (1986).
- [94] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, MA, 2005).
- [95] A. Abouelkomsan and L. Fu, First-principles AI finds crystallization of fractional quantum Hall liquids, *PRX Intell.* **1**, 013010 (2026).
- [96] H. Li and F. D. M. Haldane, Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-Abelian fractional quantum Hall effect states, *Phys. Rev. Lett.* **101**, 010504 (2008).
- [97] J. Cai, E. Anderson, C. Wang, X. Zhang, X. Liu, W. Holtzmann, Y. Zhang, F. Fan, T. Taniguchi, K. Watanabe, Y. Ran, T. Cao, L. Fu, D. Xiao, W. Yao, and X. Xu, Signatures of fractional quantum anomalous Hall states in twisted MoTe₂, *Nature (London)* **622**, 63 (2023).
- [98] H. Park, J. Cai, E. Anderson, Y. Zhang, J. Zhu, X. Liu, C. Wang, W. Holtzmann, C. Hu, Z. Liu, *et al.*, Observation of fractionally quantized anomalous Hall effect, *Nature (London)* **622**, 74 (2023).
- [99] Y. Zeng, Z. Xia, K. Kang, J. Zhu, P. Knüppel, C. Vaswani, K. Watanabe, T. Taniguchi, K. F. Mak, and J. Shan, Thermodynamic evidence of fractional Chern insulator in moiré MoTe₂, *Nature (London)* **622**, 69 (2023).
- [100] F. Xu, Z. Sun, T. Jia, C. Liu, C. Xu, C. Li, Y. Gu, K. Watanabe, T. Taniguchi, B. Tong, J. Jia, Z. Shi, S. Jiang, Y. Zhang, X. Liu, and T. Li, Observation of integer and fractional quantum anomalous Hall effects in twisted bilayer MoTe₂, *Phys. Rev. X* **13**, 031037 (2023).
- [101] Z. Lu, T. Han, Y. Yao, A. P. Reddy, J. Yang, J. Seo, K. Watanabe, T. Taniguchi, L. Fu, and L. Ju, Fractional quantum anomalous Hall effect in multilayer graphene, *Nature (London)* **626**, 759 (2024).
- [102] B. A. Bernevig, L. Fu, L. Ju, A. H. MacDonald, K. F. Mak, and J. Shan, Fractional quantization in insulators from Hall to Chern, *Nat. Phys.* **21**, 1702 (2025).
- [103] H. Lu, B.-B. Chen, H.-Q. Wu, K. Sun, and Z. Y. Meng, Thermodynamic response and neutral excitations in integer and fractional quantum anomalous Hall states emerging from correlated flat bands, *Phys. Rev. Lett.* **132**, 236502 (2024).
- [104] S. Danisch and J. Krumbiegel, Makie.jl: Flexible high-performance data visualization for Julia, *J. Open Source Softw.* **6**, 3349 (2021).

- [105] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: Composable transformations of Python+NumPy programs, 2018, <http://github.com/jax-ml/jax>.
- [106] M. Gattu, Data for “Dressing composite fermions with artificial intelligence,” GitHub repository, 2026, <https://github.com/mytraya-gattu/cf-flow-data>.
- [107] T. T. Wu and C. N. Yang, Some properties of monopole harmonics, *Phys. Rev. D* **16**, 1018 (1977).
- [108] M. Boyle, How should spin-weighted spherical functions be defined? *J. Math. Phys.* **57**, 092504 (2016).
- [109] J. K. Jain and R. K. Kamilla, Quantitative study of large composite-fermion systems, *Phys. Rev. B* **55**, R4895 (1997).
- [110] E. Wagstaff, F. Fuchs, M. Engelcke, M. A. Osborne, and I. Posner, Universal approximation of functions on sets, *J. Mach. Learn. Res.* **23**, 151 (2022).
- [111] H. Weyl, *The Classical Groups: Their Invariants and Representations*, Princeton Landmarks in Mathematics and Physics (Princeton University Press, Princeton, NJ, 1997).
- [112] M. P. do Carmo, *Riemannian Geometry, Mathematics: Theory & Applications* (Birkhäuser, Boston, 1992).
- [113] R. Rende, L. L. Viteritti, L. Bardone, F. Becca, and S. Goldt, A simple linear algebra identity to optimize large-scale neural network quantum states, *Commun. Phys.* **7**, 260 (2024).
- [114] R. B. Laughlin, Anomalous quantum Hall effect: An incompressible quantum fluid with fractionally charged excitations, *Phys. Rev. Lett.* **50**, 1395 (1983).