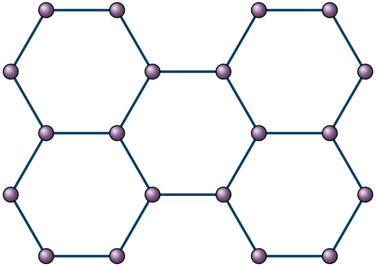


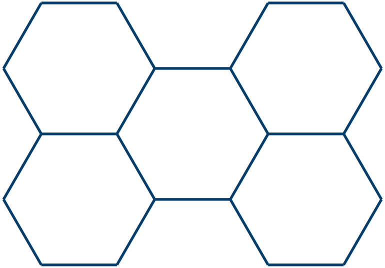
FROM THEORY TO PRACTICE:
Applying Networks to Simulate Real Systems with Sign Problem

August 3, 2023 | Marcel Rodekamp | Jülich Supercomputing Center, Forschungszentrum Jülich

2023 LATTICE

JÜLICH Forschungszentrum

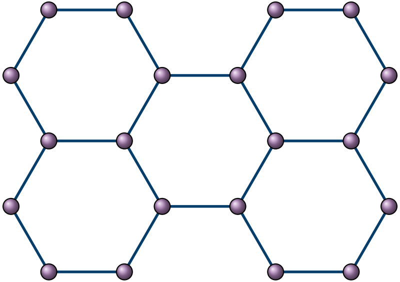
[Ostmeyer et al., 2020]



[Wynen et al., 2019]

$$\mathcal{H} = -\frac{1}{2}\sum_{x,y,\sigma}\left\{\hat{a}^{\dagger}_{x,\sigma}\kappa_{xy}\hat{a}_{y,\sigma} + \hat{a}^{\dagger}_{y,\sigma}\kappa_{xy}\hat{a}_{x,\sigma}\right\} \quad (1)$$

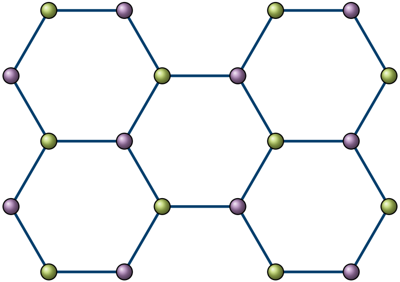
[Ostmeyer et al., 2020]



[Wynen et al., 2019]

$$\mathcal{H} = -\frac{1}{2}\sum_{x,y,\sigma}\left\{\hat{a}^\dagger_{x,\sigma}\kappa_{xy}\hat{a}_{y,\sigma} + \hat{a}^\dagger_{y,\sigma}\kappa_{xy}\hat{a}_{x,\sigma}\right\} + \frac{U}{2}\sum_{x}\left\{\hat{\eta}_{x,\uparrow} - \hat{\eta}_{x,\downarrow}\right\}^2 \tag{1}$$

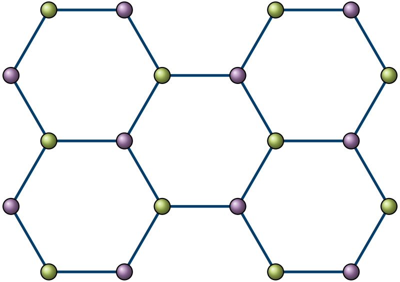
[Ostmeyer et al., 2020]



[Wynen et al., 2019]

$$\mathcal{H} = -\frac{1}{2}\sum_{x,y,\sigma}\left\{\hat{a}^\dagger_{x,\sigma}\kappa_{xy}\hat{a}_{y,\sigma} + \hat{a}^\dagger_{y,\sigma}\kappa_{xy}\hat{a}_{x,\sigma}\right\} + \frac{U}{2}\sum_{x}\{\hat{\eta}_{x,\uparrow} - \hat{\eta}_{x,\downarrow}\}^2 + \mu\sum_{x}\{\hat{\eta}_{x,\uparrow} - \hat{\eta}_{x,\downarrow}\} \qquad (1)$$

[Ostmeyer et al., 2020]

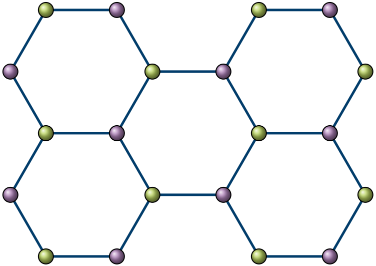


[Wynen et al., 2019]

$$S[\Phi] = \frac{1}{\delta U} \sum_{t,x} \Phi_{t,x}^2 + \log\det\left\{M[\Phi \mid \kappa, \mu] \cdot M[-\Phi \mid -\kappa, -\mu]\right\} \tag{2}$$

Perylene: $C_{20}H_{12}$ [Botoshansky et al., 2003]

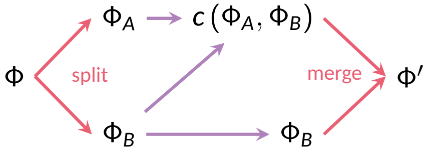
- (ML) HMC Simulation [Rodekamp et al., 2022, Gäntgen et al., 2023]:
  $U \in \{1, 1.5, \cdots, 5\}$, $\mu \in \{0, 0.1, \ldots, 1.7\}$, $Nt = 32$, $\beta = 4$
  - Single particle correlation functions
  - Total system charge
  - Single particle spectrum

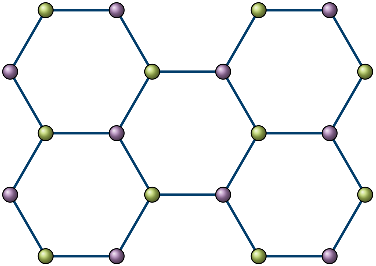


Perylene: $\mathrm{C_{20}H_{12}}$ [Botoshansky et al., 2003]

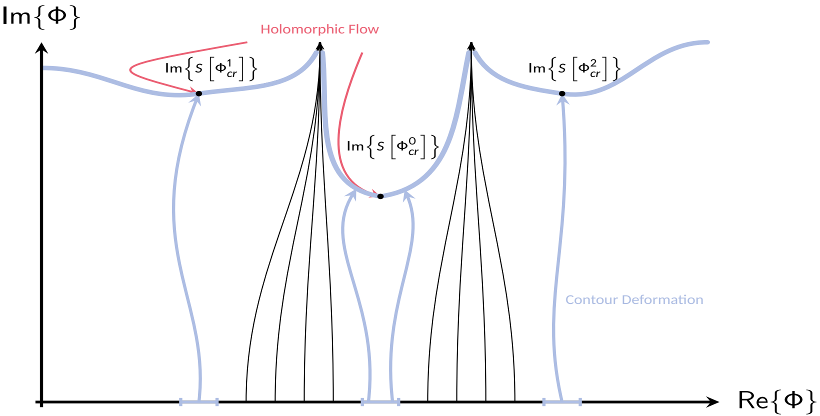
- (ML) HMC Simulation [Rodekamp et al., 2022, Gäntgen et al., 2023]:
  $U \in \{1, 1.5, \cdots, 5\}$, $\mu \in \{0, 0.1, \ldots, 1.7\}$, $Nt = 32$, $\beta = 4$
  - Single particle correlation functions
  - Total system charge
  - Single particle spectrum
- Contour deformation to mitigate sign problem [Alexandru et al., 2016]

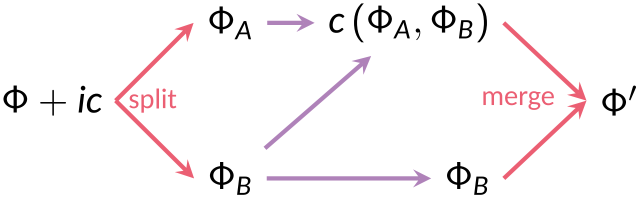
$$\langle \mathcal{O} \rangle = \frac{\left\langle \mathcal{O} e^{-\mathrm{i}\,\mathrm{Im}\{S\}} \right\rangle}{\left\langle e^{-\mathrm{i}\,\mathrm{Im}\{S\}} \right\rangle} \tag{3}$$





$$\frac{\mathrm{d}\Phi_i(\tau_f)}{\mathrm{d}\tau_f} = \pm \left[\frac{\partial S \left[\Phi(\tau_f) \right]}{\partial \Phi_i(\tau_f)} \right]^* \tag{4}$$

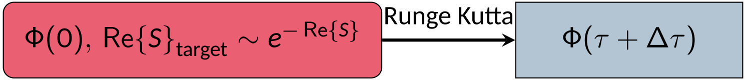


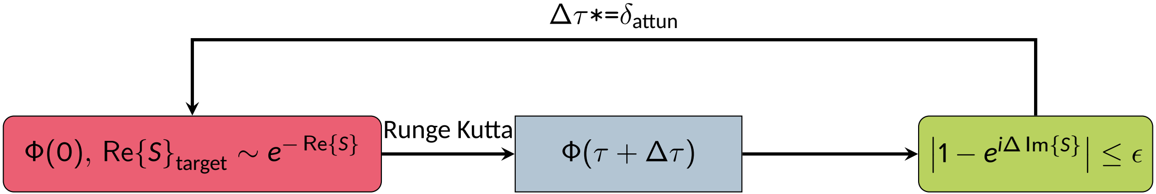


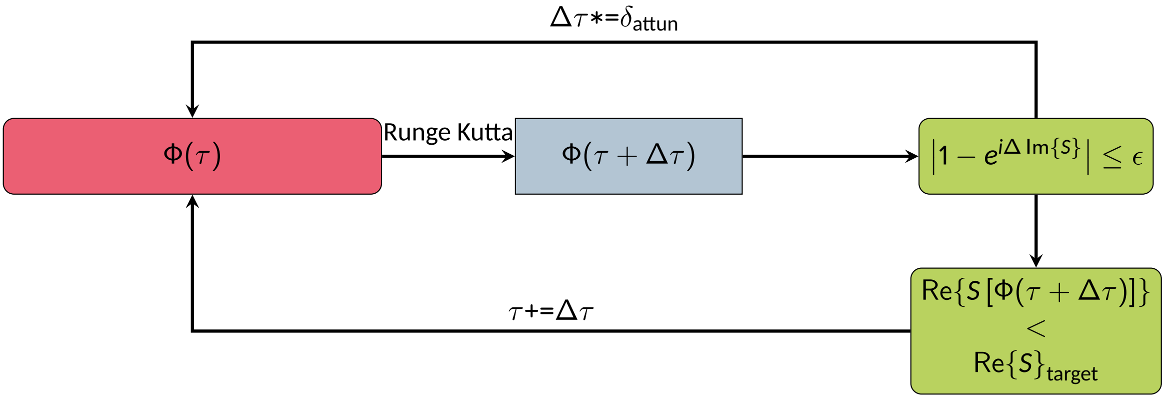
$$c(\Phi_A, \Phi_B) = e^{s(\Phi_B)} \odot \Phi_A + t(\Phi_B) \tag{5}$$

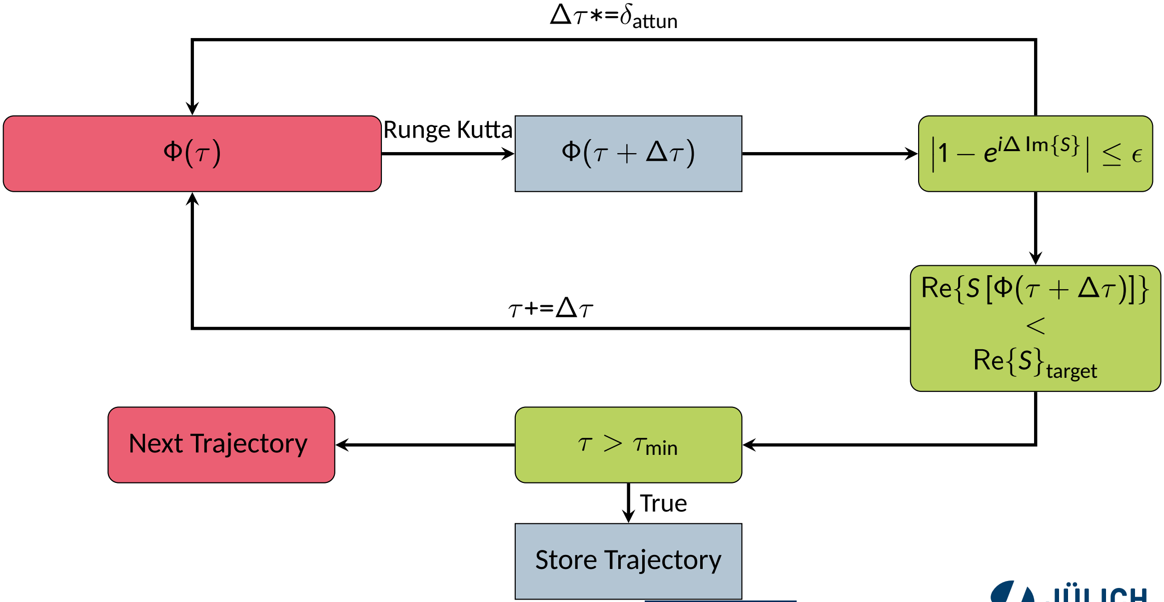
$$s, t : \mathbb{C}^{\frac{N_t}{2} N_x} \to \mathbb{C}^{\frac{N_t}{2} N_x}, \; \varphi \mapsto w' \cdot P\left[w \cdot \varphi + b\right] + b' \tag{6}$$

$$P[z] = z^3 + z^2 + z \tag{7}$$


$\Phi(0)$, $\mathrm{Re}\{S\}_{\mathrm{target}} \sim e^{-\mathrm{Re}\{S\}}$
Runge Kutta
$\Phi(\tau + \Delta\tau)$


$\Delta\tau{*}{=}\delta_{\text{attun}}$
$\Phi(0),\ \mathrm{Re}\{S\}_{\text{target}} \sim e^{-\mathrm{Re}\{S\}}$
Runge Kutta
$\Phi(\tau+\Delta\tau)$
$\left|1-e^{i\Delta\,\mathrm{Im}\{S\}}\right| \leq \epsilon$


$\Delta\tau$*=δ_{attun}
$\Phi(\tau)$
Runge Kutta
$\Phi(\tau + \Delta\tau)$
$\left|1 - e^{i\Delta\,\text{Im}\{S\}}\right| \leq \epsilon$
$\text{Re}\{S\,[\Phi(\tau + \Delta\tau)]\}$
$<$
$\text{Re}\{S\}_{\text{target}}$
τ+=$\Delta\tau$



$\Delta\tau$*=δ_{attun}
Runge Kutta
$\Phi(\tau)$
$\Phi(\tau + \Delta\tau)$
$\left|1 - e^{i\Delta\,\text{Im}\{S\}}\right| \leq \epsilon$
τ+=$\Delta\tau$
$\text{Re}\{S\,[\Phi(\tau+\Delta\tau)]\} < \text{Re}\{S\}_{\text{target}}$
Next Trajectory
$\tau > \tau_{\text{min}}$
True
Store Trajectory

1 Initialize weights $w, b, w', b' \sim U(-0.01,\ 0.01)$ [Glorot and Bengio, 2010]

1. Initialize weights $w, b, w', b' \sim U(-0.01,\ 0.01)$ [Glorot and Bengio, 2010]
2. Train with $L_2\left(\Phi(\tau),\ \mathrm{NN}\left[\Phi(0)\right]\right)$

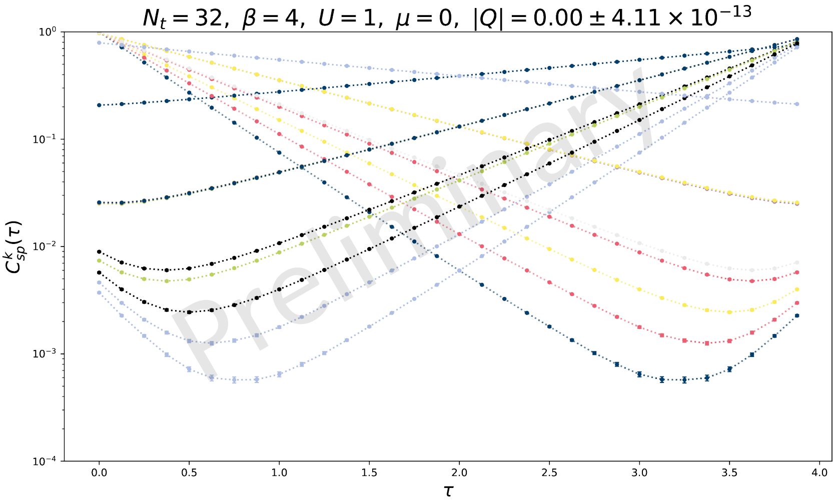
1. Initialize weights $w, b, w', b' \sim U(-0.01,\ 0.01)$ [Glorot and Bengio, 2010]
2. Train with $L_2\left(\Phi(\tau),\ \mathrm{NN}\left[\Phi(0)\right]\right)$
3. Train with $\left|\Delta \,\mathrm{Re}\{S\}\right| + \left|1 - e^{\mathrm{i}\Delta \,\mathrm{Im}\{S\}}\right|$

1. Initialize weights $w, b, w', b' \sim U(-0.01,\ 0.01)$ [Glorot and Bengio, 2010]
2. Train with $L_2\left(\Phi(\tau),\ \mathrm{NN}\left[\Phi(0)\right]\right)$
3. Train with $\left|\Delta\,\mathrm{Re}\{S\}\right| + \left|1 - e^{\mathrm{i}\Delta\,\mathrm{Im}\{S\}}\right|$

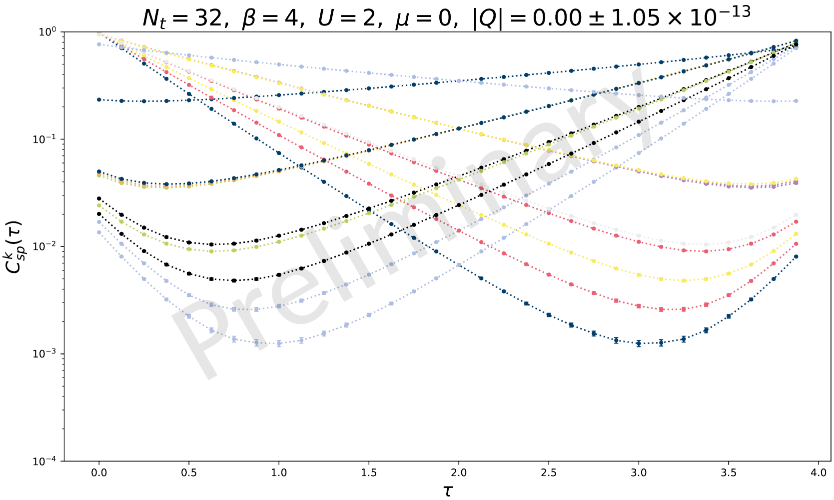
1. Initialize weights $w, b, w', b' \sim U(-0.01,\ 0.01)$ [Glorot and Bengio, 2010]
2. Train with $L_2\left(\Phi(\tau),\ \mathrm{NN}\left[\Phi(0)\right]\right)$
3. Train with $\left|\Delta\,\mathrm{Re}\{S\}\right| + \left|1 - e^{\mathrm{i}\Delta\,\mathrm{Im}\{S\}}\right|$
4. ML HMC ($N_{\mathrm{conf}} = 1000$) measure $\left|\left\langle e^{-\mathrm{i}\,\mathrm{Im}\{S\}}\right\rangle\right|$

$$C^k_{sp}(\tau) = \left\langle a^\dagger_{k\uparrow}(\tau) a_{k\uparrow}(0) \right\rangle \tag{8}$$

HMC

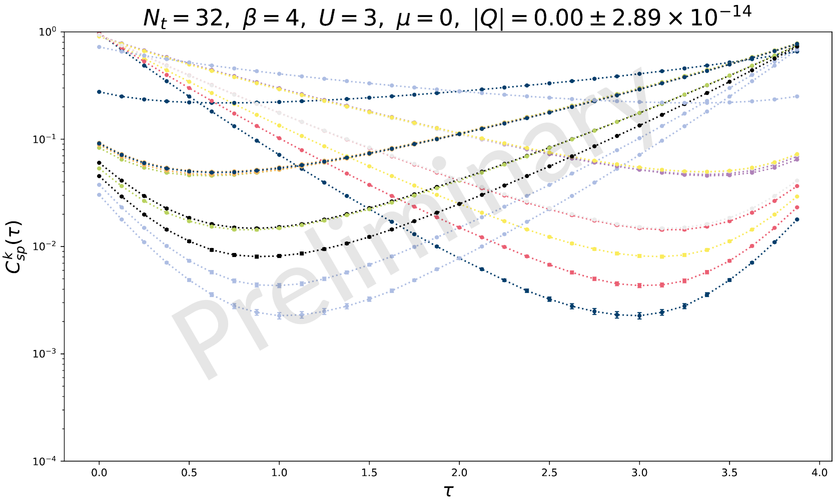

$N_t = 32,\ \beta = 4,\ U = 1,\ \mu = 0,\ |Q| = 0.00 \pm 4.11 \times 10^{-13}$
$C^k_{sp}(\tau)$
τ
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Preliminary

HMC



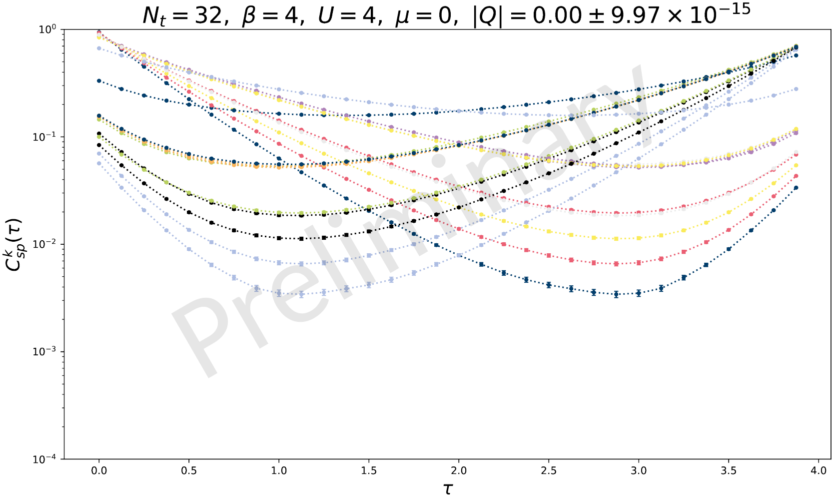
$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 0,\ |Q| = 0.00 \pm 1.05 \times 10^{-13}$
$C^k_{sp}(\tau)$
τ
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Preliminary

HMC



$N_t = 32,\ \beta = 4,\ U = 3,\ \mu = 0,\ |Q| = 0.00 \pm 2.89 \times 10^{-14}$
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$C^k_{sp}(\tau)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
Preliminary

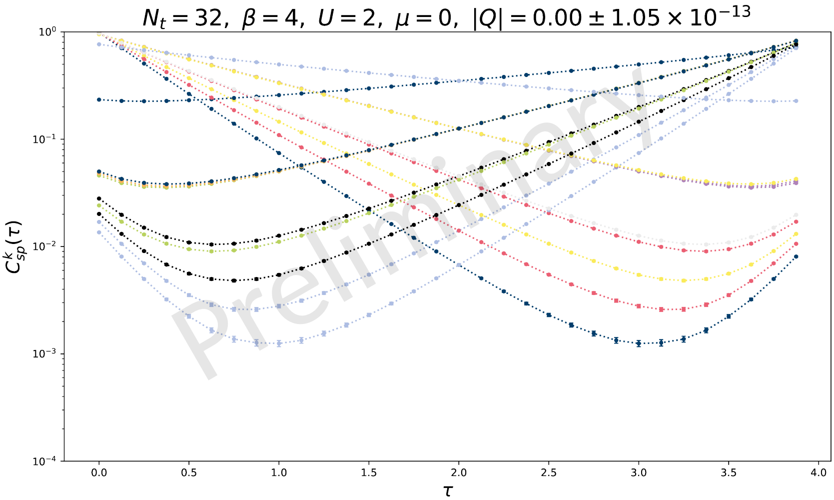
HMC



$N_t = 32,\ \beta = 4,\ U = 4,\ \mu = 0,\ |Q| = 0.00 \pm 9.97 \times 10^{-15}$
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$C^k_{sp}(\tau)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
Preliminary

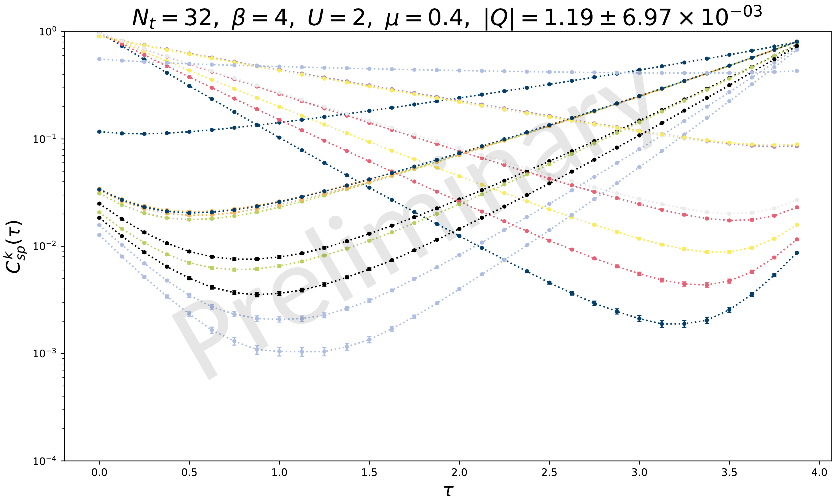
$$\mu \neq 0$$

HMC



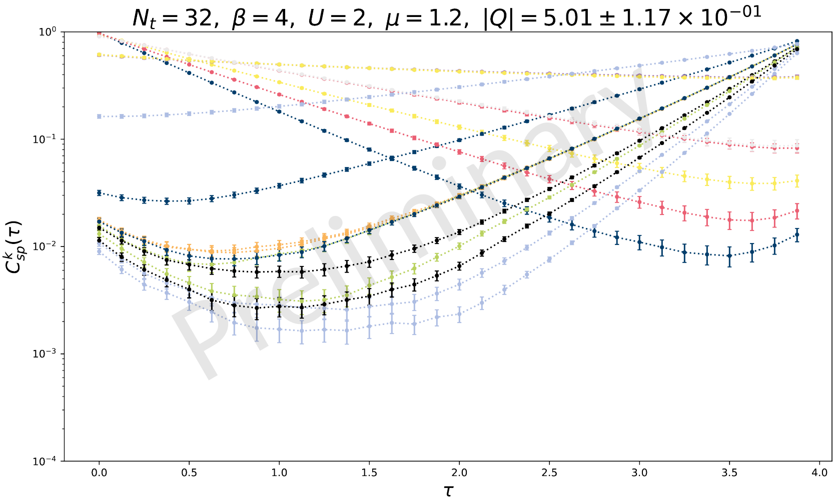
$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 0,\ |Q| = 0.00 \pm 1.05 \times 10^{-13}$
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$C^k_{sp}(\tau)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
Preliminary

HMC



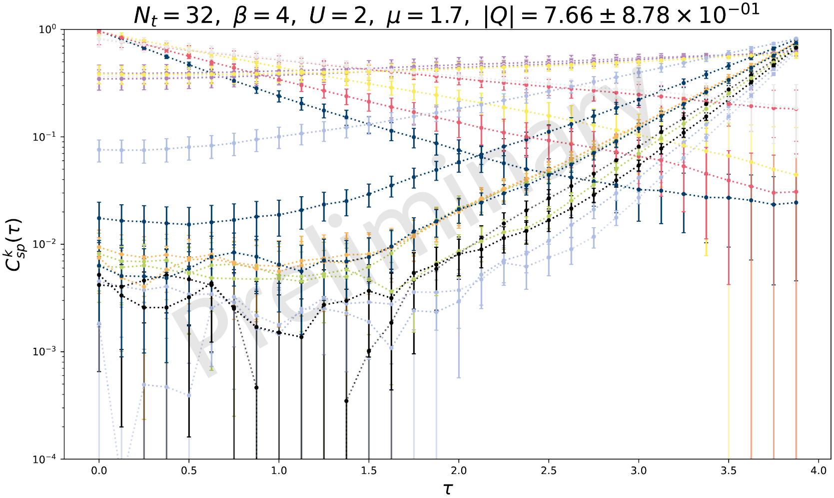
$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 0.4,\ |Q| = 1.19 \pm 6.97 \times 10^{-03}$
$C^k_{sp}(\tau)$
τ
Preliminary

HMC



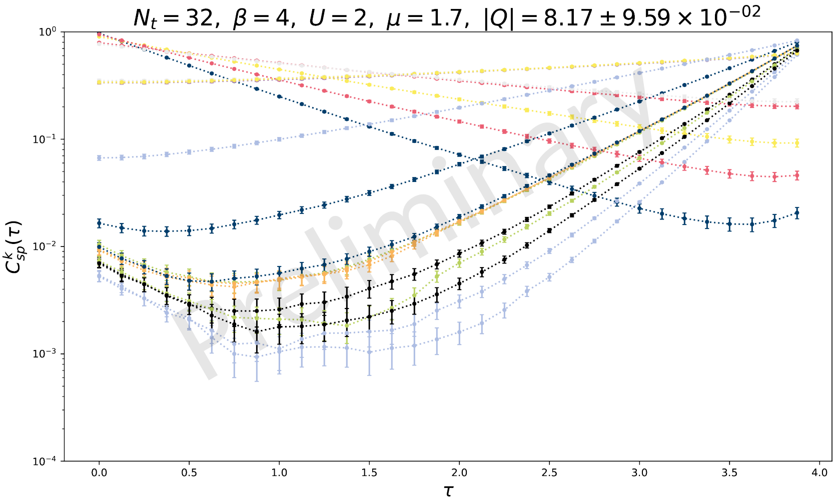
$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 1.2,\ |Q| = 5.01 \pm 1.17 \times 10^{-01}$
$C^k_{sp}(\tau)$
τ
Preliminary

HMC



$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 1.7,\ |Q| = 7.66 \pm 8.78 \times 10^{-01}$
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$C^k_{sp}(\tau)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
Preliminary

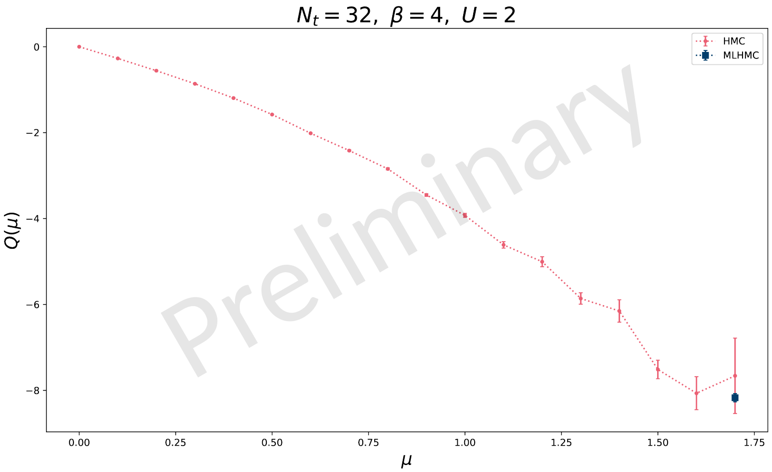
ML HMC



2023 LATTICE

JÜLICH Forschungszentrum

$$Q = \sum_{x=0}^{N_x} \langle \eta_{x,\uparrow} - \eta_{x,\downarrow} \rangle \tag{9}$$

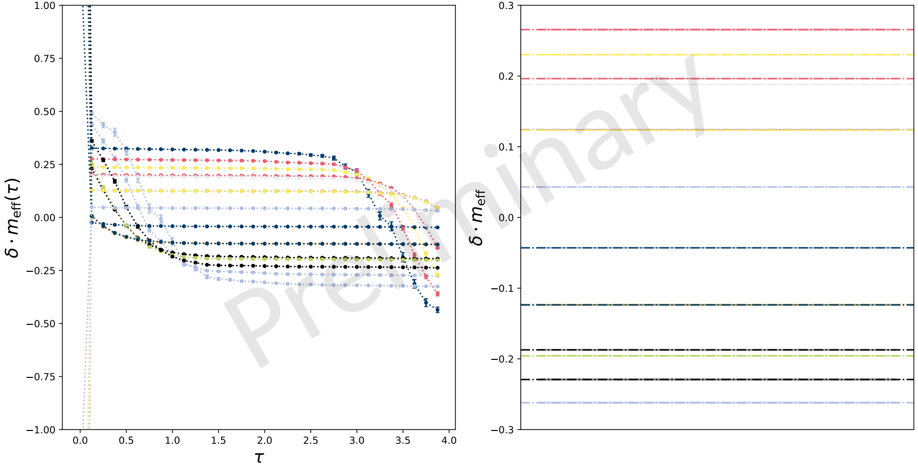


$N_t = 32,\ \beta = 4,\ U = 2$
HMC
MLHMC
Preliminary
0
−2
−4
−6
−8
$Q(\mu)$
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
μ

$$m_{\text{eff}} = \frac{\log C^k_{sp}(\tau + \delta) - \log C^k_{sp}(\tau)}{\delta} \tag{10}$$

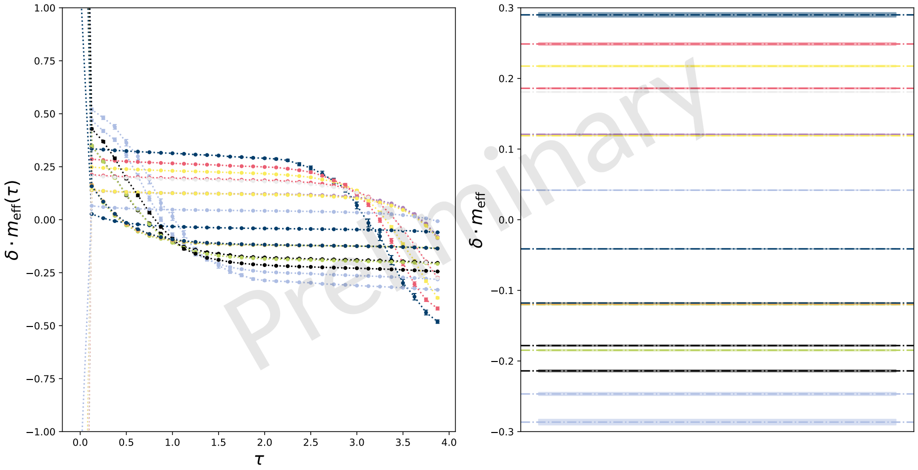
HMC

$N_t = 32,\ \beta = 4,\ U = 1,\ \mu = 0,\ |Q| = 0.00 \pm 4.11 \times 10^{-13}$



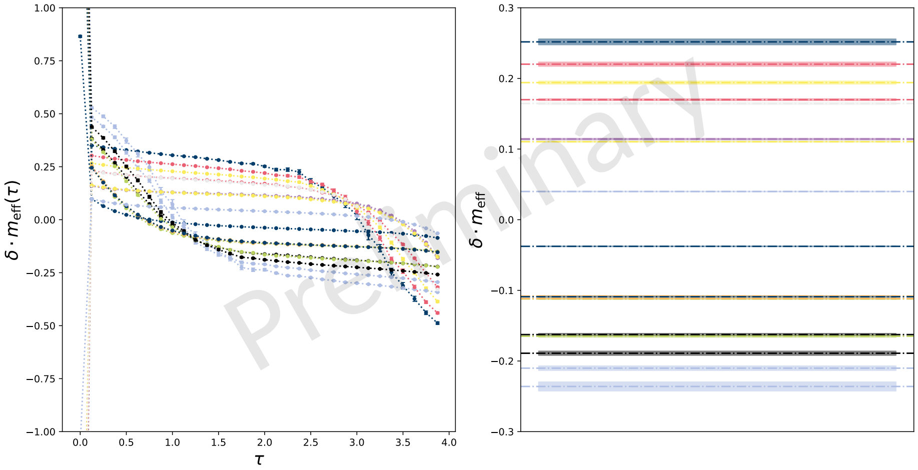
HMC

$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 0,\ |Q| = 0.00 \pm 1.05 \times 10^{-13}$



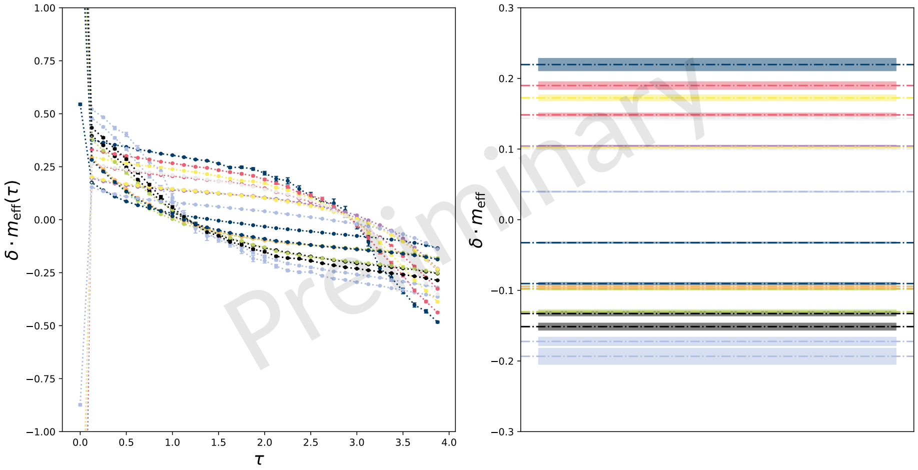
HMC

$N_t = 32,\ \beta = 4,\ U = 3,\ \mu = 0,\ |Q| = 0.00 \pm 2.89 \times 10^{-14}$



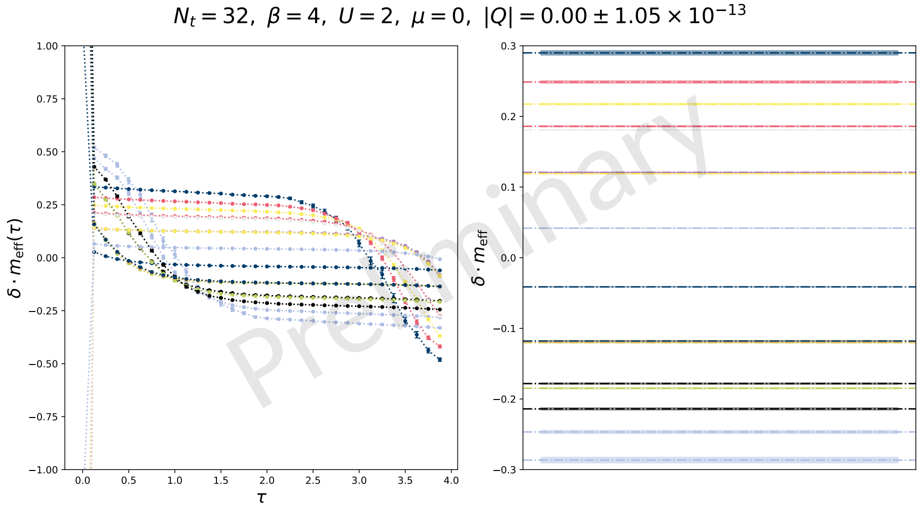
HMC

$N_t = 32,\ \beta = 4,\ U = 4,\ \mu = 0,\ |Q| = 0.00 \pm 9.97 \times 10^{-15}$



$$\mu \neq 0$$

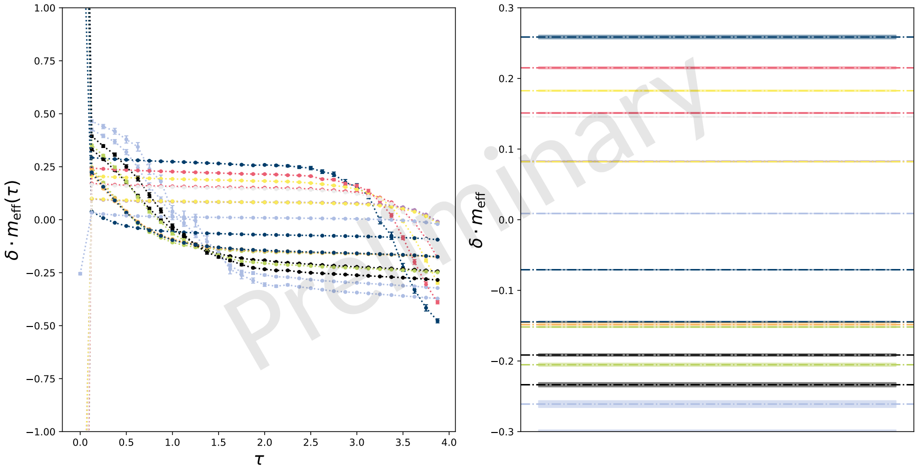
HMC



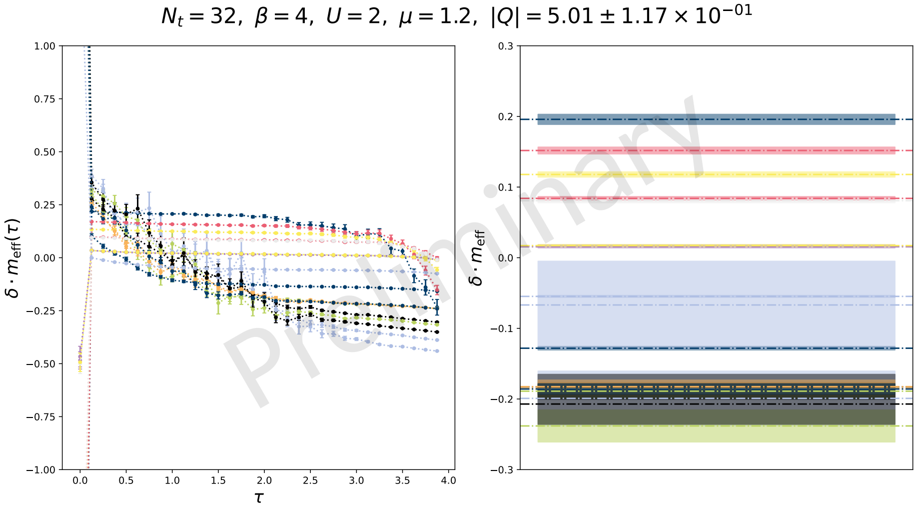
$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 0,\ |Q| = 0.00 \pm 1.05 \times 10^{-13}$
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
$\delta \cdot m_{\text{eff}}(\tau)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
$\delta \cdot m_{\text{eff}}$
Preliminary

HMC

$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 0.4,\ |Q| = 1.19 \pm 6.97 \times 10^{-03}$

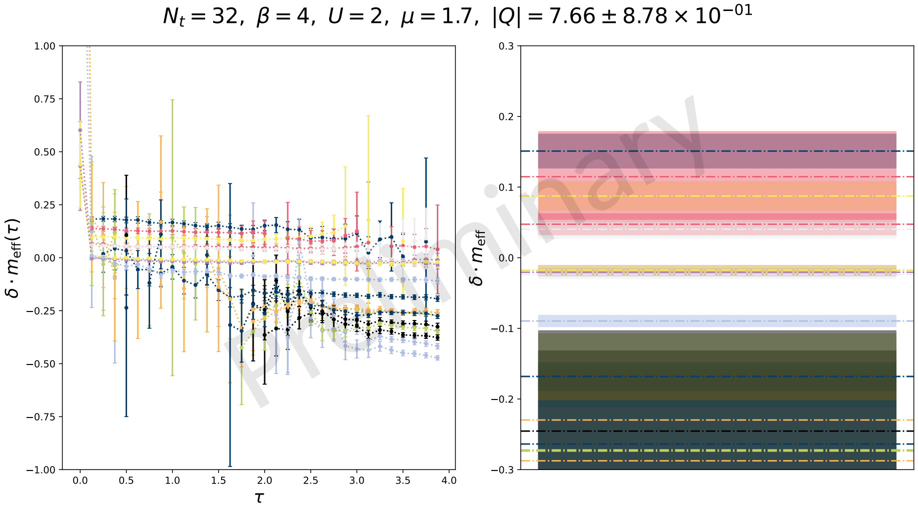


HMC



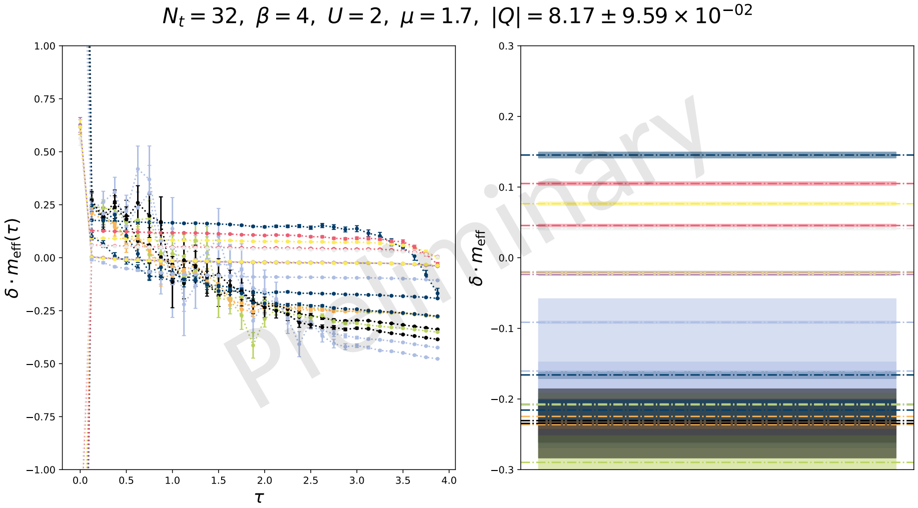
$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 1.2,\ |Q| = 5.01 \pm 1.17 \times 10^{-01}$
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
$\delta \cdot m_{\mathrm{eff}}(\tau)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
$\delta \cdot m_{\mathrm{eff}}$
Preliminary

HMC



$N_t = 32,\ \beta = 4,\ U = 2,\ \mu = 1.7,\ |Q| = 7.66 \pm 8.78 \times 10^{-01}$
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
τ
$\delta \cdot m_{\text{eff}}(\tau)$
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
$\delta \cdot m_{\text{eff}}$
Preliminary

ML HMC



2023 LATTICE

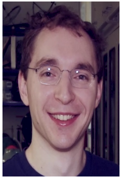
JÜLICH Forschungszentrum

Evan
Berkowitz

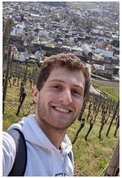
Christoph
Gäntgen

Stefan
Krieg

Tom Luu



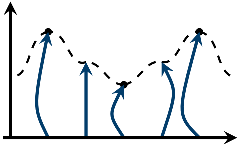
Johann
Ostmeyer

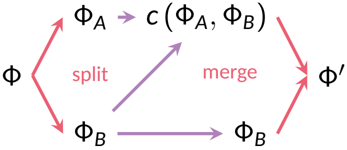


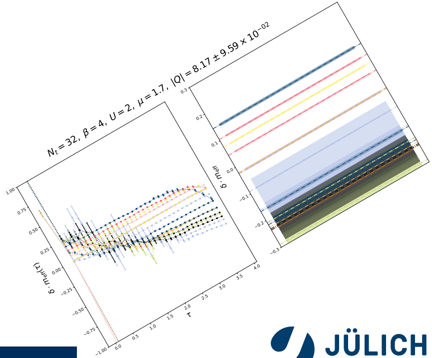
Lado
Razmadze

Aleena
Sibi

Petar
Sinilkov







References I

Alexandru, A., Başar, G., and Bedaque, P. (2016).
Monte carlo algorithm for simulating fermions on lefschetz thimbles.
Physical Review D, 93(1):014504.

Botoshansky, M., Herbstein, F. H., and Kapon, M. (2003).
Towards a Complete Description of a Polymorphic Crystal: The Example of Perylene.
Helvetica Chimica Acta, 86(4):1113–1128.

Gäntgen, C., Berkowitz, E., Luu, T., Ostmeyer, J., and Rodekamp, M. (2023).
Fermionic sign problem minimization by constant path integral contour shifts.
arXiv preprint arXiv:2307.06785.

Glorot, X. and Bengio, Y. (2010).
Understanding the difficulty of training deep feedforward neural networks.
In Teh, Y. W. and Titterington, M., 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. PMLR.

References II

Ostmeyer, J., Berkowitz, E., Krieg, S., Lähde, T. A., Luu, T., and Urbach, C. (2020).
Semimetal–mott insulator quantum phase transition of the hubbard model on the honeycomb lattice.
Physical Review B, 102(24):245105.

Rodekamp, M., Berkowitz, E., Gäntgen, C., Krieg, S., Luu, T., and Ostmeyer, J. (2022).
Mitigating the hubbard sign problem with complex-valued neural networks.
Physical Review B, 106(12):125139.

Wynen, J.-L., Berkowitz, E., Körber, C., Lähde, T. A., and Luu, T. (2019).
Avoiding ergodicity problems in lattice discretizations of the hubbard model.
Physical Review B, 100(7):075141.