Momentum space quantum Monte Carlo on twisted bilayer Graphene (Supplemental Material)

I. PROPERTIES OF THE FORM FACTOR λ AND PROOF OF THE SIGN STRUCTURE

Here we follow the discussion in band basis as in Ref.[1].
The Hermiticity condition for λ is

$$\lambda_{m,n,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G}) = \lambda^*_{n,m,\tau}(\mathbf{k}+\mathbf{q}+\mathbf{G},\mathbf{k}). \tag{1}$$

The $C_{2z}T$ symmetry operates on λ as

$$\lambda_{m,n,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G}) = \lambda^*_{m,n,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G}), \tag{2}$$

which means λ is real.
The $C_{2z}P$ symmetry operates on λ as

$$\lambda_{m,n,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G}) = m*n*\lambda_{-m,-n,-\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G}). \tag{3}$$

Based on these relations, we can derive the Statement 3 in the main text as:

$$\begin{aligned}
\delta\rho_{\mathbf{q}+\mathbf{G},-\tau} &= \sum_{\mathbf{k},m,n} \lambda_{m,n,-\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G})(d^\dagger_{\mathbf{k},m,-\tau}d_{\mathbf{k}+\mathbf{q},n,-\tau} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n}) \\
&= \sum_{\mathbf{k},m,n} -m*n*\lambda_{m,n,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G})(d_{\mathbf{k}+\mathbf{q},-n,-\tau}d^\dagger_{\mathbf{k},-m,-\tau} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n}) \\
&= \sum_{\mathbf{k},m,n} -\lambda_{m,n,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G})(\tilde{d}^\dagger_{\mathbf{k}+\mathbf{q},n,-\tau}\tilde{d}_{\mathbf{k},m,-\tau} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n}) \\
&= \sum_{\mathbf{k},m,n} -\lambda_{n,m,\tau}(\mathbf{k},\mathbf{k}-\mathbf{q}-\mathbf{G})(\tilde{d}^\dagger_{\mathbf{k},n,-\tau}\tilde{d}_{\mathbf{k}-\mathbf{q},m,-\tau} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n})
\end{aligned} \tag{4}$$

Here $\tilde{d}_{\mathbf{k},m,-\tau} = m*d^\dagger_{\mathbf{k},-m,-\tau}$, $m,n \in \{\pm 1\}$. According to the last line, single-particle matrices in the fermion determinant between two valleys satisfy $\delta\rho_{\mathbf{q}+\mathbf{G},-\tau} = -\delta\rho_{-\mathbf{q}-\mathbf{G},\tau}$. Besides, this transformation does not change kinetic terms. Since

$$\varepsilon_{m,\tau}(\mathbf{k}) = -\varepsilon_{-m,\tau}(-\mathbf{k}) = \varepsilon_{m,-\tau}(-\mathbf{k}) \tag{5}$$

So that

$$\begin{aligned}
&\varepsilon_{m,\tau}(\mathbf{k})d^\dagger_{\mathbf{k},m,\tau}d_{\mathbf{k},m,\tau} + \varepsilon_{-m,\tau}(-\mathbf{k})d^\dagger_{-\mathbf{k},-m,\tau}d_{-\mathbf{k},-m,\tau} \\
&= \varepsilon_{m,\tau}(\mathbf{k})(d^\dagger_{\mathbf{k},m,\tau}d_{\mathbf{k},m,\tau} - d^\dagger_{-\mathbf{k},-m,\tau}d_{-\mathbf{k},-m,\tau})
\end{aligned} \tag{6}$$

$$\begin{aligned}
&\varepsilon_{-m,-\tau}(\mathbf{k})d^\dagger_{\mathbf{k},-m,-\tau}d_{\mathbf{k},-m,-\tau} + \varepsilon_{m,-\tau}(-\mathbf{k})d^\dagger_{-\mathbf{k},m,-\tau}d_{-\mathbf{k},m,-\tau} \\
&= \varepsilon_{m,\tau}(\mathbf{k})(d_{\mathbf{k},-m,-\tau}d^\dagger_{\mathbf{k},-m,-\tau} - d_{-\mathbf{k},m,-\tau}d^\dagger_{-\mathbf{k},m,-\tau}) \\
&= \varepsilon_{m,\tau}(\mathbf{k})(\tilde{d}^\dagger_{\mathbf{k},m,-\tau}\tilde{d}_{\mathbf{k},m,-\tau} - \tilde{d}^\dagger_{-\mathbf{k},-m,-\tau}\tilde{d}_{-\mathbf{k},-m,-\tau})
\end{aligned} \tag{7}$$

One can see kinetic terms can be viewed as complex conjugated between two valleys since the dispersion $\varepsilon_{m,\tau}(\mathbf{k})$ is real. We note similar observation, that the TBG Hamiltonian at CNP after QMC decoupling is invariant under anti-unitary particle-hole symmetry, is also pointed out in Refs. [2, 3].

We derive the Statement 2 in another way, different from that in the main text, and then discuss the flat band kinetic terms in this case.

$$\begin{aligned}\delta\rho_{\mathbf{q}+\mathbf{G},-s} &= \sum_{\mathbf{k},m,n} \lambda_{m,n}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G})(d^\dagger_{\mathbf{k},m,-s}d_{\mathbf{k}+\mathbf{q},n,-s} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n}) \\ &= \sum_{\mathbf{k},m,n} -\lambda_{n,m}(\mathbf{k},\mathbf{k}-\mathbf{q}-\mathbf{G})(d_{\mathbf{k},n,-s}d^\dagger_{\mathbf{k}-\mathbf{q},m,-s} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n}) \\ &= \sum_{\mathbf{k},m,n} -\lambda_{n,m}(\mathbf{k},\mathbf{k}-\mathbf{q}-\mathbf{G})(\tilde{d}^\dagger_{\mathbf{k},n,-s}\tilde{d}_{\mathbf{k}-\mathbf{q},m,-s} - \frac{1}{2}\delta_{\mathbf{q},0}\delta_{m,n})\end{aligned} \tag{8}$$

In the last line we just define $\tilde{d}_{\mathbf{k},m,-s} = d^\dagger_{\mathbf{k},m,-s}$, single-particle matrices in the fermion determinant between two spins also satisfy $\delta\rho_{\mathbf{q}+\mathbf{G},-s} = -\delta\rho_{-\mathbf{q}-\mathbf{G},s}$. But this transformation add a minus sign to the kinetic terms.

$$\begin{aligned}&\varepsilon_m(\mathbf{k})d^\dagger_{\mathbf{k},m,-s}d_{\mathbf{k},m,-s} + \varepsilon_{-m}(-\mathbf{k})d^\dagger_{-\mathbf{k},-m,-s}d_{-\mathbf{k},-m,-s} \\ &= -\varepsilon_m(\mathbf{k})(d_{\mathbf{k},m,-s}d^\dagger_{\mathbf{k},m,-s} - d_{-\mathbf{k},-m,-s}d^\dagger_{-\mathbf{k},-m,-s}) \\ &= -\varepsilon_m(\mathbf{k})(\tilde{d}^\dagger_{\mathbf{k},m,-s}\tilde{d}_{\mathbf{k},m,-s} - \tilde{d}^\dagger_{-\mathbf{k},-m,-s}\tilde{d}_{-\mathbf{k},-m,-s})\end{aligned} \tag{9}$$

It is thus not obvious when consider both spin degree of freedom for a single valley, when including the kinetic energy, the QMC simulation is still absent of sign-problem.

II. MEASUREMENT OF THE GREEN'S FUNCTION IN QMC

Here we introduce the measurement of Green's funtions in QMC. Generally, average of any observables $\hat{O}$ can be written as,

$$\left\langle \hat{O} \right\rangle = \frac{\mathrm{Tr}(\hat{O}e^{-\beta H})}{\mathrm{Tr}(e^{-\beta H})} = \sum_{\{l_{|\mathbf{q}|,t}\}} \frac{P(\{l_{|\mathbf{q}|,t}\})\,\mathrm{Tr}[\prod_t \hat{B}_t(\{l_{|\mathbf{q}|,t}\})]\,\frac{\mathrm{Tr}[\hat{O}\prod_t \hat{B}_t(\{l_{|\mathbf{q}|,t}\})]}{\mathrm{Tr}[\prod_t \hat{B}_t(\{l_{|\mathbf{q}|,t}\})]}}{\sum_{\{l_{|\mathbf{q}|,t}\}} P(\{l_{|\mathbf{q}|,t}\})\,\mathrm{Tr}[\prod_t \hat{B}_t(\{l_{|\mathbf{q}|,t}\})]} \tag{10}$$

According to Eq. (4) in the main text, $P(\{l_{|\mathbf{q}|,t}\}) = \prod_t[\prod_{|\mathbf{q}+\mathbf{G}|\neq 0} \frac{1}{16}\gamma\left(l_{|\mathbf{q}|_1,t}\right)\gamma\left(l_{|\mathbf{q}|_2,t}\right)]$ and

$$\hat{B}_t(\{l_{|\mathbf{q}|,t}\}) = \prod_{|\mathbf{q}+\mathbf{G}|\neq 0} e^{i\eta\left(l_{|\mathbf{q}|_1,t}\right)A_{\mathbf{q}}(\delta\rho_{-\mathbf{q}}+\delta\rho_{\mathbf{q}})}e^{\eta\left(l_{|\mathbf{q}|_2,t}\right)A_{\mathbf{q}}(\delta\rho_{-\mathbf{q}}-\delta\rho_{\mathbf{q}})}, \tag{11}$$

respectively.

Since M_j is now of the form of fermion bilinear, one can trace out the fermion operator to obtain the determinant

$$\mathrm{Tr}[e^{M_1}e^{M_2}...e^{M_n}] = \det[I + e^{M_1}e^{M_2}...e^{M_n}], \tag{12}$$

the obtained determinant is the single particle fermion determinant for QMC. Exact constant terms in $\hat{B}(\{l_{|\mathbf{q}|,t}\})$ as $e^{-\frac{1}{2}\sum_j \mathrm{Tr}(M_j)}$ represent the left part in the single particle basis e^{M_t}, and it is from such structure

$$e^{-\frac{1}{2}\sum_j \mathrm{Tr}(M_j)}\det\left(I + e^{M_1}e^{M_2}...e^{M_n}\right) \tag{13}$$

we prove the Statement 1 in the main text.

Now consider operator $\hat{O} = d_i(\tau)d^\dagger_j(\tau)$, which is the single-particle Green's function $G(\tau,\tau)_{i,j} = \langle d_i(\tau)d^\dagger_j(\tau)\rangle$ in QMC,

$$\frac{\mathrm{Tr}[\prod_{t_1>\tau}\hat{B}_{t_1}(\{l_{|\mathbf{q}|,t}\})d_i(\tau)d^\dagger_j(\tau)\prod_{t_2\leqslant\tau}\hat{B}_{t_2}(\{l_{|\mathbf{q}|,t}\})]}{\mathrm{Tr}[\prod_t \hat{B}_t(\{l_{|\mathbf{q}|,t}\})]} = \left(I + e^{M_\tau}e^{M_{\tau-1}}...e^{M_1}\cdot e^{M_n}e^{M_{n-1}}...e^{M_{\tau+1}}\right)^{-1}_{i,j}. \tag{14}$$

Statement 4 If there is no kinetic term, Green's function $G(\tau,\tau)_{i,j}$ should always be diagonal with all non-zero elements 0.5.

Proof For a certain configuration $U_S = B_{1,s_1}B_{2,s_2}...B_{n,s_n}$, ignoring matrix commutation in each B_{t,s_1}, we can always find another configuration $U_{S'} = B_{1,-s_n}...B_{n-1,-s_2}B_{n,-s_1} = U_S^{-1}$ with the same weight

$$e^{-\frac{1}{2}\sum_j \mathrm{Tr}(M_j)} \det\left(I + e^{M_1}e^{M_2}...e^{M_n}\right) = e^{\frac{1}{2}\sum_j \mathrm{Tr}(M_j)} \det\left(I + (e^{M_1}e^{M_2}...e^{M_n})^{-1}\right). \tag{15}$$

We add Green's funstion $G(\tau,\tau)_{i,j}$ of this pair configurations together, which is $(I + e^{M_1}e^{M_2}...e^{M_n})^{-1} + (I + (e^{M_1}e^{M_2}...e^{M_n})^{-1})^{-1} = I$. Remeber we need to divide it by 2 for average of these two configurations, therefore after sampling and average, $G(\tau,\tau)_{i,j}$ should always be diagonal with all non-zero elements 0.5.

Besides, consider $G(\tau,0)_{i,j} = \langle d_i(\tau)d_j^\dagger(0)\rangle$,

$$\frac{\mathrm{Tr}[\prod_{t_1>\tau}\hat{B}_{t_1}(\{l_{|\mathbf{q}|,t}\})d_i(\tau)\prod_{t_2\leqslant\tau}\hat{B}_{t_2}(\{l_{|\mathbf{q}|,t}\})d_j^\dagger(0)]}{\mathrm{Tr}[\prod_t \hat{B}_t(\{l_{|\mathbf{q}|,t}\})]} = (G(\tau,\tau)\cdot e^{M_\tau}e^{M_{\tau-1}}...e^{M_1})_{i,j} \tag{16}$$

We also introduce a useful way below to obtain real valued measurables for other fillings, in which, the sign of the weight is always real. The definition of $\delta\rho_{\mathbf{q}+\mathbf{G}}$ should be rewritten as:

$$\delta\rho_{\mathbf{q}+\mathbf{G}} = \sum_{\mathbf{k}\in mBZ, m_1,m_2,\tau,s} \lambda_{m_1,m_2,\tau}(\mathbf{k},\mathbf{k}+\mathbf{q}+\mathbf{G})(d^\dagger_{\mathbf{k},m_1,\tau,s}d_{\mathbf{k}+\mathbf{q},m_2,\tau,s} - C\delta_{\mathbf{q},0}\delta_{m_1,m_2}) = (\delta\rho_{-\mathbf{q}-\mathbf{G}})^\dagger. \tag{17}$$

Here C can be any real number.

Assume we would like to measure $\langle F\rangle = \langle F^*\rangle = \frac{\langle F+F^*\rangle}{2}$. For any configuration S, we can define a dual configuration $-S$. For $-S$, in $\hat{B}_t(\{l_{|\mathbf{q}|,t}\}) = \prod_{|\mathbf{q}+\mathbf{G}|\neq 0} e^{i\eta\left(l_{|\mathbf{q}|_1,t}\right)A_{\mathbf{q}}(\delta\rho_{-\mathbf{q}}+\delta\rho_{\mathbf{q}})}e^{\eta\left(l_{|\mathbf{q}|_2,t}\right)A_{\mathbf{q}}(\delta\rho_{-\mathbf{q}}-\delta\rho_{\mathbf{q}})}$, all $\eta\left(l_{|\mathbf{q}|_2,t}\right)$ do not change and all other $\eta\left(l_{|\mathbf{q}|,t}\right)$ change to $-\eta\left(l_{|\mathbf{q}|,t}\right)$. Then one can see $\hat{B}_t(\{l_{|\mathbf{q}|,t}\})$ change to $\hat{B}_t^*(\{l_{|\mathbf{q}|,t}\})$, weight $P_S = e^{-C\sum_j \mathrm{Tr}(M_j)}\det\left(I + e^{M_1}e^{M_2}...e^{M_n}\right)$ changes to $P_{-S} = \left[e^{-C\sum_j \mathrm{Tr}(M_j)}\det\left(I + e^{M_1}e^{M_2}...e^{M_n}\right)\right]^* = P_S^*$. It is obvious that

$$\frac{\langle F+F^*\rangle}{2} = \sum_{\{S\}}\frac{P_SF_S + P_{-S}F^*_{-S}}{2} = \sum_{\{S\}} Re(P_S)F_S. \tag{18}$$

We conclude here, to measure real measurables for any filling, one just need to see real part of weight as an effective weight.

III. STOCHASTIC ANALYTIC CONTINUATION (SAC) METHOD

Here we use the stochastic analytic continuation (SAC) method [4–8] to obtain the spectral function $A_{\vec{k}}(\omega)$, where $A_{\vec{k}}(\omega) = -(1/\pi)\,\mathrm{Im}\,G^{ret}_{\vec{k}}(\omega)$. The relationship of $A_{\vec{k}}(\omega)$ and correlation function in imaginary time is:

$$G_{\vec{k}}(\tau) = \int_{-\infty}^{\infty} d\omega \left[\frac{e^{-\omega\tau}}{1+e^{-\beta\omega}}\right] A_{\vec{k}}(\omega) \tag{19}$$

The following quantity χ^2 is used to evaluate the quality of the fit, where we first give an very generic variational ansatz of the spectrum $A(\omega)$ and then carry out the above Laplacian transfermation and use the obtained Green's function to compare (fit) with the QMC Green's function and optimize the ansatz accroding to the χ^2 stochastically:

$$\chi^2 = \sum_{ij}\left(\bar{G}(\tau_i) - \int_{-\infty}^{\infty} d\omega\left[\frac{e^{-\omega\tau_i}}{1+e^{-\beta\omega}}\right]A(\omega)\right)\left(C^{-1}\right)_{ij}\left(\bar{G}(\tau_j) - \int_{-\infty}^{\infty} d\omega\left[\frac{e^{-\omega\tau_j}}{1+e^{-\beta\omega}}\right]A(\omega)\right) \tag{20}$$

where

$$C_{ij} = \frac{1}{N_b(N_b-1)}\sum_{b=1}^{N_b}\left(G^b(\tau_i) - \bar{G}(\tau_i)\right)\left(G^b(\tau_j) - \bar{G}(\tau_j)\right) \tag{21}$$

Here $\bar{G}(\tau_i)$ is the Monte Calro average of Green's functions of N_b bins.

In order to obtain the optimized spectral function, we perform the Monte Carlo sampling [6, 7] to improve it with the QMC Green's function. The ansatz of the spectral function is $A(\omega) = \sum_{i=1}^{N_\omega} A_i \delta(\omega - \omega_i)$ and the weight of such Monte Carlo configuration is:

$$W \sim \exp\left(-\frac{\chi^2}{2\,\Theta_T}\right) \tag{22}$$

Here Θ_T is an analogy to temperature. We carry out simulated annealing method and at different Θ_T we compute the average $\langle\chi^2\rangle$, finally choose the converged Θ_T to satisfy:

$$\langle\chi^2\rangle = \chi^2_{\min} + a\sqrt{\chi^2_{\min}} \tag{23}$$

usually we set $a = 2$, and it is from such optimized Θ and χ^2, we further compute the ensemble average of the spectra as the final ones to present in the main text.

[1] B. A. Bernevig, Z.-D. Song, N. Regnault, and B. Lian, Phys. Rev. B **103**, 205413 (2021).
[2] J. Y. Lee, J. Hofmann, E. Khalaf, A. Vishwanath, and E. Berg, APS March Meeting **S43**, 00013 (2021).
[3] J. S. Hofmann, E. Khalaf, A. Vishwanath, E. Berg, and J. Y. Lee, arXiv e-prints , arXiv:2105.12112 (2021), arXiv:2105.12112 [cond-mat.str-el].
[4] A. W. Sandvik, Phys. Rev. B **57**, 10287 (1998).
[5] K. Beach, arXiv preprint cond-mat/0403055 (2004).
[6] A. W. Sandvik, Phys. Rev. E **94**, 063308 (2016).
[7] O. F. Syljuåsen, Phys. Rev. B **78**, 174429 (2008).
[8] H. Shao, Y. Q. Qin, S. Capponi, S. Chesi, Z. Y. Meng, and A. W. Sandvik, Phys. Rev. X **7**, 041072 (2017).