Sampling Nambu-Goto theory using Normalizing Flows

Elia Cellini

Università degli Studi di Torino/Istituto Nazionale Fisica Nucleare

July 31th 2023

Lattice 2023
Fermilab

Based on:
arXiv:2307.01107
M. Caselle, E.C. and A. Nada

Outline

Nambu-Goto String

Effective String Theory

Effective string theory (EST) is a non-perturbative framework that provide an effective description of the confining flux tube in term of vibrating string. In particular, the correlator between two Polyakov loops is related to the full partition function of an EST.

$$\langle P(0)P^\dagger(R)\rangle = \int D\phi e^{-S_{eff}} \equiv Z(L, R, \sigma).$$

Effective String Theory

Effective string theory (EST) is a non-perturbative framework that provide an effective description of the confining flux tube in term of vibrating string. In particular, the correlator between two Polyakov loops is related to the full partition function of an EST.

$$\langle P(0)P^\dagger(R)\rangle = \int D\phi e^{-S_{eff}} \equiv Z(L, R, \sigma).$$

- EST anomalous at quantum level → **effective, large-distance description** of Yang-Mills theories.
- The most natural choice for S_{eff} is the **Nambu-Goto (NG)** string **[Nambu; 1974]**,**[Goto; 1971]** → universal up to term of the order R^{-5}.
- Main focus of the comunity → **theories "beyond" NG string**

For recent reviews see **[Aharony and Komargodski; 1302.6257][Brandt and Meineri; 1603.06969][Caselle; 2104.10486]**.
See also the **talk by Conghuan Luo**, 07/31 (Vacuum Structure and Confinement).

Observables

The main observables we want to compute are:

- The **free energy** $-logZ$: directly associated with the interquark potential.

$$V(R,L) = -\frac{1}{L}\log\langle P(0)P^{\dagger}(R)\rangle$$

- The **"width"** σw^2: measures the density of chromoelectric flux tube.

$$w^2 = \frac{\sum_{\vec{h}} \vec{h}^2 \langle \varphi(\vec{h}, R, L)\rangle}{\sum_{\vec{h}} \langle \varphi(\vec{h}, R, L)\rangle}$$

$$\langle \varphi(\vec{h}, R, L)\rangle = \frac{\langle P(0)P^{\dagger}(R)U_p(\vec{h})\rangle_L}{\langle P(0)P^{\dagger}(R)\rangle_L} - \langle U_p(\vec{h})\rangle_L$$

Lack of numerical methods

The free energy of the NG theory is well known, however, much less is known for the NG width and the free energy of theories beyond the NG string. Moreover, it still lacks an efficient numerical method that can be used to study EST where analytical studies are not possible.

Problems:

- **Non-linearity** of the actions.
- Direct estimation of the **partition functions**.

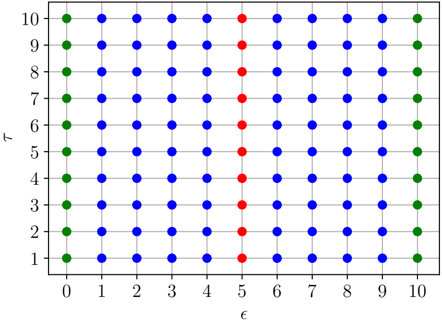
$\rightarrow$ Our proposal: **Normalizing Flows** + **Lattice regularization of EST**.

Nambu-Goto String on the Lattice

In the $d = 2 + 1$ case, using a "physical gauge" the NG action can be regularized on the lattice as **[Caselle et al.; 2023]**:

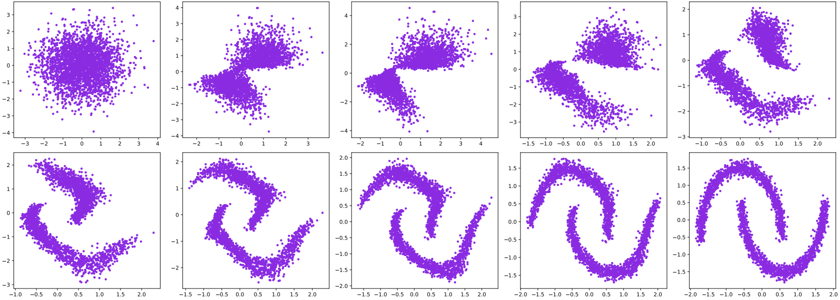
$$S_{NG}[\phi] = \sigma \sum_{x \in \Lambda} \left[\sqrt{1 + (\partial_\mu \phi)^2 / \sigma} - 1 \right]$$

where Λ is a square lattice of size $L \times R$ with step $a = 1$, $\phi(x) = \phi(\tau, \epsilon) \in \mathbb{R}$ and boundary conditions: $\phi(\tau + L, \epsilon) = \phi(\tau, \epsilon)$ and $\phi(\tau, 0) = \phi(\tau, R) = 0$.



The width of the string can be computed as: $\sigma w^2 = \langle \phi^2(\tau, R/2) \rangle_{\tau, \phi}$

Normalizing Flows



Normalizing Flows

A **Normalizing flows (NF)** **[Rezende and Mohamed; 2015]** g_θ is a parametric, invertible and differentiable function that maps an easy-to-model prior distribution $q_0(z)$, $z \in \mathbb{R}^n$, to an inferred distribution q_θ which approximate the target $p(\phi)$, $\phi \in \mathbb{R}^n$.

$$g_\theta : \; q_0 \to q_\theta \simeq p \qquad\qquad \phi = g_\theta(z) \qquad\qquad q_\theta(\phi) = q_0(g^{-1}(\phi))|J_g|^{-1}$$

Normalizing Flows

A **Normalizing flows (NF)** **[Rezende and Mohamed; 2015]** g_θ is a parametric, invertible and differentiable function that maps an easy-to-model prior distribution $q_0(z)$, $z \in \mathbb{R}^n$, to an inferred distribution q_θ which approximate the target $p(\phi)$, $\phi \in \mathbb{R}^n$.

$$g_\theta : \; q_0 \to q_\theta \simeq p \qquad\qquad \phi = g_\theta(z) \qquad\qquad q_\theta(\phi) = q_0(g^{-1}(\phi))|J_g|^{-1}$$

NFs can be trained to $q_\theta \simeq p(\phi)$ with $p(\phi) = \frac{1}{Z}\exp(-S[\phi])$ **[Albergo et al.; 2019]**,**[Noé et al.; 2019]** by minimizing the reverse Kullback-Leibler divergence:

$$D_{KL}(q_\theta||p) = \int d\phi q_\theta(\phi) \log \frac{q_\theta(\phi)}{p(\phi)} \geq 0.$$

A **Normalizing flows (NF)** **[Rezende and Mohamed; 2015]** g_θ is a parametric, invertible and differentiable function that maps an easy-to-model prior distribution $q_0(z)$, $z \in \mathbb{R}^n$, to an inferred distribution q_θ which approximate the target $p(\phi)$, $\phi \in \mathbb{R}^n$.

$$g_\theta : \ q_0 \to q_\theta \simeq p \qquad \phi = g_\theta(z) \qquad q_\theta(\phi) = q_0(g^{-1}(\phi))|J_g|^{-1}$$

NFs can be trained to $q_\theta \simeq p(\phi)$ with $p(\phi) = \frac{1}{Z}\exp(-S[\phi])$ **[Albergo et al.; 2019]**,**[Noé et al.; 2019]** by minimizing the reverse Kullback-Leibler divergence:

$$D_{KL}(q_\theta || p) = \int d\phi q_\theta(\phi) \log \frac{q_\theta(\phi)}{p(\phi)} \geq 0.$$

Observables can be computed using a re-weighting procedure also called **Importance Sampling (IS)** in machine learning field **[Nicoli et al.; 2020]**:

$$\langle \mathcal{O} \rangle_{\phi \sim p} = \frac{1}{Z} \langle \mathcal{O} \tilde{w} \rangle_{\phi \sim q_\theta} \qquad Z = \langle \tilde{w} \rangle_{\phi \sim q_\theta} \qquad \tilde{w} = \frac{e^{-S[\phi]}}{q_\theta(\phi)}$$

Architectures

1. **Continuous NFs (CNFs) [Chen et al.; 2018][Gerdes et al.; 2022][Caselle et al.; 2023]**:

$$\frac{d\phi(t)_x}{dt} = \sum_{y,d} W_{x,y,d} K(t)_d \phi(t)_y$$

where $K(t) \in \mathbb{R}^D$ is a temporal kernel of D Fourier coefficients.

2. **Physics-Informed Stochastic NFs (PI-SNFs) [Wu et al.; 2020][Caselle et al.; 2022][Abbott et al.; 2022]** (see **talk by A. Nada**):

$$\phi_0 \to g_1(\phi_0) \stackrel{P_{\eta_1}}{\to} \phi_1 \to g_2(\phi_1) \stackrel{P_{\eta_2}}{\to} \dots \stackrel{P_{\eta_N}}{\to} \phi_N = \phi$$

- $q_0 = e^{S_0}/Z_0$ and $S_0 = \frac{1}{2}\sum_{x\in\Lambda}(\partial_\mu\phi)^2$
- $S_{\eta_i} = \sigma_i \sum_{x\in\Lambda}\left[\sqrt{1+(\partial_\mu\phi)^2/\sigma_i} - 1\right]$ with $\sigma_i > \sigma_{i+1}$

Inspired by the $T\bar{T}$ integrable irrelevant perturbation **[Cavaglià et al.; 2016][Smirnov and Zamolodchikov; 2016]**

Numerical results

Goals of the numerical studies

The goals of our numerical studies are:

- Provide a **proof of concept** of the feasibility of the application of NFs as a sampler for EST with CNFs **[Caselle et al.; 2023]**.
- **Prove a conjecture** about the non-perturbative solution of the NG width in the high temperature regime **[Caselle; 2010]** using PI-SNF.

We accepted all the NFs with:

$$ESS = \frac{\langle \tilde{w} \rangle^2}{\langle \tilde{w}^2 \rangle} \geq 0.1$$

We run the models on NVIDIA Tesla V100 GPU of the Marconi100 (CINECA).

A proof of concept with CNFs

The CNFs we used in our work couldn't reach small values of σ, thus we focus on **the $\sigma \to \infty$ expansion** of the NG action:

$$S_{NG} \sim S_{FB} + O(\sigma^{-1})$$

where:

$$S_{FB} = \frac{1}{2} \sum_{x \in \Lambda} (\partial_\mu \phi)^2$$

The **finite size analytical solutions** of $-\log Z_{FB}$ can be found using a Gaussian integration **[Caselle et al.; 2023]**:

$$-\log Z_{FB} = A_{FB} RL + C_{FB} L + \log \eta(\xi)$$

where:

$$A_{FB} = -0.3358177... \quad C_{FB} = 0.478252....$$

and:

$$\eta(\xi) = q^{\frac{1}{24}} \prod_{n=1}^{\infty} (1 - q^n) \quad ; \quad q = e^{2\pi i \xi} \quad ; \quad \xi = i \frac{L}{2R}$$

Partition function at low-temperature $L \gg R$

global fit in (σ, L) with $\sigma \geq 40$

$$-\log Z = \left(a_{LT}^{(0)}(R) + \frac{a_{LT}^{(1)}(R)}{\sigma} + \frac{a_{LT}^{(2)}(R)}{\sigma^2} + \frac{a_{LT}^{(3)}(R)}{\sigma^3} \right) L$$

and then

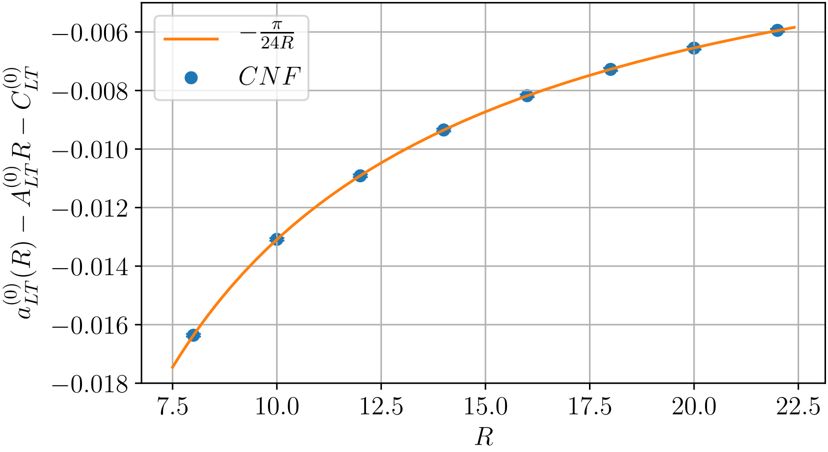
$$a_{LT}^{(0)}(R) = A_{LT}^{(0)} R + \frac{B_{LT}^{(0)}}{R} + C_{LT}^{(0)}$$

and we found:

	$A_{LT}^{(0)}$	$B_{LT}^{(0)}$	$C_{LT}^{(0)}$	$\chi^2/d.o.f.$
CNFs	$-0.335820(2)$	$-0.1309(2)$	$0.47822(4)$	0.93
Theory	$-0.3358177...$	$-0,13089969...$	$0.478252...$	

Partition function at low-temperature $L \gg R$

Plot of the Dedekind prediction $-\frac{\pi}{24R}$ (Lüscher term) compared to the numerical simulations:



See **[Caselle et al.; 2307.01107]** for other studies in the large σ region with CNFs.

Width at high temperature $R \gg L$

The analytical solution for σw^2 is well known only up to the order σ^{-1} **[Lüscher et al.; 1981][Caselle and Allais; 2009][Gliozzi, Pepe and Wiese; 2010][Caselle et al.; 2023]**:

$$\sigma w^2 = \left(1 + \frac{\pi}{6}\frac{1}{\sigma L^2}\right)\frac{R}{4L} + \dots$$

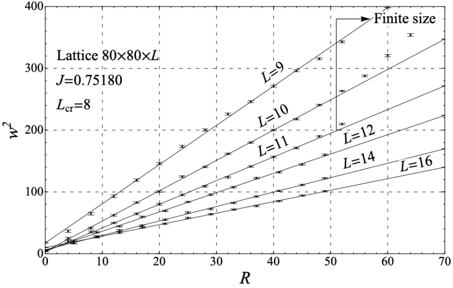


Figure 2: *Flux tube thickness as a function of the interquark distance for various values of the inverse temperature L.*

A conjecture **[Caselle; 2010]** states that the non-perturbative solution in this regime is:

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

PI-SNF prediction

We run several simulations the Nambu-Goto string in the HT regime with $\sigma = 0.1$.
Fit in R

$$\sigma w^2 = e(L)R + f(L)$$

then fit in L:

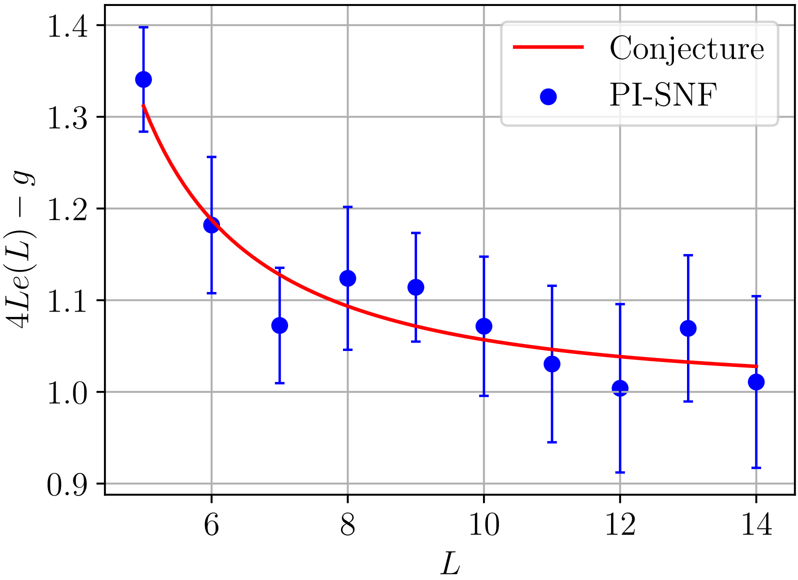
$$e(L) = \left(\frac{1}{\sqrt{1 - \frac{f}{0.1L^2}}} + g \right) \frac{1}{4L}$$

and we found:

	f	g	$\chi^2/d.o.f.$
PI-SNF	1.09(8)	12.28(2)	0.25
Theory	1.047197...		

New results for the EST field obtained with NFs!

PI-SNF prediction



Outlook

Outlook

- We showed that NFs are able to sample efficiently from the probability distribution of the Nambu-Goto EST.
- We studied the non-perturbative solution of the NG width in the HT regime.
- Toward the study of the theories beyond the Nambu-Goto string using flows-based sampler as leading numerical method!

$$S_{BNG} = S_{NG} + \gamma \mathcal{K}^4$$

Outlook

- We showed that NFs are able to sample efficiently from the probability distribution of the Nambu-Goto EST.
- We studied the non-perturbative solution of the NG width in the HT regime.
- Toward the study of the theories beyond the Nambu-Goto string using flows-based sampler as leading numerical method!

$$S_{BNG} = S_{NG} + \gamma \mathcal{K}^4$$

Outlook

- We showed that NFs are able to sample efficiently from the probability distribution of the Nambu-Goto EST.
- We studied the non-perturbative solution of the NG width in the HT regime.
- Toward the study of the theories beyond the Nambu-Goto string using flows-based sampler as leading numerical method!

$$S_{BNG} = S_{NG} + \gamma \mathcal{K}^4$$

Thank you for your attention!

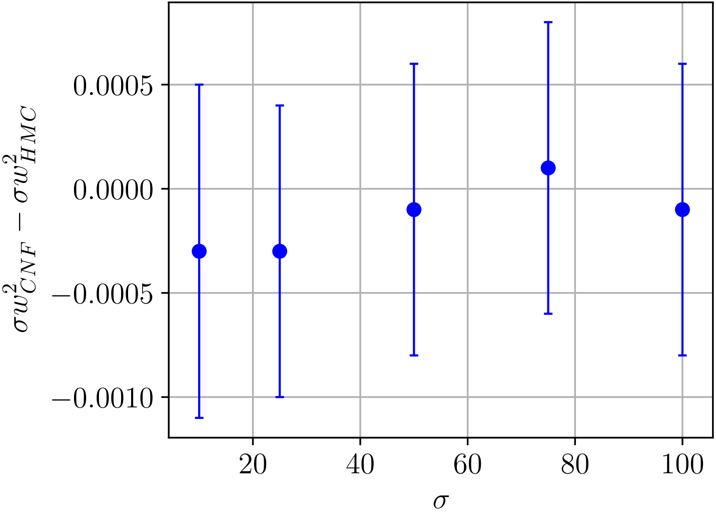
arXiv:2307.01107

CNFs notebook

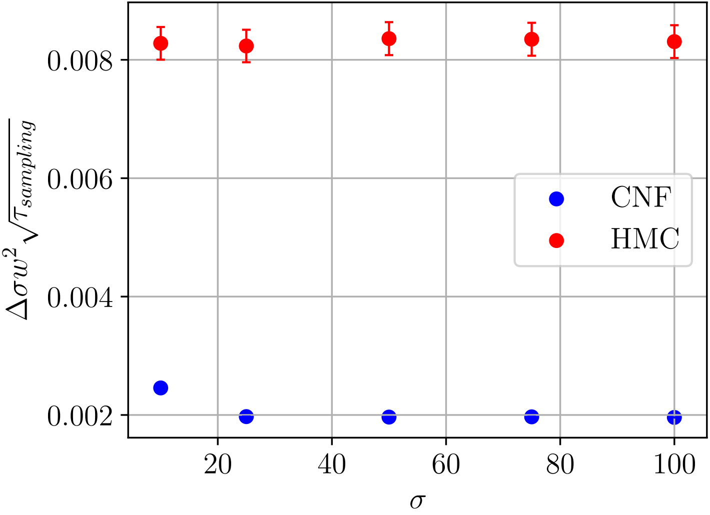
See talks by:

- 07/31: Gurtej Kanwar, Sam Foreman, Alessandro Nada, Nobuyuki Matsumoto, Rayan Abbott, Jacob Finkenrath.
- 08/01: Kim Nicoli, Joe Marsh Rossney (poster), Christopher Kirwan (poster).
- 08/03: David Albandea, Christopher Chamness, Julian Urban, Denis Boyda, Daniel Hackett.

Comparison with HMC: Bias



Comparison with HMC: Sampling time



Exploiting **Neural Ordinary Differential Equations (NODE) [Chen et al.; 2018]** is possible to build **Continuous NFs (CNFs)** in which g_θ is the solution of an ODE parameterized by a neural network V_θ:

$$\frac{d\phi(t)}{dt} = V_\theta(\phi(t), t)$$

with

$$\phi(t=0) = z \sim \mathcal{N}(0, \mathbb{1}/2) \qquad \phi = \phi(t=T) = ODESOLVER(V_\theta, \phi(0), [0, T])$$

The density of the generated samples can be computed through the ODE:

$$\frac{d \log q_\theta(\phi(t))}{dt} = -(\nabla \cdot V_\theta)(\phi(t), t)$$

We used a **linear vector field**:

$$V_\theta(\phi(t), t)_x = \sum_{y,d} W_{x,y,d} K(t)_d \phi(t)_y$$

where $K(t) \in \mathbb{R}^D$ is a temporal kernel of D Fourier coefficients.
Inspired by **[Gerdes et al.; 2022]** and used in **[Caselle et al.; 2023]**.

Physics-Informed Stochastic Normalizing Flows

Stochastic Normalizing Flows (SNFs) [Wu et al.; 2020][Caselle et al.; 2022] (see **talk by A. Nada**)

$$\phi_0 \to g_1(\phi_0) \overset{P_{\eta_1}}{\to} \phi_1 \to g_2(\phi_1) \overset{P_{\eta_2}}{\to} \dots \overset{P_{\eta_N}}{\to} \phi_N = \phi$$

$$-\ln \tilde{w} = \sum_{n=0}^{N-1} S_{\eta_{n+1}}(\phi_{n+1}) - S_{\eta_{n+1}}(g_n(\phi_n)) + \ln |\det J_n(\phi_n)|$$

where g_i are affine coupling layers and P_{η_i} are the transition probabilities of HMC updates.

Physics-informed algorithm design lead to better models **[Abbott et al.; 2022]**:

- $q_0 = e^{S_0}/Z_0$ and $S_0 = \frac{1}{2}\sum_{x\in\Lambda}(\partial_\mu \phi)^2$
- $S_{\eta_i} = \sigma_i \sum_{x\in\Lambda}\left[\sqrt{1 + (\partial_\mu \phi)^2/\sigma_i} - 1\right]$ with $\sigma_i > \sigma_{i+1}$

Inspired by the $T\bar{T}$ integrable irrelevant perturbation **[Cavaglià et al.; 2016][Smirnov and Zamolodchikov; 2016]** $\to$ **Physics-Informed SNFs (PI-SNFs)**