STOCHASTIC NORMALIZING FLOWS FOR EFFECTIVE STRING THEORY

ELIA CELLINI

31/07/2024

Lattice 2024

Liverpool

Based on:

M. Caselle, E.C., A. Nada, M. Panero

- JHEP 07 (2022) 015, arxiv:2201.08862

M. Caselle, E.C., A. Nada

- JHEP 02 (2024) 048, arxiv:2307.01107
- In prep., arxiv:2408.XXXX

UNIVERSITÀ DI TORINO

OUTLINE

1. EFFECTIVE STRING THEORY
2. STOCHASTIC NORMALIZING FLOWS
3. NUMERICAL RESULTS
4. OUTLOOKS

EFFECTIVE STRING THEORY

EFFECTIVE STRING THEORY

Correlators of Polyakov loops modelled in terms of **string partition functions**:

$$\langle P(0)P^{\dagger}(R)\rangle \sim \int DXe^{-S_{EST}[X]} \equiv Z_{EST}$$

The main choice for S_{EST} is the **Nambu-Goto** **(NG)** action:

$$S_{NG} = \sigma \int d\xi^2 \sqrt{g}$$

- **Anomalous** at quantum level → **effective, large-distance description** of Yang-Mills theories (**low-energy universality theorem)**.
- Works only up to order $1/R^5$ → first order approximation of a more general theory → **Beyond Nambu-Goto** **(BNG)**
  [Aharony and Komargodski; 1302.6257], [Brandt and Meineri; 1603.06969],[Caselle; 2104.10486]

NAMBU-GOTO STRING

Main method: **zeta-function regularization**

Main observables:

- **Partition function** → directly associated with the interquark potential. Well known at all the order.
- **Correlation functions** (e.g. width σw^2) → measure of the density of the chromoelectric flux tube

Analytical limits:

- **Correlation functions**
- **Higher order corrections** (Beyond NG)

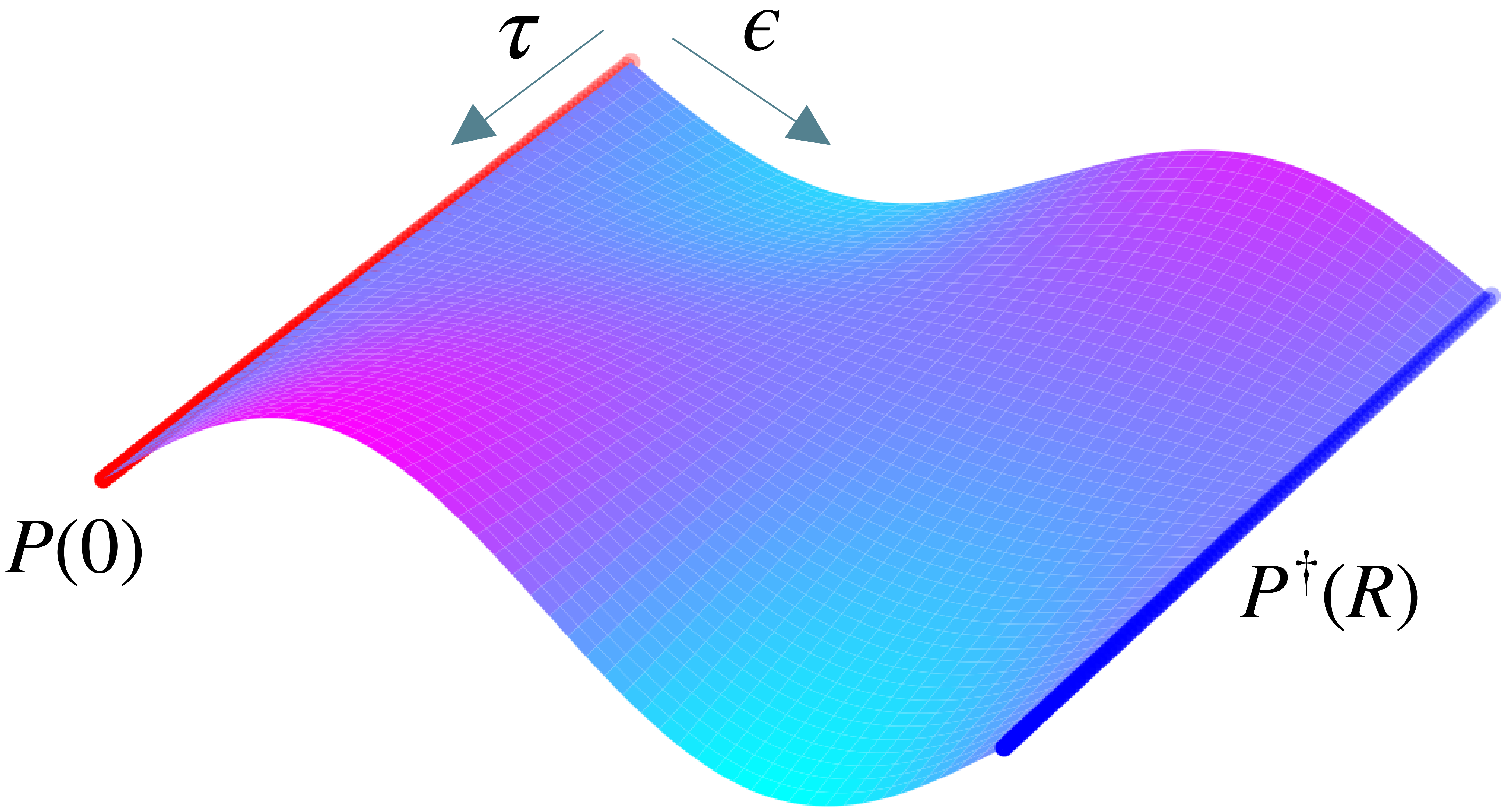
See talks by:

- D. Panfalone 29/07, 11:35
- A. Athenodorou 31/07, 11:35
- L. Verzichelli 31/07, 11:55
- M. Caselle 31/07, 12:35

LATTICE NAMBU-GOTO STRING

$$S_{NG}(\phi) = \sigma \sum_{x \in \Lambda} \left[\sqrt{1 + (\partial_\mu \phi(x))^2/\sigma} - 1 \right]$$

- $d = 2 + 1$ target Yang-Mills
- σ string tension
- Λ : square lattice of size $L \times R$, $a = 1$
- $\phi(x) = \phi(\tau, \epsilon) \in \mathbb{R}$
- $\phi(\tau + L, \epsilon) = \phi(\tau, \epsilon)$
- $\phi(\tau, 0) = \phi(\tau, R) = 0$
- $\sigma w^2 = \langle \phi(\tau, R/2)^2 \rangle_\tau$

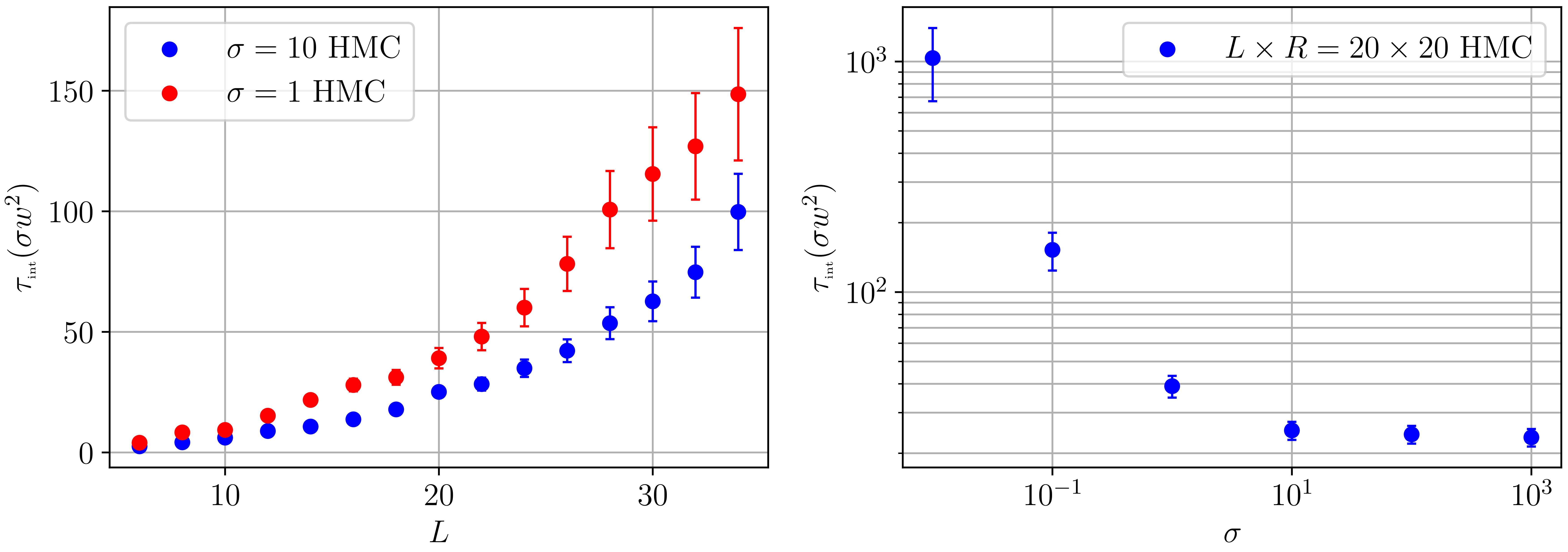


[Caselle, EC, Nada; 2307.01107]

LACKS OF NUMERICAL METHODS

Numerical problems:

- Strong **non-linearity** → **critical** theory (**Critical Slowing Down**)
- Estimation of **partition functions**



We did a proof-of-concept with continuous normalizing flows; however, we found scaling issues.

[Caselle, EC, Nada; 2307.01107]

STOCHASTIC NORMALIZING FLOWS

NON-EQUILIBRIUM MCMC (NE-MCMC)

$$q_0 \simeq e^{-S_0} \xrightarrow{P_1} e^{-S_1} \xrightarrow{P_2} \dots \xrightarrow{P_N} e^{-S_N} \simeq p$$

1. **Thermalized** q_0 "**prior**"
2. $P_i \propto \exp(-S_i)$ **change along the processes** and satisfy detailed balance.
3. $p = \exp(-S_N)/Z_N \rightarrow$ "target" distribution

Forward and **reverse** probability density:

$$q_0(\phi_0)\prod_{n=0}^{N-1} P_{i+1}[\phi_i \rightarrow \phi_{i+1}] = q_0(\phi_0)P_f[\phi_0, \dots, \phi_N]$$

$$p(\phi_N)\prod_{n=0}^{N-1} P_{i+1}[\phi_{i+1} \rightarrow \phi_i] = p(\phi_N)P_r[\phi_N, \dots, \phi_0]$$

Remark: no thermalization during the processes.

CROOKS' THEOREM

Observe that:

$$\ln \frac{q_0(\phi_0)P_f[\phi_0,\ldots,\phi_N]}{p(\phi_N)P_r[\phi_N,\ldots,\phi_0]} = \underbrace{S_N(\phi_N) - S_0(\phi_0)}_{} - Q - \Delta F = W(\phi_0,\ldots,\phi_N) - \Delta F = W_d(\phi_0,\ldots,\phi_N)$$

(dimensionless) Work W

Where:

$$Q = \underbrace{\ln \frac{P_r[\phi_N,\ldots,\phi_0]}{P_f[\phi_0,\ldots,\phi_N]} = \sum_{n=0}^{N-1} \ln \frac{q_{n+1}(\phi_{n+1})}{q_{n+1}(\phi_n)}}_{} = \sum_{n=0}^{N-1}\Big(S_{n+1}(\phi_{n+1}) - S_{n+1}(\phi_n)\Big)$$

Detailed Balance

Thus:

$$\frac{q_0(\phi_0)P_f[\phi_0,\ldots,\phi_N]}{p(\phi_N)P_r[\phi_N,\ldots,\phi_0]} = \frac{\mathscr{P}_f(W_d)}{\mathscr{P}_r(-W_d)} = e^{W_d}$$

→ **Crooks Theorem**
[Crooks; cond-mat/9901352]

JARZYNSKI'S EQUALITY

Observe also:

$$1 = \int \prod_{i=0}^{N} d\phi_i q_0(\phi_0) P_f[\phi_0, \ldots, \phi_N] \left(\frac{p(\phi_N) P_r[\phi_N, \ldots, \phi_0]}{q_0(\phi_0) P_f[\phi_0, \ldots, \phi_N]} \right) = \langle e^{-W_d} \rangle_f$$

$$\langle e^{-W} \rangle_f = e^{-\Delta F}$$

Jarzynski's equality

[Jarzynski; cond-mat/9610209]

$$\langle \mathcal{O} \rangle_{\phi \sim p} = \langle \mathcal{O} e^{-W_d} \rangle_f$$

Non-Equilibrium Ensemble

Equilibrium Quantity

Equivalent to **Annealed Importance Sampling** and widely applied to LFT

[Neal; physics/9803008]

See:

- Talk by A. Bulgarelli 29/07, 14.35
- Poster by D. Vadacchino 30/07

STOCHASTIC NORMALIZING FLOWS

Stochastic Normalizing Flows (SNFs) combine **NE-MCMC** update and **Normalizing Flows layers**:

$$\phi_0 \longrightarrow g_\theta^1(\phi_0) \xrightarrow{P_1} \phi_1 \longrightarrow g_\theta^2(\phi_1) \xrightarrow{P_2} \dots \xrightarrow{P_N} \phi_N = \phi$$

Where P_i are MCMC update and g_θ^i are Normalizing Flows (NFs) layers:

$$g_\theta^{n+1} : \; q^n \rightarrow q_\theta^{n+1} \qquad \phi_{n+1} = g_\theta(\phi_n) \qquad \ln\left(q_n(\phi_n)/q_{n+1}(\phi_{n+1})\right) = \ln|\det J_{g^{n+1}}(\phi_n)|$$

[Rezende et al.; 1505.05770],[Wu et al.; 2002.06707],[Caselle, EC, Nada, Panero; 2201.08862]

DISSIPATED WORK W_d^θ

We have now:

$$W_d^\theta = W_\theta(\phi_0, \ldots, \phi_N) - \Delta F = S_N(\phi_N) - S_0(\phi_0) - Q_\theta - \Delta F$$

Where:

$$Q_\theta = \sum_{n=0}^{N-1} \left(S_{n+1}(\phi_{n+1}) - S_{n+1}(\phi_n) + \ln |\det J_{g_\theta^n}| \right)$$

We can now train a SNF by minimizing:

$$\mathscr{L}(\theta) = \langle W_d^\theta \rangle_f = D_{KL}(q_0 P_f || p P_r) \geq 0 \quad \Rightarrow \quad \langle W_\theta \rangle_f \geq \Delta F \quad \Rightarrow$$ **<u>Second Law!</u>**

Measure how reversible the process is.

Lower W_d → numerically stabler exponential averages

[Wu+; 2002.06707],[Caselle, EC, Nada, Panero; 2201.08862]

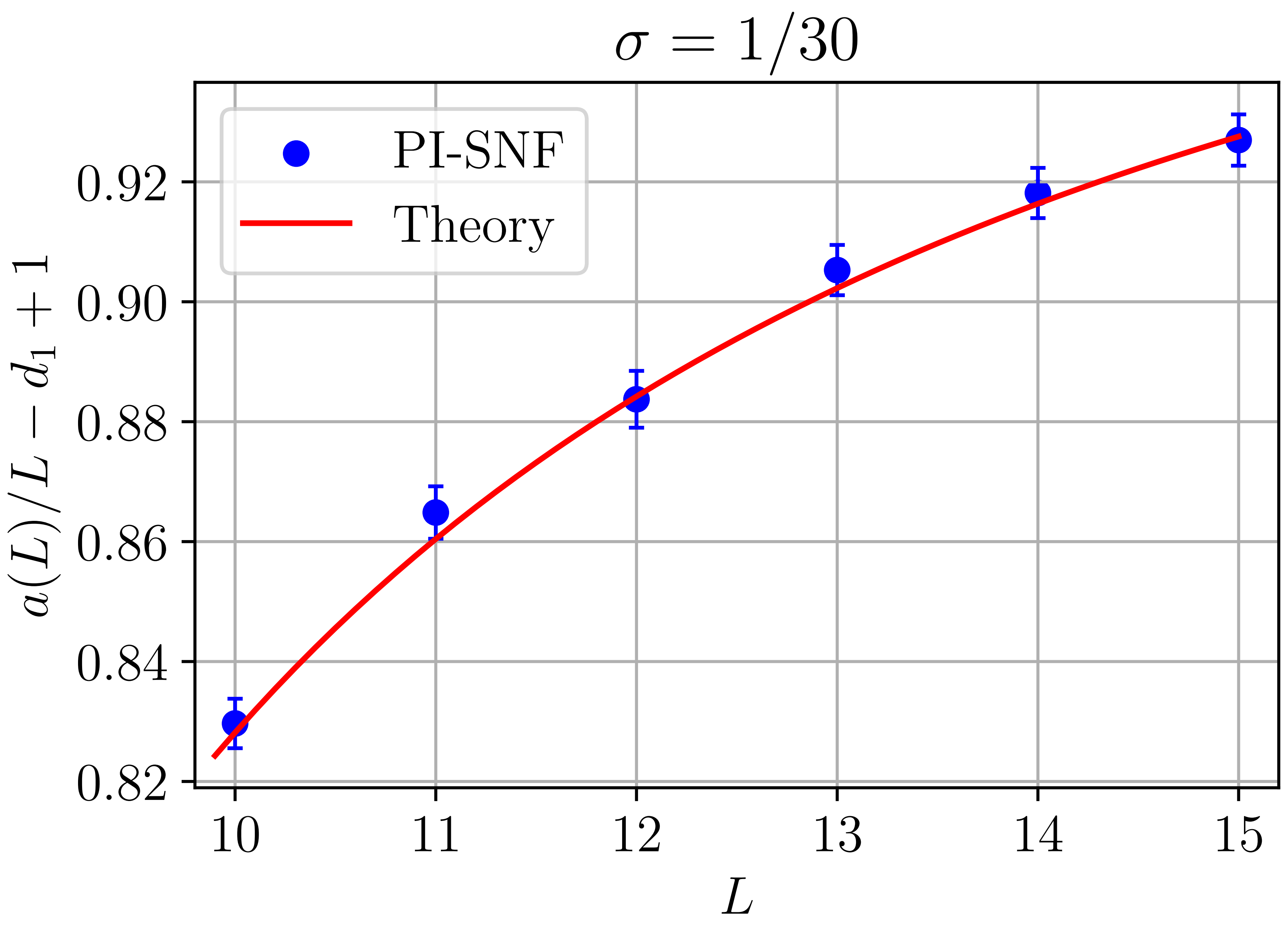
See talk by A. Nada

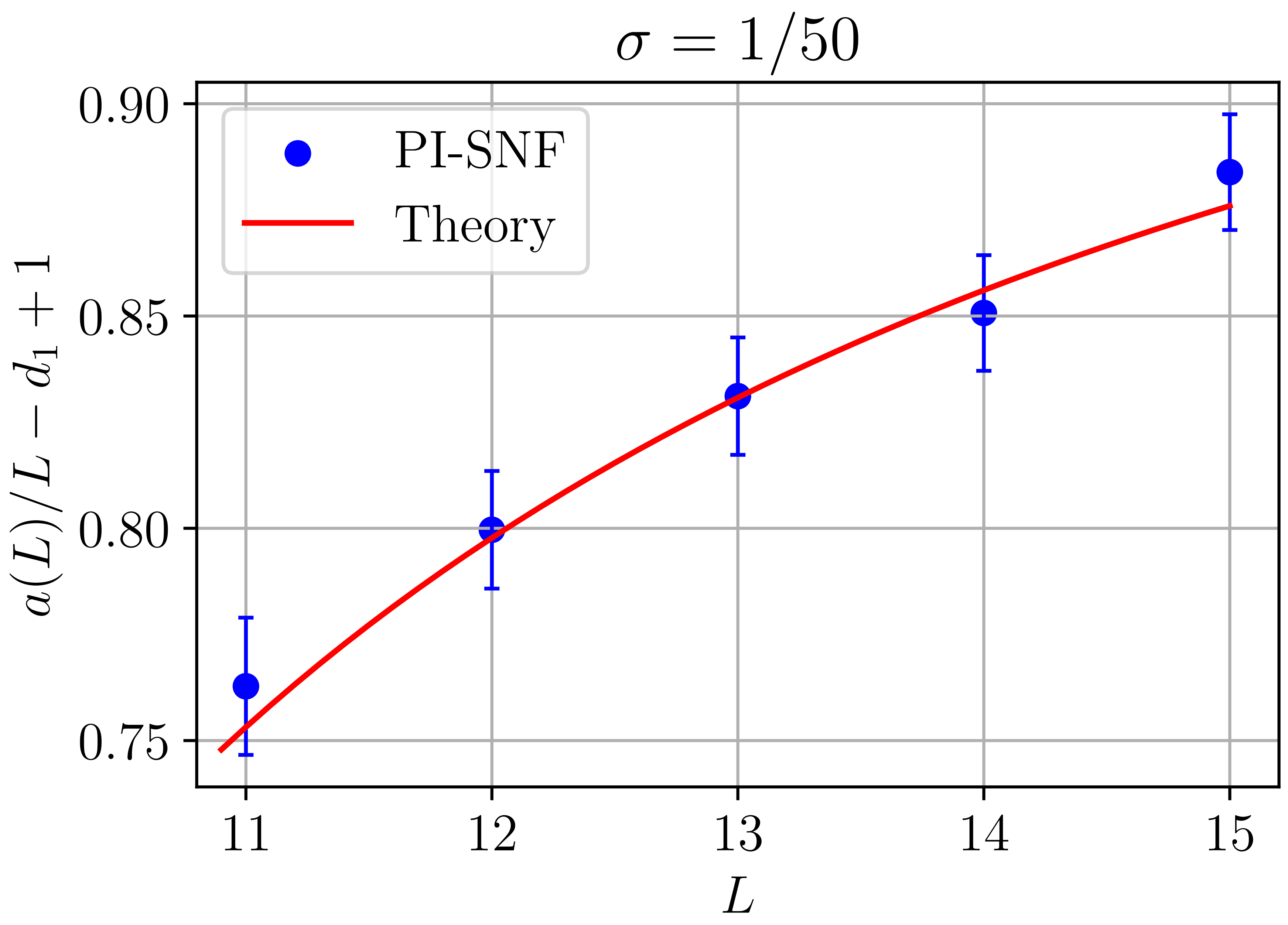
NUMERICAL RESULTS

NG FREE ENERGY $R \gg L$

$$-\log Z = \sigma R L \sqrt{1 - \frac{\pi}{3\sigma L^2}} + \cdots$$

Fitted: $-1.03(2)$,
Target: $-1.047...$





Fitted: $-1.04(7)$,
Target: $-1.047...$

NG WIDTH $R \gg L$

$\sigma = 1/10$

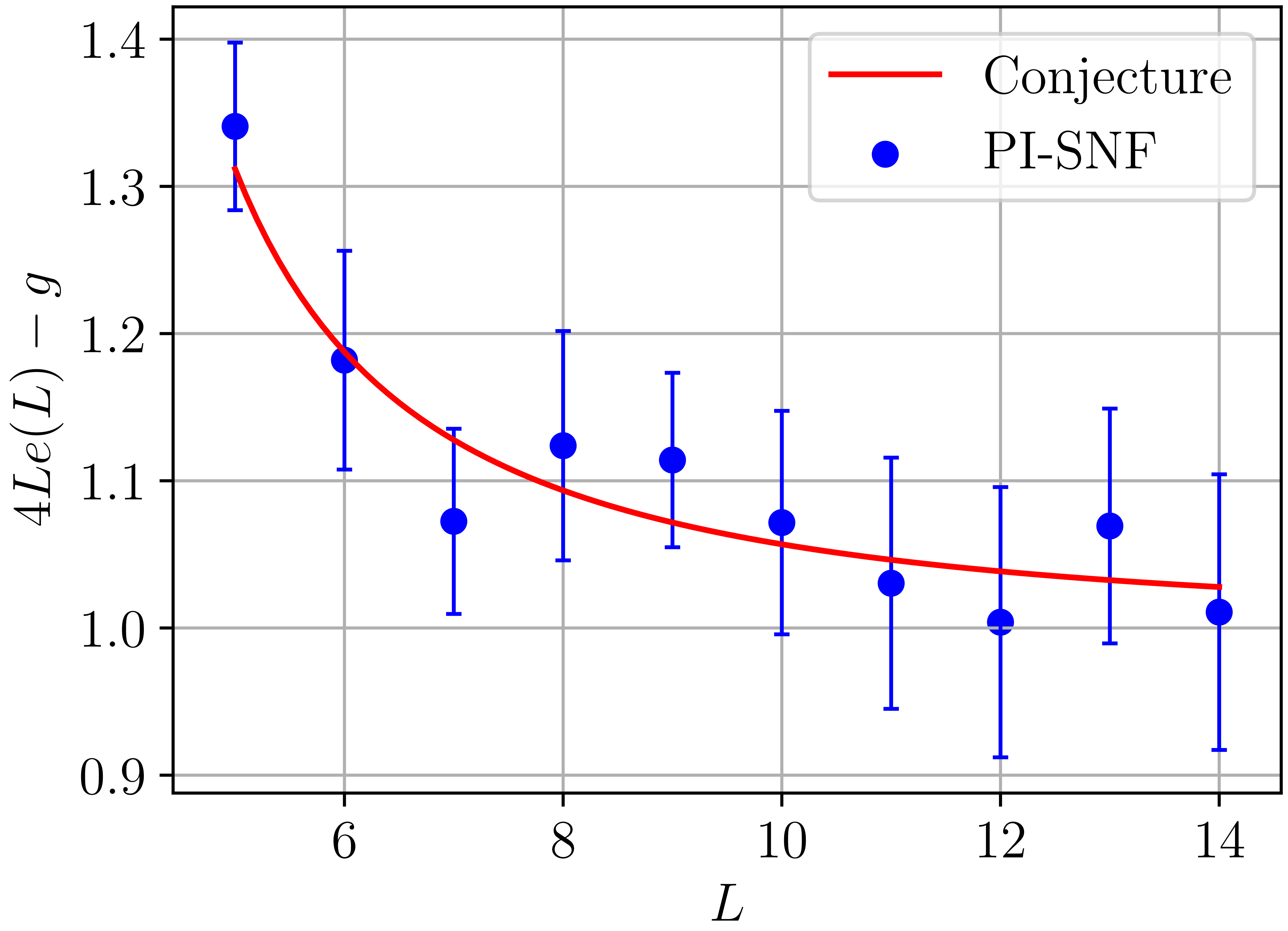
Conjecture:

$$\sigma w^2 = \frac{1}{\sqrt{1 - \frac{\pi}{3\sigma L^2}}} \frac{R}{4L} + \cdots$$

NG: $\sigma(L)/\sigma$

Gaussian solution

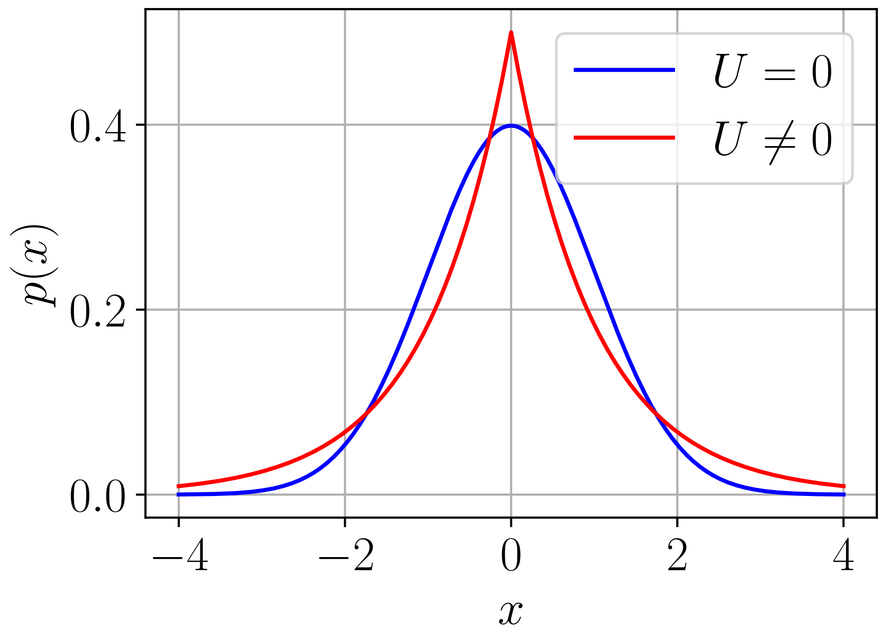
Fitted: $-1.09(8)$, target: $-1.047\ldots$



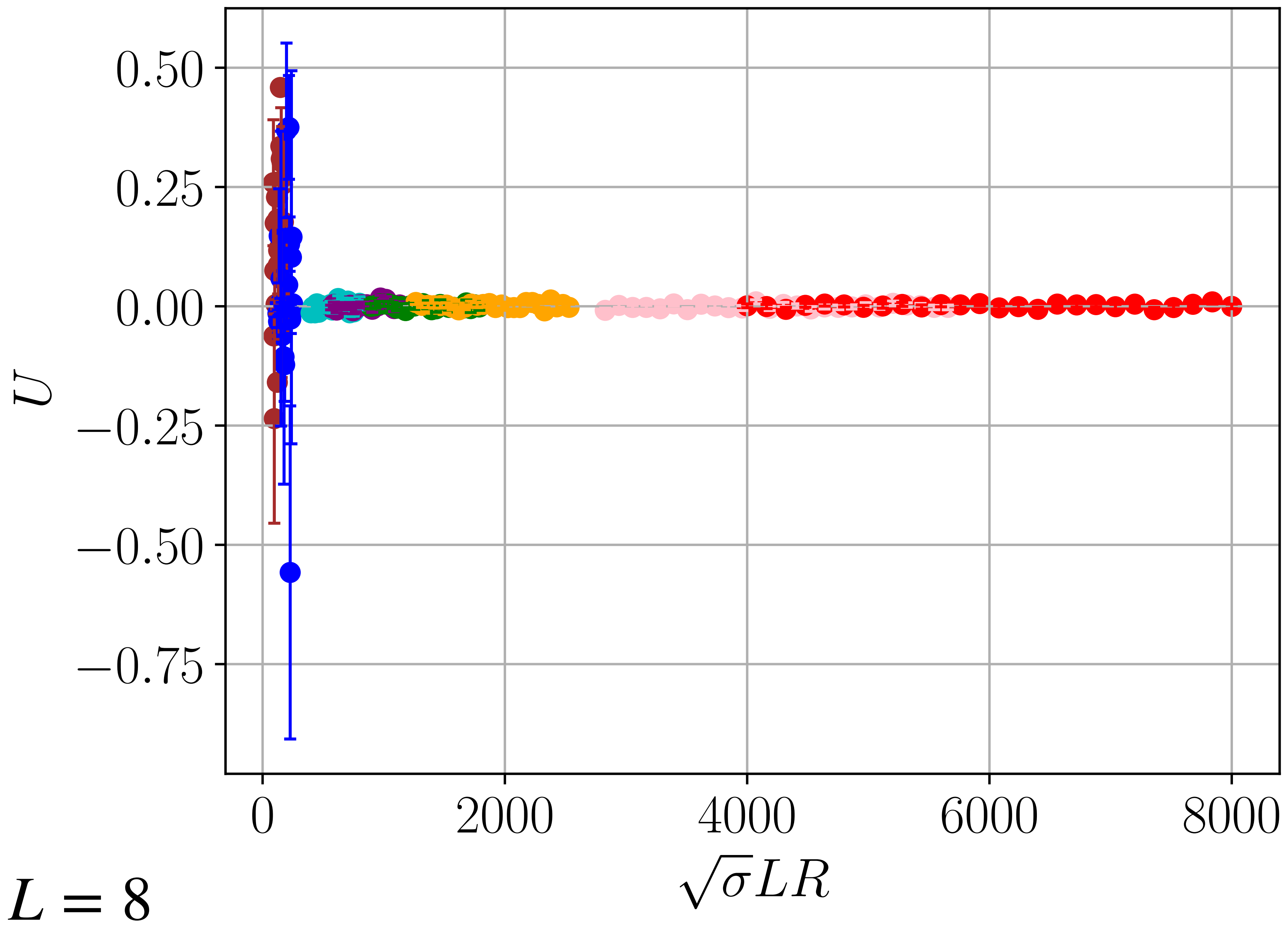
[Caselle, EC, Nada; 2309.14983][Caselle;1004.3875]

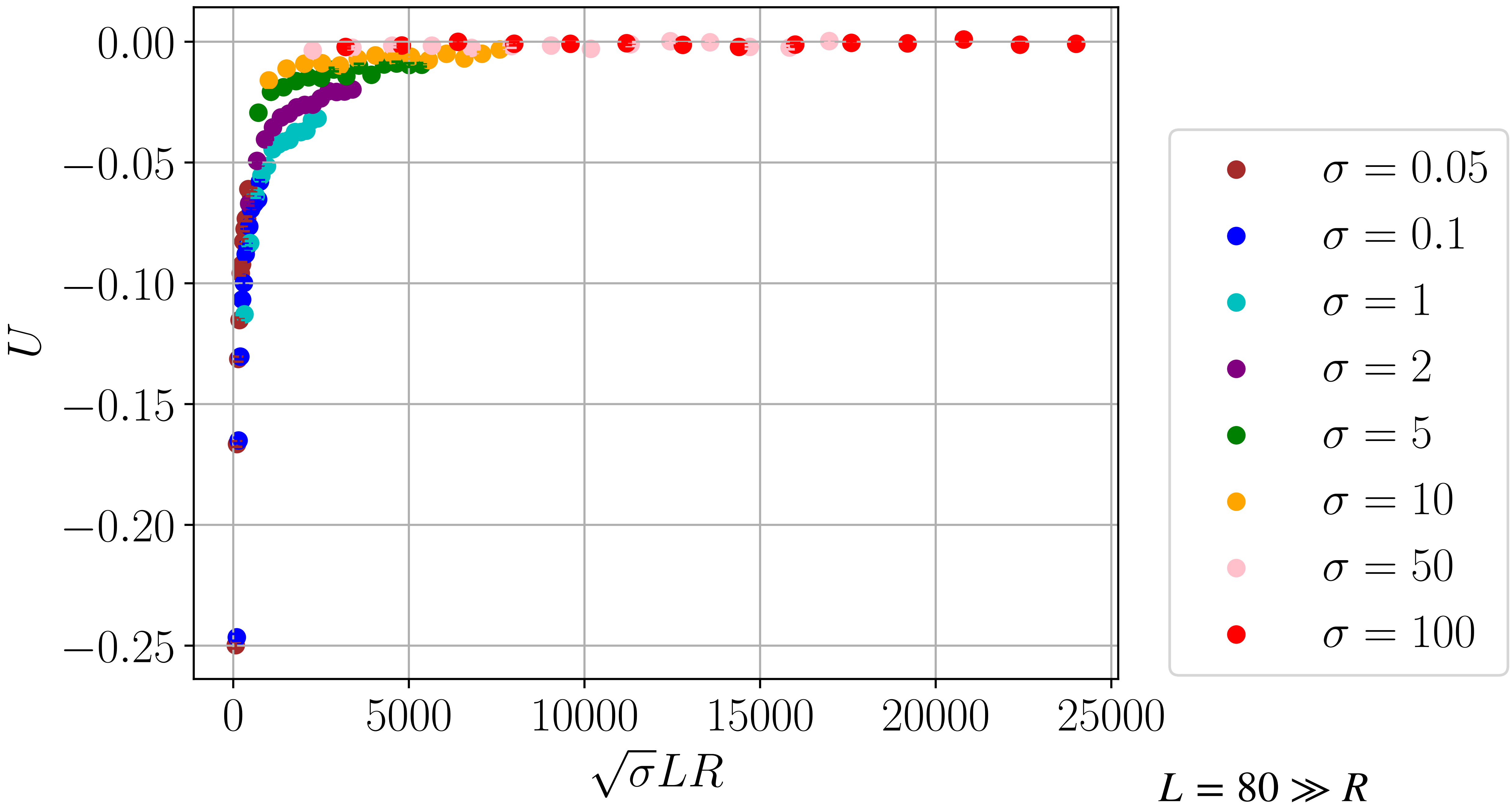
IS THE NG FLUX TUBE SHAPE GAUSSIAN?

$$U = 1 - \frac{\langle \phi^4(\tau, R/2) \rangle_\tau}{3\langle \phi^2(\tau, R/2) \rangle_\tau^2}$$



$\sigma = 100$





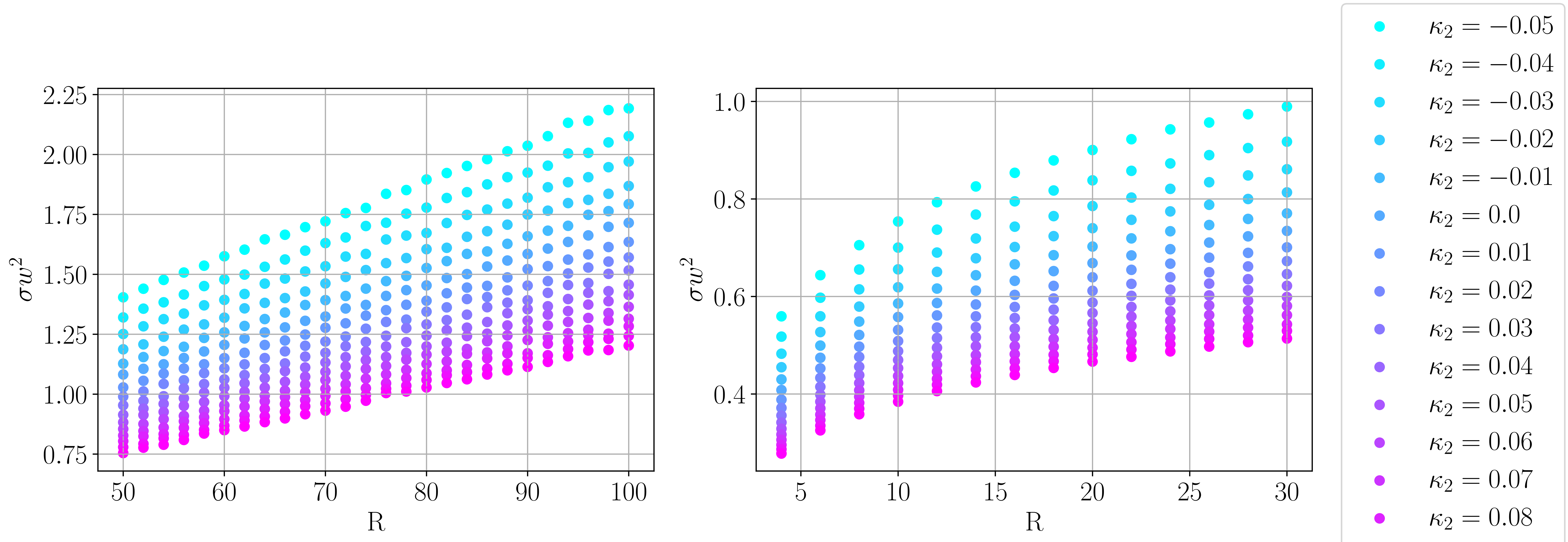
$R \gg L = 8$

$L = 80 \gg R$

BEYOND NG: WIDTH

$$S_{EST}(\phi) = S_{NG}(\phi) + \kappa_2 \mathcal{K}^2(\phi)$$

$$\mathcal{K}^2(\phi) = \sum_{(\tau,\epsilon)\in\Lambda} \mathcal{L}^2(\phi(\tau,\epsilon)) = \sum_{(\tau,\epsilon)\in\Lambda} (\partial_\tau\partial_\tau\phi(\tau,\epsilon))^2 + (\partial_\epsilon\partial_\epsilon\phi(\tau,\epsilon))^2 + 2(\partial_\tau\partial_\epsilon\phi(\tau,\epsilon))^2$$

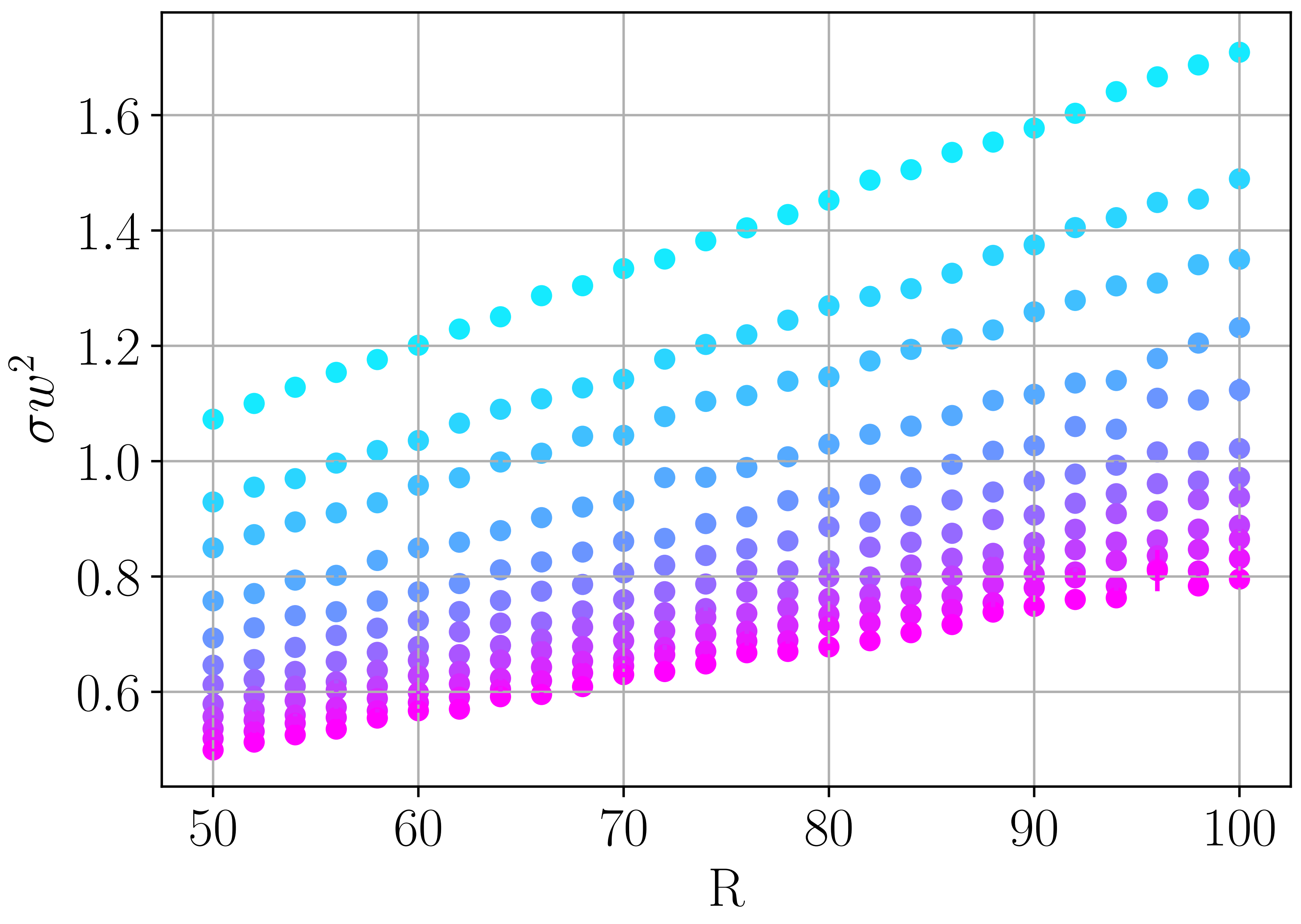


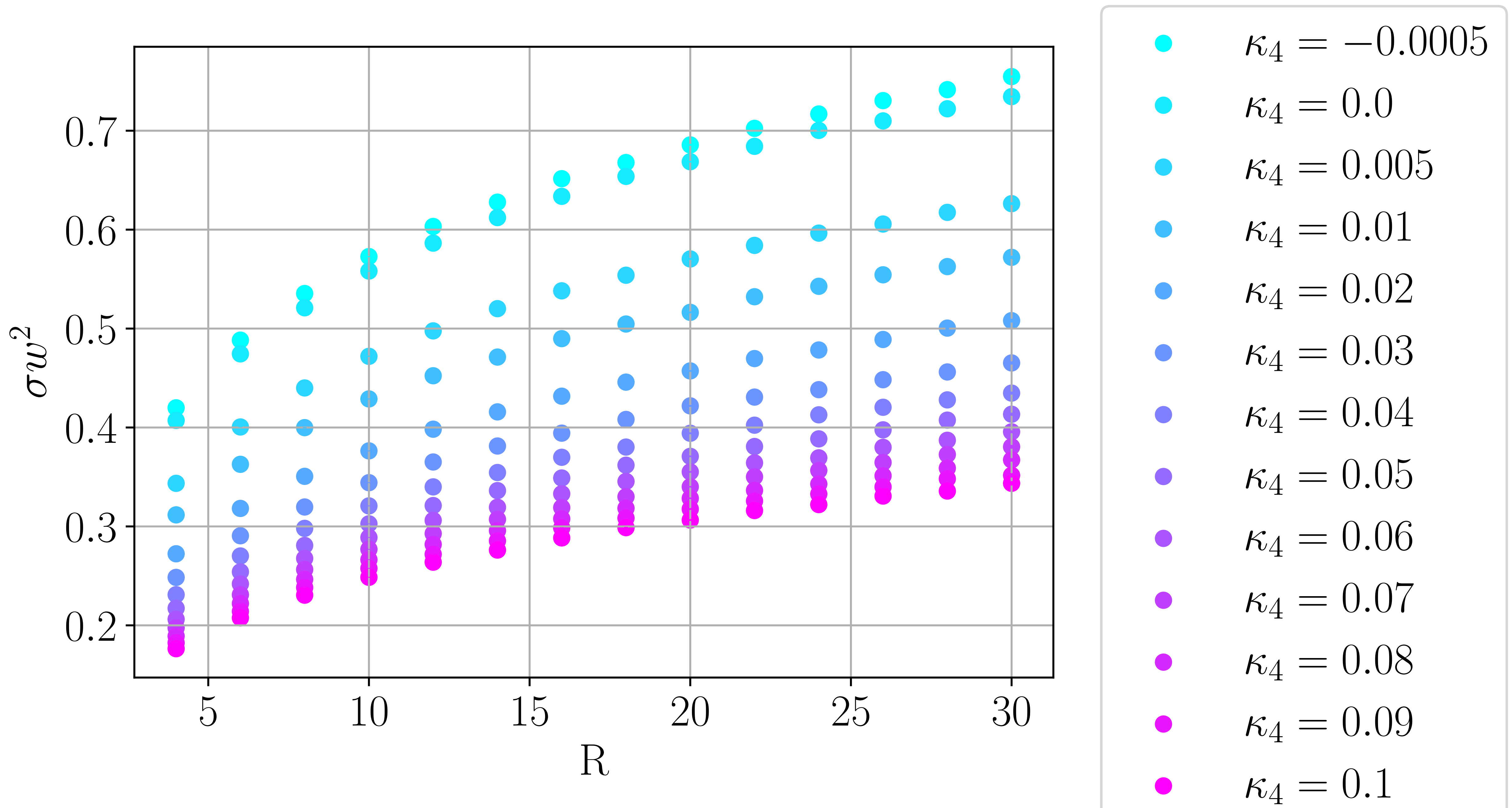
$\sigma = 100$ $\qquad R \gg L = 20$ $\qquad L = 80 \gg R$

BEYOND NG: WIDTH

$$S_{EST}(\phi) = S_{NG}(\phi) + \kappa_4 \mathcal{K}^4(\phi)$$

$$\mathcal{K}^4(\phi) = \sum_{x \in \Lambda} \left(\mathcal{L}^2(\phi(x))\right)^2$$

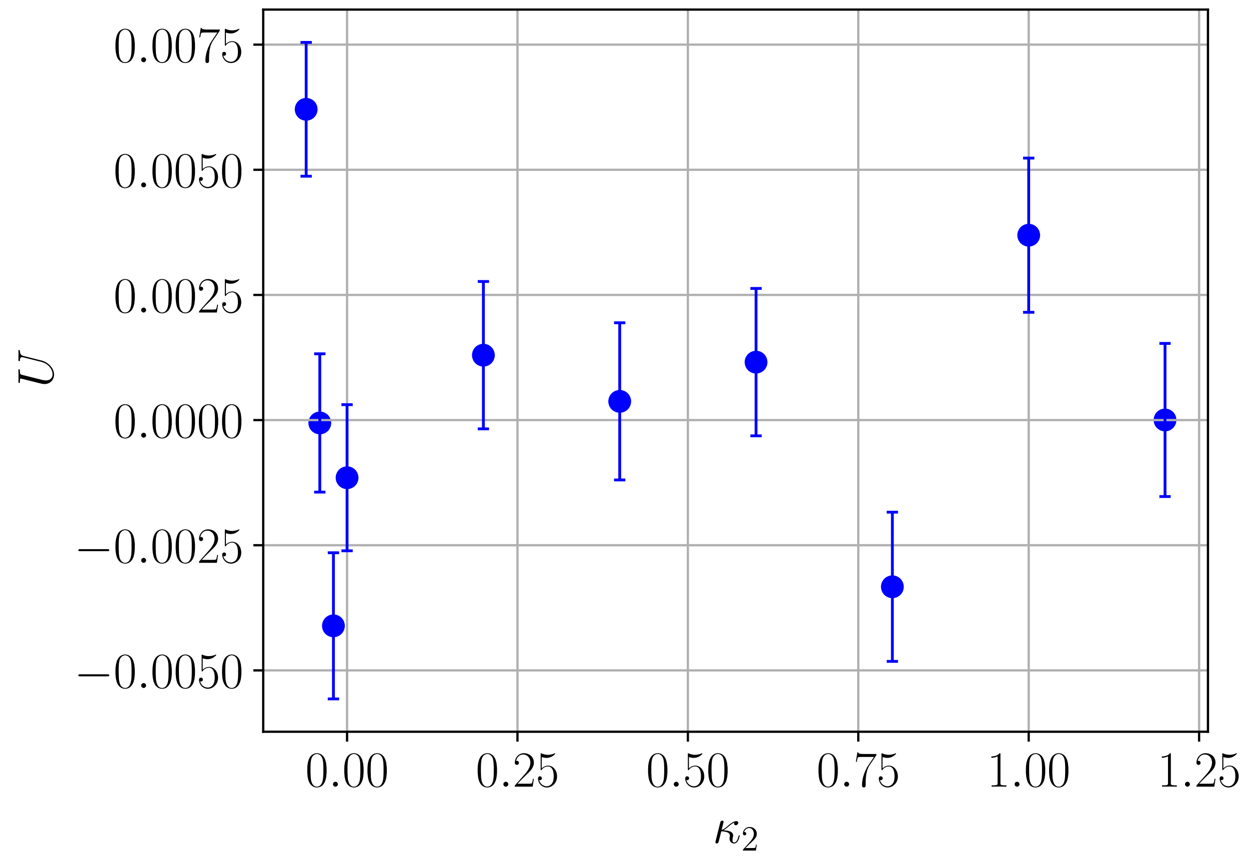


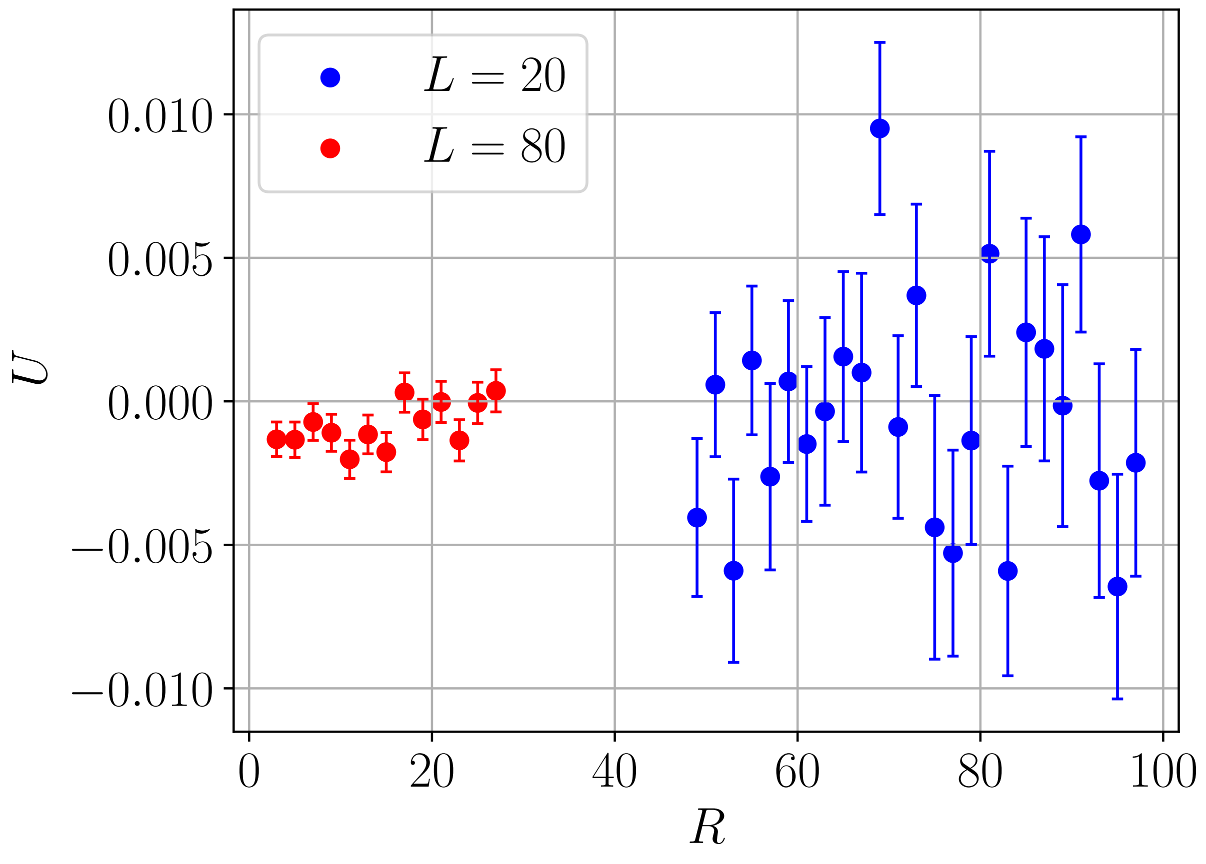


$\sigma = 100$ $R \gg L = 20$ $L = 80 \gg R$

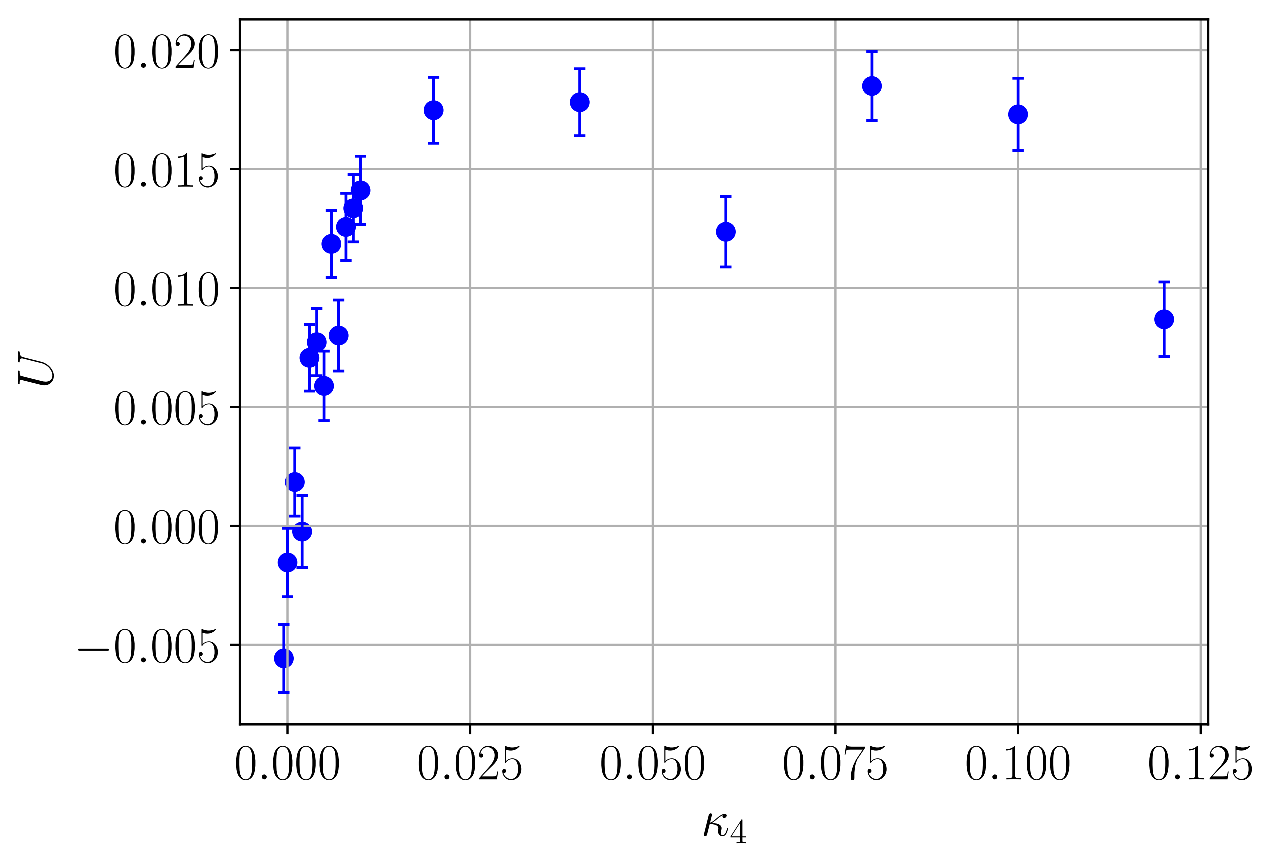
BEYOND NG: GAUSSIANITY

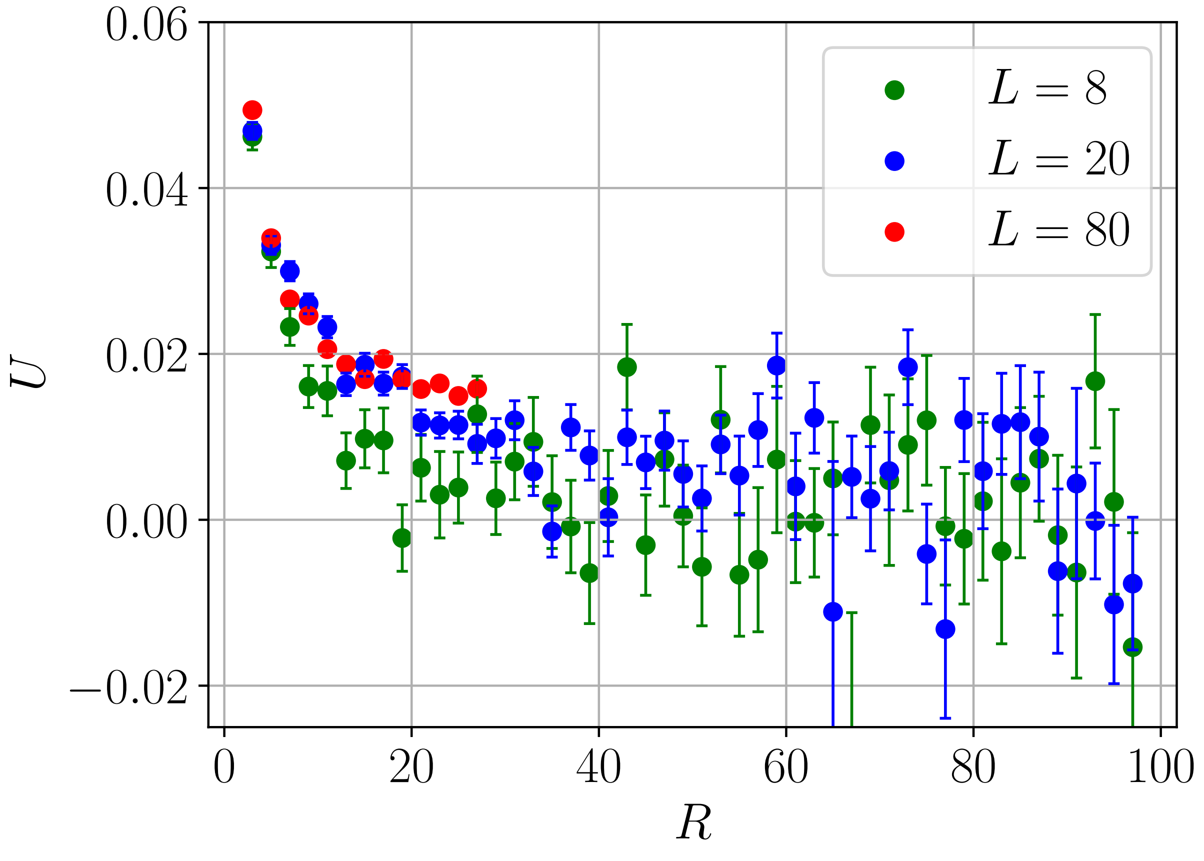
$\sigma = 100$





$$S_{NG} + \kappa_2 \mathscr{K}^2$$





$$S_{NG} + \kappa_4 \mathscr{K}^4$$

OUTLOOKS

- Flow-based sampler can be successful applied to sample Lattice EST:
  1. Numerical solution of the high temperature NG width
  2. Numerical studies of the Beyond NG EST and new interesting observable

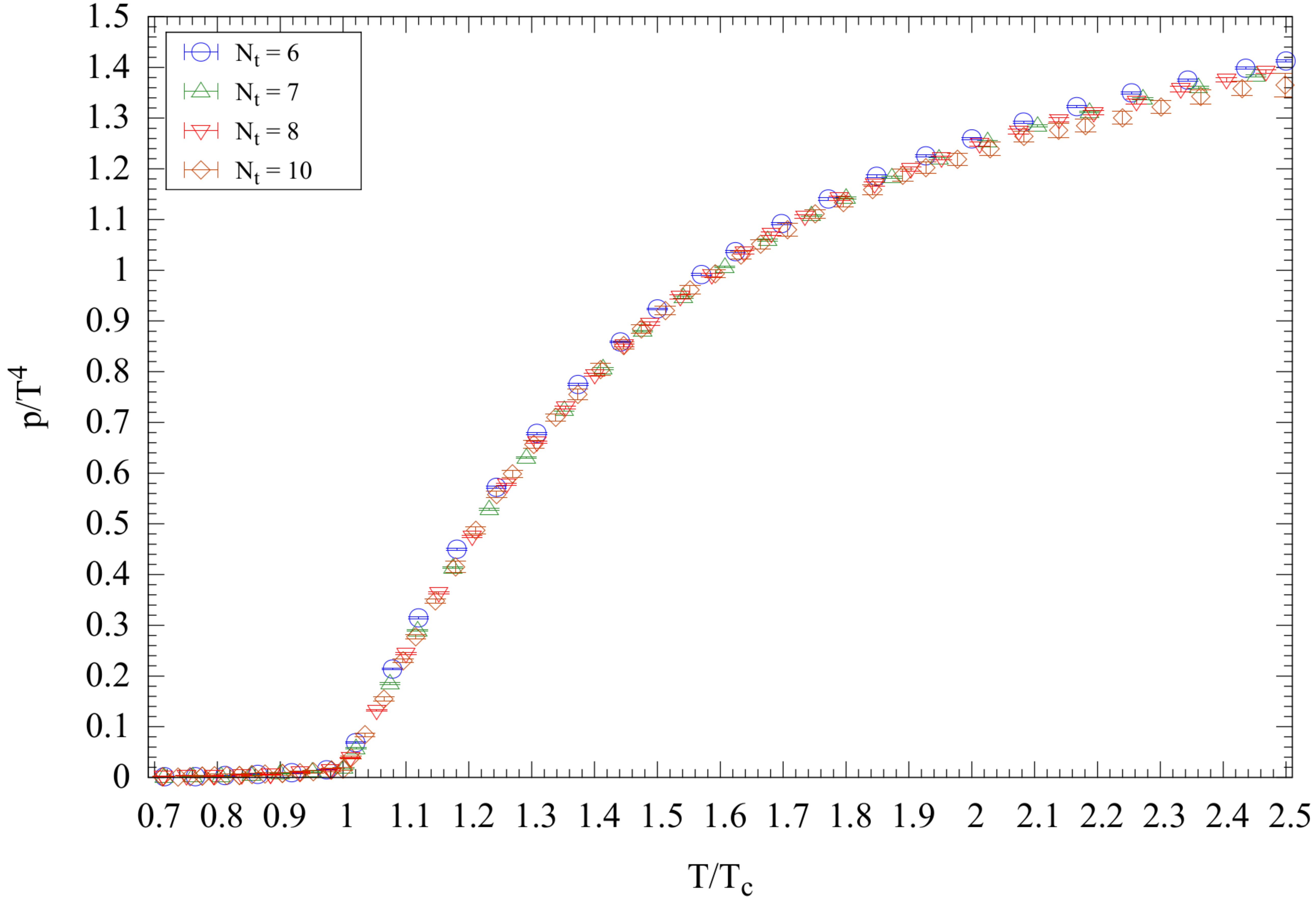
- EST provided a challenging laboratory for SNFs:
  1. Toward new applications in lattice field theory (see talk by A. Nada on SNFs for $SU(3)$)

THANK YOU FOR YOUR ATTENTION!

NE-MCMC FOR LFT

Jarzynksi's equality has been exploited to obtain state-of-the-arts results in LFT:

- Interface free energy.
  [Caselle et al.; 1604.05544]
- $SU(3)$ e.o.s.
  [Caselle et al.; 1801.03110]
- Running coupling
  [Francesconi et al.; 2003.13734]
- Entanglement entropy
  [Bulgarelli and Panero; 2304.03311, 2404.01987]
- Topological freezing
  [Bonanno et al.; 2402.06561]



STOCHASTIC NORMALIZING FLOWS: W_d

Forward and Reverse transition probabilities of NF layers can be written as:

$$P[\phi_n \to \phi_{n+1}] = \delta(\phi_{n+1} - g_\theta^n(\phi_n)) \qquad P[\phi_{n+1} \to \phi_n] = \delta(\phi_n - (g_\theta^n)^{-1}(\phi_{n+1}))$$

And satisfies:

$$\ln\big(P[\phi_{n+1} \to \phi_n]/P[\phi_n \to \phi_{n+1}]\big) = \ln\big(q_n(\phi_n)/q_{n+1}(\phi_{n+1})\big) = \ln|\det J_{g^n}(\phi_n)|$$

We have now:

$$W_d^\theta = W_\theta(\phi_0, \ldots, \phi_N) - \Delta F = S_N(\phi_N) - S_0(\phi_0) - Q_\theta - \Delta F$$

Where:

$$Q_\theta = \sum_{n=0}^{N-1} \left(S_{n+1}(\phi_{n+1}) - S_{n+1}(\phi_n) + \ln|\det J_{g_\theta^n}| \right)$$

[Wu+; 2002.06707],[Caselle, E.C., Nada, Panero; 2201.08862]

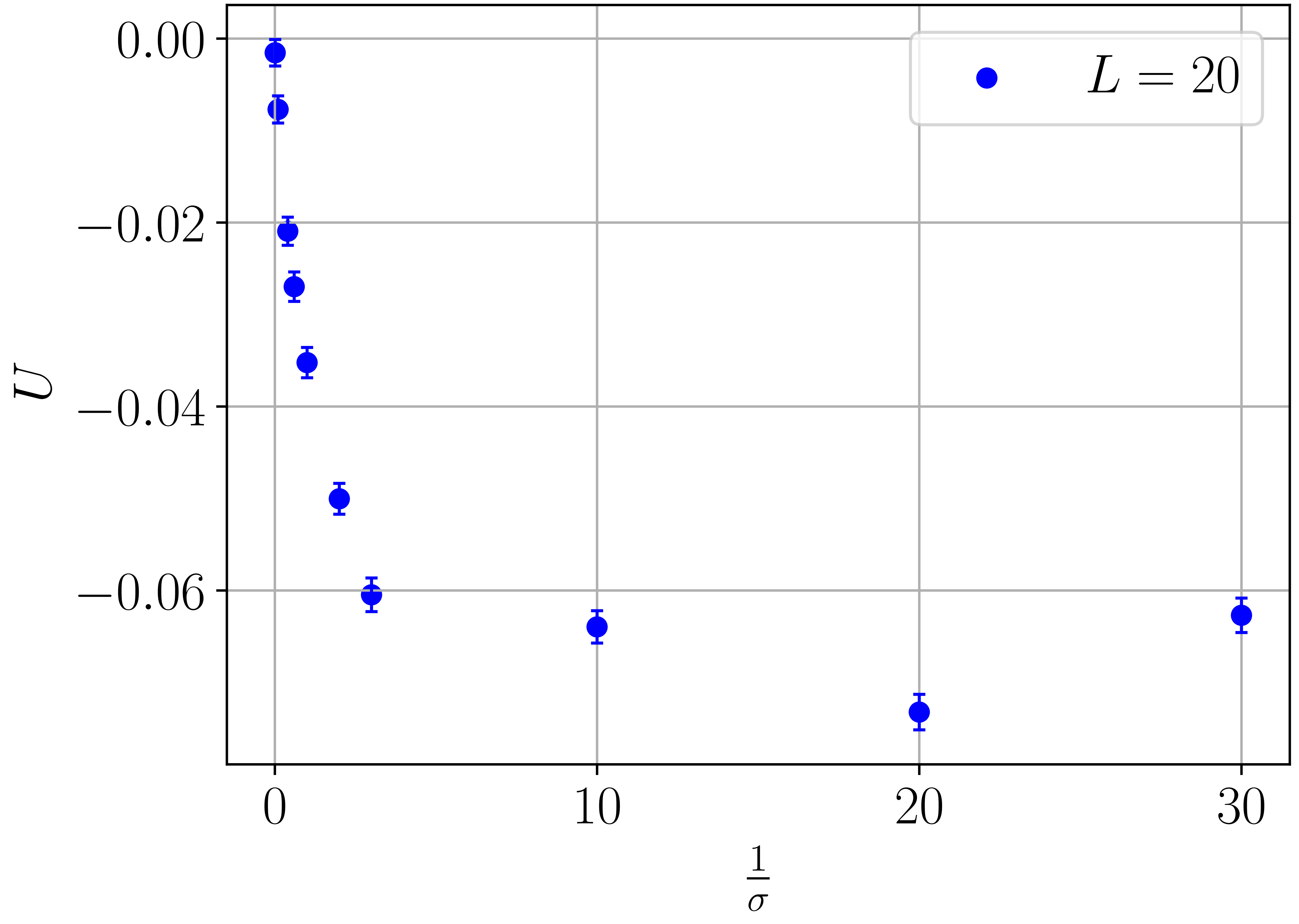
SNFS: RELATED WORKS

- Annealed Importance Sampling: Equivalent to Jarzynski's equality. Used in the original SNF paper
  [Neal; physics/9803008]
- Sequential Monte Carlo: Generalization of AIS.
  [Dai+; 2007.11936]
- SNF idea reworked in CRAFT
  [Matthews+; 2201.13117]
- An hybrid (deterministic/stochastic) approach with no neural networks has been proposed also by Jarzynksi in 2011
  [Vaikuntanathan and Jazynski; 1101.2612]
- FAB: combination of NFs and AIS.
  [Midgley+; 2208.01893]
- Exact work for discretized Langevin dynamics.
  [Sivak+; 1107.2967]

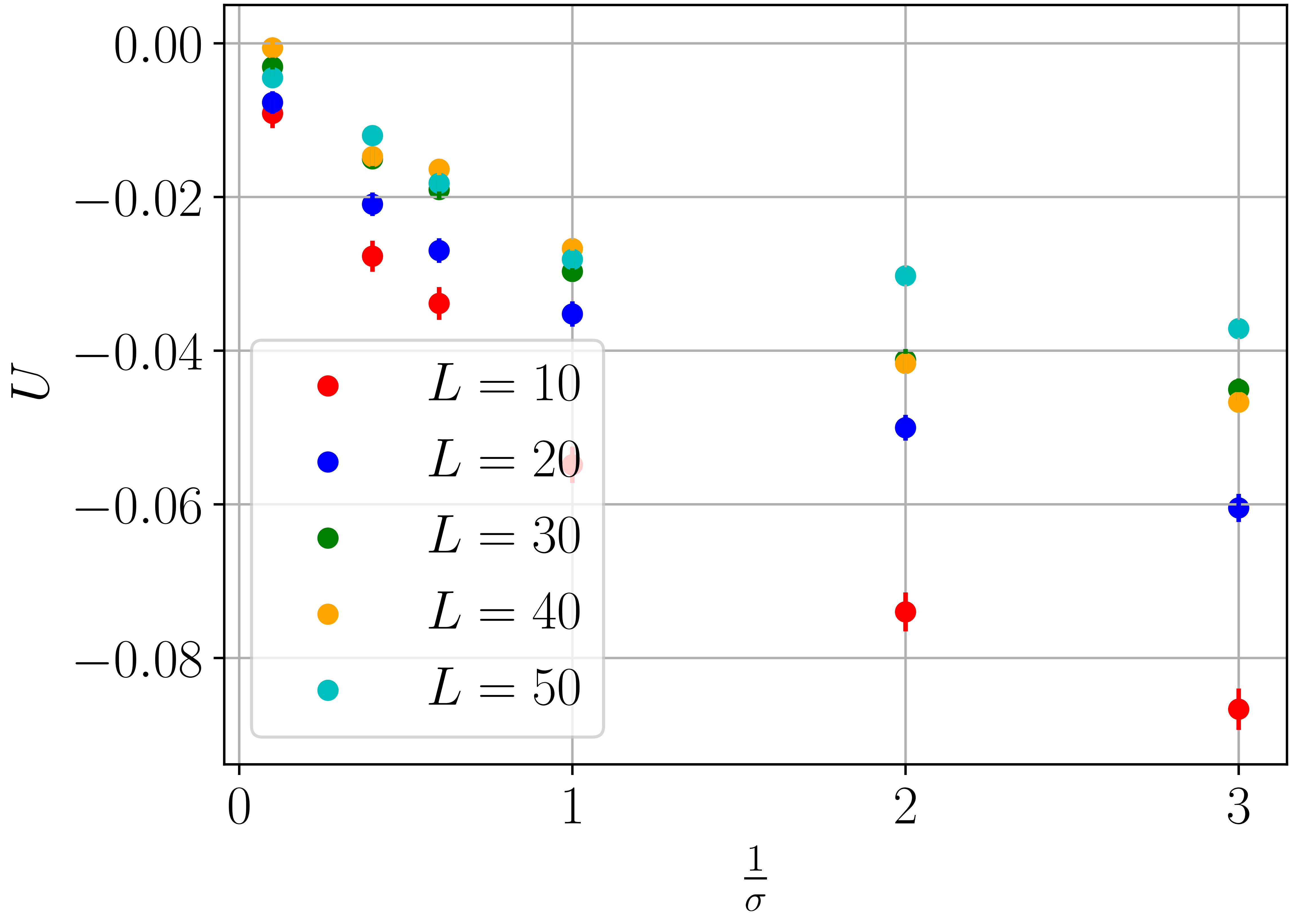
TECHNICAL DETAILS: NG

- Prior massless free boson and linear protocol in $t = 1/\sigma \rightarrow$ Inspired by Irrelevant Perturbations
- HMC for stochastic updates
- Affine coupling layers, 3 convolutional layers with $3 \times 3 \times 16$ kernels and a two channels output layer. Each blocks (even-odd) share the same network

NG BINDER 2



$L = 20$
U
$\frac{1}{\sigma}$
0.00
−0.02
−0.04
−0.06
0
10
20
30



$L = 10$
$L = 20$
$L = 30$
$L = 40$
$L = 50$
U
$\frac{1}{\sigma}$
0.00
−0.02
−0.04
−0.06
−0.08
0
1
2
3