Journal of Statistical Physics, Vol. 54, Nos. 1/2, 1989

Thermodynamics and Phase Transitions in the Overhauser Model

N. G. Duffield[1] **and J. V. Pulé**[1,2]

Received April 19, 1988; revision received August 9, 1988

We analyze the thermodynamics of the Overhauser model and demonstrate rigorously the existence of a phase transition. This is achieved by extending techniques previously developed to treat the BCS model in the quasi-spin formulation. Additionally, we compare the thermodynamics of the quasi-spin and full-trace BCS models. The results are identical up to a temperature rescaling.

KEY WORDS: Overhauser model; thermodynamic limit; phase transitions; large deviations; Berezin–Lieb inequalities; BCS model.

1. INTRODUCTION

In two previous works[1,2] we introduced a new method for calculating the thermodynamics of the full BCS model[3] (i.e., with nonconstant energies and interactions) in the quasi-spin formulation.[4] The method has already found further application: it has been used to solve the full spin-boson model.[5]

In this paper we extend the method to treat the Overhauser model[6] with reduced Hamiltonian[7]

$$H = \sum_{k,s=\pm 1} \frac{k^2}{2} a^*_{k,s} a_{k,s} - \frac{1}{V} \sum_{k,k'} U(k,k')\, a^*_{k+q/2,+1} a_{k-q/2,-1} a^*_{k'-q/2,-1} a_{k'+q/2,+1} \tag{1.1}$$

where the $a^{\#}$ are fermionic operators, k is a momentum index, q is a fixed momentum vector, and ± 1 indicates spin up or down. In ref. 7 a Hamiltonian of this form was used within an algebraic framework.

[1] Department of Mathematical Physics, University College, Belfield, Dublin 4, Ireland.
[2] Dublin Institute for Advanced Studies, Dublin, Ireland.

0022-4715/89/0100-0449$06.00/0

The Hamiltonian models electronic interactions in certain metals, most notably in chromium. Below a critical temperature the interactions give rise to coherent excitation of a spin-density wave with wave vector q. Experimentally, the amplitude of the spin-density wave has a temperature dependence strikingly similar to that of the energy gap in the BCS model. We shall see that there are indeed great mathematical similarities between the two models. Note that the Hamiltonian is of mean-field type: although the interaction is not homogeneous (i.e., U is not constant), it is scaled by a factor $1/V$.

In order to calculate the thermodynamics when $q \neq 0$, it is necessary to go beyond the quasi-spin approach of refs. 1 and 2. The reason for this is the following: in the quasi-spin approach one restricts one's attention to the subset of states in Fock space which have nonzero matrix elements with the second (interaction) term of the Hamiltonian. For the H in (1.1) this is the set of states for which all the pairs of levels $\{(k+q/2, +1), (k-q/2, -1)\}$ are singly occupied. When calculating the partition function $Z = \text{trace}\, \exp(-\beta H)$, the trace is taken over these states only. However, states in which the pairs are either full or empty also make a contribution. We shall see that this contribution is significant when $q \neq 0$.

It turns out that when $q = 0$ the free energy calculated using the full trace is essentially the same as that calculated in the quasi-spin formulation. We demonstrate this fact for the BCS model with Hamiltonian

$$H_{\text{BCS}} = H_0 + -\frac{1}{V} \sum_{k,k'} U(k, k')\, a^*_{k,+1} a^*_{-k,-1} a_{-k',-1} a_{k',+1}$$

where H_0 is quadratic in the momentum operators. As we said before, the two models are mathematically very similar. The physical difference lies in the pairing of states. The BCS interaction has nonzero matrix elements only between states for which the pairs of levels $\{(k, +1), (-k, -1)\}$ (the Cooper pairs) are either full or empty.

We now describe in more detail the methods used in this paper. The first task is to write the model in a form we can treat by using techniques from ref. 2. This is done by decomposing Fock space as the sum of a number of subspaces, each of which has a fixed number of pairs of levels either full or empty. On each of these subspaces the partial trace over singly occupied pairs is trivial to perform; the partial trace over the remaining pairs is exactly that which occurs in the calculation of the partition function for the BCS model in the quasi-spin formulation.

The decomposition of Fock space mentioned gives rise to a degeneracy in which each contribution to the partition function occurs with a certain multiplicity. These multiplicities can be used to construct a volume indexed family of probability measures. We show that this family of measures satisfies a *large-deviation principle.*[8,9] Roughly speaking, the theory of large deviations allows us to obtain a variational expression for the largest contribution to the free energy in the thermodynamic limit. Thus, this work can be seen as making rigorous the approach in ref. 10 to treating the BCS model.

We now summarize our previous treatment of the full BCS model. Our method consists in partitioning the system into smaller subsystems and approximating by Hamiltonians which are constructed from functions of total quasi-spin operators for these subsystems. The approximating Hamiltonians can be treated by an extension of the techniques in ref. 11. There, a large-deviation principle was found for the measures arising from the multiplicities of the irreducible representations in the decomposition of the total spin operator. Berezin–Lieb inequalities[12–14] were used to obtain upper and lower bounds for the free energy, bounds which coincided in the thermodynamic limit.

The layout of the remainder of this paper is as follows. In Section 2 we define the model explicity. We immediately define the approximating Hamiltonians, and prove that the approximation becomes exact in the thermodynamic limit (Theorem 1). We then carry out the two decompositions outlined above, and rewrite the free energy in terms of the corresponding family of measures. In the Appendix we prove a large-deviation principle for this family (Theorem 6). In combination with our previous results on the quasi-spin BCS model (summarized in Theorem 3), this enables us to obtain the variational principle for the free energy density in the thermodynamic limit (Theorem 7). In Section 3 we analyze the variational problem. We have attempted to calculate the free energy, but our results here are incomplete: we are unable to identify the phase transition precisely. We show that there are inverse temperatures β_t and $\tilde{\beta}_t$, $0<\tilde{\beta}_t \leqslant \beta_t \leqslant \infty$, such that if $\beta < \tilde{\beta}_t$, then the system has the free energy density of a free electron gas, while if $\beta > \beta_t$, it has a different form, corresponding to the excitation of a spin-density wave. Finally, in Section 4 we give an outline of the application of the methods of Sections 2 and 3 to the full-trace BCS model. It turns out that the variational analysis can be carried through completely, and gives (up to a temperature rescaling) the same results as for the quasi-spin formulation of the model.

2. THE THERMODYNAMIC LIMIT

2.1. The Model

We consider a slightly more general model than that described in the introduction. Let $\{\Lambda_l : l = 1, 2,...\}$ be a sequence of regions of Euclidean space $\mathbb{R}^\nu$ and denote the volume of Λ_l by V_l; we associate with the region Λ_l the sequence of momenta $\{k_l(j): j=1, 2,...\}$, where each $k_l(j)$ is in $\mathbb{R}^\nu$. We make the assumption that the sequence of measures $\{\mu_l\}$ giving the distribution of momentum states

$$\mu_l(B)=\frac{1}{V_l}\#\{j: k_l(j)\in B\}$$

for Borel subsets B of $\mathbb{R}^\nu$ converges weakly to a measure μ which is absolutely continuous with respect to Lebesgue measure. We use this property at a technical point in the variational problem. It is satisfied by, for example, the momentum distribution of free electrons in an increasing sequence of cubes with reasonable boundary conditions.[16] We shall be considering only those momenta in a *cutoff* region Ω which is a closed, bounded region in $\mathbb{R}^\nu$ and we assume that $\mu(\Omega)<\infty$.

With each Λ_l we associate a one-particle Hilbert space $\mathscr{H}_l$ for the electrons. With each $k_l(i)$ in Ω we associate a pair of vectors in $\mathscr{H}_l$ which we label $\psi_l(j, +)$ and $\psi_l(j, -)$: $\pm$ designates spin up or down. The set $\{\psi_l(j, \pm): k_l(j)\in\Omega\}$ is taken to form a complete orthonormal basis in $\mathscr{H}_l$. The physical momenta of the electrons with wavefunction $\psi_l(j, +)$ and $\psi_l(j, -)$ are $k_l(j)+\frac{1}{2}[q+\delta_l^+(j)]$ and $k_l(j)-\frac{1}{2}[q+\delta_l^-(j)]$, respectively: this *defines* the quantities $\delta_l^\pm(i)$, and $\frac{1}{2}[\delta_l^+(i)+\delta_l^-(i)]$ measures the deviation of the difference of two momenta from the fixed momentum q. With $\delta_l=\max\{\|\delta_l^\pm(j)\|: k_l(j)\in\Omega\}$ we assume that the pairing of one-particle vectors is carried out so that

$$\lim_{l\to\infty}\delta_l=0 \tag{2.1}$$

Let $\tilde{\mathscr{F}}_l$ be the antisymmetric Fock space over $\mathscr{H}_l$ and $a^*(j, \pm)$, $a(j, \pm)$ denote the creation and annihilation operators on $\tilde{\mathscr{F}}_l$ for the states $\psi_l(j, \pm)$. These satisfy the usual anticommutation relations

$$\{a(j, \mu), a^*(j', \mu')\}=\delta_{\mu,\mu'}\delta_{j,j'}$$

The Hamiltonian for the problem $\tilde{H}_l$ acts on $\tilde{\mathscr{F}}_l$ and is given by

$$\begin{aligned}\tilde{H}_l=&\frac{1}{2}\sum_{j=1}^{N_l}\left\{k_l(j)+\frac{1}{2}[q+\delta_l^+(j)]\right\}^2 a^*(j, +)\,a(j, +)\\&+\frac{1}{2}\sum_{j=1}^{N_l}\left\{k_l(j)-\frac{1}{2}[q+\delta_l^-(j)]\right\}^2 a^*(j, -)\,a(j, -)\\&-\frac{1}{V_l}\sum_{i,j=1}^{N_l}U(k_l(i), k_l(j))\,a^*(i, +)\,a(i, -)\,a^*(j, -)\,a(j, +)\end{aligned} \tag{2.2}$$

where $N_l = V_l \mu_l(\Omega)$ is the number of pairs of electron levels and $U \in C(\Omega, \Omega)$ is a symmetric function. This completes the definition of the model.

Remark. The dispersion relation $k \mapsto k^2$ is not crucial for what follows. The whole analysis could be carried through for dispersion relations $k \mapsto E_s(k)$ for spin $s = \pm 1$, where the E_s are continuous functions.

2.2. The Approximating Hamiltonians

Let $f_l(\beta)$ be the free energy density

$$f_l(\beta) = -\frac{1}{\beta V_l} \text{trace} \exp -\beta \tilde{H}_l$$

We shall prove that $f_l(\beta)$ converges as $l \to \infty$ and shall obtain a variational formula for the limit $f(\beta)$. The variational problem is analyzed in Section 3. Our method is to approximate $\tilde{H}_l$ by a Hamiltonian for which the method of ref. 11 can be used. Choose $L > 0$ such that $\Omega \subset [-L, L]^\nu$, and for each $M \in \mathbb{N}$ partition $[-L, L]$ into M^ν cubes of side $2L/M$. Denote these cubes by $\{\bar{B}_m^M : m = 1, 2, ..., M^\nu\}$ and let $B_m^M = \bar{B}_m^M \cap \Omega$. It is convenient to define the quantities

$$\eta_l^\pm(i) = \tfrac{1}{2}\{k_l(i)^2 + \tfrac{1}{4}[q + \delta_l^\pm(i)]^2\}$$

and

$$\varepsilon_l^\pm(i) = k_l(i) \cdot [q + \delta_l^\pm(i)]$$

We define the approximating Hamiltonian $\tilde{H}_l^M$ by

$$\begin{aligned} \tilde{H}_l^M = & \sum_{m=1}^{M^\nu} \left(\eta_m^M + \frac{1}{2}\varepsilon_m^M\right) \sum_{k_l(i) \in B_m^M} a^*(i, +)\, a(i, +) \\ & + \sum_{m=1}^{M^\nu} \left(\eta_m^M - \frac{1}{2}\varepsilon_m^M\right) \sum_{k_l(i) \in B_m^M} a^*(i, -)\, a(i, -) \\ & - \frac{1}{V_l} \sum_{m,m'=1}^{M_\nu} U_{m,m'}^M \left(\sum_{k_l(i) \in B_m^M} a^*(i, +)\, a(i, -) \right. \\ & \left. \times \sum_{k_l(j) \in B_{m'}^M} a^*(j, -)\, a(j, +) \right) \end{aligned} \tag{2.3}$$

where

$$\eta_m^M = \begin{cases} 0 & \text{if} \quad \mu(B_m^M) = 0 \\ \dfrac{1}{\mu(B_m^M)} \displaystyle\int_{B_m^M} \mu(dk) \frac{1}{2}\left(k^2 + \frac{1}{4} q^2\right) & \text{otherwise} \end{cases}$$

$$\varepsilon_m^M = \begin{cases} 0 & \text{if} \quad \mu(B_m^M) = 0 \\ \dfrac{1}{\mu(B_m^M)} \displaystyle\int_{B_m^M} \mu(dk) k \cdot q & \text{otherwise} \end{cases}$$

$$U_{m,m'}^M = \begin{cases} 0 \quad \text{if} \quad \mu(B_m^M)\, \mu(B_{m'}^M) = 0 \\ \dfrac{1}{\mu(B_{m'}^M)\, \mu(B_m^M)} \displaystyle\int_{B_m^M \times B_{m'}^M} \mu(dk)\, \mu(dk')\, U(k, k') & \text{otherwise} \end{cases}$$

The use of this definition is made clear by the following theorem.

Theorem 1. Let

$$f_l^M(\beta) = -\frac{1}{\beta V_l} \operatorname{trace} \exp -\beta \tilde{H}_l^M$$

Then $\lim_{M \to \infty} |f_l^M(\beta) - f_l(\beta)| \leqslant O(\delta_l)$ for sufficiently large l. Hence

$$\lim_{l \to \infty} f_l(\beta) = \lim_{l \to \infty} \lim_{M \to \infty} f_l^M(\beta) \tag{2.4}$$

when the limits on the rhs of (2.4) exist.

Proof. By the Peierls–Bogoliubov inequality,

$$|f_l^M(\beta) - f_l(\beta)| \leqslant \frac{1}{V_l} \|\tilde{H}_l^M - \tilde{H}_l\| \tag{2.5}$$

Therefore

$$\begin{aligned} &|f_l^M(\beta) - f_l(\beta)| \\ &\quad \leqslant \sum_{m=1}^{M^\nu} \mu_l(B_m^M) \left[\sup_{k_l(i) \in B_m^M, \pm} |\varepsilon_m^M - \varepsilon_l^\pm(i)| + 2 \sup_{k_l(i) \in B_m^M, \pm} |\eta_m^M - \eta_l^\pm(i)| \right] \\ &\quad\quad + \sum_{m,m'=1}^{M^\nu} \mu_l(B_m^M)\, \mu_l(B_{m'}^M) \sup_{\substack{k_l(i) \in B_m^M, \\ k_l(j) \in B_{m'}^M}} |U_{m,m'}^M - U(k_l(i), k_l(j))| \end{aligned} \tag{2.6}$$

By virtue of the continuity of U and of the function $k \mapsto k^2$

$$\lim_{M \to \infty} |f_l^M(\beta) - f_l(\beta)| \leqslant c_1 \delta_l + c_2(\delta_l)^2$$

for some positive constants c_1 and c_2, and so the first statement of the theorem follows. ∎

We next give a variational expression for the limiting approximate free energy density.

Theorem 2. $f^M(\beta) = \lim_{l \to \infty} f_l^M(\beta)$ exists and is given by

$$f^M(\beta) = -\sup\{\mathscr{S}^M(\beta; t, r, \theta, \phi): r, t \in [0, 1]^{M^\nu}, \theta \in [0, \pi]^{M^\nu}, \phi \in [0, 2\pi]^{M^\nu}\} \tag{2.7}$$

where

$$\begin{aligned}
&\mathscr{S}^M(\beta; t, r, \theta, \phi) \\
&\quad = \frac{1}{\beta} \sum_{m=1}^{M^\nu} \mu(B_m^M) \log[4 \exp(-\beta \eta_m^M) \cosh \beta \eta_m^M] \\
&\quad\quad - \frac{1}{\beta} \sum_{m=1}^{M^\nu} \mu(B_m^M)\, t_m \log \cosh \beta \eta_m^M - \frac{1}{\beta} \sum_{m=1}^{M^\nu} \mu(B_m^M)\, J(t_m) \\
&\quad\quad + \frac{1}{4} \sum_{m,m'=1}^{M^\nu} \mu(B_m^M)\, \mu(B_{m'}^M)\, U_{m,m'}^M t_m t_{m'} r_m r_{m'} \\
&\quad\quad \times \sin \theta_m \sin \theta_{m'} \cos(\phi_m - \phi_{m'}) \\
&\quad\quad + \frac{1}{2} \sum_{m=1}^{M^\nu} \mu(B_m^M)\, t_m r_m \varepsilon_m^M \cos \theta_m - \frac{1}{\beta} \sum_{m=1}^{M^\nu} \mu(B_m^M)\, t_m I(r_m)
\end{aligned} \tag{2.8}$$

where

$$J(t) = t \log t + (1 - t) \log(1 - t) + \log 2 \tag{2.9a}$$

$$I(r) = \tfrac{1}{2}[(1 + r) \log(1 + r) + (1 - r) \log(1 - r)] \tag{2.9b}$$

The remainder of this section is concerned with the proof of this result.

2.3. Reduction to the Quasi-Spin Model

We introduce another Hilbert space $\mathscr{F}_l$ and a unitary map W_l from $\tilde{\mathscr{F}}_l$ to $\mathscr{F}_l$. We choose the basis vectors in $\mathscr{F}_l$ as follows. Let

$$\begin{aligned}
\psi_l(j,1) &= 1 \in \mathbb{C} \\
\psi_l(j,2) &= \psi_l(j,+) \\
\psi_l(j,3) &= \psi_l(j,-) \\
\psi_l(j,4) &= P(\psi_l(j,-) \otimes \psi_l(j,+))
\end{aligned} \tag{2.10}$$

where P is the total antisymmetrization operator. Then our basis of vectors is $\tilde{\mathscr{F}}_l$ is the set of vectors

$$\psi_l(v_1,..., v_{N_l}) = P(\psi_l(1, v_1) \otimes \cdots \otimes \psi_l(N_l, v_{N_l})) \tag{2.11}$$

where $v_j \in \{1, 2, 3, 4\}$ for $j \in I_l = \{1, 2,..., N_l\}$. As $\mathscr{F}_l$ we choose the space $(\mathbb{C}^4)^{\otimes N_l}$. Let $\phi(v)$: $v = 1, 2, 3, 4$ be the vectors $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$, and $(0, 0, 0, 1)$, respectively, in $\mathbb{C}^4$ and define

$$\phi_l(v_1,..., v_{N_l}) = \phi(v_1) \otimes \cdots \otimes \phi(v_{N_l}) \tag{2.12}$$

The ϕ_l form a basis for $\mathscr{F}_l$, and we define the unitary transformation W_l by

$$W_l \psi_l(v_1,..., v_{N_l}) = \phi_l(v_1,..., v_{N_l}) \tag{2.13}$$

For a products of two fermion operators $a^{\#}(j, v)\, a^{\#}(j, v')$ the action of its transform under $W_l \cdot W_l^{-1}$ is simple to work out (this is not the case for single fermion operators):

$$W_l a^*(j,+)\, a(j,+)\, W_l^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_j \otimes \mathbf{1}_{I_l \setminus j} \tag{2.14a}$$

$$W_l a^*(j,-)\, a(j,-)\, W_l^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_j \otimes \mathbf{1}_{I_l \setminus j} \tag{2.14b}$$

$$W_l a^*(j,-)\, a(j,+)\, W_l^{-1} = C_j = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_j \otimes \mathbf{1}_{I_l \setminus j} \tag{2.14c}$$

where the j subscript indicates that the matrix operates on the jth component of the tensor product (2.12). We can write the transformed Hamiltonian $H_l^M = W_l \tilde{H}_l^M W_l^{-1}$ as

$$H_l^M = \sum_{m=1}^{M^\nu} \sum_{i:k_l(i)\in B_m^M} \eta_m^M \mathbf{1} + \sum_{m=1}^{M^\nu} \sum_{i:k_l(i)\in B_m^M} \left(\frac{1}{2}\varepsilon_m^M A_i + \eta_m^M B_i\right)$$
$$-\frac{1}{V_l}\sum_{m,m'=1}^{M^\nu} U_{m,m'}^M \sum_{i:k_l(i)\in B_m^M} C_i^* \sum_{j:k_l(j)\in B_{m'}^M} C_j \tag{2.15}$$

where the operators $\{A_j, B_j\colon j=1,\dots, N_l\}$ have the form

$$A_j = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_j \otimes \mathbf{1}_{I_l\setminus j}, \qquad B_j = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_j \otimes \mathbf{1}_{I_l\setminus j} \tag{2.16}$$

In order to calculate $f_l^M(\beta)$ we carry out a decomposition of the space $\mathscr{F}_l$. For each $Y \subseteq I_l$ define the projections

$$P_{l,Y} = \bigotimes_{i\in Y} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_i \tag{2.17a}$$

and

$$Q_{l,Y} = \bigotimes_{j\in I_l\setminus Y} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_j \tag{2.17b}$$

and let $\mathscr{P}_{l,Y} = P_{l,Y} \otimes Q_{l,Y}$. In $\tilde{\mathscr{F}}_l$, $W_l^{-1}\mathscr{P}_{l,Y} W_l$ is the projection onto that subspace comprising pairs of levels $\psi_l(j,+)$ and $\psi_l(j,-)$ which are half-filled when $j\in Y$ and either full or empty when $j \in I_l\setminus Y$. One demonstrates readily that

$$\sum_{Y\subseteq I_l} \mathscr{P}_{l,Y} = \mathbf{1}_{I_l}; \qquad [Q_{l,Y}, H_l^M] = [P_{l,Y}, H_l^M] = 0 \tag{2.18}$$

(In fact, the last two equations hold for H_l as well as H_l^M: the approximation is not crucial at this stage.)

Denoting by $\mathscr{F}_{l,Y}$ the range of $\mathscr{P}_{l,Y}$, we can decompose trace $\exp(-\beta H_l^M)$ over the subspaces $\{\mathscr{F}_{l,Y}\colon Y\subseteq I_l\}$. For any $Y \subseteq I_l$, define

$S_l^M(Y)=\{i\in Y: k_l(i)\in B_m^M\}$ [so that in particular $S_m^M(I_l)=\{i: k_l(i)\in B_m^M\}$]. Then

$$\begin{aligned}\text{trace}\exp-\beta H_l^M &= \sum_{Y\subseteq I_l}\text{trace}\,\mathscr{P}_{l,Y}\exp(-\beta H_l^M)\,\mathscr{P}_{l,Y}\\ &=\exp\left(-\beta\sum_{m=1}^{M^\nu}\sum_{i\in S_m^M(I_l)}\eta_m^M\right)\\ &\quad\times\sum_{Y\subseteq I_l}\text{trace}\exp-\beta H_{l,Y}^M \qquad (2.19)\end{aligned}$$

where $H_{l,Y}^M$ is the restriction of H_l^M to $\mathscr{F}_{l,Y}$. Identifying $\mathscr{F}_{l,Y}$ with $\bigotimes_{i=1}^{I_l}\mathbb{C}^2$, we can write $H_{l,Y}^M$ as

$$H_{l,Y}^M=H_Y^M\otimes\mathbf{1}_{I_l\setminus Y}+\mathbf{1}_Y\otimes\left(\sum_{m=1}^{M^\nu}-2\eta_m^M\sum_{i\in S_m^M(I_l\setminus Y)}\bar{\sigma}_i^z\right)\qquad(2.20)$$

where

$$\begin{aligned}H_Y^M=\sum_{m=1}^{M^\nu}\varepsilon_m^M\sum_{i\in S_m^M(Y)}\sigma_i^z-\frac{1}{V_l}\sum_{m,m'=1}^{M^\nu}U_{m,m'}^M\\ \times\left(\sum_{i\in S_m^M(Y)}\sigma_i^+\right)\left(\sum_{j\in S_{m'}^M(Y)}\sigma_j^-\right)\qquad(2.21)\end{aligned}$$

and

$$\sigma_i^\#=\mathbf{1}\otimes\cdots\otimes\mathbf{1}\otimes\sigma^\#\otimes\mathbf{1}\otimes\cdots\otimes\mathbf{1}\in\mathscr{L}\left(\bigotimes_{i=1}^{\#Y}\mathbb{C}^2\right)\qquad(2.22a)$$

$$\bar{\sigma}_i^\#=\mathbf{1}\otimes\cdots\otimes\mathbf{1}\otimes\sigma^\#\otimes\mathbf{1}\otimes\cdots\otimes\mathbf{1}\in\mathscr{L}\left(\bigotimes_{i=1}^{\#I_l\setminus Y}\mathbb{C}^2\right)\qquad(2.22b)$$

where $\{\sigma^\#: \#=+,-,z\}$ are the usual Pauli matrices, the i indicating that they act on the ith component of the tensor product (2.12). Using

$$\text{trace}_{\mathscr{H}_1\otimes\mathscr{H}_2}e^{A\otimes 1+1\otimes B}=\text{trace}_{\mathscr{H}_1}e^A\,\text{trace}_{\mathscr{H}_2}e^B$$

to perform the trace over the range of $Q_{l,Y}$ for each Y,

$$\begin{aligned}&\text{trace}\exp H_l^M\\ &\quad=\exp\left(-\beta\sum_{m=1}^{M^\nu}\sum_{i\in S_m^M(I_l)}\eta_m^M\right)\\ &\qquad\times\sum_{Y\subseteq I_l}\left[\exp\sum_{m=1}^{M^\nu}\log 2\cosh\beta\eta_m^M\left(\sum_{i\in S_m^M(I_l\setminus Y)}1\right)\right]\text{trace}\exp-\beta H_Y^M\\ &\qquad\qquad(2.23)\end{aligned}$$

There is a degeneracy in the decomposition (2.23) over the subsets Y of I_l. Let $N_l^m = V_l \mu_l(B_m^M)$ and denote by χ_l^M the set of all M^ν-tuples of integers $\underline{n} = (n_1,..., n_{M^\nu})$ such that $0 \leqslant n_m \leqslant N_l^m$. For each $\underline{n}$ in χ_l^M, let

$$\Xi_l^M(\underline{n}) = \{ Y \subseteq I_l : \# S_m^M(Y) = n_m, m = 1,..., M^\nu \} \tag{2.24}$$

Then the members of the set of Hamiltonians $\{H_Y^M : Y \in \Xi_l^M(\underline{n})\}$ are unitarily equivalent to $H^M(\underline{n})$ defined by

$$H^M(\underline{n}) = \sum_{m=1}^{M^\nu} \varepsilon_m^M \sum_{i=1}^{n_m} \sigma_i^z - \frac{1}{V_l} \sum_{m,m'=1}^{M^\nu} U_{m,m'}^M \left(\sum_{i=1}^{n_m} \sigma_i^+ \right) \left(\sum_{j=1}^{n_{m'}} \sigma_j^- \right) \tag{2.25}$$

Thus, given $\underline{n}$ in χ_l^M, the contributions to (2.23) for any Y in $\Xi_l^M(\underline{n})$ are identical, and occur with a multiplicity

$$\# \Xi_l^M(\underline{n}) = \prod_{m=1}^{M^\nu} \binom{N_l^m}{n_m} \tag{2.26}$$

Hence we can write

$$\begin{aligned} &\text{trace} \exp -\beta H_l^M \\ &= \sum_{\underline{n} \in \chi_l^M} \prod_{m=1}^{M^\nu} \left\{ \binom{N_l^m}{n_m} \exp(-\beta N_l^m \eta_m^M) \right. \\ &\qquad \left. \times \exp[\log 2 \cosh \beta \eta_m^M (N_l^m - n_m)] \right\} \text{trace} \exp -\beta H^M(\underline{n}) \end{aligned} \tag{2.27}$$

The traces remaining in (2.27) have been analyzed in ref. 2. $H^M(\underline{n})$ is decomposed according to the irreducible representations of $SU(2)$ and the Berezin–Lieb inequalities are used to obtain upper and lower bounds for the trace over each component. We give only a brief account here: a fuller treatment can be found in ref. 2.

Denote by Π the irreducible unitary representation of the group $SU(2)$ acting in $\mathbb{C}^2$, and define the unitary representation Π_n of $SU(2)$ in $(\mathbb{C}^2)^{\otimes n}$ by

$$\Pi_n(g) = \underbrace{\Pi(g) \otimes \cdots \otimes \Pi(g)}_{n \text{ times}}, \qquad g \in SU(2) \tag{2.28}$$

For $n > 1$, Π_n is reducible and decomposes into a direct sum

$$\Pi_n = \bigoplus_{J \in A_n} \bigoplus_{k=1}^{c(n,J)} \Pi^{J,k} \tag{2.29}$$

where $A_n = \{0, 1, ..., n/2\}$ is n is even, and $A_n = \{1/2, 3/2, ..., n/2\}$ if n is odd, and $\Pi^{J,k}$ is a copy of the irreducible representation Π^J of $SU(2)$ acting in $\mathbb{C}^{2J+1}$, and $c(n, J)$ is the multiplicity of that representation in the decomposition. Define the operator h_l^M on $(\mathbb{C}^2)^{\otimes M^\nu}$ by

$$h_l^M = \sum_{m=1}^{M^\nu} \varepsilon_m^M \sigma_m^z - \frac{1}{V_l} \sum_{m,m'=1}^{M^\nu} U_{m,m'}^M \sigma_m^+ \sigma_{m'}^- \tag{2.30}$$

Let p_n and p^J be the representations of the Lie algebra of $SU(2)$ corresponding to Π_n and Π^J, respectively. Then for $\underline{n} \in \chi_l^M$

$$H^M(\underline{n}) = p_{n_1} \otimes \cdots \otimes p_{n_{M^\nu}} h_l^M \tag{2.31}$$

and thus

$$\text{trace}\exp -\beta H^M(\underline{n}) = \sum_{J \in A_{n_1} \times \cdots \times A_{n_{M^\nu}}} c(n_1, J_1) \cdots c(n_{M^\nu}, J_{M^\nu}) \times \text{trace}\exp -\beta\{(p^{J_1} \otimes \cdots \otimes p^{J_{M^\nu}})\, h_l^M\} \tag{2.32}$$

Let $f_l^M(\beta, \cdot)$: $[0, 1]^{M^\nu} \to \mathbb{R}$ be defined by

$$f_l^M(\beta, x) = \frac{-1}{\beta V_l} \log \text{trace}\exp -\beta\{(p^{J_1} \otimes \cdots \otimes p^{J_{M^\nu}})\, h_l^M\} \tag{2.33}$$

if $x_j = 2J_j/N_l^j$, and by interpolation elsewhere. Then, using (2.27), (2.32), and (2.33)

$$\begin{aligned}\text{trace}\exp -\beta H_l^M &= \sum_{\underline{n} \in \chi_l^M} \sum_{\substack{(1/2)\, x_j N_l^j \in A_{n_j}: \\ j=1,...,M^\nu}} \prod_{m=1}^{M^\nu} \Bigg[\exp(-\beta N_l^m \eta_m^M) \\ &\times \exp[\log 2 \cosh \beta \eta_m^M (N_l^m - n_m)] \\ &\times \binom{N_l^m}{n_m} c(n_m, x_m N_l^m/2) \exp -\beta V_l f_l^M(\beta, x)\Bigg]\end{aligned} \tag{2.34}$$

If $\tilde{\chi}_l^M$ is the set of all M^ν-tuples $(t_1, t_2, ..., t_{M^\nu})$ such that $t_m = n_m/N_l^m$: $n_m = 0, 1, ..., N_l^m$, we can rewrite (2.34) as

$$\begin{aligned}\text{trace}\exp -\beta H_l^M &= \sum_{\underline{t} \in \tilde{\chi}_l^M} \sum_{\substack{(1/2)\, x_j N_l^j \in A_{t_j N_l^j}: \\ j=1,...,M^\nu}} \prod_{m=1}^{M^\nu} \Bigg[\exp(-\beta N_l^m \eta_m^M) \\ &\times \exp[\log 2 \cosh \beta \eta_m^M N_l^m (1 - t_m)] \\ &\times \binom{N_l^m}{t_m N_l^m} c(t_m N_l^m, x_m N_l^m/2) \exp -\beta V_l f_l^M(\beta, x)\Bigg]\end{aligned} \tag{2.35}$$

All necessary information about f_l^M is contained in the following theorem, which can be extracted from ref. 2.

Theorem 3.

$$\prod_{m=1}^{M^\nu}\left(1+\frac{x_m}{N_l^m}\right)\int_{(S^2)^{M^\nu}} d\Omega_{M^\mu}(\theta,\phi)\exp -\beta V_l \underline{f}_l^M(x,\theta,\phi)$$
$$\leqslant \exp -\beta V_l f_l^M(\beta, x)$$
$$\leqslant \prod_{m=1}^{M^\nu}\left(1+\frac{x_m}{N_l^m}\right)\int_{(S^2)^{M^\nu}} d\Omega_{M^\mu}(\theta,\phi)\exp -\beta V_l \bar{f}_l^M(x,\theta,\phi) \tag{2.36}$$

where

$$d\Omega_M=(4\pi)^{-M^\nu}\prod_{m=1}^{M^\nu}\sin\theta_m\, d\theta_m\, d\theta_m\, d\phi_m$$

and $\underline{f}_l^M(\beta,\cdot)$, $\bar{f}_l^M(\beta,\cdot)$ are functions differing from the function $f_{0,l}^M(\beta,\cdot)$: $[0,1]^{M^\nu}\to\mathbb{R}$ defined by

$$\begin{aligned}-f_{0,l}^M(\beta,x)=&\frac{1}{2}\sum_{m=1}^{M^\nu}\mu_l(B_m^M)\,x_m\cos\theta_m\\&+\frac{1}{4}\sum_{m,m'=1}^{M^\nu}\mu_l(B_m^M)\,\mu_l(B_{m'}^M)\,x_m x_{m'}\,U_{m,m'}^M\\&\times\sin\theta_m\sin\theta_{m'}\cos(\phi_m-\phi_{m'})\end{aligned} \tag{2.37}$$

by functions which converge to zero uniformly in x, θ, and ϕ as $l\to\infty$.

2.4. Application of the Theory of Large Deviations

In preparation for the thermodynamic limit, we introduce two families of measures, and rewrite (2.35) in terms of them. In the Appendix we summarize the theory of large deviations and prove a large-deviation principle[8,9] for these measures which will enable us to obtain a variational formula for $f_l^M(\beta)=\lim_{l\to\infty}f_l^M(\beta)$.

Proposition 4. Define a probability measure $\mathbb{L}_n$ interval $[0,1]$ by

$$\mathbb{L}_n(B)=\frac{1}{2^n}\sum_{K\in\mathbb{N}:\,K/n\in B}\binom{n}{K} \tag{2.38}$$

for $n\geqslant 1$, where B is a Borel subset of $[0,1]$. Then the sequence of measures $\{\mathbb{L}_n: n=1, 2,...\}$ satisfies a large-deviation principle with constants $\{n\}$ and rate function J, where

$$J(x) = x \log x + (1-x)\log(1-x) + \log 2 \tag{2.39}$$

is as defined in (2.9).

Proof. The proof of this simple result is in ref. 9, Example 1.1. ∎

Remark. The measures $\{\mathbb{L}_n: n=1, 2,...\}$ are the distributions of the means $\{n^{-1}S_n: n=1, 2,...\}$ of Bernoulli random variables (state space $\{0, 1\}$, probability 1/2 each). Thus, $\mathbb{L}_{N_l^m}$ is the distribution of the proportion of half-filled pairs in the box B_m^M at volume V_l.

Proposition 5. Define a probability measure $\mathbb{P}_n$ on the interval $[0, 1]$ by

$$\mathbb{P}_n(B) = \frac{1}{2^n} \sum_{K \in \mathbb{N}:\, 2K/n \in B} (2K+1)\, c(n, K) \tag{2.40}$$

for $n \geqslant 1$, where the $c(n, K)$ are as in (2.29), B is a Borel subset of $[0, 1]$, and, by $\mathbb{P}_0 = \delta_1$, the Dirac measure with support at 1. Then the sequence of measures $\{\mathbb{P}_n: n=0, 1, 2,...\}$ satisfies a large-deviation principle with constants $\{n\}$ and rate function

$$I(r) = \tfrac{1}{2}[(1+r)\log(1+r) + (1-r)\log(1-r)] \tag{2.41}$$

as defined in (2.9).

Proof. This result is proved in ref. 11, Theorem 1. ∎

Remark. The simple form of the rate function (2.41) arises, roughly speaking, because the $c(n, K)$ are derived from the n-fold convolution of the distribution of weights of the representation Π^1.

For future use we define the family of probability measures $\{\mathbb{P}_n': n=1, 2,...\}$ on $[0, 1]$, where

$$\mathbb{P}_n'(B) = \frac{1}{c(n)} \sum_{K \in \mathbb{N}:\, 2K/n \in B} c(n, K) \tag{2.42}$$

where $c(n) = \sum_{A_n} c(n, K)$.

In the Appendix we prove the following result.

Theorem 6. Define a measure $\mathbb{M}_n$ on $[0, 1] \times [0, 1]$ by its action on Borel rectangles:

$$\mathbb{M}_n[A \times B] = \int_A d\mathbb{L}_n(t) \int_B d\mathbb{P}_{nt}(r) \tag{2.43}$$

Then the sequence of measures $\{\mathbb{M}_n: n=1, 2,...\}$ satisfies a large-deviation principle with constants $\{n\}$ and rate function $K(t, r) = J(t) + tI(r)$ with I and J as defined in (2.9).

Using the measures $\mathbb{L}_n$ to count the multiplicities of the decomposition (2.23) and the measures $\mathbb{P}_n$ to count the multiplicities of the decomposition (2.29), we can write the free energy as follows:

$$-f_l^M(\beta) = \frac{1}{\beta V_l} \left\{ \log \left[\prod_{m=1}^{M^\nu} 2^{N_l^m} \int_{[0,1]} d\mathbb{L}_{N_l^m}(t_m) \right.\right.$$
$$\left.\left. \times\, c(N_l^m t_m) \int_{[0,1]} d\mathbb{P}'_{N_l^m t_m}(r_m) \right] \exp -\beta V_l f_l^M(tr) \right\} \tag{2.44}$$

where $(tr)_m = t_m r_m$. Then can use the inequalities (2.36) to write upper and lower bounds for $f_l^M(\beta)$ in the following way:

$$\frac{1}{\beta V_l} \log \int_{([0,1]^2 \times S^2)^{M^\nu}} \mathbb{K}_l^M(t, r, \theta, \phi) \exp \beta V_l \underline{g}_l^M(\beta; t, r, \theta, \phi)$$
$$\leqslant -f_l^M(\beta)$$
$$\leqslant \frac{1}{\beta V_l} \log \int_{([0,1]^2 \times S^2)^{M^\nu}} \mathbb{K}_l^M(t, r, \theta, \phi) \exp \beta V_l \bar{g}_l^M(\beta; t, r, \theta, \phi) \tag{2.45}$$

where $d\mathbb{K}_l^M(t, r, \theta, \phi) = d\mathbb{M}_l^M(t, r)\, d\Omega_M(\theta, \phi)$ and $\mathbb{M}_l^M$ is the product measure

$$\mathbb{M}_{N_l^1} \times \cdots \times \mathbb{M}_{N_l^{M^\nu}} \tag{2.46}$$

and

$$\underline{g}_l^M(\beta; t, r, \theta, \phi) = g_l^M(\beta, t) - \underline{f}_l^M(tr, \theta, \phi) \tag{2.47a}$$

$$\bar{g}_l^M(\beta; t, r, \theta, \phi) = g_l^M(\beta, t) - \bar{f}_l^M(tr, \theta, \phi) \tag{2.47b}$$

where $g_l^M(\beta, \cdot)$: $[0, 1]^{M^\nu} \to \mathbb{R}$ is the function

$$g_l^M(\beta, t) = \frac{1}{\beta} \sum_{m=1}^{M^\nu} \mu_l [\log 4 \exp(\beta \eta_m^M) + (1 - t_m) \log \cosh \beta \eta_m^M] \tag{2.48}$$

Proof of Theorem 2. By Proposition 4 the sequence of measures $\{\mathbb{M}_{N_l^m}: l=1, 2,...\}$ satisfies a large-deviation principle with constants $\{N_l^m\}$ and rate function I. Thus, this sequence also satisfies a large-deviation principle with constants $\{\beta V_l\}$ and rate function $\beta^{-1}\mu(B_m^M)I$. With respect to the constants $\{\beta V_l\}$ the measure $d\Omega_{M^\nu}$ considered as a constant sequence satisfies a large-deviation principle with zero rate function. By repeated application of the large-deviation principle for product measures (ref. 11, Appendix 1), we see that the sequence $\{\mathbb{K}_l^M: l=1, 2,...\}$ satisfies a large-deviation principle with constants $\{\beta V_l\}$ and rate function $\mathscr{I}_\beta^M$: $([0, 1]^2\times S^2)^{M^\nu}\to[0, \infty]$, where

$$\mathscr{I}_\beta^M(t, r, \theta, \phi)=\frac{1}{\beta}\sum_{m=1}^{M^\nu}\mu(B_m^M)\,K(t_m, r_m) \tag{2.49}$$

By Theorem 3, both the functions g_l^M and $\bar{g}_l^M$ converge uniformly on compact subsets of $([0, 1]^2\times S^2)^{M^\nu}$ to $g^M-f_0^M$, where

$$g^M(\beta, t)=\frac{1}{\beta}\sum_{m=1}^{M^\nu}\mu(B_m^M)[\log 4\exp(\beta\eta_m^M)+(1-t_m)\log\cosh\beta\eta_m^M] \tag{2.50}$$

and so, applying Varadhan's theorem (see the Appendix) to both sides of the inequality (2.43), we obtain

$$-f^M(\beta)=\sup_{([0,1]^2\times S^2)^{M^\nu}}\{g^M(\beta, t)-f_0^M(\beta; tr, \theta, \phi)-\mathscr{I}_\beta^M(t, r, \theta, \phi)\}\quad\blacksquare \tag{2.51}$$

We obtain a variational formula for $f(\beta)$ by combining Theorems 1 and 2.

Theorem 7. Let $\mathscr{M}=\{(t, r, \theta, \phi): t, r, \theta, \phi\in L^\infty(\Omega, \mu), 0\leqslant t(k)\leqslant 1, 0\leqslant r(k)\leqslant 1, 0\leqslant\theta(k)\leqslant\pi, 0\leqslant\phi(k)\leqslant 2\pi\}$ and define $\mathscr{S}(\beta;\cdot):\mathscr{M}\to\mathbb{R}$ by

$$\mathscr{S}(\beta; t, r, \theta, \phi)=\mathscr{S}_0(\beta)+\mathscr{S}_1(\beta, t)+\mathscr{S}_2(\beta, t, r, \theta, \phi) \tag{2.52}$$

where

$$\mathscr{S}_0(\beta)=\frac{1}{\beta}\int_\Omega\mu(dk)\log\{4\exp[-\beta\eta(k)]\cosh\beta\eta(k)\}$$

$$\mathscr{S}_1(\beta, t)=-\frac{1}{\beta}\int_\Omega\mu(dk)\,t(k)\log\cosh\beta\eta(k)-\frac{1}{\beta}\int_\Omega\mu(dk)\,J(t(k))$$

$$\begin{aligned}\mathscr{S}_2(\beta, t, r, \theta, \phi)=&\frac{1}{2}\int_\Omega\mu(dk)\,t(k)\,r(k)\,\varepsilon(k)\cos\theta(k)\\&-\frac{1}{\beta}\int_\Omega\mu(dk)\,t(k)\,I(r(k))\\&+\frac{1}{4}\int_{\Omega\times\Omega}\mu(dk)\,\mu(dk')\,U(k, k')\,t(k)\,t(k')\,r(k)\,r(k')\\&\times\sin\theta(k)\sin\theta(k')\cos[\phi(k)-\phi(k')]\end{aligned} \tag{2.53}$$

Then $f(\beta) = \lim_{M \to \infty} f^M(\beta)$ exists and is given by

$$f(\beta) = -\sup\{\mathscr{S}(\beta; t, r, \theta, \phi): (t, r, \theta, \phi) \in \mathscr{M}\} \tag{2.54}$$

Proof. We simply remark that the functional $\mathscr{S}$ is a continuum analogue of the functional $\mathscr{S}^M$ in (2.8). The proof of the theorem is very similar to that of Theorem 3 in ref. 2, so we omit it. ∎

3. THE VARIATIONAL PROBLEM

The variational problem which arises in the last theorem is very similar to that which is considered Section 3 of ref. 2. We therefore refer the reader to that paper for the proofs and here mention only the necessary modifications. We assume that $U(k, k') > 0$ for all $k, k' \in \Omega$.

Define the function $G: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$G(x, y) = \begin{cases} \left(\dfrac{\sinh(\beta/2)x}{\cosh \beta y + \cosh(\beta/2)x} \dfrac{1}{x}\right)^{1/2} & \text{if} \quad x \neq 0 \\ \left(\dfrac{\beta}{2} \dfrac{1}{1 + \cosh \beta y}\right)^{1/2} & \text{if} \quad x = 0 \end{cases} \tag{3.1}$$

Let

$$g_\beta(k) = G(|\varepsilon(k)|, \eta(k))$$
$$\tilde{g}_\beta(k) = \sup\{G(x, \eta(k)): x \geqslant |\varepsilon(k)|\}$$

Define the linear compact operators U_β on $L^2(\Omega, \mu)$ by

$$(U_\beta \psi)(k) = \int_\Omega \mu(dk')\, g_\beta(k)\, g_\beta(k')\, U(k, k')\, \psi(k') \tag{3.2}$$

and $\tilde{U}_\beta$ by

$$(\tilde{U}_\beta \psi)(k) = \int_\Omega \mu(dk')\, \tilde{g}_\beta(k)\, \tilde{g}_\beta(k')\, U(k, k')\, \psi(k') \tag{3.3}$$

Let $\lambda(\beta) = \|U_\beta\|$ and $\tilde{\lambda}(\beta) = \|\tilde{U}_\beta\|$. Let $\beta_t \in (0, \infty]$ be such that $\lambda(\beta) < 1$ if $\beta < \beta_t$ and $\lambda(\beta) > 1$ if $\beta > \beta_t$. Similarly, let $\tilde{\beta}_t \in (0, \infty]$ be such that $\tilde{\lambda}(\beta) < 1$ if $\beta < \tilde{\beta}_t$ and $\tilde{\lambda}(\beta) > 1$ if $\beta > \tilde{\beta}_t$. Since $\tilde{g}_\beta \geqslant g_\beta$, we have that $\tilde{\lambda}_\beta \geqslant \lambda_\beta$ for all β, so that $\tilde{\beta}_t \leqslant \beta_t$. For $a > 0$, $b \geqslant 1$, and $y > 0$ let

$$h(a, b; y) = \frac{\sinh[\frac{1}{2}\beta(a^2 + y^2)^{1/2}]}{b + \cosh[\frac{1}{2}\beta(a^2 + y^2)^{1/2}]} \frac{y}{(a^2 + y^2)^{1/2}} \tag{3.4}$$

Then we have the following theorem:

Theorem 8.

$$f(\beta)=-\frac{1}{\beta}\int_{\Omega}\mu(dk)\log\Bigg(2\exp[-\beta\eta(k)]$$
$$\times\left\{\cosh\beta\eta(k)+\cosh\frac{\beta}{2}\,[\varepsilon^2(k)+\Delta_\beta^2(k)]^{1/2}\right\}\Bigg)$$
$$+\frac{1}{4}\int_{\Omega}\mu(dk)\,\frac{\sinh\frac{1}{2}\beta[\varepsilon^2(k)+\Delta_\beta^2(k)]^{1/2}}{\cosh\beta\eta(k)+\cosh\frac{1}{2}\beta[\varepsilon^2(k)+\Delta_\beta^2(k)]^{1/2}}$$
$$\times\frac{\Delta_\beta^2(k)}{[\varepsilon^2(k)+\Delta_\beta^2(k)]^{1/2}} \tag{3.5}$$

where Δ_β depends on β as follows: if (a) $\beta\leqslant\tilde{\beta}_t$, then $\Delta_\beta=0$; if (b) $\beta>\beta_t$ then Δ_β is an element of $\mathscr{D}_\beta$, the set of positive solutions of the gap equation at inverse temperature β:

$$\Delta(k)=\int_{\Omega}\mu(dk)\,U(k,k')\,h(\varepsilon(k'),\cosh\beta\eta(k'),\Delta(k')) \tag{3.6}$$

while if (c) $\tilde{\beta}_t<\beta\leqslant\beta_t$, then $\Delta_\beta\in\{0\cup\mathscr{D}_\beta\}$.

Proof. We first change to the variables R, S, and t, where $R(k)=t(k)\,r(k)$ and $S(k)=t(k)\,r(k)\sin\theta(k)$. For $a>0$, $b\geqslant 1$, and $0\leqslant x\leqslant y\leqslant z\leqslant 1$ let

$$g(x,y,z;a,b)=-\frac{1}{\beta}\{z\log b+J(z)\}+\frac{1}{2}(y^2-x^2)^{1/2}-\frac{1}{\beta}zI\left(\frac{y}{z}\right) \tag{3.7}$$

Let $\mathscr{N}=\{(R,S,t)\colon R,\ S,\ t\in L^\infty(\Omega,\mu),\ 0\leqslant S(k)\leqslant R(k)\leqslant t(k)\leqslant 1\}$ and define $\mathscr{J}\colon\mathscr{N}\to\mathbb{R}$ by

$$\mathscr{J}(R,S,t)=\mathscr{S}_0+\int_{\Omega}\mu(dk)\;g(S(k),R(k),t(k);\varepsilon(k),\cosh\beta\eta(k))$$
$$+\frac{1}{4}\int_{\Omega\times\Omega}\mu(dk)\,\mu(dk')\,U(k,k')\,S(k)\,S(k') \tag{3.8}$$

Then $f(\beta)=-\sup\{\mathscr{J}(R,S,t)\colon(R,S,t)\in\mathscr{N}\}$. Now $(y,z)\mapsto g(x,y,z;a,b)$ is concave and its supremum is attained at $(R_x^{(a,b)},t_x^{(a,b)})$, where $R_x^{(a,b)}$ and $t_x^{(a,b)}$ are the unique values of y and z in $(0,1)$ which satisfy

$$\operatorname{arctanh}\frac{y}{z}=\frac{a\beta}{2}\frac{y}{(y^2-x^2)^{1/2}} \tag{3.9a}$$

and

$$\frac{z^2 - y^2}{(1-z)^2} = \frac{1}{b^2} \tag{3.9b}$$

Let $\mathscr{L} = \{S \in L^\infty(\Omega, \mu)\colon 0 \leqslant S(k) \leqslant 1\}$ and define $\mathscr{V}_\beta\colon \mathscr{L} \mapsto \mathbb{R}$ by $\mathscr{V}_\beta(S) = \mathscr{J}(R_s, S, t_s)$, where we have written R_s and t_s for $R_{s(\cdot)}^{(\varepsilon(\cdot), \cosh \beta\eta(\cdot))}$ and $t_{s(\cdot)}^{(\varepsilon(\cdot), \cosh \beta\eta(\cdot))}$, respectively. Then $f(\beta) = -\sup\{\mathscr{V}_\beta(S)\colon S \in \mathscr{L}\}$. The proof that the supremum is attained is very similar to that of Theorem 5 in ref. 2. It is sufficient to note that $(t, r) \mapsto -tI(R/t)$ is concave. We note here that the absolute continuity of μ with respect to Lebesgue measure is used in the proof of Theorem 5 of ref. 2.

For $\tilde{\lambda}(\beta) \leqslant 1$ we prove an analogue of Theorem 4 of ref. 2; the proof is similar. Letting $\psi_s(k) = (2/\beta) \operatorname{arctanh}[R_s(k)/t_s(k)]$, then from (3.9a)

$$\psi_s(k) \geqslant |\varepsilon(k)|, \qquad t_s(k) = \frac{\cosh(\beta/2)\, \psi_s(k)}{\cosh \beta\eta(k) + \cosh(\beta/2)\, \psi_s(k)} \tag{3.10}$$

Thus, the inequality

$$[r_s^2(k) - s^2(k)]^{1/2} \leqslant \tanh \frac{\beta}{2}\, |\varepsilon(k)| \tag{3.11a}$$

of Theorem 4 of ref. 2 is replaced by

$$\begin{aligned} [R_s^2(k) - S^2(k)]^{1/2} &= |\varepsilon(k)| \frac{\sinh(\beta/2)\, \psi_s(k)}{\cosh \beta + \cosh(\beta/2)\, \psi_s(k)} \frac{1}{\psi_s(k)} \\ &\leqslant |\varepsilon(k)|\, \tilde{g}_\beta^2(k) \end{aligned} \tag{3.11b}$$

We now treat the case $\lambda(\beta) > 1$. The Euler–Lagrange equation for the problem is

$$0 = \frac{|\varepsilon(k)|\, S(k)}{[R_s^2(k) - S^2(k)]^{1/2}} - \int_\Omega \mu(dk')\, U(k, k')\, S(k') \tag{3.12}$$

With

$$\Delta(k) = S(k)\, |\varepsilon(k)|\, [R_s^2(k) - S^2(k)]^{1/2} \tag{3.13}$$

this can be written in the form (3.6). Lemmas 2 and 3 remain essentially unchanged, so we conclude in a manner similar to Theorem 6 of ref. 2 that there exists a strictly positive solution of (3.6) which maximizes $\mathscr{V}_\beta$. We are unable to prove that solutions of (3.6) are unique.

For $\lambda(\beta) \leqslant 1 < \tilde{\lambda}(\beta)$ we deduce from Lemma 2 of ref. 2 that the maximizer of $\mathscr{V}_\beta$ is either zero or in the interior of $\mathscr{L}$. Since the supremum is achieved, then in the latter case, $\mathscr{D}_\beta$ must be nonempty and contain the maximizer. We have not been able to determine whether or not $\mathscr{D}_\beta$ is empty.

Finally, (3.5) follows by insertion of (3.6) and (3.13) into $\mathscr{V}_\beta$.

We are unable to identify the critical temperature exactly, but the following proposition shows that one exists.

Proposition 9. Given β, let S_β, the maximizer of $\mathscr{V}_\beta$, be strictly positive. Then for all $\beta' > \beta$, the maximizer $S_{\beta'}$ of $\mathscr{V}_{\beta'}$ is also strictly positive.

Proof.

$$\begin{aligned}
&\mathscr{V}_{\beta'}(S_{\beta'}) - \mathscr{V}_{\beta'}(0) \\
&\quad \geqslant \mathscr{V}_{\beta'}(S_\beta) - \mathscr{V}_{\beta'}(0) \\
&\quad = \mathscr{V}_\beta(S_\beta) - \mathscr{V}_\beta(0) \\
&\qquad + \left(\frac{1}{\beta} - \frac{1}{\beta'}\right) \int_\Omega \mu(dk) \left[K\left(t_{S_\beta}(k), \frac{R_{S_\beta}(k)}{t_{S_\beta}(k)}\right) - K\left(t_0(k), \frac{R_0(k)}{t_0(k)}\right) \right]
\end{aligned} \tag{3.14}$$

where the inequality is due to the fact that $S_{\beta'}$ maximizes $\mathscr{V}_{\beta'}$. Now $\mathscr{I}$ depends on t only through K, and so $K(t, R/t)$ is stationary at t_0. Since $(t, R) \to K(t, R/t)$ is convex and S_β maximizes $\mathscr{V}_\beta$, we have that

$$\mathscr{V}_{\beta'}(S_{\beta'}) - \mathscr{V}_{\beta'}(0) \geqslant \left(\frac{1}{\beta} - \frac{1}{\beta'}\right) \int_\Omega \mu(dk) [R_{S_\beta}(k) - R_0(k)] \operatorname{arctanh} \frac{R_0(k)}{t_0(k)} > 0 \tag{3.15}$$

Thus the maximizer of $\mathscr{V}_{\beta'}$ is strictly positive. ▮

4. APPLICATION TO THE BCS MODEL

The techniques used to treat the Overhauser model in Sections 2 and 3 can also be applied to calculate the free energy density in the full-trace BCS model. In the quasi-spin formulation of the BCS model, traces are taken over only those states in which the Cooper pairs are either fully occupied or empty. We show that the results obtained in ref. 2 are (up to a rescaling of temperature) unchanged when the trace is extended over those states in which the Cooper pairs are singly occupied.

The Hamiltonian in (2.2) is replaced by

$$\begin{aligned}
\tilde{H}_l = &\frac{1}{2} \sum_{j=1}^{N_l} \bar{\varepsilon}(k_l(i))(a^*(j, +)\, a(j, +) + a^*(j, -)\, a(j, -)) \\
&- \frac{1}{V_l} \sum_{i,j=1}^{N_l} U(k_l(i), k_l(j))\, a^*(k_l(i), +) \\
&\times a^*(-k_l(i), -)\, a(-k_l(j), -)\, a(k_l(j), +)
\end{aligned} \tag{4.1}$$

where $\bar{\varepsilon}$ is a continuous function on Ω. We have labeled the fermionic operators according to the spin and momentum of the particles they create or annihilate. Implicitly, the distribution of momenta is assumed to be invariant under inversion through the origin of momentum space.

The reduction to the quasi-spin model is carried out with the following choice of basis vectors [cf. (2.10)]

$$\begin{aligned}
\psi_l(j, 1) &= \psi_l(k_l(j), +) \\
\psi_l(j, 2) &= 1 \in \mathbb{C} \\
\psi_l(j, 3) &= P(\psi_l(k_l(j), +) \otimes \psi_l(-k_l(j), -)) \\
\psi_l(j, 4) &= \psi_l(-k_l(j), -)
\end{aligned} \tag{4.2}$$

The remainder of the analysis of Section 2 is identical apart from the assignments of the functions ε and η. It should be noted, however, that the origin of the σ^z terms in the quasi-spin formulation is physically different in the two models. In the Overhauser model it is due to the energy difference between the $(k+q/2, +1)$ and the $(k-q/2, -1)$ states, while for the BCS model it is due the energy difference between the full and empty Cooper pair states.

One finds that

$$f(\beta) = -\sup\{\mathscr{S}(\beta; t, r, \theta, \phi): (t, r, \theta, \phi) \in \mathscr{M}\} \tag{4.3}$$

where

$$\mathscr{S}(\beta; t, r, \theta, \phi) = \mathscr{S}_0(\beta) + \mathscr{S}_1(\beta, t) + \mathscr{S}_2(\beta, t, r, \theta, \phi) \tag{4.4}$$

and

$$\mathscr{S}_0(\beta) = \frac{1}{\beta}\int_\Omega \mu(dk)\log\left\{4\exp\left[-\frac{1}{2}\beta\bar{\varepsilon}(k)\right]\right\} \tag{4.5a}$$

$$\mathscr{S}_1(\beta, t) = -\frac{1}{\beta}\int_\Omega \mu(dk)\, J(t(k)) \tag{4.5b}$$

$$\begin{aligned}
\mathscr{S}_2(\beta, t, r, \theta, \phi) = {} & \frac{1}{2}\int_\Omega \mu(dk)\, t(k)\, r(k)\, \bar{\varepsilon}(k)\cos\theta(k) \\
& - \frac{1}{\beta}\int_\Omega \mu(dk)\, t(k)\, I(r(k)) \\
& + \frac{1}{4}\int_{\Omega\times\Omega} \mu(dk)\,\mu(dk')\, U(k, k')\, t(k)\, t(k')\, r(k)\, r(k') \\
& \times \sin\theta(k)\sin\theta(k')\cos[\phi(k) - \phi(k')]
\end{aligned} \tag{4.5c}$$

Theorem 10.

$$f(\beta)=\frac{1}{\beta}\int_{\Omega}\mu(dk)\left[\frac{\beta}{2}\bar{\varepsilon}(k)-\log 2\right]+f_{\text{q.s.}}\left(\frac{\beta}{2}\right) \tag{4.6}$$

where $f_{\text{q.s.}}$ is the free energy density calculated[2] for the quasi-spin formulation of the BCS model with energy function $\bar{\varepsilon}$ and coupling U.

Proof. We adopt the same variables $R(k)$, $S(k)$, and $t(k)$ as in Section 3. From (3.9b) with $b=1$ we find that the supremum over t is achieved when $t(k)=t_R(k)$, where

$$t_R(k)=\tfrac{1}{2}[1+R^2(k)] \tag{4.7}$$

By direct calculation

$$J(t_R(k))+t_R(k)\, I(R(k)/t_R(k))=2I(R(k)) \tag{4.8}$$

and so

$$f(\beta)=\frac{1}{\beta}\int_{\Omega}\mu(dk)\left[\frac{\beta}{2}\bar{\varepsilon}(k)\log 2\right]-\sup\{\mathscr{V}'_{\beta}(\beta;R,S):(R,S)\in\mathscr{N}'\} \tag{4.9}$$

where

$$\mathscr{N}'=\{R,S\in L^{\infty}(\Omega,\mu): 0\leqslant S(k)\leqslant R(k)\leqslant 1\}$$

and

$$\begin{aligned}\mathscr{V}'_{\beta}(\beta;R,S)=&\frac{1}{2}\int_{\Omega}\mu(dk)\,|\bar{\varepsilon}(k)|\,[R^2(k)-S^2(k)]^{1/2}\\&+\frac{1}{4}\int_{\Omega\times\Omega}\mu(dk)\,\mu(dk')\,U(k,k')\,S(k)\,S(k')\\&-\frac{2}{\beta}\int_{\Omega}\mu(dk)\,I(R(k))\end{aligned} \tag{4.10}$$

The result then follows from Theorem 7 of ref. 2. ∎

Remark. The surplus constant in (4.6) arises because a trivial constant operator was omitted from the quasi-spin Hamiltonian in ref. 2.

APPENDIX: PROOF OF THEOREM 6

In this appendix we prove Theorem 6, the large-deviation principle for the sequence of measures $\{\mathbb{M}_n: n=1,2,\ldots\}$. For the sake of completeness we first recall the results form the theory of large deviations that we shall need.[8,9]

Definition. Let $\{\mathbb{K}_n: n=1, 2,...\}$ be a sequence of probability measures on the Borel subsets of a complete separable metric space E and $\{V_n\}$ a divergent sequence of positive numbers. We say that $\{\mathbb{K}_n\}$ satisfies a large-deviation principle with constants $\{V_n\}$ and rate function $I: E \to [0, \infty]$ if the following conditions hold:

(LD1) I is lower semicontinuous.

(LD2) For each $m < \infty$, $\{x: I(x) \leqslant m\}$ is compact.

(LD3) For each closed subset C of E

$$\limsup_{n \to \infty} \frac{1}{V_n} \log \mathbb{K}_n(C) \leqslant -\inf_{x \in C} I(x)$$

(LD4) For each open subset G of E

$$\liminf_{n \to \infty} \frac{1}{V_n} \log \mathbb{K}_n(G) \geqslant -\inf_{x \in G} I(x)$$

Varadhan's Theorem. Suppose that the sequence of probability measures $\{\mathbb{K}_n\}$ on E satisfies a large-deviation principle with constants $\{V_n\}$ and rate function I. Let $\{f_n\}$ be a sequence of continuous functions $f_n: E \to \mathbb{R}$ which are uniformly bounded above, and suppose that f_n converges to $f: E \to \mathbb{R}$ uniformly on bounded sets. Then

$$\lim_{n \to \infty} \frac{1}{V_n} \log \int_E \exp[V_n f_n(x)] \, \mathbb{K}_n(dx) = \sup_E \{f(x) - I(x)\}$$

The first step in the proof is to find the cumulant generating function for the sequence $\{\mathbb{M}_n: n=1, 2,...\}$. The result is heuristically obvious.

Proposition A1. Let

$$\mathcal{Q}_n(v, u) = \frac{1}{n} \log \int_{[0,1]^2} d\mathbb{M}_n(t, r) \exp n(tv + ru), \qquad v, u \in \mathbb{R} \qquad \text{(A.1)}$$

Then

$$\mathcal{Q}(v, u) = \lim_{n \to \infty} \mathcal{Q}_n(v, u)$$

exists, and

$$\mathcal{Q}(v, u) = \sup_{(t,r) \in [0,1]^2} [vt + ur - tI(r) - J(t)] \qquad \text{(A.2)}$$

with I and J as in (2.9).

Proof. First consider the case $u=0$,

$$\mathcal{Q}_n(v, 0)=\frac{1}{n}\log\int_{[0,1]} d\mathbb{L}_n(t)\exp ntv$$

Applying Varadhan's theorem and the large-deviation principle for the sequence $\{\mathbb{L}_n: n=1, 2,...\}$, we obtain

$$\lim_{n\to\infty}\mathcal{Q}_n(v, 0)=\sup_{t\in[0,1]}[vt-J(t)] \tag{A.3}$$

which agrees with (A.2), since $-tI(r)$ has maximum value 0.

Next we take $u>0$. Let $D_n=\{0, 1/n, 2/n,..., (n-1)/n, 1\}\subset[0, 1]$ and for $t\in D_n$ define

$$F_n(t, u)=\frac{1}{n}\log\int_{[0,1]} d\mathbb{P}_{nt}(r)\exp nur \tag{A.4}$$

and extend by interpolation to the whole of $[0, 1]$. By convention on $\mathbb{P}_0$, $F_n(0, u)=u$, while for $t_n\in D_n\setminus\{0\}$ we have from (3.9), (3.11), and (4.12) of ref. 11 that

$$\log\cosh u+\frac{c^-}{n}\leqslant\frac{1}{n}\log\int_{[0,1]} d\mathbb{P}_n(r)\exp nur$$
$$\leqslant\log\cosh u+\frac{c^+ +\log(n+1)}{n} \tag{A.5}$$

where $c^\pm$ are positive constants. Replacing n by nt, u by u/t in (A.5) and multiplying by t, we find

$$F(t, u)+\frac{c^-}{n}\leqslant F_n(t, u)\leqslant F(t, u)+\frac{c^+ +\log(n+1)}{n} \tag{A.6}$$

where

$$F(t, u)=\begin{cases} t\log\cosh(u/t), & u>0,\quad 0<t\leqslant 1\\ u, & u>0,\quad t=0\end{cases} \tag{A.7}$$

Note that $t\mapsto F(t, u)$ is continuous throughout $[0, 1]$. Thus, writing

$$\mathcal{Q}_n(v, u)=\frac{1}{n}\log\int_{[0,1]^2} d\mathbb{L}_n(t)\exp n[tv+F_n(t, u)] \tag{A.8}$$

we obtain

$$\frac{1}{n}\log\int_{[0,1]} d\mathbb{L}_n(t)\exp n[tv+F(t,u)]+\frac{c^-}{n}$$
$$\leqslant \mathcal{Q}_n(v,u)$$
$$\leqslant \frac{1}{n}\log\int_{[0,1]} d\mathbb{L}_n(t)\exp n[tv+F(t,u)]+\frac{c^+ +\log(n+1)}{n} \tag{A.9}$$

Therefore, using Varadhan's theorem, we get

$$\lim_{n\to\infty}\mathcal{Q}_n(v,u)=\mathcal{Q}(v,u)=\sup_{t\in[0,1]}[vt+F(t,u)-J(t)] \tag{A.10}$$

From ref. 11 it easily follows that $F(t,u)=\sup_{r\in[0,1]}[ur-tI(r)]$, so that (A.2) is proved for $u>0$.

Finally, we consider $u<0$. For $u<0$, $F_n(t,u)\leqslant 0$ for all t in D_n and $F_n(0,u)=u$. Now we have to find a lower bound for $F_n(t,u)$ for $u<0$, $t\in D_n\setminus\{0\}$. Let

$$P=-\inf_{m>0}\frac{1}{m}\log\mathbb{P}_m(\{0\})>0$$

Given $\varepsilon>0$ and $t\in D_n\setminus\{0\}$; then

$$\int_{[0,1]} d\mathbb{P}_{nt}\exp nur\geqslant \mathbb{P}_{nt}\left(\left[0,\frac{\varepsilon}{2\,|u|}\right)\right)\exp -\frac{\varepsilon}{2} \tag{A.11}$$

and thus

$$F_n(t,u)\geqslant t\frac{1}{nt}\log\mathbb{P}_{nt}\left(\left[0,\frac{\varepsilon}{2\,|u|}\right)\right)-\frac{\varepsilon}{2} \tag{A.12}$$

But

$$\frac{1}{m}\log\mathbb{P}_m[0,a)\geqslant\frac{1}{m}\log\mathbb{P}_m(\{0\})\geqslant -P$$

Therefore

$$F_n(t,u)\geqslant -\varepsilon \qquad \text{for} \quad t<\varepsilon/2P \tag{A.13}$$

On the other hand, by Proposition 5, since $t\leqslant 1$, there is an N_0 such that

$$F_n(t,u)\geqslant -\inf_{r\in(0,\varepsilon/2|u|)} I(r)-\varepsilon=-\varepsilon \tag{A.14}$$

if $nt > N_0$. Thus

$$F_n(t, u) \geqslant -\varepsilon \qquad \text{for all} \quad t \in D_n \backslash \{0\} \qquad \text{if} \quad n > 2PN_0/\varepsilon \tag{A.15}$$

and therefore

$$\begin{aligned}
\mathcal{Q}_n(v, 0) &\geqslant \mathcal{Q}_n(v, u) \\
&\geqslant \frac{1}{n} \log \left\{ \frac{1}{2^n} \left[\exp n(u+\varepsilon) - 1 \right] + \exp n\mathcal{Q}_n(v, 0) \right\} - \varepsilon \\
&\geqslant \begin{cases} \mathcal{Q}_n(v, 0) & \text{if} \quad u \geqslant -\varepsilon \\ \mathcal{Q}_n(v, 0) + \dfrac{1}{n} \left(\dfrac{\exp n(u+\varepsilon) - 1}{2^n \exp n\mathcal{Q}_n(v, 0)} \right) - \varepsilon & \text{if} \quad u < -\varepsilon \end{cases}
\end{aligned} \tag{A.16}$$

Here we have used the inequality $\log x + y/(y+x) \leqslant \log x + y$ for $0 \leqslant x + y \leqslant x$. Since ε is arbitrary and $\mathcal{Q}_n(v, 0) \geqslant -\log 2$ for $u < 0$,

$$\begin{aligned}
\lim_{n \to \infty} \mathcal{Q}_n(v, u) = \lim_{n \to \infty} \mathcal{Q}_n(v, 0) \\
&= \sup_{t \in [0,1]} [vt - J(t)] \\
&= \sup_{(t,r) \in [0,1]^2} [vt + ur - tI(r) - J(t)]
\end{aligned} \tag{A.17}$$

since $ur - I(r)$ has maximum value 0 when $u \leqslant 0$. ∎

Lemma A2. K is strictly convex on $[0, 1]^2$.

Proof. We calculate the 2×2 Hessian matrix H of K and show that its trace and determinant, and hence its eigenvalues, are strictly positive in $(0, 1)^2$. We have

$$H(t, r) = \begin{pmatrix} J''(t) & I'(r) \\ I'(r) & tI''(r) \end{pmatrix} \tag{A.18}$$

I'', J'', and t are positive, so trace $H(t, r)$ is also positive. The determinant of $H(t, r)$ is

$$(1-t)^{-1} \cosh^2(\operatorname{arctanh} r) - \operatorname{arctanh}^2 r$$

which is positive also. ∎

Proof of Theorem 6. Extend K to the whole of $\mathbb{R}^2$ by

$$\bar{K}(t, r) = \begin{cases} K(t, r) & \text{if} \quad (t, r) \in [0, 1]^2 \\ +\infty & \text{elsewhere} \end{cases}$$

One verifies that (A.2) holds but with the supremum extended over (t, r) in the whole of $\mathbb{R}^2$. Now, $\bar{K}$ is a proper convex function,[17] so that $\bar{K}$ and $\mathcal{Q}$ are Legendre conjugates. By Lemma A2, $\bar{K}$ is essentially strictly convex, so, by ref. 17, Theorem 26.3, $\mathcal{Q}$ is essentially smooth: in particular, it is differentiable in $\mathbb{R}^2$. The minimum of $K(t, r)$ is achieved at a unique point, so all the conditions of ref. 9, Theorem II.6.1 are satisfied, from which we conclude the statement of the theorem. ∎

ACKNOWLEDGMENTS

We thank A. Verbeure and B. Nachtergaele for bringing the problem to our attention, and for making ref. 15 available prior to publication.

REFERENCES

1. N. G. Duffield and J. V. Pulé, *Lett. Math. Phys.* **14**:329–331 (1987).
2. N. G. Duffield and J. V. Pulé, *Commun. Math. Phys.* **118**:475 (1988).
3. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Phys. Rev.* **108**:1175 (1957).
4. P. W. Anderson, *Phys. Rev.* **112**:1900 (1958).
5. G. A. Raggio, Dublin Institute for Advanced Studies Preprint DIAS-STP-88-03 (1988).
6. A. W. Overhauser, *Phys. Rev.* **128**:1437 (1962).
7. G. Fano, *Nuovo Cimento* **38**:597 (1965).
8. S. R. S. Varadhan, *Commun. Pure Appl. Math.* **19**:261 (1966).
9. R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics* (Springer, 1985).
10. D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Academic Press, New York, 1961).
11. W. Cegła, J. T. Lewis, and G. A. Raggio, *Commun. Math. Phys.* **118**:337 (1988).
12. B. Simon, *Commun. Math. Phys.* **71**:247 (1980).
13. F. A. Berezin, *Math. USSR Izv.* **6**:117 (1972).
14. E. H. Lieb, *Commun. Math. Phys.* **31**:327 (1973).
15. B. Nachtergaele and A. Verbeure, Symmetry breaking in the Overhauser model.
16. L. G. Landau and I. F. Wilde, *Commun. Math. Phys.* **70**:43 (1975).
17. R. T. Rockafellar, *Convex Analysis* (Princeton University Press, 1970).