Introduction:

There is strong observational evidence for inflation. However, understanding its microscopic origin requires knowledge of very high-scale physics. Instead, most models of inflation are ad hoc effective field theories, typically involving scalars with fine-tuned potentials. Attempts to realize slow-roll inflation in string theory are plagued by difficulties that require a complex set of ingredients to circumvent. On the other hand, string theory naturally predicts the existence of a very large number of eternally inflating meta-stable phases (the “landscape” [1,2]) similar to those considered by Guth in his original “old” model of inflation [3]. Such phases also occur in many theories besides string theory, and in all cases serve as powerful attractors that dominate the system for essentially arbitrary choices of initial conditions. These phases decay via first-order phase transitions, but the bubbles produced by this process usually contain Friedmann-Robertson-Walker (FRW) cosmologies dominated by negative spatial curvature that do not reheat to a radiation dominated phase, and so old inflation was abandoned in favor of slowly rolling models [4].

One source for the vacuum energy driving false vacuum eternal inflation is higher form gauge flux (present in all string and supergravity theories). Here, we present a novel mechanism by which an eternally inflating false vacuum supported by flux can decay to a bubble containing a negatively curved FRW cosmology dominated by vacuum energy that subsequently steadily decreases over 60 or more Hubble times—much as in standard slow-roll inflation—before reheating via brane-anti-brane annihilation. This post-bubble nucleation expansion inflates away the initially large negative curvature, and at the same time produces a scale-invariant spectrum of curvature perturbations that is consistent with observational data. This occurs without the introduction of ad-hoc scalar potentials (indeed, without fundamental scalars at all), does not require any significant fine-tuning, and uses ingredients that are present in all string compactifications: higher-form flux, and at least one compact extra dimension. As such, this mechanism is both a natural candidate for the fundamental origin of slow roll inflation, and serves as a prototype for a new class of inflationary models.

Relation to previous work: The idea of using a bubble collision in a compact dimension to reheat homogeneously and isotropically was first proposed by A. Brown in [7], and the possibility of a scalar cascade in such a model was mentioned in [8]. Discharge of higher-form flux by branes was first considered in [5], and flux cascades in [6]. Another model of inflation involving compact extra dimensions to extend the field range is [9].

Mechanism:

The basic mechanism is the “flux discharge cascade” phenomenon recently discovered in [6], where a $p+2$-form electric flux threading at least one compact dimension can “unwind”, repeatedly discharging in a cascade triggered by the quantum nucleation of a charged brane, and hence steadily decreasing the effective four dimensional vacuum energy. The necessary ingredients are:

- A $(p+2)$-form field strength $F$ with $p \geq 3$, and $p$-branes that are electrically charged under $F$.
- A $D = 4 + q$ dimensional spacetime $dS_4 \times M_q$, where $dS_4$ is 4D de Sitter spacetime and $M_q$ is a stabilized compact $q$-manifold with $q \geq 1$
- $Q_0 \gg 1$ units of $F$ flux threading $dS_4$ and a $p-2$ cycle in $M_q$, supplying the $dS_4$ vacuum energy

The simplest realization—which for clarity we will focus on in this note—is a 5D space-time $dS_4 \times S_1$ with a bare cosmological constant, $Q_0$ units of initial 5-form flux, and some other source of energy (for instance Casimir energy...
of several bosonic and fermionic fields \[10\] that stabilizes the \(S_1\).

Classically, a \(D\)-form flux in \(D\) space-time dimensions (a “top form”) is simply a (positive) contribution to the cosmological constant. However, \(F\) can discharge locally by an analogue of the Schwinger process—the nucleation of an approximately spherical bubble of electrically charged brane \[5\], which we assume to be smaller than the Hubble length \(1/H\), 

given the ingredients of a curved spacetime metric and extra compact dimension(s). To study the decay, we should first find the instanton that corresponds to the nucleation of the bubble of brane. We will focus on the simplest case \(dS_4 \times S_1\), with Euclidean signature metric
\[
ds_E^2 = H^{-2} (d\xi^2 + \sin^2(\Omega_3^2) + dz^2), \quad \text{where } z \simeq z + l.
\]

Typically the dominant instanton for decay of a false vacuum state has the maximal symmetry possible. We will assume the brane is thin, and that the instanton depends only on \(\xi\) and \(z\) in accord with the symmetry of the initial state. We are interested in the case where the size \(\Delta z\) of the instanton in the \(z\) direction satisfies \(\Delta z < l\), so the periodic boundary conditions do not affect the solution (at least in the thin-wall limit). Finally, when the initial number of flux units \(Q_0 \gg 1\) the gravitational backreaction of a single bubble is small and can be ignored for the purpose of finding the instanton.

With these assumptions the instanton solution is fully characterized by the location of the wall \(z = \pm z_0(\xi)\), where \(\pm\) refers to two symmetric halves and we have chosen \(z = 0\) as the center of the bubble. To find \(z_0(\xi)\), one should minimize the action \(S_E = -\kappa \int_V \sqrt{g_E} + \sigma \int_{\partial V} \sqrt{g_E(\text{induced})}\), where \(\kappa\) is the difference in energy density on the two sides of the wall, \(\sigma\) is the tension of the wall, and \(V\) is the volume enclosed by the bubble.

Extremizing this action results in equations of motion that can be solved analytically for \(dz_0/d\xi\), with the integration constants fixed by the requirements of finite action and smoothness. For brevity we will not reproduce the result here. The coordinate shape of the instanton is oblate, in the sense that the maximum excursion in \(z\) is less than that in \(\xi\), and approaches spherical with radius \(R_0 = 4\sigma/\kappa\) in the limit \(R_0 \ll 1/H\).

To find the time evolution of the bubble immediately after it nucleates, one analytically continues the instanton solution using \(\xi \to it\), \(d\Omega_3 \to id\Omega_3\), where \(dh_3^2\) is the metric on a unit 3-hyperboloid. After this continuation the metric is de Sitter spacetime in a hyperbolic slicing, times \(S_4\) \(ds^2 = -dt^2 + \sinh^2(t) (dh_3^2 + dz^2)\), while the analytically continued instanton—that now depends only on \(t\)—describes a spherical brane that expands in both the de Sitter and \(z\) directions. As in the case of Coleman-de Luccia tunneling, the interior of the bubble can be foliated by spacelike 3-slices in the de Sitter directions that are homogeneous and isotropic and have constant curvature and energy density, and the resulting FRW cosmology is dominated by negative spatial curvature immediately after the bubble forms.

There are two crucial differences with standard Coleman-de Luccia tunneling. The first is that the bubble expands in the extra, flat \(z\) direction with a limiting asymptotic speed: \(\lim_{t \to \infty} dz_0/dt = v = 1/\sqrt{1 + (3\sigma H/\kappa)^2}\). In the models of interest \(3\sigma H/\kappa \ll 1\) and \(v\) is relativistic during the cascade. The second difference is that because \(z\) is compact, the bubble will collide with an image bubble when \(z_b = Nl/2\) (for
integer $N$). For the same reason cosmic bubble collisions in 4D preserve an $SO(2,1)$ hyperbolic symmetry in the directions transverse to the collision axis (see e.g. [11]), these collisions occur on spacelike surfaces of constant $t$ (3-hyperboloids), taking place at approximately equally spaced instants during the cascade. Crucially, brane self-collisions do not disturb the homogeneity and isotropy of the open 4D FRW cosmology inside the bubble.

After a few e-folds of de Sitter expansion, the radius of curvature of the bubble (and equivalently, that of the hyperbolic constant-$t$ slices) becomes exponentially longer than the de Sitter length $1/H$ (Fig. 1). From this time forward the space-time metric can be well approximated as $d\bar{s}^2 = -dt^2 + \exp(2Ht)d\vec{x}^2 + dz^2$, and the bubble by two disconnected parts: a planar (anti)brane extended in the $\bar{x}$ directions and located at $z = (-) + z_b(t)$, moving at close to the terminal velocity $(-) + v$. Every interval of time $\Delta t = l/2v$, the brane and anti-brane collide.

Such collisions can lead to a number of possible outcomes: perfect transmission or reflection, total annihilation, or transmission or reflection accompanied by some particle or string production. In string theory at relativistic velocities the branes transmit, because the tachyonic mode of the open string (that represents the annihilation channel) does not have time to condense as they pass through one another [6] [12]. There is some string production, which we will briefly mention later.

In general, assuming the branes transmit through one another, each collision produces a new region in between the planes where the flux has been reduced by one additional charge unit. Averaged over timescales of order $l/v$, this corresponds to a decrease in the flux $F$ that is linear in time, or a decrease in the energy density $\rho \sim F^2$ that is quadratic in time—much like $m^2\phi^2$ inflation. Other dependences are possible with more complex compact geometries [12].

**Effective Action:** In order to derive the Lorentzian equations of motion for the brane and its fluctuations, we will need a more explicit form of the action for the brane and flux. As we will see, the position $z_b$ of the brane in the extra dimension plays the role of a 4D inflaton, with a Dirac-Born-Infeld kinetic term [13] and an effective potential that is quadratic up to a small oscillating component. The brane/flux action is

$$S = \int_0^l dz \int dt d^3x \, e^{3Ht} \times$$

$$\left\{ -2\sigma \sqrt{1 - (\partial_t z_b)^2 + e^{-2Ht}(\partial_x z_b)^2} \delta(z - z_b) - \frac{F_5^2}{2 \cdot 5!} \right\},$$

where coupling between the brane and the flux determines

$$\frac{F_5^2}{5!} = \mu^5 \left( Q_0 + \sum_{j=-\infty}^{\infty} [\Theta(z - z_b + jl) - \Theta(z + z_b + jl)] \right)^2.$$

Here $\mu^{5/2}$ is the charge of the brane, and the sum arises due to the periodicity of $z$. The action can be integrated over $z$ to obtain a 4D effective action:

$$S = \int dt d^3x e^{3Ht} \left( -2\sigma \sqrt{1 - (\partial_t z_b)^2 + e^{-2Ht}(\partial_x z_b)^2} - V(z_b) \right).$$

Here $V(z_b)$ is a piecewise-linear approximation to a quadratic:

$$\frac{dV}{dz_b} = -2\mu^5 \left( Q_0 - \frac{1}{2} - \frac{2z_b}{l} \right),$$

where $[\ldots]$ denotes integer part.

The position of the brane $z_b(t)$ can be found by solving the equations of motion that result from this action, with the result that $\dot{z}_b \approx 1/\sqrt{1 + (3\sigma H/\kappa)^2}$ is nearly constant throughout the cascade. This is because the change in energy density across the brane $\kappa = \mu^5(Q - 1/2)$, while $H \approx \sqrt{(8\pi G_N/3)\rho_F} = \sqrt{(8\pi G_N/3)\mu^5Q^2/2}$ is close to linear in $Q$. Hence $v$ is approximately constant and $z_b = \int v \, dt$ is linear in time, up to oscillatory corrections with period $l/(2v)$ and amplitude suppressed by $1/Q$.

**Perturbations:** In any model of inflation, one is interested in the curvature perturbations $\zeta$ on the reheating surface. In our model, inflation proceeds until most or all of the initial flux $Q_0$ is discharged. At that point the velocity $v$ of the brane and anti-brane becomes non-relativistic, and the next collision will lead the branes to annihilate rather than transmitting each other, thereby reheating the universe and ending inflation. The reheating surface is simply the position of the brane $z_b(\vec{x})$ at the time this last collision occurs.

If the total flux discharged is $Q_t$, reheating occurs when $z_b(t) + \delta z_b(\vec{x}, t) = vt + \delta z_b = Q_t/l/2$. Solving for $t$ yields $t = Q_t/2v - \delta z_b/v \equiv t_0 + \delta t$. The curvature of this hypersurface is $a(t_0 + \delta t) = a(t_0) + \dot{a}(t_0)\delta t$, and so $\zeta = \delta a/a = H\delta t = H\delta z_b/v = H\delta z_b/z_b$, where $H$ and $\dot{z}_b$...
are evaluated at horizon crossing as usual. Therefore, to
find $\zeta$ one should compute the perturbations in the brane
direction $z_b$, for which there are several distinct sources.

Since $z_b$ is nearly massless, de Sitter quantum fluctuations
lead to perturbations in the usual way. In addition, if the brane collisions are inelastic the energy going into
particle production is a source of friction for the brane’s
velocity, and variations in the density of the produced
particles will therefore source perturbations in $z_b \ [14, 15]$. The tensor-to-scalar ratio is

$$\frac{P_T}{P_s} = \frac{H^4}{8\pi^2\sigma v^2}. \tag{2}$$

**Tensor Perturbations:** The spectrum of tensor per-
turbations due to de Sitter fluctuations is simply

$$P_h = \frac{16G_N H^2}{\pi}. \tag{3}$$

The tensor-to-scalar ratio is

$$r = \frac{128\pi G_N \sigma v^2}{H^2} = v^2 \frac{R_0}{l} \frac{Q^4}{Q^4},$$

where again $R_0 = 4\sigma/\kappa$ is the approximate radius of
the bubble when it formed. Since $R_0/l < 1$ but not
necessarily very small, and $Q \gg 10^2$, the tensor modes in
this model are potentially observable in the near future.

**Perturbations from String/Particle Production:** To compute the contribution to $\delta z$ from variations in the
number of strings or particles produced by brane colli-
sions requires a model of the brane dynamics and
degrees of freedom. In string theory, stretched open strings
will be pair-produced at each collision, because the mass
of the stretched strings changes with time as the brane
and anti-brane pass near or through each other. The
rate can be calculated at one loop in $g_s$ by computing the
imaginary part of the annulus diagram with appro-
priate boundary conditions. Given the rate, perturba-
tions in $\delta z$ can be computed using a technique similar to
that of $[14, 15]$. The result is that in a relatively broad
range of string parameter space, particularly with $p = 4$
or 5 (4- or 5-branes), string production is subdominant
to the largely model-independent de Sitter perturbations

[2, 12]. With other choices of parameters string produc-
tion is dominant, but in the interest of brevity we will
not describe the results here.

Brane collisions may also create tensor perturbations,
either directly from Bremsstrahlung or from the decay of
produced particles or strings, and the amplitude of these
perturbations might exceed the de Sitter contribution
\[13 \ 14]. It would be interesting to investigate this possibility
further.

**Tilt:** Because the variation in $v$ during the cascade
is very small, the tilt of the scalar spectrum $n_s - 1 =
dlnP/dlnk$ arises almost entirely from the $H^2$ term in

\[2]:

$$n_s - 1 \approx 4 \frac{H}{H^2} \frac{dt}{dlnk} \approx 4 \frac{H}{H^2} \approx -2/N_*, \tag{4}$$

where $N_* = \int dt = \int dH (H/H) \approx -H^2/(2H)$ is the
number of efolds from the time the quadrupole mode
crossed the horizon during inflation to the end of infla-
tion. In a model with a high reheating temperature $N_r \sim 60$, and the tilt $[4]$ is then consistent with observation.

**Non-Gaussianity:** The largest non-linearity in the ef-
fective action is the Dirac-Born-Infeld kinetic term
\[13 \ 17] which produces non-Gaussianity of equilateral type,
with an $f_{NL} \sim (1 - c_s^2)/c_s^2$. Higher dimensional gen-
eralizations of this model may also produce local non-
Gaussianity at reheating, as we will discuss below.

**Reheating:** When most or all of the flux has been dis-
charged, the Lorentz force that keeps the branes moving
around the circle is greatly reduced. At this point the
various dissipative and attractive forces on the brane and
anti-brane (for instance, those from stretched strings)
will bring them to close proximity at low velocity, at
which point they will annihilate and convert an energy
density $2\sigma$ into particles. While the details of this an-
nihilation are non-perturbative, the dynamics must be
local on Hubble scales, and so long as it does not af-
tect super horizon perturbations is largely irrelevant to
the cosmological predictions of the model. Rare regions
in which different amounts of flux have discharged may
arise, and these regions may expand, collide, and gen-
erate perturbations on larger than Hubble scales (pos-
sibly producing gravitational waves, as in [18]). While
the probability of such regions seems exponentially small
(because $\delta z_b \ll 1/2$), the dynamics of this is complex and
remains to be carefully investigated.

**Generalizations:** The model we focused on here is the
simplest possible, with one extra dimension and a $p = 5$
form field strength. However, the phenomenon of flux
cascades extends to a broad array of models with different
values of $p$ and numbers of compact dimensions $[6]$.

If the brane has co-dimension greater than 1 (i.e. if
$p < D + 2$), there may be additional light scalars that
describe the position of the brane in the transverse di-
ensions. Along with $z_b$—the radius of the brane bubble
in the compact dimension(s) the flux threads—these
additional scalars determine the time of re-heating (because to annihilate, the brane and anti-brane must come close together). Therefore, fluctuations in these scalars convert into curvature perturbations in a manner similar to that of \cite{19, 20}, and could produce a potentially significant level of local non-Gaussianity.

Another difference arises when \( p > 3 \), so that the compact cycle the flux threads is more than 1D. In such cases the geometry of the cycle is important. For instance, with \( p = 4 \) the cycle can be a \( S_2 \), in which case the brane is a string on the \( S_2 \) that repeatedly oscillates from pole to pole, wiping away one unit of flux on each pass \cite{6}. In this case the self-collisions are qualitatively different, because at least classically the brane shrinks to a point and inverts. Furthermore, the curvature of the \( S_2 \) affects the dynamics of both the brane and its perturbations. Further investigation of the higher-dimensional versions of this model may prove very interesting.

**Conclusions:**

We have presented a novel class of microscopic inflationary models that incorporates many effective models of inflation previously considered, including DBI inflation \cite{13}, trapped \cite{14} and dissipative \cite{15} inflation, and hybrid inflation. It provides a building block for new effective theories, and can potentially be naturally embedded in string theory. Our inflationary mechanism does not require significant fine-tuning of either the parameters or the initial conditions, and naturally produces a spectrum of perturbations that is consistent with observation. All models in this class predict an interesting and characteristic set of features in the power spectrum, including oscillatory components and potentially observable non-Gaussianities and tensor modes.

**Acknowledgements:** MK is supported by NSF CAREER grant PHY-0645435 and NSF grant PHY-1214302. G. D’A. is supported by a James Arthur Fellowship. The authors thank A. Brown, S. Dubovsky, G. Dvali, R. Flauger, B. Freivogel, V. Gorbenko, A. Hebecker, A. Lawrence, J. Maldacena, L. McAllister, M. Roberts, L. Senatore, T. Tanaka, T. Weigand, and M. Zaldarriaga for useful discussions.

---

[1] R. Bousso and J. Polchinski, JHEP, \textbf{0006}, 006 (2000), arXiv:hep-th/0004134 [hep-th].
[2] L. Susskind, (2003), arXiv:hep-th/0302219 [hep-th].
[3] A. H. Guth, Phys.Rev., \textbf{D23}, 347 (1981).
[4] A. D. Linde, Phys.Lett., \textbf{B108}, 389 (1982).
[5] J. D. Brown and C. Teitelboim, Nucl.Phys., \textbf{B297}, 787 (1988).
[6] M. Kleban, K. Krishnaiyengar, and M. Porrati, JHEP, \textbf{1111}, 096 (2011), arXiv:1108.6102 [hep-th].
[7] A. R. Brown, Phys.Rev.Lett., \textbf{101}, 221302 (2008), arXiv:0807.0457 [hep-th].
[8] J. T. Giblin, L. Hui, E. A. Lim, and I.-S. Yang, Phys.Rev., \textbf{D82}, 045019 (2010), arXiv:1005.3493 [hep-th].
[9] E. Silverstein and A. Westphal, Phys.Rev., \textbf{D78}, 106003 (2008), arXiv:0803.3085 [hep-th].
[10] N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and G. Villadoro, JHEP, \textbf{0706}, 078 (2007), arXiv:hep-th/0703067 [HEP-TH].
[11] M. Kleban, Class.Quant.Grav., \textbf{28}, 204008 (2011), arXiv:1107.2593 [astro-ph.CO].
[12] G. D’Amico, R. Gobbetti, M. Kleban, and M. Schillo, (To appear).
[13] M. Alishahiha, E. Silverstein, and D. Tong, Phys.Rev., \textbf{D70}, 123505 (2004), arXiv:hep-th/0404084 [hep-th].
[14] D. Green, B. Horn, L. Senatore, and E. Silverstein, Phys.Rev., \textbf{D80}, 063533 (2009), arXiv:0902.1006 [hep-th].
[15] D. Lopez Nacir, R. A. Porto, L. Senatore, and M. Zaldarriaga, JHEP, \textbf{1201}, 075 (2012), arXiv:1109.4192 [hep-th].
[16] L. Senatore, E. Silverstein, and M. Zaldarriaga, JHEP, \textbf{1109}, 054 (2011), arXiv:1109.0542 [hep-th].
[17] L. Senatore, K. M. Smith, and M. Zaldarriaga, JCAP, \textbf{1001}, 028 (2010), arXiv:0906.3746 [astro-ph.CO].
[18] A. Kosowsky, M. S. Turner, and R. Watkins, Phys.Rev.Lett., \textbf{69}, 2026 (1992).
[19] G. Dvali, A. Gruzinov, and M. Zaldarriaga, Phys.Rev., \textbf{D69}, 023505 (2004), arXiv:astro-ph/0303591 [astro-ph].
[20] L. Senatore and M. Zaldarriaga, JHEP, \textbf{1204}, 024 (2012), arXiv:1009.2093 [hep-th].