Particle-Hole Asymmetry in Doped Mott Insulators: Implications for Tunneling and Photoemission Spectroscopies

Mohit Randeria\(^{(1,2)}\), Rajdeep Sensarma\(^{(2)}\), Nandini Trivedi\(^{(1,2)}\), Fu-Chun Zhang\(^{(3)}\)

\(^{(1)}\) Physics Department, The Ohio State University, Columbus, OH 43210
\(^{(2)}\) Tata Institute of Fundamental Research, Mumbai 400 005, India
\(^{(3)}\) Physics Department, University of Hong Kong, Hong Kong

Physics Department, University of Cincinnati, Cincinnati, OH 45221

In a system with strong local repulsive interactions it should be more difficult to add an electron than to extract one. We make this idea precise by deriving various exact sum rules for the one-particle spectral function independent of the details of the Hamiltonian describing the system and of the nature of the ground state. We extend these results using a variational ansatz for the superconducting ground state and low lying excitations. Our results shed light on the striking asymmetry in the tunneling spectra of high T\(_c\) superconductors and should also be useful in estimating the local doping variations in inhomogeneous materials.

With the discovery of high temperature superconductivity in the cuprates there has been enormous interest in the properties of doped Mott insulators. One-particle spectroscopies like angle resolved photoemission (ARPES) \(^{[1,2]}\) and scanning tunneling microscopy (STM) \(^{[3,4]}\) have played a major role in our understanding of these strongly correlated materials. In this paper we examine in detail sum rule constraints on the single-electron spectral function, focusing in particular on the striking asymmetry between occupied and unoccupied spectral weights in lightly doped Mott insulators and its doping dependence. Anderson \(^{[5]}\) has recently emphasized that such an asymmetry is very unusual in conventional metallic systems, especially on low energy scales, and has suggested that it may be a characteristic signature of “projection” into a low-energy subspace where strong local Coulomb repulsion makes double occupancy at a site energetically prohibitive.

In this letter, we first make this idea precise by deriving several exact sum rules for the \(T = 0\) spectral function for both the occupied and unoccupied spectral weights in the low-energy subspace. These results are very general and do not depend either upon the details of the Hamiltonian or on any assumptions about the nature of the ground state or low-lying excitations. We show that these general results can be useful in several ways two ways. First, they help in quantifying the particle-hole asymmetry in tunneling spectroscopy: there is much more weight on the negative bias (occupied) side than on the positive bias (unoccupied) side. Second, our results can be used in present day STM experiments to estimate the local doping variations in inhomogeneous systems.

We next separate out the coherent quasiparticle (QP) and the incoherent parts of the spectral function and predict their variation with hole doping in the d-wave superconducting state obtained upon doping a Mott insulator \(^{[6]}\). This second set of results are variational in nature and require us to make assumptions about the ground state and low-lying QP excitations, and also to make the Gutzwiller approximation \(^{[10,11]}\) to obtain analytical results. These results are testable in photoemission and inverse-photoemission experiments.

**Exact Sum Rules**

Consider a system of electrons described by the Hamiltonian \(H = K + U \sum_i n_{i\uparrow} n_{i\downarrow}\) where \(K\) is the Kinetic energy operator which can be an arbitrary tight binding Hamiltonian with terms of order \(t\) (nearest neighbor hopping). We can further add to \(H\) a random one-body potential which could make the system inhomogeneous, as well as other potential energy terms such as longer range Coulomb interactions. We will always work in the limit where \(U\) is much larger than all other energy scales.

The one-electron spectral function is defined by

\[
A(r, r'; \omega) = \frac{1}{\pi} \text{Im} G(r, r'; \omega + i0^+) \quad \text{where} \quad G = \text{Green's function.}
\]

We work in real space for two reasons: first, the no-double occupancy constraint is best written in this basis, and second, this allows us to describe spatially inhomogeneous systems which is important to discuss STM experiments on the cuprates. We use the \(T = 0\) spectral representation

\[
A(r, r'; \omega) = \sum_m \left[ \langle 0| c_{r\sigma}^\dagger |m\rangle \langle m| c_{r'\sigma}|0\rangle \delta(\omega + E_m - E_0) \right] \\
+ \sum_m \left[ \langle 0| c_{r\sigma}^\dagger |m\rangle \langle m| c_{r'\sigma}|0\rangle \delta(\omega - E_m + E_0) \right]
\]

where \(|m\)'s are exact many-body eigenstates with energy \(E_m\) with \(m=0\) the ground state, and \(\omega\) is measured with respect to the chemical potential.

The large \(U\) suppresses double-occupancy at each site, and its effects on the ground state and low-lying excitations are best described using the the projection operator \(\mathcal{P} = \prod_r (1 - n_{r\uparrow} n_{r\downarrow})\). We then make the well-known unitary transformation \(^{[12]}\) \(\exp(-iS)\), such that \(\exp(iS)H\exp(-iS)\) has no matrix elements connecting states which differ in their double-
occupancy, to any given order in \( t/U \). To leading order \( iS = -(1/U) \sum_{r,r',\sigma} t_{rr'} \sigma (n_{r\sigma} c_{r\sigma}^\dagger h_{r\sigma} - h.c.) + O(t/U)^2 \), where \( h_{r\sigma} = 1 - n_{r\sigma} \). It is useful to incorporate this unitary transformation on the states, which is equivalent to transforming all operators and have introduced an equivalent to transforming this unitary transformation on the states, which is equivalent to transforming all operators and have introduced the notation: \( \hat{c}_{r\sigma} = \exp(iS)c_{r\sigma}\exp(-iS) \) and \( \hat{c}_{r\sigma}^\dagger = \exp(iS)c_{r\sigma}^\dagger\exp(-iS) \). We now use \( \sum_m |\Phi_m\rangle\langle\Phi_m| = 1 \), since the \( |\Phi_m\rangle \)'s are unprojected states, to obtain

\[
\int_0^{\Omega_L} d\omega A(r, r'; \omega) = \langle \Phi_0 | \mathcal{P}\hat{c}_{r\sigma}\mathcal{P}\hat{c}_{r'\sigma}^\dagger \mathcal{P}|\Phi_0 \rangle. \tag{4}
\]

To simplify this, we must calculate \( \mathcal{P}\hat{c}_{r\sigma}\mathcal{P} \). To order \( t/U \) we find: \( \mathcal{P}\hat{c}_{r\sigma}\mathcal{P} = h_{r\sigma}c_{r\sigma}\mathcal{P} + \frac{1}{2} \sum_{r,r',\sigma} t_{rr'}h_{r\sigma}c_{r\sigma}^\dagger c_{r'\sigma}n_{r\sigma}c_{r\sigma}^\dagger \mathcal{P} \) and \( \hat{c}_{r\sigma}^\dagger \) is the hermitian conjugate, where \( h_{r\sigma} = 1 - n_{r\sigma} \) and \( \sigma = -\sigma \).

We thus obtain the sum rule for low energy spectral weight on the unoccupied side:

\[
\int_0^{\Omega_L} d\omega N(r; \omega) = 2x(r) + 2 \langle (K(r)) | U. \tag{5}
\]

where \( \langle (K(r)) \rangle = \langle \Phi_0 | \sum_{r,\sigma}^t t_{rr'} (c_{r\sigma}^\dagger c_{r\sigma} + h.c.) \mathcal{P}|\Phi_0 \rangle \). The first term in (3) simply says that one can inject an electron into any of the \( x \) empty sites, with the factor of two for spin degeneracy. The second term gives an order \( (t/U) \) correction since the injected electron can create a temporary double occupancy and then hop off to a neighboring empty site. We note that, in contrast to this, the corresponding result in [6] to extract an electron is exact to all orders in \( t/U \).

For a translationally invariant system, we may rewrite the above result as

\[
\sum_k \int_0^{\Omega_L} d\omega A(k, \omega) = x + \langle (K) | /U. \tag{6}
\]

There is another simple result that can be obtained in the translationally invariant case. First we simplify the right hand side of (6) using the lowest order expressions for \( \mathcal{P}\hat{c}_{r\sigma}\mathcal{P} \) and \( \mathcal{P}\hat{c}_{r\sigma}\mathcal{P} \). A straightforward calculation then shows

\[
\int_0^{\Omega_L} d\omega A(r, r'; \omega) = (1 + x)/2 - \langle \hat{c}_{r\sigma}^\dagger \hat{c}_{r\sigma} \rangle + O(t/U).
\]

Fourier transforming to \( k \)-space and using standard expressions of \( n(k) \), we find that the total low-energy spectral weight is

\[
\int_{-\infty}^{\Omega_L} d\omega A(k, \omega) = \frac{1 + x}{2} + O(t/U). \tag{7}
\]

for each \( k \). Note that the deficit from unity comes from spectral weight in the “upper Hubbard band” which lies above \( \Omega_L \).

**Variational and Gutzwiller approximation results**

We emphasize that no approximations were made to obtain the above results, and we also made no assumptions about the nature of the ground state or low-lying excitations.

We now turn to translationally invariant systems and our goal is to obtain more detailed information about
the spectral function: to decompose it into its coherent and incoherent pieces and determine their doping dependence. We take the (variational) ground state to be a projected d-wave BCS state $|0\rangle = \exp(iS)P|d_{BCS}\rangle$ which has given much insight into the phenomenology of the superconducting state of the high Tc cuprates [8, 9]. Further the (variational) quasiparticle (QP) excitations [11] above this ground state are described by $|k\rangle = \exp(iS)\mathcal{P}_{\gamma^+}^{k\downarrow}|d_{BCS}\rangle$, where $\gamma^+$ is the standard Bogoliubov QP operator. The QPs lead to the coherent part of $A(k, \omega)$, i.e., delta-functions in $\omega$ at $T=0$. Finally we make the Gutzwiller approximation (GA) [10, 11] which greatly simplifies calculations of matrix elements in this strongly interacting system and gives answers which are in good agreement with exact Monte Carlo results.

Making a Gutzwiller approximation (GA) for the QP matrix elements in [11] we obtain [16]

$$A(k, \omega) = Z(k)u_k^2 \delta(\omega - E_k) + Z(k)v_k^2 \delta(\omega + E_k) + A_{inc}(k, \omega)$$

(8)

where $u_k, v_k$ and $E_k$ is standard BCS notation. The first two terms in [8] are the coherent QP pieces with the same structure as in BCS theory, except that their spectral weight is suppressed by

$$Z(k) = \frac{2x}{1 + x} + \frac{8x}{U(1 + x)^2} \sum_k \epsilon_k v_k^2$$

$$+ \frac{4x}{U(1 + x)} \epsilon_k \sum_{k'} v_{k'}^2$$

(9)

where $\epsilon_k$ is the dispersion corresponding to the bare kinetic energy $K$ in the Hamiltonian. We note that, as emphasized in [8], $Z$ vanishes as one goes to the insulating state at $x = 0$, and, in fact, the GA result [9] is in excellent quantitative agreement with the variational Monte Carlo results of Paramekanti et al. [3]. The sum over all states other than single QP's in [11] leads to the incoherent part of the spectral function denoted by $A_{inc}(k, \omega)$. Although we cannot calculate its explicit form with the minimal set of assumptions we have made, its existence is necessarily demanded by exact sum rules, as shown below, which also put constraints on $A_{inc}$.

Our use of the GA to calculate matrix elements goes beyond previous applications of this approach, which have been by-and-large restricted to ground state expectation values (see, however, the work of Laughlin [17] who uses an approximation scheme closely related to GA for spectral functions). A non-trivial consistency check is provided by $n(k)$ calculated within GA, which involves only an equal-time ground state correlation and does not depend on any assumptions about QP excited states. We find $n(k) = Z(k)v_k^2 + n_{smooth}(k)$ where $n_{smooth}(k) = (1 - x)^2/2(1 + x) + \mathcal{O}(t/U)$ is a smooth function of $k$ in the entire Brillouin zone. We omit the details of the $(t/U)$ corrections here since they involve rather long expressions [16]. We note that $n(k)$ implies that there is a jump discontinuity along the zone diagonal whose magnitude is given precisely by [16] including the $(t/U)$ corrections.

We now turn to sum rule constraints on $A_{inc}$ restricting ourselves, for the most part, to leading order results in $t/U$; the next order corrections will be presented elsewhere [16]. We begin by integrating [8] from $-\infty$ to $0$ and comparing with the GA result for $n(k)$. We thus find that for each $k$

$$\int_{-\infty}^{0} d\omega A_{inc}(k, \omega) = \frac{(1 - x)^2}{2(1 + x)} + \mathcal{O}(t/U).$$

(10)

Thus for each $k$ there is non-zero incoherent spectral weight for $\omega < 0$, whose strength relative to the coherent weight $Z(k)v_k^2$ on the occupied side, increases with underdoping (decreasing $x$).

To find the incoherent spectral weight on the unoccupied side we substitute the GA spectral function [8] in the total low-energy spectral weight sum rule [7], leading to $\int_{-\infty}^{0} d\omega A_{inc}(k, \omega) = (1 - x)^2/2(1 + x) + \mathcal{O}(t/U)$ for each $k$. This together with [16] implies that $\int_{-\infty}^{0} d\omega A_{inc}(k, \omega) = \mathcal{O}(t/U)$. Given the non-negativity of spectral weight, we find that

$$A_{inc}(k, \omega > 0) = \mathcal{O}(t/U) \ll 1.$$  

(11)

The vanishing of $A_{inc}$ for $\omega > 0$ to zeroth order in $(t/U)$ is at first sight quite surprising. Although there is very little spectral weight on the unoccupied side, as seen from [10], whatever there is, is entirely dominated by the coherent piece when $U > t$.

To gain more insight into this striking result we derive it in a completely different fashion, which also shows that it is not an artifact of the Gutzwiller approximation. To zeroth order in $(t/U)$ we can set $\exp(iS) = 1$, and using the identity $\mathcal{P}_c^{\dagger} \mathcal{P} = \mathcal{P}_c^{\dagger} r, \sigma$ we find that

$$\mathcal{P}_c^{\dagger} |0\rangle = |P_{c, k, \sigma}\rangle = constant \times \mathcal{P}_{\gamma^+}^{k, \sigma}|d_{BCS}\rangle = constant \times |k, \sigma\rangle.$$  

Thus we find that the projected creation operator acting on the ground state gives precisely the coherent QP state. Thus there is no incoherent weight in the electron creation (i.e., unoccupied) sector at least to zeroth order in $(t/U)$.

Finally, we can determine the explicit form of the order $t/U$ incoherent unoccupied spectral weight, which follows from use of eqs. [10] and [11] together with the GA result [16] for $n(k)$. We find

$$\sum_k \int_0^{\Omega_{k, \sigma}} d\omega A_{inc}(k, \omega) = \frac{2x(1 - x)}{U(1 + x)} \sum_k \epsilon_k v_k^2$$

(12)

We have also obtained sum rule constraints on the density of states contributions coming separately from the coherent and incoherent parts of the spectral function, however we omit these here. In any case, in a $k$-integrated probe like tunneling it would be very hard to
separate out the coherent contribution from the incoherent one in experiments, unlike in ARPES where it seems possible to do so.

**Implications for Experiments**

We first discuss the implications of our results for tunneling spectroscopy. The tunneling conductance in STM experiments is proportional to the local density of states $G(r; eV) = \text{constant} \times N(r; \omega = eV)$ [11], where the constant of proportionality involves tunneling matrix elements. Thus our result [12] shows that the (energy integrated) positive bias conductance is small, of order $x$, while [13] implies that the (integrated) negative bias conductance is large, of order unity. This provides a qualitative explanation for the large asymmetry seen in STM experiments which show a superconducting gap structure superimposed on a sloping “background” which decreases going from negative to positive bias. Our results predict how this asymmetry should grow with underdoping (decreasing $x$). This asymmetry is most strikingly seen in the highly underdoped non-superconducting cuprates such as Na$_{x}$Ca$_{1-x}$CuO$_2$Cl$_2$ studied by Hanaguri et al. [13]. The nature of the “zero temperature pseudogap state” in such materials is an unsolved problem, and in this context it is very important to re-emphasize that our results [12] and [13] make no assumptions about the broken symmetry in the ground state or the nature of low-lying excitations.

In order to get quantitative information from STM experiments we look at ratios in which the unknown tunneling matrix elements cancel out. Taking the ratio of the total unoccupied low-energy spectral weight [14] to the total occupied spectral weight [15] we obtain

$$\frac{\int_{0}^{\Omega_L} \omega d\omega G(r; \omega)}{\int_{-\infty}^{0} \omega d\omega G(r; \omega)} = \frac{2x(r)}{1 - x(r)} + \frac{2|\langle K(r) \rangle|}{U[1 - x(r)]} \tag{13}$$

The left hand side can now be estimated from STM data, provided one can make a reasonable choice of the positive and negative high energy cutoffs [11], and then used to infer the local hole doping $x(r)$ from the first term on the right hand side of (13). The second term of order $(xt/U)$ gives an estimate of the approximately $10\%$ error made in estimating $x$.

In the second part of the paper we derived results for the doping dependence of the coherent and incoherent parts of the spectral function. The predicted $x$ dependence of the coherent weight $Z$ of eq. (11) has already been observed in ARPES studies of nodal QP’s (see ref. [15]). The important prediction for the very small incoherent spectral weight on the unoccupied side [16] should be testable in future inverse photoemission experiments once their energy resolution is improved.

**Acknowledgments** We have greatly benefited from many stimulating conversations with P. W. Anderson, who first got us interested in this problem. We acknowledge helpful discussions with J. C. Davis and thank him for keeping us abreast of his ongoing experiments. We also thank J. C. Campuzano, O. Fischer, L. H. Greene, P. Hentges, P. A. Lee and T. M. Rice for their valuable input.

[1] A. Damascelli, Z. Hussain and Z. X. Shen, Rev. Mod. Phys. 75, 473 (2003).
[2] J. C. Campuzano, M. R. Norman and M. Randeria, in “Physics of Conventional and Unconventional Superconductors”, edited by K. H. Bennemann and J. B. Ketterson, (Springer Verlag, 2004); cond-mat/0209476.
[3] Ch. Renner et al., Phys. Rev. Lett. 80 149 (1998).
[4] S. H. Pan et al., Nature, 413 282 (2001); J. E. Hoffman et al., Science 297, 1148 (2002); K. Mc Kelroy et al., Nature 422, 592 (2003); H. Hanaguri et al., Nature 430 1001 (2004).
[5] P. W. Anderson (private communication) and P. W. Anderson and N. P. Ong, (preprint) cond-mat/0405518.
[6] See, e.g., the tunneling spectra in K. Mc Kelroy et al. and especially Fig. 1(c) and 3(e) in H. Hanaguri et al. [3].
[7] J. C. Davis, (private communication).
[8] P. W. Anderson, Science 235 1196 (1987).
[9] A. Paramekanti, M. Randeria and N. Trivedi, Phys. Rev. Lett. 87 217002 (2001); Phys. Rev. B 69 144509 (2004); Phys. Rev. B 70, 054504 (2004).
[10] P. W. Anderson, P. A. Lee, M. Randeria, T. M. Rice, N. Trivedi and F. C. Zhang, J. Phys. Cond. Mat. 16 R755 (2004).
[11] F. C. Zhang, C. Gros, T. M. Rice, and H. Shiba, Supercond. Sci. Tech. 1 36 (1988).
[12] C. Gros, R. Joynt, and T. M. Rice, Z. Phys. B 68 425 (1987); A. H. MacDonald, S. M. Girvin, and D. Yoshioka, Phys. Rev. B 37 9753 (1988).
[13] In eq. (4) the sum over intermediate states can be simply extended over all states, so that $\sum_{m}|m\rangle\langle m| = 1$, even though only states in the LHB actually contribute.
[14] M. Randeria et al., Phys. Rev. Lett. 74 4951 (1995).
[15] H. Eskes, A. M. Oles, M. B. J. Meinders, and W. Stephan, Phys. Rev. B 50 17980 (1994).
[16] R. Sensarma, M. Randeria and N. Trivedi, (unpublished).
[17] R. B. Laughlin, (preprint cond-mat/0209269).
[18] The use of the well known formula relating the tunneling conductance of planar junctions to the total density of states $G(eV) \propto \sum_{k} A(k, \omega = eV)$ makes the further assumption of $k$-independent tunneling matrix elements.
[19] In analyzing experimental data, the cutoff $\Omega_L$ should be set by the charge transfer gap $\approx 1.5 - 2.0$ eV. On the negative bias side, the ARPES dispersion in Bi2212 shows that the “bottom of the band” is $\approx -0.5$ eV, but this is only the centroid of a very broad spectral function. Other states begin to show up below about $-1$ eV in ARPES [12]. Thus replacing the “$-\infty$” by a cutoff $\approx -1.0$ eV seems reasonable.