Front. Phys., 2011, 6(4): 370–378
DOI 10.1007/s11467-011-0220-1

REVIEW ARTICLE

Mott physics, sign structure, ground state wavefunction, and high-T_c superconductivity

Zheng-Yu Weng

Institute for Advanced Study, Tsinghua University, Beijing 100084, China

E-mail: weng@tsinghua.edu.cn

Received September 29, 2011; accepted October 27, 2011

In this article I give a pedagogical illustration of why the essential problem of high-T_c superconductivity in the cuprates is about how an antiferromagnetically ordered state can be turned into a short-range state by doping. I will start with half-filling where the antiferromagnetic ground state is accurately described by the Liang–Doucot–Anderson (LDA) wavefunction. Here the effect of the Fermi statistics becomes completely irrelevant due to the no double occupancy constraint. Upon doping, the statistical signs reemerge, albeit much reduced as compared to the original Fermi statistical signs. By precisely incorporating this altered statistical sign structure at finite doping, the LDA ground state can be recast into a short-range antiferromagnetic state. Superconducting phase coherence arises after the spin correlations become short-ranged, and the superconducting phase transition is controlled by spin excitations. I will stress that the pseudogap phenomenon naturally emerges as a crossover between the antiferromagnetic and superconducting phases. As a characteristic of non Fermi liquid, the mutual statistical interaction between the spin and charge degrees of freedom will reach a maximum in a high-temperature "strange metal phase" of the doped Mott insulator.

Keywords Mott physics, high-T_c cuprates, ground state wavefunction, sign structure

PACS numbers 74.20.-z, 74.20.Mn, 74.72.-h

Contents

1 Introduction 370
2 Constructing a new superconducting ground state 371
3 Mottness and sign structure 372
4 Superconducting phase coherence 375
5 Nature of pseudogap states 376
6 Concluding remarks 377
Acknowledgements 378
References 378

1 Introduction

Right after the 1986 discovery of high-T_c superconductivity in the cuprate materials, Anderson made a seminal proposal [1] that the on-site Coulomb repulsion plays a crucial role. At half-filling, with each unit cell in the copper-oxide plane occupied by one electron from the "conduction band", the cuprate is a Mott insulator with a full gap opening up in the charge degree of freedom. Due to the odd number of $S = 1/2$ spin per unit cell, the spin degrees of freedom remain unfrozen, in contrast to a band insulator, and the localized spins antiferromagnetically interact with each other via a residual nearest-neighbor Heisenberg superexchange coupling J, to the leading order of approximation. It was then conjectured [1] that the novel superconductivity should arise by doping such a Mott insulator.

A prototype superconducting ground state incorporating the "Mottness" is known as the Gutzwiller-projected BCS state [1–3]

$$|\Psi_{\rm RVB}\rangle = \widehat{P}_{\rm G}|{\rm BCS}\rangle \tag{1}$$

where $|{\rm BCS}\rangle$ denotes an ordinary BCS superconducting state and $\widehat{P}_{\rm G}$ is a Gutzwiller projection operator enforcing the following no double occupancy constraint

$$\sum_{\sigma} c^{\dagger}_{i\sigma} c_{i\sigma} \leqslant 1 \tag{2}$$

There are two essential components in this ground state ansatz. First, the Cooper pairing in $|{\rm BCS}\rangle$ is originated

©Higher Education Press and Springer-Verlag Berlin Heidelberg 2011

from the superexchange coupling J. Second, the Gutzwiller projection $\widehat{P}_{\rm G}$ removes those doubly occupied configurations which would cost high energy due to the large on-site Coulomb repulsion. In particular, because of $\widehat{P}_{\rm G}$, the Cooper pairing in $|{\rm BCS}\rangle$ reduces to the neutralized singlet pairs of spins at half-filling, known as the resonating valence bond (RVB) state [1], which is a Mott insulator with the charge degree of freedom totally frozen out.

Hence an essential observation made in Ref. [1] is that the localized spins have a tendency to form real-space singlet pairing in order to gain the superexchange energy. Hole doping (with removing the localized spins) can then lead to high-$T_{\rm c}$ superconductivity as the charge neutral RVB pairs start to move to the empty sites and become partially charged Cooper pairs. Today, strongly motivated by this idea, searching for new Mott insulators with an RVB (spin liquid) ground state has become an important goal in material syntheses, with the hope to find new exotic superconductors upon doping.

However, the electron ground state in the half-filled cuprates has turned out to be not simply a spin liquid with short-range RVB pairing. As a matter of fact, it was quickly established experimentally [4] that the localized spins are long-range antiferromagnetically (AFM) ordered instead. Theoretically the ground state of the AFM Heisenberg model in two-dimensional (2D) square lattice is also an AFM ordered state [5] whose properties are well in accord with those found in the cuprates at half-filling. Nevertheless the AFM long-range order disappears quickly upon doping the holes into the cuprates. Consequently only short-range AFM correlations remain beyond some finite doping concentration, which may be still effectively described, as one would hope, by some kind of RVB or spin liquid state [6]. Since there is already a sufficient amount of holes present, it would then become superconducting as envisaged [1] in the original RVB proposal.

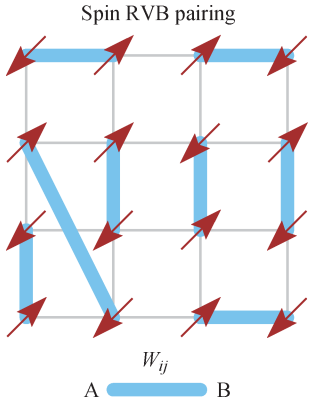


Fig. 1 In an LDA wavefunction [Eq. (3)], bosonic spins form singlet pairs with an amplitude W_{ij}, which only connects two spin partners at opposite (A and B) sublattices as indicated by a blue-colored bond. The Marshall sign rule [9] is precisely satisfied at half-filling, where $W_{ij} \propto 1/|\boldsymbol{r}_{ij}|^3$ is real and positive, and the wavefunction is very accurate to describe the ground state of the Heisenberg model with a long-range AFM order in the thermodynamic limit [8]. In the superconducting phase, such a state will reduce to a "ghost" spin liquid background with W_{ij} becoming short-ranged (cf. Fig. 2).

In this context, a spin liquid state relevant to the superconductivity, if existing in the cuprates, should be by itself an *emergent* state, arising after the long-range AFM order of an antiferromagnet gets destroyed by the motion of the doped holes. The real challenge to understanding the superconducting mechanism in the cuprates is therefore to correctly demonstrate how the antiferromagnet at half-filling can be doped into a short-range AFM state, where the doping itself plays the central role of "dynamic frustration effect". The resulting state can be either simultaneously superconducting or, more generally, a "pseudogap phase" with the superconducting phase naturally embedded inside.

This will be the central theme illustrated in this article. I will present convincing theoretical rationale of why it is absolutely necessary to carefully deal with the nonperturbative and singular nature of doping an antiferromagnet in order to correctly understand the superconducting and pseudogap physics in the cuprates. The resulting superconducting ground state [7], although bears a close resemblance to Eq. (1), is generally qualitatively different from the latter because, as it turns out, the spin and charge degrees of freedom are intrinsically entangled together by a hidden fundamental law underlying the Mottness (see Section 3), which is not manifested in Eq. (1).

2 Constructing a new superconducting ground state

As emphasized in Introduction, the superconducting ground state in a doped Mott insulator will reduce to a Mott insulator at half-filling. In particular, the localized spin state should correctly recover the AFM long-range ordered ground state of the Heisenberg model in this limit. Below we illustrate how to realize such a construction in a straightforward way.

At half-filling, although there is no exact solution in 2D, the Liang–Docout–Anderson (LDA) type bosonic RVB state [8] can give rise to the most accurate variational energy for the Heisenberg model with the correct long-range AFM order. It can be written in the form

$$|{\rm RVB}\rangle = \sum_{\{\sigma_s\}} \Phi_{\rm RVB}\left(\sigma_1, \sigma_1, \cdots, \sigma_N\right) c^{\dagger}_{1\sigma_1} c^{\dagger}_{2\sigma_2} \cdots c^{\dagger}_{N\sigma_N} |0\rangle \tag{3}$$

in the electron c-operator representation. By nature this is a *bosonic* state with a bosonic wavefunction $\Phi_{\rm RVB}\left(\{\sigma_s\}\right) \equiv \sum_{\rm partition} \prod_{(ij)} (-1)^i W_{ij}$ for each given spin configuration $\{\sigma_s\} = \sigma_1, \sigma_1, \cdots, \sigma_N$ as shown in

Fig. 1. Here the RVB pairing amplitude $(-1)^i W_{ij}$ connects two *antiparallel* spins denoted by i (up spin) and j (down spin) at two opposite sublattices of the square lattice, with the summation running over all possible pairing partitions for a given $\{\sigma_s\}$. The staggered sign $(-1)^i$ (the Marshall sign) is explicitly separated from W_{ij} such that the latter remains a smooth function of the distance between even and odd lattice sites. A real and positive W_{ij} obeying the power law at large spatial separation of ij has been shown to optimize [8] the ground state energy, with $\Phi_{\rm RVB}$ satisfying the Marshall sign rule [9].

To start with $|\text{RVB}\rangle$, which is much more precise than Eq. (1) at half-filling, a ground state ansatz for the doped case may be constructed as follows [7]:

$$|\Psi_{\rm G}\rangle = \Lambda_h \left(\sum_{ij} g_{ij} c_{i\uparrow} c_{j\downarrow} \right)^{N_h/2} |\text{RVB}\rangle \tag{4}$$

Namely, the doped holes can be introduced by annihilating N_h electrons from the "vacuum" state $|\text{RVB}\rangle$. Apparently such a new state automatically satisfies the no double occupancy constraint without invoking the Gutzwiller projection $\widehat{P}_{\rm G}$. Furthermore, the holes are created in pairs with an amplitude g_{ij} to realize the Cooper pairing, which is energetically favorable so long as the RVB pairing persists in $|\text{RVB}\rangle$ at low doping [7], as illustrated in Fig. 2.

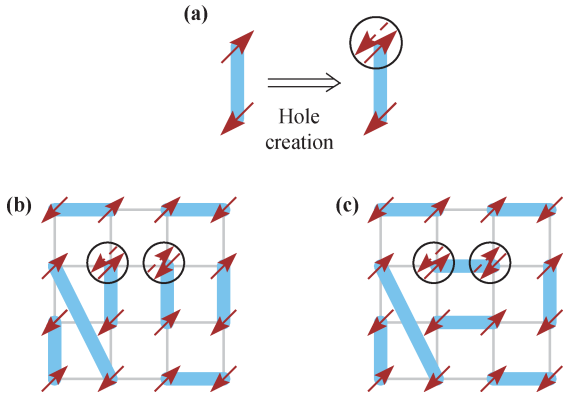


Fig. 2 Schematic illustration of the new superconducting ground state (4). **(a)** An RVB pair in Fig. 1 becomes a hole-spin pair with the hole (the circle) created by annihilating a spin (as indicated by the dashed arrow), with its RVB partner being automatically associated with the doped hole; **(b)** and **(c)** Two doped holes moving on $|\text{RVB}\rangle$ have the tendency to form a Cooper pair as driven by the RVB pairing between their spin partners. Note that the important sign structure of the wavefunction, i.e., Λ_h in Eq. (4), is not directly shown here.

Here both g_{ij} and W_{ij} can be determined variationally [7] based on the t–J model and the main results are as follows. While g_{ij} is associated with the singlet d-wave Cooper pairing, W_{ij} is found to change from a long-range power-law behavior ($\propto 1/|\boldsymbol{r}_{ij}|^3$) to an exponential behavior: $|W_{ij}| \propto \mathrm{e}^{-|\boldsymbol{r}_{ij}|^2/(2\xi^2)}$, where $\boldsymbol{r}_{ij}$ is the spatial distance and ξ is the characteristic pair size determined by the doping concentration δ: $\xi = a\sqrt{\frac{2}{\pi\delta}}$ (a is the lattice constant). This indicates that $|\text{RVB}\rangle$ indeed evolves from a long-range AFM state to a *short-ranged* spin liquid self-consistently at finite doping.

Generally speaking, one cannot smoothly connect the ground state $|\Psi_{\rm G}\rangle$ in Eq. (4) to the Gutzwiller-projected BCS state $|\Psi_{\rm RVB}\rangle$ in Eq. (1). At finite doping, the former is distinct from the latter by an explicit separation of the neutral spin RVB pairing and Cooper pairing. Thus, in contrast to the "one-component" RVB state $|\Psi_{\rm RVB}\rangle$ without explicitly distinguishing the Cooper and RVB pairings, the construction in Eq. (4) may be regarded as a "two-component" RVB structure. Such a ground state can accurately describe the AFM correlations, via a bosonic state $|\text{RVB}\rangle$, at half-filling, while it can account for the right charge degree of freedom at finite doping without an additional projection procedure.

The key in driving an AFM state into a spin liquid involves an important doping effect through Λ_h in Eq. (4). Generally the phase shift factor Λ_h emerges because the bosonic RVB pairing in $|\text{RVB}\rangle$ and the Cooper pairing of the electrons are not statistically compatible and each doped hole will create a nonlocal phase shift in the spin background. Here Λ_h will represent the most important and singular property of the Mottness as to be elaborated below.

3 Mottness and sign structure

The purely bosonic LDA wavefunction, Eq. (3), can be regarded as an important consequence of the Mott physics. Namely, the electrons are fully bosonized in the restricted Hilbert space at half-filling. In the opposite dilute electron limit, the Fermi statistics of the electrons should get recovered. Then the statistical sign structure of wavefunctions for a doped Mott insulator is expected to be generally changed when holes are doped into the system [10, 11].

Explicitly identifying this effect turns out to be crucial in determining the phase shift factor Λ_h in the wavefunction construction (4), which may be generally expressed by

$$\Lambda_h \equiv \sum_{\{l_h\}} \left(n^h_{l_1} n^h_{l_2} \cdots n^h_{l_{N_h}} \right) \varphi_h(l_1, l_2, \cdots, l_{N_h}) \cdot \mathrm{e}^{-\mathrm{i}\left(\widehat{\Omega}_{l_1} + \widehat{\Omega}_{l_2} + \cdots + \widehat{\Omega}_{l_{N_h}}\right)} \tag{5}$$

where the phase shifts $\{\widehat{\Omega}_{l_h}\}$ are associated with the holes, with $n^h_l = 1 - \sum_\sigma c^\dagger_{l\sigma} c_{l\sigma} \geqslant 0$ denoting the hole occupation number at site l, and φ_h is a bosonic wavefunction symmetric with regard to the hole coordinates $\{l_h\} = l_1, l_2, \cdots, l_{N_h}$, which is present to ensure gauge invariance of the phase shift fields.

Let us first introduce an exact theorem for the t–J model, known as the phase string effect [12–14], which

holds for a bipartite lattice in any dimension, doping and temperature. The t–J model is one of the simplest models describing doped Mott insulators in the limit that the on-site Coulomb repulsion is much larger than the hopping integral t. This model is composed of two terms, $H_{t\text{–}J} = H_t + H_J$, with the hopping term

$$H_t = -t \sum_{\langle ij\rangle\sigma} c^\dagger_{i\sigma} c_{j\sigma} + \text{H.c.}$$

and the superexchange term

$$H_J = J \sum_{\langle ij\rangle} \left(\boldsymbol{S}_i \cdot \boldsymbol{S}_j - \frac{n_i n_j}{4} \right)$$

where $\boldsymbol{S}_i$ and $n_i \equiv \sum_\sigma n_{i\sigma}$ are on-site spin and number operators, respectively, and the Hilbert space is constrained by Eq. (2) as illustrated by Fig. 3 and the basic processes are shown in Fig. 4.

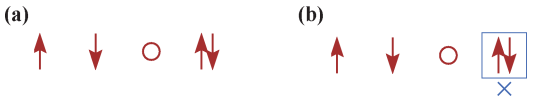


Fig. 3 **(a)** Four possible electron states at a lattice site; **(b)** The double occupancy state is forbidden under the Mottness constraint (2).

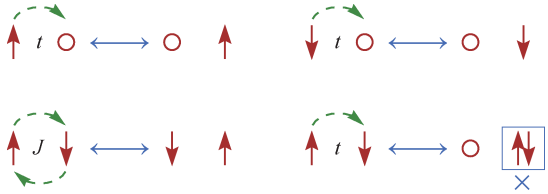


Fig. 4 The first three basic processes are allowed in the t–J model, involving the nearest-neighboring spin-hole states.

It can be rigorously demonstrated [14] that the partition function $Z_{t\text{–}J} = \text{Tr}\{\text{e}^{-\beta H_{t\text{–}J}}\}$ $[\beta = 1/(k_\text{B}T)]$ in the reduced Hilbert space (Figs. 3 and 4) can be explicitly expressed in terms of a loop summation

$$Z_{t\text{–}J} = \sum_c \tau_c \mathcal{Z}(c) \tag{6}$$

where c denotes a set of multi-loops of holes/spins at arbitrary temperature (for example, see Fig. 5). Here $\mathcal{Z}(c) \geqslant 0$ is the positive weight for each closed path c which depends on t, J and β [14].

Then τ_c collects all the nontrivial signs associated with the path c as follows:

$$\tau_c = (-1)^{N^\downarrow_h(c)} \times (-1)^{N^h_h(c)} \tag{7}$$

where $(-1)^{N^\downarrow_h(c)}$ is the so-called phase string with $N^\downarrow_h(c)$ counting the total number of the $\downarrow$-spins that the holes have "exchanged" with along the loop c, as illustrated in Fig. 5. This equation was first discovered [12] in the one-hole case of the t–J model. Physically it describes the quantum string-like spin mismatches or the disordered Marshall signs, created by the hopping of a hole in the spin background and being dynamically irreparable. At

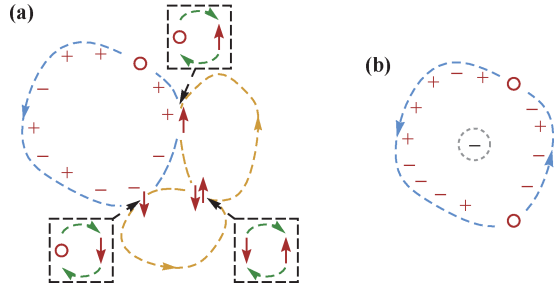


Fig. 5 **(a)** A phase string associated with a typical closed path of a hole moving on the spin background. Note that the $\pm$ signs are only created by the exchanges between the hole and spins, via τ_c in Eq. (7), while pure spin loops do not contribute to any signs. **(b)** The exchange of two holes in a close loop will lead to an additional negative sign to τ_c.

finite doping, it remains irreparable as $\mathcal{Z}(c) \geqslant 0$, with an additional sign contribution by $(-1)^{N^h_h(c)}$ where $N^h_h(c)$ counts the total number of hole–hole exchanges on the path c [cf. Fig. 5(b)].

Note that at half-filling, $\tau_c = 1$ and there is no nontrivial sign in the model, which is totally bosonized as pointed out above. In essence, although the original electrons are fermions, due to the no double occupancy constraint (2), the Fermi–Dirac statistics will no longer play a role in the restricted Hilbert space as far as the real physical process is concerned. At finite doping, the nontrivial sign does reemerge in Eq. (6) as represented by Eq. (7). Although such sign structure is solely associated with the doped holes and looks much "reduced" as compared to the full Fermi signs associated with the original electrons, the phase string effect in τ_c will lead to the very singular effect: for any loop of a hole, a mere change of one spin orientation can result in a total change of the sign in τ_c and thus the phase interference from constructive to destructive, or *vice versa*. In particular, in contrast to the conventional statistical problem involving the identical particles. Here the phase string effect is caused by the exchange between the holes and spins, and so a mutual statistical effect will emerge [13].

Figure 6 illustrates some typical phase strings picked up by a closed loop motion of a hole in a spin background which is in (a) the Nagaoka state in which spins are all polarized along the $\hat{z}$-axis; (b) an AFM state with spins forming a Néel order; (c) a spin disordered state. By noting that a closed path of the hole always involves an even number of steps in a bipartite lattice, one can easily see that only in (a) the phase interference due to phase string effect is always constructive, whereas in (b) and (c) the phase string effect becomes nontrivial and singular since even one spin flip can result in a total sign change of $(-1)^{N^\downarrow_h(c)}$. But in the partition function, the probability for (a) is extremely rare at finite J and thus the phase string effect is generally crucial in mediating the "entanglement" between the spin and charge degrees of freedom at finite doping. Figure 7 further shows how fermionic objects may arise from the bound pairs of holes

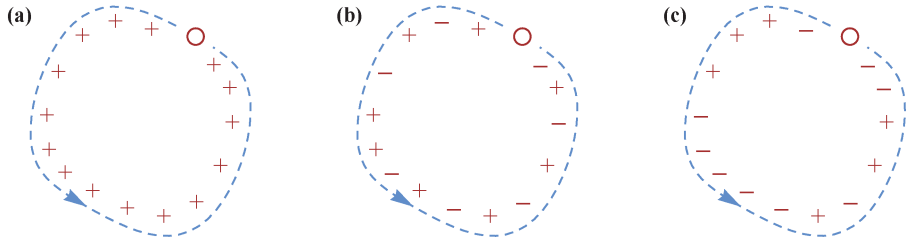


Fig. 6 There will be no frustration from the phase string effect if a hole is moving on the Nagaoka spin background where all the spins are polarized along, say, the $\hat{z}$-direction as shown in **(a)**. In **(b)**, spins are in the AFM ordered state, where even though the classical Néel order does not lead to the negative total sign, the quantum spin fluctuations can always result in a nontrivial phase string effect as one spin flip can cause the total sign change. One expects the strongest phase string frustration on the coherent motion of the hole if the spins are totally disordered as shown in **(c)**.

and spins, which are free of the phase string effect but retain the fermion signs between themselves via $(-1)^{N_h^h(c)}$ in Eq. (7). These fermions are associated with the holes and represent only a small portion of some very specially arranged loops in the summation of Eq. (6). In general, the electrons are fractionalized as shown in Fig. 5 and the holes and spins must obey a mutual statistics as dictated by the phase string effect.

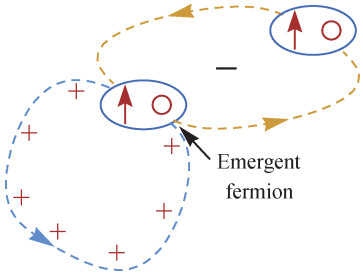


Fig. 7 Coherent fermions may emerge as tightly bound states of hole-spin where a hole hopping always involves the same spin. Note that the bound states are fermionic because of the additional sign upon the exchange of two of them according to Fig. 5(b).

Now one can introduce the precise definition of $\widehat{\Omega}_i$ to capture this mutual statistical effect. Define

$$e^{-i\widehat{\Omega}_i} = e^{-\frac{i}{2}\left(\Phi_i^s - \Phi_i^0\right)} \tag{8}$$

in which

$$\Phi_i^s \equiv \sum_{l\neq i} \theta_i(l) \left(\sum_\sigma \sigma n_{l\sigma}^b\right) \tag{9}$$

and

$$\Phi_i^0 \equiv \sum_{l\neq i} \theta_i(l) \tag{10}$$

where $\theta_i(l) = \mathrm{Im}\ln(z_i - z_l)$ (z_i is the complex coordinate of site i), and $n_{l\sigma}^b$ denotes the spin occupation number (with index σ) at site l, which always satisfies the single occupancy constraint

$$\sum_\sigma n_{l\sigma}^b = 1 \tag{11}$$

acting on the insulating spin state $|\mathrm{RVB}\rangle$.

In terms of Eq. (8), each spin in $|\mathrm{RVB}\rangle$ will contribute a $\pm\pi$ vortex via $\Phi_i^s/2$ to the doped hole at site i, with the spin itself sitting at the vortex core (see Fig. 8). *Vice versa,* each doped hole will be perceived by the spins in $|\mathrm{RVB}\rangle$ as introducing a π vortex, also via $\Phi_i^s/2$, with the hole sitting at the core. It implies that a doped hole and a neutral spin satisfy a "mutual semion statistics" as the phase shift $\widehat{\Omega}_i$ amounts to giving rise to $\pm\pi$ when one kind of species continuously circles around the other one once as illustrated in Fig. 9(c). Note that the single-valueness of Eq. (8) will be ensured by combining with $\Phi_i^0/2$. Thus the total phase shift added up in Λ_h represents a nontrivial entanglement between the doped holes and background spins, which will decide a "mutual semion statistics" sign structure in $|\Psi_\mathrm{G}\rangle$ that is fundamentally different from that in a BCS state $|\mathrm{BCS}\rangle$ satisfying the Fermi–Dirac statistics (cf. Fig. 9).

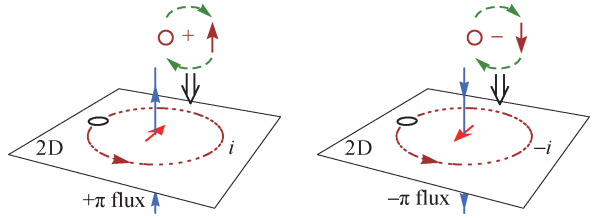


Fig. 8 The phase string effect of the t–J model can be precisely realized via the phase shift factor Λ_h in the wavefunction (2).

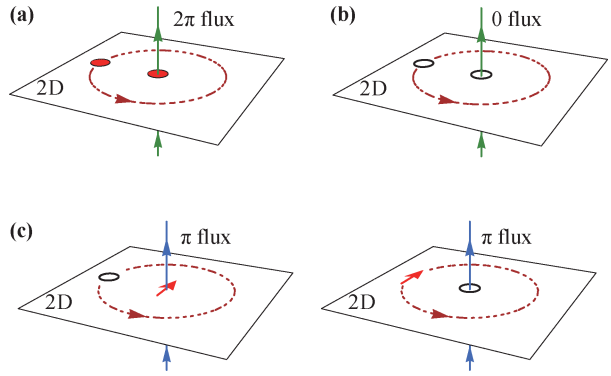


Fig. 9 Statistical transmutations by using the flux attachment based on the boson representation. **(a)** Fermionic statistics; **(b)** Bosonic statistics; **(c)** Mutual semionic statistics.

One may further examine such a sign structure by a thinking experiment in which a hole in $|\Psi_\mathrm{G}\rangle$ goes through a closed loop c. At each step of the nearest-neighbor movement of the hole, a singular phase 0 or π is generated

via a phase shift $\widehat{\Omega}_i$ in Λ_h depending on $\uparrow$ or $\downarrow$ spin that the hole "exchanges" with as shown in Fig. 8. Although Λ_h also produces other phase shifts contributed by other spins not "exchanged" with the hole, their net effect is zero after summing up the total phase for a closed-path motion of the hole. In the end, one finds

$$|\Psi_{\rm G}\rangle \to (-1)^{N_h^{\downarrow}(c)}|\Psi_{\rm G}\rangle \tag{12}$$

to reproduce the phase string sign factor. It is straightforward to verify that the fermionic signs of the doped holes created by $\left(\sum_{ij} g_{ij}c_{i\uparrow}c_{j\downarrow}\right)^{N_h/2}$ in Eq. (4) can precisely account for the additional sign factor $(-1)^{N_h^h}$ in Eq. (7). Consequently, the nontrivial sign structure τ_c identified in the t–J model is naturally incorporated into the ground state $|\Psi_{\rm G}\rangle$ in Eq. (4) since the bosonic spin background $|{\rm RVB}\rangle$ does not contribute to any additional statistical signs.

Therefore, comparing with the Gutzwiller-projected BCS state (1), one finds that the new ground state (4) has the following distinct and unique features. First, it recovers the LDA ground state at half-filling, which is known as the most accurate variational wavefunction for the Heisenberg antiferromagnet. The Fermi statistics completely disappears in this limit, and the wavefunction satisfies the well-known Marshall signs. Second, at finite doping, the new ground state (4) satisfies the mutual statistics via the phase shift field $\widehat{\Omega}_i$ defined in Eq. (8), which reproduces the precise sign structure identified in the t–J model. The form of the wavefunction is therefore generally fixed by these two unique requirements derived from the t–J model, and the adjustable parameters φ_h, g_{ij}, and W_{ij} are presumably smooth functions. For a uniform solution with φ_h = constant, a d-wave g_{ij} with a short-ranged W_{ij} can be determined variationally based on the t–J model as mentioned in the previous section, in which the phase shift field in Λ_h plays a central role in turning the long-range RVB pairing into the short-range one at finite doping [7]. In the following we illustrate how the superconducting phase coherence can be simultaneously realized once the AFM long-range order in $|{\rm RVB}\rangle$ is turned into a spin liquid.

4 Superconducting phase coherence

Generally speaking, the Cooper pairing amplitude is already preformed in Eq. (4) (with the pairing symmetry determined by g_{ij}), but the true superconducting off-diagonal-long-range-order (ODLRO) will be determined by the phase coherence through $\widehat{\Omega}_i$ in Λ_h, which sensitively depends on the spin correlation in $|{\rm RVB}\rangle$. Now we examine the condition under which the ground state has a true superconducting ODLRO.

With the pre-existence of the Cooper pairing amplitude in Eq. (4), the superconducting order will be determined by [7]

$$\langle c_{i\uparrow}c_{j\downarrow}\rangle \propto \langle {\rm RVB}|\,{\rm e}^{{\rm i}(\widehat{\Omega}_i+\widehat{\Omega}_j)}|{\rm RVB}\rangle \tag{13}$$

We further note that, due to the presence of Λ_h, injecting a hole into the ground state $|\Psi_{\rm G}(N_h)\rangle$ in Eq. (4) will also induce a phase shift by

$$c_{i\sigma}|\Psi_{\rm G}(N_h)\rangle \sim {\rm e}^{{\rm i}\widehat{\Omega}_i}|\Psi_{\rm G}(N_h+1)\rangle \tag{14}$$

Namely the wavefunction overlap between the bare hole state and the true ground state of N_h+1 holes will depend on ${\rm e}^{{\rm i}\widehat{\Omega}_i}$. Hence the phase coherence of superconductivity and the coherence of a Landau (or more precisely, Bogoliubov) quasiparticle will be simultaneously realized. And a "normal state" obtained by disordering the phase shift factor ${\rm e}^{{\rm i}\widehat{\Omega}_i}$ will be intrinsically a non Fermi liquid with vanishing quasiparticle weight.

If a long-range RVB pairing (i.e., W_{ij} is in a power-law decay at large $|\boldsymbol{r}_{ij}|$) is present, like in the AFM ordered phase, the phase coherence in Eq. (13) will be generally destroyed because the $\pm\pi$ vortices in $\widehat{\Omega}_i$ carried by the two spinon partners of an RVB pair, according to Eq. (8), do not compensate each other, which results in phase disordering. Only can a short-ranged RVB pairing lead to a vortex–antivortex binding in Eq. (13) and thus the superconducting phase coherence. In other words, in the superconducting phase $|{\rm RVB}\rangle$ has to become a spin liquid with short-range AFM correlations. As previously mentioned, the RVB amplitude W_{ij} does become short-ranged with a finite, doping-dependent ξ based on a self-consistent calculation in terms of Eq. (4).

Furthermore, the statistical sign structure Λ_h, which should be generally present in the excited states of the t–J model as well, will play a critical role to control the superconducting phase transition. Imagine a pair of spinon excitations created by breaking up an RVB pair in $|{\rm RVB}\rangle$, which through the phase shift in Λ_h will lead to two $\pm\pi$ vortices according to Eq. (8). Normally the free vortices would disorder the superconducting phase coherence, and in order to maintain the latter below $T_{\rm c}$, a pair of excited spinons will be forced to form a vortex and antivortex bound pair with the spinons sitting at the vortex cores, which are known as the spin-rotons [15] as illustrated in Fig. 10. In contrast to the singlet RVB pair in the ground state, a spin-roton is a pair of loosely bound (confined) excited spinons via a logarithmic potential with the spin quantum numbers, $S = 0$ and 1 [15].

As detailed in Ref. [15], spin-roton excitations will eventually destroy the phase coherence at $T_{\rm c}$, and a simple $T_{\rm c}$ formula has been obtained as follows:

$$T_{\rm c} \simeq \frac{E_g}{6k_{\rm B}} \tag{15}$$

with $E_g \sim \delta J$ denoting the core energy of the spin-

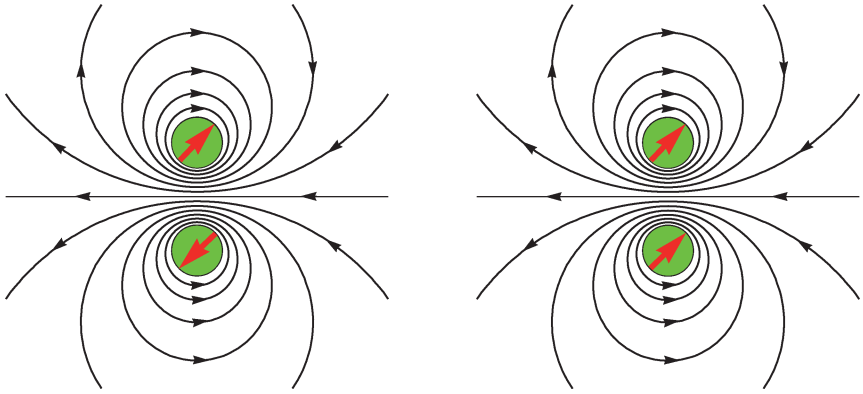


Fig. 10 Spin-rotons with $S = 0$ and $S = 1$, respectively, which are identified [15] as the most essential elementary excitations on top of the ground state (4) and decide the T_c-formula (15).

rotons, degenerate for $S = 0$ and 1. This relation is in excellent agreement with the empirical formula established in the experiment [16, 17], with $S = 0$ and $S = 1$ spin-rotons corresponding to the "resonancelike" modes observed in the Raman A_{1g} channel and neutron scattering measurements, respectively. It also provides a natural explanation of why the two modes are energetically degenerate in the experiment [16, 17].

Here the spin-rotons are the most essential elementary excitations of non-BCS-type in a superconductor of doped Mott insulators characterized by Eq. (4), which directly controls the superconducting phase coherence in Eq. (13) via a characteristic energy E_g. At $T > T_c$, spin-roton excitations (with $S = 0, 1$) will disassociate into free spinon-vortices ($S = 1/2$) as shown in Fig. 11. In the AFM long-range ordered state near half-filling, one has $E_g \to 0$ such that T_c vanishes, and the $S = 1$ spin-roton excitation will naturally reduce to the gapless spin wave with the RVB pairing becoming a long-ranged one. These non-superconducting states will be further discussed below.

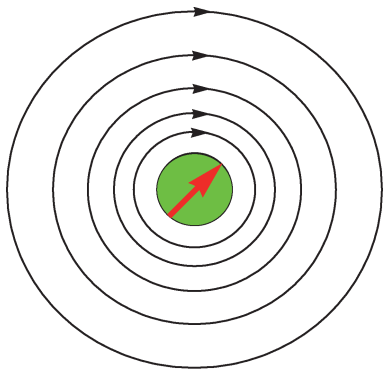


Fig. 11 A spinon-vortex is a composite object with an $S = 1/2$ free spin locking with a supercurrent vortex. It automatically appears in Eq. (4) if a single spinon is present in $|\mathrm{RVB}\rangle$. A spin-roton excitation in Fig. 10 can be regarded as a vortex-antivortex bound pair of the spinon-vortices.

5 Nature of pseudogap states

Generally, a crossover region lies between the AFM phase and superconducting phase in the phase diagram of the cuprates, which exhibits the so-called pseudogap phenomenon. Such a region is so unique to the cuprates and has been heavily investigated experimentally [18, 19] and theoretically [6, 10] in hoping to crack the high-T_c mystery.

The present superconducting ground state (4) is constructed using two criteria: It satisfies the general statistical sign structure of the t–J model and correctly reduces to the AFM state at half-filling. Therefore it can provide, in principle, an unequivocal prediction for the crossover region connecting the AFM and superconducting regimes.

By definition, a pseudogap state should be a non-superconducting version of Eq. (4) at finite doping. Indeed, the structure of Eq. (4) does suggest a series of non-superconducting "ground states" as outlined below.

Pseudogap Region I. As already pointed out in the previous section, at $T \gtrsim T_c$, spin-roton excitations will disassociate into free spinon-vortices (cf. Fig. 11) to disorder the superconducting phase coherence. The resulting vortex liquid state [10, 20, 21] provides a microscopic description of the Nernst regime discovered [22, 23] in the cuprates. Now to bring such a vortex liquid state to $T = 0$, one has to make $E_g \to 0$ in Eq. (15) or the equivalent RVB pairing in $|\mathrm{RVB}\rangle$ go to a long-ranged one to result in the right-hand-side of Eq. (13) vanishing. This is a quantum vortex ground state, in which the d-wave Cooper pair amplitude can be still maintained via g_{ij}, but the Cooper pairs apparently become incoherent because of phase disordering. On the other hand, although W_{ij} is in a power-law decay in $|\mathrm{RVB}\rangle$, the long-range AFM order is also disordered by Λ_h via the same phase shift fields due to the presence of doped holes in (incoherent) Cooper pairing. Such a peculiar "mutual duality" phase is called Bose insulator phase in Refs. [24–26], where the statistical sign structure in Λ_h has been formulated in terms of a mutual Chern–Simons gauge theory description.

In literature, a vortex liquid state has been usually considered [6, 27, 28] as the result of phase fluctua-

tions in the charge-2e d-wave pairing order parameter at low superfluid density. Pseudogap Region I apparently shares this similarity according to Eq. (13). However, I stress here that the electron fractionalization and mutual statistics further specify that each vortex has always a free $S = 1/2$ spinon sitting at the vortex core (cf. Fig. 11), and the charge-2e pairing order parameter is just a composite quantity involving both RVB and Cooper channels [7]. Namely the spin degrees of freedom control the charge behavior, including T_c [Eq. (15)] as well as the Nernst properties [10, 20, 21]. It is a unique characteristic of the mutual statistics that is clearly distinct from a conventional vortex liquid state.

Pseudogap Region II. One may obtain another non-superconducting ground state in Eq. (4) in the case that $|\mathrm{RVB}\rangle$ has become a short-ranged spin liquid with a finite E_g. In this case, the weaker Cooper pair amplitude may be made to disappear in g_{ij}, say, via strong magnetic fields, while the Zeeman energy still remains small as compared to E_g to let $|\mathrm{RVB}\rangle$ intact. Then in such a pseudogap state, a small Fermi pocket will emerge in Eq. (4) with the doped holes behave like free fermions moving in the vacuum of the spin liquid $|\mathrm{RVB}\rangle$. Because of the gap E_g in $|\mathrm{RVB}\rangle$, the phase coherence is still maintained in Λ_h at $T \ll E_g$ such that the small coherent Fermi pocket is expected to contribute to the quantum oscillation effect without involving a translational symmetry breaking. Here the pseudogap behavior will be mainly exhibited in $|\mathrm{RVB}\rangle$, which may be probed by the NMR experiments as discussed recently in Ref. [29].

It is noted that the state of Pseudogap Region II is similar to the dopon theory [30, 31] in which small Fermi pockets can also be induced [29] by strong magnetic fields. The main difference lies in the short-range RVB nature of $|\mathrm{RVB}\rangle$, which is different from a nodal fermion spin liquid state in the dopon theory [29–31]. On the other hand, a short-range bosonic spin liquid background with small Fermi pockets has been phenomenologically discussed in Ref. [32] recently.

Note that these pseudogap states generalized from Eq. (4) are not necessarily associated with any explicit symmetry-breaking orders. They emerge as a necessary crossover from the AFM state at half-filling towards the superconducting state at a finite doping. What is in common in these states is the presence of AFM correlations in $|\mathrm{RVB}\rangle$ in the form of RVB pairing. What differentiates Pseudogap Region I and II is whether there is a further preformed Cooper pairing or not in g_{ij}. Finally the pseudogap properties are gone if the RVB pairing disappears in $|\mathrm{RVB}\rangle$ at high temperatures, leading to a Curie–Weiss classical spin background [33]. The doped holes will get maximally scattered via the phase string effect in Λ_h as illustrated by Fig. 12(b), in contrast to the more coherent motion in the RVB regime [Fig. 12(a)]. As shown in Ref. [33], the resistivity in this regime will exhibit a linear-T "strange metal" behavior.

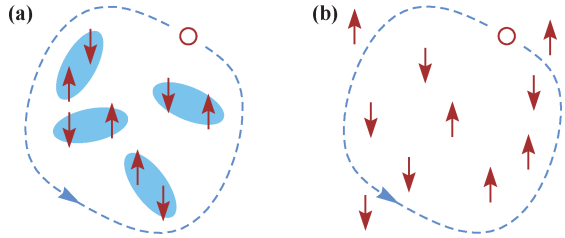


Fig. 12 **(a)** The motion of the doped hole is coherent in a short-range RVB background; **(b)** The strongest scattering between the hole and spin degrees of freedom occurs when the background spins are uncorrelated in the high-temperature "strange metal phase", where each spin carries a fictitious $\pm\pi$ fluxoids as perceived by the hole [cf. also Fig. 6(c)].

6 Concluding remarks

In the prototype RVB state (1), the Mottness simply implies the Gutzwiller projection onto a Hilbert space that excludes double occupancy of the electrons. Note that double occupancy is still allowed in $|\mathrm{BCS}\rangle$, which is basically a Fermi liquid with the Cooper channel instability. A fundamental question is if such a Gutzwiller projection $\widehat{P}_{\mathrm{G}}$ acting on a suitably chosen "ordinary" state $|\mathrm{BCS}\rangle$ (which may be considered as the ground state of an effective *local* Hamiltonian of fermions) is sufficient to capture the essential physics of the Mottness?

Through this article I have pedagogically reviewed the efforts [7, 10, 11] to reexamine this important question. The basic conclusion is that, as a matter of principle, $|\mathrm{BCS}\rangle$ cannot be "ordinary" as described above, and the singular "feedback" effect of the constrained Hilbert space should turn the effective Hamiltonian of $|\mathrm{BCS}\rangle$, if it exists, into a new paradigm of non-Fermi–Dirac statistics. In other words, we have reached the conclusion that the essential physics of the Mottness is not simply about imposing the constraint (2) on the Hilbert space. Rather, an electron fractionalization with mutual fractional statistics must take place in $|\mathrm{BCS}\rangle$ as a general consequence of the Mottness [11, 13].

The resulting new ground state [7] as given in Eq. (4) is distinct from the original RVB state [1] in Eq. (1) by that the neutral RVB and Cooper channels are clearly differentiated. It implies an electron fractionalization in which three essential types of correlations are explicitly separated and intrinsically embedded: i.e., the AFM correlations in $|\mathrm{RVB}\rangle$; the Cooper pairing of the doped holes; and the mutual influence/competition between these two channels via a mutual statistical phase in Λ_h.

As emphasized in this article, an accurate description of the AFM state is a crucial starting point. So does the singular doping effect, which brings in the fundamental change in the statistical sign structure known as the phase string effect [12–14]. Based on these two basic properties of the Mott physics, the pseudogap

phenomenon and superconductivity arise naturally as a self-organization of the three essential types of correlations mentioned above. Without holes, the long-range AFM order will always win in $|\mathrm{RVB}\rangle$, which provides a highly accurate description [8] of the ground state of the Heisenberg Hamiltonian. A sufficient concentration of doped holes will eventually turn the antiferromagnetism in $|\mathrm{RVB}\rangle$ into a true spin liquid state with short-range AFM correlations. By doing so the Cooper paired holes can gain phase coherence to realize a high-T_{c} superconductivity. The pseudogap region is merely a crossover from the long-range AFM order to superconducting order since the latter two are incompatible in nature. This incompatibility decides the complex crossover phenomenon of the pseudogap physics [10]. There is nothing mysterious about them once the essence of the Mottness is properly understood.

Deep inside the superconducting state, the spins are short-range RVB paired in $|\mathrm{RVB}\rangle$ such that the statistical signs get cancelled in Λ_h, and the state (4) resembles Eq. (1) proposed by Anderson [1] in that the Cooper pairing is driven by the spin RVB pairing upon doping. The d-wave pairing symmetry and the RVB state as a prototype pseudogap phase were actually predicted [34, 35] before experimental discoveries. In fact, in a modified version of Eq. (1), the "fugacity" has been also introduced to stress the tendency towards a separation of the neutral RVB and Cooper pairings [36]. Nevertheless, the statistical signs in Λ_h of Eq. (4) will play a critical role in dictating how the spin excitation affects the superconductivity and *vice versa*. In order to properly understand the superconducting phase transition, the nature of pseudogap physics, and even the strange metal behavior at high temperatures, the mutual statistics encoded in Λ_h cannot be omitted in a theory based on the doped Mott insulator approach.

Acknowledgements I acknowledge useful discussions with P. W. Anderson, W.-Q. Chen, Z.-C. Gu, S.-P Kou, T. Li. J.-W. Mei, V. N. Muthukumar, H. T. Nieh, N. P. Ong, X.-L. Qi, Y. Qi, D.-N. Sheng, C.-S. Tian, Y.-Y. Wang, X.-G. Wen, K. Wu, P. Ye, J. Zaanen, F. Zhou, and Y. Zhou. I am grateful to Y.-Z. You's help with the figures. This work was supported by the National Natural Science Foundation of China (Grant No. 10834003) and the National Program for Basic Research of Ministry of Science and Technology (Grant Nos. 2009CB929402 and 2010CB923003).

References

1. P. W. Anderson, Science, 1987, 235(4793): 1196
2. P. W. Anderson, P. A. Lee, M. Randeria, T. M. Rice, N. Trivedi, and F. C. Zhang, J. Phys.: Condens. Matter, 2004, 16(24): R755, and references therein
3. For a review, see: B. Edegger, V. N. Muthukumar, and C. Gros, Adv. Phys., 2007, 56(6): 927
4. D. Vaknin, S. K. Sinha, D. E. Moncton, D. C. Johnston, J. M. Newsam, C. R. Safinya, and H. King, Phys. Rev. Lett., 1987, 58(26): 2802
5. A. Auerbach, Interacting Electrons and Quantum Magnetism, New York: Springer-Verlag, 1994
6. For a review, see: P. A. Lee, N. Nagaosa, and X. G. Wen, Rev. Mod. Phys., 2006, 78(1): 17
7. Z. Y. Weng, New J. Phys., 2011, 13: 103039; arXiv: 1105.3027, 2011, and references therein
8. S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett., 1988, 61(3): 365
9. W. Marshall, Proc. R. Soc. Lond. A, 1955, 232(1188): 48
10. For a review, see: Z. Y. Weng, Int. J. Mod. Phys. B, 2007, 21: 773; arXiv:0704.2875, 2007
11. J. Zaanen and B. J. Overbosch, Phil. Trans. R. Soc. A, 2011, 369: 1599; arXiv:0911.4070, 2009
12. D. N. Sheng, Y. C. Chen, and Z. Y. Weng, Phys. Rev. Lett., 1996, 77(25): 5102
13. Z. Y. Weng, D. N. Sheng, Y.-C. Chen, and C. S. Ting, Phys. Rev. B, 1997, 55(6): 3894
14. K. Wu, Z. Y. Weng, and J. Zaanen, Phys. Rev. B, 2008, 77(15): 155102
15. J. W. Mei and Z. Y. Weng, Phys. Rev. B, 2010, 81(1): 014507
16. Y. J. Uemura, J. Phys.: Condens. Matter, 2004, 16(40): S4515
17. Y. J. Uemura, Physica B, 2006, 374-375: 1
18. T. Timusk and B. Statt, Rep. Prog. Phys., 1999, 62(1): 61
19. A. Damascelli, Z. Hussin, and Z. X. Shen, Rev. Mod. Phys., 2003, 75(2): 473
20. Z. Y. Weng and V. N. Muthukumar, Phys. Rev. B, 2002, 66(9): 094509
21. Z. Y. Weng and X. L. Qi, Phys. Rev. B, 2006, 74(14): 144518
22. Z. A. Xu, N. P. Ong, Y. Wang, T. Kakeshita, and S. Uchida, Nature, 2000, 406(6795): 486
23. Y. Wang, Z. A. Xu, T. Kakheshita, S. Uchida, S. Ono, Y. Ando, and N. Ong, Phys. Rev. B, 2001, 64(22): 224519
24. P. Ye, C. S. Tian, X. L. Qi, and Z. Y. Weng, Phys. Rev. Lett., 2011, 106(14): 147002
25. P. Ye, C. S. Tian, X. L. Qi, and Z. Y. Weng, Nucl. Phys. B, 2012, 854[FS]: 815; arXiv:1106.1223, 2011
26. S. P. Kou, X. L. Qi, and Z. Y. Weng, Phys. Rev. B, 2005, 71(23): 235102
27. V. J. Emery and S. A. Kivelson, Nature, 1995, 374(6521): 434
28. Z. Tešanović, Nat. Phys., 2008, 4: 408
29. J. W. Mei, S. Kawasaki, G. Q. Zheng, Z. Y. Weng, and X. G. Wen, arXiv:1109.0406, 2011
30. T. C. Ribeiro and X. G. Wen, Phys. Rev. B, 2006, 74(15): 155113
31. T. C. Ribeiro and X. G. Wen, Phys. Rev. Lett., 2005, 95(5): 057001
32. Y. Qi and S. Sachdev, Phys. Rev. B, 2010, 81(11): 115129
33. Z. C. Gu and Z. Y. Weng, Phys. Rev. B, 2007, 76(2): 024501
34. F. C. Zhang, C. Gros, T. M. Rice, and H. Shiba, Supercond. Sci. Technol., 1988, 1(1): 36
35. G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun., 1987, 63(11): 973
36. P. W. Anderson, arXiv:1011.2736, 2010