U(1)×SU(2) Chern-Simons gauge theory of underdoped cuprate superconductors

P. A. Marchetti
Dipartimento di Fisica ''G. Galilei'', INFN, U. of Padova, I-35131 Padova, Italy

Zhao-Bin Su
Institute of Theoretical Physics, CAS, Beijing 100080, China

Lu Yu
International Centre for Theoretical Physics, I-34100 Trieste, Italy
and Institute of Theoretical Physics, CAS, Beijing 100080, China

(Received 15 January 1998)

The Chern-Simons bosonization with U(1)×SU(2) gauge field is applied to the two-dimensional t-J model in the limit $t \gg J$, to study the normal-state properties of underdoped cuprate superconductors. We prove the existence of an upper bound on the partition function for holons in a spinon background, and we find the optimal spinon configuration saturating the upper bound on average—a coexisting flux phase and $s+id$-like resonating-valence-bond state. After neglecting the feedback of holon fluctuations on the U(1) field B and spinon fluctuations on the SU(2) field V, the holon field is a fermion and the spinon field is a hard-core boson. Within this approximation we show that the B field produces a π flux phase for the holons, converting them into Dirac-like fermions, while the V field, taking into account the feedback of holons produces a gap for the spinons vanishing in the zero-doping limit. The nonlinear-σ model with a mass term describes the crossover from the short-ranged antiferromagnetic (AF) state in doped samples to long-range AF order in reference compounds. Moreover, we derive a low-energy effective action in terms of spinons, holons and a self-generated U(1) gauge field. Neglecting the gauge fluctuations, the holons are described by the Fermi-liquid theory with a Fermi surface consisting of four ''half-pockets'' centered at $(\pm\pi/2,\pm\pi/2)$ and one reproduces the results for the electron spectral function obtained in the mean-field approximation, in agreement with the photoemission data on underdoped cuprates. The gauge fluctuations are not confining due to coupling to holons, but nevertheless yield an attractive interaction between spinons and holons leading to a bound state with electron quantum numbers. The renormalization effects due to gauge fluctuations give rise to non-Fermi-liquid behavior for the composite electron, in certain temperature range showing the linear in T resistivity. This formalism provides a new interpretation of the spin gap in the underdoped superconductors (mainly due to the short-ranged AF order) and predicts that the minimal gap for the physical electron is proportional to the square root of the doping concentration. Therefore the gap does not vanish in any direction. All these predictions can be checked explicitly in experiment. [S0163-1829(98)01233-8]

I. INTRODUCTION

A. Physical issue to be addressed

The proximity of superconductivity (SC) to antiferromagnetism (AF) in reference compounds is a distinct feature of the high-T_c superconductors. Upon doping the AF goes away, giving rise to SC. At the same time, the Fermi surface (FS) is believed to develop from small pockets around $(\pm\pi/2,\pm\pi/2)$,[1] anticipated for a doped Mott insulator, to a large one around (π,π), expected from the electronic structure calculations[2] and confirmed by the angle-resolved photoemission spectroscopy (ARPES) experiments.[3] To understand this crossover is one of the key issues in resolving the high-T_c puzzle. For this reason, the underdoped samples present particular interest due to the strong interplay of SC with AF.

There is a consensus now that these systems are strongly anisotropic, and the fundamental issue is to understand the behavior of strongly correlated electrons in the copper-oxygen plane.[4] A ''spin gap'' or ''pseudogap'' has been invoked to explain[5] the reduction of magnetic susceptibility χ below certain characteristic temperature T^* (Ref. 6) and suppression of the specific heat compared with the linear-T behavior.[7] This gap also shows up in transport properties,[8] neutron scattering,[9] and NMR relaxation rate[6] measurements. The recent ARPES experiments[10] on underdoped samples seem to indicate that the FS of these compounds is probably half-pocket-like, i.e., a small pocket near Σ point $(\pi/2,\pi/2)$, but lacking its outer part in the reduced Brillioun-zone scheme. These data show clear Fermi level crossing in the (0,0) to (π,π) direction, but no such crossing was detected in the $(0,\pi)$ to (π,π) direction.[10] The observed pseudogap above T_c is consistent with d-wave symmetry. In this paper we will be concerned with the normal-state properties of these underdoped cuprate superconductors, focusing on the implications derived from the proximity of these systems to the AF reference state.

B. A brief survey of related theoretical approaches

Theoretically there have been mainly two competing approaches: One starting from the Mott-Hubbard insulator, ad-

vocated by Anderson[11,4] using the concept of spin liquid, or the resonating-valence-bond (RVB) state, while the other starting from the more conventional Fermi liquid (FL) point of view.

One of the crucial concepts within the first approach is ''spin-charge separation'' which can be intuitively implemented by introducing ''slave bosons,''[12] namely, one rewrites the electron operator:

$$\psi_{i\sigma}=e_i^\dagger f_{i\sigma},$$

where e_i is a charged spinless (slave) boson operator (holon), while $f_{i\sigma}$ is a neutral, spin-1/2 fermion operator (spinon) satisfying constraint

$$e_i^\dagger e_i+f_{i\sigma}^\dagger f_{i\sigma}=1. \tag{1.1}$$

(Hereafter the repeated spin indices are summed over.) One can also interchange the role of boson and fermion operators, i.e., to introduce a spinless fermion to describe the charge degree of freedom, while ''spinning'' bosons to describe the spin degree of freedom. This is the ''slave-fermion'' approach.[13,14] The essential requirement for both approaches is the ''single occupancy'' constraint which is very difficult to implement. In the mean-field approximation (MFA) which satisfies the constraint (1.1) only on average, the slave-boson[15] and slave fermion[14,16] approaches gave very different phase diagrams and each of them has its own difficulties.[17] There have been several attempts to improve the situation,[18] but the basic difficulty still remains.

Moreover, in decomposing the physical electron into a product of fermion and boson, one increases the number of degrees of freedom (dof) by two. The constraint (1.1) takes care of one, but there is one extra dof which corresponds to the spinon-holon gauge field. In fact, the physical electron operator is invariant under the transformation:

$$e_j\rightarrow e_j e^{i\xi_j},\quad f_{j\sigma}\rightarrow f_{j\sigma}e^{i\xi_j},$$

so one can ''gauge-fix'' ξ_j according to the choice. This is the starting point of the gauge-field approach to the strongly correlated electron systems.[19,20] This approach has been systematically pursued by Lee and his collaborators first as the U(1) gauge theory,[21] and recently by considering the SU(2) gauge group.[22]

From the FL point of view the interplay of AF with SC, and the evolvement of the FS with doping has been elaborated by Kampt and Schrieffer,[23] and by Chubukov and his collaborators.[24] Very recently, Zhang[25] has proposed an interesting SO(5) model to consider the AF-SC interplay.

C. Basic idea of Chern-Simons (C-S) bosonization

In this paper we will use the C-S bosonization as the basic technical tool. The procedure of reformulating the fermion problem in terms of bosons was pioneered by the Jordan-Wigner transformation[26]

$$c_j^\dagger=a_j^\dagger e^{-i\pi\Sigma_{l<j}a_l^\dagger a_l}, \tag{1.2}$$

where c_j is a fermion operator, while a_j is a hard-core boson operator on a linear chain. The Abelian bosonization for one-dimensional (1D) fermions with linear dispersion[27] is similar in spirit and has found extensive applications in condensed-matter physics.[28,29] The key formula there is the Mandelstam representation (in field-theoretical jargons):[27,30]

$$\psi_\alpha(x)\sim e^{(i/\sqrt{\pi})\int_{-\infty}^{x}dy\,\dot{\phi}(y)-(-1)^\alpha i\sqrt{\pi}\phi(x)}, \tag{1.3}$$

where ψ is the fermion operator, while ϕ is the boson operator, and $\alpha=1,2$. (Here we have omitted the normal ordering of the operators.) A rigorous derivation of Eq. (1.3) in terms of path integrals was given in Ref. 31. This bosonization procedure was generalized by Witten to the non-Abelian case,[32] which reformulates the $1+1$-dimensional relativistic fermion problem with symmetry group G as the G-valued nonlinear-σ (NLσ) model. That scheme has been extensively used for the study of quantum spin systems.[33]

The Abelian bosonization procedure has been generalized to $2+1$-dimensional systems.[34] It is in some sense analogous to the Jordan-Wigner transformation. The typical relation is given as

$$\psi(x)\sim\phi(x)e^{i\int_{\gamma_x}A_\mu(y)dy^\mu}, \tag{1.4}$$

where $\psi(x)$ is the fermion field operator, while $\phi(x)$ is the boson field operator, and the two are related by the C-S U(1) gauge field operator A_μ. The integration in the exponent is taken over an arbitrary path γ_x in the 2D plane, running from x to ∞. The path integral will contain an extra factor $e^{-kS_{c.s.}}$ with the C-S action

$$S_{c.s.}=\frac{1}{4\pi i}\int d^3x\,\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho,$$

and the C-S coefficient (level) $k=1/(2l+1)$, $l=0,1,2\ldots$. For nonrelativistic fermions the C-S constraint can be solved explicitly, and the transformation (1.4) becomes[34]

$$\psi(x)\sim\phi(x)e^{i(2l+1)\int d^2y\,\Theta(x-y)\phi^\dagger(y)\phi(y)}, \tag{1.5}$$

where $\Theta(x-y)=\arctan[(x^2-y^2)/(x^1-y^1)]$. An interesting application of this formula is the analysis of the fractional quantum Hall effect at filling $\nu=1/(2l+1)$ in terms of boson liquids.[35] The statistical transmutation, implemented by the abelian C-S gauge field is a consequence of the Aharonov-Bohm effect, and it is limited to Abelian fractional statistics, characterized by the phase factor $2\pi\theta$ with $\theta\in[0,1)$.[36] In particular, $\theta=0$, 1/2 corresponds to boson and fermion, respectively, while $\theta=1/4$ corresponds to the ''semion'' case, advocated by Laughlin, as a constituent quasiparticle in high-temperature superconductors.[37]

The $2D$ analog of the Jordan-Wigner formula (1.4) and (1.5) was originally derived as an operator identity.[34] It was later on justified in the path-integral form and generalized to the non-Abelian case.[38] That paper was further extended to include the correlation functions, elaborated for the case $G=\mathrm{U(1)}\times\mathrm{SU(2)}$ and applied to the t-J model[39] (see also Ref. 40). Readers are referred to those references for a detailed presentation. To make the present paper more self-contained, we briefly outline here the basic idea of such a bosonization procedure.

Consider a system of N spin-1/2 fermions (or bosons) in two space dimensions, in an external (Abelian) gauge field A. The canonical partition function in the first quantized path-integral representation is given by[41]

$$Z(A)=\sum_{\alpha_1,\ldots,\alpha_N}\int dx_1\cdots dx_N\sum_{\pi}(-1)^{\sigma(\pi)}\times\prod_{r=1}^{N}\left(\int_{\omega_r(0)=x_r}^{\omega_r(\beta)=x_{\pi(r)}}D\omega_r(\tau)e^{-(m/2)\int\dot{\omega}_r^2(\tau)}\right)e^{i\int_{\varphi}A},\tag{1.6}$$

where $\alpha_j=1,2$ for $j=1,2,\ldots,N$ are the spin indices, $\omega_j(\tau)$ represent ''virtual trajectories'' of particles going from imaginary time $\tau=0$ to $\tau=\beta=1/T$, with T as the temperature (we set the Boltzmann constant $k_B=1$) and reaching the plane $\tau=\beta$ at the same set of points where they start at $\tau=0$, the points being arbitrarily permuted. Due to the fermion statistics of the particles, there is a factor $(-1)^{\sigma(\pi)}$, associated with each permutation π, where $\sigma(\pi)$ is the number of exchanges in the permutation π. These trajectories have vanishing probability to intersect each other for a given τ. Each set of such trajectories appearing in the first quantized representation form a link, i.e., a set of possibly interlaced loops, when the $\tau=0$ and $\tau=\beta$ planes are identified by periodicity in time. Here $\int_{\varphi}A=\int(A_l d\omega^l+A_0 d\tau)$ is a line integral in $2+1$ dimensions, $l=1,2$. The partition function for the boson system would be the same, except that $\sigma(\pi)$ is replaced with 0. On the other hand, the factor $(-1)^{\sigma(\pi)}$ is a topological invariant naturally associated with the link and, according to a general theory,[42] it can be represented as the expectation value of a ''Wilson loop'' (trace of a gauge phase factor) supported on that link in a gauge theory with a suitably chosen C-S action.

$$(-1)^{\sigma(\pi)}=\int DVe^{-kS_{c.s.}(V)}P(e^{i\int_{\varphi}V}),$$

where V is a C-S gauge field with symmetry group G, k is the C-S coefficient (level), already defined in the Abelian case, $P(\cdot)$ is the path ordering, identical to the time ordering, if ''time'' is parametrizing the path, and

$$S_{c.s.}=\frac{1}{4\pi i}\int d^3x\ \mathrm{Tr}\left[\epsilon^{\mu\nu\rho}\left(V_\mu\partial_\nu V_\rho+\frac{2}{3}V_\mu V_\nu V_\rho\right)\right].\tag{1.7}$$

As a consequence, there is a boson-fermion relation for the canonical partition function:

$$Z_F(A)=\int DVZ_B(A+V)e^{-kS_{c.s.}(V)}.$$

The bosonization formula is written in the second-quantized form for the grand-canonical partition function:

$$\Xi=\sum_N\frac{e^{\beta\mu N}}{N!}Z_N,$$

and

$$\Xi_F(A)=\int D\Psi D\Psi^* e^{-S(\Psi,\Psi^*,A)}=\int D\Phi D\Phi^* DVe^{-[S(\Phi,\Phi^*,A+V)+kS_{c.s.}(V)]},\tag{1.8}$$

where Ψ is the Grassmann variable representing the fermion field, while Φ is the complex variable representing the boson field.

There might be different choices of G, k: In particular, $G=\mathrm{U}(1)$, $k=1$ corresponds to the Abelian C-S bosonization,[34] while $G=\mathrm{U}(1)\times\mathrm{U}(1)$, $k=[\theta,(\theta+2)/(\theta+1)]$ corresponds to the ''anyon'' bosonization,[43] with θ as the statistics parameter. In this paper we will concentrate on the case $G=\mathrm{U}(1)\times\mathrm{SU}(2)$ with $k=2,1$, respectively.[38,39]

The relations for the correlation functions are given by

$$\Psi(x)\rightarrow P(e^{i\int_{\gamma_x}V})\Phi(x),\quad \Psi^*(x)\rightarrow\Phi^*(x)P(e^{-i\int_{\gamma_x}V})\tag{1.9}$$

with γ_x as a straight line in the fixed time plane joining x with ∞. This is a non-Abelian generalization of the 2D formula for the Abelian case (1.4). It is important to notice the nonlocal character of these relations.

Here we have used the first-quantized form of the path integral to identify the relation between the fermion and boson systems, while the ''working'' formula for bosonization is given in the second-quantized form of the path-integral representation. This switching from first to second quantized form and vice versa will be frequently used throughout this paper.

D. Outline of the paper content

The $\mathrm{U}(1)\times\mathrm{SU}(2)$ C-S bosonization approach has been successfully employed by us earlier[44] to calculate the critical exponents of the correlation functions in the 1D t-J model in the limit $t\gg J$. Although, in principle, all bosonization schemes should yield an exact identity between the correlation functions of the original fermionic field and corresponding bosonic correlation functions, the MFA, as mentioned earlier, gives different results in different bosonization schemes. The U(1) C-S bosonization has been shown to correspond essentially to the slave-boson and slave-fermion approaches (depending on the choice of the gauge fixing); while the non-Abelian $\mathrm{U}(1)\times\mathrm{SU}(2)$ C-S bosonization corresponds to the slave-semion approach ($\theta=1/4$).[38] We have shown[44] that the ''semion'' spin-charge separation of spinon and holon is the correct one to reproduce the exact exponents known from the Bethe ansatz solution and the Luttinger-liquid–conformal field theory calculations.[45]

We considered 1D fermion system on the background of a 2D $\mathrm{U}(1)\times\mathrm{SU}(2)$ C-S gauge field. The U(1) field is related to the charge, while the SU(2) field is related to the spin degrees of freedom. Performing the dimensional reduction and using the freedom in gauge fixing, we could analyze the original problem as an optimization process for the partition function of holons in a spinon background. We could find an upper bound for the partition function and an exact way to saturate this bound, without any approximations. Afterwards we used MFA to consider an ''averaged'' holon configuration to compute the long-time, large distance behavior of the correlation functions, reproducing the exact results. The important lesson we learned there is that the statistical (semionic) properties of the constituent particles (spinons and holons) due to gauge field fluctuations are crucial. In the operator form, the original fermion operator can be decomposed as

$$\psi_x = h_x e^{i\,\pi/2\Sigma_{l>x} h_l^* h_l} [e^{i\,\pi/2\Sigma_{l<x} b_l^* b_l} b_x], \qquad (1.10)$$

where h_x is a fermion, while b_x is a hard-core boson operator. However, only h_x together with the attached ''holon string'' (the exponentiated operator) represents a ''physical'' charged, spinless holon, whereas b_x together with the attached ''spinon string'' corresponds to the neutral, spin-$\frac{1}{2}$ spinon. Both of them satisfy the ''semionic'' equal-time commutation relations:

$$f(x^1) f(y^1) = e^{\pm i\pi/2} f(y^1) f(x^1), \quad x^1 \gtrless y^1,$$

where f represents fermion (boson) operator along with the attached string. This means that the ''semionic'' spinons and holons are ''deconfined,'' and the spin and charge are fully separated. The dynamics of the spinon filed b_x is described by an O(3) NLσ model. Due to the presence of the topological term in 1D,[28,29] the spinons are massless, and they are ''deconfined,'' i.e., the spinons themselves, not their bound states (the usual spin waves in higher dimensions) are the constituent quasiparticles. These results encourage us to explore the 2D case which is of much more physical importance. We should, however, carefully distinguish which are the generic features of the C-S gauge field theory under consideration, and what is specific for the 1D case.

In this paper we employ the U(1)×SU(2) C-S bosonization scheme to study the 2D t-J model in the underdoped regime in the limit $t \gg J$. We will try to follow as much as possible the same procedure as in 1D. The U(1) gauge field B is again related to the charge degree of freedom, while the SU(2) gauge field V is related to the spin degree of freedom. First we prove the existence of an upper bound of the partition function for holons in a spinon background, and we find the optimal, holon-dependent spinon configuration which saturates the upper bound in an average sense. The optimization arguments suggest coexistence of a flux phase[46] with an $s+id$-like RVB state,[47] where the expectation value for the Affleck-Marston (AM) bond variable of spinons is close to 1, while the $s+id$-RVB order parameter is much smaller than 1. Then we make an approximation, neglecting the feedback of holon fluctuations on the U(1) field B and spinon fluctuations on the SU(2) field V. Hence the holon field is a fermion and the spinon field a hard-core boson. Within this approximation we show that the B field produces a π-flux phase for the holons, converting them into Dirac-like fermions, while the V field, taking into account the feedback of holons produces a gap for the spinons, minimal at $(\pm\pi/2, \pm\pi/2)$. The spinons are described by a NLσ model with a mass term (gap) $\simeq \sqrt{-\delta \ln \delta}$, with δ as the doping concentration. This corresponds to a short-ranged AF order (or disordered state in the jargons of NLσ model) in doped samples, which crosses over to the long-ranged AF order in the prestine samples, when the gap vanishes. This is an attempt to include AF fluctuations self-consistently in the RVB-type approach. Moreover, we derive a low-energy effective action in terms of spinons, holons, and a self-generated U(1) gauge field. Neglecting the gauge fluctuations, the holons are described by the FL theory with a FS consisting of four ''half-pockets'' centered at $(\pm\pi/2, \pm\pi/2)$ and one reproduces the results for the electron spectral function obtained in MFA (for the coexisting π-flux and d-wave RVB state),[48] in qualitative agreement with the ARPES data[10] for underdoped cuprates. If the gauge field were coupled to the spinons alone, it would be confining, since the spinons are massive. However, due to coupling to the massless branch of holons (which are actually nonrelativistic because of a finite FS), gauge fluctuations are *not* confining, but nevertheless yield an attractive interaction between spinons and holons leading to a bound state in 2D with electron quantum numbers. This could explain why neglecting the ''semionic'' nature of spinons and holons is less dangerous in 2D than in 1D. This means that the spin and charge are not fully separated (like in 1D), showing up as bound states in low-energy phenomena. The renormalization effects due to gauge fluctuations would lead to non-FL behavior for the composite electron, in certain temperature range showing the linear in T resistivity discussed earlier.[21,49,50] This formalism provides a new interpretation of the spin gap in the underdoped superconductors (mainly due to the short-ranged AF order) and predicts that the minimal gap for the physical electron is proportional to the square root of the doping concentration. Therefore the gap does not vanish in any direction. All these predictions can be checked explicitly in experiment.

The C-S gauge field approach has also been used by Mavromatos and his collaborators[51] to study the anyon superconductivity, advocated by Laughlin.[37] To our understanding, the basic aim of their work is to construct a model exhibiting semion superconductivity without breaking the time-reversal and parity symmetry. In spite of some apparent similarities in formulas, the main issue considered and the basic physical assumptions in their work are very different from ours. They have also discussed the normal-state properties,[52] but the mechanism for a possible non-FL behavior in their paper differs from what we consider here. We should also mention that the SU(2) gauge field considered in their recent papers (quoted in Ref. 51) corresponds to a generalization of the local SU(2) symmetry at half-filling, and is not related to the spin rotational symmetry, as we discuss in this paper. Lee and his collaborators[22] have also been considering this (rather than the spin) SU(2) group.

The present paper is an extended version of the earlier short communication.[53] The rest of the paper is organized as follows. In Sec. II we summarize the U(1)×SU(2) bosonization in the context of 2D t-J model; In Sec. III we present the optimization problem for the spinon configuration; In Sec. IV we derive the spinon effective action; In Sec. V we consider the holon effective action; In Sec. VI we make some concluding remarks. The proof of the bound employed in Sec. III, is deferred to the Appendix.

II. U(1)×SU(2) CHERN-SIMONS BOSONIZATION OF THE t-J MODEL

A. The model Hamiltonian

It is widely believed that the 2D t-J model captures the essential physical properties of the Cu-O planes characterizing a large class of high-T_c superconductors.[4] The Hamiltonian of the model is given by

$$H=P_G\sum_{\langle ij\rangle}\Big[-t(\psi^\dagger_{i\alpha}\psi_{j\alpha}+\text{H.c.}) + J\psi^\dagger_{i\alpha}\frac{\boldsymbol{\sigma}_{\alpha\beta}}{2}\psi_{i\beta}\cdot\psi^\dagger_{j\gamma}\frac{\boldsymbol{\sigma}_{\gamma\delta}}{2}\psi_{j\delta}\Big]P_G, \tag{2.1}$$

where $\boldsymbol{\sigma}_{\alpha\beta}$ are the Pauli matrices, $\psi_{i\alpha}$ is the annihilation operator of a spin-1/2 electron on site i of a square lattice, corresponding to creating a hole on the Cu site, and P_G is the Gutzwiller projection eliminating double occupation, modeling the strong on-site Coulomb repulsion. To simplify notations, we introduce a two-component spinor $\psi_i=(\psi_{i\uparrow},\psi_{i\downarrow})$. Throughout this paper, the small letters will denote operators, while the capital letters will denote the corresponding complex (or Grassmann) variables in the path-integral representation, unless otherwise specified.

Using the Hubbard-Stratonovich transformation to introduce a complex gauge field $X_{\langle ij\rangle}$, the grand-canonical partition function of the t-J model at temperature $T=1/\beta$ and chemical potential μ can be rewritten as[46]

$$\Xi_{t\text{-}J}(\beta,\mu)=\int \mathcal{D}X\mathcal{D}X^*\mathcal{D}\Psi\mathcal{D}\Psi^* e^{-S_{t\text{-}J}(\Psi,\Psi^*,X,X^*)} \tag{2.2}$$

with

$$S_{t\text{-}J}(\Psi,\Psi^*,X,X^*)=\int_0^\beta dx^0\Big\{\sum_{\langle ij\rangle}\Big(\frac{2}{J}X^*_{\langle ij\rangle}X_{\langle ij\rangle}+[(-t+X_{\langle ij\rangle})\Psi^*_{i\alpha}\Psi_{j\alpha}+\text{H.c.}]\Big)+\sum_i\Psi^*_{i\alpha}(\partial_0+\mu)\Psi_{i\alpha} +\sum_{i,j}u_{i,j}\Psi^*_{i\alpha}\Psi^*_{j\beta}\Psi_{j\beta}\Psi_{i\alpha}\Big\}, \tag{2.3}$$

where the two-body potential is given by

$$u_{i,j}=\begin{cases}+\infty & i=j\\ -\dfrac{J}{4} & i,j \text{ nearest neighbors}\\ 0 & \text{otherwise.}\end{cases} \tag{2.4}$$

Hereafter we denote the Euclidean time $x^0\equiv\tau$ ($\partial_0\equiv\partial/\partial\tau$) and its dependence of the fields is not explicitly spelled out.

B. C-S bosonization

Comparing Eqs. (2.2), (2.3) with Eq. (1.8), we find that the C-S bosonization procedure, briefly introduced in Sec. IC can be applied to rewrite the grand-canonical partition function (2.2). However, there is an important difference, namely, the consideration in Sec. IC was for the 2D continuum, where the probability for two world lines (Brownian paths) to intersect each other at a given time is zero. This is not true for the lattice case we consider now, where the probability for two paths to cross each other at a given time is not vanishing. On the other hand, the model we consider contains a single-occupancy constraint, expressed in terms of the Gutzwiller projection operator P_G, or the infinite on-site repulsion $u_{i,i}$, which excludes the intersection of paths. Therefore, we can still apply the C-S bosonization scheme, leaving the C-S gauge fields in the continuum, while considering the matter field on a discrete lattice. We will introduce an Abelian U(1) gauge field B related to the charge degree of freedom and a SU(2) gauge field V related to the spin degrees of freedom. The Euclidean C-S actions for these fields are given by

$$S_{c.s.}(B)=\frac{1}{4\pi i}\int d^3x\,\epsilon^{\mu\nu\rho}B_\mu\partial_\nu B_\rho, \tag{2.5}$$

$$S_{c.s.}(V)=\frac{1}{4\pi i}\int d^3x\,\mathrm{Tr}\Big[\epsilon^{\mu\nu\rho}\Big(V_\mu\partial_\nu V_\rho+\frac{2}{3}V_\mu V_\nu V_\rho\Big)\Big],$$

where $V_\mu=V^a_\mu\sigma_a/2$, $a=1,2,3$, $\mu=0,1,2$ with σ_a as Pauli matrices.

In the fermion-boson transformation formula (1.8) $k=2$ for the U(1) field B, and $k=1$ for the SU(2) field V.[38,39] The correlation functions for the Grassmann fields $\Psi_\alpha(\Psi^*_\alpha)$ are substituted by the correlation functions of the gauge-invariant complex fields $\Phi_\alpha(\gamma_x)$ $[\Phi^*_\alpha(\gamma_x)]$, defined as

$$\Phi_\alpha(\gamma_x)=e^{i\int_{\gamma_x}B}(Pe^{i\int_{\gamma_x}V})_{\alpha\beta}\Phi_{x\beta},$$

$$\Phi^*_\alpha(\gamma_x)=\Phi^*_{x\beta}(Pe^{-i\int_{\gamma_x}V})_{\beta\alpha}e^{-i\int_{\gamma_x}B}.$$

As mentioned earlier, γ_x is a straight line in the fixed-time plane joining point x with ∞ [reaching a compensating current at ∞ (Refs. 38 and 39)] and P is the path-ordering operator. In principle, we can choose other gauge groups G to implement the bosonization scheme, but the encouraging result for the 1D t-J model, reproducing the known exact exponents of the correlation functions[44] strongly favors the U(1)×SU(2) choice.

The bosonized action is obtained via substituting the time derivative by the covariant time derivative and the spatial lattice derivative by the covariant spatial lattice derivative in the U(1)×SU(2) bosonization

$$\Psi^*_{j\alpha}\partial_0\Psi_{j\alpha}\to\Phi^*_{j\alpha}[(\partial_0+iB_0(j))\mathbb{1}+iV_0(j)]_{\alpha\beta}\Phi_{j\beta},$$

$$\Psi^*_{i\alpha}\Psi_{j\alpha}\to\Phi^*_{i\alpha}e^{i\int_{\langle ij\rangle}B}(Pe^{i\int_{\langle ij\rangle}V})_{\alpha\beta}\Phi_{j\beta}.$$

We now decompose the bosonic field as follows:

$$\Phi_{x\alpha}=\tilde{E}_x\Sigma_{x\alpha},\quad \Phi^*_{x\alpha}=\tilde{E}^*_x\Sigma^*_{x\alpha}, \tag{2.6}$$

where $\mathbb{1}$ is a unity matrix, $\tilde{E}$ is a complex scalar field, and Σ_α is a complex two-component spin-1/2 field satisfying the constraint

$$\Sigma^*_{x\alpha}\Sigma_{x\alpha}=1. \tag{2.7}$$

The gauge ambiguity involved in the decomposition (2.6) will be discussed later. The field $\tilde{E}$ is coupled to the U(1) gauge field B and it describes the charge degrees of freedom, and the field Σ_α is coupled to the SU(2) gauge field V and it describes the spin degrees of freedom of the original fermion. In this description the nature of the groups associated with charge (U(1)) and spin (SU(2)) is explicitly exhibited and the coefficients of the C-S actions are such that the charged and spin-1/2 field operators reconstructed from the gauge invariant (Euclidean) fields $[\tilde{E}_x e^{i\int_{\gamma_x}B}$ and $P(e^{i\int_{\gamma_x}V})_{\alpha\beta}\Sigma_{x\beta}]$ obey semionic statistics.[38,44]

In terms of $\tilde{E}$ and Σ the U(1)×SU(2)-bosonized action of the t-J model is given by

$$\begin{aligned}S_{t-J}(\tilde{E},\tilde{E}^*,\Sigma,\Sigma^*,X,X^*,B,V)=\int_0^\beta dx^0\Bigg\{&\sum_{\langle ij\rangle}\left(\frac{2}{J}X^*_{\langle ij\rangle}X_{\langle ij\rangle}+[(-t+X_{\langle ij\rangle})\tilde{E}^*_j e^{i\int_{\langle ij\rangle}B}\tilde{E}_i\Sigma^*_{i\alpha}(Pe^{i\int_{\langle ij\rangle}V})_{\alpha\beta}\Sigma_{j\beta}+\text{H.c.}]\right)\\&+\sum_j\left[\tilde{E}^*_j\left(\partial_0+iB_0(j)+\mu+\frac{J}{2}\right)\tilde{E}_j+\tilde{E}^*_j\tilde{E}_j\Sigma^*_{j\alpha}[\partial_0\mathbb{1}+iV_0(j)]_{\alpha\beta}\Sigma_{j\beta}\right]\\&+\sum_{i,j}u_{i,j}\tilde{E}^*_i\tilde{E}_i\tilde{E}^*_j\tilde{E}_j\Bigg\}+2S_{c.s.}(B)+S_{c.s.}(V)\end{aligned} \tag{2.8}$$

with constraint (2.7) and Coulomb gauge fixing for the U(1)×SU(2) field implemented.[44]

It is convenient to describe the charge properties in terms of a holelike field H (holon) and this can be achieved in this formalism by substitution $\tilde{E}\to H^*$, $\tilde{E}^*\to H$, with H,H^* as Grassmann fields, and changing the sign of the C-S action for the B field.[39,44] After integration over the auxiliary gauge field X, the grand-canonical partition function $\Xi(\beta,\mu)$ can be rewritten as

$$\Xi(\beta,\mu)=\int \mathcal{D}H\mathcal{D}H^*\mathcal{D}\Sigma_\alpha\mathcal{D}\Sigma^*_\alpha\mathcal{D}B\mathcal{D}V e^{-S(H,H^*,\Sigma,\Sigma^*,B,V)}\delta(\Sigma^*\Sigma-1), \tag{2.9}$$

where the Euclidean action is given by

$$\begin{aligned}S(H,H^*,\Sigma,\Sigma^*,B,V)=\int_0^\beta dx^0\Bigg(&\sum_j\left\{H^*_j\left[\partial_0-iB_0(j)-\left(\mu+\frac{J}{2}\right)\right]H_j+iB_0(j)+(1-H^*_jH_j)\Sigma^*_{j\alpha}[\partial_0\mathbb{1}+iV_0(j)]_{\alpha\beta}\Sigma_{j\beta}\right\}\\&+\sum_{\langle ij\rangle}\Bigg\{[-tH^*_j e^{i\int_{\langle ij\rangle}B}H_i\Sigma^*_{i\alpha}(Pe^{i\int_{\langle ij\rangle}V})_{\alpha\beta}\Sigma_{j\beta}+\text{H.c.}]+\frac{J}{2}(1-H^*_jH_j)(1-H^*_iH_i)\\&\times\left[|\Sigma^*_{i\alpha}(Pe^{i\int_{\langle ij\rangle}V})_{\alpha\beta}\Sigma_{j\beta}|^2-\frac{1}{2}\right]\Bigg\}\Bigg)-2S_{c.s.}(B)+S_{c.s.}(V).\end{aligned} \tag{2.10}$$

In what follows we will denote the shifted chemical potential $\mu'=\mu+J/2$ by δ, proportional to the doping concentration.

C. Gauge fixings

The action (2.10) is invariant under the local gauge transformations:

$$\begin{aligned}&\text{U(1)}:\quad H_j\to H_je^{i\Lambda(j)},\quad H^*_j\to H^*_jE^{-i\Lambda(j)},\\&B_\mu(x)\to B_\mu(x)+\partial_\mu\Lambda(x),\quad \Lambda(x)\in\mathbf{R},\\&\text{SU(2)}:\quad \Sigma_j\to R^\dagger(j)\Sigma_j,\quad \Sigma^*_j\to\Sigma^*_jR(j),\\&V_\mu(x)\to R^\dagger(x)V_\mu(x)R(x)+R^\dagger(x)\partial_\mu R(x),\quad R(x)\in\text{SU(2)},\end{aligned} \tag{2.11}$$

and an additional holon-spinon (h/s) gauge invariance arising from the ambiguity in the decomposition (2.6):

$$\begin{aligned}&h/s:\quad H_j\to H_je^{i\zeta_j},\quad H^*_j\to H^*_je^{-i\zeta_j},\\&\Sigma_{j\alpha}\to\Sigma_{j\alpha}e^{i\zeta_j},\quad \Sigma^*_{j\alpha}\to\Sigma^*_{j\alpha}e^{-i\zeta_j},\quad \zeta_j\in\mathbf{R}.\end{aligned} \tag{2.12}$$

It is important to remark that the theory in terms of $\{H,H^*,\Sigma,\Sigma^*,B,V\}$ is equivalent to the original fermionic theory only if the h/s gauge is fixed to respect U(1)×SU(2) invariance. The h/s gauge fixing will be discussed later.

We first gauge fix the U(1) symmetry imposing a Coulomb condition on B (from now on $\mu=1,2$):

$$\partial^\mu B_\mu=0. \tag{2.13}$$

To retain the bipartite lattice structure induced by the AF interactions, we gauge fix the SU(2) symmetry by a ''Néel gauge'' condition:

$$\Sigma_j=\sigma_x^{|j|}\begin{pmatrix}1\\0\end{pmatrix},\quad \Sigma_j^*=(1,0)\sigma_x^{|j|}, \tag{2.14}$$

where $|j|=j_1+j_2$. Then we split the integration over V into an integration over a field $V^{(c)}$, satisfying the Coulomb condition:

$$\partial^\mu V_\mu^{(c)}=0, \tag{2.15}$$

and its gauge transformations expressed in terms of an SU(2)-valued scalar field g (not a second-quantized operator), i.e., $V_a=g^\dagger V_a^{(c)} g+g^\dagger\partial_a g$, $a=0,1,2$.

Integrating over B_0, we obtain

$$B_\mu=\bar{B}_\mu+\delta B_\mu,\quad \delta B_\mu(x)=\frac{1}{2}\sum_j H_j^* H_j\partial_\mu \arg(x-j), \tag{2.16}$$

where $\bar{B}_\mu$ gives rise to a π-flux phase, i.e., $e^{i\int_{\partial p}\bar{B}}=-1$ for every plaquette p.

Integrating over V_0, we find

$$V_\mu^{(c)}=\sum_j (1-H_j^*H_j)\left(\sigma_x^{|j|}g_j^\dagger\frac{\sigma_a}{2}g_j\sigma_x^{|j|}\right)_{11}\partial_\mu \arg(x-j)\sigma_a, \tag{2.17}$$

where σ_a, $a=x$, y, z are the Pauli matrices. After the U(1)×SU(2) field being gauge fixed, as discussed above, the action (2.10) becomes

$$\begin{aligned}S(H,H^*,g)=\int_0^\beta dx^0\Bigg\{&\sum_j\left[H_j^*(\partial_0-\delta)H_j+(1-H_j^*H_j)(\sigma_x^{|j|}g_j^*\partial_0 g_j\sigma_x^{|j|})_{11}\right]\\&+\sum_{\langle ij\rangle}\{-tH_j^*e^{i\int_{\langle ij\rangle}\bar{B}+\delta B}H_i[\sigma_x^{|i|}g_i^\dagger(Pe^{i\int_{\langle ij\rangle}V^{(c)}})g_j\sigma_x^{|j|}]_{11}+\text{H.c.}\}\\&+\frac{J}{2}\sum_{\langle ij\rangle}(1-H_i^*H_i)(1-H_j^*H_j)\left[|[\sigma_x^{|i|}g_i^\dagger(Pe^{i\int_{\langle ij\rangle}V^{(c)}})g_j\sigma_x^{|j|}]_{11}|^2-\frac{1}{2}\right]\Bigg\}.\end{aligned} \tag{2.18}$$

Here the boundary terms are omitted and the $S^1\times R^2$ topology of the involved Euclidean space time imposes the vanishing of the topological term $\mathrm{Tr}\int_{D\times\mathbf{R}^2}(g^\dagger dg)^3$, where D is a disk of radius β.

Equation (2.18) is the starting point for our subsequent analysis. In $S(H,H^*,g)$ the charge degrees of freedom (dof) are described by H,H^* (2 dof) and the spin degrees of freedom by g (3 dof) subjected to a constraint (-1 dof) coming from the h/s gauge fixing, reproducing the correct counting of degrees of freedom of the original fermionic fields Ψ_α, Ψ_α^* ($2+2$ dof) in the Euclidean path-integral formalism.

III. OPTIMIZATION OF THE SPINON CONFIGURATION

To analyze Eq. (2.18) we first recall the strategy adopted for the 1D case.[44] We noticed that one can find an upper bound for the partition function of holons in a spinon background. Moreover, one can find explicitly the spinon configuration, exactly saturating this bound. Then one can consider the quantum fluctuations around this optimal configuration to evaluate related physical quantities. Here we will follow a similar strategy, namely to search first for the upper bound of the partition function for holons in a spinon background. It turns out that such an upper bound exists. However, unlike the 1D case, we cannot find a spinon configuration exactly saturating this upper bound. Nevertheless, we can find a holon-dependent spinon configuration g^m which is optimal, saturating the upper bound on average, and take it as the starting point to consider the spinon fluctuations.

A. Auxiliary lattice gauge field and upper bound

To find the optimal configuration we introduce an auxiliary lattice gauge field $\{A,U\}$, with $A_j\in R$ (real), $U_{\langle ij\rangle}\in C$ (complex), $|U_{\langle ij\rangle}|\leqslant 1$, and an action $S=S_1+S_2$,

$$\begin{aligned}S_1(H,H^*,A,U)=\int_0^\beta dx^0\Bigg\{&\sum_j\left[H_j^*(\partial_0-\delta)H_j\right.\\&\left.+i(1-H_j^*H_j)A_j\right]\\&+\sum_{\langle ij\rangle}(-tH_i^*U_{\langle ij\rangle}H_j+\text{H.c.})\Bigg\},\end{aligned} \tag{3.1}$$

$$\begin{aligned}S_2(H,H^*,U)=\int_0^\beta dx^0\sum_{\langle ij\rangle}&\frac{J}{2}(1-H_i^*H_i)(1-H_j^*H_j)\\&\times\left(|U_{\langle ij\rangle}|^2-\frac{1}{2}\right).\end{aligned}$$

The action (3.1) equals Eq. (2.18) if we make the identifications

$$\begin{aligned}iA_j&\sim(\sigma_x^{|j|}g_j^\dagger\partial_0 g_j\sigma_x^{|j|})_{11},\\U_{\langle ij\rangle}&\sim e^{-i\int_{\langle ij\rangle}(\bar{B}+\delta B)}[\sigma_x^{|i|}g_i^\dagger(Pe^{i\int_{\langle ij\rangle}V^{(c)}})g_j\sigma_x^{|j|}]_{11},\end{aligned} \tag{3.2}$$

(but in the derivation of the following bound these identifications are not made).

Let

$$\Xi(A,U)=\int \mathcal{D}H\mathcal{D}H^{*}e^{-S(H,H^{*},A,U)},$$

we prove in Appendix A the following upper bound:

$$|\Xi(A,U)|\leqslant \int \mathcal{D}H\mathcal{D}H^{*}e^{-[S_1(H,H^{*},0,\hat{U})+S_2(H,H^{*},0)]}, \tag{3.3}$$

where $\hat{U}$ is the time-independent U configuration maximizing

$$\int \mathcal{D}H\mathcal{D}H^{*}e^{-[S_1(H,H^{*},0,U)+S_2(H,H^{*},0)]}\big|_{\partial_0 U=0}. \tag{3.4}$$

To discuss the properties of $\hat{U}$ we first notice that the quantity optimized by $\hat{U}$ is the free energy $F(U)$ of a gas of spinless holes at temperature $T=\beta^{-1}$ with chemical potential δ ($\delta=0$ corresponding to half-filling) on a lattice, with hopping parameter on the link $\langle ij\rangle$ given by $t|U_{\langle ij\rangle}|$, in the presence of a constant (but not uniform) magnetic field where the flux through a plaquette p is given by $\arg(U_{\partial p})$ and is subjected to an attractive nearest-neighbor (NN) density-density interaction with coupling constant $J/4$. We consider the system at large β, and small δ and make the following:

Assumptions: (1) We consider negligible the density-density interaction, since the holon density (δ) is small; (2) we assume translational invariance of $\hat{U}$.

Remark: Assumption (2) appears to be reasonable in the light of the results of Ref. 54, where it has been shown that, the configuration U maximizing the determinant of the hopping matrix of the above system is translation invariant.

By gauge invariance (see Appendix A) the result of optimization depends only on $|\hat{U}_{\langle ij\rangle}|$ and $\arg(\hat{U}_{\partial p})$. As a consequence of the assumption (2), $F(U)$ is monotonically increasing in $|U|$, hence

$$|\hat{U}_{\langle ij\rangle}|=1. \tag{3.5}$$

It has been conjectured in Ref. 55 and proven in Ref. 56 that the ground-state energy at $T=0$ of the system under consideration is optimized in the magnetic field chosen as

$$\arg(\hat{U}_{\partial p})=\pi(1-\delta), \tag{3.6}$$

which is the commensurability condition for the flux. It is then natural to conjecture that this remains true for large enough β.

B. Optimal spinon configuration

In the last subsection we have stated a kind of ''theorem'' (proven in Appendix A), now we will find out the consequences of this theorem in our context, namely, assuming Eqs. (3.5) and (3.6) to hold, we attempt to find a holon-dependent spinon configuration (g^m) saturating the bound (3.3), using the identifications (3.2).

Following a strategy developed in 1D,[44] we introduce a first-quantized (Feynman-Kac) representation of the holon partition function in the presence of a g background. As in 1D, a key ingredient in the analysis is that, for a fixed link $\langle ij\rangle$, in every holon configuration the term $(\sigma_x^{|i|}g_i^{\dagger}Pe^{i\int_{\langle ij\rangle}V^{(c)}}g_j\sigma_x^{|j|})_{11}$ appears, in the first quantized formalism, either in the world lines of holons or in the Heisenberg term, but never simultaneously. This is the consequence of the single-occupancy constraint and permits a separate optimization of S_1 and S_2, as required in the bound (3.3). It turns out, however, that, contrary to the 1D results, we cannot identify a specific configuration g^m saturating the bound exactly, but only approximatively, in an appropriate average sense.

The whole procedure in 2D can then be justified in the limit $t\gg J$ because the effective mass of holes is very heavy, as a result of the large number of soft spinon fluctuations surrounding the hole in its motion[57] and in a sense our treatment can be considered as a kind of Born-Oppenheimer approximation for the spinons in the presence of the holons.

Here we do not give details of the derivation, but rather introduce notations and quote the obtained results. Those interested in further discussions are referred to Refs. 58, 38, 39, and 44.

Let Δ denote the 2D lattice Laplacian defined on a scalar lattice field f by

$$(\Delta f)_i=\sum_{j:|j-i|=1} f_j-4f_i;$$

let $d\mu(\omega)$ denote the measure on the random walks ω on the 2D lattice such that

$$(e^{\beta\Delta})_{ij}=\int_{\substack{\omega(0)=i\\ \omega(\beta)=j}} d\mu(\omega),\quad \beta>0;$$

let P_N be the group of permutations of N elements and, for $\pi\in P_N$, let $\sigma(\pi)$ denote the number of exchanges in π, then the partition functions of holons (H) in a given g background can be rewritten as

$$\Xi(g)=e^{i\Sigma_j\int_0^{\beta}dx^0A_j}\sum_{N=0}^{\infty}\frac{e^{\beta\delta N}}{N!}\sum_{\pi\in P_N}(-1)^{\sigma(\pi)}$$
$$\times\sum_{j_1\cdots j_N}\prod_{r=1}^{N}\int_{\substack{\omega_r(0)=j_r\\ \gamma_r(\beta)=j_{\pi(r)}}} d\mu(\omega_r)\prod_{\langle ij\rangle\in\underline{\omega}^{\perp}} tU_{\langle ij\rangle}e^{-i\int_{\underline{\omega}^{\parallel}}dx^0A}e^{-\Sigma_{\langle ij\rangle\cap\underline{\omega}=\varnothing}J/2\int_0^{\beta}dx^0(|U_{\langle ij\rangle}|^2-1/2)}, \tag{3.7}$$

where identifications (3.2) are understood and, for a fixed N, $\underline{\omega}=\{\omega_1,\ldots,\omega_N\}$ denote the worldlines of holon particles, $\underline{\omega}^{\perp}$ the components of $\underline{\omega}$ perpendicular and $\underline{\omega}^{\parallel}$ parallel to the time axis, respectively.

To saturate the bound (3.3) with a configuration $g^m(\underline{\omega})$ we first impose

$$iA_j=(\sigma_x^{|j|}g_j^{\dagger}\partial_0 g_j\sigma_x^{|j|})_{11}=0, \quad j\in\underline{\omega}^{\parallel}. \tag{3.8}$$

Equation (3.8) is satisfied choosing g^m constant during the period when no particle hops.

Imposing $U_{\langle ij\rangle}=0$ in S_2 [see Eq. (3.1)] corresponds in the first quantized formalism to setting:

$$(\sigma_x^{|i|}g_i^{\dagger}Pe^{i\int_{\langle ij\rangle}V^{(c)}}g_j\sigma_x^{|j|})_{11}=0, \quad \langle ij\rangle\cap\underline{\omega}=\varnothing. \tag{3.9}$$

In physical terms this means that the $s+id$ RVB order parameter[47] is very small (see the discussion at the end of Sec. IV).

We notice that if

$$g_j=\cos\ \theta_j\mathbb{1}+i\ \sin\ \theta_j\sigma_z, \quad j\notin\underline{\omega} \tag{3.10}$$

for some angle $\theta_j\in[0,2\pi)$, Eq. (3.9) is then satisfied. In fact, since $V^{(c)}$ depends only on sites where there are no holes [see Eq. (2.17)], from Eq. (3.10) it follows that correspondingly

$$V^{(c)}(x)=\sum_j\ (1-H_j^*H_j)\ \frac{(-1)^{|j|}}{2}\ \partial_\mu\ \arg(x-j)\sigma_z, \tag{3.11}$$

so that $g_i^{\dagger}Pe^{i\int_{\langle ij\rangle}V^{(c)}}g_j$ has only diagonal components.

This result shows that in the Néel gauge, quite independently of doping concentration (since the condition $U=0$ in S_2 does not depend on small doping assumption), one should expect that the physics is dominated by $V^{(c)}$ only in the U(1) subgroup of SU(2) related to the axis chosen in the Néel gauge.

The condition $|\hat{U}_{\langle ij\rangle}|=1$ in S_1 [see Eq. (3.1)] corresponds to imposing

$$|(\sigma_x^{|i|}g_i^{\dagger}Pe^{i\int_{\langle ij\rangle}V^{(c)}}g_j\sigma_x^{|j|})_{11}|=1, \quad \langle ij\rangle\in\underline{\omega}^{\perp}, \tag{3.12}$$

which means in physical terms that the AM order parameter[46] is of the order 1 (see the discussion at the end of Sec. IV).

To discuss Eq. (3.12) we recall that the paths on which $d\mu(\omega)$ is defined are left continuous,[44] so that at the jumping time $\tau,\omega_r(\tau)=\lim_{\epsilon\searrow 0}\omega_r(\tau+\epsilon)$, or, in simpler terms, one should think the holon at τ at the end of the jumping link, oriented according to the increasing world line time of the holon. As a consequence, if $\langle ij\rangle\in\underline{\omega}^{\perp}$, either $i\in\underline{\omega}$ or $j\in\underline{\omega}$, but never both. Let us assume $j\in\underline{\omega}$, then according to the previous requirements

$$g_i^{\dagger}Pe^{i\int_{\langle ij\rangle}V^{(c)}}=\cos\ \theta_{\langle ij\rangle}\mathbb{1}+i\ \sin\ \theta_{\langle ij\rangle}\sigma_z,$$

for some angle $\theta_{\langle ij\rangle}\in[0,2\pi)$. We represent

$$g_j=\cos\ \varphi_j+i\boldsymbol{\sigma}\cdot\mathbf{n}_j\ \sin\ \varphi_j,$$

for some unit vector $\mathbf{n}_j$ and angle $\varphi_j\in[0,2\pi)$.

From Eq. (3.12) one immediately obtains

$$\varphi_j=\frac{\pi}{2}, \quad n_{jz}=0. \tag{3.13}$$

Finally, let us try to impose the condition (3.6). This translation invariant condition cannot be exactly fulfilled [unlike in 1D where a configuration g^m exactly saturating the bound analogous to Eq. (3.3) can be found]. However, we notice that the B-dependent part of arg $(U_{\partial p})$ has a translation invariant mean satisfying the above condition, [see Eqs. (3.3) and (2.16)], hence it is natural to impose (on average):

$$(\sigma_x^{|i|}g_i^{\dagger}e^{i\int_{\langle ij\rangle}V^{(c)}}g_j\sigma_x^{|j|})_{11}\simeq 1, \quad \langle ij\rangle\in\underline{\omega}^{\perp}. \tag{3.14}$$

Defining

$$\overline{g}_j=\exp\left(-\frac{i}{2}\sum_{l\neq j}\ (-1)^l\sigma_z\arg(l-j)\right), \tag{3.15}$$

and choosing

$$g_j=\begin{cases}\overline{g}_j, & j\notin\underline{\omega}\\ \overline{g}_j\widetilde{g}_j, & j\in\underline{\omega},\end{cases} \tag{3.16}$$

we can kill the fast fluctuating first term in Eq. (3.11). The remaining term, denoted by

$$\overline{V}=-\sum_j\ H_j^*H_j\ \frac{(-1)^{|j|}}{2}\ \partial_\mu\arg(x-j)\sigma_z,$$

for small hole concentration is a slowly varying field yielding a contribution $O(\delta)$ to $\arg(U_{\langle ij\rangle})$, with zero translational average. The final result is that we can assume for the optimizing configuration, using Eqs. (3.13), (3.15), (3.16)

$$g_i^{\dagger}e^{i\int_{\langle ij\rangle}V^{(c)}}g_j=e^{i\int_{\langle ij\rangle}\overline{V}}\widetilde{g}_j\simeq\widetilde{g}_j=e^{i(\pi/2)(\sigma_x n_{jx}+\sigma_y n_{jy})},$$

and we immediately derive from Eq. (3.14) the condition $n_{jx}=0$, $n_{jy}=(-1)^{|j|}$.

Notice that from these definitions:

$$\widetilde{g}_j\sigma_x^{|j|}=\sigma_x^{|j|+1}. \tag{3.17}$$

In view of the optimization discussed above it appears natural to introduce a variable $R_j\in\mathrm{SU}(2)$ describing spinon fluctuations around the optimizing configuration, through the definition

$$g_j=\overline{g}_jR_j\widetilde{g}_j=e^{-(i/2)\Sigma_{l\neq j}(-1)^l\sigma_z\ \arg(l-j)}R_je^{i(\pi/2)(-1)^{|j|}\sigma_yH_j^*H_j} \tag{3.18}$$

with R_j being represented in CP^1 form as

$$R_j=\begin{pmatrix}b_{j1} & b_{j2}^*\\ b_{j2} & b_{j1}^*\end{pmatrix}, \quad b_{j\alpha}^*b_{j\alpha}=1, \tag{3.19}$$

where $b_{j\alpha}$ is a two-component complex field. The optimal configuration g^m is given by $R=\mathbb{1}$.

Using Eqs. (3.17) and the SU(2) gauge invariance of the (formal) measure $\mathcal{D}g$ to absorb $\overline{g}$, the partition function of the t-J model can be exactly rewritten in terms of the Euclidean action $S=S_h+S_s$,

$$S_h=\int_0^\beta dx^0\Big\{\sum_j H_j^*[\partial_0-(\sigma_x^{|j|}R_j^\dagger\partial_0 R_j\sigma_x^{|j|})_{11}-\delta]H_j+\sum_{\langle ij\rangle}\{-tH_j^*e^{-i\int_{\langle ij\rangle}(\bar{B}+\delta B)}H_i[\sigma_x^{|i|}R_i^\dagger P e^{i\int_{\langle ij\rangle}(\bar{V}+\delta V)}R_j\sigma_x^{|i|}]_{11}+\text{H.c.}\}\Big\}, \tag{3.20}$$

$$S_s=\int_0^\beta dx^0\Big\{\sum_j(\sigma_x^{|j|}R_j^\dagger\partial_0 R_j\sigma_x^{|j|})_{11}+\sum_{\langle ij\rangle}\frac{J}{2}(1-H_i^*H_i)(1-H_j^*H_j)\Big[|(\sigma_x^{|i|}R_i^\dagger P e^{i\int_{\langle ij\rangle}(\bar{V}+\delta V)}R_j\sigma_x^{|j|})_{11}|^2-\frac{1}{2}\Big]\Big\}, \tag{3.21}$$

where $\delta V=V^{(c)}-\bar{V}$.

Let us remark again that in Eqs. (3.20), (3.21) no approximations have been made. The point of the above analysis is that for large enough β and small enough δ we expect that R and δV describe small fluctuations.

IV. SPINON EFFECTIVE ACTION

A. The main approximation

To proceed further we make the following *Approximation* ($\delta V=\delta B=0$): we assume that the spinon fluctuations (R) are small enough (for β large, δ, J/t small) that we can neglect their back reaction on the gauge field V, i.e., we set $\delta V=0$. We expect that the main effect of the neglected fluctuations of V is to convert the gauge-invariant spinon field operator reconstructed from the Euclidean field

$$(Pe^{i\int_{\gamma_j}V})\Sigma_j=e^{i\int_{\gamma_j}(\bar{V}+\delta V)}\overline{g_j}R_j\sigma_x^{|j|}\begin{pmatrix}1\\0\end{pmatrix}$$

into a semion field operator. Retaining the fermionic nature of Ψ is then inconsistent with neglecting δV (which also introduces a fictitious parity breaking) unless we neglect also δB, responsible for the semionic nature of the gauge-invariant holon field operator reconstructed from the Euclidean field

$$e^{-i\int_{\gamma_j}(\bar{B}+\delta B)}H_j.$$

In 1D the proper account of the statistics of the holon and spinon field operators was crucial for deriving (within the C-S approach) the correct physical properties of the model, known by Luttinger liquid and conformal field theory techniques.[45] However, in 2D we believe the statistics of holons and spinons is less crucial because we expect that, contrary to 1D, they form a bound state, as will be discussed later.

To derive the low-energy spinon action let us start computing the link variable

$$R_i^\dagger e^{i\int_{\langle ij\rangle}\bar{V}}R_j=\begin{pmatrix}\alpha_{\langle ij\rangle}b_{i1}^*b_{j1}+\alpha_{\langle ij\rangle}^*b_{i2}^*b_{j2} & -\alpha_{\langle ij\rangle}b_{i1}^*b_{j2}^*+\alpha_{\langle ij\rangle}^*b_{i2}^*b_{j1}^*\\ -\alpha_{\langle ij\rangle}b_{i2}b_{j1}+\alpha_{\langle ij\rangle}^*b_{i1}b_{j2} & \alpha_{\langle ij\rangle}b_{i2}b_{j2}^*+\alpha_{\langle ij\rangle}^*b_{i1}b_{j1}^*\end{pmatrix}, \tag{4.1}$$

where $\alpha_{\langle ij\rangle}=e^{(i/2)\int_{\langle ij\rangle}\bar{V}_z}$. Looking back at Eq. (3.21) we find that in the hopping term of holons only the diagonal elements of Eq. (4.1) appear, a kind of gauge-invariant AM variable,[46] whereas in the Heisenberg term only the off-diagonal elements of Eq. (4.1) appear, a kind of gauge-invariant RVB variable.[47] According to the optimization arguments given in the previous section, the vacuum expectation value of the AM gauge variable is expected to be s like, real and close to 1 [see Eq. (3.14)], while the RVB order parameter should be rather small [see Eq. (3.9)]. These anticipations are fully confirmed by the mean-field calculations.[48,59]

B. Nonlinear-σ model with mass term

We now derive a low-energy continuum effective action for spinons by rescaling the model to a lattice spacing $\epsilon\ll 1$ and neglecting higher-order terms in ϵ. As it is standard in AF systems,[28] we assume

$$b_{j\alpha}^*\boldsymbol{\sigma}_{\alpha\rho}b_{j\beta}\sim\boldsymbol{\Omega}_j+(-1)^{|j|}\epsilon\mathbf{L}_j, \tag{4.2}$$

with $\boldsymbol{\Omega}_j^2=f\lesssim 1$, $\boldsymbol{\Omega}\cdot\mathbf{L}=0$, where $\mathbf{Q},\mathbf{L}$ are defined on a sublattice, e.g., $\boldsymbol{\Omega}_j\equiv\boldsymbol{\Omega}_{j_1+(1/2),j_2}$, $\mathbf{L}_j\equiv\mathbf{L}_{j_1+(1/2),j_2}$, $j_1=j_2$ mod 2 and they describe the AF and ferromagnetic fluctuations, respectively. It is useful to rewrite $\boldsymbol{\Omega}$ in the CP^1 form:

$$\boldsymbol{\Omega}=z_j^*\boldsymbol{\sigma}_{\alpha\beta}z_\beta,\quad z_\alpha^*z_\alpha=f, \tag{4.3}$$

with z_α, $\alpha=1,2$ as a spin-$\frac{1}{2}$ complex (hard-care) boson field. Consistently with the slowly varying nature of $\bar{V}$ for small hole concentration, we assume

$$e^{-i\int_{\langle jl\rangle}\bar{V}_z}-1\sim\epsilon(-i\bar{V}_z)(j)+\frac{\epsilon^2}{2}(-i\bar{V}_z)^2(j)+O(\epsilon^3). \tag{4.4}$$

On the rescaled lattice the Heisenberg term becomes

$$\frac{J}{2}\sum_{\langle ij\rangle}\frac{|(\sigma_x^{|x|}R_i^\dagger e^{(i/2)\int_{\langle ij\rangle}\overline{V}}R_j\sigma_x^{|j|})_{11}|^2-1/2}{\epsilon^2}=\frac{J}{2}\sum_{\langle ij\rangle}\left\{\frac{1}{2}\left(\frac{\mathbf{\Omega}_i-\mathbf{\Omega}_j}{\epsilon}\right)^2+2\mathbf{L}_j^2+\overline{V}_z^2(j)[(\Omega_{jx})^2+(\Omega_{jy})^2]\right\}+O(\epsilon). \tag{4.5}$$

For the temporal term we obtain analogously

$$-(\sigma_x^{|j|}R_j^\dagger\partial_0 R_j\sigma_x^{|j|})_{11}$$
$$=\sum_j(-1)^{|j|}z_{j\alpha}^*\partial_0 z_{j\alpha}+\frac{\epsilon}{2}\mathbf{L}_j\cdot(\mathbf{\Omega}_j\wedge\partial_0\mathbf{\Omega}_j)+O(\epsilon^2). \tag{4.6}$$

The first term in Eq. (4.6) actually vanishes, since otherwise it would produce a topological θ term which is known to be absent in 2D.[28]

In treating the holon density terms in the spinon action in MFA, we keep only the leading fluctuation terms and neglect terms of the order of $O(J\delta)$, $O(\delta^2)$. Then, integrating out $\mathbf{L}$ and taking the continuum limit, from Eqs. (4.2)–(4.6) we obtain the NLσ model effective action for spinons $S_s^\star+S_s'$,

$$S_s^\star=\int d^3x\,\frac{1}{g}\,[(\partial_0\mathbf{\Omega})^2+v_s^2(\partial_\mu\mathbf{\Omega})^2+(\mathbf{\Omega})^2(\overline{V}_z)^2],$$

$$S_s'=-\frac{1}{g}\int d^3x\Omega_z^2\overline{V}_z^2, \tag{4.7}$$

where the coupling constant g and the spin-wave velocity v_s are easily derived functions of J, δ, ϵ.

We now make the *Approximation P:* we treat the term S_s' as a perturbation. To understand the physics described by $S_s^\star$ we first notice that if $\overline{V}_z^2$ were absent, the NLσ model would be in the symmetry broken phase, since g is small ($\sim J$) at the lattice scale. For larger scales $g_{\rm eff}$ then flows towards its critical value, describing the large-distance properties of a NLσ model with an insulating ''Néel'' ground state and spin-wave Goldstone excitations.[28] To get an idea of the effect of $\overline{V}_z^2$ we replace the NLσ model constraint $\mathbf{\Omega}^2=f$ by a softened version, adding to the Lagrangian a term $\lambda(\mathbf{\Omega}^2-f)^2$, expected to produce the same low-energy behavior, and we replace $\overline{V}_z^2$, a function of the holon positions, by its statistical average $\langle\overline{V}_z^2\rangle$. By its definition, $\overline{V}_z^2$ is positive definite and we give a rough estimate of $\langle\overline{V}_z^2\rangle$ by first performing a translational average at a fixed time over a fixed holon configuration and then an average over holon configurations with mean holon density δ.

In the first step let $\underline{x}=\{x_i\}_{i=1}^N$ denote the holon positions and let $q=\{q_i\}_{i=1}^N$, $q_i=(-1)^{|x_i|+1}$ and restrict the computation to a finite volume $|\Lambda|$ and lattice cutoff ϵ. Let $\arg^\epsilon$, ∂_μ^ϵ, Δ^ϵ denote the angle function, the derivative in the μ direction and the Laplacian in the ϵ lattice, respectively, then using the equality

$$\epsilon_{\mu\nu}\partial_\nu^\epsilon\arg^\epsilon(x-y)=\partial_\mu^\epsilon(\Delta^\epsilon)^{-1}(x-y),$$

we immediately obtain

$$\frac{1}{|\Lambda|}\int_\Lambda d^2x\overline{V}_z^2(x)=-\frac{1}{|\Lambda|}\sum_{\substack{l,k\\ \underline{x}\subset\Lambda}}q_lq_k(\Delta^\epsilon)^{-1}(x_l,x_k)$$
$$\simeq-\frac{1}{|\Lambda|}\sum_{\substack{l,k\\ \underline{x}\subset\Lambda}}\frac{q_lq_k}{2\pi}\ln(|x_l-x_k|+\epsilon). \tag{4.8}$$

Equation (4.8) looks like the energy per unit volume $E_{N,\Lambda}(\underline{x})$ of a neutral two-component system of N particles with charges $\pm\sqrt{1/2\pi}$ in a volume $|\Lambda|$ interacting via 2D Coulomb potential with ultraviolet cutoff ϵ. The average of $E_{N,\Lambda}(\underline{x})$ over the positions of the charges in the limit $|\Lambda|\nearrow\mathbf{R}^2$, $N\nearrow\infty$ with fixed average density δ can be identified with the free energy of the above system at $\beta=1$ and chemical potential δ in the thermodynamic limit, and this gives our estimate of $\langle\overline{V}_z^2\rangle$. The behavior of the free energy can be obtained through a sine-Gordon transformation[60] and for small δ it is given by[61]

$$\langle\overline{V}_z^2\rangle\sim-\delta\ln\delta. \tag{4.9}$$

Hence in our crude approximation $\langle\overline{V}_z^2\rangle$ acts as a mass term increasing with δ. If we assume that the scaling limit and perturbation expansion in $\langle V_z^2\rangle$ commute with each other, we find that its effect is to drive the NLσ model at large scale to the disordered massive regime with a mass gap for $\mathbf{\Omega}$ of the order of $m_s^2(\delta)\sim-\delta\ln\delta$. Within this approximation, for a more careful analysis one should consider the renormalization-group equations for g and λ including the perturbation $\langle\overline{V}_z^2\rangle$ term from the beginning. Then the value of $m^2(\delta)$ could be renormalized.

A slightly better approximation is to consider the holons as slowly moving randomly distributed impurities, creating a random potential and analyze the behavior of the spin-wave $\mathbf{\Omega}$ in the presence of this potential (this can again be justified for the limit $J/t\ll1$ with a large effective mass of holes). At a fixed time the random potential behaves as

$$\overline{V}_z^2(w)\sim\sum_{\underline{x},\underline{y}}q_yq_x\frac{(x-w)^\mu}{|x-w|^2}\frac{(y-w)^\mu}{|y-w|^2}.$$

Hence it is positive and roughly falling like r^{-2}, where r is the distance to the closest impurity. This kind of systems have been considered in Ref. 62 and it is expected to be in the localized regime for $\mathbf{\Omega}$, where (random) averaged Green functions exhibit a mass gap $m(\delta)$ roughly proportional to the inverse mean free path. Hence up to logarithms $m_s^2(\delta)\sim\delta$ for small δ.

C. Crossover from short-range to long-range antiferromagnetic order

All the above arguments strongly suggest that the spinon system described by the action $S_s^\star$ in Eq. (4.7) exhibits a mass gap $m_s(\delta)$ increasing with δ (at least for sufficiently large, but, nevertheless $\delta \ll 1$), vanishing at $\delta=0$, thus showing an expected crossover to the insulating ''Néel state'' at zero doping.[63] As a result of $m(\delta)>0$, the constraint $\mathbf{\Omega}^2 = f$ is relaxed at large scales[28] and a Lagrange multiplier field introduced in the action to impose the constraint would just mediate short-range attractive interaction between spinons with opposite spin, with strength $0(m_s(\delta)^{-2})$ rotationally symmetric in spin space (see Ref. 64 for the discussion of an analogous situation). Another short-range interaction, but uniaxial in the spin space, is introduced by the perturbation term S_s' in Eq. (4.7), with strength $0(m_s^2(\delta)/g)$.

We can summarize the above discussion by rewriting the NLσ model effective action for spinons in the CP^1 form, neglecting short-range interactions, as

$$S_s^\star = \int_{[0,\beta]\times \mathbf{R}^2} d^3x \, \frac{1}{g} \left[|(\partial_0 - z_\beta^* \partial_j z_\beta) z_\alpha|^2 + v_s^2 |(\partial_\mu - z_\beta^* \partial_\mu z_\beta) z_\alpha|^2 + m_s^2(\delta) z_\alpha^* z_\alpha \right]. \quad (4.10)$$

In the NLσ model without mass term ($\delta=0$) the constraint $z_\alpha^* z_\alpha = f$ and the symmetry breaking condition, e.g., $\langle z_2 \rangle \neq 0$, lead to excitations described by a complex massless field $S \simeq z_1 \langle z_2^* \rangle$ with massless ''relativistic'' dispersion relations, corresponding to spin waves. In the NLσ model with mass term the absence of symmetry breaking and the effective softening of the constraint lead to excitations described by the spin-$\frac{1}{2}$ two-component field z_α, with massive ''relativistic'' dispersion relations. However, the self-generated gauge field $z_\alpha^* \partial_a z_\alpha$, $a=0,1,2$ would confine the spin-$\frac{1}{2}$ degrees of freedom into composite spin-1 spin-wave fields. On the other hand, as we shall see later, the coupling to the holons will induce deconfinement of spin-$\frac{1}{2}$ excitations.

To give a kind of ''microscopic'' interpretation of the above results using the slave-fermion picture in terms of the hard-core bosonic fields b_α, b_α^*, we follow the approach of Yoshioka.[13] First consider the case $\delta=0$; then a MF treatment with an *s*-like RVB order parameter yields an energy gap vanishing at $(\pi/2,\pi/2)$, $(-\pi/2,-\pi/2)$ in the Brillouin zone and an expansion for low momenta around these two points shows that the corresponding excitations are described by a complex massless field with relativistic dispersion relations, which can be identified with a spin-wave field *S*; the ground state is then the insulating Néel state.[13] A MF treatment with an $s+id$ RVB order parameter yields an energy gap vanishing at $(\pm\pi/2,\pm\pi/2)$ in the Brillouin zone, and an expansion for low momenta around these four points shows that the corresponding excitations are described by a two-component massless field with relativistic dispersion relation which can be identified with a spinon field z_α.[13] These excitations are turned into massive ones in our approach by the $m_s(\delta)$ term. In some sense, our result in the NLσ model corresponds to a kind of slave-fermion approach with a gauge-invariant RVB order parameter $s+id$-like and minimal gap $m_s(\delta)$ at $(\pm\pi/2,\pm\pi/2)$. We should emphasize, however, that our approach is different from the slave-fermion formalism[65] which considers long-range ordered AF state with gapless excitations, while we consider short-range ordered AF state with gapful excitations.

To the best of our knowledge, the present formulation is the first successful attempt to explicitly include the AF fluctuations into the RVB-type scheme of treating the strong correlation effects in a self-consistent manner. Upon doping the long-range AF is destroyed, being replaced by short-ranged AF order, which is physically obvious. We have obtained an explicit doping dependence of the gap value which has the correct extrapolation at zero doping (gap vanishes).[66] The doping dependence of the AF correlation length $\xi \sim m^{-1} \sim \delta^{-1/2}$, expected from our calculation, agrees very well with the neutron-scattering data.[67] Here we also propose an interpretation of the spin-gap effect in underdoped superconductors—mainly due to the short-range AF order. The recent numerical simulations[68] seem to support our interpretation.

V. HOLON EFFECTIVE ACTION

A. Holon-spinon coupling via self-generated gauge field

Now turn to holons. We use a gauge choice of the field $\bar{B}$ such that $e^{i\int_{\langle ij\rangle}\bar{B}}$ is purely imaginary and assume, following Eqs. (3.14)–(3.17) that the gauge-invariant AM parameter is $\simeq 1$, which permits us to use the equality [see Eqs. (4.7), (3.19)]

$$\langle \alpha_{\langle ij\rangle} b_{i1}^* b_{j1} + \alpha_{\langle ij\rangle}^* b_{i2} b_{j2} \rangle \simeq 1 = b_{i1}^* b_{i1} + b_{i2}^* b_{i2}. \quad (5.1)$$

We obtain a low-energy continuum effective action for holons by rescaling the lattice spacing to $\epsilon \ll 1$ and neglecting higher-order terms in ϵ. Making use of Eq. (5.1) one obtains

$$S_h = \int_0^\beta dx^0 \Bigg(\sum_j H_j^* (\partial_0 - (b_{j\alpha}^* \partial_0 b_{j\alpha})^{\#(j)} - \delta) H_j + \sum_{\langle ij\rangle} (-) t e^{i\int_{\langle ij\rangle}\bar{B}} \Bigg\{ \frac{H_i^* H_j - H_j^* H_i}{\epsilon} + (H_i^* H_j + H_j^* H_i) \Bigg[b_{i\alpha}^* \left(\frac{b_{j\alpha} - b_{i\alpha}}{\epsilon} \right) - \left(\frac{b_{i\alpha}^* - b_{j\alpha}^*}{\epsilon} \right) b_{j\alpha} \Bigg]^{\#(j)} \Bigg\} + O(\epsilon) \Bigg), \quad (5.2)$$

where $\#(j)$ denotes complex conjugation if j is even. Neglecting the b terms, the action (5.2) describes the usual two-component Dirac (''staggered'') fermions of the flux phase, with vertices of the double-cone dispersion relations in the reduced Brillouin zone, centered (in the $\epsilon=1$ lattice) at $(\pm\pi/2,\pm\pi/2)$ and chemical potential δ. To define a continuum effective action, using the standard procedure for the flux phase,[28] we first define four sublattices: (1) for j_1, j_2 even, (2) for

j_1 odd, j_2 even, (3) for j_1 even j_2 odd, (4) for j_1, j_2 odd. They can be grouped into two ''Néel'' sublattices $A=\{(1),(4)\}$, $B=\{(2),(3)\}$. The holon field H restricted to the sublattice # is denoted by $H^{(\#)}$. The holon action can then be written in the continuum limit as a bilinear form in H^* and H with kernel:

$$\begin{pmatrix} \partial_0 - z_\alpha^* \partial_0 z_\alpha - \delta, & it(\partial_1 + z_\alpha^* \partial_1 z_\alpha), & -it(\partial_2 + z_\alpha^* \partial_2 z_\alpha), & 0 \\ it(\partial_1 - z_\alpha^* \partial_1 z_\alpha), & \partial_0 + z_\alpha^* \partial_0 z_\alpha - \delta, & 0, & it(\partial_2 - z_\alpha^* \partial_2 z_\alpha) \\ -it(\partial_2 - z_\alpha^* \partial_2 z_\alpha), & 0, & \partial_0 + z_\alpha^* \partial_0 z_\alpha - \delta, & it(\partial_1 - z_\alpha^* \partial_1 z_\alpha), \\ 0, & it(\partial_2 + z_\alpha^* \partial_2 z_\alpha), & it(\partial_1 + z_\alpha^* \partial_1 z_\alpha), & \partial_0 - z_\alpha^* \partial_0 z_\alpha - \delta \end{pmatrix}, \tag{5.3}$$

where we used the decomposition of b_α carried out in the previous section [all ferromagnetic terms are $0(\epsilon)$ and so they do not appear in the continuum limit (5.3)].

Setting

$$\gamma_0 = \sigma_z, \quad \gamma_\mu = (\sigma_y, \sigma_x), \quad \not{A} = \gamma_\mu A_\mu,$$

$$\Psi_1 = \begin{pmatrix} \Psi_1^{(A)} \\ \Psi_1^{(B)} \end{pmatrix} = \begin{pmatrix} e^{-i(\pi/4)} H^{(1)} + e^{i(\pi/4)} H^{(4)} \\ e^{-i(\pi/4)} H^{(3)} + e^{i(\pi/4)} H^{(2)} \end{pmatrix},$$

$$\Psi_2 = \begin{pmatrix} \Psi_2^{(B)} \\ \Psi_2^{(A)} \end{pmatrix} = \begin{pmatrix} e^{-i(\pi/4)} H^{(2)} + e^{i(\pi/4)} H^{(3)} \\ e^{-i(\pi/4)} H^{(4)} + e^{i(\pi/4)} H^{(1)} \end{pmatrix},$$

$$\overline{\Psi}_\# = \gamma_0 \Psi_\#^\dagger, \tag{5.4}$$

and assigning charge $e_A = +1$, $e_B = -1$ to the fields corresponding to A and B sublattices, respectively, the continuum effective action for holons can be rewritten as[69]

$$S_h^\star = \int_{[0,\beta]\times \mathbf{R}^2} d^3x \sum_{r=1}^{2} \overline{\Psi}_r [\, \gamma_0 (\partial_0 - \delta - e_r z_\alpha^* \partial_0 z_\alpha) + t(\not{\partial} - e_r z_\alpha^* \not{\partial} z_\alpha)] \Psi_r \,. \tag{5.5}$$

Hence, $S_h^\star$ describes the coupling of the holon field to the spinon-generated gauge field $z_\alpha^* \partial_a z_\alpha$ in a ''relativistic'' Dirac-like form, with opposite coupling for the two Néel sublattices. From Eq. (5.5) it is clear that (neglecting the gauge terms), the upper components of $\Psi_\#$ describe gapless excitations with small FS [$\epsilon_F \simeq 0(\delta t)$], whereas the lower components describe massive excitations with a doping-dependent gap $m_h(\delta) \simeq 2\sqrt{\delta}$. In terms of lattice fields we have excitations of the two bands (gapless and gapful) supported in the reduced Brillouin zone due to the presence of the Néel sublattices. This leads to a band mixing due to the nondiagonal structure of the γ matrices and this in turn yields a ''shadow band'' effect, i.e., a reduction of the spectral weight for the outer part of the FS facing (π,π), in agreement with experimental data[10] and MF numerical simulation.[48] The FS of underdoped superconductors has also been considered in the SU(2) gauge-field theory.[22] However, there are some differences between the results obtained there and the predictions of the present model (for details see Ref. 48).

B. U(1)×SU(2) gauge invariance of the total action

Neglecting the quartic fermion term, which would produce a short-range repulsion between holons with opposite charge, the full continuum effective action can be rewritten introducing an auxiliary U(1) gauge field A (over which one integrates in the path integral):

$$S^\star = \int_{[0,\beta]\times \mathbf{R}^2} d^3x \Big\{ \frac{1}{g} [\, |(\partial_0 - A_0) z_\alpha|^2 + v_s^2 |(\partial_\mu - A_\mu) z_\alpha|^2 + m_s^2(\delta) z_\alpha^* z_\alpha] + \sum_{r=1}^{2} \overline{\Psi}_r [\, \gamma_0 (\partial_0 - e_r A_0 - \delta) + t(\not{\partial} - e_r \not{A})] \Psi_r \Big\}. \tag{5.6}$$

The action $S^\star$ is invariant under the U(1) gauge transformation

$$\begin{aligned} &\Psi_r(x) \to e^{i e_r \Lambda(x)} \Psi(x), \\ &\overline{\Psi}_r(x) \to e^{-i e_r \Lambda(x)} \overline{\Psi}(x), \\ &z_\alpha(x) \to e^{i\Lambda(x)} z_\alpha(x), \quad z_\alpha^*(x) \to e^{-i\Lambda(x)} z_\alpha^*(x), \\ &A_a(x) \to A_a(x) + \partial_a \Lambda(x), \quad \Lambda(x) \in \mathbf{R}. \end{aligned} \tag{5.7}$$

Going backwards we recognize in Eq. (5.7) the continuum limit of the h/s gauge invariance (2.12) which is gauge-fixed imposing, e.g., a Coulomb gauge condition on A_a.

Remark: One easily finds that the lattice counterpart of the A_μ gauge fixing is a gauge fixing for $\arg(\Sigma_i^* P e^{i\int_{\langle ij\rangle}^V} \Sigma_j)$ which is U(1)×SU(2) gauge invariant as required in the *Remark* in Sec. II.

In fact

$$\arg(\Sigma_i^* P e^{i\int_{\langle ij\rangle} V} \Sigma_j) \simeq \arg(\alpha_{\langle ij\rangle} b_{i1}^* b_{j1} + \alpha_{\langle ij\rangle}^* b_{i2}^* b_{j2}) \simeq z_\alpha^* \partial_\mu z_\alpha \simeq A_\mu,$$

where in the last equality we have neglected a quadratic fermion term fictitiously introduced in $S^\star$ [compare it with Eq. (5.5)].

We can use $S^\star$ to compute, as in Ref. 20, the gauge-field propagator induced by the spinon and holon vacuum polarizations. In the Coulomb gauge, denoting by $\Pi^\perp$ ($\Pi^\parallel$) the transverse (longitudinal) polarizations, respectively, we have

$$\langle A_\mu(\omega,q)A_\nu(-\omega,-q)\rangle=\left(\delta_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)\frac{1}{(\Pi_s^\perp+\Pi_h^\perp)(\omega,q)},$$

$$\langle A_0(\omega,q)A_0(-\omega,-q)\rangle=\frac{\omega^2}{q^2(\Pi_s^\parallel+\Pi_h^\parallel)(\omega,q)}.$$

Since the spinon is massive the vacuum polarization of the spinon system alone would be Maxwell-like in the absence of holons and it would induce a logarithmic confinement of spinons ($\Pi^\perp\simeq\omega^2+q^2$, $\Pi^\parallel\simeq\omega^2$ for small ω,q). However, due to their gapless component the holons induce a transverse vacuum polarization with Reizer singularity[21] ($\Pi_h^\perp\simeq|\omega/q|+\chi q^2$) and a leading behavior in $\Pi_h^\parallel\simeq\omega^2/q^2$ for small ω/q, leading to a short-range A_0 propagator, so that the full gauge interaction is *not* confining.[70]

As a result, holons and spinons are true dynamical degrees of freedom in the model. Nevertheless, since we are in 2D, the attractive force mediated by the gauge field is expected to produce bound states with the quantum numbers of the spin wave (see Ref. 64 for the discussion of a similar problem) and presumably (see, e.g., Ref. 71) of the electron. This could explain why neglecting the (expected) semionic nature of holons and spinons ($\delta V=\delta B=0$) can be justified to some extent in 2D. On the contrary, in 1D the spinons and holons are deconfined, and their statistical properties as semions are crucial for a proper account of the spin-charge separation.

VI. CONCLUSIONS

To summarize our results it is interesting to compare the features appearing in the (underdoped) 2D and 1D t-J model (for small δ, J/t) within the U(1)×SU(2) gauge approach followed here and in Ref. 44.

In both cases, applying to the t-J model the U(1)×SU(2) C-S bosonization in terms of gauge fields B, V, we separate spin and charge degrees of freedom of the electron, and describe them in terms of spinon (z) and holon (Ψ) fields.

(1) The low-energy properties of the spinons can be described by a NLσ model: in 2D it is massive, and the mass gap (vanishing at $\delta=0$) is due to the coupling to holons mediated by a nonpure gauge SU(2) field V; in 1D no such nontrivial gauge field exists and the spinons are massless. The self-generated spinon-gauge field ($z_\alpha^*\partial_a z_\alpha$) would confine the spinons into spin waves, but this is prevented in 2D by the coupling to holons, while in 1D by the presence of a topological $\theta=\pi$ term in the NLσ model, which is absent in 2D.

(2) The low-energy properties of holons are described in 2D by a Dirac action, inducing band mixing, whose appearance is due to the presence of a nonpure gauge field B (characteristic of the flux phase); the absence of such a field in 1D implies that holon are described in term of a spinless fermion action.

(3) Spinons and holons in 2D are coupled by the spinon-generated gauge field, carrying one degree of freedom (after gauge fixing) and this presumably induces binding; the gauge fixing kills the degree of freedom of an analoguous gauge field in 1D and as a consequence, spinons and holons are free. The statistics is then dictated by the U(1)×SU(2) decomposition and the corresponding (gauge-invariant) fields obey semionic statistics.

To reiterate, we outline some of the distinct features of our present study:

(i) The short-range AF order is the main reason for the existence of the spinon gap (at least in the underdoped samples) which can explain a variety of experimental observations.

(ii) Neglecting gauge fluctuations the two-point correlation function for the electron exhibits half-pocket FS in the reduced Brillouin zone, with a pseudogap minimal on the diagonal, ''shadow band'' effect and a quasiparticle peak due to the holons. These features have been demonstrated in the MF computation of Ref. 48, where a similar action for the system has been used, with a twist in holon-spinon statistics, i.e., the spinons are fermions, while the holons are bosons.

(iii) The spin-wave persists even in the region without AF long-range order, but as a short-ranged field.

(iv) We expect (this is still under investigation) that the ''composite electron'' exhibits a non-Fermi-liquid behavior due to the appearance of Reizer singularity, thanks to transverse gauge fluctuations, in the self-energy of its holon constituent.

To conclude, the C-S bosonization approach, in spite of its technical complications, is promising in providing a natural framework to describe the normal-state properties of underdoped superconductors. After having settled in this paper our framework for a discussion of the t-J model at small δ and J/t as a model for underdoped high-T_c cuprates, in the forthcoming paper we will compute correlation functions and compare the results with experimental data.[50]

ACKNOWLEDGMENTS

We would like to thank J. Fröhlich, Ch. Mudry, N. P. Ong, F. Toigo, and A. Tsvelik for stimulating discussions. The work of P.M. was partially supported by TMR Programme ERBFMRX-CT96-0045.

APPENDIX

The proof of the bound (3.3) uses techniques adapted from the proof of the diamagnetic inequality given in Ref. 72.

The main ingredients are

(1) Invariance of the action (3.1) under the gauge transformations:

$$A_j\to A_j+\partial_0\Lambda_j,\quad U_{\langle ij\rangle}\to U_{\langle ij\rangle}e^{i(\Lambda_i-\Lambda_j)},$$
$$H_j\to H_j e^{i\Lambda_j},\quad H_j^*\to H_j^* e^{-i\Lambda_j},\quad \Lambda_j\in\mathbf{R}. \tag{A1}$$

(2) Reflection (O.S.) positivity in the absence of gauge fields: define an antilinear involution θ on the holon fields supported in the positive-time lattice (the time interval is identified with $[-\beta/2,\beta/2]$) by

$$\theta{:}H_j(\tau)\to H_j^*(-\tau),\quad H_j^*(\tau)\to H_j(-\tau),$$
$$\theta(AB)=\theta(B)\,\theta(A),\quad A,B\in\mathcal{J}_+, \tag{A2}$$

where $\mathcal{J}_+$ denotes the set of functions of H,H^* with support on positive time. Reflection positivity is the following statement: $\forall F \in \mathcal{J}_+$

$$\langle F\theta F\rangle \equiv \int \mathcal{D}H\mathcal{D}H^* e^{-S(H,H^*)} F\theta F^* \geqslant 0, \tag{A3}$$

where $S(H,H^*)\equiv S(H,H^*,0,1)$ [see Eq. (3.1)].

The involution θ defines an inner product $(\cdot,\cdot)$ on $\mathcal{J}_+$ by $\langle\cdot\theta\cdot\rangle$ and as a consequence, Schwartz inequality holds: $\forall F,G \in \mathcal{J}_+$

$$|\langle F\theta G\rangle| \leqslant \langle F\theta F\rangle^{1/2}\langle G\theta G\rangle^{1/2}. \tag{A4}$$

Although the proof of Eq. (A3) is standard,[73] let us sketch it for readers' convenience and to set up a notation useful in the proof of Eq. (3.3).

We discretize the time interval introducing a time lattice with spacing $\epsilon=\beta 2^{-N}$ and we choose the time-0 plane lying halfway between lattice planes. We denote by S_+^ϵ, S_-^ϵ, and S_{+-}^ϵ the terms of the discretized action existing in the positive, negative time lattice, and coupling the two, respectively, then we have

$$S_-^\epsilon = \theta S_+^\epsilon S_{+-}^\epsilon = -\sum_j \left[H_j^*\left(\frac{\epsilon}{2}\right) H_j\left(\frac{-\epsilon}{2}\right) + H_j\left(\frac{\epsilon}{2}\right) H_j^*\left(\frac{-\epsilon}{2}\right)\right]$$
$$= -\sum_j \left[H_j^*\left(\frac{\epsilon}{2}\right)\theta H_j^*\left(\frac{\epsilon}{2}\right) + H_j\left(\frac{\epsilon}{2}\right)\theta H_j\left(\frac{\epsilon}{2}\right)\right]. \tag{A5}$$

Let P_+ denote the set of functions of H,H^* given by linear combinations with positive coefficients of elements of the form $F\theta F, F\epsilon\mathcal{J}_+$. It is easy to see that P_+ is closed under multiplication and summation with positive coefficients and for $\Sigma_i c_i F_i\theta F_i \in P_+$ we have

$$\int (\mathcal{D}H\mathcal{D}H^*)_\epsilon \sum_i c_i F_i\theta F_i = \sum_i c_i \left| \int (\mathcal{D}H\mathcal{D}H^*)_\epsilon F_i\right|^2 \geqslant 0,$$

where $(\mathcal{D}H\mathcal{D}H^*)_\epsilon$ denotes the (formal) measure on H,H^* in the path integral for the discretized time model.

From Eq. (A5) it follows that, for $F \in \mathcal{J}_+, e^{-S^\epsilon(H,H^*)} F\theta F \in P_+$, so that

$$\int (\mathcal{D}H\mathcal{D}H^*)_\epsilon e^{-S^\epsilon(H,H^*)} F\theta F \geqslant 0$$

and taking the limit $\epsilon\searrow 0$ we obtain Eq. (A3).

We turn now to the proof of Eq. (3.3). We work again with the discretized time lattice. First we use the gauge invariance (A1) to set $A=0$ on the links in S_{+-}. We denote by $\{A_1^{(1)},U_1^{(1)}\}$ ($\{A_2^{(1)},U_2^{(1)}\}$) the restriction of the gauge fields to the positive (negative) time lattice and we set

$$F_1^{(1)} = e^{-[S_+^\epsilon(H,H^*,A_1^{(1)},U_1^{(1)}) - S_+^\epsilon(H,H^*)]},$$
$$F_2^{(1)} = e^{-[S_+^\epsilon(H,H^*,-rA_2^{(1)},rU_2^*) - S_+^\epsilon(H,H^*)]}, \tag{A6}$$

where r is reflection with respect to the time-0 plane. Applying Eq. (A4) we derive the upper bound:

$$|\Xi^\epsilon(A,U)| \equiv \left| \int \mathcal{D}H\mathcal{D}H^* e^{-S^\epsilon(H,H^*,A,U)}\right| = |\langle F_1^{(1)}\theta F_2^{(1)}\rangle|$$
$$\leqslant \langle F_1^{(1)}\theta F_1^{(1)}\rangle^{1/2}\langle F_2^{(1)}\theta F_2^{(1)}\rangle^{1/2}.$$

We now choose a time plane, different from the time-0 plane, lying halfway between lattice planes, use the time translational invariance of Ξ (inherited from the antiperiodicity in time of H,H^*) to bring this plane to the position of the time-0 plane shifting the whole lattice and repeat the above procedure both for $F_1^{(1)}\theta F_1^{(1)}$ and $F_2^{(1)}\theta F_2$ (which are no more in the form $F\theta F$ with θ referred to the new time-0 plane). As a result, we obtain, with obvious notations, $F_l^{(2)}, l=1,2^2$, and then iterate. Finally we derive the bound

$$|\Xi^\epsilon(A,U)| \leqslant \prod_{l=1}^{2^N} \langle F_l^{(N)}\theta F_l^{(N)}\rangle^{2^{-N}}. \tag{A7}$$

A close inspection shows that the gauge fields appearing in $F_l^{(N)}\theta F_l^{(N)}$ have the following properties: $A=0$ and U is time independent. Using a Hamiltonian transcription, with obvious meaning of notations, looking at Eq. (3.1), we improve the bound by means of the Golden-Thompson inequality:

$$\langle F_l^{(N)}\theta F_l^{(N)}\rangle$$
$$= \mathrm{Tr}\, e^{-\beta\{[H_1+H_2(U=0)]+[H_2-H_2(U=0)]\}} e^{-\beta[H_2-H_2(U=0)]}$$
$$\leqslant \mathrm{Tr}\, e^{-\beta[H_1+H_2(U=0)]} e^{-\beta[H_2-H_2(U=0)]}$$
$$\leqslant \mathrm{Tr}\, e^{-\beta[H_1+H_2(U=0)]}, \tag{A8}$$

where we used $[H_2-H_2(U=0)]\geqslant 0$. Equation (A8) is equivalent in the path-integral formalism to bounding the right-hand side of Eq. (A7) setting everywhere $U=0$ in S_2^ϵ. Let us define

$$\hat{\Xi}^\epsilon(U) \equiv \int (\mathcal{D}H\mathcal{D}H^*)_\epsilon e^{-S_1^\epsilon(H,H^*,0,U)+S_2^\epsilon(H,H^*,0)}\big|_{\partial_0 U=0} \tag{A9}$$

and let $\hat{U}$ denote the gauge-field configuration maximizing $\hat{\Xi}^\epsilon(U)$, then the right-hand side of Eq. (A7), using Eq. (A8), is bounded by $\hat{\Xi}^\epsilon(\hat{U})$ and taking the limit $\epsilon\searrow 0$ we recover Eq. (3.3).

[1] B. I. Shraiman and E. D. Siggia, Phys. Rev. Lett. **61**, 467 (1988); **62**, 1564 (1989).

[2] See, e.g., W. E. Pickett, Rev. Mod. Phys. **61**, 433 (1989).

[3] Z. X. Shen and D. S. Dessau, Phys. Rep. **253**, 1 (1995).

[4] P. W. Anderson, *The Theory of Superconductivity in the High-T_c Cuprates* (Princeton University Press, Princeton, NJ, 1997).

[5] See, e.g., A. J. Millis *et al.*, J. Phys. Chem. Solids **56**, 1641 (1995).

[6] M. Takigawa *et al.*, Phys. Rev. B **43**, 247 (1991); M. Takigawa, *ibid.* **49**, 4158 (1994); T. Imai *et al.*, Physica C **162**, 169 (1989).

[7] J. W. Loram *et al.*, Physica C **235**, 134 (1994).

[8] See, e.g., B. Batlogg *et al.*, Physica C **235**, 130 (1994).

[9] J. Rossat-Mignod *et al.*, Physica B **169**, 58 (1991).
[10] D. S. Marshall *et al.*, Phys. Rev. Lett. **76**, 4841 (1996); A. G. Loeser *et al.*, Science **273**, 325 (1996); H. Ding *et al.*, Nature (London) **382**, 51 (1996).
[11] P. W. Anderson, Science **235**, 1196 (1987); G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).
[12] Z. Zou and P. W. Anderson, Phys. Rev. B **37**, 627 (1988); See, also, S. E. Barnes, J. Phys. F **6**, 1375 (1976); P. Coleman, Phys. Rev. B **29**, 3035 (1984); N. Read and D. M. Newns, J. Phys. C **16**, 3273 (1983).
[13] D. Yoshioka, J. Phys. Soc. Jpn. **58**, 32,1516 (1989).
[14] A. P. Arovas and A. Auerbach, Phys. Rev. B **38**, 316 (1988).
[15] See, e.g., H. Fukuyama, Prog. Theor. Phys. Suppl. **108**, 287 (1992).
[16] See, e.g., C. L. Kane *et al.*, Phys. Rev. B **41**, 2653 (1990).
[17] See, e.g., a review by Yu Lu, in *Recent Progress in Many-Body Theories*, edited by T. L. Ainsworth *et al.* (Plenum, New York, 1992), Vol. 3, p. 157.
[18] See, e.g., S. P. Feng, Z. B. Su, and L. Yu, Phys. Rev. B **49**, 2368 (1994), and references therein.
[19] G. Baskaran and P. W. Anderson, Phys. Rev. B **37**, 580 (1988).
[20] L. Ioffe and A. Larkin, Phys. Rev. B **39**, 8988 (1989).
[21] P. A. Lee and N. Nagaosa, Phys. Rev. B **46**, 5621 (1992).
[22] X. G. Wen and P. A. Lee, Phys. Rev. Lett. **76**, 503 (1996); P. A. Lee, N. Nagaosa, T. K. Ng, and X. G. Wen, Phys. Rev. B **57**, 6003 (1998).
[23] A. P. Kampf and J. R. Schrieffer, Phys. Rev. B **41**, 6399 (1990).
[24] A. V. Chubukov, D. K. Morr, and K. A. Shakhnovich, Philos. Mag. B **74**, 563 (1996); A. V. Chubukov and D. K. Morr, Phys. Rev. B **57**, 5298 (1998).
[25] S. C. Zhang, Science **275**, 1089 (1997).
[26] P. Jordan and E. Wigner, Z. Phys. **47**, 631 (1928).
[27] E. Lieb and D. Mattis, J. Math. Phys. **6**, 304 (1965); A. Luther and I. Peschel, Phys. Rev. B **12**, 3908 (1975); S. Coleman, Phys. Rev. D **11**, 2088 (1975).
[28] E. Fradkin, *Field Theories of Condensed Matter Systems* (Addison-Wesley, New York, 1991).
[29] A. O. Gogolin, A. A. Nersesyan, and A. M. Tsvelik, *The Bosonization Approach to Strongly Correlated Systems* (Cambridge University Press, Cambridge, in press).
[30] S. Mandelstam, Phys. Rev. D **11**, 3026 (1975).
[31] J. Fröhlich and P. A. Marchetti, Commun. Math. Phys. **116**, 127 (1988).
[32] E. Witten, Commun. Math. Phys. **92**, 455 (1984).
[33] I. Affleck, Nucl. Phys. B **265**, 409 (1986).
[34] E. Mele, Phys. Scr. **T27**, 82 (1988); G. W. Semënoff, Phys. Rev. Lett. **61**, 517 (1988); E. Fradkin, *ibid.* **63**, 322 (1989); G. Baskaran, in *Two-dimensional Strongly Correlated Electron Systems*, edited by Z. Z. Gan and Z. B. Su (Gordon and Breach, New York, 1989), p. 72; M. Lüscher, Nucl. Phys. B **326**, 557 (1989).
[35] S. C. Zhang, T. H. Hansson, and S. Kivelson, Phys. Rev. Lett. **62**, 82 (1989); D. H. Lee and S. C. Zhang, *ibid.* **66**, 1220 (1991).
[36] *Fractional Statistics and Anyon Superconductivity*, edited by F. Wilczek (World Scientific, Singapore, 1990).
[37] R. B. Laughlin, Science **242**, 525 (1988); Int. J. Mod. Phys. B **5**, 1507 (1991).
[38] J. Fröhlich, T. Kerler, and P. A. Marchetti, Nucl. Phys. B **374**, 511 (1992).
[39] J. Fröhlich and P. A. Marchetti, Phys. Rev. B **46**, 6535 (1992).
[40] P. A. Marchetti, Nucl. Phys. B (Proc. Suppl.) **33C**, 134 (1993).
[41] J. W. Negele and H. Orland, *Quantum Many Particle Systems* (Addison-Wesley, New York, 1988).
[42] E. Witten, Commun. Math. Phys. **121**, 351 (1989); J. Fröhlich and C. King, Int. J. Mod. Phys. A **4**, 5321 (1989).
[43] A. Lerda and S. Sciuto, Nucl. Phys. B **410**, 577 (1993).
[44] P. A. Marchetti, Zhao-Bin Su, and Lu Yu, Nucl. Phys. B **482** [FS], 731 (1996).
[45] M. Ogata and H. Shiba, Phys. Rev. B **41**, 2326 (1990); H. J. Schulz, Phys. Rev. Lett. **64**, 2831 (1990); Y. Ren and P. W. Anderson, Phys. Rev. B **48**, 16 662 (1993); H. Frahm and V. E. Korepin, Phys. Rev. B **42**, 10 553 (1990); N. Kawakami and S. K. Yang, J. Phys. C **3**, 5983 (1991).
[46] J. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774 (1988).
[47] G. Kotliar, Phys. Rev. B **37**, 3664 (1988).
[48] Xi Dai, Zhao-Bin Su, and Lu Yu, Phys. Rev. B **56**, 5583 (1997).
[49] L. B. Ioffe and P. B. Wiegmann, Phys. Rev. Lett. **65**, 653 (1990). The U(1) effective action derived in this paper looks similar to the one considered there, but those authors considered holes in the Nagaoka state, while we start from the AF state. See, also Ref. 51.
[50] P. A. Marchetti, Z. B. Su, and L. Yu (unpublished).
[51] See N. E. Mavromatos, Nucl. Phys. B (Proc. Suppl.) **33C**, 145 (1993), and references therein; see, also, K. Farakos and N. E. Mavromatos, Phys. Rev. B **57**, 3017 (1998); cond-mat/9710288 (unpublished).
[52] I. J. R. Aitchison and N. E. Mavromatos, Phys. Rev. B **53**, 9321 (1996).
[53] P. A. Marchetti, Z. B. Su, and L. Yu, Mod. Phys. Lett. B **12**, 173 (1998).
[54] E. H. Lieb and M. Loss, Duke Math. J. **71**, 337 (1993).
[55] Y. Hagesawa, P. Lederer, T. M. Rice, and P. B. Wiegmann, Phys. Rev. Lett. **63**, 907 (1989).
[56] J. Bellissard and R. Rammal, Europhys. Lett. **13**, 205 (1990).
[57] See, e.g., Lu Yu, Zhao-Bin Su, and Yanmin Li, Chin. J. Phys. **31**, 579 (1993).
[58] J. Ginibre, Commun. Math. Phys. **10**, 140 (1968); J. Ginibre and C. Gruber, *ibid.* **11**, 198 (1968).
[59] D. N. Sheng, Zhao-Bin Su, and Lu Yu, Phys. Rev. B **42**, 8732 (1990).
[60] S. F. Edwards and A. Lenard, J. Math. Phys. **3**, 778 (1962); J. M. Kosterlitz and D. V. Thouless, J. Phys. C **6**, 1181 (1973).
[61] J. Fröhlich, Commun. Math. Phys. **47**, 233 (1976).
[62] D. F. Escande and B. Souillard, Phys. Rev. Lett. **52**, 1296 (1984).
[63] To be more precise, the AF long-range order (AFLRO) does not disappear at infinitesimal dopings, and there is a finite δ_c below which the AFLRO survives. However, in our approach we have not considered in detail the mechanism by which the AFLRO disappears (we have not carried out explicitly the renormalization procedure of the NLσ model at finite dopings, etc.). Since we are mainly interested in the region where only short-range AF order exists, this should be acceptable. To make a ''partial remedy,'' probably, it is more correct to replace δ by $\delta-\delta_c$ in some of our formulas.
[64] J. Fröhlich and J. Bricmont, Nucl. Phys. B **280** [FS], 385 (1987).
[65] First of all, the standard slave-fermion formalism involves only the U(1) gauge field, whereas our model deals with both U(1) and SU(2) gauge fields. Secondly, there is an obvious difference with the slave-fermion approach, namely, the contraint in our case is: $\Sigma_\alpha b_\alpha^* b_\alpha = 1$, while for the latter it is $\Sigma_\alpha b_\alpha^* b_\alpha + h^* h = 1$, with h as the holon operator. Moreover, these two approaches are still different after a ''renormalization'' to take care of the above difference, as explained in the main text.

[66] The doping dependence of the coupling constant in the NLσ model was considered earlier [N. Dorey and N. E. Mavromatos, Phys. Rev. B **44**, 5286 (1991)]. However, the physically correct zero-doping limit, to recover the antiferromagnetic long-range order, has not been derived up to now.

[67] R. T. Birgeneau *et al.*, Phys. Rev. B **38**, 6614 (1988).

[68] T. Sakai and Y. Takahashi, cond-mat/9712104 (unpublished).

[69] The Dirac form for the holon action in the presence of π flux was considered earlier [N. Dorey and N. E. Mavromatos, Phys. Lett. B **250**, 107 (1990)]. However, there is a substantial difference between the action considered there and ours. The presence of the $\gamma_0 \delta$ term in our action S_h^* makes the problem nonrelativistic (lacking the boost symmetry due to the presence of a finite Fermi surface). One, in principle, can ''gauge away'' it by shifting A_0. However, the nonvanishing expectation value of A_0 means the same thing, because the quantum field should be expanded around this expectation value. It gives rise to Reizer singularity, leading to non-Fermi-liquid behavior.

[70] The normal-state properties of cuprates were also studied earlier using the C-S approach, Ref. 52. However, the gauge-field propagator in that formulation is still fully relativistic, so the plausible non-Fermi-liquid behavior considered there has a different origin.

[71] Y. Chen, D. Förster, and A. Larkin, Phys. Rev. B **46**, 5370 (1992); See, also, P. Beran, D. Poilblanc, and R. B. Laughlin, cond-mat/9505085 (unpublished).

[72] D. C. Brydges, J. Fröhlich, and E. Seiler, Ann. Phys. (Leipzig) **121**, 227 (1979).

[73] See, e.g., G. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View* (Springer, New York, 1981); E. Seiler, *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics* (Springer, New York, 1982).