New methods for old problems: vacua of maximal $D = 7$ supergravities

Dario Partipilo

Dipartimento di Fisica e Astronomia "Galileo Galilei", Università di Padova,
Via Marzolo 8, 35131 Padova, Italy
INFN, Sezione di Padova,
Via Marzolo 8, 35131 Padova, Italy

E-mail: dario.partipilo@icloud.com

Abstract: Finding vacua of supergravity theories is an outstanding problem which has been tackled in several ways, and with this work we add a new method to the puzzle. We analyse the scalar sector of maximal gauged supergravity theories in seven space-time dimensions. We look for vacua of the theory by varying the embedding tensor, instead of directly minimising the scalar potential. The set of quadratic constraints arising from this procedure has been solved by means of Evolution Strategies optimisation techniques, also adopted in Artificial Intelligence studies. We develop some methods to reconstruct and obtain analytical results starting from numerical outcomes, thus obtaining the complete mass spectra. In addition to some of the known vacua, we also obtain two new Minkowski vacua.

Contents

1 Introduction

Minima of potential functions have always been of uttermost importance in physics: from the search for stable trajectories in the gravitational field of a star to the spontaneous symmetry-breaking mechanism, potentials have been shown to be very useful in understanding the physical problem at hand. With the advent of String Theory, finding minima of scalar potentials was fundamental to study the String Landscape [1], in this framework Supergravity theories proved to be a conceptual testbed for this analysis. In fact, examining the structure of the vacua for supergravities is a crucial task in order to study the patterns of supersymmetry breaking, the processes behind the generation of critical points with non-vanishing cosmological constant, and in case the latter takes negative values the possible holographic dual theory. More specifically, the supergravity regime of the bulk theory in the AdS/CFT conjecture is described by gauged supergravities. Moreover, vacua with positive values of the cosmological constant have also recently received particular attention from the string theory community, due to the restrictions imposed upon them by the de Sitter conjecture [2] and other conjectures formulated in the framework of the Swampland programme. Vacua of gauged supergravities also have the desirable feature of giving masses to scalars, thus providing a solution to the moduli stabilisation problem.

In this work, we will focus on maximal ($\mathcal{N} = 4$) supersymmetric theories in $D = 7$ space-time dimensions. These theories are of particular interest because of their large amount of supersymmetry that completely fixes the field content of the theory and the Lagrangian, and in light of the limited number of possible deformations to which they can be subjected. For instance, dimensional reductions of string theories or M-theory on spheres produces theories whose vacua have maximally supersymmetric holographic duals.

Moreover, our interest in seven dimensions arises since it has been shown that CFTs exist up to six space-time dimensions [3], and it is possible to chart most of them once we know all the vacua of Supergravity in one more dimension.

Searching for minima of the scalar potentials in supergravity theories is a challenging task, and many techniques have been developed so far to tackle the problem. Progress has been achieved by pursuing three main different paths. The first consists in picking a particular theory (gaugings) and by using symmetries to truncate the field content in order to arrive at a model with a restricted number of scalars, thus making the minimisation procedure feasible. This method, introduced by N.P. Warner in [4, 5] made it possible to analytically solve different minimisation conditions, thus discovering many new vacua in 4 and 5 space-time dimensions. Another proposal is to adopt a numerical approach based on machine learning libraries, and lead to the discovery of a large amount of vacua, which also allowed the analysis of mass spectra, residual supersymmetries and residual gauge groups [6–13]. Both these techniques are based on the choice of a gauging, consequently, one has to scan one theory at a time, making them inconvenient for a complete cataloguing. Another approach has been formulated, based on the embedding tensor formalism. First discovered in the contest of the super-Higgs mechanism [14] it has been then introduced in the framework of the analysis of supergravity vacua [15], leading to a wealth of new results [16–23] as well as providing the first example of a family of gauged theories dependent on a continuous parameter [24]. This method does not pick a specific gauging choice, thus allowing for a simultaneous scan of different gauge theories. The last technique will produce a set of quadratic equations in n variables, with n the number of parameters in the embedding tensor, this problem is NP-hard, thus analytic solutions are expected to be found in very special cases, usually by restricting the number of variables, using some residual symmetry. In this work, we will build upon the last method and implement the scan using an evolution strategy algorithm, Covariance Matrix Adaptation (CMA-ES) [25], together with a Genetic Algorithm (GA). From the solutions we obtain, we can recover the analytic form of the vacua and their spectra.

2 Maximal supergravity in 7 dimensions

We recall, in this paragraph, all the notations and properties of the tensors needed in the rest of the paper. It is well known, since the seminal paper by de Wit, Samtleben and Triggiante [26], that any possible gauging of a supergravity theory is encoded in the embedding tensor $\Theta_M{}^\alpha$, where usually the indices M, N, P, transform in the vector representation and $\alpha, \beta, etc.$ in the adjoint of the duality group (for the case of 7 dimensions $M, N, etc.$ will live in the fundamental representation **5** of SL(5), so the vector representation **10** will be described by a couple of antisymmetric indices $MN = [MN]$). The job of the embedding tensor is to select the generators of the duality group, which in the case of 7 dimensions is SL(5), which will form the gauge group of the specified theory. Once this tensor is fixed, the Lagrangian and supersymmetry transformations follow, according to the analysis in [26]. We can write the new generators of the gauge group as

$$X_{MN} = \Theta_{MN,P}{}^Q t^P{}_Q, \tag{2.1}$$

where $t^P{}_Q$ are the SL(5) generators, satisfying $t^M{}_M = 0$ and the algebra

$$\left[t^M{}_N, t^P{}_Q\right] = \delta^P_N t^M{}_Q - \delta^M_Q t^P{}_N. \tag{2.2}$$

In seven space-time dimensions, the embedding tensor lives in the product of representations $\mathbf{10} \otimes \mathbf{24} = \mathbf{10} + \mathbf{15} + \overline{\mathbf{40}} + \mathbf{175}$. However, supersymmetry constrains it to lie in the $\mathbf{15} + \overline{\mathbf{40}}$, therefore it is possible to parameterize it as:

$$\Theta_{MN,P}{}^Q = \delta^Q_{[M} Y_{N]P} - 2\epsilon_{MNPRS} Z^{RS,Q} \quad . \tag{2.3}$$

Where $Y_{MN} = Y_{(MN)}$ and $Z^{MN,P} = Z^{[MN],P}$ with $Z^{[MN,P]} = 0$, that live respectiveley in the $\mathbf{15}$ and $\overline{\mathbf{40}}$ representations of SL(5). Gauge symmetry imposes further constraints, these are in particular quadratic constraints that ensure the closure of the gauge algebra. They can be expressed in terms of the Y and Z tensors as

$$Z^{MN,P} Y_{PQ} = 0 \quad , \qquad\qquad Z^{MN,P} X_{MN} = 0 \quad . \tag{2.4}$$

These $\mathbf{15} + \overline{\mathbf{40}}$ representations can also be decomposed under the maximal compact subgroup of SL(5), that is, USp(4) $\simeq$ SO(5). Under USp(4) they decompose as

$$\mathbf{15} + \overline{\mathbf{40}} \longrightarrow (\mathbf{1} + \mathbf{14}) + (\mathbf{5} + \mathbf{35}) \quad . \tag{2.5}$$

Defining the scalar matrix $\mathcal{V}_M{}^{ab}$, anti-symmetric in a and b (a,b=1,…,4), which mediates from SL(5) to USp(4), normalised by

$$\mathcal{V}_P{}^{ab} \mathcal{V}_{ab}{}^Q = \delta^Q_P, \qquad\qquad \mathcal{V}_{ab}{}^P \mathcal{V}_P{}^{cd} = \delta^{cd}_{ab} - \frac{1}{4}\Omega_{ab}\Omega^{cd}, \tag{2.6}$$

we have:

$$Y_{MN} = \mathcal{V}_M{}^{ab} \mathcal{V}_N{}^{cd} Y_{ab,cd} \,, \qquad\qquad Z^{MN,P} = \sqrt{2}\mathcal{V}_{ab}{}^M \mathcal{V}_{cd}{}^N \mathcal{V}_{ef}{}^P \Omega^{bd} Z^{(ac)[ef]}. \tag{2.7}$$

We call B, $B^{[ab]}{}_{[cd]}$, $C_{[ab]}$ and $C^{[ab]}{}_{(cd)}$ the irreducible components that live in $\mathbf{1}$, $\mathbf{14}$, $\mathbf{5}$ and $\mathbf{35}$ of USp(4), respectively. In particular, they satisfy

$$\begin{aligned}
C^{ab} &= C^{[ab]}, & \Omega_{ab}C^{ab} &= 0, \\
B^{ab}{}_{cd} &= B^{[ab]}{}_{[cd]}, & B^{ab}{}_{cb} &= 0, & \Omega_{ab}B^{ab}{}_{cd} &= 0 = \Omega^{ab}B^{cd}{}_{ab}, \\
C^{ab}{}_{cd} &= C^{[ab]}{}_{(cd)}, & C^{ab}{}_{cb} &= 0, & \Omega_{ab}C^{ab}{}_{cd} &= 0.
\end{aligned} \tag{2.8}$$

In terms of these, the Y and Z tensors become

$$Y_{ab,cd} = \frac{1}{\sqrt{2}}\left[(\Omega_{[a|[c|}\Omega_{b]d]} - \frac{1}{4}\Omega_{ab}\Omega_{cd})B + \Omega_{[a|e|}\Omega_{b]f}B^{[ef]}{}_{[cd]}\right] \quad , \tag{2.9}$$

$$Z^{(ab)[cd]} = \frac{1}{16}\Omega^{a[c}C^{d]b} + \frac{1}{16}\Omega^{b[c}C^{d]a} - \frac{1}{8}\Omega^{ae}\Omega^{bf}C^{cd}{}_{ef} \quad . \tag{2.10}$$

Mapping the Θ-tensor from SL(5) to USp(4) we get the T-tensor defined as:

$$T_{(ef)[ab]}{}^{[cd]} = \sqrt{2}\Omega^{h[c}\delta^{d]}_{(e}\mathcal{V}^M{}_{f)h}\mathcal{V}^N{}_{ab}Y_{MN} - 2\sqrt{2}\epsilon_{MNPQR}Z^{PQ,S}\mathcal{V}^M{}_{eg}\mathcal{V}^N{}_{fh}\mathcal{V}_{ab}{}^R \mathcal{V}_S{}^{cd}\Omega^{gh} \quad . \tag{2.11}$$

In terms of these new tensors, it is possible to rewrite the quadratic constraints as follows.

$$Z^{(ab)[ef]} \left[\Omega_{ce}\Omega_{df}B + \Omega_{eg}\Omega_{fh}B^{[gh]}_{[cd]} \right] = 0,$$
$$Z^{(ab)[cd]}T_{(ab)ef}^{gh} = 0.$$

(2.12)

Let us now come to the core of our analysis: the scalar potential and the mass matrices. As usually happens in gauged supergravities, the scalar potential can be expressed as the difference of the squared fermion shifts, the latter are defined to be:

$$A_1^{ab} \equiv -\frac{1}{4\sqrt{2}} \left(\frac{1}{4}B\Omega^{ab} + \frac{1}{5}C^{ab} \right) \quad,$$

(2.13)

$$A_2^{d,abc} \equiv \frac{1}{2\sqrt{2}} \left[\Omega^{ec}\Omega^{fd} \left(C^{ab}_{ef} - B^{ab}_{ef} \right) + \frac{1}{4} \left(C^{ab}\Omega^{cd} + \frac{1}{5}\Omega^{ab}C^{cd} + \frac{4}{5}\Omega^{c[a}C^{b]d} \right) \right] \quad.$$

(2.14)

Then, the scalar potential is given by:

$$V = \frac{1}{8}|A_2|^2 - 15|A_1|^2 = -\frac{1}{128} \left(15B^2 + 2C^{ab}C_{ab} - 2B^{ab}_{cd}B^{cd}_{ab} - 2C^{[ab]}_{(cd)}C_{[ab]}^{(cd)} \right) \quad.$$

(2.15)

From this we can compute the variation of the scalar potential under a variation of the scalar fields given by $\delta_\Sigma \mathcal{V}_M^{ab} = \Sigma^{ab}_{cd}\mathcal{V}_M^{cd}$, where Σ^{ab}_{cd} is a variation along the scalar manifold $SL(5)/USp(4)$, living in the **14** representation.

$$\delta_\Sigma V = -\frac{1}{16}B^{[ab]}_{[cd]}B^{[cd]}_{[ef]}\Sigma^{[ef]}_{[ab]} + \frac{1}{32}BB^{[ab]}_{[cd]}\Sigma^{[cd]}_{[ab]} - \frac{1}{64}C^{[ab]}C_{[cd]}\Sigma^{[cd]}_{[ab]}$$
$$+ \frac{1}{32}C^{[ab]}_{(ef)}C_{[cd]}^{(ef)}\Sigma^{[cd]}_{[ab]} - \frac{1}{8}C^{[ce]}_{(af)}C^{[df]}_{(be)}\Sigma^{[ab]}_{[cd]} \quad.$$

(2.16)

In order to find the vacua of the theory, we need to solve this equation of motion together with the quadratic constraints (2.12) or (2.4). Once we obtained the solutions to these quadratic equations, we studied the gauging, the residual gauge group, the residual amount of supersymmetry preserved by each vacuum, the mass spectra, and the values of the cosmological constant.

On the vacuum, residual supersymmetry is given by the number of non-vanishing ϵ^a parameters satisfying

$$A_{2a,bcd}\epsilon^a = 0 \quad.$$

(2.17)

On the other hand, for what concerns the masses, we have computed the various mass matrices. For the gravitinos, we have:

$$\mathcal{M}^{3/2}_{ij} = \frac{5}{2}A_{1ij} \quad.$$

(2.18)

The mass matrix for the fermions is given by

$$\mathcal{L}_{m-1/2} = \frac{1}{4\sqrt{2}} \left[\frac{3}{32}\delta^b_d\delta^c_e B + \frac{1}{8}\delta^b_d\Omega_{ef}C^{fc} + B^{bc}_{de} - C^{bc}_{de} \right] \overline{\chi}_{abc}\chi^{ade},$$

(2.19)

in order to compute the masses for the fermions, we must recall that χ^{abc} transforms in the **16** of Usp(4)

$$\chi^{abc} = \chi^{[ab]c}, \qquad \Omega_{ab}\chi^{abc} = 0, \qquad \chi^{[abc]} = 0,$$

(2.20)

and therefore also the mass matrix must lie in the $\mathbf{16} \times \mathbf{16}$ representation of Usp(4). The vector mass matrix is contained in the scalar kinetic term $\frac{1}{2} P_{\mu ab}{}^{cd} P^\mu{}_{cd}{}^{ab}$, where $P_{\mu ab}{}^{cd}$ is defined by the gauge covariant space-time derivative of the scalar fields

$$\mathcal{V}_{ab}{}^M (\partial_\mu \mathcal{V}_M{}^{cd} - g A_\mu^{PQ} X_{PQ,M}{}^N \mathcal{V}^{cd}) \equiv P_{\mu ab}{}^{cd} + 2 Q_{\mu[a}{}^{[c} \delta_{b]}^{d]}. \tag{2.21}$$

The $P_{\mu ab}{}^{cd}$ lies in the $\mathbf{14}$ representation of $\mathfrak{usp}(4)$, while $Q_{\mu a}{}^c$ in the $\mathbf{10}$, (these must be imposed before computing the masses). Analogously, the mass term for the 2-forms arises from the kinetic term of the vectors, namely $\Omega_{ac} \Omega_{bd} \mathcal{H}_{\mu\nu}{}^{ab} \mathcal{H}^{cd\mu\nu}$, where $\mathcal{H}_{\mu\nu}{}^{ab}$ is given by the modified field strength tensor:

$$\mathcal{H}_{\mu\nu}^{PQ} = \mathcal{F}_{\mu\nu}^{PQ} + g Z^{PQ,R} B_{\mu\nu R}, \tag{2.22}$$

Similarly, the mass term for the 3-forms, $S_{\mu\nu\rho}^N$, arises from the kinetic term of the 2-forms, $\Omega^{ac} \Omega^{bd} \mathcal{H}_{\mu\nu\rho ab} \mathcal{H}_{cd}^{\mu\nu\rho}$. The covariant field strength for the 2-forms is given by:

$$\mathcal{H}_{\mu\nu\rho M} = 3 D_{[\mu} B_{\nu\rho]M} + 6 \epsilon_{MNPQR} A_{[\mu}^{NP} \left(\partial_\nu A_\rho^{QR} + \frac{2}{3} g X_{ST,U}{}^Q A_\nu^{RU} A_{\rho]}^{ST} \right) + g Y_{MN} S_{\mu\nu\rho}^N.$$

Note that the kinetic term for the 3-forms

$$e^{-1} \mathcal{L}_{kin} = -\frac{1}{9} \epsilon^{\mu\nu\rho\lambda\sigma\tau\kappa} g Y_{MN} S_{\mu\nu\rho}^M D_\lambda S_{\sigma\tau\kappa}^N \tag{2.23}$$

is linear in the derivative, so what one really obtains from this procedure is the mass matrix, not the square mass matrix. Furthermore, this kinetic term is not canonically normalised, in fact it is in the schematic form $Y_{MN} S^M D S^N$, so the true masses are obtained once one multiplies the mass matrix that arises from the kinetic term of the 2-forms by Y_{MN}^{-1}. For the scalar masses, we note that it is possible to parameterise the scalar fields in terms of the USp(4)-invariant, symmetric unimodular matrix $\mathcal{M}_{MN}$ defined by

$$\mathcal{M}_{MN} \equiv \mathcal{V}_M{}^{ab} \mathcal{V}_N{}^{cd} \Omega_{ac} \Omega_{bd}. \tag{2.24}$$

The scalar potential, written in terms of $\mathcal{M}_{MN}$, is

$$\begin{aligned} V = &\frac{1}{64} \left(2 \mathcal{M}^{MN} Y_{NP} \mathcal{M}^{PQ} Y_{QM} - \left(\mathcal{M}^{MN} Y_{MN} \right)^2 \right) \\ &+ Z^{MN,P} Z^{QR,S} \left(\mathcal{M}_{MQ} \mathcal{M}_{NR} \mathcal{M}_{PS} - \mathcal{M}_{MQ} \mathcal{M}_{NP} \mathcal{M}_{RS} \right). \end{aligned} \tag{2.25}$$

Therefore, by means of $\delta_\Sigma \mathcal{V}_M{}^{ab} = \Sigma^{ab}{}_{cd} \mathcal{V}_M{}^{cd}$, it is possible to calculate the second variation of the potential and, therefore, the scalar mass matrix, always recalling that $\Sigma^{ab}{}_{cd}$ belongs to the $\mathbf{14}$ representation of USp(4). Eq. (2.6) has been widely used to calculate the masses. A complete study of the field content, Lagrangian, and symmetries for maximal supersymmetric gauged supergravities in 7 space-time dimensions is given in [27].

3 Scanning for vacua: methodologies

The approach we propose to chart the vacua of the theory is based on optimisation techniques, some of the methods we adopted are not unknown to the string theory community.
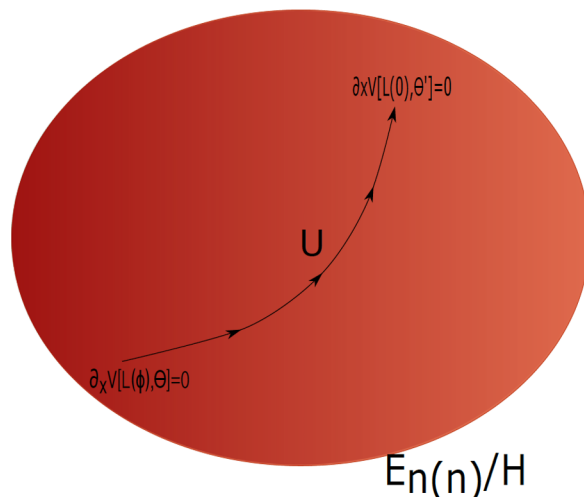
Figure 1. $E_{n(n)}$ transformations (U) mapping any point to the origin of homogeneous space $E_{n(n)}/H$.

Indeed, we used a combination of analytical and numerical tools; our starting point will be the analysis performed in [15]. There, the authors showed how every vacuum found in an explicit gauging of the theory can be mapped, by means of a duality transformation, to the origin of the scalar manifold, as shown in figure 1. This procedure is allowed by the homogeneous nature of the coset manifold, which in the case of maximal supergravities is of the form $E_{n(n)}/H$ (we do not include the Trombone symmetry in our analysis), with n the number of internal dimension and H the maximal compact subgroup of $E_{n(n)}$. Due to this method, we will be able to scan the entire space of all gaugings at once. At the origin, the physical scalars are all set to 0, but the parameters of the embedding tensor are not fixed, so in addition to considering the minimisation conditions $\delta_\Sigma V = 0$, with $\delta_\Sigma V$ given in eq. (2.16), we must also impose the quadratic constraint, eq. (2.4) or (2.12), (the linear constraint has been imposed when choosing to work with the representations $(\mathbf{1} + \mathbf{14}) + (\mathbf{5} + \mathbf{35})$ in (2.8)). The maximal compact subgroup of SL(5), SO(5), can be used to remove part of the variables, in this case 10, from the global system. In particular, we can always use local SO(5) to diagonalize Y_{MN}. This leads to the problem of solving systems of multivariate quadratic equations (($\mathcal{MQ}$) problem), which is known to be a $\mathcal{NP}$ complete task. Indeed, various cryptosystems relies on this problem in some way, $\mathcal{MQ}$ is also considered a possibility for potentially post-quantum cryptosystems, namely, it is believed not to be solved by quantum algorithms as well. Therefore, we introduce a different, related, problem: we are going to solve the systems with a certain approximation, and this can always be done in polynomial time. Indeed, optimisation techniques have been widely used by physicists and mathematicians to find solutions to systems of equations. In order to transform our problem, recalling that we have a system of homogeneous equations of the form $\sum_i^n d_i = 0$, we introduce a fit function f_{fit} defined as

$$f_{fit} = \sum_i^n (d_i)^2 \tag{3.1}$$

where i runs over the length of the system and d_i represents the polynomials in the system. Obviously, the global minima of the fit function correspond to $f_{fit} = 0$ and are the solutions to our system. This also implies that verifying that a minimisation point is also a solution takes no time. The first approach invented to solve a set of equations simultaneously is the Gradient Descent algorithm (GD), created by A. Cauchy, and since then many other attempts have been tried. Genetic Algorithms (GA), which are an instance of the broad class of Evolutionary Strategies algorithms, have already been widely used by theoretical physicists [28–32], they have the advantage to work with ill-behaved functions, while methods such as GD or stochastich GD work only on differentiable fitness functions. In this work, we are going to introduce yet another method in the set of tools used for the research of supergravity vacua, the Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES). In our search of minima of the scalar potential, we used a combination of GAs and CMA-ES, we will not dive into the explanation of GAs, since they are already known by the community, a nice introduction about these topics and other Data Science applications to String Theory can be found in [33]. Instead, we present the new technique, based on CMA-ES. CMA is an Evolutionary Strategy algorithm, therefore, it consists of a series of steps (generations), at each of them some operations are carried out before going to the next step. CMA is based on the adaptation of the covariance matrix associated with the multivariate normal distribution, from which we draw a new set of candidate solution points at each generation. The covariance matrix is not the only object that needs to be adapted at each step; the procedure implies the computation of the new mean and the overall step size as well. To compute the mean, we order the candidate solution points, based on the fitness function, and select the best μ, which results in a higher "evolution pressure". On the other hand, the Covariance Matrix is adapted in order to maximise the likelihood of precedent successful steps. Thus, CMA performs an iterated analysis of the principal components at every generation, while retaining all the principal axes. A deeper analysis of the algorithm is presented in appendix A. Let us now describe some details of the analisis. Parallelisation proved to be fundamental in making the process faster and allowing for a more global search scan. Indeed, with CMA-ES we need to choose a starting mean $\mathbf{m}^0$ and step size σ^0, together with some hyperparameters such as the time allowed to carry the computations ('timeout' parameter), the minimum value accepted for the fitness function in order to declare that a minimum was achieved ('ftarget' parameter), the precision of this result ('tolfun'), the population size λ and whether or not to activate elitist research. Parallelisation gave us the opportunity to choose more initial means $\mathbf{m}^0$ at a time, thus scanning a larger area of the parameter space. Some useful techniques have been used to adapt the algorithms to our specific problem and to render the analysis of numerical results faster.

First of all, we are dealing with systems of homogeneous quadratic equations, therefore, their solutions always pass from the origin, which helps us to restrict the area of research when setting the initial mean for the multivariate normal distribution. On the other hand, we must pay attention because the origin of the reference system is a trivial solution for any homogenous system of equations and, therefore, starting close to it may lead us always there. This could be avoided by modifying the fitness function. We found the following
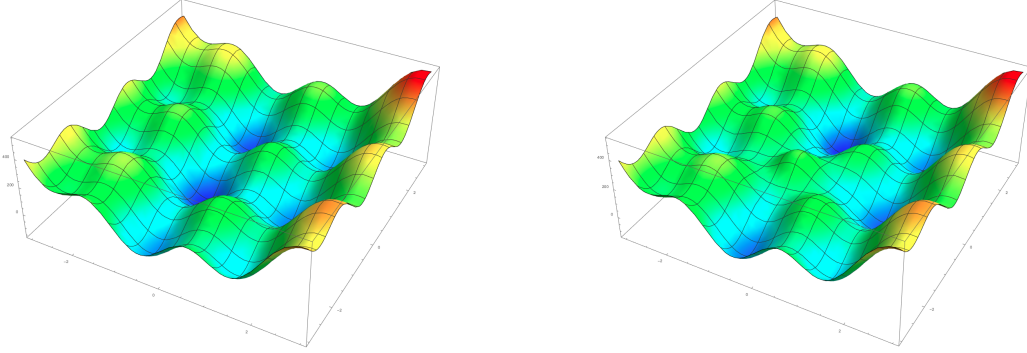
(a) Example of fitness function F_{fit} without modification.

(b) Example of fitness function with a Gaussian modification.

**Figure 2.** Modification of the fitness function after the algorithm has found a minimum.

definition to be very efficient:

$$F_{fit} = \begin{cases} f_{fit} & \text{for} \quad \sum_i^n x_i^2 > \text{threshold}, \\ 10000 & \text{for} \quad \sum_i^n x_i^2 \leq \text{threshold}. \end{cases} \tag{3.2}$$

Basically, we created a step function to avoid the algorithm from always converge in **0**. The threshold must be chosen in such a way as to leave enough parameter space near the origin to complete a full scan without hitting the barrier too often. A technique has also been implemented to avoid the algorithm from returning to the previously found minima. It consisted of adding a multivariate normal distribution function centred on the minima on top of the fitness function, as illustrated in figure 2. Basically what one does is to add a series of umbrellas on top of the fitness function to stop the algorithm from going in those directions already analised. In our case, though, the solutions to the systems are manifolds in the parameter space, so an infinite number of umbrellas would be needed to prevent the algorithm from finding again the vacuum structure under consideration. Imagine having a straight line in 2-dimensions and starting to cover it with 2-dimensional multivariate normal distributions. In 7 space-time dimensions, we were interested, above all, in (Anti)-de Sitter vacua, and this information can be used to further simplify the numerical search. Indeed, taking into account the potential in (2.15) written in terms of the fermion shifts (2.14), we can see that the potential is written as the difference between 2 squared terms. For AdS we want to impose $V = -k$, with k constant, we can always normalise k to be 1, so we need to add another equation (quadratic) to our system, analogously for dS we need to solve $V = 1$. We can always make a change of variables to completely solve this new constraint, for example, considering the case of AdS vacua, calling z_i the variables contained in $A_2^{d,abc}$ and x_i the ones in A_1^{ab}:

$$\begin{cases} z_i \to & \frac{\sinh[\psi]w_i}{\sqrt{\sum_j w_j^2}}, \\ x_i \to & \frac{\cosh[\psi]u_i}{\sqrt{\sum_j u_j^2}}. \end{cases} \tag{3.3}$$

This change of variables introduces one more variable, ψ, but solves the constraint $V = -1$, thus removing a bunch of solutions from the system. By doing so, we will be certain that the algorithm will look only for vacua with a negative cosmological constant and that the other vacua disappeared from its landscape. Analogously, we can solve the constraint $V = 1$ for dS vacua by exchanging *sinh* with *cosh* in eq. (3.3). For the Minkowski vacua we just add one more homogeneous quadratic equation ($V = 0$) to the system. The initial points have been chosen as follows, for what concerns the variables w_i and u_i in (3.3), we draw them from a normal distribution with mean **0** and standard deviation $\sigma = 4$, instead, the starting points for ψ have been evenly displaced in an interval from 0.1 to 0.5 (note that the initial variables scale exponentially with ψ, so there is no need to reach high values for the latter). We used GA techniques, such as mutation and crossing over, when the algorithm CMA got stuck in a local minimum, in order to add some noise to the candidate solution points and move them away from the local minima. Let us now describe the analysis of the results. First, we remove the solution points that are close to each other, up to a certain threshold that we set to be equal to 3σ, thus 3 times the step size, removing all candidate solutions corresponding to vacua already present in the set of solutions. Then, we studied the residual amount of supersymmetry, the gravitini masses, the signature of the Cartan matrix (providing information about the number of compact and non-compact generators of the gauge group of the theory), and the rank of $X_{PQ}{}^R$ giving us the dimension of the gauge group. The latter will provide us with everything that is needed to understand what gauging the vacuum belongs to. With all this information, it is possible to group the solution points and extract some very useful relations among them. Indeed, by plotting a variable x_i or z_i, against all other variables, it has been possible to reconstruct some analytical relations, as shown in figure 3, even by means of linear fitting. This is mainly possible when the number of unknowns is small, on the other hand, once we know the residual symmetry group of the solution, we can always use the technique of restricting the scan only to those variables which are present when that particular residual gauge symmetry has been imposed. This can lead to a new system that may be solved completely analytically, or we can always iterate the previous steps. The pseudocode for the search for AdS vacua is reported in Algorithm 1.
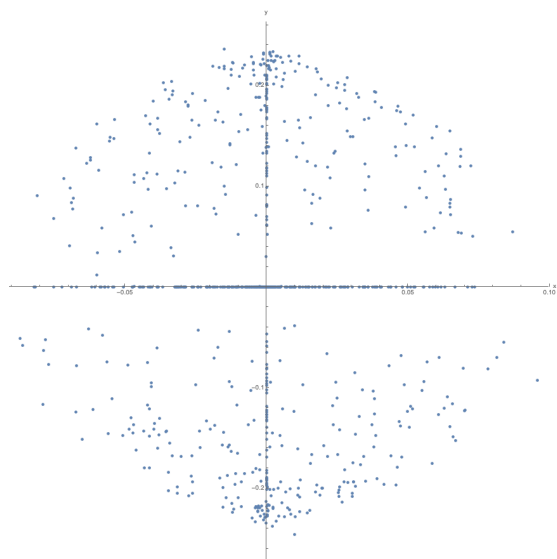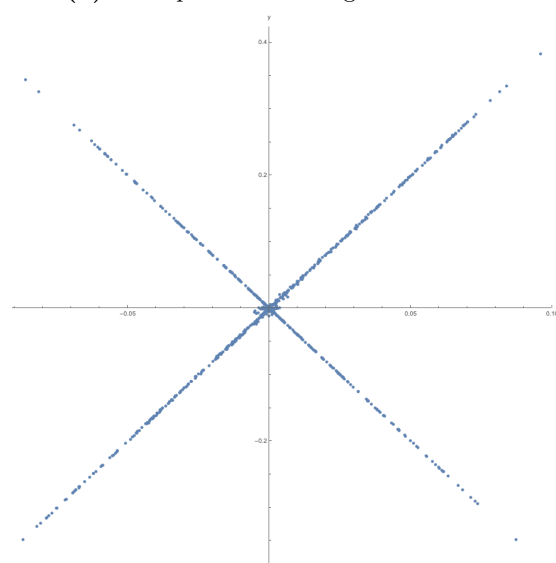
(a) Uncorrelated Variables.
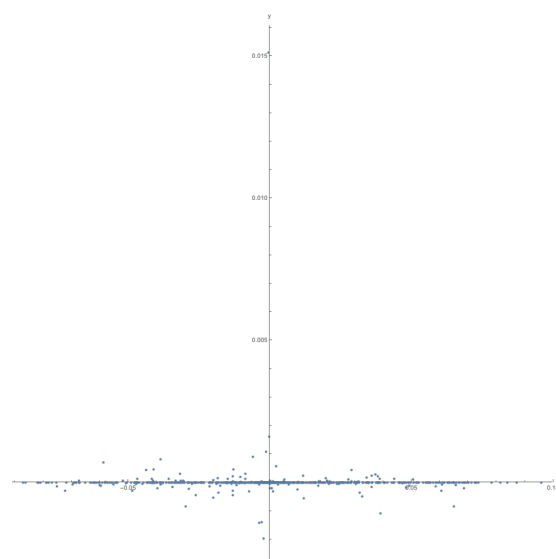
(b) No dependance among the variables.
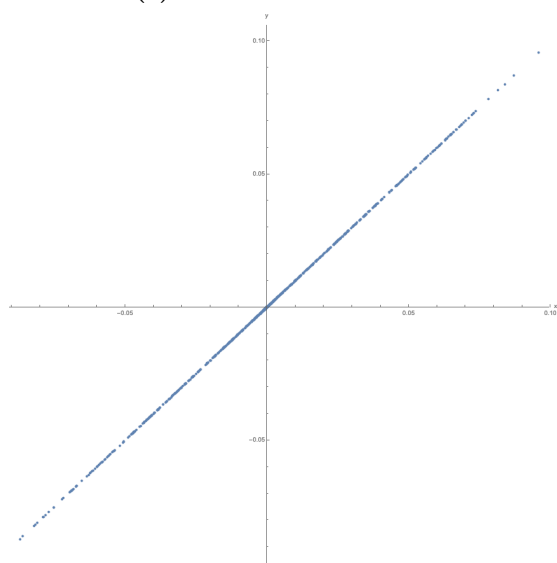
(c) variables linearly related.

(d) Galois couple.

Figure 3. Correlation among the variables.

Algorithm 1 Scanning for AdS Vacua.

▷ Initialize the number of variables, the population size, the number of values for ψ and the Step-size for CMA-ES

    Input $\# variables$, popSize , psiSize , StepSize

    $\overline{u}_i, \overline{w}_i \leftarrow \mathcal{N}(\mu = 0,\, \sigma = 3,\, size = (popSize,\, \# variables))$

    $\overline{\psi}_i \leftarrow [0.1,\, 1,\, size = psiSize]$ ▷ Linearly spaced array between 0.1 and 1

    function $f_{fit}(u_i, w_i, \psi_i)$

        return $\sum_j p_j(u_i, w_i, \psi_i)^2$

                        ▷ Sum of squared polynomials with the variables defined as in 3.3

    end function

    function $CMA\_minimization(function, \overline{u}_i, \overline{w}_i, \overline{\psi}_i,$ StepSize$)$

        return Candidate Solution Points in $\mathbb{R}^n$

          ▷ Hyperparameters, such as 'ftarget', 'seed', 'timeout', etc., must be set as well

    end function

    $Good\_\overline{u}_i,\, Good\_\overline{w}_i \leftarrow \{\}$

    while $j \leq psiSize$ **do**

        if $f_{fit}(\overline{u}_i, \overline{w}_i, \overline{\psi}_j) \leq threshold$ **then**

            $Good\_\overline{u}_i.insert(\overline{u}_i)$

            $Good\_\overline{w}_i.insert(\overline{w}_i)$

        ▷ Accept only starting points with a fitness function lower than a certain threshold

        end if

        $Candidate\_Solution \leftarrow CMA\_minimization(f_{fit}, Good\_\overline{u}_i, Good\_\overline{w}_i, \overline{\psi}_j, StepSize)$

        if $Candidate\_Solution$ is a local minimum **then**

            Use GA techniques

        end if

    end while

    for $< k = 1,\, k++,\, k \leq Length(Candidate\_Solution) >$ **do**

        for $< j = k+1,\, j++,\, j \leq Length(Candidate\_Solution) >$ **do**

            if $|Candidate\_Solution[k] - Candidate\_Solution[j]|^2 \leq 3 \cdot StepSize$ **then**

                $Candidate\_Solution.remove(j)$

            end if

        end for

    end for

    return Candidate_Solution

---

vacuum	susy	G_{gauge}	G_{res}	ref.
A1	4	SO(5)	SO(5)	[34]
A2	0	SO(5)	SO(4)	[35]
M1	0	$U(1) \ltimes \mathbb{R}^4$	U(1)	here
M2	0	$U(1) \ltimes \mathbb{R}^6$	U(1)	here

Table 1. Summary of vacua.

$L^2 m_{3/2}^2$	$\left[\frac{25}{16}\right]_{\times 4}$
$L^2 m_{\text{vec}}^2$	$[0]_{\times 10}$
$L^2 m_{2-\text{forms}}^2$	$[0]_{\times 5}$
$L^2 m_{3-\text{forms}}^2$	$[1]_{\times 5}$
$L^2 m_{1/2}^2$	$\left[\frac{9}{16}\right]_{\times 16}$
$L^2 m_{\text{scal}}^2$	$[8]_{\times 14}$

Table 2. Masses for the AdS vacuum A1.

$L^2 m_{3/2}^2$	$\left[\frac{27}{16}\right]_{\times 4}$
$L^2 m_{\text{vec}}^2$	$\left[\frac{3}{4}\right]_{\times 4}, [0]_{\times 6}$
$L^2 m_{2-\text{forms}}^2$	$[0]_{\times 5}$
$L^2 m_{3-\text{forms}}^2$	$\left[\frac{3}{4}\right]_{\times 4}, 3_{\times 1}$
$L^2 m_{1/2}^2$	$[0]_{\times 4}, \left[\frac{3}{16}\right]_{\times 12}$
$L^2 m_{\text{scal}}^2$	$[-12]_{\times 1}, [12]_{\times 9}, [0]_{\times 4}$

Table 3. Masses for the AdS vacuum A2.

4 Results

In this section we are going to present the vacua found through the methods presented before. As has been explained, we have been able to reconstruct directly or obtain analytical solutions from residual symmetry considerations starting from numerical ones. We found the two well-known AdS vacua, the maximal supersymmetric one [34] and the non-supersymmetric one [35], of which we present the mass spectrum. In addition, we found 2 new Minkowski vacua. The summary of the vacua is presented in table 1.

All AdS vacua appear in the SO(5) theory, which is a consistent truncation of M-theory compactified on S^4.

4.1 Mass spectra

In this section we present some of the masses for all the vacua listed in table 1. The masses for the AdS backgrounds, tables 2 and 3, are normalised in terms of the squared AdS radius $L^2 = |15/V|$, so that supersymmetric gravitinos have a normalised squared mass of 25/16.

The spectrum for the vacuum A1 agrees with the one found in [34], and it is determined by the representation of the superconformal algebra in six dimensions [36, 37]. The scalar masses for the A2 case are the same as those in [35], ensuring that this is the same non-supersymmetric AdS vacuum. In order to compute the masses for the fermions, we note that in the A2 case supersymmetry is completely broken, therefore a super-Higgs mechanism is in action there. Consequently, 4 fermions, which correspond to the Goldstinos, must be

$m_{3/2}^2$	$[m^2]_{\times 2}, [0]_{\times 2}$
m_{vec}^2	$[(2m)^2]_{\times 2}, [(m)^2]_{\times 2}, [0]_{\times 6}$
$m_{2-forms}^2$	$[0]_{\times 3}, [m^2]_{\times 2}$
$m_{3-\text{forms}}^2$	$[m^2]_{\times 2}, [0]_{\times 3}$
m_{scal}^2	$[-(4m)^2]_{\times 4}, [-(2m)^2]_{\times 2}, [0]_{\times 8}$

Table 4. Masses for the Minkowski vacuum M1.

$m_{3/2}^2$	$\left[\frac{1}{4}(m_1^2 - m_2^2)\right]_{\times 4}$
m_{vec}^2	$[(m_1)^2]_{\times 2}, [(m_1 \pm m_2)^2]_{\times 2}, [0]_{\times 4}$
$m_{2-\text{forms}}^2$	$[0]_{\times 3}, [m_2^2]_{\times 2}$
$m_{3-\text{forms}}^2$	$[m_1^2]_{\times 2}, [0]_{\times 3}$
m_{scal}^2	$[0]_{\times 8}, [-(4m_1)^2]_{\times 2}, [-(4m_2)^2]_{\times 2}, [-(2m_2)^2]_{\times 2}$

Table 5. Masses for the Minkowski vacuum M2.

eaten by the corrisponding gravitinos. The relevant part of the Lagrangian is

$$\mathcal{L}_{\text{mass}} = -\frac{5}{2}g A_1^{ab}\Omega_{bc}\overline{\psi}_{\mu a}\Gamma^{\mu\nu}\psi_\nu^c + \frac{1}{4}g A_2^{d,abc}\Omega_{de}\overline{\chi}_{abc}\Gamma^\mu\psi_\mu^e + \frac{g}{4}\mathcal{M}^{abc}{}_{\text{def}}\overline{\chi}_{abc}\chi^{\text{def}}. \quad (4.1)$$

where Γ^μ are the 7-dimensional Dirac Gamma matrices and

$$\mathcal{M}^{abc}{}_{\text{def}} = \frac{1}{\sqrt{2}}\delta_d^a\left[\frac{3}{32}\delta_e^b\delta_f^c B + \frac{1}{8}\delta_e^b\Omega_{fg}C^{gc} + B^{bc}{}_{ef} - C^{bc}{}_{ef}\right] \quad (4.2)$$

obtained from (2.19). We can do a field redefinition for the gravitinos

$$\psi_\mu^a \rightarrow \psi_\mu^a + \frac{1}{120}\Omega^{ab}A_{1bc}^{-1}A_2^{c,efg}\Gamma_\mu\chi_{efg} \quad (4.3)$$

in order to gauge away the interaction term between fermions and gravitinos, this will add a new term to the mass matrix of the fermions, (2.19), rendering the Goldstinos null eigenvectors. The supersymmetric AdS vacuum, (A1), is perturbatively stable, respecting the Breitenlohner-Freedman bound [38–40]

$$m^2 \geq -\frac{(d-1)^2}{4}, \quad (4.4)$$

for scalar degrees of freedom in AdS_d. On the other hand, (A2), is pertubatively unstable, as already found in [35], thus corroborating the hypothesis about the instability of non-supersymmetric AdS spacetimes formulated in the contest of the Swampland programme [41]. For the Minkowski vacua, the spectra are shown in tables 4 and 5.

In appendix B we present the form of the representations that make up the T-tensor, for each vacua we found. It is clear that this is not an exaustive analysis, indeed some vacua are missing from this scan, for example the Scherk-Schwarz Minkowski vacuum, with gauge group $U(1) \times \mathbb{R}^6$, first found in [27], did not appear in our scan. In any case, it is possible to find it, once one limits the analysis to the vacua preserving a residual symmetry, in this case $U(1)$.

D	Duality Group	$\Theta_M{}^\alpha$
9	GL(2)	$\mathbf{2}^{-3} + \mathbf{3}^{+4}$
8	SL(2)$\times$ SL(3)	$(\mathbf{2}, \mathbf{3}) + (\mathbf{2}, \mathbf{6}')$
7	SL(5)	$\mathbf{15} + \mathbf{40}'$
6	SO(5,5)	$\mathbf{144_c}$
5	$E_{6(6)}$	$\mathbf{351}'$
4	$E_{7(7)}$	$\mathbf{912}$
3	$E_{8(8)}$	$\mathbf{1 + 3875}$

Table 6. Representations of the Embedding Tensor.

5 Outlook and future directions

In section 3 we presented a new method, based on Evolutionary Strategies optimization techniques. The power of these tools combined with the thorough analysis of the numerical data allowed the reconstruction of analytical solutions and their mass spectra. The analysis has been carried out for $D = 7$ space-time dimensions, for reasons already explained in the introduction, but it would be interesting to see whether our procedure can be used also for different number of dimensions and for fewer amount of supersymmetry. The number of parameters in the embedding tensor, and therefore of variables, and the number of equations grow as the number of dimensions get lower. This is illustrated in table 6, where the representations of the embedding tensor for maximal theories in each dimension are presented, giving the number of variables in the system of equations (there is the possibility to remove some of them using the H-redundancy of the scalar coset manifold). This implies that numerical methods are necessary when dealing with lower dimensional theories, because analytical results can be achieved in very special cases, either by fixing the gauging or restricting the analysis to vacua preserving some residual symmetry. Numerical methods, instead, would allow for a more general scan. On the other hand, with a large number of non-vanishing parameters in the vacua it is no longer possible to use the reconstruction techniques presented in section 3 and more subtle ways are needed if we want to obtain analytical results. We also compared our method to some optimization tools available from TensorFlow libraries (Adam optimizer) and to other algorithms as well, finding CMA-ES and our implementations more efficient, given the cospicuous extension of the literature in Optimization it has been impossible to test our problem against many of the optimization techniques. Another interesting step for the future is to compare our methodology with other numerical approaches; maximal supergravities in $D = 7$ space-time dimensions are a good testbed due to the small number of variables and the interesting physical features.

A Covariance Matrix Adaptation — Evolutionary Strategy (CMA-ES)

In this appendix, we present some details about the CMA-ES algorithm and its operational mode. CMA-ES, in short, CMA, is an evolution strategy (ES) algorithm that, as in the case

of GA, only needs the fitness function as accessible information [25, 42, 42–48]. Therefore, differently from GD or stocastic GD, we do not require the function to be differentiable, it can also be not continuous. CMA, as the name suggests, is based on the "adaption" of a normal distribution to the fitness function under consideration (with its level curves). Let us analyse the algorithm in depth and consider a multivariate normal distribution, $\mathcal{N}(\mathbf{m}, \mathbf{C})$, which is determined by the mean $\mathbf{m} \in \mathbb{R}^n$ (in our case, n is the number of free parameters in the embedding tensor) and by the symmetric, positive definite covariance matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$. Covariance matrices are associated with the ellipsoid $\{\mathbf{x} \in \mathbb{R}^n | \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} = 1\}$, where the principal axis of the latter are the eigenvectors of $\mathbf{C}$, and the lengths of the squared axis are the eigenvalues of the covariance matrix. We can always diagonalize the covariance matrix by means of an orthogonal matrix $\mathbf{B}$ whose columns are the eigenvectors of $\mathbf{C}$ with unit length, $\mathbf{C} = \mathbf{B}(\mathbf{D})^2 \mathbf{B}^T$. Then, it is also possible to write the normal distribution as

$$\mathcal{N}(\mathbf{m}, \mathbf{C}) \sim \mathbf{m} + \mathcal{N}(\mathbf{0}, \mathbf{C}) \sim \mathbf{m} + \mathbf{C}^{-1/2} \mathcal{N}(\mathbf{0}, \mathbf{I}) \sim \mathbf{m} + \mathbf{B}\mathbf{D}\mathbf{B}^T \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{A.1}$$

with $\mathbf{I}$ the $n \times n$ identity matrix. At each step of the process, we generate a new population of points (which in GA are called offsprings) by drawing them from a multivariate normal distribution:

$$\mathbf{x}_j^{g+1} \sim \mathcal{N}\big(\mathbf{m}^g, (\sigma^g)^2 \mathbf{C}^g\big) \qquad \text{with} \quad j = 1, \dots, \lambda \tag{A.2}$$

Superscripts g, $g+1$, etc. label the generations, λ is the population size, and $\sigma^g \in \mathbb{R}^+$ is the "overall" standard deviation (step size) at generation g. Now we need to explain how the mean, the covariance matrix, and the standard deviation are computed for the next generation $g + 1$.

The mean. The new mean $\mathbf{m}^{g+1}$ is simply selected with a weighted average of the μ best points of the population:

$$\mathbf{m}^{g+1} = \sum_{j=1}^{\mu} w_j \mathbf{x}_{j:\lambda}^{g+1}, \qquad \text{with} \quad \sum_{j=1}^{\mu} w_j = 1 \text{ and } w_j > 0. \tag{A.3}$$

$w_j \in \mathbb{R}^+$ with $j = 1, \dots, \mu$ are positive ordered weights, that is, $w_1 \geq w_2 \geq \dots \geq w_\mu > 0$. If $w_j = 1/\mu$ for each j, we obtain the mean value for the best μ points. $\mathbf{x}_{j:\lambda}^{g+1}$ represents the j-th best individual of the population, meaning, $f(\mathbf{x}_{1:\lambda}^{g+1}) \leq f(\mathbf{x}_{2:\lambda}^{g+1}) \leq \dots \leq f(\mathbf{x}_{\lambda:\lambda}^{g+1})$. An essential quantity is the **variance effective selection mass**

$$\mu_{\text{eff}} = \bigg(\sum_{j=1}^{\mu} w_j^2 \bigg)^{-1}. \tag{A.4}$$

It is possible to show, from the definition of w_j that $1 \leq \mu_{\text{eff}} \leq \mu$ and that $\mu_{\text{eff}} = \mu$ only in the case where all the weights are the same and equal to $1/\mu$. Usually, $\mu \approx \lambda/2$ and $w_i \propto \mu - i + 1$.

The covariance matrix. First, let us define the empirical covariance matrix $\mathbf{C}_{emp}^{g+1}$, which is nothing more than an estimate of the covariance matrix $\mathbf{C}^g$:

$$\mathbf{C}_{emp}^{g+1} = \frac{1}{\lambda - 1} \sum_{i=1}^{\lambda} \bigg(\mathbf{x}_i^{g+1} - \frac{1}{\lambda} \sum_{j=1}^{\lambda} \mathbf{x}_j^{g+1} \bigg) \bigg(\mathbf{x}_i^{g+1} - \frac{1}{\lambda} \sum_{j=1}^{\lambda} \mathbf{x}_j^{g+1} \bigg)^T. \tag{A.5}$$

This estimator has to be modified to obtain a maximum likelihood estimator of $\mathbf{C}^g$, by defining

$$\mathbf{C}_{\lambda}^{g+1} = \frac{1}{\lambda} \sum_{j=1}^{\lambda} \left(\mathbf{x}_j^{g+1} - \mathbf{m}^g\right)\left(\mathbf{x}_j^{g+1} - \mathbf{m}^g\right)^T. \tag{A.6}$$

The difference between $\mathbf{C}_{emp}^{g+1}$ and $\mathbf{C}_{\lambda}^{g+1}$ is what is used as the mean value. The former uses the mean obtained from the entire population, thus estimating the variance of the sampled points, and the latter instead uses the mean obtained by (A.3), therefore estimating the sampled steps, $\mathbf{x}_j^{g+1} - \mathbf{m}^g$. We are going to modify this estimator again and define

$$C_{\mu}^{g+1} = \sum_{j=1}^{\mu} w_j \left(\mathbf{x}_{j:\lambda}^{g+1} - \mathbf{m}^g\right)\left(\mathbf{x}_{j:\lambda}^{g+1} - \mathbf{m}^g\right)^T. \tag{A.7}$$

$\mathbf{C}_{\mu}^{g+1}$ is an estimator of the variance of selected steps (the best / successful steps μ). We have some conditions on μ_{eff} in order for $\mathbf{C}_{\mu}^{g+1}$ to be a **reliable** estimator. Indeed, μ_{eff} has to be large enough to prevent the condition numbers (given a matrix A and a linear system Ax=b, with x unknown, they measure how sensitive the solution of the system to a change in b is, high condition numbers imply that small changes in b generate huge modifications in the solution) of $\mathbf{C}_{\mu}^{g+1}$ to be smaller than 10 for the fitness function of the sphere: $f_{sphere}(\mathbf{x}) = \sum_{i=1}^{n} x_i^2$; empirically, it is seen that $\mu_{\text{eff}} \approx 10n$ is a good choice. To avoid this problem for a small population, we will modify the update of the covariance matrix again. In order to obtain an algorithm that converges faster, we need a small population, on the other hand, to obtain a more global search the population has to increase. For a small population, also $\mu_{\text{eff}} \approx \lambda/4$ (which is the choice to take to have reasonable w_j) has to be small, then $\mathbf{C}_{\mu}^{g+1}$ is not a reliable estimator, in order to circumvent this, we define a new covariance matrix that takes into consideration the information we have from previous generations. Defining $\mathbf{C}^0 = \mathbf{I}$ and the learning rate $0 < c_{cov} \leq 1$, then

$$\begin{aligned}
\mathbf{C}^{g+1} &= (1 - c_{cov})\mathbf{C}^g + c_{cov}\left(\frac{1}{\sigma^g}\right)^2 \mathbf{C}_{\mu}^{g+1} \\
&= (1 - c_{cov})\mathbf{C}^g + c_{cov} \sum_{j=1}^{\mu} w_j \left(\frac{\mathbf{x}_{j:\lambda}^{g+1} - \mathbf{m}^g}{\sigma^g}\right)\left(\frac{\mathbf{x}_{j:\lambda}^{g+1} - \mathbf{m}^g}{\sigma^g}\right)^T.
\end{aligned} \tag{A.8}$$

Step sizes σ^g have been integrated to ensure that $\mathbf{C}_{\mu}^g$ from different generations are comparable. If $c_{cov} = 1$ the covariance matrix collapses to $\mathbf{C}_{\mu}^{g+1}$ and no information from previous generations is retained, on the other hand, if $c_{cov} = 0$, $\mathbf{C}^{g+1} = \mathbf{C}^0$ and there is no learning. This kind of update, represented in (A.8) update, is called rank-μ update, because the sum goes from 1 to μ. Eq. (A.8) is iterative and can be expanded as

$$C^{g+1} = (1 - c_{cov})^{g+1}\mathbf{C}^0 + c_{cov} \sum_{j=0}^{g} (1 - c_{cov})^{g-j}\left(\frac{1}{\sigma^j}\right)^2 \mathbf{C}_{\mu}^{j+1}. \tag{A.9}$$

Picking high values for c_{cov} leads to degenerate covariance matrices, while small values imply slow learning, a good choice is $c_{cov} \approx \mu_{\text{eff}}/n^2$. Small population sizes λ lead to a

large number of generations and therefore to a faster adaptation for the covariance matrix. A final step is necessary to update the covariance matrix: **cumulation**. In fact, information about "signs" of the steps the strategy took generation after generation has not been used so far. To do so, we introduce the **evolution path**. An evolution path is any sequence of succesive steps taken by the strategy, taking the sum of these steps is referred as cumulation; for instance, for three steps we have

$$\frac{\mathbf{m}^{g+1} - \mathbf{m}^g}{\sigma^g} + \frac{\mathbf{m}^g - \mathbf{m}^{g-1}}{\sigma^{g-1}} + \frac{\mathbf{m}^{g-1} - \mathbf{m}^{g-2}}{\sigma^{g-2}}. \tag{A.10}$$

Defining the 0-th order evolution path $\mathbf{p}_c^0 = \mathbf{0}$, we use the exponential smoothing and define iteratively

$$\mathbf{p}_c^{g+1} = (1 - c_c)\mathbf{p}_c^g + \sqrt{c_c(2 - c_c)\mu_{\text{eff}}} \frac{\mathbf{m}^{g+1} - \mathbf{m}^g}{\sigma^g}, \tag{A.11}$$

with $0 \geq c_c \leq 1$ a new learning rate for the evolution path, the normalisation factor $\sqrt{c_c(2 - c_c)\mu_{\text{eff}}}$ is dictated by the demand that $\mathbf{p}_c^{g+1}$ be extracted from a normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{C})$. When $c_c = 0$ there is no learning and $\mathbf{p}_c^g = \mathbf{0}$. Putting everything together, we obtain the update of the covariance matrix:

$$\begin{aligned}
\mathbf{C}^{g+1} = (1 - c_{cov})\mathbf{C}^g &+ \frac{c_{cov}}{\mu_{cov}}\mathbf{p}_c^{g+1}\mathbf{p}_c^{g+1\,T} \\
&+ c_{cov}\left(1 - \frac{1}{\mu_{cov}}\right)\sum_{j=1}^{\mu} w_j \left(\frac{\mathbf{x}_{j:\lambda}^{g+1} - \mathbf{m}^g}{\sigma^g}\right)\left(\frac{\mathbf{x}_{j:\lambda}^{g+1} - \mathbf{m}^g}{\sigma^g}\right)^T,
\end{aligned} \tag{A.12}$$

with $\mu_{cov} \geq 1$ and should be $\mu_{cov} = \mu_{\text{eff}}$. Eq. (A.12) reduces to eq. (A.8) in the case $\mu_{cov} \to \infty$, so information from the last generation is taken into consideration by the rank-μ update and information from previous generations, instead, is exploited by the evolution path update, which is relevant above all for small population sizes.

The step size. An evolution path is also used to update the step size σ with a method called **cumulative step size adaptation**:

- Whenever the evolution path is long, the steps are going in the same direction (approximatively), so they are correlated. Consequently, we can cover the same distance with longer but fewer steps, and the step size must be increased.

- When the evolution path is short, the steps cancel among each other, and the step size should be decreased.

- The optimal situation is that the steps are totally uncorrelated and orthogonal with respect to the previous and following ones.

We need to define what long- and short-evolution paths mean. In this respect, we compare the latter with the expected length under random selection, which means that the steps are not correlated with each other. If our strategy finds that the evolution paths are longer than the uncorrelated ones, σ has to be increased, and vice versa.

The evolution path $\mathbf{p}_c^{g+1}$ depends on its direction, therefore we define the conjugate path

$$\mathbf{p}_\sigma^{g+1} = (1 - c_\sigma)\mathbf{p}_\sigma^g + \sqrt{c_\sigma(2 - c_\sigma)\mu_{\text{eff}}}(\mathbf{C}^g)^{-\frac{1}{2}}\frac{\mathbf{m}^{g+1} - \mathbf{m}^g}{\sigma^g}, \qquad (A.13)$$

with $0 < c_\sigma < 1$ a learning rate and $(\mathbf{C}^g)^{-\frac{1}{2}} \equiv \mathbf{B}^g(\mathbf{D}^g)^{-1}\mathbf{B}^{g\,T}$. Whenever $(\mathbf{C}^g)^{-\frac{1}{2}} \neq \mathbf{I}$ it aligns step $\mathbf{m}^{g+1} - \mathbf{m}^g$ with the coordinate system produced by $\mathbf{B}^g$. In particular, $\mathbf{B}^{g\,T}$ rotates the system so that the columns of $\mathbf{B}^g$ become the axis. $(\mathbf{D}^g)^{-1}$ rescales the length of the axis so that they measure the distances in the same way. $\mathbf{B}^g$ rotates everything back, allowing us to compare the directions of the various steps. By adding the matrix $(\mathbf{C}^g)^{-\frac{1}{2}}$ in eq. (A.13) we ensure the independence of $\mathbf{p}_\sigma^{g+1}$ from the direction of the steps. Then we compare the length of $\mathbf{p}_\sigma^{g+1}$ with the expected length of the evolution path obtained from random selection $E\left[||\mathcal{N}(\mathbf{0}, \mathbf{I})||\right]$ and define the step size

$$\sigma^{g+1} = \sigma^g \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{||\mathbf{p}_\sigma^{g+1}||}{E\left[||\mathcal{N}(\mathbf{0}, \mathbf{I})||\right]} - 1\right)\right), \qquad (A.14)$$

where $d_\sigma \approx 1$ is a damping parameter and $E\left[||\mathcal{N}(\mathbf{0}, \mathbf{I})||\right] = \sqrt{2}\Gamma\left(\frac{n+1}{2}\right)/\Gamma\left(\frac{n}{2}\right) \approx \sqrt{n} + \mathcal{O}(1/n)$ is the expectation value of the Euclidean norm for a multivariate normal distribution with the identity matrix as the covariance matrix. From eq. (A.14) we can see that, whenever $||\mathbf{p}_\sigma^{g+1}|| > E\left[||\mathcal{N}(\mathbf{0}, \mathbf{I})||\right]$, σ^g increases and viceversa when $||\mathbf{p}_\sigma^{g+1}|| < E\left[||\mathcal{N}(\mathbf{0}, \mathbf{I})||\right]$.

It has been proved, in a survey about Black-Box optimizations [49], that CMA-ES outranks other 31 optimisation algorithms and that its performance is outstanding for rugged and ill-conditioned functions with large search spaces.

B   T-tensor at the vacuum in the 7 dimensional theory

In this appendix we provide an instance of the value of the irreducible USp(4) representations composing the T-tensor generating the vacua of table 1. For all the examples, we have chosen a basis where

$$\Omega = \mathbb{1} \otimes i\sigma_2. \qquad (B.1)$$

A1. The maximal supersymmetric Anti de Sitter vacuum has

$$Bs = \kappa, \qquad B^{ab}{}_{cd} = C^{ab}{}_{cd} = C^{ab} = \mathbf{0}. \qquad (B.2)$$

A2. The SO(4) non-supersymmetric AdS vacuum is given by the following representations:

$$Bs = \kappa, \qquad\qquad C^{ab}{}_{cd} = C^{ab} = \mathbf{0}, \qquad (B.3)$$

$$\begin{aligned} B^{12}{}_{12} = B^{12}{}_{43} = B^{34}{}_{21} = B^{34}{}_{34} &= \frac{\kappa}{6}, \\ B^{13}{}_{31} = B^{14}{}_{41} = B^{23}{}_{32} = B^{24}{}_{42} &= \frac{\kappa}{12}. \end{aligned} \qquad (B.4)$$

There are other non-vanishing entries in the **14** representation, which are related by symmetries in the indices and therefore have not been reported here.

M1.

$$Bs = \kappa, \qquad\qquad C^{12} = C^{43} = \frac{5}{4}\kappa, \qquad \text{(B.5)}$$

$$B^{12}{}_{21} = B^{12}{}_{34} = B^{34}{}_{12} = B^{34}{}_{43} = \frac{1}{4}\kappa,$$

$$B^{13}{}_{13} = B^{14}{}_{14} = B^{23}{}_{23} = B^{24}{}_{24} = \frac{1}{8}\kappa, \qquad\qquad \text{(B.6)}$$

$$B^{13}{}_{41} = B^{14}{}_{13} = B^{23}{}_{24} = B^{24}{}_{32} = \frac{5i}{8}\kappa,$$

$$C^{14}{}_{13} = C^{23}{}_{24} = C^{31}{}_{14} = C^{42}{}_{23} = \frac{5i}{8}\kappa. \qquad\qquad \text{(B.7)}$$

Here, one needs always to keep in mind that there are other, non-reported, entries of these tensors which are related to the ones above by symmetries of representations.

M2.

$$Bs = \kappa_1, \qquad\qquad C^{14} = C^{23} = 4i\kappa_2, \qquad \text{(B.8)}$$

$$B^{12}{}_{21} = B^{12}{}_{34} = B^{34}{}_{12} = B^{34}{}_{43} = \frac{1}{4}\kappa_1,$$

$$B^{13}{}_{13} = B^{24}{}_{24} = \frac{3}{4}\kappa_1, \qquad\qquad \text{(B.9)}$$

$$B^{14}{}_{41} = B^{23}{}_{32} = \frac{\kappa_1}{2},$$

$$C^{14}{}_{12} = C^{21}{}_{14} = C^{21}{}_{23} = C^{23}{}_{43} = C^{32}{}_{12} = C^{34}{}_{14} = C^{34}{}_{23} = C^{41}{}_{34} = i\kappa_2,$$

$$C^{13}{}_{11} = C^{13}{}_{33} = C^{42}{}_{22} = C^{42}{}_{44} = 2i\kappa_2. \qquad\qquad \text{(B.10)}$$

Again, symmetries must be imposed on these tensors.

Acknowledgments

References

[1] L. Susskind, *The Anthropic landscape of string theory*, hep-th/0302219 [INSPIRE].

[2] G. Obied, H. Ooguri, L. Spodyneiko and C. Vafa, *de Sitter Space and the Swampland*, arXiv:1806.08362 [INSPIRE].

[3] W. Nahm, *Supersymmetries and their Representations*, *Nucl. Phys. B* **135** (1978) 149 [INSPIRE].

[4] N.P. Warner, *Some Properties of the Scalar Potential in Gauged Supergravity Theories*, *Nucl. Phys. B* **231** (1984) 250 [INSPIRE].

[5] N.P. Warner, *Some New Extrema of the Scalar Potential of Gauged $N = 8$ Supergravity*, *Phys. Lett. B* **128** (1983) 169 [INSPIRE].

[6] T. Fischbacher, *Fourteen new stationary points in the scalar potential of SO(8)-gauged $N = 8$, $D = 4$ supergravity*, *JHEP* **09** (2010) 068 [arXiv:0912.1636] [INSPIRE].

[7] T. Fischbacher, *Numerical tools to validate stationary points of SO(8)-gauged $N = 8$ $D = 4$ supergravity*, *Comput. Phys. Commun.* **183** (2012) 780 [arXiv:1007.0600] [INSPIRE].

[8] T. Fischbacher, *The Encyclopedic Reference of Critical Points for SO(8)-Gauged $N=8$ Supergravity. Part 1: Cosmological Constants in the Range $-\Lambda/g^2 \in [6 : 14.7)$*, arXiv:1109.1424 [INSPIRE].

[9] I.M. Comsa, M. Firsching and T. Fischbacher, *SO(8) Supergravity and the Magic of Machine Learning*, *JHEP* **08** (2019) 057 [arXiv:1906.00207] [INSPIRE].

[10] N. Bobev, T. Fischbacher, F.F. Gautason and K. Pilch, *A cornucopia of AdS_5 vacua*, *JHEP* **07** (2020) 240 [arXiv:2003.03979] [INSPIRE].

[11] N. Bobev, T. Fischbacher, F.F. Gautason and K. Pilch, *New AdS_4 Vacua in Dyonic ISO(7) Gauged Supergravity*, arXiv:2011.08542 [INSPIRE].

[12] D. Berman, T. Fischbacher, G. Inverso and B. Scellier, *Vacua of ω-deformed SO(8) supergravity*, *JHEP* **06** (2022) 133 [arXiv:2201.04173] [INSPIRE].

[13] C. Krishnan, V. Mohan and S. Ray, *Machine Learning $\mathcal{N} = 8$, $D = 5$ Gauged Supergravity*, *Fortsch. Phys.* **68** (2020) 2000027 [arXiv:2002.12927] [INSPIRE].

[14] S.P. Li, R.B. Peschanski and C.A. Savoy, *Mass Degeneracy and the Superhiggs Mechanism*, *Nucl. Phys. B* **289** (1987) 206 [INSPIRE].

[15] G. Dall'Agata and G. Inverso, *On the Vacua of $N = 8$ Gauged Supergravity in 4 Dimensions*, *Nucl. Phys. B* **859** (2012) 70 [arXiv:1112.3345] [INSPIRE].

[16] A. Borghese, A. Guarino and D. Roest, *All G_2 invariant critical points of maximal supergravity*, *JHEP* **12** (2012) 108 [arXiv:1209.3003] [INSPIRE].

[17] A. Borghese, G. Dibitetto, A. Guarino, D. Roest and O. Varela, *The SU(3)-invariant sector of new maximal supergravity*, *JHEP* **03** (2013) 082 [arXiv:1211.5335] [INSPIRE].

[18] A. Borghese, A. Guarino and D. Roest, *Triality, Periodicity and Stability of SO(8) Gauged Supergravity*, *JHEP* **05** (2013) 107 [arXiv:1302.6057] [INSPIRE].

[19] F. Catino, G. Dall'Agata, G. Inverso and F. Zwirner, *On the moduli space of spontaneously broken $N = 8$ supergravity*, *JHEP* **09** (2013) 040 [arXiv:1307.4389] [INSPIRE].

[20] A. Gallerati, H. Samtleben and M. Trigiante, *The $\mathcal{N} > 2$ supersymmetric AdS vacua in maximal supergravity*, *JHEP* **12** (2014) 174 [arXiv:1410.0711] [INSPIRE].

[21] G. Dibitetto, A. Guarino and D. Roest, *Charting the landscape of $N = 4$ flux compactifications*, *JHEP* **03** (2011) 137 [arXiv:1102.0239] [INSPIRE].

[22] N.S. Deger, C. Eloy and H. Samtleben, *$\mathcal{N} = (8, 0)$ AdS vacua of three-dimensional supergravity*, *JHEP* **10** (2019) 145 [arXiv:1907.12764] [INSPIRE].

[23] G. Dall'agata, G. Inverso and D. Partipilo, *Old and new vacua of 5D maximal supergravity*, *JHEP* **04** (2021) 039 [arXiv:2101.04149] [INSPIRE].

[24] G. Dall'Agata, G. Inverso and M. Trigiante, *Evidence for a family of SO(8) gauged supergravity theories*, *Phys. Rev. Lett.* **109** (2012) 201301 [`arXiv:1209.0760`] [InSPIRE].

[25] N. Hansen, *The cma evolution strategy: a comparing review. Towards a new evolutionary computation*, in J.A. Lozano, P. Larrañaga, I. Inza and E. Bengoetxea, *Towards a New Evolutionary Computation*, *Studies in Fuzziness and Soft Computing* **192** (2006) 752.

[26] B. de Wit, H. Samtleben and M. Trigiante, *On Lagrangians and gaugings of maximal supergravities*, *Nucl. Phys. B* **655** (2003) 93 [`hep-th/0212239`] [InSPIRE].

[27] H. Samtleben and M. Weidner, *The Maximal D = 7 supergravities*, *Nucl. Phys. B* **725** (2005) 383 [`hep-th/0506237`] [InSPIRE].

[28] S. Abel, D.G. Cerdeño and S. Robles, *The Power of Genetic Algorithms: what remains of the pMSSM?*, `arXiv:1805.03615` [InSPIRE].

[29] S. Abel and J. Rizos, *Genetic Algorithms and the Search for Viable String Vacua*, *JHEP* **08** (2014) 010 [`arXiv:1404.7359`] [InSPIRE].

[30] A. Yamaguchi and H. Nakajima, *Landau gauge fixing supported by genetic algorithm*, *Nucl. Phys. B Proc. Suppl.* **83** (2000) 840 [`hep-lat/9909064`] [InSPIRE].

[31] J.F. Markham and T.D. Kieu, *Evolutionary algorithms applied to Landau gauge fixing*, *Nucl. Phys. B Proc. Suppl.* **73** (1999) 868 [`hep-lat/9809143`] [InSPIRE].

[32] B.C. Allanach, D. Grellscheid and F. Quevedo, *Genetic algorithms and experimental discrimination of SUSY models*, *JHEP* **07** (2004) 069 [`hep-ph/0406277`] [InSPIRE].

[33] F. Ruehle, *Data science applications to string theory*, *Phys. Rept.* **839** (2020) 1 [InSPIRE].

[34] K. Pilch, P. van Nieuwenhuizen and P.K. Townsend, *Compactification of d = 11 Supergravity on S_4 (Or 11 = 7 + 4, Too)*, *Nucl. Phys. B* **242** (1984) 377 [InSPIRE].

[35] M. Pernici, K. Pilch, P. van Nieuwenhuizen and N.P. Warner, *Noncompact Gaugings and Critical Points of Maximal Supergravity in Seven-dimensions*, *Nucl. Phys. B* **249** (1985) 381 [InSPIRE].

[36] M. Günaydin, P. van Nieuwenhuizen and N.P. Warner, *General Construction of the Unitary Representations of Anti-de Sitter Superalgebras and the Spectrum of the S^4 Compactification of Eleven-dimensional Supergravity*, *Nucl. Phys. B* **255** (1985) 63 [InSPIRE].

[37] M. Günaydin and S. Takemae, *Unitary supermultiplets of $OSp(8^*/4)$ and the AdS_7/CFT_6 duality*, *Nucl. Phys. B* **578** (2000) 405 [*Erratum ibid.* **697** (2004) 399] [`hep-th/9910110`] [InSPIRE].

[38] P. Breitenlohner and D.Z. Freedman, *Positive Energy in anti-de Sitter Backgrounds and Gauged Extended Supergravity*, *Phys. Lett. B* **115** (1982) 197 [InSPIRE].

[39] P. Breitenlohner and D.Z. Freedman, *Stability in Gauged Extended Supergravity*, *Annals Phys.* **144** (1982) 249 [InSPIRE].

[40] L. Mezincescu and P.K. Townsend, *Stability at a Local Maximum in Higher Dimensional Anti-de Sitter Space and Applications to Supergravity*, *Annals Phys.* **160** (1985) 406 [InSPIRE].

[41] H. Ooguri and C. Vafa, *Non-supersymmetric AdS and the Swampland*, *Adv. Theor. Math. Phys.* **21** (2017) 1787 [`arXiv:1610.01533`] [InSPIRE].

[42] N. Hansen and A. Ostermeier, *Completely derandomized self-adaptation in evolution strategies*, *Evolutionary Computation* **9** (2001) 159.

[43] N. Hansen, S.D. Müller and P. Koumoutsakos, *Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es)*, *Evolutionary Computation* **11** (2003) 1.

[44] N. Hansen and A. Ostermeier, *Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation*, in *Proceedings of IEEE International Conference on Evolutionary Computation*, (1996), pp. 312–317.

[45] N. Hansen, *The cma evolution strategy: A tutorial*, arXiv:1604.00772.

[46] C. Igel, N. Hansen and S. Roth, *Covariance matrix adaptation for multi-objective optimization*, *Evolutionary computation* **15** (2007) 1.

[47] N. Hansen and S. Kern, *Evaluating the cma evolution strategy on multimodal test functions*, in *International conference on parallel problem solving from nature*, Springer (2004), pp. 282–291.

[48] R. Ros and N. Hansen, *A simple modification in cma-es achieving linear time and space complexity*, in *International conference on parallel problem solving from nature*, Springer (2008), pp. 296–305.

[49] N. Hansen, A. Auger, R. Ros, S. Finck and P. Posik, *Comparing Results of 31 Algorithms from the Black-Box Optimization Benchmarking BBOB-2009*, in *ACM-GECCO Genetic and Evolutionary Computation Conference*, Portland, United States, July 2010, pp. 1689–1696.