PHYSICAL REVIEW B VOLUME 38, NUMBER 7 1 SEPTEMBER 1988

# Electronic structure of $CuO_2$ sheets and spin-driven high-$T_c$ superconductivity

E. B. Stechel and D. R. Jennison
*Solid State Theory Division 1151, Sandia National Laboratories, Albuquerque, New Mexico 87185*

(Received 22 March 1988)

Energy parameters for a $CuO_2$ sheet, taken to be prototypic of the high-temperature superconductors, are derived from semiempirical and *ab initio* sources. Intra-atomic Coulomb interactions ($U_i$) are large, but *inter*atomic Coulomb terms and direct oxygen-oxygen transfer integrals are also very important. These energies dictate a two-band extended Hubbard Hamiltonian which cannot obviously be simplified. With $Cu(d^{10})O(p^6)$ as the vacuum state, interatomic Coulomb interactions create a potential well resulting in hole localization of one hole per $CuO_2$ unit cell, so that the $Cu(d^9)$ valence is dominant. A spin-$\frac{1}{2}$ Heisenberg system thus exists independent of the presence of carriers due to the poor screening in these materials. We compute the Cu-Cu superexchange energy $J$ from the other parameters and find good agreement with empirically derived values, provided the inclusion of direct Cu-O exchange. Because of the relatively large value of $J$, we assume *local* antiferromagnetic (AF) order. Itinerent carriers exist on the oxygen sublattice because of the large Cu $U_d$ energy. The coexisting spin and carrier systems interact strongly, the most important cause being a virtual process involving the $Cu(d^{10})$ configuration, which is lowered in relative energy by Coulomb interactions with the carrier. This process can produce carrier transport with and without creating spin deviations and stabilizes holes in the oxygen $p_\sigma$ orbitals. We find that the carriers are neither weakly coupled free particles nor spin polarons, but are something new: "spin hybrids," consisting of a coherent and nonperturbative mixture of local spin-orbital electronic configurations, some of which represent deviations in the local AF order. A model Hamiltonian that describes the spin-hybrid carriers shows that the probability of finding a spin deviation associated with an isolated carrier quasiparticle is large (30–40%). We find spin-driven electronic pairing in the Cooper sense between the spin-hybrid quasiparticles. Retarded interactions occur and we also find direct attractive *un*retarded interactions at ~3–4 Cu-O spacings. These interactions cause extended *s*-wave and *d*-wave pairing and lead to a pairing Hamiltonian reminiscent of Bardeen-Cooper-Schrieffer theory which is, however, only two dimensional. The possible role of Josephson tunneling in the third dimension is also discussed.

## I. INTRODUCTION

The high-$T_c$ Cu-O superconductors $La_{2-x}(Sr,Ba)_x$-$CuO_4$ (Ref. 1), $Y_1Ba_2Cu_3O_{\sim 7}$ (Ref. 2), and the more recent Bi-Sr-Ca-Cu-O (Ref. 3) and Tl-Ba-Ca-Cu-O (Ref. 4) have electronic structures of unusual complexity, thereby raising interesting questions of valency, correlation, magnetism, and the nature of the charge carriers. Naturally, a prime goal is to determine which aspects of these materials cause high-temperature superconductivity. Our approach has been to systematically build a fundamental understanding, by relying only on those experimental and theoretical sources which are reasonably well understood. Our program therefore consists of five parts, the first four of which are realistically addressed in this paper: (1) we derive values of the important energy parameters from both *ab initio* and semiempirical sources; (2) these in turn define a Hamiltonian which contains the essential physics; (3) the origin of the spin system and the nature of the charge-carrying quasiparticles become clear, the understanding of which enables us to (4) explore the pairing interactions between the carriers. The last part, the phase transition from the normal state to the superconducting state, is not understood, since it is unclear whether the transition is Kosterlitz-Thouless[5] and thus purely a property of the two-dimensional (2D) $CuO_2$ sheets or a 3D transition perhaps driven by Josephson tunneling between the sheets. In this paper the solution of the pairing Hamiltonian is addressed but only using the mean-field approximation. This cannot realistically predict $T_c$ because in two dimensions static long-range order exists only at $T=0$.[6] Whether the phase transition is Kosterlitz-Thouless[5] or 3D driven by the weak Josephson tunneling between the sheets, the physics which determines $T_c$ is different than in the Bardeen-Cooper-Schrieffer (BCS) case.[7] We do find, however, that superconductivity involves Cooper pairs[8] (i.e., paired fermions, although coupled by an electronic instead of a phonon mechanism) as the carrier quasiparticles do not bind into bosons.

There are three fundamental points which highlight this work. The first is the coexistence of spin and carrier subsystems in a $CuO_2$ sheet: i.e., a Heisenberg spin-$\frac{1}{2}$ system exists independent of the presence of charge carriers. We argue that this is a direct consequence of poor screening which leads to large correlations between the quasiparticle holes. The second point is that because there are large interactions between the carrier and spin-$\frac{1}{2}$ subsystems, the carrier quasiparticles contain a coherent mixture of configurations with different local spin couplings. We term these carriers "spin hybrids" because they are formally similar to conventional carriers which are hybrids of different atomic orbitals. The third point is that because the spin hybrids contain large admixtures of configurations which involve deviations from the local antiferromagnetic (AF) order, there is a pairing interaction of a purely electronic origin: large matrix elements exist for the exchange of virtual magnons between quasiparticles and, in addition, there are direct interactions between adjacent carriers, thus producing an *un*retarded (in time) contribution to pairing. We therefore propose that high-temperature superconductivity is caused by the existence of the spin-hybrid quasiparticles which necessarily leads to spin-driven electronic pairing within the $CuO_2$ sheets. Furthermore, we speculate that superconducting phase coherence could be driven by Josephson tunneling in the third dimension. (The fact that electronic pairing can result from the coexistence of spin and carrier systems has, of course, been suggested before;[9-13] we will see, however, that the physical picture contained in the present paper differs considerably from previous work.) Before outlining this paper, we expand on these major points.

A most important aspect of these systems, and one which distinguishes them from many other compounds, is that the outermost atomic shell of the transition-metal ion (the Cu $3d$ shell) is nearly filled, as also of course are the $2p$ shells of the oxygen ions. Quasiparticle holes in these otherwise filled levels in a $Cu(d^{10})$-$O(p^6)$ array have large screened Coulomb interactions as a result of the inability of the system to effectively screen holes on the same or neighboring sites from each other. These interactions constitute the extended Hubbard $U_{ij}$, where $i$ and $j$ may denote different sites. As we will discuss in Sec. II, large $U$'s arise when screening charge cannot be transferred into the same principal atomic shell that contains the quasiparticles (because the shell is full in the absence of the quasiparticles). These energies (the on-site $U_i \equiv U_{ii}$ components of which are known from the interpretation of Auger and photoemission spectra from related compounds) are sufficiently large to suppress the $d^8$ configuration on the Cu ions and cause strong correlation between the quasiparticle holes. In particular, we argue that the existence of a substantial interatomic $U$ ($U_{ij}$, $i \neq j$) causes a Coulomb localization of the first hole per unit cell in the two-band extended Hubbard model which is appropriate to a $CuO_2$ sheet. (Interatomic $U$ is also thought by Varma, Schmitt-Rink, and Abrahams[14] to play an important role; however, this is in connection with a different pairing mechanism than we find.) This factor, together with the differences in the single-site energies, leads to the dominance of the $d^9$ configuration at the Cu sites (while the admixture of $d^{10}$ is significant, the $d^8$ component is small). Furthermore, this localization occurs independent of the presence of carriers. This constitutes our first major point.

We do not assume or expect that the spin subsystem is magnetic (i.e., has long-range order) in the presence of carriers. We do expect, however, that there are strong short-range antiferromagnetic correlations between the Cu spins due to the magnitude of the superexchange energy between Cu spins. The system is thus dynamic[15] in that the spin orientation changes rapidly (perhaps $\sim 10^{-14}$ sec) when compared to the time scale of spin-resonance experiments (which is why NMR works[16]), but slowly on the time scale of carrier motion ($\sim 10^{-16}$–$10^{-15}$ sec) so that the carriers experience local AF order in the Cu spins.

Additional holes, the charge carriers, are introduced into the material by doping[1] or by having a surplus of oxygen[17] in the La material or by having greater than 6.5 oxygens per unit cell in the yttrium compound.[2] (In practice, the existence of Coulomb potentials produced by dopants or other point defects in the crystal could bind a small density of carrier holes, leading to an insulating phase with a small carrier density and a insulator-metal transition with increased doping.[18] In this paper, we focus on the itinerant carriers of the metallic phase.) We find that the carrier holes in $CuO_2$ sheets have most of their amplitude in the oxygen $2p$ levels. This is a consequence of the large on-site $U$ for Cu ($\sim 9$ eV) which inhibits the $d^8$ configuration.

Our starting point for understanding charge transport is thus a model of two coexisting systems: a quasi-2D spin-$\frac{1}{2}$ Heisenberg spin system principally on the Cu sublattice and a carrier system principally on the O sublattice. We find that these two subsystems interact strongly due to the relatively small Cu-O distance and thus the large transfer integrals, the near degeneracy of the $Cu(3d)$ and $O(2p)$ levels, poor screening which causes carrier presence to enhance local $Cu(d^{10})$ valence fluctuations, and the spin-$\frac{1}{2}$ character which allows easy self-healing of spin deviations as explained below. In addition, we find that the strong coupling stabilizes carrier holes in the $p_\sigma$ orbital (which points towards the Cu ions) by 1–2 eV compared with the $p_\pi$ orbitals which have only weak coupling to the spin system. Although recent work by Guo, Langlois, and Goddard[19] and by Birgeneau, Kastner, and Aharmony[20] have proposed $p_\pi$ hole occupancy (which is found if one only considers the Madelung potential in conjunction with strong localization of the first hole per unit cell on the Cu sites), we find this is not the case upon consideration of the carrier-enhanced valence fluctuations, a factor also seen in recent simulations by Hirsch *et al.*[21]

Much of this paper explores the consequences of the interactions between the spin and carrier subsystems. These determine the nature of the quasiparticles which carry charge and hence form the basis for understanding the dynamics of transport. If there were no interactions between these subsystems, the carrier holes would occupy the topmost part of the oxygen $2p$ bands: we refer to this as the "free-hole" limit. These bands have considerable width and consist of Bloch waves made principally from the $2p_x$ and $2p_y$ oxygen orbitals which lie in the plane of the 2D square lattice formed by the $CuO_2$ sheet. (We find that the transfer integrals between these orbitals are essential to any realistic description of the carrier quasiparticles.) In general, the interactions with the spin subsystem cause the carrier quasiparticles to be mixtures (i.e., hybrids) involving several spin-orbital configurations.

It is common to think of valence electrons (or holes) as hybrid particles: the components of the hybrid are usually atomic orbitals such as the coherent $4s$ and $3d$ mixtures which describe the valence electrons of the first-row transition metals. If these particles were to interact, and if the interactions were different between $4s$-$4s$, $4s$-$3d$, and $3d$-$3d$, then the resulting net interaction would depend on the relative amplitudes of the $4s$ and $3d$ components in the carrier quasiparticle. In the present case, the components of the carrier differ not just in their atomic-orbital character, but also in the nature of the local spin couplings. For this reason, we refer to these carriers as "spin hybrids." We shall see that since the spin-hybrid components do have different pairing interactions, the net pairing is just a weighted sum of the component pairings.

The components of the carrier quasiparticle are represented by the various spin-orbital configurations, constituting all relatively low-energy arrangements of carrier and spin subsystem holes, in different atomic orbitals and/or with different spin alignments between them. The carrier quasiparticle is then a mixture of configurations which necessarily involve more than one hole. The specific mixture is determined by diagonalizing the configuration-interaction Hamiltonian matrix which contains the interactions between the spin-orbital configurations and the effects of translational symmetry. In spite of this complexity, made necessary by the coexistence of the carrier and spin subsystems, the analogy to more usual hybrid carriers is very close.

For illustration, let us consider just two of the lowest-energy local spin-orbital configurations. We may write a carrier hole which is momentarily on an oxygen site between two Cu holes in a $CuO_2$ sheet as $d^9_\uparrow p^5_\downarrow d^9_\downarrow$, where the arrows indicate the $S_z$ component of the holes. We assume this configuration to be embedded in a host which is at least locally antiferromagnetic, so that each of the $d$ holes is also approximately antialigned with three other $d$ holes which for simplicity are not shown. Because these other couplings are not disturbed by the presence of the oxygen hole, this configuration represents a "free hole" in that the oxygen carrier hole may freely hop to equivalent oxygen sites without leaving behind a spin deviation (in what follows, we disregard the $p_\pi$ orbitals for reasons stated above; thus, "$p$" implies $p_\sigma$). This spin-orbital configuration interacts and hence mixes with another in which the carrier hole has opposite spin and is antialigned to *both* neighboring Cu sites, $\underline{d}^9_\downarrow p^5_\uparrow d^9_\downarrow$. This configuration represents a carrier-associated spin deviation in the local AF order (marked by the underline). We will call such entities "flipped-hole configurations" and note that if the oxygen hole hops away, it leaves behind a spin deviation in the lattice. Classically, at zero temperature this would not happen as an isolated carrier in a perfectly ordered AF sheet would raise the energy of the system by creating spin deviations. However, in quantum systems this is not true: at all temperatures there exists a finite density of spin deviations, some of which are not associated with carriers. A carrier can make a deviation at one site and unmake a different deviation at another site, thus preserving the energy of the spin system. In addition, a spin deviation produced by one carrier can be healed by another, leading to retarded pairing interactions.[11]

The interaction between the above two configurations ($d^9_\uparrow p^5_\downarrow d^9_\downarrow$ and $\underline{d}^9_\downarrow p^5_\uparrow d^9_\downarrow$) occurs through a single matrix element by direct Cu-O exchange, $K$, which is ferromagnetic; however, the strongest ties are through virtual (intermediate) states. The lowest-energy virtuals involve the Cu($d^{10}$) configuration and are of the type $p^5_{a\uparrow} d^{10} p^5_{b\downarrow} d^9_\downarrow$, where $p_a \neq p_b$. The $3d$-$2p$ transfer integral links virtual states such as these to free-hole and flipped-hole configurations on adjacent sites.

If we only consider three isolated (Cu-O-Cu) atoms, the net result of the interactions with virtual states which involve the $d^8$ and $p^4$ configurations is that an "indirect" Cu-O exchange energy, $K_{\text{eff}} - K$, dominates over the direct Cu-O exchange energy $K$ between adjacent Cu and oxygen holes, thus causing antiferromagnetic Cu-O spin alignments to be energetically favored in spite of the relatively large ferromagnetic $K$. Nonetheless, the largest factor which produces carrier-associated spin deviations is *not* $K_{\text{eff}}$ but the above Cu($d^{10}$) valence fluctuation. This is because the latter produces a larger matrix element to simultaneously exchange and move the carrier hole by going through the virtual intermediate state.

If the above two spin-orbital configurations (free hole and flipped hole) are close in energy relative to the strength of the interactions, strong mixing is inescapable. By diagonalizing a configuration-interaction (CI) matrix representing the above configurations, a wave function for the lowest-energy state is obtained. The atomic-orbital hybridization of the carrier particle, the extra hole, is thus found, but a carrier is also obtained which is "dressed" in its spin and orbital associations, characterized by the coherent superposition of the various configurations. This superposition is the conceptual basis of the "spin hybrid," the existence of which is our second major point.

We naturally expect that the lowest-energy spin-orbital configurations will dominate the ground-state carrier-quasiparticle wave function. If the above-mentioned coupling between the free-hole and flipped-hole configurations were weak and their local energies comparable, the free-hole configuration would dominate the carrier because it can freely delocalize through the periodic array of AF-ordered Cu spins. In this case, it would be convenient to think of the quasiparticle as a free hole dressed with spin exitations in a manner similar to weak-coupling phonon dressing. For this to be true, the energy of the free-hole configuration, after being adjusted for delocalization over the oxygen sublattice, must be much lower than the energy of the flipped-hole configuration after similar adjustment for translational symmetry. However, we will show that in the case of the Cu-O materials, the delocalization energy of the free-hole configuration is not great in and of itself, principally due to the reduced dimensionality. Furthermore, because of the above-mentioned $d^{10}$ virtuals, the difference in energy between the delocalized free-hole and flipped-hole configurations (the diagonal energy difference) scales with the off-diagonal matrix elements which promote mixing. A central point then is that it is incorrect to view the flipped-hole configuration simply as a resonance which is degenerate with the free-hole band, and which therefore would form a localized state (i.e., a true spin polaron) if its exchange energy $K_{\text{eff}}$ were great enough to move it outside the band. Because of the above scaling, a localized spin polaron will not form even if $K_{\text{eff}}$ were very large, even several times the bandwidth, because $K_{\text{eff}}$ is also coupling the two configurations. (For the CuO materials $K_{\text{eff}}$ is not a large energy. In Sec. IV we compute $K_{\text{eff}} \sim 0.1$ eV.) For all reasonable energy parameters, a spin hybrid always exists in that an itinerent carrier is a strong mixture of the free-hole and flipped-hole configurations. Thus we are never in the free-hole nor spin-polaron limits in spin-$\frac{1}{2}$ CuO$_2$ sheets, and furthermore perturbation theory cannot be used to explore the interactions between the carrier and spin subsystems or (as we shall see) between the carriers themselves.

Our third major point is that the substantial spin-deviation configuration character in the spin-hybrid carrier quasiparticle leads to an attractive electronic pairing interaction because of large matrix elements for carrier-magnon interactions (other than $K_{\text{eff}}$) which produce retarded pairing through magnon exchange (a process which is analogous to the exchange of virtual phonons in Cooper pairs). In addition, however, the spin-deviation component naturally leads to a direct and thus unretarded pairing interaction. The nature of this pairing is wholly spin derived: Let us use the label $A$ for the Cu sublattice which has spin down ($\downarrow$) locally and $B$ for the spin-up ($\uparrow$) sublattice. Assume that a spin distortion, a single Cu spin flip, on the $A$ spin lattice is associated with the flipped-hole configuration of a carrier. Then the Cu-Cu superexchange energy $J$ causes real-space configurations in which spin-up ($A$) and spin-down ($B$) flipped-hole configurations with the $A$ and $B$ Cu nearest neighbors to be lower in energy than configurations where they are apart. This contributes to extended $s$-wave and $d$-wave pairing as the attraction is in real space and occurs when the quasiparticles are a fixed distance apart. In the simplest picture, two isolated Cu spin deviations cost $zJ$ each or a total of $2zJ$, where $z$ is the number of nearest neighbors which in the 2D square lattice is 4; on the other hand, it costs only $2(z-1)J$ for two nearest-neighbor deviations, a savings of $2J$.

This "simplest picture" is too simple. To actually evaluate the energy of the pairing interaction between two spin-hybrid carriers is a complex matter due to the Coulomb interactions between the carrier and Cu holes and dynamical effects from carrier motion. This would require a configuration interaction calculation with up to 10 holes and 16 sites (which involves greater than $10^7$ bases before spin and symmetry decomposition). Therefore, in this paper we will not attempt any quantitative evaluation of the retarded or unretarded pairing interactions, but only qualitatively describe the pairing which may be expected as a consequence of the existence of the spin-hybrid carriers.

To address the question of the phase transition from the normal state to the superconducting state, we speculate on the possible role of the coupling in the third dimension. This we only consider qualitatively in this paper by arguing in analogy with the Néel transition in related insulating materials. We find it appealing, *if* the phase transition is *not* Kosterlitz-Thouless,[5] that there is a formal resemblance between the superconducting and the Néel transitions, with the Josephson tunneling matrix elements in the third dimension playing the role of the superexchange perpendicular to the CuO$_2$ sheets. If so, the highest-$T_c$ materials may have the strongest tunneling between the CuO$_2$ sheets.

In the sections that follow, we provide the arguments which support each of these major points and present the mathematical basis for this qualitative discussion. This paper is necessarily long as each major point arises because of the work in preceding sections. We believe the coherence and actual simplicity of the total picture, when viewed in its entirety, argues for the validity of the basic model. Because of the length, we have attempted to make the paper slightly redundant, so that the casually interested reader can skip the more-detailed parts if he or she so chooses.

In Sec. II we discuss the *ab initio* and semiempirical values of the energies necessary for understanding these materials, including the origin of the large $U$'s which cause the breakdown of the band-structure description. We also see that the minimum relevant Hamiltonian is more complex than a simple single-band or even two-band Hubbard model. We report model calculations in Sec. III which indicate the origin of the spin subsystem and the degree of valence fluctuations on the Cu sites. We also see why the spin and carrier subsystems exist on separate sublattices. In Sec. IV, as a self-consistency check, we discuss the computation of the Cu-Cu superexchange from other energy parameters and see that it compares well to the experimental results. We also compute the effective Cu-O exchange energy, $K_{\text{eff}}$. In Sec. V we investigate the spin-$\frac{1}{2}$ Heisenberg system in the absence of carriers in order to introduce concepts and notation which are essential for understanding the nature of carrier-spin interactions and transport. In Sec. VI we write a Hamiltonian which describes the spin-hybrid carriers and allows us to study its properties over a wide range of energy parameters. The Hamiltonian which describes the pairing interactions between spin hybrids is presented in Sec. VII. The second part of Sec. VII gives a prescription for a mean-field solution of the pairing Hamiltonian. Section VIII is a qualitative discussion of the possible role of the third dimension and Josephson tunneling in the phase transition to the superconducting state. In the Discussion (Sec. IX), we compare our work to that of others and review arguments for the validity of the major points. In what follows, we emphasize the generality of the formalism.

## II. THE ENERGY PARAMETERS OF A Cu-O$_2$ SHEET

We have derived a set of reasonable energy parameters in order to guide our thinking towards a model. All that follows is based on the assumption that the energy parameters are at least approximately correct.

The spin and carrier, or two-subsystem, model for the electronic structure of these materials is a direct consequence of the strong correlations between hole quasiparticles in the Cu($3d^{10}$)-O($2p^6$) shells. Moreover, strong correlation is quite expected. The nearly closed-shell nature of both the Cu and the O ions causes screening to be inefficient, leading to large values of the screened hole-hole interactions (the generalized Hubbard $U$'s). This occurs not only in the Cu and O single-site terms, $U_d$ and $U_p$, respectively, but also for Coulomb interactions between holes on nearby atoms, i.e., the interatomic $U$ terms (which is to say that on a length scale shorter than the average carrier-carrier spacing, the material exhibits "dielectric" as opposed to metallic screening).

The basic understanding of the relationship between large $U$'s and filled shells may be found in work done quite a while ago in an attempt to explain Auger spectra from metallic Cu. Kowalczyk *et al.*[22] showed that the value of $U_d$ in Cu metal, determined from the $L_{2,3}VV$ Auger spectra, was 7.7 eV and could be simply understood in terms of "atomic" energies. [$U_d$ is 16.5 eV for a gas-phase Cu atom, obtained from the values of the first and second ionization potentials; the difference between this and the atomic $F_0$ Coulomb integral ($\sim$27 eV) has been termed "static relaxation" by Shirley[23] and involves the relaxation of the atomic electrons from their ground-state orbitals due to the presence of two holes on the atom.] In the solid, $U_d$ is reduced from the atomic value of 16.5 eV due to charge transfer between atoms and to polarization of the surrounding medium; the former dominates in the case of a metal. If one assumes complete short-range screening, this is equivalent to a local atomic configuration which is essentially $3d^8 4s^2 4p^1$ in the Auger final state. Comparing the energy of this atomic configuration to twice that of $3d^9 4s^2$ and to the ground state, Kowalczyk *et al.*[22] found a value of $U_d$ close to that observed.

In the Cu-O superconductors, charge-transfer screening is reduced from that of the metal. The reason for this is that the neighboring $2p$ oxygen shells are already full so that polarization or charge-transfer screening must involve the $4s$, $4p$ and $3s$, $3p$ shells of the Cu and O atoms, respectively. It is difficult to imagine how this could lead to better local screening than in the case of the metal since charge transfer into the $4s$ and $4p$ levels was taken there[22] to be perfect. Thus we believe that $U_d$ in the metal represents a lower bound for $U_d$ in the superconductors. For the oxygen single-site term, a $KVV$ Auger-derived value of $U_p \sim 6$ eV is typical for oxides.[24] (Because of uncertainties in the characterization of surfaces in actual high-$T_c$ compounds, note that we have used Auger-derived values only from related well-characterized materials.)

The above discussion concerns a "static" value for $U$ because it was assumed that the charge-transfer screening involves complete relaxation. This approximation was appropriate because the screening length is very short in the metal and the sharp lines seen in the Auger spectra of Cu are more than an order of magnitude narrower than the width of the Cu $4s$-$4p$ bands so that the final state is long-lived compared with screening relaxation times. A dynamical value of $U$, which is appropriate for describing the valence fluctuations in the Cu-O materials, for example, would be larger than the static $U$ because if the fluctuations were fast, screening relaxation may be incomplete. Thus the values of $U$ derived from Auger spectra represent *lower* bounds on the values appropriate for questions of quasiparticle-hole dynamics.

It is interesting to compare values of $U$ determined from the $L_{2,3}VV$ Auger spectra of Cu, Ni, and Co metals. The 7.7-eV value for Cu is reduced to $\sim$2 eV for Ni (Ref. 25) and is zero within experimental uncertainties for Co (Ref. 26), thus allowing the recovery of a one-electron description of the two-hole excitations in Co. This dramatic reduction is due to the opening of the $3d$ shell so that screening charge can be accepted (in the case of Cu screening charge is constrained to the $4s$, $4p$ shell, one principal atomic shell higher than that which contains the quasiparticle holes). Thus the large values of $U$ in the Cu-O superconductors are due to the closed shells and are an inescapable consequence of closed shells.

There have been some recent *ab initio* calculations of $U$ by Schlüter, Hybertsen, and Christensen,[27] by Harmon and co-workers,[28] and by McMahan, Martin, and Satpathy.[29] All three workers used band-structure calculations with supercells to self-consistently compute the relaxation of the medium to the presence of two holes on a Cu or an O site. Remarkably, all found $U_d \sim 7$–$11$ eV on the Cu sites which agrees well with the Auger-derived value. In addition, all also found a large $U_p$ on the oxygen sites. The semiempirical (Auger) values for Cu are close to all of the *ab initio* results; in addition, the values of $U_p$ in Refs. 27 and 29 are $\sim$6–8 eV, also close to the Auger value. Due to the difficulty of the calculations, we will adopt for use in this paper the Auger value of $U_p \sim 6$ eV as it was measured on a metal oxide. We use the calculated value, however, of $U_d \sim 9$ eV as it is slightly greater than the value for Cu metal as we expect it would be from the general arguments presented above. We also note that the value for $U_d$ is probably more certain than that of $U_p$. In spite of the above concerns for the dynamic versus static values for $U$, we will see in Sec. IV that the Auger-derived values, when used in a Hubbard-like model Hamiltonian, do indeed predict values of the superexchange, $J$, which are in agreement with experiment.

Because the $U$'s are large, the one-electron picture is an inconvenient starting point from which to analyze the electronic structure; the single-particle wave function by definition leads to fluctuations in the Cu valency which are only determined by the statistics of a single Slater determinant and not by the relative energies of the Cu-O configurations. A more-appropriate picture and one which we adopt in this work is a configuration-interaction model of valence fluctuations. In order to study the latter, in addition to the values for $U$ we need (among other energies) the transfer integrals between Cu and O sites ($t_{pd}$) and O and O sites ($t_{pp}$).

We use the band-structure calculations to estimate the transfer integrals. If one makes a tight-binding fit to the results of Mattheiss[30] in the limit of only nearest-neighbor Cu-O interactions, one obtains a value of $\sim$1.6 eV for the transfer integral $t_{pd}$. However, a calculation of the overlap of the ion orbitals between Cu-O and O-O (orbitals obtained by computing the ions in the Hartree-Fock approximation while in Madelung wells appropriate to $La_2CuO_4$) reveals that at distances appropriate to the $CuO_2$ sheet the overlaps are 0.07 and 0.04, respectively [between the $d(x^2-y^2)$ and the $p(x)$ and $p(y)$ orbitals taken as oriented in the *crystal* reference frame]. Thus the direct oxygen-oxygen transfer integral $t_{pp}$ cannot be neglected. Further support for this assertion can be found from photoemission studies of oxides such as MgO in which there are no metal $d$ levels to play a significant role in the oxygen $2p$ bandwidth, which nevertheless is $\sim 6$ eV,[31] corresponding to a tight-binding $t_{pp}$ of 0.25 eV. Since the O-O distance is $\sim 0.3$ Å (10%) smaller in the CuO$_2$ sheet than in MgO, $t_{pp}$ should be significantly larger than this value in the superconductors. Assuming a Huckel model and scaling the above $t$'s by the overlaps, one finds from the band structure and also from scaling the MgO value that $t_{pd} \sim 1.0$ eV and $t_{pp} \sim 0.6$ eV.[32] Further support for the importance of $t_{pp}$ comes from work of McMahan, Martin, and Satpathy,[29] who obtained tight-binding values of $t_{pp\sigma} \sim 1.0$ eV and $t_{pp\pi} \sim -0.3$ eV for the oxygen bands, which also results in a value of $t_{pp} \sim 0.6$ eV when rotated to the crystal reference frame.

Note especially that if $t_{pp}$ were not included, $t_{pd}$ would have to be larger in order to fit the band structure. However, we shall see in Sec. IV that the superexchange $J$ is very sensitive to $t_{pd}$ and that the inclusion of $t_{pp}$ not only provides for a more-accurate description of carrier transport, but also allows $t_{pd}$ to be small enough that the calculated $J$ is in good agreement with experiment. Many workers have left out considerations of $t_{pp}$ in their models.[11,12] We find that its inclusion is essential for any realistic description of the carrier quasiparticles, as will be seen below.

An additional energy difference needed for our calculation is that of the one-electron or diagonal energies ($\varepsilon_i$) of the $3d$ and $2p$ orbitals. Band-structure results show that the $3d$ and $2p$ energies are nearly the same, which naively suggests $\varepsilon_d \approx \varepsilon_p$. However, anticipating the results of the next section, this would cause the holes to prefer neither the Cu nor oxygen sublattices (in the absence of $t_{pp}$, including this term would cause the first hole per unit cell to favor the oxygen sublattice) even though they are localized by their Coulomb interactions. Emery has pointed out[11,33] that if we take the band-structure results as a mean-field solution of a Hubbard Hamiltonian, then the difference in energies $\Delta\varepsilon \equiv \varepsilon_d - \varepsilon_p \approx U_d/4 - U_p/8 \sim 1.5$ eV with the $3d$ orbital lower in energy. We accept these results and consider $\Delta\varepsilon$'s from 0.5 to 1.5 eV in what follows. (Very recent work[27,29] has suggested $\Delta\varepsilon \sim 2$ eV which will not qualitatively change our results.) Within this range for $\Delta\varepsilon$, a sufficient concentration of the spin-system holes resides on the Cu sublattice so as to agree with the spectroscopic results which indicate the dominance of $d^9$.

The interatomic $U_{ij}$ are important (but very uncertain) energy parameters, values for which are essential for any quantitative understanding of these materials, as illustrated in Sec. III. When $i$ and $j$ are nearest neighbors, $U_{ij}$ is in an intermediate regime between "long-range" $e^2/\epsilon_\infty r$ dielectric screening and the "short-range" single-site $U$'s which have been determined by Auger spectroscopy. Here $\epsilon_\infty$ is the optical frequency dielectric constant which is $\sim 7$ in CuO; we presume this to be a lower bound within a CuO$_2$ sheet as the electron density within the sheet is higher than average for the crystal. We shall use a value of $\epsilon_\infty \sim 10$ for the sheet. While $e^2/\epsilon_\infty r$ represents lower bounds for the interatomic $U$'s, upper bounds may be estimated based on the amount one may expect the bare $e^2/r$ to be reduced by short-range screening, found by comparing the single-site value $U_d$ between the solid state and the gas phase. With $U_d \sim 9$ eV in the solid versus 16.5 eV in the gas, the upper bounds are $\sim 3$ eV for nearest-neighbor oxygen-oxygen ($U_{pp}$) and $\sim 4$ eV for copper-oxygen ($U_{pd}$) Coulomb interactions. Now making the assumption that the holes are better screened when apart (than they are if together on the same site) leads to what actually are guesses: $U_{pd} \sim 1$–$3$ eV and $U_{pp} \sim 0.7$–$2$ eV. Calculations of these quantities are very difficult to do with accuracy, so we adopt these values for the present. Schlüter *et al.*[27] have also estimated that $U_{pd} < 4$ eV based on different reasoning, while Harmon and coworkers estimate $U_{pd} \sim 1.6$ eV, which is larger than the calculated value of 0.6–0.9 eV of McMahan *et al.*[29] The latter would indicate very rapid convergence to the dielectric limit. It is also noteworthy that a satellite has been observed in the Auger spectrum of the superconducting materials, which if assigned to a charge exchange excitation suggests a value of $U_{pd} \sim 2$ eV.[34]

Some previous work[11,12] has assumed that carrier holes on the oxygen sublattice must hop through the Cu ions in order to move; however, this is not necessary. If fact, with the above numbers $t_{pd}^2/(\Delta U_d - \Delta\varepsilon) \sim 0.3 \pm 0.15$ eV, which is less than $t_{pp}$ so that the direct oxygen-oxygen hop is the largest contribution to charge-carrier motion ($\Delta U_d$ is the difference in Coulomb energy for the carrier hole being on the Cu site versus being on the O site $\sim U_d - 2U_{pd}$). However, we shall see that there is also another substantial contribution to carrier motion, which involves the $d^{10}$ fluctuation as described in Sec. VI.

The direct Cu-O exchange energy $K$ is important for our calculation of the superexchange energy $J$ and for understanding the effective exchange energy between carrier holes and the localized Cu holes. This number was obtained from a nonorthogonal basis set from first principles, as described in Ref. 35. It was found that $K \sim -0.22$ eV; this number can be corrected for nonothogonality using other two-electron integrals[35] which increases it to $-0.17$ eV. The Cu-O singlet-triplet energy difference is then $|2K|$ and favors local ferromagnetic Cu and O spin alignments, thus competing with the "$t_{pd}^2/\Delta U$" terms[11] which require singlet Cu and O coupling. This magnitude for $K$ is much larger than in most systems (it is commonly neglected in model Hamiltonians) but is significant here due to the relatively small Cu-O distance: its magnitude falls by 50% as one increases the distance from 1.89 to 2.0 Å. Nonetheless, the net interactions favor singlet coupling with $K_{\text{eff}} \sim 0.1$–$0.2$ eV (see Sec. IV and Appendix A).

The above set of energy parameters define our model Hamiltonian. However, the model would not be substantially altered even if the energy parameters are subsequently refined, within reason. The major conclusions of this section are that the single-site and interatomic values of $U$ are large (with the former of the order of the single-particle valence bandwidth), direct oxygen-oxygen transfer integrals cannot be neglected (being about half the Cu-O values), and that direct Cu-O exchange is surprisingly large (affecting the calculated values of $J$ and $K_{\text{eff}}$).

## III. Cu VALENCE FLUCTUATIONS AND THE SPIN AND CARRIER SUBSYSTEMS

We now address whether the Heisenberg spin-$\frac{1}{2}$ system exists independently of the presence of charge carriers or whether local AF correlation is purely a consequence of a half-filled band. We also examine in this section the question of which sublattice is preferred by the carriers.

The Hamiltonian suggested on the basis of the above parameters is a two-band (for the Cu and O sublattices) extended (for the interatomic $U$) Hubbard model:

$$H_0=\sum_{i,\sigma}\varepsilon_i n_{i\sigma}+\sum_i U_i n_{i\uparrow}n_{i\downarrow}+\tfrac{1}{2}\sum_{i,\sigma_1,j,\sigma_2}{}' [U_{ij}n_{i\sigma_1}n_{j\sigma_2}+t_{ij}(\psi^\dagger_{i\sigma_1}\psi_{j\sigma_1}+\psi^\dagger_{i\sigma_2}\psi_{j\sigma_2})+K_{ij}\psi^\dagger_{j\sigma_2}\psi_{i\sigma_2}\psi^\dagger_{i\sigma_1}\psi_{j\sigma_1}] , \tag{3.1}$$

where the $n_{i\sigma}$ ($\sigma=\uparrow,\downarrow$) are occupation operators and the $\psi^\dagger_{i\sigma}$ ($\psi_{i\sigma}$) create (destroy) a hole of spin $\sigma$ on site $i$ in the otherwise filled Cu($x^2-y^2$) and O($x$ or $y$) level. The prime on the summation indicates $i\neq j$. The oxygen $2p$ orbitals are those which point towards the nearest-neighbor Cu ions (see Sec. VI and Appendix A). The sums go over all sites, where $t_{ij}=t_{pp}$ between adjacent oxygen sites in the CuO$_2$ plane, $t_{pd}$ between nearest-neighbor Cu and O sites, and zero otherwise. The explanation of the $U_{ij}$ is less obvious as the variation of $U_{ij}$ with $R_{ij}$ is uncertain. This is discussed further below. The direct ($K_{ij}=K$ if $i$ and $j$ are nearest neighbors and zero otherwise) and indirect exchange interactions will not concern us for the time being.

Numerical simulations would be useful to find the ground state of this Hamiltonian, especially in the case of one hole per unit cell, and find the variation of the AF correlation length with the number of charge carriers. Unfortunately, these calculations have not yet been done. Meanwhile, the following arguments and simple calculations suggest that the Heisenberg spin system forms.

In principle, the ground state of a finite system with one hole per unit cell could be found by mixing configurations which represent all possible placements of $N$ holes on $N$ Cu and $2N$ oxygen sites. Consider the first step in such a calculation, the evaluation of the energy (and wave function) of a single hole in an array of fixed $N-1$ holes. Anticipating that the lowest-energy configurations will have the holes uniformly distributed over the $N$ sites, we compute the site energies in the presence of a fixed $N-1$ array with these holes on all Cu sites but one. Taking the energy of the central Cu site (which is the one without a fixed hole) to be zero, the four neighboring and equivalent oxygen sites would have an energy of $\Delta\varepsilon+\Delta U$ where the latter represents the change of the total Coulomb hole-hole interaction energy in moving the hole from the central Cu site to the oxygen site. This energy is unknown as we do not know the variation of $U$ with distance, i.e., between the single-site values and the dielectric limit. If we assume that there is rapid convergence to the dielectric limit with distance, then short-range contributions to $\Delta U$ dominate and we obtain $\Delta U\sim U_{pd}$. The corrections to this value for longer-range Coulomb interactions reduces $\Delta U$, with a reasonable lower bound being approximately $U_{pd}/2$. For the hole to move further away from the central site costs more Coulomb energy: onto the next Cu atom costs between $U_d$ and $U_d-U_{pd}$ (again depending on the variation of $U_{ij}$ with distance) as the site is already occupied by a hole; however, even to move onto one of the eight equivalent next oxygen sites costs up to an additional $U_{pd}$, or $\Delta\varepsilon+2U_{pd}$ for the site energy of those oxygens. Further motion away from the central site occurs with no or only small additional Coulomb energy.

Thus the hole sits within a Coulomb potential well. If we take $\Delta\varepsilon\sim0.5-1.5$ eV and $U_{pd}\sim1-3$ eV, reasonable estimates are as follows: It costs 2–4 eV for the hole to move onto one of the adjacent oxygens and an *additional* 1–3 eV to move any further. Solving for the 2D hole wave function with these site energies and including the transfer integrals as in Eq. (3.1), we find that the hole is localized with about $75\pm10\%$ of its density on the central Cu and the remainder almost entirely on the adjacent four neighboring oxygens. Only a few percent of its wave function penetrates further from this central region.

These results are suspect because they represent a single-particle solution. For this reason, we also performed a configuration interaction calculation in which we used the Hamiltonian of Eq. (3.1) and allowed two holes to move in a correlated manner while embedded in an array of holes fixed beyond five central Cu sites. The results were almost identical to the single-particle result. Thus there was no indication that the inclusion of correlation would alter the basic conclusion of hole localization, although more rigorous calculations could address this point.

Because these calculations are simple and can be easily reproduced, we do not go into further details concerning them but simply report the observation that the Coulomb potential well causes hole localization such that each hole spends perhaps $75\pm10\%$ of its time on a Cu site and the remainder on essentially only those four oxygen nearest neighbors which surround the Cu site (thus producing a Heisenberg spin-$\frac{1}{2}$ system; see Sec. IV). If interatomic $U$ were not included, the Coulomb potential well would be "leaky" in that there would be no energy cost for moving beyond the nearest-neighbor oxygen sites, and the hole would be free to delocalize through the oxygen sublattice. In this case, unless $\Delta\varepsilon$ were larger than about half the oxygen bandwidth (in marked disagreement with the band-structure results[30,36]) the Cu sites would exist as resonances within the oxygen band and localization would not occur, thus taking the model towards the band-structure picture. A substantial interatomic $U$ thus causes the breakdown of the band-structure model and the establishment of a spin system whose existence occurs even in the presence of carriers.

One may wonder if a simpler two-band Hamiltonian, e.g., without $U_{pd}$, could achieve the same result. This of course is true; however, if we were to omit the interatomic $U$ and establish the spin system through an increase in the magnitude of $\Delta\varepsilon$, this would prevent the realistic treatment of the Cu $d^{10}$ valence fluctuation which is enhanced by carrier presence (see Sec. VI) and which is at the heart of the carrier-spin interaction. Thus we shall use the Hamiltonian of Eq. (3.1) in this paper.

If only the above qualitative arguments were available, one might hesitate to proceed without further calculations which attempt to determine the ground state of the Hamiltonian in Eq. (3.1) with one hole per unit cell, at least in a finite system. However, several experimental pieces of evidence support the existence of a spin system independent of the presence of carriers. Without carriers, we have of course the observations of antiferromagnetism in $La_2CuO_4$,[37] and more recently in $Y_1Ba_2Cu_3O_6$.[38] However, recent neutron-scattering measurements by Birgeneau *et al.*[39] show that the $Cu^{2+}$ moment is independent of carrier density in the lanthanum compound with up to 18% Sr doping. Meanwhile, polarized light scattering experiments by Lyons *et al.* on both the yttrium[40] and lanthanum[41] materials observe high energy losses with a peak at 3000 $cm^{-1}$ in the lanthanum compound and 2600 $cm^{-1}$ in the yttrium compound. In the latter case, the sharpest peak was in samples with an oxygen content of 6.0 where the system is antiferromagnetic, but a broad band of losses persisted at oxygen concentrations of 6.6 and even 7.0. This suggests that if the loss is associated with two magnon processes (which analogous studies of the well-understood antiferromagnet $K_2NiF_4$ suggest), it does not disappear but only broadens as first long-range order is lost (at 6.6) and further with an increasing number of carriers present (at 7.0). In addition, there is much spectroscopic evidence[42] which indicates that the Cu valency does not appear to change in the La material with doping, or change in the yttrium material when oxygen is increased beyond 6.5. In view of the above arguments, we assume the existence of the spin-$\frac{1}{2}$ system for what follows.

Charge carriers are introduced as the number of holes increases beyond one per unit cell. The energies derived in Sec. II indicate that the charge carriers have amplitudes which are largest on the oxygen sublattice because one localized hole per unit cell is already on the Cu sites: The large value of the Cu $U_d$ suppresses the $d^8$ configuration, since to place an additional hole on a Cu site costs 9 eV, while to place it on an oxygen site within the plane only costs $\Delta\varepsilon+2U_{pd}\sim3$–7 eV. Delocalization further favors the oxygen sublattice as the direct oxygen-oxygen transfer integrals lower the oxygen energy from the above single-site value by $\sim4t_{pp}\sim2.4$ eV, so that the comparison is actually $\sim9$ eV cost for the Cu sites versus only $\sim0.5$–4.5 eV for the oxygen sites. These carrier holes then interact with the Cu sublattice via $t_{pd}$ and direct exchange $K$.

Experimental information which supports this conclusion includes the observation of oxygen carrier holes by resonance photoemission[42] and the absence of Cu($d^8$) features even with a substantial number of carriers.[42,43] However, the physics of what follows does not change substantially as the carrier amplitude is changed between the Cu and oxygen sites within reasonable limits.

## IV. EFFECTIVE EXCHANGE ENERGIES

### A. Superexchange

The superexchange energy $J$ between Cu sites may now be discussed in light of the above energies. $J$ is defined in this work such that it costs $J$ per nearest neighbor to flip a Cu spin in an otherwise antiferromagnetic array. This energy is not an independent variable as it may be computed from the values of $t_{pd}$, $t_{pp}$, $U_{pd}$, $U_p$, $U_d$, $\Delta\varepsilon$, and $K$, the direct Cu-O exchange. We perform this calculation in an attempt to see whether the values of these parameters as determined and defined in Sec. II lead to a consistent picture of the energies, since $J$ has indirectly been determined from three experimental measurements: the spin-wave dispersion,[15] the variation of the AF correlation length with temperature,[44] and the two-magnon-loss (spin exchange) spectrum observed in polarized light scattering.[40,41,45] The spin-wave dispersion implies $J\sim37$–70 meV; the measurement of the correlation length implies that $J\sim25$–40 meV and the Raman scattering implies $J\sim67$ meV (Ref. 40) or $\sim32$ meV (Ref. 45) in $La_2CuO_4$. (A 12% smaller value than in Ref. 40 has been found for the yttrium "1:2:3" compound from two-magnon loss.[41] None of these measurements are definitive for the following reasons: (1) the spin-wave dispersion has only been determined within a factor of 2; (2) the correlation length at best determines an unknown constant times $J$; although the constant has been derived by several independent researchers, all of which more or less agree[46–48] (however, simulations on small systems may bring these theories into question[49]); finally, (3) the determination from the two-magnon-loss spectrum generally assumes that the peak is approximately $5.4J$. These estimates are based on the earlier work on $K_2NiF_4$.[50] However, similar measurements by different groups do not agree with each other. Whatever the case, there remains some uncertainty as to the precise value for $J$, although $J\sim25$–65 meV for $La_2CuO_4$ seems definitive.

We have attempted to compute $J$ without using perturbation theory. Let us first consider three sites, Cu-O-Cu, and two holes. We assume that only the $3d(x^2-y^2)$ and the $2p(x)$ orbitals are involved on these Cu and O sites, respectively. It is then simple to write the nine possible two-hole spin-orbital bases [e.g., $(d_{1\alpha}p_\beta-p_\beta d_{1\alpha})/\sqrt{2}$, $(d_{1\alpha}d_{2\beta}-d_{2\beta}d_{1\alpha})/\sqrt{2}$, etc., where the placement of the one-hole spin-orbital basis in the expression indicates coordinate $r_1$ or $r_2$], determine the Hamiltonian matrix for these bases using Eq. (3.1), diagonalize it, and find the energy splitting between the lowest two energy states, which represent singlet and triplet coupled Cu spins. The Hamiltonian can be block diagonalized using eigenstates of spin and symmetry. The triplet states block into $H_t^{(a)}(2\times2)$ and $H_t^{(s)}(1\times1)$, while the singlet states block into $H_s^{(a)}(4\times4)$ and $H_s^{(s)}(2\times2)$ where the superscripts $(a)$ and $(s)$ refer to antisymmetric and symmetric with respect to reflection about the O atom and the subscripts $t$ and $s$ refer to triplet and singlet. The two states of interest are the lowest singlet which is the minimum eigenvalue of $H_s^{(a)}$ and the lowest triplet which is the minimum eigenvalue of $H_t^{(a)}$. We expect this energy difference to be approximately $2J$. These two matrices have the following forms:

$$H_s^{(a)} = \begin{pmatrix} 0 & -\sqrt{2}t_{pd} & 0 & 0 \\ -\sqrt{2}t_{pd} & \Delta E - K & -2t_{pd} & -t_{pd} \\ 0 & -2t_{pd} & \Delta U_p & 0 \\ 0 & -t_{pd} & 0 & \Delta U_d \end{pmatrix}$$

and

$$H_t^{(a)} = \begin{pmatrix} 0 & -\sqrt{2}t_{pd} \\ -\sqrt{2}t_{pd} & \Delta E + K \end{pmatrix},$$

where $\Delta E$ is the energy cost to move a Cu hole onto an oxygen site. As discussed above, $\Delta\varepsilon + U_{pd}/2 < \Delta E < \Delta\varepsilon + U_{pd}$. $\Delta U_p$ is the energy cost to move two nearest-neighbor Cu holes onto the intervening oxygen site. This has an upper bound of $2\Delta\varepsilon + U_p$ and a lower bound of $2\Delta\varepsilon + U_p - U_{pd}$. Similarly, $\Delta U_d$ is the energy cost to move one Cu hole onto the site of a nearest-neighbor Cu hole. The upper bound for this energy is $U_d$ while the lower bound is $U_d - U_{pd}$. Recall that this uncertainty is related to the lack of knowledge of the Coulomb interaction of two holes when further apart than nearest neighbors, relative to that of two holes on nearest neighbors. Clearly, we can expect that $U_{pd} > e^2/\epsilon_\infty r_{\mathrm{CuO}}$ and $e^2/2\epsilon_\infty r_{\mathrm{CuO}} < U_{dd} < U_{pd}/2$, where $U_{dd}$ is the Coulomb repulsion of two holes on nearest-neighbor Cu's. The differences between these matrices and those from Shen *et al.*[51] are first in the diagonal energies of the intermediate states, due primarily to $U_{pd}$ and second in the matrix element from the charge-transfer fluctuation to a doubly occupied *d* orbital, i.e., the 2,4 element of $H_s^{(a)}$.

Using this three-center model, we find the energy splitting to depend strongly on the chosen values for the Cu-O transfer integral $t_{pd}$ and the Cu-O direct exchange integral $K$ (recall that $K < 0$; i.e., ferromagnetic exchange), and less strongly on $\Delta E$. With values of $t_{pd} = 1.1$ eV, $K = -0.18$ eV, $\Delta\varepsilon = 1.0$ eV, $U_{pd} = 2$ eV, $U_p = 6$ eV, and $U_d = 9$ eV ($t_{pp}$ does not enter into this calculation), the upper bound for the $\Delta$ energies gives an energy splitting ($2J$) equal to 70 meV, while the lower bound gives 160 meV. A value of 1.45 eV is determined using the energy parameters and the matrices in Ref. 51, where it is recognized that this value is larger than is realistic, reasons for which are discussed there.

We, however, have no real reason to believe that the three-center model should be adequate or that it tests the adequacy of the Heisenberg Hamiltonian as a representation of the low-lying states. Thus we also consider a cluster which has five holes and nine centers, one central Cu, its four neighboring O ions, and the four Cu ions which are also nearest neighbors to these O ions:

Cu
O
Cu O Cu O Cu .
O
Cu

In this case, we spin and symmetry decompose the Hamiltonian and we try to equate the lowest $|\mathbf{S}| = \frac{3}{2}$ state to $E_0 - 3J$, the next triply degenerate $|\mathbf{S}| = \frac{3}{2}$ states to $E_0 + J$, and the lowest $|\mathbf{S}| = \frac{5}{2}$ state to $E_0 + 2J$. This spectrum arises if the low-lying states can be modeled by the Heisenberg Hamiltonian. We take the space part of the wave function to belong to the $B_1$ representation of $C_{4v}$. If we label the central Cu to be the first Cu in the product wave functions, then the lowest $(|\mathbf{S}|, S_z) = (\frac{3}{2}, \frac{3}{2})$ state is

$$4\beta\alpha\alpha\alpha\alpha - \alpha\beta\alpha\alpha\alpha - \alpha\alpha\beta\alpha\alpha - \alpha\alpha\alpha\beta\alpha - \alpha\alpha\alpha\alpha\beta$$

with all holes on different Cu's (no holes on the oxygens). The $(|\mathbf{S}|, S_z) = (\frac{5}{2}, \frac{5}{2})$ state is, of course, $\alpha\alpha\alpha\alpha\alpha$ and the energy depends on $|\mathbf{S}|$ not on $S_z$. The three degenerate $(|\mathbf{S}|, S_z) = (\frac{3}{2}, \frac{3}{2})$ states are

$$\alpha\beta\alpha\alpha\alpha - \alpha\alpha\beta\alpha\alpha + \alpha\alpha\alpha\beta\alpha - \alpha\alpha\alpha\alpha\beta,$$

$$\alpha\beta\alpha\alpha\alpha - \alpha\alpha\beta\alpha\alpha - \alpha\alpha\alpha\beta\alpha + \alpha\alpha\alpha\alpha\beta,$$

and

$$\alpha\beta\alpha\alpha\alpha + \alpha\alpha\beta\alpha\alpha - \alpha\alpha\alpha\beta\alpha - \alpha\alpha\alpha\alpha\beta.$$

The splitting (ideally $4J$) between the $|\mathbf{S}| = \frac{3}{2}$ states is 120–240 meV ($J = 30$–$60$ meV) and the splitting (ideally $J$) between the degenerate $|\mathbf{S}| = \frac{3}{2}$ manifold and the $|\mathbf{S}| = \frac{5}{2}$ state is 32–64 meV (using $t_{pp} = 0.5$ eV, although the computed value for $J$ is quite insensitive to the input value for $t_{pp}$). This value for $J$ is in accord with the experimental determination which assumes a Heisenberg Hamiltonian.

It is encouraging that we do indeed find a splitting of the different spin states in agreement with the spectrum of the Heisenberg Hamiltonian for five spins. This increases our confidence not only in our parameters but also in the Heisenberg description for the spin system.

Note especially that the high sensitivity of $J$ to $t_{pd}$ lends support to the $\sim 1.1$ eV value of the latter. In turn, this supports a value of $t_{pp} \sim 0.5$ eV as both of these transfer integrals are then necessary in order to produce a tight-binding fit to the band-structure calculations. Thus our energy parameters do indeed lead to a self-consistent picture.

### B. Effective Cu-O exchange

The effective exchange energy between a hole on an oxygen site and a hole on a neighboring Cu site is also a dependent energy. This energy is the effective static spin coupling between holes on the Cu sites with holes on the oxygen sites again mediated by virtuals. There is experimental evidence that spin interaction energies between carrier holes and localized Cu spins may be on the order of 1 eV.[39] We compute the effective exchange energy from the three-site model [Cu(1)-O-Cu(2)] with three holes ($S_z^{\mathrm{tot}} = \frac{1}{2}$) and fit the lowest three states to the following effective spin Hamiltonian:

$$H_{\mathrm{eff}} = E_0 + 2J(1-\lambda)S_{\mathrm{Cu}(1)} \cdot S_{\mathrm{Cu}(2)} + 2K_{\mathrm{eff}} S_{\mathrm{O}} \cdot (S_{\mathrm{Cu}(1)} + S_{\mathrm{Cu}(2)}),$$

where $S_{\mathrm{Cu}(1)}$, $S_{\mathrm{Cu}(2)}$, and $S_{\mathrm{O}}$ are the spin operators for the

Cu(1), Cu(2), and O sites, respectively. The parameter $\lambda$ reflects the reduction in superexchange between Cu sites when a hole is present on the oxygen site. We expect $\lambda$ to be close to 1, which would mean that the spin coupling between Cu's is completely blocked when the oxygen site is not closed shell.

The lowest eigenvalue of $H_{\text{eff}}$ corresponds to $\alpha\alpha\beta - 2\alpha\beta\alpha + \beta\alpha\alpha$ with eigenvalue $E_0 - 2K_{\text{eff}} + \frac{1}{2}(1-\lambda)J$. The next eigenvalue corresponds to $\alpha\alpha\beta - \beta\alpha\alpha$ and has eigenvalue $E_0 - \frac{3}{2}(1-\lambda)J$. Both of these states correspond to $|\mathbf{S}^{\text{tot}}| = \frac{1}{2}$. The third eigenvalue corresponds to $|\mathbf{S}^{\text{tot}}| = \frac{3}{2}$, which for $S_z^{\text{tot}} = \frac{1}{2}$ is $\alpha\alpha\beta + \alpha\beta\alpha + \beta\alpha\alpha$ with eigenvalue $E_0 + K_{\text{eff}} + (1-\lambda)J/2$. The result of this computation with the energy parameters as used for computing $J$ suggests $\lambda = 0.945$ (thus the superexchange is indeed almost completely blocked by having an oxygen hole present) and $K_{\text{eff}} = 0.1$ eV.

There is no contradiction between the small value of $K_{\text{eff}}$ and the above-mentioned[39] large carrier-spin interaction energy. For example, when we computed the singlet-triplet splitting for two holes in five sites (a Cu site and the four surrounding O sites) we found the splitting to be 2 eV. Thus there is an effective energy of 1 eV which strongly couples the Cu spin to the spin of a carrier delocalized over four oxygen sites and which is due only in small part to $K_{\text{eff}}$. The explanation of this involves the Cu($d^{10}$) valence fluctuations as will be explained in Sec. VI and Appendix A.

## V. SPIN-$\frac{1}{2}$ HEISENBERG-NEAREST-NEIGHBOR AF COUPLING

### A. The AF spin-$\frac{1}{2}$ Heisenberg Hamiltonian

In order to describe magnetic deviations in the presence of carrier holes, we first describe them in the absence of carrier holes. This introduces some notation necessary for deriving the Hamiltonian for carrier holes coexisting and strongly interacting with the Cu spin system. We do not assume that the reader is familiar with spin-wave theory. Moreover, we have not located a reference which is specific to spin $\frac{1}{2}$. Standard spin-wave theory for arbitrary $|S|$ makes a Holstein-Primakoff transformation[52] and then expands in the operator $(2|S|)^{-1}$. This expansion is not obviously applicable to $|S| = \frac{1}{2}$.

In what follows we take the Hamiltonian in the absence of carrier holes to be that of a nearest-neighbor Heisenberg system with $|S|$ specifically equal to $\frac{1}{2}$. The applicability of the Heisenberg Hamiltonian is supported by the superexchange calculation in the preceding section as well as by the observation of antiferromagnetism[18,37,38] and the spin-wave dispersion in neutron-scattering experiments.[15] We find that we do not need to make any operator expansions when $|S|$ is $\frac{1}{2}$. In addition, since the spin Hamiltonian is a crucial part of determining the pairing Hamiltonian (see Sec. VII) this warrants complete detail.

For the present analysis we avoid unnecessary complexity and assume a two-dimensional square array, hence

$$H = J \sum_{l,\mathbf{h}} S_l \cdot S_{l+\mathbf{h}} ,$$

where $J > 0$ and $l$ labels lattice sites and $\mathbf{h}$ labels the nearest-neighbor Cu site, thus from a Cu site we add $\mathbf{h} = \mathbf{a}_x$, $-\mathbf{a}_x$, $\mathbf{a}_y$, or $-\mathbf{a}_y$. It is our aim to rewrite this Hamiltonian in terms of creation and annihilation spin deviation operators which operate on a reference state.

We take our reference state to be the classical ground state; i.e., a perfectly ordered antiferromagnetic (AF) array of spin $\frac{1}{2}$. Furthermore, we define the unit cell to have two spins, one up and one down in the reference state, where the sublattice $A$ has all spins down and sublattice $B$ all spins up. The new lattice vectors are $\mathbf{a}_1 = \mathbf{a}_x + \mathbf{a}_y$ and $\mathbf{a}_2 = \mathbf{a}_x - \mathbf{a}_y$, and the basis vectors for the two spins in sublattices $A$ and $B$ are $\frac{1}{2}\mathbf{a}_x$ and $-\frac{1}{2}\mathbf{a}_x$, respectively. Now all nearest neighbors to spins on sublattice $A$ are on sublattice $B$ and vice versa. The Hamiltonian in this representation is

$$H = 2J \sum_{l,\mathbf{h}} S_{l,A} \cdot S_{l+\mathbf{h},B} , \tag{5.1}$$

where the extra factor of 2 arises because we now sum over half as many unit cells. Next we define spin deviations from the reference state; thus a spin deviation in sublattice $A$ or $B$ at lattice point $l$ is created by the operator $S_{l,A}^{+} = [S_x]_{l,A} + i[S_y]_{l,A}$ or $S_{l,B}^{-} = [S_x]_{l,B} - i[S_y]_{l,B}$, respectively. Therefore we equate $d_{l,A}^{\dagger} = S_{l,A}^{+}$ and $d_{l,B}^{\dagger} = S_{l,B}^{-}$ and note that $[S_z]_{l,A} = -\frac{1}{2} + d_{l,A}^{\dagger} d_{l,A}$ while $[S_z]_{l,B} = \frac{1}{2} - d_{l,B}^{\dagger} d_{l,B}$. Furthermore, the $d_{l\sigma}$ and $d_{l\sigma}^{\dagger}$ operators satisfy mixed fermion-boson commutation relations. We call this particle, as it has been referred to previously,[53] a hard-core boson. Due to spin $\frac{1}{2}$, it is a fermion with respect to occupation number, that is $[d_{l\sigma}, d_{l\sigma}^{\dagger}]_{+} = 1$. However, as with all spin-raising and -lowering operators, it is a boson with respect to exchange of two such particles, that is

$$[d_{l\sigma}, d_{j\tau}^{\dagger}]_{-} = [d_{l\sigma}, d_{j\tau}]_{-} = [d_{l\sigma}^{\dagger}, d_{j\tau}^{\dagger}]_{-} = 0 ,$$

providing that $j\tau \neq l\sigma$. Substituting these new operators into the Hamiltonian, we have

$$\begin{aligned} H = {} & -6NJ + 4J \sum_{l} (d_{lA} d_{lA}^{\dagger} + d_{lB}^{\dagger} d_{lB}) \\ & + J \sum_{l,\mathbf{h}} (d_{l,A}^{\dagger} d_{l+\mathbf{h},B}^{\dagger} + d_{l+\mathbf{h},B} d_{l,A}) \\ & + 2J \sum_{l,\mathbf{h}} d_{l,A}^{\dagger} d_{l,A} d_{l+\mathbf{h},B} d_{l+\mathbf{h},B}^{\dagger} . \end{aligned} \tag{5.2}$$

Some discussion of the various terms in this Hamiltonian is useful at this stage because the origin of the terms does get slightly obscured after the Fourier transforms defined below. The first term is an irrelevant constant energy (we have taken $N$ sites per sublattice). The second term is $4J$ times the number operator for spin deviations on $B$ and the absence of spin deviations on $A$; i.e., each isolated spin deviation costs $4J$ of energy. The third term is the term which allows spin deviations to propagate through the lattice and also allows the creation of pairs of spin deviations from the vacuum and the destruction of pairs of neighboring deviations. The last term is a short-range repulsion between a spin deviation on $A$ and the lack of a spin deviation on $B$. This arises because, should a spin deviation on sublattice $A$ be nearest neighbor to a spin deviation on sublattice $B$ there is a cost of $6J$ rather than the $8J$ required to create two isolated spin deviations, thus the factor of $2J$ per nearest-neighbor pair.

We now define the Fourier transforms $b_{k,A}$ and $b_{k,B}$ (where from here forward we drop the vector notation) such that

$$b_{k,A}=e^{i\mathbf{k}\cdot\mathbf{a}_x/2}\frac{1}{\sqrt{N}}\sum_{l}e^{i\mathbf{k}\cdot l}d_{l,A}\,, \tag{5.3a}$$

$$b_{k,B}=e^{-i\mathbf{k}\cdot\mathbf{a}_x/2}\frac{1}{\sqrt{N}}\sum_{l}e^{i\mathbf{k}\cdot l}d_{l,B}\,. \tag{5.3b}$$

Once again these operators are bosons in that $b_{k_1\sigma}$ commutes with $b^{\dagger}_{k_2\tau}$ $(k_2\tau\neq k_1\sigma)$ and with $b_{k_2\tau}$. In addition, they must satisfy the following constraint:

$$\sum_{k}b_{k\sigma}b^{\dagger}_{k\sigma}+b^{\dagger}_{k\sigma}b_{k\sigma}=N\,. \tag{5.4}$$

The Hamiltonian is rewritten (ignoring the irrelevant constant energy) in terms of the $b_{k,A}$ and $b_{k,B}$ as

$$\begin{aligned}H={}&4J\sum_{G,k}(e^{iG\cdot\mathbf{a}_x/2}b_{k,A}b^{\dagger}_{k-G,A}+e^{-iG\cdot\mathbf{a}_x/2}b^{\dagger}_{k,B}b_{k-G,B})\\&+J\sum_{G,k}V_{G-k}(e^{iG\cdot\mathbf{a}_x}b^{\dagger}_{k,A}b^{\dagger}_{G-k,B}+e^{-iG\cdot\mathbf{a}_x}b_{k,A}b_{G-k,B})\\&+2J\frac{1}{N}\sum_{G,q,k_1,k_2}V_{G-q}b^{\dagger}_{k_1A}b_{(k_1-q)A}b_{-(k_2+G-q)B}b^{\dagger}_{-k_2B}\,,\end{aligned}$$

where $G$ is a reciprocal-lattice vector and in the present square lattice case

$$V_k=\sum_{h}e^{i\mathbf{k}\cdot\mathbf{h}}=2[\cos(\mathbf{k}\cdot\mathbf{a}_x)+\cos(\mathbf{k}\cdot\mathbf{a}_y)]\,. \tag{5.5}$$

We note that $V_k$ has extended $s$-wave symmetry. For simplicity we ignore umklapp processes $(G\neq 0)$ and arrive at a Hamiltonian which is reminiscent of the BCS Hamiltonian, except that the quasiparticles are more like bosons than fermions:

$$\begin{aligned}H={}&2NJ\kappa^2+4J\sum_{k}(b_{kA}b^{\dagger}_{kA}+b^{\dagger}_{kB}b_{kB}\\&+\eta J\sum_{k}V_k(b^{\dagger}_{k,A}b^{\dagger}_{-k,B}+b_{k,A}b_{-k,B})\,,\end{aligned} \tag{5.6}$$

where we have bilinearized the Hamiltonian by defining the mean-field variable $\eta=1+\kappa$ with

$$\kappa V_k\approx 2\sum_{q}V_q\langle b_{(k+q)A}b_{-(k+q)B}\rangle\,,$$

using the familiar identity that

$$\begin{aligned}\hat{A}\hat{B}&=[(\hat{A}-\langle A\rangle)+\langle A\rangle][(\hat{B}-\langle B\rangle)+\langle B\rangle]\\&=\langle A\rangle\hat{B}+\langle B\rangle\hat{A}-\langle A\rangle\langle B\rangle+(A-\langle A\rangle)(B-\langle B\rangle)\end{aligned}$$

and ignoring terms which are square in the fluctuations.

This result in itself is not profound but it does preserve the symmetry of the Hamiltonian which is invariant when $b^{\dagger}_{k\sigma}b_{k\sigma}$ goes to $b_{k\sigma}b^{\dagger}_{k\sigma}$.

Let us now use $N_{s\sigma}$ to equal the number operator for spin deviations on sublattice $\sigma$; i.e.,

$$\hat{N}_{sA}=\sum_{k}b_{kA}b^{\dagger}_{kA},\ \ \hat{N}_{sB}=\sum_{k}b^{\dagger}_{kB}b_{kB}\,.$$

The ground-state energy is minimized with the constraint that $\langle\hat{N}_{s\sigma}\rangle=\frac{1}{2}N$ (Refs. 48 and 54) (recall that $N$ is the number of sites per each sublattice) which is a requirement of no net magnetism or no long-range order as this is the case for a 2D Heisenberg system.[6] In addition, this constraint automatically satisfies the required constraint condition expressed in Eq. (5.4). Equation (5.6) is also formally similar to the magnon Hamiltonian[55] derived for large spin where the quartic term is typically dropped $(\eta=1)$ and the operators approximately satisfy true *boson* commutation relations. The magnon is then a linear combination of a creation operator of wave vector $k$ on sublattice $A$ and an annihilation operator of wave vector $-k$ on sublattice $B$ or the reverse. Here we proceed to determine the spin-wave excitation spectrum by including the quartic term in mean field (as described above). Although these bosons do not satisfy boson commutation relations in that $[b_{k\sigma},b^{\dagger}_{k\sigma}]\neq 1$, the commutator does at least equal an operator which is independent of $k$; i.e.,

$$[b_{k\sigma},b^{\dagger}_{k\sigma}]_-=\frac{1}{N}\sum_{l}[d_{l\sigma},d^{\dagger}_{l\sigma}]_-=\hat{f}_{\sigma}\,. \tag{5.7}$$

The condition that there be no long-range order implies that $\langle\hat{f}_{\sigma}\rangle=0$.

### B. Spin-wave excitation spectrum

In order to find the excitation spectrum for the spin-$\frac{1}{2}$ system with antiferromagnetic coupling, we diagonalize the Hamiltonian in Eq. (5.6) just as done in spin-wave theory (including the quartic term in mean field). The quasiparticle excitation spectrum is obtained by diagonalizing $\underline{L}_{k\sigma}$ (a matrix in $b_{k\sigma}$, $b^{\dagger}_{-k\tau}$ space) where $\tau=B$ $(A)$ when $\sigma=A$ $(B)$ and

$$L_{k\sigma}=\begin{pmatrix}(4J+g\mu_BH_A+\mu_0) & +J\eta V_k\\ -J\eta V_k & -(4J+g\mu_BH_A+\mu_0)\end{pmatrix}, \tag{5.8}$$

where we have now included (after the fact) the anisotropy field[55] $H_A$, the gyromagnetic ratio $g$, and the Bohr magneton $\mu_B$. We have also included a chemical potential $\mu_0$ to impose the constraint of no net magnetism[48,54] and the mean-field (MF) pair exchange enhancement $\eta=1+\kappa$. The normalization constraint for the new quasiparticles $(\gamma_{k\sigma})$ is that they satisfy the same commutation relations (in the mean) as the $b_{k\sigma}$ operators; i.e., $\langle[\gamma_{k\sigma},\gamma^{\dagger}_{k\sigma}]_-\rangle=\langle\hat{f}_{k\sigma}\rangle=0$. Thus the elementary spin excitations are $\gamma_{k\sigma}=b_{k\sigma}\cosh(\theta_k)+b^{\dagger}_{-k\tau}\sinh(\theta_k)$ and the excitation spectrum is

$$H_{\mathrm{MF}}+\mu_0\hat{N}_s=2NJ\kappa^2+\sum_{k}\varepsilon_k(\gamma^{\dagger}_{k\sigma}\gamma_{k\sigma}+\gamma_{-k\tau}\gamma^{\dagger}_{-k\tau})\,, \tag{5.9}$$

and

$$\varepsilon_k^2=(g\mu_BH_A+\mu_0+4J)^2-(J\eta V_k)^2\,, \tag{5.10a}$$

$$\cosh(2\theta_k)=\frac{g\mu_BH_A+\mu_0+4J}{\varepsilon_k}\,, \tag{5.10b}$$

$$\sinh(2\theta_k)=\frac{JV_k\eta}{\varepsilon_k}\,. \tag{5.10c}$$

This result is not markedly different from standard (high) spin-wave theory, albeit with a renormalized

$\tilde{J}=\eta J$ and a renormalized anisotropy field, $\mu=g\mu_B H_A+\mu_0-4J\kappa$. Note that $\mu$ is necessarily greater than zero since with these renormalized values

$$\varepsilon_k^2=(4\tilde{J}+\mu)^2-(\tilde{J}V_k)^2 . \tag{5.11}$$

As in standard spin-wave theory, the behavior of $\varepsilon_k$ near $k=0$ is

$$\varepsilon_{k\to 0}^2\approx 8\tilde{J}(\mu+\tilde{J}\,|\mathbf{k}\cdot\mathbf{a}_1|^2) ,$$

where we have assumed that $\mu\ll 4\tilde{J}$. For $|\mathbf{k}\cdot\mathbf{a}_1|^2\gg\mu/\tilde{J}$ the dispersion along the $\mathbf{a}_x$ direction is $2\tilde{J}a_1=\sqrt{32}\tilde{J}r_{\mathrm{CuO}}$. For a quoted dispersion of 0.4–0.75 eV Å (Ref. 15) and $r_{\mathrm{CuO}}=1.89$ Å this gives $\tilde{J}=37$–$70$ meV.

## VI. THE SPIN-HYBRID CARRIER QUASIPARTICLES

Based on the above results, we now write a Hamiltonian which contains the interactions between a carrier and the spin subsystems. This section consists of three parts: a formal definition of the spin-hybrid carrier Hamiltonian is followed by a discussion of how the basic physical description of the carrier can be systematically constructed. Finally, a simple band-structure model of the carrier allows us to extract the essential physical description of the quasiparticle.

### A. Formal definition of spin-hybrid Hamiltonian

There are three parts to the Hamiltonian: (1) a "static" or pure spin part which does not involve transport, (2) a transport part which does not involve the creation of spin deviations, and (3) the terms which involve both transport and the creation of spin deviations. The spin part of the Hamiltonian is the sum of two terms, the first of which reduces to the Heisenberg Hamiltonian in the absence of carriers (see Sec. V). The second term explicitly contains the exchange interactions between carrier and Cu spins. Thus

$$H^{(s)}=H_J^{(s)}+H_K^{(s)} , \tag{6.1}$$

where the superscript $(s)$ denotes spin or static (as opposed to transport).

$H_J^{(s)}$ is a modified Heisenberg Hamiltonian[56] to allow blocking of superexchange due to the presence of carriers (recall that in the preceding section we found the blocking was 94% complete). The modified Heisenberg Hamiltonian is

$$H_J^{(s)}=2J\sum_{l,h}(1-n_{lh\uparrow})(1-n_{lh\downarrow})S_{lA}\cdot S_{l+hB} , \tag{6.2a}$$

where $n_{lh\uparrow(\downarrow)}=\psi_{lh\uparrow(\downarrow)}^\dagger\psi_{lh\uparrow(\downarrow)}$ is the number operator for $\uparrow$ $(\downarrow)$ holes on the oxygen site $lh$ which sits between a Cu at $lA$ and one at $l+hB$. If the carrier is present, whether it has spin up $(n_{lh\uparrow}=1)$ or spin down $(n_{lh\downarrow}=1)$ the superexchange coupling between the neighboring Cu's essentially vanishes. In what follows we will assume that we can treat the carrier term $(1-n_{lh\uparrow})(l-n_{lh\downarrow})$ as blocking the superexchange in an average sense; i.e., we will replace $J$ by $J_\rho=(1-\rho/4)^2J$ or $J_\rho=(1-\rho/2)J$ where in the latter we assume no double occupancy for the oxygen holes. ($\rho$ is defined as the number of carrier holes per Cu site.) The difference between $(1-\rho/4)^2$ and $1-\rho/2$ for the low density of holes in the superconducting materials is too trivial to concern us. Thus in terms of the hard-core boson spin operators $(d_{l\sigma})$ we have from Eq. (5.2a)

$$\begin{aligned}H_J^{(s)}=&4J_\rho\sum_l(d_{lA}d_{lA}^\dagger+d_{lB}^\dagger d_{lB})\\&+J_\rho\sum_{l,h}(d_{lA}^\dagger d_{l+hB}^\dagger+d_{lA}d_{l+hB})\\&+2J_\rho\sum_{l,h}d_{lA}^\dagger d_{lA}d_{l+hB}d_{l+hB}^\dagger .\end{aligned} \tag{6.2b}$$

This mean-field treatment of the superexchange blocking may not be the best description. For example, if the holes were static the spin system would be better described as a percolation problem as in the treatment of dilute magnetism.[57] However, the holes are not static but move on a time scale considerably faster than the spin system. For the purposes of this paper, we do not consider further the dynamics of the spin system, thus we do not address this problem further. This is not to say that we do not consider it an important element of the problem. Quite to the contrary, we do anticipate that carrier scattering off of spin fluctuations is the strongest scattering in the system. However, this is not considered in this paper.

The carrier spin-coupling Hamiltonian is

$$H_K^{(s)}=2K_{\mathrm{eff}}\sum_{l,h}S_{lh}\cdot(S_{lA}+S_{l+hB}) , \tag{6.3a}$$

where $S_{lh}$ is the spin operator for a hole at the oxygen site $lh$. The $z$ component of $S_{lh}$ is $S_{lh}^z=\frac{1}{2}(n_{lh\uparrow}-n_{lh\downarrow})$ and the raising and lowering spin operators are $S_{lh}^+=\psi_{lh\uparrow}^\dagger\psi_{lh\downarrow}$ and $S_{lh}^-=\psi_{lh\downarrow}^\dagger\psi_{lh\uparrow}$. In terms of the Cu spin operator $(d_{l\sigma})$ and the fermion hole operators $(\psi_{lh\uparrow(\downarrow)})$ the carrier-dependent term is

$$\begin{aligned}H_K^{(s)}=K_{\mathrm{eff}}\sum_{l,h}[&(n_{lh\uparrow}-n_{lh\downarrow})(d_{lA}^\dagger d_{lA}-d_{l+hB}^\dagger d_{l+hB})\\&+\psi_{lh\downarrow}^\dagger\psi_{lh\uparrow}(d_{lA}^\dagger+d_{l+hB})\\&+\psi_{lh\uparrow}^\dagger\psi_{lh\downarrow}(d_{lA}+d_{l+hB}^\dagger)] .\end{aligned} \tag{6.3b}$$

The sum of $H_J^{(s)}+H_K^{(s)}$ constitutes the spin or static $(H^{(s)})$ term of the Hamiltonian.

We next include the transport terms of the Hamiltonian as $H^{(t)}$. There are three distinct processes by which a hole moves through the lattice, thus $H^{(t)}$ is the sum of three terms

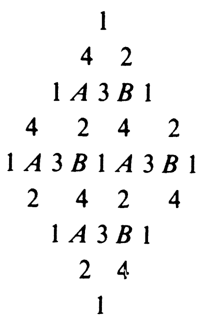
$$H^{(t)}=H_{pp}^{(t)}+H_{d^{10}}^{(t)}+H_{d^8}^{(t)} . \tag{6.4}$$

The first term $H_{pp}^{(t)}$ is the direct oxygen-oxygen hop which is one of the two largest terms in transport

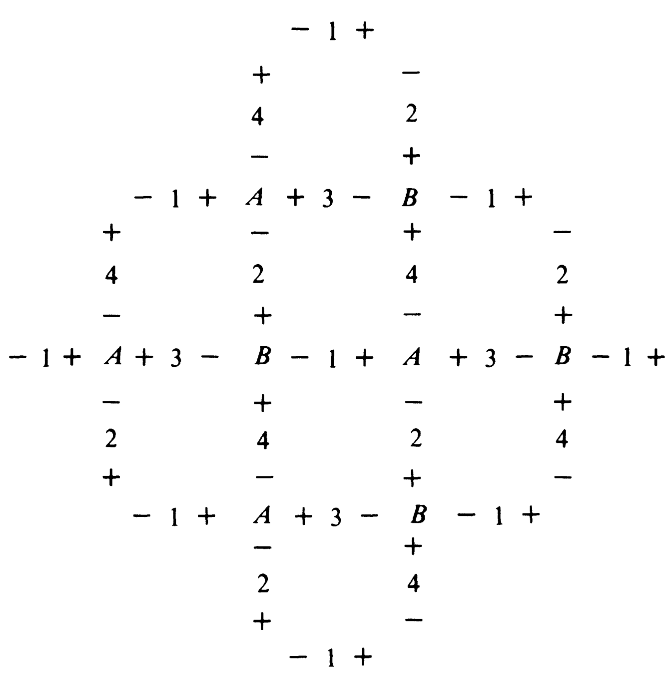
$$H_{pp}^{(t)}=\frac{1}{2}t_{pp}\sum_{\sigma,l,h_1,h_2}''(\psi_{lAh_2\sigma}^\dagger\psi_{lAh_1\sigma}+\psi_{lBh_2\sigma}^\dagger\psi_{lBh_1\sigma}) \tag{6.5a}$$

for $\sigma=\uparrow,\downarrow$, where the double prime on the summation indicates that $h_2\neq h_1$ or $-h_1$. $H_{pp}^{(t)}$ cannot change the spin of either the carrier or the localized Cu holes and is independent of the spin of the neighboring Cu holes.

We apologize for the cumbersome notation, but recall that the $CuO_2$ plane has four oxygen atoms per magnetic unit cell. If we label oxygen sites $lAh$ (where $h = -\mathbf{a}_x, -\mathbf{a}_y, \mathbf{a}_x, \mathbf{a}_y$) as $lA1$, $lA2$, $lA3$, $lA4$, respectively, then nearest neighbors to $lA1$ are $lA2$ and $lA4$, connected through space by $t_{pp}$ and through the $A$ Cu, and $lA3$ is a second nearest neighbor connected through the $A$ Cu. Similarly $lB2$ and $lB4$ are nearest neighbors to $lB1$ through the $B$ Cu ($lB3$ is again a second nearest neighbor). All other combinations are obtained from these by cyclic permutation. Thus the counting scheme looks as follows:



and one can sit at an $A$ Cu site and begin at $-\mathbf{a}_x$ *or* at a $B$ Cu site and begin at $+\mathbf{a}_x$ and count counterclockwise. In addition, we have oriented the oxygen $p_\sigma$ orbitals such that the $k = 0$ state is fully antibonding; i.e.,



where the + and − indicate the phase of the $p$ orbital. In conventional notation,[30] this is $k = X$.

The second transport route, embodied in $H_{d^{10}}^{(t)}$, is about as important as $H_{pp}^{(t)}$; however, it competes with $t_{pp}$ for transport along the $\mathbf{a}_1$ and $\mathbf{a}_2$ directions. This transport route makes use of the $d^{10}$ valence fluctuations, in the case that the Cu hole has the same spin as the oxygen hole. In this process a Cu hole nearest neighbor to an oxygen hole hops onto one of the three other nearest-neighbor oxygen sites and then the oxygen hole hops onto the Cu site. If perturbation theory were valid, we would take the effective transfer integral to be $t_{pd}^2/\Delta E$, where $\Delta E$ is the energy of the virtual ($d^{10}$) state; however, this is valid only if $t_{pd} \ll \Delta E$, which is *not* the case. In actuality, $t_{pd} \approx \Delta E$. Recall that the Cu valency fluctuates in the absence of carriers, principally between $d^9$ and $d^{10}$ in about a 3-4:1 ratio (see Sec. III). Without carriers $\Delta E$ is on the order of 3 eV. However, in the presence of a carrier hole on an adjacent oxygen site, Coulomb repulsion between the carrier hole and the localized Cu hole causes this fluctuation to be much larger ($\sim$1-2:1), since $\Delta E$ is then only about 1-2 eV. Despite the breakdown of perturbation theory, we can derive an effective transfer integral for this process which we will call $t_{d^{10}}$ ($t_{d^{10}} \sim 0.56$ eV compared to $t_{pp} \sim 0.5$ eV). We show in Appendix A how to estimate this transfer integral using a similar method as was used to compute $J$ and $K_{\text{eff}}$ in Sec. IV. This process provides transport through the following term in the Hamiltonian:

$$H_{d^{10}}^{(t)} = \sum_{l,h_1,h_2}{}' t_{d^{10}}^{21}(\theta)(d_{lA}d_{lA}^\dagger \psi_{lAh_2\downarrow}^\dagger \psi_{lAh_1\downarrow} + d_{lA}^\dagger d_{lA} \psi_{lAh_2\uparrow}^\dagger \psi_{lAh_1\uparrow} + d_{lB}d_{lB}^\dagger \psi_{lBh_2\uparrow}^\dagger \psi_{lBh_1\uparrow} + d_{lB}^\dagger d_{lB} \psi_{lBh_2\downarrow}^\dagger \psi_{lBh_1\downarrow}) , \tag{6.5b}$$

where the prime on the summation indicates that $h_2 \neq h_1$. The terms in which $h_2 = h_1$ are forbidden by the Pauli-exclusion principle. The (superscript 21) notation $t_{d^{10}}^{21}(\theta)$ indicates that this term arises from the exchange of two particles. Thus $t_{d^{10}}^{21} = -\langle p_{h_1}(1)d(2) | H_{d^{10}} | p_{h_2}(2)d(1)\rangle$. Note that $t_{d^{10}}^{12} = \langle p_{h_1}(1)d(2) | H_{d^{10}} | p_{h_2}(1)d(2)\rangle$ vanishes unless $h_2 = h_1$. The ($\theta$) notation signifies that the transfer matrix element has a nontrivial dependence on the phase between the two oxygen sites at $l\sigma h_1$ and $l\sigma h_2$. We define the interaction energy (not the transfer matrix element) in the triplet case (where the carrier and the Cu have the same spin, denoted by $T$) to be $t_{d^{10}}^T = t_{d^{10}}^{12} - t_{d^{10}}^{21}$. This energy reflects both the energy lowering of the oxygen orbital *and* the hopping integral. In Appendix A we derive the hopping integral from this interaction energy. This quantity has a different $k$ dependence than $t_{pp}$; since $t_{pp}$ does not go through a virtual the $k$ dependence is the standard and familiar one, that is $\frac{1}{2} t_{pp} \exp[i\mathbf{k} \cdot (\sigma\mathbf{h}_1 - \sigma\mathbf{h}_2)]$. In contrast, however, the $k$ dependence for transport through a virtual is, for example,

$$t_{d^{10}}^T(\mathbf{k} \cdot \mathbf{a}_1/2) = \tfrac{1}{2}[1 - \exp(i\mathbf{k} \cdot \mathbf{a}_1/2)]\sqrt{2}t_{pd} \times \tan\tfrac{1}{2}\tan^{-1}\left[\frac{\sqrt{8}t_{pd}}{|\Delta E + t_{pp}|}\right] , \tag{6.6a}$$

along the $\mathbf{a}_1$ direction. This may look complicated but it is straightforward, especially if we examine the limit of large $\Delta E$ which is the relative energy of the virtual. In this limit

$$t_{d^{10}}^T(\mathbf{k} \cdot \mathbf{a}_1/2) \approx \tfrac{1}{2}[1 - \exp(i\mathbf{k} \cdot \mathbf{a}_1/2)]\frac{2t_{pd}^2}{|\Delta E + t_{pp}|} ,$$

which obviously vanishes for $k = 0$, a direct consequence of Pauli exclusion. For the case in which the carrier and Cu spin are singlet coupled through the virtual, we have

$t^S_{d^{10}}=t^{12}_{d^{10}}+t^{21}_{d^{10}}$ which is similar except the minuses are replaced by pluses; i.e.,

$$t^S_{d^{10}}(\mathbf{k}\cdot\mathbf{a}_1/2)=\tfrac{1}{2}[1+\exp(i\mathbf{k}\cdot\mathbf{a}_1/2)]\sqrt{2}t_{pd}\times\tan\tfrac{1}{2}\tan^{-1}\left[\frac{\sqrt{8}t_{pd}}{|\Delta E-t_{pp}|}\right], \tag{6.6b}$$

which in the limit of large $\Delta E$ becomes

$$t^S_{d^{10}}(\mathbf{k}\cdot\mathbf{a}_1/2)\approx\tfrac{1}{2}[1+\exp(i\mathbf{k}\cdot\mathbf{a}_1/2)]\frac{2t^2_{pd}}{|\Delta E-t_{pp}|},$$

which now vanishes at $\mathbf{k}\cdot\mathbf{a}_1/2=\pi$. The value for $t^{T(S)}_{d^{10}}$ along the $\mathbf{a}_2$ direction is obtained by simply replacing $\mathbf{a}_1$ by $\mathbf{a}_2$. The value along $\mathbf{a}_x$ and $\mathbf{a}_y$ is obtained by setting $t_{pp}$ to zero (since there is no $t_{pp}$ transport along these directions) and replacing $\mathbf{a}_1/2$ by $\mathbf{a}_x$ or $\mathbf{a}_y$, respectively.

In summary, we have found an unusual $k$ dependence because significant transport occurs through the $d^{10}$ virtual state. There are two terms which determine the energy of a carrier state $t^{12}$ (which adds only to the diagonal energy) and $t^{21}$, because the matrix element involves multiple holes so that we have to account for indistinguishability. Thus, only $t^{21}$ contributes to transport since the hole moves only by what is effectively an exchange process.

In obtaining the expression for $H^{(t)}_{d^{10}}$ in Eq. (6.5b) we have used the identity $d_{l\sigma}d^\dagger_{l\sigma}=1-d^\dagger_{l\sigma}d_{l\sigma}$ which represents a number operator for "no spin deviations" (see Sec. V). This term contains the fact that free (i.e., without creating a spin deviation) $t^T_{d^{10}}$ transport through a particular sublattice can be shut down or opened up, depending on the carrier spin, if a spin deviation is present. The fact that $t^T_{d^{10}}$ vanishes when $k=0$ with the antibonding phase relations as illustrated above suggests that the lowest-energy state may not be $k=0$ but may involve the following phase orientation for a spin-↑ carrier depending on the ultimate values of $t_{pp}$ relative to $t^T_{d^{10}}$ (see Appendix A):

$$\begin{array}{ccccccccccccccccccc}
&&&&&&&&+&1&-&&&&&&&&\\
&&&&&&&-&&&&+&&&&&&&\\
&&&&&&&4&&&&2&&&&&&&\\
&&&&&&&+&&&&-&&&&&&&\\
&&&&+&1&-&\downarrow&-&3&+&\uparrow&-&1&+&&&&\\
&&&-&&&&+&&&&+&&&&-&&&\\
&&&4&&&&2&&&&4&&&&2&&&\\
&&&+&&&&-&&&&-&&&&+&&&\\
+&1&-&\downarrow&-&3&+&\uparrow&-&1&+&\downarrow&+&3&-&\uparrow&+&1&-\\
&&&+&&&&+&&&&-&&&&-&&&\\
&&&2&&&&4&&&&2&&&&4&&&\\
&&&-&&&&-&&&&+&&&&+&&&\\
&&&&-&1&+&\downarrow&+&3&-&\uparrow&+&1&-&&&&\\
&&&&&&&-&&&&-&&&&&&&\\
&&&&&&&2&&&&4&&&&&&&\\
&&&&&&&+&&&&+&&&&&&&\\
&&&&&&&&+&1&-&&&&&&&&
\end{array}$$

This orientation has antibonding oxygen atoms surrounding a Cu of the opposite spin to that of the carrier, but it has nonbonding oxygens about a Cu of the same spin as that of the carrier.

The final transport route $H^{(t)}_{d^8}$ is through virtual $d^8$ configurations which can only occur if the Cu spin is opposite that of the oxygen spin. This is significant though smaller than both $H^{(t)}_{pp}$ and $H^{(t)}_{d^{10}}$. Thus again using an effective transfer integral, $t^S_{d^8}=t^{12}_{d^8}+t^{21}_{d^8}\equiv 2t_{d^8}$ ($t^T_{d^8}=t^{12}_{d^8}-t^{21}_{d^8}=0$) (see Appendix A), we have the following term in the Hamiltonian:

$$H^{(t)}_{d^8}=\sum_{l,h_1,h_2}{}' t_{d^8}(\theta)(d_{lA}d^\dagger_{lA}\psi^\dagger_{lAh_2\uparrow}\psi_{lAh_1\uparrow}+d^\dagger_{lB}d_{lB}\psi^\dagger_{lBh_2\uparrow}\psi_{lAh_1\uparrow}+d^\dagger_{lA}d_{lA}\psi^\dagger_{lAh_2\downarrow}\psi_{lAh_1\downarrow}+d_{lB}d^\dagger_{lB}\psi^\dagger_{lBh_2\downarrow}\psi_{lBh_1\downarrow})\,. \tag{6.7}$$

The terms in which $h_2=h_1$ are equivalent to returning the system to an indistinguishable configuration and do contribute to the diagonal energy. The $k$ dependence of

$$t_{d^8}(\mathbf{k}\cdot\mathbf{a}_1/2)\sim\tfrac{1}{2}[1+\exp(i\mathbf{k}\cdot\mathbf{a}_1/2)]t^2_{pd}/(\Delta U+t_{pp})\,.$$

Thus when $\mathbf{k}\cdot\mathbf{a}_1/2=\pi$ the $t_{d^8}$ term vanishes.

We will treat these transport terms $H^{(t)}_{d^{10}}$ and $H^{(t)}_{d^8}$ by assuming local AF order and define $\gamma$ to be the deviation from local order. Since there is no long-range order, half the time (neglecting fluctuations about the local order) a free hole can transport through the $A$ Cu with a $t_{d^8}$ effective transfer integral and through the $B$ Cu with a $t_{d^{10}}$ effective transfer integral. The other half the time the roles of the $d^8$ and $d^{10}$ are reversed. For the purposes here we will assume that local order has $A$ Cu's down and $B$ Cu's up, hence locally and including spin fluctuations in mean field

$$H^{(t)}_{d^{10}}(\mathrm{MF})=\tfrac{1}{2}\sum_{l,h_1,h_2}{}' t^{21}_{d^{10}}(\theta)[(1-\gamma)\psi^\dagger_{lAh_2\downarrow}\psi_{lAh_1\downarrow}+(1-\gamma)\psi^\dagger_{lBh_2\uparrow}\psi_{lBh_1\uparrow}+\gamma\psi^\dagger_{lAh_2\uparrow}\psi_{lAh_1\uparrow}+\gamma\psi^\dagger_{lBh_2\downarrow}\psi_{lBh_1\downarrow}] \tag{6.8a}$$

and $H^{(t)}_{d^8}$ is similar with ↑ and ↓ interchanged

$$H^{(t)}_{d^8}(\mathrm{MF})=\tfrac{1}{2}\sum_{l,h_1,h_2}{}' t_{d^8}(\theta)[(1-\gamma)\psi^\dagger_{lAh_2\uparrow}\psi_{lAh_1\uparrow}+(1-\gamma)\psi^\dagger_{lBh_2\downarrow}\psi_{lBh_1\downarrow}+\gamma\psi^\dagger_{lAh_2\downarrow}\psi_{lAh_1\downarrow}+\gamma\psi^\dagger_{lBh_2\uparrow}\psi_{lBh_1\uparrow}]\,. \tag{6.8b}$$

When the local order changes then $A$ and $B$ are interchanged.

The $t_{d^{10}}$ and $t_{d^8}$ processes do more than enhance transport; most importantly for superconductivity, they also provide spin-flip mechanisms (through the $t^{21}_{d^{10}}$ and $t^{21}_{d^8}$ above). These spin-flip mechanisms are in addition to and in fact much more important than those from the second and third terms in $H^{(s)}_K$ [Eq. (6.3)]. Thus there are terms in the Hamiltonian which both transport charge and create spin flips in the Cu spin system and which form the essence of the origin of the spin-hybrid carriers:

$$H^{(st)}=H_{d^8}^{(st)}+H_{d^{10}}^{(st)}$$

where

$$H_{d^8}^{(st)}+H_{d^{10}}^{(st)}=\tfrac{1}{2}[t_{d^8}(\theta)+t_{d^{10}}^{2l}(\theta)]\sum_{l,h_1,h_2}{}' (d_{lA}^\dagger\psi_{lAh_2\downarrow}^\dagger\psi_{lAh_1\uparrow}+d_{lB}\psi_{lBh_2\downarrow}^\dagger\psi_{lBh_1\uparrow}+d_{lA}\psi_{lAh_2\uparrow}^\dagger\psi_{lAh_1\downarrow}+d_{lB}^\dagger\psi_{lBh_2\uparrow}^\dagger\psi_{lBh_1\downarrow})\,, \tag{6.9}$$

and where the superscript $(st)$ signifies spin flip *and* transport. Now the $h_2=h_1$ terms are equivalent to simple exchange processes, and these are included in the second and third terms of $H_K^{(s)}$ [Eq. (6.3b)].

Each term in $H_{d^{10}}^{(st)}$, such as $d_{lA}^\dagger\psi_{lh_2\downarrow}^\dagger\psi_{lh_1\uparrow}$, plays two roles depending on the circumstances. This might be a free-carrier, spin ↑, entering a region whose local AF order has sublattice $A$ spin ↓. The carrier hole can change roles with the $A$ Cu hole creating a spin-↓ carrier hole at site $lh_2$ and simultaneously creating a spin deviation in the local AF order, $d_{lA}^\dagger$. We call this a flipped-carrier configuration as the term represents the conversion of a free-carrier Cu($A$)-O-Cu($B$) spin configuration (↓↑↑) to a flipped-carrier configuration of $(\underline{\uparrow}\downarrow\uparrow)$ where the underline signifies a deviation from local order. However, if the local order has the $A$ sublattice spin ↑, then the same term restores the local AF order; i.e., $(\underline{\downarrow}\uparrow\downarrow)\rightarrow(\uparrow\downarrow\downarrow)$, flipped carrier to free carrier. Again, we will assume in Sec. VI B that the local order has sublattice $A$ spin ↓.

In summary, $H=H^{(s)}+H^{(t)}+H^{(st)}$. The latter two terms provide transport. However, the last term produces spin deviations in the local AF order at the same time as transporting the charge. This is the *dominant* term for interactions between carriers and the spin systems. The first term determines diagonal energies and creates spin deviations. It does not transport charge. The dominant spin-flip mechanism is hence *not* $K_{\text{eff}}$ in $H^{(s)}$, but most importantly $t_{d^{10}}^{2l}$ and of some importance $t_{d^8}$ in $H^{(st)}$ (we estimate $K_{\text{eff}}\sim 0.1$ eV while $t_{d^{10}}^{2l}\sim 0.36$–$0.56$ eV and $t_{d^8}$ $\sim 0.06$ eV; see Appendix A). In addition, there are three $t_{d^{10}}^{2l}$ and three $t_{d^8}$ to each $K_{\text{eff}}$.

### B. Spin-hybrid models

Needless to say, the Hamiltonian cannot be solved exactly; nonetheless, in Sec. VI C a simple carrier band-structure model will illustrate the consequences of the carrier-spin interactions for determining the nature of the carrier quasiparticle. First, however, in order to bilinearize the Hamiltonian, we define two "composite particles" composed of a fermion oxygen carrier hole and the adjacent Cu-spin deviation:

$$c_{lhA_\uparrow}=\psi_{lAh\downarrow}d_{lA}\,, \tag{6.10a}$$

$$c_{lhB_\downarrow}=\psi_{lBh\uparrow}d_{lB}\,. \tag{6.10b}$$

Recall that $d_{lA(B)}$ destroys a spin deviation at lattice site $lA(B)$, thus it takes a ↑(↓) Cu spin to a ↓(↑) Cu spin. When the local order reverses we have two different particles, namely

$$c_{lhA_\downarrow}=\psi_{lAh\uparrow}d_{lA}^\dagger\,, \tag{6.10c}$$

$$c_{lhB_\uparrow}=\psi_{lBh\downarrow}d_{lB}^\dagger\,. \tag{6.10d}$$

We note that the fermion commutation relations for this composite carrier are peculiar, in that in terms of mean-field occupations on the local ordered spin "particles" we have

$$c_{lhA_\sigma}^\dagger c_{lhA_\sigma}+c_{lhA_\sigma}c_{lhA_\sigma}^\dagger+c_{l+h,-hB_\sigma}^\dagger c_{l+h,-hB_\sigma}+c_{l+h,-hB_\sigma}c_{l+h,-hB_\sigma}^\dagger=1\,, \tag{6.11}$$

where $\sigma=\uparrow$ or ↓. Thus it seems that $c_{lhA_\uparrow}$ is only part of a particle; i.e., the one which can make a spin deviation on sublattice $A$. The other part of the particle unmakes spin deviations on sublattice $B$, which is particle $c_{l+h,-hB_\uparrow}$. This will have significance in a future paper when we consider scattering off the changing local order. However, for the purposes of this paper we assume local (denoted below by a superscript $L$) order with sublattice $A$ ($B$) spin down (up). Then the relevant "particles" are $c_{lh\sigma}$ with $\sigma=A_\uparrow$ and $B_\downarrow$) and in terms of these $H^{(t)}$ and $H^{(st)}$ become bilinear. We find

$$H^{(t;L)}=H_{pp}^{(t;L)}+H_V^{(t;L)}$$

($V$ for virtual), where

$$H_{pp}^{(t;L)}=\tfrac{1}{2}t_{pp}\sum_{l,h_1,h_2}{}''[(1-\gamma)(\psi_{lAh_2\uparrow}^\dagger\psi_{lAh_1\uparrow}+\psi_{lBh_2\downarrow}^\dagger\psi_{lBh_1\downarrow}+c_{lh_2A_\uparrow}^\dagger c_{lh_1A_\uparrow}+c_{lh_2B_\downarrow}^\dagger c_{lh_1B_\downarrow})] \tag{6.12a}$$

and

$$\begin{aligned}H_V^{(t;L)}=\tfrac{1}{2}(1-\gamma)\sum_{l,h_1,h_2}{}' t_{d^8}(\theta)[(&\psi_{lAh_2\uparrow}^\dagger\psi_{lAh_1\uparrow}+\psi_{lBh_2\downarrow}^\dagger\psi_{lBh_1\downarrow}+c_{lh_2A_\uparrow}^\dagger c_{lh_1A_\uparrow}+c_{lh_2B_\downarrow}^\dagger c_{lh_1B_\downarrow})\\&+(1-\gamma)t_{d^{10}}^{2l}(\theta)(\psi_{lAh_2\downarrow}^\dagger\psi_{lAh_1\downarrow}+\psi_{lBh_2\uparrow}^\dagger\psi_{lBh_1\uparrow})]\,.\end{aligned} \tag{6.12b}$$

Finally

$$H_V^{(st,L)}=\tfrac{1}{2}(1-\gamma)[t_{d^{10}}^{2l}(\theta)+t_{d^8}(\theta)]\sum_{l,h_1,h_2}{}'(c_{lh_2A_\uparrow}^\dagger\psi_{lAh_1\uparrow}+\psi_{lAh_1\uparrow}^\dagger c_{lh_2A_\uparrow}+c_{lh_2B_\downarrow}^\dagger\psi_{lBh_1\downarrow}+\psi_{lBh_1\downarrow}^\dagger c_{lh_2B_\downarrow})\,, \tag{6.12c}$$

where if the local order reverses then $A$ and $B$ are interchanged. Note that we have used the following identity:

$$b_{lA}^{\dagger}\psi_{lAh_2\downarrow}^{\dagger}\psi_{lAh_1\uparrow}=c_{lh_2A\uparrow}^{\dagger}\psi_{lAh_1\uparrow}=\psi_{lAh_2\downarrow}^{\dagger}c_{lh_1A\downarrow}\,. \tag{6.13}$$

Recall the use of $\theta$ in the notation is to signal an unusual $k$ dependence for these terms.

The reader may now be wondering what happened to the missing terms, for example, the remaining $\gamma(t_{d^8}+t_{d^{10}}^{21})$ in $H_V^{(st,L)}$. These terms couple to carrier configurations which are in addition to free-↑ $(\uparrow\uparrow\downarrow)$, free-↓ $(\uparrow\downarrow\downarrow)$, flipped-↑ $(\uparrow\downarrow\underline{\uparrow})$, and flipped-↓ $(\underline{\downarrow}\uparrow\downarrow)$ carrier configurations. The missing terms correspond to couplings to configurations such as Cu($A$)-O-Cu($B$)-O: $(\cdot\underline{\downarrow}\uparrow\underline{\uparrow})$, $(\underline{\downarrow}\downarrow\underline{\uparrow}\cdot)$, and $(\uparrow\underline{\downarrow}\cdot\underline{\uparrow})$ and $(\underline{\downarrow}\cdot\underline{\uparrow}\downarrow)$. There are a total of eight of these, four each for up (↑) and down (↓) carriers. We call these double-flipped carriers for obvious reasons (recall the underline indicates a deviation from local order). There are four configurations for each spin of the carrier because there are four oxygen sites neighboring the flipped Cu which has opposite spin to the carrier. In Appendix B we include these configurations in the Hamiltonian. For the remainder of this section we only consider the two lowest-energy configurations, free and flipped.

The physical description of this Hamiltonian is discussed now. If there were no interactions between the spin and carrier subsystems (i.e., $K_{\rm eff}=t_{d^8}=t_{d^{10}}^{21}=0$) then $H$ decouples, $H=H_J^{(s)}+H_{pp}^{(t)}$, and the spin and carrier systems exist essentially independently. An isolated carrier on the oxygen sublattice would delocalize through the oxygen-oxygen transfer integrals $t_{pp}$ in $H_{pp}^{(t)}$, so that its energy would be $4t_{pp}$, and its orbital would be the purely antibonding Bloch wave corresponding to the top of the oxygen $2p$ bands. This free-hole carrier then moves through the lattice without disturbing the AF order of the Cu spins except in the average sense of reducing $J$ to $J_\rho\sim(1-\rho/2)J$. Its wave function can be composed from the set of local spin-orbital configurations written as $d_\downarrow^9 p_\sigma^5 d_\uparrow^9$ for three holes (two "localized" spins and one carrier) in a Cu-O-Cu unit, which we may write in shorthand as $\downarrow\sigma\uparrow$. Its energy is independent of its spin $\sigma=\uparrow$ or $\downarrow$.

A second configuration, the flipped-carrier configuration referred to above, arises because both $H_K^{(s)}$ and $H^{(st)}$ couple the free carrier to this configuration. Of the coupling terms $H^{(st)}$ dominates since $K_{\rm eff}$ is a relatively small energy compared to $t_{d^{10}}^{21}$, and in addition there are more $t_{d^{10}}^{21}$ matrix elements than there are $K_{\rm eff}$ matrix elements. This dominance is clearly true at low carrier densities where the magnitude of $t_{d^{10}}^{21}$ is not diluted by its $k$ dependence. While it is true that $t_{d^{10}}^{21}$ decreases as we begin filling the band, it never reverses sign. Furthermore, from our estimates (see Appendix A) it would require nearly half filling the band in order for $K_{\rm eff}$ to be more important than $t_{d^{10}}^{21}$ even at the Fermi level.

The diagonalization of this Hamiltonian with these two spin-orbital configurations yields the spin-hybrid carrier wave functions. The amount of flipped-carrier configuration weight in the wave function reflects the probability of a carrier being associated with a spin deviation as a function of the wave vector $k$.

### C. Band-structure description of the spin hybrid

We now write a band-structure description of the carrier which contains our spin-hybrid model for the quasiparticle. This was derived in the normal manner of including translational symmetry by writing our basis functions (the free- and the flipped-carrier configurations) in the Bloch form (with the $k=0$ state taken as fully antibonding) and recognizing the unusual $k$ dependence for $t_{d^8}$ and $t_{d^{10}}^{21}$. A post unitary transformation assures that the final matrix is real. If we add one spin-↑ hole to an otherwise AF system of Cu spins (thus $S_z^{\rm tot}=\frac{1}{2}$), we find the Hamiltonian matrix with four oxygens per magnetic unit cell to have the following general form (before the unitary transformation):

$$H(k)=\begin{pmatrix} h_{11}(k) & h_{12}(k) & h_{13}(k) & h_{14}(k)\\ h_{12}^*(k) & h_{11}(k) & h_{14}(k) & h_{24}(k)\\ h_{13}^*(k) & h_{14}^*(k) & h_{11}(k) & h_{12}^*(k)\\ h_{14}^*(k) & h_{24}^*(k) & h_{12}(k) & h_{11}(k)\end{pmatrix}, \tag{6.14}$$

where $h_{ij}$ is a matrix which links and $i$-site oxygen with a $j$-site oxygen. After a unitary transformation this becomes

$$\begin{pmatrix} [h_{11}+\frac{1}{2}(h_{13}^R+h_{24}^R)+h_{12}^R+h_{14}^R] & \frac{1}{2}(h_{13}^I+h_{24}^I) & \frac{1}{2}(h_{13}^R-h_{24}^R) & \frac{1}{2}(h_{13}^I-h_{24}^I)+h_{12}^I+h_{14}^I\\ \frac{1}{2}(h_{13}^I+h_{24}^I) & [h_{11}-\frac{1}{2}(h_{13}^R+h_{24}^R)+h_{12}^R-h_{14}^R] & [\frac{1}{2}(h_{13}^I+h_{24}^I)-h_{12}^I-h_{14}^I] & \frac{1}{2}(h_{13}^I-h_{24}^I)\\ \frac{1}{2}(h_{13}^I-h_{24}^I) & [\frac{1}{2}(h_{13}^I+h_{24}^I)-h_{12}^I-h_{14}^I] & [h_{11}+\frac{1}{2}(h_{13}^R+h_{24}^R)-h_{12}^R-h_{14}^R] & \frac{1}{2}(h_{13}^R+h_{24}^R)\\ \frac{1}{2}(h_{13}^I-h_{24}^I)+h_{12}^I+h_{14}^I & \frac{1}{2}(h_{13}^I-h_{24}^I) & \frac{1}{2}(h_{13}^I+h_{24}^I) & [h_{11}-\frac{1}{2}(h_{13}^R+h_{24}^R)-h_{12}^R+h_{14}^R] \end{pmatrix}, \tag{6.15}$$

where we have dropped the explicit reference to the wave vector $k$ and used the superscript notation $R$ and $I$ to refer to the real and imaginary parts, respectively. Finally, after evaluating the $h_{ij}$ matrices we obtain the following relations:

$$\begin{aligned}&h_{12}+h_{14}=2B_0+2B_1c_{k_x}c_{k_y}+i2B_2s_{k_x}c_{k_y},\quad h_{12}-h_{14}=-2B_1s_{k_x}s_{k_y}+i2B_2c_{k_x}s_{k_y}\,,\\&h_{13}\pm h_{24}=(C_0\pm C_0)+C_1(c_{k_{2x}}\pm c_{k_{2y}})+iC_2(s_{k_{2x}}\pm s_{k_{2y}})\,,\end{aligned}\tag{6.16}$$

where $B_j$ and $C_j$ are constant matrices with dimension the number of configurations. For example, if we consider only free-carrier and flipped-carrier configurations then these matrices are 2×2. Note that the matrices for a spin-$\downarrow$ carrier replaces $k$ with $-k$. The notation is as follows: $c$ refers to cos, $s$ refers to sin, $k_x$ to $\mathbf{k}\cdot\hat{\mathbf{a}}_x/2$, $k_y$ to $\mathbf{k}\cdot\hat{\mathbf{a}}_y/2$, $k_{2x}$ to $\mathbf{k}\cdot\hat{\mathbf{a}}_x$, and $k_{2y}$ to $\mathbf{k}\cdot\hat{\mathbf{a}}_y$; i.e., $c_{k_x}=\cos(\mathbf{k}\cdot\hat{\mathbf{a}}_x)$, etc. Finally, the band-structure Hamiltonian has the following form:

$$\begin{pmatrix}A_1+c_{k_{2x}}C_1+2c_{k_x}c_{k_y}B_1 & \frac{1}{2}(c_{k_{2x}}-c_{k_{2y}})C_1 & s_{k_{2x}}C_2+2s_{k_x}c_{k_y}B_2 & s_{k_{2y}}C_2+2c_{k_x}s_{k_y}B_2\\ \frac{1}{2}(c_{k_{2x}}-c_{k_{2y}})C_1 & A_2+c_{k_{2y}}C_1+2c_{k_x}c_{k_y}B_1 & s_{k_{2x}}C_2-2c_{k_x}c_{k_y}B_2 & -s_{k_{2y}}C_2+2c_{k_x}c_{k_y}B_2\\ -s_{k_{2x}}C_2-2s_{k_x}c_{k_y}B_2 & s_{k_{2x}}C_2+2s_{k_x}c_{k_y}B_2 & A_3-c_{k_{2x}}C_1 & 2B_1s_{k_x}s_{k_y}\\ -s_{k_{2y}}C_2-2c_{k_x}c_{k_y}B_2 & s_{k_{2y}}C_2-2c_{k_x}c_{k_y}B_2 & 2B_1s_{k_x}s_{k_y} & A_2-c_{k_{2y}}C_1\end{pmatrix},\tag{6.17}$$

where

$$A_1=h_{11}+C_0+2B_0,\quad A_2=h_{11}-C_0,\quad A_3=h_{11}+C_0-2B_0\,.\tag{6.18a}$$

With only the two configurations in the basis, free and flipped carrier, the matrices are

$$h_{11}=\begin{pmatrix}3J & -K_{\text{eff}}\\ -K_{\text{eff}} & K_{\text{eff}}\end{pmatrix},\tag{6.18b}$$

$$B_0=\begin{pmatrix}(1-\gamma)\frac{1}{2}(t^{21}_{d^{10}}+t_{d^8}) & -(1-\gamma)\frac{1}{2}(t^{21}_{d^{10}}+t_{d^8})\\ -(1-\gamma)\frac{1}{2}(t^{21}_{d^{10}}+t_{d^8}) & \frac{1}{2}(1-\gamma)t_{d^8}\end{pmatrix},\tag{6.18c}$$

$$B_1=\begin{pmatrix}(1-\gamma)[2t_{pp}+\frac{1}{2}(t_{d^8}-t^{21}_{d^{10}})] & -(1-\gamma)\frac{1}{2}(t^{21}_{d^{10}}+t_{d^8})-\gamma t_{pp}\\ -(1-\gamma)\frac{1}{2}(t^{21}_{d^{10}}+t_{d^8})-\gamma t_{pp} & (1-\gamma)(t_{pp}+\frac{1}{2}t_{d^8})\end{pmatrix},\tag{6.18d}$$

$$B_2=\begin{pmatrix}(1-\gamma)\frac{1}{2}(t_{d^8}+t^{21}_{d^{10}}) & -(1-\gamma)\frac{1}{2}(t^{21}_{d^{10}}+t_{d^8})\\ -(1-\gamma)\frac{1}{2}(t_{d^8}+t^{21}_{d^{10}}) & (1-\gamma)(t_{pp}+\frac{1}{2}t_{d^8})\end{pmatrix}.\tag{6.18e}$$

The $C_j$ matrices are obtained from the $B_j$ matrices by setting $t_{pp}=0$. Diagonalization yields the band structure of the spin-hybrid carriers. For a low carrier concentration ($\ll$ one per unit cell), only the lowest-energy states are of interest. We would like to say that this is near $k=0$; however, as indicated above and in Appendix A the lowest-energy wave function may have a different phase around Cu's of the same spin and around Cu's of the opposite spin, as illustrated in the first part of this section. This is the case for a sufficiently small $\Delta E$. Here we ignore this complication simply because it does not alter the qualitative picture; we are focusing on the hybrid character that is determined by carriers interacting with localized Cu spins of the opposite orientation. The subtlety, however, arises from carriers interacting with Cu spins of the *same* orientation. Thus the mixing which determines the hybrid is relatively immune to this complication. For $k=0$ we have the simplifying feature that this model Hamiltonian blocks as follows:

$$\begin{pmatrix}A_1+C_1+2B_1 & 0 & 0 & 0\\ 0 & A_3+C_1-2B_1 & 0 & 0\\ 0 & 0 & A_2-C_1 & 0\\ 0 & 0 & 0 & A_2-C_1\end{pmatrix}.$$

The highest eigenvalue (lowest-energy configuration) is the largest eigenvalue from $H_0=A_1+C_1+2B_1$, where we find a most important result:

$$H_0=\begin{pmatrix}3J+(1-\gamma)3t_{d^8}+4t_{pp} & -[K_{\text{eff}}-3(1-\gamma)(t_{d^8}+t^{21}_{d^{10}})-2\gamma t_{pp}]\\ -[K_{\text{eff}}+3(1-\gamma)(t_{d^8}+t^{21}_{d^{10}})+2\gamma t_{pp}] & K_{\text{eff}}+(1-\gamma)(2t_{pp}+3t_{d^8})\end{pmatrix}.\tag{6.19}$$

From this matrix, it is easy to see the essential physics of the carrier quasiparticles: First there is strong mixing between the free-carrier and the flipped-carrier configurations; the diagonal energy difference between these two configurations is $3J-K_{\text{eff}}+(1-\gamma)2t_{pp}$, but the off-diagonal coupling is $K_{\text{eff}}+3(1-\gamma)(t_{d^8}+t_{d^{10}}^{2})+2\gamma t_{pp}$. Thus, using the energy parameters $K_{\text{eff}}=0.1$ eV (see Sec. IV B), $t_{d^{10}}^{2}=0.56$ eV, $t_{d^8}=0.06$ eV (see Appendix A), and $t_{pp}=0.5$ eV (see Sec. II) the ratio of the coupling ($\sim 1.9$ eV) to the energy difference ($\sim 1.2$ eV) is $\sim 1.6$ (with $\gamma=0$ for simplicity), which yields a mixing angle of 36°, or 65% free carrier and 35% flipped carrier, where the latter may be interpreted as the precent of time that a carrier is associated with a spin deviation on the Cu sublattice. Since $K_{\text{eff}}\sim t_{d^8}\ll t_{d^{10}}\sim t_{pp}$, we see that the $d^{10}$ fluctuations dominate the creation of spin deviations. The energy lowering due this hybridization is $\sim 1.4$ eV. This energy lowering contributes significantly to stabilizing a hole in the $p_\sigma$ band, which is the principal reason that we find the top of the $p_\sigma$ band (O $p$ orbitals along the Cu-O axis) to be significantly higher in energy than the $p_\pi$ band (O $p$ orbitals perpendicular to the Cu-O axis) even though the center of the $p_\pi$ band lies $\sim 0.8$ eV above the center of the $p_\sigma$ band is the absence of Cu-O interactions. This 0.8 eV energy is due to the difference in the Madelung potential for the two orbitals.

As $k$ increases the diagonal and the off-diagonal elements both decrease comparably, thus the mixing is strong across the entire band. In addition, we expect that these qualitative trends survive a careful treatment of the nonstatic, imperfect local order of the Cu spin system as long as the time scale for Cu spin fluctuations is significantly longer than that for carrier motion. Since Cu spin fluctuations are occurring on a time scale a factor of 5 to 10 slower in these systems, this should be sufficient.

An alternative view of this spin hybrid is that if a spin-up hole is added to the system, the resulting quasiparticle is a linear combination of a spin-up hole and a spin-down hole on the oxygen sublattice. This is the origin of an interesting subtlety of the model. As mentioned above, adding a down hole to the system produces the same band structure with $k$ replaced by $-k$. Thus we define the spin-hybrid quasiparticle as follows:

$$\begin{aligned}\Phi_{k\uparrow}&=\Psi_{k\uparrow}\cos\theta_k+C_{kA}\sin\theta_k\,,\\ \Phi_{-k\downarrow}&=\Psi_{-k\downarrow}\cos\theta_k+C_{-kB}\sin\theta_k\,,\end{aligned}\tag{6.20}$$

where $\cos^2\theta_k$ represents the probability of free-carrier character and hence $\sin^2\theta_k$ represents the probability of flipped-carrier character. There is no violation of the Pauli-exclusion principle if we populate both $\Phi_{k\uparrow}$ and $\Phi_{-k\downarrow}$. However, we cannot populate both $\Phi_{k\uparrow}$ and $\Phi_{k\downarrow}$, since this would indeed violate Pauli exclusion. In other words, we must be cognizant of the fact that the spin hybrid is a linear combination of an up and a down oxygen hole state. Thus when we compute the band structure we throw out all states with $-k\uparrow$ or $k\downarrow$.

In Appendix B we expand our basis to include additional configurations and find that the free-carrier component decreases and the flipped-carrier component increases. Thus we are able to conclude that the carriers in CuO$_2$ sheets, although they move as relatively mobile free holes, contain a strong component ($>40\%$) of spin-deviation character.

## VII. QUASIPARTICLE-QUASIPARTICLE HAMILTONIAN

In this section it is our goal to derive a Hamiltonian which describes the dynamics of quasiparticle-quasiparticle interactions for quasiparticles of opposite spin. Locally, we assume that sublattice $A$ has all spin down ($\downarrow$) in the AF reference state, hence a flipped-carrier in sublattice $A$ has Cu spin up ($\uparrow$) and oxygen spin down ($\downarrow$). For simplicity we exclude the double-flipped-carrier configurations.

The total Hamiltonian, including the chemical potential, can be written as

$$H_\mu=H_A+H_B+H_{AB}+h_{AB}\,.\tag{7.1}$$

The first two terms are "single"-particle terms and are just the quasiparticle Hamiltonians for spin $\uparrow$ ($H_A$) and for spin $\downarrow$ ($H_B$) from the preceding section. These define the energies (relative to the chemical potential $\mu$) of the two different types of "fermion" charged quasiparticles (a free carrier and a flipped carrier) on the $A$ and $B$ sublattices. Recall that we have four $k$ states for each carrier, but the band structure produces eight. The relevant four are the highest four in energy. The $k\uparrow$ states are degenerate with the $-k\downarrow$ states. The standard chemical potential $\mu$ controls the total number of fermion charged particles ($\langle\hat{N}_{c\sigma}\rangle$) where

$$\hat{N}_{c\sigma}=\sum_{l,h}c_{lhA_\sigma}^{\dagger}c_{lhA_\sigma}+\psi_{lAh\sigma}^{\dagger}\psi_{lAh\sigma}\tag{7.2}$$

for $\sigma=\uparrow,\downarrow$. (The total number of charged carriers per sublattice, $\langle\hat{N}_{c\sigma}\rangle$, is $\rho N$ where $N$ is the number of Cu sites per sublattice and $\rho$ is the density of dopant holes; e.g., La$_{2-\rho}$Sr$_\rho$CuO$_4$). The one thing that this Hamiltonian allows which technically it should not is the simultaneous occupancy of a spin flip at $lA$ and a free-spin-$\downarrow$ carrier at site $lAh$, since this combination represents a flipped carrier. We could solve this problem by imposing the additional constraints that

$$\begin{aligned}&d_{lA}^{\dagger}d_{lA}+\sum_{h,\sigma}(c_{lhA_\sigma}^{\dagger}c_{lhA_\sigma}+\psi_{lAh\sigma}^{\dagger}\psi_{lAh\sigma})+d_{lA}d_{lA}^{\dagger}=1\,,\\ &d_{lB}^{\dagger}d_{lB}+\sum_{h,\sigma}(c_{lhB_\sigma}^{\dagger}c_{lhB_\sigma}+\psi_{lBh\sigma}^{\dagger}\psi_{lBh\sigma})+d_{lB}d_{lB}^{\dagger}=1\,.\end{aligned}$$

Imposing this constraint on the mean is an equivalent formulation to that of the spin waves in Sec. V B in the absence of charge carriers.[48,54]

The third term in the Hamiltonian (the diagonal pairing term) has its origin in the energy costs (primarily due to Coulomb repulsion) and the energy savings realized by having carriers on neighboring sites out to fifth neighbor (a distance of $4r_{\text{CuO}}$). All of these terms are quartic and depend on number operators; i.e., they are diagonal energy corrections. Recall that we are including terms for quasiparticles of opposite spin, where the quasiparticle spin is not determined by the spin of the oxygen hole but by the spin of the free-carrier component. Thus a spin-up (down) particle can be a free carrier $\psi_{lAh\uparrow}$ ($\psi_{lBh\downarrow}$) or a flipped carrier $c_{lhA_\uparrow}$ ($c_{lhB_\downarrow}$) on the $A$ ($B$) sublattice. We will use the spatial identity that $lAh \equiv (l+h)B(-h)$ in order to simplify the notation. Sometimes this notational simplification leads to double counting. Accordingly, we will divide by two wherever appropriate.

We now separate $H_{AB}$ into a Coulomb and an exchange pairing term, thus

$$H_{AB} = U_{AB} + V_{AB} , \tag{7.3}$$

where $U_{AB}$ is the Coulomb repulsion between two carriers. We assume that in principle the Coulomb repulsion of all up (down) carriers with all other up (down) carriers as well as the Coulomb potential energy from all the point charges in the system have been included in $H_{A(B)}$. Thus $U_{AB}$ is just the Coulomb repulsion between an up and a down carrier; i.e.,

$$U_{AB} = \sum_{l_1,l_2,h_1,h_2} U(|l_1Ah_1 - l_2Bh_2|)(\psi^\dagger_{l_1Ah_1\uparrow}\psi_{l_1Ah_1\uparrow} + c^\dagger_{l_1h_1A_\uparrow}c_{l_1h_1A_\uparrow})(\psi^\dagger_{l_2Bh_2\downarrow}\psi_{l_2Bh_2\downarrow} + c^\dagger_{l_2h_2B_\downarrow}c_{l_2h_2B_\downarrow}) .$$

The term $U(|l_1Ah_1 - l_2Bh_2|)$ is the Coulomb potential energy between two holes, one at site $l_1Ah_1$ and the other at $l_2Bh_2$. We assume that this depends only on the distance between the holes. Thus, for holes further than a few neighbors we would take $U(r_1-r_2) = e^2/(\epsilon_\infty |r_1-r_2|)$. We rewrite $U_{AB}$ in terms of mean-field operators plus fluctuations, thus defining

$$\langle \psi^\dagger_{lAh\uparrow}\psi_{lAh\uparrow}\rangle = \langle \psi^\dagger_{lBh\downarrow}\psi_{lBh\downarrow}\rangle = n_f \text{ and } \langle c^\dagger_{lhA_\uparrow}c_{lhA_\uparrow}\rangle = \langle c^\dagger_{lhB_\downarrow}c_{lhB_\downarrow}\rangle = n_c$$

and noting that $n_f + n_c = \frac{1}{2}\rho$ we have

$$U_{AB} = U_{\rm HF} + \Delta U_{AB} ,$$

where $U_{\rm HF}$ is the Hartree-Fock or mean-field Coulomb energy which is present in the system more-or-less independent of whether we pair the particles or not. If, however, we try to evaluate $U_{\rm HF}$ as an individual term, it would, of course, diverge. Nonetheless, when combined with all the other Coulomb terms implicitly in $H_A$ and $H_B$ the net would be a finite and constant energy per carrier.

The residual carrier-carrier Coulomb potential is then

$$\Delta U_{AB} = \sum_{l_1,l_2,h_1,h_2} U(|l_1Ah_1 - l_2Bh_2|)(\psi^\dagger_{l_1Ah_1\uparrow}\psi_{l_1Ah_1\uparrow} + c^\dagger_{l_1h_1A_\uparrow}c_{l_1h_1A_\uparrow} - \tfrac{1}{2}\rho)(\psi^\dagger_{l_2Bh_2\downarrow}\psi_{l_2Bh_2\downarrow} + c^\dagger_{l_2h_2B_\downarrow}c_{l_2h_2B_\downarrow} - \tfrac{1}{2}\rho) . \tag{7.4}$$

If we bring two holes closer together than expected on average, we raise the Coulomb energy of the system. Thus for a superconducting state, we have a positive contribution to the energy from $\Delta U_{AB}$. On the other hand, we lower the Coulomb energy if we keep the holes further apart than on average. Thus if the holes are correlated so as to avoid each other we have a negative contribution to the energy from $\Delta U_{AB}$. On the other hand, both pairing and correlating also raise the kinetic energy of the particles. Nonetheless, clearly the Hartree-Fock solution will have higher energy than the correlated solution. Furthermore, if the system can support a superconducting state the Hartree-Fock solution will also be higher in energy than the superconducting state. The question will be whether the correlated state or the superconducting state will be the lowest in energy at any given density. Crude estimates can set bounds on the maximum correlation energy in the correlated solution versus the cost in Coulomb energy in the pairing solution; but only a careful calculation of both can truly estimate the net pairing energy.

The pairing potential $V_{AB}$ is the sum of four terms, two of which are attractive, and two of which are most likely to be repulsive. We write these terms as

$$\begin{aligned}
V_{AB} = & -\tfrac{1}{2}(K_{\rm eff}+J_1)\sum_{l,h_1,h_2}{}' (c^\dagger_{lh_1A_\uparrow}c_{lh_1A_\uparrow}\psi^\dagger_{lAh_2\downarrow}\psi_{lAh_2\downarrow} + c^\dagger_{lh_1B_\downarrow}c_{lh_1B_\downarrow}\psi^\dagger_{lBh_2\uparrow}\psi_{lBh_2\uparrow}) \\
& + \tfrac{1}{2}(K_{\rm eff}-J_1)\sum_{l,h_1,h_2}{}' (c^\dagger_{lh_1A_\uparrow}c_{lh_1A_\uparrow}c^\dagger_{(l+h_1)h_2B_\downarrow}c_{(l+h_1)h_2B_\downarrow} + c^\dagger_{lh_1B_\downarrow}c_{lh_1B_\downarrow}c^\dagger_{(l+h_1)h_2A_\uparrow}c_{(l+h_1)h_2A_\uparrow}) \\
& + (2K_{\rm eff}-J_1)\sum_{l,h_1} (c^\dagger_{lh_1A_\uparrow}c_{lh_1A_\uparrow}c^\dagger_{(l+h_1)-h_1B_\downarrow}c_{(l+h_1)-h_1B_\downarrow}) \\
& - J_1 \sum_{l,h_1,h_2,-h_3}{}' (c^\dagger_{lh_1A_\uparrow}c_{lh_1A_\uparrow}c^\dagger_{(l+h_2)-h_3B_\downarrow}c_{(l+h_2)-h_3B_\downarrow} + c^\dagger_{lh_1B_\downarrow}c_{lh_1B_\downarrow}c^\dagger_{(l+h_2)-h_3A_\uparrow}c_{(l+h_2)-h_3A_\uparrow}) ,
\end{aligned} \tag{7.5}$$

where in the last term the summation is over $h_2 \neq h_1$ and $h_3 \neq h_2$. The first term (for $h_1 \neq h_2$) is the energy savings ($K_{\rm eff}+J_1$, the $\frac{1}{2}$ is due to double counting) for having a flipped carrier on one sublattice and a free-carrier nearest neighbor to the flipped Cu spin, for example, Cu($B$)-O-Cu($A$)-O-Cu($B$) $(\uparrow\downarrow\underline{\uparrow}\downarrow\uparrow)$. The second term is the energy cost (assuming $J_1 < K_{\rm eff}$, it will be a savings otherwise) for having two "overlapping" flipped carriers such as Cu($B$)-O-Cu($B$)-Cu($A$) $(\underline{\uparrow}\downarrow\underline{\downarrow}\uparrow\downarrow)$. This saves an enhanced $J \equiv J_1$ (enhanced by the presence of the oxygen holes) but also costs exchange energy $K_{\rm eff}$ between the Cu and the O hole. The third term is the energy cost to completely overlap two flipped carriers; i.e., put the oxygen holes on the same site while having both neighboring Cu's flipped Cu($A$)$\underline{\uparrow}$-O($\uparrow\downarrow$)-Cu($B$)$\underline{\downarrow}$ . Besides the obvious large Coulomb cost of this configuration (contributing to

$\Delta U_{AB}$) the flipped carriers also lose their effective exchange energy, thus an additional cost of $2K_{\text{eff}}$. The last term is the only term which is likely to be net attractive after the $\Delta U_{AB}$ correction. This is the energy savings for having nearby "nonoverlapping" flipped carriers; i.e., Cu($A$)-O-Cu($B$)-O-Cu($A$)-O-Cu($B$) ($\uparrow\downarrow\underline{\uparrow}\cdot\underline{\downarrow}\uparrow\downarrow$). This configuration saves $2J_1$. (The 2 was again reduced to 1 because of double counting.)

The last term in the Hamiltonian $h_{AB}$ is a "rehybridization" term in which there are two "rehybridization" processes: (1) two nonoverlapping flipped carriers of opposite spins rehybridized to two free carriers (also of opposite spins) [Cu($A$)-O-Cu($B$)-O-Cu($A$)-O-Cu($B$): $\downarrow\uparrow\underline{\downarrow}\cdot\underline{\uparrow}\downarrow\uparrow \leftrightarrows \downarrow\uparrow\uparrow\cdot\downarrow\downarrow\uparrow$] and (2) a free ↓ carrier and a nearest-neighbor flipped ↑ carrier rehybridized to a free ↑ carrier and a free ↓ carrier [Cu($B$)-O-Cu($A$)-O-Cu($B$): $\uparrow\downarrow\underline{\uparrow}\downarrow\uparrow \rightleftarrows \uparrow\uparrow\downarrow\downarrow\uparrow$]. In this latter process the two oxygen ↑ holes have interchanged roles and the Cu has flipped its spin. Due to indistinguishability it may seem irrelevant which hole has exchanged with the Cu; however, in $k$ space this leads to matrix elements connecting $(c_{k\uparrow},\psi_{-k\downarrow})$ with $(\psi_{k\downarrow},\psi_{-k\uparrow})$ when the free carrier unflips the Cu spin associated with the flipped carrier versus $(c_{k\uparrow},\psi_{-k\downarrow})$ connected to $(\psi_{k\uparrow},\psi_{-k\downarrow})$ when the flipped carrier unflips its associated Cu spin. Thus the former does lead to definite pairing which may or may not be net attractive. Thus we have

$$h_{AB}=h_{AB}^{(1)}+h_{AB}^{(2)}\ ,$$

where

$$h_{AB}^{(1)}=\tfrac{1}{2}J_1\sum_{l,h_1,h_2,h_3}{}' (\psi_{lAh_1\uparrow}^{\dagger}\psi_{(l+h_2)B(-h_3)\downarrow}^{\dagger}c_{lh_1A\uparrow}c_{(l+h_2)-h_3B\downarrow}+c_{(l+h_2)-h_3B\downarrow}^{\dagger}c_{lh_1A\uparrow}^{\dagger}\psi_{(l+h_2)B(-h_3)\downarrow}\psi_{lAh_1\uparrow}) \tag{7.6a}$$

for $h_1\neq h_2$ and $h_3\neq h_2$. The second rehybridization term is

$$\begin{aligned}h_{AB}^{(2)}={}&K_{\text{eff}}\sum_{l,h_1,h_2}{}' (\psi_{lAh_1\downarrow}^{\dagger}\psi_{lAh_2\uparrow}^{\dagger}\psi_{lAh_2\downarrow}c_{lh_1A\uparrow}+c_{lh_1A\uparrow}^{\dagger}\psi_{lAh_2\downarrow}^{\dagger}\psi_{lAh_2\uparrow}\psi_{lAh_1\downarrow}+\psi_{lBh_1\uparrow}^{\dagger}\psi_{lBh_2\downarrow}^{\dagger}\psi_{lBh_2\uparrow}c_{lh_1B\downarrow}+c_{lh_1B\downarrow}^{\dagger}\psi_{lBh_2\uparrow}^{\dagger}\psi_{lBh_2\downarrow}\psi_{lBh_1\uparrow})\\&+(t_{d^8}+t_{d^{10}}^{21})\sum_{l,h_1,h_2,h_3}{}' (\psi_{lAh_1\downarrow}^{\dagger}\psi_{lAh_3\uparrow}^{\dagger}\psi_{lAh_2\downarrow}c_{lh_1A\uparrow}+c_{lh_1A\uparrow}^{\dagger}\psi_{lAh_2\downarrow}^{\dagger}\psi_{lAh_3\uparrow}\psi_{lAh_1\downarrow}+\psi_{lBh_1\uparrow}^{\dagger}\psi_{lBh_3\downarrow}^{\dagger}\psi_{lBh_2\uparrow}c_{lh_1B\downarrow}\\&+c_{lh_1B\downarrow}^{\dagger}\psi_{lBh_2\uparrow}^{\dagger}\psi_{lBh_3\downarrow}\psi_{lBh_1\uparrow})\ ,\end{aligned} \tag{7.6b}$$

where in the second summation $h_1\neq h_2\neq h_3\neq h_1$ and in the first summation $h_1\neq h_2$. If we consider how the expansion of our basis set will affect the pairing interactions, we note that the inclusion of bases in which the oxygen hole has hopped away from a spin deviation causes $h_{AB}$ to have retarded pairing components. Retarded pairing resulting from $h_{AB}^{(2)}$ corresponds to the exchange of a virtual magnon between carriers of opposite spin. Retarded pairing in $h_{AB}^{(1)}$ corresponds to the annihilation of two magnons emitted by two carriers of opposite spin through the quartic term of the spin Hamiltonian.

In order to bilinearize the Hamiltonian we make a mean-field approximation, first for the "normal uncorrelated" or Hartree-Fock state ($\Psi_{\text{HF}}$), from which we obtain the "normal fully-correlated" state by a Gutzwiller-like[58] projection ($\Psi_C$). Then we rebilinearize the Hamiltonian with the superconducting mean-field approximation to obtain the "superconducting" state ($\Psi_{\text{sc}}$).

To simplify the notation we define $\psi_{lAj\uparrow}=f_{l,2j-1,\uparrow}$, $c_{ljA_\uparrow}=f_{l,2j,\uparrow}$, $\psi_{lBj\downarrow}=f_{l,2j-1,\downarrow}$, $c_{ljB_\downarrow}=f_{l,2j,\downarrow}$ for $j=1,2,3,4$. Thus the free ↑ carrier on oxygen-site $lA1$ becomes particle $l1\uparrow$, the flipped ↑ carrier which has a down hole on oxygen-site $lA1$ becomes particle $l2\uparrow$, etc. The carrier particles are then $f_{li\sigma}$ for $i=1$–8, and $\sigma=\uparrow,\downarrow$. For $i$ odd (even) particles are free (flipped) carriers.

The Hamiltonian is transformed from coordinate space to momentum space by defining the following Fourier transforms:

$$g_{kj\sigma}=\frac{1}{\sqrt{N}}\sum_{l}e^{i\mathbf{k}\cdot(l+\mathbf{r}_{j\sigma})}f_{lj\sigma}\ , \tag{7.7}$$

where $r_{j\sigma}$ are vectors from the $lA$ ($lB$) Cu site for $\sigma=\uparrow$ (↓) to the oxygen orbital occupied by the carrier; i.e., for $j=1,2$, $r_{j\uparrow}=-a_x/2$ and $r_{j\downarrow}=a_x/2$, etc.

For the normal-state uncorrelated mean-field Hamiltonian we define mean-field variables $\eta_{j_1j_2\sigma}=\langle g_{kj_1\sigma}^{\dagger}g_{kj_2\sigma}\rangle$. We obtain the equivalent of the Hartree-Fock solution in self-consistent field

$$\begin{aligned}H_{\text{MF}}^{n}=E_0^{n}+\sum_{j_1,j_2,k}&(\xi_k^{(j_1j_2)}+V_k^{(j_1j_2)})\\&\times(g_{kj_1\uparrow}^{\dagger}g_{kj_2\uparrow}+g_{-kj_1\downarrow}^{\dagger}g_{-kj_2\downarrow})\ .\end{aligned} \tag{7.8}$$

For the normal-state maximally correlated Hamiltonian we assume that

$$\langle f_{l_1j_1\sigma_1}^{\dagger}f_{l_1j_1\sigma_1}f_{l_2j_2\sigma_2}^{\dagger}f_{l_2j_2\sigma_2}\rangle=0$$

for

$$|l_1+r_{j_1}-l_2+r_{j_2\sigma_2}|<|a_1|=|a_x+a_y|=\sqrt{8}r_{\text{CuO}}\ .$$

This solution is obtained from the uncorrelated normal-state solution by a Gutzwiller-like projection[58] which projects out any configurations which give finite expectations for these operators.

For the case of the "superconducting" phase we define mean-field variables $\alpha_k^{(j_1j_2)}=\langle g_{kj_1\uparrow}g_{-kj_2\downarrow}\rangle$ with which the following bilinearized Hamiltonian results:

$$\begin{aligned}H_{\text{MF}}^{\text{sc}}=E_0^{\text{sc}}+\sum_{j_1,j_2,k}&\xi_k^{(j_1j_2)}(g_{kj_1\uparrow}^{\dagger}g_{kj_2\uparrow}+g_{-kj_1\downarrow}^{\dagger}g_{-kj_2\downarrow})\\&+\sum_{j,k}\Delta_k^{(j_1j_2)}(g_{kj_1\uparrow}^{\dagger}g_{-kj_2\downarrow}^{\dagger}+g_{-kj_2\downarrow}g_{kj_1\uparrow})\ .\end{aligned} \tag{7.9}$$

where the $\xi_k^{(j_1j_2)}$ are the single-particle Hamiltonian matrix elements as in Eq. (6.14) and $\Delta_k^{(j_1,j_2)}$ comes from Fourier transforming the terms in $V_{AB}$ and $h_{AB}$. The effect of $\Delta U_{AB}$ is included self-consistently as it adds to $E_0^{\text{sc}}$. These terms are complicated functions of $k$, $j_1$ and $j_2$, and the mean-field expectations. They can be written in closed form and certainly can be calculated. In fact the problem is solvable at this level. Since we have not yet done these calculations, and as it is not difficult to derive the self-consistent expression for $\Delta_k^{(j_1 j_2)}$ and $E_0^{\mathrm{sc}}$ (although the algebra is messy), we save these unappealing expressions and calculations for a future paper. The exact expressions as well as the precise calculations are not enlightening given the uncertainty in the energy parameters, especially since this will predict static long-range order at a finite temperature $T_c$ which, of course, is not relevant to a 2D system. Thus at this level all that we could predict with any degree of confidence is whether the 2D sheets can support a superconducting phase at $T=0$. In the next section we discuss how this result, when coupled to the third dimension, could lead to a phase transition at finite $T_c$.

To conclude this section, we show that, in general, the solution of Eq. (7.9) does exhibit a gap. (The gap itself should exhibit extended $s$-wave and $d$-wave character which is obtained from the Fourier transforms of $V_{AB}$.) The elementary excitations are the eigenoperators of the following matrix:

$$L_{k\uparrow}=\begin{pmatrix} \underline{\xi}_k & \underline{\Delta}_k \\ \underline{\Delta}_k & -\underline{\xi}_k \end{pmatrix},$$

where $\underline{\xi}_k$ and $\underline{\Delta}_k$ are $8\times8$ matrices, thus $L_{k\uparrow}$ is $16\times16$. This is a matrix in the operator space of

$$\{g_{k1\uparrow},g_{k2\uparrow},\ldots,g_{k8\uparrow},g^{\dagger}_{-k1\downarrow},g^{\dagger}_{-k2\downarrow},\ldots,g^{\dagger}_{-k8\downarrow}\}\,.$$

We first diagonalize $\underline{\xi}_k$ which is the quasiparticle Hamiltonian and apply this transformation ($\underline{T}_k$) to $\underline{\Delta}_k$ ($\tilde{\underline{\Delta}}_k=\underline{T}_k^{\dagger}\underline{\Delta}_k\underline{T}$). Then the Fermi level $\mu$ and the Fermi momentum ($\varepsilon_{k_F}-\xi=\mu$) are set by particle conservation. Thus, if $[\tilde{\Delta}_k]_{11}\ll(\varepsilon_{k_F1}-\varepsilon_{k_F2})$ then the BCS spectrum near $k_F$ results from quasidegenerate perturbation theory; i.e., $E_{k1}^2=(\varepsilon_{k1}-\xi)^2+[\tilde{\Delta}_k]_{11}^2$. Relaxing the restriction $[\tilde{\Delta}_k]_{11}\ll\varepsilon_{k_F1}-\varepsilon_{k_F2}$ does not, in general, close the gap.

## VIII. THE THIRD DIMENSION

Until now this entire paper has focused only on the two-dimensional sheets common to all of the known high-$T_c$ superconductors. In this section we discuss the role of the third dimension. Common to each of the superconductors is one or two oxygens coordinated along the $z$ axis to each Cu in a sheet. Thus free carriers and flipped carriers can delocalize to some extent onto the fifth (and sixth in $La_{2-x}Sr_xCuO_4$) coordinated oxygen. In addition, the pairing term arising from a free and flipped carriers on neighboring sites includes the configuration when one or the other is on this extra oxygen. This in itself, however, cannot turn 2D superconductivity (*if* it exists only at $T=0$) into full superconductivity. If superconductivity in the 2D $CuO_2$ sheets maps onto the $XY$ Hamiltonian (i.e., with a two-component order parameter), then superconductivity can exist at finite $T$ via a Kosterlitz-Thouless phase transition.[5,59] We recognize this, but we submit that without knowing the mechanism this mapping has not been proven. Furthermore, there is some evidence that the phase transition is three dimensional[60] and may have an order parameter with greater than two components. We consider here what couples the sheets to each other which may drive a three-dimensional phase transition. Several possibilities come to mind, one of which is both elegant and appealing.

We can liken these highly anisotropic materials to weakly coupled superconductors in which the 2D sheets are separated by insulating material; i.e., sc-$I$-sc-$I$-sc-$I$-sc-$I\cdots$. Then as long as charge can tunnel across the insulating material long-range phase coherence will be established by the same mechanism by which current flows and superconductivity is maintained through an insulating barrier (typically 20 Å thick) coupling two superconductors; i.e., the proximity effect. We reproduce here a simple model (the perturbation theory method or the "Cooper-pair model"[61] illustrating this effect. We generalize to $N+1$ superconductors coupled by $N$ insulating barriers. We define a matrix element $W$ which is the transfer integral for transferring one Cooper pair from the $j$th superconductor to the $j+1$st superconductor. Next we define $\Psi_{nj}$ which is obtained from $\Psi_{0j}$ by transferring $n$ Cooper pairs between the $j$ and $j+1$st superconductor where $\Psi_{0j}$ is the unperturbed eigenstate for the $j$th superconductor. Assuming that all the superconducting sheets are at the same potential, there is no energy cost in transfering Cooper pairs between sheets, hence all $\Psi_{nj}$ are degenerate. We assume that the coupling is weak, hence only one Cooper pair is transferred at a time and only between neighboring sheets. This is equivalent to $N$ uncoupled linear chains with only nearest-neighbor interactions. Thus the eigenstates are

$$\Phi_{\{\alpha_j\}}=\prod_j\sum_{n=-\infty,\infty}\exp(in\alpha_j)\Psi_{nj}\,, \tag{8.1}$$

where each $\alpha_j$ is a continuous variable from 0 to $2\pi$. The corresponding eigenvalues are

$$\varepsilon_{\{\alpha_j\}}=-2W\sum_j\cos\alpha_j\,, \tag{8.2}$$

where we have taken the energy zero to be that of each state in the unperturbed system. Now we need to interpret the meaning of $\alpha_j$. In the case of two ($N=1$) coupled superconductors $\alpha_1$ can be identified with the phase difference between the two superconductors.[61] Thus the total energy of the system is lowered by $2W$ when the phase difference vanishes. In this case $\alpha_j$ can be similarly identified with $\phi_{j+1}-\phi_j$, and thus the lowest-energy state has phase coherence between all the sheets. This is the origin of static long-range order.

For a single superconductor, the phase of the order parameter and the number of Cooper pairs are conjugate variables. The superconducting state has a fixed value for the phase, hence for the $j$th superconductor we have

$$\chi_{\phi_j}=\sum_{\nu=-\infty,\infty}\exp(i\nu\phi_j)\psi_{\nu j}\,, \tag{8.3}$$

where $\psi_{0j}$ is an eigenstate of the unperturbed system with a definite number of carriers and $\psi_{\nu j}$ is obtained from $\psi_{0j}$ by adding $\nu$ Cooper pairs. The wave function for $N+1$

such isolated superconductors is then

$$\chi=\prod_j \chi_{\phi_j}=\sum_{\nu_1}\sum_{\nu_2}\cdots\sum_{\nu_{N+1}}\exp\left(i\sum_j \nu_j\phi_j\right)\prod_i \Psi_{\nu_i i}\,. \tag{8.4}$$

We now rewrite this as

$$\chi=\sum_M \exp\left(iM\sum_j \phi_j\right)\chi_M\,, \tag{8.5}$$

where $\chi_M$ is a state which has a definite number of Cooper pairs, a total of $M(N+1)$ in the $N+1$ superconductors

$$\chi_M=\sum_{n_1,n_2,\ldots,n_N}\exp\left(i\sum_{j=1,N} n_j(\phi_{j+1}-\phi_j)\right)\times\psi_{M-n_1,1}\psi_{M-n_2+n_1,2}\cdots\psi_{M-n_N,N+1}\,. \tag{8.6}$$

This, however, has the form of $\Phi_{\{a_j\}}$ since

$$\psi_{M-n_1,1}\psi_{M-n_2+n_1,2}\cdots\psi_{M-n_N,N+1}=\Psi_{n_1}\Psi_{n_2}\cdots\Psi_{n_N}\,.$$

That is, it is obtained from $\psi_{M,1}\psi_{M,2}\cdots\psi_{M,N+1}$ by transferring $n_1$ Cooper pairs from layer 1 to layer 2, $n_2$ from layer 2 to layer 3, $n_j$ from layer $j$ to $j+1$, and finally $n_N$ from layer $N$ to $N+1$. Thus we see that, indeed we can identify $a_j$ with $\phi_{j+1}-\phi_j$. We include this to show that we do not need pairing in the third dimension to have superconductivity, but because of the 2D sheets we may need this 3D coupling energy to drive the phase transition. From here we argue by analogy the role of $W$ in determining $T_c$ much as we expect the out-of-plane spin coupling to drive the Néel transition in the nonsuperconducting relatives.

To our knowledge it has not been rigorously shown that the Néel transition temperature in the nonsuperconducting relatives of the Cu-O perovskites satisfies the following heuristic[44,62] self-consistent equation; nonetheless, we will use it for argument's sake

$$\frac{T_N}{J_{xy}}\sim\frac{J_z}{J_{xy}}\xi^2\left(\frac{J_{xy}}{T_N}\right), \tag{8.7}$$

where $a\xi(J_{xy}/T)$ is the 2D correlation length ($a$ is the lattice constant), that is

$$\langle S(r_1)\cdot S(r_2)\rangle\propto\pm\frac{\xi}{|r_1-r_2|}\exp(-|r_1-r_2|/\xi)\,,$$

where $+$ ($-$) is for site $r_2$ on the same (opposite) sublattice as site $r_1$. $J_{xy}$ is the superexchange coupling in the 2D sheet and $J_z$ is the exchange coupling between the sheets. There have been several derivations[46–48] for $\xi(J_{xy}/T)$ which differ in method and detail, but basically agree with each other with the one exception which is a fit to simulations.[49] These differences, however, are not significant providing that the transition is driven by the third dimension as opposed to following parasitically. Since $T_N$ (Refs. 37 and 38) is greater while $J_{xy}$ (Refs. 40, 41, and 63) is smaller in $YBa_2Cu_3O_6$ than in $La_2CuO_4$, this is apparently the case. The reason that the static long-range order does not exist in 2D is that the appropriate susceptibility diverges as $1/k^2$. Thus the integral over $k$ would diverge logarithmically in two dimensions if the order parameter does not vanish.[6] For similar reasons a nonvanishing order parameter in a superconductor will not exist in 2D except again at $T=0$. Thus we expect long-range phase coherence which goes exponentially in $\Delta/T$ [if not Kosterlitz-Thouless,[6] $A\exp(\Delta/|T-T_0|^a)$] and we expect that the static long-range order is established in an analogous fashion as the antiferromagnetic phase transition. As yet, we have no theory for $\xi(\Delta/T)$ in the superconductor where $\xi$ is the number of superconducting pairs in phase coherence. This number is macroscopic when the system is a bulk superconductor. This does not affect our hypothesis that $T_c$ will be determined by a self-consistent equation similar to Eq. (8.7).

## IX. DISCUSSION AND CONCLUSION

We have seen that a central thesis of this work is that poor screening exists in a CuO$_2$ sheet due to the nearly-filled-shell nature of the Cu and O ions and that this leads to significant intra-atomic *and* interatomic screened Coulomb interactions between quasiparticles. There is evidence for this from both experimental measurements (Auger spectroscopy) and *ab initio* calculations of Schlüter *et al.*,[27] Harmon and co-workers,[28] and Mc Mahan *et al.*[29] However, further indirect support comes from both neutron[15,18,39] and polarized light scattering[40,41,45] in that both observe a Heisenberg spin system to exist in the superconducting materials (i.e., with charge carriers), a fact which we believe to be a direct consequence of poor screening and a substantial interatomic $U$ as explained in Sec. III. Cluster calculations in which correlation is included have also reached this conclusion.[19] The validity of a Heisenberg description of the spin system is supported by our semiempirical configuration-interaction calculations for spin states of a small cluster (reported in Sec. IV) as well as by *ab initio* calculations of Martin and Saxe.[63] Although the values of the interatomic $U$ are not known with any degree of confidence, we have suggested reasonable upper and lower bounds and found that only quantitative details change between these limits. The importance of interatomic $U$ has also been recognized in the work of others.[14,27–29,64] Thus we believe that our Hamiltonian is realistic since it is based on energy parameters which were derived from a variety of semiempirical and *ab initio* sources, and which form a self-consistent set as seen in Sec. IV. The relatively large values of the $U$ suggest that the theory[65] of Schrieffer, Wen, and Zhang may not be applicable to these materials. In addition, numerical simulations by Hirsch and co-workers[66] have suggested that the simplest (single-band) Hubbard Hamiltonians are incapable of explaining superconductivity. Several workers are realistically computing[27–29] or semiempirically deriving from photoemission measurements[51] the values of the energy parameters. This will provide quantitative certainty. However, it is very unlikely that the qualitative description of the carrier quasiparticle and the pairing interactions which we have presented here will be affected by refinement of the parameters. Zaanen and Oles[67] and Shen *et al.*[51] have argued that these materials represent charge-transfer (as opposed to Mott-Hubbard) insulators and mixed valence metals when doped. Furthermore, Zaanen and Oles[67] have pointed out that the presence of a substantial oxygen-oxygen transfer integral $t_{pp}$ precludes a transformation from a realistic Hamiltonian to a simple (nearest-neighbor hopping) one-band Hubbard model which is possible[68] in the absence of $t_{pp}$. This then casts doubt on the applicability of resonating valence-band (RVB) theory.[69,70] We support these conclusions of Zaanen and Oles.[67]

The strong interactions which exist between the carriers and the spin system are only due in part to the large value of the Cu-O transfer integral suggested by the band-structure calculations.[30,36] In Sec. III we found that significant valence fluctuations occur between the Cu $d^9$ and $d^{10}$ configurations. Valence fluctuations have also been recognized to be important in the work of Newns[71] and $d^{10}$ assisted transport are the $1/\Delta$ processes in the work of Zaanen and Oles.[67] It is important to recognize, however, that these fluctuations are increased by the presence of a carrier due to the Coulomb interactions which lower the relative energy of the virtual state. This decreases the energy of a carrier hole on an oxygen $p_\sigma$ orbital (along the Cu-O axis) by significantly increasing the probability of a carrier making spin deviations in the local antiferromagnetic order. This causes the oxygen orbital, which takes the carrier hole, to be $p_\sigma$ and not $p_\pi$, which we compute would be preferred by $\sim 0.8$ eV if one only considered the Madelung potential. These results are in contradiction to the conclusions of Goddard's group[19] and the assumptions of Birgeneau *et al.*[20] We find that the near degeneracy of the $p^6d^9p^5$ and $p^5d^{10}p^5$ configurations causes a dramatic stabilization of a $p_\sigma$ hole. This leads to important consequences for carrier-spin interactions since a $p_\sigma$ carrier interacts strongly, whereas a $p_\pi$ hole interacts weakly;[19] experimental evidence supports strong interactions.[39] On the other hand, the proximity in energy of the $p_\pi$ orbitals and the magnitude of the matrix element which connects the $p_\sigma$ to the $p_\pi$ between sites (which we estimate to be $\sim 0.3$ eV) necessarily produces some $p_\pi$ admixture in the carrier at $k \neq 0$ or $k \neq X$ in conventional notation[30] (at $k = 0$ the $p_\sigma$-$p_\pi$ interaction vanishes). Its effect is to dilute the strength of the carrier-spin interactions; the degree to which this occurs is the subject of current work.

Other workers have suggested spin-driven pairing in these materials such as in the early work of Emery,[11] Hirsch,[12] and Parmenter.[13] However, we saw (Sec. VII) that pairing is not driven by increased mobility,[12] a concept which also has been disputed by Trugman;[72] the charge carriers have high mobility even without the channels which produce spin deviations. (The original Emery model[11] has also been disputed.[67]) Strong-coupling arises because the configurations which involve spin deviations (the "flipped carriers") and those which do not (the "free holes") are close in energy relative to the magnitude of the matrix elements which couple them; thus they must mix strongly and nonperturbatively. This remarkable fact, represented by the model Hamiltonian matrix in Eq. (6.19), means that localized spin polarons do not form in these materials in that the carriers do not carry a spin deviation along with them as they move; instead, the carriers move as free holes since they can easily self-heal an associated spin deviation. In addition, the traditional view of a spin polaron is that of spin deviations which are stabilized by a preferred exchange coupling between the carriers and the local spins. We have found (Sec. IV) that this energy is quite small. Thus the carriers are much better described by the concept of the spin hybrid, formalized in Sec. VII.

When we state that the carriers have "high mobility" even though there is a large probability of being associated with a spin deviation (which thus increases the strength of spin-driven pairing), we are speaking locally and not considering the effects of scattering off spin deviations (produced by zero-point fluctuations, other carriers, or domain boundaries) or off Coulomb interactions with each other. All these processes are substantial and significantly reduce the mean free path of a carrier; thus these materials may be in the dirty superconductor limit.[73]

The spin-hybrid carriers do not permit the existence of boson entities because the carriers move as spin-$\frac{1}{2}$ free holes. This causes superconducting pairing in the Cooper sense,[8] as has been concluded in the Ginzberg-Landau analysis of Bardeen, Ginsberg, and Salamon,[74] and is in contrast to the spinless (boson) holes of Kivelson, Rokhsar, and Sethna,[70] which is a consequence of the RVB picture of Anderson.[69]

The exact nature of pairing is much less certain. There are obviously a variety of retarded and unretarded pairing interactions of which we cannot judge the relative importance without quantitative detail. The simplest interaction to understand is the unretarded (direct) one which has been suggested as an obvious possibility by us and others.[18] We find that this interaction is indeed expected from the existence of spin-hybrid carriers. The energy of this interaction is certainly larger than $2J \sim 0.06$–$0.12$ eV, while the Coulomb correction is less certain because of the unknown variation of $U$ with distance. Recently, Emery has suggested that the pairing is unretarded based on an analysis of muon-spin resonance data.[75]

The importance of Josephson tunneling in the determination of $T_c$ is an appealing idea because of the formal similarity with the Néel transition now being studied in detail.[48] Obviously, an important next step is to compute the tunneling rate between $CuO_2$ sheets in these materials. These calculations are underway.

We feel that the most important new result contained in this paper is the development of the spin-hybrid description of the carrier quasiparticles. Given the existence of the Cu spins and the strong coupling between the carrier and spin systems, we believe the existence of the spin hybrid to be inevitable. This in turn necessarily leads to spin-driven pairing. While the strength of the pairing and the nature of the phase transition remains to be determined, the spin-hybrid model forms a basic foundation upon which to build.

## APPENDIX A: EFFECTIVE TRANSFER AND EXCHANGE-TRANSPORT INTEGRALS

In this appendix we show how we estimate the effective transfer integrals $t^{2}_{d^{10}}$ and $t_{d^8}$ as defined in Sec. VI. We first consider the eigenstates and eigenvalues of two holes in three orbitals as we did in Sec. IV to estimate the superexchange energy $J$. However, now we examine two oxygen orbitals nearest neighbor to one Cu orbital. There are two possibilities, the oxygens along the $\mathbf{a}_1$ (or $\mathbf{a}_2$) axis (referred to below as bent O-Cu-O) and along the $\mathbf{a}_x$ (or $\mathbf{a}_y$) axes (referred to as linear). In the latter case the oxygens are not coupled by direct hopping, $t_{pp}$; in the former case they are. Again there are nine spin-orbital bases for which three are triplets, six are singlets. The basis for the triplet matrix is $\phi_1^T = 1/\sqrt{2}(p_1 d - d p_1)$, $\phi_2^T = 1/\sqrt{2}(p_1 p_2 - p_2 p_1)$, and $\phi_3^T = 1/\sqrt{2}(p_2 d - d p_2)$. The Hamiltonian matrix in this basis is

$$H^T = \begin{pmatrix} K & -t_{pd} & -t_{pp} \\ -t_{pd} & \Delta E & t_{pd} \\ -t_{pp} & t_{pd} & K \end{pmatrix} .$$

The unitary transformation which takes the antibonding ("$k=0$") $\chi_1^T = 1/\sqrt{2}(\phi_1^T + \phi_3^T)$ and bonding ("$k=\pi$") combinations $\chi_3^T = 1/\sqrt{2}(\phi_1^T - \phi_3^T)$ shows that only $\chi_3^T$ interacts with the virtual state $\phi_2^T$. The lowest two eigenvalues are

$$E_T(k=0) = -t_{pp} + K ,$$

$$E_T(k=\pi) = t_{pp} - t_{d^{10}}^T(k=\pi) + K ,$$

where

$$t_{d^{10}}^T(k=\pi) = \sqrt{2}t_{pd} \tan\tfrac{1}{2}\tan^{-1}\left[\frac{\sqrt{8}t_{pd}}{\Delta E - K - t_{pp}}\right] .$$

The linear O-Cu-O triplet Hamiltonian and eigenvalues are obtained by simply replacing $t_{pp}$ by zero. Note also that hole energies are the negative of these eigenvalues.

The basis for the singlet matrix is $\phi_1^S = 1/\sqrt{2}(p_1 d + d p_1)$, $\phi_2^S = 1/\sqrt{2}(p_1 p_2 + p_2 p_1)$, $\phi_3^S = 1/\sqrt{2}(p_2 d + d p_2)$, $\phi_4^S = dd$, $\phi_5^S = 1/\sqrt{2}(p_1 p_1 + p_2 p_2)$, and $\phi_6^S = 1/\sqrt{2}(p_1 p_1 - p_2 p_2)$. The Hamiltonian matrix in this basis is

$$H^S = \begin{pmatrix} -K & -t_{pd} & -t_{pp} & -\sqrt{2}t_{pd} & -t_{pd} & -t_{pd} \\ -t_{pd} & \Delta E & -t_{pd} & 0 & -t_{pp} & 0 \\ -t_{pp} & -t_{pd} & -K & -\sqrt{2}t_{pd} & -t_{pd} & -t_{pd} \\ -\sqrt{2}t_{pd} & 0 & -\sqrt{2}t_{pd} & \Delta U_d & 0 & 0 \\ -t_{pd} & -t_{pp} & -t_{pd} & 0 & \Delta U_p & 0 \\ -t_{pd} & 0 & -t_{pd} & 0 & 0 & \Delta U_p \end{pmatrix} .$$

The unitary transformation which takes the antibonding ($k=0$) $\chi_1^S = 1/\sqrt{2}(\phi_1^S + \phi_3^S)$ and bonding ($k=\pi$) combinations $\chi_3^S = 1/\sqrt{2}(\phi_1^S - \phi_3^S)$ shows now that only $\chi_1^S$ interacts with the virtual states ($\phi_2^S, \phi_4^S, \phi_5^S$). The lowest two eigenvalues are

$$E_S(k=0) = -t_{pp} - t_{d^{10}}^S(k=0) - t_{d^8}^S(k=0) - t_{p^4}^S(k=0) - K ,$$

$$E_S(k=\pi) = t_{pp} - K ,$$

where

$$t_{d^{10}}^S(k=0) \approx \sqrt{2}t_{pd} \tan\tfrac{1}{2}\tan^{-1}\left[\frac{\sqrt{8}t_{pd}}{\Delta E + K + t_{pp}}\right] ,$$

$$t_{d^8}^S(k=0) \approx 2t_{pd} \tan\tfrac{1}{2}\tan^{-1}\left[\frac{4t_{pd}}{\Delta U_d + K + t_{pp}}\right] ,$$

$$t_{p^4}^S(k=0) \approx t_{pd} \tan\tfrac{1}{2}\tan^{-1}\left[\frac{2t_{pd}}{\Delta U_p + K + t_{pp}}\right] .$$

We now rewrite these energies (we drop the $p^4$ since it is a very weak term and does not add to the physics) in the following form:

$$\begin{aligned} E_S(k=0) = &-[t_{d^{10}}^{12}(k=0) + t_{d^{10}}^{12}(k=\pi) + t_{d^8}^{12}(k=0) + t_{d^8}^{12}(k=\pi)] - t_{pp} - [t_{d^{10}}^{12}(k=0) - t_{d^{10}}^{12}(k=\pi)] \\ &- [t_{d^8}^{12}(k=0) - t_{d^8}^{12}(k=\pi)] + t_{d^{10}}^{21}(k=0) + t_{d^8}^{21}(k=0) , \\ E_S(k=\pi) = &-[t_{d^{10}}^{12}(k=0) + t_{d^{10}}^{12}(k=\pi) + t_{d^8}^{12}(k=0) + t_{d^8}^{12}(k=\pi)] + t_{pp} - [t_{d^{10}}^{12}(k=\pi) - t_{d^{10}}^{12}(k=0)] \\ &- [t_{d^8}^{12}(k=\pi) - t_{d^8}^{12}(k=0)] + t_{d^{10}}^{21}(k=\pi) + t_{d^8}^{21}(k=\pi) , \\ E_T(k=0) = &-[t_{d^{10}}^{12}(k=0) + t_{d^{10}}^{12}(k=\pi) + t_{d^8}^{12}(k=0) + t_{d^8}^{12}(k=\pi)] - t_{pp} - [t_{d^{10}}^{12}(k=0) - t_{d^{10}}^{12}(k=\pi)] \\ &- [t_{d^8}^{12}(k=0) - t_{d^8}^{12}(k=\pi)] - t_{d^{10}}^{21}(k=0) - t_{d^8}^{21}(k=0) , \\ E_T(k=\pi) = &-[t_{d^{10}}^{12}(k=0) + t_{d^{10}}^{12}(k=\pi) + t_{d^8}^{12}(k=0) + t_{d^8}^{12}(k=\pi)] + t_{pp} - [t_{d^{10}}^{12}(k=\pi) - t_{d^{10}}^{12}(k=0)] \\ &- [t_{d^8}^{12}(k=\pi) - t_{d^8}^{12}(k=0)] - t_{d^{10}}^{21}(k=\pi) - t_{d^8}^{21}(k=\pi) . \end{aligned}$$

Using these results, we now define an effective Hamiltonian such that $\underline{U}\,\underline{E}\,\underline{U}^T = \underline{H}_{\text{eff}}$ with $\underline{U}$ the unitary transformation to singlet-antibonding, singlet-bonding, triplet-antibonding, and triplet-bonding combinations of two holes in two oxygen orbitals, and $\underline{E}$ is a diagonal matrix of the four lowest eigenvalues of the coupled system, thus

$$\underline{U} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix}$$

and $\underline{E}$ is

$$\underline{E}=\begin{bmatrix} E_S(k=0) & 0 & 0 & 0 \\ 0 & E_S(k=\pi) & 0 & 0 \\ 0 & 0 & E_T(k=0) & 0 \\ 0 & 0 & 0 & E_T(k=\pi) \end{bmatrix} .$$

Then we find that the diagonal energy (the diagonal elements of $\underline{H}_{\mathrm{eff}}$) is

$$[H_{\mathrm{eff}}]_{jj}=\tfrac{1}{4}[E_S(k=0)+E_S(k=\pi)+E_T(k=0)+E_T(k=\pi)]=-\tfrac{1}{4}[t_{d^{10}}^{S}(k=0)+t_{d^{8}}^{S}(k=0)+t_{d^{10}}^{T}(k=\pi)] .$$

The 12 (and 34) element which represents transport and no exchange is

$$[H_{\mathrm{eff}}]_{12}=\tfrac{1}{4}[E_S(k=\pi)+E_T(k=0)-E_S(k=0)-E_T(k=\pi)]$$
$$=-t_{pp}-\tfrac{1}{4}[t_{d^{10}}^{S}(k=0)-t_{d^{10}}^{T}(k=\pi)-t_{d^{8}}^{S}(k=0)]\approx-t_{pp}-\tfrac{1}{4}t_{d^{8}}^{S}(k=0) ,$$

which is reasonable; i.e., there is no transport without exchange through $t_{d^{10}}$ but there is transport without exchange through $t_{d^{8}}$ and of course $t_{pp}$. There is some interference due to $t_{d^{10}}$ fluctuations but it is not significant.

The 13 (and 24) element which represents pure exchange is

$$[H_{\mathrm{eff}}]_{13}=\tfrac{1}{4}[E_T(k=0)+E_T(k=\pi)-E_S(k=0)-E_T(k=\pi)]$$
$$=K+\tfrac{1}{4}[t_{d^{10}}^{S}(k=0)-t_{d^{10}}^{T}(k=\pi)+t_{d^{8}}^{S}(k=0)]\approx K+\tfrac{1}{4}t_{d^{8}}^{S}\approx K_{\mathrm{eff}} .$$

Recall that there are also some $p^4$ contributions to $K_{\mathrm{eff}}$ that have been dropped. See. Sec. IV for the more-precise determination of $K_{\mathrm{eff}}$ from three holes in three sites Cu-O-Cu.

Finally, the 14 (and 23) element which represents simultaneous transport and exchange is

$$[H_{\mathrm{eff}}]_{14}=\tfrac{1}{4}[E_T(k=0)-E_S(k=0)-E_T(k=\pi)+E_S(k=\pi)]+\tfrac{1}{4}[t_{d^{10}}^{S}(k=0)+t_{d^{10}}^{T}(k=\pi)+t_{d^{8}}^{S}(k=0)] .$$

From these results we now identify

$$t_{d^{10}}^{21}\equiv\tfrac{1}{4}[t_{d^{10}}^{S}(k=0)+t_{d^{10}}^{T}(k=\pi)]$$

and

$$t_{d^{8}}\equiv\tfrac{1}{4}t_{d^{8}}^{S}(k=0) .$$

Calculations yield $t_{d^{10}}^{21}\approx0.56$ eV for $\Delta E=1$ eV, 0.48 eV for $\Delta E=1.5$ eV, 0.42 eV for $\Delta E=2$ eV, and finally 0.36 eV for $\Delta E=2.5$ eV. The triplet bonding ($k=\pi$) is lower in energy than the triplet antibonding ($k=0$) for $\Delta E$ up to about 1.6 eV, thus actually it is a close call. Since this $t_{d^{10}}^{21}$ averages $t_{d^{10}}^{21}$ from the singlet antibonding with that from the triplet bonding, we get a lower crossover energy at about 1 eV. Whatever the case the triplet bonding and the triplet antibonding combinations are very nearly degenerate for the relative phase of two oxygens along the $\mathbf{a}_1$ or $\mathbf{a}_2$ axes. Calculations yield $t_{d^{8}}=0.063$ eV for $\Delta E=1$ eV ($\Delta U=5$ eV), 0.072 eV for $\Delta E=1.5$ eV ($\Delta U=4.5$ eV), 0.084 eV for $\Delta E=2$ eV ($\Delta U=4$ eV), and finally 0.096 eV for $\Delta E=2.5$ eV ($\Delta U=3.5$ eV).

## APPENDIX B: EXTENSION OF SPIN-HYBRID MODEL

In this appendix we illustrate the expansion towards a more-detailed model Hamiltonian from which the complete hybridization of the carrier quasiparticle may be obtained using our parameters from Secs. II and IV. We first consider only six local spin configurations. From here the generalization to include still more local spin configurations will be obvious. In addition to the above free-carrier ($\downarrow\uparrow\uparrow$) and flipped-carrier ($\underline{\uparrow}\downarrow\uparrow$) configurations, we consider four others, $d^9_\uparrow p^5_\uparrow d^9_\downarrow$, or ($\underline{\uparrow}\uparrow\underline{\downarrow}$) in shorthand, and $d^9_\uparrow p^6 d^9_\downarrow p^5_\uparrow$ ($\underline{\uparrow}\cdot\underline{\downarrow}\uparrow$) where the $p^5_\uparrow$ is this latter configuration may be any one of the three oxygen orbitals neighboring the $d^9_\downarrow$ but not neighboring the $d^9_\uparrow$. We call these configurations the "double-flipped-carrier " as it represents the exchange of the two Cu spins (two spin deviations) in association with a carrier hole. We will see, however, that the double-flipped-carrier components are weak relative to the free-carrier and the flipped-carrier components; however, they do serve to slightly enhance the flipped-carrier component.

The energy of the double-flipped-carrier with the carrier hole between the exchange Cu spins is $-3J$, since there are six unfavorable Cu-Cu spin pairs. The remaining three double-flipped-carrier configurations have energy $K_{\mathrm{eff}}-2.5J$ since the carrier hole is not blocking the favorable exchange between the two flipped Cu's and is instead blocking one of the unfavorable spin pairs created by the spin flip. In terms of these six configurations, the submatrices are 6×6. (For the purposes here we take the simplifying case of $\gamma=0$). Thus

$$h_{11}=\begin{pmatrix} 3J & -K_{\rm eff} & 0 & -J & -J & -J \\ -K_{\rm eff} & K_{\rm eff} & -K_{\rm eff} & 0 & 0 & 0 \\ 0 & -K_{\rm eff} & -3J & 0 & 0 & 0 \\ -J & 0 & 0 & -\frac{5}{2}J+K_{\rm eff} & 0 & 0 \\ -J & 0 & 0 & 0 & -\frac{5}{2}J+K_{\rm eff} & 0 \\ -J & 0 & 0 & 0 & 0 & -\frac{5}{2}J+K_{\rm eff} \end{pmatrix} .$$

Recall from Sec. VI, Eq. (6.16) that

$$h_{12}+h_{14}=2B_0+2B_1c_{k_x}c_{k_y}+i2B_2s_{k_x}c_{k_y} ,$$

$$h_{12}-h_{14}=-2B_1s_{k_x}s_{k_y}+i2B_2c_{k_x}s_{k_y} ,$$

$$h_{13}\pm h_{24}=(C_0\pm C_0)+C_1(c_{k_{2x}}\pm c_{k_{2y}})+iC_2(s_{k_{2x}}\pm s_{k_{2y}}) ,$$

where $B_j$ and $C_j$ are constant matrices with dimension the number of configurations. For notational convenience we define $u\equiv\frac{1}{2}(t_{d^{10}}^{21}+t_{d^8})$, $v\equiv t_{pp}+\frac{1}{2}t_{d^8}$, $w\equiv\frac{1}{2}t_{d^8}$, $x\equiv 2t_{pp}+\frac{1}{2}(t_{d^8}-t_{d^{10}}^{21})$, and $y=\frac{1}{2}(t_{d^8}-t_{d^{10}}^{21})$.

$$B_0=\begin{pmatrix} u & -u & 0 & 0 & 0 & 0 \\ -u & w & 0 & -u & 0 & -u \\ 0 & 0 & 0 & w & 0 & w \\ 0 & -u & w & 0 & w & 0 \\ 0 & 0 & 0 & w & 0 & w \\ 0 & -u & 0 & 0 & w & 0 \end{pmatrix} ,\quad B_1=\begin{pmatrix} x & -u & 0 & 0 & 0 & 0 \\ -u & v & 0 & -u & 0 & -u \\ 0 & 0 & 0 & v & 0 & v \\ 0 & -u & v & 0 & v & 0 \\ 0 & 0 & 0 & v & 0 & v \\ 0 & -u & v & 0 & v & 0 \end{pmatrix} ,\quad B_2=\begin{pmatrix} u & -u & 0 & 0 & 0 & 0 \\ -u & v & 0 & -u & 0 & -u \\ 0 & 0 & 0 & v & 0 & v \\ 0 & -u & v & 0 & v & 0 \\ 0 & 0 & 0 & v & 0 & v \\ 0 & -u & v & 0 & v & 0 \end{pmatrix} ,$$

$$C_0=\begin{pmatrix} u & -u & 0 & 0 & 0 & 0 \\ -u & w & 0 & 0 & -u & 0 \\ 0 & 0 & 0 & 0 & w & 0 \\ 0 & 0 & 0 & 0 & 0 & w \\ 0 & -u & w & 0 & 0 & 0 \\ 0 & 0 & 0 & w & 0 & 0 \end{pmatrix} ,\quad C_1=\begin{pmatrix} y & -u & 0 & 0 & 0 & 0 \\ -u & w & 0 & 0 & -u & 0 \\ 0 & 0 & 0 & 0 & w & 0 \\ 0 & 0 & 0 & 0 & 0 & w \\ 0 & -u & w & 0 & 0 & 0 \\ 0 & 0 & 0 & w & 0 & 0 \end{pmatrix} ,\quad C_2=\begin{pmatrix} u & -u & 0 & 0 & 0 & 0 \\ -u & w & 0 & 0 & -u & 0 \\ 0 & 0 & 0 & 0 & w & 0 \\ 0 & 0 & 0 & 0 & 0 & w \\ 0 & -u & w & 0 & 0 & 0 \\ 0 & 0 & 0 & w & 0 & 0 \end{pmatrix} .$$

It is not difficult to include a finite $\gamma$; however, the additional complexity does not add anything to the physics. Thus for $k=0$ we have

$$H_0=\begin{pmatrix} 3J+4t_{pp}+3t_{d^8} & -K_{\rm eff}-6u & 0 & -J & -J & -J \\ -K_{\rm eff}-6u & K_{\rm eff}+3t_{d^8}+2t_{pp} & -K_{\rm eff} & -4u & -2u & -4u \\ 0 & -K_{\rm eff} & -3J & 2(t_{pp}+t_{d^8}) & t_{d^8} & 2(t_{pp}+t_{d^8}) \\ -J & -4u & 2(t_{pp}+t_{d^8}) & -\frac{5}{2}J+K_{\rm eff} & 2(t_{pp}+t_{d^8}) & t_{d^8} \\ -J & -2u & t_{d^8} & 2(t_{pp}+t_{d^8}) & -\frac{5}{2}J+K_{\rm eff} & 2(t_{pp}+t_{d^8}) \\ -J & -4u & 2(t_{pp}+t_{d^8}) & t_{d^8} & 2(t_{pp}+t_{d^8}) & -\frac{5}{2}J+K_{\rm eff} \end{pmatrix} .$$

We now diagonalize this matrix to find the spin hybrid given a choice of model parameters consistent with those from the earlier sections ($t_{pp}=0.53$ eV, $J=0.06$ eV, $K_{\rm eff}=0.1$ eV, $t_{d^8}=0.06$ eV, $t_{d^{10}}^{21}=0.56$ eV); we find only 36% of the lowest-energy band is free-carrier configuration, 41% is flipped-carrier configuration and 2.6% is double-flipped-carrier with the carrier hole between the exchange Cu's, and finally 20.5% is the remaining double-flipped-carrier configurations. Use of a finite $\gamma=0.1$ makes about a 1% change in the results.

Other configurations which we have not yet discussed include spin deviations $d_\sigma$ ($\sigma=\uparrow,\downarrow$) that are removed in real space from the associated carrier at $p_\tau$ ($\tau=\downarrow,\uparrow$). Although carriers can create spin deviations by direct and indirect exchange, as has been pointed out by others,[11,12] the carrier will not just leave, because to do so would raise the energy of the system until such a time that a created spin flip on sublattice $A$ can find and annihilate a spin flip on sublattice $B$, hence restoring order to the lattice. Another process that can restore order to the lattice is creating a spin flip on sublattice $A$ and later annihilating a spin flip on sublattice $B$, since spin flips exist from zero-point fluctuations. These two processes of returning the system to an equal energy configuration compete with retarded pairing, since a third process which can restore order to the lattice is one carrier creating the spin flip and a second carrier annihilating the spin flip at some later time. These first two processes are difficult to include in the theory. Including configurations with the hole removed from the spin deviation will increase both the flipped and double-flipped components. However, we do not expect configurations in which the hole is greatly removed from the spin deviation to contribute substantially. Preliminary calculations found the component of the free-carrier configuration dropping to 30%, the flipped-carrier component increasing to 46%, and the double-flipped-carrier configurations to a total of 24%. A future publication will report extensive calculations which more precisely (than the crude estimates presented in Appendix A) determine $t_{d^8}$ and $t_{d^{10}}^{21}$ as well as the effect of including configurations where the hole is removed from the spin deviations.

## ACKNOWLEDGMENTS

We have benefited from conversations with many people including R. Birgeneau, K. Schweizer, P. Feibelman, R. Martin, M. Schlüter, V. Emery, B. Harmon, A. Mc Mahan, J. Allen, D. Newns, D. Scalapino, J. Schirber, E. Venturini, and J. Zaanen, and wish to recognize the encouraging support of V. Narayanamurti. This work was supported by the U.S. Department of Energy under Contract No. DE-AC04-76DP00789.

---

[1] J. G. Bednorz and K. A. Müller, Z. Phys. B **64**, 189 (1986).

[2] M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Z. Wang, and C. W. Chu, Phys. Rev. Lett. **58**, 908 (1987).

[3] H. Maeda, Y. Tanaka, M. Fukutomi, and T. Asano, Jpn. J. Appl. Phys. Lett. (to be published); R. M. Hazen, C. T. Prewitt, R. J. Angel, N. L. Ross, L. W. Finger, C. G. Hadidiacos, D. R. Veblen, P. J. Heaney, P. H. Hor, R. L. Meng, Y. Y. Sun, Y. Q. Wang, Y. Y. Xue, Z. J. Huang, L. Gao, J. Bechtold, and C. W. Chu (unpublished).

[4] A. Herman and Z. Sheng, Superconductor Week, February 22, 1988, p. 5; D. S. Ginley, E. L. Venturini, J. F. Kwak, R. J. Baughman, M. L. Carr, P. F. Hlava, J. E. Schirber, and B. Morosin (unpublished).

[5] J. Kosterlitz and D. Thouless, J. Phys. C **6**, 1181 (1973); J. Kosterlitz, *ibid.* **7**, 1046 (1974).

[6] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **22**, 1133 (1966); *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions,* edited by D. Forster (Benjamin/Cummings, Reading, MA, 1975).

[7] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[8] L. N. Cooper, Phys. Rev. **104**, 1189 (1956).

[9] D. J. Scalapino, E. Loh, and J. E. Hirsch, Phys. Rev. B **34**, 8190 (1986).

[10] J. Miyake, S. Schmitt-Rink, and C. Varma, Phys. Rev. B **34**, 6554 (1986).

[11] V. Emery, Phys. Rev. Lett. **58**, 2794 (1987).

[12] J. E. Hirsch, Phys. Rev. Lett. **59**, 228 (1987).

[13] R. H. Parmenter, Phys. Rev. Lett. **59**, 923 (1987).

[14] C. M. Varma, S. Schmitt-Rink, and E. Abrahams, Solid State Commun. **62**, 681 (1987).

[15] G. Shirane, Y. Endoh, R. J. Birgeneau, M. A. Kastner, Y. Hidaka, M. Oda, M. Suzuki, and T. Murakami, Phys. Rev. Lett. **59**, 1613 (1987).

[16] W. W. Warren, Jr., R. E. Walstedt, G. F. Brennert, G. P. Espinosa, and J. P. Remeika, Phys. Rev. Lett. **59**, 1860 (1987).

[17] J. E. Schirber, B. Morosin, R. M. Merrill, P. F. Hlava, E. L. Venturini, J. F. Kwak, P. J. Nigrey, R. J. Baughman, and D. S. Ginley, Physica C **152**, 121 (1988).

[18] A. Aharony, R. J. Birgeneau, A. Coniglio, M. A. Kastner, and H. E. Stanley (unpublished).

[19] Y. Guo, J-M. Langlois, and W. A. Goddard III, Science **239**, 896 (1988).

[20] R. J. Birgeneau, M. A. Kastner, and A. Aharmony, Z Phys. (to be published).

[21] J. E. Hirsch, S. Tang, E. Loh, Jr., and D. J. Scalapino, Phys. Rev. Lett. **60**, 1668 (1988).

[22] S. P. Kowalczyk, R. A. Pollak, F. R. McFeely, L. Ley, and D. A. Shirley, Phys. Rev. B **8**, 2387 (1973).

[23] D. A. Shirley, Phys. Rev. A **7**, 1520 (1973).

[24] M. L. Knotek and P. J. Feibelman, Phys. Rev. Lett. **40**, 964 (1978).

[25] T. Jach and C. J. Powell, Phys. Rev. Lett. **46**, 953 (1981).

[26] D. M. Zehner, J. R. Noonan, and H. H. Maden, J. Vac. Sci. Technol. **20**, 859 (1982).

[27] M. Schlüter, M. S. Hybertsen, and N. E. Christensen, in *Proceedings of the International Conference on High-$T_c$ Superconductors: Materials and Mechanisms of Superconductivity, Interlaken, Switzerland, 1988,* edited by J. Muller and J. L. Olsen [Physica C (to be published)].

[28] C. F. Chen, X. W. Wang, T. C. Leung, and B. N. Harmon (unpublished).

[29] A. K. McMahan, R. M. Martin, and S. Satpathy (unpublished).

[30] L. F. Mattheiss, Phys. Rev. Lett. **58**, 1028 (1987).

[31] Y. Fukuda and I. Toyoshima, Surf. Sci **158**, 482 (1985).

[32] One can show that for the $CuO_2$ planar system that half the bandwidth is given by

$$W/2 = 4t_{pp} + \sqrt{8}t_{pd}\tan[\tfrac{1}{2}\tan^{-1}(\sqrt{2}t_{pd}/t_{pp})]$$

which is given to a good approximation by $8(t_{pp}+t_{pd})/3$ for $0 \leq t_{pp}/t_{pd} \leq 1$.

[33] V. Emery (private communication).

[34] D. E. Ramaker, N. H. Turner, J. S. Murday, L. E. Toth, M. Osofsky, and F. L. Hutson, Phys. Rev. B **36**, 5672 (1987).

[35] The ion orbitals were calculated self-consistently at the Hartree-Fock double-$\zeta$ level by embedding the ion in a Madelung potential produced by an array of point charges appropriate to the La compound. The resulting wave functions were used to compute two-electron Coulomb integrals between Cu and O ions at 1.87-Å separation. These in turn were used to compute the exchange interaction between symmetrically orthogonalized Cu $3d$ and O $2p_\sigma$ orbitals.

[36] J. Redinger, J. Yu, A. J. Freeman, and P. Weinberger, Phys. Lett. A **124**, 463 (1987).

[37] D. Vaknin, S. K. Sinha, D. E. Moncton, D. C. Johnston, J. M. Newsam, C. R. Safinya, and H. E. King, Jr., Phys. Rev. Lett. **58**, 2802 (1987).

[38] J. M. Tranquada *et al.*, Phys. Rev. Lett. **60**, 156 (1988); N. Nishida *et al.*, Jpn. J. Appl. Phys. Pt. 2, **26**, L1856 (1987); J. Phys. Soc. Jpn. **57**, 599 (1988).

[39] R. J. Birgeneau *et al.* (unpublished).

[40] K. B. Lyons, P. A. Fleury, L. F. Schneemeyer, and J. V. Waszczak, Phys. Rev. Lett. **60**, 732 (1988).

[41] K. B. Lyons, P. A. Fleury, J. P. Remeika, and T. J. Negran, Phys. Rev. B **37**, 2353 (1988).

[42] J. M. Traquada *et al.*, Phys. Rev. B **36**, 5263 (1987); Z.-X. Shen *et al.*, *ibid.* **36**, 8414 (1987); N. Nucker, J. Fink, J. C. Fuggle, P. J. Durham, and W. M. Temmerman, *ibid.* **37**, 5158 (1988).

[43] M. Onellion, Y. Chang, D. W. Niles, R. Joynt, G. Margaritondo, N. G. Stoffel, and J. M. Tarascon, Phys. Rev. B **36**, 819 (1987); R. L. Kurtz, R. L. Stockbauer, D. Mueller, A. Shih, L. E. Toth, M. Osofsky, and S. A. Wolf, *ibid.* **35**, 8818 (1987).

[44] Y. Endoh, K. Yamada, R. J. Birgeneau, M. A. Kastner, Y. Hidaka, M. Oda, M. Suzuki, and T. Murakami, Phys. Rev. Lett. **59**, 1613 (1987).

[45] I. Ohana, Y. C. Liu, P. J. Picone, A. Lusnikov, M. S. Dresselhaus, G. Dresselhaus, H. P. Jenssen, D. R. Gabbe, H. J. Zeiger, and A. J. Strauss (unpublished).

[46] S. Chakravarty, B. I. Halperin, and D. R. Nelson, Phys. Rev. Lett. **60**, 1057 (1988).

[47] D. P. Arovas and A. Auerbach (unpublished).

[48] E. B. Stechel and K. S. Schweizer (unpublished).

[49] E. Manousakis and R. Salvador, Phys. Rev. Lett. **60**, 840 (1988).

[50] J. B. Parkinson, J. Phys. C **2**, 2012 (1969).

[51] Z. Shen, J. W. Allen, J. J. Yeh, J.-S. Kang, W. Ellis, W. Spicer, I. Lindau, M. B. Maple, Y. D. Dalichaouch, M. S. Torikachvili, J. Z. Sun, and T. H. Geballe, Phys. Rev. B **36**, 8414 (1987).

[52] T. Holstein and H. Primakoff, Phys. Rev. **58**, 1098 (1940).

[53] T. Matsubara, Prog. Theor. Phys. (Kyoto) **14**, 351 (1955).

[54] M. Takahashi, Prog. Theor. Phys. Suppl. **87**, 233 (1986).

[55] See, for example, J. Callaway, *Quantum Theory of the Solid State* (Academic, New York, 1974).

[56] K. S. Schweizer (unpublished).

[57] R. B. Stinchcombe, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and J. L. Lebowitz (Academic, New York, 1983), Vol. 7, p. 151, and references therein.

[58] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).

[59] D. R. Nelson, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and J. L. Lebowitz (Academic, New York, 1983), Vol. 7, p. 1, and references therein.

[60] S. E. Inderhees, M. B. Salamon, N. Goldenfeld, J. P. Rice, B. G. Pazol, and D. M. Ginsberg, Phys. Rev. Lett. **60**, 1178 (1988).

[61] R. A. Ferrell and R. E. Prange, Phys. Rev. Lett. **10**, 479 (1963); also reproduced by B. D. Josephson, in *Superconductivity,* edited by R. D. Parks (Marcel Dekker, New York, 1969), p. 423.

[62] R. J. Birgeneau, G. Shirane, and H. J. Guggenheim, Phys. Rev. Lett. **22**, 720 (1969); Phys. Rev. B **1**, 2211 (1970); R. J. Birgeneau, J. Skalyo, Jr., and G. Shirane, J. Appl. Phys. **41**, 1303 (1970).

[63] R. L. Martin and P. W. Saxe (unpublished).

[64] J. E. Hirsch, S. Tang, E. Loh, and D. J. Scalapino (unpublished).

[65] J. R. Schrieffer, X.-G. Wen, and S.-C. Zhang, Phys. Rev. Lett. **60**, 944 (1988).

[66] J. E. Hirsch and H. Q. Lin (unpublished); H. Q. Lin, J. E. Hirsch, and D. J. Scalapino (unpublished).

[67] J. Zaanen and A. M. Oles, Phys. Rev. B **37**, 9423 (1988); A. M. Oles and J. Zaanen, in *Proceedings of the International Conference on High-$T_c$ Superconductors: Materials and Mechanisms of Superconductivity, Interlaken, Switzerland, 1988,* edited by J. Muller and J. L. Olsen [Physica C (to be published)].

[68] F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).

[69] P. W. Anderson, Science **235**, 1196 (1987); G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).

[70] S. A. Kivelson, D. S. Rokhsar, and J. P. Sethna, Phys. Rev. B **35**, 8865 (1987).

[71] D. M. Newns, Phys. Rev. B **36**, 5595 (1987).

[72] S. A. Trugman, Phys. Rev. B **37**, 1597 (1988).

[73] P. W. Anderson, J. Phys. Chem. Solids **11**, 26 (1959).

[74] J. Bardeen, D. M. Ginsberg, and M. B. Salamon, in Proceedings of the Conference on Novel Mechanisms of Superconductivity, Berkeley, California, 1987 (unpublished).

[75] V. Emery, Nature **328**, 756 (1987).