

Fermion Bag Approach to Lattice Hamiltonian Field Theories

Emilie Huffman^{1,*}

¹Duke University

Abstract. Using a model in the Gross-Neveu Ising universality class, we show how the fermion bag idea can be applied to develop algorithms to Hamiltonian lattice field theories. We argue that fermion world lines suggest an alternative method to the traditional techniques for calculating ratios of determinants in a stable manner. We show the power behind these ideas by extracting the physics of the model on large lattices.

1 Introduction

Traditionally lattice field theories have been studied by the lattice community using the Lagrangian path integral approach. However, the Hamiltonian formalism does offer some compelling reasons for its application. The first reason has to do with additional symmetries that can be found in Hamiltonian theories and may be relevant to phase transitions. As an example theory, the spectrum of Dirac fermions can be obtained through the following Hamiltonian,

$$H_0 = -t \sum_{\langle xy \rangle} (c_{x,\sigma}^\dagger c_{y,\sigma} + c_{y,\sigma}^\dagger c_{x,\sigma}), \quad (1)$$

when placed on a honeycomb lattice. Here $\langle x, y \rangle$ denotes nearest neighbor sites, and σ is a flavor index.

Now specifically taking $\sigma = \{\uparrow, \downarrow\}$, and $H = H_0 + H_{\text{int}}$ with H_{int} as the Hubbard interaction,

$$H_{\text{int}} = U \sum_{\sigma} (n_{i,\uparrow} - 1/2)(n_{i,\downarrow} - 1/2), \quad (2)$$

the model has an $SU(2)$ symmetry that is broken at a critical coupling $U = U_c$ when the fermions transition from a massless Dirac fermion phase to an antiferromagnetic phase. Its critical behavior belongs to the Gross-Neveu Heisenberg universality class, corresponding to the Heisenberg $SU(2)$ symmetry broken in the effective continuous Gross-Neveu theory.[1]

However, it is not clear how to reach this same universality class using a four-fermion Lagrangian lattice theory. When staggered fermions are combined with a four-fermion term, as in [2], the same symmetry is not broken at the critical coupling and the universality class is different from that of the Hubbard model, as evidenced by different critical exponents in the calculation.

*e-mail: emilie.huffman@duke.edu. Work done in collaboration with Shailesh Chandrasekharan

Another related reason to use the Hamiltonian formalism is that sometimes a Hamiltonian is sign-problem-free, but the proposed Lagrangian path integral analog is not. For the following Hamiltonian model,

$$H = \sum_{x,d} -\eta_{x,d}t \left(c_x^\dagger c_{x+\hat{d}} + c_{x+\hat{d}}^\dagger c_x \right) + V \sum_{x,d} (n_x - 1/2) (n_{x+\hat{d}} - 1/2), \quad (3)$$

where x is a spatial lattice site, \hat{d} labels the directions such that x, d labels a unique nearest neighbor bond, and η_{xy} is defined as $\eta_{x,1} = 1$, $\eta_{x,2} = (-1)^{x_1}$, the bilinear terms in the Hamiltonian describe one flavor of four-component Dirac fermions at low energy. We can get one-flavor of massless Dirac fermions using reduced staggered fermions in the Lagrangian approach, but the interaction term will cause a sign problem.

On the other hand, the Hamiltonian version does not have a sign problem, as seen in [3, 4]. This model exhibits a second-order phase transition from massless fermions to a massive charge-density-wave (CDW) ordered phase, with a broken \mathbb{Z}_2 symmetry. This puts it in the Gross-Neveu Ising universality class.

The model in (3) will be our focus in the demonstration of a new continuous-time Hamiltonian Quantum Monte Carlo algorithm, which we call the Hamiltonian fermion bag algorithm. Since the solution was found, multiple calculations have been done to determine the critical exponents for the model (3), but there have been discrepancies between them, as summarized in Table 1. (Some of these calculations are actually for the model (3) on a honeycomb lattice with $\eta_{x,d} = 1$, which has the same critical behavior.) While the continuous time algorithms seem to have produced η -values of around .3, the discrete time MQMC algorithm is significantly different with an η value of .45. One might be tempted to assume the difference was due to discrete versus continuous time, but the discrete MQMC algorithm also used larger lattices than did the three continuous time results. In our application of the

Critical Exponents from Quantum Monte Carlo				
Past Work	ν	η	Critical Coupling $(V/t)_c$	Largest lattice (sites)
CT-INT (GS) [5, 6]	.80(3)	.302(7)	1.304	225
MQMC (GS) [7]	.77(3)	.45(2)	1.296	576
CT-INT (Finite-T) [8]	.74(4)	.275(25)	-	484

Table 1. Summary of earlier QMC results.

fermion bag algorithm, we manage to go to even larger lattices (our largest lattice size used has 2304 lattice sites) than the other calculations and find an η value that is larger (.54(6)) than that of the other lattice calculations. We conclude that larger lattices are necessary to get the critical behavior of model (3), and demonstrate that we have a new algorithm capable of doing calculations at larger lattices than before. Our calculations for lattices with 2304 lattice sites are on par with the largest lattices ever used for these spatially two-dimensional Hamiltonian models (2592 sites have been reached for the Hubbard model [1, 9]), and additionally we show stable behavior for the algorithm at lattices of 4096 sites.

2 Fermion Bag Method

2.1 Defining the Fermion Bags

A review of the application of fermion bags to Lattice Field Theories can be found in [10]. A key idea is to expand out the exponentiated fermionic terms in the path integral sum, and for each term

to divide the lattice space into smaller groups of sites, known as *fermion bags*, that correspond to the configuration.

For example, an exponentiated fermion-interaction term can be expanded as

$$e^{U \sum_{x,y} \bar{\psi}_x \psi_x \bar{\psi}_y \psi_y} = \sum_{\{n_{xy}\}} \prod_{x,y} (U \bar{\psi}_x \psi_x \bar{\psi}_y \psi_y)^{n_{xy}} \quad (4)$$

where $\bar{\psi}_x$ and ψ_x are Grassmann fields, and $\{n_{xy}\}$ represents a specific configuration of $n_{xy} = \pm 1$ values. In this way there would be no more fermion interaction terms in the exponent.

The fermion bags could then be defined for each term in the expansion. For example, suppose we had a term in the expansion for some model that was of the form [10]

$$I = \int [d\bar{\psi}d\psi] e^{-\bar{\psi}M\psi} (-U_1 \bar{\psi}_{x_1} \psi_{x_1}) (-U_2 \bar{\psi}_{x_2} \psi_{x_2}) (-U_3 \bar{\psi}_{x_3} \psi_{x_3} \bar{\psi}_{x_4} \psi_{x_4}). \quad (5)$$

For this particular term there would be four groups of sites that could be treated independently: (x_1) , (x_2) , (x_3, x_4) , and all the remaining sites. Grassmann integration could be performed over each of the four groups independently. The fermion bag idea has two main potential advantages: 1. sometimes sign problems in these alternative expansions are solved, and 2. these groupings of fermions into bags can be helpful in terms of allowing for fast algorithmic updates [11].

We now turn our attention back to the Hamiltonian formalism, and develop a Hamiltonian fermion bag algorithm based on similar principles. We assume a Hamiltonian that may be written in as a sum of terms of the following form

$$H_{x,d} = -\omega_{\langle x,d \rangle} e^{2\alpha_{\langle x,d \rangle} \sum_{a=1}^{N_f} (c_x^{a\dagger} c_{x+d}^a + c_{x+d}^{a\dagger} c_x^a)}, \quad (6)$$

where x is a spatial lattice site, and \hat{d} labels the directions such that x, d labels a unique nearest neighbor bond. The operators $c_x^{a\dagger}$ and c_x^a are fermionic creation and annihilation operators at the site x with a flavor $a = 1, 2, \dots, N_f$. The couplings of the model are defined through the real constants $\delta_{\langle x,d \rangle} > 0$ and $\alpha_{\langle x,d \rangle}$. To get model (3) from this model, up to a constant, we can set $\omega_{\langle x,d \rangle} = t^2 / (V(1 - (V/2t)^2))$, and $\alpha_{\langle x,d \rangle} = \alpha \eta_{x,d}$ where $\cosh 2\alpha = (1 + (V/2t)^2) / (1 - (V/2t)^2)$, and $\sinh 2\alpha = (V/t) / (1 - (V/2t)^2)$ [12]. Terms of the form (6) contain two key properties that are essential to this algorithm. First, they are *local*, since each only pertains to two site bond. Second, they are *exponentiated fermionic bilinears*, which will be important for the application of the Blankenbecler-Scalapino-Sugar (BSS) formula, given later on.

Assuming that $H = \sum_{s,d} H_{s,d}$ we use an expansion similar to the Stochastic Series Expansion (SSE). Here the partition function is expanded as

$$Z = \text{Tr} \left(e^{-\beta H} \right) = \sum_{k, [x,d]} \int [dt] (-1)^k \text{Tr} (H_{x_k d_k} \dots H_{x_2 d_2} H_{x_1 d_1}), \quad (7)$$

where there are k insertions of the bond Hamiltonian $H_{x,d}$ inside the trace at times $t_1 \leq t_2 \leq \dots \leq t_k$. The symbol $[dt]$ represents the k time-ordered integrals and $[x, d] = \{x_1, d_1, x_2, d_2, \dots, x_k, d_k\}$ represents the configuration of bonds at different times. Since a configuration of bonds also requires the information of the times where the bonds are inserted we label the configuration as $[x, d, t]$. A configuration of bonds can be represented pictorially, and we do so in the left image of Figure 1. Each bond represents the operator $H_{x,d}$ that is present inside the trace in (7). The horizontal axis labels the sites, and the vertical axis is imaginary time (inverse temperature). Now we define the concept of fermion bags in this Hamiltonian context. We can think of $H_{x,d}$ as creating quantum entanglement between

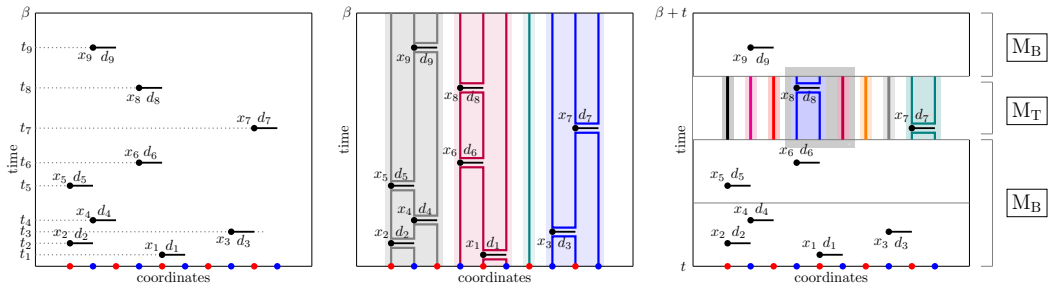


Figure 1. Images to illustrate the fermion bag concept. *Left:* One bond configuration in the expansion (7). *Middle:* The fermion bags for this same configuration where all of β is one timeslice. *Right:* The separation of the configuration into timeslices, and subsequent splitting into additional fermion bags. An update block is shaded gray, and definitions for M_B and M_T are also given.

the fermions at x and $x + \hat{d}$. Thus, all spatial sites connected by bonds to each other at various times become entangled with each other. Such a group of entangled sites can be defined as a fermion bag. Returning to Figure 1, the middle image shows the same configuration as the left image, but with the fermion bags highlighted. There are four fermion bags in the figure: three with three sites each, and one with a single site since no bonds touch it. When two bonds $\langle x, d \rangle$ and $\langle x', d' \rangle$ do not share a site between them the bond Hamiltonians commute, i.e., $[H_{x,d}, H_{x',d'}] = 0$. This property is key to a factorization that allows for faster updates.

However, once the temperature is low enough (and thus β is large enough), we should expect to have so many bonds that all the sites end up entangled into a single bag, leaving us with no ability to take advantage of small bags. We can circumvent this problem by dividing the low temperature space into several high temperature spaces. Since the space-time density of bonds is a physical quantity related to the energy density of the system [12], we expect a fixed density of bonds for a particular set of model parameters. At high temperatures we will have fewer bonds and many small fermion bags, and as mentioned before when the temperature is lowered fermion bags will begin to merge to form a single large fermion bag. This suggests that at some optimal temperature the fermion bags may efficiently break up the system into smaller regions that do not depend on the system size. Even at low temperatures, we may be able to divide the imaginary time axis into many time-slices and update a single time-slice efficiently. This is illustrated in the right image of Figure 1, where the imaginary time extent is divided into four-time slices and in the shaded time-slice there are eight fermion bags, instead of the four shown in the middle image.

A nice property is that the sizes of these fermion bags appear independent of the lattice size. In Figure 2 we took model 3 near its critical point, and taking $\beta = 4.0$ we divided the imaginary time direction into 16 time-slices and studied the fermion bag size as a function of the lattice size. For equilibrated configurations of $L = 48, 64$ and 100 , the average maximum fermion bag size within a time slice was about 30— independent of L (see Figure 2). Further tests suggests that the optimal temperature is roughly 0.25. Since bond insertions in different fermion bags commute with each other, we can efficiently update fermion bags in space-time blocks (shown as a box in the shaded time slice in the right image of Figure 1) involving 30 to 60 spatial sites within each time slice.

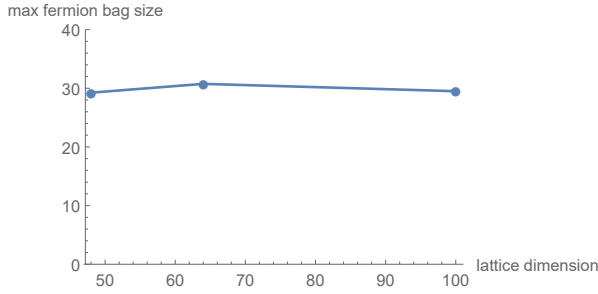


Figure 2. The maximum fermion bag size in a timeslice of .25 as a function of lattice size. We see that this size remains constant even for large lattices so long as the timeslice remains the same size.

2.2 Mathematical Details

To study the critical behavior of a model using the expansion (7), we need a mathematical expression for

$$\Omega_n([x, d, t]; t_0) = \text{Tr} [H_{x_k, d_k} \dots C_n \dots H_{x_2, d_2} H_{x_1, d_1}]. \quad (8)$$

Here $0 \leq t_0 \leq \beta$ is a time where the operator C_n is introduced. In the partition function sector $C_0 = I$ (the identity operator) and in the observable sector $C_1 = \hat{O}$, where \hat{O} is operator for the order parameter we would like to measure. The operator O must be written in terms of exponentiated bilinear operators.

Because every factor in (8) is an exponentiated bilinear factor, we can apply the BSS formula to (8) and get

$$\Omega_n([x, t, b]; t_0) = \det(\mathbb{1}_N + B_{x_k, d_k} \dots O_n \dots B_{x_2, d_2} B_{x_1, d_1}), \quad (9)$$

where $B_{x,d}$ and O_n are $N \times N$ matrices, where N is the number of sites, and now the matrices have a commutation property analogous to that of the $H_{x,d}$ operators in that $[B_{x,d}, B_{x',d'}] = 0$ if the corresponding bonds share no sites in common.

For model (3) in particular, B_{x_i, d_i} is the identity matrix except in a 2×2 block labeled by the rows and columns of the sites that touch the bond $\langle x_i, d_i \rangle$. Within this block, B_{x_i, d_i} takes the form

$$B_{x,d} = \begin{pmatrix} \cosh 2\alpha & \eta_{x,d} \sinh 2\alpha \\ \eta_{x,d} \sinh 2\alpha & \cosh 2\alpha \end{pmatrix}. \quad (10)$$

Also, the matrix O_0 is the identity, whereas O_1 comes from the order parameter operator $C = (-1)^{L/2} (n_{(0,0)} - 1/2) (n_{(L/2,0)} - 1/2)$ and is given by $(O_1)_{x,y} = \delta_{x,y} - 2\delta_{x,(0,0)} - 2\delta_{x,(L/2,0)}$.

Now we have what we need to make fast updates according to the fermion bags. As mentioned earlier, we divide the configuration space into smaller timeslices—for model (3) we used timeslices of width 0.25, with t_0 chosen to be at the beginning of the first time slice. We then update bonds within each time-slice sequentially. During the update of a time-slice we define two $N \times N$ matrices: the *background matrix* M_B (which is a product of all of the $B_{x,d}$ matrices outside the selected time-slice and O_n), and the *time-slice matrix* M_T , which is the product of all the $B_{x,d}$ matrices within the time-slice being updated. The right image of Figure 1 shows what contributes to M_B and M_T . When the configuration of bonds within the time-slice is changed then only M_T changes to M'_T . The ratio R is given by

$$R = \frac{\det(\mathbb{1}_N + M_B M'_T)}{\det(\mathbb{1}_N + M_B M_T)} = \det(\mathbb{1}_N + G_B \Delta), \quad (11)$$

where we have defined two new $N \times N$ matrices: $G_B = (\mathbb{1}_N + M_B M_T)^{-1} M_B M_T$ and $\Delta = (M_T^{-1} M_T' - \mathbb{1}_N)$. Since the bond matrices $B_{x,d}$ in different fermion bags commute, it is easy to verify that Δ is non-zero only within a block which contains spatial sites connected to fermion bags that change. If we randomly choose a spatial block containing about 30 – 60 sites and focus on updating the bonds only within that block, during such a *block-update* the size of the matrix Δ cannot be greater than the sum of the sites in the fermion bags that touch the sites within the block. We refer to this set of sites, which can be larger than the block size, as a *super-bag* and denote its size as s . Since Δ is non-zero only in an $s \times s$ block, it is easy to show that the computation of R (the ratio of the weight of the current configuration with that of the background configuration that existed at the time when the block update began) using (11), reduces to the computation of the determinant of an $s \times s$ matrix. Since G_B and M_T are fixed matrices during the entire *block-update*, they can be computed and stored. All proposals to update the current configuration within the block reduce to taking the determinant of an $s \times s$ matrix, independent of system size.

Beyond this main savings, there is one other aspect where the fermion bags allow us to save time, and that is for the updates to G_B . First note that given two matrices, M_1 and M_2 , we have the identity

$$(\mathbb{1} + M_1 M_2)^{-1} = (\mathbb{1} + M_2)^{-1} \left((\mathbb{1} + M_1)^{-1} (\mathbb{1} + M_2)^{-1} + (\mathbb{1} + M_1)^{-1} M_1 M_2 (\mathbb{1} + M_2)^{-1} \right)^{-1} (\mathbb{1} + M_1)^{-1}. \quad (12)$$

Further it is convenient that $(\mathbb{1} + M)^{-1} M = \mathbb{1} - (\mathbb{1} + M)^{-1}$. This means we can build G_B from partial versions labeled as $G_T = (\mathbb{1} + M_T)^{-1} M_T$ associated with each time slice T . This is important to keep the calculation of G_B stable. The matrix G_T in turn is obtained by combining the bond-specific $G_{x,d} = (\mathbb{1} + B_{x,d})^{-1} B_{x,d}$ matrices within a time slice. We can do this efficiently using the idea of fermion bags. For each fermion bag in a given time-slice, we first construct G_f from the B_{x_1,d_1} to B_{x_f,d_f} matrices that belong to a specific fermion bag. Thus the non-identity block of each G_f is a matrix with f rows and f columns corresponding to the fermion bag sites. We can then combine the G_f matrices into a G_T matrix, which has distinct blocks according to the fermion bags. Thus, while G_B is an $N \times N$ matrix, we can use the idea of fermion bags along with the identity (12) to reduce the number of $O(N^3)$ operations. While this does not reduce the scaling of the algorithm, it does significantly reduce the prefactor.

These are the main aspects of the algorithmic updates that are helped by the fermion bags. The full details of the algorithm are given in [13].

3 Results

The results at the time of Lattice 2017 are summarized by the plots in Figure 3. The left image demonstrates the stability of the algorithm by showing the high-temperature equilibration of 100×100 lattices (10,000 sites) at $\beta = 4.0$. Now low temperature equilibrations of 64×64 lattices and 100×100 lattices can be found in [13]. For $\beta = L$ sweeps on 48×48 , 64×64 , and 100×100 lattices on a single core are 4 hours, 30 hours, and 30 days respectively.

The middle figure and left figure give our calculations that were done at the couplings $V = 1.304t$ and $V = 1.296t$. These were the critical couplings found by [6] and [7]. At the critical coupling, the power law behavior of

$$\langle C \rangle = \frac{a}{L^{1+\eta}}, \quad (13)$$

is expected. Our calculations include lattices up to size 48×48 (2304 sites) and fit the data to (13). For $V = 1.304t$, the power law fit was poor and the η value was only .149(9). However, when we

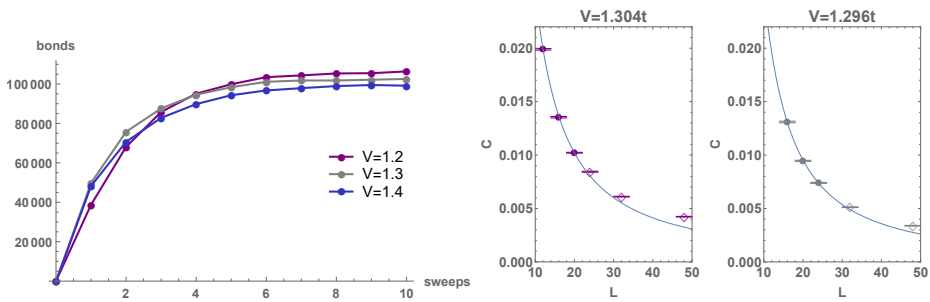


Figure 3. *Left:* High temperature ($\beta = 4.0$) equilibration of large lattices for several couplings. *Middle,Right:* Data for the couplings $V = 1.304t$ and $V = 1.296t$, which were the critical couplings found in [6, 7].

dropped the larger lattices and only kept up to lattice size $L = 20$, we could achieve a good fit with $\eta = .31(2)$, consistent with the previous result of $\eta = .318(8)$ that went up to $L = 20$ for specifically model (3).[5, 6, 8] For $V = 1.296t$, the power law fit was again poor with an η value of $.35(1)$. Again dropping the larger lattices and this time keeping lattices up to $L = 24$ led to a good fit with $\eta = .41(4)$, consistent with the previous result of $\eta = .43(2)$ that went up to $L = 22$ for specifically model (3).[7] Both of these results suggest that larger lattices make a substantial difference in the calculation and therefore are important in determining the critical behavior. In [13] we combine our data from several different couplings (using lattices up to $L = 48$) and use a combined fit to calculate exponents of $\eta = .54(6)$, $\nu = .88(2)$, and $V_c = 1.279(3)t$.

4 Conclusion

We have shown how the fermion bag idea may be adapted into an algorithm suited to a Hamiltonian context and given an example that shows the stability as well as potential of this algorithm. By using lattices that were five times larger than those used in previous calculations we have shown the both the algorithm's power and the necessity of larger lattice calculations in determining the critical behavior for this Gross-Neveu Ising universality class.

References

- [1] Y. Otsuka, S. Yunoki, S. Sorella, Phys. Rev. X **6**, 011029 (2016)
- [2] S. Chandrasekharan, A. Li, Phys. Rev. D **88**, 021701 (2013)
- [3] E.F. Huffman, S. Chandrasekharan, Phys. Rev. **B89**, 111101 (2014), 1311.0034
- [4] Z.X. Li, Y.F. Jiang, H. Yao, Phys. Rev. B **91**, 241117 (2015)
- [5] L. Wang, P. Corboz, M. Troyer, New Journal of Physics **16**, 103008 (2014)
- [6] L. Wang, M. Iazzi, P. Corboz, M. Troyer, Phys. Rev. **B91**, 235151 (2015), 1501.00986
- [7] Z.X. Li, Y.F. Jiang, H. Yao, New Journal of Physics **17**, 085003 (2015)
- [8] S. Hesselmann, S. Wessel, Phys. Rev. **B93**, 155157 (2016), 1602.02096
- [9] S. Sorella, Y. Otsuka, S. Yunoki, **2**, 992 EP (2012), article
- [10] S. Chandrasekharan, The European Physical Journal A **49**, 90 (2013)
- [11] V. Ayyar, S. Chandrasekharan, Phys. Rev. **D93**, 081701 (2016), 1511.09071
- [12] L. Wang, Y.H. Liu, M. Troyer, Phys. Rev. B **93**, 155117 (2016)

- [13] E. Huffman, S. Chandrasekharan, [1709.03578] (2017), **1709.03578**