10-1-1995

Anisotropic Spin Hamiltonians Due to Spin-Orbit and Coulomb Exchange Interactions

Taner Yildirim
University of Pennsylvania, taner@seas.upenn.edu

A. Brooks Harris
University of Pennsylvania, harris@sas.upenn.edu

Amnon Aharony

Ora Entin-Wohlman

Recommended Citation
Yildirim, T., Harris, A., Aharony, A., & Entin-Wohlman, O. (1995). Anisotropic Spin Hamiltonians Due to Spin-Orbit and Coulomb Exchange Interactions. *Physical Review B, 52* (14), 10239-10267. http://dx.doi.org/10.1103/PhysRevB.52.10239

Anisotropic Spin Hamiltonians Due to Spin-Orbit and Coulomb Exchange Interactions

Abstract

Here we correct, extend, and clarify results concerning the spin Hamiltonian $\mathcal{H}_S$ used to describe the ground manifold of Hubbard models for magnetic insulators in the presence of spin-orbit interactions. Most of our explicit results are for a tetragonal lattice as applied to some of the copper oxide lamellar systems and are obtained within the approximation that $\mathcal{H}_S$ consists of a sum of nearest-neighbor bond Hamiltonians. We consider both a "generic" model in which hopping takes place from one copper ion to another and a "real" model in which holes can hop from a copper ion to an intervening oxygen 2*p* band. Both models include orbitally dependent direct and exchange Coulomb interactions involving two orbitals. Our analytic results have been confirmed by numerical diagonalizations for two holes occupying any of the 3*d* states and, if applicable, the oxygen 2*p* states. An extension of the perturbative scheme used by Moriya is used to obtain analytic results for $\mathcal{H}_S$ up to order $\mathbf{t}^2$ ($\mathbf{t}$ is the matrix of hopping coefficients) for arbitrary crystal symmetry for both the "generic" and "real" models. With only direct orbitally independent Coulomb interactions, our results reduce to Moriya's apart from some minor modifications. For the tetragonal case, we show to all orders in $\mathbf{t}$ and λ, the spin-orbit coupling constant, that $\mathcal{H}_S$ is isotropic in the absence of Coulomb exchange terms and assuming only nearest-neighbor hopping. In the presence of Coulomb exchange, scaled by *K*, the anisotropy in $\mathcal{H}_S$ is biaxial and is shown to be of order $Kt^2\lambda^2$. Even when $\mathbf{K}=0$, for systems of sufficiently low symmetry, the anisotropy in $\mathcal{H}_S$ is proportional to $t^6\lambda^2$ when the direct on-site Coulomb interaction *U* is independent of the orbitals involved and of order $t^2\lambda^2$ otherwise. These latter results apply to the orthorhombic phase of La_2CuO_4.

Disciplines

Physics

Anisotropic spin Hamiltonians due to spin-orbit and Coulomb exchange interactions

T. Yildirim
University of Maryland, College Park, Maryland 20742;
National Institute of Standards and Technology, Gaithersburg, Maryland 20899;
and Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6369

A. B. Harris
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6369

Amnon Aharony and O. Entin-Wohlman
School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

(Received 29 March 1995)

Here we correct, extend, and clarify results concerning the spin Hamiltonian $\mathcal{H}_S$ used to describe the ground manifold of Hubbard models for magnetic insulators in the presence of spin-orbit interactions. Most of our explicit results are for a tetragonal lattice as applied to some of the copper oxide lamellar systems and are obtained within the approximation that $\mathcal{H}_S$ consists of a sum of nearest-neighbor bond Hamiltonians. We consider both a "generic" model in which hopping takes place from one copper ion to another and a "real" model in which holes can hop from a copper ion to an intervening oxygen $2p$ band. Both models include orbitally dependent direct and exchange Coulomb interactions involving two orbitals. Our analytic results have been confirmed by numerical diagonalizations for two holes occupying any of the $3d$ states and, if applicable, the oxygen $2p$ states. An extension of the perturbative scheme used by Moriya is used to obtain analytic results for $\mathcal{H}_S$ up to order $\mathbf{t}^2$ ($\mathbf{t}$ is the matrix of hopping coefficients) for arbitrary crystal symmetry for both the "generic" and "real" models. With only direct orbitally independent Coulomb interactions, our results reduce to Moriya's apart from some minor modifications. For the tetragonal case, we show to all orders in $\mathbf{t}$ and λ, the spin-orbit coupling constant, that $\mathcal{H}_S$ is isotropic in the absence of Coulomb exchange terms and assuming only nearest-neighbor hopping. In the presence of Coulomb exchange, scaled by K, the anisotropy in $\mathcal{H}_S$ is biaxial and is shown to be of order $Kt^2\lambda^2$. Even when $\mathbf{K}=0$, for systems of sufficiently low symmetry, the anisotropy in $\mathcal{H}_S$ is proportional to $t^6\lambda^2$ when the direct on-site Coulomb interaction U is independent of the orbitals involved and of order $t^2\lambda^2$ otherwise. These latter results apply to the orthorhombic phase of La_2CuO_4.

I. INTRODUCTION

A longstanding problem which has attracted much interest recently concerns the mechanism whereby spin-orbit interactions give rise to magnetic anisotropy in magnetic insulators. This subject, which was extensively investigated three decades ago,[1,2] has recently been the object of renewed attention due to interest in the lamellar copper oxide systems.[3] The first of these to be extensively investigated, La_2CuO_4, has a small orthorhombic distortion away from a tetragonal structure[4] and the above mechanism was shown[5–7] to give rise to an anisotropic exchange, including that of the antisymmetric Dzyaloshinskii-Moriya type. In that system there are two anisotropy energies.[8,9] One of these, the out-of-plane anisotropy, is of the form αN_z^2, where N_z is the z component of the staggered magnetization, the z axis is taken to be perpendicular to the copper oxide plane, and α is an anisotropy constant. This energy causes the spins to lie in the basal plane. There is also an in-plane anisotropy energy which selects the orientation of the spins within the basal plane. Until recently the discussions of the origins of anisotropy were confined to the orthorhombic structure. However, more recently a family of copper oxide materials of similar structure, but which are actually tetragonal, have been studied[10,11] and found to have roughly the same out-of-plane anisotropy as La_2CuO_4. The earlier studies[5–7] did not predict any anisotropy in the tetragonal limit. Accordingly, a reanalysis of anisotropy for the tetragonal systems ought to show a common origin of the out-of-plane anisotropy which does not rely on the orthorhombic distortion. That is the main purpose of this paper. However, in the course of this work, we have found that a number of general questions concerning both the results and the methodology required some clarification, which this paper is intended to provide.

A microscopic basis for superexchange between magnetic ions was first given almost forty years ago by Anderson.[12] In the language of a Hubbard model,[13] his calculation started from an orbitally nondegenerate band in which there is one electron per site in the limit of large Coulomb interaction U whenever two electrons occupy the same site. If the kinetic energy is completely neglected, each electron (or hole) may be characterized by its spin. When kinetic energy (described by hopping) is included perturbatively, one finds a spin Hamiltonian, which in low-order perturbation theory can be expressed as the sum of contributions $\mathcal{H}(i,j)$ from each bond (i,j). This spin Hamiltonian describes the per-

0163-1829/95/52(14)/10239(29)/$06.00

turbative removal of degeneracy. In higher order in the hopping, one encounters contributions to the spin Hamiltonian from plaquettes (at order t^4/U^3, where t is a hopping matrix element) and eventually from even higher-order clusters. Ignoring higher-order contributions, Anderson obtained an isotropic exchange interaction between nearest-neighbor spins,

$$\mathcal{H}(i,j) = J(i,j)\mathbf{S}(i)\cdot\mathbf{S}(j) \ , \tag{1}$$

where $J(i,j) = 4t_{ij}^2/U$ and t_{ij} is the hopping matrix element between sites i and j.

Soon afterwards Moriya[1] used Anderson's formalism to study the effect of spin-orbit interactions on superexchange between magnetic ions. He showed that for sufficiently low symmetry the most general effective spin Hamiltonian for two spin-$\frac{1}{2}$ magnetic ions, such as Cu^{++}, is of the form

$$\mathcal{H}(i,j) = J(i,j)\mathbf{S}(i)\cdot\mathbf{S}(j) + \mathbf{D}(i,j)\cdot\mathbf{S}(i)\times\mathbf{S}(j) + \mathbf{S}(i)\cdot\mathbf{M}(i,j)\cdot\mathbf{S}(j) \ , \tag{2}$$

where $\mathbf{M}(i,j)$ is a symmetric 3×3 tensor. The first term represents the isotropic symmetric exchange. The second and third terms represent the antisymmetric and symmetric anisotropies, respectively. Moriya's results were obtained to second order in the hopping perturbation, but in principle provided a framework in which the spin-orbit interaction could be included to arbitrary order. Convenient explicit results were given to lowest nontrivial order in the spin-orbit coupling constant, λ.

Much more recently, Thio *et al.*[4] found that La_2CuO_4 is described by Eq. (2). Consequently Coffey and coworkers[5] invoked this Hamiltonian to describe the CuO planes in the cuprates. They found that $\mathbf{D}(i,j)$ cannot be the same for all bonds $\langle ij\rangle$, as was assumed by a number of previous authors. The form of the $\mathbf{D}(i,j)$ is determined by the symmetry properties of the crystal structure. The first attempt at a microscopic calculation of the vectors $\mathbf{D}(i,j)$ was made by Coffey, Rice, and Zhang[6] in the framework of the Moriya theory of the anisotropic superexchange interactions. Within this theory, $\mathbf{D}(i,j)$ is of order λ, whereas $\mathbf{M}(i,j)$ is of order λ^2. Therefore, many authors neglected $\mathbf{M}$. Naively, one expected a gap in the spin-wave spectrum due to anisotropy, and this is what one finds when $\mathbf{M}(i,j)$ is neglected. Subsequently, Shekhtman, Entin-Wohlman, and Aharony (SEA)[7] have shown that $\mathbf{M}(i,j)$ can *never* be neglected. Most interestingly, when $\mathbf{M}(i,j)$ is included, they found a hidden symmetry in $\mathcal{H}(i,j)$, as a result of which inclusion of spin-orbit interactions did not reduce the degeneracy of the ground state of the pair of spins (i,j).[14] Their result was that $\mathcal{H}(i,j)$ could be written in the following form:

$$\mathcal{H}(i,j) = \left(J - \frac{D^2}{4J}\right)\mathbf{S}(i)\cdot\mathbf{S}(j) + \mathbf{D}(i,j)\cdot\mathbf{S}(i)\times\mathbf{S}(j) + \mathbf{S}(i)\cdot\frac{\mathbf{D}\otimes\mathbf{D}}{2J}\cdot\mathbf{S}(j) \ , \tag{3}$$

where the vector $\mathbf{D}(i,j)$ is bond dependent and $[\mathbf{A}\otimes\mathbf{B}]_{\mu\nu} = A_\mu B_\nu$. As SEA show, the result (3) indicates that although the pair interaction is not of the isotropic form of Eq. (1), it is rotationally invariant and hence the energy level spectrum of the pair interactions consists of a singlet and a triplet, just as it would in the absence of spin-orbit interactions. In previous work the terms in $\mathbf{D}(i,j)$ and those in $M_{\mu,\nu}(i,j) \equiv D_\mu(i,j)D_\nu(i,j)/(2J)$ were not treated on an equal footing, and therefore this hidden symmetry was never noticed. Furthermore, SEA showed that even though each individual bond might have this hidden symmetry, the crystal as a whole could have anisotropy because of the frustration caused by the competition between exchange interactions of different bonds.[14] In particular, for La_2CuO_4 they found that the anisotropy was a result of this frustration.

All the work cited so far relied on the idea, introduced by Moriya, that the effect of spin-orbit interactions could be taken into account by a gauge transformation on the hopping between sites. As used by Moriya to obtain results up to order t^2/U, this formulation is correct and convenient. However, this formulation does not form a correct basis for calculations to higher order in t/U. Thus, as we shall see, the hidden symmetry of SEA, although maintained at order t^2 for constant U, is broken at order t^6 for constant U or at order t^2 for nonconstant U. (Here constant U means that the Coulomb interaction between holes in two orbitals does not depend on which orbitals are involved.) In addition, the calculations of Shekhtman, Aharony, and Entin-Wohlman[15] or Bonesteel[16] for the anisotropy of the cuprates were based on terms requiring the existence of a distortion from tetragonal symmetry. However, the easy plane anisotropy is observed[10,11] to have similar magnitudes in both the orthorhombic and tetragonal cuprates isostructural to La_2CuO_4. The main reason for the failure of the previous calculations to give anisotropy for the tetragonal cuprates was the fact that these calculations neglected the Coulomb exchange interaction. From the results of Barriquand and Sawatzky[17] (BS) one can see that they partially included such interactions. However, it remained unclear which aspects of the BS results would persist when the calculation was pursued more systematically. In fact, in Ref. 18 it was shown that for tetragonal symmetry Coulomb exchange interactions played a crucial role in determining the anisotropy.

In view of the above history, the following points remained to be clarified and are addressed in the present paper. (1) One should generalize Moriya's results for $\mathcal{H}(i,j)$ to the case of nonconstant U. Having done that, we find that when reduced to the case of constant U, our present results differ in a small way from those of Moriya, who overlooked some λ-dependent contributions to the magnitude of the isotropic exchange interaction, J. We also give general results for superexchange interactions, i.e., for the case when the copper ions are separated by an intervening oxygen ion. However, the results are given in a general form which can equally apply to systems of ions other than Cu, as long as their ground state is orbitally nondegenerate. (2) Since earlier calculations for the cuprates omitted hopping between excited states of the Cu ions, we have reanalyzed the role of symmetry at arbitrary order in the matrix elements $t_{\mathrm{Cu-Cu}} \equiv t$ which describe the effective hopping between copper ions. We

find that in the absence of Coulomb exchange interactions [i.e., for $\mathbf{K}=0$ in Eq. (4), below], one recovers isotropic exchange for a simplified "generic" model which describes the complete $3d$ band for copper ions on a simple tetragonal Bravais lattice. This isotropy is the result of the high symmetry of the crystal field levels and the resulting high symmetry of the hopping matrix elements. This result shows that in the absence of Coulomb exchange terms, one retains isotropy in $\mathcal{H}(i,j)$ to all orders in both λ and t_{ij}/U and thus that inclusion of Coulomb exchange interactions is essential to obtain anisotropic exchange interactions in the tetragonal case. (3) For a tetragonal lattice we find that this accidental isotropy in $\mathcal{H}(i,j)$ is removed at order $t^2\lambda^2 K$ when Coulomb exchange interactions are allowed and we give detailed expressions for the exchange anisotropy in terms of the hopping matrix elements and the matrix elements of the Coulomb interaction. (4) For a crystal with arbitrarily low symmetry (i.e., when the crystal field states have no special symmetry), we expect to (and do) find a removal of degeneracy of the spin triplet. This breaking of rotational invariance occurs at order t^6 for the case of constant U and at order t^2 when U is nonconstant. These results modify the conclusion given in Ref. 7. (5) In contrast to all previous work, we also found an in-plane anisotropy originating from the anisotropy of the spin-wave zero-point energy.

Most of the above results have been obtained analytically, both for the "generic" model (with only Cu ions) and for the "real" model (in which the Cu ions are separated by oxygen ions). Furthermore, we have corroborated our results by comparing them to results obtained by numerically diagonalizing the Hamiltonian which describes all possible states of two holes on one bond. For the "generic" model, there are 20 single-particle orbitals, 10 on each copper ion, so that in all there are 190 two-hole states. For the "real" model there are six additional $2p$ states on the oxygen ion, so there are 325 two-hole states in all. In the tetragonal case, where we know that the exchange interaction matrix $J_{\mu\nu}(i,j)$ [see Eq. (13), below] is diagonal, its values may be deduced from the values of the energy splittings of the ground manifold, as is discussed in Appendix A.

Briefly, this paper is organized as follows. In Sec. II, we first introduce "generic" and "real" Hamiltonians and then discuss the perturbative framework we use to calculate the spin Hamiltonian $\mathcal{H}_S$. The actual perturbative calculations of $\mathcal{H}_S$ are described in Sec. III, although many of the details are relegated to Appendixes. Here we give expressions for $J(i,j)$, $\mathbf{D}(i,j)$, and $\mathbf{M}(i,j)$ up to order $t^2\lambda^2\Delta\mathcal{H}_c$ for the "generic" model and analogously for the "real" model, where $\Delta\mathcal{H}_c$ represents Coulomb interactions beyond the approximation in which the Coulomb exchange $\mathbf{K}$ is zero and $\mathbf{U}$ is constant. The case of tetragonal site symmetry is discussed in Sec. IV for both the "generic" and "real" models. There we prove a theorem, valid to all orders in the hopping matrix elements and spin-orbit coupling, which says that for nearest-neighbor hopping, the complete spin Hamiltonian is isotropic when Coulomb exchange is absent. There we display explicitly the leading contribution to the anisotropic exchange when Coulomb exchange is treated perturbatively. In Sec. V we discuss the experimental consequences of these results. In particular we estimate the anisotropies and spin-wave gaps which our work would predict. In Sec. VI we study the case of arbitrarily low symmetry for the "generic" model and show that the anisotropy in $J_{\mu\nu}(i,j)$ is of order $t^6\lambda^2$ for the case of constant U and of order $t^2\lambda^2$ when U is not constant. Finally, in Sec. VII we summarize the conclusions of this work. A brief summary of our major conclusions has been given previously.[19,18]

II. HUBBARD HAMILTONIAN AND SYMMETRY OF EXCHANGE

A. Generic model

In this section we introduce a general Hamiltonian, versions of which will be studied in this paper. We start from the following generic model,[20] which captures the symmetries of the cuprates. For holes which reside only on the Cu ions, this model is given by

$$\begin{aligned}\mathcal{H}=&\sum_{i,\alpha,\sigma}\epsilon_{i\alpha}d^\dagger_{i\alpha\sigma}d_{i\alpha\sigma}+\sum_{\text{holes,h}}\lambda\mathbf{L}(h)\cdot\mathbf{S}(h)+\sum_{\substack{\alpha,\beta,\sigma\\ i\neq j}}t_{i\alpha,j\beta}(d^\dagger_{i\alpha\sigma}d_{j\beta\sigma}+d^\dagger_{j\beta\sigma}d_{i\alpha\sigma})\\&+\frac{1}{2}\sum_{\substack{i,\alpha,\alpha'\\ \sigma,s}}U_{i\alpha,i\alpha'}d^\dagger_{i\alpha\sigma}d^\dagger_{i\alpha' s}d_{i\alpha' s}d_{i\alpha\sigma}+\frac{1}{2}\sum_{\substack{i,\alpha\neq\alpha'\\ \sigma,s}}K_{i\alpha,i\alpha'}d^\dagger_{i\alpha\sigma}d^\dagger_{i\alpha' s}d_{i\alpha s}d_{i\alpha'\sigma}\\&+\frac{1}{2}\sum_{\substack{\alpha,\beta,\sigma,s\\ i\neq j}}V_{i\alpha,j\beta}(d^\dagger_{i\alpha\sigma}d^\dagger_{j\beta s}d_{j\beta s}d_{i\alpha\sigma}+d^\dagger_{j\beta\sigma}d^\dagger_{i\alpha s}d_{i\alpha s}d_{j\beta\sigma})\\&+\frac{1}{2}\sum_{\substack{\alpha,\beta,\sigma,s\\ i\neq j}}N_{i\alpha,j\beta}(d^\dagger_{i\alpha\sigma}d^\dagger_{j\beta s}d_{i\alpha s}d_{j\beta\sigma}+d^\dagger_{j\beta\sigma}d^\dagger_{i\alpha s}d_{j\beta s}d_{i\alpha\sigma})\ .\end{aligned}\tag{4}$$

Here $d^\dagger_{i\alpha\sigma}$ creates a hole in the αth spatial orbital, whose single-particle energy is $\epsilon_{i\alpha}$, with z component of spin σ on the Cu ion at site i. In general we allow hopping with matrix elements $t_{i\alpha,j\beta}$ between the α orbital on site i and the β orbital on site j. This Hamiltonian also includes direct Coulomb interactions between electrons on the same site (scaled by $\mathbf{U}$) and on different sites (scaled by $\mathbf{V}$) and exchange Coulomb interactions between electrons on the

same site (scaled by **K**) and on different sites (scaled by **N**). Our numerical work indicates that when $\mathbf{t} \neq 0$, the effects of **V** and **N** are not qualitatively different from those of **U** and **K**, respectively. Since the latter are often dominant, we shall neglect **V** and **N**. In principle one should also include Coulomb terms with four states, $U_{\alpha\beta\gamma\delta} d^\dagger_{i\alpha\sigma} d^\dagger_{i\beta s} d_{i\gamma s} d_{i\delta\sigma}$. Here we follow most of the literature and start with the simpler Eq. (4), which involves only the Hartree-like terms **U** and the simple Coulomb exchange terms **K**. Equation (4) can easily be extended to the "real" model in which we include p states on the oxygen ions, with hopping between them and the d states on the nearest-neighboring Cu ions.

B. Single site Hamiltonian for noninteracting holes

In this subsection we briefly discuss the basis states used in the perturbative scheme described in the next subsection. We first consider ions within a single-particle picture. We therefore start by considering the effects of the crystal field Hamiltonian, $\mathcal{H}_x$, and the spin-orbit interaction, $\mathcal{H}_{\rm so}$. The former is constructed so as to give the observed ionic levels. Including only such energies the single-particle Hamiltonian is

$$\mathcal{H}_x + \mathcal{H}_{\rm so} = \sum_{i\alpha\sigma} \epsilon_{i\alpha} d^\dagger_{i\alpha\sigma} d_{i\alpha\sigma} + \lambda \sum_{\substack{i\alpha\beta \\ \sigma\sigma'}} [\omega_i(\alpha,\beta)]_{\sigma\sigma'} d^\dagger_{i\alpha\sigma} d_{i\beta\sigma'} \ , \tag{5}$$

where

$$[\omega_i(\alpha,\beta)]_{\sigma\sigma'} \equiv \frac{1}{2} \sum_\mu \langle i\alpha \mid L_\mu \mid i\beta\rangle (\sigma_\mu)_{\sigma\sigma'}, \tag{6}$$

in which $\langle i\alpha \mid L_\mu \mid i\beta\rangle \equiv L^\mu_{\alpha\beta}$ is the matrix element of the μ component of the orbital angular momentum between the two single-particle states, and σ_μ is the Pauli matrix. We shall often present results for i-independent matrix elements of **L**.

For many purposes it is convenient to diagonalize the single-particle, single-site Hamiltonian $\mathcal{H}_x + \mathcal{H}_{\rm so}$. We may choose the wave functions $|i\alpha\rangle$ to be real, in which case the matrix elements $L^\mu_{\alpha\beta}$ are purely imaginary. As a result, every single-particle energy of $\mathcal{H}_x + \mathcal{H}_{\rm so}$ is at least doubly degenerate. That is, the two linearly independent wave functions which are related to one another by time reversal,

$$\begin{aligned} \psi_a &= \sum_\alpha (y_{\alpha a} \mid \alpha\uparrow\rangle + z_{\alpha a} \mid \alpha\downarrow\rangle), \\ \phi_a &= \sum_\alpha (-z^*_{\alpha a} \mid \alpha\uparrow\rangle + y^*_{\alpha a} \mid \alpha\downarrow\rangle), \end{aligned} \tag{7}$$

belong to the same energy. We use greek indices to label the crystal field states in the absence of spin-orbit interactions and roman ones for the eigenstates of $\mathcal{H}_x + \mathcal{H}_{\rm so}$. The latter can be characterized by pseudospin quantum numbers, $\sigma = \pm 1$, and are associated with the creation operators $c^\dagger_{ia\sigma}$. These operators are related to the $d^\dagger_{i\alpha\sigma}$'s via

$$d^\dagger_{i\alpha\sigma} = \sum_{a\sigma_1} (m^i_{\alpha a})^*_{\sigma\sigma_1} c^\dagger_{ia\sigma_1}, \tag{8}$$

where the unitary matrix $\mathbf{m}^i_{\alpha a}$ is

$$\mathbf{m}^i_{\alpha a} \equiv \begin{pmatrix} y^i_{\alpha a} & -(z^i_{\alpha a})^* \\ z^i_{\alpha a} & (y^i_{\alpha a})^* \end{pmatrix} \equiv u^i_{\alpha a}\mathbf{I} + i\mathbf{v}^i_{\alpha a}\cdot\vec{\sigma} \ , \tag{9}$$

where $u^i_{\alpha a}$ is a real scalar, $\mathbf{v}^i_{\alpha a}$ is a real vector, and **I** is the 2×2 unit matrix. This leads to

$$\mathcal{H}_x + \mathcal{H}_{\rm so} \equiv \sum_{ia\sigma} E_{ia} c^\dagger_{ia\sigma} c_{ia\sigma}, \tag{10}$$

where

$$\left\{ \sum_\alpha \epsilon_{i\alpha} (\mathbf{m}^i_{\alpha a})^\dagger \mathbf{m}^i_{\alpha b} + \lambda \sum_{\alpha\beta} (\mathbf{m}^i_{\alpha a})^\dagger \omega_i(\alpha,\beta) \mathbf{m}^i_{\beta b} \right\}_{\sigma_1\sigma_2} = E_{ia}\delta_{ab}\delta_{\sigma_1\sigma_2} \ . \tag{11}$$

[Here the dagger operation on $\mathbf{m}^i_{\alpha a}$ operates only in terms of the 2×2 matrices as in Eq. (9) and is not to be applied to the scripts i, α, or a.] The transformation $\mathbf{m}_{\alpha a}$ that diagonalizes the single-particle Hamiltonian is in general different for each site. Consequently, the single-site energies may depend on the site index. However, in certain situations, for example, in the presence of the tetragonal to orthorhombic distortion in La_2CuO_4, it is possible to define the transformation such that the single-particle energies are site independent. This will be the case for some of the explicit calculations which are presented below for the cuprates.

C. Formulation of perturbation theory

For Cu^{++} ions in a d^9 configuration we are dealing with an ionic ground state having one $3d$ hole whose spin is arbitrary. When we include the oxygen ions in the model, those ions have filled $2p$ bands in their ground state. In either case, in the absence of hopping, i.e., for $\mathbf{t} = 0$, the many-electron ground state manifold is one in which one hole of arbitrary spin resides on each copper ion. The energy levels within this ground manifold, when the remaining terms in the Hamiltonian, especially hopping, are considered, are the object of our study.

When hopping is introduced as a perturbation, the splitting of the hitherto degenerate ground state manifold can be described by a spin Hamiltonian, $\mathcal{H}_S$. In view of time reversal invariance $\mathcal{H}_S$ will consist of two-spin interactions (between nearest and further neighbors), four-spin interactions, and so forth. In the present paper most of our results will be for the nearest-neighbor two-spin coupling constants, except for the general theorem of Sec. IV, which makes no assumptions about the specific form of $\mathcal{H}_S$. If we only consider two-spin interactions between nearest-neighboring spins, we effectively write

$$\mathcal{H}_S = \sum_{\langle ij\rangle} \mathcal{H}(i,j) \ , \tag{12}$$

where $\langle ij \rangle$ indicates a sum over pairs of nearest-neighboring sites and for spins-$\frac{1}{2}$

$$\mathcal{H}(i,j) = \sum_{\mu\nu} J_{\mu\nu}(i,j) S_\mu(i) S_\nu(j) \ , \tag{13}$$

where μ and ν label Cartesian components. We refer to the case when $J_{\mu,\nu}(i,j) = J(i,j)\delta_{\mu,\nu}$, where δ is the Kronecker delta function, as isotropic exchange. [To avoid confusion between the two kinds of exchange, the terms in Eq. (4) proportional to $\mathbf{K}$ are referred to as Coulomb exchange.] Appendix A contains a discussion of the possible anisotropies in $\mathcal{H}(i,j)$.

The major objective of this paper is to discuss the symmetry of the matrix $\mathbf{J}(i,j)$ and develop perturbative expressions for it on the basis of the generic Hamiltonian of Eq. (4) and its generalization to include the intervening oxygen ions. From our point of view the most important early work was that of Moriya,[1] who studied a simplified version of the above model. The most significant simplifications necessary to obtain Moriya's main result were to neglect the Coulomb exchange, $\mathbf{K}$, and to assume constant $\mathbf{U}$, i.e., to assume that $U_{i\alpha,i\beta}$ did not depend on either the site index i or the orbital indices α and β. In particular, when $U_{i\alpha,i\beta}$ is independent of α and β, the wave functions for the two-hole states are Slater determinants of the one-hole states as obtained by the canonical transformation of Eq. (8). In other words, in this very special case, the exact eigenstates of the Hamiltonian $\mathcal{H}_x + \mathcal{H}_{\rm so}$ also diagonalize the Coulomb interaction, $\mathcal{H}_c$. In terms of these new single-particle states the transformed hopping Hamiltonian now assumes the form

$$\mathcal{H}_{\rm hop} = \sum_{i,j} T_{ij}, \tag{14a}$$

where

$$T_{ij} = \sum_{\substack{ab \\ \sigma\sigma'}} (\tilde{t}^{ij}_{ab})_{\sigma\sigma'} c^\dagger_{ia\sigma} c_{jb\sigma'} \tag{14b}$$

represents hops from site j to site i, and $(\tilde{\mathbf{t}}^{ij}_{ab})$ is the 2×2 matrix

$$\begin{aligned}(\tilde{\mathbf{t}}^{ij}_{ab}) &\equiv \sum_{\alpha\beta} t_{i\alpha,j\beta} (\mathbf{m}^i_{\alpha a})^\dagger \mathbf{m}^j_{\beta b} \\ &\equiv A^{ij}_{ab}\mathbf{I} + i\mathbf{B}^{ij}_{ab}\cdot\vec{\sigma} \ ,\end{aligned} \tag{14c}$$

in which A^{ij}_{ab} ($\mathbf{B}^{ij}_{ab}$) is a real scalar (vector), that can be found using Eq. (9) and the representation in which $t_{i\alpha,j\beta}$ is real. By Hermiticity these coefficients obey

$$A^{ij}_{ab} = A^{ji}_{ba} \ , \quad \mathbf{B}^{ij}_{ab} = -\mathbf{B}^{ji}_{ba} \ . \tag{15}$$

Actually (and this seems to have caused much subsequent confusion), Moriya did not write down Eqs. (14a)–(14c). Instead, in a further simplification, he truncated $\mathbf{t}$ to include only hopping between the $\mathbf{t} = 0$ ground states. Even for his calculations at order t^2, this simplification is slightly incorrect. However, we should emphasize that this truncation is totally inappropriate for a discussion of effects of order higher than t^2, since hopping between excited states then comes into play. Also, when $\mathbf{U}$ is not constant, hopping between exact eigenstates of $\mathcal{H}_x + \mathcal{H}_{\rm so} + \mathcal{H}_c$ is no longer a single-particle interaction. To see this, note that there are matrix elements between an initial state, in which both holes are in their ground states on different ions, and a final state in which, for instance, both holes are in excited states of one ion. Such a process explicitly relies on the fact that the two-hole states are not simply obtained from single-hole states. Thus, in this case, when "final-state interactions" are present, the hopping perturbation involves four electron operators.

Accordingly, to study the case when $\mathbf{U}$ is not constant and when Coulomb exchange is not neglected, we write $\mathcal{H}_c = \mathcal{H}_{c0} + \Delta\mathcal{H}_c$, where

$$\begin{aligned}\mathcal{H}_{c0} &= \frac{1}{2}U_0 \sum_{\substack{ia\beta \\ \sigma\sigma'}} d^\dagger_{ia\sigma} d^\dagger_{i\beta\sigma'} d_{i\beta\sigma'} d_{ia\sigma} \\ &\equiv \frac{1}{2}U_0 \sum_{\substack{iab \\ \sigma\sigma'}} c^\dagger_{ia\sigma} c^\dagger_{ib\sigma'} c_{ib\sigma'} c_{ia\sigma},\end{aligned} \tag{16}$$

and the additional Coulomb terms resulting from non-constant $\mathbf{U}$ and $\mathbf{K}$ take the form

$$\begin{aligned}\Delta\mathcal{H}_c = \frac{1}{2}\sum_i \sum_{\substack{aba'b' \\ \sigma\sigma'\sigma_1\sigma_1'}} &[\Delta\tilde{U}_{\sigma\sigma'\sigma_1\sigma_1'}(i;abb'a') \\ &+\tilde{K}_{\sigma\sigma'\sigma_1\sigma_1'}(i;abb'a')] c^\dagger_{ia\sigma} c^\dagger_{ib\sigma'} c_{ib'\sigma_1} c_{ia'\sigma_1'},\end{aligned} \tag{17}$$

with

$$\begin{aligned}\Delta\tilde{U}_{\sigma\sigma'\sigma_1\sigma_1'}(i;abb'a') \equiv \sum_{\alpha\alpha'} &\Delta U^i_{\alpha\alpha'} [(\mathbf{m}^i_{\alpha a})^\dagger \mathbf{m}^i_{\alpha a'}]_{\sigma\sigma_1'} \\ &\times [(\mathbf{m}^i_{\alpha' b})^\dagger \mathbf{m}^i_{\alpha' b'}]_{\sigma'\sigma_1},\end{aligned} \tag{18a}$$

$$\begin{aligned}\tilde{K}_{\sigma\sigma'\sigma_1\sigma_1'}(i;abb'a') \equiv \sum_{\alpha\alpha'} &K^i_{\alpha\alpha'} [(\mathbf{m}^i_{\alpha a})^\dagger \mathbf{m}^i_{\alpha' a'}]_{\sigma\sigma_1'} \\ &\times [(\mathbf{m}^i_{\alpha' b})^\dagger \mathbf{m}^i_{\alpha b'}]_{\sigma'\sigma_1} \ .\end{aligned} \tag{18b}$$

Expressions for $\Delta\mathbf{U}$ and $\mathbf{K}$ for tetragonal crystal field states in terms of Racah parameters are given in Appendix B.

In the following, we will calculate the effective spin Hamiltonian using perturbation theory in which we take the unperturbed Hamiltonian to be

$$\begin{aligned}\mathcal{H}_0 &= \mathcal{H}_x + \mathcal{H}_{\rm so} + \mathcal{H}_{c0} \\ &= \sum_{ia\sigma} E_{ia} c^\dagger_{ia\sigma} c_{ia\sigma} + \frac{1}{2}U_0 \sum_{\substack{iab \\ \sigma\sigma'}} c^\dagger_{ia\sigma} c^\dagger_{ib\sigma'} c_{ib\sigma'} c_{ia\sigma}\end{aligned} \tag{19}$$

and the perturbation to be

$$V = \mathcal{H}_{\rm hop} + \Delta\mathcal{H}_c \ , \tag{20}$$

where these quantities are given in Eqs. (14) and (17).

III. PERTURBATIVE CONTRIBUTIONS TO $\mathcal{H}(i,j)$

A. Contributions of order t^2

The lowest-order contributions to $\mathcal{H}(i,j)$ are second order in $\mathbf{t}$. At this order in $\mathbf{t}$ in the absence of the Coulombic perturbation $\Delta\mathcal{H}_c$, we can use the result of Eq. (C2) in Appendix C to evaluate

$$\mathcal{H}^{(2)}(i,j) = -2\left\langle \psi_0' \middle| T_{ij}\frac{1}{\mathcal{H}_0}T_{ji}\middle|\psi_0\right\rangle , \tag{21}$$

where the factor of 2 accounts for the similar term when the hopping is in the reverse direction. Here $|\psi_0\rangle$ and $|\psi_0'\rangle$ are states in the ground manifold with one hole per site, and the superscript (2) indicates a result which is second order in $\mathbf{t}$. In using the result in Appendix C we must truncate the matrix element so that it remains within this manifold. Also, in evaluating this expression it is convenient to use the identity

$$c_{i0\sigma}^{\dagger}c_{i0\sigma'} \equiv \left[\frac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}\right]_{\sigma'\sigma} , \tag{22}$$

whereby we obtain the result

$$\mathcal{H}^{(2)}(i,j) = -\sum_b\left(\mathrm{Tr}\left\{\tilde{\mathbf{t}}_{0b}^{ij}\tilde{\mathbf{t}}_{b0}^{ji}\left[\frac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}\right]\right\}/(U_0+E_{jb})+(i\leftrightarrow j)\right)$$
$$+2\,\mathrm{Tr}\left\{\tilde{\mathbf{t}}_{00}^{ij}\left[\frac{1}{2}+\mathbf{S}(j)\cdot\vec{\sigma}\right]\tilde{\mathbf{t}}_{00}^{ji}\left[\frac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}\right]\right\}/U_0 , \tag{23}$$

where the traces are over the 2×2 matrices in σ space and $(i\leftrightarrow j)$ denotes the sum of all previous terms with i and j interchanged.

The first term in (23), which only involves hopping of a single hole (from site i to j and back), is easily shown to be independent of the spins at i and j. [This follows directly from the identities of Eq. (15), or more simply from time reversal invariance.] Therefore, this term contributes a spin-independent constant, and does not affect the splitting of the ground state. Similarly, the terms coming from the factors of $1/2$ inside the square brackets in the second term also give constants. To order t^2 we have thus arrived at an effective magnetic Hamiltonian of the form of Eq. (12), with

$$\mathcal{H}^{(2)}(i,j) = \frac{2}{U_0}\mathrm{Tr}\{\tilde{\mathbf{t}}_{00}^{ij}[\mathbf{S}(j)\cdot\vec{\sigma}]\tilde{\mathbf{t}}_{00}^{ji}[\mathbf{S}(i)\cdot\vec{\sigma}]\}. \tag{24}$$

In view of Eq. (14c), this becomes

$$\mathcal{H}^{(2)}(i,j) = \frac{2}{U_0}\mathrm{Tr}\{[A_{00}^{ij}+i\mathbf{B}_{00}^{ij}\cdot\vec{\sigma}][\mathbf{S}(j)\cdot\vec{\sigma}][A_{00}^{ji}+i\mathbf{B}_{00}^{ji}\cdot\vec{\sigma}][\mathbf{S}(i)\cdot\vec{\sigma}]\}. \tag{25}$$

The symmetry of this form is further discussed in Appendix D, where we show that in fact $\mathcal{H}^{(2)}(i,j)$ is of the isotropic form of Eq. (1).

B. Contribution of order $t^2\Delta\mathcal{H}_c$

To calculate the contributions of the Coulomb terms of Eq. (17) to the magnetic exchange we need to carry out third-order perturbation theory. By taking two factors of the hopping matrix element we generate terms of order $\tilde{\mathbf{t}}^2$. We must include an additional factor of $\Delta\mathcal{H}_c$. This factor is only relevant in the intermediate state when there are two holes on the same site. The relevant matrix element for third-order perturbation theory is written in Eq. (C4) of Appendix C. In using this result it is convenient to use the identity of Eq. (22). Then we obtain the correction to the energy at second order in $\tilde{\mathbf{t}}$ including perturbatively the leading Coulombic contributions [which we indicate by the superscript "(2,c)"]:

$$\mathcal{H}^{(2,c)}(i,j) = -\Bigg[\sum_{\substack{\sigma\sigma_1\sigma_2\sigma_3\\ ss'}}\sum_{ab}\frac{1}{(U_0+E_{ia})}\frac{1}{(U_0+E_{ib})}(\tilde{t}_{b0}^{ij})_{\sigma_2\sigma_1}(\tilde{t}_{0a}^{ji})_{\sigma_3 s}$$
$$\times[\Delta\tilde{U}_{ss'\sigma_2\sigma}(i;a0b0)+\tilde{K}_{ss'\sigma_2\sigma}(i;a0b0)-\Delta\tilde{U}_{ss'\sigma\sigma_2}(i;a00b)-\tilde{K}_{ss'\sigma\sigma_2}(i;a00b)]$$
$$\times\left(\frac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}\right)_{\sigma s'}\left(\frac{1}{2}+\mathbf{S}(j)\cdot\vec{\sigma}\right)_{\sigma_1\sigma_3}+(i\leftrightarrow j)\Bigg] , \tag{26}$$

where we have used the property $\Delta\tilde{U}_{\sigma\sigma'\sigma_1\sigma_1'}(i;abb'a') = \Delta\tilde{U}_{\sigma'\sigma\sigma_1'\sigma_1}(i;baa'b')$, $\tilde{K}_{\sigma\sigma'\sigma_1\sigma_1'}(i;abb'a') = \tilde{K}_{\sigma'\sigma\sigma_1'\sigma_1}(i;baa'b')$. (In writing the above result we set $E_{i,0} = E_{j,0} = 0$ for simplicity.) In order to carry out the spin summations, we insert here the explicit expressions for $\Delta\tilde{U}$ and $\tilde{K}$, Eqs. (18). This leads to

$$\mathcal{H}^{(2,c)}(i,j) = -\sum_{\alpha\alpha'}[(\Delta U_{\alpha\alpha'}[\mathrm{Tr}\{[\tfrac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}](\mathbf{x}^{ji}_{\alpha'\alpha'})^{\dagger}(\tfrac{1}{2}+\mathbf{S}(j)\cdot\vec{\sigma})\mathbf{x}^{ji}_{\alpha\alpha}\}$$
$$-\mathrm{Tr}\{[\tfrac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}](\mathbf{m}^{i}_{\alpha'0})^{\dagger}\mathbf{m}^{i}_{\alpha'0}\}\mathrm{Tr}\{[\tfrac{1}{2}+\mathbf{S}(j)\cdot\vec{\sigma}]\mathbf{w}^{ji}_{\alpha\alpha}\}]$$
$$+K_{\alpha\alpha'}[\mathrm{Tr}\{[\tfrac{1}{2}+\mathbf{S}_i\cdot\vec{\sigma}](\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}[\tfrac{1}{2}+\mathbf{S}(j)\cdot\vec{\sigma}]\mathbf{x}^{ji}_{\alpha\alpha'}\}$$
$$-\mathrm{Tr}\{[\tfrac{1}{2}+\mathbf{S}(i)\cdot\vec{\sigma}](\mathbf{m}^{i}_{\alpha'0})^{\dagger}\mathbf{m}^{i}_{\alpha0}\}\mathrm{Tr}\{[\tfrac{1}{2}+\mathbf{S}(j)\cdot\vec{\sigma}]\mathbf{w}^{ji}_{\alpha\alpha'}\})+(i\leftrightarrow j)], \tag{27}$$

where

$$\mathbf{x}^{ji}_{\alpha\alpha'} = \sum_{a}\frac{1}{U_0+E_{ia}}\tilde{\mathbf{t}}^{ji}_{0a}(\mathbf{m}^{i}_{\alpha a})^{\dagger}\mathbf{m}^{i}_{\alpha'0} \equiv X^{ji}_{\alpha\alpha'}\mathbf{I}+i\mathbf{Y}^{ji}_{\alpha\alpha'}\cdot\vec{\sigma}, \tag{28a}$$

and

$$\mathbf{w}^{ji}_{\alpha\alpha'} = \sum_{ab}\frac{1}{(U_0+E_{ia})}\frac{1}{(U_0+E_{ib})}\tilde{\mathbf{t}}^{ji}_{0a}(\mathbf{m}^{i}_{\alpha a})^{\dagger}\mathbf{m}^{i}_{\alpha' b}\tilde{\mathbf{t}}^{ij}_{b0} \equiv W^{ji}_{\alpha\alpha'}\mathbf{I}+i\mathbf{Z}^{ji}_{\alpha\alpha'}\cdot\vec{\sigma}, \tag{28b}$$

in which $X^{ji}_{\alpha\alpha'}$ and $W^{ji}_{\alpha\alpha'}$ are real scalars and $\mathbf{Y}^{ji}_{\alpha\alpha'}$ and $\mathbf{Z}^{ji}_{\alpha\alpha'}$ are real vectors. It is straightforward to verify that in Eq. (27) the terms which involve two traces, as well as those coming from the factors of 1/2 and involving one spin variable, do not contribute to spin dependence in the spin Hamiltonian. This follows by noting that (a) $(\mathbf{m}^{i}_{\alpha'0})^{\dagger}\mathbf{m}^{i}_{\alpha'0}$ and $\mathbf{w}^{ji}_{\alpha\alpha}$ are proportional to the unit matrix, and (b) one can interchange α and α' in the sums.

The full effective magnetic Hamiltonian, to order $\mathbf{t}^2$, is obtained by combining $\mathcal{H}^{(2)}(i,j)$, Eq. (24), with $\mathcal{H}^{(2,c)}(i,j)$, Eq. (27). The result has the form of Eq. (2), with

$$J(i,j) = \frac{2}{U_0}\mathrm{Tr}\{(\tilde{\mathbf{t}}^{ji}_{00})(\tilde{\mathbf{t}}^{ij}_{00})\}-\sum_{\alpha\alpha'}\Delta U_{\alpha\alpha'}\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha'\alpha'})^{\dagger}(\mathbf{x}^{ji}_{\alpha\alpha})^{\dagger}+(\mathbf{x}^{ij}_{\alpha'\alpha'})^{\dagger}(\mathbf{x}^{ij}_{\alpha\alpha})^{\dagger}\}$$
$$-\sum_{\alpha\alpha'}K_{\alpha\alpha'}\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}+(\mathbf{x}^{ij}_{\alpha\alpha'})^{\dagger}(\mathbf{x}^{ij}_{\alpha\alpha'})^{\dagger}\}, \tag{29a}$$

$$\mathbf{D}(i,j) = -\frac{i}{U_0}[\mathrm{Tr}\{\tilde{\mathbf{t}}^{ji}_{00}\}\mathrm{Tr}\{\tilde{\mathbf{t}}^{ij}_{00}\vec{\sigma}\}-(i\leftrightarrow j)]$$
$$+\frac{i}{2}\sum_{\alpha\alpha'}\Delta U_{\alpha\alpha'}\{[\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha}\}\mathrm{Tr}\{[\mathbf{x}^{ji}_{\alpha'\alpha'}]^{\dagger}\vec{\sigma}\}-\mathrm{Tr}\{[\mathbf{x}^{ji}_{\alpha'\alpha'}]^{\dagger}\}\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha}\vec{\sigma}\}]-(i\leftrightarrow j)\}$$
$$+\frac{i}{2}\sum_{\alpha\alpha'}K_{\alpha\alpha'}\{[\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha'}\}\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}\vec{\sigma}\}-\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}\}\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha'}\vec{\sigma}\}]-(i\leftrightarrow j)\}\ , \tag{29b}$$

$$\mathbf{M}(i,j) = \frac{1}{U_0}[\mathrm{Tr}\{\tilde{\mathbf{t}}^{ji}_{00}\vec{\sigma}\}\otimes\mathrm{Tr}\{\tilde{\mathbf{t}}^{ij}_{00}\vec{\sigma}\}+(i\leftrightarrow j)]$$
$$-\frac{1}{2}\sum_{\alpha\alpha'}\Delta U_{\alpha\alpha'}\{[\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha}\vec{\sigma}\}\otimes\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha'\alpha'})^{\dagger}\vec{\sigma}\}+\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha'\alpha'})^{\dagger}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha}\vec{\sigma}\}]+(i\leftrightarrow j)\}$$
$$-\frac{1}{2}\sum_{\alpha\alpha'}K_{\alpha\alpha'}\{[\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha'}\vec{\sigma}\}\otimes\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}\vec{\sigma}\}+\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha'}\vec{\sigma}\}]+(i\leftrightarrow j)\}$$
$$+\sum_{\alpha\alpha'}K_{\alpha\alpha'}[\mathrm{Tr}\{(\mathbf{m}^{i}_{\alpha'0})^{\dagger}\mathbf{m}^{i}_{\alpha0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{w}^{ji}_{\alpha\alpha'}\vec{\sigma}\}+(i\leftrightarrow j)]\ . \tag{29c}$$

One notes that when the contributions of $\Delta U_{\alpha\alpha'}$ and $K_{\alpha\alpha'}$ are ignored, Eqs. (29) reproduce Eq. (3), with $\mathbf{D}(i,j) = -i[\mathrm{Tr}\{\tilde{\mathbf{t}}^{ji}_{00}\}\mathrm{Tr}\{\tilde{\mathbf{t}}^{ij}_{00}\vec{\sigma}\}/U_0-(i\leftrightarrow j)]$, and $J = 2\,\mathrm{Tr}\{\tilde{\mathbf{t}}^{ji}_{00}\tilde{\mathbf{t}}^{ij}_{00}\}/U_0$. The results (29) hold for general site symmetry, and to all orders in the spin-orbit coupling. They become particularly simple in the special case of tetragonal symmetry, as is discussed in Sec. IV. In Eq. (29a) we see that even when $\mathbf{U}$ is a constant and $\mathbf{K}=0$, $J(i,j)$ does depend on λ. Moriya's expression for $J(i,j)$ is only correct to zeroth order in λ.

C. The copper—oxygen—copper bond

Here we derive the effective magnetic Hamiltonian of the copper spins for the bond Cu—O—Cu. The spin-orbit interaction on the oxygen is much smaller than that on the copper,[21] and therefore may be neglected. Then the microscopic Hamiltonian (4) is modified as follows. First, the kinetic energy now represents hopping between the oxygen and the copper ions. That is, in place of Eqs. (14) we now have

$$\mathcal{H}_{\rm hop} = \sum_{iq} T_{qi} + \text{H.c.} \ , \tag{30}$$

where

$$T_{qi} = \sum_{an} \sum_{\sigma\sigma'} (\bar{t}^{qi}_{na})_{\sigma\sigma'} p^{\dagger}_{qn\sigma} c_{ia\sigma'} \ , \tag{31}$$

in which $p_{qn\sigma}$ ($p^{\dagger}_{qn\sigma}$) are the destruction (creation) operators for a hole on one of the states (n) of the qth oxygen, and

$$(\bar{t}^{qi}_{na})_{\sigma\sigma'} = \sum_{\alpha} t^{qi}_{n\alpha} (m^{i}_{\alpha a})_{\sigma\sigma'} \ . \tag{32}$$

Here $t^{qi}_{n\alpha}$ describes hopping from the αth orbital on the ith copper ion to the nth orbital on the qth oxygen ion. The matrix element $(\bar{t}^{qi}_{na})_{\sigma\sigma'}$ describes hopping between the copper states [see Eq. (8)] which diagonalize $\mathcal{H}_x + \mathcal{H}_{\rm so}$ and the qth oxygen ion.

Second, we add to the Hamiltonian the on-site single-particle energies and the Coulomb interactions on the oxygen. These terms are written in the form

$$\begin{aligned} \mathcal{H}_p &= \sum_{qn\sigma} \epsilon_n p^{\dagger}_{qn\sigma} p_{qn\sigma} + \frac{1}{2} \sum_{qnn'} \sum_{\sigma\sigma'} U^{(q)}_{p} p^{\dagger}_{qn\sigma} p^{\dagger}_{qn'\sigma'} p_{qn'\sigma'} p_{qn\sigma} \\ &\quad + \frac{1}{2} \sum_{qnn'} \sum_{\sigma\sigma'} \Delta U^{(q)}_{nn'} p^{\dagger}_{qn\sigma} p^{\dagger}_{qn'\sigma'} p_{qn'\sigma'} p_{qn\sigma} + \frac{1}{2} \sum_{qnn'} \sum_{\sigma\sigma'} K^{(q)}_{nn'} p^{\dagger}_{qn\sigma} p^{\dagger}_{qn'\sigma'} p_{qn\sigma'} p_{qn'\sigma} \\ &\equiv \mathcal{H}^{(p)}_0 + \Delta\mathcal{H}^{(p)}_c \ , \end{aligned} \tag{33}$$

where $\mathcal{H}^{(p)}_0$ is the first line of this equation and $\Delta\mathcal{H}^{(p)}_c$ is the second line. Thus, the total Hamiltonian for this case is taken to be $\mathcal{H}_0 + V$, where

$$\mathcal{H}_0 = \mathcal{H}_x + \mathcal{H}_{\rm so} + \mathcal{H}_{c0} + \mathcal{H}^{(p)}_0 \tag{34}$$

and V, which we treat perturbatively, is

$$V = \mathcal{H}_{\rm hop} + \Delta\mathcal{H}_c + \Delta\mathcal{H}^{(p)}_c \ . \tag{35}$$

In the above, the index q distinguishes between oxygens on the bond along the x and y directions from the copper ion in question. However, the perturbation expansion gives results in the form of contributions summed over all pairs of single bonds between nearest-neighboring copper ions i and j. Then, the index q is fixed once the values of i and j are specified, as one sees from Fig. 1. Accordingly, we henceforth omit the index q, so that, for instance, $\bar{\mathbf{t}}^{qi}_{na} \to \bar{\mathbf{t}}^{i}_{na}$, $\bar{\mathbf{t}}^{iq}_{an} \to \bar{\mathbf{t}}^{i}_{an}$, $T_{qi} \to T_i$, and $p^{\dagger}_{qn\sigma} \to p^{\dagger}_{n\sigma}$.

We now turn to the perturbation expansion, from which we obtain the magnetic Hamiltonian of the copper spins. It is clear that the lowest-order contribution to the effective interaction between two copper spins is of order $\mathbf{t}^4$. There are two possible channels in this order, which we denote by a and b. In channel a, the hole is transferred from one of the coppers to the oxygen, then to the second copper, and then back to the first copper via the oxygen. Hence in this channel there are two holes on the *copper* in the intermediate state. In channel b, the hole is transferred from one of the Cu ions to the oxygen, and then a second hole is taken from the second copper to the same oxygen. Afterwards the two holes return to the coppers, i.e., back to the ground state in which there is one hole on each Cu ion. Thus in channel b there are two holes on the *oxygen* in the intermediate state. When the terms coming from the Coulomb interactions $\Delta U_{nn'}$ and $K_{nn'}$ for the oxygen ions are included, then their effect will appear only in channel b, in which the two holes have a state where both are on the oxygen.

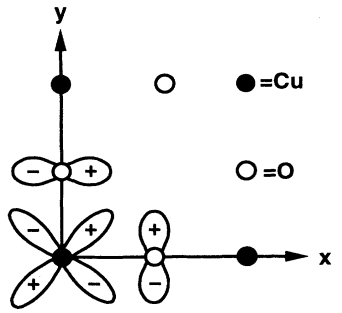

FIG. 1. A CuO plaquette. Here we distinguish between "y" oxygen ions (on y-directed bonds) and "x" oxygen ions (on x-directed bonds.) In each case we show a p orbital on the oxygen ion to which we give the symmetry label, z, since these orbitals can only hop to an orbital on a copper ion with that same symmetry label, i.e., to $\psi_z \sim xy$, which is also shown.

It turns out that for channel a all our previous expressions, derived for the Cu—Cu bond, hold with the replacement

$$\tilde{\mathbf{t}}^{ij}_{ab} = \sum_{n} \frac{\bar{\mathbf{t}}^{i}_{an} \bar{\mathbf{t}}^{j}_{nb}}{\epsilon_n} \ . \tag{36}$$

We show this explicitly in Appendix E for the $\bar{\mathbf{t}}^4$ process. Similar arguments hold for the processes of order $\bar{\mathbf{t}}^4 \Delta\tilde{\mathbf{U}}$ and $\bar{\mathbf{t}}^4 \tilde{\mathbf{K}}$, where $\Delta\tilde{\mathbf{U}}$ and $\tilde{\mathbf{K}}$ are the Coulomb interactions on the copper ion. (Note that for this channel $\Delta\tilde{U}$ and $\tilde{K}$ represent the Coulomb interactions on the copper.)

It thus remains to investigate the perturbation expansion in channel b. Applying once to ψ_0 the term in the Hamiltonian of Eq. (30), which describes hopping from the copper ions to the intervening oxygen ion, one obtains

$$\begin{aligned} |\,\psi_1\rangle &\equiv \sum_{\sigma\sigma_1} (T_i + T_j)\, c^{\dagger}_{i0\sigma} c^{\dagger}_{j0\sigma_1} c_{j0\sigma_1} c_{i0\sigma} \,|\,\psi_0\rangle \\ &= \sum_{n} \sum_{\sigma\sigma_1\sigma_2} [(\bar{t}^{i}_{n0})_{\sigma_2\sigma} p^{\dagger}_{n\sigma_2} c^{\dagger}_{j0\sigma_1} \\ &\quad + (\bar{t}^{j}_{n0})_{\sigma_2\sigma_1} c^{\dagger}_{i0\sigma} p^{\dagger}_{n\sigma_2}] c_{j0\sigma_1} c_{i0\sigma} \,|\,\psi_0\rangle \ , \end{aligned} \tag{37}$$

which represents virtual states with energy ϵ_n. In the next order, the second hole is put on the same oxygen. This leads to

$$|\psi_{2b}\rangle \equiv \sum_{\sigma\sigma_1}\left(\frac{1}{\mathcal{H}_0}T_j\frac{1}{\mathcal{H}_0}T_i+\frac{1}{\mathcal{H}_0}T_i\frac{1}{\mathcal{H}_0}T_j\right)c^\dagger_{i0\sigma}c^\dagger_{j0\sigma_1}c_{j0\sigma_1}c_{i0\sigma}|\psi_0\rangle$$
$$=\sum_{nn'}\sum_{\sigma\sigma_1}\sum_{\sigma_2\sigma_3}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}(\bar{t}^i_{n0})_{\sigma_2\sigma}(\bar{t}^j_{n'0})_{\sigma_3\sigma_1}p^\dagger_{n\sigma_2}p^\dagger_{n'\sigma_3}c_{j0\sigma_1}c_{i\sigma\sigma}\mid\psi_0\rangle\ . \tag{38}$$

In order to return to the ground state two more powers of the hopping are needed. This gives

$$\mathcal{H}^{(2b)}(i,j)=-\sum_{nn'}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)^2\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}[\mathrm{Tr}\{[\tfrac{1}{2}+\vec{\sigma}\cdot\mathbf{S}(i)]\bar{\mathbf{t}}^i_{0n}\bar{\mathbf{t}}^i_{n0}\}$$
$$\times\mathrm{Tr}\{[\tfrac{1}{2}+\vec{\sigma}\cdot\mathbf{S}(j)]\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^j_{n'0}\}-\mathrm{Tr}\{[\tfrac{1}{2}+\vec{\sigma}\cdot\mathbf{S}(i)]\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}[\tfrac{1}{2}+\vec{\sigma}\cdot\mathbf{S}(j)]\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\}]\ , \tag{39}$$

where we used the identity Eq. (22). (We labeled this contribution with a superscript 2 because even though it is fourth order in the $\bar{\mathbf{t}}$'s, it is really a second-order process in terms of a renormalized Cu-Cu hopping interaction. The superscript "b" indicates a contribution from channel b.) In a similar way to the arguments given after Eq. (23), one can convince oneself that the first term in Eq. (39) as well as the terms coming from the factors of $1/2$ in the second term do not contribute to the spin Hamiltonian. Thus, to order $\bar{\mathbf{t}}^4$, the contribution of channel b is

$$\mathcal{H}^{(2b)}(i,j)=\sum_{nn'}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)^2\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}\mathrm{Tr}\{\vec{\sigma}\cdot\mathbf{S}(i)\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\cdot\mathbf{S}(j)\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\}\ . \tag{40}$$

Next we calculate the effect of the Coulomb terms $\Delta U_{nn'}$ and $K_{nn'}$ of Eq. (33). To this end we apply them to the state $|\psi_{2b}\rangle$ of Eq. (38). The result is

$$-\frac{1}{\mathcal{H}_0}(\Delta\mathbf{U}+\mathbf{K})|\psi_{2b}\rangle=-\sum_{nn'}\sum_{\sigma\sigma_1}\sum_{\sigma_2\sigma_3}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)\left(\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}\right)^2(\bar{t}^i_{n0})_{\sigma_2\sigma}(\bar{t}^j_{n'0})_{\sigma_3\sigma_1}$$
$$\times[\Delta U_{nn'}p^\dagger_{n\sigma_2}p^\dagger_{n'\sigma_3}+K_{nn'}p^\dagger_{n'\sigma_2}p^\dagger_{n\sigma_3}]c_{j0\sigma_1}c_{i0\sigma}\mid\psi_0\rangle. \tag{41}$$

Finally we apply two factors of the hopping which bring the holes back to the ground state. This leads to

$$\mathcal{H}^{(2b,c)}(i,j)=\sum_{nn'}\left[\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}\right]^2[-\Delta U_{nn'}\mathrm{Tr}\{\vec{\sigma}\cdot\mathbf{S}(i)\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\cdot\mathbf{S}(j)\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\}$$
$$+K_{nn'}\mathrm{Tr}\{\vec{\sigma}\cdot\mathbf{S}(i)\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^i_{n0}\}\mathrm{Tr}\{\vec{\sigma}\cdot\mathbf{S}(j)\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^j_{n'0}\}-K_{nn'}\mathrm{Tr}\{\vec{\sigma}\cdot\mathbf{S}(i)\bar{\mathbf{t}}^i_{0n}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\cdot\mathbf{S}(j)\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^i_{n0}\}]. \tag{42}$$

Combining Eqs. (40) and (42) we obtain the magnetic interaction arising from channel b in the form of Eq. (2), with the nearest-neighbor interactions

$$J^{(b)}(i,j)=\sum_{nn'}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)^2\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}\left(1-\frac{\Delta U_{nn'}}{\epsilon_n+\epsilon_{n'}+U_p}\right)\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}(\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0})^\dagger\}$$
$$-\sum_{nn'}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)^2\frac{K_{nn'}}{(\epsilon_n+\epsilon_{n'}+U_p)^2}\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n}\bar{\mathbf{t}}^j_{n'0}(\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^i_{n0})^\dagger\}\ , \tag{43a}$$

$$\mathbf{D}^{(b)}(i,j)=-\frac{i}{2}\sum_{nn'}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)^2\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}\left[\left(1-\frac{\Delta U_{nn'}}{\epsilon_n+\epsilon_{n'}+U_p}\right)\right.$$
$$\times(\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\}\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\}-\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}\}\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\})$$
$$\left.-\frac{K_{nn'}}{\epsilon_n+\epsilon_{n'}+U_p}(\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^i_{n0}\}\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\}-\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n}\bar{\mathbf{t}}^j_{n'0}\}\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\})\right], \tag{43b}$$

$$\mathbf{M}^{(b)}(i,j)=\frac{1}{2}\sum_{nn'}\left(\frac{1}{\epsilon_n}+\frac{1}{\epsilon_{n'}}\right)^2\frac{1}{\epsilon_n+\epsilon_{n'}+U_p}\left[\left(1-\frac{\Delta U_{nn'}}{\epsilon_n+\epsilon_{n'}+U_p}\right)\right.$$
$$\times(\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\}+\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\})$$
$$+\frac{K_{nn'}}{\epsilon_n+\epsilon_{n'}+U_p}\left(2\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n'}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\}\otimes\mathrm{Tr}\left\{\bar{\mathbf{t}}^j_{0n}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\right\}\right.$$
$$\left.\left.-\mathrm{Tr}\{\bar{\mathbf{t}}^i_{0n}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\}-\mathrm{Tr}\{\bar{\mathbf{t}}^j_{0n'}\bar{\mathbf{t}}^i_{n0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\bar{\mathbf{t}}^i_{on}\bar{\mathbf{t}}^j_{n'0}\vec{\sigma}\}\right)\right]. \tag{43c}$$

The full magnetic Hamiltonian for the copper spins of the Cu—O—Cu bond is obtained by combining the results of Eq. (43) for channel b, with those for channel a given by Eqs. (29), in conjunction with the identification of Eq. (36). These results generalize those of Refs. 15, 16, and 22, which were obtained in the absence of the Coulomb terms $\Delta\mathbf{U}$ and $\mathbf{K}$.

IV. TETRAGONAL SYMMETRY

This section consists of three subsections. In Sec. IV A, we apply a canonical transformation to show that without Coulomb exchange interactions the effective spin Hamiltonian is isotropic at all orders of t and λ. In Sec. IV B, we use this theorem to isolate the most important contribution to the anisotropy, namely, that involving the Coulomb exchange energy. In Sec. IV C we corroborate our analytical results of perturbation theory by numerical solutions for single-bond clusters: Cu—Cu and Cu—O—Cu.

A. Canonical transformation

We start by proving our strongest result, namely, that the spin Hamiltonian $\mathcal{H}_S$ arising from the generic model is isotropic for a wide class of models in the absence of Coulomb exchange interactions. In particular, this result holds for a commonly used model of the cuprates, in which interionic Coulomb interactions, $\mathbf{V}$ and $\mathbf{M}$ in Eq. (4), and the Coulomb exchange terms $\mathbf{K}$ are neglected, hopping is between nearest-neighboring Cu ions, and the site symmetry is tetragonal. Strictly speaking, the only use we make of site symmetry is that it has to be high enough so that the $3d$ spatial orbitals which diagonalize the crystal field Hamiltonian are $\psi_0(\mathbf{r}) \sim x^2 - y^2$, $\psi_1(\mathbf{r}) \sim 3z^2 - r^2$, $\psi_x(\mathbf{r}) \sim yz$, $\psi_y(\mathbf{r}) \sim xz$, and $\psi_z(\mathbf{r}) \sim xy$. Here the z axis coincides with the tetragonal c axis and the x and y axes coincide with the nearest-neighbor directions in the plane perpendicular to the c axis, as shown in Fig. 1. These symmetry labels are chosen so that $\psi_\alpha(\mathbf{r})$ transforms (under the operations of tetragonal symmetry) like L_α for $\alpha = x, y, z$ and $\psi_0(\mathbf{r})$ and $\psi_1(\mathbf{r})$ transform like scalars.

An important observation is that the orbital angular momentum operator $\mathbf{L}$ has matrix elements only between states of specific symmetry. For instance, L_x connects $\psi_x(\mathbf{r})$ only to the states $\psi_0(\mathbf{r})$ and $\psi_1(\mathbf{r})$ and it connects $\psi_y(\mathbf{r})$ to $\psi_z(\mathbf{r})$ and vice versa. Similar statements can be made about the other components of $\mathbf{L}$. We now introduce a transformation in spin space (to pseudospin) such that the spin-orbit interaction is diagonal with respect to pseudospin. For that purpose we introduce pseudospin $\vec{\mu}$ as follows:

$$|\alpha,\mu\rangle = \psi_\alpha(\mathbf{r}) \sum_\eta [\sigma_\alpha]^*_{\mu,\eta} \phi_\eta \equiv f^\dagger_{\alpha,\mu}|\text{vac}\rangle \ , \tag{44}$$

where ϕ_η is a spin function for spin "up" if $\eta = 1/2$ and for spin "down" if $\eta = -1/2$, and $|\text{vac}\rangle$ denotes the vacuum state. Here σ_α for $\alpha = x, y, z$ are the Pauli matrices, and $\sigma_0 = \sigma_1 = \mathbf{I}$ is the unit matrix. As discussed in Appendix F, the above transformation is such that the spin-orbit interaction is diagonal in pseudospin:

$$\mathcal{H}_{\text{so}} = \lambda \sum_{k=i,j} \sum_{\alpha,\beta,\mu} W(k,\alpha,\beta) f^\dagger_{k\alpha\mu} f_{k\beta\mu} \ , \tag{45}$$

where $W(k,\alpha,\beta)$ is a spin-orbit matrix element. [The transformation of Eq. (44) should not be confused with Eq. (8). The latter involves an exact diagonalization and requires a knowledge of all the parameters. In contrast, the transformation of Eq. (44) is independent of the interaction parameters. It merely brings the Hamiltonian into block diagonal form in which there are two identical blocks, one for $\mu = \frac{1}{2}$ and one for $\mu = -\frac{1}{2}$.] Because the unitary transformation of Eq. (44) does not mix spatial states, it does not affect the form of the Coulomb interactions scaled by $\mathbf{U}$. Furthermore, in view of the lattice symmetry hopping can only involve holes moving from one site to a neighboring site *without changing their symmetry*. Thus holes in a state ψ_α on one ion, where $\alpha = x, y$, or z, can only hop to states of the same α on a nearest-neighboring ion. Likewise, holes in states $\psi_0(\mathbf{r})$ or $\psi_1(\mathbf{r})$ on one ion can only hop to states $\psi_0(\mathbf{r})$ or $\psi_1(\mathbf{r})$ on an adjacent ion. Since states α and β which are connected by hopping must be states of the *same* symmetry, we have

$$T_{ij} = \sum_{\alpha,\beta,\mu} t_{i\alpha,j\beta} d^\dagger_{i\alpha\mu} d_{j\beta\mu} = \sum_{\alpha,\beta,\mu,\rho,\tau} t_{i\alpha,j\beta} [\sigma_\alpha]_{\tau\mu} f^\dagger_{i\alpha\tau} [\sigma_\beta]^*_{\rho\mu} f_{j\beta\rho} = \sum_{\alpha,\beta,\rho} t_{i\alpha,j\beta} f^\dagger_{i\alpha\rho} f_{j\beta\rho} \ . \tag{46}$$

In other words, the total Hamiltonian (for $\mathbf{K} = 0$) can be written in the form

$$\mathcal{H} = \sum_k \sum_{\alpha,\mu} \epsilon_{k\alpha} f^\dagger_{k\alpha\mu} f_{k\alpha\mu} + \lambda \sum_k \sum_{\alpha,\beta,\mu} W(k,\alpha,\beta) f^\dagger_{k\alpha\mu} f_{k\beta\mu} + \sum_{i,j,\alpha,\beta,\mu} t_{i\alpha,j\beta} f^\dagger_{i\alpha\mu} f_{j\beta\mu}$$
$$+ \frac{1}{2} \sum_k \sum_{\alpha,\alpha',\mu,\mu'} U_{k\alpha,k\alpha'} f^\dagger_{k\alpha\mu} f^\dagger_{k\alpha'\mu'} f_{k\alpha'\mu'} f_{k\alpha\mu} . \tag{47}$$

Thus this Hamiltonian can be written[23] in terms of the quantities

$$Q_{\alpha\beta}(i,j) \equiv \sum_\mu f^\dagger_{i\alpha\mu} f_{j\beta\mu} \ , \tag{48}$$

which themselves are invariant under rotations in pseudospin space. Therefore $\mathcal{H}$ is invariant under rotations in pseudospin space. To construct the effective spin Hamiltonian $\mathcal{H}(i,j)$ involves using degenerate perturba-

tion theory to eliminate the excited states. Accordingly, it is clear that the resulting spin Hamiltonian will be rotationally invariant in pseudospin space. Since we have defined pseudospin so that in the ground state (in which all the holes are in the state ψ_0) pseudospin and real spin are identical, it follows that for the tetragonal case with no Coulomb exchange interactions the spin Hamiltonian is also rotationally invariant. This theorem indicates that even though $\mathcal{H}_S$ may include further-neighbor two-spin interactions, four-spin interactions, etc., it is nevertheless rotationally invariant, so that the spin-wave spectrum cannot have a gap at zero wave vector in the absence of Coulomb exchange terms.

For the case of the nearest-neighbor exchange interaction we can make some further explicit statements. For tetragonal symmetry, $J_{\mu\nu}(i,j)$ must also be diagonal (with its principal axes along the tetragonal axes). Thus, $\mathcal{H}(i,j)$ is an isotropic Heisenberg model.

The above theorem can be generalized to include the intervening oxygen ions. Here we consider a Hubbard model which includes the three $2p$ spatial orbitals. Now we introduce different unitary transformations for oxygen ions on y-directed and x-directed bonds (see Fig. 1). For those on y-directed bonds we set $\psi_0(\mathbf{r}) = |2p_y\rangle$, $\psi_x(\mathbf{r}) = |2p_z\rangle$, and $\psi_z(\mathbf{r}) = |2p_x\rangle$. We then introduce states $|\alpha,\mu\rangle$ by

$$|\alpha,\mu\rangle = \psi_\alpha(\mathbf{r}) \sum_\eta [\sigma_\alpha]^*_{\mu,\eta} \phi_\eta(\sigma) \ , \tag{49}$$

similar to Eq. (44) which was used for the Cu d states. We need to examine how the hopping and spin-orbit interactions are affected by this transformation. Note that hopping along the y direction can only take place between Cu states like x^2-y^2 and oxygen $2p_y$ states. These are both associated with symmetry 0 or 1. Likewise, an oxygen $2p_z$ orbital can only hop to a copper yz state, both of which have symmetry label x. Also an oxygen $2p_x$ orbital can only hop to a copper xy state, both of which have symmetry label z. Thus with this labeling of states, hopping occurs only between states of the same symmetry label and the canonical transformation has no effect on the hopping, just as in Eq. (46). One can verify that the spin-orbit interaction on the oxygen ions does conserve pseudospin. Oxygen ions on the x-directed bonds are treated analogously. For them we write $\psi_0(\mathbf{r}) = |2p_x\rangle$, $\psi_y(\mathbf{r}) = |2p_z\rangle$, and $\psi_z(\mathbf{r}) = |2p_y\rangle$ and we again use Eq. (49). Then, we conclude that the Hamiltonian of the entire lattice can be expressed in terms of the quantities $Q_{\alpha\beta}(i,j)$. Thus the theorem holds with intervening oxygen ions: in the absence of Coulomb exchange, this model gives no anisotropy in $\mathcal{H}_S$ for a tetragonal lattice.

The fact that this conclusion is demonstrated to all orders in perturbation theory represents an important new result. The low-order perturbation result of BS is in accord with this theorem. As mentioned there, this conclusion modifies the conventional wisdom that the anisotropy in the exchange interaction is trivially related to the anisotropy of the g tensor. Finally, we emphasize that this theorem depends crucially on the fact that the eigenstates of the crystal field are those of tetragonal site symmetry and that hopping is only between nearest neighbors. In addition, the theorem is only valid for Coulomb terms which have the Hartree form, i.e., those which only involve two orbitals, as in Eq. (4).

B. Anisotropy

Anisotropy can occur via various mechanisms. One such mechanism is to introduce Coulomb exchange interactions, as done implicitly by BS. In the context of the above discussion we note that exchange interactions compete with spin-orbit interactions in the following sense. With only the former interactions the eigenstates of a single ion are states of total real spin 1 or 0. With no exchange but with spin-orbit interactions, the eigenstates of a single ion are states of total pseudospin 1 or 0. In both cases, our numerical evaluation of the energy levels gave singlets and triplets as this argument requires. However, the wave functions are different, of course. When both interactions are present, the degeneracies are removed because wave functions cannot be simultaneous eigenfunctions of both real spin and pseudospin. From our calculation, treating hopping, spin-orbit, ΔU, and exchange interactions as perturbations, we reach the following conclusions. For the "generic" model, anisotropic exchange appears at order $\mathbf{t}^2\lambda^2\mathbf{K}$. An efficient way to perform this calculation is to use the hopping matrix elements, $\tilde{\mathbf{t}}$, of Eq. (14b) and work to order $\tilde{\mathbf{t}}^2\mathbf{K}$, as we did in Sec. III. The perturbation expansion yields the general expression, given in Eq. (27). In Appendix G we analyze this expression for tetragonal symmetry and find that it agrees with the result given previously,[18] namely,

$$J^{\text{anis}}_{\mu\mu} = -2\lambda^2 \left\{ \frac{|\, L^\mu_{0,\mu}\,|^2\, t^2_{0,1} K_{1,\mu}}{(\epsilon_\mu + \epsilon_1 + U_{1,\mu})^2} \left[\frac{1}{\epsilon_\mu} + \frac{1}{\epsilon_1 + U_{0,1}} \right]^2 + \frac{K_{0,\mu}}{(\epsilon_\mu + U_{0,\mu})^2} \left| \frac{(t_{\mu,\mu} - t_{0,0}) L^\mu_{0,\mu}}{\epsilon_\mu} + \frac{t_{0,1} L^\mu_{1,\mu}}{\epsilon_1 + U_{0,1}} \right|^2 \right\}, \tag{50}$$

where $L^\mu_{\alpha\beta}$ denotes the orbital angular momentum matrix element, $\langle\alpha|L_\mu|\beta\rangle$, μ assumes the values x, y, and z, and the superscript "anis" indicates that we have arbitrarily omitted isotropic (i.e., μ-independent) contributions. The same expression is also derived directly from perturbation theory in $\mathbf{t}$, λ, and $\mathbf{K}$ in Appendix H.

Now we briefly discuss the implications of the above result. First of all, note that within tetragonal symmetry the result does display the expected full anisotropy for a single bond, under which $J_{xx} \equiv J_\parallel$, $J_{yy} \equiv J_\perp$, and J_{zz} are all different. To get biaxiality ($J_\parallel \neq J_\perp$) requires either $t_{01} \neq 0$ or $t_{xx} \neq t_{yy}$; see, Fig. 2. Of course, in tetragonal symmetry single-site quantities cannot differentiate between the x ($\parallel$) and y ($\perp$) directions. To understand why $t_{01} \neq 0$ introduces biaxiality, note that

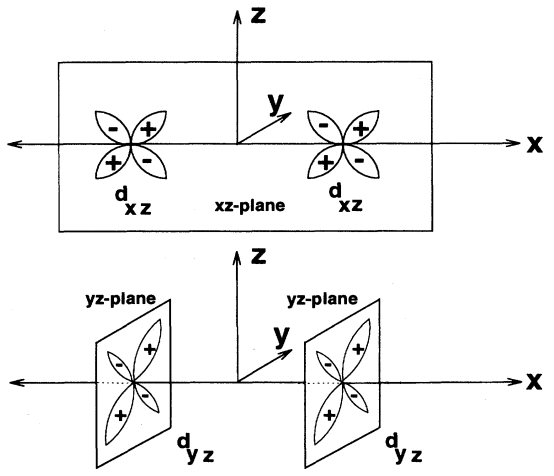

FIG. 2. Schematic view of $|y\rangle = d_{xz}$ and $|x\rangle = d_{yz}$ of two Cu ions when they are on the x axis. Note that while d_{xz} orbitals are in the same plane, the plane of d_{yz} orbitals are parallel to each other. Hence there is no reason that $t_{x,x}$ should be equal to $t_{y,y}$.

t_{01} changes sign when the local x coordinates are rotated into the y coordinates. Also note that even in the limit when U is considered to be very large, the result still does depend on the hopping between excited levels through $t_{\mu\mu}$. Finally, we remark that these expressions differ in several respects from those of BS. This point is discussed in Appendix I, where we give the results more explicitly.

C. With oxygens

Turning now to the Cu—O—Cu bond in the tetragonal symmetry, we again discuss separately the anisotropy resulting from channel a (the two holes occupy the same copper in the intermediate state) and that coming from channel b (the two holes are on the oxygen in the intermediate state).

For channel a, we use the transformation of Eq. (36). Using Eqs. (G5) and (32) we obtain

$$\left(\tilde{t}^{ij}_{ab}\right)_{\sigma_1\sigma_2} = \sum_{n\alpha\beta} \frac{1}{\epsilon_n} t^i_{\alpha n} t^j_{n\beta} \bar{m}^*_{\alpha a} \bar{m}_{\beta b} (\sigma^\dagger_\alpha \sigma_\beta)_{\sigma_1\sigma_2} , \quad (51)$$

where $\bar{m}_{\alpha a}$ are scalars [see Eq. (G5)]. For tetragonal symmetry, α and β belong to the same symmetry class (e.g., $\alpha = \beta$ for $\alpha = x, y$, or z, or α and β are 0 or 1). Hence $\tilde{\mathbf{t}}^{ij}_{ab}$ becomes the unit matrix times a scalar given by

$$\tilde{t}^{ij}_{ab} = \sum_{\alpha\beta} t^{ij}_{\alpha\beta} \bar{m}^*_{\alpha a} \bar{m}_{\beta b} , \quad t^{ij}_{\alpha\beta} = \sum_n \frac{1}{\epsilon_n} t^i_{\alpha n} t^j_{n\beta} , \quad (52)$$

in which α and β refer to tetragonal d states. The calculation can now proceed exactly as for the "generic" model described in Appendix G (or H), with the effective hopping matrix elements $t_{\alpha\beta}$ given by

$$t_{00} = t^2_{0p_x}/\epsilon_{p_x}, \quad t_{01} = t_{0p_x} t_{p_x 1}/\epsilon_{p_x},$$
$$t_{xx} = 0, \quad t_{yy} = t^2_{yp_z}/\epsilon_{p_z}, \quad t_{zz} = t^2_{zp_y}/\epsilon_{p_y}, \quad (53)$$

where p_x, p_y, and p_z represent the p states on the oxygen. Thus the contribution of channel a to the symmetric anisotropy of the spin Hamiltonian for a bond along the x direction is reproduced by Eq. (50), with the replacements (53). Analogous expressions hold for a Cu—O—Cu bond along the y direction.

Now let us consider the magnetic anisotropy in channel b. Inspection of Eqs. (43b) and (43c) shows that we need to examine the 2×2 matrices $\bar{\mathbf{t}}_{0n}\bar{\mathbf{t}}_{n'0}$. (We omit the site indices which are irrelevant for the tetragonal symmetry.) Using Eq. (32) and tetragonal symmetry, we write

$$\bar{\mathbf{t}}_{0n}\bar{\mathbf{t}}_{n'0} = \sum_{\alpha\beta} t_{\alpha n} t_{n'\beta} \mathbf{m}^\dagger_{\alpha 0} \mathbf{m}_{\beta 0}$$
$$= \sum_\alpha t_{\alpha n} t_{n'\alpha} \mathbf{m}^\dagger_{\alpha 0} \mathbf{m}_{\alpha 0} . \quad (54)$$

Therefore, they are proportional to the unit matrix. As a result, there is no contribution to the magnetic anisotropy in channel b in order $\mathbf{t}^4$. The reasons are that there is no spin-orbit coupling on the oxygen, and that in this channel (and to this order) the excited states on the copper are not visited at all. Therefore, just as is the case for the Cu—Cu bond when those states are ignored [cf. Eq. (24)] the magnetic Hamiltonian resulting from this channel is isotropic.

D. Numerical study

We have checked our analytical results of perturbation theory against results (shown in Fig. 3) obtained from exact diagonalization for the four lowest levels out of the

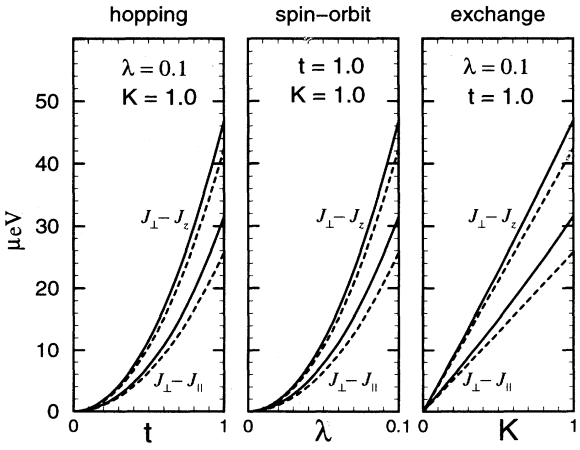

FIG. 3. Comparison of perturbation results (dotted line) with the exact results (solid line). Here $J_\parallel$, $J_\perp$, and J_z correspond to J_{xx}, J_{yy}, and J_{zz} of Eq. (50). The hopping matrix elements $t_{\alpha,\beta}$ are estimated from Eq. (53) as explained in the text. In the left and right panel the $t_{\alpha,\beta}$ and $K_{\alpha,\beta}$ are replaced by $t\ t_{\alpha,\beta}$ and $K\ K_{\alpha,\beta}$, respectively. The values of λ (in eV), t, and K are given in the panels.

TABLE I. Values in eV of parameters used.

U_0[a]	A[b]	B[b,c]	C[b,c]	$(pd\sigma)$[d]	ϵ_1[e]	$\epsilon_x \equiv \epsilon_y$[e]	ϵ_z[e]	ϵ_{p_x}[f]	ϵ_{p_y}[f]	ϵ_{p_z}[f]
9.34	7.00	0.15	0.58	1.5	1.8	1.8	1.8	3.25	3.25	3.25

[a]Refs. 25 and 26 use $U_0 = 8.8$. Local density calculations of Refs. 28 and 27 give 10.5 and 9.4, respectively.
[b]The Racah coefficients, A, B, and C are defined in Ref. 29.
[c]For the solid, the values of B and C are appropriately taken from the free-ion optical values of Ref. 30, as is discussed by Eskes, Tjeng, and Sawatsky (Ref. 31).
[d]See Ref. 32.
[e]See Ref. 31.
[f]Reference 32 gives $\epsilon_p = 3.5$, but smaller values of ϵ_p are plausible (Ref. 17).

190 possible two-hole states for a pair of Cu sites. The relations between the exchange constants and the four lowest levels are obtained in Appendix A.

For our numerical results shown in Fig. 3, we used the values of the parameters listed in Table I,[25–32] together with $\lambda = 0.1$. The hopping matrix elements are related to $(pd\sigma)$ as follows: $t_{0,p_x} = -\sqrt{3}t_{1,p_x} = \frac{\sqrt{3}}{2}(pd\sigma)$ and $t_{y,p_z} = t_{z,p_y} = (pd\pi)$, with $(pd\pi) \approx -\frac{1}{2}(pd\sigma)$.[33] The expressions for $U_{\alpha,\beta}$ and $K_{\alpha,\beta}$ in terms of the Racah parameters were taken from Ref. 29 and are listed in Appendix B. We also checked that the $J^{\rm anis}_{\mu\mu}$, shown in Fig. 3, agree to within about 10% with those obtained from the full 325 site Hamiltonian for the Cu—O—Cu cluster. Very crudely, as A increases above 7 eV, the $J^{\rm anis}_{\mu\mu}$ are inversely proportional to A^{-2} and are proportional to a linear combination of B and C. (When $B = C = 0$, our theorem indicates that there is no anisotropy in $J_{\mu\mu}$.) Thus our results are not highly sensitive to increasing the value of A. As A is decreased below about 6 eV, perturbation theory rapidly becomes increasingly inaccurate. Now we discuss briefly the numerical values of the Racah parameters. We took the values of B and C from Ref. 31. Then, fixing the value of A is equivalent to fixing the value of $U_0 = A + 4B + 3C$. Recently proposed values for U_0 are 8.8, 8.8, 9.4, and 10.5 eV from Refs. 25, 26, 27, and 28, respectively. As a compromise, we took $U_0 = 9.34$ or $A = 7$ eV. Our parameters yield an anisotropy in J of order 0.03 meV and, as we shall see in the next section, give an out-of-plane gap in the spin-wave spectrum within 10% of the experimental[9] value 5 meV.

V. SPIN WAVE SPECTRUM OF THE EFFECTIVE SPIN HAMILTONIAN IN TETRAGONAL SYMMETRY

Given Eqs. (13) and (50) for single bonds, the classical ground state of the effective spin Hamiltonian is rotationally invariant in the basal plane. The out-of-plane anisotropy $\alpha_{XY} \approx \Delta J/J_0$ is positive (see Fig. 3), and therefore the spins order in that plane, as is well established. In the absence of spin-wave fluctuations, the in-plane gap is zero. However, the classical rotational invariance within the basal plane is broken by the dependence of the spin-wave energies on the angle θ between the staggered magnetization and the crystal x axis. The purpose of this section is to study this anisotropy and show that it leads to a nonvanishing in-plane gap in the spin-wave spectrum.[18,19]

In order to show this, we start with the following general Hamiltonian for the CuO_2 plane in a tetragonal system

$$H_{\rm eff} = \sum_{\langle ij \rangle} H_{ij} \ , \tag{55}$$

where for $\langle ij \rangle$ along the x direction, H_{ij} is

$$H_{ij} = J_{\parallel} S_i^x S_j^x + J_{\perp} S_i^y S_j^y + J_z S_i^z S_j^z , \tag{56}$$

and for $\langle ij \rangle$ along the y direction H_{ij} is

$$H_{ij} = J_{\perp} S_i^x S_j^x + J_{\parallel} S_i^y S_j^y + J_z S_i^z S_j^z \ . \tag{57}$$

We will now calculate the spin-wave spectrum of this Hamiltonian and then the first quantum correction to the classical ground state energy. We consider the case where the spins lie in the xy plane and are ordered antiferromagnetically ($J_{\parallel}, J_{\perp} > J_z > 0$). Assuming the staggered magnetization moment makes an angle θ with the positive x axis, we use the following transformation so that spins are parallel to the new z axis:

$$\mathbf{S}_i = \begin{pmatrix} 0 & -\sin\theta & \cos\theta \\ 0 & \cos\theta & \sin\theta \\ -1 & 0 & 0 \end{pmatrix} \mathbf{S}'_i \ . \tag{58}$$

Defining sublattice A to have up spins (in the rotated frame) and sublattice B to have down spins (in the rotated frame), we have the following bosonic spin representation:

$$S'^x_i = \sqrt{\frac{S}{2}}\,[a_i + a_i^+] \ , \quad S'^y_i = -i\sqrt{\frac{S}{2}}\,[a_i - a_i^+] \ ,$$
$$S'^z_i = S - a_i^+ a_i \tag{59}$$

for sublattice A, and

$$S'^x_j = \sqrt{\frac{S}{2}}\,[b_j + b_j^+] \ , \quad S'^y_j = i\sqrt{\frac{S}{2}}\,[b_j - b_j^+] \ ,$$
$$S'^z_j = -S + b_j^+ b_j \tag{60}$$

for sublattice B. For later convenience we consider the case of general spin, although in the end we set $S = \frac{1}{2}$. Using Eqs. (58)–(60) we may write the effective spin Hamiltonian $H_{\rm eff}$ given in Eq. (55) in momentum space as

$$H_{\rm eff} = E_0 + 4J_{\rm av} S \sum_{\mathbf{q}} [a^\dagger_{\mathbf{q}} a_{\mathbf{q}} + b^\dagger_{\mathbf{q}} b_{\mathbf{q}} + (A_{\mathbf{q}} a_{\mathbf{q}} b_{-\mathbf{q}} + B_{\mathbf{q}} a_{\mathbf{q}} b^\dagger_{\mathbf{q}} + \text{H.c.})] \ , \tag{61}$$

where $\mathbf{q}$ is summed over the first Brillouin zone of the magnetic reciprocal lattice and

$$E_0 = -2J_{\rm av}NS^2 \ , \quad J_{\rm av} = \tfrac{1}{2}(J_\parallel + J_\perp) \ ,$$
$$A_{\mathbf{q}} = \frac{1}{4J_{\rm av}}[J_1\cos(q_x a) + J_2\cos(q_y a)] \ ,$$
$$B_{\mathbf{q}} = -\frac{1}{4J_{\rm av}}[J_3\cos(q_x a) + J_4\cos(q_y a)] \ . \qquad (62)$$

Here N is the total number of spins and

$$J_1 = J_\parallel \sin^2\theta + J_\perp\cos^2\theta + J_z \ ,$$
$$J_2 = J_\parallel \cos^2\theta + J_\perp\sin^2\theta + J_z \ ,$$
$$J_3 = J_\parallel \sin^2\theta + J_\perp\cos^2\theta - J_z \ ,$$
$$J_4 = J_\parallel \cos^2\theta + J_\perp\sin^2\theta - J_z. \qquad (63)$$

Henceforth we will set the lattice constant a to unity. Note that our conventions imply that $\sum_{\mathbf{q}} 1 = N/2$. As one expects, the classical ground state energy E_0 does not depend on θ and thus we have complete degeneracy with respect to θ. However, diagonalization of the Hamiltonian in Eq. (61) leads to the result

$$H_{\rm eff} = E_0' + \sum_{\mathbf{q}}\{\omega_+(\mathbf{q})a'^\dagger_{\mathbf{q}}a'_{\mathbf{q}} + \omega_-(\mathbf{q})b'^\dagger_{\mathbf{q}}b'_{\mathbf{q}}\} \ , \qquad (64)$$

where the new ground state energy E_0' is now

$$E_0' = -2\left(1+\frac{1}{S}\right)NJ_{\rm av}S^2 + \frac{1}{2}\sum_{\mathbf{q}}\{\omega_+(\mathbf{q}) + \omega_-(\mathbf{q})\} \ , \qquad (65)$$

and thus does depend on θ. This dependence on θ arises because the zero-point motion contribution [which is the sum of spin-wave energies $\omega_+(\mathbf{q}) + \omega_-(\mathbf{q})$ over the Brillouin zone] depends on θ. The spin-wave energies are

$$\omega_+(\mathbf{q}) = 4J_{\rm av}S\sqrt{(1-B_{\mathbf{q}})^2 - A_{\mathbf{q}}^2} \ ,$$
$$\omega_-(\mathbf{q}) = 4J_{\rm av}S\sqrt{(1+B_{\mathbf{q}})^2 - A_{\mathbf{q}}^2} \ . \qquad (66)$$

Note that when $J_\parallel = J_\perp = J_z$, $B_{\mathbf{q}}$ is zero and thus we have two degenerate spin modes as usual. When $J_\parallel, J_\perp$ and J_z are different, the two modes are no longer degenerate. This remains true when $\mathbf{q} \to 0$:

$$\omega_+(0) = 4S\sqrt{2J_{\rm av}(J_{\rm av} - J_z)} \ , \quad \omega_-(0) = 0 \ . \qquad (67)$$

This result shows that we have only one gap in the noninteracting spin-wave picture even though the ground state energy is anisotropic and therefore selects[34] a value of θ.

In Fig. 4 we plot the noninteracting spin-wave spectrum according to Eq. (66) along different directions in the Brillouin zone. For illustrative purposes we arbitrarily chose values of the J's which correspond to much larger anisotropy than we have for the cuprates. An indication in the spin-wave spectrum that $\delta J \equiv J_\parallel - J_\perp$ is nonzero is the removal of degeneracy[35] between $\omega_+(\mathbf{q})$ and $\omega_-(\mathbf{q})$ on the boundary of the Brillouin zone (where $q_x + q_y = \pi$). This effect is illustrated in Fig. 4. Even though noninteracting spin-wave theory does not lead to two gaps at zero wave vector when $J_\parallel \neq J_\perp$, one can obtain the second gap by calculating the spin-wave spectrum including higher orders in $1/S$. However, below we will estimate this in-plane gap without explicitly invoking spin-wave interactions.

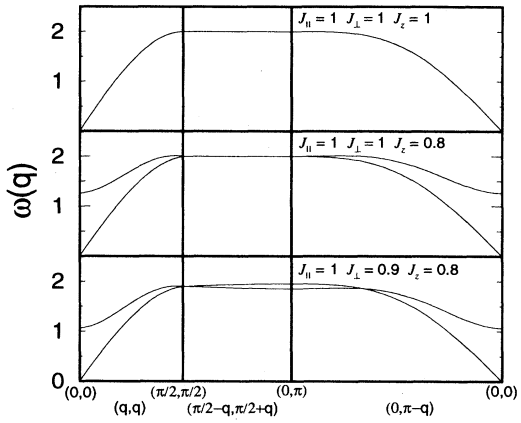

FIG. 4. Noninteracting spin-wave spectrum along different high symmetry lines in the Brillouin zone according to Eq. (66) for $\theta = 0$. For this plot the lattice constant a is set equal to unity. The J's and ω are all in the same arbitrary units.

For this purpose we study the quantum zero-point energy (per spin) in detail. It is given by

$$E_Z(\theta) = \frac{1}{2N}\sum_{\mathbf{q}}[\omega_+(\mathbf{q}) + \omega_-(\mathbf{q})] \ . \qquad (68)$$

From Eqs. (62), (63), and (66) one can write

$$\omega_\pm(\mathbf{q}) = 4J_{\rm av}S\left[(f\pm g) + (h\pm k)\cos(2\theta)\left(\frac{\delta J}{J_{\rm av}}\right)\right]^{1/2} , \qquad (69)$$

where

$$f = 1 - \frac{J_z C_+^2}{4J_{\rm av}}, \quad h = \frac{J_z C_+ C_-}{8J_{\rm av}},$$
$$g = \frac{(J_{\rm av} - J_z)C_+}{2J_{\rm av}}, \quad k = -\frac{C_-}{4} \ , \quad \delta J = J_\parallel - J_\perp \ , \qquad (70)$$

with C_+, C_- given by

$$C_\pm = \cos(q_x a) \pm \cos(q_y a) \ . \qquad (71)$$

To obtain the leading θ dependence of the mode energies, we expand $\omega_\pm(\mathbf{q})$ up to second order in powers of $(\delta J/J_{\rm av})$:

$$\omega_\pm(\mathbf{q}) = 4J_{\rm av}S(f\pm g)^{1/2}\left\{1 + \frac{1}{2}\left(\frac{h\pm k}{f\pm g}\right)\cos(2\theta)\left(\frac{\delta J}{J_{\rm av}}\right) - \frac{1}{8}\left(\frac{h\pm k}{f\pm g}\right)^2\cos^2(2\theta)\left(\frac{\delta J}{J_{\rm av}}\right)^2\right\}. \qquad (72)$$

By using this in Eq. (68) we can obtain the leading θ dependence of the quantum zero-point energy,

$$E_Z(\theta) = 2J_{\rm av}S\left[C_0 + C_1\cos(2\theta)\frac{\delta J}{J_{\rm av}} - C_2\cos^2(2\theta)\frac{(\delta J)^2}{J_{\rm av}^2}\right], \quad (73)$$

where the numerical constants are

$$C_0 = \tfrac{1}{N}\sum_{\mathbf{q}}\left\{\sqrt{f+g}+\sqrt{f-g}\right\}$$
$$C_1 = \frac{1}{2N}\sum_{\mathbf{q}}\left\{\frac{(h+k)}{\sqrt{f+g}}+\frac{(h-k)}{\sqrt{f-g}}\right\} = 0+0=0,$$
$$C_2 = \frac{1}{8N}\sum_{\mathbf{q}}\left\{\frac{(h+k)^2}{(f+g)^{3/2}}+\frac{(h-k)^2}{(f-g)^{3/2}}\right\}. \quad (74)$$

Note that coefficients of odd powers of δJ vanish due to the fact that these terms include odd power of C_- which changes sign under $q_x \leftrightarrow q_y$ while the other expressions are invariant under this operation.

In Fig. 5 we show $E_Z(\theta)$ from Eq. (73) and from the exact sum given in Eq. (68) for $J_{\|}=1, J_\perp = 0.9$, and $J_z = 0.8$ for which $C_0 = 0.44 + 0.39 = 0.83$ and $C_2 = 2.95\times10^{-3} + 0.7\times10^{-2} \approx 1\times10^{-2}$, where the first and second numbers are the contribution from out-of-plane and in-plane modes, respectively. Note that the in-plane mode contributes almost twice as much as the out-of-plane mode. The agreement between the exact and approximate results is excellent even though we have taken $\delta J/J_{\rm av} \approx 0.1$. Since in many real systems this ratio is extremely small, Eq. (73) should give nearly the exact value. For $J_{\|} = J_\perp = J_z = J$ we have

$$C_0 = 0.842\ , \quad C_2 = 1\times10^{-2}\ . \quad (75)$$

Note that the zero-point fluctuation energy favors the staggered magnetization to point along a $[1,0]$ direction within the easy[36,37] plane. Experiments[38] indicate that this may be the case for $YBa_2Cu_3O_6$, where the dipolar energy does not select a value of θ,[19] although it is not easy to distinguish the direction of the staggered magnetization in such systems.[39] For other tetragonal cuprates, the magnetic structure in the ground state is determined by the competition between $E_Z(\theta)$ and other anisotropies which result from interplane interactions.[19]

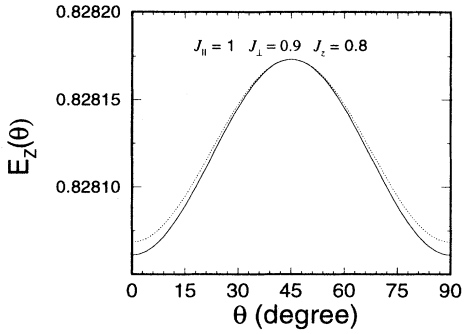

FIG. 5. Variation of the zero-point energy as a function of the angle θ between the staggered magnetization and a $[1,0]$ direction in the easy plane. Dotted and solid lines are from the approximate [Eq. (73)] and exact [Eq. (68)] expressions for $E_Z(\theta)$, respectively. The J's and $E_Z(\theta)$ are all in the same arbitrary units.

We are now ready to estimate the in-plane gap due to the anisotropy of quantum zero-point energy shown in Fig. 5. To do this we assume that the quantum zero-point energy is equivalent to an effective Hamiltonian for general S of the form

$$H_{\rm QZPE} = \sum_i \gamma_{\rm in} S(S_i^{x\,2}S_i^{y\,2}/S^4)\ . \quad (76)$$

Since we are interested in the region where $\mathbf{q}\approx 0$, this effective interaction is probably adequate to approximate the dependence of $E_Z(\theta)$ on θ_i even when θ has a slow nonzero spatial variation. Note that $H_{\rm QZPE}$ is of order S because $E_Z(\theta)$ in Eq. (68) is of that order. By comparing the angular dependence of Eq. (73) and Eq. (76), one obtains

$$\gamma_{\rm in} = 8C_2\frac{(\delta J)^2}{J_{\rm av}} \equiv 4J_{\rm av}\delta_{\rm in}\ . \quad (77)$$

Transforming S_x and S_y into the local quantization axis by using Eq. (58) (with $\theta = 0$) and Eqs. (59) and (60), and keeping only the terms at order of $1/S^0 = 1$, we find that

$$H_{\rm QZPE} = 4NJ_{\rm av}\delta_{\rm in} + 4J_{\rm av}\delta_{\rm in}\sum_i\{a_i^\dagger a_i + b_i^\dagger b_i - \tfrac{1}{2}[a_i^2 + b_i^2 + \text{H.c.}]\}\ . \quad (78)$$

In momentum space, $H_{\rm QZPE}$ is

$$H_{\rm QZPE} = 4NJ_{\rm av}\delta_{\rm in} + 4J_{\rm av}\delta_{\rm in}\sum_{\mathbf{q}}\{a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + b_{\mathbf{q}}^\dagger b_{\mathbf{q}} - \tfrac{1}{2}[a_{\mathbf{q}}a_{-\mathbf{q}} + b_{\mathbf{q}}b_{-\mathbf{q}} + \text{H.c.}]\}\ . \quad (79)$$

Hence the total Hamiltonian $H_{\rm tot} = H_{\rm QZPE} + H_{\rm eff}$, where $H_{\rm eff}$ is given in Eq. (61), is

$$H_{\rm tot} = E_{\rm tot} + 4J_{\rm av}S\Bigg[\sum_{\mathbf{q}}(1+S^{-1}\delta_{\rm in})(a^\dagger_{\mathbf{q}}a_{\mathbf{q}} + b^\dagger_{\mathbf{q}}b_{\mathbf{q}}) - \sum_{\mathbf{q}}\frac{1}{2}S^{-1}\delta_{\rm in}(a^\dagger_{\mathbf{q}}a^\dagger_{-\mathbf{q}} + a_{\mathbf{q}}a_{-\mathbf{q}} + b^\dagger_{\mathbf{q}}b^\dagger_{-\mathbf{q}} + b_{\mathbf{q}}b_{-\mathbf{q}}) + \sum_{\mathbf{q}}A_{\mathbf{q}}(a^\dagger_{\mathbf{q}}b^\dagger_{-\mathbf{q}} + a_{\mathbf{q}}b_{-\mathbf{q}}) + \sum_{\mathbf{q}}B_{\mathbf{q}}(a_{\mathbf{q}}b^\dagger_{\mathbf{q}} + a^\dagger_{\mathbf{q}}b_{\mathbf{q}})\Bigg] . \tag{80}$$

The spin-wave energies $\omega_\pm$ are given as

$$\omega^2_+(\mathbf{q}) = \left(4J_{\rm av}S\right)^2\left(1-A_{\mathbf{q}}-B_{\mathbf{q}}\right)\left(1+2S^{-1}\delta_{\rm in}+A_{\mathbf{q}}-B_{\mathbf{q}}\right),$$
$$\omega^2_-(\mathbf{q}) = \left(4J_{\rm av}S\right)^2\left(1+A_{\mathbf{q}}+B_{\mathbf{q}}\right)\left(1+2S^{-1}\delta_{\rm in}-A_{\mathbf{q}}+B_{\mathbf{q}}\right). \tag{81}$$

To get the in-plane and out-of-plane gaps, we set $\mathbf{q}=0$, in which case

$$\omega_+(q=0) = 4J_{\rm av}S\sqrt{2[1-(J_z/J_{\rm av})][1+S^{-1}\delta_{\rm in}]},$$
$$\omega_-(q=0) = 4J_{\rm av}S\sqrt{2[1+(J_z/J_{\rm av})]S^{-1}\delta_{\rm in}}, \tag{82}$$

where $\omega_-(q=0)$ is the in-plane-gap due to the quantum zero-point energy that we are looking for. Note that $\omega \propto \sqrt{S}$, as was originally found[40] in a similar situation where the gap is due to quantum zero-point effects. Thus we see that the noninteracting result, plotted in Fig. 4, which gives one gapless mode, needs to be modified as we have just done.

We now give a numerical evaluation of the gaps $\omega_\pm(q=0)$. For that purpose we approximate the result of Eq. (82) as

$$\omega_+(q=0) = 4S\sqrt{2J_{\rm av}(J_{\rm av}-J_z)},$$
$$\omega_-(q=0) = 8J_{\rm av}S\sqrt{\delta_{\rm in}/S} = 8\delta J\sqrt{(2S)C_2} \approx 0.8\mid J_{\parallel}-J_{\perp}\mid, \tag{83}$$

where we have used Eqs. (75) and (77) and set $S=\frac{1}{2}$. To evaluate $\omega_+(q=0)$ we use the experimental value[9,41] $J_{\rm av}=130$ meV and take $J_{\rm av}-J_z=(J_\perp-J_z)-(J_\perp-J_\parallel)2=30$ μeV from Fig. 3, in which case

$$\omega_+(q=0) = 2S(5.9\ \text{meV}) . \tag{84}$$

Measurements[11] show that zero-point fluctuations reduce $2S$ to about 0.8. Using this value, we get $\omega_+(q=0)=4.7$ meV, which compares favorably with the experimental[9] value of 5 meV. From the data shown in Fig. 3 we see that the in-plane gap, $\omega_-(q=0)$, should be about 25 μeV. It would be interesting to observe this via an infrared absorption experiment. Because the theoretical estimate of the frequency range is uncertain, it might be useful to locate the mode at high magnetic field and follow it back to zero applied field.

VI. THE LOWEST SYMMETRY MODEL

In Sec. IV we showed that for the model of Eq. (4) in tetragonal symmetry the anisotropy vanishes in the absence of Coulomb exchange. This was due to the fact that we have only hopping between orbitals of the same symmetry. However, the theorem breaks down when we have nonzero hopping between orbitals of different symmetry. Thus in this section we consider a system with lower symmetry to show that we can have anisotropy without Coulomb exchange.

Here we again consider the effective spin Hamiltonian for two copper ions, but now we do not assume any particular symmetry. Thus the orbitals localized on the two Cu ions which diagonalize $\mathcal{H}_x$ are no longer the same and will be some arbitrary linear combinations of $x^2-y^2, 3z^2-r^2, xy, yz, zx$, respectively. We write these orbitals as

$$\tilde{d}^\dagger_{i,\alpha} = \sum_\beta R^{-1}_{\alpha,\beta}(i)d^\dagger_\beta, \tag{85}$$

where $R_{\alpha,\beta}(i)$ is the matrix element of the orthogonal matrix which gives the new states in terms of the undistorted d orbitals for the ith copper ion.

Within these orbitals the hopping matrix elements now are

$$\tilde{t}_{i\alpha,j\beta} = \sum_{\gamma,\eta}R_{\alpha,\gamma}(i)R_{\beta,\eta}(j)t_{\gamma,\eta}, \tag{86}$$

where $t_{\gamma,\eta}$ is the usual overlap integral between the undistorted d orbitals, listed above. Similarly, the matrix elements of angular momentum in this new basis are

$$\tilde{\mathbf{L}}_{i\alpha,i\beta} = \sum_{\gamma,\eta}R_{\alpha,\gamma}(i)R_{\beta,\eta}(i)\mathbf{L}_{\gamma,\eta} . \tag{87}$$

A. Numerical study

We now present our numerical results for the effective spin Hamiltonian when we use these new hopping and angular momentum matrices for two arbitrarily chosen matrices $\mathbf{R}(i)$ and $\mathbf{R}(j)$. In Fig. 6 we show the anisotropy (energy differences between triplet states) for two different situations: (1) the on-site Coulomb repul-

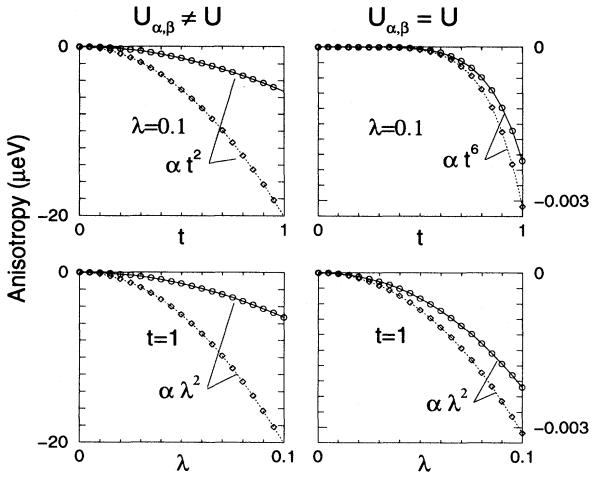

FIG. 6. Anisotropies (energy differences between triplet states) as a function of t and λ for nonconstant (left) and constant (right) on-site Coulomb interaction. If the energies of the triplet states are $\lambda_1 < \lambda_2 < \lambda_3$, the data points are λ_1-λ_2 (circles) and λ_1-λ_3 (diamonds). The solid and dashed lines are power-law fits, as indicated.

sive interactions depend on the orbitals ($U_{\alpha,\beta} \neq U$) and (2) orbital-independent (constant) Coulomb interactions, $U_{\alpha,\beta} = U$. As in the tetragonal case (but now with no Coulomb exchange interaction), we have full anisotropy for both cases. By fitting the numerical results shown in Fig. 6 we showed that anisotropy is proportional to $t^2\lambda^2$ for nonconstant **U** and to $t^6\lambda^2$ for constant **U**. In the next section we give an analytic proof that for constant **U** the anisotropy vanishes up to order $\mathbf{t}^4$.

B. Order $\tilde{t}^4$ results for constant U and $\mathbf{K} = \mathbf{0}$

With only (nearest-neighbor) (NN) hopping on the square lattice, there are no contributions at order $\tilde{t}^3$. The calculations at order $\tilde{t}^4$ will generate two types of contributions:[42] one a four-spin interaction, the other two-spin interactions either between nearest neighbors or between next-nearest neighbors. The first type is generated when a hole hops around a closed loop, i.e., from site 1 to 2, the hole which had been earlier on site 2 hops to 3 and so on, until the hole from site 4 hops to 1. The second type of interaction is generated both by closed loop processes and by various arrangements of four hops involving two or three sites.

In this paper we are mainly concerned with the evaluation of the NN pair exchange interactions. In particular we have concentrated mostly on the anisotropy of these interactions due to spin-orbit interactions. To study the contributions to this anisotropy from repeated hopping within a single bond to order $\tilde{t}^4$, we use Eq. (C2) of Appendix C and apply to it two more hopping terms, $T_{ij}\mathcal{H}_0^{-1}T_{ji}\mathcal{H}_0^{-1}$, ending at a ground state. After some algebra we obtain

$$\Delta\mathcal{H}(i,j) = \mathrm{Tr}\left\{(\vec{\sigma}\cdot\mathbf{S}_j)\left(\sum_{ab}\tilde{t}_{0b}^{ji}\tilde{t}_{ba}^{ij}\tilde{t}_{a0}^{ji}\right)(\vec{\sigma}\cdot\mathbf{S}_i)(\tilde{t}_{00}^{ij})\right\}$$
$$\times\frac{2}{U_0}\left[\frac{1}{E_{ja}(U_0+E_{ib})}+\frac{1}{E_{ib}(U_0+E_{ja})}-\frac{1}{(U_0+E_{ja})(U_0+E_{ib})}\right]+(i\leftrightarrow j)\ . \tag{88}$$

Combining Eqs. (24) and (88), we end up with Eq. (D4), in which both $(A_1 + i\mathbf{B}_1\cdot\vec{\sigma})$ and $(A_2 + i\mathbf{B}_2\cdot\vec{\sigma})$ are of the form $\tilde{t}_{00}^{ij} + O(\tilde{t}^3)$. Since at order $\tilde{t}^2$ we had $\mathbf{D}_2 = 0$, Eq. (D5) now yields $\mathbf{D}_2 = O(\tilde{t}^4)$, and thus the energy splitting of the triplet due to $\mathbf{D}_2$ is of order [cf. Eqs. (D6) and (D7)] $\mathbf{D}_2^2/A_1A_2 = O(\tilde{t}^6)$, irrespective of the details of Eq. (88). Thus, the NN magnetic exchange interaction becomes anisotropic only at order $\tilde{t}^6$, and this is correct to all orders in the spin-orbit coupling λ and for all lattice symmetries. This result is indeed confirmed by our single-bond numerical diagonalization, as we showed in Fig. 6.

We end this discussion with two comments. First, note that the separation of $(A_i + i\mathbf{B}_i\cdot\vec{\sigma})$ into a sum of terms of orders $\tilde{t}$ and $\tilde{t}^3$ was only possible because the sums over a and b in Eq. (88) all appeared within one matrix [which appears between $(\vec{\sigma}\cdot\mathbf{S}_j)$ and $(\vec{\sigma}\cdot\mathbf{S}_i)$]. This would not have been possible if we had contributions of the kind $\sum_a \mathrm{Tr}\{(\vec{\sigma}\cdot\mathbf{S}_j)T_1^a(\vec{\sigma}\cdot\mathbf{S}_i)T_2^a\}$, representing interference between different hopping paths. Such contributions arise at order $\tilde{t}^6$, and generate further anisotropy. (Without them, the analysis of Appendix D indicates that the energy levels would be two singlets and a doublet.) Second, note that the symmetry contained in Eq. (25) would persist to all orders, had we ignored excited states, allowing only $\tilde{t}_{00}^{ij}$.

VII. DISCUSSION AND CONCLUSIONS

A. Discussion

It is clear that the role of spin-orbit interactions in causing anisotropy in the exchange interaction is an interesting and subtle one. In particular, there has been much controversy concerning the way the exchange interaction $J_{\mu\nu}$ depends on the crystal symmetry and under what conditions one expects to find a gap in the spin-wave spectrum. For a long time after Moriya's seminal paper it was thought that one could neglect **M** in Eq. (2) and that the spin Hamiltonian for a single bond would be anisotropic if the Dzyaloshinskii vector **D** were nonva-

nishing. It was then observed by Kaplan[14] and by SEA[7] that although **M** is of order λ^2 and **D** is of order λ, one must nevertheless keep both terms when discussing the anisotropy or the gap in the spin-wave spectrum. Two other conclusions of these authors were (1) the single-bond spin Hamiltonian $\mathcal{H}_S(i,j)$ was rotationally invariant and (2) the overall anisotropy of the Cu-O plane resulted from a frustration between bonds with different values of **D**. In view of the results of the present paper we are in a position to state clearly the conditions under which the first conclusion is valid. In particular, our results show that rotational invariance (when the Coulomb exchange **K** is zero) of the single-bond spin Hamiltonian to all orders in t is only to be expected when hopping between excited states is ignored, as the SEA argument does implicitly. We remark that since the spin-orbit interaction involves coupling to excited orbital states, it is only non-negligible when the energies of the excited states involved are finite. This being the case, strictly speaking, it is not totally consistent to neglect hopping between such states, especially since the associated hopping matrix elements are comparable to those involving hopping to or from the orbital ground state. Nonetheless, as we have seen, the departures from the SEA rotational invariance theorem (due to hopping between excited states) are numerically quite small in most cases of physical interest. In fact, for the case of constant **U** considered by SEA, the deviations from rotational invariance only enter at order t^6.

In Moriya's original work to order t^2 it was correct to ignore hopping between excited states because he considered the case when **U** was a constant. In this case, as our results in Sec. VI show, rotational invariance only breaks down at order t^6, because one has to go to that high order for hopping between excited states to come into play. When **U** is nonconstant, hopping between excited states leads to anisotropy in $\mathcal{H}_S(i,j)$ at order t^2, as our results in Sec. VI demonstrate. These results thus represent a generalization of those by Moriya and by most of the literature which followed him and assumed constant **U**.

From this discussion one might now conclude that spin-orbit interactions would lead to anisotropy for the Cu-O planes as long as one includes hopping between excited Cu states. However, the theorem of Sec. IV shows that for the special case when the Cu sites have tetragonal symmetry, the generic model of Eq. (4) with only **t**, λ, and **U** nonzero does not yield nonzero anisotropy. The same result also applies to the "real" model including oxygen ions. This theorem explains why most previous calculations give no anisotropy for tetragonal site symmetry and it emphasizes the importance of including Coulomb exchange terms, **K**. It is then clear why the exchange anisotropy is so small, especially (as noted by BS) when compared to the anisotropy in the g tensor. We thus find that for each bond the exchange interaction has biaxial anisotropy ($J_\parallel$, $J_\perp$, and J_z are all different), where the anisotropy in J is of order $t^2_{\mathrm{Cu-Cu}}\lambda^2 K$, or more correctly, it is of order $t^2_{\mathrm{Cu-Cu}}\lambda^2 B$ or $t^2_{\mathrm{Cu-Cu}}\lambda^2 C$, where B and C are the Racah parameters which represent deviations from the simple constant **U** Hartree term.

Even though the single-bond exchange has biaxial anisotropy, the classical ground state energy, because it is averaged over bonds along [1,0,0] and [0,1,0], does not select an orientation of the staggered magnetization within the easy plane. As we have shown, the anisotropy within the easy plane results from quantum zero-point fluctuations. In summary, a complete discussion of the anisotropy of the Cu-O planes requires an interesting study of several novel symmetries and the way they are broken by fluctuations.

B. Conclusions

We may summarize our conclusions as follows.

(1) For tetragonal site symmetry, with only Hartree-like direct Coulomb terms, the effective spin Hamiltonian is isotropic at any order in the parameters t and λ. Inclusion of Coulomb exchange breaks this degeneracy at order $t^2_{\mathrm{Cu-Cu}}\lambda^2 K$ for our generic model and at order $t^4_{\mathrm{Cu-O}}\lambda^2 K$ for the cuprate system with an oxygen ion between the copper ions.

(2) Since the easy-plane anisotropy (observed via the "out-of-plane" spin-wave gap at zero wave vector) has comparable magnitudes for many orthorhombic and tetragonal cuprates, it cannot depend significantly on the orthorhombic distortion. Our result, Eq. (50), yields a biaxial anisotropy in the exchange interaction of order $t^4_{\mathrm{Cu-O}}\lambda^2 K$ which can explain the observed[9] out-of-plane spin-wave gap.

(3) In the tetragonal case, with the exchange interactions having biaxial anisotropy given by Eq. (50), the ground state does not depend on the orientation of the staggered magnetization within the easy plane. (As shown in Ref. 19, this remains true when dipolar interactions are included.) However, as we show,[19] quantum zero-point fluctuations cause an anisotropy within the easy plane which leads to ordering of the spins along the (1,0) axes, as indeed was claimed to be observed in $YBa_2Cu_3O_6$.[38] A rough estimate yields a resulting "in-plane" spin-wave gap of about 25 μeV. An experimental measurement of this gap would be very desirable.

(4) In real crystals, the three-dimensional ordering of the spins is determined by a competition between the anisotropies treated in the present paper and several other mechanisms, such as interplane hopping and interactions, as discussed recently in Refs. 19, 43, and 44. These may also affect the estimate of the in-plane gap given in conclusion (3).

(5) For sufficiently low symmetry and without exchange interactions, the rotational invariance of the single-bond Hamiltonian is broken at order $t^6\lambda^2$ for constant **U**. For arbitrary $U_{\alpha,\beta}$ and sufficiently low symmetry, the single-bond Hamiltonian is not rotationally invariant even at order $t^2\lambda^2$.

(6) We have given results for arbitrary symmetry for the effective spin Hamiltonian at order $\mathbf{t}^2$ including, for the first time, the effects of realistic Coulomb interactions. These expressions are valid for the orthorhombic phases of La_2CuO_4.

(7) In view of the controversies in the literature concerning the results which include spin-orbit interactions

we have implemented several checks of our perturbative results. First of all, we compared the results given in Eqs. (29) and (43) with expressions obtained by treating both the hopping *and* the spin-orbit interactions as perturbations. In addition, we subjected our analytic results for the tetragonal symmetry case to numerical verification as follows. We diagonalized exactly the Hamiltonian within the basis of two holes on either a Cu—Cu cluster or a Cu-O-Cu cluster. Then we compared the splittings of the ground state manifold (in this case, the lowest four states) with those predicted on the basis of our analytic evaluation of the perturbative contributions to the spin Hamiltonian. This comparison (see Fig. 3) was made with small enough values of the perturbative parameters that we can easily check how the results depend on the parameters.

ACKNOWLEDGMENTS

We acknowledge stimulating discussions with R. J. Birgeneau and M. A. Kastner. Work at the University of Pennsylvania was partly supported by the National Science Foundation MRL Program under Grant No. DMR-91-22784 and that at Tel Aviv by the U.S.-Israel Binational Science Foundation and the USIEF.

APPENDIX A: EXCHANGE AND FOUR LOWEST ENERGY LEVELS

In this appendix, we explain what we mean by anisotropy and how we study it by identifying the eigenvalues of $\mathcal{H}$ obtained numerically with those of a general spin Hamiltonian such as that given in Eq. (13).

The most general effective spin Hamiltonian for a single bond can be written as

$$H_{\rm eff} = E_0 + \mathbf{S}_1 \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{pmatrix} \mathbf{S}_2, \tag{A1}$$

where the matrix $\mathbf{J}$ is a 3×3 real matrix. There exist two transformations $\mathbf{R}_1$ and $\mathbf{R}_2$ which transform $\mathbf{J}$ into diagonal form if we rotate the spins: $\mathbf{S}_1 = \mathbf{R}_1\mathbf{S}_1'$ and $\mathbf{S}_2 = \mathbf{R}_2\mathbf{S}_2'$. To obtain $\mathbf{R}_1$ and $\mathbf{R}_2$ we first obtain the orthogonal matrix $\mathbf{O}$ which diagonalizes $\mathbf{J}\mathbf{J}^t$, where the subscript "t" indicates transpose:

$$\mathbf{O}^t(\mathbf{J}\mathbf{J}^t)\mathbf{O} = \tilde{\mathbf{J}}^2 , \tag{A2}$$

where $\tilde{\mathbf{J}}^2$ is a diagonal matrix with non-negative entries. For simplicity we assume that all its entries are actually positive. Then we define $\tilde{\mathbf{J}}$ and $\tilde{\mathbf{J}}^{-1}$ to be the corresponding diagonal matrices with positive entries.

Then we set

$$\mathbf{R}_1 = \mathbf{O} , \quad \mathbf{R}_2 = \mathbf{J}^t\mathbf{O}\tilde{\mathbf{J}}^{-1}\sigma , \tag{A3}$$

where $\sigma = {\rm Det}\mathbf{J}/|{\rm Det}\mathbf{J}|$. In terms of the transformed spins the Hamiltonian is

$$H'_{\rm eff} = E'_0 + \sigma\mathbf{S}'_1 \begin{pmatrix} \tilde{J}_x & 0 & 0 \\ 0 & \tilde{J}_y & 0 \\ 0 & 0 & \tilde{J}_z \end{pmatrix} \mathbf{S}'_2 , \tag{A4}$$

where all the $\tilde{J}_\alpha$'s are positive. Obviously, a further rotation could be made to change the sign of any two components of $\mathbf{S}'_1$ (or $\mathbf{S}'_2$). So the energy level scheme must be invariant under such a change of signs. It also has to be invariant under permutations of the $\tilde{J}_\alpha$. We find the four energy levels to be

$$\begin{aligned} \lambda_S &= E'_0 - \sigma(\tilde{J}_x + \tilde{J}_y + \tilde{J}_z)/4 , \\ \lambda_\alpha &= E'_0 + \sigma(\tilde{J}_x + \tilde{J}_y + \tilde{J}_z - 2\tilde{J}_\alpha)/4 , \quad \alpha = x,\ y,\ z. \end{aligned} \tag{A5}$$

This set of energies has the proper invariance under change of the signs of any two $\tilde{J}_\alpha$'s. For the case of arbitrary low symmetry, we did not try to identify the principal axes, but tabulated anisotropies, defined to be $\lambda_1 - \lambda_2$ and $\lambda_1 - \lambda_3$, where $\lambda_1 < \lambda_2 < \lambda_3$ are the three eigenvalues of the set $\{\lambda_\alpha\}$.

For D_{2h} bond symmetry, $\mathbf{J}$ must be diagonal, so that

$$\mathcal{H}_{\rm eff} = E_0 + \sum_\alpha J_\alpha S_{1,\alpha} S_{2,\alpha} . \tag{A6}$$

In this case the eigenvalues are given by Eq. (A5) with $\sigma = 1$ and the J_α have whatever signs they have in Eq. (A6). The identification of the J's from the set of eigenvalues of Eq. (A5) is not unique, because either permuting the J's or changing two of their signs leaves the set of eigenvalues invariant. Thus, identification of the J's with coordinate directions requires consideration of the eigenfunctions. For this purpose we write them explicitly:

$$\begin{aligned} \psi_S &= |\uparrow\downarrow\ - \downarrow\uparrow\rangle/\sqrt{2} , \quad \psi_z = |\uparrow\downarrow\ + \downarrow\uparrow\rangle/\sqrt{2} , \\ \psi_x &= |\uparrow\uparrow\ - \downarrow\downarrow\rangle/\sqrt{2} , \quad \psi_y = |\uparrow\uparrow\ + \downarrow\downarrow\rangle/\sqrt{2} . \end{aligned} \tag{A7}$$

Then the eigenfunctions are distinguished by their expectation values:

$$\begin{aligned} &\psi_S\text{: } \langle S_{1,z}S_{2,z}\rangle = \langle S_{1,x}S_{2,x}\rangle = \langle S_{1,y}S_{2,y}\rangle = -1/4 \\ &\psi_z\text{: } -\langle S_{1,z}S_{2,z}\rangle = \langle S_{1,x}S_{2,x}\rangle = \langle S_{1,y}S_{2,y}\rangle = 1/4 , \end{aligned} \tag{A8}$$

and so forth for the other ψ_α. Having identified which wave functions (coming out of the diagonalization of the 190×190 matrix) are which, one can easily deduce the values of the J_α. For instance

$$J_z = -\lambda_S - \lambda_z + \lambda_x + \lambda_y . \tag{A9}$$

APPENDIX B: COULOMB INTERACTION PARAMETERS IN TERMS OF THE RACAH COEFFICIENTS

Here we list the Coulomb interaction parameters for the tetragonal symmetry crystal field states for a d^8 configuration in terms of the Racah parameters.[29] Here $\Delta U_{\alpha\beta} = -2K_{\alpha\beta}$, where $U_{\alpha\alpha'} = U_0 + \Delta U_{\alpha\alpha'}$, with $U_0 = A + 4B + 3C$. In terms of the triplet and singlet energies given in Ref. 25 one has $\mathbf{U}^s = \mathbf{U} + \mathbf{K}$ and $\mathbf{U}^t = \mathbf{U} - \mathbf{K}$

$$\mathbf{K} = \begin{array}{c|ccccc} & d_{x^2-y^2} & d_{3z^2-r^2} & d_{xy} & d_{yz} & d_{zx} \\ \hline d_{x^2-y^2} & 0 & 4B+C & C & 3B+C & 3B+C \\ d_{3z^2-r^2} & 4B+C & 0 & 4B+C & B+C & B+C \\ d_{xy} & C & 4B+C & 0 & 3B+C & 3B+C \\ d_{yz} & 3B+C & B+C & 3B+C & 0 & 3B+C \\ d_{zx} & 3B+C & B+C & 3B+C & 3B+C & 0 \end{array} . \tag{B1}$$

In the numerical calculations we used (see Table I) $A = 7.00$, $B = 0.15$, and $C = 0.58$, so that $U_0 = 9.34$, all in eV.

APPENDIX C: MATRIX ELEMENTS NEEDED FOR PERTURBATION THEORY

Here we record some of the matrix elements needed to implement perturbation theory. If $|\psi_0\rangle$ represents any state in the ground state manifold (having one hole per site), then we may write

$$\begin{aligned} |\psi_1\rangle &\equiv \frac{1}{\mathcal{H}_0} T_{ji} |\psi_0\rangle \\ &\equiv \sum_{\sigma_1\sigma_2} \frac{1}{\mathcal{H}_0} T_{ji} c^\dagger_{i0\sigma_1} c^\dagger_{j0\sigma_2} c_{j0\sigma_2} c_{i0\sigma_1} |\psi_0\rangle \\ &= \sum_{\sigma_1\sigma_2\sigma_3 b} \frac{(\tilde{t}^{ji}_{b0})_{\sigma_3\sigma_1}}{E_{jb}+U_0} c^\dagger_{jb\sigma_3} c^\dagger_{j0\sigma_2} c_{j0\sigma_2} c_{i0\sigma_1} |\psi_0\rangle , \end{aligned} \tag{C1}$$

where $\mathcal{H}_0$ and T_{ij} are defined in Eqs. (19) and (14b), respectively, and where we set the ground state energy of $\mathcal{H}_0$ to zero. For results to second and fourth order in $\mathbf{t}$ with no Coulombic perturbations we need to generate the following matrix element:

$$\begin{aligned} |\psi_2\rangle &\equiv T_{ij} |\psi_1\rangle \\ &= \sum_{\substack{\sigma_1\sigma_2\sigma_3\sigma_4 \\ bc}} [(\tilde{t}^{ji}_{b0})_{\sigma_3\sigma_1} (\tilde{t}^{ij}_{cb})_{\sigma_4\sigma_3} c^\dagger_{ic\sigma_4} c^\dagger_{j0\sigma_2} - (\tilde{t}^{ji}_{b0})_{\sigma_3\sigma_1} (\tilde{t}^{ij}_{c0})_{\sigma_4\sigma_2} c^\dagger_{ic\sigma_4} c^\dagger_{jb\sigma_3}] (E_{jb}+U_0)^{-1} c_{j0\sigma_2} c_{i0\sigma_1} |\psi_0\rangle . \end{aligned} \tag{C2}$$

We also need

$$\begin{aligned} |\psi_3\rangle &\equiv \frac{1}{\mathcal{H}_0} \Delta\mathcal{H}_c \frac{1}{\mathcal{H}_0} T_{ji} |\psi_0\rangle \\ &= \sum_{\substack{\sigma_1\sigma_2\sigma_3 \\ ss' a_1 a_2 b}} \left\{ \frac{(\tilde{t}^{ji}_{b0})_{\sigma_3\sigma_1}}{(E_{jb}+U_0)(E_{ja_1}+E_{ja_2}+U_0)} [\Delta\tilde{U}_{ss'\sigma_2\sigma_3}(j;a_1a_20b) + \tilde{K}_{ss'\sigma_2\sigma_3}(j;a_1a_20b) \right. \\ &\quad \left. - \Delta\tilde{U}_{ss'\sigma_3\sigma_2}(j;a_1a_2b0) - \tilde{K}_{ss'\sigma_3\sigma_2}(j;a_1a_2b0)] c^\dagger_{ja_1s} c^\dagger_{ja_2s'} \right\} c_{j0\sigma_2} c_{i0\sigma_1} |\psi_0\rangle . \end{aligned} \tag{C3}$$

Finally, to get the energy at order $\mathbf{t}^2\Delta\mathcal{H}_c$ we need

$$\begin{aligned} |\psi_4\rangle &\equiv T_{ij} |\psi_3\rangle \\ &= \sum_{\substack{\sigma_1\sigma_2\sigma_3\sigma_4 \\ ss' a_1 a_2 b}} \left\{ \frac{(\tilde{t}^{ji}_{b0})_{\sigma_3\sigma_1}}{(E_{jb}+U_0)(E_{ja_1}+E_{ja_2}+U_0)} \right. \\ &\quad \times [\Delta\tilde{U}_{ss'\sigma_2\sigma_3}(j;a_1a_20b) + \tilde{K}_{ss'\sigma_2\sigma_3}(j;a_1a_20b) - \Delta\tilde{U}_{ss'\sigma_3\sigma_2}(j;a_1a_2b0) - \tilde{K}_{ss'\sigma_3\sigma_2}(j;a_1a_2b0)] \\ &\quad \left. \times [(\tilde{t}^{ij}_{0a_1})_{\sigma_4 s} c^\dagger_{i0\sigma_4} c^\dagger_{j0s'} \delta_{a_2,0} - (\tilde{t}^{ij}_{0a_2})_{\sigma_4 s'} c^\dagger_{i0\sigma_4} c^\dagger_{j0s} \delta_{a_1,0}] \right\} c_{j0\sigma_2} c_{i0\sigma_1} |\psi_0\rangle . \end{aligned} \tag{C4}$$

In order to make sure this matrix element connects to the ground state, we had to insert the factors $\delta_{a_2,0}$ and $\delta_{a_1,0}$.

APPENDIX D: SYMMETRY OF THE MAGNETIC HAMILTONIAN

In this appendix we analyze the eigenvalue spectrum of a system of two spins-$\frac{1}{2}$ with coupling which is arbitrary except that, for simplicity, we consider the isotropic interaction to be dominant. In the presence of antisymmetric exchange interactions one can always put the Hamiltonian into the following canonical form:

$$\mathcal{H}(i,j) = \alpha' \mathbf{S}(i)\cdot\mathbf{S}(j) + \beta\hat{n}\cdot(\mathbf{S}(i)\times\mathbf{S}(j)) + \gamma'\mathbf{S}(i)\cdot\mathbf{M}\cdot\mathbf{S}(j) , \tag{D1}$$

where $\hat{n}$ is a unit vector specifying the orientation of the Dzyaloshinskii vector and $\mathbf{M}$ is a symmetric matrix. Here

we will show that the eigenvalue spectrum of this Hamiltonian consists of a singlet and a triplet, if and only if the matrix $\mathbf{M}$ is such that $\mathcal{H}(i,j)$ can be written in the form

$$\mathcal{H}(i,j) = \alpha \mathbf{S}(i)\cdot\mathbf{S}(j) + \beta\hat{n}\cdot(\mathbf{S}(i)\times\mathbf{S}(j)) + \gamma(\mathbf{S}(i)\cdot\hat{n})(\hat{n}\cdot\mathbf{S}(j)) \ , \tag{D2}$$

where the coefficients obey the relation

$$\gamma = -\alpha + \alpha \mid 1 + (\beta/\alpha)^2 \mid^{1/2} \ , \tag{D3}$$

for finite α/β. [The reason for phrasing the condition in terms of Eq. (D2) rather than Eq. (D1) is that the former, unlike the latter, is a unique representation.] In this case, as we shall see, the spins can be rotated (about the same axis, but through opposite angles) so that in terms of the rotated spins the Hamiltonian looks isotropic. This result shows that for this relation between the parameters the Hamiltonian is rotationally invariant, even if it is not isotropic. (By isotropic, we mean $\beta = \gamma = 0$.)

As seen in the text, many of the perturbative results have the form

$$\mathcal{H}(i,j) = \frac{2}{U_0}\mathrm{Tr}\{(A_1 + i\mathbf{B}_1\cdot\vec{\sigma})(\mathbf{S}(j)\cdot\vec{\sigma})(A_2 + i\mathbf{B}_2\cdot\vec{\sigma})(\mathbf{S}(i)\cdot\vec{\sigma})\}. \tag{D4}$$

Defining the vectors $\mathbf{D}$ and $\mathbf{D}_2$,

$$\mathbf{D} = A_1\mathbf{B}_2 - A_2\mathbf{B}_1, \quad \mathbf{D}_2 = A_1\mathbf{B}_2 + A_2\mathbf{B}_1, \tag{D5}$$

this becomes

$$\begin{aligned}\mathcal{H}(i,j) &= \frac{2}{U_0}A_1A_2\mathrm{Tr}\left\{\left(1 - i\frac{\mathbf{D}-\mathbf{D}_2}{2A_1A_2}\cdot\vec{\sigma}\right)(\mathbf{S}(j)\cdot\vec{\sigma})\left(1 + i\frac{\mathbf{D}+\mathbf{D}_2}{2A_1A_2}\cdot\vec{\sigma}\right)(\mathbf{S}(i)\cdot\vec{\sigma})\right\} \\ &= \frac{4A_1A_2}{U_0}\left\{\left(1 + \frac{D_2^2 - D^2}{(2A_1A_2)^2}\right)\mathbf{S}(j)\cdot\mathbf{S}(i) + \frac{1}{A_1A_2}\mathbf{D}\cdot\mathbf{S}(j)\times\mathbf{S}(i)\right. \\ &\quad \left. + \frac{1}{2(A_1A_2)^2}[(\mathbf{S}(j)\cdot\mathbf{D})(\mathbf{S}(i)\cdot\mathbf{D}) - (\mathbf{S}(j)\cdot\mathbf{D}_2)(\mathbf{S}(i)\cdot\mathbf{D}_2)]\right\}.\end{aligned} \tag{D6}$$

This is clearly of the general form (2), with the Dzyaloshinskii vector $2\mathbf{D}/U_0$ and the symmetric anisotropy matrix $2(\mathbf{D}\otimes\mathbf{D} - \mathbf{D}_2\otimes\mathbf{D}_2)/(U_0A_1A_2)$. The most general form for $\mathbf{M}$ would involve introducing a third linearly independent vector $\mathbf{D}_3$.

We now show that the eigenvalues of the Hamiltonian of Eq. (D6) are a singlet and a triplet if and only if $\mathbf{D}_2$ vanishes. To see this we study its eigenvalue equation, which, after some algebra, can be cast into the form

$$(\lambda - x)^2[(\lambda - x)(\lambda + 3x) - 4 \mid \mathbf{V}_2 \mid^2] = 0 \ , \tag{D7}$$

where

$$\begin{aligned} x &= \frac{4}{U_0}A_1A_2\left(1 + \frac{D^2 - D_2^2}{(2A_1A_2)^2}\right), \\ \mathbf{V}_2 &= \frac{4}{U_0}\left(\mathbf{D}_2 - \frac{1}{2A_1A_2}\mathbf{D}\times\mathbf{D}_2\right) . \end{aligned} \tag{D8}$$

It is clear that a triplet occurs if and only if $\mathbf{V}_2 = 0$, which, in turn, happens if and only if $\mathbf{D}_2 = 0$. Q.E.D. Furthermore, we see that in the presence of nonzero $\mathbf{D}_2$, the triplet is split into a doublet and a singlet. To remove all degeneracy it is necessary to introduce a third vector $\mathbf{D}_3$.

We make some further remarks about the case when $\mathbf{D}_2 = 0$. One can easily verify that the conditions of Eq. (D3) and (D2) are equivalent to requiring that $\mathbf{D}_2$ in Eq. (D6) vanish. We further show now that when $\mathbf{D}_2 = 0$, the Hamiltonian is rotationally invariant. For this purpose note that

$$\begin{aligned} 1 + i\frac{\mathbf{D}}{2A_1A_2}\cdot\vec{\sigma} &= \frac{1}{\cos\theta}e^{i\theta\hat{\mathbf{d}}\cdot\vec{\sigma}}, \quad \hat{\mathbf{d}} = \frac{\mathbf{D}}{\mid D \mid}, \\ \tan\theta &= \frac{\mid D \mid}{2A_1A_2}. \end{aligned} \tag{D9}$$

The Hamiltonian (D6) with $\mathbf{D}_2 = 0$ then becomes

$$\begin{aligned}\mathcal{H}(i,j) &= \frac{2}{U_0}A_1A_2\left(1 + \frac{D^2}{(2A_1A_2)^2}\right) \\ &\quad \times\mathrm{Tr}\{(\vec{\sigma}\cdot\mathbf{S}'(j))(\vec{\sigma}\cdot\mathbf{S}'(i))\} \\ &= \frac{4}{U_0}A_1A_2\left(1 + \frac{D^2}{(2A_1A_2)^2}\right)\mathbf{S}'(i)\cdot\mathbf{S}'(j) \ ,\end{aligned} \tag{D10}$$

in terms of rotated variables [equivalent to those of SEA (Ref. 7)]:

$$\begin{aligned} \vec{\sigma}\cdot\mathbf{S}(j) &= e^{i\frac{\theta}{2}\hat{\mathbf{d}}\cdot\vec{\sigma}}\vec{\sigma}\cdot\mathbf{S}'(j)e^{-i\frac{\theta}{2}\hat{\mathbf{d}}\cdot\vec{\sigma}}, \\ \vec{\sigma}\cdot\mathbf{S}(i) &= e^{-i\frac{\theta}{2}\hat{\mathbf{d}}\cdot\vec{\sigma}}\vec{\sigma}\cdot\mathbf{S}'(i)e^{i\frac{\theta}{2}\hat{\mathbf{d}}\cdot\vec{\sigma}}. \end{aligned} \tag{D11}$$

When $\mathbf{D}_2$ is finite, the triplet splits into a singlet and a doublet. One may ask whether it is possible to perform rotations of the spins such that the antisymmetric Dzyaloshinskii term will be eliminated and the Hamiltonian will contain only the symmetric anisotropy. This is in general not the case. Returning to the Hamiltonian (D4), we put

$$\vec{\sigma}\cdot\mathbf{S}(j) = e^{i\frac{\alpha}{2}\hat{\mathbf{a}}\cdot\vec{\sigma}}\vec{\sigma}\cdot\mathbf{S}'(j)e^{-i\frac{\alpha}{2}\hat{\mathbf{a}}\cdot\vec{\sigma}},$$
$$\vec{\sigma}\cdot\mathbf{S}(i) = e^{-i\frac{\alpha}{2}\hat{\mathbf{a}}\cdot\vec{\sigma}}\vec{\sigma}\cdot\mathbf{S}'(i)e^{i\frac{\alpha}{2}\hat{\mathbf{a}}\cdot\vec{\sigma}}, \tag{D12}$$

where the unit vector $\hat{\mathbf{a}}$ and the angle α are yet to be determined. The Hamiltonian then takes the form (D4), with the replacements $\mathbf{S}(i)\to\mathbf{S}'(i)$, $\mathbf{S}(j)\to\mathbf{S}'(j)$, and $A_i\to A_i'$, $\mathbf{B}_i\to\mathbf{B}_i'$, $(i=1,2)$ with

$$A_{\frac{1}{2}}' = A_{\frac{1}{2}}\cos\alpha \mp \sin\alpha(\mathbf{B}_{\frac{2}{1}}\cdot\hat{\mathbf{a}}),$$
$$\mathbf{B}_{\frac{1}{2}}' = \mathbf{B}_{\frac{1}{2}} \pm A_{\frac{1}{2}}\hat{\mathbf{a}}\sin\alpha - 2\sin^2\frac{\alpha}{2}(\hat{\mathbf{a}}\cdot\mathbf{B}_{\frac{1}{2}})\hat{\mathbf{a}}. \tag{D13}$$

The condition that the Dzyaloshinskii term vanish is therefore [cf. Eqs. (D5) and (D6)]

$$[A_1\cos\alpha - \sin\alpha(\mathbf{B}_1\cdot\hat{\mathbf{a}})][\mathbf{B}_2 - \hat{\mathbf{a}}(\hat{\mathbf{a}}\cdot\mathbf{B}_2)] - [A_2\cos\alpha + \sin\alpha(\mathbf{B}_2\cdot\hat{\mathbf{a}})][\mathbf{B}_1 - \hat{\mathbf{a}}(\hat{\mathbf{a}}\cdot\mathbf{B}_1)]$$
$$= \hat{\mathbf{a}}[\sin 2\alpha(A_1A_2 + (\hat{\mathbf{a}}\cdot\mathbf{B}_1)(\hat{\mathbf{a}}\cdot\mathbf{B}_2)) + \cos 2\alpha((A_2\mathbf{B}_1 - A_1\mathbf{B}_2)\cdot\hat{\mathbf{a}})]. \tag{D14}$$

Since the vector on the left-hand side is orthogonal to $\hat{\mathbf{a}}$, Eq. (D14) yields

$$\tan 2\alpha = \frac{(A_1\mathbf{B}_2 - A_2\mathbf{B}_1)\cdot\hat{\mathbf{a}}}{A_1A_2 + (\hat{\mathbf{a}}\cdot\mathbf{B}_1)(\hat{\mathbf{a}}\cdot\mathbf{B}_2)},$$
$$[A_1\cos\alpha - (\hat{\mathbf{a}}\cdot\mathbf{B}_1)\sin\alpha](\mathbf{B}_2 - \hat{\mathbf{a}}(\hat{\mathbf{a}}\cdot\mathbf{B}_2)) = [A_2\cos\alpha + (\hat{\mathbf{a}}\cdot\mathbf{B}_2)\sin\alpha](\mathbf{B}_1 - \hat{\mathbf{a}}(\hat{\mathbf{a}}\cdot\mathbf{B}_1)) \ . \tag{D15}$$

It can be shown that these two equations can be satisfied only when the vectors $\mathbf{B}_1$ and $\mathbf{B}_2$ are parallel. Hence the antisymmetric anisotropy can be eliminated from the Hamiltonian only for specific configurations. Moreover, the criterion for complete rotational invariance of $\mathcal{H}(i,j)$ is that $\mathbf{D}_2 = A_1\mathbf{B}_2 + A_2\mathbf{B}_1 = 0$, or equivalently, that $\mathbf{B}_2/A_2 = -\mathbf{B}_1/A_1$, which is equivalent to the condition that

$$(A_1 + i\mathbf{B}_1\cdot\vec{\sigma})^\dagger = \frac{A_1}{A_2}(A_2 + i\mathbf{B}_2\cdot\vec{\sigma}). \tag{D16}$$

Returning to Eq. (24) we note that $(\tilde{\mathbf{t}}_{00}^{ij})^\dagger = \tilde{\mathbf{t}}_{00}^{ji}$, and therefore that $\mathcal{H}^{(2)}(i,j)$ is indeed rotationally invariant. This represents an alternative proof for the SEA result,[7] which holds to order t^2, to all orders in the spin-orbit coupling λ and for all site symmetries providing $U_{\alpha\beta} = U$ and $\mathbf{K} = 0$.

APPENDIX E: PERTURBATION THEORY INCLUDING OXYGEN ORBITALS

Here we show that the perturbation theory results for the Cu—O—Cu bond, through the intermediate state in which the two holes are on the copper ion (channel a), are obtained from those of the Cu—Cu bond, with the replacement (36).

We start from the state $|\psi_1\rangle$ in Eq. (37). As explained in the text, the index q that labels oxygen ions on the bonds along x and along y may be omitted for simplicity. Applying again the hopping Hamiltonian yields

$$|\psi_{2a}\rangle = \sum_{\sigma\sigma_1}\frac{1}{\mathcal{H}_0}\left(T_j^\dagger\frac{1}{\mathcal{H}_0}T_i + T_i^\dagger\frac{1}{\mathcal{H}_0}T_j\right)c_{i0\sigma}^\dagger c_{j0\sigma_1}^\dagger c_{j0\sigma_1}c_{i0\sigma}|\psi_0\rangle$$
$$= \sum_n\sum_{\substack{\sigma\sigma_1\\ \sigma_2\sigma_3}}\frac{1}{\epsilon_n}\left[\frac{(\vec{t}_{n0}^{\,i})_{\sigma_2\sigma}(\vec{t}_{an}^{\,j})_{\sigma_3\sigma_2}}{U_0 + E_{ja}}c_{ja\sigma_3}^\dagger c_{j0\sigma_1}^\dagger\right.$$
$$\left. + \frac{(\vec{t}_{n0}^{\,j})_{\sigma_2\sigma_1}(\vec{t}_{an}^{\,i})_{\sigma_3\sigma_2}}{U_0 + E_{ia}}c_{i\sigma}^\dagger c_{ia\sigma_3}^\dagger\right]c_{j0\sigma_1}c_{i0\sigma} \mid \psi_0\rangle \ , \tag{E1}$$

where we have written the energy denominators explicitly. We now concentrate on the terms that will eventually contribute to the spin Hamiltonian. To order t^4, these are obtained by applying two more factors of the hopping that bring the holes back to the ground state. The result is

$$\left(T_i^\dagger\frac{1}{\mathcal{H}_0}T_j + T_j^\dagger\frac{1}{\mathcal{H}_0}T_i\right)|\psi_{2a}\rangle = -\sum_{nn'}\sum_{\sigma\sigma_1\sigma_2}\sum_{\sigma_3\sigma_4\sigma_5}\frac{1}{\epsilon_n\epsilon_{n'}U_0}[(\vec{t}_{n0}^{\,i})_{\sigma_2\sigma}(\vec{t}_{0n}^{\,j})_{\sigma_3\sigma_2}(\vec{t}_{n'0}^{\,j})_{\sigma_4\sigma_1}(\vec{t}_{0n'}^{\,i})_{\sigma_5\sigma_4}c_{j0\sigma_3}^\dagger c_{i0\sigma_5}^\dagger$$
$$+(\vec{t}_{n0}^{\,j})_{\sigma_2\sigma_1}(\vec{t}_{0n}^{\,i})_{\sigma_3\sigma_2}(\vec{t}_{n'0}^{\,i})_{\sigma_4\sigma}(\vec{t}_{0n'}^{\,j})_{\sigma_5\sigma_4}c_{j0\sigma_5}^\dagger c_{i0\sigma_3}^\dagger]c_{j0\sigma_1}c_{i0\sigma} \mid \psi_0\rangle \ , \tag{E2}$$

from which it is clear that using Eq. (36) one arrives at Eq. (24).

APPENDIX F: DIAGONALIZATION OF THE SPIN-ORBIT INTERACTION

In this appendix we show that the transformation of Eq. (44) does indeed make the spin-orbit interaction diagonal in pseudospin space. We wish to show that

$$\langle \beta, \nu | \mathbf{L} \cdot \vec{\sigma} | \alpha, \mu \rangle \tag{F1}$$

vanishes unless $\mu = \nu$, in which case it is independent of μ. First of all, note that $\mathbf{L}$ has zero diagonal matrix elements, i.e., the above matrix element vanishes when $\alpha = \beta$. There are now three cases to consider: (i) $\alpha = 0$ and $\beta = 1$ or $\alpha = 1$ and $\beta = 0$; (ii) $\alpha = 0, 1$ and $\beta = x, y, z$; and (iii) $\alpha \neq \beta$ but both are x, y, or z. In case (i) the matrix element of $\mathbf{L}$ is again zero, so this case is as desired. In case (ii) with $\alpha = 0$ ($\alpha = 1$ is similar) we express the above matrix element as

$$\sum_{\gamma,\rho} \langle \psi_\beta(\mathbf{r}) | L_\gamma | \psi_0(\mathbf{r}) \rangle (\sigma_\beta)_{\nu,\rho} \langle \phi_\rho(\sigma) | \sigma_\gamma | \phi_\mu(\sigma) \rangle = \langle \psi_\beta(\mathbf{r}) | L_\beta | \psi_0(\mathbf{r}) \rangle \delta_{\mu,\nu} \ . \tag{F2}$$

In the last step we used the fact that the orbital matrix element is only nonzero when $\gamma = \beta$. So case (ii) is as desired. In case (iii) we write the matrix element as

$$\begin{aligned} \sum_{\gamma,\rho,\tau} \langle \psi_\beta(\mathbf{r}) | L_\gamma | \psi_\alpha(\mathbf{r}) \rangle (\sigma_\alpha)^*_{\mu,\tau} (\sigma_\beta)_{\nu,\rho} \langle \phi_\rho(\sigma) | \sigma_\gamma | \phi_\tau(\sigma) \rangle &= \sum_{\gamma,\rho,\tau} \langle \psi_\beta(\mathbf{r}) | L_\gamma | \psi_\alpha(\mathbf{r}) \rangle (\sigma_\alpha)_{\tau,\mu} (\sigma_\beta)_{\nu,\rho} (\sigma_\gamma)_{\rho,\tau} \\ &= \langle \psi_\beta(\mathbf{r}) | L_\gamma | \psi_\alpha(\mathbf{r}) \rangle i \epsilon_{\beta\gamma\alpha} \delta_{\mu,\nu} \ , \end{aligned} \tag{F3}$$

where $\epsilon_{\alpha\beta\gamma}$ is the totally antisymmetric tensor. In the last equality we used the fact that α, β, and γ are Cartesian indices which are all different. Thus all the types of matrix elements are diagonal and independent of pseudospin, as asserted.

APPENDIX G: RESULTS FOR TETRAGONAL SYMMETRY

The only nonzero matrix elements of the angular momentum within the manifold of normalized tetragonal d states, $|\ 0\rangle = d_{x^2-y^2}$, $|\ 1\rangle = d_{3z^2-r^2}$, $|\ z\rangle = d_{xy}$, $|\ x\rangle = d_{yz}$, and $|\ y\rangle = d_{xz}$ are

$$\begin{aligned} \langle x^2 - y^2 | L_x | yz \rangle &= \langle zx | L_x | xy \rangle = -\langle xy | L_x | zx \rangle = -\langle yz | L_x | x^2 - y^2 \rangle = \langle xy | L_y | yz \rangle \\ &= \langle x^2 - y^2 | L_y | zx \rangle = -\langle yz | L_y | xy \rangle = -\langle zx | L_y | x^2 - y^2 \rangle = \langle yz | L_z | zx \rangle = -\langle zx | L_z | xy \rangle = i \ , \end{aligned} \tag{G1}$$

$$\langle 3z^2 - r^2 | L_x | yz \rangle = -\langle yz | L_x | 3z^2 - r^2 \rangle = -\langle 3z^2 - r^2 | L_y | zx \rangle = \langle zx | L_y | 3z^2 - r^2 \rangle = i\sqrt{3} \ , \tag{G2}$$

$$\langle xy | L_z | x^2 - y^2 \rangle = -\langle x^2 - y^2 | L_z | xy \rangle = 2i \ . \tag{G3}$$

From Appendix F and the pseudospin transformation (44) it follows that

$$[\omega(\alpha, \beta)]_{\sigma\sigma'} = \bar{\omega}(\alpha, \beta)(\sigma_\alpha \sigma_\beta)_{\sigma,\sigma'} \ , \tag{G4}$$

where ω was defined in Eq. (6), and $\bar{\omega}(\alpha, \beta)$ is a scalar. Turning now to the diagonalization of the single-particle, single-site Hamiltonian $\mathcal{H}_x + \mathcal{H}_{so}$ of the Cu—Cu bond, one finds that this can be accomplished by putting

$$\mathbf{m}_{\alpha a} = \sigma_\alpha \bar{m}_{\alpha a} \ , \tag{G5}$$

where $\bar{m}_{\alpha a}$ are scalars which are determined by Eq. (11), which now becomes

$$\sum_\alpha \epsilon_\alpha \bar{m}^*_{\alpha a} \bar{m}_{\alpha b} + \lambda \sum_{\alpha\beta} \bar{\omega}(\alpha, \beta) \bar{m}^*_{\alpha a} \bar{m}_{\beta b} = \delta_{ab} E_a, \quad \sum_a \bar{m}_{\alpha a} \bar{m}^*_{\beta a} = \delta_{\alpha\beta}. \tag{G6}$$

The solution of this equation, to second order in λ, reads

$$\bar{m}_{\alpha a} = \delta_{\alpha a} \left(1 - \frac{\lambda^2}{2} \sum_\gamma \frac{\bar{\omega}(\alpha, \gamma) \bar{\omega}(\gamma, \alpha)}{(\epsilon_\alpha - \epsilon_\gamma)^2} \right) + (1 - \delta_{\alpha a}) \left(\lambda \frac{\bar{\omega}(\alpha, a)}{\epsilon_a - \epsilon_\alpha} + \lambda^2 \sum_\gamma \frac{\bar{\omega}(\alpha, \gamma) \bar{\omega}(\gamma, a)}{(\epsilon_\alpha - \epsilon_a)(\epsilon_\gamma - \epsilon_a)} \right). \tag{G7}$$

The hopping matrix elements are diagonal in the tetragonal symmetry, except for the states $|0\rangle$ and $|1\rangle$, and are independent of the site indices i and j. It follows that $\tilde{t}^{ij}_{ab}$ are scalars independent of the site indices as well, and that the hopping is not accompanied by a pseudospin flip. This implies that $\mathbf{x}_{\alpha\alpha}$ [Eq. (28a)] is diagonal in that space, and that $\mathbf{x}_{\alpha\alpha'}$ is independent of i and j. Thus the vector $\mathbf{D}_{ij}$, Eq. (29b), vanishes and the matrix $\mathbf{M}(i,j)$, Eq. (29c), is given solely by the $K_{\alpha\alpha'}$ terms, and is independent of i and j.

From these arguments and using Eqs. (28), one obtains that the symmetric matrix $\mathbf{M}(i,j)$ (which is diagonal in tetragonal symmetry) is given by

$$\begin{aligned}\mathbf{M}(i,j) &= 2\sum_{\alpha\alpha'} K_{\alpha\alpha'}(\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha'0}\mathbf{m}_{\alpha0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{w}^{ji}_{\alpha\alpha'}\vec{\sigma}\}\\ &\quad -\tfrac{1}{2}\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha'}\vec{\sigma}\}\otimes\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}\vec{\sigma}\}-\tfrac{1}{2}\mathrm{Tr}\{(\mathbf{x}^{ji}_{\alpha\alpha'})^{\dagger}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{x}^{ji}_{\alpha\alpha'}\vec{\sigma}\})\\ &= 2\sum_{ab}\frac{1}{(U_0+E_a)}\frac{1}{(U_0+E_b)}\tilde{t}^{ji}_{0a}\tilde{t}^{ij}_{b0}\sum_{\alpha\alpha'}K_{\alpha\alpha'}(\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha'0}\mathbf{m}_{\alpha0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha a}\mathbf{m}_{\alpha' b}\vec{\sigma}\}\\ &\quad -\tfrac{1}{2}\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha a}\mathbf{m}_{\alpha'0}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha'0}\mathbf{m}_{\alpha b}\vec{\sigma}\}-\tfrac{1}{2}\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha'0}\mathbf{m}_{\alpha b}\vec{\sigma}\}\otimes\mathrm{Tr}\{\mathbf{m}^{\dagger}_{\alpha a}\mathbf{m}_{\alpha'0}\vec{\sigma}\}). \end{aligned} \tag{G8}$$

Performing the sums over α and α', this yields

$$\begin{aligned}M_{xx}(i,j) &= 8\sum_{ab}\frac{1}{(U_0+E_a)}\frac{1}{(U_0+E_b)}\tilde{t}^{ji}_{0a}\tilde{t}^{ij}_{b0}[K_{0x}(\bar{m}^{*}_{x0}\bar{m}^{*}_{0a}-\bar{m}^{*}_{00}\bar{m}^{*}_{xa})(\bar{m}_{00}\bar{m}_{xb}-\bar{m}_{x0}\bar{m}_{0b})\\ &\quad +K_{1x}(\bar{m}^{*}_{x0}\bar{m}^{*}_{1a}-\bar{m}^{*}_{10}\bar{m}^{*}_{xa})(\bar{m}_{10}\bar{m}_{xb}-\bar{m}_{x0}\bar{m}_{1b})\\ &\quad +K_{yz}(\bar{m}^{*}_{z0}\bar{m}^{*}_{ya}-\bar{m}^{*}_{y0}\bar{m}^{*}_{za})(\bar{m}_{y0}\bar{m}_{zb}-\bar{m}_{z0}\bar{m}_{yb})]\ , \end{aligned} \tag{G9}$$

with analogous expressions for the yy and zz entries of $\mathbf{M}$. The next step is to write $\bar{m}_{\alpha a}$ in terms of the spin-orbit matrix elements, Eq. (G7). In doing this we keep in mind that both states a and b cannot be the ground state 0, as they refer to intermediate states of the perturbation theory. Therefore, it is sufficient to retain for the coefficient of K_{0x} the terms

$$-\delta_{xa}\delta_{xb}-\lambda\delta_{xa}\frac{\bar{\omega}(x,b)}{\epsilon_b-\epsilon_x}+\lambda\delta_{xb}\frac{\bar{\omega}(a,x)}{\epsilon_x-\epsilon_a}-\lambda^2\frac{\bar{\omega}(a,x)\bar{\omega}(x,b)}{(\epsilon_x-\epsilon_b)(\epsilon_x-\epsilon_a)}\ . \tag{G10}$$

Similarly, the coefficient of K_{1x} is

$$-\lambda^2\frac{\bar{\omega}(x,0)\bar{\omega}(0,x)}{\epsilon_x^2}\delta_{1a}\delta_{1b}\ . \tag{G11}$$

The leading order of the coefficient of K_{yz} is of order $\lambda^2\delta_{yb}\delta_{za}$, etc. But then $\tilde{t}_{0z}\tilde{t}_{y0}$ will be proportional to λ^2 too. Therefore, to order λ^2, the terms arising from K_{yz} do not contribute. Collecting terms we find the contribution to the anisotropic exchange as

$$\begin{aligned}J^{\mathrm{anis}}_{xx}(i,j) &= -8K_{0x}\left[\frac{1}{(U_0+\epsilon_x)^2}\tilde{t}_{0x}\tilde{t}_{x0}+\lambda\frac{1}{U_0+\epsilon_x}\sum_a\left(\frac{\bar{\omega}(x,a)}{\epsilon_a-\epsilon_x}\tilde{t}_{0x}\tilde{t}_{a0}-\frac{\bar{\omega}(a,x)}{\epsilon_x-\epsilon_a}\tilde{t}_{0a}\tilde{t}_{x0}\right)\right.\\ &\quad \left.+\lambda^2\sum_{ab}\frac{1}{U_0+\epsilon_a}\frac{1}{U_0+\epsilon_b}\frac{\bar{\omega}(a,x)\bar{\omega}(x,b)}{(\epsilon_x-\epsilon_b)(\epsilon_x-\epsilon_a)}\tilde{t}_{0a}\tilde{t}_{b0}\right]\\ &\quad -8K_{1x}\lambda^2\frac{1}{(U_0+\epsilon_1)^2}\tilde{t}_{01}\tilde{t}_{01}\frac{\bar{\omega}(x,0)\bar{\omega}(0,x)}{\epsilon_x^2}, \end{aligned} \tag{G12}$$

where we have retained terms up to order λ^2. Finally we write, using Eqs. (14c) and (G8)

$$\tilde{t}_{0x}=\lambda\frac{\bar{\omega}(0,x)}{\epsilon_x}(t_{00}-t_{xx})-\lambda t_{01}\frac{\bar{\omega}(1,x)}{\epsilon_1-\epsilon_x}\ , \tag{G13}$$

and put a and b in the sums of (G12) equal to 1, with $\tilde{t}_{01}=t_{01}$. (These are the only possible contributions up to order λ^2). It then follows that

$$J^{\mathrm{anis}}_{xx}(i,j)=-8\lambda^2\left[K_{0x}\frac{1}{(U_0+\epsilon_x)^2}\left|\frac{\bar{\omega}(0,x)}{\epsilon_x}(t_{00}-t_{xx})-\frac{\bar{\omega}(1,x)}{U_0+\epsilon_1}t_{01}\right|^2+K_{1x}\left|\frac{1}{U_0+\epsilon_1}t_{10}\frac{\bar{\omega}(x,0)}{\epsilon_x}\right|^2\right]\ . \tag{G14}$$

This result reproduces that of Ref. 18. This result differs slightly from that given in Ref. 18 and Eq. (50) in the

text, in that the denominators include just the constant part U_0 of the Coulomb interactions. The corresponding expression in Ref. 18 [and Eq. (50)] includes instead $U_{0,\mu}$, $U_{0.1}$, and $U_{1,\mu}$ and thus represents an expansion, described in Appendix H, in $\mathbf{K}$ but not in $\Delta\mathbf{U}$.

APPENDIX H: PERTURBATION THEORY FOR THE TETRAGONAL CASE

In this Appendix we give an alternative derivation of Eq. (50) based on conventional perturbation theory in which we treat hopping, $\mathcal{H}_{\rm hop}$, spin-orbit, $\mathcal{H}_{\rm so}$, and the Coulomb exchange interactions, $\mathcal{H}_{\rm ex}$, as perturbations. For the "generic" model, anisotropic exchange appears at order $\mathcal{H}_{\rm hop}^2\mathcal{H}_{\rm so}^2\mathcal{H}_{\rm ex}$. To perform this calculation we therefore need to work to fifth-order perturbation theory and will arbitrarily omit contributions to the isotropic exchange. Of course, $\mathcal{H}_{\rm ex}$ can only exist when there are two holes on the same ion, so the five perturbations must be arranged so that $\mathcal{H}_{\rm hop}$ and $\mathcal{H}_{\rm ex}$ occur in the order $\mathcal{H}_{\rm hop}\mathcal{H}_{\rm ex}\mathcal{H}_{\rm hop}$. In principle, there are ten ways to insert the two factors of $\mathcal{H}_{\rm so}$. But some study shows that only if the two powers of $\mathcal{H}_{\rm so}$ are separated by $\mathcal{H}_{\rm ex}$ does the result lead to anisotropy. So the relevant fifth-order terms in the effective Hamiltonian are

$$\mathcal{H}(i,j) = \left[\mathcal{H}_{\rm so}\frac{1}{\mathcal{E}}\mathcal{H}_{\rm hop} + \mathcal{H}_{\rm hop}\frac{1}{\mathcal{E}}\mathcal{H}_{\rm so}\right]\frac{1}{\mathcal{E}}\mathcal{H}_{\rm ex}\frac{1}{\mathcal{E}}\left[\mathcal{H}_{\rm so}\frac{1}{\mathcal{E}}\mathcal{H}_{\rm hop} + \mathcal{H}_{\rm hop}\frac{1}{\mathcal{E}}\mathcal{H}_{\rm so}\right] , \tag{H1}$$

where $\mathcal{E}$ is the appropriate energy denominator. If we write the spin-orbit perturbation as

$$\mathcal{H}_{\rm so} = \lambda\sum_\alpha\left(\sum_{\rm holes,h} L_\alpha(h)s_\alpha(h)\right) \equiv \sum_\alpha V_\alpha , \tag{H2}$$

then it is easy to see that there are no cross terms, i.e., terms involving $V_\alpha V_\beta$ with $\alpha \neq \beta$. In addition, hopping from site i to site j and back will give the same result as the reverse process. So if hopping from site j to site i is denoted T_{ij}, then we may write

$$\begin{aligned}\mathcal{H}(i,j) &= 2\sum_\alpha\left[V_\alpha\frac{1}{\mathcal{E}}T_{ji} + T_{ji}\frac{1}{\mathcal{E}}V_\alpha\right]\frac{1}{\mathcal{E}}\mathcal{H}_{\rm ex}\frac{1}{\mathcal{E}}\left[V_\alpha\frac{1}{\mathcal{E}}T_{ij} + T_{ij}\frac{1}{\mathcal{E}}V_\alpha\right] \\ &\equiv 2\sum_\alpha \mathbf{Q}_\alpha^\dagger\frac{1}{\mathcal{E}}\mathcal{H}_{\rm ex}\frac{1}{\mathcal{E}}\mathbf{Q}_\alpha,\end{aligned} \tag{H3}$$

where the operator $\mathbf{Q}_\alpha$ that we need to evaluate is simply

$$\mathbf{Q}_\alpha = \left[V_\alpha\frac{1}{\mathcal{E}}T_{ij} + T_{ij}\frac{1}{\mathcal{E}}V_\alpha\right] . \tag{H4}$$

There are two channels to be considered for the intermediate state in which $\mathcal{H}_{\rm ex}$ operates. Channel "0" is one in which site j has orbitals $|0\rangle$ and $|\alpha\rangle$ occupied, whereas channel "1" is one in which site j has orbitals $|1\rangle$ and $|\alpha\rangle$ occupied. Then we may define

$$\left[Q_\alpha^{(\gamma)}\right]_{\sigma,\eta;\sigma',\eta'} = \langle 0|d_{i,\gamma,\sigma}d_{i,\alpha,\eta}\left[V_\alpha\frac{1}{\mathcal{E}}T_{ij} + T_{ij}\frac{1}{\mathcal{E}}V_\alpha\right]d_{i,0,\eta'}^\dagger d_{j,0,\sigma'}^\dagger|0\rangle , \tag{H5}$$

where $\gamma = 0$ or 1. Then

$$\mathcal{H}(i,j) = 2\sum_\alpha\sum_\gamma\left[\mathbf{Q}_\alpha^{(\gamma)}\right]^\dagger\mathcal{H}_{{\rm ex},\alpha}^\gamma\mathbf{Q}_\alpha^{(\gamma)}(\epsilon_\alpha + \epsilon_\gamma + U_{\alpha\gamma})^{-2} , \tag{H6}$$

where

$$\mathcal{H}_{{\rm ex},\alpha}^\gamma = -\tfrac{1}{2}\mathcal{K}_{\alpha\gamma}\left[\mathcal{I}\mathcal{I} + \vec{\sigma}\cdot\vec{\sigma}\right] . \tag{H7}$$

Here $\mathcal{I}$ is the identity operator and $\sigma\cdot\sigma$ denotes the sum over direct products, $\sum_\alpha\sigma_\alpha\sigma_\alpha$. Also, each matrix [$\mathbf{Q}^{(\gamma)}$ or $\mathcal{H}_{\rm ex}^{(\gamma)}$] is a matrix in the direct product of the two spin variables. Any operator in this space can be written as a linear combination of direct product operators. We define $\mathcal{AB}$ via

$$[\mathcal{AB}]_{\sigma,\eta;\sigma',\eta'} = A_{\sigma,\sigma'}B_{\eta,\eta'} . \tag{H8}$$

Explicit calculation of the processes shown in Fig. 7 shows that

$$\left[Q_\alpha^{(0)}\right]_{\sigma,\eta;\sigma',\eta'} = \frac{t_{00}\lambda\langle\alpha|L_\alpha|0\rangle}{2}\left(\frac{\delta_{\sigma,\sigma'}[\sigma_\alpha]_{\eta,\eta'}}{\epsilon_\alpha} - \frac{\delta_{\eta',\sigma}[\sigma_\alpha]_{\eta,\sigma'}}{U_{00}} + \frac{\delta_{\sigma,\sigma'}[\sigma_\alpha]_{\eta,\eta'}}{U_{00}}\right)$$
$$-\frac{\lambda\langle\alpha|L_\alpha|0\rangle t_{\alpha\alpha}}{2\epsilon_\alpha}\delta_{\eta',\sigma}[\sigma_\alpha]_{\eta,\sigma'} - \frac{\lambda\langle\alpha|L_\alpha|1\rangle t_{01}}{2(\epsilon_1+U_{10})}\delta_{\eta',\sigma}[\sigma_\alpha]_{\eta,\sigma'} \ . \tag{H9}$$

To write this in operator form, note that $\frac{1}{2}[\mathcal{I}\mathcal{I}+\vec{\sigma}\cdot\vec{\sigma}]_{\sigma,\eta;\sigma',\eta'}$ is unity if $\sigma=\eta'$ and $\eta=\sigma'$ and is zero otherwise. Thus

$$\left[\mathbf{Q}_\alpha^{(0)}\right] = C_1\left[\mathcal{I}\sigma_\alpha\right] + \tfrac{1}{2}C_2\left[\mathcal{I}\sigma_\alpha\right]\left[\mathcal{I}\mathcal{I}+\vec{\sigma}\cdot\vec{\sigma}\right]$$
$$\equiv (C_1+C_2)\left[\mathcal{I}\sigma_\alpha\right] - C_2\left[\mathcal{I}\sigma_\alpha\right]\left[\mathcal{O}\right] \ , \tag{H10}$$

where $[\mathcal{O}] = [\mathcal{I}\mathcal{I}-\vec{\sigma}\cdot\vec{\sigma}]/2$ and

$$C_1 = \frac{\lambda}{2}\left\{\frac{t_{00}\langle\alpha|L_\alpha|0\rangle}{\epsilon_\alpha} + \frac{t_{00}\langle\alpha|L_\alpha|0\rangle}{U_{00}}\right\} , \tag{H11}$$

$$C_2 = -\frac{\lambda}{2}\left\{\frac{t_{00}\langle\alpha|L_\alpha|0\rangle}{U_{00}} + \frac{t_{\alpha\alpha}\langle\alpha|L_\alpha|0\rangle}{\epsilon_\alpha} + \frac{t_{01}\langle\alpha|L_\alpha|1\rangle}{(\epsilon_1+U_{10})}\right\} . \tag{H12}$$

Also

$$\left[Q_\alpha^{(1)}\right]_{\sigma,\eta;\sigma',\eta'} = \frac{\lambda t_{01}\langle\alpha|L_\alpha|0\rangle}{2}\left(\frac{1}{\epsilon_\alpha}+\frac{1}{(\epsilon_1+U_{10})}\right)\delta_{\sigma,\sigma'}[\sigma_\alpha]_{\eta,\eta'} \ , \tag{H13}$$

so that

$$\left[\mathbf{Q}_\alpha^{(1)}\right] = \frac{\lambda t_{01}\langle\alpha|L_\alpha|0\rangle}{2}\left(\frac{1}{\epsilon_\alpha}+\frac{1}{(\epsilon_1+U_{10})}\right)\left[\mathcal{I}\sigma_\alpha\right] \equiv C_3\left[\mathcal{I}\sigma_\alpha\right] \ . \tag{H14}$$

Thus we have the result

$$\mathcal{H}(i,j) = -\sum_\alpha\left\{\frac{K_{0\alpha}}{(\epsilon_\alpha+U_{0\alpha})^2}\left[C_1^*\mathcal{I}\mathcal{I}+C_2^*\mathcal{I}\mathcal{I}-C_2^*\mathcal{O}\right]\left[\mathcal{I}\sigma_\alpha\right]\left[\mathcal{I}\mathcal{I}+\vec{\sigma}\cdot\vec{\sigma}\right]\left[\mathcal{I}\sigma_\alpha\right]\right.$$
$$\left.\times\left[C_1\mathcal{I}\mathcal{I}+C_2\mathcal{I}\mathcal{I}-C_2\mathcal{O}\right] + \frac{K_{1\alpha}|C_3|^2}{(\epsilon_\alpha+\epsilon_1+U_{1\alpha})^2}\left[\mathcal{I}\sigma_\alpha\right]\left[\mathcal{I}\mathcal{I}+\vec{\sigma}\cdot\vec{\sigma}\right]\left[\mathcal{I}\sigma_\alpha\right]\right\} . \tag{H15}$$

To simplify the above result we use the identity for Pauli matrices,

$$\left[\mathcal{I}\sigma_\alpha\right]\left[\mathcal{I}\mathcal{I}+\vec{\sigma}\cdot\vec{\sigma}\right]\left[\mathcal{I}\sigma_\alpha\right] = \left[\mathcal{I}\mathcal{I}+2\sigma_\alpha\sigma_\alpha-\vec{\sigma}\cdot\vec{\sigma}\right] \ . \tag{H16}$$

From the form of Eq. (H15) we see that all the anisotropic contributions come from the term $\sigma_\alpha\sigma_\alpha$ in Eq. (H16). Keeping only such terms we have

$$\mathcal{H}(i,j) = -\sum_\alpha\left\{\frac{2K_{0\alpha}}{(\epsilon_\alpha+U_{0\alpha})^2}\left[C_1^*\mathcal{I}\mathcal{I}+C_2^*\mathcal{I}\mathcal{I}-C_2^*\mathcal{O}\right]\left[\sigma_\alpha\sigma_\alpha\right]\left[C_1\mathcal{I}\mathcal{I}+C_2\mathcal{I}\mathcal{I}-C_2\mathcal{O}\right] + \frac{2K_{1\alpha}|C_3|^2\left[\sigma_\alpha\sigma_\alpha\right]}{(\epsilon_\alpha+\epsilon_1+U_{1\alpha})^2}\right\} . \tag{H17}$$

The terms involving the operator $\mathcal{O}$ give only isotropic terms. This can be seen by using the equality

$$2\left[\sigma_\alpha\sigma_\alpha\right]\left[\mathcal{O}\right] = \left[\sigma_\alpha\sigma_\alpha\right]\left[\mathcal{I}\mathcal{I}-\vec{\sigma}\cdot\vec{\sigma}\right] = \left[\vec{\sigma}\cdot\vec{\sigma}-\mathcal{I}\mathcal{I}\right] , \tag{H18}$$

which is isotropic. Thus the anisotropic exchange terms are correctly given by $\mathcal{H}(i,j) = (1/4)\sum_\mu J_{\mu\mu}^{\rm anis}\sigma_\mu\sigma_\mu = \sum_\mu J_{\mu\mu}^{\rm anis}S_\mu(i)S_\mu(j)$, with

$$J_{\mu\mu}^{\rm anis} = -2\lambda^2\left\{\frac{|\ L_{0,\mu}^\mu\ |^2\, t_{0,1}^2 K_{1,\mu}}{(\epsilon_\mu+\epsilon_1+U_{1,\mu})^2}\left[\frac{1}{\epsilon_\mu}+\frac{1}{\epsilon_1+U_{0,1}}\right]^2\right.$$
$$\left.+\frac{K_{0,\mu}}{(\epsilon_\mu+U_{0,\mu})^2}\left|\frac{(t_{\mu,\mu}-t_{0,0})L_{0,\mu}^\mu}{\epsilon_\mu}+\frac{t_{0,1}L_{1,\mu}^\mu}{\epsilon_1+U_{0,1}}\right|^2\right\} , \tag{H19}$$

where $L_{\alpha\beta}^\mu$ denotes the orbital angular momentum matrix element, $\langle\alpha|L^\mu|\beta\rangle$.

APPENDIX I: EXPRESSIONS FOR $J_{\mu\mu}^{\rm anis}$

In this appendix we give expressions for $J_{\mu\mu}^{\rm anis}$ [see Eq. (50)] assuming the relations for the hopping matrix elements implied by Eq. (53) and the relations involving $(pd\sigma)$ and $(pd\pi)$ listed in Sec. IV D:

$$t_{xx}=0\ , \quad t_{yy}=t_{zz}=t_{00}/3, \quad t_{01}=-t_{00}/\sqrt{3}\ . \tag{I1}$$

Also, we use the identifications of the Racah coefficients given in Appendix B. Thereby we obtain

$$J_{xx}^{\rm anis}=-2\lambda^2 t_{00}^2\left[\frac{B+C}{3(\epsilon_x+\epsilon_1+A+2B+C)^2}+\frac{3B+C}{(\epsilon_x+A-2B+C)^2}\right]\left[\frac{1}{\epsilon_x}+\frac{1}{\epsilon_1+A-4B+C}\right]^2\ , \tag{I2}$$

$$J_{yy}^{\rm anis}=-2\lambda^2 t_{00}^2\left\{\frac{B+C}{3(\epsilon_x+\epsilon_1+A+2B+C)^2}\left[\frac{1}{\epsilon_x}+\frac{1}{\epsilon_1+A-4B+C}\right]^2 +\frac{3B+C}{(\epsilon_x+A-2B+C)^2}\left[-\frac{2}{3\epsilon_x}+\frac{1}{\epsilon_1+A-4B+C}\right]^2\right\}\ , \tag{I3}$$

$$J_{zz}=-8\lambda^2 t_{00}^2\left\{\frac{4B+C}{3(\epsilon_z+\epsilon_1+A-4B+C)^2}\left[\frac{1}{\epsilon_z}+\frac{1}{\epsilon_1+A-4B+C}\right]^2+\frac{C}{(\epsilon_z+A+4B+C)^2}\left[\frac{2}{3\epsilon_z}\right]^2\right\}\ , \tag{I4}$$

where we set $\epsilon_y=\epsilon_x$ for tetragonal symmetry.

We now compare our results with those of Eq. (14) of BS. In that equation the only hopping matrix element that was included was that between the ground state orbitals of the Cu ions. If we keep only such terms in Eq. (50), we obtain

$$J_{xx}^{\rm anis}=J_{yy}^{\rm anis}=-2\lambda^2 t_{00}^2\frac{3B+C}{\epsilon_x^2(\epsilon_x+A-2B+C)^2}\approx-\frac{\lambda^2 t_{0p_x}^4}{\epsilon_{p_x}^2\epsilon_x^2}\left[\frac{1}{\epsilon_x+A-5B}-\frac{1}{\epsilon_x+A+B+2C}\right]\ , \tag{I5}$$

$$J_{zz}^{\rm anis}=-8\lambda^2 t_{00}^2\frac{C}{\epsilon_z^2(\epsilon_z+A+4B+C)^2}\approx-4\frac{\lambda^2 t_{0,p_x}^4}{\epsilon_{p_x}^2\epsilon_z^2}\left[\frac{1}{\epsilon_z+A+4B}-\frac{1}{\epsilon_z+A+4B+2C}\right]\ , \tag{I6}$$

where we used Eq. (53) to set $t_{00}=t_{0,p_x}^2/\epsilon_{p_x}$ and wrote our expression in terms of singlet and triplet energy denominators to facilitate comparison with BS. We note the following differences between their results and ours.

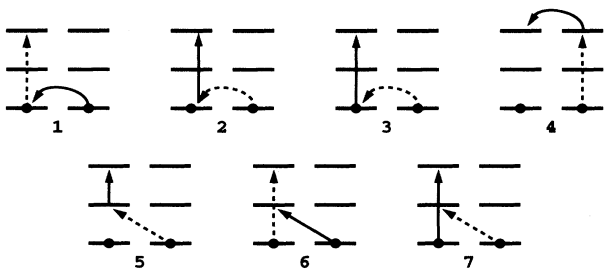

FIG. 7. Processes (1–5) which contribute to $Q_\alpha^{(0)}$ [in the order written in Eq. (H9)] and those (6–7) which contribute to $Q_\alpha^{(1)}$ [in the order written in (H13)]. Here the left site is the site i and the right site is site j. The dashed line depicts the first matrix element and the full line the second. The orbitals are the ground state lowest, the state $|1\rangle$ next, and the state $|\alpha\rangle$ highest. In term 3 the second process promotes the left-hand hole to an excited state, whereas in term 2 the left-hand hole remains in the ground state.

(1) Our results are smaller by an overall factor of 2. (2) Instead of our evaluation in which $t_{00}=t_{pd}^2/\Delta$ (to use their notation), they use $t_{00}=t_{pd}^2/(\Delta+E_{xy})$. (3) The last energy denominator in their J_{yy} is wrong: their $4B$ should be replaced by B. With respect to the first difference we would note that, as described in the text, we did compare results from the full diagonalization with those using the spin Hamiltonian. Thus an error by a factor of 2 in our calculations is extremely unlikely. Difference 2 comes about because BS do not sum over all processes. In particular, consider process 1 of Fig. 7. When the hop actually consists of two hops, one from a Cu to an O and another from an O to a Cu, this process corresponds to two orderings of the three perturbations, one in which the spin-orbit interactions comes first and one in which it comes second. (The case when it comes third should be identified with processes 2 or 3 of Fig. 7.) Summing over these two orderings converts the denominator $(\Delta+E_{xy})$ of BS into ours. The correct energy denominator can also be obtained from Eqs. (E1) and (E2).

In principle, we ought also to compare with Eq. (20) of BS, where contributions involving $\psi_1(\mathbf{r})$ are claimed to be included. Here their results are so different from ours that we cannot identify their terms with ours. In particular, we note the following. (1) Although they claim

to include the effects of $\psi_1(\mathbf{r})$, their expressions do not include any energy denominators which depend on the associated crystal field energy ϵ_1. Obviously, when $\epsilon_1 = 0$, the ground manifold would be described by a totally different spin Hamiltonian to remove the spin *and* orbital degeneracy. (2) In our Eq. (50) (Ref. 18), t_{01} enters in several places and thus gives rise to many more terms than appear in BS. (3) When $t_{01} \neq 0$, as we have noted, each bond has biaxial anisotropy in contrast to the axial anisotropy they implicitly assume. (4) We include hopping between excited crystal field states. These give contributions which are of the same order of magnitude as those involving hopping into the ground state. (5) We have not included covalency corrections. It may indeed be a good idea to include such corrections, but at present the parameters are themselves so uncertain that we regard this correction as a refinement.

[1] T. Moriya, Phys. Rev. **120**, 91 (1960).

[2] I. Dzyaloshinskii, J. Phys. Chem. Solids **4**, 241 (1958).

[3] R. J. Birgeneau and G. Shirane, in *Physical Properties of High Temperature Superconductors I*, edited by D. M. Ginsberg (World Scientific, Singapore, 1989), and references therein.

[4] T. Thio, T. R. Thurston, N. W. Preyer, P. J. Picone, M. A. Kastner, H. P. Henssen, D. R. Gabbe, C. Y. Chen, R. J. Birgeneau, and A. Aharony, Phys. Rev. B **38**, 905 (1988).

[5] D. Coffey, K. S. Bedell, and S. A. Trugman, Phys. Rev. B **42**, 6509 (1990).

[6] D. Coffey, T. M. Rice, and F. C. Zhang, Phys. Rev. B **44**, 10112 (1991).

[7] L. Shekhtman, O. Entin-Wohlman, and A. Aharony, Phys. Rev. Lett. **69**, 836 (1992).

[8] C. J. Peters, R. J. Birgeneau, M. A. Kastner, H. Yoshizawa, Y. Endoh, J. Tranquada, G. Shirane, Y. Hidaka, M. Oda, M. Suzuki, and T. Murakami, Phys. Rev. B **37**, 9761 (1988).

[9] B. Keimer, R. J. Birgeneau, A. Cassanho, Y. Endoh, M. Greven, M. A. Kastner, and G. Shirane, Z. Phys. B **91**, 373 (1993).

[10] For $Sr_2CuO_2Cl_2$, see D. Vaknin, S. K. Sinha, C. Stassis, L. L. Miller, and D. C. Johnston, Phys. Rev. B **41**, 1926 (1990); M. Greven, R. J. Birgeneau, Y. Endoh, M. A. Kastner, M. Matsuda, and G. Shirane, Phys. Rev. Lett. **74**, 1096 (1994).

[11] For tetragonal rare-earth cuprates R_2CuO_4 (R=Nd,Pr, etc.) see, for instance, M. Matsuda, K. Yamada, K. Kakurai, H. Kadowaki, T. R. Thurston, Y. Endoh, Y. Hidaka, R. J. Birgeneau, M. A. Kastner, P. M. Gehring, A. H. Moudden, and G. Shirane, Phys. Rev. B **42**, 10 098 (1990).

[12] P. W. Anderson, Phys. Rev. **115**, 2 (1959).

[13] J. Hubbard, Proc. R. Soc. A **276**, 238 (1963); A **281**, 401 (1964).

[14] In fact, this result was already noticed by T. A. Kaplan [T. A. Kaplan, Z. Phys. B **49**, 313 (1983)], in the context of a linear chain. We thank T. A. Kaplan for giving us this reference.

[15] L. Shekhtman, A. Aharony, and O. Entin-Wohlman, Phys. Rev. B **47**, 174 (1992).

[16] N. E. Bonesteel, Phys. Rev. B **47**, 9144 (1993).

[17] F. Barriquand and G. A. Sawatzky, Phys. Rev. B **50**, 16 649 (1994).

[18] T. Yildirim, A. B. Harris, O. Entin-Wohlman, and A. Aharony, Phys. Rev. Lett. **73**, 2919 (1994).

[19] T. Yildirim, A. B. Harris, O. Entin-Wohlman, and A. Aharony, Phys. Rev. Lett **72**, 3710 (1994). The spin structures for cuprates containing magnetic rare-earth ions, such as Nd_2CuO_4, turn out to be more complicated than assumed in this paper due to the single-ion anisotropy of the rare-earth ion. For more discussion see Refs. 43 and 44.

[20] In Ref. 18 there are some misprints in the Hamiltonian of Eq. (3). The last two lines there are rewritten correctly here in Eq. (4).

[21] H. J. Zeiger and G. W. Pratt, *Magnetic Interactions In Solids* (Clarendon, Oxford, 1973), p. 85.

[22] W. Koshibae, Y. Ohta, and S. Maekawa, Phys. Rev. B **47**, 3391 (1993).

[23] In Eq. (47) when $\alpha \neq \alpha'$, $f^{\dagger}_{k\alpha\mu}$ can be anticommuted to the right of $f_{k\alpha'\mu'}$. When $\alpha = \alpha'$ the fermion operators can be written as $(n^2_{k\alpha} - n_{k\alpha})/2$, where $n_{k\alpha} = Q_{\alpha\alpha}(k,k)$.

[24] W. Koshibae, Y. Ohta, and S. Maekawa, Phys. Rev. Lett. **71**, 467 (1993).

[25] H. Eskes and J. H. Jefferson, Phys. Rev. B **48**, 9788 (1993).

[26] H. Eskes and G. Sawatzky, Phys. Rev. B **44**, 9656 (1991).

[27] A. F. McMahan, J. F. Annett, and R. M. Martin, Phys. Rev. B **42**, 6268 (1990).

[28] M. S. Hybertsen, M. Schluter, and N. E. Christensen, Phys. Rev. B **39**, 9028 (1989).

[29] J. S. Griffith, *The Theory of Transition Metal Ions* (Cambridge University Press, Cambridge, 1961).

[30] C. E. Moore, *Atomic Energy Levels*, Natl. Bur. Stand. (U.S.) Circ. No. 467 (U.S. GPO, Washington, D.C., 1958), Vols. 1–3.

[31] H. Eskes, L. H. Tjeng, and G. A. Sawatzky, Phys. Rev. B **41**, 288 (1990).

[32] H. Eskes, G. A. Sawatzky, and L. F. Ferner, Physica C **160**, 424 (1989).

[33] L. F. Matheiss, Phys. Rev. B **5**, 290 (1972).

[34] C. L. Henley, J. Appl. Phys. **61**, 3962 (1987).

[35] This degeneracy is also removed by second-neighbor interactions.

[36] J. M. Tranquada, D. E. Cox, W. Kunnmann, H. Moudden, G. Shirane, M. Suenaga, P. Zolliker, D. Vaknin, S. K. Sinha, M. S. Alvarez, A. J. Jacobson, and D. C. Johnston, Phys. Rev. Lett. **60**, 156 (1988).

[37] P. Burlet, C. Vettier, M. J. G. M. Jurgens, J. Y. Henry, J. Rossat-Mignod, H. Noel, M. Potel, P. Gougeon, and J. C. Levet, Physica C **153–155**, 1115 (1988).

[38] J. Rossat-Mignod, L. P. Regnault, P. Bourges, P. Burlet, C. Vettier, and J. Y. Henry, in *Selected Topics in Superconductivity*, edited by L. C. Gupta and M. S. Multani (World Scientific, Singapore, 1993), p. 265.

[39] G. Shirane, Acta Crystallogr. **12**, 282 (1959).

[40] E. Shender, Sov. Phys. JETP **56**, 178 (1982).

[41] Our results for the isotropic component of exchange, $J_{\rm av} = (J_{xx} + J_{yy} + J_{zz})/3$, for the Cu—O—Cu cluster are in qual-

itative agreement with those of Eskes and Jefferson (Ref. 25). For the parameters used in Fig. 3 our numerical results for the Cu—O—Cu cluster give $J_{\mathrm{av}} \approx 0.06$ eV, a value which is smaller than the perturbative result by a factor of about 2. (Perturbation theory does not work well for J_{av} because $t_{\mathrm{Cu-O}}/e_{p_\alpha}$ is not small.) However, we reiterate that the results obtained from perturbation theory and those from numerical diagonalization of the Cu—Cu and Cu—O—Cu cluster for the *anisotropy* of the exchange agree within about 10%.

[42] The spin Hamiltonian for a square plaquette in the absence of spin-orbit interactions is discussed in M. Takahashi, J. Phys. C **10**, 1289 (1977). We thank Dr. J. H. Jefferson for supplying this reference.

[43] S. Skanthakumar, J. W. Lynn, and I. W. Sumarlin, Phys. Rev. Lett. **74**, 2842 (1995).

[44] T. Yildirim, A. B. Harris, O. Entin-Wohlman, and A. Aharony, Phys. Rev. Lett. **74**, 2843 (1995).