PAPER • OPEN ACCESS

Low-temperature pseudogap phenomenon: precursor of high-T_c superconductivity

New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG | IOP Institute of Physics

Low-temperature pseudogap phenomenon: precursor of high-T_c superconductivity

Yao Ma[1], Peng Ye[2] and Zheng-Yu Weng[1,3]

[1] Institute for Advanced Study, Tsinghua University, Beijing, 100084, People's Republic of China
[2] Perimeter Institute for Theoretical Physics, Waterloo, ON, N2L 2Y5, Canada
[3] Collaborative Innovation Center of Quantum Matter, Tsinghua University, Beijing, 100084, People's Republic of China
E-mail: weng@tsinghua.edu.cn

Received 10 March 2014, revised 18 June 2014
Accepted for publication 21 July 2014
Published 26 August 2014

New Journal of Physics **16** (2014) 083039

doi:10.1088/1367-2630/16/8/083039

Abstract

In this paper, we try to understand the pseudogap phenomenon observed in the cuprate superconductor through a model study. Specifically, we explore the so-called low-temperature pseudogap state by turning off the superconducting (SC) off diagonal long range order in an ansatz state for the t–J model (Weng 2011 *New J. Phys.* **13** 103039). Besides strong non-Gaussian SC fluctuations, the resulting state also exhibits a systematic pseudogap behavior in both spin and charge degrees of freedom, manifested in the uniform spin susceptibility, specific heat, non-Drude resistivity, Nernst effect, as well as the quantum oscillation associated with small Fermi pockets emerging in strong magnetic fields, etc. These anomalous 'normal state' properties are found in qualitative consistency with experimental measurements in the cuprates. Such a model study establishes an intrinsic connection between the peculiar pseudogap properties and the non-BCS nature of the SC ground state. A critical comparison with other approaches to the doped Mott insulator is also made.

Keywords: doped Mott insulators, superconductivity, pseudogap physics

1. Introduction

It has been well established [1–4] experimentally that the high-T_c superconductivity in the cuprates shares the same off-diagonal long-range order (ODLRO), in a form of electron Cooper pairing, with a conventional d-wave BCS superconductor. In the BCS theory, turning off such an ODLRO, say, by increasing temperature or applying a strong magnetic field, will lead to a Landau's Fermi liquid description of the normal state [5]. The issue under debate is what the 'normal state' is for a cuprate superconductor.

Experimentally it has been observed that the so-called *pseudogap regime* is always present at temperatures higher than the superconducting (SC) transition temperature T_c, at least in the underdoped regime [2, 6–8]. The pseudogap phenomenon has been regarded as one of the most unique physical properties of cuprate superconductors. Whether the pseudogap phase sets the stage for superconductivity to emerge as its low-temperature instability [8–18] or it simply competes with superconductivity as an independent phase [19–22], has been a crucial issue.

It has been a very challenging quest to understand the nature of pseudogap physics purely phenomenologically. Alternatively, it is desirable to lay down a general theoretical framework or organizing principles first and then critically examine the experimental facts. Once the self-consistency of a microscopic theory is established, its comparison with the rich experimental observations, even at a qualitative level, can either lend strong support for it or simply falsify it.

It has been widely accepted that the cuprate superconductors are doped Mott insulators of strong correlations [9]. At half-filling, the low-energy spin degrees of freedom are properly described by an antiferromagnetic (AF) Heisenberg model, while the charge degree of freedom is gapped due to the strong on-site Coulomb repulsion. The low-lying charge degree of freedom is introduced by doping holes into the system [8–10]. The SC state appears at low doping, where the AF long-range order (AFLRO) gets destroyed by the motion of doped holes. Experimentally, the Mott gap seems to remain finite and large [23, 24], which guarantees that the charge carriers are the doped holes, while the majority spins in the background are still neutral. Namely, at least in the underdoped cuprates, the superconductivity occurs in a doped Mott insulator regime [8–10].

Some main issues concerning the ground state physics are as follows. Firstly, how the AFLRO at half-filling gets destroyed by doping; secondly, how the superconductivity arises at finite doping; thirdly, how the superconductivity begins to diminish at overdoping. Once these are understood, then the next question is, after the SC coherence is destroyed by raising temperature or applying strong magnetic fields, what will be the normal state? In particular, if the pseudogap phase is a natural normal state above the SC dome in a doped Mott insulator?

Hence, the pseudogap physics can be used as a direct probe into a hypothesized SC ground state. The Gutzwiller-projected BCS ground state or the so-called 'plain vanilla' resonating-valence-bond (RVB) state has been previously proposed [9, 13] to describe the superconductivity in the simplest doped Mott insulator, i.e., the two-dimensional (2D) square lattice t–J model. The pseudogap properties based on this ground state ansatz has been intensively studied in the literature [8, 15].

Recently, a new SC ground state ansatz has been proposed [25] for the same t–J model. It is distinct from the aforementioned 'plain vanilla' RVB ground state by a two-component RVB

order instead of the one-component one. In contrast to the 'plain vanilla' RVB state, the new ground state can recover an accurate description of the AFLRO state of the Heisenberg model at the half-filling (zero doping) limit, while predicts a different (non-BCS) superconductivity at finite doping [25].

Most importantly, the new ansatz state can explicitly keep track of a singular sign structure (altered Fermi statistical signs) of the t–J model. Specifically, the conventional fermion signs of non-interacting electron gas are completely removed by strong on-site repulsion in the t–J model at half-filling, which is described by the Heisenberg model with the ground state precisely satisfying the so-called Marshall sign rule [26] that is a non-statistical trivial sign structure. The non-trivial sign structure (known as the phase string effect [27, 28]) only emerges upon doping, but is much sparser without bearing any similarity with the conventional fermion signs of the underlying electrons.

Based on such a new SC ground state [25], we shall present a theoretical description of the low-temperature pseudogap phase in this paper.

Some general properties unique to the present psuedogap phase are found as follows. The state is characterized by a generalized electron fractionalization. Here the correlated electrons are described as if there are three subsystems: holons for doped holes of charge $+e$; neutral spinons of $S = 1/2$; and the backflow spinons accompanying the hopping process of the holons. What is particularly simplifying in such a low-temperature pseudogap phase is that the bosonic holons are always Bose-condensed, while the neutral spinons and backflow spinons remain in two-component RVB pairing. In other words, the three subsystems are respectively all in the ODLRO states of their own.

It is well known that ODLRO in a condensed matter system breaks a global symmetry, resulting in a generalized 'rigidity' and great simplications for a many-body system. For these subsystems, the involved symmetries are actually the associated emerging U(1) gauge degrees of freedom upon fractionalization [25]. The 'Meissner' effects due to these ODLROs then suppress the gauge fluctuations, which otherwise would strongly fluctuate and confine these fractionalized particles back to the electrons [8]. Thus, the hidden ODLROs of the subsystems protect the fractionalization of the electrons self-consistently.

These hidden ODLROs do not necessarily break the *true* global symmetries of the system in general. As it turns out, the aforementioned non-trivial sign structure or phase string effect dictates the above peculiar fractionalization of the electrons and mediates the so-called mutual Chern–Simons gauge interaction between the holons and neutral spinons. In contrast to the U(1) gauge fluctuations associated with the fractionalization, such topological gauge fields cannot be 'Higgsed' by the hidden ODRLOs and play a crucial role in distroying the SC phase coherence in the pseudogap phase.

The pseudogap behaviors are actually the explicit consequences of these hidden ODLROs. For example, the RVB ordering of the spinons is responsible for the pseudogap properties in the uniform magnetic susceptibility and specific heat capacity over a characteristic temperature T_0 set by a renormalized superexchange coupling (which decreases with doping), known as the upper pseudogap phase (UPP) [29]. It is in agreement with the early experiments [30–36].

The holon condensation further defines a lower pseudogap phase (LPP) at lower temperatures. Once it happens at a finite doping, the true AFLRO ceases to develop because a doping-dependent, small gap is induced by the mutual Chern–Simons gauge fields in the spin excitation spectrum. So the LPP is a 'spin liquid' (or short-ranged RVB state). It sets a stage for superconductivity to emerge, but the phase coherence is still disordered by thermally excited

spin excitations, again via the mutual Chern–Simons gauge fields. Such an LPP is thus featured by strong non-Gaussian SC fluctuations. The longitudinal resistivity is of non-Drude behavior, which is supplemented with a strong Nernst effect. All of these are clear at variance with the usual Gaussian fluctuations in a narrow critical regime of a BCS superconductor [5].

These anomalous pseudogap properties in the LPP are qualitatively consistent with the experimental observations in the cuprates [30–33, 35, 37–44]. Finally, the true SC instability happens when temperature is sufficiently lower than the spin gap, where the fractionalized spin excitations become 'confined' via the mutual Chern–Simons gauge fields, and their novel phase disordering effect on the SC coherence gets screened out below T_c [45].

It is noted that the LPP as a spin liquid/vortex liquid state has been previously studied by effective theory approaches [46–48]. The present study further provides a microscopic framework based on the wave function description, which handles the short-range correlations more carefully. The main distinction is an emergent two-component RVB structure in the spin background [25]. Namely, the spin degrees of freedom are now composed of two: a spin liquid always pinned at half-filling together with the so-called *backflow* spinons describing the hopping effect on the spin background.

Such a two-component RVB structure further predicts another non-SC state, which may be obtained with the turning off the ODLRO of the backflow spinons by strong magnetic fields, say, in the magnetic vortex core region. This core state is to be called the LPP-II, in which the backflow spinons become charged with coherent Fermi pockets, whose Luttinger volume is commensurate with the doped holes. It is responsible for a novel quantum oscillation and the Pauli behavior of the spin susceptibility, and provides a consistent explanation for the experiments [49–51].

The rest of the paper is organized as follows. In section 2, by starting with a new SC ground state ansatz for the doped t–J model, we introduce the LPP. An effective Hamiltonian and the corresponding phase diagram from mean field self-consistent calculation will be presented. In section 3 a self-consistent phenomenology for the LPP will be presented, based on the microscopic effective theory and a comparison with the experiments in the cuprates will be made. In section 4, a critical comparison of the present approach with the slave-boson approach to the t–J model will be made. Finally, section 5 will be devoted to a discussion. An effective topological field theory description of the LPP known as the compact mutual Chern–Simons gauge theory will be also outlined in appendix A.

2. Low-temperature pseudogap phase as precursor of superconductivity

As emphasized in section 1, in this work we will explore the LPP as a 'normal state' with the SC ODLRO being turned off. In other words, it is non-SC but is most closely related to the SC ground state. To characterize such an LPP microscopically, in the following, we start with a new SC ground state ansatz, which has been recently proposed [25] for the t–J model.

2.1. Motivation: SC ground state ansatz

In a doped Mott insulator, the doubly occupied sites are in a high-energy sector due to a large on-site repulsion U, which exists only in virtual processes to mediate the so-called AF superexchange coupling $J = 4t^2/U$ between the nearest neighboring (NN) spins (t is the bare NN hopping integral). At half-filling, with each lattice site occupied by one electron, the

relevant degrees of freedom are localized neutral spins. At finite doping, the neutral spins remain at singly occupied sites, and doped charge carriers are at those sites where the electron numbers deviate from the single occupancy. The simplest model to describe the cuprate system as a doped Mott insulator is the so-called t–J model [9, 11, 52]

$$H_{t-J} = -t\sum_{\langle ij\rangle\sigma}\hat{c}^{\dagger}_{i\sigma}\hat{c}_{j\sigma} + \text{h.c.} + J\sum_{\langle ij\rangle}\left(\hat{S}_i\cdot\hat{S}_j - \tfrac{1}{4}\hat{n}_i\hat{n}_j\right), \tag{1}$$

where i, j represent the sites in a 2D square lattice, 'h.c.' represents the Hermitian conjugate of the forward term, and $\hat{c}_{i\sigma}$ is the annihilation operator of an electron at site i with spin index σ. $\hat{\mathbf{S}}_i$ and $\hat{n}_i$ are the spin and number operators, respectively, at site i. In equation (1), a no-double-occupancy constraint must be always satisfied: $\hat{n}_i \leqslant 1$ for each site i for the hole-doped case. Defined in this restricted Hilbert space, $\hat{c}_{i\sigma}$ as a bare hole creation operator is not equivalent to the original electron annihilation operator anymore [8, 10].

An SC ground state has been recently constructed [25] based on the hole-doped 2D t–J model, which is of the following peculiar form of electron fractionalization

$$\left|\Phi_{\mathrm{G}}\right\rangle \equiv C\hat{\mathcal{P}}\left(\left|\Phi_h\right\rangle\otimes\left|\Phi_a\right\rangle\otimes\left|\Phi_b\right\rangle\right), \tag{2}$$

with C as the normalization factor.

Here, one does not see the electron creation or annihilation operators directly. Instead, the ground state is composed of three subsystems. The $+e$ charge sector of doped holes is described by

$$\left|\Phi_h\right\rangle \equiv \sum_{\{l_h\}}\varphi_h(l_1, l_2,...)h^{\dagger}_{l_1}h^{\dagger}_{l_2}...|0\rangle_h, \tag{3}$$

where the bosonic wave function $\varphi_h \simeq$ constant, which defines a Bose-condensed 'holon' state with $h^{\dagger}_l$ acting on a vacuum $|0\rangle_h$. The motion of doped holes will also generate spin backflows, which are described by 'itinerant' fermionic a-spinons as

$$\left|\Phi_a\right\rangle \equiv \exp\left(\sum_{ij}\tilde{g}_{ij}a^{\dagger}_{i\downarrow}a^{\dagger}_{j\uparrow}\right)|0\rangle_a, \tag{4}$$

where $a^{\dagger}_{i\sigma}$ acts on a vacuum $|0\rangle_a$ and the a-spinons are paired with an RVB amplitude $\tilde{g}_{ij}$.

The main spin background is described by

$$\left|\Phi_b\right\rangle = \exp\left(\sum_{ij}W_{ij}b^{\dagger}_{i\uparrow}b^{\dagger}_{j\downarrow}\right)|0\rangle_b \tag{5}$$

in which the neutral 'spinons' are RVB-paired with an amplitude W_{ij} in equation (5), where $b^{\dagger}_{i\sigma}$ as a bosonic b-spinon creation operator acts on a vacuum $|0\rangle_b$.

In the SC ground state (2), the no-double-occupancy constraint $\hat{n}_i \leqslant 1$ is enforced by the projection operator

$$\hat{\mathcal{P}} \equiv \hat{P}_{\mathrm{B}}\hat{P}_s, \tag{6}$$

where $\hat{P}_s$ enforces the single-occupancy constraint $\sum_\sigma n^b_{i\sigma} = 1$ ($n^b_{i\sigma} \equiv b^{\dagger}_{i\sigma}b_{i\sigma}$). Namely, the state

$$|\mathrm{RVB}\rangle \equiv \hat{P}_s\left|\Phi_b\right\rangle, \tag{7}$$

is a neutral spin background which is always pinned at half-filling. The localized spin background $|\mathrm{RVB}\rangle$ is a Liang–Docout–Anderson (LDA) [54] type of bosonic RVB state. At half-filling, $|\Psi_{\mathrm{G}}\rangle = C\,|\mathrm{RVB}\rangle$, and the LDA state can describe the AF ground state very accurately in this limit [25, 54].

$\hat{P}_{\mathrm{B}}$ in equation (6) further enforces the constraint that the h-holon and a-spinon always satisfy

$$n_{i\bar{\sigma}}^{a} = n_{i}^{h} n_{i\sigma}^{b} \tag{8}$$

such that each a-spinon always coincides with a holon as $\sum_{\sigma} n_{i\bar{\sigma}}^{a} = n_{i}^{h}$ (here $n_{i\bar{\sigma}}^{a} \equiv a_{i\bar{\sigma}}^{\dagger} a_{i\bar{\sigma}}$ and $n_{i}^{h} \equiv h_{i}^{\dagger} h_{i}$ with $\bar{\sigma} \equiv -\sigma$). Namely, at the hole site, the a-spinon will compensate the neutral b-spinon in $|\mathrm{RVB}\rangle$ which is an empty site physically. The a-spinons will accompany the hopping of the holons, and play a crucial role in keeping track of the effect of the itinerant motion of the doped holes on the spin degrees of freedom.

In the fractionalized form equation (2), there is no trace of the original electrons. In fact, the original electron $\hat{c}$-operator, acting on the ground state (2), can be expressed as follows: [25]:

$$\hat{c}_{i\sigma} = \hat{\mathcal{P}} h_{i}^{\dagger} a_{i\bar{\sigma}}^{\dagger} (-\sigma)^{i} e^{i\hat{\Omega}_{i}} \hat{\mathcal{P}}, \tag{9}$$

where $(-\sigma)^{i}$ is a staggered sign factor introduced for convenience and the phase shift operator $\hat{\Omega}_{i}$ will sensitively depend on spin correlations in $|\Phi_{b}\rangle$, as defined by [25]

$$\hat{\Omega}_{i} = \frac{1}{2}\left(\Phi_{i}^{s} - \Phi_{i}^{0}\right), \tag{10}$$

where

$$\Phi_{i}^{s} = \sum_{l \neq i} \theta_{i}(l) \left(\sum_{\sigma} \sigma n_{l\sigma}^{b} \right), \tag{11}$$

and

$$\Phi_{i}^{0} = \sum_{l \neq i} \theta_{i}(l), \tag{12}$$

in which $\theta_{i}(l) = \mathrm{Im}\ln\,(z_{i} - z_{l})$ (z_{i} is the complex coordinate of site i).

Here the $\hat{c}$-operator defined in equation (9) is only a hole-creation operator. The corresponding SU(2) spin operators associated with doped holes are given by

$$\hat{S}_{i}^{az} \equiv \frac{1}{2} \sum_{\sigma} \sigma a_{i\sigma}^{\dagger} a_{i\sigma}, \tag{13}$$

and

$$\hat{S}_{i}^{a+} \equiv -(-1)^{i} a_{i\uparrow}^{\dagger} a_{i\downarrow} \quad \hat{S}_{i}^{a-} \equiv -(-1)^{i} a_{i\downarrow}^{\dagger} a_{i\uparrow}. \tag{14}$$

On the other hand, acting on the neutral spin state in equation (2), the SU(2) spin operators for b-spinons are given as follows:

$$\hat{S}_{i}^{bz} \equiv \frac{1}{2} \sum_{\sigma} \sigma b_{i\sigma}^{\dagger} b_{i\sigma}, \tag{15}$$

and

$$\hat{S}_i^{b+} \equiv (-1)^i b_{i\uparrow}^\dagger b_{i\downarrow} \mathrm{e}^{\mathrm{i}\Phi_i^h}, \quad \hat{S}_i^{b-} \equiv (-1)^i b_{i\downarrow}^\dagger b_{i\uparrow} e^{-\mathrm{i}\Phi_i^h}, \tag{16}$$

from the original Schwinger-boson representation of spin operators with, say, $\hat{S}_i^{b+} \equiv (-1)^i b_{i\uparrow}^\dagger b_{i\downarrow}$. Here

$$\Phi_i^h = \sum_{l\neq i} \theta_i(l) n_l^h, \tag{17}$$

which entangles the b-spinons with the h-holons non-locally. Finally, the total local spin operators are given by

$$\hat{\mathbf{S}}_i = \hat{\mathcal{P}}\left(\hat{\mathbf{S}}_i^b + \hat{\mathbf{S}}_i^a\right)\hat{\mathcal{P}}. \tag{18}$$

Hence, in such a fractionalized doped-Mott-insulator superconductor, the fundamental spin and charge degrees of freedom are described by three subsystems, all possessing ODLROs of their own. The charge carriers as bosons experience a Bose condensation in $|\Phi_h\rangle$. The spins are effectively characterized by a two-fluid state, which shares similarities with other proposals in different contexts [56–59]. Here one type is of local moment character for a Mott insulator: they form a bosonic RVB pairing in $|\Phi_b\rangle$, which can recover the correct description of the antiferromagnetism in the half-filling limit [25, 54], while become a short-range AF (spin liquid) state in the SC phase; the other type is of itinerant character associated with doping: they are fermions forming a BCS-like pairing. Such an SC state will thus show distinctive 'rigidity' associated with different ODLROs hidden in its charge and spin components.

2.2. d-wave superconductivity

2.2.1. SC order parameter. Now let us examine the SC ODLRO in equation (9). The electron singlet pair operator

$$\begin{aligned}\hat{\Delta}_{ij}^{\mathrm{SC}} &\equiv \sum_\sigma \sigma \hat{c}_{i\sigma}\hat{c}_{j\bar{\sigma}} \\ &= \hat{\Delta}_{ij}^0 F_{ij},\end{aligned} \tag{19}$$

in which the second line acts on the fractionalized state (2).

Here the phase part of $\hat{\Delta}_{ij}^{\mathrm{SC}}$ is defined by

$$F_{ij} \equiv \mathrm{e}^{\mathrm{i}(1/2)\left[\Phi_i^s(j)+\Phi_j^s(i)\right]}. \tag{20}$$

The amplitude part of the Cooper pairing is given by

$$\hat{\Delta}_{ij}^0 \equiv \hat{\mathcal{P}}\left(h_i^\dagger h_j^\dagger \hat{\Delta}_{ij}^a e^{-\mathrm{i}\phi_{ij}^0}\right)(-1)^j e^{-\mathrm{i}\Phi_j^0}\hat{\mathcal{P}}, \tag{21}$$

where the pairing operator of the a-spinon is defined by

$$\hat{\Delta}_{ij}^a \equiv \sum_\sigma \sigma a_{i\sigma}^\dagger a_{j\bar{\sigma}}^\dagger. \tag{22}$$

ϕ_{ij}^0 is a non-dynamic π flux per plaquette in a square lattice which is defined as

$$\phi_{ij}^0 \equiv \frac{1}{2}\sum_{l\neq i,j}\Big[\theta_i(l) - \theta_j(l)\Big]. \tag{23}$$

In the following, for simplicity, we shall focus on the case of the NN sites, i.e., $i = NN(j)$, in examining the SC order parameter. In the ground state of equation (2), the holon condensation and a-spinon pairing $\langle \hat{\Delta}_{ij}^{a} e^{-\mathrm{i}\phi_{ij}^0}\rangle \neq 0$ will lead to an s-wave $\langle \hat{\Delta}_{ij}^{0}\rangle$ = constant. The phase factor $(-1)^j e^{-\mathrm{i}\Phi_j^0}$ is a constant independent of site index j, which may be easily shown by noting that $(-1)^j e^{-\mathrm{i}\Phi_j^0} \times \Big[(-1)^i e^{-\mathrm{i}\Phi_i^0}\Big]^* = \mathrm{e}^{\mathrm{i}2\phi_{ij}^0} = 1$ for $i = NN(j)$ with a proper gauge choice of ϕ_{ij}^0 in evaluating $\langle \hat{\Delta}_{ij}^{a} e^{-\mathrm{i}\phi_{ij}^0}\rangle$ (see below).

The SC phase coherence will be determined by

$$\Big\langle F_{ij}\Big\rangle \neq 0. \tag{24}$$

Note that here $\Phi_i^s(j)$ is different from Φ_i^s defined in equation (11) by that $l = j$ should be removed in the summation over site l:

$$\Phi_i^s(j) \equiv \sum_{l\neq i,j}\theta_i(l)\left(\sum_\sigma \sigma n_{l\sigma}^b\right). \tag{25}$$

According to equation (24), each unpaired spinon will contribute to a 2π vortex to the phase of the SC order parameter and their confinement will result in the SC phase coherence.

With the s-wave amplitude $\langle \hat{\Delta}_{ij}^{a} e^{-\mathrm{i}\phi_{ij}^0}\rangle$ and phase coherence $\langle F_{ij}\rangle \neq 0$, the d-wave pairing symmetry of the order parameter can be further identified by comparing the phase difference between the two NN bonds, i.e., i, $i + \hat{x}$ and i, $i + \hat{y}$, as follows. First, one finds

$$F_{ii+\hat{x}}F_{ii+\hat{y}}^* = \left(\mathrm{e}^{\mathrm{i}\sum_\Delta A^s}\right)\mathrm{e}^{\mathrm{i}\theta_{i+\hat{x}}(i+\hat{y})\sum_\sigma \sigma n_{i+\hat{y}\sigma}^b - \theta_{i+\hat{y}}(i+\hat{x})\sum_\sigma \sigma n_{i+\hat{x}\sigma}^b}, \tag{26}$$

where, on the right-hand side of the first line, the subscript Δ denotes a summation over the closed path of the links $(i, i + \hat{x})$, $(i + \hat{x}, i + \hat{y})$ and $(i + \hat{y}, i)$ for the gauge field A^s.

Because of the presence of short-range AF order in $|\mathrm{RVB}\rangle$, such that $\sum_\sigma \sigma n_{i+\hat{y}\sigma}^b \simeq \sum_\sigma \sigma n_{i+\hat{x}\sigma}^b$, one has

$$\mathrm{e}^{\mathrm{i}\theta_{i+\hat{x}}(i+\hat{y})\sum_\sigma \sigma n_{i+\hat{y}\sigma}^b - \theta_{i+\hat{y}}(i+\hat{x})\sum_\sigma \sigma n_{i+\hat{x}\sigma}^b} \simeq \mathrm{e}^{\mathrm{i}\left[\theta_{i+\hat{x}}(i+\hat{y}) - \theta_{i+\hat{y}}(i+\hat{x})\right]\sum_\sigma \sigma n_{i+\hat{y}\sigma}^b} = -1 \tag{27}$$

in the second line of equation (26) (by noting $\theta_{i+\hat{x}}(i + \hat{y}) - \theta_{i+\hat{y}}(i + \hat{x}) = \pm\pi$).

Consequently

$$\Big\langle F_{ii+\hat{x}}F_{ii+\hat{y}}^*\Big\rangle \simeq -\left\langle \mathrm{e}^{\mathrm{i}\sum_\Delta A^s}\right\rangle < 0 \tag{28}$$

(generally the flux produced by $A^s \simeq 0$ within the small loop Δ in the bracket on the right-hand side of equation (27) is vanishingly small and the average of such a phase factor will not change the overall sign) and one thus expects a negative sign difference between $\langle F_{ii+\hat{x}}\rangle$ and $\langle F_{ii+\hat{y}}\rangle$. Therefore, the electron pairing order parameter is generally d-wave like, which originates from the phase string effect *and* short-range AF correlations as was pointed out previously in [55].

2.2.2. Emergent quasiparticle. Just as annihilating a Cooper pair will produce a non-local phase factor F_{ij} in equation (19), the electron decomposition form (9) also implies that injecting

a bare hole into the system will induce a global (non-local) phase shift [25]. It means that a bare hole state created by $\hat{c}$ would be different from a low-lying hole excited state by a phase shift $e^{i\hat{\Omega}_i}$, i.e.

$$\begin{aligned}\hat{c}_{i\sigma}\left|\Psi_{\mathrm{G}}\right\rangle = & \ e^{i\hat{\Omega}_i}\left(\hat{c}_{i\sigma}e^{-i\hat{\Omega}_i}\right)\left|\Psi_{\mathrm{G}}\right\rangle \\ \rightarrow & \ e^{i\hat{\Omega}_i}\times|\mathrm{low}-\mathrm{lying\ excitation}\rangle. \end{aligned} \tag{29}$$

In the SC phase with $\langle e^{i\hat{\Omega}_i}\rangle \neq 0$, based on the electron fractionalization (9), an injected bare hole has an intrinsic fractionalization by

$$\begin{aligned}\hat{c}_{i\sigma}\left|\Psi_{\mathrm{G}}\right\rangle \rightarrow & \ \hat{\mathcal{P}}\left(h_i^{\dagger}\left|\Phi_h\right\rangle\otimes a_{i\bar{\sigma}}^{\dagger}\left|\Phi_a\right\rangle\otimes e^{i\hat{\Omega}_i}\left|\Phi_b\right\rangle\right) \\ \propto & \ \left\langle h_i^{\dagger}\right\rangle\left\langle e^{i\hat{\Omega}_i}\right\rangle\left(\left|\Phi_h\right\rangle\otimes a_{i\bar{\sigma}}^{\dagger}\left|\Phi_a\right\rangle\otimes\left|\Phi_b\right\rangle\right). \end{aligned} \tag{30}$$

Namely, in the presence of holon condensation and phase coherence, a bare hole injected into the SC ground state may decay into an a-spinon excitation. In the single-particle spectral function, such a low-lying sharp mode will appear in the antinodal region of $(\pm\pi, 0)$ and $(0, \pm\pi)$ with a spectral weight proportional to the density of holon condensate $|\langle h_i^{\dagger}\rangle|^2$ and vanishing above T_c when $\langle e^{i\hat{\Omega}_i}\rangle = 0$ [25].

On the other hand, such a fractionalized quasiparticle mode may not be stable in other momentum region. As a matter of fact, a quasiparticle excitation as created by the c-operator can become a stable mode as a bound state of the $h^{\dagger}$, $a^{\dagger}$ and the phase shift factor $e^{i\hat{\Omega}}$ near the nodal region, with a BCS-like nodal energy spectral given by

$$E_{\mathbf{k}} = \sqrt{(\epsilon_{\mathbf{k}}-\mu)^2+\left(\Delta_{\mathbf{k}}\right)^2}\,. \tag{31}$$

Here $\epsilon_{\mathbf{k}}$ is a band spectrum of the original electron with a renormalized hopping integral $t_{\mathrm{eff}} \propto t(1+\delta)/2$ [25], μ is the chemical potential of electrons and a d-wave gap function

$$\Delta_{\mathbf{k}} \equiv 2J\sum_{\mathbf{q}}\left(\cos q_x + \cos q_y\right)\Delta_{\mathbf{k}+\mathbf{q}}^{SC}. \tag{32}$$

Such a quasiparicle mode is presumably coherent in the nodal region at $E_{\mathbf{k}}$ which is lower than the gap of an a-spinon excitation [25]. Therefore the present ground state predicts a dichotomy of quasiparticle excitations between the nodal and antinodal regions.

2.3. Definition of the LPP

Based on the SC ground state (2), its normal state can be defined by switching off the SC coherence in the wave function. Instead of a conventional Fermi liquid state, new states of matter will emerge in the underdoped regime and exhibit pseudogap behaviors as to be explored in the following.

According to equation (19), the SC order parameter

$$\left\langle\hat{\Delta}_{ij}^{\mathrm{SC}}\right\rangle = \left\langle\hat{\Delta}_{ij}^{0}\right\rangle\left\langle e^{i(1/2)\left[\Phi_i^s(j)+\Phi_j^s(i)\right]}\right\rangle, \tag{33}$$

requires the simultaneous formation of both pairing amplitude and phase coherence. Hence, the LPP can be defined as a normal state with either $\langle e^{i(1/2)\left[\Phi_i^s(j)+\Phi_j^s(i)\right]}\rangle = 0$ or $\langle \hat{\Delta}_{ij}^{0} \rangle = 0$. In the following we discuss the two cases separately.

2.3.1. LPP I. The LPP with the pairing amplitude $\langle \hat{\Delta}_{ij}^{0} \rangle \neq 0$ but

$$\left\langle e^{i(1/2)\left[\Phi_i^s(j)+\Phi_j^s(i)\right]} \right\rangle = 0 \tag{34}$$

will be defined as the LPP-I.

An LPP-I state can be naturally obtained as

$$\left|\Phi_{\mathrm{LPP-I}}\right\rangle = \hat{\mathcal{P}}\left(\left|\Phi_h\right\rangle \otimes \left|\Phi_a\right\rangle \otimes \left|\Phi_b\right\rangle\right)_{T>0}. \tag{35}$$

Namely, it is a $T > 0$ version of the SC ground state (2). In other words, the LPP-I state (35) is non-SC at $T > 0$, but reduces to the SC ground state at $T = 0$. Here $|\Phi_h\rangle$ still describes the holon condensation, $|\Phi_a\rangle$ describes the a-spinons in BCS-like pairing and $|\Phi_b\rangle$ describes the b-spinons in bosonic RVB pairing, as to be determined as the mean-field solution below, just like the ground states of equations (3)–(5).

It is straightforward to understand why the SC phase coherence is thermally disordered in equation (35) at $T > 0$. Note that the phase Φ_i^s as defined in equation (11), is composed of vortices locking with single b-spinons, which are thermally excited in $|\Phi_b\rangle$ once $T > 0$, and consequently lead to disordering the phase coherence in equation (34). Later we shall show that the LPP-I state actually has an SC instability at a *finite* T_c. At $T < T_c$, the thermally excited spinons in $|\Phi_b\rangle$ will further form loosely bound pairs, due to the confinement force generated by the vortices locking with the b-spinons. This will go beyond the mean-field solution (35).

2.3.2. LPP II. At $T = 0$, equation (35) naturally recovers SC phase coherence even at the mean-field level. This is due to the fact that all the b-spinons form *short-range* RVB pairs in the ground state, which ensures the phase coherence in equation (2). In order to make phase disordering at $T = 0$, a natural case is that the b-spinons form *long-range* RVB pairings such that free neutral spinons can be spontaneously generated without energy cost, which may happen in the dilute doping boundary with the AFLRO starting to recover [60, 61].

But in the following, we shall consider another case of non-SC ground state at finite doping, in the *absence* of the AFLRO. It corresponds to the case that the pairing amplitude $\langle \hat{\Delta}_{ij}^{0} \rangle$ vanishes, while the phase coherence (24) is still maintained. Here, the pairing of the a-spinons is destroyed in $|\Phi_a\rangle$ to result in a vanishing amplitude of the Cooper pairing at $T = 0$. Such a non-SC ground state may be realized by, say, applying strong magnetic fields in the underdoped regime. It will be called the LPP-II, which is described by

$$\left|\Phi_{\mathrm{LPP-II}}\right\rangle = \hat{\mathcal{P}}\left(\left|\Phi_h\right\rangle \otimes \left|\Phi_a\right\rangle_{\Delta^a=0} \otimes \left|\Phi_b\right\rangle\right). \tag{36}$$

In the ground state of the LPP-II, the holon and b-spinon states, $|\Phi_h\rangle \otimes |\Phi_b\rangle$, will remain essentially the same as in the SC state. But the a-spinon pairing disappears in $|\Phi_a\rangle_{\Delta^a=0}$. Consequently, a free Fermi gas state of $|\Phi_a\rangle_{\Delta^a=0}$ will dominate the low-energy physics in an LPP-II state, as to be detailed below.

2.4. Effective Hamiltonian

In order to get the effective Hamiltonian for the variational wave function (2), we reexpress the original *t*–*J* model in equation (1) in terms of the fractionalization formalism (9) as follows:

$$H_{t-J} = \hat{\mathcal{P}}\left(\tilde{H}_t + \tilde{H}_J\right)\hat{\mathcal{P}}, \tag{37}$$

where [25]

$$\tilde{H}_t = -t\sum_{\langle ij\rangle\sigma}\left(h_i^{\dagger}h_j \mathrm{e}^{\mathrm{i}A_{ij}^{s}+ieA_{ij}^{e}}\right)\left(a_{i\sigma}^{\dagger}a_{j\sigma}e^{-\mathrm{i}\phi_{ij}^{0}}\right) + \mathrm{h.c.}, \tag{38}$$

and

$$\tilde{H}_J = -\frac{J}{2}\sum_{\langle ij\rangle}\left[\left(1-n_i^{h}\right)\left(1-n_j^{h}\right)\left(\hat{\Delta}_{ij}^{s}\right)^{\dagger}\hat{\Delta}_{ij}^{s}\right], \tag{39}$$

in which the bosonic RVB order parameter

$$\hat{\Delta}_{ij}^{s} \equiv \sum_{\sigma} e^{-\mathrm{i}\sigma A_{ij}^{h}} b_{i\sigma}b_{j\bar{\sigma}}. \tag{40}$$

Here, the *h*-holon field formally carries charge $+e$ and couples to the external electromagnetic field A_{ij}^{e} in equation (38). The *h*-holons and *b*-spinons are further mutually coupled to each other via the U(1)⊗U(1) gauge fields, A_{ij}^{s} and A_{ij}^{h}, respectively, in equations (38) and (39), which are topological (mutual Chern–Simons) fields as their gauge-invariant flux strengths in an arbitrary closed (oriented) loop *c* are constrained to the numbers of spinons and holons within the enclosed area Σ_{c}:

$$\sum_{c}A_{ij}^{s} = \pi\sum_{l\in\Sigma_c}\left(n_{l\uparrow}^{b}-n_{l\downarrow}^{b}\right), \tag{41}$$

and

$$\sum_{c}A_{ij}^{h} = \pi\sum_{l\in\Sigma_c}n_{l}^{h}. \tag{42}$$

The origin of such mutual Chern–Simons gauge fields can be traced [25, 27] back to the large gauge (mutual duality) transformation $\mathrm{e}^{\mathrm{i}\hat{\Theta}}$ in equation (86), which precisely incorporates the non-local topological effect of the phase string sign structure hidden in the *t*–*J* model.

Now we introduce the most essential mean-field order parameter [14, 25, 55] for the *t*–*J* model in the representation of (38) and (39):

$$\Delta^{s} \equiv \left\langle \hat{\Delta}_{ij}^{s}\right\rangle, \tag{43}$$

with $i = NN(j)$. It is a bosonic RVB order parameter characterizing the spin singlet background, which reduces to the original Schwinger-boson mean-field order parameter [53] at half-filling, where it well describes quantum AF spin correlations over a wide temperature regime $T_0 \sim J/k_{\mathrm{B}}$ (k_{B} is the Boltzmann coefficient). A finite Δ^{s} will persist into the underdoped regime to define a pseudogap phase known as the UPP [29], which covers both the SC state as well as the LPP states discussed in this work.

According to equation (9), one expects an additional U(1) gauge symmetry between the *h*-holon and *a*-spinon: $h_i^{\dagger}\rightarrow \mathrm{e}^{\mathrm{i}\theta_i}h_i^{\dagger}$ and $a_{i\bar{\sigma}}^{\dagger}\rightarrow e^{-\mathrm{i}\theta_i}a_{i\bar{\sigma}}^{\dagger}$. The presence of this gauge symmetry

implies that a new U(1) gauge field denoted by A_{ij}^{a} is minimally coupled to h-holons and a-spinons via $+1$ and -1 gauge charge, respectively. Because of the RVB pairing of the a-spinons in the ground state of LPP-I as we will discuss later, this gauge field is generally 'Higgsed'. In fact, the projection operator $\hat{\mathcal{P}}$ in equation (37) will result in a general relation [25]: $(\hat{\Delta}_{ij}^{\ a})^{\dagger}\hat{\Delta}_{ij}^{\ a} = n_i^{\ h} n_j^{\ h}(\hat{\Delta}_{ij}^{\ s})^{\dagger}\hat{\Delta}_{ij}^{\ s}$, which ties the RVB pairing of the b-spinons with that of the backflow a-spinons. If one imposes this constraint by introducing a Lagrangian multiplier γ, then a fractionalized effective Hamiltonian can be finally written down as follows:

$$\tilde{H}_{\text{eff}} = \tilde{H}_h + \tilde{H}_s + \tilde{H}_a, \tag{44}$$

with

$$\tilde{H}_h = -t_h \sum_{\langle ij\rangle} h_i^{\dagger} h_j \mathrm{e}^{\mathrm{i}\left(A_{ij}^{s} + eA_{ij}^{e}\right)} + \text{h.c.} + \lambda_h \left(\sum_i h_i^{\dagger} h_i - \delta N \right), \tag{45}$$

$$\tilde{H}_s = -J_s \sum_{\langle ij\rangle} \hat{\Delta}_{ij}^{\ s} + \text{h.c.} + \lambda_b \left(\sum_{i\sigma} b_{i\sigma}^{\dagger} b_{i\sigma} - N \right), \tag{46}$$

$$\tilde{H}_a = -t_a \sum_{\langle ij\rangle\sigma} a_{i\sigma}^{\dagger} a_{j\sigma} e^{-\mathrm{i}\phi_{ij}^{0}} + \text{h.c.} - \gamma \sum_{\langle ij\rangle} \left(\hat{\Delta}_{ij}^{\ a}\right)^{\dagger} \hat{\Delta}_{ij}^{\ a} + \lambda_a \left(\sum_{i\sigma} a_{i\sigma}^{\dagger} a_{i\sigma} - \delta N \right), \tag{47}$$

where

$$J_s = J_{\text{eff}} \Delta^{s}/2, \tag{48}$$

$$J_{\text{eff}} = J(1-\delta)^2 - 2\gamma\delta^2, \tag{49}$$

and $\lambda_{h/b/a}$ represents the chemical potential for the degrees of freedom of the holon/the b-spinon/the backflow a-spinon.

Note that slightly different from [25], here the pairing amplitude of the a-spinons is introduced via the Lagrangian multiplier γ by implementing the average constraint: $(\hat{\Delta}_{ij}^{\ a})^{\dagger}\hat{\Delta}_{ij}^{\ a} \simeq \delta^2 |\Delta^{s}|^2$. The fluctuations going beyond this mean-field equality can be expressed [25] by $J\sum_{\langle ij\rangle}\left(\hat{\mathbf{S}}_i^{b}\cdot\hat{\mathbf{S}}_j^{a} + \hat{\mathbf{S}}_i^{a}\cdot\hat{\mathbf{S}}_j^{b}\right)$. Such a term is to be omitted in the following since we shall be mainly concerned with the mean-field description of the LPP at lower temperature, where at least one of the degrees of freedom is gapped.

The effective coupling constants, t_h and t_a, in $\tilde{H}_{\text{eff}}$ can be determined either variationally or by mean-field approximation, which depend on the bare t, J and the doping concentration, as well as the projection operator $\hat{\mathcal{P}}$. In fact, based on the renormalized Gutzwiller approximation scheme [12], we have approximately doping-independent $t_a \approx t_h \approx t$. But since the basic sign structure of the t–J model has been rigorously captured by the mutual Chern–Simons gauge fields together with the statistics of the constituent particles, the basic physical behavior that we are concerned with at long wavelength, low energy, should not be qualitatively sensitive to the choices of these effective coupling constants.

Therefore, the hidden ODLRO of $\Delta^{s} \neq 0$, without explicitly breaking symmetries, provides the necessary 'rigidity' for the present fractionalization to occur. It defines the underdoped regime of the t–J model and ensures the validity of the above effective Hamiltonian in the so-called UPP (see below).

2.5. Fractionalized states as mean-field solutions

There are three subsystems in the fractionalized ground state (2), which can be determined at the mean-field level by the effective Hamiltonian $\tilde{H}_{\rm eff}$ in equation (44). Since the LPP states defined at the beginning of this section are closely related to these mean-field solutions, in the following we discuss each degree of freedom as well as their interplay one by one.

2.5.1. The holon degree of freedom. The charge degree of freedom is characterized by the bosonic h-holons in $|\Phi_h\rangle$ in equation (3). In both the SC and LPP states, the holons as bosons will further experience a Bose-condensation (i.e., $\langle h\rangle \neq 0$) according to the definition. Such a holon condensation will provide another hidden rigidity in addition to the two-component RVB pairings of the spinons. As a result, the corresponding low-lying excitation will be 'supercurrents' generated from the holon condensate, which lead to the following unique observable consequences.

Based on $\tilde{H}_h$ in equation (45), the supercurrents contributed by the holon condensate are described by a generalized London equation [45–47]

$$\mathbf{J}_h(\mathbf{r}) = \rho_s\left(\nabla\phi + \mathbf{A}^s + e\mathbf{A}^e\right). \tag{50}$$

Here, the superfluid stiffness $\rho_s \equiv \frac{\rho_h}{m_h}$, where $\rho_{\rm h}$ is the superfluid density of the holons with an effective mass $m_h = \frac{\hbar^2}{2a^2 t_h}$ (a is the lattice constant of the square lattice). Reflecting the Mott physics, $\rho_s \to 0$ in the half-filling limit. $\nabla\phi$ ensures the U(1) gauge invariance and satisfies

$$\oint_c \mathrm{d}\mathbf{r}\cdot\nabla\phi = 2\pi\times\text{integer} \tag{51}$$

under the requirement of single-valueness of the holon field.

What is special in equation (50) is the presence of an emergent 'electromagnetic field' vector $\mathbf{A}^s$ in additional to the true external electromagnetic field $\mathbf{A}^e$. Its gauge-invariant field strength is given by equation (41) in a lattice version, or in the following continuum version

$$\oint_c \mathrm{d}\mathbf{r}\cdot\mathbf{A}^s(\mathbf{r}) = \pi\int_{\Sigma_c}\mathrm{d}^2\mathbf{r}\left[n^b_\uparrow(\mathbf{r}) - n^b_\downarrow(\mathbf{r})\right], \tag{52}$$

where the flux of $\mathbf{A}^s$ within an arbitrary loop c on the left-hand side is constrained to the enclosed spinon numbers on the right-hand side, as if a $\pm\pi$ flux-tube is attached to each individual spinon. Here $n^b_{\uparrow,\downarrow}(\mathbf{r})$ denotes the local density of spinons.

Based on equation (50), each unpaired b-spinon will automatically generate a supercurrent vortex, known as a spinon–vortex composite [46, 47], as follows:

$$\oint_c \mathrm{d}\mathbf{r}\cdot\mathbf{J}_h(\mathbf{r}) = \rho_s(\pm\pi)\,, \tag{53}$$

where the loop c encloses a single unpaired spinon (for $\mathbf{A}^e = 0$). In other words, besides conventional minimal 2π-type vortices given in equation (51), the holon condensate can sustain a minimal π-type vortex, in which a b-spinon has to be nucleated at the vortex core. This 'cheap vortex' excitation is one of the most important elementary excitations in the LPP-I.

The supercurrents in equation (50) will generally cost a kinetic energy according to $\tilde{H}_h$, which is given by [45–47]

$$L_h = \frac{1}{2\rho_s}\int \mathrm{d}^2\mathbf{r}\mathbf{J}_h(\mathbf{r})^2. \tag{54}$$

At low concentration of spinon–vortex excitations, such a 'London action' will provide a logarithmic potential between the vortices to make vortex–antivortex binding and thus SC phase coherence. Beyond a critical concentration, the proliferation of these spinon-vortices will effectively screen out the logarithmic interaction in the same fashion as a Kosterlitz–Thouless (KT) type transition, resulting in the intrinsic LPP-I state at $T > T_c$.

Finally, to ensure the holon condensation, the fluctuations of $\mathbf{A}^s$ in equation (50) should be under control. Indeed, according to its definition in equation (41), if the underlying neutral b-spinons are all in short-ranged RVB-pairing in equation (5), $\mathbf{A}^s$ can get cancelled out significantly in favor of Bose condensation of the holons as well as the mean-field decoupling between the holon and b-spinon states in the ground state (2). In the following subsection, we shall see that, as a matter of fact, a short-range RVB state in equation (5) will be in turn caused by the holon condensation self-consistently.

2.5.2. The b-spinon degree of freedom. The building block of the neutral spin background $|\Phi_b\rangle$ is the b-spinon, as shown in equation (5) at $T = 0$. It is governed by $\tilde{H}_s$ in equation (46), in which such neutral spin degrees of freedom are influenced by the doped holes mainly through the lattice gauge field A_{ij}^h as well as J_{eff}. As the holons remain Bose condensed in both the SC and LPP states, its gauge-invariant field strength in equation (42) is basically determined by the local superfluid density ρ_{h} in the following continuum version

$$\oint_c \mathrm{d}\mathbf{r}\cdot\mathbf{A}^h(\mathbf{r}) = \pi\int_{\Sigma_c}\mathrm{d}^2\mathbf{r}\,\rho_h(\mathbf{r}). \tag{55}$$

Consequently, A_{ij}^h becomes a non-dynamic field and $\tilde{H}_s$ in equation (46) can be easily diagonalized, resulting in a mean-field solution [25, 62] $|\Phi_b\rangle$ in equation (5) (more details are presented in appendix B).

At $T = 0$, the mean-field RVB pair amplitude W_{ij} in the state $|\Phi_b\rangle$ of equation (5) has been obtained previously as follows [25]: $|W_{ij}| \propto e^{-|\mathbf{r}_{ij}|^2/2\xi^2}$ if i and j belong to opposite sublattices and $W_{ij} = 0$ for two sites on the same sublattice. Here $\xi = a\sqrt{2/\pi\delta}$ is the corresponding spin–spin correlation length of $|\Phi_b\rangle$. Namely, with a finite hole concentration δ, $|\Phi_b\rangle$ describes a short-range AF state with ξ essentially determined by the average hole–hole distance. It is a 'ghost' (neutral) spin liquid state always pinned at half-filling, with the spin excitation gapped at $E_g \propto \delta J$ in the spin-1 excitation spectrum [25, 62]. Note that ξ diverges at $\delta = 0$, where $W_{ij} \propto 1/|\mathbf{r}_{ij}|^3$ actually becomes quasi long-ranged. Correspondingly $|\Phi_b\rangle$ exhibits an AF long-range order with $|\mathrm{RVB}\rangle = P_s|\Phi_b\rangle$ reproducing [25] a highly accurate variational ground state energy for the t–J model at half-filling [54].

Thus, the ground state (2) reduces to $|\mathrm{RVB}\rangle$ at half-filling, which is of the same form as the LDA wave function [54] and naturally restores the AFLRO state of the Heisenberg model. In the SC regime, $|\mathrm{RVB}\rangle$ becomes a spin liquid state, which in turn ensures the phase coherence of equation (24) for the Cooper pairs moving on the 'vacuum' $|\mathrm{RVB}\rangle$.

As discussed above, $|\Phi_b\rangle$ will remain the same mean-field solution of $\tilde{H}_s$ in the LPP-I at a finite temperature. The gapped thermally excited b-spinons will then decide the basic thermodynamic properties of the LPP-I as to be detailed in section 3.

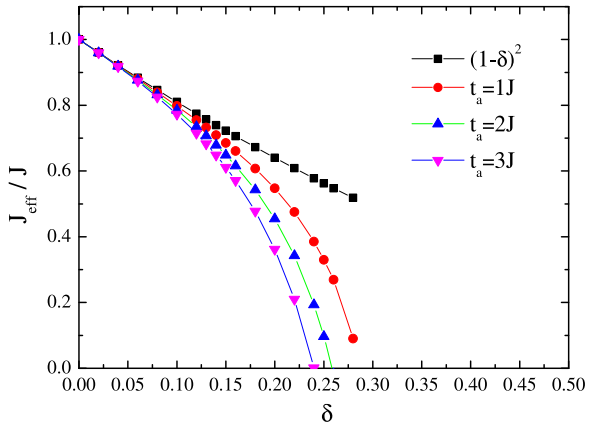

Figure 1. The doping dependence of the renormalized superexchange coupling $J_{\rm eff}$ at zero temperature, obtained by solving the mean-field self-consistent equations at different values of the parameter t_a.

Besides the gauge field A_{ij}^{h}, the doping effect will further influence the neutral spinon background by strongly renormalizing the effective superexchange coupling $J_{\rm eff}$ or $J_{\rm s}$ in $\tilde{H}_s$ (equation (46)). In a previous approach [29], an empirical fitting of $J_{\rm eff}$ as a function of doping has been used. Here $J_{\rm eff}$ can be determined self-consistently with the Lagrangian multiplier γ after the consideration of the backflow a-spinon subsystem (see below). Based on equations (46) and (47), the magnitude and doping dependence of $J_{\rm eff}$ at zero temperature are obtained from the mean-field self-consistent equations as shown in figure 1 at different choices of the parameter t_a. Clearly, $J_{\rm eff}$ decreases monotonically from the bare J with the increase of doping through the interplay between the localized spins (i.e., b-spinons) and the itinerant spins (i.e., the backflow a-spinons). The latter's density is commensurate with the doping concentration. Once $J_{\rm eff}$ is known, the transition temperature T_0 for the bosonic RVB order $\Delta^{\rm s}$ can be determined by

$$T_0 = \frac{J_{\rm eff}}{k_B \ln 3}, \tag{56}$$

which defines the UPP boundary [29][4]. By approximately using $J_{\rm eff}$ calculated at $T = 0$ with $t_a = 2J$ and $J = 120$ meV, the crossover temperature T_0 (as noted before, $\Delta^{\rm s}$ does not correspond to a real symmetry breaking) for the UPP is shown in figure 2 (the curve marked by triangles). The critical doping at vanishing $J_{\rm eff}$ will separate the underdoped regime from the overdoped regime in the present doped Mott insulator. With $\Delta^s = 0$ in the 'overdoped' regime, the electron fractionalization discussed so far will no longer be stable at the mean-field level (a Fermi liquid like state may become stabilized at low temperatures as to be discussed later).

2.5.3. The spinon–vortex composite: an elementary excitation. Although the SC ground state is explicitly fractionalized in terms of three mean-field-type subsystems in equation (2), the aforementioned elementary excitations of two subsystems, i.e., the holon condensate and b-

[4] Note that here the expression of T_0 in equation (56) is slightly different from equation (15) given in [29] because in the latter a slightly different constraint for b-spinons, i.e., $\sum_{i\sigma} b_{i\sigma}^{\dagger} b_{i\sigma} = N(1-\delta)$, is used in the mean-field theory.

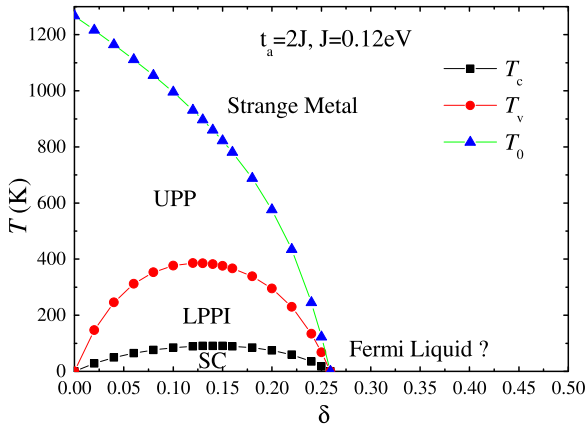

Figure 2. The characteristic temperature scales for the UPP (T_0) and the LPP-I (T_v) as well as the superconducting (SC) phase (T_c) are marked based on the mean field theory of the effective Hamiltonian (44). Note that in this phase diagram, the AFLRO state at half-filling actually can persist [60, 61] over a small but finite doping concentration, which will be further investigated elsewhere. The transport and charge dynamics in the so-called strange metal regime at $T > T_0$ have been previously explored in [63] based on the effective Hamiltonians (45) and (46). In the overdoped regime with $T_0 \to 0$, a possible Fermi-liquid-like instability may occur at low temperature.

spinon RVB background, will be essentially 'entangled' by mutual Chern–Simons gauge fields according to equations (45) and (46), resulting in a unique novel excitation: the spinon–vortex composite [46, 47]. Such a spinon–vortex will be a crucial elementary excitation in characterizing the LPP-I state.

Specifically, the b-spinon excitations are created by breaking up RVB pairs in the mean-field spin liquid state $|\Phi_b\rangle$. They are responsible for the pseudogap behavior in the spin degrees of freedom as will be shown in the next section. According to the generalized London equation (50), a superfluid current vortex of vorticity $\pm\pi$ (cf equation (53)) will be spontaneously generated around each b-spinon excitation. Then, accompanying the b-spinon excitations that are charge-neutral, the associated vortices will play a fundamental role to result in non-Gaussian-type SC fluctuations in the LPP-I state.

Generally, the spinon–vortex excitations are described by $\tilde{H}_h + \tilde{H}_s$ in equations (45) and (46), whose low-energy effective description is a mutual Chern–Simons gauge theory [48, 60], outlined in appendix A. In section 3.1 we will present the basic phenomenology of the LPP-I state governed by such elementary excitations.

In the global phase diagram shown in figure 2, we present a characteristic temperature T_v for the LPP-I, as estimated based on the criterion previously obtained [47], i.e., the holon condensation is totally destroyed when the concentration n_v of excited spinon-vortices becomes equal to the concentration δ of the holons. In the same phase diagram, the SC phase transition T_c is determined [45] by equation (77), at which the free spinon-vortices form bound pairs (cf section IV below), leading to the so-called spinon confinement transition. In table 1, different phases in the global phase diagram of figure 2 are marked by their corresponding 'hidden' ODLROs, where '1' represents a non-zero value of the order parameter and '0' denotes a vanishing value.

Table 1. Hidden ODLROs in the fractionalized degrees of freedom and the corresponding characteristics in the global phase diagram illustrated in figure 2. Here '1' represents a non-zero value of the corresponding order parameter and '0' indicates a zero value.

Phase	Δ^s	$\langle h\rangle$	$\langle F_{ij}\rangle$	$\langle \hat{\Delta}_{ij}^{a}\rangle$
SC	1	1	1	1
LPP-I	1	1	0	1
UPP	1	0	0	1
LPP-II	1	1	1	0
Strange metal	0	0	0	0
Fermi liquid?	0	0	0	0

2.5.4. The backflow a-spinon. An emergent fermionic spinon, i.e., the a-spinon, is an important component of the ground state (2). It is described by $\tilde{H}_a$ in equation (47). The corresponding ground state (4) can be obtained as the mean-field solution with $\langle \hat{\Delta}_{ij}^{a}\rangle \neq 0$, which is similar to a conventional BCS state, but does not carry charge due to its 'Meissner' response to the gauge field as pointed out before. Such an itinerant neutral spinon serves as a spin backflow accompanying the hopping of a holon, which describes the hopping effect, in addition to the phase string effect via the mutual Chern–Simons gauge field A^h, on the spin degrees of freedom [25].

In terms of $\tilde{H}_a$ in equation (47), one may write down the corresponding mean-field Hamiltonian as follows:

$$\begin{aligned}\tilde{H}_a^{\mathrm{MF}} = & -\left(t_a+\gamma\chi^a\right)\sum_{\langle ij\rangle\sigma} e^{-\mathrm{i}\phi_{ij}^0} a_{i\sigma}^{\dagger} a_{j\sigma} - \gamma\left(\Delta^a\right)^*\sum_{\langle ij\rangle\sigma} e^{-\mathrm{i}\phi_{ij}^0}\sigma a_{i\sigma}^{\dagger} a_{j\bar{\sigma}}^{\dagger} + \mathrm{h.c.} \\ & + \gamma\sum_{\langle ij\rangle}\left(\left|\chi^a\right|^2+\left|\Delta^a\right|^2\right) + \lambda_a\left(\sum_{i\sigma} a_{i\sigma}^{\dagger} a_{i\sigma} - \delta N\right),\end{aligned} \tag{57}$$

where $\chi^a \equiv \langle e^{-\mathrm{i}\phi_{ij}^0}\sum_\sigma a_{i\sigma}^{\dagger} a_{j\sigma}\rangle = (\chi^a)^*$ and $\Delta^a \equiv \langle e^{-\mathrm{i}\phi_{ij}^0}\sum_\sigma \sigma a_{i\sigma}^{\dagger} a_{j\bar{\sigma}}^{\dagger}\rangle = (\Delta^a)^*$. Note that in the presence of a π flux depicted by ϕ_{ij}^0, we have found that the s-wave solution of $\langle \hat{\Delta}_{ij}^{a} e^{-\mathrm{i}\phi_{ij}^0}\rangle$ is always more stable than the d-wave one at low doping, in contrast to [25]. Here, for convenience, the gauge of ϕ_{ij}^0 will be chosen such that

$$e^{-\mathrm{i}\phi_{i,i+\hat{x}}^0} = (-1)^{i_y+1}, \quad \mathrm{e}^{-\mathrm{i}\phi_{i,i+\hat{y}}^0} = 1 \tag{58}$$

so there are two sites in a unit cell, and the two sublattices are defined by:

$$i = \begin{cases} A & \text{if} \quad i_y \in \text{odd} \\ B & \text{if} \quad i_y \in \text{even}. \end{cases} \tag{59}$$

By the Fourier transformation:

$$a_{I\sigma}^{A/B} = \frac{1}{\sqrt{N/2}}\sum_{\mathbf{k}} \exp\left(\mathrm{i}\mathbf{k}\cdot\mathbf{R}_I^{A/B}\right) a_{\mathbf{k}\sigma}^{A/B}. \tag{60}$$

Here $\mathbf{R}_I^{A/B}$ represents the position of A/B site in the unit cell 'I'. Then we get

$$\begin{aligned}\tilde{H}_a^{\mathrm{MF}} = & \sum_{\mathbf{k}} \Psi_{\mathbf{k}}^{\dagger} \mathcal{M}_{\mathbf{k}} \Psi_{\mathbf{k}} + \gamma\left(\left|\chi^a\right|^2 + \left|\Delta^a\right|^2\right) \times 2N \\ & + \lambda_a N(1-\delta),\end{aligned} \tag{61}$$

where $\Psi_{\mathbf{k}} \equiv (\Psi_{\mathbf{k}}^A, \Psi_{\mathbf{k}}^B)^T$, $\Psi_{\mathbf{k}}^A \equiv (a_{\mathbf{k}\uparrow}^A, a_{-\mathbf{k}\downarrow}^{A\dagger})^T$, $\Psi_{\mathbf{k}}^B \equiv (a_{\mathbf{k}\uparrow}^B, a_{-\mathbf{k}\downarrow}^{B\dagger})^T$, and the matrix $\mathcal{M}_{\mathbf{k}}$ is

$$\begin{aligned}&\mathcal{M}_{\mathbf{k}} \\ &= \begin{bmatrix} \left(-2\tilde{t}_a \cos k_x a + \lambda_a\right)\sigma_z - 2\gamma\Delta^a \cos k_x a \sigma_x & -2\tilde{t}_a \cos k_y a \sigma_z - 2\gamma\Delta^a \cos k_y a \sigma_x \\ -2\tilde{t}_a \cos k_y a \sigma_z - 2\gamma\Delta^a \cos k_y a \sigma_x & \left(2\tilde{t}_a \cos k_x a + \lambda_a\right)\sigma_z + 2\gamma\Delta^a \cos k_x a \sigma_x \end{bmatrix},\end{aligned} \tag{62}$$

in which $\tilde{t}_a \equiv t_a + \gamma\chi^a$, σ_{x} and σ_{z} are the Pauli matrices. Then it is straightforward to diagonalize the mean-field Hamiltonian equation (61) and obtain the energy dispersions, $\pm\epsilon_{\mathbf{k}1}^a$ and $\pm\epsilon_{\mathbf{k}2}^a$, by

$$\epsilon_{\mathbf{k}1}^a = \sqrt{\left(\xi_{\mathbf{k}1}^a\right)^2 + \left(\Delta_{\mathbf{k}}^a\right)^2}, \quad \epsilon_{\mathbf{k}2}^a = \sqrt{\left(\xi_{\mathbf{k}2}^a\right)^2 + \left(\Delta_{\mathbf{k}}^a\right)^2}, \tag{63}$$

where $\xi_{\mathbf{k}1}^a = -2\tilde{t}_a\sqrt{\cos^2 k_x a + \cos^2 k_y a} + \lambda_a$, $\xi_{\mathbf{k}2}^a = 2\tilde{t}_a\sqrt{\cos^2 k_x a + \cos^2 k_y a} + \lambda_a$, $\Delta_{\mathbf{k}}^a = 2\gamma\Delta^a\sqrt{\cos^2 k_x a + \cos^2 k_y a}$ due to the π-flux. The mean-field free energy reads

$$\begin{aligned}\tilde{F}_a^{\mathrm{MF}} = & -\frac{2}{\beta}\sum_{\mathbf{k},\alpha=1}^{\alpha=2} \ln\left[2\cosh\left(\frac{\beta\epsilon_{\mathbf{k}\alpha}^a}{2}\right)\right] \\ & + \gamma\left(\left|\chi^a\right|^2 + \left|\Delta^a\right|^2\right) \times 2N + \lambda_a N(1-\delta),\end{aligned} \tag{64}$$

where $\beta \equiv 1/k_B T$. Next, by minimizing this mean-field free energy, i.e.

$$\frac{\partial\tilde{F}_a^{\mathrm{MF}}}{\partial\Delta^a} = \frac{\partial\tilde{F}_a^{\mathrm{MF}}}{\partial\chi^a} = \frac{\partial\tilde{F}_a^{\mathrm{MF}}}{\partial\lambda_a} = 0, \tag{65}$$

we get the self-consistent equations:

$$\begin{aligned}&\frac{\gamma}{N}\sum_{\mathbf{k},\alpha=1}^{\alpha=2} A_{\mathbf{k}} B_{\mathbf{k}\alpha} = 1, \\ &\frac{1}{N}\sum_{\mathbf{k},\alpha=1}^{\alpha=2} (-1)^{\alpha}\sqrt{A_{\mathbf{k}}} B_{\mathbf{k}\alpha}\xi_{\mathbf{k}\alpha}^a = 2\chi^a, \\ &\frac{1}{N}\sum_{\mathbf{k},\alpha=1}^{\alpha=2} B_{\mathbf{k}\alpha}\xi_{\mathbf{k}\alpha}^a = 1-\delta\ ,\end{aligned} \tag{66}$$

where $A_{\mathbf{k}}$ and $B_{\mathbf{k}\alpha}$ are defined as

$$A_{\mathbf{k}} \equiv \cos^2 k_x a + \cos^2 k_y a,\ B_{\mathbf{k}\alpha} \equiv \frac{\tanh\left(\beta\epsilon_{\mathbf{k}\alpha}^a/2\right)}{\epsilon_{\mathbf{k}\alpha}^a}. \tag{67}$$

Hence, the backflow a-spinons form a BCS-like pairing state in equation (4). Due to the s-wave nature, in the SC state and LPP-I state, they will not contribute to the low-lying dynamics and thermodynamics significantly except for providing a finite hopping integral t_h for the holons

and renormalizing $J_{\rm eff}$. In other words, the a-spinons constitute the backbone of the unique fractionalization in equations (2) and (35) with a rigidity against the internal U(1) gauge fluctuations. They will also contribute to some unique finite energy dynamics [25], which are not the focus of the present work.

Finally, in the LPP-II state defined at the beginning of this section, a special case has been considered, in which the pairing of the a-spinons gets suppressed, say, in magnetic vortex cores by strong magnetic fields at low temperature. Here $|\Phi_a\rangle$ in equation (4) will reduce to a gapless (Fermi liquid) nomal state as a local LPP-II state defined in equation (36). With vanishing $\Delta^{\rm a}$ inside the vortex core, the 'Meissner' effect or the rigidity due to the a-spinon pairing gets destroyed, while the holons still remain Bose condensed in equation (36). Then the external electromagnetic field A^e will be transferred, via the internal U(1) gauge degree of freedom, from the holon part in equation (45) to solely act on the a-spinons. In other words, the a-spinons will become *charged* and exposed to the probe of external electromagnetic fields in the LPP-II. On the other hand, with the holon condensation, the internal U(1) gauge fluctuations are still 'Higgsed', and thus the a-fermions should be quite coherent without feeling strong gauge scattering.

Therefore, a Fermi liquid composed of the a-spinons will emerge as a new state of matter in the LPP-II, which violates the Luttinger theorem for the original electrons without explicitly breaking a global symmetry. Due to the fractionalization in equation (36), such a new Fermi liquid state is embeded in the backdrop of a SC/pseudogap background where the majority of the spin degrees of freedom are still governed by the b-spinons. Only in the overdoped regime with $J_{\rm eff} \to 0$, would a different non-SC state appear which is beyond the scope of the present work.

3. Phenomenology of LPPs: experimental consequences

In the previous section, we have shown that the low-temperature pseudogap states, i.e., the LPP-I and -II, can be naturally connected to the SC ground state (2) as its normal states. In the following, we shall further study the generic spin and charge properties based on the elementary excitations associated with the fractionalized degrees of freedom in the LPP. On one hand, such anomalous properties can be directly compared to the experimental observations in the cuprates. On the other hand, the unique behaviors of the LPP can reveal the intrinsic non-BCS nature of the SC ground state.

3.1. LPP I

According to the discussion in the previous section, the LPP-I is characterized by three hidden ODLROs, with the SC phase coherence destroyed by the thermally excited spinon-vortices.

In the LPP-I, the spinon–vortex, as a composite of a b-spinon binding with a holon supercurrent vortex, plays the essential role in dictating the basic properties. In the following we first focus on the b-spinon excitations based on the mean-field description, which determine the spin pseudogap phenomenon.

3.1.1. Uniform spin susceptibility and specific heat capacity. The spin uniform susceptibility $\chi_u^{\rm b}$ contributed by the b-spinons is shown in figure 3 at $\delta = 0.1$. It is obtained based on $\tilde{H}_s$ in equation (46) (cf appendix B):

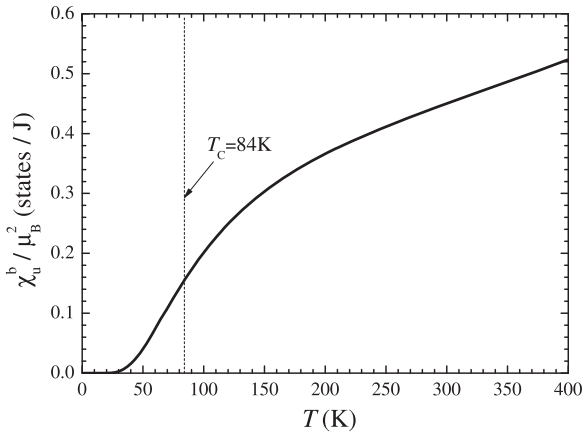

Figure 3. The pseudogap behavior shown by the temperature dependence of the uniform spin susceptibility $\chi_u^{\rm b}$ contributed by the b-spinons at $\delta = 0.1$, obtained with $t_a = 2J$ and $J = 120$ meV. The dashed vertical line marks the characteric temperature $T_{\rm c}$, below which the LPP-I is no longer stable (see text).

$$\chi_u^b = \frac{2\mu_B^2\beta}{N}\sum_m n_B(E_m)\Big[n_B(E_m)+1\Big], \tag{68}$$

where E_m is the eigen energy of the b-spinon excitation, μ_B is the Bohr magneton and $n_{\rm b}(x)$ is the Bose function: $n_{\rm B}(x) \equiv [\exp(\beta x) - 1]^{-1}$.

$\chi_u^{\rm b}$ exhibits a continuous suppression in magnitude with decreasing temperature over the whole regime of the pseudogap phase ($T < T_0$). In particular, the vanishing $\chi_u^{\rm b}$ at low temperature limit is due to a spin gap E_g openned up in the LPP-I. Note that a dotted vertical line in figure 3 marks the SC instability of the LPP-I at $T_{\rm c}$, which is also closely correlated with E_g (see below).

Similar pseudogap behavior is also exhibited in the spin specific heat capacity of the b-spinons, which are shown in figure 4(a) at the same doping concentration as in figure 3. The spin specific heat capacity $\gamma^{\rm b}$ can be expressed by (cf appendix B)

$$\gamma^b = \frac{1}{N}\sum_m \frac{2E_m^2}{k_B T^3} n_{\rm B}(E_m)\Big[n_{\rm B}(E_m)+1\Big]. \tag{69}$$

It has been noted that both $\chi_u^{\rm b}$ (figure 3) and $\gamma^{\rm b}$ (figure 4(a)) exhibit two distinct 'pseudogap' behaviors: a slow general decrease with temperature over a wide range down from T_0 (defining the UPP as shown in figure 2) versus the much steeper suppression at sufficiently low temperatures. The former is due to the formation of the spin RVB pairing (i.e., $\Delta^s \neq 0$), which is already encoded in the mean-field Hamiltonian (46) and is present even at half-filling, indicating the enhanced AF correlations with reducing T. On the other hand, the latter suppression is due to the fact that a true small spin gap E_g opens up in the LPP-I. It is a direct consequence of the charge condensation in the LPP-I, driven through A_{ij}^h in $\tilde{H}_s$. As discussed before, the LPP-I and UPP are distinguished (cf table 1) by that in the latter the bosonic charge carriers (holons) are no longer condensed due to the strong fluctuations of A_{ij}^s in $\tilde{H}_h$, where the generalized London equation (50) is not valid anymore. In figures 3 and 4, the instability of the LPP-I at $T_{\rm c}$ is marked by the dotted vertical line, which is also related to E_g by equation (77) as

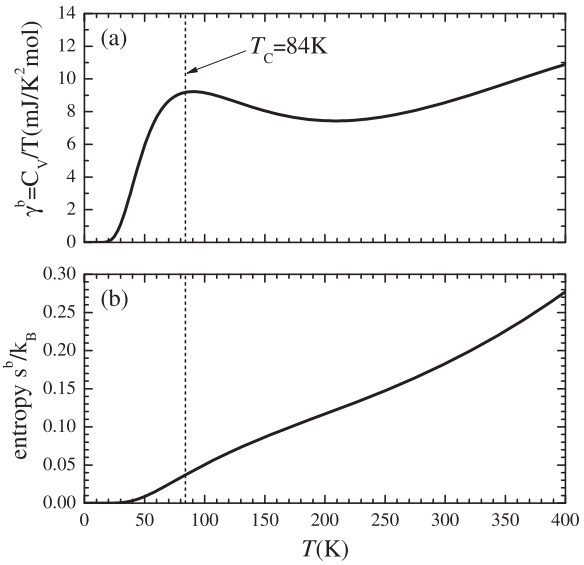

Figure 4. The pseudogap behavior shown by the temperature dependence of (a) the specific heat cofficient γ^b and (b) the corresponding entropy per site s^b contributed by the b-spinon at $\delta = 0.1$, with $t_a = 2J$ and $J = 120$ meV. The dashed vertical line marks the characteric temperature T_c, below which the LPP-I is no longer stable (see text). The magnitude of γ^b is $\sim$10mJ/K^{2}mol in the LPP-I above T_c, which is quite comparable to the experimental data around the optimal doping [35, 36].

to be discussed later. Moreover, the corresponding entropy per site s^b contributed by the b-spinons is shown in figure 4(b), which will be further discussed in section 4.

3.1.2. Longitudinal resistivity. The longitudinal resistivity ρ_e in the LPP-I will not be described by a Drude formula, since the quasiparticle excitations are no longer coherent due to the electron fractionalization [25]. By contrast, the motion of spinon-vortices will generate a distinct dissipation, which is essentially governed by the dynamics of b-spinons [46].

This is a very unique property in the charge transport described by the mutual Chern–Simons theory [60]. In contrast with the Ioffe–Larkin rule in the U(1) gauge theory [8], the so-called non-Ioffe–Larkin rule has been previously obtained [60]:

$$\rho_e(\mathbf{q}, \omega) = \frac{1}{e^2}\left[\sigma_h^{-1}(\mathbf{q}, \omega) + \pi^2\hbar^2\sigma_s(\mathbf{q}, \omega)\right]. \tag{70}$$

Here σ_h represents the longitudinal holon conductivity, σ_s represents the longitudinal b-spinon conductivity with using the SI units: $[\sigma_h] = [\sigma_s] = [\hbar]^{-1}$ (cf appendix C). At any temperature, the static conductivity may be obtained by taking the limits, $\mathbf{q} \to 0$ first and then $\omega \to 0$. Notice that there is no contribution of the backflow a-spinon in equation (70) which is gauge neutral with regard to the mutual Chern–Simons fields and is in a 'BCS' state with regard to the external electromagnetic field.

Due to the condensation of the holon in the LPP-I, we further have $\sigma_h^{-1} = 0$ such that

$$\rho_e = \frac{\pi^2\hbar^2}{e^2}\sigma_s(\mathbf{q} = 0, \omega \to 0), \tag{71}$$

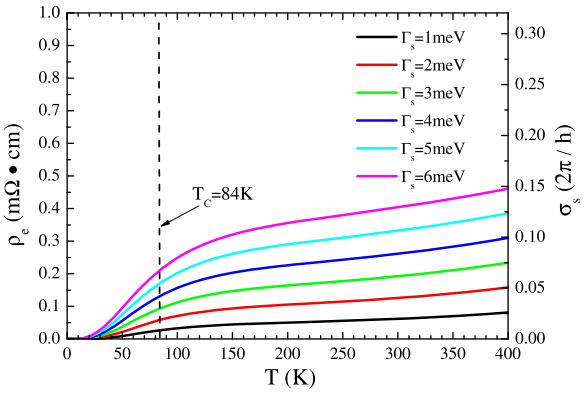

Figure 5. The longitudinal resistivity ρ_e in the LPP-I is determined in a non-Drude formula (71) by the b-spinon conductivity σ_s contributed by the b-spinons. Here $\delta = 0.1$, $t_a = 2J$ and $J = 120$ meV. The parameter $\Gamma_s \ll E_g$ specifies the broadening of the spinon spectrum. In order to make comparision with the cuprates, ρ_e is obtained by multiplying the 2D resistance by a lattice constant along the c axis: $d = 7.7$Å. Here the magnitude of ρ_e above T_c is in the range of $0.1\,\mathrm{m}\Omega\,\mathrm{cm} \sim 0.5\,\mathrm{m}\,\Omega\,\mathrm{cm}$, which is comparable to the experimental data [31].

where $\sigma_s(\mathbf{q}, \omega)$ denotes the b-spinon conductivity. The underlying physical meaning of equation (71) can be understood by that each spinon excitation behaves like a supercurrent vortex, namely, the spinon–vortex [46, 47].

At $T = 0$, $\rho_e \to 0$ with $\sigma_s(\mathbf{q} = 0, \omega = 0) = 0$ as there are no free spinon excitations due to the spin gap E_g. This corresponds to the SC ground state. At a finite T, the thermally excited b-spinons will make σ_s and thus ρ_e in equation (71) finite, meaning that a non-SC phase is naturally realized via the vortex fluctuations associated with the thermal b-spinon excitations. Here the SC phase coherence disappears, unless the excited b-spinons remain 'confined' within $T < T_c$, paired up via the logarithmic interaction introduced in L_h by equation (54) [45]. While a finite T_c will be discussed in the next subsection, in the following we shall simply assume that such an interaction has been screened such that the LPP-I persists over the whole low-temperature regime at $T > 0$, as described by a finite resistivity in equation (71).

In the LPP-I, the b-spinons are deconfined and described by the mean-field $\tilde{H}_s$. The spinon conductivity σ_s can be calculated by the Kubo formula as given by equation (D.20) in appendix D. Figure 5 shows the temperature dependence of σ_s and thus ρ_e at $\delta = 0.1$. (Here ρ_e is obainted by the one-layer resistivity multiplied by the lattice constant along the c axis by $d = 7.7$Å.) The magnitude of the resistivity is quite comparable to the experimental data around the optimal doping [31, 39]. To obtain this result, we fix the parameters at $t_a = 2J$ and $J = 120$ meV. In addition, a small broadening, $\Gamma_s \ll E_g$, is introduced in the spectral function for the mean-field spinon energy level (cf appendix D):

$$A(m, \omega) = \frac{1}{\pi} \frac{\Gamma_s}{(\omega - E_m)^2 + \Gamma_s^2}. \tag{72}$$

In figure 5, ρ_e (σ_s) versus T at different choices of Γ_s are shown.

3.1.3. Nernst effect. Another peculiar transport phenomenon for the LPP-I in the presence of spinon-vortices is a large Nernst signal [46–48]. Physically, the spinon-vortices will move along

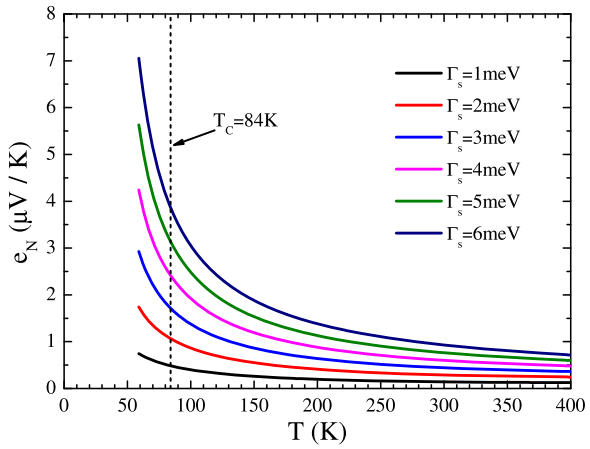

Figure 6. The temperature dependence of the Nernst signal e_N at $\delta = 0.1$ and $B = 20\ T$, with $t_a = 2J$, $J = 120$ meV, and different choices of Γ_s (cf figure 5). Here the magnitude of e_N around $\sim 1\ \mu\text{V K}^{-1}$ above T_c is comparable to the experimental data in the same doping and temperature regime [40, 41, 43].

an applied temperature gradient, driven by the entropy associated with the spin-1/2 free moments centered at vortex cores. Because of the motion of supercurrent vortices, transverse electric voltage will be spontaneously established, if those vortices have a net vorticity polarized by the perpendicular magnetic filed, which is known as the Nernst effect.

The Nernst effect is therefore an important signature of the LPP-I state due to the presence of spontaneous spinon-vortices [46], which are thermally excited to destroy the SC phase coherence.

Based on the generalized London equation (50), the Nernst coefficient can be expressed by [46]

$$e_N = \alpha_{xy}\rho_e, \tag{73}$$

where

$$\alpha_{xy} = \frac{Bs_\phi}{\phi_0^2 n_v}. \tag{74}$$

Here B denotes the magnetic field strength, $\phi_0 \equiv hc/2e$ is the flux quantum. The 'transport entropy' s_ϕ comes from the the spinon with a free S = 1/2 moment locking with a supercurrent vortex, given by [47]

$$s_\phi = k_B \left\{ \ln\left[2\cosh\left(\beta\mu_B B\right)\right] - \beta\mu_B B \tanh\left(\beta\mu_B B\right)\right\}. \tag{75}$$

The temperature dependence of Nernst signal e_N at $\delta = 0.1$ and $B = 20\ T$ is shown in figure 6. The magnitude of the Nernst coefficient is quantitatively comparable to the experimental data [40, 41, 43].

3.1.4. Spin Hall effect. As a unique signature for the presence of spinon-vortices, a dissipationless spin Hall effect has been predicted [64] for the LPP-I. Physically, spinon–vortex composites can be driven to move by a perpendicular electric field E_y^e and consequently a spin current J_x^s is simultaneously generated if the free moments at the centers of the vortex cores are

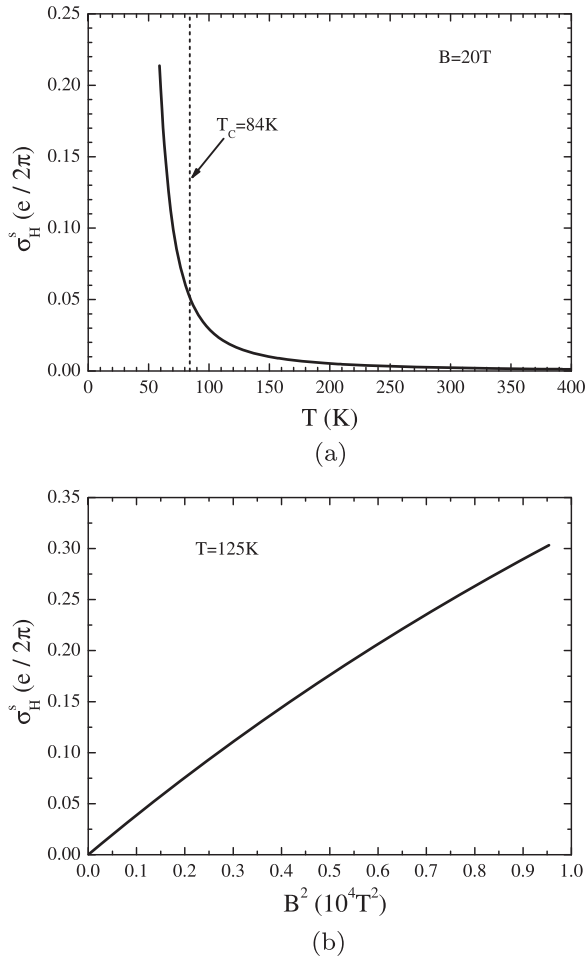

Figure 7. The temperature (a) and magnetic field (b) dependence of the spin Hall conductivity σ_H^s at $\delta = 0.1$ with $t_a = 2J$ and $J = 120$ meV.

polarized by an external magnetic field b along the $\hat{z}$-axis. A quantitative prediction is $J_x^s = \sigma_H^s E_y^e$, and based on the generalized London equation (50), the spin Hall conductivity can be expressed by [64]

$$\sigma_H^s = \frac{\hbar \chi_u^b}{g\mu_B}\left(\frac{B}{n_v \phi_0}\right)^2, \tag{76}$$

where, the electron g-factor ≈2. The temperature dependence σ_H^s at $\delta = 0.1$ and $B = 20\ T$ is shown in figure 7(a) and the magnetic field dependence at $T = 125\ K > T_c$ is shown in figure 7(b).

3.1.5. SC instability. As shown in figures 5 and 6, the resistivity is quickly diminished with a divergent Nernst signal as $T \to 0$. Such strong non-Gaussian fluctuations in the LPP-I indicate that an intrinsic SC instability may happen at a low temperature.

Note that equation (50) will reduce to a conventional London-like equation describing a SC state if the internal gauge field $\mathbf{A}^s = 0$. In fact, the London action (54) will provide a logarithmic 'confinement force' for the spinon-vortices to pair up at sufficiently low

temperatures to make $\mathbf{A}^s = 0$ at a large length scale. In other words, the true SC condensation is signalled by the confinement of the b-spinons below a critical temperature T_c. Correspondingly, the phase coherence condition in equation (33) is achieved by a vortex–antivortex binding associated with the spinon confinement transition [45–48]. Here one may see the similarity of the current SC phase transition to the traditional KT transition for a 2D superfluid system [65]. As a matter of fact, the T_c formula can be similarly obtained as follows [45] (cf appendix B)

$$k_B T_c = \frac{E_g}{\kappa}, \tag{77}$$

which is controlled by the spin gap E_g with $\kappa \sim 6$. Such a T_c has been marked by vertical dotted lines in figures 5 and 6.

There are several important remarks that concern the nature of the SC phase as given below. First, an ordinary KT transition is driven by conventional 2π vortices (in the field of ϕ of equation (50)). But here the spinon-vortices involve a vorticity π. In contrast to the former, the spinon–vortex–antivortex pairing will not annihilate each other at $T = 0$, because of the conserved spinon numbers. Instead, they form tightly bound vortex–antivortex pairs as the corresponding spinons form RVB pairs in the ground state (5). When single spinon-vortices are created by breaking up such RVB pairs, the minimal excitation energy essentially measures the b-spinon excitation spectrum (without creating vortices as they are already there in the ground state). Therefore, the spinon–vortex excitation is a 'cheap vortex' not only because of the lower vorticity (π instead of 2π), but most importantly because of the fact that it still exists as a vortex in the ground state. One can estimate [47] the lowest energy to create a pair of spinon-vortices from the ground state as $\simeq E_g \propto \delta J_{\rm eff}$, which controls T_c as shown in equation (77).

Second, corresponding to a finite spin gap E_g, the RVB background of the b-spinons has a finite spin–spin correlation length $\xi \sim 1/\sqrt{\delta}$. The contribution of those RVB paired spinons to $\mathbf{A}^s$ is thus cancelled out at a length scale larger than ξ or in other words the ground state is a spin liquid state and at the same time a vortex–antivortex binding state. Thermally excited spinons in the LPP-I are spontaneous vortices which form a vortex liquid [46–48]. Then, when temperature is substantially lower than E_g/k_B, only a very small amount of free spinons get thermally excited. One finds that the logarithmic potential provided by the London action (54) is sufficient to cause the confinement of these free spinons and make $\oint_c d\mathbf{r} \cdot (\nabla\phi + \mathbf{A}^s) = 0$ at length scales much larger than that of the spin–spin correlation. Subsequently the SC phase coherence is realized. So the precursor of superconductivity in the LPP-I is closely related to the opening up of the spin gap E_g, concomitant with the holon condensation at T_v in figure 2.

Third, an important distinction of equation (50) from the conventional London equation for a BCS superconductor is that a charge $+e$ instead of $2e$ condensate couples to the electromagnetic field $\mathbf{A}^e$ here. Nevertheless, a minimal magnetic flux quantization at $hc/2e$ can be still expected in the present SC state [46]. This is because the flux quantization condition is now given by

$$\oint_c d\mathbf{r} \cdot \mathbf{J}_h = \oint_c dr \cdot \left(\nabla\phi + \mathbf{A}^s + e\mathbf{A}^e\right) = 0, \tag{78}$$

where, according to equation (52), the unit flux quanta of $\oint_c d\mathbf{r} \cdot \mathbf{A}^e = \pm\, hc/2e$ (with restoring the full units of $\hbar$ and c) can be still found. The prediction [46] is that each magnetic vortex core must trap a non-trivial zero mode: a free spinon, which leads to $\oint_c d\mathbf{r} \cdot \mathbf{A}^s = \pm\, \pi$.

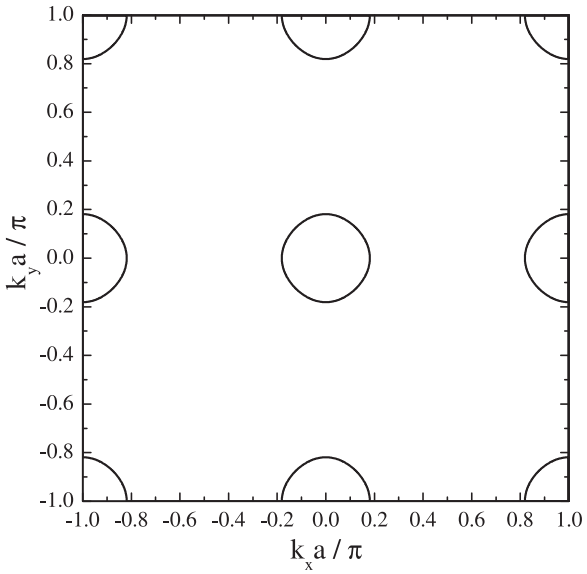

Figure 8. The emergent Fermi surfaces of the a-spinon in the LPP-II state, after the superconducting state is suppressed with $\Delta^a = 0$ by a strong external magnetic field ($\delta = 0.1$ and $t_a = 2J$).

3.2. LPP II: quantum oscillation

In the above, we have discussed a novel magnetic vortex core which traps a free b-spinon. At a finite temperature close to T_c, spinon-vortices are more easily nucleated by external magnetic fields, right before the thermally excited spinon-vortices destroy the SC phase coherence. On the other hand, at sufficiently low-temperature: $T \ll E_g/k_B$, such a novel magnetic vortex may no longer energetically competitive, due to the minimal spin gap E_g in breaking up an RVB pair, as compared to a conventional magnetic vortex of quantization $hc/2e$, which is realized by that the external magnetic field penetrates the a-spinon subsystem, thanks to the U(1) gauge freedom associated with decomposing the holon and a-spinon as discussed in section 3. In this case, one finds $\Delta^a \rightarrow 0$ at the vortex magnetic core. Namely, inside the vortex core, one has the gapless a-spinon state, i.e., the LPP-II state, instead of trapping a b-spinon in the LPP-I. In other words, two types of magnetic vortices are predicted for this non-BCS superconductor, which may appear in different temperatures and doping regimes.

In the following, instead of justifying its stability, we shall explore the LPP-II state at the mean-field level, which is obtained by turning off Δ^{a} in $\tilde{H}_a^{\mathrm{MF}}$ (equation (57)). In the SC phase, the a-spinons are fully gapped due to its s-wave pairing with $\Delta^a \neq 0$. On the other hand, according to the definition of the LPP-II in equation (36), the a-spinons in the LPP-II state will become gapless with $\Delta^a = 0$. As shown in figure 8, the mean-field state discussed in section 2.5.4 for the a-spinons will reduce to two Fermi pockets around both (0, 0) and (π, 0).

As a matter of fact, in the LPP-II, the DC transport will be solely carried by the a-spinons that are in a Landau–Fermi liquid state. To see this clearly, let us first generalize the non-Ioffe–Larkin rule in equation (70), in which the contribution from the a-spinon is not included because the latter remains in the BCS-pairing state in the LPP-I.

By taking account of the internal gauge field A^a that is minimally coupled to h-holons and s-spinons via $+1$ and -1 gauge charge respectively, we may end up with a general combination

rule for transport properties (cf appendix E):

$$\rho_e(\mathbf{q},\,\omega) = \frac{1}{e^2}\Big[\sigma_a^{-1}(\mathbf{q},\,\omega) + \sigma_h^{-1}(\mathbf{q},\,\omega) + \pi^2\hbar^2\sigma_s(\mathbf{q},\,\omega)\Big]. \tag{79}$$

Here $\sigma_{\rm a}$ is the 'conductivity' of a-spinons and this general formula is applicable to the whole phase diagram where $J_{\rm eff}$ is non-zero. In the LPP-I, $\sigma_a = \infty$ (at zero frequency) due to BCS pairings of a-spinons, which sends equation (79) back to equation (70). Then the aforementioned U(1) gauge field between a-spinons and h-holons is 'Higgsed'. However, in the LPP-II considered here, a-spinons form a Fermi liquid since strong magnetic field breaks the BCS-pairing of a-spinons inside the vortex core with a finite σ_a. By noting that $\sigma_h = \infty$ (holon condensation) and $\sigma_s = 0$ (RVB pairing of the b-spinons) at zero temperature, we have the following formula specific to the LPP-II:

$$\rho_e = \frac{1}{e^2}\sigma_a^{-1} \tag{80}$$

which indicates that the physical electric transport is merely carried by the a-spinons, Namely, the a-spinons become charged with vanishing $\Delta^{\rm a}$ inside the vortex core, which is already discussed in section 2.5.4.

At low doping $\delta \ll 1$, the energy dispersion around $\mathbf{k} = (0,\,0)$ at the lower band may be approximatively expressed by

$$\xi^a_{\mathbf{k}1} \approx \frac{\hbar^2\mathbf{k}^2}{2m_a} - \lambda_a{}', \tag{81}$$

where the effective mass $m_a \equiv \hbar^2/\sqrt{2}\,a^2\tilde{t}_a$ and the effective chemical potential is defined by $\lambda_a{}' \equiv 2\sqrt{2}\,\tilde{t}_a - \lambda_a$.

In 2D, the density of states of the a-spinon is given by $N_a = m_a/2\pi\hbar^2$, and the chemical potential can be obtained from the constraint $\left\langle \sum_{I\sigma} a^{A\dagger}_{I\sigma} a^{A}_{I\sigma} + \sum_{I\sigma} a^{B\dagger}_{I\sigma} a^{B}_{I\sigma} \right\rangle = \delta N$. The 2D area (denoted by $\mathcal{A}_{\rm F}$) expanded by each of the Fermi surfaces of the a-spinons is

$$\mathcal{A}_{\rm F} = \mathcal{A}_{\rm BZ}\frac{\delta}{4} \tag{82}$$

where $\mathcal{A}_{\rm BZ} = 4\pi^2/a^2$. According to the Onsager relation, we can get the frequency of the quantum oscillation by

$$\mathcal{F} = \frac{\hbar c}{2\pi e}\mathcal{A}_{\rm F} = \frac{\phi_0}{2a^2}\delta. \tag{83}$$

If $\delta = 0.1$, we find the frequency of the quantum oscillation $\mathcal{F} \approx 697\,{\rm T}$, and the magnitude of this result is comparable to the experimental data: $\mathcal{F}^{\rm exp} = (530 \pm 20)\,{\rm T}$ at the similar empiracal doping concentration [49].

Apart from the quantum oscillation, the emergent Fermi pockets of the a-spinons provide a qualitative explanation for some other experimental consequences in strong magnetic fields at low temperature. For example, a finite value of uniform susceptibility and a linear-T specific heat capacity at extremely low temperature corresponding to the finite density of state on the Fermi energy have been observed after the SC state is suppressed by a strong external magnetic field [50, 51]. In particular, here the low-T Fermi liquid behavior has been found to be embedded in a larger pseudogap background presumably from the b-spinons in the present approach.

4. Critical comparison with the 'plain vanilla' RVB theory and the slave-boson approach

In the previous sections, we have explored the LPP physics based on the SC ground state ansatz (2) and shown a systematic agreement with the experiments in the cuprate. In particular, we have emphasized throughout the paper that the pseudog physics has truthfully reflected the non-BCS nature of the SC ground state. Since such SC ground state as well as the LPP are obtained based on the t–J model, it is very meaningful to make a critical comparison of the present approach with the standard 'plain vanilla' RVB theory and the slave-boson approach to the same model.

4.1. The present SC ground state versus the Gutzwiller projected BCS state

An alternative ansatz for the SC ground state of the t–J model is the well-known Gutzwiller projected BCS state proposed by Anderson [9]. It can be written as

$$\left|\Psi_{\mathrm{RVB}}\right\rangle = \hat{P}_{\mathrm{G}}\left|d-\mathrm{BCS}\right\rangle, \tag{84}$$

where $|\mathrm{d\text{-}BCS}\rangle$ denotes an ordinary d-wave BCS state and $\hat{P}_{\mathrm{G}}$ is a Gutzwiller projection operator enforcing the following no double occupancy constraint $\hat{n}_i \leqslant 1$. Because of $\hat{P}_{\mathrm{G}}$, the Cooper pairing in $|\mathrm{d\text{-}BCS}\rangle$ reduces to the neutralized RVB pairing [9] at half-filling.

Mathematically, in order to implement the no-double-occupancy constraint, there are many choices for a formal fractionalization. For example, one may treat the spinon as fermion and the holon as boson, in the so-called slave-boson decomposition [8, 66, 67] or vice versa in the so-called slave-fermion decomposition [53, 68, 69]. In the literature, a popular electron fractionalization is the slave-boson approach [8], in which the ground state is obtained with the neutral *fermionic* spinons forming a d-wave RVB state $|\Phi_f\rangle$ and the holons being in a Bose-condensed state $|\Phi_h\rangle$. Namely,

$$\left|\Psi_{\mathrm{RVB}}\right\rangle = C\hat{P}_{hf}\left(\left|\Phi_h\right\rangle \otimes \left|\Phi_f\right\rangle\right), \tag{85}$$

where the Gutzwiller projection operator $\hat{P}_{hf}$ implements the no double occupancy constraint on the bosonic holons and fermionic spinons: $n_i^h + n_i^f = 1$. In fact, $\hat{P}_{hf}\left|\Phi_f\right\rangle \rightarrow \hat{P}_{\mathrm{G}}\left|d-\mathrm{BCS}\right\rangle$ due to the holon condensation, where the RVB and Cooper pairings are not explicitly distinguished at finite doping. Namely, the 'plain vanilla' RVB state is equivalent to the electron fractionalization in the slave-boson formalism.

By contrast, the present SC ansatz (2) involves a quite different electron fractionalization from the usual slave-particle decomposition, which is given in equation (9). One may reexpress equation (2) in terms of equation (9) as follows:

$$\left|\Psi_{\mathrm{G}}\right\rangle = \mathrm{e}^{\mathrm{i}\hat{\Theta}}\left|\Phi_{\mathrm{G}}\right\rangle, \tag{86}$$

where

$$\hat{\Theta} \equiv -\sum_i n_i^h \hat{\Omega}_i, \tag{87}$$

and

$$\left|\Phi_{\rm G}\right\rangle \equiv C\ \exp\left(\sum_{ij} g_{ij}\hat{c}_{i\uparrow}\hat{c}_{j\downarrow}\right)|{\rm RVB}\rangle, \tag{88}$$

with $g_{ij} = (-1)^i \tilde{g}_{ij}$.

Firstly, the bosonic RVB state $|{\rm RVB}\rangle$ in equation (88) remains always at half-filling, describing an RVB or neutral spin liquid background defined in equation (7). Then doped holes are further introduced by the electron annihilating operators, which are paired in equation (88) with a pairing amplitude g_{ij}. Namely, the neutral RVB pairing of spins and the charge BCS-pairing are explicitly separated in equation (88), in contrast to the Gutzwiller projected BCS ground state in equation (84) where the two are not distinguished. Here the no double occupancy constraint is automatically enforced so long as $|{\rm RVB}\rangle$ remains singly occupied.

Secondly, the above distinction between the neutral spins and doped holes makes the definition of a non-local unitary transformation, i.e., $e^{i\hat{\Theta}}$ in equation (86), possible. Here with n_i^h in equation (87) as the hole number operator, each doped hole will generate a non-local *phase shift* $\hat{\Omega}_i$ via equation (87). It plays a crucial role to regulate the singular sign structure of the *t*–*J* model at the lattice scale by transforming it into a large-scale geometric/topological phase shift [27, 28]. Consequently, in a mean-field-type treatment of $|\Phi_{\rm G}\rangle$, this important sign structure can be accurately retained.

To understand such a sign structure, now imagine a given hole moving through a closed path and then count the geometric (Berry) phase contributed by $e^{i\hat{\Theta}}$ in equation (86). Note that, combining with Φ_i^0, an ↑-spin will contribute totally nothing but a ↓-spin will give rise to a 2π phase vortex to $\hat{\Omega}_i$ in equation (10). It is then easy to see that all the ↓ spins enclosed within the loop will each contribute to a $\pm 2\pi$ phase while 0 outside the loop. As for those ↓-spins right on the hole loop, meaning those ↓ spins exchanged with the hole during the thinking experiment, will each give rise to a $\pm\pi$ phase, resulting in a string of signs as a non-trivial geometric phase given by

$$e^{i\hat{\Theta}} \to (-1)^{N_h^{\downarrow}(c)} \times e^{i\hat{\Theta}}, \tag{89}$$

in which $N_h^{\downarrow}(c)$ denotes the total number of hole-↓-spin exchanges on the closed loop c. Furthermore, additional statistical signs can be contributed by the fermionic $\hat{c}$-operators in $|\Phi_{\rm G}\rangle$ of equation (88), i.e.,

$$\left|\Phi_{\rm G}\right\rangle \to (-1)^{N_h^h(c)} \times \left|\Phi_{\rm G}\right\rangle, \tag{90}$$

where $N_h^h(c)$ denotes the total number of exchanges between doped holes in a set of close path c by which holes are exchanged. Here it is noted that the half-filled spin background $|{\rm RVB}\rangle$ will not produce any statistical signs as described by *bosonic* wavefunction in equation (5). Then, combining equations (89) and (90), the so-called phase string sign structure [27, 28] of the *t*–*J* model is precisely reproduced, which are both geometric and topological as identified previously for arbitrary doping and temperature on a bipartite lattice of any dimensions [28], and in this procedure, the unitary-transformed representation $|\Phi_{\rm G}\rangle$ becomes 'smooth' and

locally singular-free to allow for a mean-field treatment. The unitary transformation in equation (86) may be also called a mutual-duality transformation.

4.2. Nature of Mott physics

It has been well appreciated that the no double occupancy constraint in the t–J model is a key ingredient reflecting the basic physics of the doped Mott insulator. In both the 'plain vanilla' RVB state in equations (84) and (85) and the present one in equations (86) and (2), such a constraint is implemented.

However, in the present approach, the Mott physics for a doped Mott insulator further means the following: the fermionic statistical sign structure of the original electrons has been completely changed to the phase string sign structure in the restricted Hilbert space. In short, the Mott physics should be understood as the no double occupancy constraint plus a non-Fermi sign structure.

Consequently, different from the conventional slave-particle scheme, a new electron fractionalization (9) is found. It leads to the construction of the new class of ground state (2) or (86), which precisely satisfies the sign structure. Moreover, the ground state naturally reduces to the most accurate AF state $|\mathrm{RVB}\rangle$ (the LDA state [54]) at half-filling, where long-range AF correlations are correctly recovered. It serves as an appropriate starting point to understand the doping problem.

The two SC ansatz states in equations (84) and (86), or in the fractionalized form equations (85) and (2), are further distinguished by their dramatically different elemenary excitations.

First of all, they both have bosonic holons in a Bose-condensed state $|\Phi_h\rangle$. However, other than this similarity, the rest is so drastically distinct. In particular, the holon condensation in the slave-boson approach as given in equation (85) should be destroyed at T_c to result in a pseudogap phase. Correspondingly the charge degree of freedom is characterized by a Bose metal [8, 70] with uncondensed bosonic holons in $|\Phi_h\rangle$. By contrast, the holons still remain condensed in the present LPP.

Secondly, in the slave-boson approach, the low-lying elementary excitation (nodal Bogoliubov quasiparticle) is reduced to the f-spinon in the SC phase based on the fermionic RVB state in $|\Phi_f\rangle$ of equation (85). By contrast, the Bogoliubov quasiparicle is emergent as a bound state of the holon and a-spinon as given in section 2.2.2.

Thirdly, in contrast to the fermionic f-spinon in the slave-boson approach, there is a bosonic b-spinon in the present state. Such a neutral spin excitation will always induce a supercurrent vortex to form a spinon–vortex composite. In particular, a pair of them form the so-called spin-roton excitation as the unique excitation in the SC phase, which determines the SC phase transition at a lower temperature than the characteristic temperature of the holon condensation, which decides the LPP-I.

Fourthly, a gapped fermionic a-spinon is predicted as a unique feature of the two-component RVB state of the present case in equation (2) or (86), associated with the spin backflow of the holon hopping. In contrast to the f-spinon in the slave-boson approach where its number is equal to the total electron number, here the number of a-spinons is commensurate with the holon number.

Finally, we point out that although the LPP state has been expressed in terms of three fractionalized particles, the holon, b-spinon and a-spinon, the total entropy contributed by them

is not expected to be overcounted because each of the subsystems is in an ODLRO state, in which the entropy is generally suppressed. However, all the hidden ODLROs, including the holon condensation and the spinon pairs, will be melted in a high-temperature (strange metal) regime at $T > T_0$ (cf figure 2). It is a natural question if the existence of the three fractionalized particles would lead to an overcounting of the entropy, as is the case for the slave-boson mean-field state in comparison with the high-temperature series expansion results [71]. We emphasize that different from a Fermi liquid description of the *f*-spinons in the slave-boson approach, which is the main reason for the overcounting there, in the present case, the *b*-spinons in equation (88) are localized free moments satisfying the Curie–Weiss behavior with the entropy per site bounded by $k_B \ln 2$. Furthermore, the contribution of the *a*-spinons as fermions will be reduced by $\delta/(1-\delta)$ as compared to that of the *f*-spinons in the slave-boson approach. However, a more quantitative comparison to the high-temperature numerical results for the *t*–*J* model, which is beyond the present scope focusing on the LPP at low temperatures, will be discussed elsewhere.

4.3. Nature of pseudogap physics

As noted already, the pseudogap phase in the slave-boson approach is basically a Bose metal, with the fermionic spinons remaining in a d-wave RVB state. Namely, the fermionic RVB state $|\Phi_f\rangle$ in equation (85) is responsible for the pseudogap properties in the spin degrees of freedom, which should be similar to the SC state. On the other hand, the charge dynamics will be governed by the uncondensed holons of a Bose metal, which is subject to further investigations. Eventually, a strange metal phase is expected at higher temperatures/doping concentrations, after the RVB pairing is destroyed.

In the present approach, the UPP, which has not been discussed in the present paper, corresponds to the above pseudogap phase in the slave-boson approach. However, the LPP, which has been explored in this work, has no correspondence in the slave-boson approach. Specifically, the holons still remain Bose-condensed in the LPP. In fact, all the three subsystems in equation (35) are still in the ODLROs in the LPP-I as emphasized before.

Hence, the LPP is something unique, as predicted by the SC ground state (2) or (86)–(88). Here it is distinguished from the SC state by the thermally excited unpaired spinon–vortex excitations in the LPP-I or by vanishing RVB pairing of the *a*-spinons in the LPP-II.

The most essential characteristic of the LPP is the opening up of a doping-dependent spin gap E_g as indicated in figures 3 and 4 at low temperature. Such a spin gap in $|\Phi_b\rangle$ of equation (35) describes a spin liquid with a finite spin correlation length. In particular, E_g vanishes in the dilute hole limit to result in an AFLRO in $|\mathrm{RVB}\rangle = \hat{\mathcal{P}}\,|\Phi_b\rangle$, which at half-filling becomes a very accurate variational ground state of the Heisenberg Hamiltonian.

A finite E_g is caused by the holon condensation via the mutual Chern–Simons gauge field $\mathbf{A}^h$, which is due to the altered statistical sign structure of the *t*–*J* model explicitly formulated in equation (86) as the mutual duality transformation. This is absent in the slave-boson approach.

Another important prediction of the mutual duality is that the neutral spin excitations (*b*-spinons) will strongly affect the charge condensate by creating supercurrent vortices, i.e., the spinon-vortices. In the LPP-I, their thermal excitations disorder the SC phase coherence, resulting in a large non-Drude resistivity and strong Nernst effect as the characteristics of non-Gaussian-like SC fluctuations as illustrated in figures 5 and 6. Note the reduction of the

resistivity in figure 5 and the divergence of the Nernst siginal in figure 6 as the temperature is lowered below E_g/k_B.

Eventually, at a sufficiently low temperature, with the thermally excited spinon-vortices greatly reduced in number due to the spin gap, the confinement of them into vortex–antivortex pairs becomes possible, in a fashion of KT-type transition, which results in a true SC phase coherence below $T \leqslant T_c$ as controled by the spin gap E_g in equation (77).

Finally, even in the zero temperature, a non-SC state, i.e., an LPP-II state, can be also realized when the BCS-like pairing of the *a*-spinons in $|\Phi_a\rangle$ of equation (4) is destroyed, say, by strong magnetic fields *before* the occurrence of phase disordering by thermally excited *b*-spinon–vortex excitations (cf table 1). Correspondingly, the Cooper pairing amplitude vanishes to result in a non-SC normal state, at least in the magnetic vortex core region.

Note that the *a*-spinon in the SC phase and the LPP-I state is charge-neutral as well as gauge-neutral, immune from the mutual Chern–Simons gauge force between the *b*-spinons and holons. But once in the LPP-II, with vanishing RVB pairing, the *a*-spinons will carry the full charge as the holons are still condensed, and the corresponding Fermi pockets of the fermionic *a*-spinons give rise to quantum oscillation, a Pauli susceptibility and a linear-*T* specific heat just like in a typical Fermi liquid.

Therefore, the LPP studied in the present work is a unique low-temperature pseudogap phenomenon, which is not present in the simple slave-boson approach. It is physically related to the rigidity associated with the hidden ODLROs in the fractionalized degrees of freedom in the ground state. Due to the sharp distinctions between the SC ground states as well as elementary excitations in the slave-boson and present fractionalization schemes, the nature of the low-temperature pseudogap physics thus differ strongly.

5. Discussion

In this work, we have intended to understand the pseudogap phenomenon observed in the cuprate superconductor through a model study. Namely, we have explored the so-called low-temperature pseudogap state in a doped Mott insulator based on the *t*–*J* model. In addition to its intrinsic SC instability, such a state exhibits a systematic pseudogap behavior in both spin and charge degrees of freedom, as shown by the uniform spin susceptibility, specific heat, non-Drude resistivity, Nernst effect, as well as the quantum oscillation in strong magnetic fields, etc. These anomalous properties in the low-temperature pseudogap phase are found to be qualitatively consistent with the experimental measurements in the cuprates.

As an important lesson that we learned from this study, these pseudogap properties unveil the most essential non-BCS nature of the SC state. Namely, they are hidden in the SC ground state as an integral part of it, and start to explicitly manifest once the SC ODLRO is turned off by temperature, magnetic field or other means. In a conventional BCS state, the SC ground state is composed of the Bloch electrons filling up a Fermi sea and forming the Cooper pairs close to the Fermi energy. The non-BCS superconductivity means that the normal state is no longer a conventional Fermi liquid dominated by the low-lying Landau quasiparticle excitations. In the present ground state, while the Cooper pairing of the electrons as the true SC ODLRO is still present, the quantum numbers of the electrons are in fact all fractionalized with a peculiar composite structure. The Mottness, namely, the strong on-site Coulomb repulsion, is the fundamental driving force behind such fractionalization.

We point out that the SC state of the doped Mott insulator, either the Gutzwille-projected BCS state (84) or the present one (86), is a natural ground state of pure electrons, without needing an extra 'gluon'-like phonon in a BCS superconductor. In the latter, a Fermi liquid state is a natural ground state for purely electronic degrees of freedom, which sets in as a 'normal state' once the Cooper pairing mediated by phonons is turned off [5].

In this sense, the proposed SC ground states in the doped Mott insulator are the stable infrared fixed point states that essentially control all the anomalous pseudogap behaviors at finite temperature above T_c. In other words, the basic correlations exhibited in high-temperature 'normal state' regimes are already encoded in the ground states, in the specific forms of electron fractionalization, as shown in equations (85) and (2), respectively.

With both the SC ansatz states mentioned above satisfying the no double occupancy constraint of the t–J model, the present one has two advantages: (I) it naturally reduces to the most accurate AF state (the LDA state) at half-filling; (II) it precisely keeps track of the altered statistic signs (known as the phase string sign structure) of the t–J model at finite doping. Then the specific fractionalization dictated by the new sign structure of the doped Mott insulator leads to a peculiar non-BCS SC ground state, which manifects the unique low-temperature pseudogap behavior once the SC coherence is removed. The pseudogap phenomenon here is thus physically related to the rigidity associated with the hidden ODLROs in the fractionalized degrees of freedom in the ground state, which does not necessarily correspond to any explicit spontaneous symmetry breaking.

Several important issues have not yet been explored in the present work. (A) How does the SC ODLRO terminate at a finite but sufficiently low doping? We have pointed out that at half-filling, the AF LDA state is naturally recovered as the ground state. But the AF order is expected to persist over some very dilute amount of doped holes before the SC ground state sets in at zero temperature. The doped holes have been predicted to be self-localized [72, 73] in this non-SC regime, and the transitions between the AF and SC phases have been studied in the framework of mutual Chern–Simons gauge theory [60]. But more detailed properties like the fate of the backflow fermionic spinons remain to be investigated and compared with experiment. (B) In the overdoped regime with vanishing $J_{\rm eff}$, does the low-temperature pseudogap state eventually become unstable towards a Fermi liquid state at low temperatures? If the answer is yes, then how can this picture can be reconciled with the non-Fermi sign structure of the t–J model [17]? If the answer is no, then what would be the non-Fermi-liquid ground state after the superconductivity disappears beyond a sufficiently large doping in the t–J model? In particular, if the Fermi liquid behavior of the overdoped cuprates corresponds to the so-called Mott collapse [17, 18] due to a finite Hubbard U, it should be already beyond the scope of the t–J model. Then what would be the reliable doping regime that the t–J model may be relevant to the experiment? (C) In the low-temperature pseudogap state studied in the present work, the detailed behavior of the quasiparticle excitation remains to be investigated. With the vanishing d-wave order parameter due to the proliferation of the spinon-vortices, the quasiparticles are expected [74] to become incoherent with the Fermi arc feature observed in the ARPES spectral function [2]. But a quantitative study is still absent here. (D) The fractionalized structure should not only be exhibited in the pseudogap phase, but also be present in the SC state, for instance, in the normal core of the magnetic vortex as well as in the excitation spectra of the bulk. While the fate of the backflow fermionic spinons in a normal core has been studied as in the LPP-II, the contributions of such spinon excitations, emerging at finite doping, to the dynamic spin susceptibility function and the single-particle spectral function need further study.

Acknowledgements

We would like to thank V N Muthukumar, X-L Qi, S-P Kou, C-S Tian and L Zhang for the previous collaborations and discussions. This work was supported by the NBRPC grant no. 2010CB923003. P Y is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development & Innovation.

Appendix A. Compact mutual Chern–Simons gauge theory description

In the main text, the LPP state has been discussed in terms of the effective Hamiltonian (44) at the mean-field level. To go beyond the mean-field treatment [48], an effective topological field theory description known as the compact mutual Chern–Simons gauge theory [60] will be needed. In the following, such a field-theory description for the low-energy physics of equations (45) and (46) is presented.

The holons and b-spinons are generally coupled via a pair of mutual Chern–Simons gauge fields in the effective Hamiltonians in equations (45) and (46), which represent the most fundamental force originated from the phase string sign structure of the doped Mott insulator as emphasized in the main text. In the following Lagrangian formulation [60], these two degrees of freedom can be expressed generally as

$$\mathcal{L}_h\left[h^\dagger,\, h;\, A^s_\mu\right] = h^\dagger_I\left(\mathrm{d}_0 - iA^s_0 + \lambda_h\right)h_I - t_h h^\dagger_I \sum_\alpha \mathrm{e}^{\mathrm{i}A^s_\alpha} h_{I-\alpha}, \tag{A.1}$$

$$\mathcal{L}_s\left[b^\dagger,\, b;\, A^h_\mu\right] = \sum_\sigma b^\dagger_{i\sigma}\left(\mathrm{d}_0 - i\sigma A^h_0 + \lambda_b\right)b_{i\sigma} - J_s\sum_{\alpha\sigma}\left(\mathrm{e}^{\mathrm{i}\sigma A^h_{i+\hat{\alpha},i}} b^\dagger_{i+\hat{\alpha}\sigma} b^\dagger_{i\bar{\sigma}} + \mathrm{h.c.}\right), \tag{A.2}$$

$$\mathcal{L}_{CS}\left[A^s_\mu,\, \mathcal{N}^s_\mu;\, A^h_\lambda,\, \mathcal{N}^h_\lambda\right] = \frac{\mathrm{i}}{\pi}\epsilon^{\mu\nu\lambda}\left(A^s_\mu - 2\pi\mathcal{N}^s_\mu\right)d_\nu\left(A^h_\lambda - 2\pi\mathcal{N}^h_\lambda\right), \tag{A.3}$$

where the bosonic matter field h/b representing the holon/spinon field is coupled to the statistic gauge field A^s_μ/A^h_μ, respectively. The two gauge fields are entangled by the mutual-Chern–Simons term $\mathcal{L}_{CS}$, where $\mathcal{N}^s_\mu/\mathcal{N}^h_\mu$ is an integer field in the compact mutual-Chern–Simons theory. Here $A^{s,\,h}_\alpha$ are the compact link variables with $A^{s,\,h}_\alpha \in [-\pi,\, \pi)$ and $A^{s,\,h}_0 \in \mathbb{R}$. Notice that α/β in the superscript or subscript represents the direction, i.e., $\hat{x}$ or $\hat{y}$ in real space. The total Lagrangian is apparently invariant under the local $U(1) \otimes U(1)$ gauge transformation up to *mod* 2π.

The LPP state has been defined by the holon condensation $\langle h\rangle \neq 0$. Define $h_I = \sqrt{n_h}\,\mathrm{e}^{\mathrm{i}\phi(I)}$ with $n_h = \rho_h a^2$. One has

$$\mathcal{L}_h = -iA^s_0 n_h + t_h n_h\left(A^s_\alpha\right)^2. \tag{A.4}$$

Here the change of variable: $A^s_\mu \to A^s_\mu + d_\mu\phi$, has been made and after this shift, the vector field $A^s_\mu \in \mathbb{R}$ instead of $[-\pi,\, \pi)$, which ensures the correctness of subsequent Gaussian integral.

Summing up the intger field $\mathcal{N}^s_0$, one gets the quantization of the gauge field $\frac{\mathrm{i}}{\pi}\epsilon^{0\alpha\beta}d_\alpha(A^h_\beta - 2\pi\mathcal{N}^h_\beta)$ and the following effective Lagrangian

$$\mathcal{L}_{eff} = \mathcal{L}_s + t_h n_h \left(A_\alpha^s\right)^2 + \frac{\mathrm{i}}{\pi}\left[\epsilon^{0\alpha\beta}\mathrm{d}_\alpha\left(A_\beta^h - 2\pi\mathcal{N}_\beta^h\right) - \pi n_h\right]A_0^s + \frac{\mathrm{i}}{\pi}\epsilon^{\alpha\mu\nu}\left(A_\alpha^s - 2\pi\mathcal{N}_\alpha^s\right)\mathrm{d}_\mu\left(A_\nu^h - 2\pi\mathcal{N}_\nu^h\right). \tag{A.5}$$

In the LPP, we may separate the spatial components of the gauge field A_α^h into two parts: $A_\alpha^h = \bar{A}_\alpha^h + \delta A_\alpha^h$, where $\bar{A}_\alpha^h$ depicts the background component which satisfies

$$\epsilon^{0\alpha\beta}\mathrm{d}_\alpha \bar{A}_\beta^h = \pi n_h, \tag{A.6}$$

and δA_α^h represents the fluctuating component. Next we combine the original gauge field with corresponding integer field: $\tilde{A}_\alpha^h = \delta A_\alpha^h - 2\pi\mathcal{N}_\alpha^h$ and $\tilde{A}_0^h = A_0^h - 2\pi\mathcal{N}_0^h$, and thus the new defined field $\tilde{A}_\mu^h \in \mathbb{R}$, which ensures the correctness of subsequent Gaussian integral in the resulting effective Lagrangian

$$\mathcal{L}_{eff} = \mathcal{L}_s + t_h n_h \left(A_\alpha^s\right)^2 + \frac{\mathrm{i}}{\pi}A_0^s\epsilon^{0\alpha\beta}\mathrm{d}_\alpha\tilde{A}_\beta^h + \frac{\mathrm{i}}{\pi}\epsilon^{\alpha\mu\nu}\left(A_\alpha^s - 2\pi\mathcal{N}_\alpha^s\right)\mathrm{d}_\mu\tilde{A}_\nu^h. \tag{A.7}$$

After intergrating out A_μ^s, one obtains

$$\mathcal{L}_{\mathrm{eff}} = \mathcal{L}_s + \frac{1}{4\pi^2 n_h t_h}\left(\epsilon^{\alpha\mu\nu}\mathrm{d}_\mu\tilde{A}_\nu^h\right)^2 - 2\mathrm{i}\epsilon^{\alpha\mu\nu}\mathcal{N}_\alpha^s\mathrm{d}_\mu\tilde{A}_\nu^h, \tag{A.8}$$

where a constraint on $\tilde{A}_\alpha^h$ is $\epsilon^{0\alpha\beta}d_\alpha\tilde{A}_\beta^h = 0$.

Here we may ignore the imaginary time-dependence of $\tilde{A}_\alpha^h$. Then we arrive at

$$\mathcal{L}_{\mathrm{eff}} = \frac{1}{4\pi^2 n_h t_h}\left(\epsilon^{\alpha\beta 0}\mathrm{d}_\beta\tilde{A}_0^h\right)^2 + \mathcal{L}_s\left(\tilde{A}_0^h = 0; \tilde{A}_\alpha^h\right) - \mathrm{i}\tilde{A}_0^h\left(n_s + 2\epsilon^{0\alpha\beta}\mathrm{d}_\alpha\mathcal{N}_\beta^s\right), \tag{A.9}$$

where $n_s(\mathbf{r}_i) \equiv \sum_\sigma \sigma n_{i\sigma}^b = \sum_\sigma \sigma b_{i\sigma}^\dagger b_{i\sigma}$. Define the spinon vorticity

$$q_{sv}(\mathbf{r}_i) \equiv n_s(\mathbf{r}_i) + 2\epsilon^{0\alpha\beta}\mathrm{d}_\alpha\mathcal{N}_\beta^s(\mathbf{r}_i). \tag{A.10}$$

After integrating out $\tilde{A}_0^h$, we finally obtain the effective action which works in both the LPP and SC phases:

$$S_{\mathrm{eff}} = S_s\left(\tilde{A}_\mu^h = 0\right) + S_{\mathrm{sv}}, \tag{A.11}$$

where $S_s(\tilde{A}_\mu^h = 0) = \sum_x \mathcal{L}_s(\tilde{A}_\mu^h = 0)$ and the second term is the effective action for the interacting spinon vortices:

$$S_{\mathrm{sv}} = \int_0^{1/k_B T}\mathrm{d}\tau\frac{\pi n_h t_h}{2}\ln\left(\frac{R}{a}\right)\left[\sum_i q_{\mathrm{sv}}(\mathbf{r}_i)\right]^2 - \int_0^{1/k_B T}\mathrm{d}\tau\frac{\pi n_h t_h}{2}\sum_{i\neq j} q_{\mathrm{sv}}(\mathbf{r}_i)\ln\left(\frac{\left|\mathbf{r}_i - \mathbf{r}_j\right|}{a}\right)q_{\mathrm{sv}}(\mathbf{r}_j). \tag{A.12}$$

The intrinsic SC instability of the LPP-I state at low temperature will be further discussed based on the above mutual Chern–Simons gauge theory formulation in appendix B.

Appendix B. Pseudogap behavior and SC instability

The link variables, A_{ij}^{s} and A_{ij}^{h}, can be regarded as mediating the mutual statistics coupling between the charge and spin degrees of freedom, i.e., the 'mutual semion statistics' entanglement. But in the ground state, these two subsystems can be effectively 'disentangled', with each in a condensed state with an ODLRO of its own. Consequently, the fluctuations around such a 'saddle-point' state will become well controlled and well behaved, just as in all the conventional systems with an ODLRO where the emergent 'rigidity' suppresses the violent fluctuations of the many-body degrees of freedom.

B.1. Mean-field solution of the b-spinon

The Hamiltonian $\tilde{H}_s$ in equation (46) with the mean field approximation can be diagonalized by a Bogoliubov transformation [55, 62]

$$\mathrm{b}_{i\sigma} = \sum_{m}\left(u_m\gamma_{m\sigma} - v_m\gamma_{m\bar{\sigma}}^{\dagger}\right)w_{m\sigma}(\mathrm{i}), \tag{B.1}$$

which results in

$$\tilde{H}_s^{\mathrm{MF}} = \sum_{m\sigma}E_m\gamma_{m\sigma}^{\dagger}\gamma_{m\sigma} + J_{\mathrm{eff}}\left|\Delta^s\right|^2 N - 2\lambda_b N + \sum_{m}E_m, \tag{B.2}$$

with

$$u_m = \frac{1}{\sqrt{2}}\sqrt{\frac{\lambda}{E_m}+1}\,, \qquad v_m = \mathrm{sgn}\left(\xi_m\right)\frac{1}{\sqrt{2}}\sqrt{\frac{\lambda}{E_m}-1}\,, \tag{B.3}$$

and

$$E_m = \sqrt{\lambda_b^2 - \xi_m^2}\,, \tag{B.4}$$

where ξ_{m} is the eigenvalue of $w_{m\sigma}(i)$ which is the eigenstate of the equation

$$\xi_m w_{m\sigma}(i) = -J_s\sum_{j=\mathrm{NN}(i)}\mathrm{e}^{\mathrm{i}\sigma A_{ij}^{h}}w_{m\sigma}(j), \tag{B.5}$$

with $J_s \equiv J_{\mathrm{eff}}\Delta^s/2$. In determining the spinon excitation spectrum E_m, the self-consistent conditions: $\langle\hat{\Delta}_{ij}^{s}\rangle = \Delta^s$ and $\left\langle\sum_{i\sigma}b_{i\sigma}^{\dagger}b_{i\sigma}\right\rangle = N$ have to be used.

Based on the above mean-field solution, the ground state of the b-spinons, which is effectively decoupled from the other fractional particles, is given in equation (5), in which one finds that the RVB pairing amplitude $W_{ij} = 0$ if both i and j belong to the same sublattice and decays exponentially at large spatial separations for opposite sublattice sites i and j [25]

$$\left|W_{ij}\right| \propto e^{-\frac{\left|\mathbf{r}_{ij}\right|^2}{2\xi^2}}. \tag{B.6}$$

Here $\mathbf{r}_{ij}$ is the spatial distance and ξ is the characteristic pair size determined by the doping concentration: $\xi = a\sqrt{2/\pi\delta}$. As pointed out above, once the b-spinons are all short-range paired up in $|\Phi_b\rangle$, the fluctuations of A_{ij}^{s} would become negligible and the two subsystems of the holons and b-spinons are decoupled as depicted by $|\Phi_h\rangle\otimes|\Phi_b\rangle$. Note that at half-filling where

$\rho_h = 0$, $\tilde{H}_s$ in equation (46) reduces to the Schwinger-boson mean-field Hamiltonian, which well captures the AF correlations including the long-range AF order at $T = 0$ [25, 54].

B.2. Spin degrees of freedom

Uniform spin susceptibility is defined by

$$\chi_u^b = \left.\frac{M}{NB}\right|_{B\to 0}, \tag{B.7}$$

where B is the strength of external magnetic field and the b-spinon magnetization is given by

$$M = \mu_B \int \mathrm{d}\omega n_B(\omega) \sum_{m\sigma} \sigma A_\sigma(m, \omega), \tag{B.8}$$

with $A_\sigma(m, \omega) = \frac{1}{\pi}\frac{\Gamma_s}{(\omega - E_{m\sigma})^2 + \Gamma_s^2}$ and $E_{m\sigma} = E_m - \sigma\mu_B B$. One gets

$$\chi_u^b \approx \frac{2\mu_B^2\beta}{N}\sum_m n_B(E_m)\left[n_B(E_m) + 1\right] \tag{B.9}$$

at $\Gamma_s \ll E_g$.

Furthermore, the contribution of the b-spinon to the specific heat can be evaluated as follows:

$$\gamma^b \equiv \frac{C_V}{T} = -\frac{1}{N}\frac{\partial^2}{\partial T^2}\tilde{F}_s^{\mathrm{MF}}, \tag{B.10}$$

where the mean field free energy $\tilde{F}_s^{\mathrm{MF}}$ is given by

$$\tilde{F}_s^{\mathrm{MF}} = \frac{2}{\beta}\sum_m \ln\left(1 - e^{-\beta E_m}\right) + J_{\mathrm{eff}}\left|\Delta^s\right|^2 N - 2\lambda_b N + \sum_m E_m. \tag{B.11}$$

Consequently

$$\begin{aligned}\gamma^b = & \frac{2}{N}\int \mathrm{d}\omega \frac{\omega^2}{k_B T^3} n_B(\omega)\left[n_B(\omega) + 1\right]\sum_m A(m, \omega) \\ \approx & \frac{2}{N}\sum_m \frac{E_m^2}{k_B T^3} n_B(E_m)\left[n_B(E_m) + 1\right].\end{aligned} \tag{B.12}$$

B.3. SC instability

In the LPP, the holons are always condensed such that the holon conductivity $\sigma_h = 0$. Thus, according to the non-Ioffe–Larkin rule in equation (70), the dc resistivity is essentially determined by the b-spinon conductivity σ_s, which will be evaluated based on the above mean-field solution in appendix D.

Due to the interaction term in equation (A.12), the residual interaction between the b-spinon-vortices can lead to their 'confinement' at low temperatures. Namely, at a sufficiently low temperature, the dilute spinon-vortices and spinon-antivortices tend to form bound pairs, which then leads to the true SC phase coherence as discussed in the main text. Indeed, such a spinon confinement will make the b-spinon conductivity σ_s vanishing such that

$$\rho_e = \frac{\pi^2\hbar^2}{e^2}\sigma_s(\mathbf{q}=0,\,\omega\to 0) \propto \sigma_s(\mathbf{q}=0,\,\omega\to 0) = 0. \tag{B.13}$$

In the following, we briefly discuss such a KT-like vortex–antivortex binding transition based on the mutual Chern–Simons gauge theory outlined in appendix A.

Note that in the above mean-field solution, an eigen state of the b-spinon has a wave-packet wave function like [45]

$$\left|w_{m\sigma}(\mathbf{r}_i)\right|^2 \simeq \frac{a^2}{2\pi a_c^2}\exp\left(-\frac{\left|\mathbf{r}_i-\mathbf{R}_m\right|}{2a_c^2}\right), \tag{B.14}$$

with a 'cyclotron length' $a_c \equiv a/\sqrt{\pi\delta}$. Here the degenerate levels are labeled by the coordinates $\mathbf{R}_m$, the centers of the spinon wave packet, which form a von Neumann lattice with a lattice constant $\xi_0 = \sqrt{2\pi}\,a_c$. So the effective Lagrangian S_{eff} in appendix A can be further simplified as

$$\begin{aligned} S_{\text{eff}} &\simeq S_s\left(\tilde{A}_\mu^h = 0\right) + S_{\text{sv}} \\ &\simeq \int_0^{1/k_BT} \mathrm{d}\tau \frac{E_{\text{g}}}{2}\sum_m \left|q_{\text{sv}}(\mathbf{R}_m)\right| \\ &\quad - \int_0^{1/k_BT} \mathrm{d}\tau \frac{\pi n_h t_h}{2}\sum_{\mathbf{R}_m\neq\mathbf{R}_{m'}} q_{\text{sv}}(\mathbf{R}_m)\ln\left(\frac{\left|\mathbf{R}_m-\mathbf{R}_{m'}\right|}{\xi_0}\right)q_{\text{sv}}(\mathbf{R}_{m'}), \end{aligned} \tag{B.15}$$

where E_g denotes the minimal energy gap of the spin-1 excitation. Finally by noting $n_h = \rho_h a^2$, $t_h = \hbar^2/2m_h a^2$ and the spin stiffness $\rho_s \equiv \rho_h/m_h$, the effective action is rewritten as

$$S_{\text{eff}} = \frac{E_{\text{g}}}{2k_BT}\sum_m \left|q_{\text{sv}}(\mathbf{R}_m)\right| - \frac{\pi}{4}\frac{\rho_s}{k_BT}\sum_{m\neq m'} q_{\text{sv}}(\mathbf{R}_m)\ln\left(\frac{\left|\mathbf{R}_m-\mathbf{R}_{m'}\right|}{\xi_0}\right)q_{\text{sv}}(\mathbf{R}_{m'}). \tag{B.16}$$

Next one can take a standard procedure in dealing with a conventional KT transition [45, 65]. Define the reduced stiffness $K = \rho_s/k_BT$ and the effective fugacity of each spinon vortex $y \equiv e^{-E_g/2k_BT}$.

Finally, the differential renormalization group (RG) equations are obtained by [45]

$$\frac{\mathrm{d}K^{-1}}{\mathrm{d}l} = g^2\pi^3 y^2 + O\left(y^4\right), \tag{B.17}$$

$$\frac{\mathrm{d}y}{\mathrm{d}l} = \left(2-\frac{\pi}{4}K\right)y + O\left(y^3\right), \tag{B.18}$$

where, $g = 4$ is degeneracy for each site $\mathbf{R}_m$ in the von Neumann lattice due to the time reversal and bipartite lattice symmetries. It is easy to find that the two RG equations above could also be obtained if we replace (K, y) by $(K/4, gy)$ in the RG equations of conventional KT transition. Here K is replaced by $K/4$ because the unit vorticity of each spinon vortex is π instead of 2π of a conventional vortex; and y is replaced by gy because of the g degeneracy for each site $\mathbf{R}_m$ in the von Neumann lattice.

Therefore, the RG flow in the present case is the same as in a conventional KT transition if we replace (K,y) in the latter by $(K/4,\ gy)$. The RG equations result in a fixed point at $K^* = 8/\pi$

and $y^* = 0$, and there is a separatrix passing through the critical point $K^{-1} = \pi/8$, $y(l) = 0$. Points above this separatrix flow towards large values of K^{-1} and large values of y, in other words, toward the phase with unbound spinon vortices. Points exactly on the separatrix with $K^{-1} < \pi/8$ flow to the critical point. The starting point of flows is on the line $y = \exp(-E_g/2k_BT) = \exp(-E_gK/2\rho_s)$. The transition temperature is then determined by the intersection of this line with the separatrix. The flow for $T < T_c$ is towards the line $y = 0$, which means no spinon excitation is allowed below T_c, which corresponds to the spinon confinement in the SC phase.

$$\left\langle q_{sv}(\mathbf{R}_m) q_{sv}(\mathbf{R}_{m'}) \right\rangle = -2\, g^2 y^2 \left[\frac{|\mathbf{R}_m - \mathbf{R}_{m'}|}{\xi_0} \right]^{-\pi K/2} \to 0, \tag{B.19}$$

where, the fugacity y is renormalized to zero when $T < T_c$. The transition temperature T_c determined [45] is given in equation (77) in the main text.

Appendix C. Definitions and the units of spinon and holon conductivities

In the compact mutual Chern–Simons theory, the b-spinon/holon/a-spinon conductivity [60] is defined by

$$\mathbf{J}_s = \sigma_s \mathbf{E}_h, \tag{C.1}$$

$$\mathbf{J}_h = \sigma_h (\mathbf{E}_s + e\mathbf{E}_e + \mathbf{E}_a), \tag{C.2}$$

$$\mathbf{J}_a = -\sigma_a \mathbf{E}_a. \tag{C.3}$$

Here $\sigma_{s/h/a}$ represents the b-spinon/h-holon/a-spinon conductivity, and the vector field $\mathbf{J}_{s/h/a}$ represents the corresponding current. Equation (C.1) is due to b-spinon, which is coupled to $\mathbf{A}^h$ in equations (40) and (46). Equation (C.2) is due to h-holon, which is charged and also coupled to $\mathbf{A}^s$ in equation (45), moreover, there is another internal U(1) gauge field $\mathbf{A}^a$ between h-holon and a-spinon in equation (9) (cf the discussion in the paragraph just above equation (44)), i.e. $\mathbf{E}_a$ in equation (C.2); in equation (9), the U(1) gauge charges of a-spinon and h-holon should have opposite sign, and this is the origin of minus sign in equation (C.3).

On the one hand, we have the conservation equation of $j_s^\mu = (\rho_{\text{spin}}, \mathbf{J}_s)$:

$$\nabla \cdot \mathbf{J}_s + \partial_t \rho_{\text{spin}} = 0, \tag{C.4}$$

where $[\rho_{\text{spin}}] = [L]^{-2}$. Thus $[\mathbf{J}_a] = [L]^{-1}[T]^{-1}$. On the other hand, $\mathbf{E}_h = \partial_t \mathbf{A}^h$ and $\nabla \times \mathbf{A}^h = \pi\hbar\rho_h$. $[\rho_h] = [L]^{-2}$, thus $[\mathbf{A}^h] = [\hbar][L]^{-1}$ and $[\mathbf{E}_h] = [\mathbf{A}^h][T]^{-1} = [\hbar][L]^{-1}[T]^{-1}$.

Finally, one gets the unit of b-spinon conductivity

$$[\sigma_s] = [\mathbf{J}_s][\mathbf{E}_h]^{-1} = [\hbar]^{-1}, \tag{C.5}$$

and similarly, $[\sigma_h] = [\sigma_a] = [\sigma_s] = [\hbar]^{-1}$.

Appendix D. The calculation of the *b*-spinon conductivity σ_s

In the LPP, the excited b-spinons are deconfined and free, which will decide the longitudinal resistivity via the non-Ioffe–Larking rule (71).

The b-spinon conductivity $\sigma_{\rm s}$ can be calculated by the Kubo formula as follows:

$$\sigma_s^{\alpha\beta}(\omega) = \frac{\mathrm{i}}{\omega}\Pi_s^{\alpha\beta}\left(\mathbf{q} = 0,\, i\omega_n \to \omega + i0^+\right). \tag{D.1}$$

Here the polarization tensor in the real space/imaginary time is given by

$$\Pi_s^{\alpha\beta}(i,\, i';\, \tau) \equiv \Pi_{\rm curr.}^{\alpha\beta}(i,\, i';\, \tau) + \Pi_{\rm diam.}^{\alpha\beta}, \tag{D.2}$$

where $\Pi_{\rm diam.}$ represents the diamagnetic term of the polarization tensor, and $\Pi_{\rm curr.}^{\alpha\beta}(i,\, i';\, \tau)$ denotes the spinon current–current correlation function:

$$\Pi_{\rm curr.}^{\alpha\beta}(i,\, i';\, \tau) \equiv -\left\langle T_\tau J_s^{i\,+\,\hat{\alpha},\, i}(\tau) J_s^{i'\,+\,\hat{\beta},\, i'}(0)\right\rangle, \tag{D.3}$$

in which the spinon current density $J_s^{i\,+\,\alpha,\, i}$ is defined by

$$J_s^{i\,+\,\hat{\alpha},\, i} = \left(-iJ_s \sum_\sigma \sigma \mathrm{e}^{\mathrm{i}\sigma\bar{A}^{\,h}_{i+\hat{\alpha},i}} b^\dagger_{i\,+\,\hat{\alpha}\sigma} b^\dagger_{i\bar{\sigma}} + \mathrm{h.c.}\right)\hat{\alpha}. \tag{D.4}$$

Define the Matsubara Green's function for the Bogoliubov quasiparticle of the b-spinons:

$$G_\sigma(m,\, \tau) = -\left\langle T_\tau \gamma_{m\sigma}(\tau)\gamma^\dagger_{m\sigma}\right\rangle. \tag{D.5}$$

After taking the Fourier transformation, the mean-field solution is given by

$$G_\sigma^0\left(m,\, \mathrm{i}\omega_n\right) = \frac{1}{i\omega_n - E_m}, \tag{D.6}$$

where $\mathrm{i}\omega_{\rm n}$ is the bosonic Matsubara frequency $\omega_n = 2n\pi k_B T$. One may further introduce the spectral function $A_\sigma(m,\, \omega)$ such that

$$G_\sigma\left(m,\, \mathrm{i}\omega_n\right) = \int \mathrm{d}\omega \frac{A_\sigma(m,\, \omega)}{i\omega_n - \omega}, \tag{D.7}$$

where

$$A_\sigma(m,\, \omega) \equiv -\frac{1}{\pi}\,\mathrm{Im}\, G_\sigma\left(m,\, i\omega_n \to \omega + i0^+\right). \tag{D.8}$$

At the mean-field level, the spectral function simply reduces to $A_\sigma^{\,0}(m,\, \omega) = \delta(\omega - E_{m\sigma})$. Then the momentum-frequency representation of the spinon current–current correlation can be obtained:

$$\Pi_{\rm curr.}^{\alpha\beta}\left(\mathbf{q},\, i\omega_n\right) = J_s^2 \sum_{mm'} F_{mm'}\left(i\omega_n\right) G_{mm'}^{\alpha\beta}(\mathbf{q}), \tag{D.9}$$

where

$$\begin{aligned} G_{mm'}^{\alpha\beta}(\mathbf{q}) = \; & \frac{1}{N}\sum_\sigma \left[\sum_i \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_i}\mathrm{e}^{\mathrm{i}\sigma\bar{A}^{\,h}_{i+\hat{\alpha},i}} w^*_{m\sigma}\left(i\,+\,\hat{\alpha}\right) w_{m'\sigma}(i)\right]\left[\sum_{i'} \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_{i'}}\mathrm{e}^{\mathrm{i}\sigma\bar{A}^{\,h}_{i'+\hat{\beta},i'}} w^*_{m'\sigma}\left(i'\,+\,\hat{\beta}\right) w_{m\sigma}(i')\right. \\ & \left. -\sum_{i'} \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_{i'}}\mathrm{e}^{-\mathrm{i}\sigma\bar{A}^{\,h}_{i'+\hat{\beta},i'}} w_{m\sigma}\left(i'\,+\,\hat{\beta}\right) w^*_{m'\sigma}(i')\right]. \end{aligned} \tag{D.10}$$

Here it is easy to verify that $G^{\alpha\beta}_{mm'}(\mathbf{q}) = \left[G^{\alpha\beta}_{mm'}(-\mathbf{q})\right]^*$, and one may define a real number $G^{\alpha\beta}_{mm'} \equiv G^{\alpha\beta}_{mm'}(\mathbf{q}=0) \in \mathbb{R}$. Then

$$\begin{aligned} F_{mm'}(i\omega_n) = \; & \frac{1}{\beta}(u_m u_{m'} + v_m v_{m'})^2 \sum_{i\omega_m}\Big[G(m, i\omega_m)G(m', -i\omega_n - i\omega_m) \\ & + G(m, i\omega_m)G(m', i\omega_n - i\omega_m)\Big] + \frac{1}{\beta}(u_m v_{m'} + u_m v_{m'})^2 \\ & \times \sum_{i\omega_m}\Big[G(m, i\omega_m)G(m', i\omega_n + i\omega_m) + G(m, i\omega_m)G(m', -i\omega_n + i\omega_m)\Big]. \end{aligned} \tag{D.11}$$

After summing over the Matsubara frequency, we get

$$\begin{aligned} F_{mm'}(i\omega_n) = \; & (u_m u_{m'} + v_m v_{m'})^2 \int \mathrm{d}\omega A(m, \omega) \int \mathrm{d}\omega' A(m', \omega')\left[n_B(\omega) - n_B(-\omega')\right] \\ & \times \left(\frac{1}{i\omega_n + \omega + \omega'} - \frac{1}{i\omega_n - \omega - \omega'}\right) + \\ & (u_m v_{m'} + u_m v_{m'})^2 \int \mathrm{d}\omega A(m, \omega) \int \mathrm{d}\omega' A(m', \omega')\left[n_B(\omega') - n_B(\omega)\right] \\ & \times \left(\frac{1}{i\omega_n + \omega - \omega'} - \frac{1}{i\omega_n - \omega + \omega'}\right), \end{aligned} \tag{D.12}$$

where

$$A(m, \omega) = \frac{1}{\pi}\frac{\mathrm{Im}\,\Sigma}{\left(\omega - E_m - \mathrm{Re}\,\Sigma\right)^2 + (\mathrm{Im}\,\Sigma)^2}, \tag{D.13}$$

in which $\mathrm{Re}\,\Sigma$ and $\mathrm{Im}\,\Sigma$ denote the real part and imaginary part, respectively, of the self-energy of the b-spinon.

Substituting the mean-field result $A(m, \omega) = \delta(\omega - E_m)$, one obtains

$$\begin{aligned} F_{mm'}(i\omega_n) = \; & (u_m u_{m'} + v_m v_{m'})^2\left[n_B(E_m) - n_B(-E_{m'})\right] \\ & \times \left(\frac{1}{i\omega_n + E_m + E_{m'}} - \frac{1}{i\omega_n - E_m - E_{m'}}\right) \\ & + (u_m v_{m'} + u_m v_{m'})^2\left[n_B(E_{m'}) - n_B(E_m)\right] \\ & \times \left(\frac{1}{i\omega_n + E_m - E_{m'}} - \frac{1}{i\omega_n - E_m + E_{m'}}\right). \end{aligned} \tag{D.14}$$

The diamagnetic term of the polarization tensor is given by

$$\Pi^{\alpha\beta}_{MF\,\mathrm{diam.}}(i - i', \tau) = 2J_s\Delta^s\delta_{\alpha\beta}\delta_{ii'}\delta(\tau). \tag{D.15}$$

Numerically, we have checked that the diamagnetic term of the polarization tensor gets precisely canceled:

$$\mathrm{Re}\,\Pi^{\alpha\alpha}_{\mathrm{curr.}}(\mathbf{q}=0, i\omega_n=0) = -2J_s\Delta^s = -\Pi^{\alpha\alpha}_{MF\,\mathrm{diam.}}(\mathbf{q}=0, i\omega_n=0). \tag{D.16}$$

Now we consider

$$\begin{aligned}\mathrm{Re}\,\sigma_s^{\alpha\beta}(\omega=0) &= -\left.\frac{\mathrm{Im}\,\Pi_s^{\alpha\beta}\left(\mathbf{q}=0,\, i\omega_n\to\omega+i0^+\right)}{\omega}\right|_{\omega\to 0}\\ &= -J_s^2\sum_{mm'}G_{mm'}^{\alpha\beta}\left.\frac{\mathrm{Im}\,F_{mm'}\left(i\omega_n\to\omega+i0^+\right)}{\omega}\right|_{\omega\to 0},\end{aligned} \tag{D.17}$$

where

$$\begin{aligned}\left.\frac{\mathrm{Im}\,F_{mm'}\left(i\omega_n\to\omega+i0^+\right)}{\omega}\right|_{\omega\to 0} &= 2\pi\left(u_m u_{m'}+v_m v_{m'}\right)^2\int \mathrm{d}\tilde{\omega}A(m,\,\tilde{\omega})A\left(m',\,-\tilde{\omega}\right)\frac{\partial n_B(\tilde{\omega})}{\partial\tilde{\omega}}\\ &\quad -2\pi\left(u_m v_{m'}+v_m u_{m'}\right)^2\int \mathrm{d}A(m,\,\tilde{\omega})\\ &\quad \times A\left(m',\,\tilde{\omega}\right)\frac{\partial n_B(\tilde{\omega})}{\partial\tilde{\omega}}.\end{aligned} \tag{D.18}$$

Finally, we arrive at

$$\mathrm{Re}\,\sigma_s^{\alpha\beta}(\omega=0)=2\pi J_s^2\sum_{mm'}\left(u_m v_{m'}+v_m u_{m'}\right)^2 G_{mm'}^{\alpha\beta}\int \mathrm{d}\tilde{\omega}A(m,\,\tilde{\omega})A\left(m',\,\tilde{\omega}\right)\frac{\partial n_B(\tilde{\omega})}{\partial\tilde{\omega}}, \tag{D.19}$$

Or, after recovering the SI unit, i.e. $[\sigma_s]=[\hbar]^{-1}$,

$$\mathrm{Re}\,\sigma_s^{\alpha\beta}(\omega=0)=\frac{2\pi}{\hbar}J_s^2\sum_{mm'}\left(u_m v_{m'}+v_m u_{m'}\right)^2 G_{mm'}^{\alpha\beta}\int \mathrm{d}\tilde{\omega}A(m,\,\tilde{\omega})A\left(m',\,\tilde{\omega}\right)\frac{\partial n_B(\tilde{\omega})}{\partial\tilde{\omega}}. \tag{D.20}$$

Appendix E. Derivation of equation (79)

First of all, we may list the following useful formulas:

$$\mathbf{J}_e=\sigma_e\mathbf{E}_e, \tag{E.1}$$

$$\mathbf{J}_s=\sigma_s\mathbf{E}_h, \tag{E.2}$$

$$\mathbf{J}_h=\sigma_h\left(\mathbf{E}_s+e\mathbf{E}_e+\mathbf{E}_a\right), \tag{E.3}$$

$$\mathbf{J}_a=-\sigma_a\mathbf{E}_a, \tag{E.4}$$

$$\mathbf{J}_s=\frac{1}{\pi}\epsilon\cdot\mathbf{E}_s, \tag{E.5}$$

$$\mathbf{J}_h=\frac{1}{\pi}\epsilon\cdot\mathbf{E}_h, \tag{E.6}$$

$$\mathbf{J}_e=e\mathbf{J}_h=e\mathbf{J}_a. \tag{E.7}$$

The notations are defined as follows: $\sigma_{s/h/a}$ denote the 'conductivity' of b-spinons / h-holons / a-spinons. $\mathbf{J}_{s/h/a}$ denote their currents. $\epsilon\equiv\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$ is a matrix acting on $\hat{x}$- and $\hat{y}$- coordinates,

$\epsilon^{12} = \epsilon^{xy} = 1$, $\epsilon^{21} = \epsilon^{yx} = -1$. $\mathbf{J}_e$ is the electric current that is physically detected in transport experiments. $\mathbf{E}_{s/h/a} = -\partial_t \mathbf{A}^{s/h/a}$ are electric fields formed by gauge fields $\mathbf{A}^{s/h/a}$.

The physical pictures of the seven formulas are explained as follows. Equation (E.1) is the definition of the well-known Ohm's law. $\mathbf{A}^e$ is external electromagnetic field. Equation (E.2) is a response formula in analog to 'Ohm's law', meaning that 'electric field' $\mathbf{E}_h$ formed by $\mathbf{A}^h$ generates a b-spinon current owing to the minimal coupling between $\mathbf{A}^h$ and b-spinons. This minimal coupling can be found in equations (40) and (46). Likewise, equation (E.3) is a response formula about h-holon current. h-holons simultaneously couple to three gauge fields, namely, $\mathbf{A}^s$ (i.e. equation (45)), $\mathbf{A}^e$ (i.e. equation (45)) and, $\mathbf{A}^a$ (see the discussion above equation (44)). Due to opposite $\mathbf{A}^a$ gauge charges carried by h-holons and a-spinons, equation (E.4) that describe the linear response of a-spinons can also be easily understood.

Equations (E.5) and (E.6) can be understood via equations (41) and (42). For instance, equation (42) indicate that h-holon particle density is the source of the magnetic flux of $\mathbf{A}^h$ gauge field. Therefore, once h-holons moves and thereby there is a holon current $\mathbf{J}_h$, h-holons will necessarily generate electric field of $\mathbf{A}^h$ along the transverse direction. More rigorous derivation of equations (E.5) and (E.6) can be performed in the mutual Chern–Simons gauge field theory which has space-time covariant form as shown in [60].

The first identity in equation (E.7) is obvious since each h-holon carries a fundamental electric charge while a-spinons and b-spinons are charge-neutral in our fractionalization framework. The second identity in equation (E.7) can be understood as a consequence of the internal gauge field $\mathbf{A}^a$. More pictorially, h-holons and a-spinons are created and annihilated together implied by equation (9). In the following, we may apply these seven formulas to derive equation (79).

$\mathbf{E}_s$ in equation (E.5) can be expressed as: $\mathbf{E}_s = \epsilon^{-1} \cdot \mathbf{J}_s \pi = -\epsilon \cdot \mathbf{J}_s \pi$. $\mathbf{E}_h$ in equation (E.6) can be expressed as: $\mathbf{E}_h = \epsilon^{-1} \cdot \mathbf{J}_h \pi = -\epsilon \cdot \mathbf{J}_h \pi$. Further considering (E.2), we end up with:

$$\mathbf{E}_s = -\epsilon \cdot \mathbf{J}_s \pi = -\epsilon \cdot \mathbf{E}_h \sigma_s \pi = \epsilon \cdot \epsilon \cdot \mathbf{J}_h \sigma_s \pi^2 = -\mathbf{J}_h \sigma_s \pi^2. \tag{E.8}$$

Considering equations (E.7) and (E.1), we have

$$\mathbf{E}_s = -\mathbf{J}_h \sigma_s \pi^2 = -\mathbf{J}_e \frac{\sigma_s \pi^2}{e} = -\mathbf{E}_e \frac{\sigma_e \sigma_s \pi^2}{e}. \tag{E.9}$$

Considering equations (E.7), (E.1) and (E.4), we have

$$\mathbf{E}_a = -\sigma_a^{-1} \mathbf{J}_a = -\frac{\sigma_a^{-1}}{e} \mathbf{J}_e = -\frac{\sigma_a^{-1} \sigma_e}{e} \mathbf{E}_e. \tag{E.10}$$

Substituting (E.9) and (E.10) into equation (E.3), we end up with

$$\mathbf{J}_h = \sigma_h \left(e\mathbf{E}_e - \frac{\sigma_e \sigma_s \pi^2}{e} \mathbf{E}_e - \frac{\sigma_a^{-1} \sigma_e}{e} \mathbf{E}_e \right). \tag{E.11}$$

By further considering $e\mathbf{J}_h = \mathbf{J}_e = \sigma_e \mathbf{E}_e$, for any $\mathbf{E}_e$, the following identity is valid:

$$\sigma_e = \sigma_h \left(e^2 - \sigma_e \sigma_s \pi^2 - \sigma_a^{-1} \sigma_e \right), \tag{E.12}$$

which is identical to:

$$\sigma_e^{-1} = \frac{1}{e^2}\left(\sigma_h^{-1} + \sigma_a^{-1} + \pi^2\sigma_s\right). \tag{E.13}$$

Once the SI unit is recovered, equation (79) is obtained:

$$\sigma_e^{-1} = \frac{1}{e^2}\left(\sigma_h^{-1} + \sigma_a^{-1} + \pi^2\hbar^2\sigma_s\right). \tag{E.14}$$

References

[1] Tsuei C C and Kirtley J R 2000 *Rev. Mod. Phys.* **72** 969
[2] Damascelli A, Hussain Z and Shen Z-X 2003 *Rev. Mod. Phys.* **75** 473 and the references therein
[3] Devereaux T P and Hackl R 2007 *Rev. Mod. Phys.* **79** 175
[4] Fischer Ø *et al* 2007 *Rev. Mod. Phys.* **79** 353
[5] de Gennes P G 1966 *Superconductivity of Metals and Alloys* (New York: Benjamin)
[6] Timusk T and Statt B 1999 *Rep. Prog. Phys* **62** 61 and the references therein
[7] Schrieffer J R and Brooks J S 2007 *Handbook of High-Temperature Superconductivity* (Berlin: Springer)
[8] Lee P A, Nagaosa N and Wen X-G 2006 *Rev. Mod. Phys.* **78** 17 and the references therein
[9] Anderson P W 1987 *Science* **235** 1196
[10] Anderson P W 1997 *The Theory of Superconductivity in the High T_c Cuprates* (Princeton, NJ: Princeton University Press)
[11] Baskaran G, Zou Z and Anderson P W 1987 *Solid State Commun.* **63** 973
[12] Zhang F C, Gros C, Rice T M and Shiba H 1988 *Supercond. Sci. Technol.* **1** 36
[13] Anderson P W, Lee P A, Randeria M, Rice T M, Trivedi N and Zhang F C 2004 *J. Phys.: Condens. Matter* **16** R755
[14] Weng Z Y 2007 *Int. J. Mod. Phys.* B **21** 773 and the references therein
[15] Edegger B, Muthukumar V N and Gros C 2007 *Adv. Phys.* **56** 927 and the references therein
[16] Tešanović Z 2008 *Nat. Phys.* **4** 408
[17] Zaanen J and Overbosch B J 2011 *Phil. Trans. R. Soc.* A **369** 1599 and the references therein
[18] Phillips P 2011 *Phil. Trans. R. Soc.* A **369** 1574 and the references therein
[19] Carlson E W, Emery V J, Kivelson S A and Orgad D 2003 *The Physics of Superconductivity: Conventional and Unconventional* ed K H Benneman and J B Ketterson vol 2 (Berlin: Springer)
[20] Zhang S C 1997 *Science* **275** 1089
[21] Demler E, Sachdev S and Zhang Y 2001 *Phys. Rev. Lett.* **87** 067202
Zhang Y, Demler E and Sachdev S 2002 *Phys. Rev.* B **66** 094501
[22] Chakravarty S 2011 *Rep. Prog. Phys.* **74** 022501
[23] Hasan M Z *et al* 2000 *Science* **288** 1811
[24] Ye C *et al* 2013 *Nat. Commun.* **4** 1365
[25] Weng Z Y 2011 *New J. Phys.* **13** 103039
Weng Z Y, Zhou Y and Muthukumar V N 2005 *Phys. Rev.* B **72** 014503
[26] Marshall W 1955 *Proc. R. Soc.* A **232** 48
[27] Weng Z Y, Sheng D N, Chen Y and Ting C S 1997 *Phys. Rev.* B **55** 3894
Sheng D N, Chen Y C and Weng Z Y 1996 *Phys. Rev. Lett.* **77** 5102
[28] Wu K, Weng Z Y and Zaanen J 2008 *Phys. Rev.* B **77** 155102
[29] Gu Z C and Weng Z Y 2005 *Phys. Rev.* B **72** 104520
[30] Johnston D C 1989 *Phys. Rev. Lett.* **62** 957
[31] Takagi H *et al* 1992 *Phys. Rev. Lett.* **69** 2975

[32] Imai T, Slichter C P, Yoshimura K and Kosuge K 1993 *Phys. Rev. Lett.* **70** 1002
[33] Nakano T, Oda M, Manabe C, Momono N, Miura Y and Ido M 1994 *Phys. Rev.* B **49** 16000
[34] Shibauchi T, Krusin-Elbaum L, Li M, Maley M P and Kes P H 2001 *Phys. Rev. Lett.* **86** 5763
Krusin-Elbaum L, Shibauchi T and Mielke C H 2004 *Phys. Rev. Lett.* **92** 097005
[35] Loram J W, Luo J, Cooper J R, Liang W Y and Tallon J L 2001 *J. Phys. Chem. Solid* **62** 59
[36] Momono N, Matsuzaki T, Oda M and Ido M 2002 *J. Phys. Soc. Jpn.* **71** 2832–5
[37] Corson J *et al* 1999 *Nature* **398** 221
[38] Takigawa M *et al* 1991 *Phys. Rev.* B **43** 247
[39] Ando Y, Komiya S, Segawa K, Ono S and Kurita Y 2004 *Phys. Rev. Lett.* **93** 267001
[40] Xu Z A, Ong N P, Wang Y, Kakeshita T and Uchida S 2000 *Nature* **406** 486
[41] Wang Y Y, Xu Z A, Kakeshita T, Uchida S, Ono S, Yoichi Ando and Ong N P 2001 *Phys. Rev.* B **64** 224519
[42] Wang Y Y, Li L, Naughton M J, Gu G D, Uchida S and Ong N P 2005 *Phys. Rev. Lett.* **95** 247002
Li L, Wang Y Y, Naughton M J, Ono S, Ando Y and Ong N P 2005 *Europhys. Lett.* **72** 451
[43] Wang Y Y, Li L and Ong N P 2006 *Phys. Rev.* B **73** 024510
[44] Honma T and Hor P H 2008 *Phys. Rev.* B **77** 184520
[45] Mei J W and Weng Z Y 2010 *Phys. Rev.* B **81** 014507
Shaw M, Weng Z Y and Ting C S 2003 *Phys. Rev.* B **68** 014511
[46] Muthukumar V N and Weng Z Y 2002 *Phys. Rev.* B **65** 174511
Weng Z Y and Muthukumar V N 2002 *Phys. Rev.* B **66** 094509
[47] Weng Z Y and Qi X L 2006 *Phys. Rev.* B **74** 144518
[48] Qi X L and Weng Z Y 2007 *Phys. Rev.* B **76** 104502
[49] Doiron-Leyraud N, Proust C, LeBoeuf D, Levallois J, Bonnemaison J-B, Liang R, Bonn D A, Hardy W N and Taillefer L 2007 *Nature* **447** 565
[50] Kawasaki S *et al* 2010 *Phys. Rev. Lett.* **105** 137002
[51] Mei J W, Kawasaki S, Zheng G Q, Weng Z Y and Wen X-G 2012 *Phys. Rev.* B **85** 134519
[52] Zhang F C and Rice T M 1988 *Phys. Rev.* B **37** 3759
[53] Auerbach A and Arovas D P 1988 *Phys. Rev. Lett.* **61** 617
Sarker S, Jayaprakash C, Krishnamurthy H R and Ma M 1989 *Phys. Rev.* B **40** 5028
Yoshioka D 1989 *J. Phys. Soc. Japan* **58** 32
[54] Liang S, Doucot B and Anderson P W 1988 *Phys. Rev. Lett.* **61** 365
[55] Weng Z Y, Sheng D N and Ting C S 1999 *Phys. Rev.* B **59** 8943
Weng Z Y, Sheng D N and Ting C S 1998 *Phys. Rev. Lett.* **80** 5401
[56] Ribeiro T C and Wen X-G 2006 *Phys. Rev.* B **74** 155113
Ribeiro T C and Wen X-G 2005 *Phys. Rev. Lett.* **95** 057001
[57] Barzykina V and Pines D 2009 *Adv. Phys.* **58** 1
[58] Kou S P, Li T and Weng Z Y 2009 *Europhys. Lett.* **88** 17010
[59] You Y Z and Weng Z Y 2014 *New J. Phys.* **16** 023001
[60] Ye P, Tian C S, Qi X L and Weng Z Y 2012 *Nucl. Phys.* B **854** 815
Ye P, Tian C S, Qi X L and Weng Z Y 2011 *Phys. Rev. Lett.* **106** 147002
Ye P, Zhang L and Weng Z Y 2012 *Phys. Rev.* B **85** 205142
Kou S P, Qi X L and Weng Z Y 2005 *Phys. Rev.* B **71** 235102
[61] Kou S P and Weng Z Y 2003 *Phys. Rev. Lett.* **90** 157003
[62] Chen W Q and Weng Z Y 2005 *Phys. Rev.* B **71** 134516
[63] Gu Z C and Weng Z Y 2007 *Phys. Rev.* B **76** 024501
[64] Kou S P, Qi X L and Weng Z Y 2005 *Phys. Rev.* B **72** 165114
[65] Chaikin P M and Lubensky T C 2000 *Principles of condensed matter physics* (Cambridge: Cambridge University Press)
[66] Barnes S E 1976 *J. Phys. F: Met. Phys.* **6** 1375
[67] Coleman P 1984 *Phys. Rev.* B **29** 3035

[68] Shraiman B I and Siggia E D 1989 *Phys. Rev. Lett.* **62** 1564
[69] Lee P A 1989 *Phys. Rev. Lett.* **63** 680
[70] Jiang H C, Block M S, Mishmash R V, Garrison J R, Sheng D N, Motrunich O I and Fisher M P A 2013 *Nature* **493** 39
[71] Hlubina R, Putikka V O and Rice T M 1992 *Phys. Rev.* B **46** 11224
[72] Zhu Z, Jiang H C, Qi Y, Tian C S and Weng Z Y 2013 *Sci. Rep.* **3** 2586
[73] Ye P and Wang Q-R 2013 *Nucl. Phys.* B **874** 386
[74] Senthil T and Lee P A 2009 *Phys. Rev.* B **79** 245116