Practical applications of machine-learned flows on gauge fields

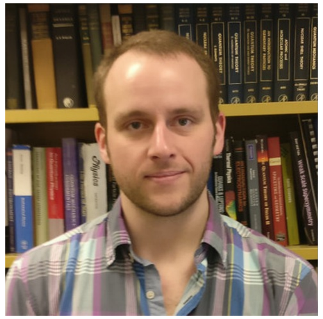
Dan Hackett (MIT / IAIFI)

Lattice 2023

July 3, 2023

Collaborators (non-exhaustive)


Phiala Shanahan


Dan Hackett


Denis Boyda


Ryan Abbott

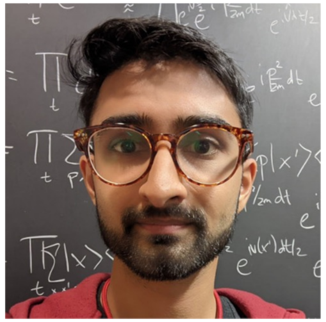

Julian Urban


Fernando Romero-López

NEW YORK UNIVERSITY


Michael Albergo

UNIVERSITÄT BERN


Gurtej Kanwar


Kyle Cranmer

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

DeepMind


Sébastien Racanière


Danilo Rezende

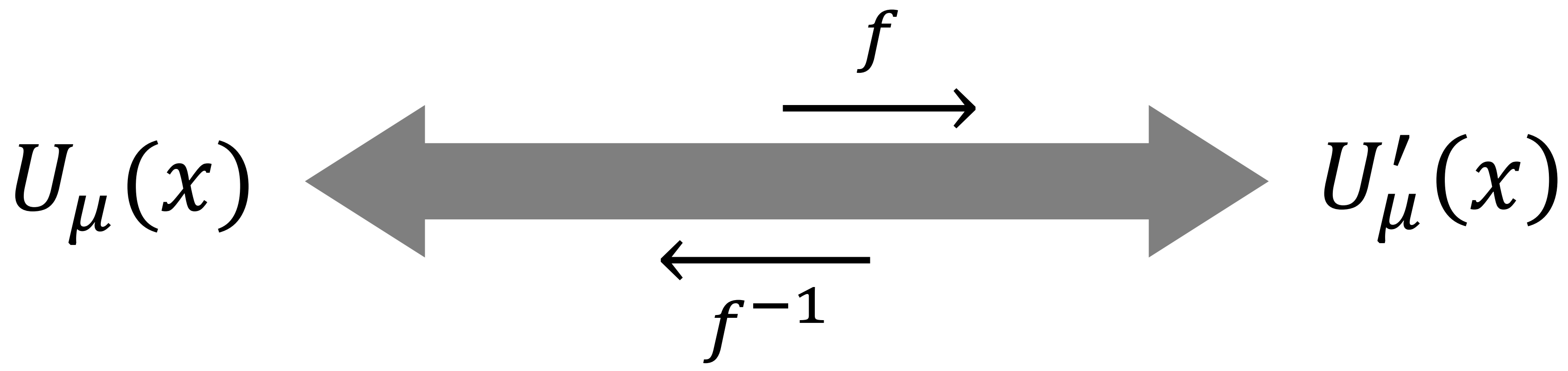

Alex Matthews


Aleksandar Botev


Ali Razavi

Flows

f: learned, invertible map between gauge fields



…with a tractable Jacobian determinant $J_f(U) = \left|\det \frac{\partial f(U)}{\partial U}\right|$

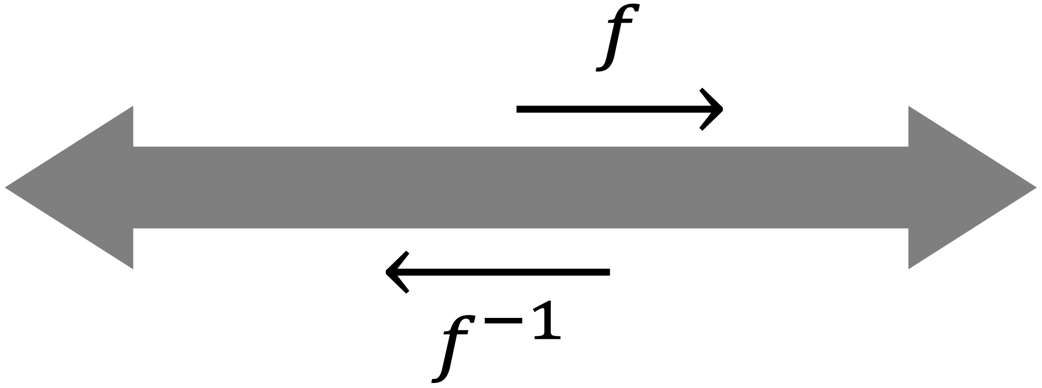
…maybe equivariant w/r/t symmetries g of interest $f\big(g(U)\big) = g(f(U))$

…with many tunable parameters

Flows

Flows are "bridges" between different distributions/theories/actions

$$r(U) = \frac{e^{-S_r(U)}}{Z_r}$$



$$q(U') = \frac{r(U)}{J_f(U)}$$

Exact bridge between r and q

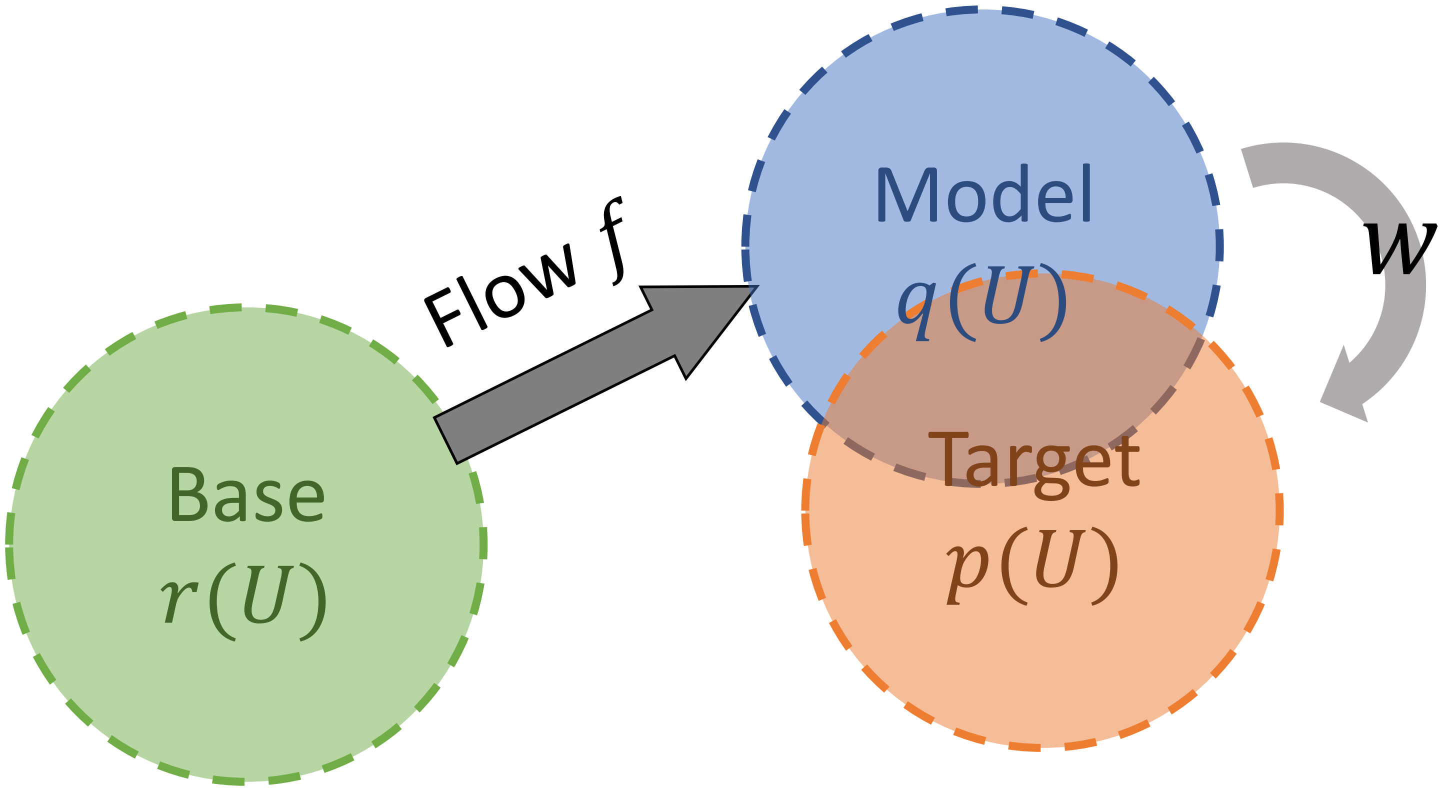
Choose r, but flow induces q

For sampling applications: variationally optimize f so $q \approx p \propto e^{-S_p}$

→ *Approximate* bridge between r and p

(Approximate) direct sampling with flows

Apply f to Haar uniform to get model q, tune f so $q \approx p \propto e^{-S_{\text{target}}}$



Reweight from $q \to p$

$w(U) = p(U) \,/\, q(U)$

$\langle O \rangle_p = \langle w O \rangle_q$

Measure of sampling quality:

$\text{ESS} = 1/\langle w^2 \rangle_q \in [0,1]$

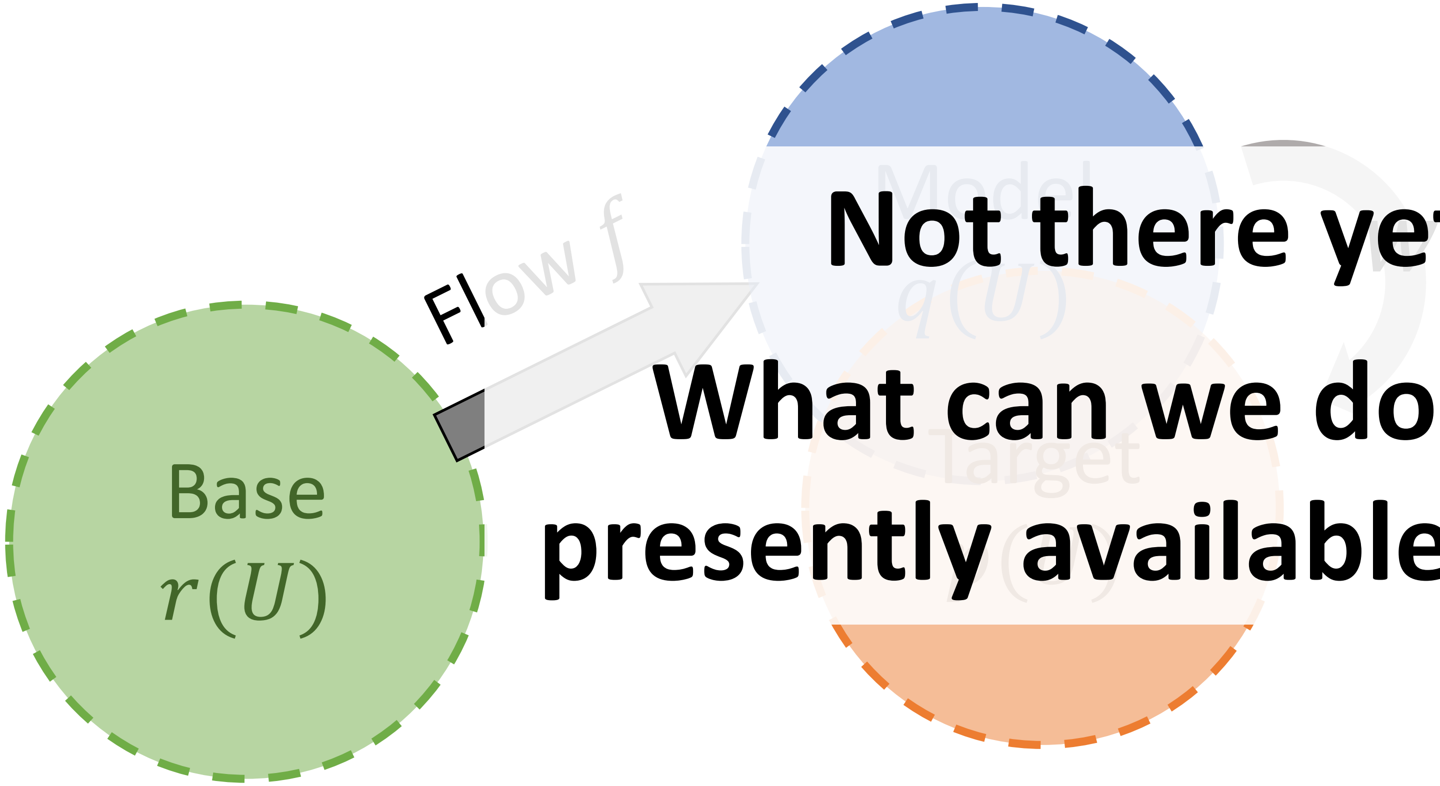
Intuition:

$\text{Var}_q[wO] \approx \text{Var}_p[O]/\text{ESS}$

$\text{ESS} \sim 1/2\, \tau_{AC}$

(Approximate) direct sampling with flows

Apply f to Haar uniform to get model q, tune f so $q \approx p \propto e^{-S_{\text{target}}}$



Not there yet!

What can we do with presently available flows?

Reweight from $q \to p$

$w(U) = p(U) / q(U)$

$\langle O \rangle_p = \langle wO \rangle_q$

Measure of sampling quality:

$\text{ESS} = 1/\langle w^2 \rangle_q \in [0,1]$

Intuition:

$\text{Var}_q[wO] \approx \text{Var}_p[O]/\text{ESS}$

$\text{ESS} \sim 1/2\,\tau_{AC}$

Numerical details

Wilson pure gauge $\mathrm{SU}(3)$

Sample w/ heatbath (HB) + overrelaxation (OR)

$$S = -\frac{\beta}{N_c}\sum_{\mu<\nu} P_{\mu\nu}$$

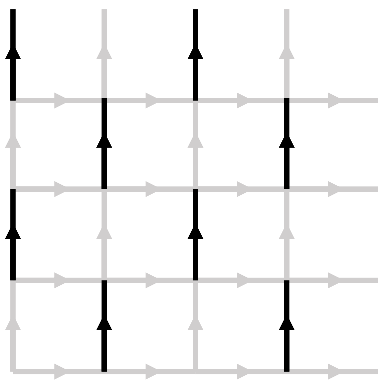
Topological charges

Wilson flow to $t/a^2 = 2$, compute w/ clover definition, round to integer

Residual flows [2305.02402]

Each layer applies a step of gradient flow w/r/t a learned action to a subset of links

→ tractable/inexpensive exact Jacobian



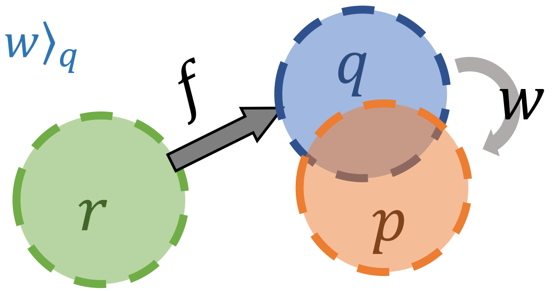
Reverse KL self-training

Sample base distribution r w/ HB+OR

Train on smaller volumes, then transfer

Minimize:

$D_{KL}(q||p) = \langle -\log w\rangle_q$



App 1: Correlated ensembles

Flow an ensemble

$\rightarrow$ $\{U\}$ and $\{f(U)\}$ are correlated

This is useful!

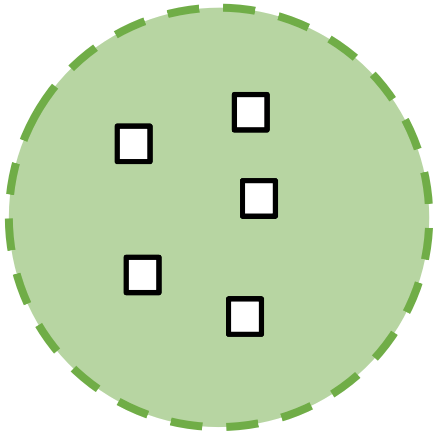
e.g. for noise cancellation in differences

$$\langle O \rangle_p - \langle O \rangle_r$$

$$= \langle wO \rangle_q - \langle O \rangle_r$$

$$= \langle w(f(U))\, O(f(U)) - O(U) \rangle_{U\sim r}$$

$\{U\} \sim r$

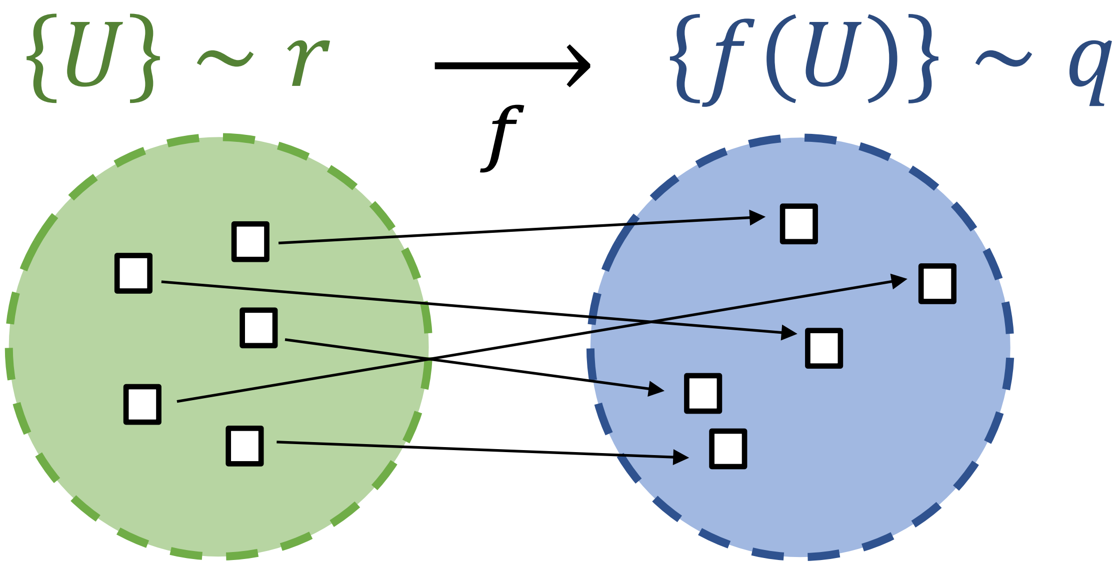


Application: Feynman-Hellmann $\quad S \rightarrow S + \lambda\, O \quad \left.\frac{\partial E_h}{\partial \lambda}\right|_{\lambda=0} \sim \langle h|O|h\rangle$

(Complication: involves fits for E_h, but same idea)

See also [Bacchio 2305.07932]

App 1: Correlated ensembles



Flow an ensemble

$\rightarrow$ $\{U\}$ and $\{f(U)\}$ are correlated

This is useful!

e.g. for noise cancellation in differences

$$\langle O \rangle_p - \langle O \rangle_r$$
$$= \langle wO \rangle_q - \langle O \rangle_r$$
$$= \big\langle w\big(f(U)\big)\, O\big(f(U)\big) - O(U) \big\rangle_{U\sim r}$$

Application: Feynman-Hellmann $\quad S \rightarrow S + \lambda\, O \quad \left.\frac{\partial E_h}{\partial \lambda}\right|_{\lambda=0} \sim \langle h|O|h\rangle$

(Complication: involves fits for E_h, but same idea)

See also [Bacchio 2305.07932]

App 1: Correlated ensembles

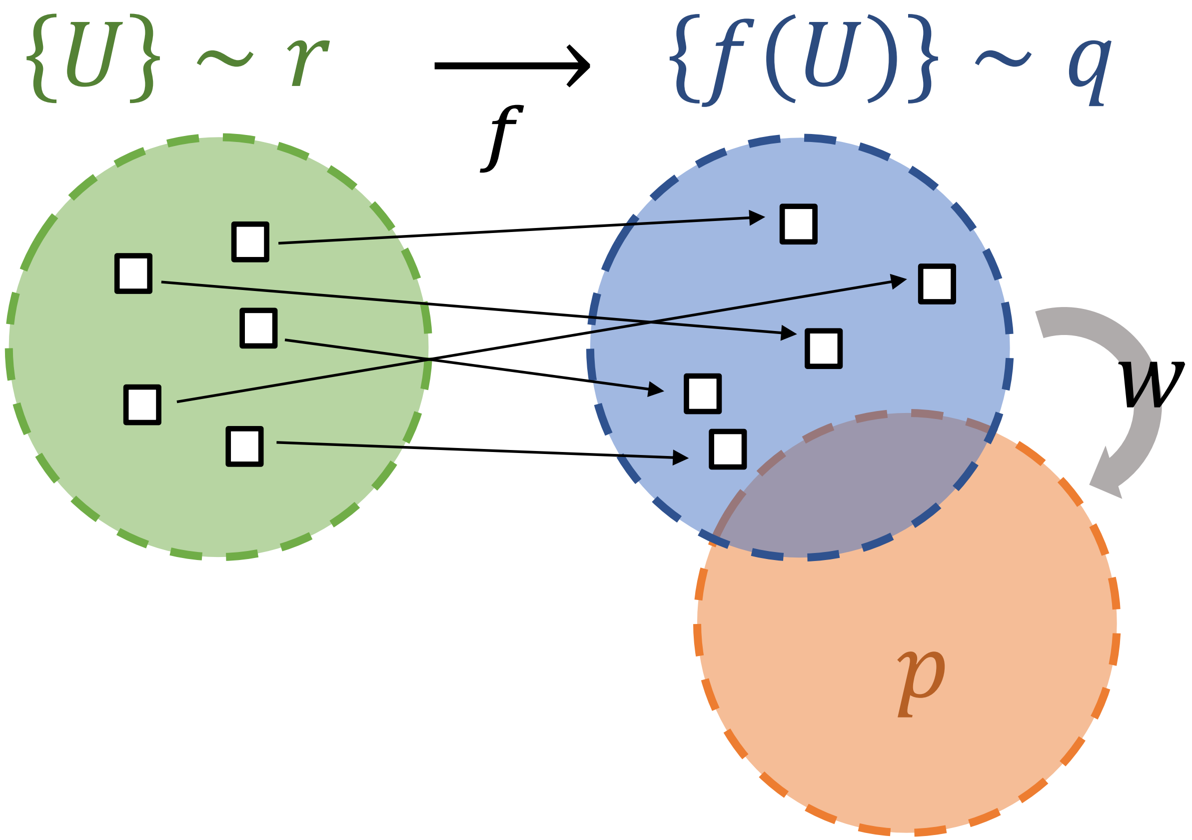
Flow an ensemble

→ $\{U\}$ and $\{f(U)\}$ are correlated

This is useful!

e.g. for noise cancellation in differences

$$\langle O\rangle_p - \langle O\rangle_r$$
$$= \langle wO\rangle_q - \langle O\rangle_r$$
$$= \langle w(f(U))\, O(f(U)) - O(U)\rangle_{U\sim r}$$



Application: Feynman-Hellmann $\qquad S \to S + \lambda\, O \qquad \left.\frac{\partial E_h}{\partial \lambda}\right|_{\lambda=0} \sim \langle h|O|h\rangle$

(Complication: involves fits for E_h, but same idea)

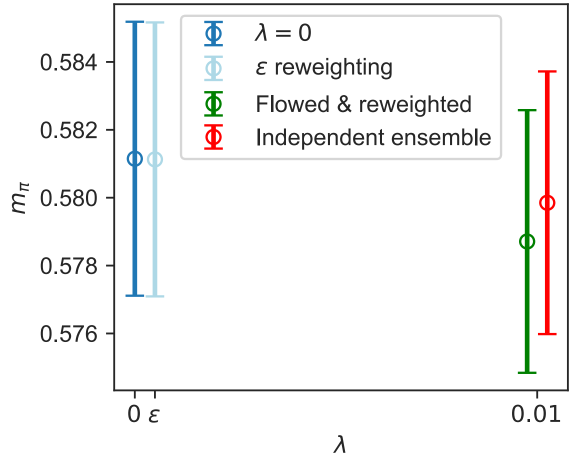
See also [Bacchio 2305.07932]

App 1: Pion $\langle x \rangle_g$ w/ flowed Feynman-Hellmann

[QCDSF-UKQCD 1205.6410]

$$\delta S = -\lambda \frac{\beta}{N_c} \left[\sum_i P_{ti} - \sum_{i<j} P_{ij} \right]$$

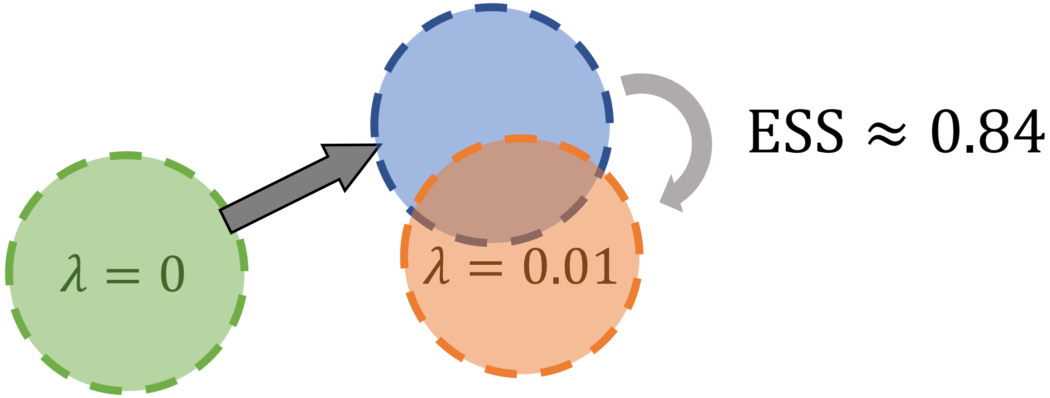
$$\langle x \rangle_g^{\mathrm{lat}} = -\frac{2}{3m} \frac{\partial m}{\partial \lambda}\bigg|_{\lambda=0} \approx -\frac{2}{3m} \frac{1}{\lambda} [m(\lambda) - m(0)]$$

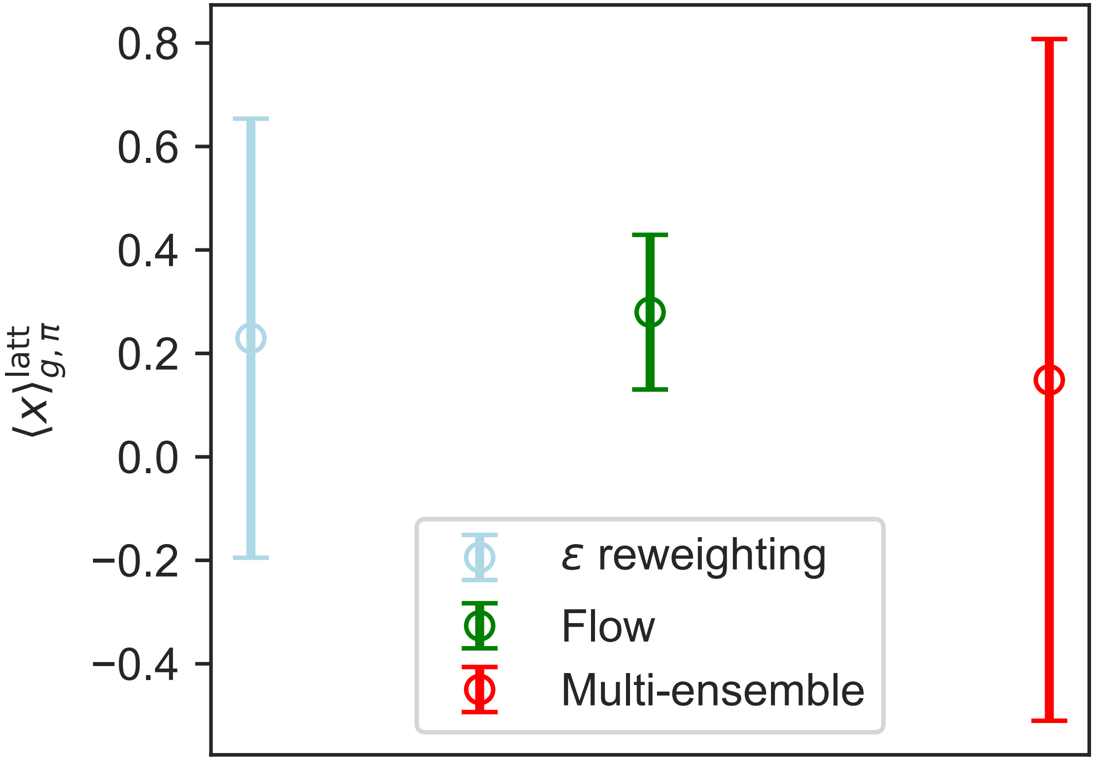


Parameters:

$8^3 \times 16 \quad \beta = 6 \quad \kappa = 0.132$ (quenched)

Flow:





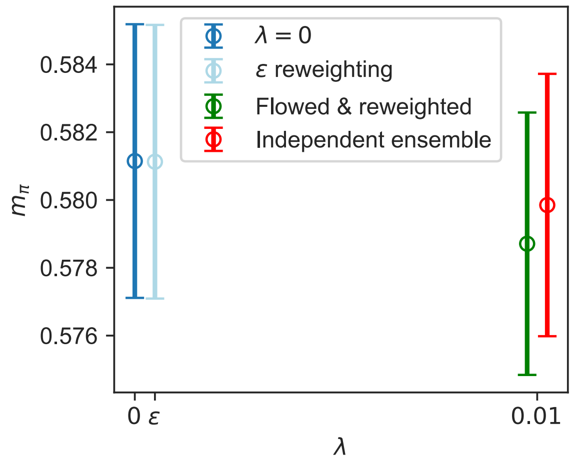
Compute $m(\lambda)$ from $\langle w \, C^{\mathrm{2pt}} \rangle_q$

App 1: Pion $\langle x \rangle_g$ w/ flowed Feynman-Hellmann

[QCDSF-UKQCD 1205.6410]

$$\delta S = -\lambda \frac{\beta}{N_c}\left[\sum_i P_{ti} - \sum_{i<j} P_{ij}\right]$$

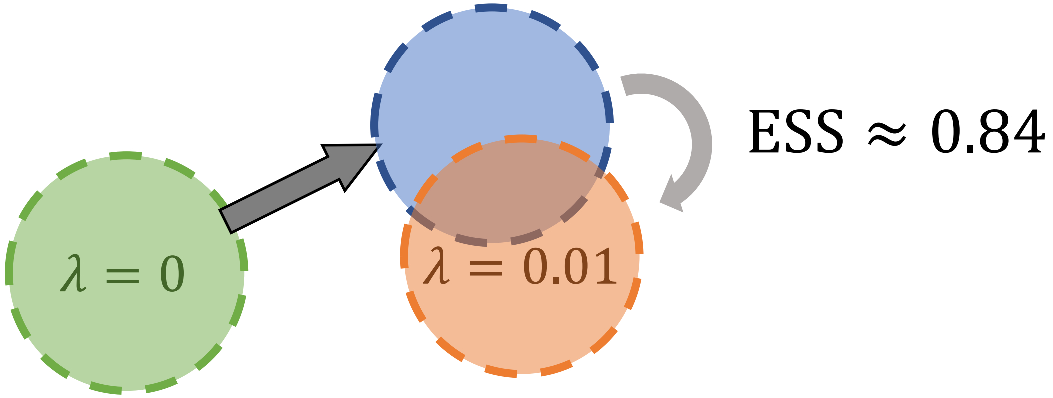
$$\langle x \rangle_g^{\text{lat}} = -\frac{2}{3m}\frac{\partial m}{\partial \lambda}\bigg|_{\lambda=0} \approx -\frac{2}{3m}\frac{1}{\lambda}[m(\lambda) - m(0)]$$

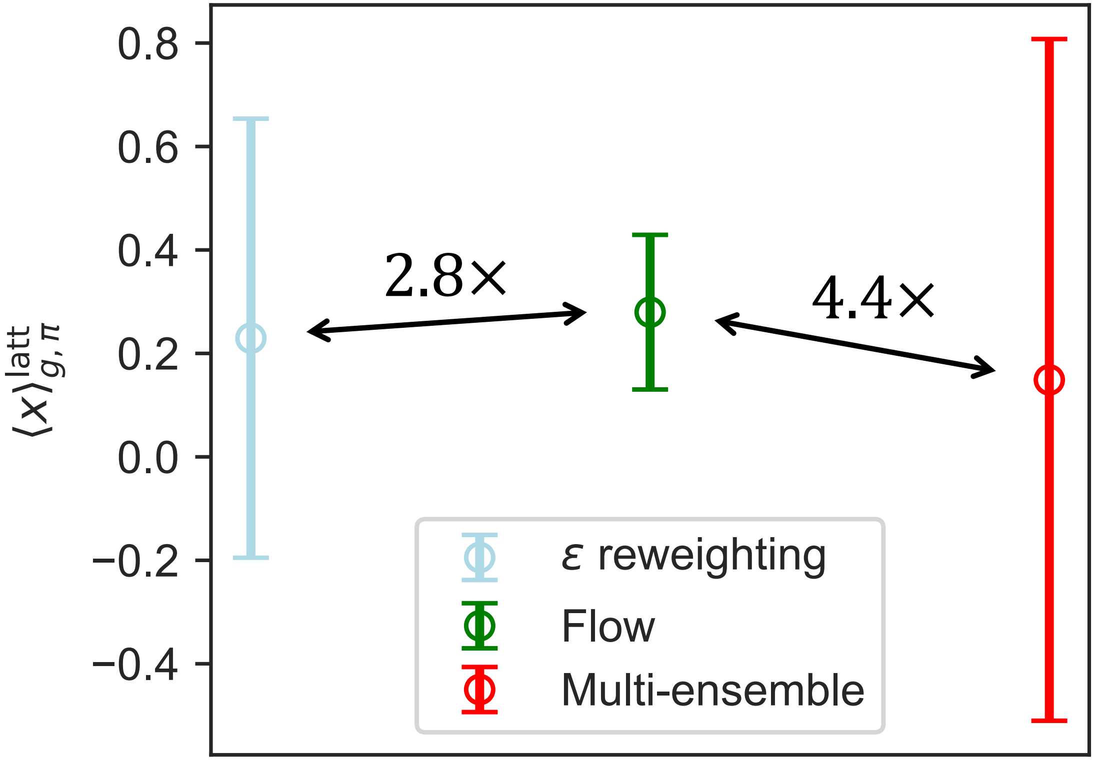


Parameters:

$8^3 \times 16 \quad \beta = 6 \quad \kappa = 0.132$ (quenched)

Flow:



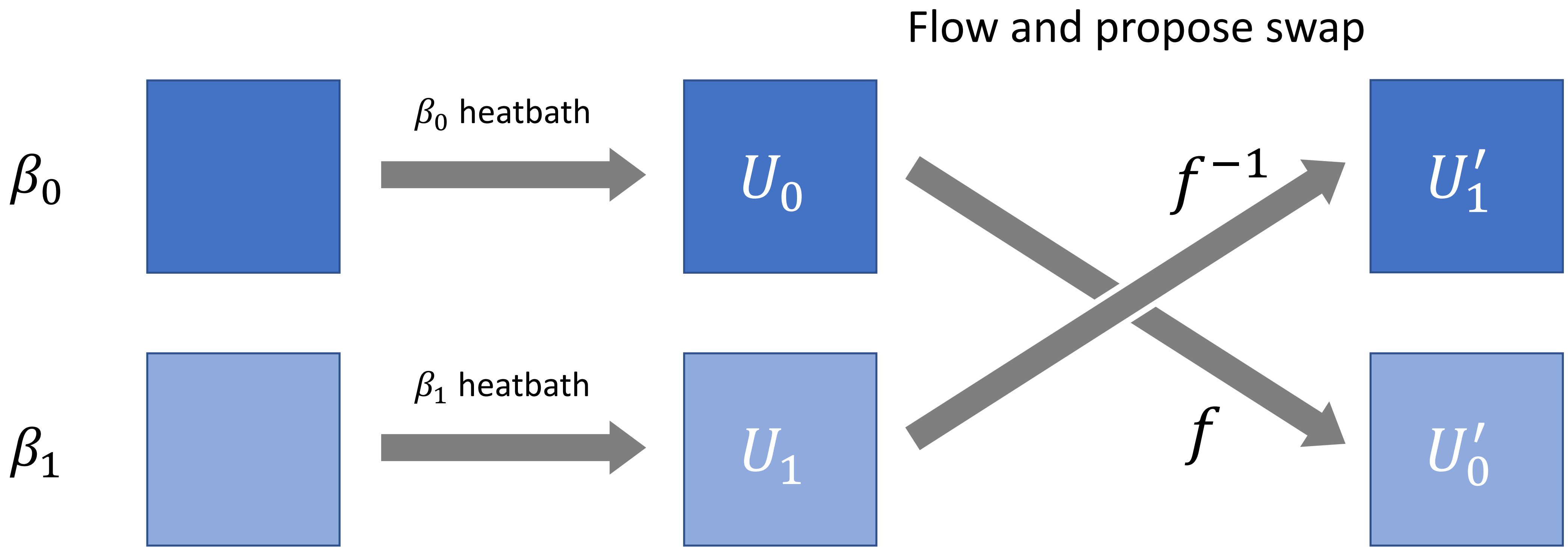


Compute $m(\lambda)$ from $\langle w \, C^{2\text{pt}} \rangle_q$

App 2: Transformed Replica EXchange (T-REX)

(REX a.k.a. parallel tempering) [Invernizzi Krämer Clemente Noé 2210.14104]

Simultaneously sample chains for different targets



$$p_{\text{acc}} = \min\left[1, \frac{p_0(U_1')\, p_1(U_0')}{p_0(U_0)\, p_1(U_1)}\; J_f(U_0)\, J_{f^{-1}}(U_1)\right]$$

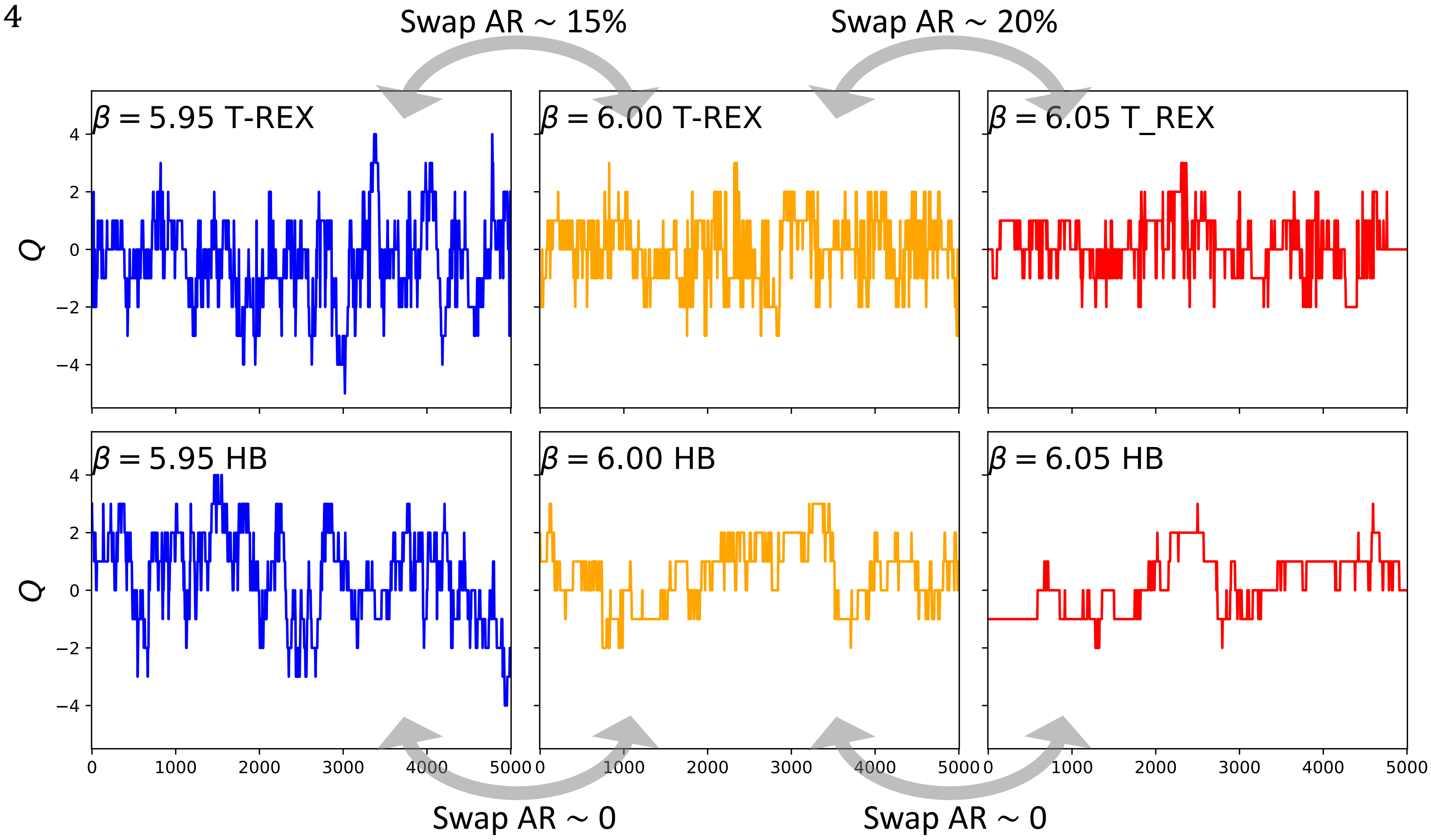
App 2: T-REX Results

Three target βs on 12^4

Two different flows

$$5.95 \leftrightarrow 6$$

$$6 \leftrightarrow 6.05$$



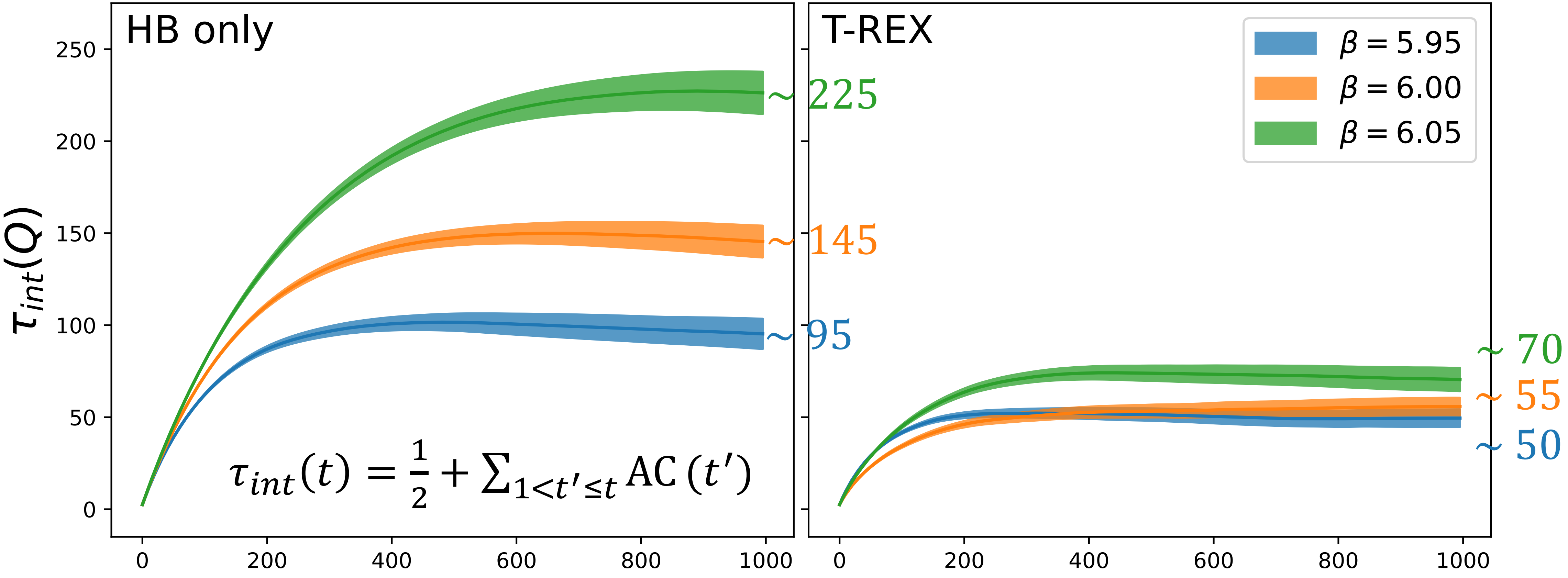
1 step = 5 HB + 2 OR, propose swaps every 5 steps

App 2: T-REX Results

Speed-up for multi-ensemble calculation

T-REX streams correlated (useful!)

Neglecting flow costs!



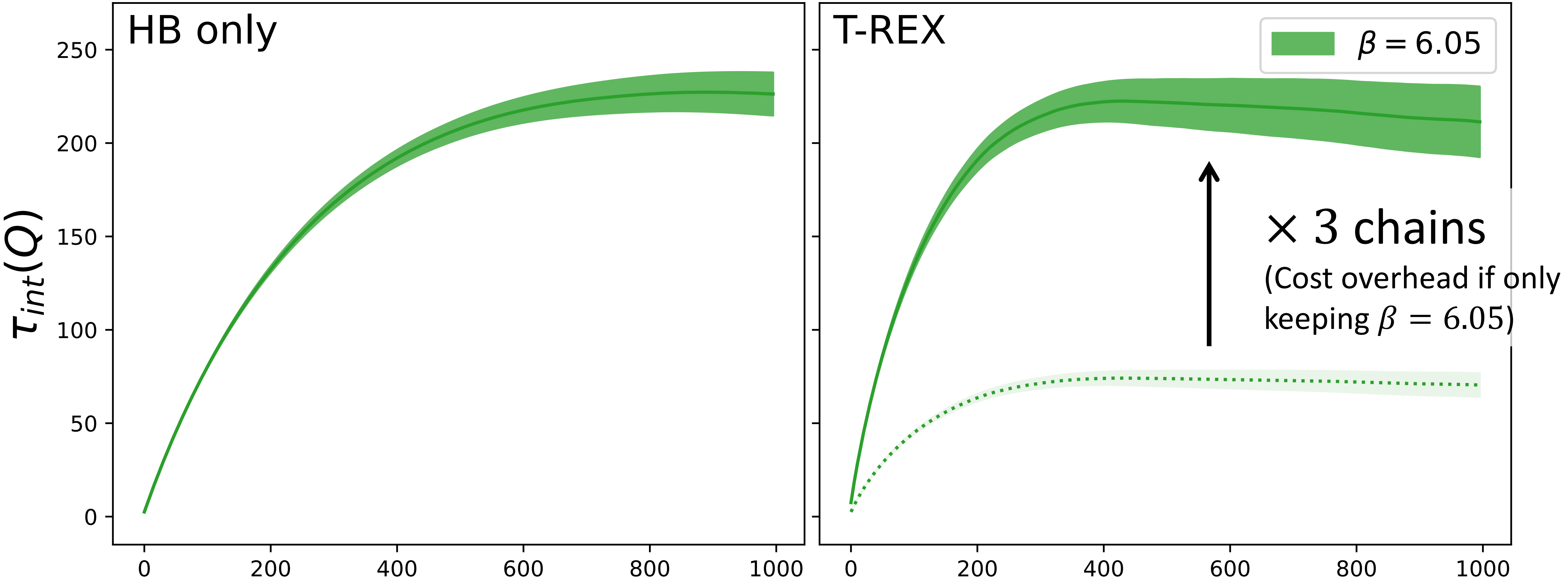
App 2: T-REX Results

Speed-up for multi-ensemble calculation

T-REX streams correlated (useful!)

Break-even for sampling $\beta = 6.05$

} Neglecting flow costs!



App 3: Parallel Tempering on Boundary Conditions (PTBC)

[Hasenbusch 1706.04443] [Bonnano Bonati D'Elia 2012.14000]

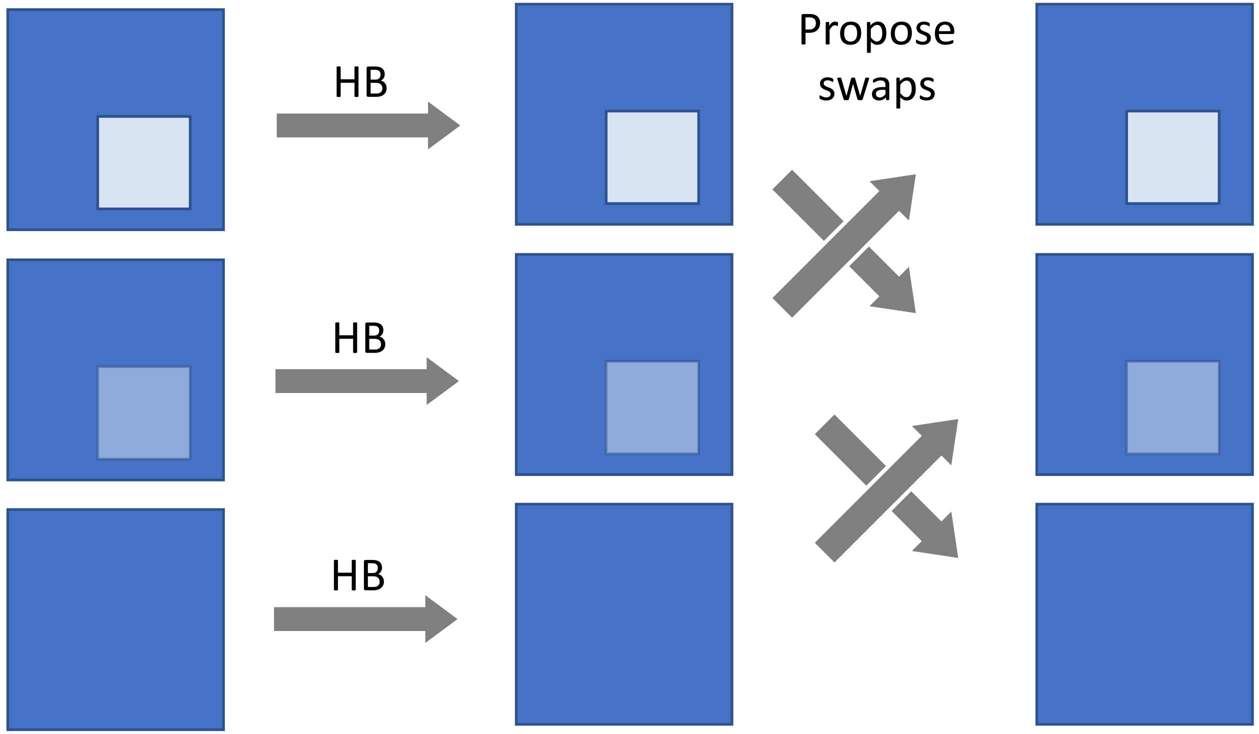
[Boyle T 9:00] [Nada M 14:30]

Introduce localized OBC defect

“Poke a hole in the boundary”

→ Faster topological mixing

Remove defect w/ REX



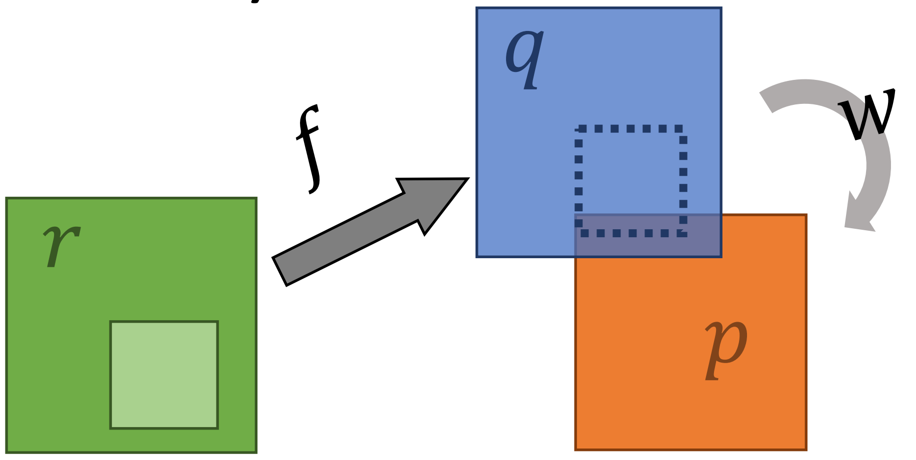
Note: ∃ other options than OBC defects

Other geometries

$\beta_{\text{defect}} > 0$

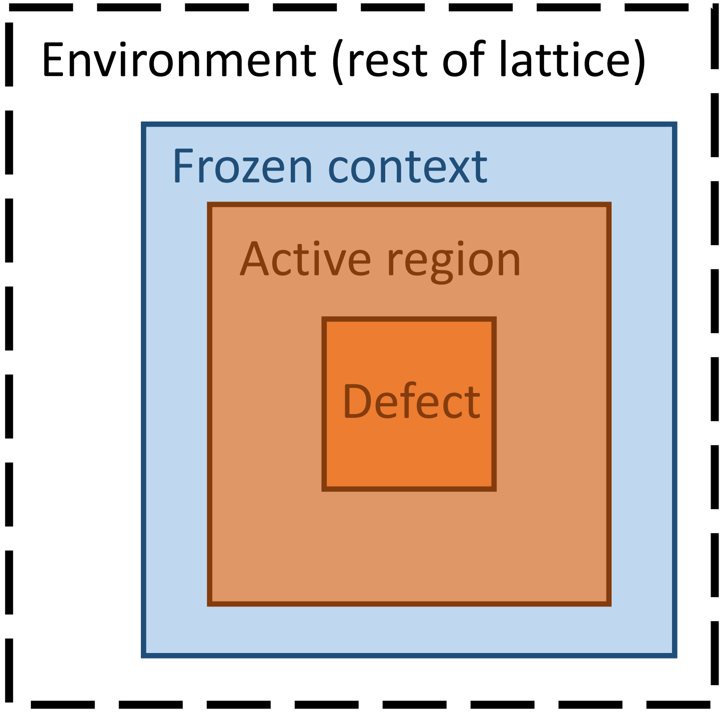
May work better(?)

App 3: Defect Repair Replica EXchange (DR-REX)



Train flow to repair defect

(Or, multiple flows for several steps of partial repair)



Defect has localized physical effects

Flow acts on subvolume

→ No ESS volume scaling

→ Volume-independent computational cost

App 3: DR-REX Results

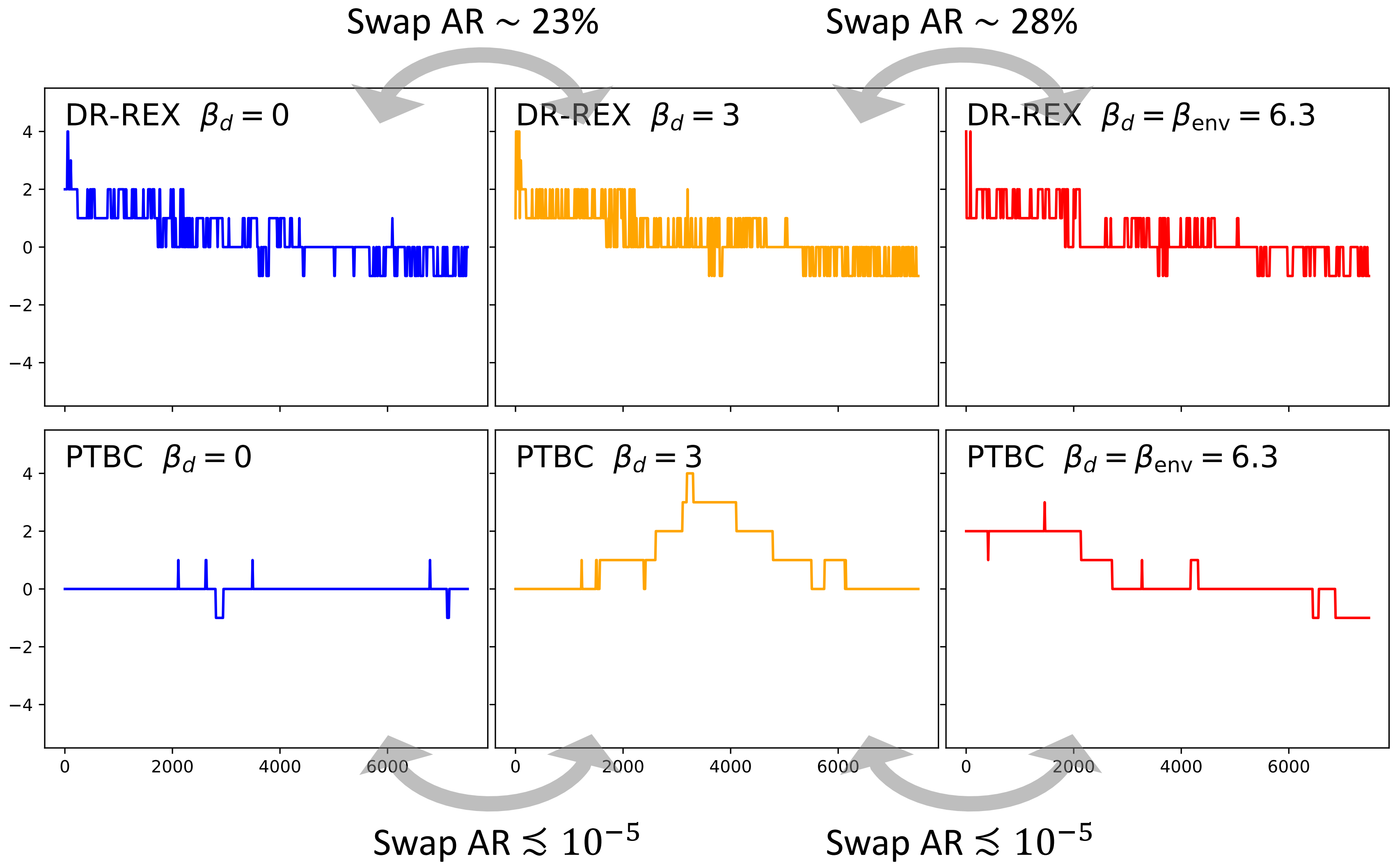
Target: $\beta = 6.3$ on 16^4

2^3 OBC defect

Two flows to repair

$$\beta_d = 0 \rightarrow 3 \rightarrow 6.3$$

Flows act on 8^4 subvolume



Similar swap AR w/o flows requires 7-8 chains

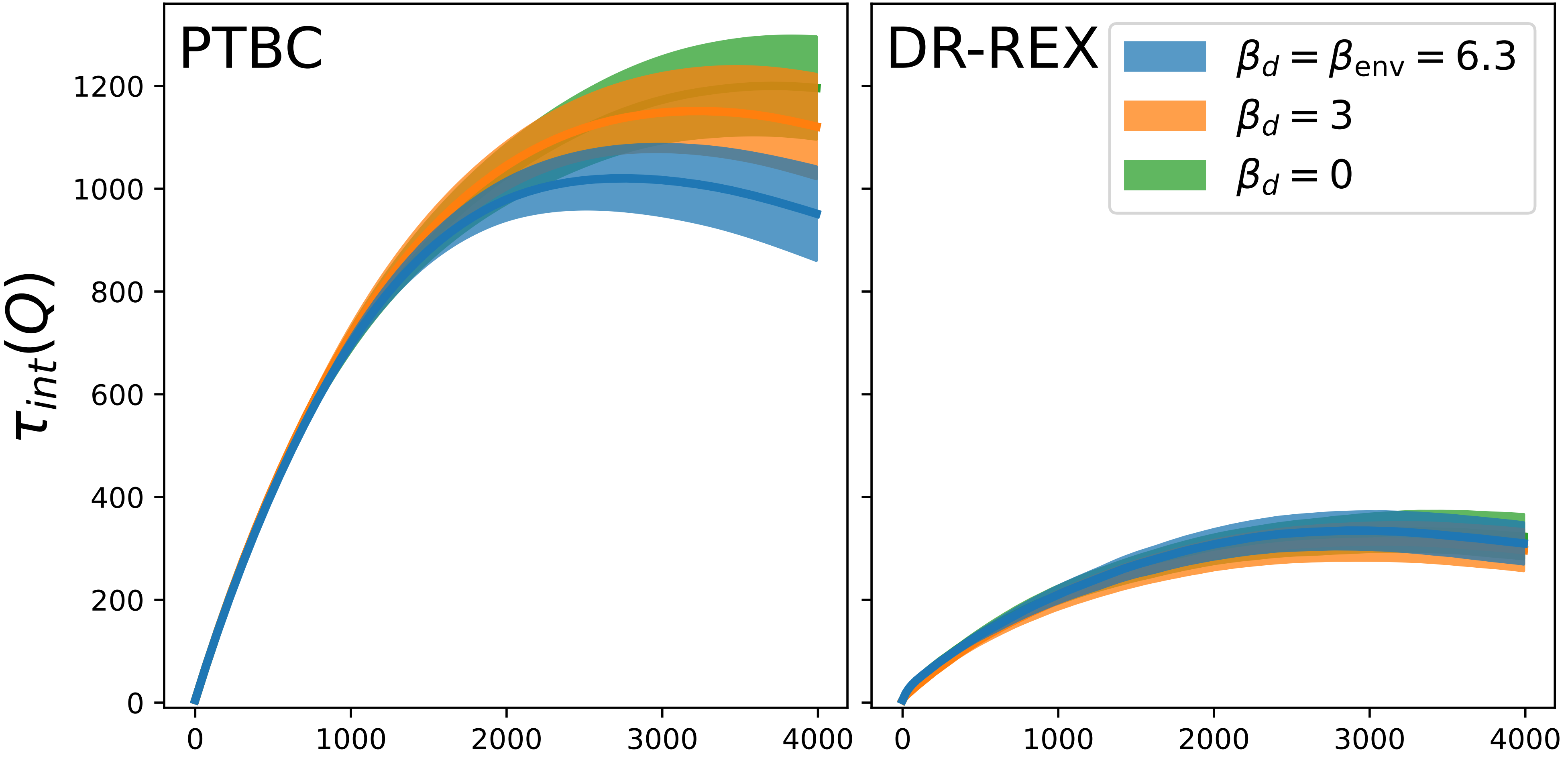
1 step = 1 HB + 5 OR, propose swaps every 10 steps

App 3: DR-REX Results

Target: $\beta = 6.3$ on 16^4

Two flows to repair a 2^3 OBC defect $\beta_d = 0 \to 3 \to 6.3$

Flows act on 8^4 subvolume

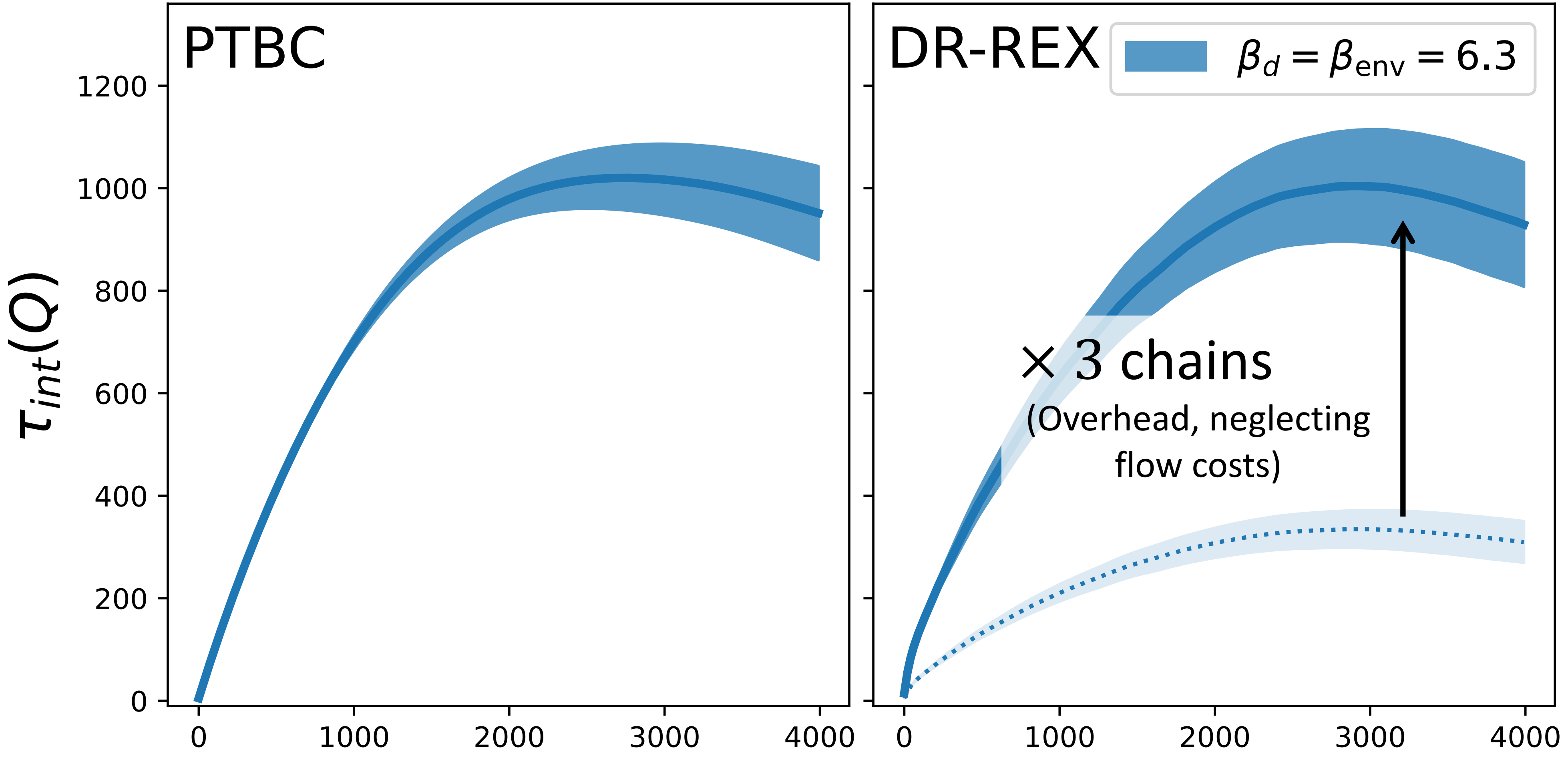


App 3: DR-REX Results

Target: $\beta = 6.3$ on 16^4

Two flows to repair a 2^3 OBC defect $\beta_d = 0 \rightarrow 3 \rightarrow 6.3$

Flows act on 8^4 subvolume

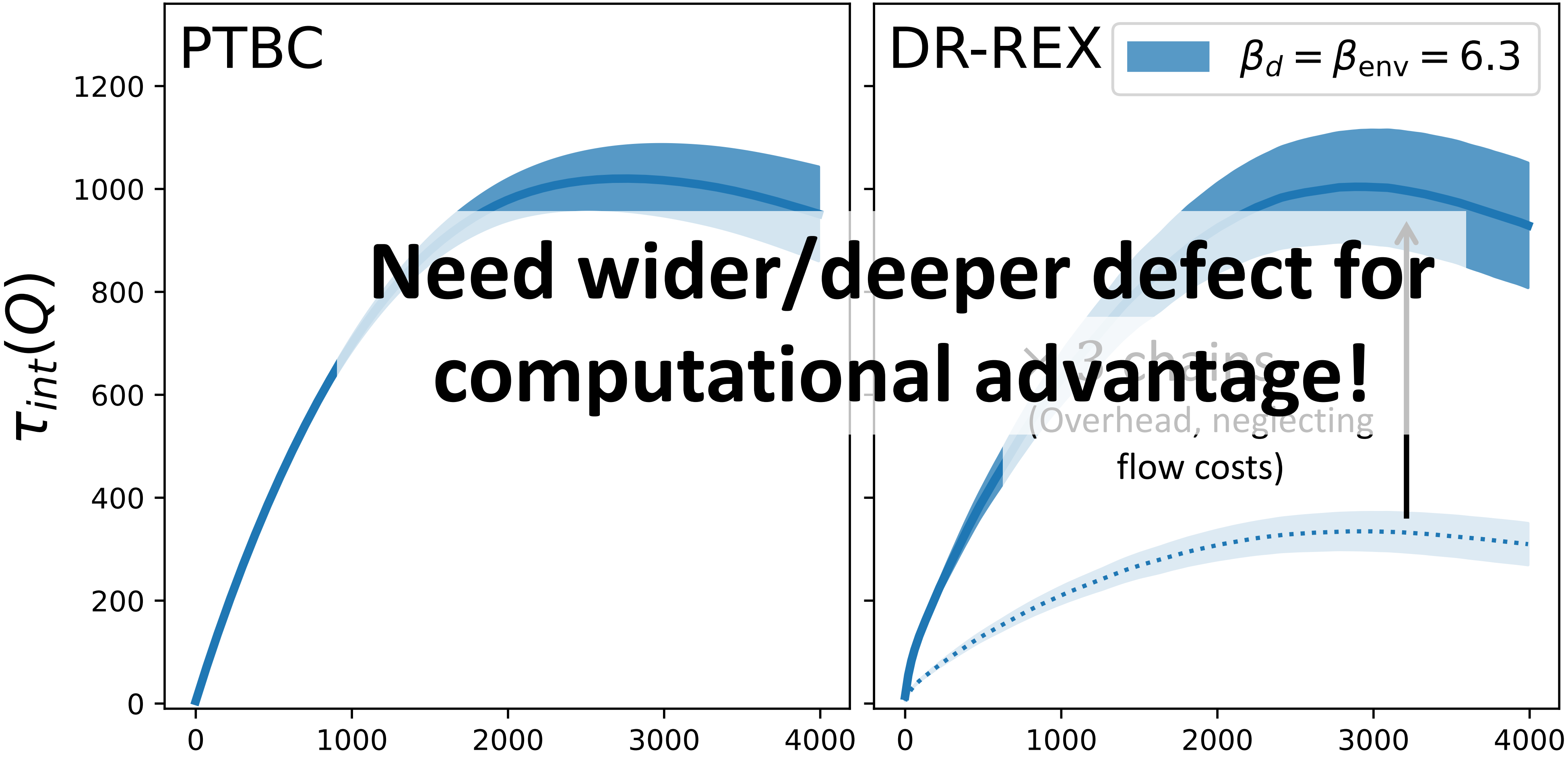


App 3: DR-REX Results

Target: $\beta = 6.3$ on 16^4

Two flows to repair a 2^3 OBC defect $\beta_d = 0 \to 3 \to 6.3$

Flows act on 8^4 subvolume



Need wider/deeper defect for computational advantage!

Closing Thoughts

Flows can give useful correlations between ensembles

- Feynman-Hellmann, continuum/chiral limits, …

REX makes natural use of flows

- Limiting case: direct sampling
  - See also: CRAFT [Matthews Arbel Rezende Doucet 2201.13117] → SNFs [Nada M 14:30]

Straightforward generalizations to (pseudo)fermions / QCD

- Flows for fermions [2106.05934]   PFs for gauge fields [2207.08945]   QCD [2207.08945]

PTBC → DR-REX: very promising

- Flows can make deeper/wider defects practical than w/ REX alone

Stay tuned!