Here I would like to introduce Quantum Mento Carlo (QMC) routine and sign problem first, then introduce the model we use, and why the average sign is limited.

1. Quantum Mento Carlo

Imagine we have a fermion system with Hamiltonian,

$$H_I = \sum_q V(q)\rho_{-q}\rho_q = \sum_q V(q)\rho_q^\dagger \rho_q$$

$$\rho_q = \sum_{i,j} \lambda_{i,j}(q) c_i^\dagger c_j - \frac{1}{2}\mu$$

We require $V(q) = V(-q) > 0$, μ is real and $Tr\left(\lambda_{i,j}(q)\right) = \mu$. In physics, it is a general form for Coulomb repulsion in density channel with chemical potential μ for half-filled and flat band limit.

Next, we would like to introduce HS transformation to decouple 4-fermion interaction to 2-fermion coupled with an auxiliary field. Here is the general HS transformation and its discrete version, one can see it as a Gaussian integral and parameter fitting after Taylor expansion.

$$e^{\alpha\hat{O}^2} = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}x^2} e^{\sqrt{2\alpha}x\hat{O}} dx$$

$$e^{\alpha\hat{O}^2} = \frac{1}{4} \sum_{l=\pm1,\pm2} \gamma(l) e^{\sqrt{\alpha}\eta(l)\hat{O}} + O(\alpha^4)$$

$$\gamma(\pm1) = 1 + \frac{\sqrt{6}}{3}, \gamma(\pm2) = 1 - \frac{\sqrt{6}}{3}, \eta(\pm1) = \pm\sqrt{2(3-\sqrt{6})}, \eta(\pm2) = \pm\sqrt{2(3+\sqrt{6})}$$

Write down its partition function, to get a small α, we need Trotter decomposition first and do HS transformation for each decomposed slice.

$$Z = Tr(e^{-\beta H_I}) = Tr\left(\prod_t e^{-\Delta\tau H_I(t)}\right)$$

$$= Tr\left(\prod_t e^{-\Delta\tau \sum_q V(q)\left[(\rho_{-q}+\rho_q)^2 - (\rho_{-q}-\rho_q)^2\right]}\right)$$

$$\approx \sum_{\{l_{|q|,t}\}} \prod_t \left[\prod_{|q|} \frac{1}{16} \gamma(l_{|q_1|,t}) \gamma(l_{|q_2|,t})\right] Tr\left\{\prod_t \left[\prod_{|q|} e^{i\eta(l_{|q_1|,t})A_q(\rho_{-q}+\rho_q)} e^{\eta(l_{|q_2|,t})A_q(\rho_{-q}-\rho_q)}\right]\right\}$$

$$A_q = \sqrt{\Delta\tau V(q)}$$

For any observable $\hat{O}$, its ensemble average is

$$\langle\hat{O}\rangle = \frac{Tr(\hat{O}e^{-\beta H_I})}{Tr(e^{-\beta H_I})} = \sum_{\{l_{|q|,t}\}} \frac{P(\{l_{|q|,t}\}) Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})] \frac{Tr[\hat{O}\prod_t \hat{B}_t(\{l_{|q|,t}\})]}{Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]}}{\sum_{\{l_{|q|,t}\}} P(\{l_{|q|,t}\}) Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]}$$

$$P(\{l_{|q|,t}\}) = \prod_t \left[\prod_{|q|} \frac{1}{16} \gamma(l_{|q_1|,t}) \gamma(l_{|q_2|,t}) \right]$$

$$\hat{B}_t(\{l_{|q|,t}\}) = \prod_{|q|} e^{i\eta(l_{|q_1|,t})A_q(\rho_{-q}+\rho_q)} e^{\eta(l_{|q_2|,t})A_q(\rho_{-q}-\rho_q)}$$

One can see orange part as a possibility distribution and blue part as sample value at certain configuration $l_{|q|,t}$. Then Mento Carlo can finish the sample according to this distribution and get.

2. Sign problem

One can see $P(\{l_{|q|,t}\})$ is always positive, but we cannot make sure $Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]$ is always non-negative for all configuration, generally. If it is the case $Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]$ is not always non-negative, we carry out our simulation by equation below.

$$\langle \hat{O} \rangle = \frac{Tr(\hat{O}e^{-\beta H_I})}{Tr(e^{-\beta H_I})} = \sum_{\{S\}} \frac{P(S)O(S)}{\sum_{\{S\}} P(S)} = \frac{\dfrac{\sum_{\{S\}} |P(S)| \frac{P(S)O(S)}{|P(S)|}}{\sum_{\{S\}} |P(S)|}}{\dfrac{\sum_{\{S\}} P(S)}{\sum_{\{S\}} |P(S)|}}$$

The numerator part is nothing but a Mento Carlo simulation with a non-negative distribution and an observable modified with a phase. The problem here is the denominator $\langle sign \rangle = \frac{\sum_{\{S\}} P(S)}{\sum_{\{S\}} |P(S)|}$. If this one is small, the fluctuation is huge, and we hardly get a meaningful result.

3. Why $Tr\left(\lambda_{i,j}(q)\right) = \mu$ is useful?

In our last paper computing Twisted Bilayer Graphene (TBG), we proved at this situation, $Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]$ is always real. Besides, we observed that $\langle sign \rangle$ is a finite at low temperature, and not going to zero very fast with system size, which we can hardly give a reasonable explanation. Here I will introduce our proof that $Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]$ is always real first.

According to equation below, where M_i are linear combination of 2-fermion operators.

$$Tr(e^{M_1} e^{M_2} \cdots e^{M_n}) = \det(I + e^{M_1} e^{M_2} \cdots e^{M_n})$$

We can see

$$\hat{B}_t(\{l_{|q|,t}\}) = \prod_{|q|} e^{i\eta(l_{|q_1|,t})A_q(\rho_{-q}+\rho_q)} e^{\eta(l_{|q_2|,t})A_q(\rho_{-q}-\rho_q)}$$

$$Tr\left[\prod_t \hat{B}_t(\{l_{|q|,t}\}) \right] = e^{-\frac{1}{2}\sum_j Tr(M_j)} \det(I + e^{M_1} e^{M_2} \cdots e^{M_n})$$

By noticing M_i are all anti-Hermitian so that e^{M_i} are all unitary, we define a unitary operator $U = e^{M_1} e^{M_2} \cdots e^{M_n}$, $\det(U) = e^{\sum_j Tr(M_j)} = e^{\sum_\alpha i\lambda_\alpha} = e^{i\Gamma}$. Here $e^{i\lambda_\alpha}$ are eigenvalues of U.

$$e^{-i\frac{\Gamma}{2}} \det(I + U) = e^{-i\frac{\Gamma}{2}} \prod_\alpha (1 + e^{i\lambda_\alpha})$$

For any term $e^{i\left(\sum_{k\in A} \lambda_k - \frac{\Gamma}{2}\right)}$, there is always a term $e^{i\left(\sum_{k\notin A} \lambda_k - \frac{\Gamma}{2}\right)} = e^{i\left(-\sum_{k\in A} \lambda_k + \frac{\Gamma}{2}\right)}$. Add all terms together,

$$Tr\left[\prod_t \hat{B}_t(\{l_{|q|,t}\})\right] = e^{-i\frac{\Gamma}{2}}\det(I+U) = \sum_A 2\cos\left(\sum_{k\in A}\lambda_k - \frac{\Gamma}{2}\right)$$

So, we proved $P(\{l_{|q|,t}\})Tr[\prod_t \hat{B}_t(\{l_{|q|,t}\})]$ is real.

4. Why $\langle sign\rangle$ is finite?

We will give a very rough range of $\langle sign\rangle$ in this page and state $\langle sign\rangle$ will not continue decay with β for finite matrix size D. Here we assume $Tr\left(\lambda_{i,j}(q)\right) = \mu = 0$.

The Hamiltonian before decoupled is in a form below,

$$H_I = \sum_q V_q\rho_q^\dagger\rho_q = \sum_{|q|}\frac{V_{|q|}}{2}\left[\left(\rho_q^\dagger+\rho_q\right)^2 - \left(\rho_q^\dagger-\rho_q\right)^2\right]$$

Any $\langle H_I\rangle$ is no less than 0, so the ground states satisfy $\rho_q|\varphi_0\rangle = 0$. $|\varphi_0\rangle$ should also be the eigenstate of total particle number N (particle number conservation).

Since there is no constant term in $\rho_q = \sum_{i,j}\lambda_{i,j}(q)c_i^\dagger c_j$, there are at least two $|\varphi_0\rangle$ satisfy $\rho_q|\varphi_0\rangle = 0$ that is full occupy and empty situations.

Now we would like to know

$$Z = \lim_{N_b\to\infty} Tr((e^{-\Delta t H})^{N_b})$$

It is then obvious that in this limit, only ground state energy will contribute to Z.

$$Z = \lim_{\beta\to\infty} ke^{-\beta E_0} = k$$

Here k is degeneracy of ground state, and $E_0 = 0$ exactly. From discussion above we know k is no less than 2.

Now let us consider another ρ_q where dimension doubled with the same copy in original ρ_q,

$$\rho_q \to \rho_{q,1} + \rho_{q,2} = \sum \lambda_{i,j}(q)c_{i,1}^\dagger c_{j,1} + \sum \lambda_{i,j}(q)c_{i,2}^\dagger c_{j,2}$$

Since $Tr(\rho_{q,1}) = Tr(\rho_{q,2}) = 0$, one should notice that for n-particle state $|\varphi_{0,1}\rangle$ which makes $\rho_{q,1}|\varphi_{0,1}\rangle = \lambda_n|\varphi_{0,1}\rangle$, there is always another $D-n$ particles state satisfying $\rho_{q,1}|\varphi'_{0,1}\rangle = -\lambda_n|\varphi'_{0,1}\rangle$. So, $|\varphi_{0,1}\rangle\otimes|\varphi'_{0,2}\rangle$ is a new ground state for $\rho_q^\dagger\rho_q$. This will lift degeneracy of ground state from k to some value m, here $k^2 < m < k^2 + 2^D - 2$ (at least m is independent with β and finite). $2^D - 2$ comes from $\sum_{i=1}^{D-1} C_D^i = 2^D - 2$, and $<$ comes from those new ground states are not necessarily ground state for all q.

In QMC decoupled Hamiltonian, the argument above means,

$$Z = \sum_{\{\phi\}} P(\phi)D(\phi) = k$$

$$Z_2 = \sum_{\{\phi\}} P(\phi)D^2(\phi) = m$$

Since we have discussed for this kind of ρ_q, determinants $D(\phi)$ are always real, by defining

$$Z_r = \sum_{\{\phi\}} P(\phi)|D(\phi)| = \langle|D(\phi)|\rangle > 0$$

We would expect

$$\langle D(\phi)\rangle^2 \le \langle |D(\phi)|\rangle^2 \le \langle D^2(\phi)\rangle$$
$$k \le Z_r \le \sqrt{m}$$

Finally, for $\langle sign\rangle$

$$\frac{k}{\sqrt{k^2+2^D-2}} \le \frac{k}{\sqrt{m}} \le \langle sign\rangle \le 1$$

This is the result we get before. Since 2^D grows exponentially with D, even though we controlled this range of $\langle sign\rangle$, this control is not good enough to explain much larger $\langle sign\rangle$ than $\frac{k}{\sqrt{k^2+2^D-2}}$ observed in simulation.

Very interestingly, I find another operator which commute with Hamiltonian after we introduce $\rho_q \to \rho_{q,1}+\rho_{q,2}$, so that the up limit of degeneracy m can be smaller than k^2+2^D-2. The operator is defined below,

$$\Delta = \sum_{i'} c^\dagger_{i',1} c_{i',2}$$

One can check,

$$[\Delta, \rho_{q,1}+\rho_{q,2}] = \left[\sum_{i'} c^\dagger_{i',1} c_{i',2}, \sum \lambda_{i,j}(q) c^\dagger_{i,1} c_{j,1} + \sum \lambda_{i,j}(q) c^\dagger_{i,2} c_{j,2}\right]$$
$$= \sum -\lambda_{i,j}(q) c^\dagger_{i,1} c_{j,2} + \sum \lambda_{i,j}(q) c^\dagger_{i,1} c_{j,2} = 0$$

So, if $|\psi_0\rangle$ is a ground state, $\Delta|\psi_0\rangle$ is also a ground state with adding a particle in subspace 1 and reducing a particle in subspace 2. For most general $\lambda_{i,j}(q)$ (say generated randomly), $k=2$ for full occupy and empty ground states $|\psi_F\rangle$ and $|\psi_E\rangle$. After $\rho_q \to \rho_{q,1}+\rho_{q,2}$, we can see $|\psi_{F,1}\rangle \otimes |\psi_{F,2}\rangle$ and $|\psi_{E,1}\rangle \otimes |\psi_{E,2}\rangle$ are 2 ground states. Besides, $(\Delta)^n |\psi_{E,1}\rangle \otimes |\psi_{F,2}\rangle$ also gives $D+1$ ground states. Then we have $m = D+3$ for $k=2$.

$$\frac{2}{\sqrt{D+3}} = \frac{k}{\sqrt{m}} \le \langle sign\rangle \le 1$$

More general case ($k>2$ or $\mu \ne 0$)…

General $\mu \ne 0$ seems can not promise $\langle sign\rangle$ is finite large, one also need there is at least one zero energy ground state. Because we know space doubled situation always has zero energy ground state (Also described by $(\Delta)^n |\psi_{E,1}\rangle \otimes |\psi_{F,2}\rangle$ with $D+1$ ground states), if there is no zero-energy ground state for original one, partition function will always decrease exponentially with temperature. While for spin-polarized valley-polarized TBG at chiral limit, there are 2 ground states, and after doubling we can construct two groups of independent Δ like operators. Since for each Δ like operator, it will give $\frac{D}{2}+1$ degenerate states, totally ground states after doubling should be $\left(\frac{D}{2}+1\right)^2$. $\langle sign\rangle \ge \frac{4}{D+2}$ in this case. The proof in detail is written below.

$$H_I = \sum_q V(q)\rho_{-q}\rho_q = \sum_q V(q)\rho_q^\dagger \rho_q$$

$$\rho_q = \sum_{i,j,m,n} \lambda_{i,m,j,n}(q) c^\dagger_{i,m} c_{j,n} - \frac{1}{2}\mu$$

Here m, n are band indexes. $\lambda_{i,m,j,n}(q) = \lambda_{i,-m,j,-n}(q)$ is true for chiral limit after gauge fixing. And we call H_{I2} Hamiltonian after doubling. Define

$$\Delta_1^\dagger = \sum_{i',m} c_{i',m,1}^\dagger c_{i',m,2}$$

$$\Delta_2^\dagger = \sum_{i',m} c_{i',m,1}^\dagger c_{i',-m,2}$$

One can see $[\Delta_1^\dagger, H_{I2}] = [\Delta_2^\dagger, H_{I2}] = [\Delta_1^\dagger + \Delta_2^\dagger, \Delta_1^\dagger - \Delta_2^\dagger] = [\Delta_1^\dagger + \Delta_2^\dagger, \Delta_1 - \Delta_2] = 0$.

Then we can see two independent boson-like operators $D_1^\dagger = \Delta_1^\dagger + \Delta_2^\dagger, D_2^\dagger = \Delta_1^\dagger - \Delta_2^\dagger$. One can apply each operator $\frac{D}{2}$ times at most, so that $\left(\frac{D}{2} + 1\right)^2$ ground states conclusion can be derived.

Get $Z_r = \sum_{\{\phi\}} P(\phi)|D(\phi)| = \langle |D(\phi)| \rangle$ exactly…