The Generalized Gradient Approximation Applied to the Three-Dimensional Hubbard Model in a Trap: A CAP and Challenge Project

J.K. Freericks and K. Mikelsons
Department of Physics, Georgetown University, Washington, DC
{freericks, karlis}@physics.georgetown.edu

H.R. Krishnamurthy
Centre for Condensed Matter Theory, Department of Physics, Indian Institute of Science and Condensed Matter Theory Unit, Jawarharlal Nehru Centre for Advanced Scientific Research, Bangalore, India
hrkrish@physics.iisc.ernet.in

Abstract

We describe a generalization of the local density approximation for ultra-cold atomic systems, which takes into account the gradients of the Green's functions with respect to position when solving the local Dyson equation and, hence, is sensitive to the local variation of the system within the trap. We use this technique along with dynamical mean-field theory to simulate the Hubbard model data of the Swiss group, which measured the Mott transition in a three-dimensional optical lattice with ultra-cold ^{40}K atoms. We simulate the same size system as used in experiment, which includes about 8 million lattice sites and up to 300,000 atoms. We describe our computational algorithm in detail, along with modifications needed for the communications to make it work on the Cray XE6. We show strong scaling curves up to approximately 43,500 cores. The code performs remarkably well with perfect linear scaling up to the largest systems we ran on. The results that we present are a comparison of the dynamical mean-field theory results within the generalized gradient approximation to the strong-coupling perturbation theory, which is known to be accurate for high-temperature and large interaction strength. This work is part of the Defense Advanced Research Projects Agency (DARPA) optical lattice emulator program.

1. Introduction

The Department of Defense (DoD) is always interested in the newest materials which can outperform those that already exist. Currently, there is a dearth of design principles that can create a material with a specific set of desired properties. Instead, we usually follow an Edisonian approach, where we try many different kinds of materials and examine their properties to see if they are better than ones already discovered. While nearly all materials of interest that have been used in electronic devices, sensors, magnetic devices, etc. have been discovered in this fashion, we are now beginning to enter a new age where this might change. Conventional high performance computers allow for the computation of the materials properties of a wide-range of materials using techniques like density-functional theory, and there certainly have been a number of successful materials designed from first-principles, including electrodes for rechargeable lithium ion batteries, high-performing ferroelectrics for sonar applications, and so on. One class of materials, however, defies this approach because even these state-of-the-art techniques are not sufficient to properly describe their behavior. These materials are called strongly-correlated materials, and their complication comes from the fact that the electrons in them interact so strongly with each other that one needs to know what all other electrons are doing in order to determine how any particular electron will act. Only a very small percentage of all materials fall into this strongly-correlated electron class, but they display some of the most interesting behavior from high-temperature superconductivity, to ultra-strong permanent magnets, to materials with strongly renormalized energy scales like the heavy fermion materials. Because they display complex phenomena like metal-insulator transitions, complex magnetic ordering, and the so-called Kondo effect (marked by a loss of magnetism at low-temperature), it is believed that these materials may outperform more conventional materials in a number of different categories, and hence are a very fruitful area for new materials research.

The challenge, of course, is to be able to predict the properties of these materials from first-principles, which has become an extremely difficult problem. Some success has occurred by combining density-functional theory with dynamical mean-field theory, which is a new technique that can solve the quantum many-body problem exactly if one is in a universe

with a very large number of spatial dimensions (it turns out that three dimensions is close to this large-dimensional limit). Even so, there still remains a wide class of materials that have not yet yielded to a successful description with conventional computation, or the computational time is too long to use them for new materials design and discovery.

Enter Richard Feynman (Feynman, 1982) who suggested an alternative way to solve this problem by using an analog quantum computer. The basic idea is that nature itself is the most efficient computer for solving the quantum many-body problem. Hence, if we can set up a controlled system of quantum particles that obey the properties of the material we are trying to calculate properties for, then we just let the system evolve as a function of time and determine its properties by measurement. While this might sound like just what an experimental physicist would do in a laboratory, the key difference here is that Feynman was proposing creating an analog computer where we could rapidly "dial-in" the different materials properties and investigate a wide-class of materials over a short period of time. While some might argue that this would still be an Edisonian approach, it is a much more efficient Edisonian approach than trying to make a wide-class of new materials and measure all of their properties. We are not yet at the stage where such a Feynman analog computer is available, but we have started to construct simpler prototypes as part of the Defense Advanced Research Projects Agency (DARPA) program (Abo-Shaeer, 2010). One class of analog computers is made from ultra-cold neutral atoms placed in an optical lattice.

These optical lattice emulators are constructed by placing ultra-cold clouds of alkali atoms in the egg-carton-like potential energy surface created by retro-reflected laser beams in each spatial direction, which have a lattice spacing that is typically about 0.5 micron. This lattice spacing is about two hundred times larger than the lattice spacing in a material, which produces a huge advantage—we can directly image the atoms as a function-of-time, and see how they evolve due to their quantum-mechanical behavior. This would be the equivalent of having an ultrafast microscope which can image the electrons deep inside the bulk on a time scale short, compared to the time for an electron to move from one lattice site to the next. Unfortunately, we can only examine the simplest of models of strong electron correlation with these first-generation optical lattice emulators.

In order to verify that these optical lattice emulators are working properly, we need to also benchmark their accuracy. This approach is further complicated by the fact that these atomic systems involve clouds of atoms which are trapped in space at some particular location via a harmonic trap. The trap makes the system inhomogeneous; hence, one needs to use caution in comparing their behavior to those of quantum particles in a perfectly periodic lattice, which is needed to describe a real material. It is precisely the role of high performance computing to determine the properties of these inhomogeneous systems, and thereby verify and validate the analog optical lattice emulators. Since the exact problem is too difficult to solve, even with today's computers, we need to make approximations of one form or another. We try our best to do so in such a way that the approximations are faithful to the exact solutions and will yield interesting information about the behavior of the system.

Our numerical approach uses dynamical mean-field theory, which we already described as a formal technique to solve the many-body problem which becomes exact in the limit of an infinite number of spatial dimensions. While the approach is approximate in three dimensions, it is believed that the approximation is quite good for many systems. One huge advantage of this approach is that it is the only numerically-tractable method that can describe fermionic systems at low-temperature without facing the so-called sign problem, and is also efficient enough that it can model systems that are as large as the actual experimental systems being studied. We must generalize the original formulation of dynamical mean-field theory to inhomogeneous systems (called IDMFT) to take into account the effect of the trap. The approach works with objects called Green's functions which describe the amplitude and phase for creating a particle at one lattice site at time t_1 and removing the particle from another lattice site at time t_2. The Green's function can be parameterized by an object called the self-energy, which describes how the energy of the excitations is modified by the interactions between the particles, and how long-lived they are. In a conventional many-body problem, the self-energy is a function of both momentum and frequency. In an inhomogeneous system, momentum is no longer a good quantum number, so the momentum dependence would become a matrix-valued function of two momenta. Fortunately, in the IDMFT approach, the self-energy is actually independent of momenta, so we need only determine a (different) scalar function of frequency for each lattice site. The algorithm for solving IDMFT involves the following steps: 1) start with a guess for the self-energy on each lattice site; 2) calculate all the local Green's functions (for each frequency) by finding the diagonal of the inverse of a general complex and sparse matrix (whose dimension is the number of lattice sites) which includes the self-energy and trapping potential for each lattice site on the diagonal, and the connections between neighboring lattice sites on the off-diagonal elements; 3) extract the local effective medium by using the local Dyson's equation to remove the local self-energy from the local Green's function; 4) solve a quantum impurity problem in the extracted local effective medium for each lattice site to find the new local Green's function; 5) use the local Dyson's equation with the new local Green's function and the old effective medium to find the new self-energy; and 6) use the new self-energy in step (2) to continue the iterative algorithm. One repeats these

steps until the functions have converged, which typically takes from 10 to 1,000 iterations. The two numerically-intensive parts of the algorithm are finding the diagonal of the inverse of the large sparse matrix (step 2) and solving the local impurity problem on each lattice site (step 4). The former problem grows as N^3 with N the number of lattice sites, while the latter is also slow for complicated quantum models (like the Hubbard model), but grows linearly with the system size N. It turns out that finding the diagonal of the inverse of a few million by few million matrixes is not feasible with current computational resources which employ dense matrix algorithms. One can, instead use the sparsity of the matrix to really speed up the inversion and to reduce the storage requirements. We are currently working on developing such an algorithm, but it is not yet complete (Freericks, et al., 2010; Carrier, et al., 2011).

Instead, there is an alternative approach which is called the local density approximation (LDA). In this approach, we assume the system changes slowly from one lattice site to another, so we can replace the problem in the inhomogeneous system by a series of problems in homogeneous systems, but with a chemical potential for the homogeneous problem set by the global chemical potential of the inhomogeneous system minus the local potential due to the trap $\mu_{homog.}=\mu-V_i$. This technique requires us to essentially solve N independent homogeneous problems and then "stitch" them together into an approximation for the inhomogeneous system. This technique has long been used for these ultra-cold atomic systems in a trap, but it isn't obvious how accurate these techniques actually are. A systematic way to improve upon the LDA is needed. If the corrections to the LDA are small, then one can deduce that the LDA is a good approximation, and if the corrections are large, then it is not. In density-functional theory this same problem has been addressed many years ago, and it is now routine to include gradient corrections on top of the LDA. We can do the same thing here, and we similarly call this approach the generalized-gradient approximation (GGA) to make contact with the same ideas which arose from the density-functional theory community.

The derivation of the GGA equations is rather straightforward, but is somewhat lengthy and will be omitted here. We do describe; however, the basic ideas that are used. The key computational step which is difficult to achieve is the solution of the Dyson equation, which involves finding the diagonal of the inverse of an $N\times N$ matrix. The LDA approximation involves replacing the inhomogeneous problem, where along the diagonal, the local chemical potential (chemical potential minus trap potential) minus the local self-energy appears, by a homogeneous one, where every diagonal element is equal to the current diagonal element. By taking the difference of these two Dyson equations, we can find a formula for the gradient of the Green's function, which involves the difference of the local Green's function at one site and the local Green's functions at its nearest-neighbors. Solving this set of equations, where we discard the differences between the Green's function and the second-neighbors, yields a set of equations for the GGA (which is the correction of the LDA result by gradients of the Green's functions at the given site, and its nearest-neighbors). The modification of the IDMFT algorithm then rests solely with changes to step (2) of the above algorithm. For the LDA, we replace every diagonal by the diagonal element of the ith lattice site, and find the inverse by Fourier transformation. For the GGA, we need to solve a self-consistent equation which requires the LDA solutions at the given lattice site, at the nearest-neighbors, the nearest-neighbor Green's function (in the LDA) and the local GGA Green's function and self-energy. Because this step is embedded in the self-consistent solution of the IDMFT algorithm, we need to solve N impurity problems as well, at each iteration. For the GGA approach, it is the solution of these impurity problems which takes the most computational effort. For the full IDMFT approach, it is the computation of the diagonal of the inverse of the large sparse matrix that will likely be the most computationally-intensive step.

We also need to solve the impurity problem, which can be mapped onto an effective single-impurity Anderson model. No exact solution for this problem exists (when we have an arbitrary time-dependent source-field), but it can be solved with a number of different numerical techniques ranging from explicit perturbation theory, to quantum Monte Carlo (QMC) approaches, to the numerical renormalization group. We will use a QMC approach here, because the experiments are at high-temperature and the numerical renormalization group is best at lower-temperatures.

The QMC algorithm that we use is a weak-coupling version of the continuous-time quantum Monte Carlo algorithm (Rombouts, 1999; Rubtsov, et al., 2005). This approach uses stochastic sampling to evaluate Feynman diagrams in imaginary-time to solve the problem. The algorithm sums over a random collection of diagrams at different orders, and adjusts the order of the diagram based on importance sampling via a Metropolis strategy. The approach is most accurate at high-temperature and small interaction strength. As the temperature or the interaction is reduced, the average order of the calculation increases, and the integration range in imaginary-time also increases. This technique is currently believed to be the most accurate state-of-the-art approach for determining properties that can be found at finite-temperature using Green's functions evaluated along the imaginary-time axis (which includes the particle-density at each lattice site, the double-occupancy, the entropy-per-particle, and the order parameter if the system goes into an ordered anti-ferromagnetic phase). Details for how this algorithm is implemented can be found elsewhere (Jarrell, et al., 2008; Gull, et al., 2011).

We end this section by describing the model that is used, which is the Hubbard model in a trap (Hubbard, 1963). The Hubbard model includes three terms:

$$H = -\sum_{\langle i,j\rangle\sigma} t_{ij}\left(c_{i\sigma}^{+}c_{j\sigma} + c_{j\sigma}^{+}c_{i\sigma}\right) - \sum_{i}\left(\mu - V_i\right)c_{i\sigma}^{+}c_{i\sigma} + U\sum_{i} n_{i\uparrow}n_{i\downarrow} \tag{1}$$

Where $c_{i\sigma}^{+}$ and $c_{i\sigma}$ are fermionic creation and annihilation operators at site *i* which satisfy the canonical anti-commutation relations $\left\{c_{i\sigma}^{+}, c_{j\sigma'}\right\} = \delta_{ij}\delta_{\sigma\sigma'}$. The number operator is $n_{i\sigma} = c_{i\sigma}^{+}c_{i\sigma}$. The first term in Equation 1 is the fermion kinetic energy, which has an overall scale set by the hopping integral (*t*) and which involves a hopping of the fermions (of either spin) from one site to any of its nearest-neighbors. The second term is the local chemical potential, which is a site-energy given by the global chemical potential minus the trap potential at the given lattice site (which is taken to increase quadratically away from the origin, but with a different curvature along each of the different spatial axes, just like in experiment). The third term is the interaction term, which includes a Coulomb repulsion (*U*) multiplied by the double-occupancy, which is the number of up-spin particles multiplied by the number of down-spin particles at the lattice site. All of the parameters are taken directly from the experiment (Jördens, et al., 2008; Jördens, et al., 2010).

2. Scaling of the Numerical Algorithm

Since this project was selected for a Capability Applications Project (CAP) on the XE6 machines, we started by running scaling studies on the Cray XE6 machine at Engineering Research Development Center (ERDC) [garnet] which has approximately 22,000 cores. An initial porting of our code failed to run on the XE6. It turned out this was related to buffer issues for send and receive calls under Message Passing Interface (MPI). In the original code, we did not organize the communications to guarantee that they would be non-blocking. On most high-performance machines, such sloppiness is allowed because they receive and send buffers are large enough to wait for the expected calls and to receive them out of order. Not so on the XE6, so we were forced to engage in the more proper practice of ensuring we had non-blocking communications. Making these modifications was fairly straightforward to do, and once we had completed it, the code ran very well.

Before reporting our strong scaling analysis, we need to discuss a bit about how we chose the lattice sites in our system. We begin with an $M^{\times}M^{\times}M$ cubic lattice (with M in the range from 100 to 250). We calculate the lattice potential at each lattice site, and include only those sites whose potential is smaller than the smallest value of the potential at the edge boundary of the cube. Then, using a fact that our system is symmetric with respect to mirror reflections along each axis, we can further reduce the number of sites for which we need to solve the QMC impurity problem, by solving it only for inequivalent lattice sites (all lattice sites must be used, however, in the Dyson equation). Since we cannot store the data for the Green's functions and self-energies on just one core, we have to distribute the memory. Our code is set up to distribute the memory on a number of cores which is equal to a multiple of two (because we use a binary bifurcation to assign lattice sites to a given core), or to the number of lattice sites divided by 2. In order to perform scaling studies, we need to vary the number of cores by increasing by a factor of 2 or by choosing them to equal *N*/2 (with N the number of inequivalent lattice sites now). Since these machines do not have the number of cores equal to a multiple of 2, we have to choose our lattice size appropriately to be able to map the same number of lattice sites to each core for the maximal number of cores, so that the system can equally distribute the labor for solving the QMC impurity problem (it turns out that the computational time is not the same, because it depends on the particle density at a given site, so there always is a spread in the computational time for different lattice sites and different runs). Using this approach, we can have an optimal way of testing the scaling of the code up to the maximal number of cores that we can run on.

We report results for the initial scaling analysis on garnet in Figure 1. This data is for the computational part of the code, not the input/output part, and the scaling studies only ran for a few iterations, rather than the number that would be used in a production run. One can see that the initial scaling is quite linear, and we fit the computational speed (inverse wall-clock time as measured by MPITIME) versus number of cores to a linear curve with nearly zero intercept for the lowest few runs. This curve is then extrapolated to the larger number of cores where we also run the code. One can see that the linear scaling starts to tail off for 16,384 slave cores, but then recovers for 20,000 cores. This result occurs because the lattice was set up so that we would have two impurity problems to solve on each core for the 20,000 core case, but then the 16,384 case will have some cores solving two and some solving three impurity problems, hence there is a degradation of the scaling for that case.

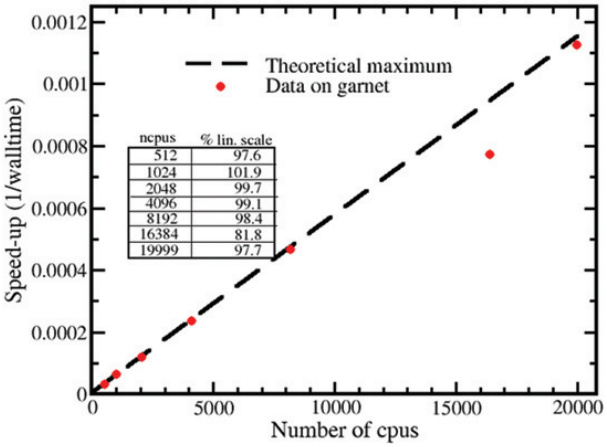

Figure 1. Strong scaling results for the algorithm on garnet (Cray XE6 with approximately 22,000 cores). We report the speed-up (inverse of the wall-time) for the different number of cores used in the scaling run. Note how there is a dip in the curve at 16,384 cores. This occurs because there are some cores with 3 impurity solvers required and some with 2. When we go to the commensurate limit of 2 solvers per core at 20,000 cores, we find the speed-up recovers, as expected. The inset is a table showing the percentage of linear scaling of the speed-up.

After completing the scaling analysis, we were selected for a CAP Phase II run on the Arctic Region Supercomputing Center (ARSC), machine chugach, which has about 11,000 cores. We ran our code on 8,192 slaves (plus one master) in a continuous fashion for approximately six weeks, generating a large portion of our data. We then performed a scaling analysis on the raptor machine at the US Air Force Research Laboratory (AFRL), which has about 44,000 cores. There we were able to run a scaling analysis with up to 43,500 cores, and we found that the system seems to show a super-linear scaling (see Figure 2)! Most likely this is just arising from some issues with run-to-run variations on the machine and to the possibility that our fit to a linear curve at small core counts is not perfectly accurate (although those points are quite linear, with a correlation coefficient of 0.99994). We compare the normalized speed-up curves on garnet and raptor in Figure 3. This needs to be done because the size of the problems we ran was different on the two machines. One can see that the behavior is quite similar on the two machines. Overall, we used about 10 million CPU-hours for the scaling runs and for the production runs on the machines during the CAP. We also use the same codes, running now on 4,097 cores for our Challenge Project.

In conclusion, we have shown excellent scaling behavior of this algorithm. This occurs primarily because the code has limited communications between nodes during a given iteration, and MPI can handle those communications very efficiently. In addition, the workload for the impurity solvers seems to divide well between the different cores as we increase the core count. We are currently working on understanding the differences in the workload and searching for means to balance the loads amongst the cores "on-the-fly" during the computation. It is likely that such an approach will require us to use a global addressing approach for the memory, which is available with either co-array FORTRAN or UPC.

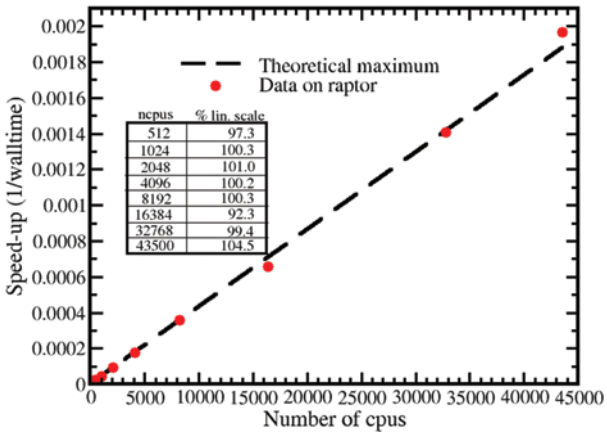

Figure 2. Same as Figure 1, except on the AFRL machine raptor, which has about 44,000 cores. Here, we surprisingly see super-linear scaling up the maximum size we could run on the machine.

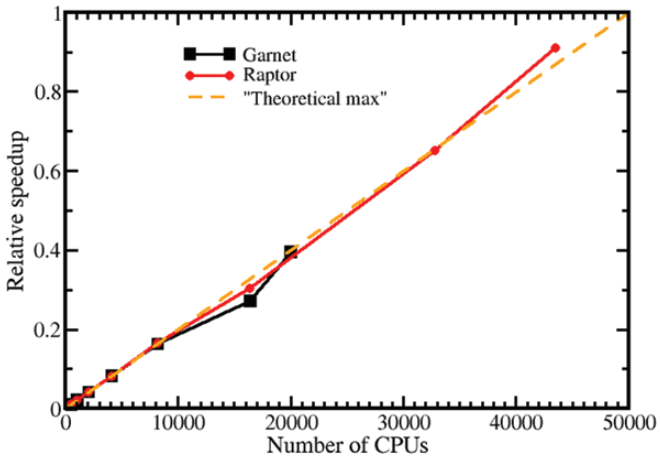

Figure 3. Comparison of the renormalized speed-ups on the two machines (we need to normalize to unit slope since the job sizes run on the different machines were different). One can immediately see the excellent scaling of the code which is showing no signs of stopping up to 43,500 cores.

3. Results

We begin by discussing the experimental data of the Swiss group (Jördens, 2010). They work with the fermionic isotope of K, ^{40}K, placed on an optical lattice. They cool their system down to an entropy per particle of about $1.3k_B$. After their experiment is finished, they drop the optical lattice potential and measure the entropy per particle again, and find it to be equal to $2.5k_B$. Hence, they bracket the entropy per particle of the system within the lattice to be between these two extremes. They use a clever method to detect the double-occupancy of the system. First, they perform an absorption image to count the total number of particles in the optical lattice. Next, they use a radio frequency pulse to transfer a K atom to a new hyperfine state if and only if two atoms sit at the same site, then drop the lattice and perform an absorption image to count the number of atoms in the new hyperfine state, which equals the number of double-occupancies. This gives them the total number of particles and the total number of double-occupancies. These two quantities allow one to see behavior that gives rise to the Mott transition.

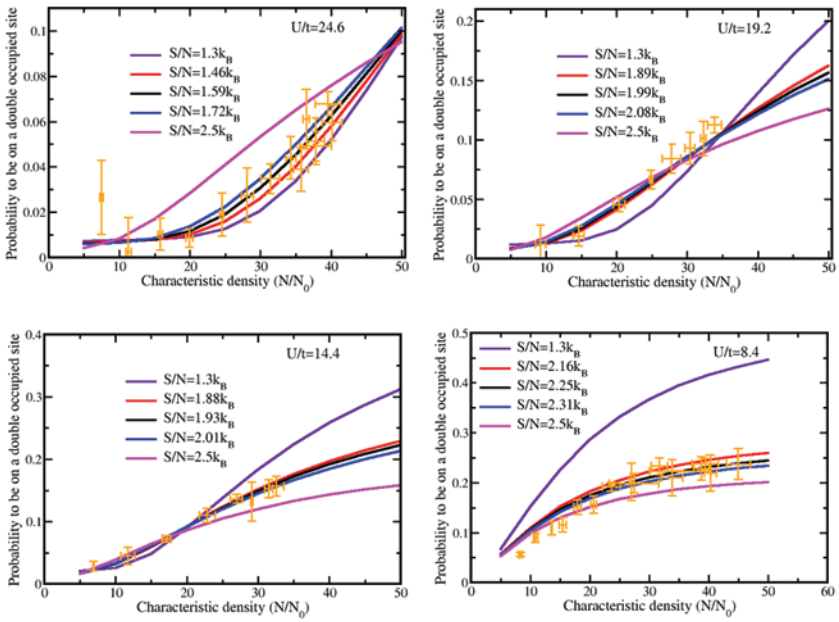

Figure 4. Fitting results for constant entropy curves of the double-occupancy versus particle number and the experimental data reported with error bars for the four different interaction strengths. Note how the fits are in general excellent, except for the lowest value of interaction strength (*U*/*t*=8.4), where the fit is poorer, especially at the lower densities, which also correspond to lower temperatures.

The way to see this is as follows: As we start filling fermions onto our lattice, we initially can fill them in the lowest available energy levels, which are typically spread throughout the lattice. As we keep filling in levels, the fermions begin to interact with each other as they become denser. If the repulsion term, U, is large enough, it will forbid two fermions of opposite spin from occupying the same lattice site at the same time (two fermions of the same spin are already forbidden due to the Pauli exclusion principle). Hence, the larger the interaction is, the lower the double-occupancy will be, until we have too many particles in our system that we have to increase the filling to be larger than one per site in the center of the trap. Then the double-occupancy will sharply increase as sites are filled with two fermions. This transition to the sharp increase in the double-occupancy will become smoother as the interaction is reduced. The ratio of the interaction to the hopping (U/t) is controlled by varying the depth of the lattice potential, which primarily controls the magnitude of the hopping t, and by tuning the interaction strength by using a Feshbach resonance to vary the two-particle scattering length, which changes U. Four values are considered in the experiments; i) U/t=24.6; ii) U/t=19.2; iii) U/t=14.4; and iv) U/t=8.4, where the lattice depth is fixed at 7 potassium recoil energies (t=174 Hz). The optical lattice is a simple cubic lattice in three dimensions with a trap that is harmonic, but with different curvatures along the three different axes, yielding constant energy shells which are ellipsoidal spheroids in shape (the trap frequencies are 49.4, 52.6, and 133 Hz).

Since the optical lattice is turned on slowly, it is believed to be a nearly adiabatic process, which implies that the entropy will be conserved. Of course it cannot be completely adiabatic; otherwise the final entropy per particle would be equal to the initial entropy. But, if we believe the adiabatic processes are governed by the speed that the lattice is turned on, then we expect the entropy of the fermions in the lattice to be the same for all cases with the same interaction strength, and hence this is the criterion we will use to fit the data. The data is plotted as the probability for an atom to be in a doubly-occupied state (two times the number of double-occupancies divided by the total number of particles) versus the total number of particles (plotted relative to the characteristic density N_0=7,393). In the experiment the number of particles varies from about 40,000 to 300,000, and the diameter of the cloud is always less than about 250 lattice sites.

The simplest way to solve this problem is to solve it in the high-temperature limit when U/t is large using what is called the strong-coupling expansion (Scarola, 2009). This produces a set of algebraic equations that need to be solved to determine the density of particles at each site as well as the double occupancy and the entropy per particle. This technique is only approximate, but is very efficient, requiring only a few hours on a single-processor machine for the total analysis. Our fitting procedure is slightly different than the one performed by the Swiss group, so we describe it in detail here. What we do is minimize the error for the double-occupancy versus the number of particles for a fixed-value of the entropy, taking into account the error bars of the double-occupancy in computing the goodness of fit. We compute the best-fit (black), the curves for the extremal values where the entropy per particle is fixed at the largest or smallest allowed values (purple and magenta), and then for the results where we shift the number of particles up or down (red and blue) by the amount of one standard deviation and recalculate the best-fit. Results are shown in Figure 4 for the four different values of the interaction strength. In general, the fits are excellent, but one can clearly see deviations for the lowest number of particles in the smallest interaction strength case U/t=8.4. Hence, we focus our efforts on trying to fix this discrepancy by performing quantum Monte Carlo simulation for this case.

We approximately solve the IDMFT problem using the QMC+GGA algorithm sketched above. Just like in experiment, we will determine the total number of particles, the total number of double-occupancies, and the entropy per particle. The former two results come straight out of the QMC solution for each lattice site, where the particle filling and the probability to be on a doubly-occupied site both are outputs of our impurity solver. The entropy per particle is more complicated to solve. We start at a high-temperature of about twenty-times the hopping, where the strong-coupling calculation will accurately determine the entropy for us. Then, we reduce the temperature, keeping the number of particles fixed. The entropy can then be calculated as an integral over the inverse of the temperature of the energy per particle, which can also be calculated directly by the QMC solver for each lattice site (Werner, 2005). Armed with this data, we can then recalculate the fit to the double-occupancy versus particle number plots at constant entropy. We have not yet completed this exercise, since we are still generating QMC data for some of the different experimental cases. But we can describe to what extent the IDMFT(QMC+GGA) approach differs from the strong-coupling approach, and also we can discuss the differences between the GGA and the LDA. We begin with the latter. We find that for nearly all cases we have examined that the LDA and the GGA agree very well with each other, which supports the notion that the LDA is very accurate when we are above the ordering transition temperature.

We also find that the strong-coupling approach is quite accurate until we start to get into the low-temperature regimes, where the IDMFT results clearly lie above the strong-coupling results for the entropy per particle. This is illustrated as a function of inverse temperature for two values of particle numbers: one with N=61,455 and one with N=202,780 in Figure 5. We see the deviation between the two entropies is more significant at the lower densities. In fact, we are probably reaching to temperatures where the IDMFT results themselves may no longer be so accurate, since they neglect the momentum

dependence of the self-energy which gives rise to quantum fluctuations which reduce the entropy per particle (the IDMFT cannot have the entropy reduced below ln 2 at half-filling until the system orders into an anti-ferromagnetic order, which we have not examined yet with our code; away from half-filling it can be reduced somewhat more). We expect that there is a range of temperature where the true entropy per particle lies in between the IDMFT and strong-coupling results.

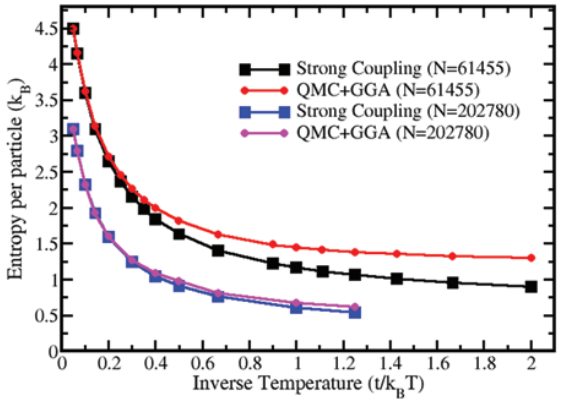

Figure 5. Entropy per particle versus inverse temperature for two different numbers of particles with *U*/*t*=8.4. Note how the entropy per particle deviates more from each other for the more dilute system. The deviations do become large enough to have a significant effect at the lower temperatures (toward the right of the plot).

We finally compare the radial particle density profiles to the radial double-occupancy profiles at two temperatures for the two different numbers of particles in Figure 6 (using the QMC_GGA algorithm). The low particle number case never achieves too high a density because the Fermi pressure makes the cloud size large enough that one never has any sites which have, on average two particles per site; hence the total number of double-occupancies is small. The higher particle number has an appreciable core region with two particles per site; hence the double-occupancy is sharply increased. The change in curvature of the density profile near unit density occurs due to the initial signatures of Mott-like physics entering at the lower temperatures (the interaction strength is too low for the system to be a true Mott insulator). The high-temperature results do not appear to have much, if any quantum degeneracy effects, while the lower temperature results certainly do.

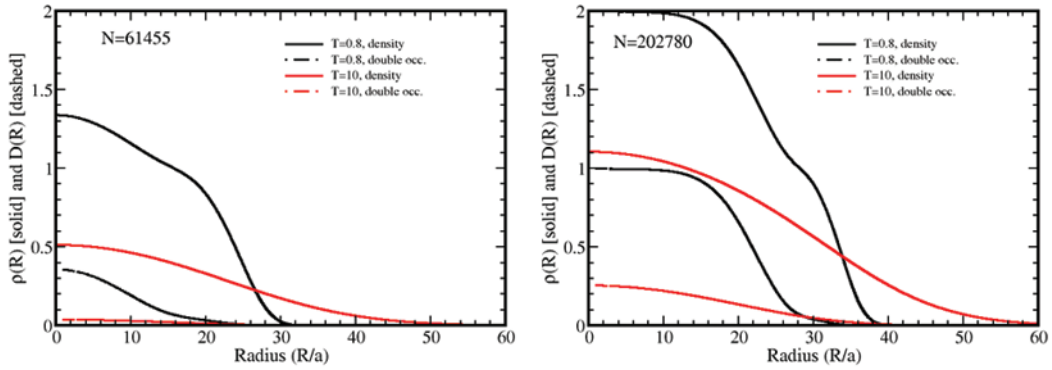

Figure 6. Radial density (solid) and double-occupancy (dashed) profiles for high (k_BT=10t, red) and low (k_BT=0.8t, black) temperatures for the two different cases of particle numbers shown in the previous figure (*U*/*t*=8.4). The change in curvature of the radial density near a filling of one at low-temperature arises from the beginning signals of the formation of a Mott-insulating state. These signatures only begin to be seen at the lower temperatures.

4. Significance to DoD

DARPA's interest in this problem is to ultimately build a materials science emulator out of ultra-cold atomic atoms. One can then hypothesize a particular material, program the emulator to simulate its properties, see if these properties are an improvement over currently-known materials, and then devise a way to make the new materials with these targeted properties. We are still far away from this goal, but have made much progress with being able to make simpler quantum many-body problem emulators in a variety of different platforms and for a variety of different models. These simpler emulators can still be benchmarked with conventional high performance computing, which we do here.

5. Conclusion

In this work, we have shown how one can go beyond the LDA to include gradient corrections. We find at high-temperature these gradient corrections are quite small. As the temperature is lowered, they become larger, and certainly can have some effect on the entropy per particle at moderate temperatures (the difference between the DMFT entropy and the strong-coupling entropy is much more significant). We expect both the LDA and GGA approximations to be less accurate once the system begins to display anti-ferromagnetic order, because the transition temperature for both approximate methods is closely-tied to the transition temperature in the bulk, while a full IDMFT solution would probably have a suppressed transition temperature due to inhomogeneity effects and proximity effects that are not present in the current generation of codes. We are currently working on an efficient implementation of the full IDMFT algorithm.

Acknowledgements

JKF acknowledges support from the Army Research Office Grant Number W911NF0710576, with funds from the DARPA OLE program (for the computations that are used to benchmark the Swiss experiment), and from the National Science Foundation under grant number OCI-0904597 (for development of the MPI algorithm). KM was supported by the AFOSR under the MURI program from grant number FA9559-09-1-0617. HRK acknowledges support from the DST (India). The collaboration between the US and India was supported by grant number JC-18-2009 of the Indo-US Science and Technology Forum. DoD HPC computer time was provided on Cray XE6 machines located at the Arctic Region Supercomputer Center (ARSC), the US Air Force Engineering and Research and Development Center (AFRL) and the US Army Engineering Research and Development Center (ERDC). This project was supported primarily by Challenge Project DARPA-C4J and a CAP in fiscal year 2010–2011. We thank the Esslinger group for providing us with their experimental data and for conversations with Niels Strohmaier and Leticia Tarruell that helped clarify the data.

References

Abo-Shaeer, J., website for the optical lattice emulator program, http://www.darpa.mil/dso/thrusts/physci/funphys/ole/index.htm, 2010.

Carrier, P., J.M. Tang, Y. Saad, and J.K. Freericks, "Lanczos-based low-rank correction method for solving the Dyson equation in inhomogeneous dynamical mean-field theory", *Physics Procedia*, 2011 (submitted); arxiv:1102.5738.

Feynman, R.P., "Simulating physics with computers", *International Journal of Theoretical Physics.*, 21, pp. 467–488, 1982.

Freericks, J.K., H.R. Krishnamurthy, P. Carrier, and Y. Saad, "Efficiently generalizing ultra-cold atomic simulations via inhomogeneous dynamical mean-field theory from two- to three-dimensions", *Proceedings of the HPCMP Users Group Conference 2010*, Shaumburg, IL, edited by D.E. Post, IEEE Computer Society, Los Alamitos, CA, 2010.

Gull, E., A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, "Continuous-time Monte Carlo methods for quantum impurity problems", *Rev. Mod. Phys.*, 83, pp. 349–404, 2011.

Hubbard, J., "Electron correlations in narrow energy bands", *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences*, 276, pp. 238–257, 1963.

Jarrell, M., A. Macridin, K. Mikelsons, D.G.S.P. Doluweera, and J.E. Gubernatis, "The dynamical cluster approximation with quantum Monte Carlo cluster solvers", Lectures on the physics of strongly-correlated systems XII, *American Institute of Physics Conference Proceedings*, 1014, pp. 34–106, 2008.

Jördens, R,. N. Strohmaier, K. Günter, H. Moritz, and T. Esslinger, "A Mott insulator of fermionic atoms in an optical lattice", *Nature*, 455, pp. 204–207, 2008.

Jördens, R., L. Taruell, D. Greif, T. Uehlinger, N. Strohmaier, H. Moritz, T. Esslinger, L. De Leo, C. Kollath, A. Georges, V. Scarola, L. Pollet, E. Burovski, E. Kozik, and M. Troyer, "Quantitative determination of temperature in the approach to magnetic order of ultra-cold fermions in an optical lattice", *Physical Review Letters*, 104, p. 180401, 2010.

Rombouts, S.M.A., K. Heyde, and N. Jachowicz, "Quantum Monte Carlo method for fermions, free of discretization errors", *Physical Review Letters*, 82, pp. 4155–4159, 1999.

Rubtsov, A.N., V.V. Savkin, and A.I. Lichtenstein, "Continuous-time quantum Monte Carlo method for fermions", *Physical Review B*, 72, p. 035122, 2005.

Scarola, V.W., L. Pollet, J. Oitmaa, and M. Troyer, "Discerning incompressible and compressible phases of cold atoms in optical lattices", Physical Review Letters, 102, pp. 135302–1-4, 2009; Erratum, *Physical Review Letters*, 103, p. 189901(E), 2009.

Werner, F., O. Parcollet, A. Georges, and S. Hassan, "Interaction-induced adiabatic cooling and anti-ferromagnetism of cold atoms in optical lattices", *Physical Review Letters*, 95, p. 056401, 2005.