Flows for Fermionic Lattice Field Theories

Michael Albergo, Gurtej Kanwar, Sébastien Racanière, Danilo J. Rezende, Julian M. Urban, Denis Boyda, Kyle Cranmer, Daniel C. Hackett, and Phiala Shanahan

Lots of interesting ML talks !!

Check them out: h/t Tej Kanwar

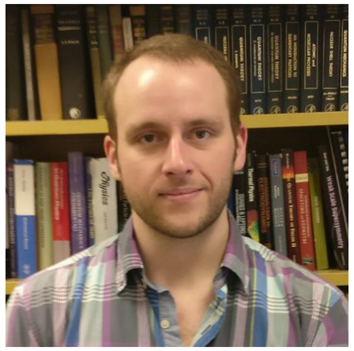
ML talks

Name	Date (EDT)	Title
Sébastien Racanière (Plenary)	@July 28, 2021 9:00	Perspectives on the current status of and future prospects for ML in lattice QFT
Akio Tomiya	@July 27, 2021 13:00	Smearing is a neural network
Neill Warrington	@July 27, 2021 13:15	Contour Deformations for Lattice Field Theory
Gurtej Kanwar	@July 27, 2021 13:30	Observifolds: Taming the observable signal-to-noise problem via path integral contour deformations
Shuzhe Shi	@July 27, 2021 13:45	From lattice QCD to heavy-flavor in-medium potential via deep learning
Sunkyu Lee	@July 27, 2021 22:45	Deep learning study on the Dirac eigenvalue spectrum of staggered quarks
Yukari Yamauchi	@July 28, 2021 21:00	Normalizing flows for the real-time sign problem
Chen ShiYang	@July 28, 2021 22:00	Machine learning Hadron Spectral Functions in Lattice QCD
Fu-Jiun Jiang	@July 28, 2021 22:15	A universal neural network for learning phases and criticalities
Dimitrios Bachtis	@July 29, 2021 5:00	Machine learning with quantum field theories
David Muller	@July 29, 2021 5:15	Lattice Gauge Symmetry in Neural Networks
Matteo Favoni	@July 29, 2021 5:30	Generalization capabilities of neural networks in lattice applications
Gert Aarts	@July 29, 2021 5:45	Interpreting machine learning functions as physical observables
Kim Nicoli	@July 29, 2021 6:00	Machine Learning for Thermodynamic Observables
Marina Marinkovic	@July 29, 2021 6:15	Machine learning phase transitions in a scalable manner
Michael Albergo	@July 29, 2021 13:30	Flow-based sampling for fermionic field theories
Xiao-Yong Jin	@July 29, 2021 13:45	Neural Network Field Transformation and Its Application in HMC
Denis Boyda	@July 29, 2021 14:00	Sampling lattice gauge theory in four dimensions with normalizing flows
Sam Foreman	@July 29, 2021 14:45	LeapFrogLayers: A Trainable Framework for Effective Topological Sampling
Boram Yoon	@July 29, 2021 22:15	Prediction and compression of lattice QCD data using machine learning algorithms on quantum annealer

Thanks to collaborators!

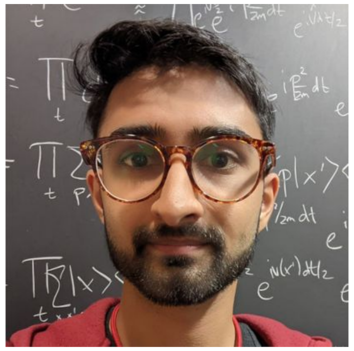
Massachusetts Institute of Technology

Phiala Shanahan



Dan Hackett

Denis Boyda



Gurtej Kanwar

Google DeepMind

Sébastien Racanière

Danilo Rezende

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Julian Urban

NEW YORK UNIVERSITY

Kyle Cranmer

Agenda

Fermions and Monte Carlo Sampling

Exact sampling schemes for theories with fermions

Quick intro to flow-based sampling

Constructing flows for these samplers, applications to Yukawa theory

Fermionic Path Integrals and the Determinant

Standard LFT protocol: numerically evaluate path integral expectation

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \mathcal{D}\phi \, e^{-S(\phi)} \mathcal{O}(\phi) \quad \longrightarrow \quad \begin{cases} \langle \mathcal{O} \rangle \approx \frac{1}{N} \sum_{\phi \sim p} \mathcal{O}(\phi) \\ p(\phi) = \frac{1}{Z} e^{-S(\phi)} \end{cases}$$

Monte Carlo (HMC)

Fermions:

$$S(\psi, \bar{\psi}, \phi) = S_B(\phi) + S_F(\psi, \bar{\psi}, \phi)$$

$$S_F(\psi, \bar{\psi}, \phi) = \sum_{f=1}^{N_f} \bar{\psi}_f D_f(\phi) \psi_f$$

$$\int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F(\psi, \bar{\psi}, \phi)} = \prod_{f=1}^{N_f} \det D_f(\phi) \quad \rightarrow \quad p(\phi) = \frac{1}{Z} e^{-S_B} \prod_{f=1}^{N_f} \det D_f$$

Can flows help?

$\det D$ scales $O(V^3)$
sampling this way intractable ✗

Common workarounds

Determinant estimators

- Hutchinson: $\operatorname{tr} D_f = \mathbb{E}_{x\sim\mathcal{N}(0,1)}\left[x^T D_f x\right]$
- Sohl-Dickstein: $|\det(D_f)|^{-1} = \mathbb{E}_{s\sim S^{n-1}}[||D_f s||^{-n}]$

Operator/action preconditioning

- even-odd conditioning: effective volume $V \to V/2$
- Hasenbusch action separation

Pseudofermions

auxiliary boson field

"pseudofermion"

$$\det \mathcal{M} = \frac{1}{Z_{\mathcal{N}}}\int \mathcal{D}\left[\varphi_R, \varphi_I\right] e^{-\varphi^\dagger \mathcal{M}^{-1}\varphi}$$

For two degenerate fermions, replace D_f with $\mathcal{M} = D_{f_1} D_{f_1}^\dagger$ (positive definite)

$$S_{PF}(\phi,\varphi) = \varphi^\dagger \mathcal{M}^{-1}(\phi)\varphi \equiv \sum_{k=1}^{N_{\text{pf}}} \varphi_k^\dagger \mathcal{M}_k^{-1}(\phi)\varphi_k$$

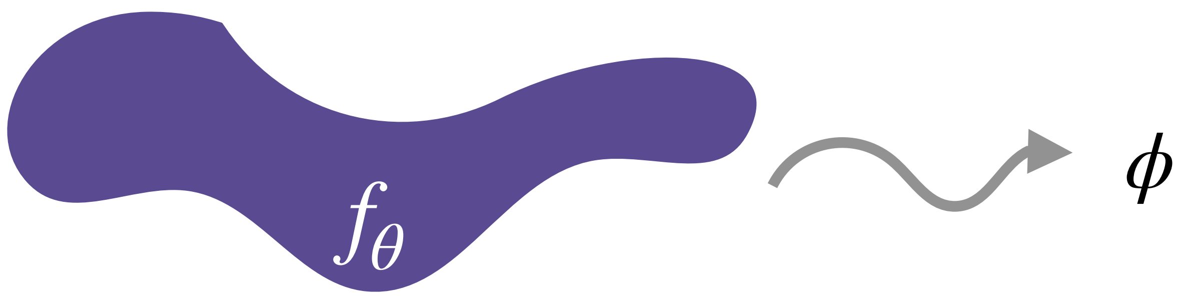
$$p(\phi,\varphi) = \frac{1}{Z} e^{-S_B(\phi) - S_{PF}(\phi,\varphi)}$$

S_{PF} encodes presence of fermions w/o determinant

new **joint** distribution

Aside: exact sampling with generative models

Generative models (in our words)



emits model $q(\phi)$ that approximates target $p(\phi)$

function or sampler that models the data generation process

If $q(\phi)$ is tractable to compute, correctable with Markov chains/reweighting schemes

$$A(\phi \to \phi') = \min\left(1, \frac{q(\phi)}{p(\phi)}\frac{p(\phi')}{q(\phi')}\right)$$

$$\langle O \rangle \approx \frac{1}{Z}\frac{1}{n}\sum_i^n O(\phi_i) w_i$$

where $w_i = \dfrac{e^{-S(\phi_i)}}{q(\phi_i)}$

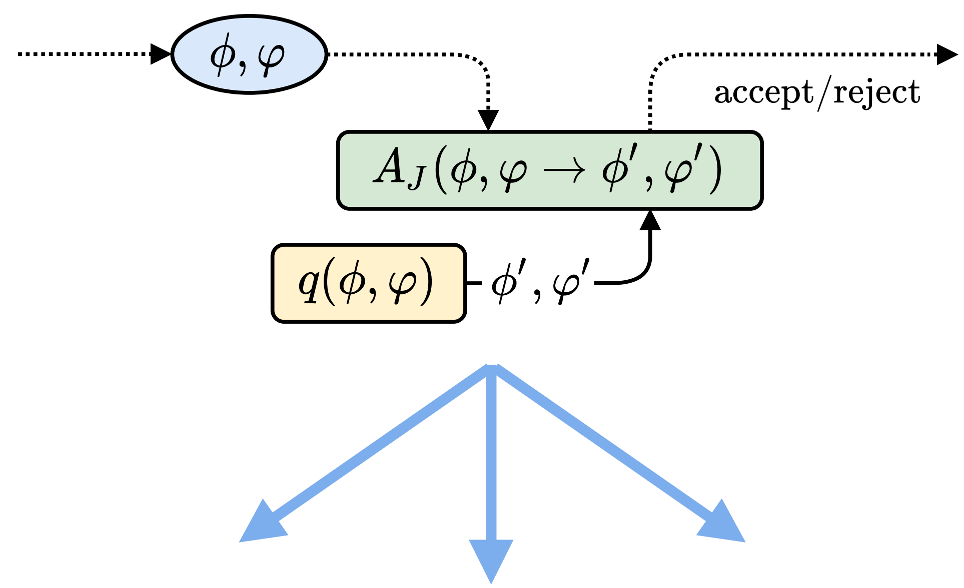
Given our target densities $p(\phi), p(\phi, \varphi)$, what exact sampling chains can we construct?

Sampling schemes derived from the joint

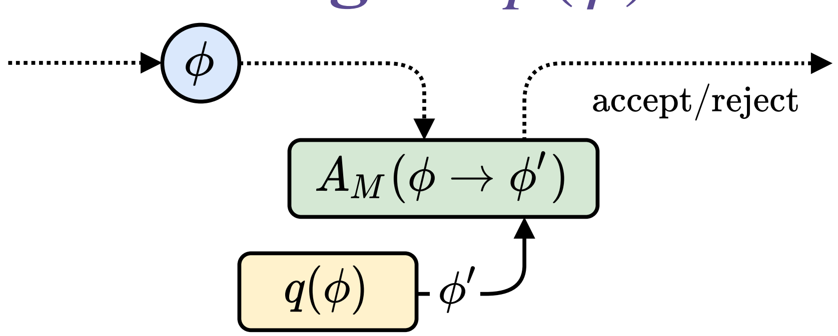
Joint $p(\phi, \varphi)$

$$p(\phi,\varphi) = \frac{1}{Z}\exp\left(-S_B(\phi) - \varphi^\dagger[\mathcal{M}(\phi)]^{-1}\varphi\right)$$

$$A_J\left(\phi,\varphi \to \phi',\varphi'\right) = \min\left(1, \frac{p\left(\phi',\varphi'\right)}{p(\phi,\varphi)}\frac{q(\phi,\varphi)}{q\left(\phi',\varphi'\right)}\right)$$



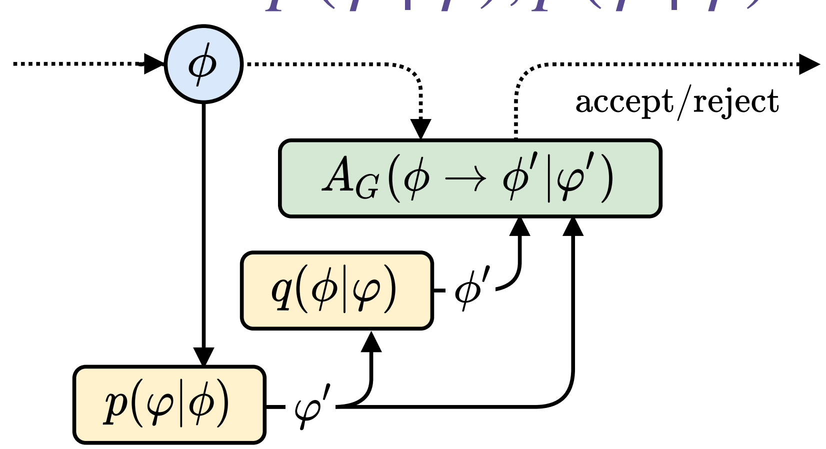
Marginal $p(\phi)$



$$p(\phi) = \frac{Z_{\mathcal{N}}}{Z}\exp\left(-S_B(\phi)\right)\det\mathcal{M}(\phi)$$

$$A_M\left(\phi \to \phi'\right) = \min\left(1, \frac{e^{-S_B(\phi')}\det\mathcal{M}\left(\phi'\right)}{e^{-S_B(\phi)}\det\mathcal{M}(\phi)}\frac{q(\phi)}{q\left(\phi'\right)}\right)$$

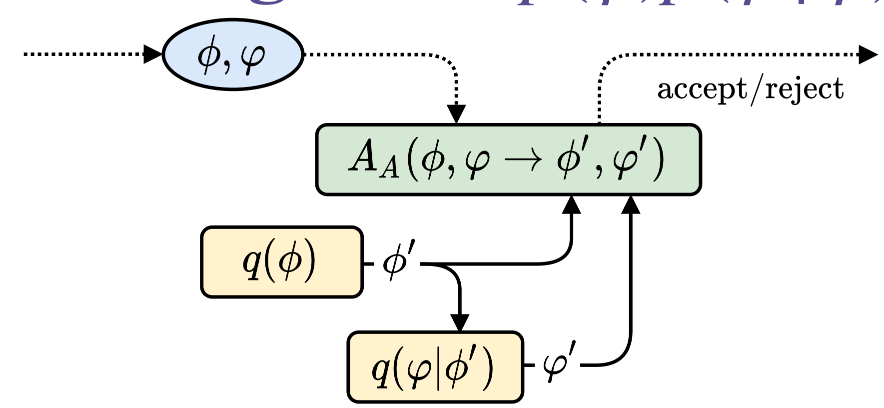
Gibbs $p(\varphi \,|\, \phi), p(\phi \,|\, \varphi)$



$$p(\phi \mid \varphi) = \frac{\exp\left(-S_B(\phi) - \varphi^\dagger[\mathcal{M}(\phi)]^{-1}\varphi\right)}{\int d\phi \exp\left(-S_B(\phi) - \varphi^\dagger[\mathcal{M}(\phi)]^{-1}\varphi\right)}$$

$$A_G\left(\phi \to \phi' \mid \varphi'\right) = \min\left(1, \frac{p\left(\phi',\varphi'\right)}{p\left(\phi,\varphi'\right)}\frac{q\left(\phi \mid \varphi'\right)}{q\left(\phi' \mid \varphi'\right)}\right)$$

Autoregressive $p(\phi)p(\varphi \,|\, \phi)$



$$p(\varphi \mid \phi) = \frac{1}{Z_{\mathcal{N}}\det\mathcal{M}(\phi)}\exp\left(-\varphi^\dagger[\mathcal{M}(\phi)]^{-1}\varphi\right)$$

$$A_A\left(\phi,\varphi \to \phi',\varphi'\right) = \min\left(1, \frac{p\left(\phi',\varphi'\right)}{p(\phi,\varphi)}\frac{q(\phi)q(\varphi|\phi)}{q\left(\phi'\right)q\left(\varphi'|\phi'\right)}\right)$$

Flow-based generative models

How do we realize a model $q(\phi)$?

Let $f : \mathscr{X} \to \mathscr{X}$ such that $\phi = f(z)$ and $z = f^{-1}(\phi)$.

then

$$q(\phi) = r_z(f^{-1}(\phi)) \left| \det \frac{\partial f}{\partial z} \right|^{-1}$$

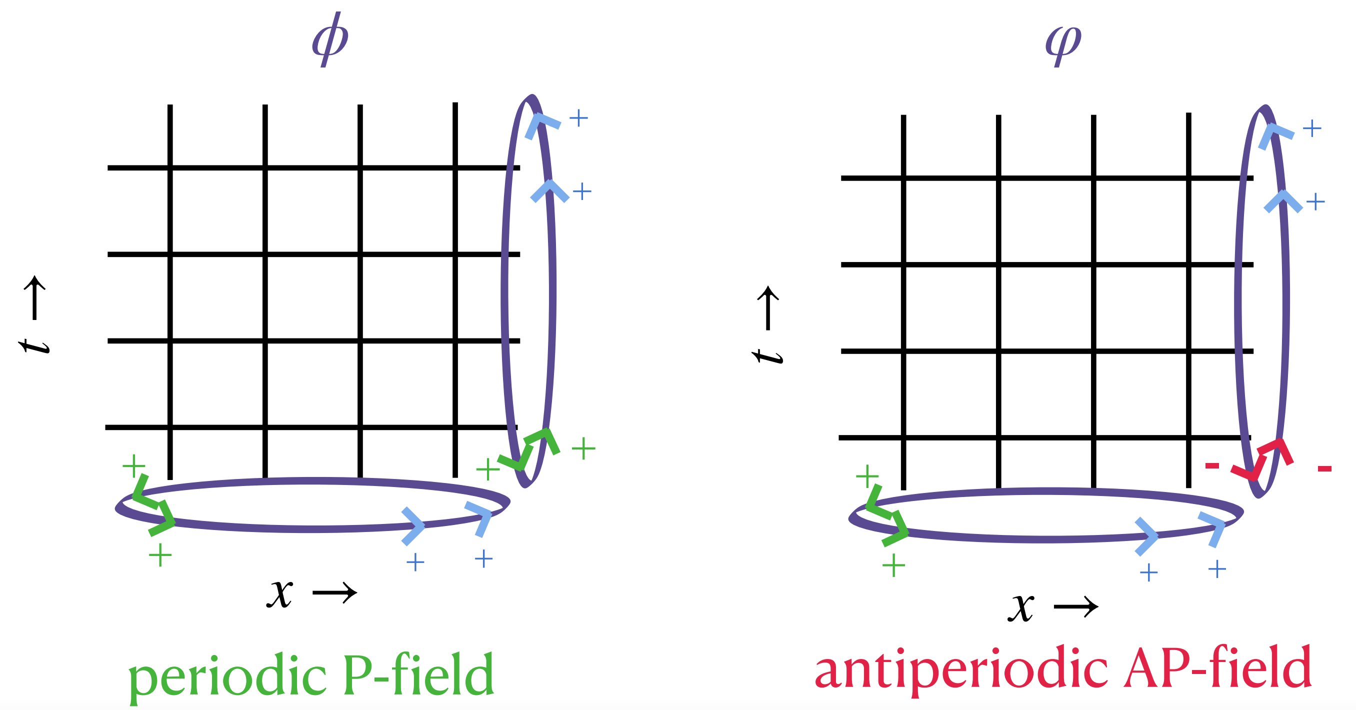
tells us how to reweight a sample ϕ under prior to define new density

optimal $q(\phi)$?

$$q^* = \arg\min_{q \in Q} D_{\mathrm{KL}}(q||p)$$

Efficient parameterization: f should be equivariant to symmetries $p(\phi)$

For staggered pseudofermions



need translation equivariance in f_φ

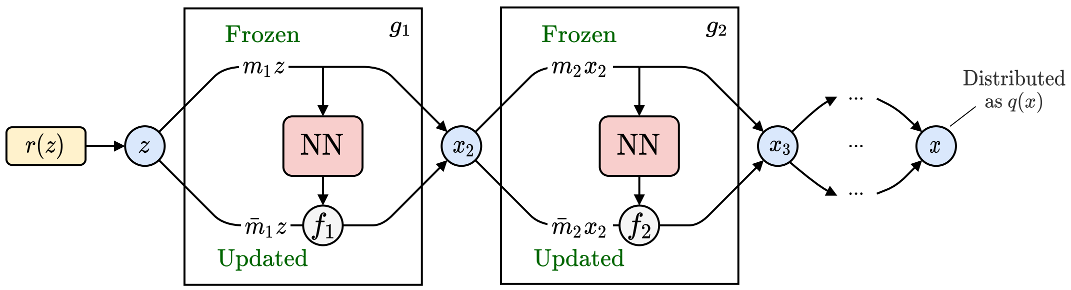
Constructing equivariant fermionic flows

We build $\begin{bmatrix} \text{convolutional couplings } g_i \\ \text{linear operators } \mathscr{W}_i \\ \text{convex potentials } u_i \end{bmatrix}$ that transform:

P-fields → P-fields

AP-fields → AP-fields

g_i



$\mathscr{W}_i$

linear operator for which $\mathscr{W}\mathscr{W}^\dagger \approx DD^\dagger$

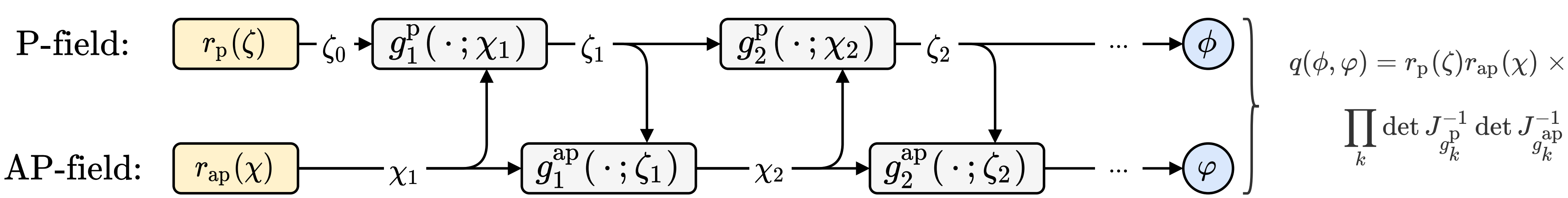
$$\mathcal{W}_i = \begin{bmatrix} a_1 & & & & \pm b_1 \\ b_2 & a_2 & & 0 & \\ & \cdots & \cdots & & \\ & 0 & & & \\ & & & b_L & a_L \end{bmatrix}$$

∇u_i

$f_i = \nabla u_i$ transform block is the gradient of strongly convex function

Example:

flow for $p(\phi, \varphi)$



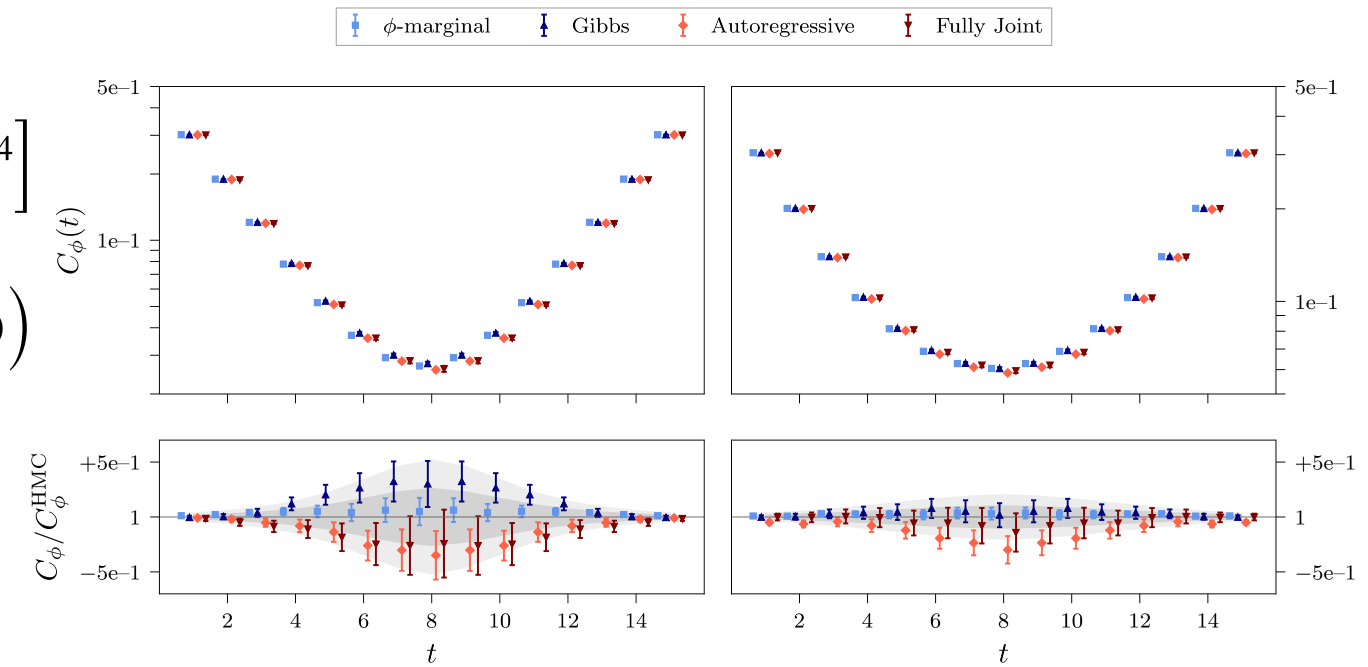
see backup slides for the rest

Application: 1+1 Yukawa

Action $S = S_B + S_F$

$$S_B(\phi) = \sum_{x\in\Lambda}[-2\sum_{\mu=1}^{d}\phi(x)\phi(x+\hat{\mu}) + \left(m^2+2d\right)\phi(x)^2 + \lambda\phi(x)^4]$$

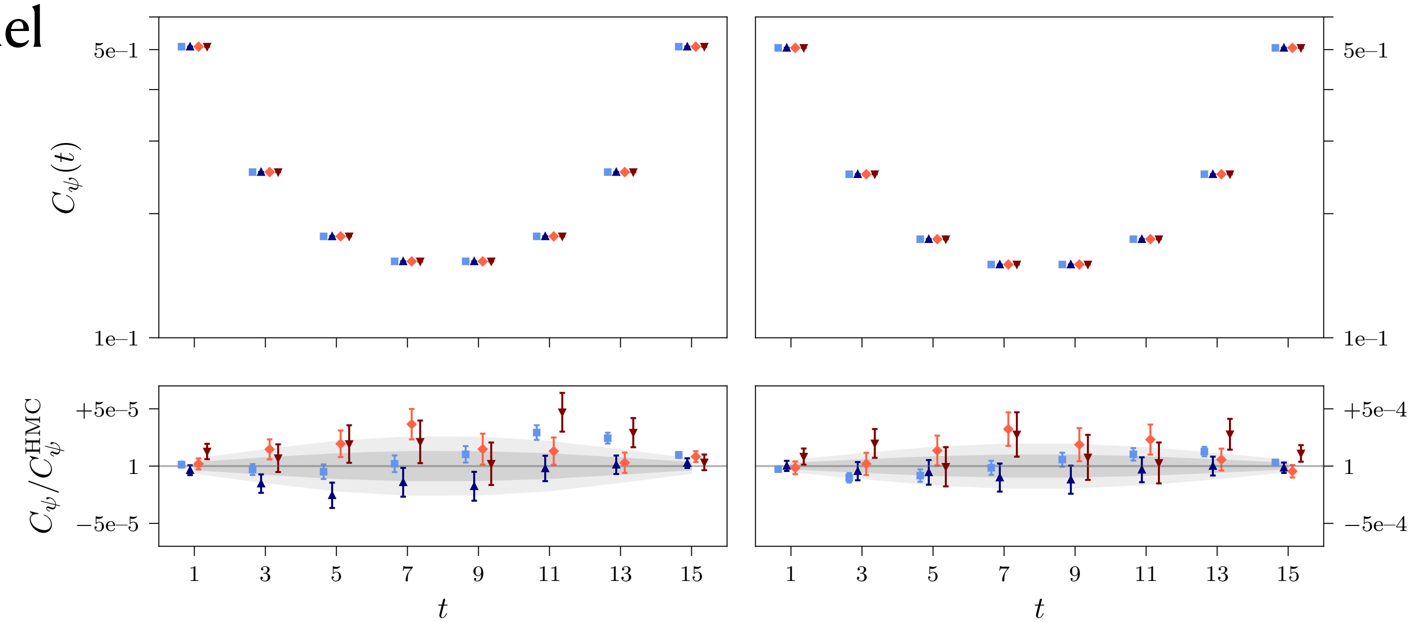
$$D_{xy} = \sum_{\mu=1}^{d}\eta_\mu(x)\frac{\delta(x-y+\hat{\mu})-\delta(x-y-\hat{\mu})}{2} + \delta(x-y)\left(m_f + g\phi(x)\right)$$



On L=16, t=16 lattice

scalar correlator: $C_\phi(t) = \frac{1}{V}\sum_{x}\sum_{\vec{y}} C(x, x+(\vec{y}, t))$

- τ_{int}'s range from 0.7 -> 8.7 depending on model
- accept rates range from 30% to 92%



fermionic correlator

Takeaways

A suite of methods for flows and fermions that we hope you'll help us explore further. Thanks !!

Backup Slides

Even-odd conditioning

$$D_f = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \qquad \det D_f = \det\left(\mathcal{D}\mathcal{B}^{-1}\mathcal{A} - \mathcal{C}\right)\det(\mathcal{B})$$

effective volume $V \to V/2$

Flow architectures for each target distribution