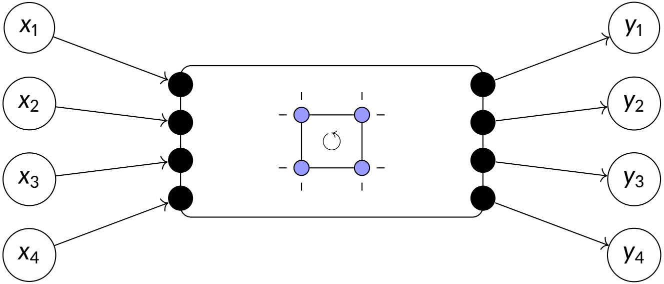


EXPLORING GROUP CONVOLUTIONAL NETWORKS

for Sign Problem Mitigation via Contour Deformation

July 30, 2024 | Christoph Gäntgen | Forschungszentrum Jülich

⟨NuMeriQS⟩

JÜLICH Forschungszentrum

Motivation

Motivation

- To access new physics we need better simulations

Motivation

- To access new physics we need better simulations
- Reduce the Sign-Problem

Motivation

- To access new physics we need better simulations
- Reduce the Sign-Problem
- Find beneficial contour deformation

Motivation

- To access new physics we need better simulations
- Reduce the Sign-Problem
- Find beneficial contour deformation
- And make it cheap

Motivation

- To access new physics we need better simulations
- Reduce the Sign-Problem
- Find beneficial contour deformation
- And make it cheap

Neural Networks

Motivation

- To access new physics we need better simulations
- Reduce the Sign-Problem
- Find beneficial contour deformation
- And make it cheap

Neural Networks

- Risk of unphysical results

Motivation

- To access new physics we need better simulations
- Reduce the Sign-Problem
- Find beneficial contour deformation
- And make it cheap

Neural Networks

- Risk of unphysical results
- Waste resources to learn known features

Hubbard Model

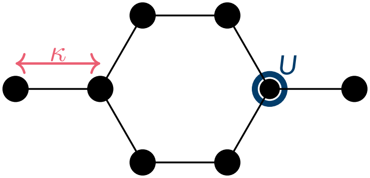
$$H = -\sum_{x,y} \kappa_{x,y} \left(a^\dagger_{x\uparrow} a_{y\uparrow} + a^\dagger_{x\downarrow} a_{y\downarrow} \right) - \frac{U}{2} \sum_{x} (n_{x\uparrow} - n_{x\downarrow})^2 - \mu \sum_{x} (n_{x\uparrow} + n_{x\downarrow})$$

- The Hubbard model is used to approximate solid state systems [Hubbard, 1963]

  κ: Nearest-neighbour hopping (tight binding)

  U: On-site interaction

  μ: Chemical potential



The Sign-Problem

Expectation value

$$\langle \hat{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \, \hat{O}\,[\phi]\, e^{-S[\phi]}\,, \qquad \mathcal{Z} = \int \mathcal{D}\phi \, e^{-S[\phi]}$$

$$\Rightarrow \quad \langle \hat{O} \rangle = \int \mathcal{D}\phi \, \hat{O}\,[\phi]\,\rho\,[\phi] \approx \frac{1}{N}\sum_{n=0}^{N} \hat{O}\,[\phi_n]$$

with $\phi_n \sim \rho\,[\phi_n] = \frac{1}{\mathcal{Z}} e^{-S[\phi_n]}$

Action

$$S = \sum_{x,t} \frac{\phi_{x,t}^2}{2\tilde{U}} - \log\det(M[\phi,\tilde{\kappa},\tilde{\mu}]M[-\phi,-\tilde{\kappa},-\tilde{\mu}])$$

The Sign-Problem

Expectation value

$$\langle \hat{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \, \hat{O}\left[\phi\right] e^{-S[\phi]} \,, \qquad \mathcal{Z} = \int \mathcal{D}\phi \, e^{-S[\phi]}$$

$$\Rightarrow \quad \langle \hat{O} \rangle = \int \mathcal{D}\phi \, \hat{O}\left[\phi\right] \rho\left[\phi\right] \approx \frac{1}{N} \sum_{n=0}^{N} \hat{O}\left[\phi_n\right]$$

with $\phi_n \sim \rho\left[\phi_n\right] = \frac{1}{\mathcal{Z}} e^{-S[\phi_n]}$

Action

$$S = \sum_{x,t} \frac{\phi_{x,t}^2}{2\tilde{U}} - \log\det(M[\phi, \tilde{\kappa}, \tilde{\mu}]M[-\phi, -\tilde{\kappa}, -\tilde{\mu}]) \in \mathbb{C}$$

The Sign-Problem

- $S[\phi] \rightarrow \mathrm{Re}\{S[\phi]\} + \mathrm{i}\,\mathrm{Im}\{S[\phi]\}$
- $\rho[\phi] \rightarrow e^{-\mathrm{Re}\{S[\phi]\}} \in \mathbb{R}$
- $\hat{O}[\phi] \rightarrow \hat{O}e^{-\mathrm{i}\,\mathrm{Im}\{S[\phi]\}} \in \mathbb{C}$

$$\Rightarrow \quad \langle \hat{O} \rangle = \frac{\langle \hat{O} e^{-\mathrm{i}S_I} \rangle_R}{\langle e^{-\mathrm{i}S_I} \rangle_R} \approx \frac{\sum_{n=0}^{N} \hat{O}\,[\phi_n]\, e^{-\mathrm{i}S_I[\phi_n]}}{\sum_{n=0}^{N} e^{-\mathrm{i}S_I[\phi_n]}}$$

$$\text{with } \phi_n \sim e^{-S_R}$$

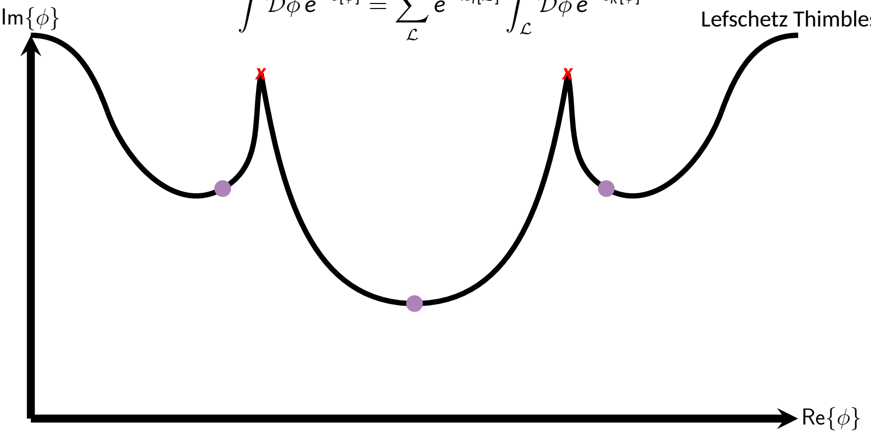
Statistical Power

$$\Sigma = \left| \langle e^{-\mathrm{i}S_I} \rangle_R \right| \qquad \left(N_{\mathrm{eff}} = \Sigma^2 \times N_{\mathrm{cfg}} \right)$$

Contour Deformation

Lefschetz Thimbles

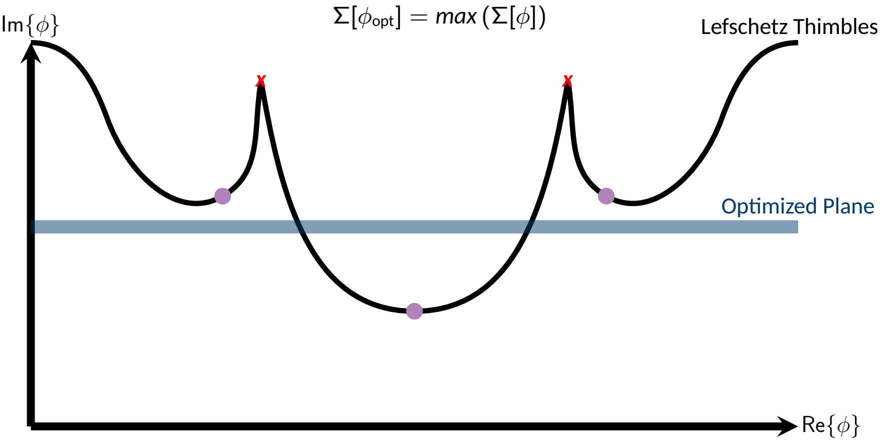
$$\int \mathcal{D}\phi\, e^{-S[\phi]} = \sum_{\mathcal{L}} e^{-iS_I[\mathcal{L}]} \int_{\mathcal{L}} \mathcal{D}\phi\, e^{-S_R[\phi]}$$



[Alexandru et al., 2017, Ulybyshev et al., 2019, Wynen et al., 2020, Rodekamp et al., 2022]

Contour Deformation

Constant Offsets - Optimized Offset



[Gäntgen et al., 2023, Gäntgen et al., 2024]

Neural Networks

Neural Networks

But which one?

Group equivariant Convolutional NNs (G-CNNs)

T. S. Cohen & M. Welling (arXiv: 1602.07576)

Group equivariant Convolutional NNs (G-CNNs)

T. S. Cohen & M. Welling (arXiv: 1602.07576)

- generalization of convolutional NN

Group equivariant Convolutional NNs (G-CNNs)

T. S. Cohen & M. Welling (arXiv: 1602.07576)

- generalization of convolutional NN
- reduces sample complexity by exploiting symmetries

Group equivariant Convolutional NNs (G-CNNs)

T. S. Cohen & M. Welling (arXiv: 1602.07576)

- generalization of convolutional NN
- reduces sample complexity by exploiting symmetries
- high degree of weight sharing

Group equivariant Convolutional NNs (G-CNNs)

T. S. Cohen & M. Welling (arXiv: 1602.07576)

- generalization of convolutional NN
- reduces sample complexity by exploiting symmetries
- high degree of weight sharing
- increase the expressive capacity of the network without increasing the number of parameters

Group equivariant Convolutional NNs (G-CNNs)

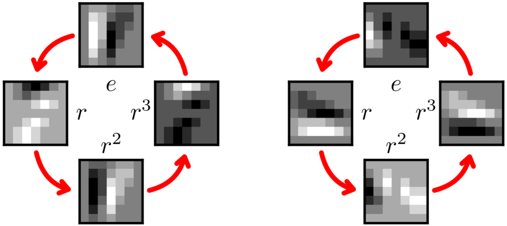
T. S. Cohen & M. Welling (arXiv: 1602.07576)

- generalization of convolutional NN
- reduces sample complexity by exploiting symmetries
- high degree of weight sharing
- increase the expressive capacity of the network without increasing the number of parameters
- can be implemented with negligible computational overhead for discrete groups

Group equivariant Convolutional NNs (G-CNNs)

T. S. Cohen & M. Welling (arXiv: 1602.07576)

$$
\begin{aligned}
[[L_u f] \star \psi](g) &= \sum_{h \in G} \sum_k f_k(u^{-1}h)\psi(g^{-1}h) \\
&= \sum_{h \in G} \sum_k f_k(h)\psi(g^{-1}uh) \\
&= \sum_{h \in G} \sum_k f_k(h)\psi((u^{-1}g)^{-1}h) \\
&= [L_u[f \star \psi]](g)
\end{aligned}
$$



Feature map with rotation [Cohen and Welling, 2016]

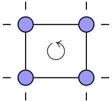
Group equivariant Convolutional NNs (G-CNNs)

for the Hubbard Model

Symmetries of the Hubbard Action:

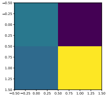
- geometric symmetries of the spatial lattice
  $\rightarrow$ translation (assuming periodic boundaries), rotation, mirroring
- symmetries of the temporal lattice
  $\rightarrow$ translation, reversal
- sign flips (real valued)
  $\rightarrow \mathbb{Z}_2$

Example: four-site square lattice with periodic boundaries



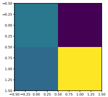
Group equivariant Convolutional NNs (G-CNNs)

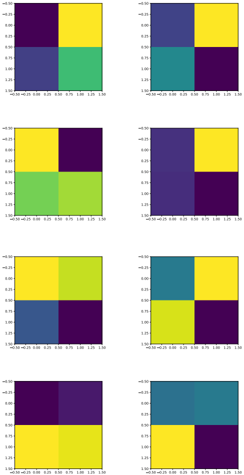
Visualization: Square



Group equivariant Convolutional NNs (G-CNNs)

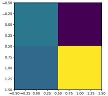
Visualization: Square

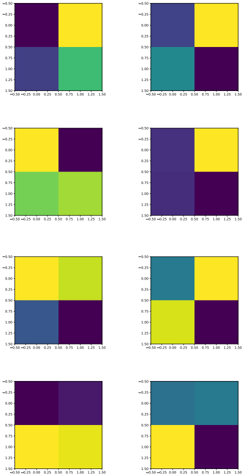




Group equivariant Convolutional NNs (G-CNNs)

Visualization: Square

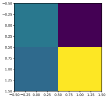


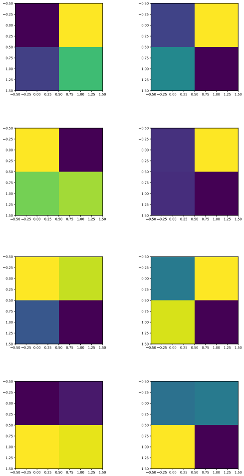


$f(x)$

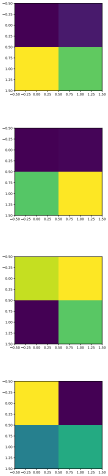
Group equivariant Convolutional NNs (G-CNNs)

Visualization: Square



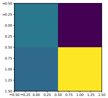


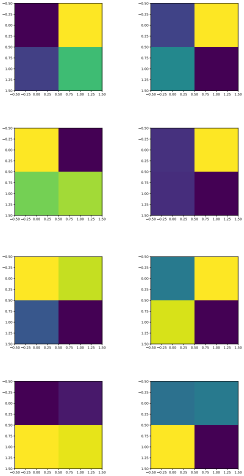
$f(x)$



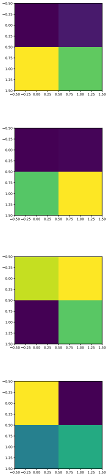
Group equivariant Convolutional NNs (G-CNNs)

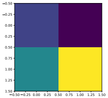
Visualization: Square





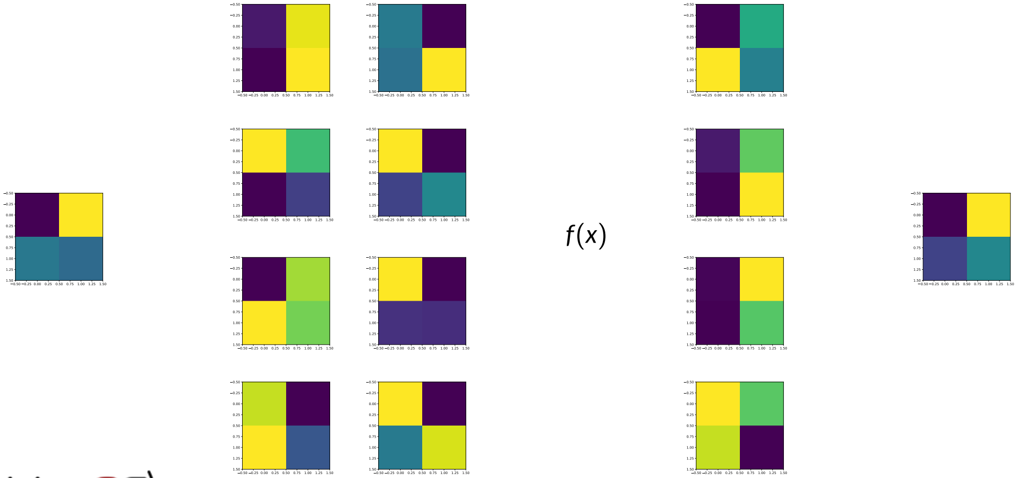
$f(x)$





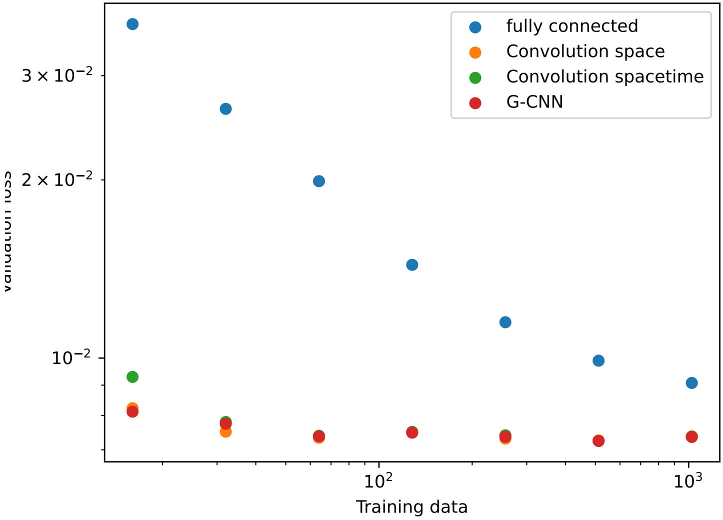
Group equivariant Convolutional NNs (G-CNNs)

Visualization: Square



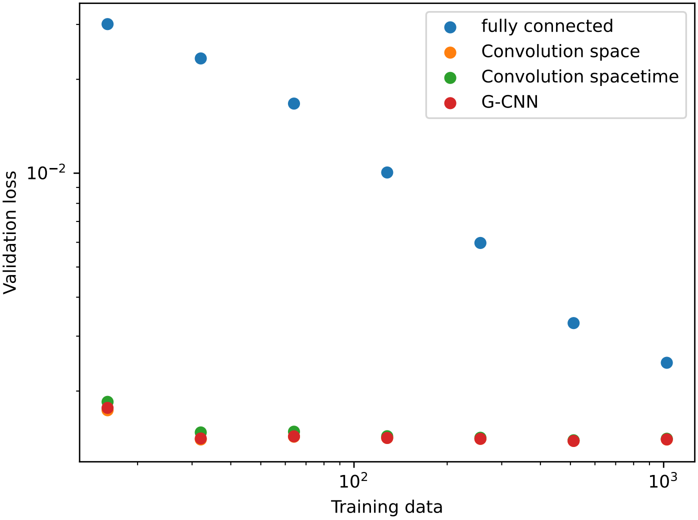
Results

Structure 2x2



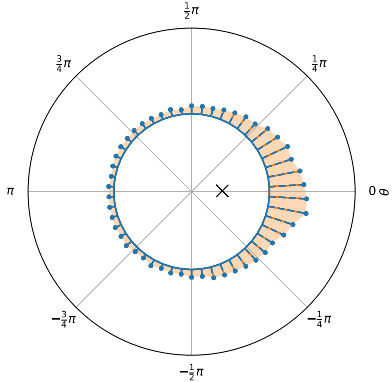
Results

Structure 4x4



Results

HMC



Polar distribution of imaginary phase

Transfer Learning

Outlook: Transfer to larger lattice

Idea: Convolutional networks have no fixed input size

$\rightarrow$ Scale to larger N_t.
$\rightarrow$ Maybe even greater N_x when boundaries are periodic.

Summary

Summary

- Need a fraction of the training data

Summary

- Need a fraction of the training data
- Output is more stable

Summary

- Need a fraction of the training data
- Output is more stable
- Better scaling of needed parameters ?

Summary

- Need a fraction of the training data
- Output is more stable
- Better scaling of needed parameters ?
- Potential for transfer learning

Thank You for your attention

Collaborators:

Tom Luu

Marcel Rodekamp

References I

Alexandru, A., Bedaque, P., Lamm, H., and Lawrence, S. (2017).
Deep learning beyond lefschetz thimbles.
Phys. Rev. D 96, 094505 (2017), 96(9):094505.

Cohen, T. S. and Welling, M. (2016).
Group equivariant convolutional networks.
Proceedings of the International Conference on Machine Learning (ICML), 2016.

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

Gäntgen, C., Berkowitz, E., Luu, T., Ostmeyer, J., and Rodekamp, M. (2024).
Reducing the sign problem with simple contour deformation.
Proceedings of Science Volume 453 (LATTICE2023).

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

References II

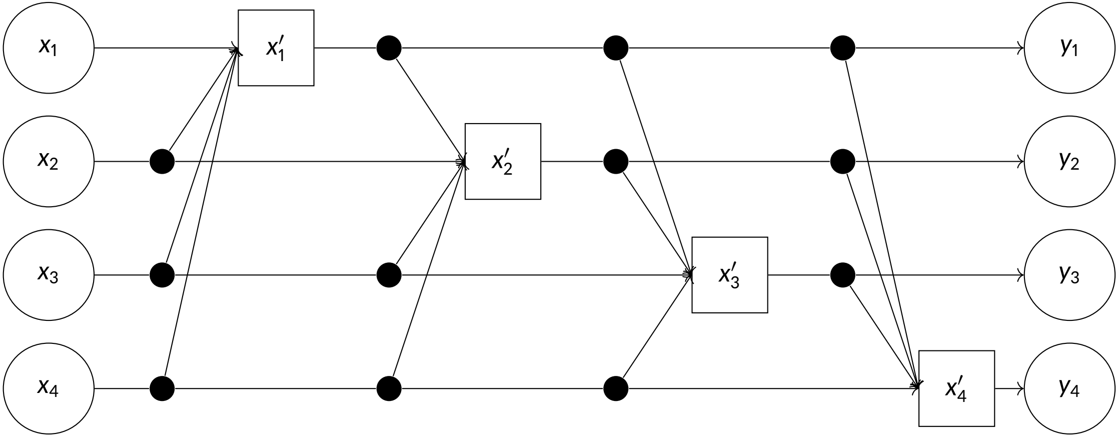
Rodekamp, M., Berkowitz, E., Gäntgen, C., Krieg, S., Luu, T., and Ostmeyer, J. (2022).
Mitigating the hubbard sign problem with complex-valued neural networks.
Physical Review B, 106(12):125139.

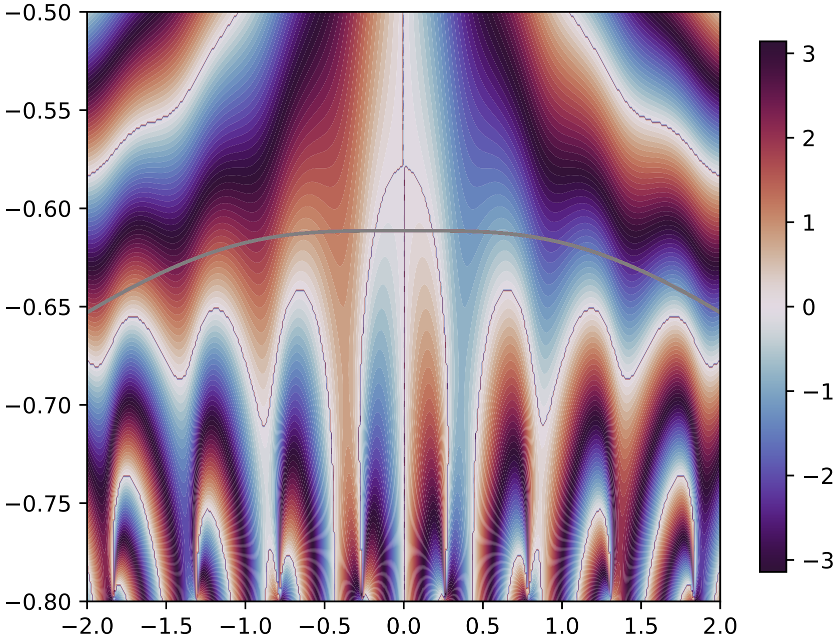
Ulybyshev, M., Winterowd, C., and Zafeiropoulos, S. (2019).
Lefschetz thimbles decomposition for the hubbard model on the hexagonal lattice.
Phys. Rev. D 101, 014508 (2020), 101(1):014508.

Wynen, J.-L., Berkowitz, E., Krieg, S., Luu, T., and Ostmeyer, J. (2020).
Leveraging machine learning to alleviate hubbard model sign problems.
Phys. Rev. B 103, 125153 (2021), 103(12):125153.

Backup

Coupling Structure




−0.50
−0.55
−0.60
−0.65
−0.70
−0.75
−0.80
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
3
2
1
0
−1
−2
−3