Notes on conformal manifolds

November 19, 2024

Workflow as of November 19, 2024:

1. Generate larger clouds of points for the 7- and 13-param potentials (Gabriel).

2. For these two models, get statistics of autoencoders for several attempts per dimension of latent layer (Bastien).

3. Go up to full 32-param SO(8,4) potential.

Contents

1 Comments on Flat Directions and n-Hessians

Given a real function $V : M \to \mathbb{R}$, its critical points are defined as the zeroes of the gradient

$$x^* : H_m(x^*) \equiv \partial_m V\big|_{x^*} = 0 \,, \tag{1.1}$$

with the n-Hessian of $V(x)$ defined as

$$H_{m_1 \ldots m_n} \equiv \partial_{m_1} \ldots \partial_{m_n} V \,. \tag{1.2}$$

Flat directions around the critical point x^* are curves $\phi : I \to M$ such that

$$H_m(\phi(t)) = 0 \,, \qquad \text{with} \quad \phi(0) = x^* \,. \tag{1.3}$$

Close to x^*, the curve can be expanded as

$$\phi(t) = x^* + \sum_{k=1}^{\infty} \tfrac{1}{k!} \xi_k\, t^k\,, \qquad \text{with} \quad \xi_k = \frac{d^k \phi}{dt^k}\bigg|_{x^*}\,, \tag{1.4}$$

and, accordingly,

$$H_m(\phi(t)) = \sum_{k=1}^{\infty} H_{mn_1...n_k} \sum_{a \in \lambda_k} t^{\sum_i^{|a|} i a_i} \left[\prod_{i=1}^{|a|} \frac{1}{a_i!} \frac{1}{(i!)^{a_i}} \xi_i^{\otimes a_i} \right]^{n_1...n_k}\,, \tag{1.5}$$

with all the n-Hessians on the right-hand side evaluated at x^*. Here, λ_k denotes the integer partitions of k (e.g. $\lambda_3 = \{(3),(2,1),(1,1,1)\}$), $|a|$ is the number of elements of a specific partition $a \in \lambda_k$ (e.g. $|(2,1)| = 2$ and $|(1,1,1)| = 3$), and a_i each of its elements. Equation (1.3) needs to be satisfied at every order in t, and therefore the existence of flat directions amounts to the existence solutions of the system

$$O(t^p): \quad \sum_{\{a_i\}} H_{mn_1...n_k} \prod_i^p \left[\frac{1}{a_i!} \frac{1}{(i!)^{a_i}} \xi_i^{\otimes a_i} \right]^{n_1...n_k} = 0\,, \tag{1.6}$$

with the set of coefficients $\{a_i\}$ determined as

$$\sum_i^p i a_i = p\,, \qquad a_i \in \mathbb{N}\,, \tag{1.7}$$

and k in (1.6) a shorthand for $k = \sum_i^p a_i$ for the solutions of (1.7). For lowest p, this is **Note: Can** $\Longleftarrow$ **map to CFT n-point functions à la Bertolini?**

$$0 = H_{mn}\xi_1^n\,,$$

$$0 = H_{mn}\xi_2^n + H_{mnp}\xi_1^n\xi_1^p\,,$$

$$0 = H_{mn}\xi_3^n + 3H_{mnp}\xi_2^n\xi_1^p + H_{mnpq}\xi_1^n\xi_1^p\xi_1^q\,, \tag{1.8}$$

$$0 = H_{mn}\xi_4^n + 3H_{mnp}\xi_2^n\xi_2^p + 4H_{mnp}\xi_3^n\xi_1^p + 6H_{mnpq}\xi_2^n\xi_1^p\xi_1^q + H_{mnpqr}\xi_1^n\xi_1^p\xi_1^q\xi_1^r\,,$$

$$\dots$$

Note that the non-linearity of (1.8) in the expansion coefficients of $\phi(t)$ implies that the number of flat directions is not easily related to the dimension of the null space of the 2-Hessian at x^* (i.e. the number of massless scalars if V is the scalar potential). A necessary condition for (1.6) to have a solution is

$$\det H_{mn} = 0\,, \tag{1.9}$$

but there can be more flat directions than zero eigenvalues and vice versa. These features will be studied in the following.
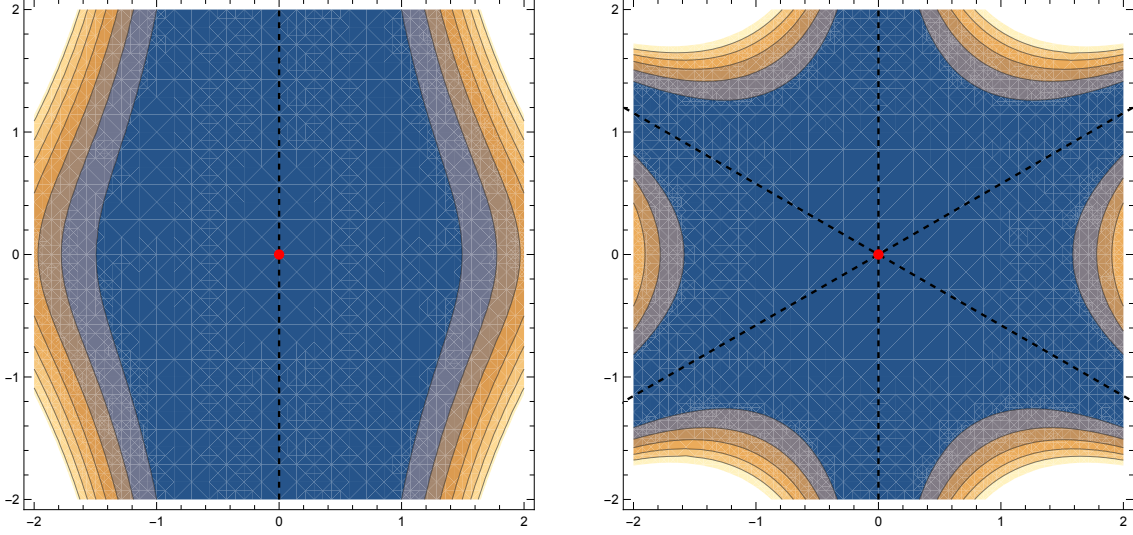
Fig. 1 Contour plots for the potentials V_1 (left) and V_3 (right) in (1.12) with flat directions and $x^* = 0$ locus superimposed.

1.1 Flat flat directions

When flat directions can be given as straight lines $\phi(t) = x^* + \xi t$ for constant ξ, the set of equations (1.8) simplifies to the condition $\xi \in K_k$ for all $n \geq 1$, with the kernels defined as

$$K_k = \{\xi : \xi^{n_1} \dots \xi^{n_k} H_{mn_1 \dots n_k}(x^*) = 0\}, \qquad \forall n \geq 1. \tag{1.10}$$

This condition is required for $\partial_m V$ to be constant along the integral curve of ξ^m, since

$$\xi^{n_k} \partial_{n_k} \left(\xi^{n_1} \dots \xi^{n_{k-1}} H_{mn_1 \dots n_{k-1}} \right) \big|_{x^*} = \xi^{n_1} \dots \xi^{n_k} H_{mn_1 \dots n_k}(x^*) \tag{1.11}$$

must hold for all $k \geq 1$.

Two functions that have flat flat directions and exhibit the subtleties mentioned before are given by

$$V_1(x,y) = x^4(1 + y^2), \qquad \text{and} \qquad V_3(x,y) = y^2(x^2 - 3y^2)^2, \tag{1.12}$$

which are plotted in Fig. 1. In both cases, the 2-Hessian vanishes identically at $x^* = 0$, but V_1 only has one flat direction given by $(x,y) = (0, \zeta)$, whereas V_3 has three independent flat curves:

$$(x,y) = (0, \zeta_1), \qquad (x,y) = (\sqrt{3}\,\zeta_2, \zeta_2), \qquad (x,y) = (-\sqrt{3}\,\zeta_3, \zeta_3). \tag{1.13}$$

⇐

Note: How is the second compatible with the standard holography lore? Can this type of potentials arise in consistent truncations? ⇐

Note: We should remark that conformal manifolds are not smooth manifolds in general, but algebraic varieties. Apparently, people doing quivers know this [look at papers by Hanany]

1.2 Curved flat directions
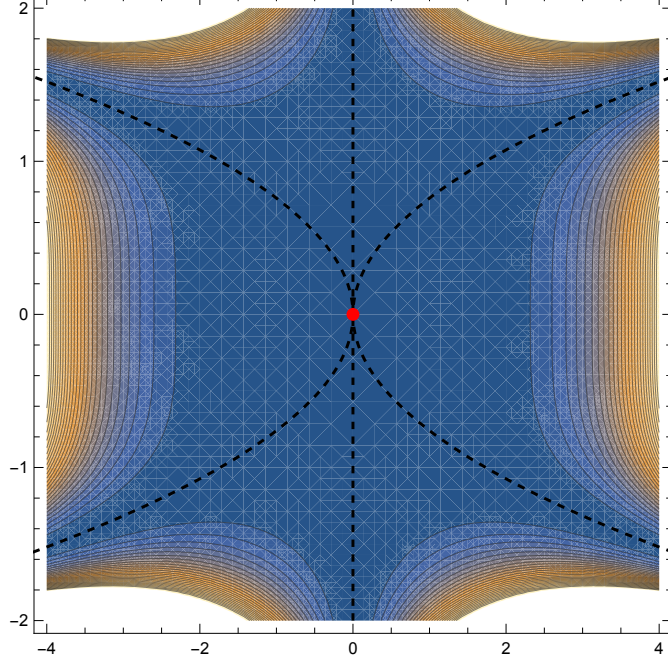
(alternative name: curved flat curves)

Fig. 2 Contour plot for the potential in (1.14) with flat directions in (1.15) and $x^* = 0$ locus superimposed.

The condition (1.10) does not capture curves with a non-linear dependence on t. For instance,

$$V(x, y) = x^2(x^2 - 3y^4)^2 \tag{1.14}$$

has flat directions

$$(x, y) = (0, \zeta_1), \qquad (x, y) = (\sqrt{3}\,\zeta_2^2,\, \zeta_2), \qquad (x, y) = (-\sqrt{3}\,\zeta_3^2,\, \zeta_3), \tag{1.15}$$

of which only the first is captured by (1.10). For the other two, the higher order terms in (1.6) also contribute. To gauge-fix the reparameterisation invariance of (1.3), we demand without loss of generality that the lowest non-vanishing term in (1.4) is linear in t. Doing so, (1.6) yields (1.15) unambiguously. We note that for (1.14) the first non-trivial equation arises at order $O(t^5)$, the following at $O(t^{10})$, and from then on we get a condition on every level.

This method has been checked with three other functions with flat directions:

$$\begin{aligned}
V(x, y) &= (y - \cos x)^2, \\
V(x, y) &= (x^2 + y^2 - 1)^2, \\
V(x, y) &= x^2(x^2 + y^2 - 1)^2,
\end{aligned} \tag{1.16}$$

recovering the analytic result in all instances and converging more quickly than for (1.14) (as their series have non-zero contributions at lower orders). ⟸

Note: The distinction between the different curves in (1.15) disappears at linear level, and this is very reminiscent of the TsT vs Wilson loop families in [1].
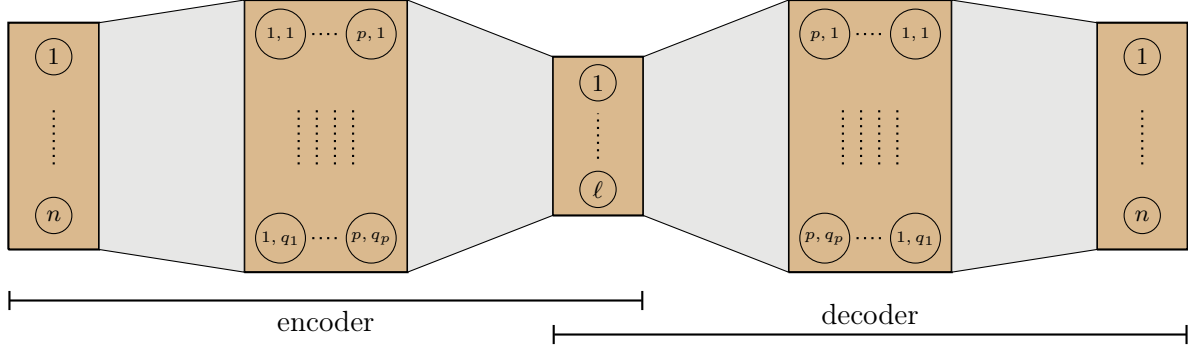
Fig. 3 Architecture of a fully-connected symmetric autoencoder with input/output of dimension n, a latent layer of dimension ℓ. The hidden encoder and decoder subnetworks are arranged specularly, and have p layers of width q_p each. This information will be denoted by a list $[n, q_1, \ldots, q_p, \ell]$.

2 Machine Learning Approach

In the following, we obtain and analyse numerically the critical points of different functions $V : \mathbb{R}^n \to \mathbb{R}$. In this approach, a cloud of points x_i^m, $m \in [\![1, n]\!]$ and $i \in [\![1, N]\!]$, is generated so that the loss function

$$L_V = \sum_i^N |\nabla_m V(x_i)|^2 \tag{2.1}$$

is minimised. For this minimisation, we have employed an protocol based on the `TensorFlow` Adam optimiser which restarts the optimiser with the initial learning rate after a specified number of epochs until a specified loss value is attained. Afterwards and until the pre-determined number of epochs is reached, the learning rate decreases by a factor 10 every few epochs (in the `grad_descent_module.py` package, the original learning rate us used by default for the first 2500 epochs with restarts every 200 steps, and then updated every 2500 steps. These parameters need to be adapted to the potential under consideration).

The structure of the resulting cloud of points minimising (2.1) is subseqently learnt using autoencoders (AEs), which are non-linear endomorphisms constructed via a neural network with an intermediate layer of dimension $\ell < n$, as schematically depicted in figure 3. The loss function for the autoencoder is chosen as

$$L_{\mathrm{AE}}(\ell) = \sum_i^N \left| x_i^{(\mathrm{in})} - x_i^{(\mathrm{out})} \right|^2 \tag{2.2}$$

with $x_i^{(\mathrm{out})}$ being the output of the decoder for a given input $x_i^{(\mathrm{in})}$. These outputs depend on the value of ℓ, and the loss function (2.2) displays a sharp transition from a steep decline into a plateau when ℓ reaches the dimension of the conformal manifold, if the latter does not have separated components. In the latter scenario, a less pronounced transition occurs, and inspection of the values at the latent layer reveals clusters of data associated to the different components.

In the following, we introduce these techniques on a set of toy models of low dimensionality, and then apply these tools to the scalar potential (2.26) in [2].

2.1 Toy examples

Higgs potential   The potential

$$V(x^1, x^2, x^3, x^4) = \left[(x^1)^2 + (x^2)^2 - 1\right]^2 + (x^3)^2 + (x^4)^2, \tag{2.3}$$

is a mexican-hat potential for the variables x^1 and x^2, and a harmonic potential for the variables x^3 and x^4. Its vacua are given by two different families of solutions,

$$\begin{aligned}
&\text{family 1} \quad : \quad x^1 = 0, \qquad x^2 = 0, \qquad x^3 = 0, \quad x^4 = 0, \\
&\text{family 2} \quad : \quad x^1 = \cos t, \quad x^2 = \sin t, \quad x^3 = 0, \quad x^4 = 0, \qquad t \in [0, 2\pi).
\end{aligned} \tag{2.4}$$

We can capture this structure by performing a gradient descent with loss function (2.1) for a cloud of 10^5 randomly generated points in the region $[-0.75, 0.75]^4$. See figure 4 for the attained cloud of points and the corresponding loss function during the gradient descent.

To ascertain the dimension of the geometric structures behind the shapes in figure 4, we have employed an AE with the following architectures

$$[n, q_1, q_2, q_3, \ell] = [4, 32, 16, 8, \ell], \qquad \ell = 1, 2, 3, \tag{2.5}$$

following the notation of figure 3. For this example, $L_{\text{AE}}(\ell)$ in figure 5(a) would naively suggest that the underpinning structures are two-dimensional, instead of the zero- and one-dimensional loci in (2.4). However, close inspection of the output of the encoder half of the autoencoder (2.5) for $\ell = 2$ reveals that the data is cleanly separated into two distinct structures: one point and one (broken) line, as can be observed in figure 5(b). In fact, filtering using the decoder to label the data corresponding to the isolated point in figure 5(b), it can be verified that they correspond to solutions in family 1. Filtering out these points, the analytical structure of the complementary set can in fact be found to by using the $\ell = 1$ decoder as a non-linear function ...   **Note: how were the details of this procedure?** $\Longleftarrow$

Scrambled Higgs potential   The example above was somewhat trivial given that one could immediately recognise the relevant structure by glancing at figure 4, as the variables in which the solutions in (2.4) are most naturally expressed coincide with the variables that appear in the potential. In supergravity this will no longer be the case, and to exemplify how the autoencoder is useful in this context, let us consider the same potential as above,

$$V(\tilde{x}^1, \tilde{x}^2, \tilde{x}^3, \tilde{x}^4) = \left[(\tilde{x}^1)^2 + (\tilde{x}^2)^2 - 1\right]^2 + (\tilde{x}^3)^2 + (\tilde{x}^4)^2, \tag{2.6}$$

with the variables $\tilde{x}^i$ given by   **Note: need to decide a non-singular and bijective change of** $\Longleftarrow$
coordinates.

$$\tilde{x}^1 = \dots . \tag{2.7}$$

Constructing the conformal manifold in terms of x^i results in a could for which the structure of the solutions is much less clear, as can be appreciated in figure **??**. However, the information in an autoencoder given again by (2.5) results in loss functions and encoded data in precise correspondence with figure 5.
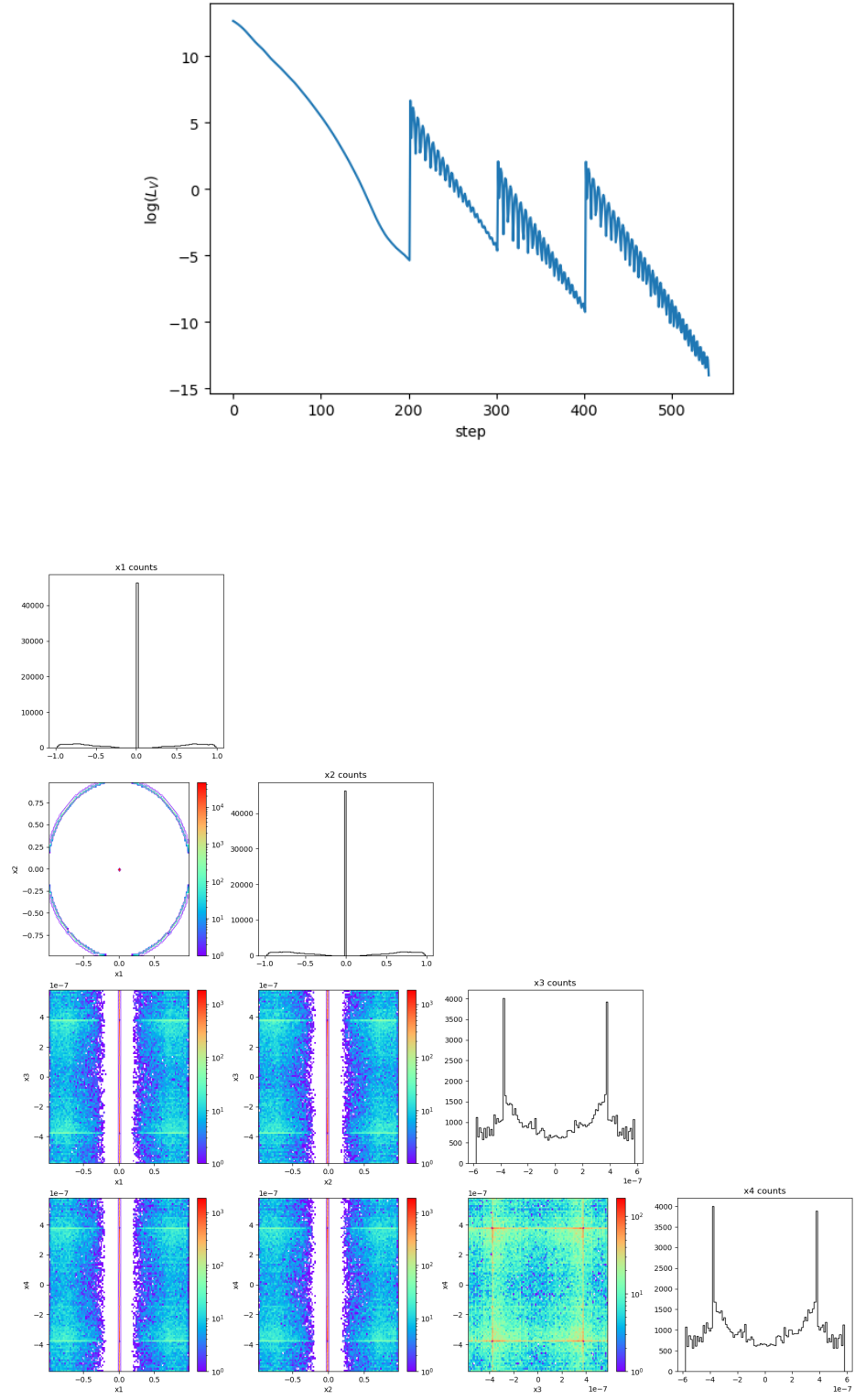
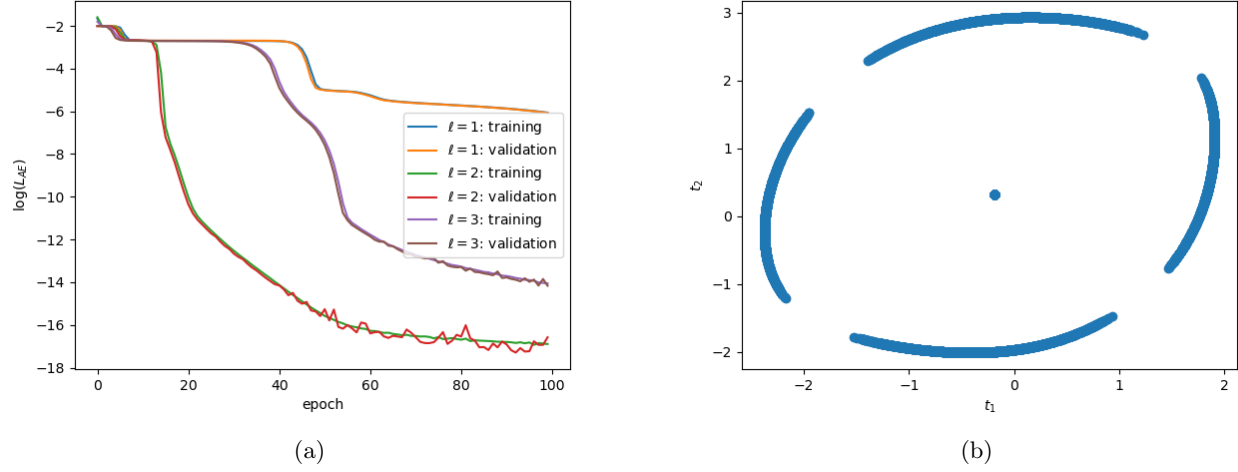**Fig. 4** Loss function and triangular plot of the numerical vacua of (2.3).

(a)

(b)

Fig. 5 Loss function encoder output for the autoencoder (2.5) and the entire set of data in figure 4.

2.2 SO(8,4) supergravity

We have

1. Looked at 3-, 5-, 7- and 13-parameter truncations.

2. There was qualitatively different behaviour when using logs vs exps.

3. There are spurious structures (forks) that the AE erases (why does it?).

4. The spurious structures can be also be broadened using random walks.

References

[1] C. Eloy and G. Larios, *Charting the Conformal Manifold of Holographic CFT_2's*, `arXiv:2405.17542`.

[2] C. Eloy, G. Larios, and H. Samtleben, *Triality and the consistent reductions on $AdS_3 \times S^3$*, *JHEP* **01** (2022) 055, [`arXiv:2111.01167`].