# Deep Finite Temperature Bootstrap

Vasilis Niarchos - University of Crete

based on 2508.08560

with **C. Papageorgakis, A. Stratoudakis, M. Woolley**

Progress of Theoretical Bootstrap, YITP, Kyoto

12 November, 2025

# Motivation

In many (CFT) bootstrap problems we need to analyze sum rules of the form

$$\sum_J \sum_\Delta \vec{x}_{\Delta,J} = 0$$

which contain an infinite number of CFT data labeled by Δ, J.

e.g. 4-point bootstrap

$$\sum_{\mathcal{O}} \mathrm{C}_{\mathcal{O}_{\Delta,J}} \mathrm{F}_{\mathcal{O}_{\Delta,J}}(z,\bar{z}) = 0$$

Constructing explicit solutions and scanning over them is extremely hard!

The **linear functional** method [Rattazi-Rychkov-Tonni-Vichi '08] is a very successful workaround.

It yields **feasibility constraints** that can be analysed with a powerful (**convex**) semi-definite programming algorithm (eg SDPB, [Simmons-Duffin, '15])

**Nevertheless,** novel approaches aiming at **constructive solutions** might be **useful** for the following reasons:

- **Exploration** and feedback to more rigorous methods

  (high-dimensional searches, intuition into non-extremal solutions…)

- **Re-organization** and development of novel search strategies [see D. Poland talk]

  (focus on new CFT quantities or discovery of new types of analytic results that can drive the searches more efficiently: **analytics**)

- Discovery of new **efficient numerical techniques**: **numerics**

  (scaling the multi-correlator bootstrap, improving numerical stability…)

- New, currently inaccessible situations **without positivity**

  (finite-T, defects, higher-point bootstrap, non-unitary theories….)

**Constructive solutions** are hard because it is difficult to tame the infinite dimensionality without losing control of the systematic error!

**Previous approaches:**

- Hard truncations: drop high-energy tail contributions [Gliozzi '13, …]

  typically uncontrolled *(but see recent boundary bootstrap in 3d O(N) vector model by **Hu-Li, '05**)*

- Soft truncations: model high-energy tail contributions

  [e.g. VN-Papageorgakis-Richmond-Stapleton-Woolley '23 [1d defects] *(effective operators…)*,

  Marchetto-Miscioscia-Pomoni '23 [thermal]

  *(Tauberian theorems…)*]

  better, but also hard to control

SDPB (~100-1000 digit precision)

| $C_1^2$ | $C_2^2$ | $C_3^2$ |
|---|---|---|
| $0.294014873 \pm 4.88{\cdot}10^{-8}$ | $0.039788 \pm 4.10{\cdot}10^{-4}$ | $0.146757 \pm 5.82{\cdot}10^{-4}$ |
| $0.294014228 \pm 6.77{\cdot}10^{-7}$ | $0.041832 \pm 1.86{\cdot}10^{-3}$ | $0.144100 \pm 2.39{\cdot}10^{-3}$ |

$\lambda = (4\pi)^2 \simeq 157.91$

soft truncation (16 digit precision)

**New proposed approach:** bootstrap the **tails of high-energy data**

(reconstruct (an approximation of) the full correlation function)

General idea

$$\sum_J \sum_\Delta \vec{\mathcal{X}}_{\Delta,J} = 0$$

Finite spectrum: CFT data **D'**
**(as small as you like!)**

$$\sum_{J>J_*} \sum_\Delta \vec{\mathcal{X}}_{\Delta,J}$$

`High'-spin part

Express with **dispersion relations**

**Approximation** with finite spectrum **D''**

$$\sum_{J\leq J_*} \sum_\Delta \vec{\mathcal{X}}_{\Delta,J}$$

`Low'-spin part

$$\sum_{J\leq J_*} \sum_{\Delta\leq\Delta_*(J)} \vec{\mathcal{X}}_{\Delta,J}$$

$$\sum_{J\leq J_*} \sum_{\Delta>\Delta_*(J)} \vec{\mathcal{X}}_{\Delta,J}$$

Low-spin,
high-energy spectrum

## The low-spin, high-energy part

$$\sum_{J\leq J_*}\sum_{\Delta>\Delta_*(J)}\vec{\mathcal{X}}_{\Delta,J}$$

$$\sum_{J<J_*}A_J(r)C_J^{(d)}(\cos\theta)$$

Finite number of unknown 1d functions $A_J(r)$

Call these unknown functional data $\mathbf{F}$

## Goal

**Approximations**

There is a finite-dimensional vector $\sum_J \sum_\Delta \vec{\mathscr{X}}_{\Delta,J}[\mathtt{D},\mathtt{F}]$ $(\mathtt{D} = \mathtt{D}' \cup \mathtt{D}'')$

- Minimize a semi-positive-definite functional (loss) $\mathscr{L}[\mathtt{D},\mathtt{F}]$ to determine solutions $(\mathtt{D}_*,\mathtt{F}_*)$ of the crossing equation

$$(\mathtt{D}_*,\mathtt{F}_*) = \min_{\mathtt{D},\mathtt{F}} \mathscr{L}(\mathtt{D},\mathtt{F}) \simeq \min_{\mathtt{D},\vec{\theta}} \mathscr{L}[\mathtt{D},\vec{\theta}]$$

- That requires a flexible way to vary over the functional data $\mathtt{F}$.

In what follows, we parametrize functions (1d vector fields) using **Neural Networks** (expressed in terms of a finite-dimensional vector $\vec{\theta}$ –typically, in thousands or millions of dimensions)

To summarize:

*we use a combination of dispersion relations + optimizable functions to parametrize unknown correlation functions (and the CFT data they entail). We try to fix these functions by imposing suitable constraints (crossing symmetry + other desired sum-rules or constraint)*

Approaches with some similarities have appeared in the recent literature of the

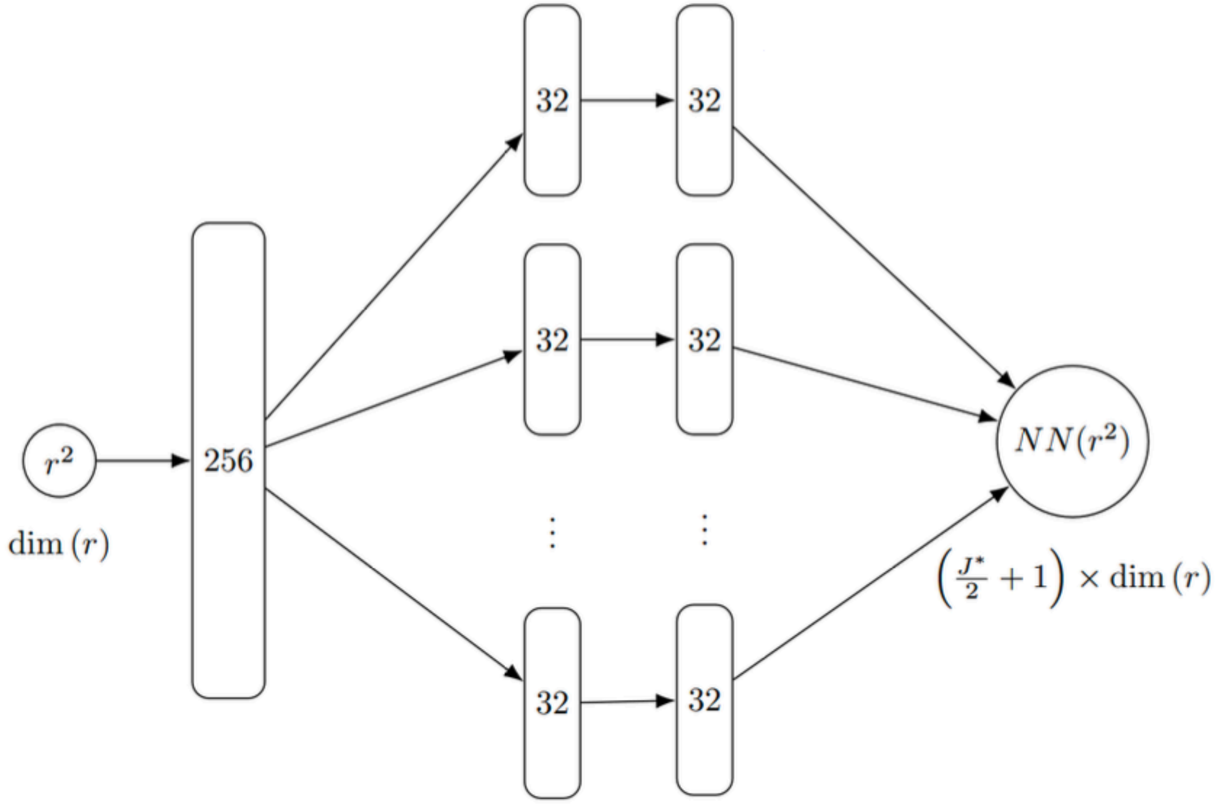
**(primal) S-matrix bootstrap**

## Note on Neural Networks: Multi-branch Multi-Layer-Perceptron (MLP) models

$$f_{\mathrm{MLP}}(\vec{x}) = (\mathbf{W}_{L-1} \circ \sigma \circ \mathbf{W}_{L-2} \circ \sigma \circ \cdots \mathbf{W}_1 \circ \sigma \circ \mathbf{W}_0)\, \vec{x}$$

$$x_n^{(\ell+1)} = W_n^m x_m^{(\ell)} + b_n$$

$\sigma$ is some non-linear activation function like tanh



## Note on loss functions

- $L_p$-norm (a rather obvious choice): If $\vec{\mathscr{F}}$ is the crossing vector we try to set to 0, then

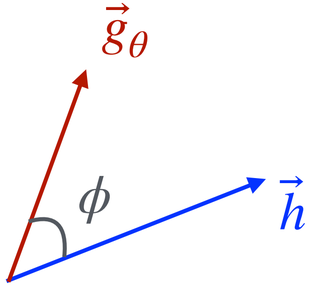
$$\mathscr{L}(\vec{\mathscr{F}}) := \left( \sum_{n=1}^{\dim \vec{\mathscr{F}}} \left| \mathscr{F}_n \right|^p \right)^{1/p} \qquad \text{for} \quad p \geq 1 \qquad\qquad (\textbf{abs-loss} \text{ for } p = 1)$$

- A (novel) alternative that works better: **dot-loss**

V1: D is empty (**no exposed CFT data**) and $\vec{\mathscr{F}} = \vec{g}_\theta - \vec{h}$ ($\vec{g}_\theta$ optimizable, $\vec{h}$ known)

$$\mathscr{L}(\vec{\mathscr{F}}(\theta)) = 1 - \frac{|\vec{g}_\theta \cdot \vec{h}|}{|\vec{g}_\theta||\vec{h}|} + \left| 1 - \frac{\vec{g}_\theta \cdot \vec{h}}{|\vec{h}|^2} \right|$$

$$0 \leq 1 - |\cos(\phi)| \leq 1$$



V2: D has one element (**1 exposed CFT datum a**) and

$$\vec{\mathcal{F}}(a,\vec{\theta}) = a\vec{f} + \vec{g}_\theta - \vec{h} \qquad (\vec{g}_\theta \text{ optimizable, } \vec{h} \text{ known, } \vec{f} \text{ known crossed block})$$

In this case:

$$\mathscr{L}(\vec{\mathcal{F}}(\theta)) = 1 - \frac{|(\vec{g}_\theta - \vec{h})\cdot\vec{f}|}{|\vec{g}_\theta - \vec{h}||\vec{f}|}$$

you optimize only with respect to the unknown tail functions and you get the coefficient a for free at the end

$$a = -\frac{(\vec{g}_{\theta_*} - \vec{h})\cdot\vec{f}}{|\vec{f}|^2}$$

# An application: Thermal bootstrap

## 1. The equations of thermal bootstrap

We consider 2-point functions of scalar primary operators in d-dimensional CFTs on $S^1_\beta \times \mathbb{R}^{d-1}$

$$g(\tau, |x|) = \langle \phi(x)\phi(0)\rangle_\beta \ , \quad |x| = \sqrt{\tau^2 + \vec{x}^2}$$

Convenient notation: for $\vec{x} = (\sigma,0,\ldots,0)$ set $z := \tau + i\sigma = rw$ , $\bar{z} := \tau - i\sigma = rw^{-1}$

Kubo-Martin-Schwinger condition: $g(\tau, r) = g(1-\tau, r) \quad (\beta = 1)$

or $$g(1-z,1-\bar{z}) = g(z,\bar{z})$$

The KMS condition is a non-trivial constraint for the thermal 1-point functions in thermal OPE:

$$g(rw, rw^{-1}) = \sum_{\mathcal{O}_{\Delta,J} \in \phi \times \phi} a_{\mathcal{O}_{\Delta,J}} C_J^{(\frac{d-2}{2})} \left( \frac{1}{2}(w + w^{-1}) \right) r^{\Delta - 2\Delta_\phi}$$

$$a_{\mathcal{O}} := \frac{f_{\phi\phi\mathcal{O}} b_{\mathcal{O}}}{c_{\mathcal{O}}} \frac{J!}{2^J (\frac{d-2}{2})_J}$$

One would like to use the KMS condition to determine the thermal OPE coefficients $a_{\mathcal{O}}$ assuming knowledge of the spectrum

In the context of the above low-high spin/low-high scaling dimension split the 2-point function becomes

$$
\begin{aligned}
g(rw, rw^{-1}) \;=\; & \sum_{J=0}^{J_*} \sum_{\Delta \le \Delta_*(J)} a_{\Delta,J}\, C_J^{(\nu)}\left(\frac{1}{2}(w+w^{-1})\right) r^{\Delta-2\Delta_\phi} \\
& + \sum_{J=0}^{J_*} A_{\Delta_*(J),J}(r)\, C_J^{(\nu)}\left(\frac{1}{2}(w+w^{-1})\right) \\
& + 2\left(\int_{-\infty}^{-r^{-1}} + \int_{r^{-1}}^{\infty}\right) dw'\, \mathcal{K}_{J_*}(w,w')\, \mathrm{Disc}\left[g(rw', rw'^{-1})\right] \;.
\end{aligned}
\tag{3.10}
$$

exposed data

low-spin/high-Δ data

tail functions

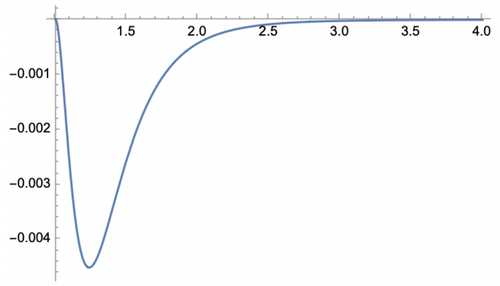
$$A_{\Delta_*(J),J}(r) := \sum_{\Delta > \Delta_*(J)} a_{\Delta,J}\, r^{\Delta-2\Delta_\phi}$$

high-spin data

$J_*$-dependent kernel

$$\mathcal{K}_{J_*}(w,w') := \frac{1}{2}\mathcal{K}(w,w') - w'^{-1}(w'-w'^{-1})^{2\nu}\left[\sum_{J=0}^{J_*} K_J\, C_J^{(\nu)}\left(\frac{1}{2}(w+w^{-1})\right) F_J(w'^{-1})\right],$$

where




$$\mathcal{K}(w,w') := \frac{1}{2\pi w'}\frac{w'^2-1}{(w'-w)(w'-w^{-1})}\;,$$

$$K_J := \frac{\Gamma(J+1)\Gamma(\nu)}{4\pi\Gamma(J+\nu)}\;,$$

$$F_J(w) = w^{J+d-2}\;{}_2F_1\left(J+d-2, \frac{d}{2}-1, J+\frac{d}{2}, w^2\right)\;.$$

Finite-T dispersion relations

Alday-Kologlu-Zhiboedov ’20

Stratoudakis ’24

Barrat-Bozkurt-Marchetto -Miscioscia-Pomoni ’25

## Comments

- $J_*$ is arbitrary. High $J_*$ suppresses the discontinuity and puts all the info into the tails

- The disc will be approximated using the crossed-channel OPE. This introduces errors:

  (a) Part of the $w'$-integral is outside the common region of convergence of the s- & t-channel OPEs

  (b) The crossed-channel OPE is truncated

  *Both issues are mitigated by the kernel $K_{J_*}(w,w')$ which suppresses the integral away from the branch points at $\pm r^{-1}$, as well as the higher-twist contributions*

- If the discontinuity is (partially) known, the corresponding information acts as a source guiding the search towards specific solutions

# Learning how Deep Thermal Bootstrap works

## 1. Generalized Free Fields (GFFs) as a testing ground

Simple case where everything is analytically known. For a scalar GFF with scaling dimension $\Delta_\phi$

$$g(z,\bar{z}) = \sum_{m=-\infty}^{\infty} \frac{1}{\left[(m-z)(m-\bar{z})\right]^{\Delta_\phi}}$$

$$[\phi\phi]_{n,J}, \quad n = 0, 1, ..., \quad J = 2\ell \quad \ell = 0, 1, ...$$

$$\Delta_{n,J} = 2\Delta_\phi + 2n + 2\ell$$

$$g(rw, rw^{-1}) = r^{-2\Delta_\phi} + \sum_{n=0}^{\infty}\sum_{\ell=0}^{\infty} a_{n,2\ell}\, C_{2\ell}^{(\nu)}\left(\frac{1}{2}(w+w^{-1})\right) r^{2(n+\ell)}$$

$$a_{n,J} = 2\zeta(2\Delta_\phi + 2n + J)\frac{(J+\nu)(\Delta_\phi)_{J+n}(\Delta_\phi-\nu)_n}{n!(\nu)_{J+n+1}}$$

$$\begin{aligned}\mathbb{T}_{\text{Disc}}^{(\text{approx})}[J_*; rw, rw^{-1}] \;=\; & 4\sin(\pi\Delta_\phi)\Big[\int_{r^{-1}}^{2r^{-1}} dw'\, \mathcal{K}_{J_*}(w,w')(rw'-1)^{-\Delta_\phi}(1-rw'^{-1})^{-\Delta_\phi} \\ & - \int_{-2r^{-1}}^{-r^{-1}} dw'\, \mathcal{K}_{J_*}(w,w')(-rw'-1)^{-\Delta_\phi}(1+rw'^{-1})^{-\Delta_\phi}\Big]\end{aligned}$$

**The (approximate) KMS condition** (only identity exposed)

$$r^{-2\Delta_\phi} - \tilde{r}^{-2\Delta_\phi} + \sum_{J=0}^{J_*} \left[ A_{\Delta_*(J),J;\boldsymbol{\theta}}(r)\, C_J^{(\nu)}\left(\frac{1}{2}(w+w^{-1}\right) - A_{\Delta_*(J),J;\boldsymbol{\theta}}(\tilde{r})\, C_J^{(\nu)}\left(\frac{1}{2}(\tilde{w}+\tilde{w}^{-1}\right) \right]$$

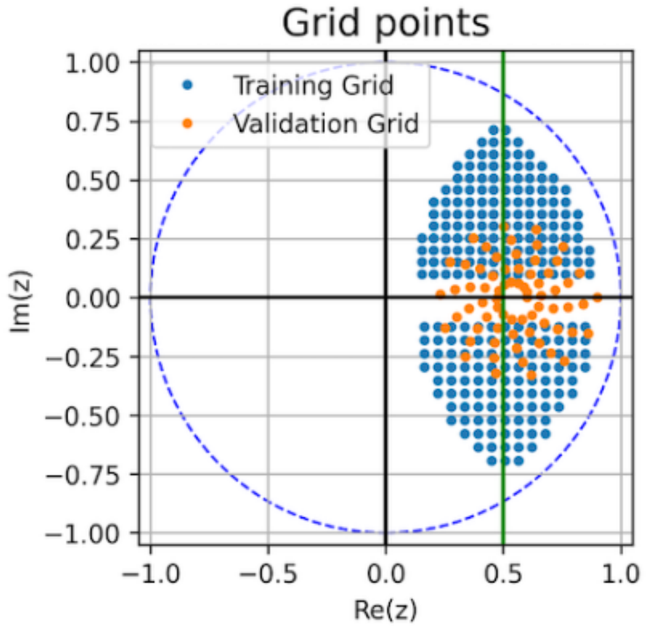
$$+\mathbb{T}^{(\text{approx})}_{\text{Disc}}[J_*; rw, rw^{-1}] - \mathbb{T}^{(\text{approx})}_{\text{Disc}}[J_*; \tilde{r}\tilde{w}, \tilde{r}\tilde{w}^{-1}] = 0 \ .$$

**(Non-zero spatial separation!)**

**Parameters**

- $d = 4,\ \Delta_\phi = 1.68$
- A grid of 243 points inside

$$|z| < 0.95, \quad |1-z| < 0.95$$



**Test 1. How good is the approximate KMS equation?**

Substitute the exact analytic GFF solution to the equation and check the loss as a function of $J_*$

| $J_*$ | $\mathcal{L}_{\overline{\text{abs}}}$ | $\mathcal{L}_{\text{dot}(0)}$ |
|---|---|---|
| 0 | 0.1177 | 0.0125 |
| 2 | 0.0300 | $1.3 \times 10^{-5}$ |
| 4 | 0.0067 | $1.6 \times 10^{-7}$ |
| 6 | 0.0013 | $3.3 \times 10^{-9}$ |
| 8 | 0.0002 | $0.8 \times 10^{-10}$ |

Now let's expose one operator

$[\phi\phi]_{0,0} = \phi^2$ does not contribute to the KMS condition

Interestingly, $[\phi\phi]_{1,0}$ comes together with $[\phi\phi]_{0,2}$!

$$
\begin{aligned}
&r^2 C_0^{(\nu)}\left(\frac{1}{2}(w+w^{-1})\right) - \tilde{r}^2 C_0^{(\nu)}\left(\frac{1}{2}(\tilde{w}+\tilde{w}^{-1})\right) \\
&= \frac{1}{\nu(2\nu+1)}\left[r^2 C_2^{(\nu)}\left(\frac{1}{2}(w+w^{-1})\right) - \tilde{r}^2 C_2^{(\nu)}\left(\frac{1}{2}(\tilde{w}+\tilde{w}^{-1})\right)\right] \\
&= r(w+w^{-1}) - 1 \ ,
\end{aligned}
$$

We can only consider the combination $a_{\Delta=2\Delta_\phi+2} := a_{1,0} + \nu(2\nu+1)\, a_{0,2}, \qquad \nu := \dfrac{d-2}{2}$

$$
\begin{aligned}
&r^{-2\Delta_\phi} - \tilde{r}^{-2\Delta_\phi} + \left(a_{1,0} + 3a_{0,2}\right)\left(r(w+w^{-1}) - 1\right) \\
&+ \sum_{J=0}^{J_*}\left[A_{\Delta_*(J),J;\boldsymbol{\theta}}(r)\, C_J^{(\nu)}\left(\frac{1}{2}(w+w^{-1}\right) - A_{\Delta_*(J),J;\boldsymbol{\theta}}(\tilde{r})\, C_J^{(\nu)}\left(\frac{1}{2}(\tilde{w}+\tilde{w}^{-1}\right)\right] \\
&+ \mathbb{T}^{(\text{approx})}_{\text{Disc}}[J_*; rw, rw^{-1}] - \mathbb{T}^{(\text{approx})}_{\text{Disc}}[J_*; \tilde{r}\tilde{w}, \tilde{r}\tilde{w}^{-1}] = 0
\end{aligned}
$$

$d = 4$

Inserting the exact analytic solution for the tail functions to the dot-loss we get

| $J_*$ | $\mathcal{L}_{\mathrm{dot}(1)}$ | $a_{1,0} + 3\,a_{2,0}$ from dot loss |
|---|---|---|
| 2 | $9.5 \times 10^{-6}$ | 15.16582 |
| 4 | $1.4 \times 10^{-6}$ | 15.07614 |
| 6 | $7.8 \times 10^{-8}$ | 15.06252 |
| 8 | $3.5 \times 10^{-9}$ | 15.06049 |
| | | Exact value: 15.06013 |

We obtain systematically better approximations of the full KMS condition with increasing $J_*$

## Test 2. Tail bootstrap without prior knowledge of the analytic GFF solution

- **No exposed operators**

- Use dot-loss $\mathscr{L}_{\text{dot}(0)} + \mathscr{L}_{\text{BC}:r=0.9999}$

$$\Delta T_{rel} := \frac{|T_{0+2;predicted}(r,w) - T_{0+2;analytic}(r,w)|}{|T_{0+2;predicted}(r,w)| + |T_{0+2;analytic}(r,w)|}$$

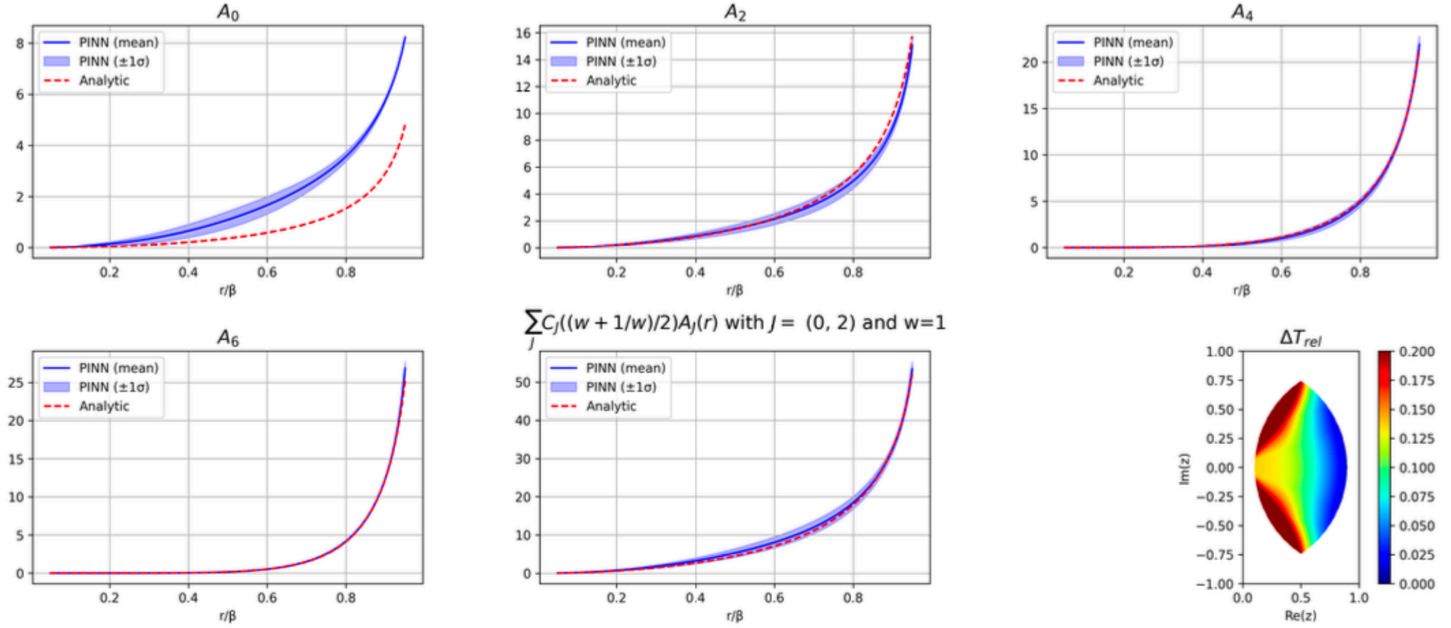



**Fig. 3:** Plots of the predicted tail functions obtained with $\mathcal{L}_{\text{dot}(0)} + \mathcal{L}_{\text{BC}}$ for $J_* = 6$ in the $d = 4$ GFF theory for $\Delta_\phi = 1.68$. The obtained results are based on the 10 lowest-loss configurations for 1K independent runs of 50K epochs. The mean loss, $3.54 \times 10^{-9} \pm 6.4 \times 10^{-10}$, should be compared to the loss of the analytic GFF solution $3.35 \times 10^{-9}$. The middle plot in the second line represents the combined contribution to the 2-point function of the $A_0$, $A_2$ tails at $w = 1$. The heatmap depicts the relative difference (5.20) inside the training region.

# Test 3. Tail bootstrap without any prior knowledge of the analytic GFF solution

- **1 exposed operator:** $a_{1,0} + 3a_{0,2}$

- Use dot-loss $\mathscr{L}_{\text{dot}(1)} + \mathscr{L}_{\text{BC}:r=0.9999}$

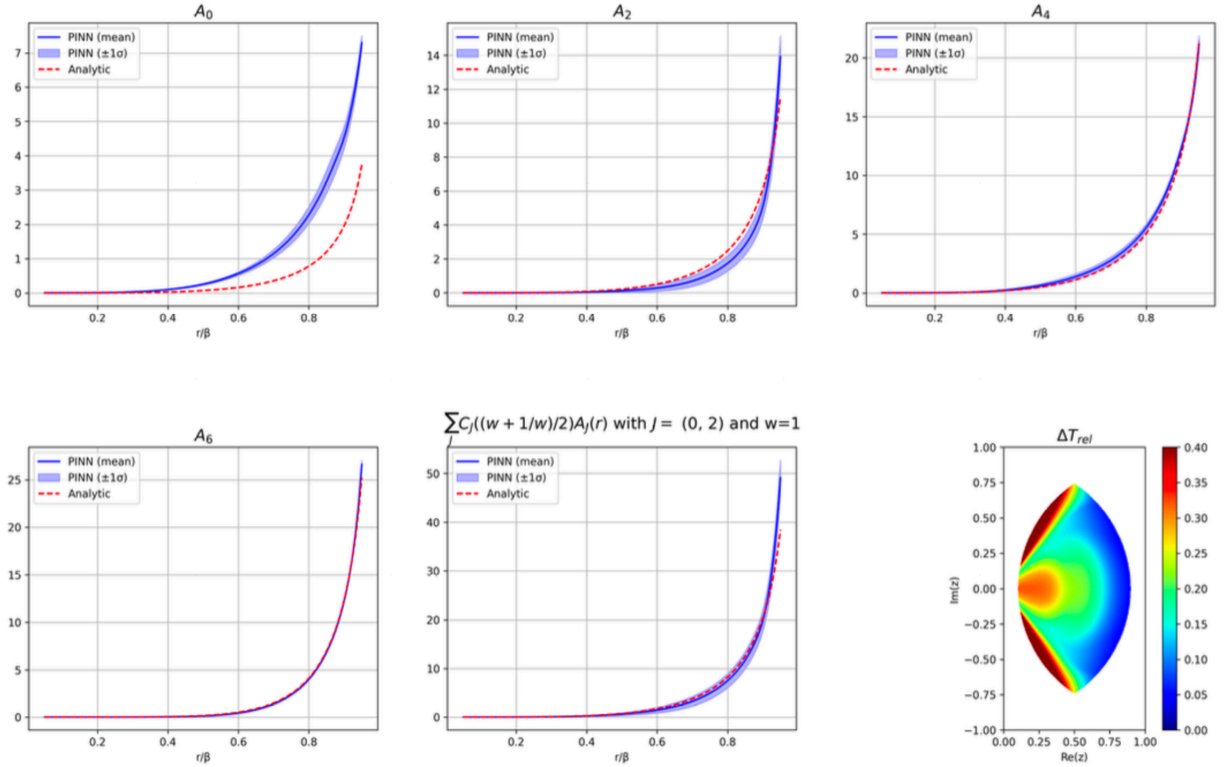



**Fig. 6:** Results obtained with the loss $\mathcal{L}_{\text{dot}(1)} + \mathcal{L}_{\text{BC}}$ and the CFT datum $a_{1,0} + 3a_{0,2}$ exposed. $\mathcal{L}_{\text{BC}}$ in Eq. (5.18) is set once again at $r = 0.9999$ using information about the asymptotic behavior of the tails in the vicinity of $r = 1$. The mean loss of the predicted functions is $6.12 \times 10^{-7} \pm 1.26 \times 10^{-7}$ with the loss of the analytic solution at $7.78 \times 10^{-8}$. The predicted coefficient is $a_{1,0} + 3a_{0,2} = 13.29 \pm 2.82$ with the analytic value at 15.06013. The first 4 plots (from top left to right) depict the predicted tail functions. The 5th plot and the heatmap provide information about the combined contribution of the tail functions $A_0, A_2$ to the conformal block expansion at $w = 1$ and the full training region respectively.

**Difficulties**

- There is an approximate reconstruction of the analytic correlator (which supports the claim that tails can be bootstrapped), but the result is not very accurate

- There is sensitivity on the $r \sim 1$ BC and without it the optimization is **not stable**.

  There are a lot of false minima!

  *The missing information from the region near the boundaries of the s/t-convergence region is important!*

  *Can we overcome these deficiencies with additional selection rules?*

**Observation 1.** $A_J(r_i) = A_J(r_i)\Big|_{\text{analytic}}$ **yields stability and accurate results**

**no exposed operators**

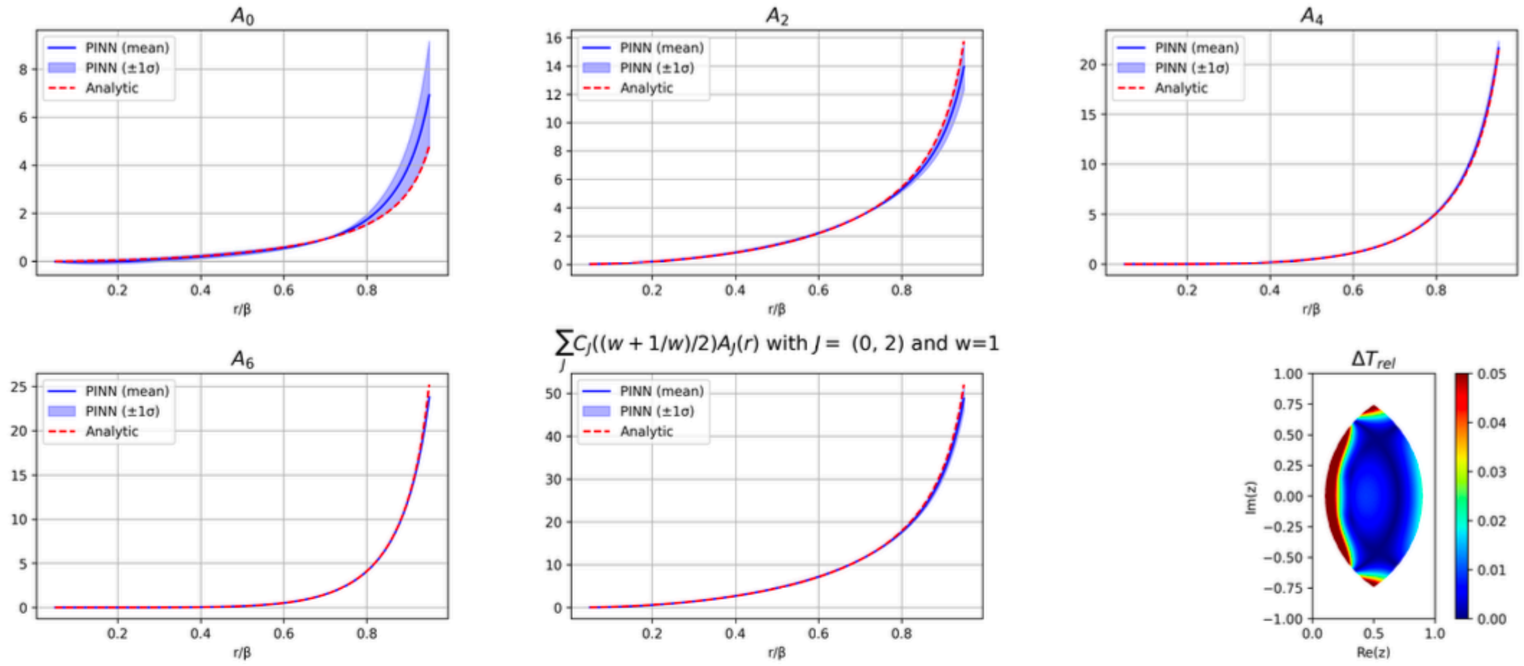



**Fig. 5:** Results obtained with the dot-loss $\mathcal{L}_{\text{dot}(0)}$, performing the same runs as in Fig. 3 after replacing the condition on the asymptotic values of the tail functions at $r = 0.9999$ with a condition that sets the values of the tail functions at $r_i = 0.7$ to the analytic values of the GFF solution. Notice that the maximum of the color bar scale in the heatmap is now set to 0.05, compared to the higher value of 0.20 in the previous Figs. 3, 4.

**Observation 2.** $A_J(r_i) = A_J(r_i)\Big|_{\text{analytic}}$ **yields stability and accurate results**

$a_{1,0} + 3a_{0,2}$ **exposed**

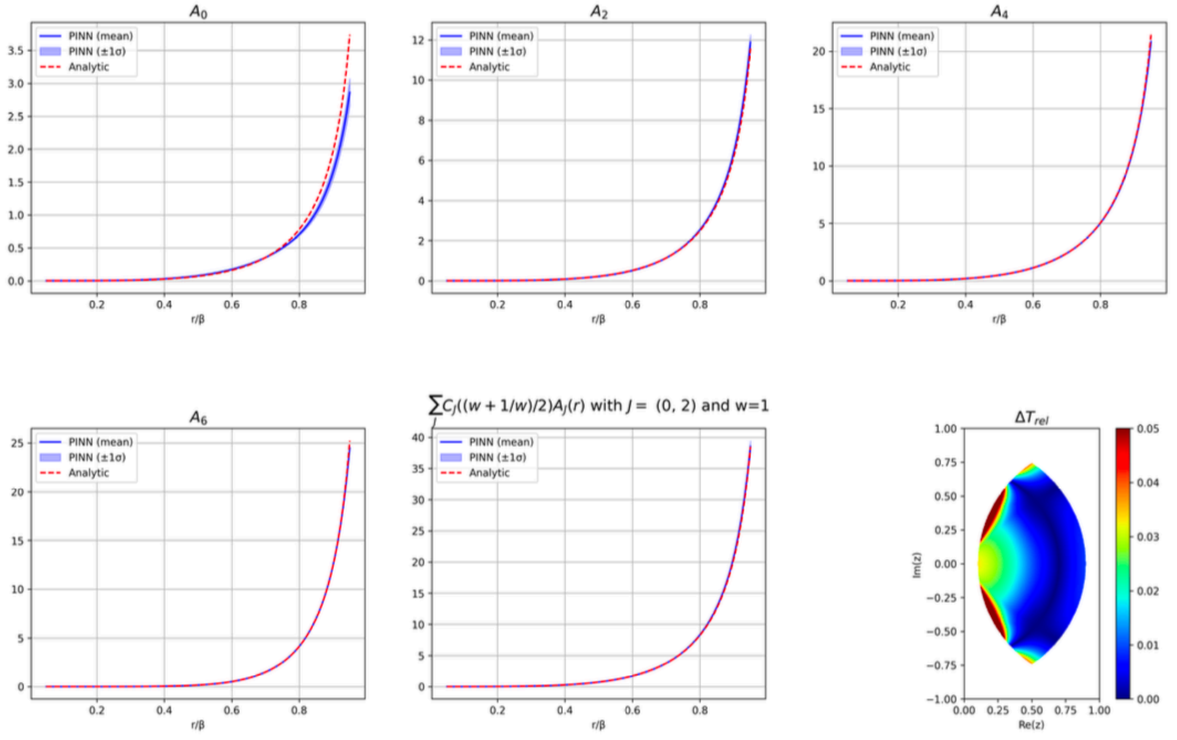



**Fig. 7:** Results obtained with the dot-loss $\mathcal{L}_{\text{dot}(1)}$ and the CFT datum $a_{1,0} + 3a_{0,2}$ exposed. In this case, the conditions on the asymptotic values of the tail functions at $r = 0.9999$ have been replaced by conditions that set the values of the tail functions at $r = 0.7$ to the analytic values of the GFF solution. The mean loss of the predicted functions is $9.49 \times 10^{-9} \pm 6.27 \times 10^{-10}$ with the loss of the analytic solution at $7.78 \times 10^{-8}$. The predicted coefficient is $a_{1,0} + 3a_{0,2} = 15.0647 \pm 0.0291$ with the analytic value at 15.06013. The first 4 plots (from top left to right) depict the predicted tail functions. The 5th plot and the heatmap provide information about the combined contribution of the tail functions $A_0, A_2$ to the conformal block expansion at $w = 1$ and the training region respectively.

## **Observation 3.** $A_J(r_i) =$ **(random non-GFF value) destroys stability**

This raises the exciting possibility that we can recover the exact solution by identifying an island of stability in the space of the fixed vector $A_J(r_i)$

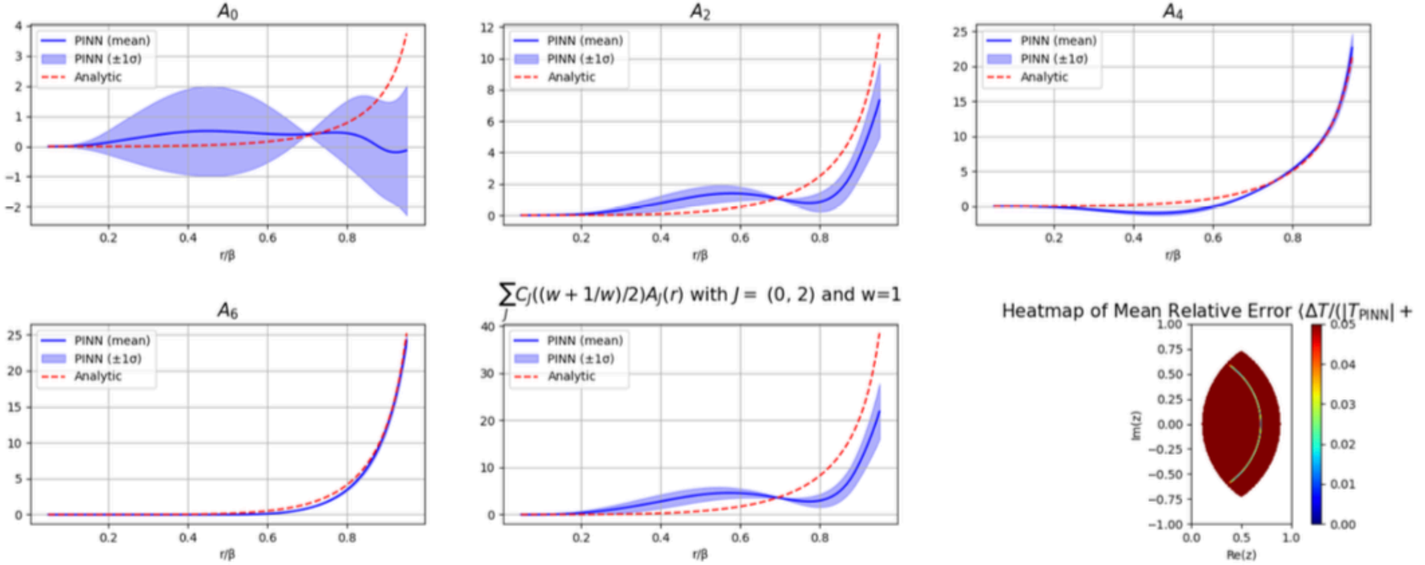



**Fig. 8:** Results obtained with the dot-loss $\mathcal{L}_{\text{dot}(1)}$ and the CFT datum $a_{1,0} + 3a_{0,2}$ exposed. In this case, we impose conditions on the tail functions at $r = 0.7$ (as in Figure 6), but to values that differ from the analytic values of the GFF solution. The mean loss of the predicted functions is $6.64 \times 10^{-9} \pm 8.67 \times 10^{-9}$ (to be compared against the loss of the analytic solution at $7.78 \times 10^{-8}$). The predicted coefficient, $a_{1,0} + 3a_{0,2} = 28.784 \pm 4.083$, is now significantly farther away from the analytic value at 15.06013. The first 4 plots (from top left to right) depict the predicted tail functions. The 5th plot and the heatmap provide information about the combined contribution of the tail functions $A_0, A_2$ to the conformal block expansion at $w = 1$ and the training region respectively. The thin yellowish line within the heatmap represents the $r = 0.7$ curve, where we impose the condition on $A_J(0.7)$. Since these values are relatively close to the GFF values, $\Delta T_{rel}$ is suppressed there.

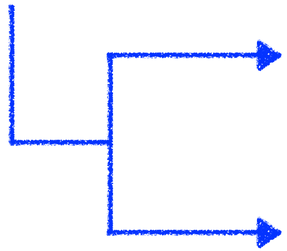
## 2. Holographic correlators

**Setup:** Thermal 2-point functions $\langle\phi\phi\rangle_\beta$ of scalar operators in a large-c holographic CFT with a (super)-gravity dual

**Gravity side:** The following spectrum contributes to the OPE

**(same spectrum, infinite KMS solutions provided by arbitrary higher-derivative gravities!)**

only energy-momentum sector contributes to discontinuity



| Operator | Scaling dimension $\Delta$ | Spin $J$ | $a_{\Delta,J}$ |
|---|---|---|---|
| Identity | $0$ | $0$ | $a_{\mathbf{1}}=1$ |
| $T_{\mu\nu}$ | $d$ | $2$ | $a_T$ |
| $[\phi\phi]_{n,J}$ | $2(\Delta_\phi+n+\ell)$ | $0\leq 2\ell$ | $a_{n,J}$ |
| $[T^k]_J$ | $d\,k$ | $0\leq 2\ell\leq 2k$ | $a_J^{(k)}$ |

In holography, the 2-point function $\langle\phi\phi\rangle_\beta$ on $S^1_\beta \times \mathbb{R}^{d-1}$ can be computed by solving the wave equation on the black brane background

- The OPE expansions give the unknown CFT data $a_T, a_{n,J}, a_J^{(k)}$

- Energy-momentum data easily determined from the asymptotics at infinity [Fitzpatrick-Huang '19]

- Double-twist data $a_{n,J}$ require the solution everywhere in the black brane geometry, and that has proven to be difficult

- **Universality relations**: [Fitzpatrick-Huang '19] observed that the lowest-twist energy-momentum multi-trace data $a_{2k}^{(k)}$ are all determined by $a_T$

**What can one ask from a bootstrap computation?**

1. To detect the multiple solutions

2. To detect the universality relations

3. To recover double-twist data from complete (or partial) energy-momentum data

   [see recent analysis of Burić-Gusev-Parnachev ’25

   using 0-spatial separation correlators & Borel-Padé approximations]

**Let’s explore point (3) with the deep finite-T bootstrap approach…**

**Preliminary results on double-twist thermal 1-point functions**

As an illustration (& to compare with previous literature) we fix $\Delta_\phi = 1.5$

- Set $J_* = 6$
- Approximate the discontinuity with only 7 data from Einstein gravity (up to twist 8)

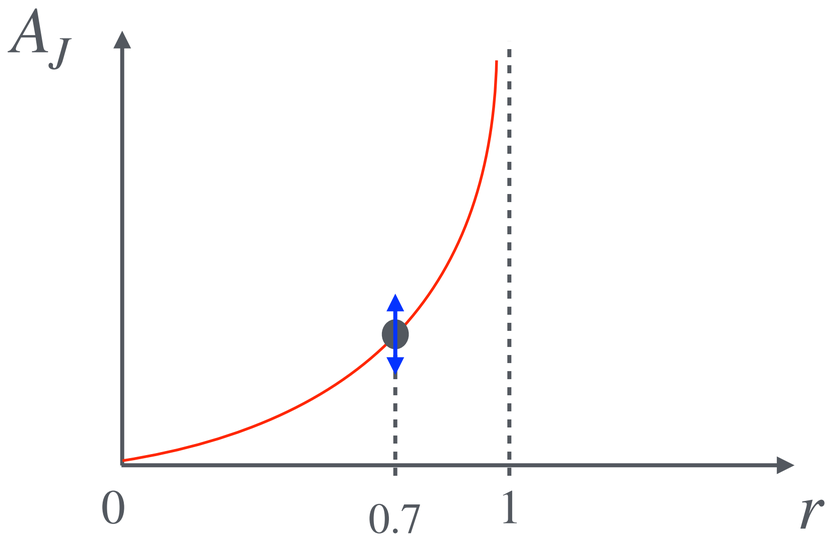
$$\text{identity} + T_{\mu\nu} \;+\; [T^2]_0 \;+\; [T^2]_2 \;+\; [T^2]_4 \;+\; [T^3]_4 \;+\; [T^3]_6 \;+\; [T^4]_8$$

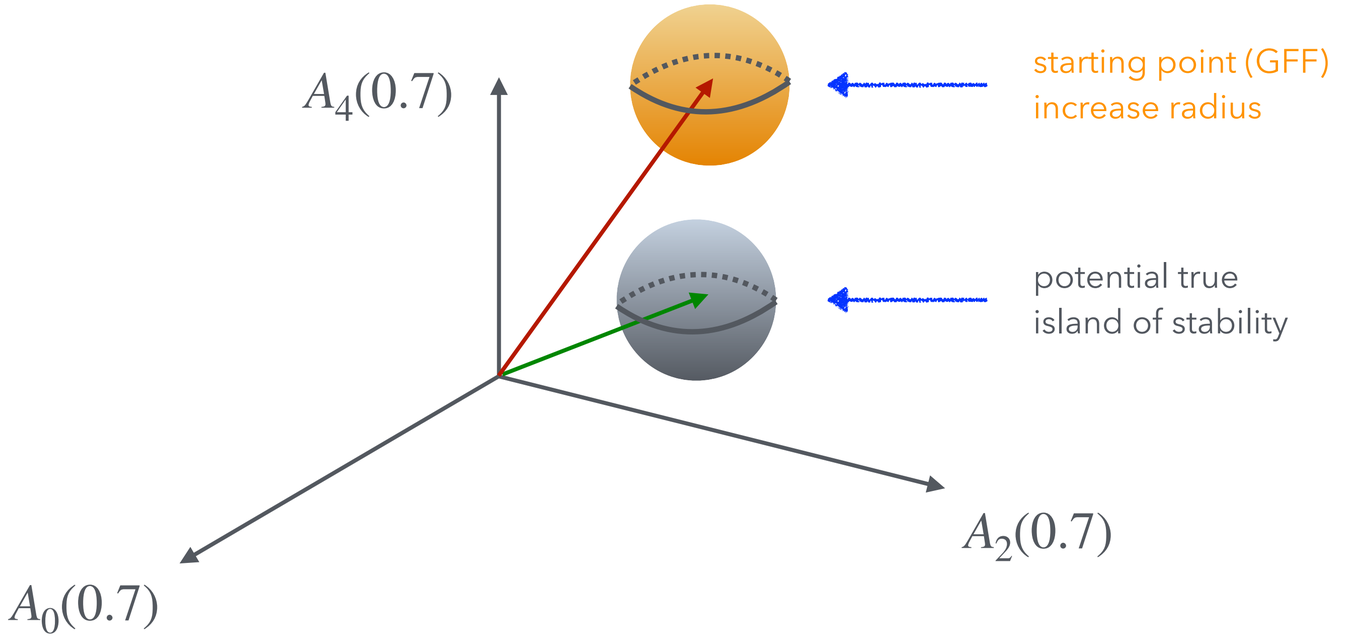
$$a_{T,\mathrm{GR}} = 1.21761,\;\; a^{(2)}_{0,\mathrm{GR}} = -1.37668,\;\; a^{(2)}_{2,\mathrm{GR}} = 1.58848,\;\; a^{(2)}_{4,\mathrm{GR}} = -4.05945,$$
$$a^{(3)}_{4,\mathrm{GR}} = 1.77035,\;\; a^{(3)}_{6,\mathrm{GR}} = 8.52362,\;\; a^{(4)}_{8,\mathrm{GR}} = -15.9641.$$

We expose the operators $[\phi\phi]_{1,0}$, $[\phi\phi]_{0,2}$ with combined thermal OPE coefficient in KMS

$$a_{1,0} + 3a_{0,2}$$

To mitigate the difficulties of the missing information from the boundaries of the s/t-convergence region we search for an **island of stability** in the space of fixed vectors $A_J(r_i)$ (we set $r_i = 0.7$)



We search for a potential island of stability on $A_J(0.7)$ by adding to the loss function an extra term that allows $A_J(0.7)$ to vary in selected region

We gradually increase that region until we hit the point of stability

$$\mathcal{L}_{\text{ReLU}} = \frac{1}{\frac{J_*}{2}+1} \sum_J \text{ReLU}\left( \left| \mathcal{A}_J(r_i) - \mathcal{A}_J(r_i)|_{\text{GFF}} \right| - \mathtt{p} \times \left| \mathcal{A}_J(r_i)|_{\text{GFF}} \right| \right)$$

$$\text{ReLU}(x) = \max(0, x)$$

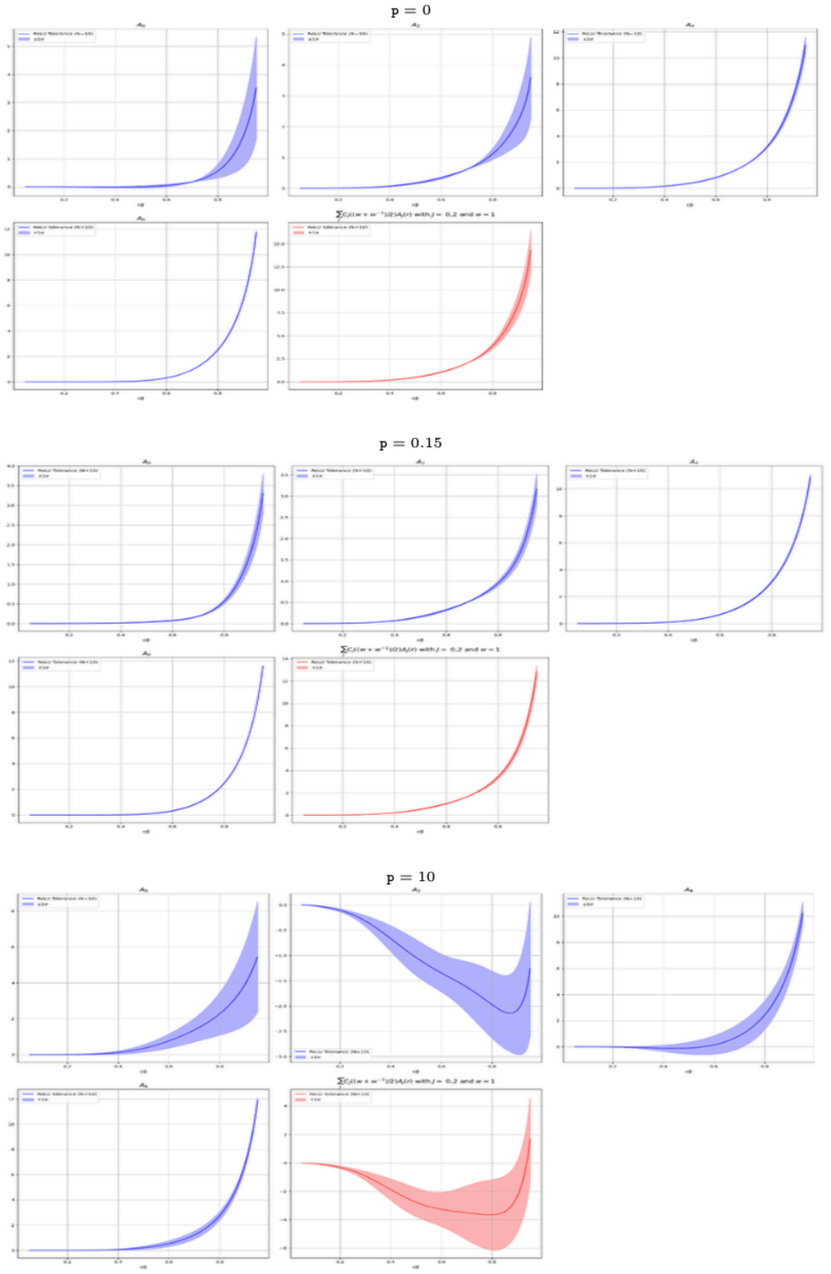
Total loss function

$$\mathcal{L} = \mathcal{L}_{\text{KMS}} + \mathcal{L}_{\text{ReLU}}$$

$$\mathrm{p} = 0.20 \ : \ a_{10} + 3a_{02} = 9.37 \pm 0.44$$

The more robust approach of
Burić-Gusev-Parnachev 2508.08373
(based on a 0-spatial-separation analysis
and around 40 energy-momentum data)
yields $a_{10} + 3a_{02} = 7.686$

The independent numerical sol of the PDE
gives $a_{10} + 3a_{02} \in [8.84, 10.97]$
Burić-Gusev-Parnachev 2505.10277



# Outlook

Take home message: evidence that *high-energy OPE tails can be bootstrapped*

Various elements need to be combined in this type of bootstrap

The approach is generic and opens up the road to many new explorations

I focused on the thermal bootstrap, but similar things can be done in various other contexts, where positivity constraints may or may not be available

**In the context of the thermal bootstrap…**

- Our approach is non-convex and allows a more complete exploration of the KMS conditions
- It has the analytic structure built in (it incorporates the input of dispersion relations)
- It is the **only** approach currently that can study thermal 2-point functions at **non-zero spatial separation**. It can produce predictions for spin-dependent thermal coefficients that no other method in the current literature can
- It is more systematic, flexible and economical compared to previous truncation schemes. The spectrum can also be varied and KMS can be combined with other sum rules

The practical implementation can be improved and there are aspects that require better understanding, e.g.

- The difficulties associated with the boundaries of the s/t-convergence region and the role of the stability islands
- The discontinuity approximations and the overall convergence of the scheme (e.g. further improve the double-twist predictions in holographic CFTs)
- More efficient optimization with dynamical discontinuities
- More applications (at zero-T, finite-T, holographic, non-holographic, with/without defects…)

**Supplementary slides**

In the above results we used

$$\mathcal{L} = \mathcal{L}_{\mathrm{KMS}} + \mathcal{L}_{\mathrm{ReLU}}$$

Using an additional condition on the asymptotic value of the tail functions near r=1 we would have

$$\mathcal{L} = \mathcal{L}_{\mathrm{KMS}} + \mathcal{L}_{\mathrm{ReLU}} + \mathcal{L}_{\mathrm{BC}}$$

BC near r~1

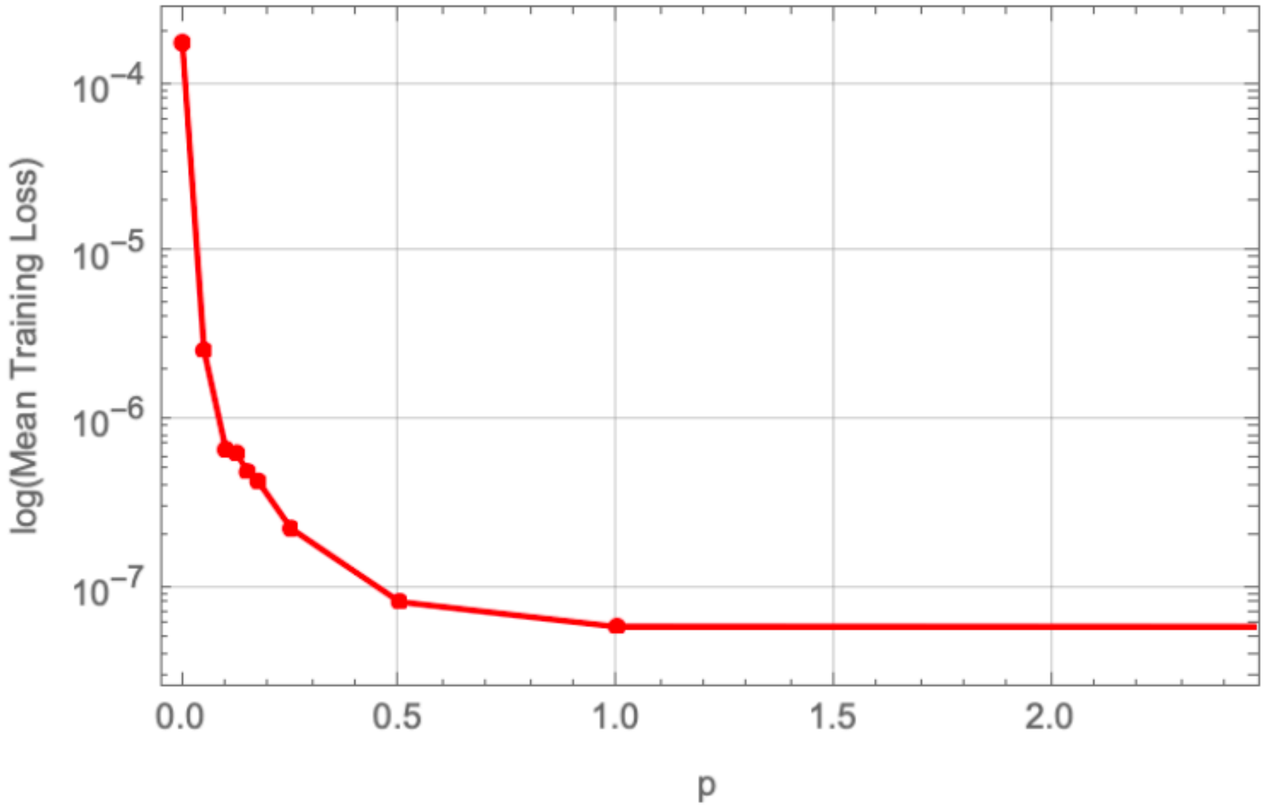



**Fig. 9:** Mean training loss for independent ReLU runs at different values of the factor $\mathbf{p}$. The mean is computed for the 10 lowest-loss configurations of each cluster run. The $y$-axis scale is logarithmic. There is a concentration of data points near the value $\mathbf{p} = 0.15$, where we observed outcomes with the smallest statistical variation.