From Theory to Practice: Applying Neural Networks to Simulate Real Systems with Sign Problems

Marcel Rodekamp,[a,b,c,d,*] **Evan Berkowitz,**[a,b,d] **Maria Dincă,**[e] **Christoph Gäntgen,**[b,c,d] **Stefan Krieg**[a,b,c,d] **and Thomas Luu**[b,c,d]

[a] *JARA & Jülich Supercomputing Center, Forschungszentrum Jülich, 54245 Jülich, Germany*

[b] *Institute for Advanced Simulation, Forschungszentrum Jülich, 54245 Jülich, Germany*

[c] *Helmholtz-Institut für Strahlen- und Kernphysik, Rheinische Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany*

[d] *Center for Advanced Simulation and Analytics (CASA), Forschungszentrum Jülich, 52425 Jülich, Germany*

[e] *Department of Physics, Condensed Matter Physics sub-department, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*

E-mail: m.rodekamp@fz-juelich.de, e.berkowitz@fz-juelich.de, maria.dinca0501@gmail.com, c.gaentgen@fz-juelich.de, s.krieg@fz-juelich.de, t.luu@fz-juelich.de

The numerical sign problem poses a seemingly insurmountable barrier to the simulation of many fascinating systems. We apply neural networks to deform the region of integration, mitigating the sign problem of systems with strongly correlated electrons. In this talk we present our latest architectural developments as applied to contour deformation. We also demonstrate its applicability to real systems, namely perylene.

*Speaker
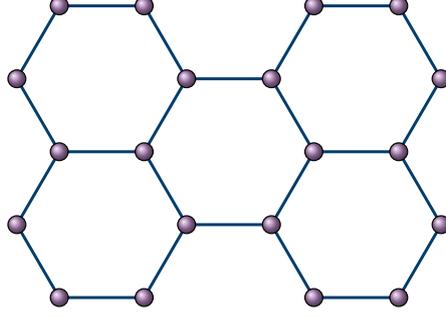
Figure 1: Graphical representation of the perylene molecule. The sites represent carbon-ions, while links indicate allowed hopping. External hydrogen atoms are not drawn.

1. Introduction

The challenge of the computational sign problem poses obstacles for effective importance sampling when dealing with complex-valued distributions in Markov-Chain Monte-Carlo algorithms, like Hybrid Monte Carlo (HMC). This challenge is pervasive when sampling from the configuration space of many physical systems, including but not limited to lattice QCD at finite baryon chemical potential, doped condensed matter systems in equilibrium, and the real-time evolution of quantum systems. A method to substantially alleviate the sign problem involves deforming the original (real) manifold of integration into a sign improved (complex) one [1–4]. Formal developments have spurred exploration into leveraging Lefschetz thimbles [5–11] —these are high-dimensional counterparts to contours of steepest descent and can be located through holomorphic flow. However, pinpointing the location of each relevant thimble's saddle point, or critical point, and determining the relevant sampling 'direction' around these points is prohibitive. An alternative approach is to employ neural networks to learn the mapping from an initial manifold to a beneficial one, including one that approximates the contributing thimbles to the integral [12–15].

In this work, we explore the use of this method for perylene, $C_{20}H_{12}$ [16]. The molecule is (almost) planar and sp^2-hybridized making it a perfect target for the Hubbard model [17]. We show some preliminary results for a chemical potential (doping) scan. Practical applications range from organic semiconductors [18], over organic light emitting diodes (OLEDs) [19], to organic solar cells [20]. Perylene also falls under the class of so-called polycyclic aromatic hydrocarbon (PAH) molecules thought to be ubiquitous in interstellar gases and nebulae [21]. A detailed accounting of electronic properties of PAHs, like perylene, could help constrain interstellar gas models [22].

2. Formalism

Given a fixed spatial arrangement of ions X, the Hubbard model in particle-hole basis reads [23]

$$\mathcal{H}[K,V,\mu] = -\sum_{x,y \in X} \left(p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \right) + \frac{1}{2} \sum_{x,y \in X} \rho_x V^{xy} \rho_y + \mu \sum_{x \in X} \rho_x, \qquad (1)$$

where the amplitudes in K encode the hopping of fermionic particles p and holes h, interactions between these are modelled by the potential V, $\rho_x = p_x^\dagger p_x - h_x^\dagger h_x$ is the net charge per site, and the

chemical potential μ incentivizes charge. We restrict our attention to the case where $K = \kappa\delta_{\langle xy\rangle}$ encodes the structure of perylene, compare figure 1, with nearest-neighbor hopping, and an on-site interaction $V = U\delta_{xy}$. In this case, when $\mu = 0$, the Hamiltonian in (1) corresponds to the neutral, 'half-filled' system.

Trotterizing the thermal trace into N_t timeslices, inserting Grassmannian resolutions of the identity, and linearizing the interaction via the Hubbard-Stratonovich transformation [24] exposes the Hubbard action

$$S\left[\Phi \mid K, V, \mu\right] = \frac{1}{2}\sum_{\substack{x,y\in X \\ t\in[0,N_t-1]}} \Phi_{tx}(\delta V^{-1})^{xy}\Phi_{ty} - \log\det\{M\left[\Phi \mid K, \mu\right]M\left[-\Phi \mid -K, -\mu\right]\}, \tag{2}$$

where $\Phi \in \mathbb{R}^{|\Lambda|}$ is the (auxiliary) hubbard field on the spacetime lattice $\Lambda = [0, N_t - 1] \otimes X$ and $\delta = \beta/N_t$. We use the exponential discretization [25] for the fermion matrices

$$M\left[\Phi \mid K, \mu\right]_{x't';xt} = \delta_{x'x}\delta_{t't} - \left(e^{\delta(K+\mu)}\right)_{x'x}e^{+i\Phi_{xt}}\delta_{t'(t+1)} \tag{3}$$

with antiperiodic boundary conditions in time. With this we can express the thermal trace as path integral

$$\langle O\rangle = \frac{1}{Z}\text{Tr}\left\{Oe^{-\beta\mathcal{H}}\right\} = \frac{1}{Z}\int \mathcal{D}\left[\Phi\right]e^{-S[\Phi]}O\left[\Phi\right]. \tag{4}$$

The partition function Z is the trace/integral without the observable. On a bipartite lattice we may replace the $-K$ in the holes' fermion matrix with $+K$ leading to a real and positive determinant at vanishing chemical potential. When μ is finite the determinant is complex valued and results in a sign problem.

3. Method

As in our previous work [14, 15, 26], for a Monte-Carlo algorithm with complex action we can separate the real and imaginary parts, $S = \text{Re}\{S\} + i\,\text{Im}\{S\}$, and rewrite the path integral as

$$Z = \int \mathcal{D}\left[\Phi\right]e^{-S} = \int \mathcal{D}\left[\Phi\right]e^{-\text{Re}\{S\}}e^{-i\,\text{Im}\{S\}} \propto \left\langle e^{-i\,\text{Im}\{S\}}\right\rangle_{\text{Re}\{S\}} \equiv \Sigma \tag{5}$$

where $|\Sigma| \in [0, 1]$ is the statistical power. Sampling according to the real part of the action and then applying reweighting,

$$\langle O\rangle = \frac{\left\langle e^{-i\,\text{Im}\{S\}}O\right\rangle_{\text{Re}\{S\}}}{\left\langle e^{-i\,\text{Im}\{S\}}\right\rangle_{\text{Re}\{S\}}} = \frac{1}{\Sigma}\left\langle e^{-i\,\text{Im}\{S\}}O\right\rangle_{\text{Re}\{S\}}, \tag{6}$$

allows the computation of observables. This procedure fails if the statistical power $|\Sigma|$ cannot be reliably distinguished from zero [6, 14, 27, 28].

Combining reweighting with path deformation techniques offers a promising approach to mitigating the sign problem. The primary objective is to expand the accessible parameter space for computation, ultimately allowing extrapolations to the continuum limit, $\delta \to 0$, and/or the zero-temperature limit, $\beta \to \infty$. The core idea involves manipulating the integration contour in

a way that enhances the statistical power. One noteworthy method in this context is the use of Lefschetz thimbles, which eliminate fluctuations in the imaginary part of the action. However, it is essential to acknowledge that these thimbles are notoriously challenging to compute in practical simulations. As a result, the focus shifts towards harnessing a more versatile transformation to a sign-optimized manifold, denoted as $\tilde{\mathcal{M}}$. Crucially, Cauchy's theorem assures us that expectation values of holomorphic observables remains unchanged under this transformation[1]:

$$
\begin{aligned}
\langle O \rangle &= \frac{1}{Z} \int_{\tilde{\mathcal{M}}} \mathcal{D}\left[\tilde{\Phi}\right] e^{-S\left[\tilde{\Phi}\right]} O\left[\tilde{\Phi}\right] \\
&= \frac{1}{Z} \int \mathcal{D}\left[\Phi\right] e^{-S\left[\tilde{\Phi}(\Phi)\right] + \log \det\left\{J_{\tilde{\Phi}}[\Phi]\right\}} O\left[\tilde{\Phi}(\Phi)\right]
\end{aligned}
\tag{7}
$$

A simple transformation is the mapping to the tangent plane $\tilde{\Phi}[\Phi] = \Phi + \mathrm{i}\,\phi_0$. As outlined in [26] a transcendental equation

$$
\phi_0/\delta = -\frac{U}{N_x} \sum_k \tanh\left(\frac{\beta}{2}\left[\epsilon_k + \mu + \phi_0/\delta\right]\right)
\tag{8}
$$

can be used to identify this plane. The sum is over the eigenvalues ϵ_k of the hopping matrix K. This adds no additional computational cost to the simulation and thus forms our starting point for further transformations.

Based on our earlier machine learning approach [15], we can use a neural network to map to a sign optimized manifold $\tilde{\mathcal{M}}$, i.e. $\tilde{\Phi}[\Phi] = \mathrm{NN}[\Phi + i\phi_0]$. We continue to use coupling networks to enable a tractable Jacobian determinant. However, we changed the activation function to be holomorphic compared to our earlier approach which simplified the training procedure significantly. A single layer thus reads

$$
\mathrm{NN}_l\left[\Psi\right]_{t,x} = \begin{cases} e^{s_l(\Psi_B)} \odot \Psi_A + t_l(\Psi_B) & (t,x) \in A \\ \Psi_B & (t,x) \in B \end{cases}
\tag{9}
$$

where $s, t : \mathbb{C}^{|\Lambda|/2} \to \mathbb{C}^{|\Lambda|/2}$, $\Psi_B \mapsto w' \cdot P\left[w \cdot \Psi_B + b\right] + b'$ are linear networks; the activation $P[\cdot]$ is a polynomial of degree 3. We always pair two of these layers to form a single transformation that allows us to change every value of the input configuration. Notice, this network is unbounded at $\Psi \to \pm\infty$ and thus escapes Liouville's theorem [29, 30]. Furthermore, the homology class is completely determined by the asymptotic behavior of the network. This means that any parametrization of the contour $\tilde{\mathcal{M}}$ preserves the homology class if it approaches a constant asymptotically. To achieve this, we implement a projection layer

$$
\mathcal{P}_\sigma\left[\Psi\right]_i = \mathrm{Re}\{\Psi\}_i + \mathrm{i}\,\mathrm{Im}\{\Psi\}_i e^{-\frac{|\Psi_i|^2}{\sigma^2}}
\tag{10}
$$

which pushes the imaginary part down towards the real manifold. The parameter σ can be chosen arbitrarily. Choosing $\sigma \gg 1$ but finite, the projection virtually becomes the identity up to $O(\sigma^{-2})$ and the Jacobian remains unchanged $J_{\tilde{\Phi}}[\Phi] = \det\{\mathrm{NN}[\Phi]\} + O(\sigma^{-2})$. Care has to be taken if the

---

[1]Strictly speaking, this holds when no singularities are crossed when performing the contour deformation, which is the case in our problem.
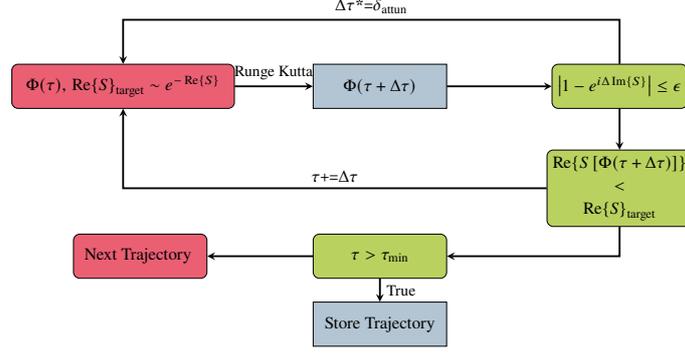
Figure 2: Algorithm to generate training data. Each step integrates the flow equation for a time $\Delta\tau$ and checks that the imaginary part of the action is preserved for the step size. The step size becomes updated accordingly. Further, only data that stems to be relevant for training, i.e. is not flowed to little and is not in neverland, is accepted.

configuration value NN $[\Phi]$ becomes comparable to σ. Naturally, such configuration values have small weights; thus, they practically never appear.

We train the parameters – $w, w' \in \mathrm{Mat}_{|\Lambda|/2}[\mathbb{C}]$, $b, b' \in \mathbb{C}^{|\Lambda|/2}$ per layer – of the network using training data sets of flowed configurations [14]. The training data is generated by integrating the holomorphic flow equation, starting from configurations sampled with a normal-distribution $\mathcal{N}(\mu = 0, \sigma = \delta U)$ and mapped to the tangent plane, with a Runge Kutta and accepting only configurations that preserve the imaginary part of the action, are flowed enough, and are not in neverland [2]. A graphical representation of this can be found in figure 2. Once the data is gathered we initialize the network with random parameters $\sim U(0.01, 0.01)$ [31], and train it using the Adam optimizer [32] with a learning rate of 10^{-3}. For the first iterations we use a plain L_2-loss to measure the distance between the training data and the network output. After that a couple of iterations are trained by comparing the real part of the action plus the preservation of the imaginary part, $\mathrm{L}\left(\tilde{\Phi}, \mathrm{NN}\left[\Phi + i\,\phi_0\right]\right) = |\Delta\,\mathrm{Re}\{S\}| + \left|1 - e^{i\,\Delta\,\mathrm{Im}\{S\}}\right|$. Once a desired precision $(< 1e - 4)$ is reached, we perform a short HMC simulation, $\mathrm{N}_{\mathrm{cfg}} = 1000$, and measure the statistical power $|\Sigma|$. If the statistical power is increased, we accept the new network parameters and continue the HMC simulation resulting in the observables presented in the next section.

4. Results

4.1 Correlation Functions

The basic building block of Monte-Carlo simulations for physical systems is the correlation function

$$C^{x,y}(\tau) = \left\langle p_x(\tau)p_y^\dagger(0)\right\rangle = \left\langle M^{-1}_{(\tau,x);(0,y)}\left[\Phi \mid K, \mu\right]\right\rangle, \qquad (11)$$

as it encodes the physics of a single excitation-annihilation process. In the following we block diagonalize the correlation matrix using the symmetry eigenspaces of the ion lattice X

$$C^k(\tau) = \delta^{k,k'}\bar{\mathfrak{u}}_{k,x}C^{x,y}(\tau)\mathfrak{u}_{y,k'}, \qquad (12)$$

---

[2]By neverland we denote regions of infinite $\mathrm{Re}\{S\}$. Points where Lefshetz thimbles meet have diverging $\mathrm{Re}\{S\}$ and are not relevant for the training.

where the unitary matrix $\mathfrak{u}$ is determined by the irreducible representations of the point group D_{2h} of X. In principle, a further diagonalization within the irreducible blocks is required, however, we found that the diagonal terms are dominant in the cases we have studied thus far and so only take the diagonal correlators for study in these proceedings. To classify the molecule and understand the entirety of its (single particle) spectrum a full parameter scan in the interaction strength and its doping needs to be performed as well as the continuum and zero temperature limit ($\beta, N_t \rightarrow \infty$). In this preliminary results we will focus on the application of the neural network to the sign problem. Hence, we fix $N_t = 32$ and $\beta = 4$, while the interaction and chemical potential strength are indicated at the specific results. In Figures 3 we compare correlator measurements from HMC on the tangent plane, on the left, and from machine learning enhanced HMC (MLHMC), on the right. These are computed at fixed $U = 2$ and $\mu = 1.7$. As expected the MLHMC resolves the correlator much better than the HMC on the tangent plane. This chemical potential is the first where the computational extra cost of the neural network can be justified.
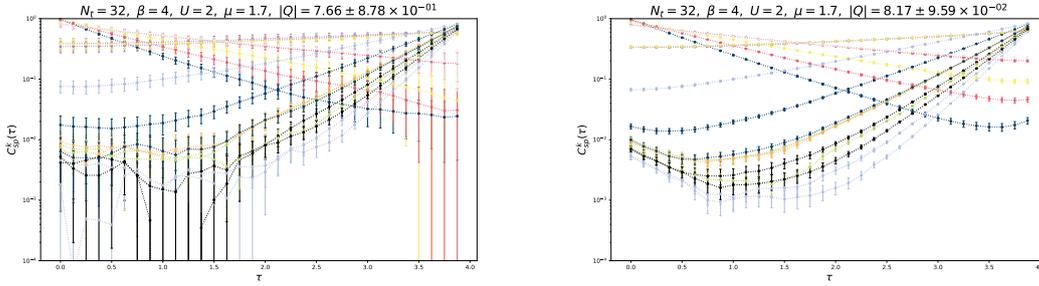
Figure 3: Perylene correlator for $U = 2$, $\mu = 1.7$ simulated using HMC on the tangent plane (top) and machine learning guided HMC (bottom)

4.2 Charge

The chemical potential is an effective description of the doping of the system. We translate this by measuring the total system charge

$$Q = \sum_{x=0}^{N_x} \langle \rho_x \rangle = \sum_{x=0}^{N_x} C^{x,x}(\tau = 0). \tag{13}$$

In Figure 4 we present the total charge as a function of the chemical potential at fixed $U = 2$. The interaction strength is the same as for the correlators in Figure 3. We can identify, that the statistical error at $\mu = 1.7$ is much smaller for the MLHMC. At cold temperatures, $\beta \rightarrow \infty$, we expect a stepwise behaviour for integer charge changes at certain chemical potentials. This is washed out here due to the small value of β. To fully classify the doping of the molecule a zero-tempreature limit must be taken which will increase the use of the neural network as the sign problem exponentially becomes worse with increasing β.

4.3 Effective Masses

To study the molecule for various applications like energy production in organic solar cells it is required to get an understanding of the low lying energy spectrum of a particle-hole excitation. At
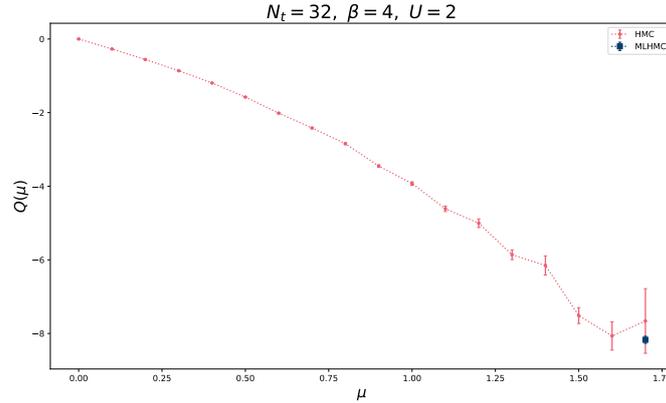
Figure 4: Perylene total charge μ-scan at $U = 2$

non-interacting systems $U = 0$, this can be done analytically as the single-particle energy is given by the eigenvalues of the hopping matrix $E_{sp}^k = \epsilon^k$. Doping simply shifts these $E_{sp/sh}^k = \epsilon^k \pm \mu$. Further, if particles and holes do not interact the exciton energy is simply the sum $E^k = E_{sp}^k + E_{sh}^k$, which, noticeably, is independent of doping. As soon as interactions are turned on this is not true anymore and the exciton energy needs to be measured explicitly involving computations of 4 point functions. For this proceedings we take a first step into that direction and measure the single particle energy at various interaction strengths and chemical potentials. In the low-temperature spectral decomposition $C^k(\tau) = A_0 e^{-E_0^k \tau} + O\left(e^{-E_1^k \tau}\right)$ one expects ground state contributions in the large τ limit, defining the effective mass

$$m_{\text{eff}}^k\big|_\tau = \frac{\log C^k(\tau + \delta) - \log C^k(\tau)}{\delta}. \tag{14}$$

The effective mass serves as a simple estimator for the energy $E_{sp}^k \approx m_{\text{eff}}^k\big|_{\tau=\beta/2}$ that neglects excited states. In figures 5a to 5d we present a selection of effective mass plots for two interaction strengths $U = 2, 4$ (5a,5b) at zero chemical potential and two chemical potentials $\mu = 0.4, 1.7$ at fixed $U = 2$ (5c,5d). These plots suggest that the interaction strength has a great effect on the excited states, as the plateau becomes less pronounced, but only little on the ground state energy, i.e. the smallest positive energy value. Further, we can identify that the gap between states opens as the interaction strength is increased. Additionally, the change of chemical potential clearly shifts the ground state gap, which, together indicate a rich energy spectrum of the molecule.

5. Summary

The numerical sign problem affects Monte-Carlo simulations of various systems. In this contribution we modelled the electronic structure of the sp^2-hybridized molecule $C_{20}H_{12}$ (perylene) with a contact-interaction Hubbard model which is then simulated using the Hybird/Hamilton Monte-Carlo algorithm. At finite chemical potential, a parameter to control doping, a sign problem appears which we tackle using reweighting and contour deformations. Starting with simple deformations, the tangent plane, and continuing with neural networks at larger μ we are able to resolve the total

(a) Perylene effective mass $U = 2$, $\mu = 0$ (HMC)

(b) Perylene effective mass $U = 4$, $\mu = 0$ (HMC)

(c) Perylene effective mass $U = 2$, $\mu = 0.4$ (HMC)

(d) Perylene effective mass $U = 2$, $\mu = 1.7$ (MLHMC)
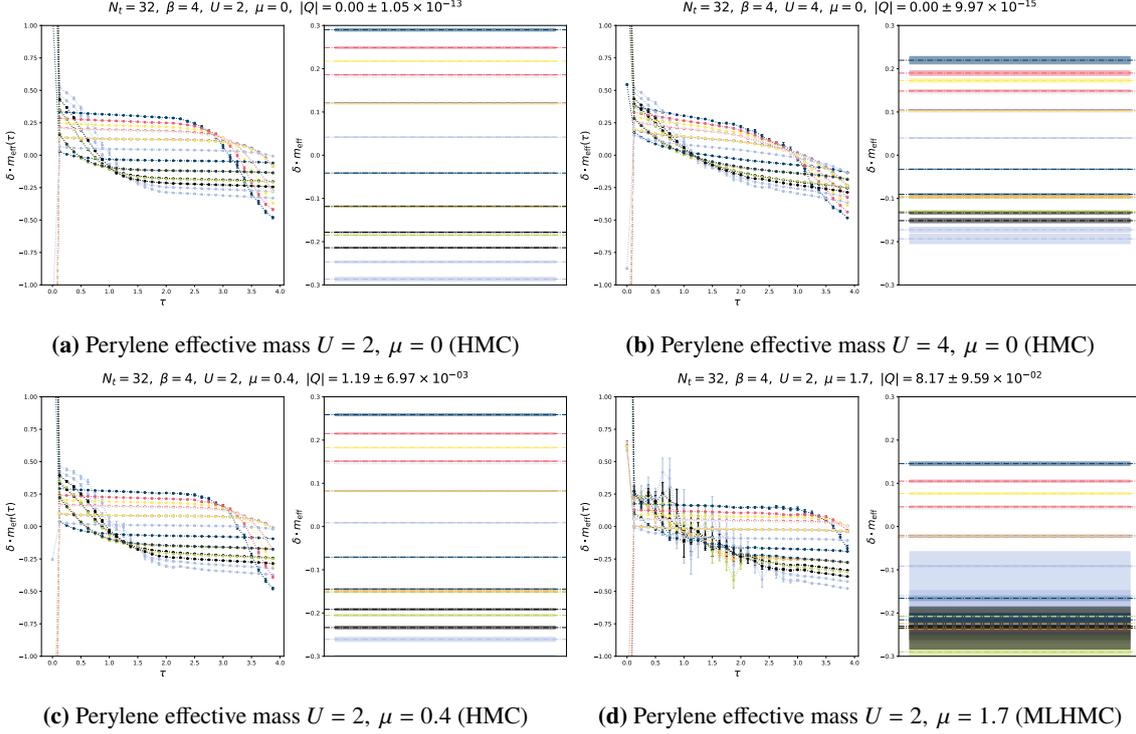
system charge and single particle energy spectrum of perylene. To limit thermal contamination of the energy states, we require that $\beta E_0 \gg 1$, where E_0 is the lowest single particle excitation energy in the system. In our case we have $\beta E_0 \sim 1.5$, which should be increased in future calculations. We plan to extend the parameter space and perform a more detailed analysis of the system, including the study of particle-hole excitations which are relevant for practical application of the molecule.

Acknowledgments

References

[1] Kouji Kashiwa, Yuto Mori, and Akira Ohnishi. Controlling the model sign problem via the path optimization method: Monte Carlo approach to a QCD effective model with Polyakov loop. *Phys. Rev. D*, 99(1):014033, 2019.

[2] Andrei Alexandru, Gokce Basar, Paulo F Bedaque, and Neill C Warrington. Complex paths around the sign problem. *arXiv preprint arXiv:2007.05436*, 2020.

[3] William Detmold, Gurtej Kanwar, Michael L. Wagman, and Neill C. Warrington. Path integral contour deformations for noisy observables. *Phys. Rev. D*, 102(1):014514, 2020.

[4] William Detmold, Gurtej Kanwar, Henry Lamm, Michael L. Wagman, and Neill C. Warrington. Path integral contour deformations for observables in $SU(N)$ gauge theory. *Phys. Rev. D*, 103(9):094517, 2021.

[5] Solomon Lefschetz. On certain numerical invariants of algebraic varieties with application to abelian varieties. *Transactions of the American Mathematical Society*, 22(3):327–406, 1921.

[6] Andrei Alexandru, Gökçe Başar, and Paulo Bedaque. Monte carlo algorithm for simulating fermions on lefschetz thimbles. *Physical Review D*, 93(1):014504, 2016.

[7] M. Cristoforetti, F. Di Renzo, G. Eruzzi, A. Mukherjee, C. Schmidt, L. Scorzato, and C. Torrero. An efficient method to compute the residual phase on a Lefschetz thimble. *Phys. Rev.*, D89(11):114505, 2014.

[8] Marco Cristoforetti, Francesco Di Renzo, Abhishek Mukherjee, and Luigi Scorzato. Monte Carlo simulations on the Lefschetz thimble: Taming the sign problem. *Phys. Rev.*, D88(5):051501, 2013.

[9] Abhishek Mukherjee, Marco Cristoforetti, and Luigi Scorzato. Metropolis Monte Carlo integration on the Lefschetz thimble: Application to a one-plaquette model. *Phys. Rev.*, D88(5):051502, 2013.

[10] Takuya Kanazawa and Yuya Tanizaki. Structure of Lefschetz thimbles in simple fermionic systems. *JHEP*, 03:044, 2015.

[11] Yuya Tanizaki, Yoshimasa Hidaka, and Tomoya Hayata. Lefschetz-thimble analysis of the sign problem in one-site fermion model. *New Journal of Physics*, 18(3):033002, mar 2016.

[12] Andrei Alexandru, Paulo F Bedaque, Henry Lamm, and Scott Lawrence. Deep learning beyond lefschetz thimbles. *Physical Review D*, 96(9):094505, 2017.

[13] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. Application of a neural network to the sign problem via the path optimization method. *PTEP*, 2018(2):023B04, 2018.

[14] Jan-Lukas Wynen, Evan Berkowitz, Stefan Krieg, Thomas Luu, and Johann Ostmeyer. Machine learning to alleviate hubbard-model sign problems. *Physical Review B*, 103(12):125153, 2021.

[15] Marcel Rodekamp, Evan Berkowitz, Christoph Gäntgen, Stefan Krieg, Thomas Luu, and Johann Ostmeyer. Mitigating the hubbard sign problem with complex-valued neural networks. *Phys. Rev. B*, 106:125139, Sep 2022.

[16] Mark Botoshansky, Frank H. Herbstein, and Moshe Kapon. Towards a Complete Description of a Polymorphic Crystal: The Example of Perylene. *Helvetica Chimica Acta*, 86(4):1113–1128, April 2003.

[17] A. Camerman and J Trotter. The crystal and molecular structure of perylene. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 220(1142):311–321, December 1953.

[18] Ananth Dodabalapur, Howard E Katz, and Luisa Torsi. Molecular orbital energy level engineering in organic transistors. *Advanced Materials*, 8(10):853–855, 1996.

[19] Yoshiharu Sato, Shoko Ichinosawa, and Hiroyuki Kanai. Operation characteristics and degradation of organic electroluminescent devices. *IEEE Journal of Selected Topics in Quantum Electronics*, 4(1):40–48, 1998.

[20] Ching W Tang. Two-layer organic photovoltaic cell. *Applied physics letters*, 48(2):183–185, 1986.

[21] Farid Salama. Pahs in astronomy - a review. *Proceedings of the International Astronomical Union*, 4(S251):357–366, 2008.

[22] Aigen Li. Spitzer's perspective of polycyclic aromatic hydrocarbons in galaxies. *Nature Astronomy*, 4(4):339–351, 2020.

[23] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 276(1365):238–257, 1963.

[24] J. Hubbard. Calculation of partition functions. *Phys. Rev. Lett.*, 3(2):77–78, 1959.

[25] Jan-Lukas Wynen, Evan Berkowitz, Christopher Körber, Timo A Lähde, and Thomas Luu. Avoiding ergodicity problems in lattice discretizations of the hubbard model. *Physical Review B*, 100(7):075141, 2019.

[26] Christoph Gäntgen, Evan Berkowitz, Thomas Luu, Johann Ostmeyer, and Marcel Rodekamp. Fermionic sign problem minimization by constant path integral contour shifts. *arXiv preprint arXiv:2307.06785*, 2023.

[27] Casey E Berger, Lukas Rammelmüller, Andrew C Loheac, Florian Ehmann, Jens Braun, and Joaquın E Drut. Complex langevin and other approaches to the sign problem in quantum many-body physics. *Physics Reports*, 892:1–54, 2021.

[28] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. Lefschetz thimbles in fermionic effective models with repulsive vector-field. *Physics Letters B*, 781:688–693, 2018.

[29] Akira Hirose. *Complex-valued neural networks*, volume 400. Springer Science & Business Media, 2012.

[30] ChiYan Lee, Hideyuki Hasegawa, and Shangce Gao. Complex-valued neural networks: A comprehensive survey. *IEEE/CAA Journal of Automatica Sinica*, 9(8):1406–1426, 2022.

[31] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In Yee Whye Teh and Mike Titterington, editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 249–256, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.

[32] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.