Modeling Polarons in Transition-Metal Oxides

by

Bayo Lau

B.A.Sc, University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES
(Physics)

The University Of British Columbia
(Vancouver)

April 2011

© Bayo Lau, 2011
Abstract

This work is a series of reports on progress in the description of electron-electron and electron-lattice interactions in transition-metal oxides, with an emphasis to the class of high-$T_C$ superconducting cuprates. Combinations of numerical and analytical approaches were devised and developed to study large-scale models, which distinguish cation and anion sites of the realistic lattice structure. The many results range from incremental deviation to significant difference compared to those of the widely-accepted simple models without such distinction.

A previously proposed numerical scheme and a perturbation approach were adapted to study the one-dimensional breathing-mode model, which describes a charge carrier interacting with vibrating anions. The effort yielded the first accurate benchmark result for all parameters values of the model.

Based on a physical insight about the spin-1/2 Heisenberg antiferromagnet on a two-dimensional square lattice, an octapartite approach was devised to model the low-energy states. The efficiency of the resulting numerical approach was showcased with the explicit solution to a record-breaking 64-spin torus.

A spin-polaron model was derived to model holes injected into cuprate’s quasi two-dimensional copper-oxygen layer. Total-spin-resolved exact diagonalization was performed for a single fermionic hole in a record-breaking cluster with 32 copper and 64 oxygen sites. The solutions revealed important physics missed by previous studies.

The octapartite approach was then developed to solve the spin-polaron model with two extra holes in the same cluster. The accuracy and efficiency of the method were established. Enhanced singlet correlation between two holes was observed. The preliminary results justify the need for an in-depth study.
Preface

A version of Chapter 2 has been published as B. Lau, M. Berciu, and G. A. Sawatzky. “Single-polaron properties of the one-dimensional breathing-mode Hamiltonian.” Physical Review B 76 174305 (2007). The effort was initiated by Prof. M. Berciu and Prof. G. A. Sawatzky. I was responsible for all computational and theoretical work. Further physical interpretations and the writing of the manuscript were carried out with suggestions from the co-authors.

A version of Chapter 3 has been published as B. Lau, M. Berciu, and G. A. Sawatzky. “Computational scheme for the spin-\(\frac{1}{2}\) Heisenberg antiferromagnet based on an octapartite description of the square lattice.” Physical Review B 81 172401 (2010). I was responsible for the initial effort as well as all computational and theoretical work. Further physical interpretations and the writing of the manuscript were carried out with suggestions from the co-authors.

A version of Chapter 4 has been published as B, Lau, M. Berciu, and G. A. Sawatzky. “High-Spin Polaron in Lightly Doped CuO\(_2\) Planes.” Physical Review Letter 106 036401 (2011). I was responsible for the initial effort as well as all computational and theoretical work. Further physical interpretations and the writing of the manuscript were carried out with suggestions from the co-authors.

A manuscript for a refereed journal is being prepared from the results in Chapter 5. I was responsible for the initial effort as well as all computational and theoretical work. M. Berciu and G. A. Sawatzky have been involved in the interpretation of the results.
# Table of Contents

Abstract ............................................................... ii
Preface ................................................................. iii
Table of Contents ..................................................... iv
List of Tables .......................................................... vii
List of Figures .......................................................... viii
Glossary ............................................................... xiii
Acknowledgments ......................................................... xiv

1 Introduction ............................................................ 1
   1.1 Hole-Doped Cuprates ........................................ 3
   1.2 The t-J Model .................................................... 8
   1.3 Towards Realistic Oxide Structures ......................... 11
   1.4 Scope .......................................................... 12

2 Lattice Polaron on the Oxide Chain ............................... 15
   2.1 Introduction ..................................................... 15
   2.2 The Holstein and Breathing-Mode Models .................. 16
   2.3 Methodology .................................................... 18
       2.3.1 Strong-Coupling Perturbation Results ............... 18
       2.3.2 Matrix Computation .................................... 21
       2.3.3 Green’s Function Computation ...................... 23
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>Low-Energy Two-Hole Solutions</td>
<td>75</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Energetics</td>
<td>76</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Wavefunction Analysis</td>
<td>78</td>
</tr>
<tr>
<td>5.5</td>
<td>Speculations</td>
<td>85</td>
</tr>
<tr>
<td>5.6</td>
<td>Conclusions</td>
<td>90</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions</td>
<td>92</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary</td>
<td>92</td>
</tr>
<tr>
<td>6.2</td>
<td>Opportunities for Further Developments</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Appendices</td>
<td>103</td>
</tr>
<tr>
<td>A</td>
<td>Software Integrity Checks</td>
<td>103</td>
</tr>
<tr>
<td>B</td>
<td>Computation Development</td>
<td>104</td>
</tr>
<tr>
<td>B.1</td>
<td>Software Optimizations</td>
<td>104</td>
</tr>
<tr>
<td>B.2</td>
<td>Details of the Octapartite Approach</td>
<td>105</td>
</tr>
<tr>
<td>B.2.1</td>
<td>General Observations</td>
<td>106</td>
</tr>
<tr>
<td>B.2.2</td>
<td>Implementations</td>
<td>108</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1 Types of electronic models for the CuO$_2$ plane, their assumptions, and the Fork space of n extra holes injected into a half-filled system with N unit cells. ....................... 7

Table 2.1 Most-relaxed cut-off condition for the 1D breathing-mode polaron Green’s function computation. ......................... 23

Table 2.2 $\alpha_{n_-, n_+}$ vs. $n_+, n_-$ for the ground state ......................... 29

Table 2.3 $\alpha_{n_-, n_+}$ vs. $n_+, n_-$ for the first-excited state ......................... 29

Table 3.1 Size of the $C_\frac{1}{4}$ subspace as compared to that of the commonly used basis. ......................... 47

Table 5.1 (N=32) The $S=\frac{1}{2}$ $k = (\frac{\pi}{2}, \frac{\pi}{2})$ single-hole ground state’s weight in subspaces of particular $S_A \otimes S_B$. Numbers are percentage adding up to 100%. Two lines are drawn to highlight the $C_S = \frac{1}{4}$ truncation which discards states with $S_A, S_B < 6$ and yield an energy within 3.6% of the exact value. ......................... 66

Table 5.2 Single-hole eigenstates of $H_{pd}$. $p_\sigma^\dagger$ creates an oxygen hole and the arrows in the ket indicate the spins of the two copper sites neighboring the oxygen hole. ......................... 81
List of Figures

Figure 1.1 Phase diagram of hole-doped cuprates (Temperature vs electron-removal per unit-cell)[3] .................................................. 3

Figure 1.2 (top) A CuO$_2$ layer embedded in certain chemical host[4]. (bottom) Lattice structure of single-layer cuprate HgBa$_2$CuO$_4$[12]. ... 5

Figure 1.3 Two adjacent unit cells of the CuO$_2$ plane. The orbitals kept in the 3-band model of Eq. (1.1) are shown, with white/shaded for positive/negative signs. The two $\varepsilon$ vectors (solid arrow) and the four $\delta$ vectors (dashed arrow) are also shown. ........ 6

Figure 1.4 Mean-field phase diagram for the t-J model[5]. ................. 10

Figure 2.1 (a) GS energy and (b) GS $qp$ weight as a function of the corresponding dimensionless coupling parameter. The dashed line corresponds to the Holstein model, while the breathing-mode results are shown by circles (line is a guide to the eye). Parameters are $t = 2, \Omega = 1$. .......................................................... 25

Figure 2.2 (a) Energy of the first excited $K = 0$ state, measured from $E_{GS}$, and (b) its $qp$ weight. The dashed line corresponds to the Holstein model, while the breathing-mode results are shown by circles (line is a guide to the eye). Parameters are $t = 2, \Omega = 1$. ... 26

Figure 2.3 P(0) for $t = 2, \Omega = 1$ for the breathing-mode (symbols) and Holstein (line) models, respectively. The solid and dashed lines correspond to ground state and first excited state, respectively. ................... 27
Figure 2.4  The $k$-dependent (a) energy and (b) $qp$ weight of the GS (circles) and first bound state (square) for the 1D breathing mode model, for $t = 2, \Omega = 1, \lambda_B = 1.07$.  

Figure 2.5  Ratio of effective polaron mass to that of the free electron. Circles and squares show breathing-mode results for GS and first bound state, respectively. The other lines correspond to the Holstein model GS (full) and second bound state (dashed). Parameters are $t = 2, \Omega = 1$.  

Figure 2.6  GS $qp$ weights for the breathing-mode (circles) and Holstein (squares) models. Full symbols correspond to $\alpha = 8$. For comparison purposes, the empty symbols show the $\alpha = 4$ results of Fig. 2.1(b). ($\Omega = 1$).  

Figure 2.7  The spectral function of the 1D breathing-mode model, on linear (left) and logarithmic (right) scales. The curves are shifted for better viewing, and correspond (top to bottom) to $\lambda_B = 1.5; 1.25; 1.0$ and $0.75$. Vertical lines indicates $E_{GS}$ and $E_{GS} + \Omega$. $K = 0, t = 2, \Omega = 1, \eta = 0.004\Omega$.  

Figure 2.8  $A(k, \omega)$ vs $\omega$, for $K/\pi = 0; 0.25; 0.5; 0.75$ and $1$ (top to bottom) for intermediate coupling $t = 2, \Omega = 1, \lambda_B = 1.25$. $\eta/4 = \Delta E = 0.001\Omega$. The height of the spectral weight is plotted in linear scale on the left and logarithmic scale on the right. The two vertical lines indicate $E_{GS}$ and $E_{GS} + \Omega$.  

Figure 3.1  Square lattice divided in octads (shaded areas). Sites positioned similarly inside octads have the same label. Each is surrounded by neighbors with different labels.  

Figure 3.2  (a) Overlap $|\langle GS, C_s | GS, C_s + \frac{\lambda}{N} \rangle|^2$ between ground-state wave functions corresponding to adjacent values of $C_s$ for $N = 16, 32, 64$. Note that for $N = 64$, the overlap is already 95% for $C_s = \frac{1}{4}$; (b) Fractional change $\frac{1}{E_{GS}} \frac{dE_{GS}}{dC_s}$ in ground-state energy corresponding to different $C_s$ cutoffs. The lines are linear fits to the data.
Figure 3.3  GS energy computed at specific $C_s$ values (symbols) and the extrapolation to $C_s = 1$ (solid lines) for $N = 16, 32, 64$. The horizontal dashed lines are the lowest values from Ref. [37] inset: expectation values of the non-commuting operators $\sqrt{\langle m^2 \rangle}$ and $\langle \vec{S}_r \cdot \vec{S}_{r+\left(\frac{1}{2}, \frac{1}{2}\right)} \rangle$.  

49

Figure 4.1 (a) Two adjacent unit cells of the CuO$_2$ plane. The orbitals kept in the 3-band model of Eq. (4.1) are shown, with white/shaded for positive/negative signs. The two $\varepsilon$ vectors (solid arrow) and the four $\delta$ vectors (dashed arrow) are also shown. (b) Sketch of a virtual process of $T_{\text{swap}}$.  

54

Figure 4.2 a) Energy and b) quasiparticle weight (bottom) for the lowest eigenstates with $S_T = \frac{1}{2}$ and $\frac{3}{2}$ vs. momentum. Different sets are shifted so as to have the same GS energy.  

58

Figure 4.3 $\langle C_x(\delta, \Delta) \rangle$ for the lowest energy state at (a) $(\frac{\pi}{2}, \frac{\pi}{2})$ with $S_T = \frac{1}{2}$, and (b) at $(\pi, \pi)$ with $S_T = \frac{3}{2}$. The darkly-shaded bullet denotes the oxygen position at $l + e_x$. Each bullet shows the correlation value between the two sandwiching Cu sites. The central 12 Cu sites are shown; the correlations between the other 20 Cu spins converge fast towards the AFM value of $\sim$0.33. $\langle C_x(\delta, a) \rangle$ is the $\vec{P}_{x+y}$ reflection.  

60

Figure 4.4 $\langle C_{x/y}(\delta, a) \rangle$ for lowest state at $k=(0, \pi)$ with $S_T=1/2$.  

62

Figure 5.1 Convergence of the one-hole ground state. (left) GS energy calculated for increasing $C_S$. The value approaches rapidly to the exact value marked by the dotted line. (right) fractional change in GS energy for the next increment of $C_s$. Solid lines are extrapolation and linear fits.  

66

Figure 5.2 Lowest two-hole states in the $C_S = \frac{1}{4}$ subspace.  

76

Figure 5.3 Convergence of the $S_T = 0$ lowest energy states. The two-hole energy level corresponding to $\Delta E_b = 0$ is indicated by the horizontal dash line.  

77
Figure 5.4  (top) Probability of charge separation, \(\langle C_{\text{charge}}(R)\rangle\) for the lowest states in the \(C_S = \frac{1}{2}\) subspace. (bottom) The difference from a random distribution. ................................................................. 80

Figure 5.5  (top) Probability of singlet three-spin polaron (3SP) pair, \(\langle C_{3\text{SP},\text{singlet}}(R)\rangle\) for the lowest states in the \(C_S = \frac{1}{2}\) subspace. (bottom) The difference from a randomly distributed paramagnetic configuration. 83

Figure 5.6  \(c_{3\text{SP},\text{singlet}}(\delta) \times 10^4\) for one of the two \(K = (\pi, \pi)\) GS in the \(C_S = \frac{1}{2}\) subspace. Cu and O are marked by circles and rectangles, respectively. Values for the other state are obtained by 90° rotation. (top, black→red) 28 x-y, (bottom, black→black) 16 x-x and (bottom, red→red) 16 y-y 3SP pair configurations are shown; 6 O-O configurations, which involve the two O holes sharing the same copper site, have low probability and are omitted. For example, the bottom figure shows that \(c_{3\text{SP},\text{singlet}}(2a,0) = 0.0109\) for two 3SPs both centered on x-rung O sites. .................. 84

Figure 5.7  \(c_{3\text{SP},\text{singlet}}(\delta) \times 10^4\) for the lowest \(K = (\pi,0)\) state in the \(C_S = \frac{1}{2}\) subspace. Cu and O are marked by circles and rectangles, respectively. Values for the \(K = (0, \pi)\) state are obtained by 90° rotation. (top, black→red) 28 x-y, (bottom, black→black) 16 x-x and (bottom, red→red) 16 y-y 3SP pair configurations are shown; 6 O-O configurations, which involve the two O holes sharing the same copper site, have low probability and are omitted. For example, the bottom figure shows that \(c_{3\text{SP},\text{singlet}}(2a,0) = 0.0265\) for two 3SPs both centered on x-rung O sites. .................. 86

Figure 5.8  \(c_{\text{charge}}(\delta) \times 10^4\) for the lowest \(K = (\pi,0)\) state in the \(C_S = \frac{1}{2}\) subspace. Cu and O are marked by circles and rectangles, respectively. Values for the \(K = (0, \pi)\) state are obtained by 90° rotation. (top, black→red) 32 x-y, (bottom, black→black) 17 x-x and (bottom, red→red) 17 y-y 3SP pair configurations are shown. For example, the bottom figure shows that \(c_{\text{charge}}(2a,0) = 0.0289\) for two 3SPs both centered on x-rung O sites. .................. 87
Figure 5.9 \(c_{3SP,singlet}(\delta) \times 10^4\) for the lowest \(K = (0,0)\) \(S_T = 0\) state in the \(C_S = \frac{1}{2}\) subspace. Cu and O are marked by circles and rectangles, respectively. (top, black→red) 28 x-y, (bottom, black→black) 16 x-x and (bottom,red→red) 16 y-y 3SP pair configurations are shown; 6 O-O configurations, which involve the two O holes sharing the same copper site, have low probability and are omitted. For example, the bottom figure shows that \(c_{3SP,singlet}(0,-2a) = 0.0367\) for two 3SPs both centered on x-rung O sites.
## Glossary

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZRS</td>
<td>Zhang-Rice singlet</td>
</tr>
<tr>
<td>AFM</td>
<td>antiferromagnet</td>
</tr>
<tr>
<td>ARPES</td>
<td>Angular-resolved photoemission spectroscopy</td>
</tr>
<tr>
<td>RVB</td>
<td>resonating valence bond</td>
</tr>
<tr>
<td>BCS</td>
<td>Bardeen, Cooper, and Schrieffer</td>
</tr>
<tr>
<td>3SP</td>
<td>three-spin polaron</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to thank my supervisor, George. A. Sawatzky, who guided my growth as a scientist and a problem solver, particularly with many messages which seem trivial at first but proved to be very insightful and stimulating. I am grateful for him always making himself available, looking past my blunders, and most importantly, showing tremendous sympathy and support during the hardest time.

I would also like to thank my PhD committee for making themselves available. In particular, M. Berciu has played a far greater role than required, as is evident in her co-authorship in my publications. I would like to thank her for the many insights and for going out of her way to support me during hardships. Before my graduate studies, A. Damascelli mentioned “One thing [I] can’t do is to say ‘It’s too hard!’ and give up.”, which has since become a motto of my life.

While our research never overlapped, I. Elfimov has taught me many peculiar aspects about scientific computation and has also spent many extra hours in supporting the computing platform I used. M. Choptuik, M. Feeley, C. Greif, O. Toader, and A. Wagner have also provided computational advices for my work.

The foundation for my research was laid by the excellent Canadian education system. I would like to acknowledge Kitsilano Secondary School’s teachers, Mr. Sarna and Ms. Kolson, without whom I would have never received a higher education. I would also like to praise the UBC Engineering Physics program, which provided me the multidisciplinary training that made this work possible.

Last but not least, I would like to thank my family who has been supportive. My parents and my brother have always been there, praying for me, during the ups and the downs. My other half Kimberley has been undyingly supportive and has also tolerated many missed moments such as birthdays and valentines.
Chapter 1

Introduction

When the temperature drops below the melting point of a material, atoms break the continuous spatial symmetries of the gaseous and liquid phase and condense into an energetically favorable solid state. Electrons are more than a thousand times lighter than the nuclei which form the lattice, so the electronic structure is a prominent characteristic. The most primitive theoretical description starts with the scenario of an electron placed in a periodic lattice potential with the potential of the inner core electrons screening part of the bare nuclear potential. For many materials such as normal metals and band insulators, a good description can be obtained by correcting, mostly by perturbation theory, the independent-particle picture with the effects of electron-electron and electron-lattice interactions\[1, 2\]; however, this way of treating many-body correlation cannot explain the properties of all materials, and the community has put forth continuous effort in the understanding of the anomalous cases such as high-temperature superconductors\[3–7\] and colossal magneto resistors\[8\].

The single-particle band theory predicts that the $E \sim k^2$ energy dispersion of an electron in free space is broken into energy bands with different E-vs-k dispersion periodic in the Brillouin zone. The full many-body quantum state is then constructed by assigning the many electrons into these bands of single-particle levels according to Fermi statistics which forbids multiple electrons to have the same set of quantum numbers. If all bands are either empty or completely occupied, the material is predicted to be an insulator because electrons must excite to higher
energy band to acquire net momentum for charge flow and electrical conduction. If the valence band is only partially occupied, its electrons can acquire momentum by exciting to an unoccupied level in the same band with only an infinitesimal cost of energy. These low-lying unoccupied levels provide a channel for scattering due to electron-electron interaction, resulting in states different from the single-particle states in the band theory. The Fermi liquid theory considers such scattering and predicts a one-to-one correspondence between the original single electron state and the “quasiparticles” formed as a result of electron-electron interactions; that is, the single-particle description remains intact, but the states near the Fermi level are quasiparticles, or dressed electrons since they are accompanied by a polarization cloud, instead of free electrons. The Fermi liquid theory predicts such systems with strong electron-electron interaction to be metallic with a resistivity that scales as \( \sim T^2 \) at low temperature. The electrons can also interact with phonons, quantized modes of lattice vibrations, which induces an effective attraction among the quasiparticles near the Fermi level. This introduces an instability against the Fermi sea, leading to a superconducting groundstate, from which excitation of certain fermion pairs of momenta \( \pm k \) is characterized by a Bardeen, Cooper, and Schrieffer (BCS) energy gap \( \Delta(k) = \Delta_0 \) which is constant.

The preceding picture is taught in standard condensed matter courses, but such description of electronic structure is not universal. One of the most striking counterexamples is the class of cuprate compounds, whose properties contradict all of the above statements[4]. The undoped “parent” cuprate compound is an insulator with an odd number of electrons per unit cell. This is not the insulating scenario predicted by band theory, which requires an even number of electrons to fill the up- and down-spin states of a full band for insulation. Removing about one electron per five unit cells creates the “optimally doped” compounds, which exhibit a superconducting transition temperature \( T_C \). Above \( T_C \), the materials were found to have a resistivity that scales as \( \sim T \) instead of \( \sim T^2 \) as predicted by the Fermi-liquid theory. Below \( T_C \), the system superconducts with a k-dependent BCS gap \( \Delta(k) \) of \( d_{x^2-y^2} \) symmetry, which is different from the s-symmetric constant predicted due to electron-phonon interaction in conventional theory. For the “underdoped” compounds with electron concentration between that of the “parent” and “optimally doped” compounds, an extra excitation energy scale \( \Delta_{PG} > \Delta_{SC} \) has been measured
by various experiments. This is another deviation from the phonon-driven superconductor theory, which lacks an extra higher energy scale separating the normal and superconducting state.

### 1.1 Hole-Doped Cuprates

Since their discovery in 1986 [9], cuprates have been classified as a high-$T_C$ superconductor, but their importance to condensed matter theory lies in the many anomalous properties, including the aforementioned contradictions to the conventional theory, when the material is tuned away from the superconducting phase. The class of hole-doped cuprates, which allows electron removal from the parent compound, is of particular interest because of the existence of pseudogap phenomena, in addition to antiferromagnetism, superconductivity, Fermi liquid and non-Fermi liquid physics in different regions of the phase diagram as shown in Fig. 1.1. The undoped antiferromagnet (AFM) insulating phase is well understood in terms of the superexchange physics discussed below. The relatively low $\sim$300K magnetic ordering temperature compared to a 3D antiferromagnet is due to the weak coupling
between the strongly correlated 2D AFM planes. The superconducting phase centered at ~0.15% doping is characterized by a BCS gap $\Delta(k)$ of d-symmetry, which is zero for $k_x = k_y$. Between the AFM phase and superconducting phase lies the region of pseudogap phenomena, which is a cross-over region instead of a definite phase. Spin-sensitive measurements showed a gapped excitation behavior but in-plane conductivity showed no gapped behavior. Angular-resolved photoemission spectroscopy (ARPES) measurements show excitations $\Delta_{PG}(k)$ of the same trend as $\Delta(k)$ with maximum at $(0, \pi)$ and zero for $k_x = k_y$[6]. Breaking of spatial symmetries have also been observed[10]. The connection between different regions has not been elucidated[11].

The mysteries stem from cuprates’ unique structure, which differs greatly from the three-dimensional isotropic lattice assumed by conventional theories. The common feature among cuprates is the copper-oxygen layer embedded in different chemical hosts as shown in Fig. 1.2. One of the simplest examples is the HgBa$_2$CuO$_4$ compound with tetragonal crystal structure and one copper-oxygen layer per unit cell[12]. Although in general $T_C$ is higher for complicated materials with multiple copper-oxygen planes and chains per unit cell, a full understanding of a single layer has not been achieved.

While some compounds, such as La$_{1-x}$Sr$_x$CaO$_4$, exhibit lattice buckling at low temperature[4], HgBa$_2$CuO$_4$ does not exhibit lattice distortion[12]. Lattice distortion is not a universal property, and the layer structure as shown in Fig. 1.2 is believed to be an adequate representation of the system. For the undoped case, each Cu$^{2+}$ ([Ar]d$^9$) is surrounded by an octahedron of six O$^{2-}$ ([Ne]2s$^2$2p$^6$). Four “in-plane” oxygens are linked to other Cu$^{2+}$’s. The other two apical out-of-plane oxygens are further away from the copper than the in-plane ones are. Crystal field theory predicts the splitting of the ten 3d states into four energy levels with different spatial symmetries: $E(d_{xy}) < E(d_{xz}) = E(d_{yz}) < E(d_{x^2-y^2})$[4]. The nine d-electrons thus fully fill all but the $d_{x^2-y^2}$ orbital. With one electron (or hole) per $d_{x^2-y^2}$ orbital, this configuration is the so-called half-filled limit because the other four d-orbitals are commonly treated as inert.

Doping is achieved by changing the chemical host outside of the CuO$_2$ plane shown in Fig. 1.2. An electron removal does not simply introduce a hole in the $d_{x^2-y^2}$ band because the oxygen 2p level is closer to the chemical potential than the
Figure 1.2: (top) A CuO$_2$ layer embedded in certain chemical host[4], (bottom) Lattice structure of single-layer cuprate HgBa$_2$CuO$_4$[12].
copper $d_{x^2-y^2}$ orbital. The material is in the charge-transfer regime in the Zaanen-Sawatzky-Allen classification scheme, and a hole would have strong character of the oxygen ligands\[13\]. Although there are three orthogonal orientations per 2p oxygen level, the $d_{x^2-y^2}$ wavefunction overlaps predominantly with the $\sigma$ orientation of the four surrounding in-plane oxygens.

The most widely studied model with one copper and two oxygen sites per unit cell is the $p-d$ model whose unit cell consists of one $d_{x^2-y^2}$ and two $2p_\sigma$ orbitals\[14, 15\]. For copper sites located at $l$, some integral multiples of the lattice parameters $a_x = (a,0)$ and $a_y = (0,a)$, the oxygen sites are located at $l + \varepsilon_{x/y}$, where $\varepsilon_{x/y} = \frac{1}{2}a_{x/y}$. Since there are more electrons than vacancies in the problem, the model is specified with fermion operators $p_{l+\varepsilon,\sigma}^\dagger$ and $d_{l,\sigma}^\dagger$ creating holes of spin $\sigma$ on a vacuum with one hole per unit cell. For the orbital setting illustrated in Fig.1.3, the model reads

$$H_{3B} = T_{pd} + T_{pp} + \Delta_{pd} \sum n_{l+\varepsilon,\sigma} + U_{pp} \sum n_{l+\varepsilon,\uparrow} n_{l+\varepsilon,\downarrow} + U_{dd} \sum n_{l,\uparrow} n_{l,\downarrow},$$

\textit{Figure 1.3:} Two adjacent unit cells of the CuO$_2$ plane. The orbitals kept in the 3-band model of Eq. (1.1) are shown, with white/shaded for positive/negative signs. The two $\varepsilon$ vectors (solid arrow) and the four $\delta$ vectors (dashed arrow) are also shown.
Table 1.1: Types of electronic models for the CuO$_2$ plane, their assumptions, and the Fork space of n extra holes injected into a half-filled system with N unit cells.

<table>
<thead>
<tr>
<th>Model</th>
<th>Approximations</th>
<th>Number of Fermions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-band</td>
<td>$d_{x^2-y^2}$ on Cu and two $p_{\sigma}$ bands</td>
<td>$N+n$</td>
</tr>
<tr>
<td>2-band</td>
<td>$d_{x^2-y^2}$ and one $p_{\sigma}$ band</td>
<td>$N+n$</td>
</tr>
<tr>
<td>Ch. 3-5</td>
<td>$\langle n_{d^\uparrow} + n_{d^\downarrow} \rangle = 1$ on each Cu and two $p_{\sigma}$ bands</td>
<td>$N+n$</td>
</tr>
<tr>
<td>1-band</td>
<td>$d_{x^2-y^2}$ band</td>
<td>$N+n$</td>
</tr>
<tr>
<td>t-J</td>
<td>each $p_{\sigma}$ hole is locked to a Cu spin as a ZRS</td>
<td>$N-n$</td>
</tr>
</tbody>
</table>

with

\[
\begin{align*}
T_{pd} &= t_{pd} \sum (\delta p_{l+\epsilon,\sigma}^\dagger + p_{l-\epsilon,\sigma}^\dagger) d_{l,\sigma} + h.c. \\
T_{pp} &= t_{pp} \sum \delta \delta p_{l+\epsilon+\delta,\sigma}^\dagger p_{l+\epsilon,\sigma} - t_{pp'} \sum p_{l-\epsilon,\sigma}^\dagger + p_{l+3\epsilon,\sigma} p_{l+\epsilon,\sigma} \\
 n_{l,\sigma} &= d_{l,\sigma}^\dagger d_{l,\sigma} \\
 n_{l+\epsilon,\sigma} &= p_{l+\epsilon,\sigma}^\dagger p_{l+\epsilon,\sigma}.
\end{align*}
\]

$T_{pd}$ is the direct oxygen-copper hopping. $T_{pp}$ contains the direct $t_{pp}$ oxygen-oxygen nearest-neighbor hopping as well as the $t_{pp'}$ oxygen-oxygen next-nearest-neighbor hopping mediated by the copper 4s orbitals. The sign of the matrix elements are determined by the phases of the overlapping orbital wavefunctions in Fig.1.3a. In particular, for upper-right/lower-left (upper-left/lower-right) O-O hopping, $s_\delta = 1$ ($s_\delta = -1$). $\Delta_{pd}$ is the charge transfer energy, which is the energy cost of the $d^9 \rightarrow d^{10}L$ Ligand hole excitation. $U_{dd/pp}$ are repulsive on-site Hubbard interactions, neglecting interaction of longer range. The magnitude of parameters have been established to be $t_{pp} < t_{pd} < \Delta_{pd} < U_{pp} < U_{dd}[5]$. In practice, neighbor Coulomb interactions such as $U_{pd} d^\dagger d p^\dagger p$ should be considered; however, such considerations do not add any new matrix elements into the effective Hamiltonian of interest (Sec. 4.2.1).

While \textit{ab initio} density functional theory made important contributions in determining physical parameters, its applicability is based on groundstate energy and electron density, and the description of strongly correlated many-body physics would require models such as the aforementioned three-band model[4]. In this
context, microscopic modeling has been carried out using models with one \([5–7, 16–19]\), two \([20–28]\), three \([14, 15]\), or more \([29–32]\) bands. Table 1.1 gives a comparison of the models. While exact analytical solutions seem to be out of reach, numerical studies are also carried out with compromises such as the use of small clusters or variational approaches, classical Néel state plus spin-waves, etc \([33–35]\). Given these difficulties and the drive to find the simplest model, the t-J model is unsurprisingly the most studied \([5–7]\).

### 1.2 The t-J Model

Recently, the two-dimensional t-J model has emerged as the most well understood microscopic model\([5–7]\). This route is based on the super-exchange physics and the formation of Zhang-Rice singlet (ZRS) upon electron removal.

In the half-filled limit with no hopping, the groundstate of Eq. (1.1) has one hole per copper orbital. When hopping increases from zero, a pair of neighboring copper holes with opposite spins can hop onto the oxygen orbital shared by both sites, but such process is blocked by the Pauli principle if the two holes have the same spin. The net effect leads to superexchange\([36]\), described by the Heisenberg quantum AFM coupling spins of neighboring copper sites of a 2D square lattice:

\[
H_J = J \sum_{\langle i,j \rangle} \left[ S_i^+ S_j^- + \frac{1}{2} \left( S_i^+ S_j^- + S_i^- S_j^+ \right) \right] = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j \tag{1.2}
\]

Although the 2D Hamiltonian has not been solved exactly, intensive studies over the years have shown that the undoped AFM insulating phase in Fig. 1.1 is well-described by the super-exchange mechanism\([37]\).

The ZRS mechanism was proposed to extend the undoped AFM description away from half filling. Although the lowest energy electron removal happens at the oxygen orbitals, Zhang and Rice pointed out that a specific linear combination of single oxygen hole states can form a singlet, the ZRS, with a copper hole to significantly lower local energy\([16]\). The injection of an extra hole can be effectively described by removing a spin from \(H_J\). It was shown that the resulting vacancy can propagate, which is equivalent to having the electron at the destination site hopping to the originally vacant site. The effective physics leads to the t-J model, which can
also be derived from the $U \to \infty$ limit of the one-band Hubbard model\cite{5}.

$$H_{tJ} = \hat{P} \left[ \sum_{i,j} t_{ij} \left( c_{j,\sigma}^\dagger c_{i,\sigma} + \text{h.c.} \right) \right] \hat{P} + J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j$$  \hspace{1cm} (1.3)$$

where the $c_{i,\sigma}^\dagger$ creates a spin-$\sigma$ electron at site $i$ and $\hat{P}$ projects out states with double electron occupancy. In the literature, t-J model refers to the original case of $t_{i,j} = 0$ for hopping range longer than the nearest neighbor. The community soon realized that further hopping is needed for better agreement with experiments, especially the dispersion observed by ARPES for the undoped parent insulating compound\cite{38,39}. The case with next-nearest-neighbor hopping is known as t-t'-J model, and further corrections in hopping are denoted as t-t'-t''-J and so forth.

The t-J model is more manageable compared to other electronic models because each additional hole removes a spin from the problem instead of adding one and because the effective Hamiltonian acts on a square lattice (Tab. 1.1). The (N+n)-fermion problem is reduced to a (N-n)-spin problem with a no-double-occupancy constraint. Equation 1.3 is deceptively simple, as it took many years of effort from the community to gain a satisfactory understanding of the model.

Although exact analytical and numerical solutions are still lacking, the combination of mean-field treatment and variational Monte-Carlo calculation makes the t-J model the most understood microscopic model\cite{5-7}. The most prominent mean-field description evolved from the connection between the resonating valence bond (RVB) ansatz to an undoped AFM and the BCS wavefunction\cite{40}. In the slave-boson approach to the t-J model, an electron creation is equivalent to a simultaneous creation of a fermionic spin and destruction of a bosonic vacancy

$$c_{i,\sigma}^\dagger = f_{i,\sigma}^\dagger b_i$$  \hspace{1cm} (1.4)$$

if the restriction of

$$f_{i,\uparrow}^\dagger f_{i,\downarrow} + f_{i,\downarrow}^\dagger f_{i,\uparrow} + b_i^\dagger b_i = 1$$  \hspace{1cm} (1.5)$$

is enforced at each site. The boson $b_i^\dagger$ takes on the meaning of a ZRS, which is conveniently a $S=0$ singlet. The t-J model can thus be viewed as spinons $f^\dagger$ and holons $b^\dagger$ coupled via Eq. 1.3 plus the constraint in Eq. 1.5. Treating the constraint
approximately, the mean-field theory yields the phase diagram shown in Fig. 1.4. Holons Bose condense below $T_B$. Spinons acquire coherent motion below $T_D$ and pairs into singlets to form a RVB state with d-symmetry below $T_{RVB}$. In the left region for $T_B < T < T_{RVB}$, the RVB state is formed without Bose condensation, and the pseudogap region in Fig. 1.1 is explained in terms of the formation of spin singlets. The superconducting region for $T < T_{RVB}, T_B$ is explained as a RVB spin singlet state with Bose condensation of the charge carrying holons. To the right for $T_{RVB} < T < T_B, T_D$, the spinons no longer form into singlets, and the remaining coherent motion leads to Fermi-liquid like behavior. In the region above the superconducting area for $T_{RVB}, T_B < T < T_D$, the non-Fermi-liquid region is attributed to spinons coherence without singlet formation. The highest $T_c$ predicted by mean-field theory is about a factor of two to four higher than the highest known $T_c$ for single-layer compounds[5, 6]. Remarkably, the non-Fermi-liquid $\sim T$ resistivity was found in accordance with experiment.

The most significant difference between the mean-field phase diagram and the general one in Fig. 1.1 is the superconducting region which extends all the way to half-filling as well as the lack of separate AFM phase in that region. In this
low-doping regime, variational Monte-Carlo calculation showed that the RVB-type ansatz does not yield the lowest energy. A combination of RVB and AFM character in the variational ansatz can yield a lower energy state, signaling the coexistence of superconductivity and antiferromagnetism at low-doping and low temperature[41, 42].

Another inconsistency between the t-J model and experiment at low doping is that the one-hole solution predicts a band of coherent quasiparticles with sharp peak in the spectral weight; however, ARPES measurement at 300K reported a wide peak of 300meV in width, which is as broad as the band dispersion, decreasing linearly with temperature. It has been suggested that coupling to phonons might be the cause of this broadening[43], and that phonons might be relevant to cuprate physics, but in a way different from conventional superconductor. This sparked the study of carrier-phonon interactions for models relevant to the cuprate lattice structure[44]. For an electron-phonon coupling parameter much larger than that predicted by density functional theory, a wide peak broadening that decreases linearly with temperature was found[45]. However, the issue has not been elucidated because phonon theories predict that the linear decrease ceases at around 200K, but no experimental data is available below that temperature due to technical limitations. It was also pointed out that phonons can at least enhance spin-mediated superconductivity.

1.3 Towards Realistic Oxide Structures

Holistically speaking, a full understanding of the physics of a CuO$_2$ layer doped with a few holes has still not been achieved despite continuous effort[11, 46, 47]. In particular, the role of phonons[44, 48], time-reversal symmetry breaking[49], and spatial symmetry breaking[50] have not been elucidated.

Recent experiments have raised further questions regarding the understanding of cuprates, in addition to the ARPES line shape broadening [43, 51–54] discussed in the last section. Neutron experiments have been performed in the pseudogap phase and reported magnetic response throughout the Brillouin zone, not restricted to the region of the much discussed magnetic resonance[55, 56]. The neutron results have been used to rule out one-band models such as the t-J model.
These and other issues including the broken local fourfold symmetry, which is taken for granted by the t-J model in Eq. 1.3, seen in scanning tunneling probe (STM)[57] and x-ray scattering [58] remain either open questions or are controversial. While oxygen-bond-specific characteristics observed by XAS, EELS, and STM [10, 59, 60] cannot be described using one band, the significance of omitting other bands cannot be quantified without a comparison to unbiased solutions of more detailed models.

Aside from adding more parameters such as long range hopping and carrier-phonon coupling for better agreement with experiments, a logical step is to study models more faithful to the realistic structure of the CuO$_2$ plane to verify the main concepts of the t-J paradigm. For example, even though the RVB nature is compatible with the BCS wavefunction[40], the robustness of spin-charge separation $c_i^\dagger \sigma = f_i^\dagger \sigma b_i$, where the holon $b_i^\dagger$ takes on the ZRS’s identity, should be verified in more detailed scenarios.

In particular, cuprates exhibit charge-transfer band-gap behavior with mobile holes located mainly on anion ligands and unpaired electrons on cation d orbitals [13]. One trade-off for the t-J model’s elegance is the use of momentum-independent effective parameters, even though it is well known that the ZRS state has a strong $k$-dependent renormalization [16]. The impact of such approximations must be verified with models that distinguish anion and cation sites.

Carrier-phonon interactions have been studied almost exclusively in the framework of the Holstein model[61]. The Holstein coupling is compatible with the t-J model because it does not distinguish anion and cation sites. The effect of such distinction should be studied as well.

1.4 Scope

Given how the conventional understanding of electron-electron and electron-lattice interactions fails to transcribe into cuprate, and how the most developed microscopic description relies on significant assumptions, it is important to carefully examine these interactions in detail. The description of fermions interacting with other degrees of freedom is challenging, especially when models retain the specifics of realistic material structures. One powerful concept is the formation of polaron —
a quasiparticle composed of the propagating carrier accompanied by some changes in the otherwise undisturbed background. While the quasiparticle description is certainly robust in the dilute limit, intuition about the many-body background is crucial in the actual solution process and is also the central theme of this dissertation. Physical insights about the many-body background were exploited to advance the treatment of models relevant to the transition metal oxide structures.

Chapter 2 discusses the modeling of lattice polarons. A previously proposed numerical scheme is adapted to study the breathing-mode Hamiltonian which details the interaction between a carrier and a background of vibrating oxygen ions. Results are compared with those of the Holstein model which is the most primitive starting point of carrier-phonon interaction.

The “parent” cuprate compound is an AFM insulator (Eq. 1.2). The many electrons order antiferromagnetically with an energy scale of $J \sim 150\text{meV}$. The AFM behavior is expected to be robust at low doping at low temperature, so an injected charge carrier can be thought of as a spin polaron of the carrier with a local disturbance to the AFM background. This spin polaron problem does not enjoy an exactly known background like the bosonic $b^{-}\ket{0} = 0$ in a carrier-phonon interaction scenario. Chapter 3 discusses progress in dealing with this issue. A novel octapartite numerical approach is devised to capture the two-dimensional Heisenberg antiferromagnetic spin background with breakthrough efficiency.

A surprising feature is revealed during the application process of the aforementioned scheme to the doped copper-oxygen plane. Contrary to the common belief of ZRS formation and to the intuition that spin-$\frac{1}{2}$ quasiparticles would result from a spin-$\frac{1}{2}$ hole interacting with a spin-0 anti-ferromagnetic background, Chapter 4 shows that, when details are treated properly, spin-$\frac{1}{2}$ and spin-$\frac{3}{2}$ polarons become energetically favorable in different regions of the Brillouin zone. The chapter covers the derivation of a spin polaron model relevant to cuprates, the solution process, and the nature of the low-energy wavefunctions.

Chapter 5 studies two fermionic holes in the aforementioned spin-polaron framework. The computational approach from Ch. 3 is utilized to bypass various technical barriers in modeling large systems. The result serves as a proof of concept, demonstrating the scheme’s prowess in capturing the AFM background disturbed by multiple holes. The results show enhanced singlet correlation.
Lastly, the conclusion is presented in Ch. 6. It summarizes the progress achieved in this series of works and presents various opportunities for further developments.

Appendix A outlines the extensive efforts spent in ensuring the integrity of these novel numerical approaches. Computational optimizations that enabled these calculations are listed in Appendix B.
Chapter 2

Lattice Polaron on the Oxide Chain

2.1 Introduction

In a solid-state system, the interaction between a charge carrier and phonons (quantized lattice vibrations) leads to the formation of polarons. This mechanism is a key ingredient in the physics of the manganites\[8\], Bechgaard salts\[62]\[63\], and, possibly, of the cuprates\[43]\[64\]. There are various model Hamiltonians describing the coupling of the particle and bosonic degrees of freedom. The asymptotic limits of weak or strong coupling can be investigated analytically using perturbation theory; however, the intermediate-coupling regime generally requires numerical simulations. Recently, investigations of basic model Hamiltonians have progressed rapidly thanks to the development of efficient analytical and computational tools, and we are now able to begin to study more and more realistic models.

We studied numerically various low-energy properties and the spectral function of the single polaron in the 1D breathing-mode Hamiltonian which captures some details of the transition metal oxide structure. The results are compared with the relevant results for the single Holstein 1D polaron, allowing us to contrast the behavior of the polarons in the two models. The chapter is organized as follows: Sec. 2.2 introduces the two models of interest. Sec. 2.3 discusses relevant asymptotic results and describes the numerical methods we use to calculate low-energy
properties and the spectral functions for both models. Sec. 2.4 presents the results. Conclusions are made in Sec. 2.5.

### 2.2 The Holstein and Breathing-Mode Models

The simplest electron-phonon coupling is described by the Holstein Hamiltonian.\[61\] It is essentially the tight-binding model with an on-site energy proportional to the lattice displacement \(X_i = \frac{1}{\sqrt{2M\Omega}} (b_i + b_i^\dagger)\):

\[
H_H = -t \sum_{<ij>} \left( c_i^\dagger c_j + c_j^\dagger c_i \right) + \Omega \sum_i b_i^\dagger b_i + g \sum_i n_i X_i
\] (2.1)

Here \(c_i\) is the annihilation operator for an electron at site \(i\) (since we only consider the single electron case, the spin is irrelevant and we drop its index. We also set \(\hbar = 1\)). \(t\) is the hopping matrix, \(n_i = c_i^\dagger c_i\). For the Einstein phonons, \(b_i\) is the annihilation operator at site \(i\), \(\Omega\) is the frequency, \(M\) is the atomic mass, and \(g\) is the electron-phonon coupling strength. The model has been widely studied numerically by Monte-Carlo calculations[65–76], variational methods[77][78–89], and exact diagonalization[90–95]. Analytic approximations, such as the momentum-average approximation, have also progressed over the years[96–99].

For some materials, a more appropriate model is provided by the breathing-mode Hamiltonian. For example, consider a half-filled 2D copper-oxygen plane of a parent cuprate compound. Injection of an additional hole should fill an oxygen 2\(p\) orbital. Due to hybridization between oxygen 2\(p\) and copper 3\(d_{x^2−y^2}\) orbitals, the hole resides in fact in a ZRS with a binding energy proportional to \(-8t_{dp}^2\), where \(t_{dp}\) is the hopping between neighboring O and Cu orbitals. The dynamics of the ZRS can be described by an effective one-band model with orbitals centered on the copper sub-lattice.[16] If lattice vibrations are considered, the motion of the lighter oxygen ions, which live on the bonds connecting Cu sites, are the most relevant. The hopping integral \(t_{dp}\) and charge-transfer gap between Cu and O orbitals are now modulated as the oxygen moves closer or further from its neighboring Cu atom. Both the on-site energy and hopping integral are modulated in the effective one-band model, but the former has been shown to be dominant[100][101]. The breathing-mode Hamiltonian describes the physics of the linear modulation of on-
site energy.

While this breathing-mode model is motivated as a 2D model, here we investigate numerically only its 1D version, relevant, e.g. for CuO chains. In 1D we can investigate accurately and efficiently not only ground-state (GS) properties, but also some excited state properties. For the Holstein model, it was found that polaron properties are qualitatively similar in different dimensions,\cite{83, 98} but with a sharper large-to-small polaron crossover in higher dimensions. We will show that a sharp crossover is already present in the 1D model, and we expect less of a dimensionality effects in the breathing-mode Hamiltonian.

Quasi-1D systems, such as the Bechgaard salts, also involve electron-phonon coupling which is more complicated than the Holstein model; however, these systems have a rather complicated structure and involve both strong electron-electron and electron-phonon interaction. The electron-phonon coupling also modulates the intermolecular hopping integrals in addition to the on-site energy. This adds a considerable degree of complication to the calculations and will be a subject of future studies.

The 1D breathing-mode Hamiltonian that we investigate here is described by:

\[
H_B = -t \sum_i \left( c_i^\dagger c_{i+1} + c_{i+1}^\dagger c_i \right) + \Omega \sum_i b_{i+\frac{1}{2}}^\dagger b_{i+\frac{1}{2}} + \frac{g}{\sqrt{2M\Omega}} \sum_i n_i \left( b_{i+\frac{1}{2}}^\dagger + b_{i+\frac{1}{2}} - b_{i-\frac{1}{2}}^\dagger + b_{i-\frac{1}{2}} \right) \tag{2.2}
\]

The notation is the same as before, except that now the phonons live on an interlaced lattice. The difference between the two models is more apparent in momentum space. The Holstein model has constant coupling to all phonon modes

\[
V_H = \frac{g}{\sqrt{N\sqrt{2M\Omega}}} \sum_{kq} c_{k-q}^\dagger c_k \left( b_q^\dagger + b_{-q} \right),
\]

whereas the breathing-mode model has a coupling strength that increases monotonically with increasing phonon momentum

\[
V_B = \frac{2ig}{\sqrt{N\sqrt{2M\Omega}}} \sum_{kq} \sin \frac{q}{2} c_{k-q}^\dagger c_k \left( b_q^\dagger + b_{-q} \right).
\]
Here \( N \) is the number of lattice sites, and becomes infinite in the thermodynamic limit. The momenta \( k, q \) are restricted to the first Brillouin zone \( (-\pi, \pi] \) (we take the lattice constant \( a = 1 \)).

While numerical and analytical studies of the Holstein polaron abound, there is much less known about models with \( g(q) \) coupling. In particular, there is no detailed numerical study of the 1D breathing-mode model, apart from an exact diagonalization of a simplified t-J-breathing-mode model in restricted basis [102], and an investigation based on the self-consistent Born approximation [103], which is known to become inaccurate for intermediate and strong couplings.

### 2.3 Methodology

#### 2.3.1 Strong-Coupling Perturbation Results

Perturbational results for the strong-coupling limit \( g \gg t \) provide a good intuitive picture of the problem even for the intermediate coupling regime. In the absence of hopping, \( t = 0 \), both Hamiltonians can be diagonalized by the Lang-Firsov transformation[104]

\[
\hat{O} = e^S O e^{-S}.
\]  

Using

\[
S_H = \frac{g}{\Omega \sqrt{2M \Omega}} \sum_i n_i (b_i^\dagger - b_i),
\]

and

\[
S_B = \frac{g}{\Omega \sqrt{2M \Omega}} \sum_i n_i (-b_{i-\frac{1}{2}}^\dagger + b_{i-\frac{1}{2}} + b_{i+\frac{1}{2}}^\dagger - b_{i+\frac{1}{2}})
\]

respectively, the diagonal forms of the Hamiltonians are, in terms of the original (undressed) operators:

\[
\hat{H}_H = \tilde{T}_H + \Omega \sum_i b_i^\dagger b_i - \frac{g^2}{2M \Omega^2} \sum_i n_i^2
\]

\[
\hat{H}_B = \tilde{T}_B + \Omega \sum_i b_{i,\frac{1}{2}}^\dagger b_{i+\frac{1}{2}} - \frac{2g^2}{2M \Omega^2} \sum_i n_i (n_i - n_{i+1})
\]
where the kinetic energies are:

\[ T_H = -te^{-\frac{z^2}{2M\Omega^2}} \sum_i c_{i+1}^\dagger c_i e^{\frac{g(-b_{i+1}^\dagger + b_i^\dagger)}{\sqrt{2M\Omega}}} e^{-\frac{g(-b_{i+1} + b_i)}{\sqrt{2M\Omega}}} + h.c. \]

\[ T_B = -te^{-3\frac{z^2}{2M\Omega^2}} \sum_i c_{i+1}^\dagger c_i e^{\frac{g(-b_{i+1}^\dagger + b_i^\dagger)}{\sqrt{2M\Omega}}} e^{-\frac{g(-b_{i+1} + b_i)}{\sqrt{2M\Omega}}} \times \]

\[ e^{-\frac{g}{\sqrt{2M\Omega}}(b_{i+\frac{1}{2}} - 2b_i + b_{i-\frac{1}{2}})} + h.c. \]

For a \( d \)-dimensional lattice, \( \tilde{T}_B \) is modified by i) extra creation and annihilation operators of phonons in the direction transverse to hopping, and ii) change of the \(-3\) factor in the exponent to \(-(z+1), z = 2d\). The third term in Eq. (2.4) and in Eq. (2.5) signifies that the mere presence of an electron would induce a lattice deformation, leading to the formation of a polaron to lower the system’s energy. For a single polaron, the lattice deformation energy is proportional to the number of nearest phonon sites (one for the Holstein model and \( z \) for the breathing-mode model). For \( t = 0 \), the ground-state energy is degenerate over momentum-space:

\[ E_H^{(0)}(k) = -\frac{g^2}{2M\Omega^2} \]

\[ E_B^{(0)}(k) = -z\frac{g^2}{2M\Omega^2} \]

Each model has three energy scales, therefore the parameter space can be characterized by two dimensionless ratios. It is natural to define the dimensionless (effective) coupling as the ratio of the lattice deformation energy to the free-electron ground-state energy \(-zt\):

\[ \lambda_H = \frac{g^2}{2M\Omega^2} \frac{1}{zt} \]  

\[ \lambda_B = \frac{g^2}{2M\Omega^2} \frac{1}{t} \]  

where \( z \) is also the number of nearest neighbors in the electron sublattice. It should be noted that since \( \Omega \sim 1/\sqrt{M} \), the \( \lambda \)'s do not depend on the ion mass, \( M \). \( \lambda \) has
been shown to be a good parameter to describe the large-to-small polaron crossover in the Holstein model. It will be shown in later sections that the definition also works well for the breathing-mode model. The other parameter is the adiabatic ratio which appears naturally from the perturbation in $t$:

$$\alpha = \frac{zt}{\Omega}. \quad (2.8)$$

Using standard perturbation theory[95], the first-order corrections to the energy of the lowest state of momentum $k$ are:

$$E^{(1)}_H(k) = -2te^{-\alpha\lambda_H} \cos(k), \quad (2.9)$$
$$E^{(1)}_B(k) = -2te^{-\frac{1}{2}\alpha\lambda_B} \cos(k), \quad (2.10)$$

showing that the polaron bandwidth is exponentially suppressed in the strong coupling limit. As is well known, this is due to the many-phonon clouds created on the electron site (Holstein) or on the two phonon sites bracketing the electron site (breathing-mode model). As the polaron moves from one site to the next, the overlaps between the corresponding clouds become vanishingly small and therefore $t_{\text{eff}} \to 0$. To first order in $t$, the suppression is stronger for the breathing-mode model simply because the overlap integral involves phonon clouds on 3 sites instead of just 2, as in the case for Holstein. The second-order corrections are:

$$E^{(2)}_H(k) = -2t^2 \frac{\Omega}{e^{-2\alpha\lambda_H} f_H(k, \alpha, \lambda_H)},$$
$$E^{(2)}_B(k) = -2t^2 \frac{\Omega}{e^{-3\alpha\lambda_B} f_B(k, \alpha, \lambda_B)}.$$

The functions $f$ can be written in the form

$$f_{H,B} = A_{H,B} + B_{H,B} \cos(2k)$$

with

$$A_H = Ei(2\alpha\lambda_H) - \ln(2\alpha\lambda_H) - \gamma$$
$$B_H = Ei(\alpha\lambda_H) - \ln(\alpha\lambda_H) - \gamma$$

20
Here, \( \gamma \) is the Euler-Mascheroni constant and \( Ei(x) \) is the exponential integral with the series expansion

\[
Ei(x) = \gamma + \ln(x) + \sum_{n=1}^{\infty} \frac{x^n}{n!n}.
\]

The result can be further simplified in the limit \( \alpha \lambda_H, \alpha \lambda_B \gg 1 \) using \( \lim_{x \to \infty} \sum_{n=1}^{\infty} \frac{e^n}{n!n} \sim e^x/x \). This leads to the simplified expressions:

\[
E_H^{(2)}(k) \sim -2 \frac{t^2}{\alpha \lambda_H} \left[ \frac{1}{2} + e^{-\alpha \lambda_H} \cos(2k) \right]
\]

(2.11)

\[
E_B^{(2)}(k) \sim -2 \frac{t^2}{\alpha \lambda_B} \left[ \frac{1}{3} + e^{-\alpha \lambda_B} \frac{1}{2} \cos(2k) \right]
\]

(2.12)

Thus, the breathing-mode model’s ground state energy is slightly higher for any finite \( t \). For the Holstein model, the dispersion is monotonic, since the second order \( \cos(2k) \) contribution is suppressed more strongly than the first order \( \cos(ka) \) contribution. However, a quick comparison between Eq. (2.10) and (2.12) shows that in the breathing-mode model, the second order \( \cos(2k) \) contribution becomes dominant at large enough coupling. As a result, at strong couplings we expect the breathing-mode polaron energy to exhibit a maximum at a finite \( k < \pi \), and then to fold back down.

### 2.3.2 Matrix Computation

The computation method we use is a direct generalization of the method introduced by Bonca et al. for the Holstein model, in Ref. [77]. This approach requires sparse matrix computations to solve the problem. It gives us a systematic way to compute excited state properties, which would be more difficult to achieve using other numerical methods.

The idea is to use a suitable basis in which to represent the Hamiltonian as
a sparse matrix. For the Holstein model, this basis contains states of the general form:

\[ |S, K⟩ = \frac{1}{\sqrt{N}} \sum_j e^{iK_j c_j^+} \prod_{m \in \{s\}} \frac{b_{j+m}^{\dagger}}{\sqrt{n_m}} |0⟩, \] (2.13)

where \( K \) is the total momentum and \( S \) denotes a particular phonon configuration, with sets of \( n_m \) phonons located at a distance \( m \) away from the electron. For the Holstein model, \( m \) are integers, since the phonons are located on the same lattice as the electrons. The generalization for the breathing-mode is simple: here \( m \) are half-integers, since here the phonons live on the interlaced sublattice. All states in either basis can be obtained by repeatedly applying the Hamiltonian to the free electron state which has all \( n_m = 0 \). The possible matrix elements are \(-te^{\pm iK}\), \( \Omega n \), and \( \pm \frac{g}{\sqrt{2M}} \sqrt{n} \), where \( n \) are integers related to the numbers of phonons.

Since the Hilbert space of the problem is infinite, this basis must be truncated for computation. The original cut-off scheme in Ref. [77] was optimized for computation of ground-state properties of the Holstein model, by restricting the number of matrix elements between any included state and the free-electron state. Getting the higher-energy states is more involved, as it is evident from Eq. (2.3) that each state \( |ϕ⟩_{LF} \) in the Lang-Firsov basis correspond to a state

\[ |ϕ⟩_{R} = e^{-S} |ϕ⟩_{LF} \] (2.14)

in real space. With our choice for the \( S \) operators, this reverse transformation induces a phonon coherent state structure at the electron site (Holstein), respectively the two bracketing phonon sites (breathing-mode). The phonon statistics of the coherent state obeys the Poisson distribution. In the anti-adiabatic regime \( (zt > \Omega) \), the splitting due to the hopping (off-diagonal hopping matrix elements) is significant compared to the diagonal matrix elements proportional to \( \Omega \). The underlying Lang-Firsov structure needs to be modeled by the hopping of the electron away from the coherent state structure created by the \( e^{-S} \) operator. To capture these characteristics, the basis is divided into subspaces with fixed numbers of phonons. Each subspace is enlarged by the addition of states with phonons further and further away from the electron site (increase of maximum value of \( m \) in Eq. (2.13)), until convergence is reached. This procedure allows for efficient generation of all
basis states required to model the higher-energy states.

Matrices of dimension up to $10^6 - 10^7$ were needed to compute the two lowest-energy states accurately. These two states were calculated numerically using the Lanczos method with QR shift, [105, 106] which works efficiently for the low energy bound states. For the larger $n$ values used in the Green’s function calculation (see below), SMP machines were used for efficient computation. Table 2.1 lists the most-relaxed cut-off condition we used to calculate the Green’s function (see next section) for the 1D breathing-mode model. We relaxed the cut-off condition rather roughly until convergence was observed. As a result, the size of these matrices is certainly much larger than it has to be.

### 2.3.3 Green’s Function Computation

Computation of higher-energy properties requires much larger matrices. The memory and fflops needed for such computations are formidable, especially to extensively investigate the multi-dimensional parameter space $(\lambda, \alpha, K)$. Furthermore, one characteristic of the single polaron problem is a continuum of states starting at one phonon quantum above the $K = 0$ groundstate. Lanczos-type methods typically are problematic in dealing with bands of eigenvalues with small separation. Therefore, in order to study higher energy states, we calculate directly the Green’s function:[107]

$$G(\omega, k) = \langle k | \frac{1}{\omega - H + i\eta} | k \rangle,$$

(2.15)
where $|k\rangle = c_k^\dagger |0\rangle$. This can be written as the solution of a linear system of equations:

$$G(\omega, k) = \langle k|y\rangle$$

$$(\omega - H + i\eta)|y\rangle = |k\rangle$$

One can iteratively tri-diagonalize $H$ by the vanilla Lanczos process:[108]

$$H = QTQ^\dagger$$

$$(\omega + i\eta - T)Q^\dagger |y\rangle = Q^\dagger |k\rangle$$

If the RHS is of the form $[100...]^T$, Cramer’s rule can be used to express $G = \langle k|QQ^\dagger |y\rangle$ as a continuous fraction in terms of the matrix elements of the tri-diagonal matrix $(\omega + i\eta - T)$. In particular, this condition is achieved by picking the initial Lanczos vector to be $|k\rangle$. This method is efficient because it does not require the complete solution of the linear system nor of the eigenvalue problem. It is well known that this type of iterative process suffers from numerical instability, which leads to the loss of orthogonality in $Q$ and incorrect eigenvalue multiplicity in $T$[109]. We perform the vanilla Lanczos tri-diagonalization and re-orthogonalize each state in $Q$ against the starting vector $|k\rangle$ to validate the continuous fraction expansion. Then, numerical errors may come from the fact that $T$ may have the wrong eigenvalue multiplicity. However, this will not affect the location of poles in the spectral weight; that is, the eigenenergies are accurate.

### 2.4 Results

#### 2.4.1 Low-Energy States

The ground state energy $E_{GS}$ and quasiparticle ($qp$) weight $Z_0 = |\langle \Phi_{GS} | c_{k=0}^\dagger |0\rangle|^2$, where $|\Phi_{GS}\rangle$ is the ground-state eigenfunction, are shown in Fig. 2.1 for the 1D breathing-mode and Holstein models. For a fixed value of $\alpha$, we see the expected crossover from a large polaron (at weak coupling $\lambda_{B,H}$) to a small polaron (at strong coupling $\lambda_{B,H}$), signaled by the collapse of the $qp$ weight.
The ground state energy of both models decreases monotonically with increasing coupling, but that of the Holstein polaron is lower. This is in agreement with the second order strong-coupling perturbation results in Eq. (2.11) and (2.12). Unlike the rather gradual decrease in the quasiparticle weight of the 1D Holstein polaron, the 1D breathing-mode polaron shows a large $Z_0$ at weak couplings, followed by a much sharper collapse at the crossover near $\lambda_B \approx 1$. The reason for the enhanced $Z_0$ at weak couplings is straightforward to understand. Here, the wave-function is well described by a superposition of the free electron and electron-plus-one-phonon states. Given the conservation of the total polaron momentum $K = 0 = k + q$ and the large electron bandwidth $t$, states with high electron and phonon momenta have high energies and thus contribute little to the weak-coupling polaron ground-state. On the other hand, the coupling $g(q) \sim \sin(q/2)$ to the low-energy states with low electron and phonon momenta is very small for the breathing-mode model. This explains the slower transfer of spectral weight at weak coupling for breathing-mode
Figure 2.2: (a) Energy of the first excited $K = 0$ state, measured from $E_{GS}$, and (b) its $qp$ weight. The dashed line corresponds to the Holstein model, while the breathing-mode results are shown by circles (line is a guide to the eye). Parameters are $t = 2, \Omega = 1$.

versus the Holstein polaron.

The energy (measured from $E_{GS}$) and $qp$ weight of the first excited $K = 0$ state are shown in Fig. 2.2. For both models, at weak-coupling this state is precisely at $\Omega$ above the ground-state energy, at the lower edge of the polaron-plus-one-phonon continuum. As the coupling increases above a critical value, a second bound state gets pushed below the continuum. This second bound state is absent in SCBA calculations[103]. The separation between the two lowest energy states now first decreases and then increases back towards $\Omega$ as $\lambda_{H, B} \to \infty$. This behavior is well-known for the Holstein polaron[77]. The breathing-mode polaron shows the same qualitative behavior. Note that below the critical coupling, the computed energy of the first excited state is slightly larger than $\Omega$. The reason is a systematic error that can be reduced by increasing the number of one-phonon basis states, in order to better simulate the delocalized phonon that appears in this state. The $qp$ weight of the first excited state is zero below the critical coupling, due to the crossing
between on-site and off-site phonon states[77].

The nature of these states is revealed by checking the locality of the phonon cloud. We define the projection operator

\[ P(K) = \sum_{S_{local}} \langle S, K \rangle \langle S, K \rangle \]

where the summation is over all states with \( m_H = 0 \) and \( m_B = \pm 0.5 \) in Eq. (2.13). Comparison with Eq. (2.14) shows that this operator selects only basis states with phonons only on the electron site (Holstein) and only on the two phonon sites bracketing the electron site (breathing-mode). Note that the coherent state structure is expressed as phonon creation operators in an exponential function, which includes states with zero number of phonons. Fig. 2.3 shows the expectation value of this operator for the two lowest eigenstates of both models. For both groundstates \( <P(0)> \sim 1 \), indicating that here most phonons are nearest to the electron. However, at weak coupling the first excited state (which is here the band-edge of
the polaron + free phonon continuum) has $\langle P(0) \rangle \to 0$, precisely because the free phonon can be anywhere in the system. When the second bound states form, $\langle P(0) \rangle$ becomes large, showing that phonons in these states are primarily localized near the electron.[77] While there appears to be a crossing between the ground-state and first-excited state values, we emphasize that $P(K)$ measures the locality of the phonon cloud, not its structure.

For the breathing-mode model, these results suggest the possibility to describe them using the on-site coherent-state structure. That is, a Lang-Firsov state with $n_-$ and $n_+$ number of phonons excited to the left and right of the electron site, mapped to real space by Eq. (2.14). We note that we are no longer in the strong coupling regime and the transformation cannot be determined by $g$ and $\Omega$ alone; therefore, we seek an effective transformation with

$$\tilde{S}(\Delta) = S_B |_{\Omega = \Delta}$$

The computed eigenstates $|\phi\rangle$ are projected into such structure $\alpha_{n_-, n_+}$ by

$$\frac{1}{\sqrt{N}} \sum_l e^{iKl} c_l^\dagger \sum_{n_-, n_+ = 1}^{\infty} \alpha_{n_-, n_+} \frac{(b_{l-\frac{1}{2}}^\dagger)^{n_-} (b_{l+\frac{1}{2}}^\dagger)^{n_+}}{\sqrt{n_-!} \sqrt{n_+!}} |0\rangle = e^{\tilde{S}(\Delta)} P(K) |\phi\rangle.$$

Tables 2.2 and 2.3 show the results of such projections for the ground state and for the first excited bound state. It is clear that they can be rather well described as $|GS\rangle \sim e^{-\tilde{S}(t, g, \Omega, K)} \frac{1}{\sqrt{N}} \sum_l e^{iKl} c_l^\dagger |0\rangle$, respectively $|1\rangle_{\text{bound}} \sim e^{-\tilde{S}(t, g, \Omega, K)} \frac{1}{\sqrt{N}} \sum_l e^{iKl} c_l^\dagger (e^{i\theta} b_{l-\frac{1}{2}}^\dagger - e^{-i\theta} b_{l+\frac{1}{2}}^\dagger) |0\rangle$ for some phase $\theta(t, g, \Omega, K)$ needed to satisfy time-reversal symmetry. These states no longer have definite parity symmetry like those of the Holstein model. The symmetry is broken by the anti-symmetric coupling term in the model. If the free-electron component is non-zero for an eigenstate, its components with odd/even number of phonons should have odd/even parity. For increasing momentum, this description is valid as long as the excited state remains bound, with energy less than $E_{GS} + \Omega$.

Fig. 2.4 shows momentum dependent energy and $qp$ weight for the two lowest eigenstates of the breathing-mode model, for an intermediate coupling strength.
Table 2.2: $\alpha_{n_+,n_-}$ vs. $n_+,n_-$ for the ground state

<table>
<thead>
<tr>
<th>$n_+,n_-$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

$t = 2\, \Omega = 1\, g = 0\, \Delta = 0$

<table>
<thead>
<tr>
<th>$n_+,n_-$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.9150</td>
<td>0.0584</td>
<td>0.1372</td>
<td>-0.0477</td>
</tr>
<tr>
<td>1</td>
<td>-0.0584</td>
<td>-0.1681</td>
<td>0.0520</td>
<td>-0.0536</td>
</tr>
<tr>
<td>2</td>
<td>0.1372</td>
<td>-0.0520</td>
<td>0.0549</td>
<td>-0.3300</td>
</tr>
<tr>
<td>3</td>
<td>0.0477</td>
<td>-0.0536</td>
<td>0.0330</td>
<td>-0.0255</td>
</tr>
</tbody>
</table>

$t = 2\, \Omega = 1\, g = 1.5\, \Delta = 1.05$

<table>
<thead>
<tr>
<th>$n_+,n_-$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.9876</td>
<td>-0.0509</td>
<td>0.0082</td>
<td>-0.0025</td>
</tr>
<tr>
<td>1</td>
<td>0.0509</td>
<td>-0.0100</td>
<td>0.0034</td>
<td>-0.0020</td>
</tr>
<tr>
<td>2</td>
<td>0.0082</td>
<td>-0.0034</td>
<td>0.0022</td>
<td>-0.0019</td>
</tr>
<tr>
<td>3</td>
<td>0.0025</td>
<td>-0.0020</td>
<td>0.0019</td>
<td>-0.0019</td>
</tr>
</tbody>
</table>

$t = 2\, \Omega = 1\, g = 1.964\, \Delta = 1.964$

Table 2.3: $\alpha_{n_-,n_+}$ vs. $n_+,n_-$ for the first-excited state

<table>
<thead>
<tr>
<th>$n_+,n_-$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0228</td>
<td>-0.6515</td>
<td>0.0152</td>
<td>0.1151</td>
</tr>
<tr>
<td>1</td>
<td>0.6515</td>
<td>0.0062</td>
<td>0.1700</td>
<td>-0.0473</td>
</tr>
<tr>
<td>2</td>
<td>0.0152</td>
<td>-0.1700</td>
<td>0.0505</td>
<td>-0.0456</td>
</tr>
<tr>
<td>3</td>
<td>0.1151</td>
<td>-0.0473</td>
<td>0.0456</td>
<td>-0.0220</td>
</tr>
</tbody>
</table>

$t = 2\, \Omega = 1\, g = 1.5\, \Delta = 1.05$

<table>
<thead>
<tr>
<th>$n_+,n_-$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.0859</td>
<td>-0.6675</td>
<td>0.0630</td>
<td>-0.0164</td>
</tr>
<tr>
<td>1</td>
<td>0.6675</td>
<td>-0.0864</td>
<td>0.0246</td>
<td>-0.0120</td>
</tr>
<tr>
<td>2</td>
<td>0.0630</td>
<td>-0.0246</td>
<td>0.0134</td>
<td>-0.0099</td>
</tr>
<tr>
<td>3</td>
<td>0.0164</td>
<td>-0.0120</td>
<td>0.0099</td>
<td>-0.0089</td>
</tr>
</tbody>
</table>

$t = 2\, \Omega = 1\, g = 1.964\, \Delta = 1.964$

Table 2.3: $\alpha_{n_-,n_+}$ vs. $n_+,n_-$ for the first-excited state
\( \lambda_B = 1.07 \). The polaron band has a maximum at \( k \sim 0.4 \pi \), in qualitative agreement with the strong-coupling perturbation theory results. This behavior is not captured by SCBA which is only accurate for low coupling strength[103]. For the Holstein polaron, the polaron dispersion is a monotonic function of momentum[77]. Even though the \( q_p \) weights remain moderately high at zero momentum, the weights collapse towards zero with increasing momentum, similar to the well-known Holstein case. This is due to the fact that at large total momentum, the significant contribution to the eigenstate comes from states with at least one or more phonons. The free electron state has a large energy for large momentum, and contributes very little to the lowest energy eigenstates, so indeed \( Z \to 0 \).

The effective masses for the two lowest eigenstates of both models are shown in Fig. 2.5, as a function of \( \lambda_{H,B} \). These were calculated from the second derivative of the energy at momentum \( K = 0 \). For the GS of both models, the effective mass increases monotonically with \( \lambda_{H,B} \). At weak couplings, the breathing-mode polaron is lighter than the Holstein polaron. As already discussed, this is due to
Figure 2.5: Ratio of effective polaron mass to that of the free electron. Circles and squares show breathing-mode results for GS and first bound state, respectively. The other lines correspond to the Holstein model GS (full) and second bound state (dashed). Parameters are $t = 2, \Omega = 1$.

The vanishingly weak coupling to low-momentum phonons. At strong coupling, however, the effective mass is larger for the breathing-mode polaron. This is in agreement with predictions of the strong-coupling perturbation theory, and results from the fact that the hopping of a breathing-mode polaron involves phonon clouds on $2z - 1$ phonon sites, whereas hopping of a Holstein polaron involves phonons at only two sites.

The effective mass of the first excited state can only be defined once this state has split-off from the continuum. It has non-monotonic behavior, first decreasing and then increasing with increasing $\lambda_{H,B}$. This can be understood through the link of the effective mass and the $qp$ weight. In terms of derivatives of the self-energy $\Sigma(k, \omega)$, the effective mass $m^*$ is given by:

$$\frac{m^*}{m} = \left(1 - \frac{\partial \Sigma}{\partial \omega}\right) \cdot \left(1 + \frac{m}{\hbar^2} \frac{\partial^2 \Sigma}{\partial k^2}\right)^{-1},$$
Figure 2.6: GS $qp$ weights for the breathing-mode (circles) and Holstein (squares) models. Full symbols correspond to $\alpha = 8$. For comparison purposes, the empty symbols show the $\alpha = 4$ results of Fig. 2.1(b). ($\Omega = 1$)

where derivatives are evaluated at $K = 0$ and at the corresponding eigenenergy. The first term is linked to the $qp$ weight, $Z = (1 - \frac{\partial \Sigma}{\partial \omega})^{-1}$, so that $m^* \sim 1/Z$. As shown in Fig. 2.2(b), the $qp$ weight of the first excited state has non-monotonic behavior, leading to the non-monotonic behavior of the effective mass.

All the results shown so far were for $\alpha = 4$. For higher $\alpha$ (lower $\Omega$ and/or larger $t$), the difference between the two models can be grasped from Fig. 2.6. Similar to the Holstein model, the large-to-small polaron cross-over occurs at lower $\lambda$ for increasing $\alpha$[83, 98]. At weak and moderate coupling, the $qp$ weight and the effective mass (not shown) in the breathing-mode model are much less sensitive to an increase in $\alpha$ than is the case for the Holstein polaron. This suggests that breathing-mode polarons should be better charge carriers than the Holstein polarons in this regime.
2.4.2 Spectral Function

The spectral function is proportional to the imaginary part of the Green’s function:

\[ A(k, \omega) = -\frac{1}{\pi} \text{Im} G(k, \omega) \]  

(2.16)

In terms of single electron eigenstates and eigenfunctions \( H|n\rangle = E_n|n\rangle \), we obtain the Lehmann representation:

\[ A(k, \omega) = \sum_n |\langle n|c_{k}^\dagger|0\rangle|^2 \delta(\omega - E_n) \]

Of course, since we use a finite small \( \eta \) in numerical calculations, the \( \delta \)-functions are replaced by Lorentzians of width \( \eta \) [see Eq. (2.15)]. We calculate the Green’s functions as discussed in the previous section.

Fig. 2.7 shows the spectral function for zero momentum as a function of en-
Figure 2.8: $A(k, \omega)$ vs $\omega$, for $K/\pi=0; 0.25; 0.5; 0.75$ and 1 (top to bottom) for intermediate coupling $t=2, \Omega=1, \lambda_B=1.25$. $\eta/4 = \Delta E = 0.001 \Omega$. The height of the spectral weight is plotted in linear scale on the left and logarithmic scale on the right. The two vertical lines indicate $E_{GS}$ and $E_{GS} + \Omega$.

Energy. Results corresponding to four different coupling strengths $\lambda_B$ from near the crossover region are shown for the breathing-mode polaron. We note that there is always a continuum starting at one phonon quantum above the ground state energy. This is more clearly visible in the right panel, where the spectral weight is shown on a logarithmic scale, and vertical lines mark the position of the ground-state energy $E_{GS}$, respectively of $E_{GS} + \Omega$. As the coupling $\lambda_B$ increases, we see the appearance of the second bound state below the continuum. We only find at most one extra bound state in this energy range for all coupling strengths. As $\lambda_B$ increases, the spectral weight of the first continuum decreases dramatically. Other bound states form above it, followed by higher energy continua whose weight is also systematically suppressed. This is qualitatively similar to the behavior exhibited by Holstein polaron[99].

Fig. 2.8 illustrates the momentum dependence of the breathing-mode polaron’s spectral weight. The results correspond to a coupling above the critical value,
where there is a second bound state. The majority of the spectral weight is transferred to much higher energies as the momentum increases, and a broad feature develops at roughly the position of the free-electron energy for that momentum. This spectral weight transfer is also qualitatively similar to what is observed for Holstein polarons. Our results have a high-enough resolution to clearly show the continuum at \( E_{GS} + \Omega \) for all values of \( K \). This is part of the kink-like structure reported in Ref. [103]. The logarithmic plot clearly reveals a non-monotonic dispersion of the ground-state like in Fig. 2.4, characteristic for the breathing-mode polaron.

Fig. 2.8 shows only one peak located between the ground-state and the polaron-plus-one-phonon continuum, even though in fact we believe that there are more eigenstates within this region. We found, from eigenvalue computation, additional energy states below the continuum; however, computation of exact energy values requires prohibitively long computation time due to the clustering of eigenvalues. By observing the convergence behavior due to increasing basis size, we can conclude that additional bound states do exist below the continuum. The lack of their contribution to the spectral function can be understood by the fact that the single particle Green’s function only contains information about eigenstates with finite \( qp \) weight, \( |\langle \phi | c_K^\dagger |0\rangle|^2 > 0 \), see Eq. (2.15). These eigenstates must have components corresponding to some Lang-Firsov eigenstate with no off-site phonons (Eq. (2.14)). Also, the wave-function of these states must have a peculiar space inversion symmetry: S-symmetric for all even-phonon-number components and P-symmetric for all odd-phonon-number components. The ground-state always satisfies this requirement, but above a critical coupling, only one other state below the phonon threshold satisfies this requirement.

2.5 Conclusions

In summary, we have reported here the first accurate numerical study of the 1D breathing-mode polaron. A previous study[103] based on the self-consistent Born approximation proves to be inadequate to describe correctly the behavior for medium and strong couplings, as expected on general grounds.

Comparison with the Holstein model results, which correspond to a coupling
\( g(q) = \text{const.} \), reveals some of the similarities and differences of the two models. The breathing-mode polaron is much more robust (has a much larger \( qp \) weight, and less variation with parameters) at weak couplings. This is a direct consequence of the fact that coupling to low-momentum phonons, which is relevant here, becomes vanishingly small \( g(q) \sim \sin(q/2) \rightarrow 0 \). Similar behavior is expected for any other \( g(q) \) model if \( \lim_{q \rightarrow 0} g(q) \rightarrow 0 \). On the other hand, at strong couplings the breathing-mode polaron is much heavier and has a lower \( qp \) weight than the Holstein polaron. This also results from strong-coupling perturbation results, and is due to the fact that in order to move from site \( i \) to site \( i + 1 \), a small breathing-mode polaron must (i) create a new polaron cloud at site \( i + \frac{3}{2} \); (ii) rearrange the polaron cloud at site \( i + \frac{1}{2} \), so that its displacement is now pointing towards site \( i + 1 \), not towards site \( i \); and (iii) relax (remove) the phonon cloud at site \( i - \frac{1}{2} \). This results in a suppressed polaron kinetic energy, and an enhanced effective mass. Because of the larger \( Z \) at weak coupling, and the lower \( Z \) at strong couplings, the crossover from large to small polaron is much sharper for the breathing-mode polaron. Another interesting observation is that the polaron dispersion becomes non-monotonic with momentum \( k \) for medium and large couplings. This can be understood in terms of strong-coupling perturbation theory, which shows that the second-order \( \cos(2k) \) correction is larger than the first-order \( \cos(k) \) correction for large-enough couplings.

Similarities with the Holstein behavior regard the appearance of the polaron-plus-free-phonon continuum at \( E_{GS} + \Omega \), and the appearance of a second bound state with finite \( qp \) weight for large-enough couplings. The convergence of numerics points to the existence of additional bound states, whose absence from the spectral function can be explained by symmetry or missing free-electron components in the wavefunction; however, this issue is not fully settled. Also, the importance of such states for the physical properties is not known. The general aspect of the higher energy spectral weight at strong couplings, as a succession of bound states with large spectral weight and continua with less and less spectral weight, is also reminiscent of the Holstein polaron results.

This chapter considers the propagation of a single fermionic carrier interacting with a many-body phonon background. The effect of the phonon background on the realistic strongly-correlated many-fermion scenario (Ch. 1) can be different.
As will be shown in the rest of the thesis, treating the details of the fermionic part already presents significant challenges and leads to remarkable properties. The following chapters are devoted into the modeling the strongly-correlated fermionic part, leaving the many-fermion-many-phonon scenario for future exploration.
Chapter 3

Octapartite Description of the Antiferromagnetic Square Lattice

3.1 Introduction

The d-wave superconducting gap in cuprates is inconsistent with the s-wave gap due solely to electron-phonon interactions in normal metals. Because of the belief that the dominant physics lies in the two-dimensional copper-oxygen plane and because such planes are antiferromagnetic (AFM) in the undoped limit, much effort has been spent on understanding of charged fermions moving in a two-dimensional background of antiferromagnetically coupled spins[6][7]. In spite of recent efforts that consider carriers interacting with both phonon- and spin-degrees of freedom, an exact analytical solution for the spin part of the problem — doped or not — remains elusive to the best of my knowledge. This adds another layer to the puzzle as compared to the carrier-phonon problem, which takes for granted an exactly known bosonic background, \[ b^- |0\rangle = 0 \text{ and } b^+ |n\rangle = \sqrt{n + 1} |n + 1\rangle. \] Numerical modeling of large-scale, strongly correlated spin systems is thus of extreme importance and has been pursued in many different ways, each with their own advantages and disadvantages.

For example, the powerful Monte Carlo (MC) methods give extremely accurate information for the undoped AFM, providing the benchmark for the ground state (GS) energies and correlation functions of systems with up to thousands.
of spins[110–114]. However, they are limited by the sign problem if additional fermions are present. This led to work on alternative optimized ways of treating the undoped AFM, with the latest development in MC sampling being the use of variational RVB-type ansätze[5, 40–42, 115–117]. In the same context, density matrix renormalization group (DMRG) has also progressed over the years[118, 119].

Since the interactions with the fermions are often comparable or bigger than the AFM’s energy scale, most such AFM-specialized methods require very non-trivial modifications to adapt to doped systems, due to technicalities of the RVB ansatz or the DMRG special boundary conditions. In fact, recent studies still perform exact diagonalization (ED) of the full Hilbert space to complement these somewhat biased schemes[17, 120, 121]. ED studies have the huge advantage that they provide the GS wavefunction, from which any GS properties, including all correlation functions, can be calculated. This would make ED the numerical method of choice, were it not for the extreme restriction on the sizes of doped systems that can be currently treated[17, 121].

The motivation behind this chapter is the belief that some additional physical insights about the AFM background can help in the modeling of the aforementioned unresolved issues. The accuracy and efficiency of the phonon basis, Eq. 2.13 hinges on the insights about the phonon coherent state background Eq. 2.14. Although Bonca et. al. did successfully adapt the idea for a carrier in the spin background[34], the method has left something to be desired because its Néel starting point cannot capture a real AFM background. The strong quantum fluctuations leading to strong deviations from a Néel-like background are believed to be an essential part of the low-energy physics of the doped systems. It is therefore necessary to find accurate yet efficient ways to describe the AFM background, which can also be straightforwardly extended to the doped problem.

3.2 The spin-\(\frac{1}{2}\) Heisenberg Antiferromagnet on the Square Lattice

The dominant physics in an undoped copper-oxide plane is due to the superexchange mechanism[36], which leads to AFM Heisenberg exchange among nearest neighbor \(S = \frac{1}{2}\) copper spins arranged in the square lattice.
$$H_{AFM} = \sum_{\langle i,j \rangle} \left[ S_i^z S_j^z + \frac{1}{2} \left( S_i^+ S_j^- + S_i^- S_j^+ \right) \right] = \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j.$$  \hspace{1cm} (3.1)

Here, a $-2^N$ constant is omitted and the characteristic energy scale $J_{dd} \approx 150 meV$ has been set to unity. The “natural” basis has a dimension of $2^N$ and a wavefunction can be written as a superposition of different spin configurations with $|\sigma_i\rangle_l$ specifying the $z$-projection of spin at the copper site $l$.

$$\sum_{\sigma} c_{\sigma} \prod_l |\sigma_l\rangle_l$$  \hspace{1cm} (3.2)

The total spin $S_T \in [0, \frac{N}{2}]$ is conserved with degenerate $z$-projections so the full spectrum can thus be obtained from the $S_T^z = \sum_l \sigma_l = 0$ basis, which contains $N!/(\frac{N}{2}!)^2$ states without exploiting spatial symmetries. Efficient matrix-vector multiplication can be performed using Algorithm 151 [122] as discussed in Appendix B.

Although the exact ground state wavefunction remains unknown, Marshall showed very early that the groundstate wavefunction must have $S_T = 0$ and that its $c_{\sigma}$ in Eq. 3.2 obeys

$$c_{\sigma} = (-1)^{n_{\sigma}}|c_{\sigma}|$$  \hspace{1cm} (3.3)

where $n_{\sigma}$ is the number of up spins in one sub-lattice[123]. This is now known as the “Marshall sign rule” in the community.

Among the many routes in treating an undoped AFM[37], two intuitions have been prominently linked to the doped case.

### 3.2.1 The Neel Picture

The model is bi-partite because the square lattice can be divided into two interlaced sub-lattices such that spins within each sub-lattice are not coupled to each other by $H_{AFM}$. The bi-partite nature and the dot-product interpretation in Eq. 3.1 motivated the popular “up-down-up-down” classical Neel state starting point. In this picture, the spin-wave theory justified in the $S \gg 1$ limit worked surprisingly well for $S = \frac{1}{2}$[37]. Starting from a Neel background, the $z$-projection of a spin at site $l_u$ in the spin-up or at $l_d$ in the spin-down sublattice is mapped to bosonic number operators.
\[ n_{a} = a_{a}^{+} a_{a} = S - S_{a}^{c} \]
\[ n_{b} = b_{b}^{+} b_{b} = S + S_{b}^{c} \]  \hspace{1cm} (3.4)

Eq. 3.1 can then be expanded in powers of \( \frac{a_{j}}{2S} \) due to the \( \sqrt{n} \) matrix elements of bosonic operators. The collection of O(1) zeroth-order terms is known as the linear spin-wave theory, which can be diagonalized as

\[ H_{LSW} = -0.658N + 2 \sum_{q} \sqrt{1 - \frac{1}{4} (\cos(q_{x}) + \cos(q_{y}))^{2}} \left( \alpha_{q}^{\dagger} \alpha_{q} + \beta_{q}^{\dagger} \beta_{q} \right) \]  \hspace{1cm} (3.5)

where \( \alpha^{\dagger} \) and \( \beta^{\dagger} \) are S=1 “magnon” excitations related to \( a^{\dagger} \) and \( b^{\dagger} \) by a unitary transformation. The prediction of gapless excitation is correct. The groundstate energy-per-site of -0.658 is a bit higher than the best known value of -0.6692. The magnon bandwidth is grossly underestimated to 2 compared to the actual bandwidth of \( \sim 2.5 \). The next order 1/S correction in spin-wave theory lowers the groundstate energy to -0.6705, increases the bandwidth to 2.36 but also introduces three magnon-magnon interaction terms: \( \alpha_{p}^{\dagger} \alpha_{p} \alpha_{q}^{\dagger} \alpha_{q}, \beta_{p}^{\dagger} \beta_{p} \beta_{q}^{\dagger} \beta_{q}, \) and \( \alpha_{p}^{\dagger} \alpha_{p} \beta_{q}^{\dagger} \beta_{q}. \) Even though the results are excellent, the major challenge against the spin-wave theory is that the \( S \gg 1 \) bosonic representation in Eq. 3.4 simply fails for a \( S = \frac{1}{2} \) system because the Hilbert space allows only zero or one occupation. For the undoped case, this occupation constraint can be treated by less-transparent bi-partite methods which cannot be easily generalized to the doped case[37].

The classical Neel state has a finite overlap with some low-energy states of the model, and correcting a Neel state starting point, for example with off-diagonal terms in Eq 3.1, can eventually lead to a decent approximation of the low-energy sector; however, strictly speaking, the zero-temperature ground state do not have a finite Neel characteristic. After all, each \( S \cdot S \) term acting on a Neel state yields \( \pm \frac{1}{4} \) as the diagonal matrix element but \( \pm \frac{1}{2} \) as the off-diagonal matrix element. An upper-bound of the overlap between the classical Neel state and the true ground state wavefunction is derived here. In the bipartite picture, the Neel state has two sub-lattices of opposite spins.
Knowing that the ground state has $S_T = 0$, one can immediately conclude that the Neel state overlaps with only one $S_T = 0$ basis state from the Clebsch-Gordon coefficients:

$$
\langle S_A; m_A | S_B; m_B | S = 0, S_z = 0 \rangle = \left( -1 \right)^{S_A - m_A} \frac{1}{\sqrt{2S_A + 1}} \delta_{S_A, S_B} \delta_{m_A, -m_B} \tag{3.7}
$$

The ground state’s weight in a Neel state is thus at most \( \frac{1}{N^{1/2+1}} \) which vanishes in the \( N \to \infty \) limit. Even though the solution works out in the undoped case, when deviations from the Neel background are equal everywhere, one should be careful when injecting extra fermions into the otherwise homogeneous system. To be specific, the idea of modeling “local flipped spins” without correcting all spins from the Neel background [34] can yield an incomplete solution because a few holes can hardly change the low-range behavior in the \( N \to \infty \) limit. For example, the high-spin polarons in Chapter 4 have not been reported until we performed calculation with specific \( S_T \), which cannot be properly differentiated if some spins remain uncorrected from the Neel background.

### 3.2.2 The Resonating Valence Bond Picture

Anderson has pointed out a vastly different starting point in dealing with Eq. 3.1 by the class of resonating valence bond (RVB) singlet liquid wave functions. The original observation was that a superposition of states with \( \frac{N}{2} \) pairs of singlets can be energetically competitive in low-dimensional and/or frustrated systems. The crucial idea is that a RVB state can be constructed by first removing all \( N \) spins from the lattice, and then projecting a mean-field, BCS wavefunction back to the lattice with the \( \frac{N}{2} \) pairs and no-double-occupancy constrains[40].

$$
|RVB \rangle = \prod_l (1 - n_{l\uparrow} n_{l\downarrow}) P_{N/2} |BCS \rangle \tag{3.8}
$$

Over the years, the RVB framework has been implemented for an AFM torus by two approaches. The bottom-up route consists of the explicit optimization of
the bond amplitudes[115][116]. The top-down approach is by energy minimization by tweaking the unprojected wavefunction \( |BCS\rangle \) in Eq. 3.8 to other related ansatz[41][117][42][5]. Because the optimal variational ansatz is determined as the mean-field solution to some Hamiltonian with optimal values for parameters such as superconducting, antiferromagnetic, and density-wave orders, the real strength of this route is to test a priori postulations about the nature of the solution. The RVB starting point can indeed lead to the correct ground state wavefunction, but the projective, variational nature of the solution procedure hampers its neutrality and transparency.

### 3.3 Octapartite Computational Scheme

One message from section 3.2 is that, in the lightly-doped scenario, the Neel starting point eases the description of local interactions at the cost of accuracy in dealing with the spin background, whereas RVB excels in implementing holistical insights with little details about the local picture. Here, I discuss how a physical insight can be exploited to bridge the gap between these two starting points.

#### 3.3.1 Physical Insights

Knowing that the total spin of the ground state must be zero[123], one can start with the bipartite division and write down all relevant basis states as Clebsch-Gordon series

\[
|S = 0, S^c = 0\rangle = \sum \frac{(-1)^{S_A - m_A}}{\sqrt{2S_A + 1}} \delta_{S_A, S_B} \delta_{m_A, -m_B} |S_A, m_A\rangle |S_B, m_B\rangle \tag{3.9}
\]

for \( S_{A/B} \in [0, \frac{N}{4}], m_{A/B} \in [-S_{A/B}, S_{A/B}] \). Note that the Marshall sign rule Eq. 3.3 is automatically satisfied so all of Marshall’s observations have been exploited. The Hilbert space contains a huge number of such singlets. They can be labeled by an index \( \alpha \), and the GS wave function can be expressed as a superposition

\[
|GS\rangle = \sum_\alpha c_\alpha |0, 0\rangle_\alpha \tag{3.10}
\]

I then noticed one characteristic shared by all \( \alpha \) spin states with high values of
This can be illustrated by the staggered magnetization, which is a well-studied observable of the ground-state wave function. For a bi-partite lattice in a staggered magnetic field, this quantity is formally defined as the difference between $\langle m_A \rangle$ and $\langle m_B \rangle$ in the zero-field limit. In the absence of a staggered field, this quantity has alternative definitions, such as:

$$\hat{m}^2 = \frac{1}{N^2} \left( \sum_r (-1)^{|r|} \hat{S}_r \right)^2 = \frac{(\hat{S}_A - \hat{S}_B)^2}{N^2}, \quad (3.11)$$

where $\hat{S}_{A/B}$ are the total sublattice spin operators. In the AFM GS, $m = \langle \hat{m}^2 \rangle^{\frac{1}{2}}$ has been extrapolated to $\sim 0.3$ as $N \to \infty$, and increases like $1/\sqrt{N}$ to $\sim 0.45$ for $N = 32[37]$. In terms of the total spin $\hat{S} = \hat{S}_A + \hat{S}_B$, we have:

$$\hat{m}^2 = \frac{1}{N^2} (2\hat{S}_A^2 + 2\hat{S}_B^2 - \hat{S}^2) \quad (3.12)$$

The ground state is a singlet, $S = 0$. It follows that the wave function must allocate significant weight to sectors with high values of $S_A = S_B$ in order to yield such large values of $m$. In fact, for the RHS of Eq. 3.12 to be larger than the expected value, the sublattice spins $S_{A/B}$ have to be within $\frac{N}{16}$ of their maximum values. The physical meaning is that $\frac{N}{8}$ spins add to a total spin of zero while the rest adds to the maximum $3\frac{N}{16}$.

### 3.3.2 The Octapartite Approach

There are many ways to add spins quantum mechanically, but not all are viable because the truncation error needs to be bounded systematically and the computational effort must be small. There are two extremes: if the original single-site basis of Eq. 3.1 is used, no basis transformation is needed but enforcing truncation based on Eq. 3.12 is costly. However, if a random basis tabulated according to values of $S_A$ is used, transforming Eq. 3.1 into the new basis would be costly due to the many Clebsch-Gordon series needed. We propose a parametrization which is a good compromise between these two extremes.

Starting from the case where all $\frac{N}{4}$ sublattice spins add to the maximum of $\frac{N}{4}$, if we want to include states with spin down to $3\frac{N}{16} = \frac{N}{4} - \frac{N}{16}$, the basis must
allow many new configurations, including ones where $\frac{N}{8}$ spins have a total spin of 0 while the other $\frac{3N}{8}$ spins have a total spin of $\frac{3N}{16}$. This suggests that groups of $\frac{N}{8}$ or fewer spins must be allowed to take all possible spin quantum numbers. Since a larger group would overshoot the sublattice spin below the $\frac{3N}{16}$ “threshold”, while a smaller group would introduce extra “enforcement” costs as discussed above, we divide the full lattice into groups of 8 sites, see Fig. 3.1. The resulting octads repeat periodically with translational vectors $2a(1, \pm 1)$. Each spin is identified by the octad it belongs to, and by its position inside the octad. A sublattice is composed from 4 groups of spins indexed with the same label, e.g. $\hat{S}_A = \hat{S}_0 + \hat{S}_1 + \hat{S}_2 + \hat{S}_3$. With this arrangement, each spin interacts with spins from all the 4 groups of the other sublattice, e.g. a group 0 spin is always to the west/east/south/north of a spin in group 4/5/6/7. This partition is the minimum division that permits such a description and is a fundamental building block for the wave function in the model. Even though we arrived at it by looking to optimize the computation, our partition has been used in other contexts[124].

Figure 3.1: Square lattice divided in octads (shaded areas). Sites positioned similarly inside octads have the same label. Each is surrounded by neighbors with different labels.
We can now generalize Eq. 3.9 and write singlets of the total lattice as:

\[ |0,0,f(\{S_i\})\rangle = \sum c_{\{S_i,m_i\}} \prod_{i=0}^{7} |S_i,m_i\rangle \]  

(3.13)

where \( S_i,m_i \) are the quantum numbers for \( \hat{S}_i \), i.e. the total spin for group \( i = 0,7 \). Here, \( c_{\{S_i,m_i\}} \) is the product of the appropriate Clebsch-Gordon coefficients for the 8 pairs of quantum numbers \( (S_i,m_i) \). The \( f(\{S_i\}) \) index on the LHS keeps track of the many ways in which these spins can be added to form a singlet.

If all these singlets are kept into the computational basis, the GS calculation is exact. The total spin of each group takes values \( S_i \in [0, \frac{N}{16}] \), and there is no reason to restrict them. However, our previous discussion suggests that while all possible \( S_i \) values must be allowed, the singlets with highest GS weight are those whose total sublattice spin is within \( \frac{N}{16} \) of the maximum value.

We therefore parameterize the singlet Hilbert space in terms of a completeness parameter, \( C_s \in [0,1] \), and include in the singlet basis only states for which \( S_A/B \geq \frac{N}{4}(1-C_s) \). For \( C_s = 1 \), the calculation is thus exact. For \( C_s = \frac{1}{4} \), this means that the maximum number of anti-aligned spins in the sublattice is \( \frac{N}{8} \). Based on the discussion for the staggered magnetization, we expect this to already be a good variational basis for the GS. For any \( C_s \neq 1 \), this formulation allows all possible values of \( S_i \) for each \( i = 0,7 \), but restricts the ways in which they can combine to give the total sublattice spin. The number of discarded states is combinatorially large due to the large degeneracy of states with low sublattice spins.

Equation 3.13 is invariant under spin rotations; thus, the anti-alignment is enforced here in terms of angular momentum addition, very different from the classical Néel picture. This approach also maintains the full translational symmetry of the Hamiltonian. If the RVB bond-amplitude optimization is a bottom-up way of adding AFM order to a singlet[115, 116], our method is a compatible top-down approach without the need of projecting an ansatz.

### 3.3.3 Numerical Implementation

The cost of this truncation scheme is the one-time computation of matrix elements between these basis states. This is acceptable because conventional diagonalization


Table 3.1: Size of the $C_s = \frac{1}{4}$ subspace as compared to that of the commonly used basis.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$C_s = \frac{1}{4}$</th>
<th>$S = 0$</th>
<th>$S_z = 0$</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>50</td>
<td>1430</td>
<td>12870</td>
<td>$2^{16}$</td>
</tr>
<tr>
<td>32</td>
<td>11042</td>
<td>35357670</td>
<td>601080390</td>
<td>$2^{32}$</td>
</tr>
<tr>
<td>64</td>
<td>$9.8 \times 10^8$</td>
<td>$5.55 \times 10^{16}$</td>
<td>$1.83 \times 10^{18}$</td>
<td>$2^{64}$</td>
</tr>
</tbody>
</table>

schemes are limited by storage, not processor speed, and most importantly, the time for matrix element computations was less than the amount of time required to develop the actual numerical solver.

The octapartite division of Fig. 3.1 is valuable for greatly minimizing the number of non-zero matrix elements of the Hamiltonian in the basis states of Eq. 3.13. In particular, the unique neighboring property dictates that each $\vec{S}_i \cdot \vec{S}_j$ term in Eq. 3.1 changes only 2 out of 8 kets in Eq. 3.13. This relatively small change in the wavefunction can by exploited by building the N-spin $S=0$ basis by mixing two identical $\frac{N}{2}$-spin subspaces, which are in turn built from two identical $\frac{N}{4}$-spin subspaces, which are formed by two identical $\frac{N}{8}$-spin subspaces. The matrix element, $\langle \alpha | \vec{S}_i \cdot \vec{S}_j | \alpha' \rangle$, between two N-spin configurations $\alpha$ and $\alpha'$ can be calculated from the $S^i_\uparrow$, $S^i_\downarrow$, and $S^i_z$ matrix elements within the $\frac{N}{8}$ subspace and the Clebsch-Gordon coefficients in the mix hierarchy. The viability of the computation hinges on the fact that $S^i_\uparrow$, $S^i_\downarrow$, and $S^i_z$ yield extremely sparse matrices among the $\frac{N}{8}$ subspace. This efficiency was complemented by tracing the spin-mix hierarchy according to the delta functions in the Clebsch-Gordon coefficients.

While the full $S_z = 0$ basis can be enumerated by, for example, algorithm 151[122], enumerating states of the form Eq. 3.13 is tricky, more so when the $C_s$ condition is enforced and when the above computation needs to be efficient. Increasing $C_s$ increases the basis by lowering the minimum $S_A = S_B$ in the $\frac{N}{2}$-spin enlarges so the basis is enumerated as consecutive blocks of decreasing $S_A$. The number of states within each block is \( \left( \frac{n_{S_A}}{2S_A+1} \right)^2 \), where $n_{S_A}$ is the number of unique $\frac{N}{2}$-spin $S_A$ sectors (with $2S_A + 1$ z-projections) obtained by mixing lower level subspaces. To speed up the matrix element computation, the $\frac{N}{2}$ basis is also sorted in blocks of decreasing $S_{A/B}$. Each of those blocks is further divided into sectors of $2S_A + 1$ containing consecutive z-projection of each unique $S_A$ sector. Details of
the implementation is provided in Appendix B.

The overall matrix can be constructed column-by-column and term-by-term and is extremely sparse. The practical implication is that the grand problem is decomposed into many small independent computations which can be massively parallelized (See Appendix B). Table 3.1 compares the $C_s = \frac{1}{4}$ basis size to those of other bases commonly used by non-iterative methods. At $C_s = \frac{1}{4}$, the $N = 32$ system was solved by a single CPU within seconds using a generic Lanczos routine.

To convince the critics, the single-processor computation was modified to calculate an explicit wavefunction for $N = 64$ to break the N=40 limit which required computer clusters. The numerical vector was about 8GB in size, and the matrix was too large for the physical memory of a single workstation. The remedy was the use of one processor core in fetching data from fast hard drive storages while another core performs floating-point operations. Starting from the $C_S = \frac{1}{8}$ solution, the GS vector was obtained after 15 matrix-vector products, each taking less than 30 minutes without any optimizations.

Figure 3.2: (a) Overlap $|\langle GS, C_s | GS, C_s + \frac{2}{N} \rangle|^2$ between ground-state wave functions corresponding to adjacent values of $C_s$ for $N = 16, 32, 64$. Note that for $N = 64$, the overlap is already 95% for $C_s = \frac{1}{4}$; (b) Fractional change $\frac{1}{E_{GS}} \frac{dE_{GS}}{dC_s}$ in ground-state energy corresponding to different $C_s$ cutoffs. The lines are linear fits to the data.
Figure 3.3: GS energy computed at specific $C_s$ values (symbols) and the extrapolation to $C_s = 1$ (solid lines) for $N = 16, 32, 64$. The horizontal dashed lines are the lowest values from Ref. [37]. Inset: expectation values of the non-commuting operators $\sqrt{\langle m^2 \rangle}$ and $\langle S_r \cdot S_{r+\left(\frac{L}{2},\frac{L}{2}\right)} \rangle$.

3.4 Results

We have performed computations for the full basis for $N = 16$ as a benchmark for larger $N$ values. For $N = 32$ (64), we computed up to $C_s = \frac{1}{2}$ ($C_s = \frac{1}{4}$). While the computations with $C_s < 1$ are variational, the goal here is to demonstrate that $C_s \sim \frac{1}{4}$ is a good rule-of-thumb value to capture well the AFM ground-state.

The stability of this formulation is demonstrated in Fig. 3.2(a) which shows the overlap between the GS wave functions for consecutive values of $C_s$. The deviation decreases exponentially with increasing $C_s$, and for $C_s = \frac{1}{4}$ the overlap is $\approx 95\%$. We conclude that the GS vector is already pointed in the correct direction even for small $C_s$, and subsequent increments of $C_s$ merely result in minor improvements. This is reinforced by the convergence of the GS energy demonstrated in Fig. 3.2(b), which supports the scaling law $\frac{dE_{GS}}{dC_s} \sim e^{-\alpha C}$. The error decays exponentially, with a rate that increases with $N$. The $N = 32, 64$ lines cross at $C_s \sim \frac{1}{4}$ where the fractional error is $\sim 10\%$. This exponential efficiency is amplified by the fact that a linear decrease in $C_s$ leads to a combinatorial decrease of the basis size because of the numerous low-spin combinations removed from the basis.

These GS energies are not as accurate as for established iterative methods;
however, a single-pass computation at $C_s = 1/4$ already has less than 2% error. Using a linear fit of $\frac{dy}{dx} \sim e^{-x}$, an estimate of $E_{GS}$ at $C_s = 1$ is obtained by simple integration (see Fig. 3.3). From the $C_s = 1$ estimates and the $N^{-\frac{3}{2}}$ scaling, we extrapolate $\lim_{N \to \infty} \frac{E_{GS}}{N} = -0.6671$, within 0.5% of the best published value of -0.6692[37], even though it is achieved at a significantly reduced computational cost.

We have thus far demonstrated the exponential convergence of the GS eigenvalue and eigenvector with increasing $C_s$, which validates our proposed variational basis. While a systematic approximation of the wavefunction can provides estimates of all quantities – more accurately so for those commuting with the Hamiltonian – derivable from wavefunctions, our approach is not directly competitive with established methods in evaluating particular correlation functions of interest. For example, Monte-Carlo sampling is inherently one of the best, if not the best, way of evaluating numerical integrals. The undoped system is predominantly characterized by correlation operators which do not commute with the Hamiltonian. The insets of Fig. 3.3 show expectation values of two such operators (the staggered magnetization of Eq. 3.12 and the spin-spin correlation at maximum distance) evaluated from our eigenvector for a given value of $C_s$. Although these operators are non-commuting, their matrix representation is diagonal in our basis and therefore trivial to compute. The convergence of these quantities is essentially linear for $C_s < \frac{1}{4}$. For greater values of $C_s$, $\frac{d}{dC_s} \langle \vec{S}_r \cdot \vec{S}_{r+\frac{1}{2}} \rangle$ is suppressed faster than $\sim e^{-C_s}$ so the post-threshold convergence is even better than for the GS energy. On the other hand, because $m$ is exploited to discard low-spin parts of the Hilbert space, its approximated value is always higher than the true value; convergence is only asymptotical for $C_s > \frac{1}{4}$. Thus, the computational discount for the GS wavefunction at a small $C_s$ is achieved at the cost of lowered accuracy for one of many definitions of staggered magnetization. As discussed in Sec. 3.1, the polaronic description of doped cuprate systems is difficult due the non-bosonic nature of the undoped background. The ultimate goal of the approach is to model the wavefunction of the AFM background with and without doped hole. The application to lightly doped AFM system will be demonstrated in Ch. 4 and 5.
3.5 Conclusions

In summary, one dominant characteristic of the AFM background is that, within each bipartite sublattice, $\frac{N}{8} \text{ spins add to a total spin of zero while the other } \frac{3}{8}N \text{ add to the maximum value of } \frac{3}{16}N$. The octa-partition provides a natural, efficient means of exploiting this knowledge in modeling the background.

In particular, $C_s$ was introduced to systematically parameterize the resulting Hilbert space, and the $\frac{N}{8}$-spin idea was used as an apriori argument in postulating that $C_s \sim \frac{1}{4}$ is the minimal value that captures the GS accurately. The resulting basis is combinatorially smaller than in any other schemes, apriorily known, and computationally very efficient. Numerical results confirmed the postulation. The stability of the formulation was demonstrated in Fig. 3.2(a) and the sources of error were identified. Using a single thread on a commodity computer, explicit calculation of the AFM wavefunction is feasible for systems with up to $N = 64$ spins, significantly exceeding the current full ED record of $N = 40$[121].

This breakthrough was achieved without transforming the problem away from the real-space basis of the square lattice with periodic boundary conditions, in which other interactions are defined naturally. Neither translational nor point-group symmetries are exploited. Thus, it can be used as an excellent starting point for studying doped models where these symmetries are broken, or as an efficient kernel for iterative methods, such as renormalization and quantum cluster theory[125]. Of course, this insight might also be adapted in devising more efficient variational Monte-Carlo formulations.

From a holistic point of view, the $\frac{N}{8}$-spin observation contributes to the description of the AFM background in analogy to the $b^-|0\rangle = 0$ statement taken for granted by a phonon problem. Excited states analogous to $b^+|0\rangle$ are indeed important when carriers are added to the system. Because the total spin of the AFM plus carriers is conserved while AFM order is expected to persist for low doping at low temperature, arguments such as Eq. 3.12 can still shed light on the AFM phase space spanned by the solution. This will be explored in Ch. 4 and 5.
Chapter 4

Spin Polaron on the Copper-Oxygen Plane

4.1 Introduction

It is widely believed that a proper description of holes in a spin-$\frac{1}{2}$ 2D antiferromagnet (AFM) with full quantum fluctuations could provide the answers to many unsolved puzzles in cuprates. Consideration of more exotic scenarios [44, 45, 48, 126] are indeed exciting developments; however, a detailed modeling of the hole+AFM is a crucial first step to understanding the significance of such additions. This problem is difficult because of the complicated nature of the 2D AFM background, whose quantum fluctuations in the presence of extra holes were never fully captured for a large CuO$_2$ lattice.

As discussed in Ch. 1, the simplest model required to describe a CuO$_2$ plane is the three-band $p-d$ model whose unit cell consists of one $d_{x^2-y^2}$ and two $2p_{\sigma}$ orbitals[14, 15]. For copper sites located at $l$, some integral multiples of the lattice parameters $a_x = (a,0)$ and $a_y = (0,a)$, the oxygen sites are located at $l + \epsilon_{x/y}$, where $\epsilon_{x/y} = \frac{1}{2}a_{x/y}$. Since there are more electrons than vacancies in the problem, the model is specified with fermion operators $p_{l+x,\sigma}^\dagger$ and $d_{l,\sigma}^\dagger$ creating holes of spin $\sigma$ on a vacuum with one hole per unit cell. For the orbital setting illustrated in Fig.4.1a, the model reads
\[ H_{3B} = T_{pd} + T_{pp} + \Delta_{pd} \sum n_{l+\varepsilon,\sigma} + U_{pp} \sum n_{l+\varepsilon,\uparrow} n_{l+\varepsilon,\downarrow} + U_{dd} \sum n_{l,\uparrow} n_{l,\downarrow}, \quad (4.1) \]

with
\[
T_{pd} = t_{pd} \sum (-p_{l+\varepsilon,\sigma}^\dagger + p_{l-\varepsilon,\sigma}^\dagger) d_{l,\sigma} + h.c.
\]
\[
T_{pp} = t_{pp} \sum s_{\delta} p_{l+\varepsilon+\delta,\sigma}^\dagger p_{l+\varepsilon,\sigma} - t'_{pp} \sum p_{l-\varepsilon,\sigma}^\dagger + p_{l+3\varepsilon,\sigma}^\dagger p_{l+\varepsilon,\sigma}
\]
\[
n_{l,\sigma} = d_{l,\sigma}^\dagger d_{l,\sigma}
\]
\[
n_{l+\varepsilon,\sigma} = p_{l+\varepsilon,\sigma}^\dagger p_{l+\varepsilon,\sigma}.
\]

\(T_{pd}\) is the direct oxygen-copper hopping. \(T_{pp}\) contains the direct \(t_{pp}\) oxygen-oxygen nearest-neighbor hopping as well as the \(t'_{pp}\) oxygen-oxygen next-nearest-neighbor hopping mediated by the copper 4s orbitals. The sign of the matrix elements are determined by the phases of the overlapping orbital wavefunctions in Fig.4.1a. In particular, for upper-right/lower-left (upper-left/lower-right) O-O hopping, \(s_\delta = 1\) (\(s_\delta = -1\)).

\(\Delta_{pd}\) is the charge transfer energy, which is the energy cost of the \(d^9 \rightarrow d^{10}\) Ligand hole excitation. In reality, this energy decreases with the distance between the \(d^{10}\) site and the \(\bar{L}\) Ligand hole, and a nearest neighbor Hubbard interaction should be added to the three-band model to capture this deviation. However, the following procedure is sensitive only to the local charge transfer energy, and such inclusion would merely lead to minor variations of effective parameters with no significant effects on the results, at least in the single hole scenario. \(U_{dd/\text{pp}}\) are repulsive Hubbard interactions. Although one issue in cuprate physics is the lack of significantly small parameters, the magnitude of parameters have been established to be \(t_{pp} < t_{pd} < \Delta_{pd} < U_{pp} < U_{dd}\)\([5]\). In practice, neighbor Coulomb interactions such as \(U_{pd} d^3 \tilde{d} p^\dagger p\) should be considered; however, such considerations do not add any new matrix elements into the effective Hamiltonian of interest (Sec. 4.2.1).

This three-band model has not been solved exactly. Unbiased numerical solution has also been limited due to the combinatorially large Hilbert space and the sign problem in quantum Monte-Carlo integration.

To make progress, this chapter derives an effective spin polaron model which
Figure 4.1: (a) Two adjacent unit cells of the CuO$_2$ plane. The orbitals kept in the 3-band model of Eq. (4.1) are shown, with white/shaded for positive/negative signs. The two $\epsilon$ vectors (solid arrow) and the four $\delta$ vectors (dashed arrow) are also shown. (b) Sketch of a virtual process of $T_{\text{swap}}$.

captures the anion–cation nature of the CuO$_2$ plane in the context of superexchange. The model is then solved exactly in a 32-unit-cell CuO$_2$ cluster with the aid of the optimization in Ch. 3.

4.2 The Microscopic Model

4.2.1 Derivation of the Spin Polaron Model

Noting that all $T_{pd}$ virtual processes in Eq. 4.1 increase energy by either $U$ and/or $\Delta$, and that high-order terms in $t_{pp}$ always come with an even power of $t_{pd}$, we derive an effective Hamiltonian for the states $p_{l+\epsilon,\sigma}^\dagger \prod l |\sigma_l\rangle$ using degenerate Rayleigh-Schroedinger perturbation theory[127]

$$H_{eff} = P_{gs} [H + T \frac{1 - P_{gs}}{E_0 - H_0} T + \cdots] P_{gs}$$

$$P_{gs} = \prod l \sum \sigma d_{l,\sigma}^\dagger d_{l,\sigma}$$

using $P_{gs}$ as the projector into this subspace of degenerate states with all $N$ Cu spins intact. The resulting Hamiltonian has the form

$$H_{eff} = T_{pp} + T_{\text{swap}} + H_{Jpd} + H_{Jdd}$$

(4.2)
where $T_{pp}$ is the oxygen-oxygen hopping in Eq. 4.3. The most straightforward derivation can be given by considering the $U_{dd} \rightarrow \infty$ limit. The second-order $t_{pd}$ virtual processes shown in Fig 4.1b lead to an additional spin swap O-O hopping term:

$$T_{swap} = -t_{sw} \sum s_{\eta} p_{l+\epsilon+\eta,\sigma} p_{l+\epsilon,\sigma'} |\sigma_{l,\eta}^{\prime}\rangle \langle \sigma_{l,\eta}|$$

where $t_{sw} = t_{pd}^2 / \Delta_{pd}$ and $l_{\epsilon,\eta} = l + \epsilon + (\eta \cdot \epsilon / |\eta \cdot \epsilon|) \epsilon$. This comes about when a hole from a neighboring Cu site located at $l_{\epsilon,\eta}$ first hops to one of its other 3 hole-free neighbor O sites, followed by the original O hole falling into the now-empty Cu site. The end result is either NN O-O hopping if $\eta = \delta$ (Fig 4.1b), or NNN O-O hopping if $\eta = \pm 2 \epsilon$. In the latter case $s_{\eta} = 1$. For $\eta = \delta$, the sign $s_{\eta}$ is opposite to that of the direct NN O-O hopping, due to the phases of the $d_{x^2-y^2}$ orbital. As shown below, such processes do not simply renormalize $t_{pp}$ or $t_{pp}'$.

If the original Cu and O holes have antiparallel spins, it is also possible for the Cu hole to hop onto the occupied O orbital, followed by one of the O holes falling back on the empty Cu site. This leads to Heisenberg exchange:

$$H_{Jpd} = J_{pd} \sum \mathcal{S}_l \cdot \mathcal{S}_{l \pm \epsilon}$$

where $J_{pd} = 2t_{pd}^2 / [U_{pp} + \Delta_{pd}]$ and $\mathcal{S}_{l \pm \epsilon}, \mathcal{S}_l$ are respectively the spin operators at O and Cu sites.

Finally, there is superexchange between NN Cu holes, except for the bonds “blocked” by p-holes:

$$H_{Jdd} = J_{dd} \sum \mathcal{S}_{l_{\pm 2 \epsilon}} \cdot \mathcal{S}_l \Pi_{\sigma}(1 - n_{l \pm \epsilon, \sigma})$$

where $J_{dd} = 8t_{pd}^4 / [\Delta_{pd}^2 (U_{pp} + 2\Delta_{pd})]$. There are also processes involving a Cu hole hopping to an adjacent empty O and then back. These are analogous to lollipop diagrams and renormalize all the above terms by roughly equal amounts when all orders are included, hence we ignore them. We also discard a third-order $t_{pp} t_{pd}^2$ “Kondo-hopping” term for reasons discussed below.

While we find that the three-spin polaron (3SP) plays an important role, our approach is different from previous work [20–24] by recognizing (i) $T_{pp}$’s signif-
icant contribution to polaron coherence (as discussed in the next session), (ii) its complementing process $T_{\text{swap}}$ (Fig. 4.1b), and (iii) suppression of superexchange along the bond inhabited by the hole.

### 4.2.2 Polaron Coherence and Other Considerations

To illustrate the importance of the terms we kept, consider the GS of $H_{pd}$, which has energy $-J_{pd}$. It consists of two out-of-phase terms, each with the O hole plus a triplet of two $d^9$ Cu spin neighbors. This is the 3SP

$$|3SP\rangle = \sqrt{\frac{1}{6}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) |\sigma\rangle_h - \sqrt{\frac{2}{3}} |\sigma\sigma\rangle |-\sigma\rangle_h,$$

(4.3)

first studied by Emery and Reiter [25]. The other eigenvalues are 0 (Cu singlet plus O hole) and $J_{pd}/2$ (in-phase terms of Cu triplets plus O hole).

The Cu-Cu exchange in $H_{dd}$ is different from that of other models as it accounts for the lack of exchange between Cu spins bridged by an O hole. This lowers the 3SP by a further $\sim -J_{dd}$. Basically, the O hole activates an effective Cu-Cu FM coupling to frustrate the bi-partite description (which, however, is not a requirement for long-range AFM order [115, 128]). This is qualitatively different from the $t-J$ model, where a ZRS breaks 4 AFM bonds without frustrating the bi-partite lattice.

Due to the sign difference between $T_{pp}$ and $T_{\text{swap}}$, the effective hopping of the O hole is reduced to $t_{pp} - t_{sw}$ for a Cu-O triplet (with the “central” Cu at $l_{\epsilon,\eta}$), but is enhanced to $t_{pp} + t_{sw}$ for a singlet. The interference is completely kinetic in nature, and thus conceptually different from the ZRS mechanism which is due to intra-plaquette phase coherence [16]. The 3SP can also be written as a superposition of two Cu-O singlets: $|3SP\rangle \sim (|\uparrow,\sigma\rangle|\downarrow\rangle_h - |\downarrow,\sigma\rangle|\uparrow\rangle_h) + (|\sigma,\uparrow\rangle|\downarrow\rangle_h - |\sigma,\downarrow\rangle|\uparrow\rangle_h)$. Thus, the 3SP already stabilized by $J_{pd}$ and $J_{dd}$ is also expected to have an increased bandwidth because of $T_{pp} + T_{\text{swap}}$.

We have considered corrections to $H_{\text{eff}}$ numerically and have found that they do not affect our results. The reasons for this can be understood as follows.

A third-order term in the previous section and finite $U_{dd}$ yield an additional
"Kondo-hopping" term

\[
T_{\text{Kondo}} = -t_k \sum (-1)^{\eta_n} p_{l+\epsilon+\eta, \alpha}^\dagger p_{l+\epsilon, \alpha} (S_l+\epsilon \cdot S_{l+\eta} - \frac{1}{4})
\]

\[
t_k = \left( \frac{t_{pp}^2 \Delta}{\Delta (\Delta + U_{pp})} + \frac{t_{pd}^2}{(U_{dd} - \Delta_{pd})} \right)
\]

considered previously [19–24]. \(T_{\text{Kondo}}\) is non-zero only for a Cu-O singlet so it is merely a \(\sim O(t_{pd}^2/U_{dd})\) reinforcement to the above terms.

A finite \(U_{dd}\) also modifies matrix elements by some \(t_{pd}^2/(U_{dd} - \Delta_{pd})\) corrections. The effects are marginal because we scale the model in units of \(J_{dd}\) which itself contains such a correction. There are also fourth-order corrections to \(H_{\text{eff}}\). These are either reinforcement to \(H_{pd}\) or \(S^+S^-\) terms. These terms are small, especially against the coherence discussed above, due to the lack of the factor of 4 as in superexchange, and because their denominator involves \(U_{dd}\) and \(\Delta + U_{pp}\).

### 4.3 Results

Using \(t_{pd} = 1.3eV, t_{pp} = 0.65eV, t'_{pp} = 0.58t_{pp}, \Delta_{pd} = 3.6eV, \) and \(U_{pp} = 4eV\) [5], we scale the parameters in units of \(J_{dd}\) to find their dimensionless values to be \(t_{pp} = 4.13, t_{sw} = 2.98, \) and \(J_{pd} = 2.83\).

We push the computational limit to perform total-spin-resolved exact diagonalization (ED) of a topologically superior [129] cluster of \(N = 32\) CuO\(_2\) unit cells, treating the AFM background exactly. ED provides the transparency, flexibility, and neutrality to support new results. The price for a systematic mapping of the excited states is the limited cluster size, denying us access to higher \(k\)-space resolution. \(N = 16\) results are provided to check for size dependence.

All low-energy eigenstates have a total spin of either \(S_T = \frac{1}{2}\) or \(\frac{3}{2}\). The \(z\) projections for each \(S_T\) are degenerate. The \(S_T = \frac{1}{2}\) subspace is due to the \(s = \frac{1}{2}\) hole mixing with various \(S = 0\) background states, including the AFM GS, or mixing with the \(S = 1\) background states, including the "single-magnon" states. The \(s = \frac{1}{2}\) carrier can also mix with \(S = 1\) or 2 background states to yield the \(S_T = \frac{3}{2}\) subspace. The partition of the \(S_T^c\) subspace into separate \(S_T\) sectors was managed by the optimizations from Ch.3. Unlike there, no basis truncation was employed here.
Figure 4.2: a) Energy and b) quasiparticle weight (bottom) for the lowest eigenstates with $S_T = \frac{1}{2}$ and $\frac{3}{2}$ vs. momentum. Different sets are shifted so as to have the same GS energy.

for rigorous results. The $(k, S_T = \frac{3}{2}, S_z = \frac{1}{2})$ sector contains $\sim 0.44 \times 10^9$ states.

4.3.1 Energetics

Fig. 4.2(a) shows the lowest eigenenergies. The GS has $k = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $S_T = \frac{1}{2}$, and is consistent with the 3-spin polaron (Eq. 4.3). Remarkably, we find similar dispersion along $(0,0)\rightarrow(\pi,\pi)$ and $(0,\pi)\rightarrow(\pi,0)$ without having to add longer range hopping or fine-tune parameters as is needed in one band models. The biggest surprise, though, are the low-lying $S_T = \frac{3}{2}$ states which go below the $\frac{1}{2}$ states near $(0,0)$ and $(\pi,\pi)$. The solution’s robustness is supported by the difference $E_{\frac{3}{2}} - E_{\frac{1}{2}}$ which decreases with increasing $N$ at each $k$. We note that the magnon spectrum is gapped as $\sim 1/N$ for finite lattices [37]. Going from $N = \infty \rightarrow 32$, excitation energies increase at all $q$, e.g. from $0\rightarrow 0.28J$ at $q = 0$ and $2.4J\rightarrow 2.6J$ at the BZ.
Given this finite-size scaling of the “free magnon” energy, we find that $S_T = \frac{3}{2}$ states at $(0,0), (\frac{\pi}{4}, \frac{\pi}{4}), (\frac{3\pi}{4}, \frac{3\pi}{4})$ and $(\pi, \pi)$ have much lower energy than $\min_q \{E_\frac{1}{2}(k-q)+\Omega_q\}$ for a $S_T = \frac{1}{2}$ state plus a free magnon. At other $k$-points, the $S_T = \frac{3}{2}$ states are within $\sim 0.2J$ of this value, so we cannot draw definite conclusions without access to larger $N$ data. It follows that the $S_T = \frac{3}{2}$ states are stable polarons at least in the regions marked by thick solid lines in Fig. 4.2(a). Thus, a $S_T = \frac{1}{2}$ quasiparticle cannot describe the low-energy states throughout the BZ.

4.3.2 Wavefunction Analysis

To compare the lowest energy states on both sides of the crossing, we note that the lowest $k_x = k_y$ eigenstates have odd parity upon a $\hat{P}_{x\leftrightarrow y}$ reflection (Fig. 4.1a) so they can be expressed as $2^{-1/2}(1 + \hat{P}_{x\leftrightarrow y}) \sum e^{ikl} p_{l+e_x,\sigma}^l |\sigma, l\rangle$. The band-crossing results in noticeable change in the expectation values of the correlation function:

$$\langle \hat{C}_x(\delta, a) \rangle = 2 \sum_{l,\sigma} \delta \cdot S_{l+\delta} n_{l+e_x,\sigma}$$

(4.4)

which measures the correlation between two Cu sites separated by a lattice constant at a certain distance away from the hole. $\langle \hat{C}_y \rangle$ is a reflection with $\hat{P}_{x\leftrightarrow y}$ for $k_x = k_y$. $\langle C \rangle$ ranges from $-\frac{3}{4}$ for singlet, to $\sim -0.33$ for 2D AFM GS, to $\frac{1}{4}$ for triplet.

Fig. 4.3a shows the $\langle \hat{C}_x \rangle$ correlations between neighbor Cu spins, when the hole is located at the darkly shaded bullet, in the GS: $k = (\frac{\pi}{2}, \frac{\pi}{2}), S_T = \frac{1}{2}$. The hole affects the AFM order in its vicinity. Because of the hole-spin exchange $H_{J_{pd}}$ and the blocked superexchange between the two Cu spins neighboring the hole, these “central” spins have triplet correlations, of $\sim 0.13$. Also, $\langle H_{J_{pd}} \rangle \sim -0.9J_{pd}$, showing that locally this is consistent with the 3SP solution (Eq. 4.3). More interesting are the correlations with the other 3 neighbors of each of these central Cu spins: with two of them, there are robust AFM correlations of $\sim 0.22$, while with the third the correlation nearly vanishes (lightly shaded bullet). This is counterintuitive if one views the system as a fluctuating Neél background, where a spin-flip would change the spin-spin correlation to all four neighbors. Although the two central Cu spins have $\frac{2}{3}$ weight in triplet configuration which is hardly bi-partite, long-range
AFM order cannot be automatically discounted \([115, 128]\). Indeed, the correlations we find are consistent with such order, except for the zigzag of 3 bonds shown by shaded bullets. This strange shape is dictated by the hopping mechanism. For a Bloch wave, oxygen-oxygen hole hopping in the upper-left/lower-right direction yields a phase shift of \(e^{i0}/e^{i(k_x-k_y)}\) and hence constructive interference if \(k_x = k_y\).

In contrast, hopping in the upper-right/lower-left direction yields a phase shift of \(e^{ik_x}/e^{-ik_y}\), so the interference is scaled down by \(\cos(k_x/k_y)\). For \(k_x = k_y = \frac{\pi}{2}\), having a mixture of singlets and triplets upper-left/lower-right to the O hole lowers energy with the least disturbance to AFM order. Thus, the two outside zigzag bonds are triplet “disturbance tails” pointing orthogonal to the momentum direction. This is different from the ZRS, which freezes a copper spin by the intraplaquette coherence of four oxygen sites.

Fig. 4.3b shows the correlation values when the \(S_T = \frac{3}{2}\) polaron becomes the lowest energy state at \((\pi, \pi)\). The results look similar at \((0,0)\). \(\langle H_{pd}\rangle\) remains \(~ -0.9J_{pd}\), but there are now four more heavily disturbed bonds. This further supports this being a stable polaron with an extra magnon locally bound close to the O hole. We stress here that this \(\frac{3}{2}\) polaron is formed by a spin disturbance around the 3SP. This is very different from the \(S = \frac{1}{2}\) excitation local to \(H_{pd}\) with energy \(+\frac{J_{pd}}{2}\)\([29]\).
Before we proceed further, let us note that the T=0 single-hole Green’s function
\[ \langle \text{AFM} | p e^{iHt} p^\dagger | \text{AFM} \rangle \]
contains information only of the states in the Krylov subspace which is the subspace spanned by the class of wavefunction \( H^n p^\dagger | \text{AFM} \rangle \), for all positive integers \( n \). It follows that the spectral weight of first electron removal \( Z(k) \), being the imaginary part of \( G(k, \omega) \), is exactly zero if the lowest hole state is not in the Krylov space of the electron removal state.

Fig. 4.2b shows the quasiparticle weight \( Z(k) \) for the first electron removal state. The major difference from other models is that \( Z(k) = 0 \) in three regions:
a) \( Z(0, 0) = Z(\pi, \pi) = 0 \) because here the lowest eigenstate has \( S_T = \frac{3}{2} \) which due to spin-conservation is not in the Krylov space of any \( S_T = \frac{1}{2} \) state, and b) \( Z(0, \pi) = 0 \) because the lowest eigenstate is not in the Krylov space \( H^n p^\dagger | \text{AFM} \rangle \). The t-t’-t”-J model restricts a spin degree of freedom, leading to \( Z(0, 0) \sim 0.1 \) and a finite \( Z(\pi, \pi) \) [18] instead of zero as dictated by a \( \frac{3}{2} \) polaron. Unlike the t-t’-t”-J model [18], \( Z(k) \) is concave down along \( (\pi/2, \pi/2) \rightarrow (0, \pi) \) in closer agreement with ARPES which measured a maximum between \( (\frac{\pi}{2}, \frac{\pi}{2}) \) and \( (\frac{\pi}{4}, \frac{3\pi}{4}) \) [52, 53]; however, this observation might be sensitive to finite-size effects. Our \( Z(k) \) is smaller than that of the t-J model, suggesting less “free particle” nature of the polaron.

The lowest energy state at \( k = (0, \pi) \) has \( S_T = \frac{1}{2} \), but the spectral weight of first electron removal state is exactly zero. This is confirmed by solving for the lowest energy state in the full basis, the Krylov subspace of \( \sum e^{i kl} p^\dagger_{l+\xi} | \text{AFM} \rangle \) and Krylov subspace of \( \sum e^{i kl} p^\dagger_{l+\xi} | \text{AFM} \rangle \). The \( k = (0, \pi) \) lowest state of the two latter cases are at least 0.006J higher than the first. The lowest energy state is not in the Krylov space of electron removal states so \( Z \) is exactly zero. This seems to be due to symmetry, although we do not yet fully understand this. Their close existence above the \( Z=0 \) lowest state may be related to pseudogap phenomena in this region; however, this needs to be investigated in more detail. Fig. 4.4 shows the correlation for the lowest eigenstate. Compared to the GS (see Fig. 4.3a), there are two more disturbed bonds as required by the reflection parity about the momentum direction. This larger disturbance range is accompanied by more negative (AFM) correlation values.
4.3.3 Further Experimental Implications

Although we are restricted to rather low momentum resolution, more can be said about the \( E_{\frac{3}{2}} - E_{\frac{1}{2}} = 0 \) band crossings in the nodal direction by looking at the \( k \) points between which the difference switch signs. The observation in Fig. 4.2a is that, going away from the \( (\frac{\pi}{2}, \frac{\pi}{2}) \) GS, \( E_{\frac{3}{2}} - E_{\frac{1}{2}} \) is larger towards \((0,0)\) than towards \((\pi, \pi)\). The \( \frac{1}{2} \rightarrow \frac{3}{2} \) band crossing would induce an abrupt change in \( Z(k) \) from non-zero to exactly zero, irrespective of the \( Z(k) \) value on the finite side. The larger \( E_{\frac{3}{2}} - E_{\frac{1}{2}} \) towards \( k = (0,0) \) would suggest the non-zero region is larger towards \( k = (0,0) \) than towards \( k = (\pi, \pi) \). Fig. 4.2a also shows that the \( \frac{3}{2} \) states get pushed further down as \( N \rightarrow \infty \) so the crossing is expected to be closer to the \( (\frac{\pi}{2}, \frac{\pi}{2}) \) GS. This is consistent with ARPES which indeed observed an abrupt peak suppression in the nodal direction as well as the peaks surviving longer towards \( k = (0,0) \) \([52, 53]\).

Even when the \( S_T = \frac{3}{2} \) states are not lowest in energy, they hug the \( S_T = \frac{1}{2} \) band. This provides a \( \lesssim J_{dd}/2 \) energy scale for spin excitations. At finite \( T \), as magnons become thermally activated, these \( \frac{3}{2} \) states become “visible” to ARPES. This suggests a \( T \)-dependent broadening mechanism of \( \lesssim J_{dd}/2 \) scale. Coincidentally, this is the same as the energy scale linked to phonons \([43, 45, 54]\).

Recent neutron experiments on samples at higher doping reveal \( \sim 50 \text{ meV} \) magnetic response centered at \( q = 0 \), away the AFM resonance momentum \([55, 56]\). The bottom of the single-particle band structure in Fig 2(a) indeed has a
$q = 0 \frac{1}{2}$-to-$\frac{3}{2}$ excitation of this energy scale (the computed energy gap is larger than the $\frac{1}{N}$ S=1 gap for N=32 lattice). It has also been pointed out that the $q = 0$ magnetic excitation can be explained by ordering multiple spins on oxygen sites in the higher-doping scenario [130].

In addition to the \(3/2\) band, there are internal excitations of the local 3SP since \(H_{pd}\) also has a \(S = \frac{1}{2}\) doublet and a \(S = \frac{3}{2}\) quartet separated in energy by \(J_{pd}\) and \(3J_{pd}/2\) from the 3SP. Magnetic excitations at these energy scales have been observed via inelastic resonant x-ray scattering for highly doped samples [131].

### 4.4 Conclusions

We derived and solved a model which includes the O sites and takes full account of the AFM quantum fluctuations, for large \(N = 32\) clusters. The phases of the \(p\) and \(d\) orbitals lead to phase coherence via \(T_{pp} + T_{\text{swap}}\). This is reenforced by \(H_{pd}\) and the blocking of the AFM superexchange, making corrections such as \(T_{\text{Kondo}}\) negligible. The dispersion is similar to that measured by ARPES; however, the lifting of the Cu-O singlet restriction present in ZRS-based models leads to wave functions of a different nature, namely the 3SP where the O hole correlates with both its neighbor Cu sites. This model also provides low-energy channels for \(S = 1\) excitations. \(Z(k)\) was found to be identically zero in certain regions of the BZ for two reasons: (1) the spin-$\frac{3}{2}$ of the lowest states close to \((0,0)\) and \((\pi,\pi)\); and (2) around the antinodal region because of the lowest state there being exactly orthogonal to the single electron removal state.
Chapter 5

Two Holes in the Spin Polaron Model

5.1 An Intuition about the Doped Antiferromagnet

The work presented in Ch. 3 and Ch. 4 was actually motivated by the drive to study multiple holes injected into the CuO$_2$ layer. The Hilbert space of the doped cuprate is combinatorially large. Brute-force Monte-Carlo simulations are limited by the Fermion sign problem, and zero-temperature computations have been carried out with variational Monte-Carlo and NRG, which employ various approximation or special boundary conditions. The largest two-hole system studied by an unbiased method is the N=32 t-t'-J model. Because each additional hole removes a spin instead of adding one in this one-band model, Leung et. at. managed to use exact diagonalization to show that a realistic value of t' destroys short-range hole-hole attraction[132]. The result contradicts other studies using variational or mean-field approaches which predict binding. The discrepancy between this exact N=32 solution and approximated solutions on larger lattices has been attributed to a finite size issue.

As discussed in Ch. 4, the solution of the spin polaron model (Eq. 4.2) solved in a cluster with 32 Cu and 64 O orbitals reveals properties completely missed by previous studies which employed various approximation or small clusters. In fact, 12 unit cells were required to illustrate the difference between the $S_T = \frac{1}{2}$ and
The two-hole solution for a larger cluster could be different from previous studies limited to N=4[24].

I realized early that, even after exploiting translational and spin-projection symmetries, the two holes injected into the N=32 spin polaron model would have a Hilbert space with \(0.154 \times 10^{12}\) states, and larger N would involve literally an astronomical number of states. A system of this size challenges the capability of unbiased methods, and such an endeavor is extremely high-risk, especially considering the many unknowns that come with the novelty of the spin polaron model. With the drive to solve for the zero-temperature ground state with 2-dimensional periodic boundary condition, I came up with a rather radical plan. In the \(N \to \infty\) limit, two additional holes should not be able to drastically change the undoped AFM groundstate; therefore, the compact yet systematic way of modeling AFM in Ch. 3 should be a stable, systematic way of modeling the two-hole scenario. The octapartite scheme with \(C_S\) truncation discussed in Ch. 3 was formulated under this belief. Its first application has not been publicized, but the \(C_S = \frac{1}{4}\) truncation was crucial in the development of the work in Ch. 4. The formulation enabled an approximate solution to Eq. 4.2 for N=32 within seconds, allowing me to scan the parameter space extensively and note various properties missed by previous studies. Nevertheless, full exact diagonalization was performed to prove the new findings are not mere numerical artifacts.

The convergence of the one-hole ground state is illustrated in Fig. 5.1. The energy computed for \(C_S = \frac{1}{4}\) and \(\frac{1}{2}\) is within 3.6% and 0.5%, respectively, of the exact value. Table 5.1 shows the exact wavefunction’s probability in specific subspaces \(S_A \otimes S_B\) of sub-lattice total-spin \(S_{A/B}\). While the \(C_S = \frac{1}{4}\) truncation captures \(\sim 95\%\) of the undoped wavefunction as discussed in Ch. 3, Table 5.1 shows that \(C_S = \frac{1}{4}\) captures \(\sim 73\%\) of the one-hole groundstate. The next two increments of \(C_S\) contains 17% and 6%, respectively, of remaining weight. The addition of a hole couples the spin background to adjacent values of \(S_A \otimes S_B\), and states added by increasing \(C_S\) have decreasing importance in the wavefunction; therefore, it is reasonable to expect an increment of \(\sim dC_S \sim \frac{1}{N}\) from \(C_S = \frac{1}{4}\) to suffice. At least in the very dilute limit, my postulate holds, and the formulation in Ch. 3 does provide a good starting point to systematically capture the low energy state. These observations provide reasonable merits for the application of the scheme to
Figure 5.1: Convergence of the one-hole ground state. (left) GS energy calculated for increasing $C_S$. The value approaches rapidly to the exact value marked by the dotted line. (right) fractional change in GS energy for the next increment of $C_s$. Solid lines are extrapolation and linear fits.

Table 5.1: (N=32) The $S_1 = \frac{1}{2}$ $k = (\frac{\pi}{2}, \frac{\pi}{2})$ single-hole ground state’s weight in subspaces of particular $S_A \otimes S_B$. Numbers are percentage adding up to 100%. Two lines are drawn to highlight the $C_S = \frac{1}{4}$ truncation which discards states with $S_A, S_B < 6$ and yield an energy within 3.6% of the exact value.
the two-hole scenario, whose groundstate should be captured in the subspace of 
\( C_S \sim \frac{1}{4} + O(dC_S) \) for \( N \to \infty \). To provide a conservative error analysis though, I 
would aim for a capability of up to \( C_S \sim \frac{1}{2} \), which in turn limits the cluster size to 
\( N=32 \).

This chapter reports the application of the octapartite scheme discussed in Ch. 
3 in solving the two-hole scenario in the spin polaron framework discussed in Ch. 
4. In particular, two-body interactions are derived for the two-hole scenario. The 
Hilbert space formulation and truncation is then documented in detail. Numerical 
results are presented. Due to the novelty of the model and the numerical scheme 
as well as the lack of any exact solution, many details remain open questions and 
this effort is on-going at the time of writing. Nevertheless, the numerical approach 
proves to be successful in capturing doped antiferromagnet and the solution shows 
enhanced singlet correlations.

5.2 The Spin Polaron Model for Two Holes

An effective model for two holes interacting with an antiferromagnetic background 
can be derived using the same procedure discussed in Sec. 4.2.1. When hopping 
vanesishes in the three-band model (Eq. 4.1), the two-hole GS is 
\[ p_{l+\pm,\sigma}^{\dagger} p_{l'+\pm',\sigma'}^{\dagger} \prod |\sigma_{l''}\rangle, \]
\[ \text{with a degeneracy of } \frac{1}{2} N^2 2^{N+2}. \]

The Rayleigh-Schrodinger expansion in Sec. 4.2.1 operates in the \( N+2 \) hole sector of the Fock space, and the outcome is slightly dif-

ferent from the single-hole scenario. In particular, the projector used in the expan-
sion becomes 
\[ P_{2h} = \prod (1 - n_{l+\pm,\sigma} n_{l'+\pm',\sigma'}) \prod (n_{l,\uparrow} + n_{l,\downarrow} - 2n_{l,\uparrow} n_{l,\downarrow}), \] 
which allows only states with a full lattice of copper spins and no double-occupancies 
(due to \( U_{pp/dd} \)). The Hamiltonian can be written as two parts 
\[ H_{eff} = P_{2h} H_1 P_{2h} + H_2, \]
where \( H_1 \) is Eq. 4.2 derived for the single-hole scenario and \( H_2 \) is the correction 
when the two holes are close to one another.

The idea here is to start with the established single-hole scenario and correct
$P_{2h}H_1P_{2h}$ from second- to forth-order with $H_2 = H_2^{(2)} + H_2^{(3)} + H_2^{(4)}$ in the presence of a second hole. In the Rayleigh-Schrodinger expansion with projector $P_{2h}$, the appearance of $(1 - P_{2h})$ dictates that all terms in $H_2$ must have an even power of $t_{pd}$ because the final states must have one copper spin per unit cell. Because $H_2$ accounts for only 2-hole corrections, one can deduce that all terms in the second-order $H_2^{(2)}$ share a factor of $t_{pp}^2$. All terms in the third-order $H_2^{(3)}$ share a $t_{pp}^2t_{pd}$ factor because the second step of any three-step $t_{pp}^3$ process would yield states projected out by $(1 - P_{2h})$ in the perturbation. The forth-order $H_2^{(4)}$ can have terms with a prefactor of $t_{pp}^4$, $t_{pp}^2t_{pd}$ and $t_{pp}^4$.

The second order corrections are the bare hole-hole correlations.

$$H_2^{(2)} = -\frac{2t_{pp}^2}{U_{pp}} \sum (-1)^{\delta_1 \cdot \delta_2} \left( \frac{\mathbf{S} \cdot \mathbf{S}}{4} - \frac{t_{pp}^2}{2} \right) \alpha' \beta' \alpha P_{l+\epsilon+\delta_1, \alpha} P_{l+\epsilon+\delta_1, \alpha} P_{l+\epsilon+\delta_1, \beta} P_{l+\epsilon+\beta}$$

(5.3)

Here, $\delta_{1,2}$ sum over $(\pm \epsilon_x, \pm \epsilon_y)$. The matrix element is non-zero only if the two holes are $|\delta_1| = \frac{a}{\sqrt{2}}$ apart. The sign of the matrix element is positive when $\delta_1 \cdot \delta_2 = 0$. There are 8 non-zero matrix elements in this situation. Two of these correspond to $\delta_1 + \delta_2 = 0$ so the static AFM exchange is $2 \times \frac{2t_{pp}^4}{U_{pp}}$. There are also 6 other ways for one hole to “skip” over the other with such a Heisenberg factor. The factor of $\frac{1}{4}$ plays an important role here in setting the matrix element to zero whenever the two oxygen holes are in a triplet configuration.

There is an abundance of terms in third- and forth-order corrections which are a sub-set of terms studied in the literature of 2-band models[20–24]. To provide a simple physical picture, we consider terms that are greater or equal to $J_{dd}$, which is roughly $\frac{8 \cdot 16 \cdot t_{pp}^4}{31 \cdot U_{pp}} \sim \frac{16 \cdot t_{pp}}{31 \cdot t_{pp}}$ for $t_{pd} \sim 2t_{pp}$ and $\Delta_{pd} \sim U_{pp} \sim \frac{t_{pp}}{3}$. The basis of this approximation is the observation that the dominant short-range physics should be already captured by the numerous low-order terms with an energy scale of $t_{pp}$ and $T_{\text{swap}}, J_{pd}, J_{pp} \sim 0.66t_{pp}$, which are already greater than the long-range physics at the scale of $J_{dd} \sim 0.2t_{pp}$. Adding the relevant short-range corrections with magnitude greater than $0.2t_{pp}$ should be more than adequate to capture the physics. We also discards terms that can be factored into $t_{pp}, T_{\text{swap}}, J_{pd}, J_{pp}$ and the identity processes, which merely contribute some combinatorial constants scaling in the
infinite-order perturbation.

As discussed above, all third-order corrections have a prefactor of $t_{pp}t_{pd}^2$. The perturbation goes through two virtual states so the largest possible magnitude is

$$t_{pp}t_{pd}^2 \sim \frac{t_{pp}t_{pd}^2}{\Delta_{pd}U_{pp}} \sim \frac{1}{9}t_{pp}.$$  

The splitting due to such a pair of Hermitian matrix elements is $\frac{2}{9}t_{pp}$ which is the same as the $J_{dd}$ splitting. Evidently, all these processes can be factored into consecutive low-order processes and are discarded. Because $U_{pp} \sim \Delta_{pd}$, other processes would have denominators that are at least a factor of two greater; that is, constructive quantum interference is required for any terms to be non-negligible compared to the superexchange. Constructive interference among $t_{pp}t_{pd}^2$ processes requires the same orbital occupancy in the initial and final state, and this can happen only if the two oxygen holes are $\delta = (\pm e_x, \pm e_y)$ apart.

The transition of interest is then

$$p_{\pm e_x}^\dagger \alpha p_{\pm e_y}^\dagger \beta d_{i,y}^\dagger \rightarrow p_{\pm e_x}^\dagger \alpha p_{\pm e_y}^\dagger \beta d_{i,y}^\dagger.$$  

The effective matrix elements have the form $\langle \alpha', \beta', \gamma' | H_2^{(3)} | \alpha, \beta, \gamma \rangle$, a local three-spin ring involving the copper spin sandwiched by the oxygen holes. The correction term can be expressed as a summation over each copper spin with two additional vector $\Delta_{x/y}$ summed over $\pm \varepsilon_{x/y}$. Noting that all 3-step hopping would give an overall phase of $-1$ (Fig. 4.1) and virtual states have double-occupancy, it’s not surprising that all operators contain a shifted Heisenberg factor.
Therefore we can approximate this term by raising the energy of oxygen-oxygen singlet pair and 0 for triplet pair. The last four terms give an eigenvalue of \( J_{\text{dd}} \). No constructive interference is possible. We set 

\[
H_2^{(4)} = \frac{2t_{pd}^2}{U_{pp}(\Delta_{pd} + U_{pp})} \sum \left( \begin{array}{c}
\langle S \cdot S - \frac{1}{4} \rangle \beta' \alpha' \beta_\alpha \sum p_{l+\Delta, \beta} p_{l+\Delta, \alpha} \gamma p_{l+\Delta, \beta} | l, \alpha' \rangle \langle l, \gamma | \\
+ \langle S \cdot S - \frac{1}{4} \rangle \alpha' \beta \beta_\alpha \sum p_{l+\Delta, \gamma} p_{l+\Delta, \alpha} \beta p_{l+\Delta, \beta} | l, \gamma \rangle \langle l, \beta | \\
+ \langle S \cdot S - \frac{1}{4} \rangle \gamma \alpha' \beta \beta_\alpha \sum p_{l+\Delta, \gamma} p_{l+\Delta, \alpha} \beta p_{l+\Delta, \beta} | l, \alpha \rangle \langle l, \gamma | \\
+ \langle S \cdot S - \frac{1}{4} \rangle \beta' \gamma \beta_\alpha \sum p_{l+\Delta, \beta} p_{l+\Delta, \alpha} \gamma p_{l+\Delta, \beta} | l, \beta \rangle \langle l, \gamma | \\
\end{array} \right)
\]

These terms are finite if there is a 2-spin singlet formation among the 3 spins. The first four terms give an eigenvalue of \( +4 \frac{t_{pd}^2}{U_{pp}(\Delta_{pd} + U_{pp})} \sim 0.2t_{pp} \) for oxygen-oxygen singlet pair and 0 for triplet pair. The last four terms give an eigenvalue of \( -3 \frac{t_{pd}^2}{U_{pp}(\Delta_{pd} + U_{pp})} \sim -0.06t_{pp} \) for singlet pairs and 0, -3 \( \frac{t_{pd}^2}{U_{pp}(\Delta_{pd} + U_{pp})} \sim -0.06t_{pp} \) for triplets. Therefore we can approximate this term by raising the energy of oxygen-oxygen singlet pair appropriately.

In the forth-order correction, all \( t_{pd}^4 \) processes can be factored into two \( T_{\text{swap}} \) or \( J_{pd} \) processes. The \( t_{pd}^2 \) processes are smaller than \( J_{dd} \) by a factor of \( \sim 4 \times 4 \) and no constructive interference is possible. \( 2t_{pp}^2 \) processes are even smaller. Therefore, we set \( H_2^{(4)} \sim 0 \). In summary, the 2-hole correction is in essence.
\[
H_2 \sim + J_{pp}^{(2)} \sum (-1) \delta_1 \delta_2 \left( \mathbf{S} \cdot \mathbf{S} - \frac{1}{4} \right) \alpha' \beta' P_{l+\epsilon+\delta_1}^\dagger \alpha' P_{l+\epsilon+\delta_1}^\dagger \alpha P_{l+\epsilon+\delta_1+\delta_2}^\dagger \beta P_{l+\epsilon+\beta} \\
- J_{pp}^{(3)} \sum \left( \mathbf{S} \cdot \mathbf{S} - \frac{1}{4} \right) \alpha' \beta' P_{l+\epsilon+\delta_1}^\dagger \alpha' P_{l+\epsilon+\delta_1+\delta_2}^\dagger \alpha P_{l+\epsilon+\beta} P_{l+\epsilon+\beta} 
\]

with \( J_{pp}^{(2)} = \frac{2t_{pp}^2}{U_{pp}} \) and \( J_{pp}^{(3)} = \frac{4t_{pp}^2}{U_{pp}(\Delta_{pd}+U_{pp})} \) for hole-hole interaction due to second and third order corrections.

### 5.3 Hilbert Space Formulation

The Hilbert space of the two-hole spin polaron model (Eq. 5.2 with Eq. 4.2 and Eq. 5.5) contains \( \left( \frac{2N}{2} \right)^2 \) states of the form \( P_{l+\epsilon}^\dagger \sigma P_{l+\epsilon}^\dagger \sigma' \prod |\sigma\rangle \), with \( l + \epsilon \neq l' + \epsilon' \). To perform computations for large \( N \), this basis must be reformulated to take advantage of translational symmetry, total-spin symmetry, total-spin-projection symmetry, and, most importantly, the truncation scheme discussed in Chapter 3.

Point-group symmetries were not exploited to reduce the basis. The Hilbert space formulation discussed below is non-trivial; the matrix transformation required to implement the basis and the truncation is a convoluted process. The code written for the project totaled \( \sim 0.5 \text{MB} \) in size, containing \( \sim 65536 \) characters. Software development of this scale is prone to minor but fatal mistakes such as a typo between “-” and “+” of adjacent keys on the keyboard. The point-group operators are used as a consistency check at various stages of the computation as discussed in Appendix A. Nevertheless, point-group symmetries are not compatible with all directions of total momentum \( K \) and the basis reduction is negligible to the combinatorial reduction by the octapartite scheme as shown in Table 3.1. Leaving this redundancy is a small price to pay for a consistency check.
5.3.1 Translational and Total-Spin Symmetries

First, I’ll define singlet and triplet creation operators for the two oxygen holes.

\[ s_{l,l'}^+ = \frac{1}{\sqrt{2}} (p_{l\uparrow}^+ p_{l'\downarrow}^+ - p_{l\downarrow}^+ p_{l'\uparrow}^+) \]

\[ t_{-1,l,l'}^+ = p_{l\downarrow}^+ p_{l'\downarrow}^+ \]

\[ t_{0,l,l'}^+ = \frac{1}{\sqrt{2}} (p_{l\uparrow}^+ p_{l'\downarrow}^+ + p_{l\downarrow}^+ p_{l'\uparrow}^+) \]

\[ t_{1,l,l'}^+ = p_{l\uparrow}^+ p_{l'\uparrow}^+ \]  

(5.6)

According to quantum mechanical angular momentum addition, \( N+2 \) spins can add up to a total spin \( S_T \) by mixing a two-hole singlet to a background with total spin \( S_{T,N} = S_T \) and also a two-hole triplet mixing with a background with total spin \( S_{T,N} = S_T, S_T \pm 1 \). Taking \( S_T^z = 0 \), any spin background can be specified orthonormally with respect to a position \( l \).

\[ \mathbb{D}_{l,l'}^\dagger |\alpha\rangle_I \equiv \begin{cases} 
  s_{l,l'}^+ |\alpha, S^z = 0\rangle_I \\
  \sum_{\ell = -1}^1 c(z, S_{T,N}, S_T) t_{\ell,l,l'}^+ |\alpha, S^z = -\ell\rangle_I 
\end{cases} 
\]

(5.7)

\( \alpha \) denotes a particular group of \( 2S_{T,N} + 1 \) spin configurations related by the total spin raising and lowering operators \( S_T^{+/-} = \sum_l S_l^{+/-} \), summed over all Cu sites. The total spin of this \( \alpha \) group can be \( S_{T,N} = S_T \) for two-hole singlet and \( S_{T,N} = S_T, S_T \pm 1 \) for two-hole triplets. Due to the choice of the overall projection \( S_T^z = 0 \), the two-hole singlet would mix only with backgrounds with \( S_{T,N}^z = 0 \). The two-hole triplets would mix with three different projections \( |\alpha, S^z = -1\rangle_I, |\alpha, S^z = 0\rangle_I \), and \( |\alpha, S^z = 1\rangle_I \) from the group \( \alpha \). The weight \( c(z, S_{T,N}, S_T) \) is the Clebsch-Gordon coefficients for mixing these three states with the three two-hole triplets to achieve a state with total spin \( S_T \) and \( S_T^z = 0 \).

Exploitation of the translational symmetry is performed by the use of a Fourier series of the form

\[ \sim \sum e^{ikl} \mathbb{D}_{l+e,l'+e}^\dagger |\alpha\rangle_I \]  

however, care must be taken to ensure orthonormality. Due to the commutation relation of the triplet and singlet (Eq. 5.6), the Fourier series is not straightforward.
for $\varepsilon = \varepsilon' = \varepsilon_{x/y}$ when both oxygen holes occupy $x$- or $y$-rung oxygen orbitals (see Fig. 4.1a). The specification of these two-hole configurations requires the two orthogonal periodic lattice vector $L_0$ and $L_1$ of length $\sqrt{N}$ for the 2D, N-unit-cell lattice. Defining

$$\delta l = (l_0' - l_0, l_1' - l_1), \quad (5.8)$$

most hole-hole configurations can be classified in the region

$$0 \leq \delta l_0 < \frac{L_0}{2}, \quad 0 < \delta l_1 < \frac{L_1}{2}$$

$$0 < \delta l_0 < \frac{L_0}{2}, \quad \delta l_1 = \frac{L_1}{2}$$

$$0 < \delta l_0 < \frac{L_0}{2}, \quad -\frac{L_1}{2} < \delta l_1 \leq 0$$

$$\delta l_0 = \frac{L_0}{2}, \quad -\frac{L_1}{2} < \delta l_1 < 0. \quad (5.9)$$

These states can be expressed using N-term Fourier series

$$|\square_{xx/yy}, \delta l, \alpha, K\rangle = \frac{1}{\sqrt{N}} \sum e^{iKl_0} \hat{\square}^{\dagger}_{l+\delta l_0, l+\delta l_1} |\alpha\rangle_l \quad (5.10)$$

Because the two oxygen holes are indistinguishable fermions (Eq. 5.6), there are three remaining $\delta l$ values which require special attention due to the periodic boundary condition.

$$\delta l_0 = \frac{L_0}{2}, \quad \delta l_1 = 0$$

$$\delta l_0 = 0, \quad \delta l_1 = \frac{L_1}{2}$$

$$\delta l_0 = \frac{L_0}{2}, \quad \delta l_1 = \frac{L_1}{2} \quad (5.11)$$

The number of terms in the Fourier series depends on the spin background translated by $\delta l$: $T_{\delta l} |\alpha\rangle_l$. If such a translation yields an orthogonal state, $\langle \alpha | T_{\delta l} |\alpha\rangle_l = 1$.
0, the Fourier series still has \( N \) terms. For the example of \( \delta l = \left( \frac{L_0}{2}, 0 \right) \),

\[
|\square_{xx/yy}, \delta l, \alpha, K\rangle = \frac{1}{\sqrt{N}} \sum_{l_1=0}^{L_z-1} e^{iK_{l_1}} \sum_{l_0=0}^{L_z-1} e^{iK_{l_0}} \sum_{l_{+e/\alpha}} e^{iK_{l_{+e/\alpha}}} \left( 1 + s_{\square} e^{iK_{l_{+e/\alpha}}} T_{l_{+e/\alpha}} \right) |\alpha\rangle_I \tag{5.12}
\]

where \( s_{\square} \) is the sign change due to hole-swapping in the singlet or triplet (\( \square_{a,b}^\dagger = s_{\square} \square_{a,b} \)). For the case of \( \langle \alpha | T_{\delta l} | \alpha \rangle_I = \pm 1 \), the above expansion makes clear that there can be only \( \frac{N}{2} \) terms in the Fourier series due to the term \( 1 + s_{\square} e^{iK_{l_{+e/\alpha}}} T_{l_{+e/\alpha}} \).

For this case the series has the form

\[
|\square_{xx/yy}, \delta l, \alpha, K\rangle = \sqrt{\frac{2}{N}} \sum_{l_1=0}^{L_z-1} e^{iK_{l_1}} \sum_{l_0=0}^{L_z-1} e^{iK_{l_0}} \sum_{l_{+e/\alpha}} e^{iK_{l_{+e/\alpha}}} \left( 1 + s_{\square} e^{iK_{l_{+e/\alpha}}} T_{l_{+e/\alpha}} \right) |\alpha\rangle_I \tag{5.13}
\]

The formulation for the scenario in which one hole occupies a \( p_{l+e_x} \) and the other occupies \( p_{l+e_y} \) is straight forward. All values of \( \delta l \) are unique and the Fourier series has the form

\[
|\square_{xy}, \delta l, \alpha, K\rangle = \frac{1}{\sqrt{N}} \sum e^{iK_{l}} \sum_{l_{+e_x/\alpha}} e^{iK_{l_{+e_x/\alpha}}} |\alpha\rangle_I. \tag{5.14}
\]

In summary, a full orthonormal Hilbert space can be specified by the states \(|\square_{xy}, \delta l, \alpha, K\rangle\) and \(|\square_{xx/yy}, \delta l, \alpha, K\rangle\). Translational symmetry is specified by the quantum number \( K \). Total spin and its projection are specified by the singlet/triplet nature of the two oxygen holes (\( \square \)) (singlet-triplet) in conjunction to the spin background \( \alpha \).

### 5.3.2 Enumeration

The reason behind the formulation in the previous section is to take advantage of the combinatorial truncation of the octapartite scheme discussed in Chapter 3. For a given completeness value \( C_S \) (see section 3.3), the number of different spin configurations \(|\square, \alpha\rangle\) (Eq. 5.7) is reduced to an integer \( n_{\text{SpinRow}} \). I would also define a number \( n_{2h} \) for the total number of oxygen-oxygen hole configur-
tions. For N=32, there are N=32 distinct δl’s for \( \square_{l+e_x,l+\delta l+e_y} \), 17 distinct δl’s for \( \square_{l+e_x,l+\delta l+e_y} \), and 17 distinct δl’s for \( \square_{l+e_x,l+\delta l+e_y} \). So \( n2h=66 \). The total number of states \( |\square_{x\text{y},\delta l,\alpha,K}\rangle \) and \( |\square_{x\text{y},\delta l,\alpha,K}\rangle \) is thus \( n2h\times n\text{SpinRow} \).

Because the matrix elements of the Cu-Cu superexchange \( \sum_{\ell,\ell',S} \delta_{\ell}\cdot\delta_{\ell'} \) is the same irrespective of the location of oxygen holes, the numerical basis was chosen such that an arbitrary wavefunction can be expressed in terms of a vector \( \vec{c} \) of complex numbers.

\[
|\Psi\rangle = \sum_{i=0}^{(n2h-1)} \sum_{j=0}^{(n\text{SpinRow}-1)} c_{n\text{SpinRow}\times i+j} |i,j\rangle
\] (5.15)

where \( j \) specifies a particular combination of singlet/triplet and background (Eq. 5.7) and \( i \) specifies the oxygen-oxygen configuration, i.e. \( x\text{-y}, x\text{-x}, y\text{-y} \) occupancies and separation \( \delta l \). Each \( |i,j\rangle \) corresponds to a Fourier series of a definite set of \( (\square, \delta l, K, S_T) \) as discussed in the previous section. This is an orthonormal numerical basis

\[
\langle i,j|j',j'\rangle = \delta_{i,j'} \delta_{j,j'}
\] (5.16)

\[
\sum_{m=0}^{n\text{SpinRow}\times n2h-1} |c_m|^2 = 1
\] (5.17)

Of course, the values of vector elements are the probability amplitudes of a quantum mechanical description. For example, the probability of the two oxygen holes in one of the \( n2h \) configurations is

\[
P(i) = \sum_{j}^{n\text{SpinRow}-1} |c_{n\text{SpinRow}\times i+j}|^2
\] (5.18)

### 5.4 Low-Energy Two-Hole Solutions

Exact diagonalization in the subspace of \( C_S = 0 \) to \( \frac{1}{2} \) was performed for the two-hole spin polaron model (Eq. 5.2 with Eq. 4.2 and Eq. 5.5), with parameters \( t_{pp} = 4.1293, t'_{pp} = 0, t_{sw} = 2.9822, J_{pd} = 2.8253, J_{pp}^{(2)} = 1.3420, \) and \( J_{pp}^{(3)} = 0.9182 \). \( t'_{pp} \) was set to zero to simplify the model. Another reason is that the \( J_{pp} \) corrections
due to $t'_{pp}$ would hop a hole half-way across the $N = 32$ cluster, worsening finite size effects. The one-hole GS was recalculated with $t'_{pp} = 0$ for comparison.

### 5.4.1 Energetics

Fig. 5.2 shows the dispersion of the lowest two-hole states for different total-momentum $K$ and total-spin $S_T$. The $K = (\pi, \pi)$ $S_T = 0$ groundstate is doubly degenerate. The $K = (\pi, 0)$ and $K = (0, \pi)$ states are $\sim 0.17J_{dd}$ higher in energy. The $S = 1$ band crosses below the $S = 0$ band in the region around $K = 0$. These states are higher than the ground state by at least $\sim 0.3J_{dd}$, which is roughly the smallest difference between the $\frac{1}{2}$ and $\frac{3}{2}$ single-hole states as shown in Fig. 4.3. The convergence of the lowest states at high-symmetry points is shown in Fig. 5.3. The trend of exponential convergence is similar to the undoped (Fig. 3.3) and single hole case (Fig. 5.1), suggesting that the dominant part of the Hilbert space has been captured.

The binding energy of two holes is the energy difference between the two-hole energy minus twice the one-hole energy, shifted by the energy of the undoped
Figure 5.3: Convergence of the $S_T = 0$ lowest energy states. The two-hole energy level corresponding to $\Delta E_b = 0$ is indicated by the horizontal dash line.

The energy of the $K = (\pi, \pi)$ groundstate was extrapolated to be $-69.167 J_{dd}$, giving $\Delta E_b = +0.18J_{dd}$ which is non-binding. Of course, this value is sensitive to the extrapolation. If the result of the smallest $C_S = 0$ subspace is excluded from the linear fit of $\frac{\Delta E}{E}$ in Fig. 5.3, the extrapolated groundstate energy lowers to $-70.1228 J_{dd}$. This lowest estimate gives a binding of $\Delta E_b = -0.0252 J_{dd}$. This is roughly 44K, which is three times smaller than the extrapolated pseudo gap temperature of $T^* \sim 145 K$ at 2/32 = 6.25% doping. The finding here is that the groundstate of two holes does not have significant, if any, binding energy in the spin polaron model for a N=32 cluster. However, one should note that for a N=32 cluster, the undoped AFM has a $\frac{1}{N} \sim 0.3 J_{dd}$ q=0 magnon excitation, so a rigorous conclusion cannot be drawn because $|\Delta E_b|$ is smaller than this finite-size effect.
5.4.2 Wavefunction Analysis

The $K = (\pi, \pi)$ GS was found to be degenerate. It is well known that Lanczos-type diagonalization suffers greatly from degeneracy due to numerical noise. In this case, neither the generic ARPACK solver or my custom solver managed to produce a pair of orthogonal eigenvectors. As a quick fix, I modified my plain Lanczos solver to first calculate one GS vector $|GS_0\rangle$, then extract the GS of $(I - |GS_0\rangle\langle GS_0|)H(I - |GS_0\rangle\langle GS_0|)$. The crucial point here is to make sure the starting vectors of the two runs are not the same. If the two starting vectors are the same, for example because the pseudo-random number generator uses the same seed, the projected Hamiltonian would not have the other GS in the Krylov space and the computation would yield the lowest excited state instead. To get an orthogonal $|GS_1\rangle$, I start with the same pseudo-random vector in both runs, but before starting the second Lanczos procedure, I maximize orthogonality of the starting vectors by rotating pairs of vector components by the block-diagonal matrix of $2 \times 2$ rotation

$$
\begin{pmatrix}
0 & 1 \\
-1 & 0 \\
0 & \ddots & 0 \\
\vdots & & 0 & \ddots
\end{pmatrix}
$$

(5.20)

The resulting pair of groundstate wavefunctions satisfies

$$
\langle GS_0|GS_0\rangle = \langle GS_1|GS_1\rangle = 1 \\
\langle GS_0|GS_1\rangle = 0 \\
\langle GS_0|P_{xy}|GS_0\rangle = -\langle GS_1|P_{xy}|GS_1\rangle = \cos^2\theta - \sin^2\theta \\
\begin{pmatrix} |GS+\rangle \\
|GS-\rangle
\end{pmatrix} =
\begin{pmatrix}
\cos\theta & \sin\theta \\
-\sin\theta & \cos\theta
\end{pmatrix}
\begin{pmatrix} |GS_0\rangle \\
|GS_1\rangle
\end{pmatrix}
$$

(5.21)

where $P_{xy}$ and $P_{x'y'}$ are reflections about the diagonal directions $(a,a)$ or $(-a,a)$ in terms of lattice parameters (Fig. 4.1a). The total momentum $K = (\pi, \pi)$ is a high
symmetry point invariant to \( P_{xy} \) and \( P_{x'y'} \). The eigenvalues are

\[
\langle GS \pm | P_{xy} | GS \pm \rangle = -\langle GS \pm | P_{x'y'} | GS \pm \rangle = \pm 1
\]  

(5.22)

which indicate \( p \)-symmetry.

The \( K = (\pi, 0) \) and \( K = (0, \pi) \) states are merely \( \sim 0.17 J_{dd} \) higher in energy. These two momenta are invariant to \( P_x \) and \( P_y \), the reflection about \((a, 0)\) and \((0, a)\), respectively. The lowest \( K = 0 \) state is a doublet of total spin of \( S = 1 \) (6 states in total counting all \( z \)-projections) and should be an \( S=1 \) excitation from the \( K = (\pi, \pi) \) GS doublet. Disregarding such \( S = 1 \) excitations, one can consider the \( K = 0 \) \( S = 0 \) state. This energy level is not degenerate, and the wavefunction can be classified using \( P_{xy} \) and \( P_{x'y'} \). The expectation value of both operators are -1 so this state has a \( d \) symmetry. This is different from the \( t-t' \)-J model diagonalized in the \( N=32 \) cluster, wherein \( s \)-state is lower than the \( d \)-state.

The charge correlation as function of hole-hole separation can be illustrated by the expectation value of

\[
\hat{c}_{\text{charge}}(\delta) = \sum_{\delta} \hat{c}_{\text{charge}}(\delta) = \sum_{R} \hat{C}_{\text{charge}}(R)
\]

(5.23)

In a finite cluster, the number of hole-hole configuration separated by a distance \( R \), is limited by the periodic boundary condition to be smaller than that of \( L^2 = N \rightarrow \infty \) limit if \( R > \frac{L}{2} \). Therefore, such correlations should be compared to \( \sum_R P(R) = 1 \), the probability of two randomly distributed holes to be separated by \( R \) in the same finite cluster. For this purpose, the correlations can be gauged by

\[
\Delta \hat{C}_{\text{charge}}(R) = \hat{C}_{\text{charge}}(R) - P(R)
\]

(5.26)

where the correlation is the same as that of a random distribution if \( \Delta C(R) = 0 \).

The expectation values as a function of hole-hole separation are shown in
Figure 5.4: (top) Probability of charge separation, $\langle C^{\text{charge}}(R) \rangle$ for the lowest states in the $C_S = \frac{1}{2}$ subspace. (bottom) The difference from a random distribution.

Fig. 5.4. $\Delta C^{\text{charge}}(R)$ indicates charge repulsion at $K = (\pi, \pi)$, a small peak of correlation at $R=2$ for $(\pi,0)$, and attraction for $K = (0,0)$. However, one should note that two holes in a N=32 lattice corresponds to 6.25\% doping, which could be in the pseudogap regime without the simultaneous charge and phase coherence needed for superconductivity. It is thus important to also study the spin correlation between the two charges.

For the single-hole scenario in Chapter 4, the mobile carrier was found to be the $S = \frac{1}{2}$ 3$\text{SP}$, which is a composite of the oxygen spin and two neighboring copper spins. The two-hole solutions yielded $\langle H_{J_{pd}} \rangle \sim 2 \times -0.9 J_{pd}$, meaning $H_{J_{pd}}$ can still provide a good characterization. Determining the spin-spin correlation between the two polarons requires phase information among the four copper spins. This is not a straightforward exercise because annihilating copper spins would take the wavefunction outside of the $\langle \sum d_{i+\sigma}^+ d_{i+\sigma} \rangle = N$ numerical basis formulated in Sec. 5.3. It is instructive to first consider the eight single-hole states of $H_{J_{pd}}$ listed in Tab. 5.2. In the two-hole scenario, the oxygen holes can neighbor the same copper spin to form a 5-spin object, but $C^{\text{charge}}(R)$ shown in Fig. 5.4 shows that the probability is low. Ignoring these common-copper configurations, the wavefunction contains two
Table 5.2: Single-hole eigenstates of $H_{pd}$. $p_{\sigma}^\dagger$ creates an oxygen hole and the arrows in the ket indicate the spins of the two copper sites neighboring the oxygen hole.

polarons involving a total of six spins. Noting that $\langle H_{pd} \rangle \sim -1.8J_{pd}$, the single-polaron levels in Tab. 5.2 suggest that the dominant part of the wavefunction contains four possible 3SP pairs with $\langle H_{pd} \rangle = -2J_{pd}$: $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle|\downarrow\uparrow\rangle \pm |\uparrow\downarrow\rangle|\uparrow\downarrow\rangle)$, and $|\downarrow\downarrow\rangle|\downarrow\downarrow\rangle)$. Other pair configurations would have $\frac{\langle H_{pd} \rangle}{J_{pd}} \in \{0, \pm \frac{1}{2}, \pm 1\}$. The numerical wavefunction can be projected into the subspace of 3SP pairs by

$$P_{3SP} = \prod_{\Delta \lambda \in \{0, \pm \frac{1}{2}, \pm 1\}} \frac{\langle H_{pd} \rangle / J_{pd} - \Delta \lambda}{-2 - \Delta \lambda}$$

(5.27)

wherein each term in the product zeros out the eigenvalue of a particular level while scaling the $\frac{\langle H_{pd} \rangle}{J_{pd}} = -2$ level to unity. The probability of having two 3SP with any spin-spin correlation and at any separation is thus $\langle P_{3SP} \rangle \sum_{\delta \sigma} \langle c_{\text{charge}}(\delta) | P_{3SP} \rangle$. In a six-spin problem, the projected wavefunction is a superposition of the four possible 3SP pairs

$|6\text{spin}\rangle = a|\uparrow\downarrow\rangle|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle|\downarrow\downarrow\rangle + b|\uparrow\uparrow\rangle|\downarrow\downarrow\rangle + c|\downarrow\downarrow\rangle|\downarrow\downarrow\rangle + d|\downarrow\downarrow\rangle|\downarrow\downarrow\rangle$. (5.28)
$p_{L,\sigma}^\dagger$ and $p_{R,\sigma}^\dagger$ respectively, is

$$\langle \frac{p_{L\uparrow}^\dagger p_{R\downarrow}^\dagger - p_{L\downarrow}^\dagger p_{R\uparrow}^\dagger}{\sqrt{2}} - \frac{p_{L\downarrow}^\dagger p_{R\uparrow}^\dagger - p_{L\uparrow}^\dagger p_{R\downarrow}^\dagger}{\sqrt{2}}\rangle = \frac{3}{9}a^2 + \frac{2}{9}(b^2 + c^2 + d^2) \quad (5.29)$$

Rearranging for $a^2$ and generalizing the expression to different hole-hole separation $\delta$, the nature of singlet 3SP pair can be gauged by

$$c_{3SP;\text{singlet}}(\delta) = \frac{\langle P_{3SP} (9\hat{c}_{\text{singlet}}(\delta) - 2\hat{c}_{\text{charge}}(\delta)) P_{3SP} \rangle}{\langle P_{3SP} \sum_{\delta'} \hat{c}_{\text{charge}}(\delta') P_{3SP} \rangle} \quad (5.30)$$

$$\hat{c}_{\text{singlet}}(\delta) = \sum s_{l+\varepsilon,j+\varepsilon+\delta} s_{l+\varepsilon,j+\varepsilon+\delta} \quad (5.31)$$

where $s^\dagger$ is the singlet creation operator defined in Eq. 5.6. $c_{3SP;\text{singlet}}(\delta)$ is the probability, within the projected subspace, of the two oxygen holes and the AFM background forming a 3SP singlet pair separated at distance $\delta$. The value ranges from zero for no singlet nature to unity for pure singlet at hole-hole separation of $\delta$; however, these two extreme values are not possible due to, for example, the spatial spreading required to lower the kinetic energy of $T_{pp}$ and $T_{swap}$. The measure can be summed as a function of hole-hole separation

$$C_{3SP;\text{singlet}}(R) = \sum_{|\delta|=R} c_{3SP;\text{singlet}}(\delta) \quad (5.32)$$

The corresponding triplet measure is

$$C_{3SP;\text{triplet}}(R) = \sum_{|\delta|=R} \frac{\langle P_{3SP} \hat{c}_{\text{charge}}(\delta) P_{3SP} \rangle - c_{3SP;\text{singlet}}(\delta)}{\langle P_{3SP} \sum_{\delta'} \hat{c}_{\text{charge}}(\delta') P_{3SP} \rangle} \quad (5.33)$$

$$1 = \sum_R C_{3SP;\text{singlet}}(R) + C_{3SP;\text{triplet}}(R) \quad (5.34)$$

The singlet correlation should be compared to $\sum_\delta P'(R) = 1$, the random distribution of two 3SPs spreading over the 64 oxygen sites, with no common copper spin. The probability of singlet correlation is $\frac{1}{4}$ in a paramagnetic state. The difference of interest is thus

$$\Delta C_{3SP;\text{singlet}}(R) = C_{3SP;\text{singlet}}(R) - \frac{P'(R)}{4}. \quad (5.35)$$
Figure 5.5: (top) Probability of singlet 3SP pair, $\langle C_{3SP,singlet}(R) \rangle$ for the lowest states in the $C_S = \frac{1}{2}$ subspace. (bottom) The difference from a randomly distributed paramagnetic configuration.

The low-energy states all have $\langle P_{3SP} \sum \delta \hat{c}_{\text{charge}}(\delta) P_{3SP} \rangle \sim 0.8$. The rest of the weight are distributed towards the sector with one or two $|0\pm\rangle$ (Tab. 5.2). Figure 5.5 shows the resulting $C_{3SP,singlet}(R)$ and $\Delta C_{3SP,singlet}(R)$. It is evident that all states have enhanced short-range singlet nature compared to the random distribution – a feature absent from previous studies using small cluster or the t-J model which does not distinguish oxygen and copper sites. This short-range nature is robust against finite-size issues because long-range nature can indeed be found as for the charge correlation Fig. 5.4.

The $K = (\pi, \pi)$ GS has enhanced short-range singlet correlation but the overall charge distribution is repulsive as shown in Fig. 5.4. This implies that the most of the long-range charge correlations belongs to the sector of non-3SP-pair nature. The individual $c_{3SP,singlet}(\delta)$ values are shown in Fig. 5.6. The local maximum occurs when the two 3SP are $R/a \sim 1.5 - 2$ apart. I stress here that, although the fluctuation of the values over space is merely about a factor of five, such fluctuation is not present in the t-J model solved for two holes propagating in a cluster with the same geometry[132]. This enhanced correlation in the singlet nature must be an anomalous signature because, for example, the singlet enhancement for $R/a \leq 2$
Figure 5.6: \( c_{3\text{SP,singlet}}(\delta) \times 10^4 \) for one of the two \( K = (\pi, \pi) \) GS in the \( C_S = \frac{1}{2} \) subspace. Cu and O are marked by circles and rectangles, respectively. Values for the other state are obtained by 90° rotation. (top, black→red) 28 x-y, (bottom, black→black) 16 x-x and (bottom,red→red) 16 y-y 3SP pair configurations are shown; 6 O-O configurations, which involve the two O holes sharing the same copper site, have low probability and are omitted. For example, the bottom figure shows that \( c_{3\text{SP,singlet}}(2a,0) = 0.0109 \) for two 3sp's both centered on x-rung O sites.
is observed even though the charge measure shows no such tendency (Fig. 5.4). In fact, a singlet correlation that is not confined in real space implies that the singlet correlation should be more confined in momentum space; however, k-space features have not been extracted because each 3S\text{P} involves two background copper spin in real space, and k-space features would require Fourier transforming the local spins while leaving the rest intact.

The $K = (\pi, 0)$ state is slightly higher in energy. Figure 5.5 shows enhanced short-range singlet correlation. However, the charge correlation in Fig. 5.4 shows a local maximum at $R/a = 2$ unlike the $K = (\pi, \pi)$ state. The spatial distribution of the singlet and charge are shown in Fig. 5.7 and Fig. 5.8. The observation is that the charge and singlet maxima happens when both holes are aligned along the momentum direction. This alignment can potentially be pinned by external interactions as stripes, but a much larger system with more holes are required to verify this.

Disregarding the $S=1$ excitation from the above states, the $K = (0, 0)$, $S_T = 0$ state is much higher in energy and has d-symmetry, unlike the s-symmetric state found in N=32 exact diagonalization of the t-t'-J model\cite{132}. The singlet correlation is shown in Fig. 5.9. The two polarons are $R/a = 1$ apart without sharing a copper site.

### 5.5 Speculations

The results presented above are fairly unbiased interpretations of the numerical solutions. This section makes some speculations about larger systems with more holes. The goal is to list several potential features of interest for future developers.

Although we have broken some technological barriers in extracting explicit multi-hole wavefunctions for large lattice at zero temperature, finite system modeling is inherently limited by cluster size. Two holes in a periodic N=32 system is often interpreted as an approximation to 6.25% doping, but the solution might or might not be representative of a real system because inhomogeneous concentration is not allowed. To be precise, if the holes have tendency to cluster, two holes in a N=32 system should be interpreted as $\lim_{N \to \infty} \frac{2}{N}$ doping because a nominal 32-unit-cell patch in a real 6.25% doping system would often not contain two holes.


Figure 5.7: \( c_{3SP,\text{singlet}}(\delta) \times 10^4 \) for the lowest \( K = (\pi, 0) \) state in the \( C_S = \frac{1}{2} \) subspace. Cu and O are marked by circles and rectangles, respectively. Values for the \( K = (0, \pi) \) state are obtained by 90° rotation. (top, black→red) 28 x-y, (bottom, black→black) 16 x-x and (bottom, red→red) 16 y-y 3SP pair configurations are shown; 6 O-O configurations, which involve the two O holes sharing the same copper site, have low probability and are omitted. For example, the bottom figure shows that \( c_{3SP,\text{singlet}}(2a, 0) = 0.0265 \) for two 3SPs both centered on x-rung O sites.
Figure 5.8: $c_{\text{charge}}(\delta) \times 10^4$ for the lowest $K = (\pi, 0)$ state in the $C_S = \frac{1}{2}$ subspace. Cu and O are marked by circles and rectangles, respectively. Values for the $K = (0, \pi)$ state are obtained by $90^\circ$ rotation. (top, black→red) 32 x-y, (bottom, black→black) 17 x-x and (bottom, red→red) 17 y-y 3SP pair configurations are shown. For example, the bottom figure shows that $c_{\text{charge}}(2a, 0) = 0.0289$ for two 3SPs both centered on x-rung O sites.
Figure 5.9: \( c_{3SP,\text{singlet}}(\delta) \times 10^4 \) for the lowest \( K = (0,0) \) \( S_T = 0 \) state in the \( C_S = \frac{1}{2} \) subspace. Cu and O are marked by circles and rectangles, respectively. (top, black→red) 28 x-y, (bottom, black→black) 16 x-x and (bottom, red→red) 16 y-y 3SP pair configurations are shown; 6 O-O configurations, which involve the two O holes sharing the same copper site, have low probability and are omitted. For example, the bottom figure shows that \( c_{3SP,\text{singlet}}(0, -2a) = 0.0367 \) for two 3SPs both centered on x-rung O sites.
Although in the last section we noted that the enhanced singlet correlations are not confined in real space and should therefore be confined in momentum space, one can speculate on the origin of the singlet tendency by considering the 6-spin wavefunctions \( \frac{1}{\sqrt{2}} (|\uparrow\rangle |\downarrow\rangle \pm |\downarrow\rangle |\uparrow\rangle) \). By expanding and taking the \( \mathbf{S} \cdot \mathbf{S} \) between a copper spin from the first 3SP and one from the second 3SP, the expectation value is \(-\frac{5}{18}J_{dd}\) if they are singlet and \(+\frac{1}{18}J_{dd}\) if they are triplet. Therefore, two 3SPs can take advantage of this nature by forming a singlet if they are separated by an empty oxygen site, which mediate an ordinary \( \mathbf{S} \cdot \mathbf{S} \) Heisenberg AFM bond. The two-hole numerical solution indeed shows a local maximum when \( R/a \leq 2 \) (Fig. 5.5). As doping increases, the average spacing between oxygen holes would decrease if they do not cluster. This can be checked by expanding the 12-spin wavefunctions obtained as the direct product of two 3SP singlet pairs: \( \frac{1}{2} (|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle) \otimes (|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle) \). The expectation value of \( \mathbf{S} \cdot \mathbf{S} \) between a copper spin from the first pair and one from the second pair is exactly zero. Therefore, different pairs tend to be separated by more than one empty oxygen sites for some negative correlations. There is no immediate clustering tendency beyond two holes, at least according to this back-of-the-envelope calculation. This provides some merit for interpreting the finite-size results as a realistic doped scenario with little clustering.

Having noted the lack of immediate clustering mechanism, I would like to point out that three features of the numerical solution are likely to prevail in larger systems with more holes: i) the kinetic and potential energy leads to the formation of 3SPs, ii) 3SP destroys AFM correlation in its vicinity (Fig. 4.3 and 4.4), and iii) two 3SPs tend to have enhanced singlet correlation compared to the pure random scenario (Fig. 5.5).

Noting these three points, the above back-of-the-envelope calculations suggests a possible binding mechanism for a real doped system. The \(-\frac{5}{18}J_{dd}\) energy for a 3SP singlet pair separated by one empty oxygen site is of course higher than the \(-0.33J_{dd}\) copper-copper correlation of an undoped system; however, the \(-\frac{5}{18}J_{dd}\) is a local AFM correlation which is independent of doping, and the correlations between copper spins away from the 3SP are raised on average due to point ii). If pair-pair clustering is indeed absent, doping would reduce the average distance between oxygen holes while increasing the average copper-copper corre-
lation. At certain doping, the $-\frac{5}{18}J_{dd}$ short-range AFM correlation can become energetically important and the holes form bipolaron pairs. This pairing tendency cannot increase indefinitely because the pairs want to be more than one empty oxygen apart. Evidently, $\sim 25\%$ is the highest doping level which allows $3SP$ singlet pairs without having pairs within one empty oxygen site.

I stress again that this section is merely providing a list of speculations based on the two-hole solution. The two-hole solution has not been obtained when previous speculations of the same type were made. In particular, the numerical two-hole wavefunctions are not node-less s-symmetric as suggested by earlier approximations by Emery and Reiter[14, 25]. The temperature scale of the mechanism is likely to be much lower than $J_{dd}$ because the binding tendency is determined by the difference between the $-\frac{5}{18}J_{dd}$ short-range AFM $3SP$-$3SP$ correlation and the other copper-copper spin correlation, which increases from the undoped value of $\sim -0.33J_{dd}$. The stability of the $3SP$ itself increases and also the pairing tendency increases as long range antiferromagnetic correlations are destroyed by further doping. The temperature cannot increase indefinitely because overpopulation would destroy the tendency of bipolaron singlet formation. If the pairing physics were indeed due to this singlet binding mechanism, the transition from the superconducting regime to the Fermi-liquid regime at $\sim 25\%$ doping (Fig. 1.1) can be explained as the spatial saturation of $3SP$ singlet pairs. The copper spin background can no longer be correlated into $3SP$ with the oxygen holes, and the Fermi-liquid behavior is due to the “free” oxygen holes. This speculative picture should be explored when solutions of larger systems become available.

5.6 Conclusions

The octapartite scheme has been successfully applied to the scenario of two holes propagating in the spin-polaron model with 32 Cu and 64 O sites, breaking the technological barrier. Exponential convergence similar to the undoped and one-hole scenarios has been established.

While the energetics of the low energy states yield inconclusive results, one should be reminded about the danger of relying solely on ground-state energy, which is the “‘Quantum Chemist’ Fallacy No. 2” coined by by Anderson[133].
The perturbative derivation of the spin polaron model in Sec. 4.2.1 is bound to leave out terms that would slightly affect quantities such as binding energy. In fact, the pseudo gap temperature of $T^* \sim 150K$ is less than one-tenth of $J_{dd}$, smaller than the $\propto \frac{1}{N} \sim 0.3J_{dd}$ $q = 0$ magnon excitation for an $N = 32$ undoped AFM as well as any energy difference in the one- or two-hole energy spectrum (Fig. 4.2 and Fig. 5.2). It would be fruitful to refine the octapartite scheme to treat larger lattices.

Careful analysis of the wavefunctions at $K = (0, 0), (\pi, 0), (\pi, \pi)$ shows that all states exhibit anomalous enhanced hole-hole singlet correlation, compared to the ED results of the t-t'-J model solved on a N=32 cluster with the same geometry[132]. Some interesting speculations have been made about the behavior of larger systems, providing the merits for further developments.
Chapter 6

Conclusions

6.1 Summary

This work describes progress in the treatment of electron-electron and electron-phonon interactions. Several numerical approaches were devised, developed, and adapted to accurately solve large models with details. These technological developments were complemented by analytical approaches to reveal new physics, stressing the importance of distinguishing anions and cations in transition-metal oxides.

Chapter 2 explores the breathing-mode model which describes a charge carrier centered at cation sites interacting with quantized modes of vibration at anion sites. In this more realistic scenario, the coupling amplitude becomes momentum-dependent unlike the widely-studied Holstein model. This momentum dependence directly leads to non-monotonic energy dispersion, sharper large-to-small polaron crossover as a function of coupling strength, and lighter effective mass in the low-coupling regime compared to the Holstein model. The result serves as the first available accurate benchmark for the model for all coupling regime.

Chapter 3 points out a low-energy characteristic of the much-studied spin-$\frac{1}{2}$ Heisenberg antiferromagnet on a 2D square lattice, which is relevant to the undoped cuprate compounds. The octapartite approach is then developed to take advantage of this intuition. The groundstate wave function can be solved explicitly for a record breaking 64-spin torus. The error can be systematically reduced, and
the convergence of the approach is established to be exponential with respect to the completeness parameter $C_S \in [0, 1]$.

A spin polaron model is derived in Chapter 4 to capture the coherence of holes doped into the copper-oxygen layer in cuprate compounds. The solution of a single hole is solved via a novel total-spin-resolved exact diagonalization approach for a record-breaking cluster of 32 copper and 64 oxygen sites. The combination of the new model and large cluster reveals important physics, missed by previous studies, with direct experimental consequences.

Chapter 5 combines Ch. 3 and 4 to bypass several computational restrictions, such as the sign problem and special periodic boundary condition, in dealing with two holes propagating in a cluster with 32 copper and 64 oxygen sites. The results prove that the octapartite scheme can indeed accurately capture the disturbed spin background. The solutions exhibit enhanced singlet correlation compared to the simplest t-J model solved in the same cluster.

6.2 Opportunities for Further Developments

The promising results in the two-hole scenario provide merits for further developments. Theoretically, the model can be tweaked by adding other terms, including coupling to phonons. Larger clusters are within reach because of various ways of improving the numerical scheme. Technologically speaking, specialized optimization can be worthwhile because the octapartite approach has proven to be reliable. In particular, the current implementation requires a pre-transformed matrix for computation. It would be fruitful to develop an algorithm fast enough to by-pass this step. Implementing this concept via variational Monte-Carlo sampling or renormalization can significantly extend its prowess.

In a more general context, the octapartite approach (Ch. 3) is devised to capture an AFM background on the square lattice complicated by the presence of doped holes. Orthogonal extensions to the square lattice AFM problem include the frustrations due to non-nearest-neighbor AFM bonds and due to other lattice structures. Because the current approach is formulated to capture the essence of doped cuprates, the method is not expected to be directly applicable to these frustrated scenarios. Nevertheless, the use of nonlocal spin partitions to capture important
physical essences is a holistic way which could lead to powerful case-by-case solutions in dealing with these peculiar spin systems.
Bibliography


[106] I have reported a software bug in the parallel version of the ARPACK package, but the package has not been updated as of to date.


[131] Keimer, B. to be published.


Appendix A

Software Integrity Checks

- Trial computations were performed by zeroing parameters such that the non-zero parts are known analytically.
- Trial computations were performed without assuming $H^\dagger = H$.
- All computations were checked by different combinations of at least two numerical solvers and two algorithmic routines developed independently.
- Solutions are verified against those computed from larger Hilbert space without reduction due to translational, total-spin, total-spin-projection, or point-group symmetries.
- The symmetries of solutions were always explicitly characterized.
- Symmetries were checked whenever possible during intermediate steps.
- For all operators that preserve 2-norm, $\sum_f |\langle f|\hat{O}|i\rangle|^2 = 1$ was checked explicitly for an initial state $|i\rangle$ and all final states $|f\rangle$.
- For some cases, $HP \pm PH$ was verified explicitly for a transformed Hamiltonian, $H$, and known symmetry operators, $P$.
- $H^\dagger - H = 0$ was verified explicitly.
- When possible, floating-point variables were initialized as NaN, and unsigned integers were initialized as $0x\text{FFFFFFFDFFFFFD}$. 
Appendix B

Computation Development

This appendix first lists the software optimizations employed to perform record-breaking numerical computations on relatively modest hardwares. The technical reasons behind the novel octapartite approach (Ch. 3-5) as well as its implementation are then discussed in details.

B.1 Software Optimizations

Software was developed using C++ for the x86 and, in some cases, powerPC, platforms. Some legacy F77 libraries were integrated using a C interface. Here is a list of optimizations utilized throughout the course of this work.

- Multi-thread parallelization was implemented using OpenMP due to the ease of synchronization.
- Multi-node parallelization was implemented using MPI. This includes the use of processor farm for pipe-line procedures.
- Performance critical part of the computation was written with in-line assembly instructions. In particular, data parallelism was exploited using SSE instructions with the 128-bit registers.
- Matrix-vector product can be optimized using compressed sparse row (CSR) formulation. CSR compression can be implemented efficiently using the
\_gnu\_cxx\_::hash\_map template with the floating-point type overloaded with a value-zeroing default constructor.

- Spin configurations of the form $\prod | \sigma \rangle$ can be represented by 0 and 1 of an unsigned integer. Translation, spin flip, and spin-spin correlations can be performed in parallel with bitwise operators such as $<<$, $>>$, $\&$, $\mid$, and $\hat{\}$. 

- Before peculiar transformations and truncations, spin and phonon basis can be enumerated using Algorithm 151[122].

- Computations with precomputed matrix transformations were optimized by the parallel use of separate computation core, write-only core, and read-only core. The throughput of the read-only process was optimized by reading from a RAID00 configuration of SSD drives connected via PCIe slots.

- In general, the precomputed approach is better than on-the-fly matrix-vector product only if the basis reduction allows solution to system size that is otherwise impossible. Also, the preliminary computation should be massively parallelizable. The strategy is viable only if the code does not loop over the matrix elements that are identically zero. This can be accomplished by paying close attention to, for example, the delta functions of the Clebsch-Gordan coefficients.

### B.2 Details of the Octapartite Approach

As discussed in Ch. 3-5, the “octapartite approach” is motivated by the need of efficient modeling of large AFM background with and without the presence of extra charge carriers. The idea is to transform the spin background into a suitable basis such that the essence can be captured by a manageable number of states.

There might be more efficient ways of doing so, but the current implementation was the sure way of producing interesting results for a set period of time for a pioneering project. The aim was to demonstrate an overall idea which combinatorially decreases the number of states required to model the interesting part of the Hilbert space. This reduction is exactly the reason why, even without the most sophisticated implementation, technological barriers were immediately broken for
combinatorially large system when available computational power advances only exponentially over time according to Morse’s Law.

In particular, the implementation partitions the computation into consecutive stages for two main reasons. First, the Hilbert space truncation for picking out the \( \frac{N}{8} \) spins adding to zero characteristic (Sec. 3.3) requires a peculiar quantum-mechanical reformulation with no apparent way of enumeration. Developing matrix-vector product within such strange basis is highly prone to errors, and stage-by-stage consistency checks are required anyways in the development process. Second, the strategy allowed me to simultaneously run one stage of computation while coding for the next. By doing so, the “precomputation time” mentioned in the last section is effectively zero because coding and quality assurance took longer than actual run time.

The rest of this appendix will point out some general observations of the problem then describe the current implementation. Future developers are encourage to consider some general observations before either duplicating the current implementation or developing their own.

**B.2.1 General Observations**

There are several challenges in working within a basis formulated to single out the \( \frac{N}{8} \) spins adding to zero subspace. First, the truncation of less important states should be controllable systematically and flexible enough to adapt to the many unknowns of doped systems. Second, within the truncated basis transformed away from the natural \( z \)-projected description, there should be a fast way of indexing initial and final states upon Hamiltonian operations. Third, one should have apriori knowledge about the identically-zero matrix elements in order to avoid them for sparse matrices’ performance scaling.

Let’s review several general aspects of an arbitrary basis for the spin problem in Ch. 3-5. For a small number of oxygen holes and a copper spin background, an arbitrary translationally invariant basis can be written as

\[
\frac{1}{\sqrt{N}} \sum e^{iKl} p_{l,\sigma_l}^+ \left( \prod_{\Delta l} p_{l+\Delta l,\sigma_{\Delta l}}^+ \right) |\overline{\sigma}\rangle_l
\]  

(B.1)
with one oxygen hole and a spin background centered at position \( l \) and all other oxygen holes positioned \( \Delta l \) away from the reference point \( l \). The Hamiltonian acted upon an initial state would produce some final states scaled by matrix elements. Each term in the Hamiltonian is a summation over all unit cells; therefore, the final states always come in the form of a Fourier sum, \( \sum e^{iKl} \). The overlap between the final state and orthonormal basis states is thus some complex constant with a sum, \( \frac{1}{N} \sum \delta \), wherein delta functions arise from the phases, the fermion operators and the bra-ket of the spin backgrounds. On top of a particular spin background \( \sigma_l \), the different configurations of oxygen holes, \( p_{\Delta l}^\dagger \sigma_{\delta l} \), can be enumerated as shown in Sec. 5.3. The remaining part is then the connections between different spin backgrounds.

A spin polaron model of the form introduced in Ch. 4 and 5 would yield several types of connections. Direct oxygen-oxygen hopping \( T_{pp} \) moves either the “non-reference” oxygen holes or the oxygen hole at the reference position, \( l \). The former is simply changing the \( \Delta l \) tabulation while the later result in no \( p^\dagger \) occupation at the reference point \( l \). This simply means that the final state can be specified in terms of a new reference point \( l' \) occupied by a \( p^\dagger \) hole. The matrix element is thus complex with a phase \( e^{iK(l-l')} \) and a spin background shifted by \( l - l' \):

\[
e^{iK(l-l')}(\sigma'|T_{l-l'}|\sigma),
\]

which requires the knowledge of how one spin configuration connect to another upon translation.

Noting that the oxygen holes \( p^\dagger \) are already taken care of, the \( J_{pd} \) exchange requires knowledge of \( S_{\Delta L}^{c,\pm} \). \( S^\pm \) is computed using

\[
\langle \sigma'|S_{\Delta L}^c|\sigma \rangle = \frac{1}{2} \langle \sigma'|d_{\Delta L,\uparrow}^\dagger d_{\Delta L,\downarrow} - d_{\Delta L,\downarrow}^\dagger d_{\Delta L,\uparrow}|\sigma \rangle.
\]

\( S^\pm \) requires information about kicking out an original spin, \( \alpha \), then injecting a new spin \( \beta = -\alpha \):

\[
\langle \sigma|K_{\Delta L}(\alpha, \beta)|\sigma \rangle = \langle \sigma|d_{\Delta L,\beta}^\dagger d_{\Delta L,\alpha}|\sigma \rangle.
\]

The \( T_{swap} \) hopping is the combination of the direct oxygen-oxygen hopping, plus kicking process \( K_{\Delta L}(\alpha, \beta) \) with a \( \alpha \) to \( \beta \) spin swapping. Note that \( \beta \) is now
the original spin of the oxygen hole with no definite sign relative to $\alpha$. $J_{dd}$ matrix elements involve two copper spins and cannot be tabulated as easily. For nearest neighbor separation $a$, one would require the overlaps

$$
\langle \sigma | S_{\Delta L}^z S_{\Delta L+a}^z | \bar{\sigma} \rangle \quad (B.5)
$$

$$
\langle \sigma | K_{\Delta L}(\alpha, -\alpha) K_{\Delta L+a}(\alpha, \alpha) | \bar{\sigma} \rangle \quad (B.6)
$$

Note that for a full Hilbert space specified in terms of the $z$-projection of copper spins, all these connections can be computed trivially by flipping bits of an unsigned integer and using Algorithm 151\[122\] to enumerate the states of the basis. However, such full Hilbert space is unmanageable for the large system sizes of interest (Sec. 5.1) and an alternative way is needed. The observations made in Ch. 3 are specified in terms of $s_{A/B}$, the total spin of the sublattices; therefore, a formulation in terms of total quantum mechanical spin is required to reduce the dimension of the Hilbert space. The task is to find one such formulation which also allows easy computation of the above matrix elements. The following section discusses a way of doing so.

### B.2.2 Implementations

The previous section pointed out that the matrix representation of an effective spin polaron model requires several “basic” operations on the spin background: $K_{\Delta L}(\alpha, \beta)$, $T_{l-l'}$, and $S_{\Delta L}^z$.

There are two immediate advantages of building the full spin background in terms as a Clebsch-Gordan series of total-sub-lattice-spin states (Eq. 3.9). First, it is apparently easier to truncate out the low-sub-lattice-spin part of the Hilbert space according the $C_S$ criterion (Ch. 3.3.2). Second, note that processes $T_{pp/\text{swap}}$ and $J_{pd}$ requires a single operation of one of the three “basic” operators, and the $J_{dd}$ copper-copper exchange requires two consecutive operations of either $K_{\Delta L}$ or $S_{\Delta L}^z$. $J_{dd}$ is bi-partite in nature so the pair of consecutive operations is decomposed into two simultaneous operations on orthogonal part of the Hilbert space, for example

$$
\langle \sigma' | K_{\Delta L}(\alpha, -\alpha) K_{\Delta L+a}(\alpha, +\alpha) | \sigma \rangle \rightarrow \langle \sigma'_A | K_{\Delta L}(\alpha, -\alpha) | \sigma_A \rangle \langle \sigma'_B | K_{\Delta L+a}(\alpha, \alpha) | \sigma_B \rangle.
$$

(B.7)
It was found that the choice of sub-lattice basis yields extremely sparse matrices. The consecutive matrix multiplications within the final overall basis are reduced to matrix elements look up within a sparse matrix of a much smaller basis.

Having established the usefulness of bi-partite separation in dealing with truncation as well as the the evaluation of the $J_{dd}$ term, the remaining question is how to formulate such spin basis. The truncation according to $s_{A/B}$ requires the sub-lattice spins to be added quantum mechanically, e.g. in terms of Clebsch-Gordan series. However, the hopping processes require knowledge of how each of the spin background relate to one another upon the 2D translation $T_{l-l'}$, which has no immediately apparent relationship with a spin addition exercise for a large number of spins. For example, the apparent way of determinationing the overall sign of a matrix element is to follow the signs of Clebsch-Gordan addition for both the initial and final states.

As discussed there, the square lattice is partitioned into eight sublattices such that the sites of a particular sublattice are connected by the vectors $\frac{1}{\sqrt{2}}(a, \pm a)$ (Fig. 3.1). For a group of $\frac{N}{8}$ spins, I’d start with the $z$-projected representation of $2^\frac{N}{8}$ states and transform into a Clebsch-Gordan basis |$s_{N/8}, s_{z N/8}$⟩. |$s_{N/8}, s_{z N/8}$⟩ states can be done by adding $i$ spins with $(\frac{N}{8} - i)$ spins. One way to proceed is to add the group of spins neighbored by the oxygen holes with the rest, but such approach would complicate the determination of $T(l, l')$ among spin background. For example, each oxygen configuration, $\prod_{P} l + \delta l', \sigma l'$, would have a different basis for the spin background centered at $l$. While this approach could potentially be useful for detailed truncation of the disturbed background, I decided to use a simple coarse-grain approach (Sec. 3.3.2) to eradicate this complication.

The tabulation of the $\frac{N}{8}$ sublattice spins into a basis of |$s_{A/B}, s_{z AB}$⟩ can be done by adding $i$ spins with $(\frac{N}{8} - i)$ spins. One way to proceed is to add the group of spins neighbored by the oxygen holes with the rest, but such approach would complicate the determination of $T(l, l')$ among spin background. For example, each oxygen configuration, $\prod_{P} l + \delta l', \sigma l'$, would have a different basis for the spin background centered at $l$. While this approach could potentially be useful for detailed truncation of the disturbed background, I decided to use a simple coarse-grain approach (Sec. 3.3.2) to eradicate this complication.

To build |$s_{N/8}, s_{z N/8}$⟩ according to Clebsch-Gordan addition. Then with two identical enumerations of |$s_{N/8}, s_{z N/8}$⟩, I’d build a single enumeration of |$s_{N/8}, s_{z N/8}$⟩. Finally, I can similarly enumerate the overall background |$s_{N}, s_{z N}$⟩ from the enumeration of |$s_{N/8}, s_{z N/8}$⟩. The enumeration can be specified by a non-negative integer, state_index. The state represented by a particular index value can be derived from the “parent” basis.
with half the number of spins, for example using the following loop structure. It is apparent that truncation can be performed at each stage of mixing by simply tuning the values of \( \min 2S, \max 2S, \min \text{Sub} 2S, \text{and } \max \text{Sub} 2S. \)

```cpp
size_t state_index = 0;
// in general total spin can be odd multiples of 1/2, so I worked with 2S
const int \( \min 2S = \min 2 \times \text{total spin} \);
const int \( \max 2S = \text{number of spins} \);
const int \( \min \text{Sub} 2S = \min 2 \times \text{sub lattice spin} \);
const int \( \max \text{Sub} 2S = \text{number of sub lattice spins} \);
// states with higher total spin have higher indices
for (int \( ss = \min 2S; ss < \max 2S; ss += 2 \))
  // loop over the left ket
  for (int \( ssa = 0; ssa < \max \text{Sub} 2S; ssa += 2 \))
    // loop over the right ket
    for (int \( ssb = 0; ssb < \max \text{Sub} 2S; ssb += 2 \))
      if (Clebsch_Gordan_coefficients == 0) continue;
      const size_t \( na = \text{number of } (ssa + 1) \text{ blocks with total spin } ssa \);
      const size_t \( nb = \text{number of } (ssb + 1) \text{ blocks with total spin } ssb \);
      for (size_t \( aa = 0; aa < na; ++aa \))
        for (size_t \( bb = 0; bb < nb; ++bb \))
          // loop over the \( ssa + 1 \) components of \( z \)-projection
          for (int \( zz = -ss; zz < ss; zz += 2; ++state_index \))
            fprintf(stderr, "State %lu has s=%4f and sz=%+4f, product of", state_index, 0.5 * ss, 0.5 * zz);
            fprintf(stderr, "block %lu/%lu of the s=%4f sector of left ket and ", aa, na, 0.5 * ssa);
            fprintf(stderr, "block %lu/%lu of the s=%4f sector of right ket", bb, nb, 0.5 * ssb);
          }
      } // zz
    } // bb
  } // aa
} // ssa
} // ss
```

Each state indexed by \( \text{state index} \) is associated with two blocks of states in the smaller parent basis. Each block contains the \( 2S + 1 \) projections related by \( S^+/- \). This is done to accommodate the state’s Clebsch-Gordan series, which has a delta function, \( \delta_{\lambda' \lambda} = (\lambda' + \lambda)^+ \) in the summation over \( z \)-projections. This blocking scheme immediately reveals the “parent” basis states which are relevant to that particular
Clebsch-Gordan series.

The benefit of this enumeration is an instantaneous “reverse lookup”. When performing Hamiltonian operations, one is really interested in the non-zero overlap between the outgoing states and those in the orthonormal basis. The naive way is to compute the dot product againsts all basis states, but this adds a $O(N)$ layer on top everything else and is detrimental in the case of large systems. Under the above looping scheme, an increasing state index is associated with increasing values of $ssa, aa, ssb, and bb$. Any arbitrary ket, $|\sigma_a; \sigma_z a\rangle |\sigma_b; \sigma_z b\rangle$, is trivially associated with these four indices so states with non-zero overlap is known immediately, with an $O(1)$ reverse lookup operation (in analogy to Algorithm 151[122]).

This hierarchy allows an efficient computation of the overall matrix in terms of $K_{\Delta L}(\alpha; \beta), T^{-}, and S_{\Delta L}^z$. The initial information needed is some small matrices $\hat{O}$ within the smallest $N_8$ spin basis. The matrices for the larger children basis required two general operations. For $K_{\Delta L}(\alpha; \beta)$ and $S_{\Delta L}^z$, one needs to evaluate

$$\langle \sigma_a, \sigma_z^a | O_L | \sigma_a, \sigma_z^a \rangle \times \langle \sigma_b, \sigma_z^b | \sigma_b, \sigma_z^b \rangle \quad (B.8)$$

$$\langle \sigma_a, \sigma_z^a | \sigma_a, \sigma_z^a \rangle \times \langle \sigma_b, \sigma_z^b | O_R | \sigma_b, \sigma_z^b \rangle \quad (B.9)$$

for the $N_8$-spin basis, use the result for the same procedure for the $N_4$-spin basis, and eventually propagated to the $N$-spin basis. The operation needed for each stage of the heirarchy is the same and the computation is essentially the aforementioned trivial “reverse lookup” after having the indices and value of one ket modified by $O$ of the smaller “parent” basis. For $J_{dd}$ in the largest $N$-spin basis, the indices and values of both left and right kets are modified:

$$\langle \sigma_a, \sigma_z^a | O_L | \sigma_a, \sigma_z^a \rangle \langle \sigma_b, \sigma_z^b | O_R | \sigma_b, \sigma_z^b \rangle,$$ (B.10)

and again these are just trivial lookup operations. The octapartite division groups the sites such that sites from a particular group are neighbored by sites of only one other group in a particular direction (Fig.3.1). If one expand the spin basis in terms of the eight sublattice (Eq. 3.13), one can immediately see that the special neighboring pattern leads to simple computation which is the above simultaneous
$O_{L/R}$ lookup followed by a possible swap of left and right ket.

\[ \langle \sigma_u, \sigma^+_u | O_R | \sigma_b, \sigma^+_b \rangle \langle \sigma_b, \sigma^+_b | O_L | \sigma_u, \sigma^+_u \rangle, \]  

(B.11)

Due to the fact that the reverse lookups are performed according to the indices $s_a, a, s_b, b$, the need of swapping the left and right ket does not introduce any complication.

Finally, I note that the computation for $K_{ML}(\alpha, \beta), T_{l-l'}, S_{ML}$, and each bond of the $J_{dd}$ terms are independent of one another and thus trivially parallelizable. The bottleneck of the overall computation is actually the vector size that can be accommodated by the eigen-solvers. The coarse-grain truncation discussed here leads to vector size of $< 10 GB$ for the undoped case with 64 spins and the two-hole case with 32 $CuO_2$ unit cells.