Neural Control Variates at LO and NLO

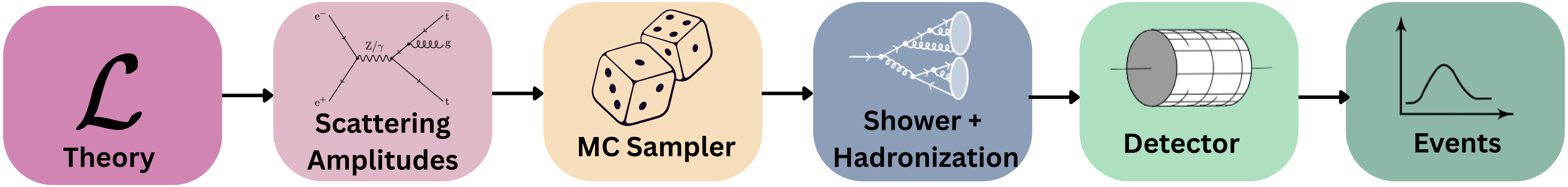
ML4Jets 2026 Vienna

Sophia Vent | Institute for Theoretical Physics Heidelberg
Theo Heimel, Tilman Plehn, Rebecca Revelli, Ramon Winterhalder
vent@thphys.uni-heidelberg.de

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Motivation

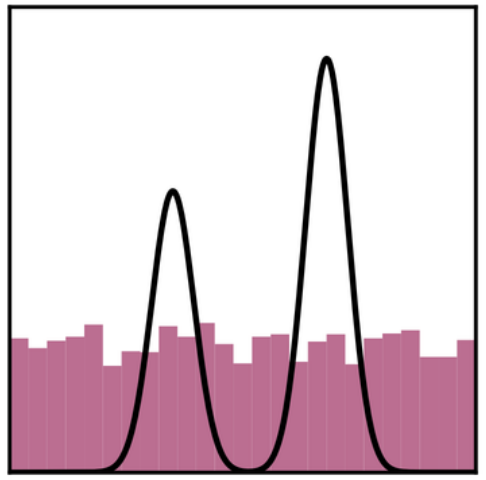
[credit: Ramon Winterhalder]



- NLO corrections in QCD are already expensive to compute
- HL-LHC and FCC demand NNLO — orders of magnitude worse

Monte Carlo 101

$$\sigma = \int f(x)\, dx$$



multiply by 1

$$\int \frac{f(x)}{g_\phi(x)} g_\phi(x)\, dx$$

neural importance sampling

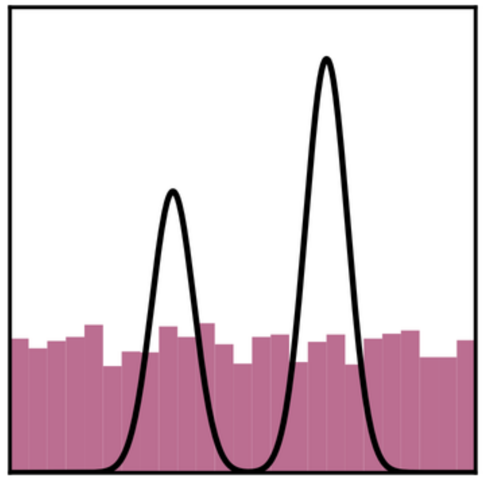
add a zero

$$\int \big(f(x) - c_\theta(x)\big)dx + \int c_\theta(x)\, dx$$

neural control variates

Monte Carlo 101

$$\sigma = \int f(x)\, dx$$



multiply by 1

$$\boxed{\int \frac{f(x)}{g_\phi(x)} g_\phi(x)\, dx}$$

neural importance sampling

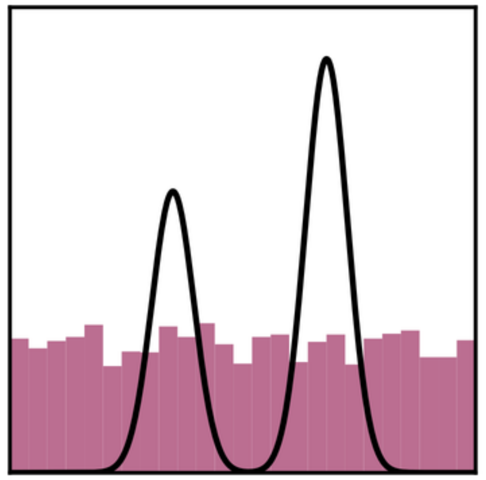
add a zero

$$\boxed{\int \big(f(x) - c_\theta(x)\big) dx + \int c_\theta(x)\, dx}$$

neural control variates

Monte Carlo 101

$$\sigma = \int f(x)\, dx$$



multiply by 1

$$\boxed{\int \frac{f(x)}{g_\phi(x)} g_\phi(x)\, dx}$$

neural importance sampling

add a zero

$$\boxed{\int \big(f(x) - c_\theta(x)\big) dx + \int c_\theta(x)\, dx}$$

neural control variates

Monte Carlo 101

neural importance sampling

$$\int \frac{f(x)}{g_\phi(x)} g_\phi(x)\, dx$$

- changes integral **measure**
- **optimal sampling**
- must be able to sample from g_ϕ

neural control variates

$$\int \big(f(x) - c_\theta(x)\big) dx + \int c_\theta(x)\, dx$$

- changes the **integrand**
- **uniform sampling**
- must know analytic integral of c_ϕ

Best of Both Worlds

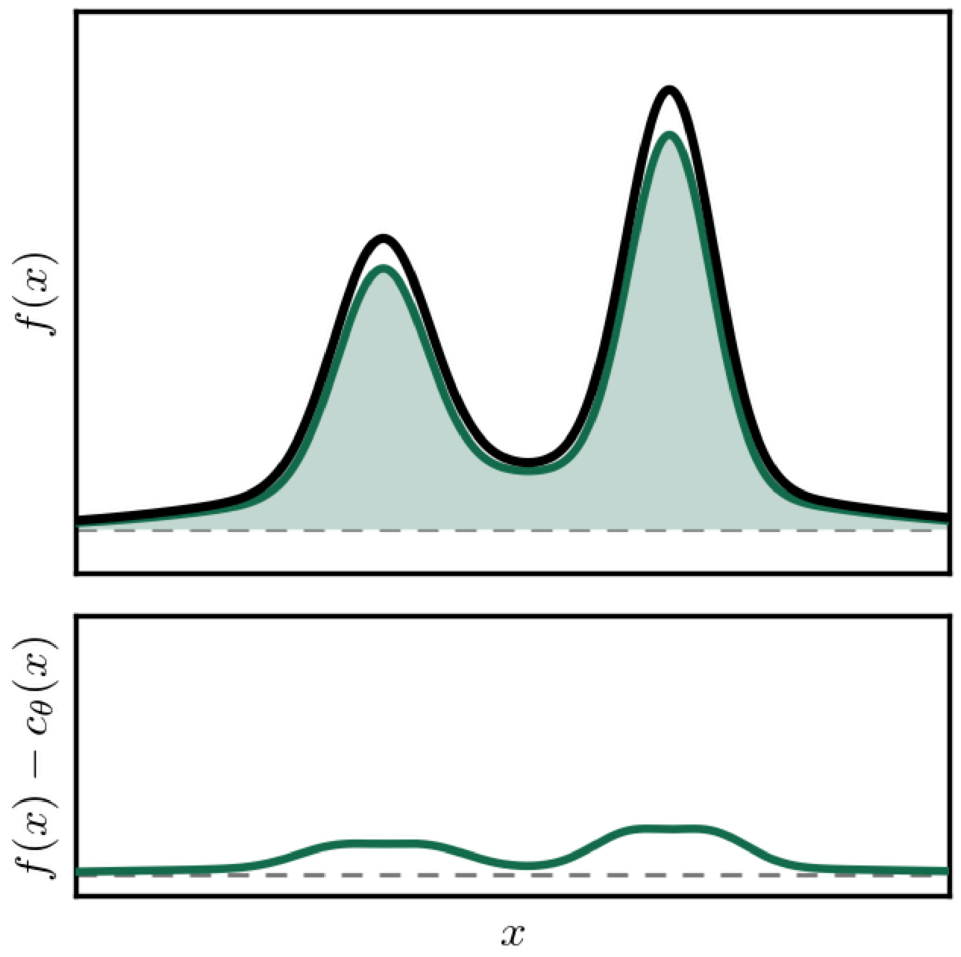
$$\sigma = \int \left(f(x) - c_\theta(x)\right) dx \; + \int c_\theta(x)\, dx$$

parametrize with **normalizing flow x constant**

$$\boxed{\sigma = \int \Big(f(x) - C_\theta p_\theta(x)\Big) dx \; + \langle C_\theta \rangle}$$

MadNIS on residual integral

$$\boxed{\sigma = \left\langle \frac{f(x) - C_\theta p_\theta(x)}{g_\phi} \right\rangle_{x\sim g_\phi} + \langle C_\theta \rangle}$$



Best of Both Worlds

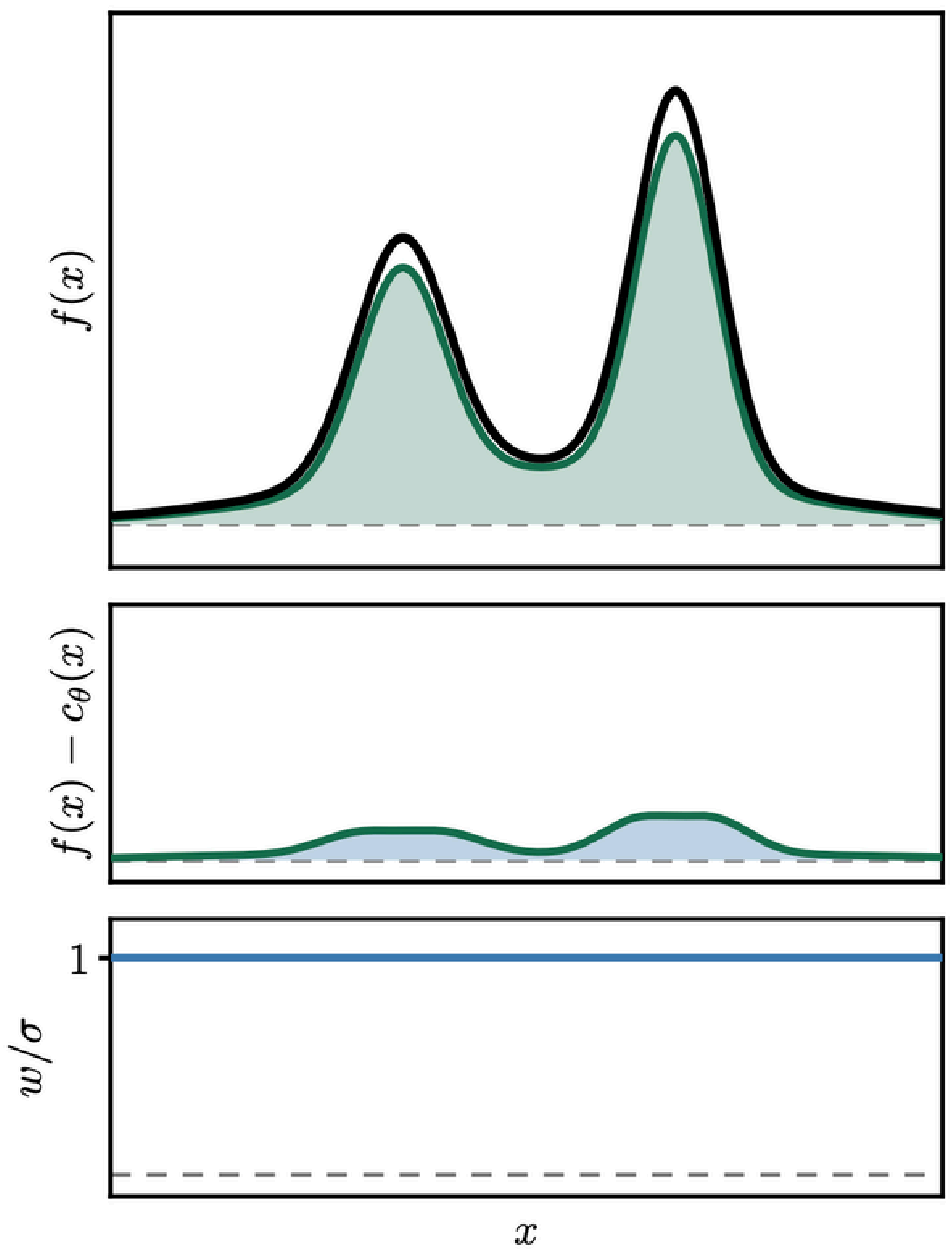
$$\sigma = \int \left(f(x) - c_\theta(x)\right)\, dx \; + \int c_\theta(x)\, dx$$

↓ parametrize with **normalizing flow x constant**

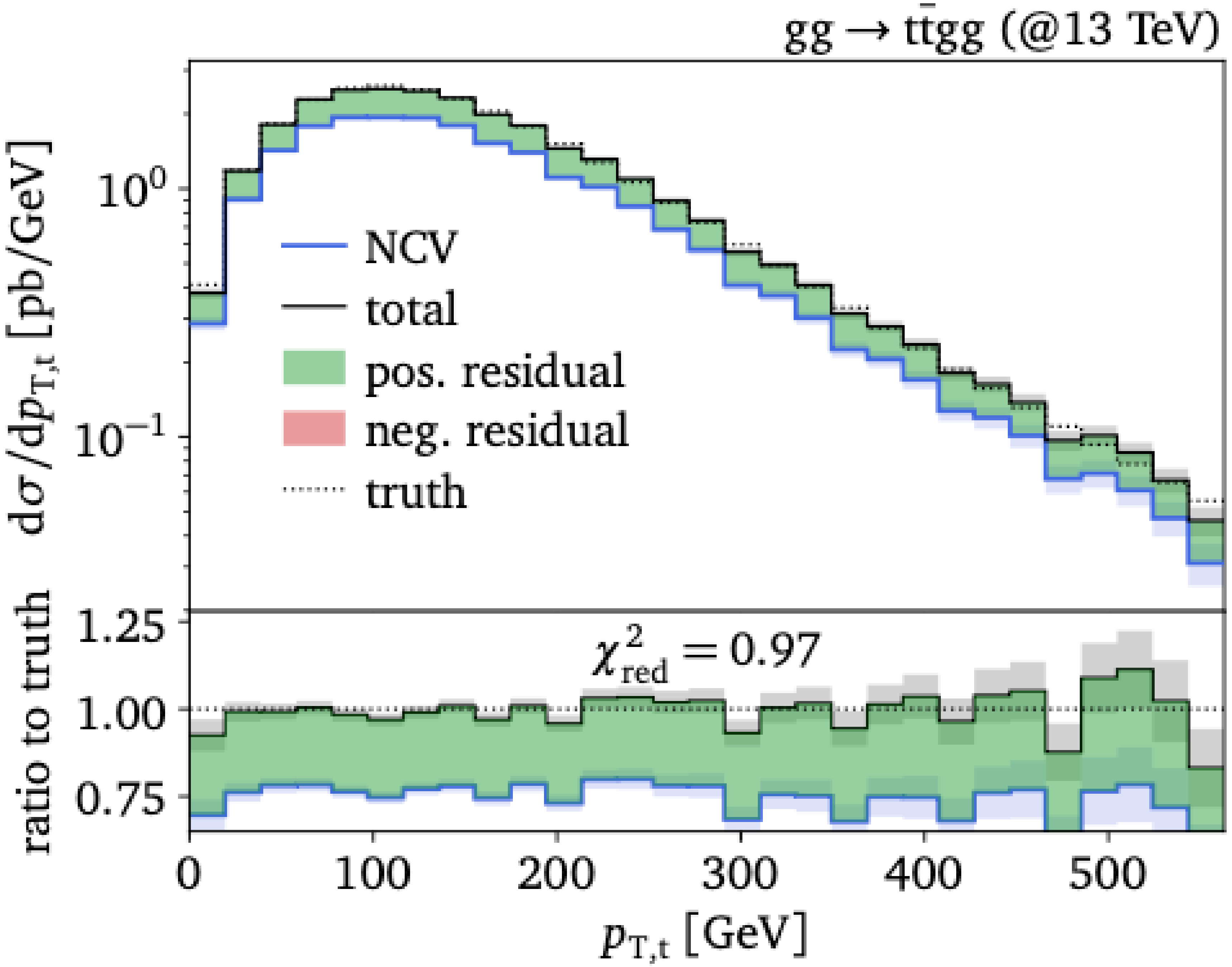
$$\boxed{\sigma = \int \left(f(x) - C_\theta p_\theta(x)\right) dx \; + \langle C_\theta \rangle}$$

↓ **MadNIS on residual integral**

$$\boxed{\sigma = \left\langle \frac{f(x) - C_\theta p_\theta(x)}{g_\phi} \right\rangle_{x \sim g_\phi} + \langle C_\theta \rangle}$$



LO Results

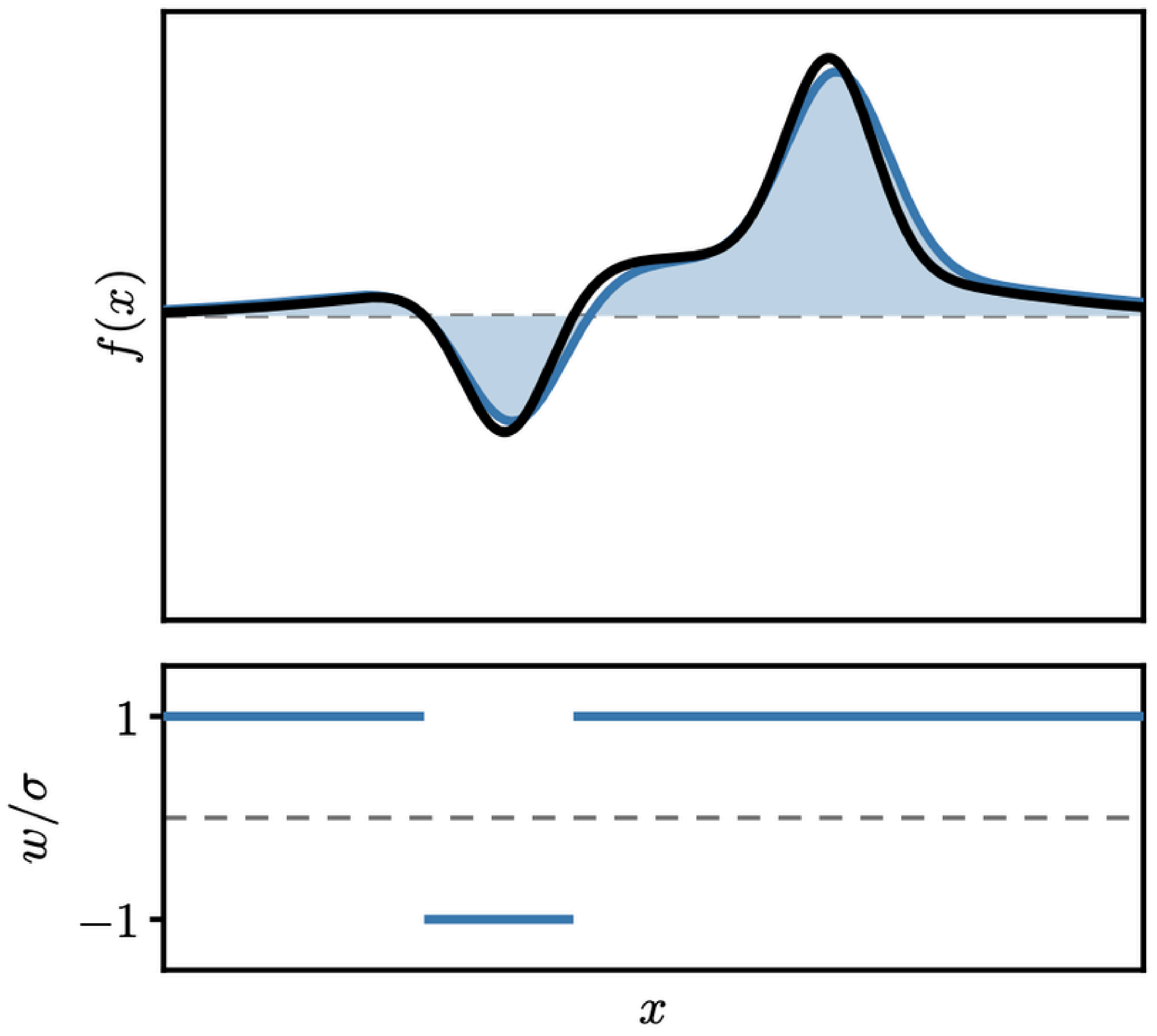


$gg \to t\bar{t}gg$

$\sim \times 20$ speed up using MadNIS

$\sim \times 3$ speed up compared to MadNIS!

The Problem with Negative Weights



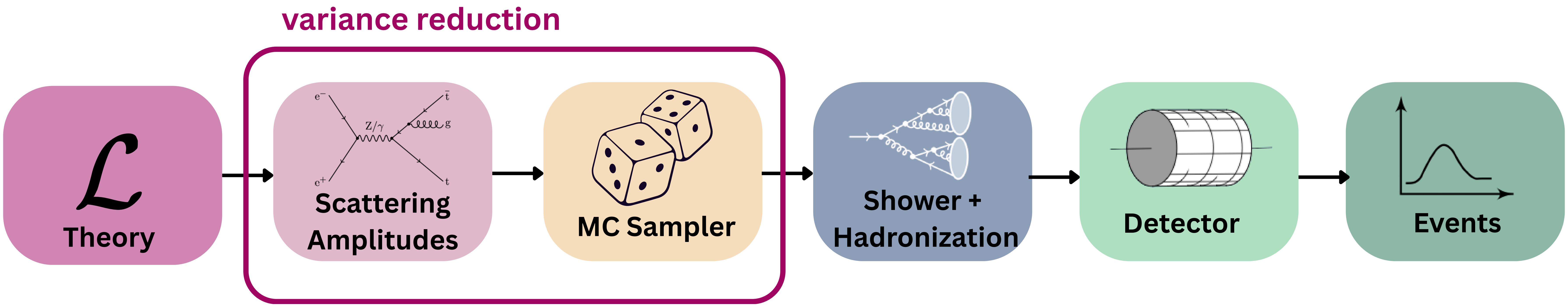
Inherent lower bound on the relative variance:

$$\mathrm{RV} \geq \mathrm{RV}_{\mathrm{min}} = \frac{\mathrm{Var}(w)}{\sigma^2} = \frac{4\rho(1-\rho)}{(2\rho-1)^2}$$

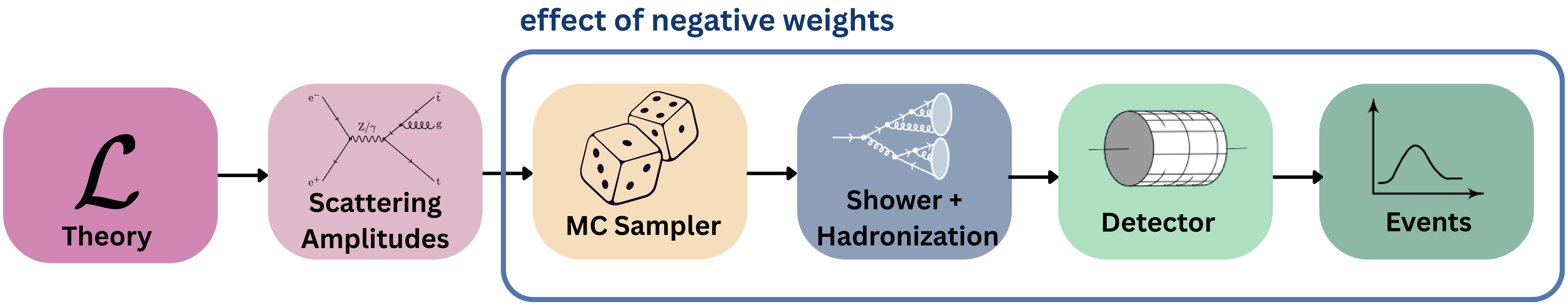
with $\quad \rho = \dfrac{\sigma^+}{\sigma^+ + \sigma^-}$

even perfect importance sampling cannot move this bound!

The Problem with Negative Weights



The Problem with Negative Weights



negative events cost statistics:

$$N_{\text{eff}} = N(1-2r)^2 \qquad r = \frac{N_{\text{neg}}}{N}$$

eg. 25% of negative events : **4x more events through the whole simulation chain**

Next-to-Leading Order Calculations

At NLO the cross section is:

$$\sigma_{\mathrm{NLO}} = \int d\Phi_n \left[B(\Phi_n) + V(\Phi_n)\right] + \int d\Phi_{n+1}\, R(\Phi_{n+1}) \,,$$

contain soft and collinear divergences

Subtraction scheme: **Catani-Seymour Dipoles**, FKS, Antenna

$$\sigma_{\mathrm{NLO}} = \int d\Phi_n \left[B(\Phi_n) + V(\Phi_n) + I(\Phi_n)\right] + \int d\Phi_{n+1} \left[R(\Phi_{n+1}) - S(\Phi_{n+1})\right]$$

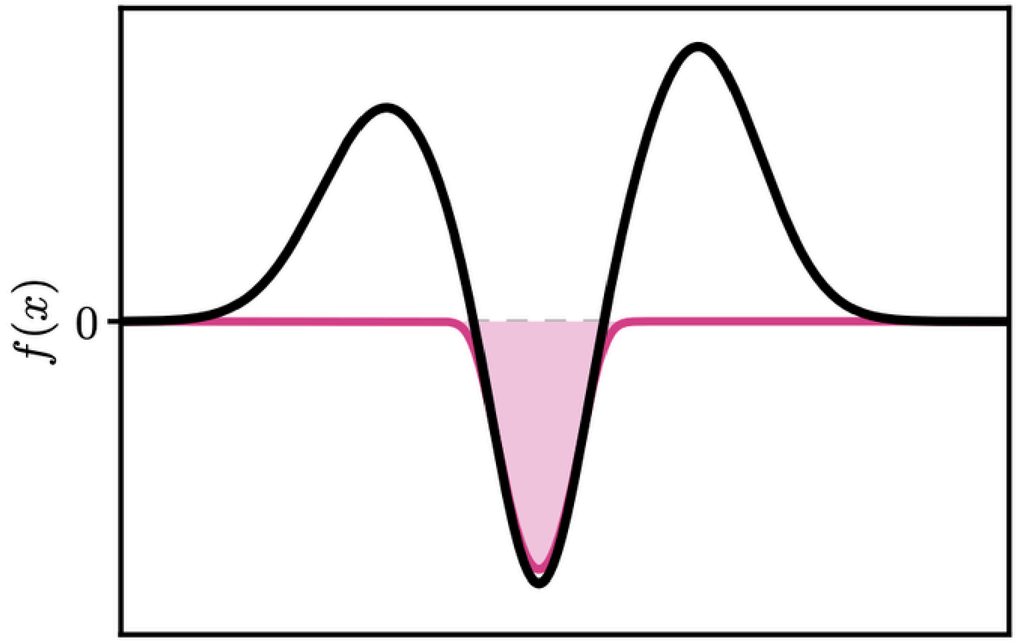
sign-changing **sign-changing**

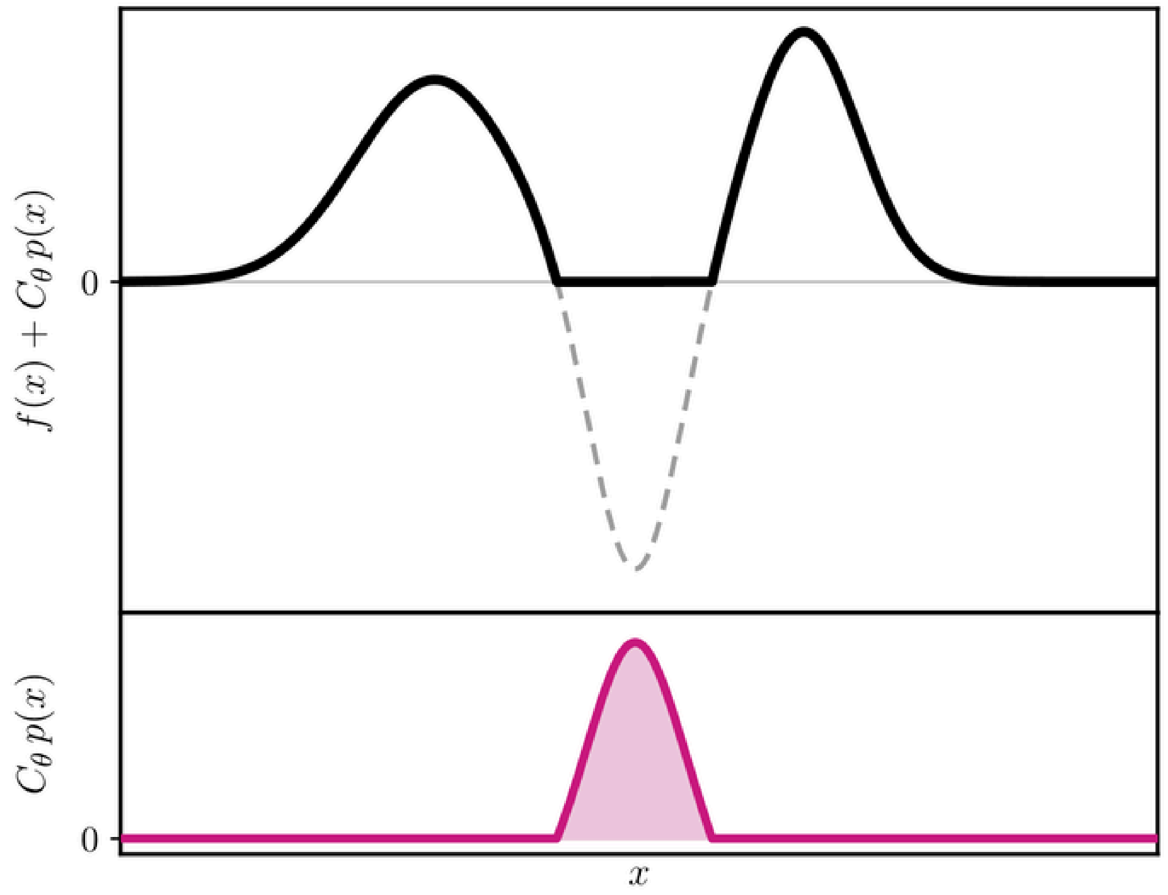
Neural Control Variates for NLO

$$\sigma = \int \big(f(x) + C_\theta p_\theta(x)\big) dx \; - \langle C_\theta \rangle$$

parametrize with *signed-NCV*

$$\sigma = \langle f(x) + C_\theta^- p(x) \rangle \; - \langle C_\theta^- \rangle$$





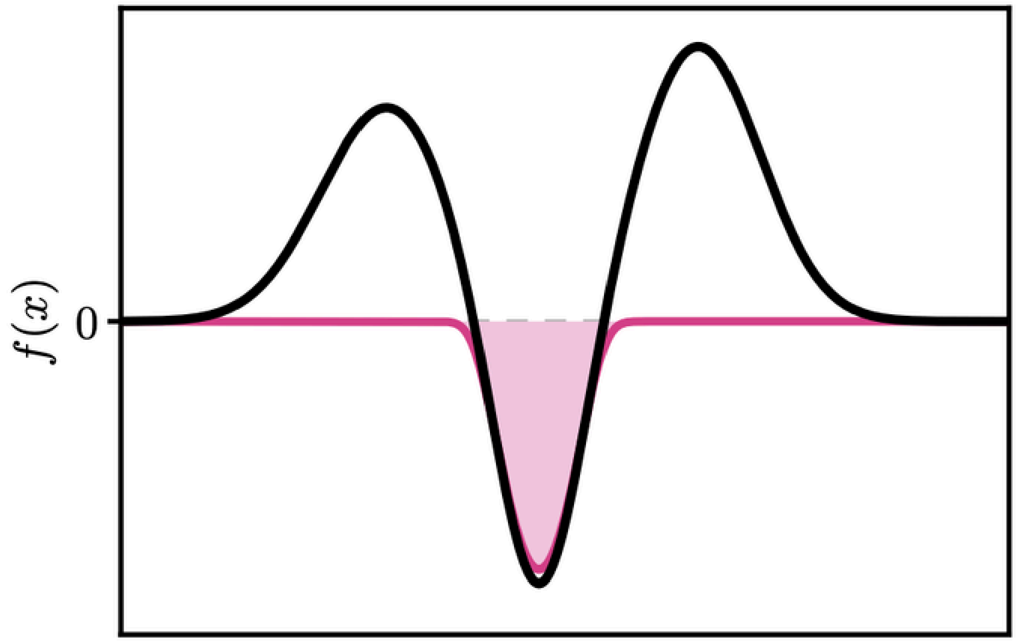
Neural Control Variates for NLO

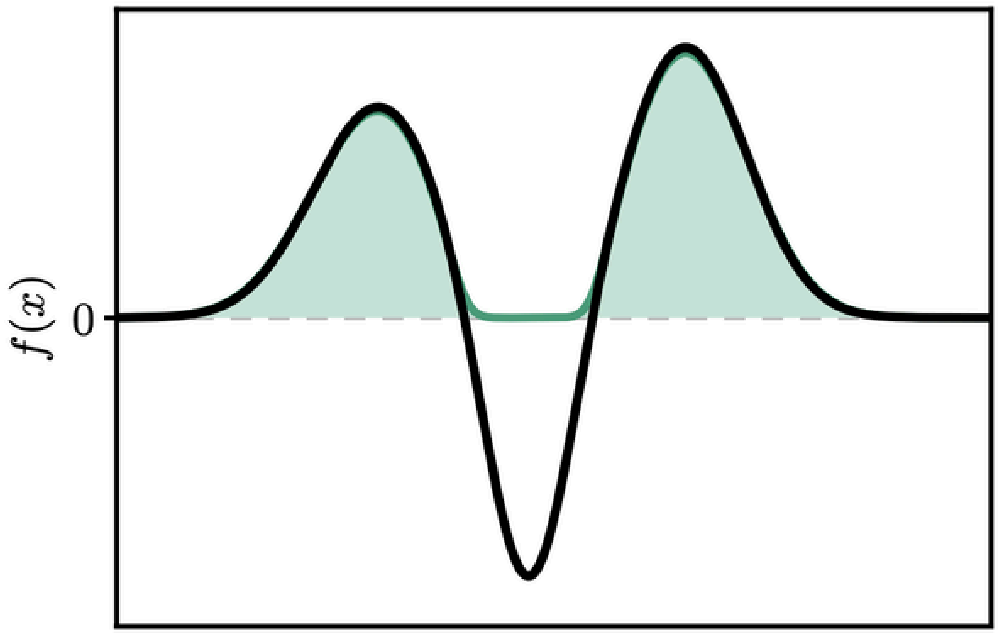
$$\sigma = \int \big(f(x) + C_\theta p_\theta(x)\big) dx \; - \langle C_\theta \rangle$$

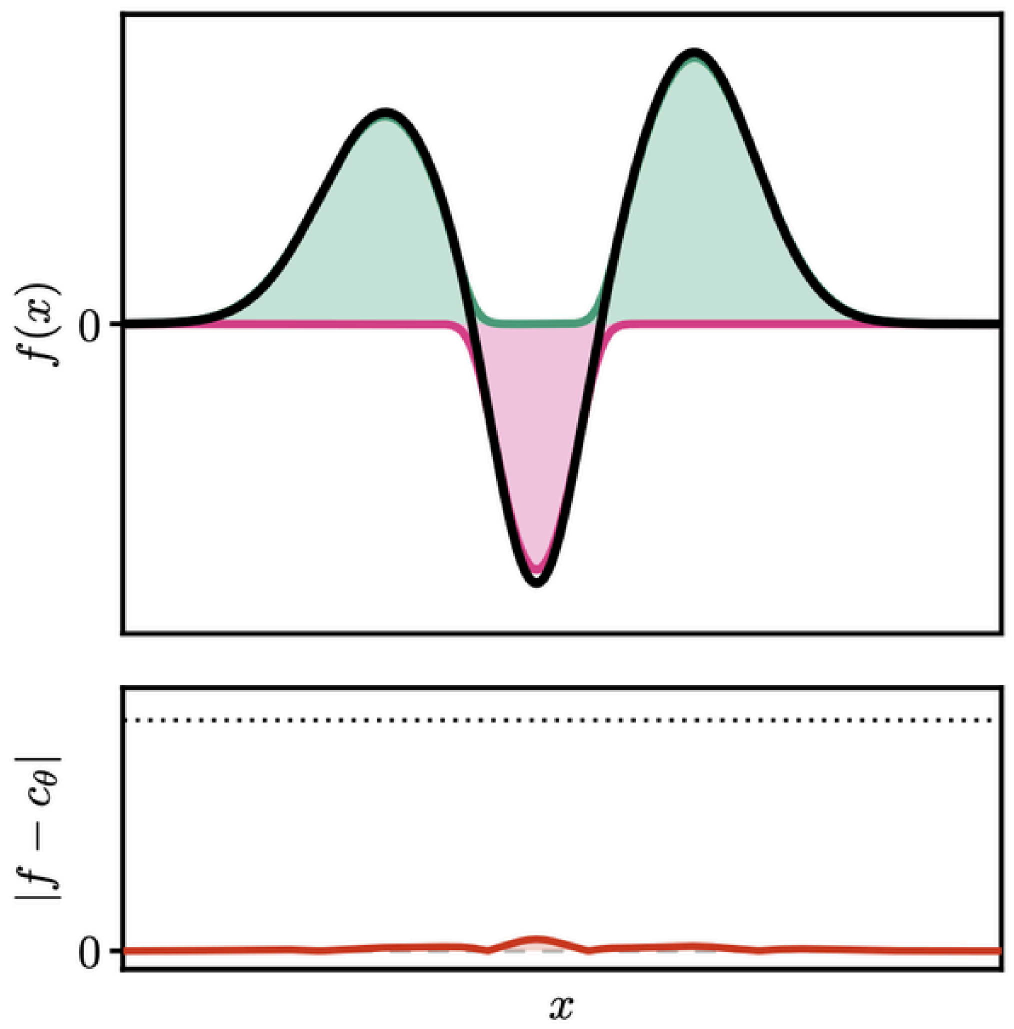
parametrize with *signed-NCV*

1 positive NCV + 1 negative NCV

$$\sigma = \langle f(x) + C_\theta^- p(x) - C_\theta^+ p(x) \rangle - \langle C_\theta^- \rangle + \langle C_\theta^+ \rangle$$







Neural Control Variates for NLO

$$\sigma = \int \big(f(x) + C_\theta p_\theta(x)\big)dx \; - \langle C_\theta \rangle$$

parametrize with *signed-NCV*

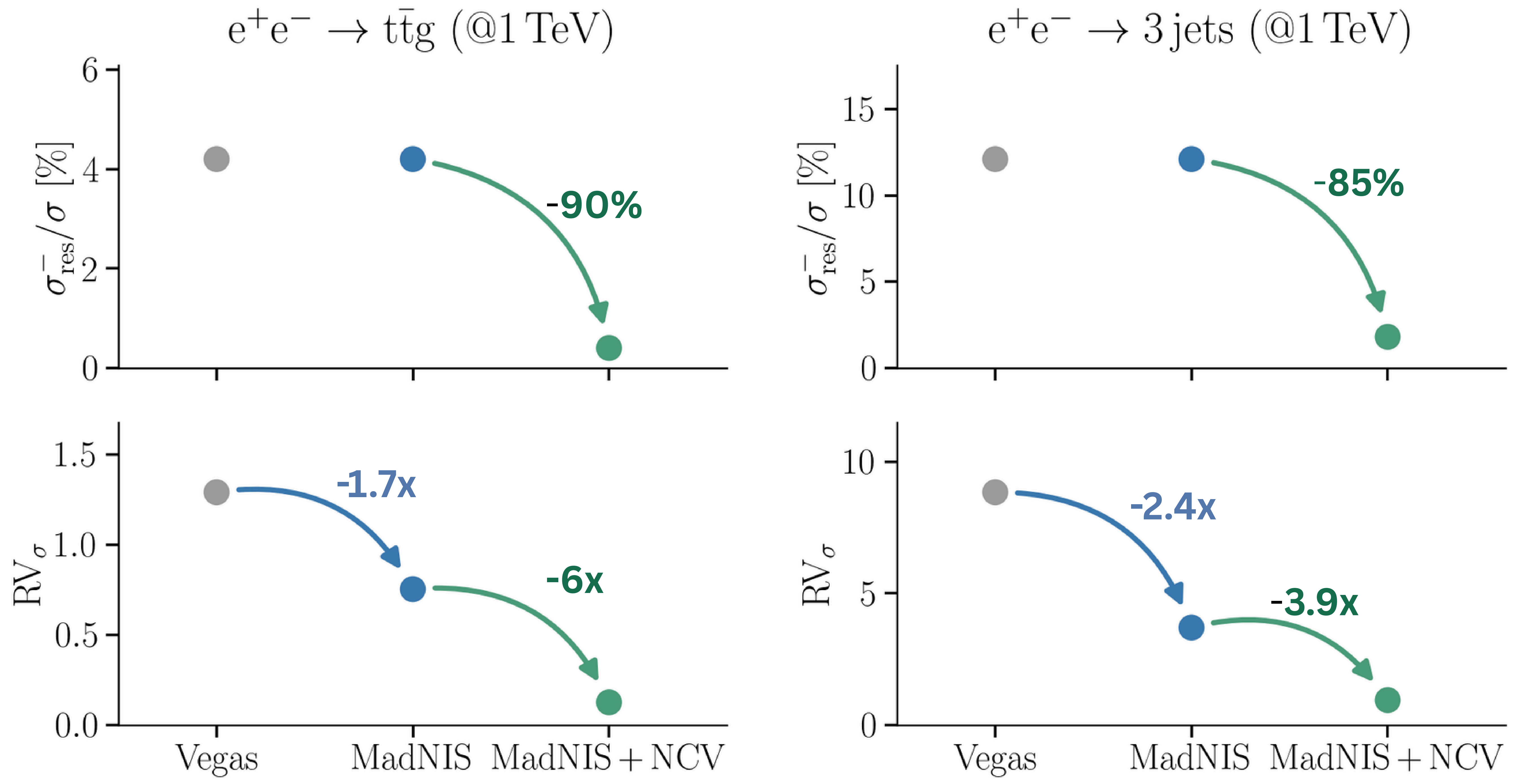
1 positive NCV + 1 negative NCV

$$\sigma = \langle f(x) + C_\theta^- p(x) - C_\theta^+ p(x)\rangle - \langle C_\theta^- \rangle + \langle C_\theta^+ \rangle$$

MadNIS on residual integral

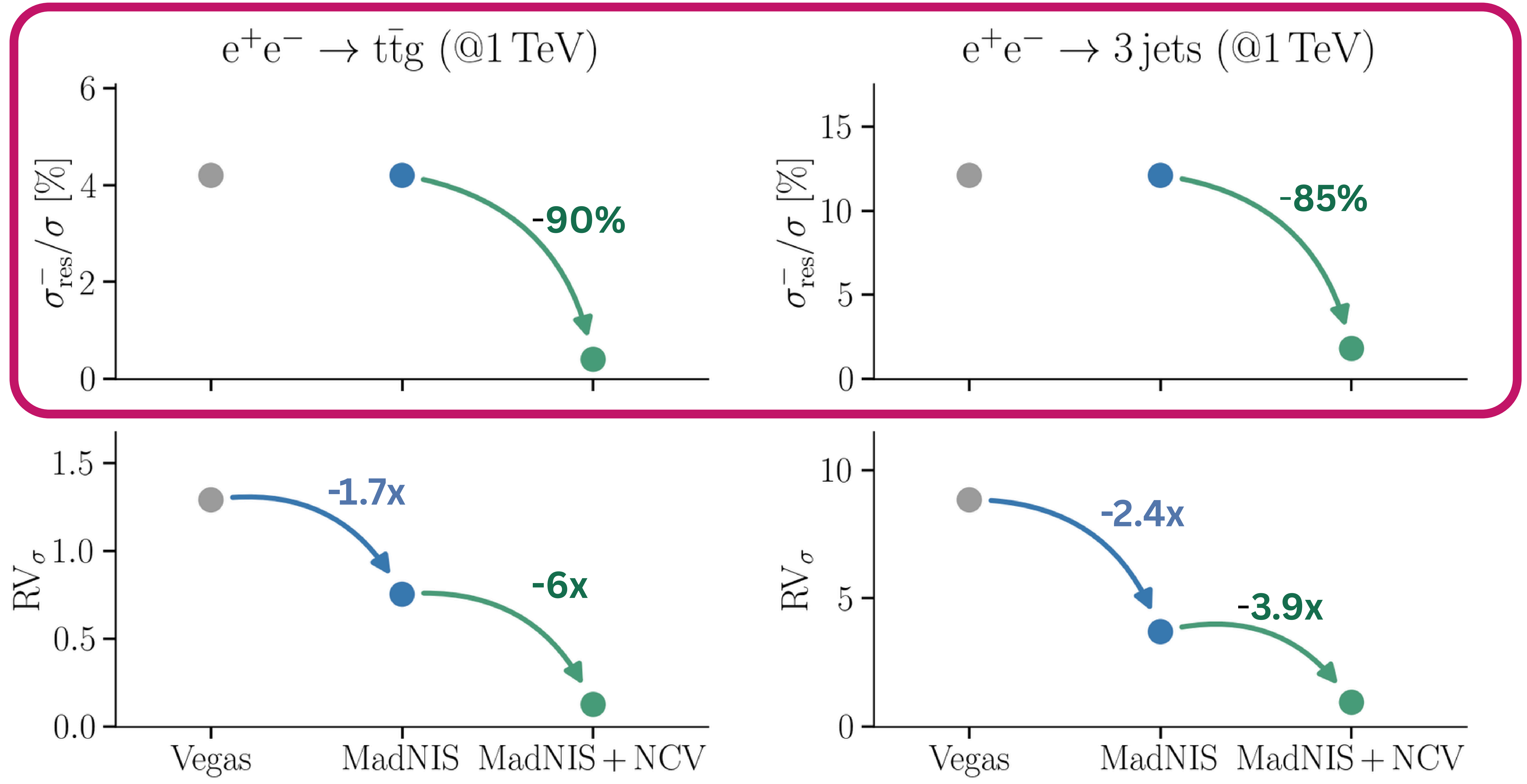
$$\boxed{\sigma = \left\langle \frac{f(x) + C_\theta^- p(x) - C_\theta^+ p(x)\rangle}{g_\phi(x)} \right\rangle_{x\sim g_\phi} - \langle C_\theta^- \rangle + \langle C_\theta^+ \rangle}$$

NLO Results



$e^+e^- \to t\bar{t}g$ (@1 TeV)
$e^+e^- \to 3\,\text{jets}$ (@1 TeV)
$\sigma^-_{\text{res}}/\sigma$ [%]
RV_σ
-90%
-85%
-1.7x
-6x
-2.4x
-3.9x
Vegas
MadNIS
MadNIS + NCV

NLO Results



$e^+e^- \to t\bar{t}g$ (@1 TeV)
$e^+e^- \to 3$ jets (@1 TeV)
$\sigma^-_{\mathrm{res}}/\sigma$ [%]
-90%
-85%
RV_σ
-1.7x
-6x
-2.4x
-3.9x
Vegas
MadNIS
MadNIS + NCV

Let’s Talk!

NCVs

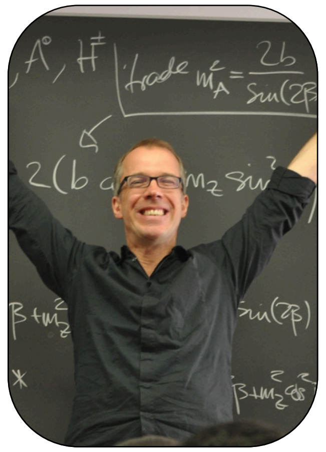
- improve variance reduction for (N)LO predictions
- effectively reduce negative weights
- can reshape the integrand

Rebecca

Theo

Ramon

Sophia



Tilman

read the paper!

[2607.23591]