The dependence of observables on action parameters

[arXiv:2307.15406]

Guilherme Catumba
Alberto Ramos; Bryan Zaldívar

IFIC – Valencia

August 2, 2023

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^{N} \sim e^{-S(x;\theta)}$$

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^{N} \sim e^{-S(x;\theta)}$$

$$\langle f(\theta) \rangle_{S_\theta} = \frac{1}{N} \sum_{\alpha=1}^{N} f(x^\alpha; \theta)$$

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^{N} \sim e^{-S(x;\theta)}$$

$$\langle f(\theta) \rangle_{S_\theta} = \frac{1}{N} \sum_{\alpha=1}^{N} f(x^\alpha;\theta)$$

$$\frac{\partial}{\partial \theta_i} \langle f(x;\theta) \rangle_{S_\theta}$$

$$\frac{\partial^2}{\partial \theta_i \partial \theta_j} \langle f(x;\theta) \rangle_{S_\theta}$$

$$\ldots$$

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^{N} \sim e^{-S(x;\theta)}$$

$$\langle f(\theta) \rangle_{S_\theta} = \frac{1}{N} \sum_{\alpha=1}^{N} f(x^\alpha; \theta)$$

$$\frac{\partial}{\partial \theta_i} \langle f(x;\theta) \rangle_{S_\theta}$$

$$\frac{\partial^2}{\partial \theta_i \partial \theta_j} \langle f(x;\theta) \rangle_{S_\theta}$$

$$\ldots$$

Conventional Automatic Differentiation ruled out due to Stochastic elements

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$$

$$\langle f(\theta)\rangle_{S_\theta} = \frac{1}{N}\sum_{\alpha=1}^N f(x^\alpha;\theta)$$

$$\frac{\partial}{\partial\theta_i}\langle f(x;\theta)\rangle_{S_\theta}$$

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\langle f(x;\theta)\rangle_{S_\theta}$$

$$\dots$$

Conventional Automatic Differentiation ruled out due to Stochastic elements

- Dependence on bare couplings
  - Reweighting
  - QCD+QED – ∂_{e^2}

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$$

$$\langle f(\theta)\rangle_{S_\theta} = \frac{1}{N}\sum_{\alpha=1}^{N} f(x^\alpha;\theta)$$

$$\frac{\partial}{\partial\theta_i}\langle f(x;\theta)\rangle_{S_\theta}$$

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\langle f(x;\theta)\rangle_{S_\theta}$$

$$\dots$$

Conventional Automatic Differentiation ruled out due to Stochastic elements

- Dependence on bare couplings
  - Reweighting
  - QCD+QED – ∂_{e^2}
- Optimization problems

Differentiation Problem

$$\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$$

$$\langle f(\theta)\rangle_{S_\theta} = \frac{1}{N}\sum_{\alpha=1}^N f(x^\alpha;\theta)$$

$$\frac{\partial}{\partial\theta_i}\langle f(x;\theta)\rangle_{S_\theta}$$

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\langle f(x;\theta)\rangle_{S_\theta}$$

$$\dots$$

Conventional Automatic Differentiation ruled out due to Stochastic elements

- Dependence on bare couplings
  - Reweighting
  - QCD+QED – ∂_{e^2}
- Optimization problems
- Bayesian inference – Hyperparameters
- ...

[G.C., A. Ramos, B. Zaldivar, arXiv:2307.15406]

Automatic Differentiation – truncated polynomials

- Power series $\mathcal{O}\left(\varepsilon^K\right)$

$$\tilde{x} \equiv x_0 + x_1\varepsilon + x_2\varepsilon^2 + \cdots + x_K\varepsilon^K$$

Automatic Differentiation – truncated polynomials

- Power series $\mathcal{O}\left(\varepsilon^K\right)$

$$\tilde{x} \equiv x_0 + x_1\varepsilon + x_2\varepsilon^2 + \cdots + x_K\varepsilon^K$$

- Define operations & elementary functions – exact at each order

$$\tilde{x}\tilde{y} = x_0y_0 + (x_0y_1 + x_1y_0)\varepsilon + (x_0y_2 + 2x_1y_1 + x_2y_0)\varepsilon^2 + \dots$$
$$\exp(\tilde{x}) = e^{x_0} + e^{x_0}x_1\varepsilon + e^{x_0}(x_1^2/2 + x_2)\varepsilon^2 + \dots$$
$$\dots$$

Automatic Differentiation – truncated polynomials

- Power series $\mathcal{O}\left(\varepsilon^K\right)$

$$\tilde{x} \equiv x_0 + x_1\varepsilon + x_2\varepsilon^2 + \cdots + x_K\varepsilon^K$$

- Define operations & elementary functions – exact at each order

$$\tilde{x}\tilde{y} = x_0y_0 + (x_0y_1 + x_1y_0)\varepsilon + (x_0y_2 + 2x_1y_1 + x_2y_0)\varepsilon^2 + \dots$$
$$\exp(\tilde{x}) = e^{x_0} + e^{x_0}x_1\varepsilon + e^{x_0}(x_1^2/2 + x_2)\varepsilon^2 + \dots$$
$$\dots$$

Evaluate function at $\tilde{x} = x_0 + \varepsilon$ (Taylor theorem)

$$f(\tilde{x}) = f(x_0) + f'(x_0)\varepsilon + \frac{1}{2}f''(x_0)\varepsilon^2$$

Automatic Differentiation – truncated polynomials

- Power series $\mathcal{O}\left(\varepsilon^K\right)$

$$\tilde{x} \equiv x_0 + x_1\varepsilon + x_2\varepsilon^2 + \cdots + x_K\varepsilon^K$$

- Define operations & elementary functions – exact at each order

$$\tilde{x}\tilde{y} = x_0y_0 + (x_0y_1 + x_1y_0)\varepsilon + (x_0y_2 + 2x_1y_1 + x_2y_0)\varepsilon^2 + \ldots$$
$$\exp(\tilde{x}) = e^{x_0} + e^{x_0}x_1\varepsilon + e^{x_0}(x_1^2/2 + x_2)\varepsilon^2 + \ldots$$
$$\ldots$$

Evaluate function at $\tilde{x} = x_0 + \varepsilon$ (Taylor theorem)

$$f(\tilde{x}) = f(x_0) + f'(x_0)\varepsilon + \frac{1}{2}f''(x_0)\varepsilon^2$$

- Chain rule
- Expansion in multiple variables!

deterministic function

Reweighting & Automatic Differentiation

Samples

$$\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$$

Expectation values w.r.t

$$e^{-S(x;\theta')}$$

Reweighting & Automatic Differentiation

Samples

$$\{x^\alpha\}_{\alpha=1}^{N} \sim e^{-S(x;\theta)}$$

Expectation values w.r.t

$$e^{-S(x;\theta')}$$

- Conventional Reweighting

$$\langle f(x)\rangle_{S'} = \frac{\left\langle e^{S-S'} f(x)\right\rangle_S}{\left\langle e^{S-S'}\right\rangle_S}$$

Samples

$$\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$$

Expectation values w.r.t

$$e^{-S(x;\theta')}$$

- Conventional Reweighting

$$\langle f(x) \rangle_{S'} = \frac{\left\langle e^{S-S'} f(x) \right\rangle_S}{\left\langle e^{S-S'} \right\rangle_S}$$

- Introduce truncated polynomials with $\tilde{\theta} = \theta + \varepsilon$

$$w^\alpha = e^{S(x,\theta) - S(x,\tilde{\theta})} = 1 + (\ \dots\ )\varepsilon + (\ \dots\ )\varepsilon^2 + \dots$$

Samples

$$\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$$

Expectation values w.r.t

$$e^{-S(x;\theta')}$$

- Conventional Reweighting

$$\langle f(x)\rangle_{S'} = \frac{\left\langle e^{S-S'} f(x)\right\rangle_S}{\langle e^{S-S'}\rangle_S}$$

- Introduce truncated polynomials with $\tilde{\theta} = \theta + \varepsilon$

$$w^\alpha = e^{S(x,\theta)-S(x,\tilde{\theta})} = 1 + (\ \dots\ )\varepsilon + (\ \dots\ )\varepsilon^2 + \dots$$

- Taylor series coefficients

$$\frac{\sum_\alpha e^{S(x^\alpha,\theta)-S(x^\alpha,\tilde{\theta})} f(x)}{\sum_\alpha e^{S(x^\alpha,\theta)-S(x^\alpha,\tilde{\theta})}} = \sum_{n=0}^K f_n \varepsilon^n, \qquad f_n = \frac{1}{n!}\frac{\partial^n}{\partial\theta^n}\,\langle f(x)\rangle\Big|_\theta$$

1. Fictitious momenta π conjugate to x

$$H(x, \pi) = \frac{1}{2}\pi^2 + S(x; \theta)$$

2. Solve EoM with initial random momenta $\pi(t = 0) \sim \mathcal{N}(0, 1)$

$$\dot{x} = \frac{\partial H}{\partial \pi} = \pi, \qquad \dot{\pi} = -\frac{\partial H}{\partial x}$$

$$(x(0), \pi(0)) \longrightarrow (x(t), \pi(t))$$

3. Metropolis: Acc./Rej. with probability $e^{-\Delta H}$

4. Repeat

1. Fictitious momenta π conjugate to x

$$H(x,\pi) = \frac{1}{2}\pi^2 + S(x;\theta)$$

2. Solve EoM with initial random momenta $\pi(t=0) \sim \mathcal{N}(0,1)$

$$\dot{x} = \frac{\partial H}{\partial \pi} = \pi, \qquad \dot{\pi} = -\frac{\partial H}{\partial x}$$

$$(x(0), \pi(0)) \longrightarrow (x(t), \pi(t))$$

3. Metropolis: Acc./Rej. with probability $e^{-\Delta H}$

4. Repeat

Promote $\theta \longrightarrow \tilde{\theta} = \theta + \varepsilon$
(also $\pi,\ x$)

HMC & Automatic Differentiation – NSPT

1. Fictitious momenta π conjugate to x

$$H(x, \pi) = \frac{1}{2}\pi^2 + S(x; \theta)$$

2. Solve EoM with initial random momenta $\pi(t=0) \sim \mathcal{N}(0, 1)$

$$\dot{x} = \frac{\partial H}{\partial \pi} = \pi, \qquad \dot{\pi} = -\frac{\partial H}{\partial x}$$

$$(x(0), \pi(0)) \longrightarrow (x(t), \pi(t))$$

3. Metropolis: Acc./Rej. with probability $e^{-\Delta H}$

4. Repeat

Promote $\theta \longrightarrow \tilde{\theta} = \theta + \varepsilon$
(also $\pi,\ x$)

Solve EoM at each order

$x \longrightarrow \tilde{x}$

1. Fictitious momenta π conjugate to x

$$H(x,\pi) = \frac{1}{2}\pi^2 + S(x;\theta)$$

2. Solve EoM with initial random momenta $\pi(t=0) \sim \mathcal{N}(0,1)$

$$\dot{x} = \frac{\partial H}{\partial \pi} = \pi, \qquad \dot{\pi} = -\frac{\partial H}{\partial x}$$

$$(x(0),\pi(0)) \longrightarrow (x(t),\pi(t))$$

3. ~~Metropolis: Acc./Rej. with probability $e^{-\Delta H}$~~

4. Repeat

Promote $\theta \longrightarrow \tilde{\theta} = \theta + \varepsilon$
(also $\pi,\ x$)

Solve EoM at each order

$x \longrightarrow \tilde{x}$

ΔH as Taylor series
Cannot Acc./Rej.

HMC & Automatic Differentiation – NSPT

1. Fictitious momenta π conjugate to x

$$H(x,\pi) = \frac{1}{2}\pi^2 + S(x;\theta)$$

2. Solve EoM with initial random momenta $\pi(t=0) \sim \mathcal{N}(0,1)$

$$\dot{x} = \frac{\partial H}{\partial \pi} = \pi, \qquad \dot{\pi} = -\frac{\partial H}{\partial x}$$

$$(x(0),\pi(0)) \longrightarrow (x(t),\pi(t))$$

3. ~~Metropolis: Acc./Rej. with probability $e^{-\Delta H}$~~

4. Repeat

Promote $\theta \longrightarrow \tilde{\theta} = \theta + \varepsilon$
(also $\pi,\ x$)

Solve EoM at each order

$x \longrightarrow \tilde{x}$

ΔH as Taylor series
Cannot Acc./Rej.

MC average as Taylor series

$$\langle f(x) \rangle = \frac{1}{N}\sum_{\alpha} f(\tilde{x}) = \sum_{n=0}^{K} f_n \varepsilon^n$$

Comparison & Change of variables

Reweighting

- $\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$
- Weights w^α take into account dependence on parameters θ

HMC

- $\{\tilde{x}^\alpha\}_{\alpha=1}^N \sim e^{-S(\tilde{x};\tilde{\theta})}$
- Samples carry dependence on the parameters $\tilde{\theta}$

How are these methods related?

Comparison & Change of variables

Reweighting

- $\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$
- Weights w^α take into account dependence on parameters θ

HMC

- $\{\tilde{x}^\alpha\}_{\alpha=1}^N \sim e^{-S(\tilde{x};\tilde{\theta})}$
- Samples carry dependence on the parameters $\tilde{\theta}$

How are these methods related?

Toy model: $$p_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \qquad \{x^\alpha\} \sim p_{\sigma^*=1},$$

Comparison & Change of variables

Reweighting

- $\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$
- Weights w^α take into account dependence on parameters θ

HMC

- $\{\tilde{x}^\alpha\}_{\alpha=1}^N \sim e^{-S(\tilde{x};\tilde{\theta})}$
- Samples carry dependence on the parameters $\tilde{\theta}$

How are these methods related?

Toy model: $p_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \qquad \{x^\alpha\} \sim p_{\sigma^*=1},$

$\sigma^* = 1 \longrightarrow \sigma \qquad w^\alpha(x) = x - \text{dependent}$

Comparison & Change of variables

Reweighting

- $\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$
- Weights w^α take into account dependence on parameters θ

HMC

- $\{\tilde{x}^\alpha\}_{\alpha=1}^N \sim e^{-S(\tilde{x};\tilde{\theta})}$
- Samples carry dependence on the parameters $\tilde{\theta}$

How are these methods related?

Toy model: $p_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$, $\{x^\alpha\} \sim p_{\sigma^*=1}$,

$\sigma^* = 1 \longrightarrow \sigma$ $w^\alpha(x) = x - \text{dependent}$

Transformation: $y^\alpha = \sigma x^\alpha$, $\{y^\alpha\} \sim p_\sigma$

Comparison & Change of variables

Reweighting

- $\{x^\alpha\}_{\alpha=1}^N \sim e^{-S(x;\theta)}$
- Weights w^α take into account dependence on parameters θ

HMC

- $\{\tilde{x}^\alpha\}_{\alpha=1}^N \sim e^{-S(\tilde{x};\tilde{\theta})}$
- Samples carry dependence on the parameters $\tilde{\theta}$

How are these methods related?

Toy model: $p_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$, $\{x^\alpha\} \sim p_{\sigma^*=1}$,

$\sigma^* = 1 \longrightarrow \sigma$ $w^\alpha(x) = x -$ dependent

Transformation: $y^\alpha = \sigma x^\alpha$, $\{y^\alpha\} \sim p_\sigma$

$\sigma^* = 1 \longrightarrow \sigma$ $w^\alpha(y) = y -$ independent 'No Reweighting'

Comparison & Change of variables

$$y^\alpha = \sigma x^\alpha \qquad \tilde{y}^\alpha = x^\alpha + x^\alpha \varepsilon \qquad \varepsilon = \sigma - 1$$

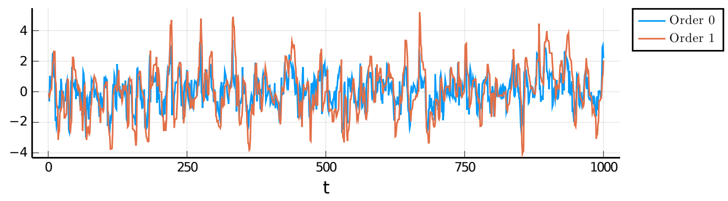
Comparison & Change of variables

$$y^\alpha = \sigma x^\alpha \qquad \tilde{y}^\alpha = x^\alpha + x^\alpha \varepsilon \qquad \varepsilon = \sigma - 1$$

What about HMC?
Equations of motion

$$\ddot{x}_0 = -\frac{x_0}{\sigma^2}, \qquad \langle x_0 \rangle$$

$$\ddot{x}_1 = -\frac{x_1}{\sigma^2} + 2\frac{x_0}{\sigma^3}, \qquad \langle x_1 \rangle = \langle x_0 \rangle$$

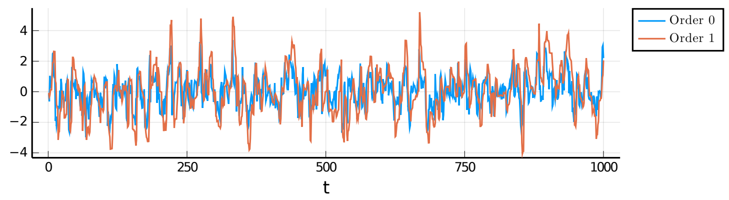


$$y^\alpha = \sigma x^\alpha \qquad \tilde{y}^\alpha = x^\alpha + x^\alpha \varepsilon \qquad \varepsilon = \sigma - 1$$

What about HMC?
Equations of motion

$$\ddot{x}_0 = -\frac{x_0}{\sigma^2}, \qquad \langle x_0 \rangle$$

$$\ddot{x}_1 = -\frac{x_1}{\sigma^2} + 2\frac{x_0}{\sigma^3}, \qquad \langle x_1 \rangle = \langle x_0 \rangle$$



- Hamiltonian method finds the change of variables $x \to \tilde{y}$ that lead to constant reweighting factors

4D Scalar theory on the lattice: $\lambda - \phi^4$

$$S_{\text{latt}}(\hat{\phi}; m, \lambda) = \sum_x \left\{ \frac{1}{2} \sum_\mu [\phi(x+\mu) - \phi(x)]^2 + \frac{m^2}{2} \phi^2(x) + \lambda \phi^4(x) \right\}$$

$$\left\langle \phi^2(x) \right\rangle$$

$$\langle s(x; m, \lambda) \rangle = \frac{1}{2} \langle [\phi(x+\mu) - \phi(x)]^2 \rangle + \frac{m^2}{2} \langle \phi^2(x) \rangle + \lambda \langle \phi^4(x) \rangle \,.$$

Connected VS Disconnected contributions

$$\partial \langle \mathcal{O} \rangle = \langle \partial \mathcal{O} \rangle + [\langle \mathcal{O} \rangle \langle \partial S \rangle - \langle \mathcal{O} \partial S \rangle]$$

4D Scalar theory on the lattice: $\lambda - \phi^4$

$$S_{\text{latt}}(\hat{\phi}; m, \lambda) = \sum_x \left\{ \frac{1}{2} \sum_\mu [\phi(x+\mu) - \phi(x)]^2 + \frac{m^2}{2} \phi^2(x) + \lambda \phi^4(x) \right\}$$

$$\left\langle \phi^2(x) \right\rangle$$

$$\langle s(x; m, \lambda) \rangle = \frac{1}{2} \langle [\phi(x+\mu) - \phi(x)]^2 \rangle + \frac{m^2}{2} \langle \phi^2(x) \rangle + \lambda \langle \phi^4(x) \rangle \,.$$

Connected VS Disconnected contributions

$$\partial \langle \mathcal{O} \rangle = \langle \partial \mathcal{O} \rangle + [\langle \mathcal{O} \rangle \langle \partial S \rangle - \langle \mathcal{O} \partial S \rangle]$$

Dependence on the couplings:

$$\frac{\partial}{\partial m^2} \qquad \frac{\partial}{\partial \lambda} \qquad \frac{\partial^2}{\partial m^2 \partial \lambda}$$

$\lambda - \phi^4$: Reweighting VS Hamiltonian Expansion

			λ				
			0.0	0.1	0.2	0.3	0.4
$\langle\phi^2\rangle$	$\partial_{\hat{m}^2}$	RW	-0.0428(20)	-0.0328(14)	-0.0270(13)	-0.0241(12)	-0.0220(11)
		HAD	-0.042526(41)	-0.030880(14)	-0.026273(10)	-0.0233672(82)	-0.0212721(72)
	∂_λ	RW	-0.0779(22)	-0.05227(94)	-0.04370(89)	-0.03534(61)	-0.03169(50)
		HAD	-0.077816(79)	-0.052499(24)	-0.042218(19)	-0.035830(14)	-0.031323(11)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	0.43(43)	0.03(16)	0.16(14)	-0.10(11)	0.116(77)
		HAD	0.2733(22)	0.061593(99)	0.035082(69)	0.024240(42)	0.018263(31)
$\langle s\rangle$	$\partial_{\hat{m}^2}$	RW	-0.0025(42)	-0.0006(34)	0.0027(35)	0.0028(36)	0.0057(32)
		HAD	-0.000003(22)	0.002623(16)	0.004218(20)	0.005397(14)	0.006265(16)
	∂_λ	RW	-0.0765(46)	-0.0567(26)	-0.0538(25)	-0.0447(23)	-0.0400(17)
		HAD	-0.069774(49)	-0.057738(34)	-0.050128(40)	-0.044530(27)	-0.040250(27)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	-1.9(2.8)	-0.16(39)	0.36(43)	-0.43(32)	0.16(26)
		HAD	0.038864(96)	0.019197(66)	0.013405(69)	0.010126(50)	0.007860(47)

TableL^4 lattice with $L/a = 32$ and $\hat{m}^2 = 0.05$.

$\lambda - \phi^4$: Reweighting VS Hamiltonian Expansion

			λ				
			0.0	0.1	0.2	0.3	0.4
$\langle \phi^2 \rangle$	$\partial_{\hat{m}^2}$	RW	-0.0428(20)	-0.0328(14)	-0.0270(13)	-0.0241(12)	-0.0220(11)
		HAD	-0.042526(41)	-0.030880(14)	-0.026273(10)	-0.0233672(82)	-0.0212721(72)
	∂_{λ}	RW	-0.0779(22)	-0.05227(94)	-0.04370(89)	-0.03534(61)	-0.03169(50)
		HAD	-0.077816(79)	-0.052499(24)	-0.042218(19)	-0.035830(14)	-0.031323(11)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	0.43(43)	0.03(16)	0.16(14)	-0.10(11)	0.116(77)
		HAD	0.2733(22)	0.061593(99)	0.035082(69)	0.024240(42)	0.018263(31)
$\langle s \rangle$	$\partial_{\hat{m}^2}$	RW	-0.0025(42)	-0.0006(34)	0.0027(35)	0.0028(36)	0.0057(32)
		HAD	-0.000003(22)	0.002623(16)	0.004218(20)	0.005397(14)	0.006265(16)
	∂_{λ}	RW	-0.0765(46)	-0.0567(26)	-0.0538(25)	-0.0447(23)	-0.0400(17)
		HAD	-0.069774(49)	-0.057738(34)	-0.050128(40)	-0.044530(27)	-0.040250(27)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	-1.9(2.8)	-0.16(39)	0.36(43)	-0.43(32)	0.16(26)
		HAD	0.038864(96)	0.019197(66)	0.013405(69)	0.010126(50)	0.007860(47)

TableL^4 lattice with $L/a = 32$ and $\hat{m}^2 = 0.05$.

- Same statistics
- Hamiltonian expansion is 100x! more precise
- Reweighting shows no signal for cross derivatives

Method Comparison

Reweighting+AD

- Re-utilize conventional samples $\{x^\alpha\}_{\alpha=1}^N$
- Weighted Expectation values

$$\langle f(x)\rangle = \frac{\sum_\alpha \tilde{w}(x^\alpha) f(x^\alpha)}{\sum_\alpha \tilde{w}(x^\alpha)}$$

- Noisy disconnected contributions – larger errors
- Rules out cross derivatives

HMC+AD

- Samples are power series $\{\tilde{x}^\alpha\}_{\alpha=1}^N$ – carry information about the dependence on parameters
- Normal expectation value

$$\langle f(x)\rangle = \sum_\alpha f(\tilde{x}^\alpha)$$

- HMC finds exact reparametrization
- Only connected contributions

Can we improve reweighting?

Standard RW:

$$w^\alpha = e^{S(x^\alpha;m,\lambda)-S(x^\alpha;\tilde{m},\tilde{\lambda})}$$

- Weights w^α are potentially large
- Samples $\{x^\alpha\}_{\alpha=1}^N$ do not describe ∂_{m^2}
- Large variance

Can we improve reweighting?

Standard RW:

$$w^\alpha = e^{S(x^\alpha;m,\lambda)-S(x^\alpha;\tilde{m},\tilde{\lambda})}$$

- Weights w^α are potentially large
- Samples $\{x^\alpha\}_{\alpha=1}^N$ do not describe ∂_{m^2}
- Large variance

Change variables first!

$$x \longrightarrow \tilde{y} = x + f(x)\varepsilon$$

Reweighting after change of variables

$$w^\alpha = e^{S(x^\alpha;m,\lambda)-S(\tilde{y}^\alpha;\tilde{m},\tilde{\lambda})-\log\left|\tilde{J}\right|}$$

Find transformation that gives constant weights

[S. Bacchio, arXiv:2305.07932] – similar concept with Gradient Flow transformation

Free Theory case

Lattice action, momentum space

$$S_{\text{latt}}(\phi_p; m) = \sum_p \phi_p^* \left[\sum_\mu \hat{p}_\mu^2 + m^2 \right] \phi_p, \qquad \hat{p} = 2\sin(ap/2)$$

- Transformation $\tilde{\varphi}_p = \phi_p + f_p \varepsilon$ (expansion around m)

Free Theory case

Lattice action, momentum space

$$S_{\text{latt}}(\phi_p; m) = \sum_p \phi_p^* \left[\sum_\mu \hat{p}_\mu^2 + m^2 \right] \phi_p, \qquad \hat{p} = 2\sin(ap/2)$$

- Transformation $\tilde{\varphi}_p = \phi_p + f_p \varepsilon$ (expansion around m)
- RW factor

$$S(\phi_p; m) - S(\tilde{\varphi}_p; \tilde{m}) = -\varepsilon \left[(\hat{p}^2 + \tilde{m}^2)(\phi_p^* f_p + \phi_p f_p^*) + \phi_p^* \phi_p \right]$$

Free Theory case

Lattice action, momentum space

$$S_{\text{latt}}(\phi_p; m) = \sum_p \phi_p^* \left[\sum_\mu \hat{p}_\mu^2 + m^2 \right] \phi_p, \qquad \hat{p} = 2\sin(ap/2)$$

- Transformation $\tilde{\varphi}_p = \phi_p + f_p \varepsilon$ (expansion around m)
- RW factor

$$S(\phi_p; m) - S(\tilde{\varphi}_p; \tilde{m}) = -\varepsilon \left[(\hat{p}^2 + \tilde{m}^2)(\phi_p^* f_p + \phi_p f_p^*) + \phi_p^* \phi_p \right]$$

- $-\Delta S = 0$ with

$$f_p = -\frac{\phi_p}{2(\hat{p}^2 + m^2)}$$

Lattice action, momentum space

$$S_{\text{latt}}(\phi_p; m) = \sum_p \phi_p^* \left[\sum_\mu \hat{p}_\mu^2 + m^2 \right] \phi_p, \qquad \hat{p} = 2\sin(ap/2)$$

- Transformation $\tilde{\varphi}_p = \phi_p + f_p \varepsilon$ (expansion around m)
- RW factor

$$S(\phi_p; m) - S(\tilde{\varphi}_p; \tilde{m}) = -\varepsilon \left[(\hat{p}^2 + \tilde{m}^2)(\phi_p^* f_p + \phi_p f_p^*) + \phi_p^* \phi_p \right]$$

- $-\Delta S = 0$ with

$$f_p = -\frac{\phi_p}{2(\hat{p}^2 + m^2)}$$

- Constant weights (from Jacobian only)

$$w(\phi_p; m) = \prod_p -\frac{1}{2(\hat{p}^2 + m^2)}$$

Reweighting, Transformed Reweighting & HMC

			λ				
			0.0	0.1	0.2	0.3	0.4
$\langle \phi^2 \rangle$	$\partial_{\hat{m}^2}$	RW	-0.0428(20)	-0.0328(14)	-0.0270(13)	-0.0241(12)	-0.0220(11)
		TRW	-0.042528(30)	-0.03069(32)	-0.02604(37)	-0.02328(54)	-0.02079(56)
		HAD	-0.042526(41)	-0.030880(14)	-0.026273(10)	-0.0233672(82)	-0.0212721(72)
	∂_{λ}	RW	-0.0779(22)	-0.05227(94)	-0.04370(89)	-0.03534(61)	-0.03169(50)
		HAD	-0.077816(79)	-0.052499(24)	-0.042218(19)	-0.035830(14)	-0.031323(11)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	0.43(43)	0.03(16)	0.16(14)	-0.10(11)	0.116(77)
		HAD	0.2733(22)	0.061593(99)	0.035082(69)	0.024240(42)	0.018263(31)
$\langle s \rangle$	$\partial_{\hat{m}^2}$	RW	-0.0025(42)	-0.0006(34)	0.0027(35)	0.0028(36)	0.0057(32)
		TRW	$3(28)\times10^{-19}$	0.00344(85)	0.0052(11)	0.0052(17)	0.0072(15)
		HAD	-0.000003(22)	0.002623(16)	0.004218(20)	0.005397(14)	0.006265(16)
	∂_{λ}	RW	-0.0765(46)	-0.0567(26)	-0.0538(25)	-0.0447(23)	-0.0400(17)
		HAD	-0.069774(49)	-0.057738(34)	-0.050128(40)	-0.044530(27)	-0.040250(27)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	-1.9(2.8)	-0.16(39)	0.36(43)	-0.43(32)	0.16(26)
		HAD	0.038864(96)	0.019197(66)	0.013405(69)	0.010126(50)	0.007860(47)

Reweighting, Transformed Reweighting & HMC

			λ				
			0.0	0.1	0.2	0.3	0.4
$\langle\phi^2\rangle$	$\partial_{\hat{m}^2}$	RW	-0.0428(20)	-0.0328(14)	-0.0270(13)	-0.0241(12)	-0.0220(11)
		TRW	-0.042528(30)	-0.03069(32)	-0.02604(37)	-0.02328(54)	-0.02079(56)
		HAD	-0.042526(41)	-0.030880(14)	-0.026273(10)	-0.0233672(82)	-0.0212721(72)
	∂_λ	RW	-0.0779(22)	-0.05227(94)	-0.04370(89)	-0.03534(61)	-0.03169(50)
		HAD	-0.077816(79)	-0.052499(24)	-0.042218(19)	-0.035830(14)	-0.031323(11)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	0.43(43)	0.03(16)	0.16(14)	-0.10(11)	0.116(77)
		HAD	0.2733(22)	0.061593(99)	0.035082(69)	0.024240(42)	0.018263(31)
$\langle s\rangle$	$\partial_{\hat{m}^2}$	RW	-0.0025(42)	-0.0006(34)	0.0027(35)	0.0028(36)	0.0057(32)
		TRW	$3(28)\times10^{-19}$	0.00344(85)	0.0052(11)	0.0052(17)	0.0072(15)
		HAD	-0.000003(22)	0.002623(16)	0.004218(20)	0.005397(14)	0.006265(16)
	∂_λ	RW	-0.0765(46)	-0.0567(26)	-0.0538(25)	-0.0447(23)	-0.0400(17)
		HAD	-0.069774(49)	-0.057738(34)	-0.050128(40)	-0.044530(27)	-0.040250(27)
	$\partial^2_{\hat{m}^2,\lambda}$	RW	-1.9(2.8)	-0.16(39)	0.36(43)	-0.43(32)	0.16(26)
		HAD	0.038864(96)	0.019197(66)	0.013405(69)	0.010126(50)	0.007860(47)

- Exact Transformation $\lambda = 0$
- Precision degrades with λ
- Improved w.r.t. reweighting even for large couplings (4D!)

Second Derivative – Reweighting VS Transformed Reweighting

			λ				
			0.0	0.1	0.2	0.3	0.4
$\langle\phi^2\rangle$	$\partial^2_{\hat{m}^4}$	RW	0.11(25)	0.13(15)	0.17(12)	0.16(13)	0.15(11)
		TRW	0.0669(34)	0.0297(97)	0.018(13)	0.022(27)	0.030(29)
$\langle s\rangle$	$\partial^2_{\hat{m}^4}$	RW	0.59(56)	0.23(35)	0.51(34)	0.29(37)	0.54(32)
		TRW	0.0015(74)	0.066(28)	-0.018(40)	0.066(73)	0.103(83)

TableL^4 lattice with $L/a = 32$ and $\hat{m}^2 = 0.05$.

- Transformed reweighting allows higher order derivatives

Conclusions & Outlook

- AD is an important tool – no simple extension for stochastic processes
- HMC (or similar)
  - NSPT inspired solution
  - Model dependent convergence
  - Good precision
- Reweighting for generic MC process
  - Reutilize samples (sampler agnostic)
  - Larger variance – worse precision
- Apt for arbitrarily complicated observables
- RW open for improvement: ML techniques to find transformation $f(x)$
  - Reduce variance from disconnected contributions
  - LQCD state of the art use simple reweighting