

Why the cosmological constant is so small? A string theory perspective

S.-H. HENRY TYE
(WITH YOSKE SUMITOMO AND SAM WONG)

With no free parameter (except the string scale M_S), dynamical flux compactification in Type IIB string theory determines both the cosmological constant (vacuum energy density) Λ and the Planck mass M_P in terms of M_S , thus yielding their relation. Following elementary probability theory, we find that a good fraction of the meta-stable de Sitter vacua in the cosmic string theory landscape tend to have an exponentially small cosmological constant Λ compared to either the string scale M_S or the Planck scale M_P , i.e., $\Lambda \ll M_S^4 \ll M_P^4$. Here we illustrate the basic stringy idea with a simple scalar field ϕ^3 (or ϕ^4) model coupled with fluxes to show how this may happen and how the usual radiative instability problem is bypassed (since there are no parameters to be fine-tuned). These low lying semi-classical de Sitter vacua tend to be accompanied by light scalar bosons/axions, so the Higgs boson mass hierarchy problem may be ameliorated as well.

1. Introduction

Cosmological data strongly indicates that our universe has a vanishingly small positive cosmological constant Λ (or vacuum energy density) as the dark energy,

$$(1.1) \quad \Lambda \sim 10^{-122} M_P^4$$

where the Planck mass $M_P = G_N^{-1/2} \simeq 10^{19}$ GeV. Such a small Λ is a major puzzle in physics.¹ In general relativity, Λ is a free arbitrary parameter one can introduce, so its smallness can be accommodated but not explained within quantum field theory. On the other hand, string theory has only a

¹If the dark energy is due to some other mechanism, e.g., quintessence, then the cosmological constant may have to be even smaller, or a more fine-tuned cancellation/correlation must be present.

single parameter, namely the string scale $M_S = 1/\sqrt{2\pi\alpha'}$, so everything else should be calculable for each string theory solution. String theory has 9 spatial dimensions, 6 of them must be dynamically compactified to describe our universe. Since both M_P and Λ are calculable, Λ can be determined in terms of M_P dynamically in each local minimum compactification solution. This offers the possibility that we may find an explanation for a very small positive Λ . This happens if a good fraction of the meta-stable deSitter (dS) vacua in the landscape tend to have a very small Λ , as is the case in the recent studies in flux compactification in string theory. Here, instead of reporting the main result of our work in the past few years [1–5], I like to introduce the basic stringy idea in a simple ϕ^3 or ϕ^4 quantum field theory model for illustration.

Although string theory has no free parameter (i.e., masses and couplings), it has fluxes and moduli in flux compactification from 10 spacetime dimensions to 4 (or any lower than 10) spacetime dimensions. In low energies, the moduli (plus the dilaton) are scalar fields describing the shape and size of the compactified manifold. Fluxes come from anti-symmetric P-form field strengths $F_{\mu_1\mu_2,\dots,\mu_P}$ where $\mu_0, \mu_1, \mu_2, \mu_3$ are the 4-dimensional spacetime indices and the rest stands for internal dimensions. (For $P < 4$, consider its dual.) so the effective 4-dimensional $F_i = F_i^{0123}$, $i = 1, 2, \dots, N$ takes constant values in 4-dimensional spacetime. They contribute to the vacuum energy density. Here i refers to the collection of internal indices.

Bousso and Polchinski show that each F_i takes only quantized values at the local minima in string theory [6]. For enough number of such fluxes (say $N > 14$), with each F_i sweeping through a range of discrete values, the spacing between allowed vacuum energy density values is small enough so a small Λ like that observed (1.1) can be one of the allowed values. They call this "dense discretuum". Our approach is to include the moduli coupled to the fluxes and find the vacuum energy density Λ at every local minimum. Sweeping over all choices of the flux values, we then find the distribution of the Λ values. Here, dynamics is brought in beyond a simple counting exercise, since not all choices of fluxes have a local minimum, while some flux choices yield multiple local minima.

Start with the four-dimensional low energy (supergravity) effective potential $V(F_i, \phi_j)$ obtained from flux compactification in string theory, where F_i are the 4-form field strengths and ϕ_j are the complex moduli (and dilaton) describing the size and shape of the compactified manifold as well as the couplings. In the search of classical minima, this flux quantization property allows us to rewrite $V(F_i, \phi_j)$ as a function of the quantized values $q_i n_i$

of the fluxes present,

$$V(F_i, \phi_j) \rightarrow V(q_i n_i, \phi_j), \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, K$$

where the charges q_i are determined by the compactification dynamics. (To simplify the discussion, we shall suppress the q_i so n_i takes discrete instead of integer values.) Since string theory has no continuous free parameter, there is no arbitrary free parameter in $V(n_i, \phi_j)$, though it does contain (in principle) calculable quantities like α' corrections, loop and non-perturbative corrections, and geometric quantities like Euler index χ etc..

For a given set of discrete flux parameters $\{n_i\}$, we can solve $V(n_i, \phi_j)$ for its meta-stable (classically stable) vacuum solutions via finding the values $\phi_{j,\min}(n_i)$ at each solution and determine its vacuum energy density $\Lambda = \Lambda(n_i, \phi_{j,\min}(n_i)) = \Lambda(n_i)$. Since we are considering the physical ϕ_j , it is the physical Λ we are determining. Since a typical flux parameter n_i can take a large range of discrete values, we may simply treat each n_i as an independent random variable with some distribution $P_i(n_i)$. Collecting all such solutions, we can next find the probability distribution $P(\Lambda)$ of Λ of these meta-stable solutions as we sweep through all the discrete flux values n_i . That is putting $P_i(n_i)$ and $\Lambda(n_i)$ together yields $P(\Lambda)$,

$$(1.2) \quad P(\Lambda) = \sum_{n_i} \delta(\Lambda - \Lambda(n_i)) \Pi_i P_i(n_i)$$

so $\sum_{n_i} P_i(n_i) = 1$ for each i implies that $\int P(\Lambda) d\Lambda = 1$. For large enough ranges for n_i , we may treat each $P_i(n_i)$ as a smooth continuous function over an appropriate range of values.

Simple probability properties show that $P(\Lambda)$ easily peaks and diverges at $\Lambda = 0$ [1], implying that a small Λ is statistically preferred. For an exponentially small Λ , the statistical preference for $\Lambda \simeq 0$ has to be overwhelmingly strong, that is, $P(\Lambda)$ has to diverge (i.e., peak) sharply at $\Lambda = 0$. Such an analysis has been applied to the Kähler uplift scenario, where $P(\Lambda)$ is so peaked at $\Lambda = 0$ that the the median Λ matches the observed Λ (1.1) if the number of complex structure moduli $h^{2,1} \sim \mathcal{O}(100)$ [2]. Such a value for $h^{2,1}$ is quite reasonable for a typical manifold considered in string theory. That is, an overwhelmingly large number of meta-stable vacua have an exponentially small Λ , so statistically, we should end up in one of them. In other words, a very small Λ is quite natural. The preference for an exponentially small Λ has also been observed in the racetrack Kähler uplift scenario [4].

In usual quantum field theory, the parameters of any model (masses, couplings and Λ) that include the standard model of strong and electroweak

interactions have to be fine-tuned to satisfy Eq.(1.1). However, quantum corrections are typically orders-of-magnitude bigger than the observed value. So the renormalized parameters have to be re-fine-tuned after each order of radiative correction. This is the radiative instability problem. Since there are no free parameters to be fine-tuned in string theory, the radiative instability problem is simply absent here.

This leads us to conjecture that

Most meta-stable vacua in regions of the cosmic stringy landscape have $|\Lambda| \ll M_P^4$.

If true, this may (1) provide an explanation why the observed Λ is so small, and (2) after inflation, why the universe is not trapped in a relatively high Λ vacuum.

Instead of reviewing the string theory models, let us discuss in some detail an illustrative ϕ^3/ϕ^4 model, where there is no uncoupled sector and all couplings/parameters are treated as if they are flux parameters so they will take random values within some reasonable ranges. There, the physical (loop corrected) $V(n_i, \phi_j)$ yields $P(\Lambda_{\text{ph}})$ for the physical Λ_{ph} while the tree (or bare) $V(n_i, \phi_j)$ yields $P(\Lambda_0)$ for the tree Λ_0 . We find that $P(\Lambda_{\text{ph}})$ hardly differs from $P(\Lambda_0)$. Both $P(\Lambda)$ s peak (i.e., diverge) at $\Lambda = 0$, and the two sets of statistical preferred flux values for $\Lambda \sim 0$ are in general only slightly different. In fact, up to two-loops, $P(\Lambda_{\text{ph}})$ is essentially identical to the tree $P(\Lambda_0)$. As a result, although radiative instability may be present for any fixed flux choice, the statistical preference approach actually evades or bypasses this radiative instability problem. We like to convince the readers that this phenomenon of bypassing the radiative instability problem stays true in more complicated models, as well as when applied to very light scalar boson masses (if present).

There are a vast number of very small Λ dS vacua in the cosmic string landscape. Lest one may think the accumulation of $\Lambda \simeq 0^+$ is due to energetics (i.e., small positive Λ s are energetically preferred over not so small positive Λ s), we note that the same accumulation happens for AdS vacua as well; that is, $P(\Lambda)$ peaks (diverges) as $\Lambda \rightarrow 0^-$ [3]. Our universe rolling down the landscape after inflation is unlikely to be trapped by a relatively high dS vacuum, since there is hardly any around. However, since it has to pass through the positive Λ region first, it is likely to be trapped at a small positive Λ vacuum (as there are many of them) before reaching the sea of AdS vacua with small negative Λ .

2. An illustrative ϕ^3/ϕ^4 toy model

The statistical preference for a small Λ follows if the low energy effective potential has no continuous free parameter and all sectors are connected via interactions, as is the case in string theory; that is, it is a function of only scalar fields or moduli, quantized flux values, discrete values like topological indices, and calculable quantities like loop and string corrections, with no disconnected sectors. To get some feeling on some of these features, let us review the single scalar field polynomial model discussed. In this model, gravity and so M_P is absent. So the statistical preference for a small Λ shows up only as the (properly normalized) probability distribution $P(\Lambda)$ peaks at $\Lambda = 0$, in particular when $P(\Lambda)$ diverges there, i.e.,

$$(2.1) \quad \lim_{\Lambda \rightarrow 0^+} P(\Lambda) \rightarrow \infty$$

The divergence of $P(\Lambda = 0)$ in this toy model is rather mild here, so it is far from enough to explain the very small observed value of Λ (1.1); but it does allow us to explain a few properties that are relevant for later discussions. Consider the tree level potential,

$$(2.2) \quad V_0(\phi) = a_1\phi + \frac{a_2}{2}\phi^2 + \frac{a_3}{3!}\phi^3 + \frac{a_4}{4!}\phi^4$$

where ϕ is a real field, mimicking a modulus.

Imposing the constraint that the tree level V_0 has no continuous free parameter except some scale M_s , the parameters a_1 , a_2 , a_3 and a_4 mimic the flux parameters that take only discrete values of order of the M_s scale, thus spanning a “mini-landscape”. Let them take only real values for simplicity. We may also choose units so $M_s = 1$. For a dense enough discretuum for each flux parameter, a flux parameter may be treated as a random variable with continuous value over some range. Let us look for dS solutions with flux parameters $a_1, a_2, a_3, a_4 \in [-1, 1]$ or some other reasonable ranges. We start with the tree-level properties and then discuss the multi-loop corrections.

Starting with the tree-level effective potential $V_0(\phi)$ (2.2), we impose the stability $M^2 = \partial_\phi^2 V_0|_{v_0} > 0$ at the extremal points given by $\partial_\phi V_0|_{v_0} = 0$, with each vacuum expectation value v_0 yielding $\Lambda_0(v_0) = V_0(v_0)$ and

$$(2.3) \quad M_0^2 = \left. \frac{\partial^2 V_0}{\partial \phi^2} \right|_{v_0} = +a_2 + a_3 v_0 + a_4 v_0^2 / 2 > 0$$

where $\lambda = \frac{\partial M^2}{\partial v} = a_3 + a_4 v_0$. Not all choices of fluxes yield a local minimum, while some choices yield more than one minimum. We study three case: the

ϕ^3 model with $a_3 = 1$ and with random a_3 , and the ϕ^4 case with random flux parameters $\{a_i\}$. Using Eq.(1.2), we find $P(\Lambda)$. The ϕ^4 case is shown in Figure 1. $P(\Lambda)$ clearly peaks (i.e, diverges logarithmically) at $\Lambda = 0$. (Similar divergence of $P(\Lambda)$ emerges for the ϕ^3 case.)

A few comments are in order here:

- $V_0(\phi)$ (2.2) has no free parameter, as demanded by string theory. Here we have chosen a flat distribution for each flux parameter. We do expect the distributions $P_i(a_i)$ to be smooth and include $a_i = 0$, since zero flux value must be permitted. Actually, the peaking of $P(\Lambda)$ is independent of the precise distributions of $P_i(a_i)$, as long as they are relatively smooth.

- We are not allowed to introduce a “constant” or an independent flux parameter a_0 term by itself in $V_0(\phi)$ (2.2), since it will be disconnected to the ϕ terms in $V_0(\phi)$. In string theory, everything is coupled to everything else, directly or indirectly, at least via the closed string sector, which includes gravity. For example, the following possibility is dis-allowed (i.e., in the swampland),

$$V(\phi) = V_1(\phi_1) + V_2(\phi_2)$$

if V_1 and V_2 are totally uncoupled. Here we can minimize $V_2(\phi_2)$ (for fixed flux values in $V_2(\phi_2)$) to obtain Λ_2 so

$$V(\phi) = V(\phi_1) + \Lambda_2$$

where Λ_2 is simply a constant term in $V(\phi)$ un-connected to ϕ_1 or the flux parameters in $V_1(\phi_1)$. In this case (dis-allowed by string theory), even if both $P_1(\Lambda_1)$ and $P_2(\Lambda_2)$ peak at their respective zeros, $P(\Lambda)$ does not peak at $\Lambda = \Lambda_1 + \Lambda_2 = 0$.

- It is easy to show that $P(\Lambda)$ similarly peaks at $\Lambda = 0$ if we include higher terms (e.g., up to a $a_6\phi^6/6!$ term) in $V_0(\phi)$ (2.2).

- For vacuum energy, we may choose to introduce $F_i^2/2$ terms to $V_0(\phi)$ (2.2); for example,

$$E = V(\phi) + \frac{a_1^2}{2}$$

in the ϕ^3 case. We see that $P(\Lambda)$ again peaks at $\Lambda = 0$. Notice that such flux terms are coupled to ϕ . In actual string theory models, $F_i = q_i n_i$ where, although n_i is an integer, the ”charge q_i ” is determined by dynamics, probably some function of the moduli. So a pure F_i^2 implies that some (presumably heavier) moduli have already been stabilized to yield the charge q_i . Adding more flux terms into E may destroy the peaking behavior.

- Here we find the accumulation of vacua with $\Lambda \simeq 0^+$. We note that $P(\Lambda)$ also peaks (diverges) as $\Lambda \rightarrow 0^-$ in this toy model. In fact, typical AdS

solutions of $V(n_i, \phi_j)$ involve 2 branches: supersymmetric vacua and non-supersymmetric vacua, where the latter set mirrors the dS solutions (see e.g., [3]). So for a given range of small $|\Lambda|$, we expect more AdS vacua than dS vacua; that happens even before we relax the constraint to allow light tachyons which do not destabilize the AdS vacua.

• Let us discuss the multi-loop contributions to the effective potential $V(\phi)$. We see that the n -th loop contribution to $V(\phi)$, namely $V_n(\phi)$, is a function of $M^2(\phi)$, $\lambda(\phi)$ (2.3), and a_4 only. At the one-loop level, $V_1(\phi)$ is a function of $M^2(\phi)$ only. Simple dimensional reasoning yields

$$(2.4) \quad V(\phi) = V_0 + \sum_{n \geq 1} V_n = M^4 F \left(\frac{\lambda^2}{M^2}, \ln(M^2), a_4 \right),$$

$$M^2(\phi) = V_0''(\phi), \quad \lambda(\phi) = V_0'''(\phi), \quad a_4(\phi) = V_0''''(\phi),$$

where each prime stands for a derivative with respect to ϕ and $F(\lambda^2/M^2, \ln(M^2), a_4)$ is a polynomial in the dimensionless parameters λ^2/M^2 , $\ln(M^2)$ and a_4 . More precisely, for $n \geq 1$,

$$V_n = \frac{M^4}{(4\pi)^{2n}} f_n(\lambda^2/M^2, \ln(M^2), a_4)$$

where f_n is a polynomial up to n -th power in $\ln(M^2)$, and $(n-1)$ -th (combined) power in λ^2/M^2 and a_4 , with n -dependent coefficients which grow much slower than the $(4\pi)^{2n}$ factor.

The key of a naturally small Λ_{ph} depends on its functional dependence on the flux values, which is different from that for Λ_0 . Here we consider explicitly the one- and two-loop contributions to Λ and find the $P(\Lambda_{\text{ph}})$ for the one- and two-loop corrected cases, namely $P(\Lambda_1)$ and $P(\Lambda_2)$. At least up to two-loops, $P(\Lambda_{\text{ph}})$ continues to peak (diverge) at $\Lambda_{\text{ph}} = 0$. In fact, the loop corrected $P(\Lambda)$ are essentially indistinguishable from the tree $P(\Lambda_0)$, as shown in FIG 1. This is despite the fact that, for a specific choice of $\{a_i\}$ that yields $\Lambda_0 = 0$, $\Lambda_1 \neq 0$ and $\Lambda_2 \neq 0$ in general.

To summarize, the statistical preference for $\Lambda = 0$ is robust, for either the tree-level Λ_0 or the loop-corrected Λ_{ph} . Although the functional dependence of Λ on the flux parameters are different for Λ_0 and Λ_{ph} , nevertheless, given the same probability distributions for the flux parameters, we see that $P(\Lambda_{\text{ph}})$ is essentially the same as $P(\Lambda_0)$. It will be nice to investigate the above properties for more general quantum field theory models that satisfy the stringy conditions: no free parameters except flux parameters and no uncoupled sectors. Of course, the cases we are really interested in are the

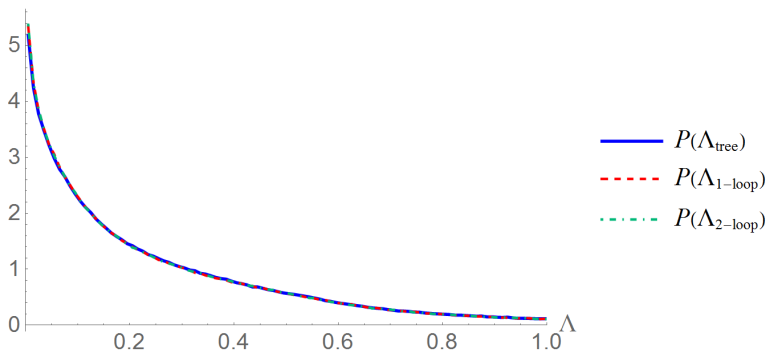


Figure 1: Probability distribution $P(\Lambda)$ for the ϕ^4 model (similar behaviors for the ϕ^3 model) for the tree-level and the one- and two-loop corrected cases. The blue solid curve is for the tree-level $P(\Lambda_0)$, the red dashed curve is for the one-loop corrected $P(\Lambda_1)$ and the green dot-dash curve is for the two-loop corrected $P(\Lambda_2)$. We see that the loop-corrected and the tree $P(\Lambda)$ s are essentially on top of each other, showing that loop corrections have little impact on the distribution $P(\Lambda)$. In particular, the peaking behavior of $P(\Lambda)$ at $\Lambda = 0$ remains intact.

flux compactifications in string theory. However, we do gain some intuitive understanding from examining this relatively simple model.

Actually we are interested only in the preferred values of the physical Λ . However, including quantum effects fully is in general a very challenging problem in any theory. Fortunately, if one can argue that the peaking behavior of $P(\Lambda)$ is hardly modified by quantum corrections, as this model suggests, a simpler tree-level result provides valuable information on the statistical preference of a small physical Λ . For ground states in string theory, an effective potential description may be sufficient to capture the physics of the value of Λ in some region of the landscape. We may hope that stringy corrections will not qualitatively disrupt the statistical preference approach adopted here.

2.1. Scalar masses in the toy model

We also like to propose that

A dS vacuum with a naturally small Λ tends to be accompanied by very light scalar bosons.

Consider the 4-dimensional effective action

$$(2.5) \quad S = \int dx^4 \sqrt{-g} \left[-\Lambda + \frac{M_P^2}{16\pi} R - \frac{m_H^2}{2} \Phi_H^2 + \dots \right]$$

where we have displayed all the relevant operators that are known to be present in nature. If we ignore the Λ (the most relevant operator) term, then we have two scales, $M_P \gg m_H$. Why the Higgs mass m_H is so much smaller than the Planck mass M_P poses the well-known mass hierarchy problem. Now knowing that a very small Λ is present in nature, we like to know its origin. If its value arises via fine-tuning or by pure accident, we have to consider M_P as more fundamental and so are led back to the original mass hierarchy problem. However, if the smallness of Λ arises naturally, in that most of the de Sitter vacua in string theory tend to have a very small Λ , we should expect some scalar masses comparable to the Λ scale, as is the case in the string models examined. Following this viewpoint, we may instead wonder why the Higgs mass is so much bigger than Λ , i.e., $m_H^2 \gg \Lambda/M_P^2$ (an inverted mass hierarchy problem). Surely, we should re-examine the mass hierarchy problem in this new light.

Along this direction, we show that the following scenario can easily happen: the physical mass-squared probability distribution $P_j(m_j^2)$ for some scalar field ϕ_j may be peaked at $m_j^2 = 0$ but the peaking is less strong than that for Λ . If the Higgs boson is such a particle, i.e., $\Phi_H = \phi_j$, then it is natural for

$$(2.6) \quad \Lambda/M_P^2 \ll m_H^2 \ll M_S^2 \ll M_P^2$$

This statistical preference approach allows us to circumvent the original mass hierarchy problem; that is, a small Higgs mass is natural, not just technically natural.

It is straightforward to find that $P(m^2)$ does not peak at $m^2 = 0$ in the above ϕ^3/ϕ^4 model, in every case considered, loop corrected or not. In more realistic models, $P(\Lambda_{\text{ph}})$ has to diverge at $\Lambda_{\text{ph}} = 0$ much more sharply than the logarithmical divergence shown in this model. In the non-trivial models in string theory studied so far, we see that both $P(\Lambda_{\text{ph}})$ for Λ_{ph} and $P(m_{\text{ph}}^2)$ for the some bosons prefer small values, while the peaking in $P(\Lambda_{\text{ph}})$ is much stronger than that in $P(m_{\text{ph}}^2)$. If one applies this to the Higgs boson in a phenomenological model, the observed situation (2.6) can follow from their statistical preferences.

The way of bypassing the radiative instability problem should also apply to the masses as well when the probability distribution $P(m^2)$ for some

scalar mass also peaks at $m^2 = 0$. Furthermore, one may convince oneself that this statistical preference for a small Λ also bypasses the disruptions caused by phase transitions during the evolution of the early universe, as the universe rolls down the landscape in search of a meta-stable minimum. (See Discussion below.)

3. A Kähler uplift model of flux compactification

Here we summarize the results of a flux compactification model where the AdS vacua are Kähler uplifted to dS vacua via the presence of an α^3 correction plus a non-perturbative term. Using reasonable probability distributions for the flux values, it has been shown that the probability distribution $P(\Lambda)$ peaks sharply at $\Lambda = 0$, resulting in a median Λ comparable to the observed value if the number of complex structure moduli $h^{2,1} \sim \mathcal{O}(100)$.

To be specific, consider a Calabi-Yau-like three-fold M with a single ($h^{1,1} = 1$) Kähler modulus and a relatively large $h^{2,1}$ number of complex structure moduli, so the manifold M has Euler number $\chi(M) = 2(h^{1,1} - h^{2,1}) < 0$. The simplified model of interest was first studied in Ref[7], where earlier references can be found. Setting $M_P = 1$,

$$\begin{aligned}
 (3.1) \quad V &= e^K \left(K^{I\bar{J}} D_I W D_{\bar{J}} \bar{W} - 3 |W|^2 \right), \\
 K &= K_K + K_d + K_{cs} \\
 &= -2 \ln \left(\mathcal{V} + \frac{\hat{\xi}}{2} \right) - \ln (S + \bar{S}) - \sum_{i=1}^{h^{2,1}} \ln (U_i + \bar{U}_i), \\
 \mathcal{V} &\equiv \frac{\text{vol}}{\alpha'^3} = (T + \bar{T})^{3/2}, \quad \hat{\xi} = -\frac{\zeta(3)}{4\sqrt{2}(2\pi)^3} \chi(M) (S + \bar{S})^{3/2} > 0, \\
 W &= W_0(U_i, S) + A e^{-aT}, \\
 W_0(U_i, S) &= c_1 + \sum_{i=1}^{h^{2,1}} b_i U_i - S \left(c_2 + \sum_{i=1}^{h^{2,1}} d_i U_i \right) + \sum_{i,j}^{h^{2,1}} \alpha_{ij} U_i U_j,
 \end{aligned}$$

The original model has $h^{2,1} = 3$ complex structure moduli. We simply take this form and straight-forwardly generalize it to an arbitrary number of complex structure moduli. In known models for $h^{2,1} > 3$, the potential is actually somewhat different. However, this straightforward generalization, though naive, allows us to solve the model semi-analytically, which is crucial to obtain any numerical properties when $h^{2,1}$ is large [4]. In this sense, our model is at best semi-realistic.

The flux contribution to $W_0(U_i, S)$ depends on the dilation S and the $h^{2,1}$ complex structure moduli U_i ($i = 1, 2, \dots, h^{2,1}$), while the non-perturbative term for the Kähler modulus T is introduced in the superpotential W . The dependence of A on U_i, S are suppressed. The model also includes the α' -correction (the $\hat{\xi}$ term) to the Kähler potential where c_i, b_i, d_i and the non-geometric $\alpha_{ij} = \alpha_{ji}$ are (real) flux parameters that may be treated as independent random variables with smooth probability distributions that allow the zero values. Here we are interested in the physical Λ (instead of, say, the bare Λ), so the model should include all appropriate non-perturbative effects, α' corrections as well as radiative corrections. We see that the above simplified model (3.1) includes a non-perturbative A term to stabilize the Kähler modulus and the α' correction $\hat{\xi}$ term to lift the solution to de-Sitter space. In the same spirit, all parameters in the model, in particular the coupling parameters c_i, b_i, d_i and α_{ij} in W_0 (3.1), should be treated as physical parameters that have included all relevant corrections.

Now we sweep through the flux values c_i, b_i and d_i treating them as independent random variables to find the probability distribution $P(\Lambda)$. The ranges of flux values are constrained by our weak coupling approximation (i.e., $s > 1$) et. al.. For any reasonable probability distributions $P_i(c_i), P_i(b_i)$ and $P_i(d_i)$, we find that $P(\Lambda)$ peaks (and diverges) at $\Lambda = 0$. To quantify this peaking behavior, it is convenient to summarize the result by looking at $\Lambda_{Y\%}$. That is, there is $Y\%$ probability that $\Lambda_{Y\%} \geq \Lambda \geq 0$. So $\Lambda_{50\%}$ is simply the median. In Ref.[2], we find that (with $\alpha_{ij} = 0$), as a function of the number $h^{2,1}$ of complex structure moduli, for $h^{2,1} > 5$ and $\Lambda \geq 0$,

$$\begin{aligned}
 \Lambda_{50\%} &\simeq 10^{-h^{2,1}-2} M_P^4 \\
 \Lambda_{10\%} &\simeq 10^{-1.3h^{2,1}-3} M_P^4 \\
 \langle \Lambda \rangle &\simeq 10^{-0.03h^{2,1}-6} M_P^4
 \end{aligned}
 \tag{3.2}$$

where we have also given $\Lambda_{10\%}$. We see that the average $\langle \Lambda \rangle$ does not drop much, since a few relatively large Λ s dominate the average value. A typical flux compactification can have dozens or even hundreds of $h^{2,1}$, so we see that a Λ as small as that observed in nature can be dynamically preferred. Scalar masses are of order [5],

$$m^2 M_P^2 / \Lambda \sim \mathcal{O}(1)$$

while axions are massless before non-perturbative interactions are included. Presumably there are heavier scalar bosons that have been integrated out in this model. Although their masses are by definition much heavier than

the moduli masses just discussed, some of them can easily have masses that are much smaller than the string scale M_S .

4. Modulus masses in Racetrack Kähler uplift

The Kähler uplift model studied in the last section has a single non-perturbative term in the superpotential W . To relax the constraint on the volume size, we generalize the model to include two non-perturbative terms in W , i.e., the *racetrack* model. (See Ref.[4] for details.) Unlike the Kähler uplift model studied previously, the α' -correction is more controllable for the meta-stable de-Sitter vacua in the racetrack case since the constraint on the compactified volume size is very much relaxed. So the model admits solutions with a large adjustable volume.

Interestingly, in this Racetrack Kähler uplift model, the stability condition for both the real and imaginary sectors requires that the minima of the potential V always exist for $\Lambda \geq 0$ at large volumes. Further, the cosmological constant Λ is naturally exponentially suppressed as a function of the volume size, and the resultant probability distribution $P(\Lambda)$ for Λ gets a sharply peaked behavior toward $\Lambda \rightarrow 0$, which can be highly divergent. This peaked behavior of $P(\Lambda)$ can be much sharper than that of the previous Kähler Uplift model with a single non-perturbative term studied in [2]. So an exponentially small median for Λ is natural.

The racetrack Kähler uplift model is similar to the above Kähler Uplift model, but with one major addition. The super-potential W now has two non-perturbative terms for the Kähler modulus $T = t + i\tau$ instead of one,

$$(4.1) \quad W = W_0(U_i, S) + W_{\text{NP}} = W_0(U_i, S) + Ae^{-aT} + Be^{-bT}$$

where the coefficients $a = 2\pi/N_1$ for $SU(N_1)$ gauge symmetry and $b = 2\pi/N_2$ for $SU(N_2)$ gauge symmetry. In the large volume region and in units where $M_P = 1$, the resulting potential may be approximated to

$$(4.2) \quad V \simeq \left(-\frac{a^3 A W_0}{2} \right) \lambda(x, y),$$

$$\lambda(x, y) = -\frac{e^{-x}}{x^2} \cos y - \frac{\beta e^{-\beta x}}{z x^2} \cos(\beta y) + \frac{\hat{C}}{x^{9/2}},$$

$$x = at, \quad y = a\tau, \quad z = A/B,$$

$$\beta = b/a = N_1/N_2 > 1, \quad \hat{C} = -\frac{3a^{3/2}W_0 \xi}{32\sqrt{2}A},$$

The extremal conditions $\partial_t V = \partial_\tau V = 0$ may be solved. The typical values of a , β and x are $\mathcal{O}(2\pi/16)$, $\mathcal{O}(1)$ and $\mathcal{O}(100)$ respectively, and the e^{-x} factor suggests very small Λ as well as moduli masses. After randomizing W_0 , A and B , we collect the solutions and find that the probability distribution $P(\Lambda)$ for small positive Λ is approximately given by [4],

$$(4.3) \quad P(\Lambda) \stackrel{\Lambda \rightarrow 0}{\sim} \frac{243\beta^{1/2}}{16(\beta - 1)} \frac{1}{\Lambda^{\frac{\beta+1}{2\beta}} (-\ln \Lambda)^{5/2}}.$$

So for $\beta \gtrsim 1$, we see that the diverging behavior of $P(\Lambda)$ is very peaked as $\Lambda \rightarrow 0$. Since $(\beta + 1)/2\beta < 1$, $P(\Lambda)$ is normalizable, i.e., $\int P(\Lambda)d\Lambda = 1$. For illustration, we have

$$(4.4) \quad \beta = 1.10 : \quad \Lambda_{50} = 7.08 \times 10^{-10}, \quad \Lambda_{10} = 3.61 \times 10^{-24}$$

$$(4.5) \quad \beta = 1.04 : \quad \Lambda_{50} = 5.47 \times 10^{-19}, \quad \Lambda_{10} = 2.83 \times 10^{-54}$$

where $\beta = N_1/N_2 = 26/25 = 1.04$. Choosing a gauge group larger than $N_2 = 25$, $\beta = N_1/N_2 = (N_2 + 1)/N_2$ is closer to unity and the median Λ will take values much closer to the the observed value (1.1). We also see that both t and τ masses are exponentially suppressed. By using the small value of Λ , we can obtain bounds on both masses [5],

$$(4.6) \quad \begin{aligned} \frac{m_t^2}{\Lambda} &= \frac{\partial_t^2 V}{2K_{T\bar{T}}\Lambda} \leq \frac{9\beta x + 30(\beta + 1)}{a^4(2\beta x - 5(\beta + 1))}, \\ \frac{m_\tau^2}{\Lambda} &= \frac{\partial_\tau^2 V}{2K_{T\bar{T}}\Lambda} \leq \frac{6x(3\beta x + 10(\beta + 1))}{a^4(4\beta x^2 - 10(\beta + 1)x + 35)}. \end{aligned}$$

Solving for $x \sim \mathcal{O}(100)$, we see that the Kähler modulus masses are exponentially small unless one fine-tunes one of the denominating factor to a very small value. Presumably there are heavier scalar bosons that have been integrated out in this simple model. Although their masses are much heavier than t and τ , some of them can easily have masses that are much smaller than the string or Planck scale.

5. Discussions

So far, we have a few looks at the global picture of some corners of the string landscape. As illustrated by the Kähler uplift models and the racetrack model discussed, we see hints that, of the meta-stable solutions, most of them have very small Λ , while each such vacuum has very light bosons. Here we like to discuss a few issues related to this property.

Since $P(\Lambda)$ is properly normalized, even with a divergent $P(\Lambda = 0) = \infty$, the probability of having a universe with $\Lambda \equiv 0$ is zero. Furthermore, the flux values are discrete, so the probability distribution $P(a)$ of flux a is non-zero only at discrete values of a . So Λ can be exponentially small, but not exactly zero.

The string theory models studied in this paper are admittedly relatively simple. Nonetheless, they incorporate known stringy properties in a consistent fashion so they are non-trivial enough for us to learn about the structure and dynamics of flux compactification in string theory. They clearly illustrate that a statistical preference for a very small physical Λ in the cosmic landscape as a solution to the cosmological constant problem is a distinct possibility. This way to solve the cosmological constant problem bypasses the radiative instability problem. Associated with the very small Λ are very light moduli masses. So this offers the possibility of having light bosons via statistical preference as well. It is important to point out that this solution or explanation is possible because of the existence of the landscape. Comparing to the earlier works [6] where explicit interactions among the moduli and fluxes are not taken into account, we see that the statistical preference for a small Λ (and at times some scalar masses) emerges only when couplings are included. Intuitively, in examining the models studied (albeit a rather limited sample), more fluxes and moduli with more couplings among them tend to enhance or at least maintain the divergence of $P(\Lambda)$ at $\Lambda = 0$. This is encouraging, since higher order corrections and more realistic (and so more complicated) models are very challenging to study.

What happens when finite temperature and phase transition appear? Suppose the Universe starts out at a random point somewhere high up in the landscape, at zero temperature (for zero temperature, we mean zero thermal temperature, not the Gibbons-Hawking temperature $H/2\pi = \sqrt{2V/3\pi}/M_P$, which is assumed to be negligible here). It rolls down and ends up in a local minimum. Because it starts from a random point, this minimum may be considered to be randomly chosen. If most of the vacua have a small Λ , it is likely that this minimum is one of these small Λ vacua. What happens if we turn on a finite temperature T ? We have essentially the same landscape, but is starting from a different point up in the landscape, so the evolution of the Universe will be different and possibly ending at a different local minimum, also randomly chosen. As temperature $T \rightarrow 0$, we find that the chosen local vacuum at T probably turns out to have a small Λ at $T = 0$, because most vacua at $T = 0$ have a small Λ . If the chosen local vacuum has a critical temperature $T_c < T$, phase transition happens as T drops below T_c . If this is a second order phase transition, then the Universe will roll away to another

local minimum, which is likely to have a small Λ as $T \rightarrow 0$, because most vacua at $T = 0$ have a small Λ . If it is a first order phase transition, the Universe will stay at this vacuum as $T \rightarrow 0$ (before tunneling). This vacuum should have a small Λ , because most vacua at $T = 0$ have a small Λ . In all cases, we see that the Universe most likely end up in a vacuum with a small Λ . It is possible that this same vacuum has a relatively large Λ at finite T .

In terms of cosmology, one may wonder why the dark energy is so large, contributing to about 70% of the content of our universe. However, from the fundamental physics point of view, the puzzle is why it is so small, when we assume that the scale of gravity is dictated by the Planck scale M_P which is so much bigger. In our new picture, once we are willing to accept that the smallness of Λ has a fundamental explanation like the statistical preference employed here, the question is again reversed. For example, in the viewpoint adopted here, we see that typical moduli mass scales are guided by Λ , not M_P . That is, some of the bosonic/axionic masses are small enough to play the role of fuzzy dark matter.

Once we accept that both Λ and M_P have their respective places in the theory (that is, generated by string theory dynamics, with string scale M_S , not via fine-tuning), the presence of some intermediate mass scales such as the Higgs boson mass should not be so surprising. We see that the probability distribution $P(m^2)$ of bosonic mass m^2 does not peak at $m^2 = 0$ in the ϕ^3/ϕ^4 model. In the string theory models, one envisions scenarios where some bosonic masses have a statistical preference for small values, but such preference is not as strong as that for Λ . So the Higgs mass $m_H = 125$ GeV may fit in in such a scenario, thus evading the usual mass hierarchy problem for the Higgs boson. The scenario also offers the possibility that very light bosons can be present as the dark matter in our universe. In fact, any small number (e.g., the θ angle, light quark or neutrino masses in the standard electroweak model) in nature may be due to some level of a statistical preference without fine-tuning. Of course, we expect that heavier scalar bosons are generically present in a realistic string model, but they may have been integrated out in the low energy effective theory studied in the literature. One can imagine that some of these "heavy" scalar bosons have masses order-of-magnitude smaller than the string or the Planck scales.

The string theory models considered so far are necessarily relatively simple, to allow semi-analytic studies. It will be important to consider more realistic versions (for example, the form of the Kähler potential and couplings among moduli) to see if such statistical preference for small Λ and small bosonic masses are robust. In the search for the standard model within string theory, it may be fruitful to narrow the search of the three family standard

model only in the region of the landscape where order of magnitude mass scales as well as Λ come out in the correct range.

References

- [1] Y. Sumitomo and S.-H. H. Tye, *A stringy mechanism for a small cosmological constant*, JCAP **1208** (2012) 032. [arXiv:1204.5177 \[hep-th\]](#).
- [2] Y. Sumitomo and S. H. H. Tye, *A stringy mechanism for a small cosmological constant – multi-moduli cases –*, JCAP **1302** (2013) 006. [arXiv:1209.5086 \[hep-th\]](#).
- [3] Y. Sumitomo and S.-H. H. Tye, *Preference for a vanishingly small cosmological constant in supersymmetric vacua in a type IIB string theory model*, Phys. Lett. B **723** (2013) 406. [arXiv:1211.6858 \[hep-th\]](#).
- [4] Y. Sumitomo, S. H. H. Tye, and S. S. C. Wong, *Statistical distribution of the vacuum energy density in Racetrack Kähler uplift models in string theory*, JHEP **1307** (2013) 052. [arXiv:1305.0753 \[hep-th\]](#).
- [5] S.-H. H. Tye and S. S. C. Wong, *Linking light scalar modes with a small positive cosmological constant in string theory*, JHEP **1706** (2017) 094. [arXiv:1611.05786 \[hep-th\]](#).
- [6] R. Bousso and J. Polchinski, *Quantization of four form fluxes and dynamical neutralization of the cosmological constant*, JHEP **0006** (2000) 006. [arXiv:hep-th/0004134](#).
- [7] M. Rummel and A. Westphal, *A sufficient condition for de Sitter vacua in type IIB string theory*, JHEP **1201** (2012) 020. [arXiv:1107.2115 \[hep-th\]](#).

JOCKEY CLUB INSTITUTE FOR ADVANCED STUDY AND DEPARTMENT OF PHYSICS
HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY, HONG KONG
AND LABORATORY FOR ELEMENTARY-PARTICLE PHYSICS
CORNELL UNIVERSITY, ITHACA, NY 14853, USA
E-mail address: iastye@ust.hk