Effect of the Coulomb repulsion and oxygen level on charge distribution and superconductivity in the Emery model for cuprates superconductors.

Louis-Bernard St-Cyr[1] and David Sénéchal[1]

[1]Département de physique and Institut quantique, Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1

(Dated: April 30, 2025)

The Emery model (aka the three-band Hubbard model) offers a simplified description of the copper-oxide planes that form the building blocks of high-temperature superconductors. By contrast with the even simpler one-band Hubbard model, it differentiates between copper and oxygen orbitals and thus between oxygen occupation (n_p) and copper occupation (n_d). Here we demonstrate, using cluster dynamical mean field theory, how the two occupations are related to the on-site Coulomb repulsion U on the copper orbital and to the energy difference ϵ_p between oxygen and copper orbitals. Since the occupations (n_p and n_d) have been estimated from NMR for a few materials (LCO, YBCO and NCCO), this allows us to estimate the value of $U - \epsilon_p$ for these materials, within this model. We compute the density of states for these and the effect of (U, ϵ_p) on the n_d-n_p curve, superconductivity, and antiferromagnetism.

I. INTRODUCTION

High-temperature superconductors, nearly forty years after their discovery, still constitute a theoretical challenge. The dozens of materials in this category all share a common feature: planes of copper oxide (CuO_2). Thus one would naturally expect a universal explanation of superconductivity in cuprates based on a model of these planes only. The one-band Hubbard model has long been the focus of research in this direction, in part because of its relative simplicity: a few hopping parameters, the Coulomb repulsion U and the chemical potential controlling the electron density. That model features a single band crossing the Fermi level, argued to be a combination of copper $d_{x^2-y^2}$ and oxygen p_x and p_y orbitals, forming excitations above what we call the Zhang-Rice singlet [1]. Various theoretical approaches applied to this model have provided evidence for a dome of d-wave superconductivity away from half-filling, in addition to a robust antiferromagnetic phase close to half-filling [2–6], although there is also evidence, in some parameter range, that superconductivity is not the lowest-energy state [7, 8].

A more realistic description of the CuO_2 planes is provided by the Emery-VSA (Varma-Schmitt-Rink-Abrahams) model [9, 10], which involves the copper $d_{x^2-y^2}$ orbital and two oxygen p-orbitals per unit cell. Within that model, the question of the distribution of holes between the copper and oxygen orbitals now has meaning. Since the occupation of both orbitals can be estimated from nuclear magnetic resonance (NMR) experiments [11], a comparison with NMR data can in principle be used to impose constraints on some of the parameters of that model, namely the on-site Coulomb interaction on the copper orbitals and the energy levels of both orbitals. This is what this paper is mainly about. This will also lead us to revisit the relation of superconductivity with U and the matter of localized vs itinerant antiferromagnetism in NCCO.

This paper is organized as ~~follow~~follows: In Sect. II we review the definition of the Emery model and the method used (Cluster dynamical mean-field theory or CDMFT). The latter is mostly discussed in the Appendix. In Sect. III we
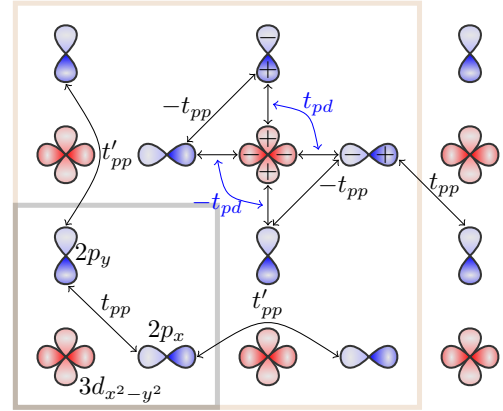
Figure 1. Schematic representation of the Emery model used in this work: Copper $d_{x^2-y^2}$ orbitals are shown in red and Oxygen p_x,p_y orbitals in blue. Nonzero hopping parameters are indicated. The gray square defines the unit cell, and the beige square the super unit cell used in CDMFT (see appendix).

present our numerical results for the relative hole and electron distribution among the Cu and O orbitals as a function of doping, interaction strength U and oxygen energy level ϵ_p. We also show the effect of these parameters on the superconducting order parameter at zero temperature and we compare the charge-distribution curves with those of real materials obtained from NMR. Finally, in Sect. IV we discuss the applicability of the Emery model to these materials and conclude.

II. MODEL AND METHOD

A. The Emery model

Let us first review the basics of the Emery model. It is schematically illustrated in Fig. 1. In second-quantized language, the kinetic energy $H_{\text{kin.}}$ and interaction $H_{\text{int.}}$ are ex-

pressed as follows:

$$
\begin{aligned}
H_{\text{kin.}} = \sum_{\mathbf{r},\sigma} \Big\{ &(\epsilon_p - \mu)(n_{\mathbf{r}+\mathbf{a}_x} + n_{\mathbf{r}+\mathbf{a}_y}) - \mu n_{\mathbf{r}} \\
&+ t_{pd} \sum_{\mathbf{r}} \sum_{\mathbf{a}=\mathbf{a}_x,\mathbf{a}_y} \sum_{\nu=\pm 1} \nu d^{\dagger}_{\mathbf{r},\sigma} p_{\mathbf{r}+\nu\mathbf{a},\sigma} + \text{H.c} \\
&+ t_{pp} \sum_{\mathbf{r}} \sum_{\nu,\nu'=\pm 1} \nu\nu' p^{\dagger}_{\mathbf{r}+\nu\mathbf{a}_x,\sigma} p_{\mathbf{r}+\nu'\mathbf{a}_y,\sigma} + \text{H.c} \\
&+ t'_{pp} \sum_{\mathbf{r}} \sum_{\mathbf{a}=\mathbf{a}_x,\mathbf{a}_y} p^{\dagger}_{\mathbf{r}-\mathbf{a},\sigma} p_{\mathbf{r}+\mathbf{a},\sigma} + \text{H.c} \Big\}
\end{aligned} \tag{1}
$$

$$
H_{\text{int.}} = U \sum_{\mathbf{r}} n_{\mathbf{r}\uparrow} n_{\mathbf{r}\downarrow} \ , \tag{2}
$$

where $\mathbf{r}$ labels the position of the copper atoms and $2\mathbf{a}_x = 2\mathbf{a}_y$ are the Bravais square lattice vectors. The operator $d_{\mathbf{r},\sigma}$ annihilates an electron of spin σ on the copper $3d_{x^2-y^2}$ orbital located at the position $\mathbf{r}$. The operator $p_{\mathbf{r}+\mathbf{a},\sigma}$ annihilates an electron of spin σ on the oxygen $2p_x$ or $2p_y$ orbital located at $\mathbf{r}+\mathbf{a}_{x,y}$. $n_{\mathbf{r},\sigma} = d^{\dagger}_{\mathbf{r},\sigma} d_{\mathbf{r},\sigma}$ is the number of electrons of spin σ in the Cu orbital at site $\mathbf{r}$, whereas $n_{\mathbf{r}+\mathbf{a}_x}$ and $n_{\mathbf{r}+\mathbf{a}_y}$ are the number of electrons (summed over spins) on each of the oxygen orbitals. ϵ_p represents the difference

in energy between the oxygen and copper orbital and the energy ϵ_d of the copper orbital is used as reference. t_{pd} is the first-neighbor hopping amplitude between a copper and the oxygen orbitals that are directly around it. t_{pp} is the hopping amplitude between two nearest-neighbor oxygen orbitals (diagonally) and t'_{pp} is the third-neighbor hopping amplitude between two oxygen orbitals over a copper orbital. The sign convention for the hopping parameters used in this work is illustrated in Fig. 1. U represents the energy cost of double occupancy caused by the Coulomb repulsion between electrons located on the same copper orbital. The corresponding local repulsion U_p on the oxygen orbitals is neglected because of its smaller value, as justified by density functional theory (DFT)[12], and also because the oxygen orbitals are nearly filled in the cases of interest. The chemical potential μ allows to control the filling (electron or hole doping) of the system.

We can express the kinetic part $H_{\text{kin.}}$ more compactly in reciprocal space:

$$
H_{\text{kin.}} = \sum_{\mathbf{k}\sigma} \Phi^{\dagger}_{\mathbf{k},\sigma} \mathbf{h}_{\mathbf{k}} \Phi_{\mathbf{k},\sigma} \tag{3}
$$

where the different annihilation operators are assembled in the multiplet $\Phi_{\mathbf{k},\sigma} = (d_{\mathbf{k}\sigma}, p^x_{\mathbf{k}\sigma}, p^y_{\mathbf{k}\sigma})$ and the 3×3 matrix $\mathbf{h}_{\mathbf{k}}$ is

$$
\mathbf{h}_{\mathbf{k}} = \begin{pmatrix} 0 & t_{pd}(1-e^{-ik_x}) & t_{pd}(1-e^{-ik_y}) \\ t_{pd}(1-e^{ik_x}) & \epsilon_p + 2t'_{pp}\cos k_x & t_{pp}(1-e^{ik_x})(1-e^{-ik_y}) \\ t_{pd}(1-e^{ik_y}) & t_{pp}(1-e^{-ik_x})(1-e^{ik_y}) & \epsilon_p + 2t'_{pp}\cos k_y \end{pmatrix} . \tag{4}
$$

Note that we did not renormalize ϵ_p by including the $-2t_{pp}$ contribution of the oxygen orbital energy, in contrast with Ref. [13].

Since the unit cell contains three orbitals, the maximum occupation is $n = 6$. The "half-filling" state is set to five electrons per unit cell, with the O orbitals ~~being~~ fully filled and the Cu orbital half-filled, in the zero-hopping limit. At finite hopping the O orbitals are slightly less than fully filled and the Cu orbital is slightly more than half-filled [11]. Such an undoped state describes a charge transfer insulator (CTI) when the interaction U is large enough to split the Cu energy band into an upper (UHB) and lower (LHB) Hubbard band [14]. At those large U values, the UHB is pushed past the energy of the oxygen band, also called the charge-transfer band (CTB). Partial holes on the O orbitals cause a loss of electronic weight in the CTB and give rise to the Zhang-Rice singlet band (ZRB) [1] made of low-energy states formed by the mixing of Cu and neighboring O orbitals. In a CTI, an insulating gap, called a charge transfer gap (CTG), is formed between the ZRB and the UHB upon increasing U. This CTG in the Emery model is the analog of the Mott gap in a Mott insulator (MI). The different spectral features of the Emery model are illustrated in the density of

states of Fig. 6a below.

Finally, a note on *double counting*. The relative energy level ϵ_p between O and Cu orbitals computed from DFT is subject to a renormalization when the Coulomb interaction U is introduced. Indeed, Coulomb interactions are already taken into account at the DFT level and impact the value of ϵ_p. When U is introduced in an explicit way in the Emery model, the DFT effect of that local Coulomb interaction should be subtracted in order to avoid counting it twice. This of course is rather difficult to evaluate exactly. A rough estimate of this double counting can be obtained by computing the effect of U at the mean-field (Hartree) level, i.e., by replacing

$$
U n_{\mathbf{r}\uparrow} n_{\mathbf{r}\downarrow} \quad \text{by} \quad U\left(\bar{n}_{\mathbf{r}\uparrow} n_{\mathbf{r}\downarrow} + n_{\mathbf{r}\uparrow}\bar{n}_{\mathbf{r}\downarrow} - \bar{n}_{\mathbf{r}\uparrow}\bar{n}_{\mathbf{r}\downarrow}\right) \ , \tag{5}
$$

where $\bar{n}_{\mathbf{r}\sigma}$ is the occupation of the Cu orbital of spin σ at site $\mathbf{r}$. In the paramagnetic state, $\bar{n}_{\mathbf{r}\uparrow} = \bar{n}_{\mathbf{r}\downarrow}$ and this brings a contribution $\frac{1}{2}U\bar{n}_{\mathbf{r}}$ to the copper orbital energy ϵ_d that must be subtracted, which leads to a corresponding correction $\frac{1}{2}U\bar{n}_{\mathbf{r}}$ to ϵ_p since ϵ_d is used as the reference energy. Such a correction was already applied in Refs [15, 16] but manifestly depends on U. Hence one cannot set U in the model without also playing with ϵ_p. This is why in the following we

study both the effect of U and ϵ_p on the properties of the model.

B. Cluster dynamical mean-field theory

In this work we approximately solve the Emery model using cluster dynamical mean field theory (CDMFT) [17, 18] with an exact diagonalization (ED) impurity solver at zero temperature, implemented in the open-source library Pyqcm [19]. CDMFT is an extension of dynamical mean field theory (DMFT). DMFT is an embedding procedure that incorporates the correlations of a single correlated orbital with its lattice environment into a hybridization of that orbital with a set of non-interacting orbitals (the "bath"). DMFT assumes that the self-energy is purely local, i.e., its momentum dependence is neglected. To include the latter, which is essential to describe $d_{x^2-y^2}$ superconductivity in cuprates, we need to upgrade the single-site impurity problem (DMFT) into a cluster impurity problem (CDMFT), in which a four-site cluster of Cu atoms is embedded into an effective environment. In this work the decomposition of the infinite lattice into clusters is done in the same way as in Ref. [20]: the lattice is tiled with identical units, each of which consisting of a cluster of $N_d = 4$ correlated Cu sites and another, overlapping cluster of $N_O = 8$ uncorrelated oxygen sites, as shown in Fig. 1. The correlated Cu cluster is augmented by an 8-orbital bath that represents that cluster's environment, and defines an "impurity problem". See Appendix A for technical details. Ideally, one would prefer an infinite number of bath orbitals, but as we are using an ED impurity solver the computational complexity increases exponentially with the number of bath orbitals. It was shown in [21] that the hybridization of the impurity problem is very well approximated with a discrete bath of only 8 orbitals.
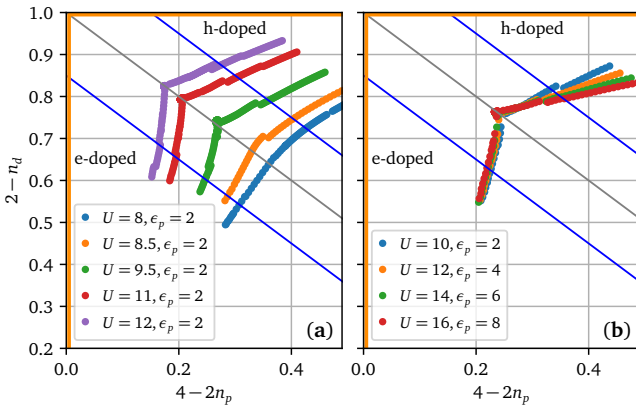
Figure 2. Relative filling n_d of copper and n_p of oxygen in a model for Bi$_2$Sr$_2$Ca2Cu$_3$O$_{10}$. Panel (a): the effect of U for a fixed value of ϵ_p. Panel (b): the effect of U when the nominal gap $U - \epsilon_p$ is kept constant. The gray line sits at half-filling and the blue lines at ±15% doping. The orange lines show the zero-hopping limit.

Table I. Hopping parameters used in this work, from Refs [15, 16].

Material	t_{pd}/t_{pp}	t'_{pp}/t_{pp}	$(\epsilon_p - \epsilon_d)/t_{pp}$
BSSCO	2.139	0.258	2.01
LSCO	2.172	0.1609	4.08
YBCO	1.902	0.223	3.05
NCCO	2.148	0.389	2.98

III. RESULTS

A. Doping curves

The effect of U and ϵ_p on the distribution of holes and electrons in the Emery model can be seen by plotting the occupation of O orbitals against that of the Cu orbital. We have computed the distribution of electrons and holes for a range of values of U and ϵ_p in a model of Bi$_2$Sr$_2$Ca$_2$Cu$_3$O$_{10}$ (BSCCO) as shown in Fig. 2. The hopping parameters are taken from ~~ref~~Ref. [15]. Note that we also obtained solutions on the "wrong" side of half-filling for this material (and for others in this paper), i.e, for both hole and electron doping, irrespective of its actual hole- or electron-doped character.

Let n_d (n_p) denote the average occupation of copper (oxygen) orbitals. In Fig. 2, one notices a discontinuity in the slope $\frac{1}{2}dn_d/dn_p$ of the curves at half-filling, i.e., along the diagonal gray line defined by $n_d + 2n_p = 5$. This discontinuity occurs for all curves with ~~$U > 9$~~ $U > 8.5$ on the figure and is a signature of the charge transfer gap (CTG). On the hole-doped side, the smaller slope indicates that the hole density increases faster on the O orbitals than on the Cu orbital when removing electrons. By contrast, the slope on the electron-doped side indicates that the doped electrons mostly go to the Cu orbital. Note from Fig. 2a that the slope discontinuity between the electron- and hole-doped sides of the curve increases with U, that is, as U gets stronger, the doped hole (electrons) go increasingly towards the O (Cu) orbitals. Both of these observations are in agreement with experimental data [22, 23] (see Fig. 5) and with theoretical calculations using a different impurity solver [24, 25]. In the atomic limit, i.e., when setting all hopping terms to zero (but keeping ϵ_p), the corresponding curve would follow the $n_p = 2$ axis on the electron-doped side, and the $n_d = 1$ axis on the hole-doped side (orange lines ~~on~~ in the figure). At lower values of U, crossing the half-filling axis has no effect on the slope of the curve, corresponding to a continuously hybridized Cu-O band at the Fermi level.

In the zero-hopping limit, the CTG gap is precisely $\Delta_n = U - \epsilon_p$. We will call this the *nominal gap*, as opposed to the real CTG gap obtained from the computed density of states when hopping is present. The actual CTG is much smaller than $U - \epsilon_p$ because of the finite bandwidth in the presence of hopping. From Fig. 2b, one sees that increasing ϵ_p while keeping the nominal gap $U - \epsilon_p$ constant barely affects the position of the curve along the half-filling line and makes the curves flatter on the hole-doped side while hav-

ing no effect on the electron-doped side. Indeed, increasing ϵ_p pushes the lower Hubbard band to even lower energies, thus making the charge transfer band even more dominated by the oxygens and removing copper content from the ZRS band. Hence added holes go more and more on the oxygen atoms. On the other hand, that increase has apparently much less effect on the oxygen content of the upper Hubbard band, because the relative location of the UHB and the CTB remains roughly the same, and so does the actual CTG. We conclude from this that the ZRS band gets its Cu content from both the lower and upper Hubbard bands. Let us point out that increasing ϵ_p at constant U has the same effect as decreasing U at constant ϵ_p, like in Fig. 2a: This causes the CTG to shrink and a smaller discontinuity of the doping curves across half-filling. The metallic or insulating character of the solutions thus depends on the relative value U and ϵ_p, not on each of them separately.
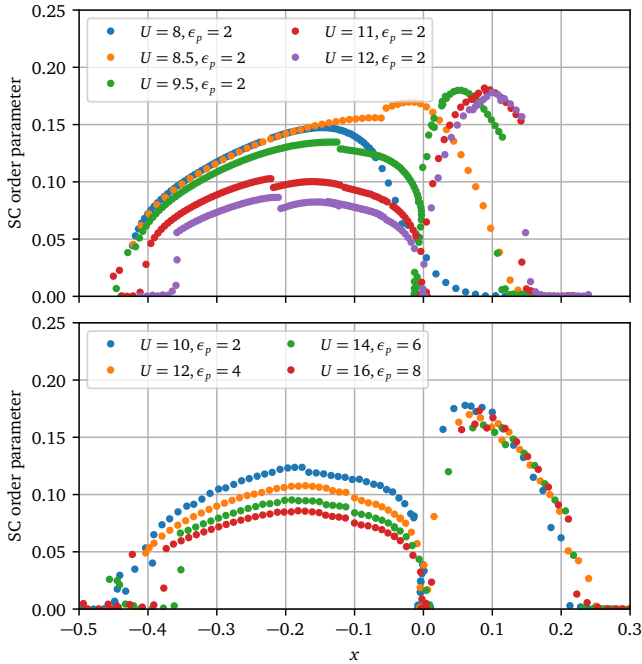
Figure 3. d-wave superconducting order parameter vs electron-doping x in a models for BSCCO with varying U at fixed ϵ_p (top) and varying U at constant CTG, i.e., varying ϵ_p (bottom).

B. Superconductivity

Fig. 3 shows the d-wave superconducting order parameter Ψ, the average of the pairing operator (A11), as a function of electron doping x (negative for hole doping) for the same values of U and ϵ_p as in Fig. 2. On the top panel, U is varied while ϵ_p is kept constant, whereas in the bottom panel U is varied while $U - \epsilon_p$ is kept constant. We first note that superconductivity exists even when the CTG vanishes ($U = 8$ and ~~$U = 9$~~ $U = 8.5$). It is even at its maximum at half-filling just under the Mott transition (~~$U = 9$~~ $U = 8.5$),
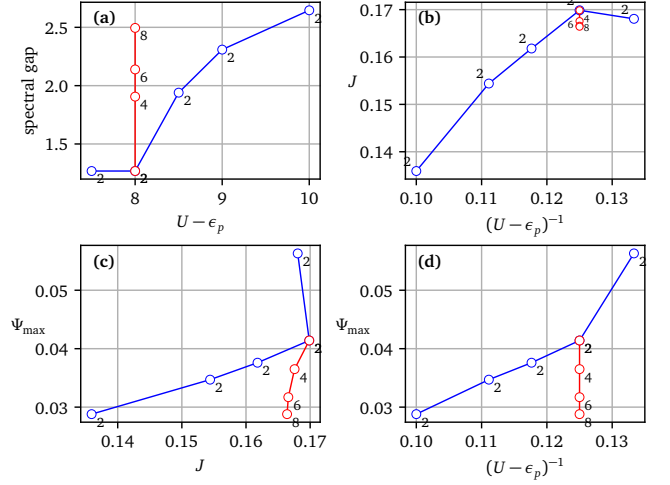
Figure 4. Various properties of the Emery model computed from CDMFT and the parameters used for BSCCO in Fig. 2. Panel (a) shows the CTG computed from the density of states, as a function of $U - \epsilon_p$. Panel (b) shows the superexchange estimate J as a function of $(U - \epsilon_p)^{-1}$. Panel (c) shows the maximum superconducting order parameter on the hole-doped side (see Fig. 3) as a function of the superexchange J at half-filling, and Panel (d) the same quantity as a function of $(U - \epsilon_p)^{-1}$. The value of ϵ_p is written next to each data point.

but shifts towards the hole-doped region as U is lowered to $U = 8$. These two curves are thus not in the CTI regime, making them less relevant to the study of actual cuprates. Note that these calculations ignore the presence of antiferromagnetism near half-filling. In all cases, the maximum order parameter Ψ_{max} on the electron-doped side is not affected by U or ϵ_p, whereas it decreases with U on the hole-doped side, although less markedly if the nominal gap $U - \epsilon_p$ is kept constant.

In Ref. [20] it was argued that the main factor influencing Ψ_{max} was the effective super-exchange J computed at half-filling. This quantity is computed on the impurity (the four-site plaquette) and is the position of the first (dominant) pole in the dynamic antiferromagnetic spin susceptibility as a function of frequency [20, 26]. Fig. 4b shows how J depends on the inverse nominal gap, whereas Fig. 4c shows how Ψ_{max} is related to J. As can be seen, the relation between the two is roughly linear for fixed $\epsilon_p = 2$ until the CTG almost closes. The linear relation also holds as a function of the inverse nominal gap (Fig. 4d). Note that the data shown in red, with a constant nominal gap, corresponds to a single value of J, but a sizeable variation of Ψ_{max}, demonstrating that J is not the only factor at play. When the CTG is very small, as for the data shown in red, the linear relationship between the super-exchange J and Ψ_{max} is lost. In this small CTG regime, it is the size of the spectral gap that has the biggest impact on Ψ_{max}: A smaller spectral gap leads to a larger value of Ψ_{max}. Furthermore, when the CTG completely closes, the relationship between J and Ψ_{max} changes even more dramatically: An increase in

J could very well lead to a decrease in $\Psi_{\max}$, as is the case for the curves $U = 8$, $\epsilon_p = 2$ and $U = 8.5$, $\epsilon_p = 2$ in Fig. 3.

Interestingly, all curves of Ψ with a ~~nonzero~~ non-zero CTG also show a discontinuity in the hole-doped region ~~,~~ that appears at optimal doping, as illustrated in Fig. 3. This discontinuity within the superconducting phase was also observed in Refs [20, 27] and marks the boundary between a superconducting phase with and without pseudogap.
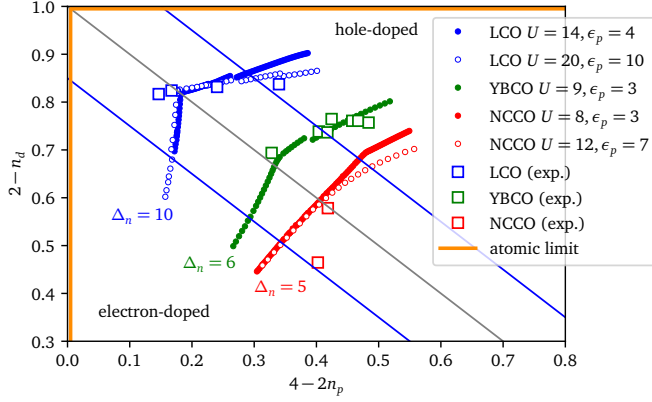
Figure 5. Relative filling n_d of Cu n_p of O orbitals. NMR data from [11] (squares) compared with our CDMFT results for LCO, YBCO and NCCO. The orange lines are the zero-hopping limit of the Emery model. The values of the nominal gap $\Delta_n = U - \epsilon_p$ are indicated.

C. Comparison with NMR

We studied three different cuprate compounds and computed the hole distribution between oxygen and copper sites in the superconducting phase, with the goal of estimating the strengths of the Coulomb interaction U and of the orbital energy ϵ_p in those materials. We looked at two hole-doped cuprates, La_2CuO_4 (LCO) and $YBa_2Cu_3O_7$ (YBCO), using the hopping parameters from Ref. [15] given in Table I, and varying also ϵ_p from the values of that table. We also looked at one electron-doped cuprate, Nd_2CuO_4 (NCCO), using the band parameters from Ref. [16]. We compare in Fig. 5 the hole distributions of these three materials computed from CDMFT with the nuclear magnetic resonance (NMR) data from Ref. [11]. For each of the three materials, we computed the doping curves for many values of U and the value of ϵ_p of Table I. We show in Fig. 5 only those curves that cross half-filling closest to their respective experimental data points. We also show doping curves (small open symbols) for larger values of U and correspondingly larger values of ϵ_p for LCO and NCCO, thus keeping the nominal gap $\Delta_n = U - \epsilon_p$ constant.

Fig. 6 shows the computed density of states (DoS) at half-filling for each of the three materials studied, with the values of U corresponding to the full circles in Fig. 5. We see that LCO has a large CTG of approximately $3.5t_{pp}$ (panel (a)), whereas in YBCO this gap is only about $1.0t_{pp}$ (panel

(b)) and in NCCO the CTG has disappeared (panel (c)). The absence of gap at the Fermi level for NCCO was also predicted by a hybrid DFT-DMFT study [16] in the case where long range magnetic order is not allowed, like here.
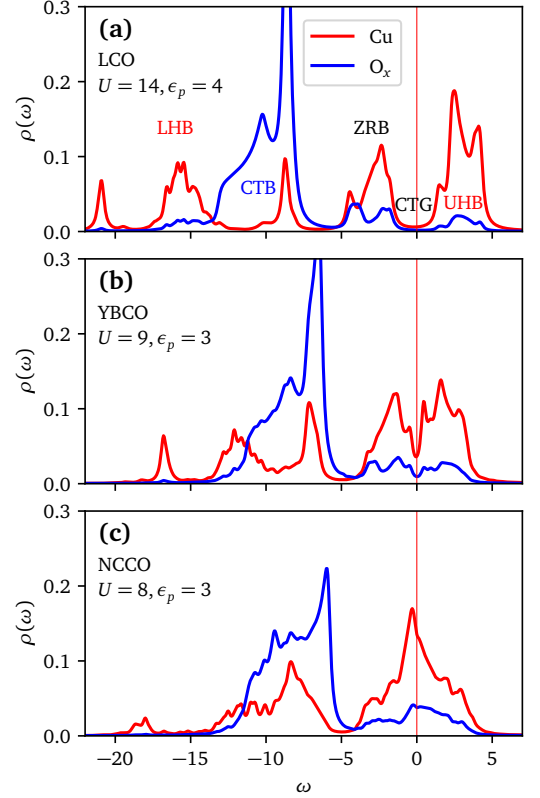
Figure 6. Density of states for the solutions crossing the half-filling axis of Fig. 5 for each of the three cuprate compounds. The Fermi level is the vertical red line. The Cu contribution is shown in red and the contribution of one of the oxygen orbitals in blue (the two oxygens contribute equal amounts).

We do not observe a superconducting dome on the electron doped side in NCCO, rather like the $U = 8$ curve for BSCCO in Fig. 3a, even though there should be one as the material is superconducting. As we mentioned before, DMFT+LDA computations on a similar model showed that by allowing ~~long range~~ long-range magnetic order a gap should open up around the Fermi level at half-filling [16]. We modified the impurity model (A2) in order to allow for antiferromagnetic order, to see ~~wether~~ whether a CTG would appear for NCCO at $U = 8.0$. To measure the strength of the antiferromagnetic order we defined the order parameter

$$M = \sum_{\mathbf{r}} e^{\mathbf{Q}\cdot\mathbf{r}}(n_{\mathbf{r}\uparrow} - n_{\mathbf{r}\downarrow}) \tag{6}$$

with the antiferromagnetic wave-vector $\mathbf{Q} = (\pi, \pi)$. At $(U, \epsilon_p) = (8.0, 3)$ only a weak antiferromagnetic order was observed, which was not strong enough to open a CTG. We then studied the relationship between the ~~value of interaction~~ interaction value U and the antiferromagnetic

order parameter, as shown in Fig. 7. We notice a discontinuity in the value of $\langle M \rangle$ in between $U \approx 9.06$ and $U \approx 8.90$. This discontinuity is characterized by the opening of an antiferromagnetic order-induced CTG at the Fermi level, as can be seen ~~from~~ in Fig. 8. The nature of antiferromagnetism goes from local to itinerant as U ~~is decreased across~~ decreases throughout this discontinuity.
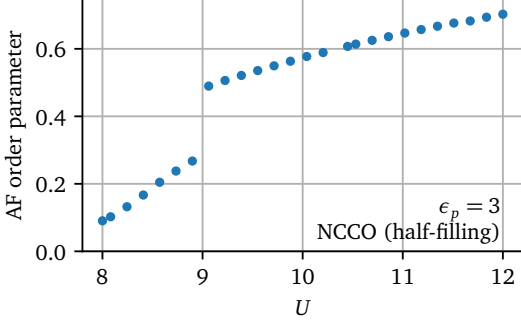
Figure 7. Amplitude of the antiferromagnetic order parameter of NCCO at half-filling in relation to the interaction strength U.
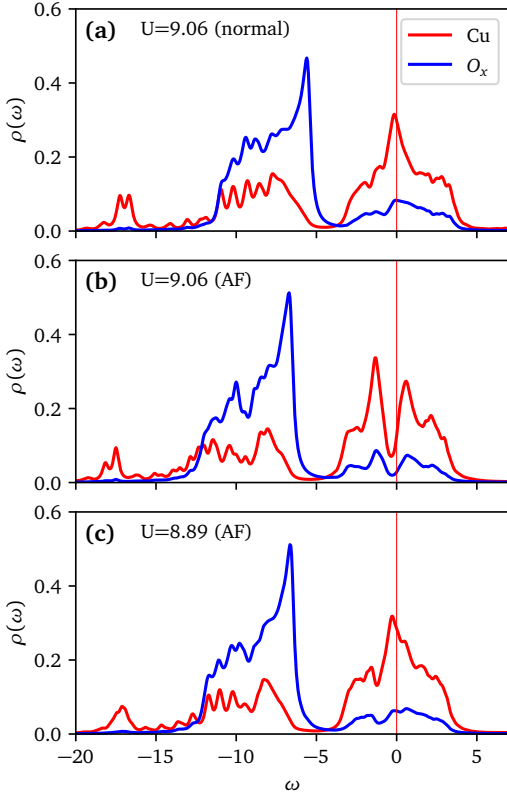
Figure 8. DoS of NCCO in the normal phase in which we allowed for long range antiferromagnetic order. The two values of interaction U shown are the values on the border of the discontinuity in Fig. 7.

IV. DISCUSSION AND CONCLUSION

Comparing the hole distribution data from NMR and CDMFT should in principle allow us to set the interaction strength U and the relative band energy ϵ_p in the Emery model. Knowing the band parameters from first-principles calculations, one can produce curves similar to Fig. 2 and approximately identify the right $U - \epsilon_p$ by looking at the intersection of their curves with the half-filling axis and compare it with a NMR measurement on the undoped compound, ~~like~~ as in Fig. 5. Ideally, (n_d, n_p) NMR data and the CDMFT results would have the same slope away from half-filling. From Fig. 5 one sees that this is not exactly the case, although the agreement for YBCO is satisfactory given the dispersion of NMR data.

In LCO, the slope from the 4 NMR data points is smaller than predicted by CDMFT for $\epsilon_p = 4$. This behavior indicates that the O and Cu orbitals are actually less hybridized than the DFT-derived parameters of the Emery model lead us to believe. There could be many causes. Maybe the value of ϵ_p is actually larger than predicted by DFT. Indeed, our results are closer to the mark if we adopt the more extreme values $U = 20$ and $\epsilon_p = 10$ (open blue circles in Fig. 5): A larger value of ϵ_p lowers the hybridization of the Cu and O bands and produces a flatter curve. Maybe also our neglect of the local Coulomb interactions on the oxygens (U_p) is to blame, as its importance will grow with hole doping. Finally, this may just be an inherent insufficiency of the Emery model to describe the actual material, which only a treatment of all orbitals could resolve, in the line of Ref. [28].

More serious is the disagreement between the data for NCCO (open red squares on Fig. 5) and the model's predictions on the electron-doped side. The NMR data consists of two points only, but these two points hint at a rather large value of dn_d/dn_p, as if the upper Hubbard band contained little oxygen, even though the values of U and ϵ_p are such that the CTG vanishes.

The value $U - \epsilon_p = 5$ for NCCO, as determined by the position of (n_p, n_d) at half-filling, was not large enough to open a CTG even when allowing ~~long range~~ long-range antiferromagnetic order. Figs 7 and 8 show that a CTG opens up at the Fermi level between $U = 8.90$ and $U = 9.06$ for $\epsilon_p = 3$, leading to a jump in the AF order parameter $\langle M \rangle$. Knowing that NCCO is a superconductor, we should observe a non-negligible value of the SC order parameter Ψ at $U - \epsilon_p = 5$ if it is indeed the right value to describe NCCO in the Emery model, but we don't (not shown). We are clearly underestimating the value of $U - \epsilon_p$ in that case. For hole-doped compounds, which are commonly believed to possess greater U values than their electron-doped counterparts [16, 29], this underestimation is less of a problem since they stand further from the transition between a CTI and a metal. As for the electron-doped cuprates, some argue that they are simply not in the CTI regime [16]. This also points towards a possible (small) disconnect between the computation of densities and that of the SC order parameter, the latter being more affected by the finite size of the cluster since it is less of a local quantity: neglecting

inter-cluster components of the anomalous self-energy has more consequences when computing Ψ.

Let us point out some deficiencies of the Emery model or of our treatment thereof:

1. In principle, one could construct a tight-binding model that fits the DFT band structure perfectly, with a ~~large-enough~~ large enough number of orbitals. The Emery model keeps only three of them, hence some hybridization with other bands is lost and this will affect the occupation numbers n_d and n_p, especially if we venture too far from half-filling, the usual operating point at which the model parameters are derived.

2. In the model itself, we have neglected all interactions but the on-site interaction U on copper orbitals, effectively treating other orbitals in the mean-field approximation. Since the occupation n_p of oxygens is close to 2, this is not such a bad approximation (fluctuations of these interactions are small) but it is less and less valid as hole-doping increases.

3. The value of ϵ_p, as insisted upon above, is strongly renormalized by the interaction, so much so that it is better to treat it as an adjustable parameter. However, that renormalization is doping dependent, and this was ignored here.

Despite these shortcomings, the Emery model achieves a lot [20]: it reproduces the correlation between superconductivity and superexchange[30], charge transfer gap [31, 32] and oxygen hole doping [23].

To conclude, studying the relation between oxygen occupation n_p and copper occupation n_d can help determine the degree of correlation $U - \epsilon_p$ in the model and set this parameter combination from observations. This works reasonably well for the hole-doped cuprates LSCO ($U - \epsilon_p = 10$) and YBCO ($U - \epsilon_p = 6$), but less so for the electron-doped cuprate NCCO. In the latter case, $U - \epsilon_p$ may be strongly underestimated, or the weaker correlations make the CDMFT approach used here less accurate.

ACKNOWLEDGMENTS

Fruitful conversations with A.-M. S. Tremblay are gratefully acknowledged. This work has been financially supported by the Natural Sciences and Engineering Research Council of Canada (Grants No. RGPIN-2020-05060) and by the Canada First Research Excellence Fund. Computational resources were provided by the Digital Research Alliance of Canada and Calcul Québec.

Appendix A: Cluster dynamical mean-field theory

In the appendix we review some technical details of cluster dynamical mean field theory (CDMFT) [33, 34] as applied to our problem. General reviews of CDMFT may be found in Refs [19, 35]. We define an impurity problem consisting of 4 unit cells of the Emery model, which we further decompose into a 4-site, correlated cluster of Cu $d_{x^2-y^2}$ orbitals, plus an 8-site, uncorrelated cluster of O p orbitals. The correlated Cu cluster is augmented by an 8-orbital bath that represents that cluster's environment, and defines an "impurity problem" whose Hamiltonian is

$$H_{\text{imp}} = H_{\text{clus}} + H_{\text{bath}} + H_{\text{hyb}}, \tag{A1}$$

with the following partial Hamiltonians:

$$
\begin{aligned}
H_{\text{clus}} &= U \sum_{\mathbf{r}} n_{\mathbf{r}\uparrow} n_{\mathbf{r}\downarrow} - \sum_{\mathbf{r},\mathbf{r}',\sigma} t_{\mathbf{r}\mathbf{r}'} c_{\mathbf{r}\sigma}^\dagger c_{\mathbf{r}\sigma} \\
H_{\text{bath}} &= \sum_{\alpha,\sigma} \varepsilon_\alpha b_{\alpha\sigma}^\dagger b_{\alpha\sigma} \\
H_{\text{hyb}} &= \sum_{\mathbf{r},\alpha,\sigma} \theta_{\mathbf{r}\alpha} (c_{\mathbf{r}\sigma}^\dagger b_{\alpha\sigma} + \text{H.c})
\end{aligned}
\tag{A2}
$$

The index $\mathbf{r}$ stands for site within the cluster and α is the bath orbital index. The $t_{\mathbf{r}\mathbf{r}'}$ are elements of the cluster hopping matrix $\mathbf{t}$, $c_{\mathbf{r}\sigma}$ and $b_{\alpha\sigma}$ are destruction operators on cluster and bath orbitals, respectively, for spin $\sigma = \uparrow, \downarrow$. ε_α is the energy of a bath orbital and ~~the~~ $\theta_{\mathbf{r}\alpha}$ are elements of the hybridization matrix that allows electrons to hop between cluster and bath orbitals. The effect of the bath environment on the electron Green function is contained in the hybridization function:

$$\Gamma_{\mathbf{r}\mathbf{r}'}(\omega) = \sum_\alpha \frac{\theta_{\mathbf{r}\alpha} \theta_{\mathbf{r}'\alpha}^*}{\omega - \varepsilon_\alpha} \tag{A3}$$

We can extract the correlated cluster's self-energy $\mathbf{\Sigma}_c$ from the Cu orbital Green function obtained via ED through Dyson's equation:

$$\mathbf{G}_c^{-1}(\omega) = \omega - \mathbf{t} - \mathbf{\Gamma}(\omega) - \mathbf{\Sigma}_c(\omega). \tag{A4}$$

Here $\mathbf{G}_c$ is the exact 8×8 Green function of the Cu orbitals on the correlated cluster (8 because of spin). The self-energy is solely associated to the Cu cluster since we neglect correlation on the O orbitals. From this self-energy, we construct an approximation to the Green function of the infinite system

$$\mathbf{G}^{-1}(\tilde{\mathbf{k}}, \omega) = \mathbf{G}_0^{-1}(\tilde{\mathbf{k}}, \omega) - \mathbf{\Sigma}_c(\omega). \tag{A5}$$

Where $\mathbf{G}_0^{-1}$ is the non-interacting part of the lattice Green function and $\tilde{\mathbf{k}}$ is the reduced wave vector in the Brillouin zone of the super-lattice of repeated clusters. Here $\mathbf{G}_0^{-1}(\tilde{\mathbf{k}})$ is a 24×24 matrix, since it also contains the contribution of the oxygen orbitals, and $\mathbf{\Sigma}_c$ has been padded with zeros from 8×8 to 24×24 to include the null contribution of the oxygen orbitals to the self-energy. From this expression, we can in turn calculate the local Green function:

$$\bar{\mathbf{G}}(\omega) = \frac{L}{N} \sum_{\tilde{\mathbf{k}}} \frac{1}{\mathbf{G}_0^{-1}(\tilde{\mathbf{k}}) - \mathbf{\Sigma}_c(\omega)}. \tag{A6}$$

The goal of CDMFT is then, through self-consistency, to obtain a local Green function that is equal to the exactly-solved Green function of the cluster: $\mathbf{G}_c(\omega) = \bar{\mathbf{G}}(\omega)$, when restricted to the correlated (Cu) degrees of freedom. With a finite number of bath orbitals, it is impossible for those two functions to coincide perfectly at all frequencies. A distance function is then defined to minimize the difference between the two solutions:

$$d = \sum_{i\omega_n, \nu, \nu'} W_n \left| (\mathbf{G}_c^{-1}(i\omega_n) - \bar{\mathbf{G}}^{-1}(i\omega_n))_{\nu\nu'} \right|^2 \quad \text{(A7)}$$

(again, in the above equation, $\bar{\mathbf{G}}^{-1}$ is restricted to the Cu orbitals only). The sum is taken over a suitable set of Matsubara frequencies at some fictitious temperature $1/\beta$, with weights W_n that are usually taken to be constant up to some cutoff frequency ω_c. The distance function is minimized with the help of some minimization method (here we use the Nelder-Mead algorithm [36]), varying the bath parameters, ε_α and $\theta_{\mu\alpha}$. The newly obtained bath parameters are then used to update the impurity model, and the procedure is iterated until convergence.

Once a self-consistent local Green function has been found, different observables of the system can be computed. Consider for instance a general one-body operator $\hat{S}$:

$$\hat{S} = \sum_{\alpha\beta} s_{\alpha\beta} c_\alpha^\dagger c_\beta \quad . \quad \text{(A8)}$$

The ground state expectation value of such an operator is [19]

$$\bar{S} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{d^2\tilde{\mathbf{k}}}{(2\pi)^2} \left\{ \text{tr}\, \mathbf{s}(\tilde{\mathbf{k}}) \mathbf{G}(\tilde{\mathbf{k}}, \omega) - \frac{\text{tr}\, \mathbf{s}(\tilde{\mathbf{k}})}{i\omega - p} \right\} \quad , \quad \text{(A9)}$$

where convergence is ensured by the subtraction of a pole p located on the positive real axis, if $\text{tr}\, \mathbf{s} \neq 0$. The expression

(A9) is mainly used in this work to compute the electronic (hole) density on Cu and O sites in the unit cell of the model illustrated in Fig. 1.

In this work, we also probe the superconducting state. In that case the impurity model (A2) needs to be upgraded by adding pairing between cluster and bath:

$$H_{\text{anom}} = \sum_{\mathbf{r},\alpha} s_{\mathbf{r}\alpha}(c_{\mathbf{r}\uparrow} b_{\alpha\downarrow} - c_{\mathbf{r}\downarrow} b_{\alpha\uparrow} + \text{H.c}) \quad . \quad \text{(A10)}$$

We then use the Nambu formalism, which here amounts to a particle-hole transformation on the spin-down part of the problem and allows to formulate the problem in the same Green function language as in the normal solution. The Nambu off-diagonal blocks of the Green function and self-energy then contain the anomalous parts of these quantities. A superconducting order parameter with $d_{x^2-y^2}$ symmetry is then defined as $\Psi = \langle \hat{\Delta} \rangle / N$ with N the number of unit cells in the lattice and the singlet pairing operator $\hat{\Delta}$ defined as:

$$\hat{\Delta} = \sum_{\langle \mathbf{rr'} \rangle_x} (d_{\mathbf{r}\uparrow} d_{\mathbf{r'}\downarrow} - d_{\mathbf{r}\downarrow} d_{\mathbf{r'}\uparrow}) - \sum_{\langle \mathbf{rr'} \rangle_y} (d_{\mathbf{r}\uparrow} d_{\mathbf{r'}\downarrow} - d_{\mathbf{r}\downarrow} d_{\mathbf{r'}\uparrow}) + \text{H.c} \,,$$
$$\text{(A11)}$$

with $\langle \mathbf{rr'} \rangle_x$ and $\langle \mathbf{rr'} \rangle_y$ nearest neighbor copper orbitals in the x and y directions, respectively.

Finally, let us point out that the use of an impurity solver based on exact diagonalization (ED), even though limited to small systems sizes, and in particular a discrete bath, has the advantage over quantum Monte Carlo methods that it is not plagued by the fermion sign problem. The latter only gets worse while studying covalent (low values of ϵ_p) compounds compared to ionic compounds. Since cuprates are definitey covalent [15], the use of an ED solver is of great convenience.

[1] F. C. Zhang and T. M. Rice, Effective Hamiltonian for the superconducting Cu oxides, Phys. Rev. B **37**, 3759 (1988).

[2] E. Gull and A. J. Millis, Numerical models come of age, Nature Physics **11**, 808 (2015).

[3] E. Pavarini and I. for Advanced Simulation, eds., *Emergent phenomena in correlated matter: lecture notes of the Autumn School on Correlated Electrons 2013 at Forschungszentrum Jülich, 23 27 September 2013*, Schriften des Forschungszentrums Jülich Reihe Modeling and Simulation No. 3 (Forschungszentrum Jülich, Jülich, 2013) meeting Name: Autumn School Correlated Electrons OCLC: 862816584.

[4] H. Alloul, What is the simplest model that captures the basic experimental facts of the physics of underdoped cuprates?, Comptes Rendus Physique **15**, 519524 (2014).

[5] D. Sénéchal, P. L. Lavertu, M. a. Marois, and A.M.. Tremblay, Competition between antiferromagnetism and superconductivity in high-T_c cuprates, Physical Review Letters **94**, 156404 (2005).

[6] A. Foley, S. Verret, A.-M. S. Tremblay, and D. Sénéchal, Coexistence of superconductivity and antiferromagnetism in the Hubbard model for cuprates, Physical Review B **99**, 184510 (2019).

[7] E. Fradkin, S. A. Kivelson, and J. M. Tranquada, Colloquium: Theory of intertwined orders in high temperature superconductors, Reviews of Modern Physics **87**, 457 (2015).

[8] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, and S. S. C. o. t. M.-E. P. Zhang, Absence of superconductivity in the pure two-dimensional Hubbard model, Physical Review X **10**, 031016 (2020).

[9] V. J. Emery, Theory of high-T_c superconductivity in oxides, Phys. Rev. Lett. **58**, 2794 (1987).

[10] C. Varma, S. Schmitt-Rink, and E. Abrahams, Charge transfer excitations and superconductivity in ionic metals, Solid State Communications **62**, 681 (1987).

[11] M. Jurkutat, D. Rybicki, O. P. Sushkov, G. V. M. Williams, A. Erb, and J. Haase, Distribution of electrons and holes in cuprate superconductors as determined from ^{17}O and

^{63}Cu nuclear magnetic resonance, Phys. Rev. B **90**, 140504 (2014).

[12] M. S. Hybertsen, M. Schlüter, and N. E. Christensen, Calculation of Coulomb-interaction parameters for La_2CuO_4 using a constrained-density-functional approach, Phys. Rev. B **39**, 9028 (1989).

[13] O. Andersen, A. Liechtenstein, O. Jepsen, and F. Paulsen, LDA energy bands, low-energy Hamiltonians, t', t'', $t_{\perp}(\mathbf{k})$, and $J_{\perp}$, Journal of Physics and Chemistry of Solids **56**, 1573 (1995), proceedings of the Conference on Spectroscopies in Novel Superconductors.

[14] J. Zaanen, G. A. Sawatzky, and J. W. Allen, Band gaps and electronic structure of transition-metal compounds, Phys. Rev. Lett. **55**, 418 (1985).

[15] C. Weber, C. Yee, K. Haule, and G. Kotliar, Scaling of the transition temperature of hole-doped cuprate superconductors with the charge-transfer energy, Europhysics Letters **100**, 37001 (2012).

[16] C. Weber, K. Haule, and G. Kotliar, Strength of correlations in electron- and hole-doped cuprates, Nature Physics **6**, 574 (2010).

[17] A. I. Lichtenstein and M. I. Katsnelson, Antiferromagnetism and d-wave superconductivity in cuprates: A cluster dynamical mean-field theory, Phys. Rev. B **62**, R9283 (2000).

[18] G. Kotliar, S. Y. Savrasov, G. Pálsson, and G. Biroli, Cellular dynamical mean field approach to strongly correlated systems, Physical Review Letters **87**, 186401 (2001).

[19] T. N. Dionne, A. Foley, M. Rousseau, and D. Sénéchal, Pyqcm: An open-source Python library for quantum cluster methods, SciPost Phys. Codebases , 23 (2023).

[20] N. Kowalski, S. S. Dash, P. Sémon, D. Sénéchal, and A.-M. Tremblay, Oxygen hole content, charge-transfer gap, covalency, and cuprate superconductivity, Proceedings of the National Academy of Sciences **118**, e2106476118 (2021).

[21] B. Bacq-Labreuil, B. Lacasse, A.-M. S. Tremblay, D. Sénéchal, and K. Haule, Towards an ab initio theory of high-temperature superconductors: a study of multilayer cuprates (2024), arXiv:2410.10019 [cond-mat.str-el].

[22] J. M. Tranquada, S. M. Heald, and A. R. Moodenbaugh, X-ray-absorption near-edge-structure study of $La_{2-x}(Ba, Sr_x)CuO_{4-y}$ superconductors, Phys. Rev. B **36**, 5263 (1987).

[23] D. Rybicki, M. Jurkutat, S. Reichardt, C. Kapusta, and J. Haase, Perspective on the phase diagram of cuprate high-temperature superconductors, Nature Communications **7**, 11413 (2016).

[24] G. Sordi, G. L. Reaney, N. Kowalski, P. Sémon, and A.-M. S. Tremblay, Ambipolar doping of a charge-transfer insulator in the Emery model, Physical Review B **111**, 045117 (2025).

[25] N. Kowalski, *Dopage, température critique et étude du modèle de Hubbard à trois bandes*, Ph.D. thesis, Université de Sherbrooke (2021).

[26] B. Kyung, D. Sénéchal, and A.-M. S. Tremblay, Pairing dynamics in strongly correlated superconductivity, Physical Review B **80**, 205109 (2009).

[27] S. S. Dash and D. Sénéchal, Pseudogap transition within the superconducting phase in the three-band Hubbard model, Physical Review B **100**, 214509 (2019).

[28] B. Bacq-Labreuil, B. Lacasse, A.-M. S. Tremblay, D. Sénéchal, and K. Haule, Towards an ab initio theory of high-temperature superconductors: a study of multilayer cuprates, arXiv preprint arXiv:2410.10019 (2024).

[29] D. Sénéchal and AM. Tremblay, Hot spots and pseudogaps for hole- and electron-doped high-temperature superconductors, Physical Review Letters **92**, 126401 (2004).

[30] L. Wang, G. He, Z. Yang, M. Garcia-Fernandez, A. Nag, K. Zhou, M. Minola, M. L. Tacon, B. Keimer, Y. Peng, *et al.*, Paramagnons and high-temperature superconductivity in a model family of cuprates, Nature Communications **13**, 3163 (2022).

[31] S. M. OMahony, W. Ren, W. Chen, Y. X. Chong, X. Liu, H. Eisaki, S.-i. Uchida, M. Hamidian, and J. S. Davis, On the electron pairing mechanism of copper-oxide high temperature superconductivity, Proceedings of the National Academy of Sciences **119**, e2207449119 (2022).

[32] W. Ruan, C. Hu, J. Zhao, P. Cai, Y. Peng, C. Ye, R. Yu, X. Li, Z. Hao, C. Jin, *et al.*, Relationship between the parent charge transfer gap and maximum transition temperature in cuprates, Science bulletin **61**, 1826 (2016).

[33] A. I. Lichtenstein and M. I. Katsnelson, Antiferromagnetism and d-wave superconductivity in cuprates: A cluster dynamical mean-field theory, Phys. Rev. B **62**, R9283 (2000).

[34] G. Kotliar, S. Y. Savrasov, G. Pálsson, and G. Biroli, Cellular Dynamical Mean Field Approach to Strongly Correlated Systems, Phys. Rev. Lett. **87**, 186401 (2001).

[35] D. Sénéchal, Quantum cluster methods: CPT and CDMFT, in *Many-Body Physics: From Kondo to Hubbard*, Vol. 5, edited by E. Pavarini, E. Koch, and P. Coleman (Forschungszentrum Jülich, 2015) pp. 13.1–13.22.

[36] J. A. Nelder and R. Mead, A Simplex Method for Function Minimization, The Computer Journal **7**, 308 (1965).