Revisiting the no-boundary proposal with a scalar field

Caroline Jonas[1,*], Jean-Luc Lehners[1,†] and Vincent Meyer[2,‡]

[1]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Potsdam 14476, Germany*
[2]*Humboldt-Universität zu Berlin, Zum Grossen Windkanal 2, Berlin 12489, Germany*

Recent works have suggested that the no-boundary proposal should be defined as a sum over regular, not necessarily compact, metrics. We show that such a prescription can be implemented in the presence of a scalar field. For concreteness, we consider the model of Garay *et al.*, in which the potential is a sum of exponentials, and which lends itself to an analytical treatment. Compared to the earlier implementation, we find that saddle points with unstable fluctuations can be eliminated by imposition of an appropriate regularity condition. This leads to the appearance of additional saddle points, corresponding to unclosed geometries. We argue that such saddles will occur generically, though we also find in our example that they are subdominant to the closed, Hartle-Hawking, saddle points. When the potential is positive, classical spacetime is only predicted for inflationary histories. When the potential is negative, we recover the anti–de Sitter gravitational path integral, with a stable scalar field included. One puzzle that we find is that in general the path integral must be restricted to sum only over specific, discrete and late time dependent initial values of the scalar field. Only when the scalar is required to take real values is this puzzle eliminated, a situation that moreover leads to advantageous phenomenological characteristics.

I. INTRODUCTION

The no-boundary proposal is an attempt to describe the quantum wave function of the Universe. It was first introduced in the 1980s by Hartle and Hawking [1,2], and has recently been the focus of renewed interest, see e.g. [3–7]. This is because, on the one hand, it has proven remarkably difficult to construct competing theories of initial conditions and, on the other hand, there is a certain naturalness to the no-boundary proposal that makes it compelling. Recent works have reinforced this last point: for instance, if one studies gravitational path integrals in the presence of a negative cosmological constant, one is led to the equivalent of no-boundary conditions, but in spacetimes that are asymptotically anti–de Sitter (AdS) [8]. What is remarkable about this setting is that via AdS/CFT it is known what result to expect for the wave function [9] and thus we can be rather confident that, in this setting, the no-boundary prescription is correct. Then, via analytic continuation in the cosmological constant, one recovers the standard cosmological no-boundary proposal [10]. It has further been shown that the no-boundary proposal is robust to the inclusion of higher order terms in the Riemann curvature, which are expected to arise via quantum gravitational corrections [11,12]. Finally, the case has been made that, possibly together with loitering/emergent scenarios, the no-boundary proposal constitutes the only currently known way of rendering cosmological amplitudes finite and well defined [13]. Despite these appealing features we should point out that a crucial issue, namely whether or not the no-boundary proposal is in agreement with observations, is still open, for a recent discussion see [7]. This is partly due to the fact that the predictions of the no-boundary proposal depend on the dynamical theory under consideration, and partly because the no-boundary proposal explains aspects of early Universe theories that are otherwise already assumed, such as the classicality of spacetime and the ground state of fluctuations [14,15], so that it may be difficult to crystallize out truly new predictions. Still, we will have a little more to say about observational predictions in the discussion section.

Traditionally, the no-boundary proposal was defined as a path integral summing over geometries that are both compact and regular [2]. However, this definition is in tension with the uncertainty principle, as compactness is a condition on the scale factor and regularity a condition on its conjugate momentum (i.e. the expansion rate of the Universe). If one imposes compactness alone, then in general one finds that the path integral contains two types of saddle points: those

*caroline.jonas@aei.mpg.de
†jlehners@aei.mpg.de
‡vmeyer@aei.mpg.de

with stable fluctuations, and their complex conjugates with unstable fluctuations. This hinders a clear-cut definition of the no-boundary proposal, at least in certain simple minisuperspace models [16,17]. By contrast, if one imposes regularity alone, then one may eliminate unstable saddle points [6,18] (early suggestions to use a momentum condition, though with somewhat different motivations, include [19,20], and momentum conditions were also used in [5]). The price to pay is that it is not guaranteed that the saddle points will also be compact, which is certainly a condition for interpreting the no-boundary wave function as describing the creation of the Universe. For pure gravity and isotropic metrics, this has been shown to work, and the saddle points indeed end up being both regular and compact [6,18].

However, the link to observations is done using a scalar field, either with an inflationary [21] or an ekpyrotic [22] potential. In both cases, it has been shown that the no-boundary proposal can explain the nucleation of a classical spacetime from nothing, as both mechanisms drive the wave function to a semiclassical Wentzel-Kramers-Brillouin (WKB) form [15,23]. For ekpyrosis, one faces the challenge of understanding the bounce to the expanding phase. In phenomenological models, such a bounce can be included [24]. But on a fundamental level bounces remain ill understood, as it remains unclear whether quantum gravity can allow for effective null energy violations on large scales [25]. Hence we will not consider ekpyrotic models further in this work.

We will rather focus on inflationary potentials here, revisiting a minisuperspace model considered previously by Garay *et al.* [26] and that has the property that it can be treated analytically to a large extent. Garay *et al.* defined the no-boundary wave function by a sum over compact metrics. Since they did not discuss cosmological fluctuations in their work, they were not concerned by the fact that, with this definition, stable and unstable saddle points are typically linked by steepest descent contours, making it impossible to avoid the unstable saddles [17]. Our goal is to see whether the proposal to define the no-boundary wave function by a sum over regular metrics is viable in this case.

We find that, just as for the pure gravity case, imposing regularity is equivalent to requiring the absence of momentum flow at the creation of the Universe [10]. However, in contrast to the pure gravity case where this condition was unique, here it leads to a family of possible initial conditions, with a parameter γ that may loosely speaking be thought of as the initial value of the scalar field (this interpretation is not always valid, as we will discuss). By construction, the saddle points are always regular, but only for certain values of γ can the saddle points also be compact. Even then, we find that closed saddle point geometries appear alongside unclosed geometries, which we argue to be generically present. Still, an important result is that the unclosed geometries are always subdominant to the closed ones, so that the no-boundary wave function ends up being dominated by closed,

regular geometries, as desired. As previously noted by Garay *et al.* [26], the dominant saddle points do not lead to a WKB wave function in all regions of phase space, and in fact a classical spacetime is only predicted in inflationary regions of the potential. We also extend our analysis to include a negative potential, in which case we recover geometries that are asymptotically AdS, but with a dynamical scalar field.

An important question highlighted by our work is how to define the no-boundary wave function in general. As we will see, the regularity condition (which may also be seen as the absence of momentum flow) does not in and of itself uniquely fix the initial conditions. Further input is required, such as the condition that compact saddle points exist. This issue, and the resulting puzzles that it implies, are discussed in more detail in the discussion section. However, let us point out already that recent works proposing to define quantum field theories on complex spacetimes using specific convergence conditions offer a clue as to how the puzzle may ultimately be resolved.

II. MODEL AND PATH INTEGRAL

We are interested in the no-boundary wave function of the Universe $\Psi(h_{ij}, \Phi_f)$, a functional of a three-dimensional hypersurface h_{ij} on which a scalar field configuration Φ_f resides. The arguments may be thought of as a possible current state of the Universe, with the wave function providing probabilities for inequivalent histories of the Universe [15]. The wave function is formally defined as a path integral

$$\Psi(h_{ij}, \Phi_f) = \int Dg_{\mu\nu} D\Phi e^{\frac{i}{\hbar}S}, \tag{1}$$

where the integration runs over a restricted class of 4-manifolds with metrics $g_{\mu\nu}$ and scalar field configurations. A precise definition of the above path integral remains an important goal of quantum cosmology. In the present paper, we will investigate how to define this integral in a minisuperspace setting, where we restrict the metric to be of Friedmann-Lemaitre-Robertson-Walker (FLRW) form and the scalar to be homogeneous.

The action is taken to be that for general relativity minimally coupled to a scalar field. A Gibbons-Hawking-York term is added on the final boundary, but not on the initial boundary, since we would like to impose a regularity (Neumann) condition there, rather than fixing the metric (which would have required a Dirichlet condition),

$$S = \frac{1}{16\pi G} \int_M d^4x \sqrt{-g} [R - 8\pi G(g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi + \tilde{V}(\Phi))]$$
$$- \frac{1}{8\pi G} \int_{\partial M_f} d^3x \sqrt{h} K. \tag{2}$$

The scalar potential $\tilde{V}(\Phi)$ will be specified shortly. We work in minisuperspace with an FLRW ansatz for the metric. More

specifically, we consider the Universe to be homogeneous and isotropic, with scale factor a:

$$ds^2 = \frac{2G}{3\pi}\left[-\frac{N^2(\tau)}{a^2(\tau)}d\tau^2 + a^2(\tau)d\Omega_3^2\right]. \quad (3)$$

The time coordinate has been defined in a somewhat unusual way for later convenience [26]. The scalar field is also assumed to be homogeneous, $\Phi = \Phi(\tau)$. One can rescale the scalar field and the potential to remove numerical factors from the action, to find

$$\Phi(\tau) = \sqrt{\frac{3}{4\pi G}}\phi(\tau), \qquad \tilde{V}(\Phi) = \left(\frac{9}{8G^2}\right)V(\phi), \quad (4)$$

$$\Rightarrow S = \frac{1}{2}\int d\tau\left[N - \frac{a^2\dot{a}^2}{N} + \frac{a^4\dot{\phi}^2}{N} - a^2 NV(\phi)\right] - \frac{1}{2}\left[\frac{a^3\dot{a}}{N}\right]^0. \quad (5)$$

The last term is a surface term at $\tau = 0$. The surface term at $\tau = 1$ has disappeared after integration by parts.

As in [26], we use the following scalar potential

$$V(\phi) = \alpha\cosh(2\phi) + \beta\sinh(2\phi), \quad (6)$$

which renders the system analytically solvable for arbitrary real parameters α and β. This becomes obvious by applying the following change of variables:

$$x = a^2\cosh(2\phi), y = a^2\sinh(2\phi), \quad (7)$$

which transforms the action into

$$2S[x,y] = \int d\tau N\left[\frac{\dot{y}^2 - \dot{x}^2}{4N^2} + 1 - \alpha x - \beta y\right] - \left[\frac{x\dot{x} - y\dot{y}}{2N}\right]^0. \quad (8)$$

We define the momenta conjugate to x, y by

$$\Pi_x(\tau) = -\frac{\dot{x}}{2N}, \qquad \Pi_y(\tau) = \frac{\dot{y}}{2N}. \quad (9)$$

The equations of motion and constraint deriving from the action are very simple,

$$\ddot{x} = 2N^2\alpha, \qquad \ddot{y} = -2N^2\beta, \quad (10)$$

$$\frac{1}{4N^2}(\dot{x}^2 - \dot{y}^2) + 1 - \alpha x - \beta y = 0. \quad (11)$$

Before proceeding, we must first discuss the boundary conditions that should be imposed on the fields. There are two complementary ways of doing this, and fortunately both agree. The first is to continue with the action integral,

and look at the variational problem that it implies. Varying with respect to the fields x and y we obtain the surface terms

$$\delta(2S) \supset \left[-x\delta\left(\frac{\dot{x}}{2N}\right) + y\delta\left(\frac{\dot{y}}{2N}\right)\right]^0 + \left[\frac{1}{2N}(-\dot{x}\delta x + \dot{y}\delta y)\right]^1. \quad (12)$$

Thus at $\tau = 0$ we may impose a condition on the momenta $\Pi_x \equiv -\Pi_x(\tau = 0)$, $\Pi_y \equiv \Pi_y(\tau = 0)$. (Note that, for later calculational convenience, we define Π_x with an additional minus sign.) The appropriate regularity condition is obtained by looking at the constraint (11) when $x, y \to 0$, implying that we should impose

$$\Pi_x^2 - \Pi_y^2 = -1 \quad \text{(regularity condition)}. \quad (13)$$

At $\tau = 1$ we can simply fix the field values $x(1) \equiv x_f$, $y(1) \equiv y_f$.

A second approach is to look at the Wheeler-DeWitt (WDW) equation, which is the quantum version of the classical constraint (11). As discussed in [10], the no-boundary proposal can be seen as the requirement that the Universe is entirely self-contained, and that consequently no momentum flows into it from "outside" at the nucleation of the Universe. In momentum space, the WDW equation is obtained as the operator version of the constraint (11), replacing $x \to -i\frac{d}{d\Pi_x}, y \to i\frac{d}{d\Pi_y}$,

$$\left(\Pi_x^2 - \Pi_y^2 + 1 + i\alpha\frac{d}{d\Pi_x} - i\beta\frac{d}{d\Pi_y}\right)\Psi = 0. \quad (14)$$

The absence of momentum flow, $\frac{d\Psi}{d\Pi_{x,y}} = 0$, thus yields the exact same condition, $\Pi_x^2 - \Pi_y^2 = -1$.

It will be useful to parametrize the initial momenta as

$$\Pi_x = i\cosh(2\gamma), \qquad \Pi_y = i\sinh(2\gamma). \quad (15)$$

This means that we still have a free parameter γ specifying the initial conditions. Determining and interpreting γ will be one of the main tasks of the present work. Before continuing, let us remark that one condition on γ is familiar from earlier works [6,8,18], namely that the sign of the imaginary part of Π_x must be chosen such that perturbations of the geometry and matter are stable, rather than unstable.[1] The appropriate choice is

$$\text{Im}[\Pi_x] = -\text{Im}[\Pi_{x(0)}] > 0 \quad \text{(stability condition)}. \quad (16)$$

---

[1] This condition may also be seen as choosing the Hartle-Hawking wave function [2], rather than Vilenkin's tunneling wave function [27].

With the boundary conditions specified so far (fixing the momenta at $\tau = 0$ and the field values at $\tau = 1$), it is straightforward to find solutions to the equations of motion for x and y,

$$\bar{x}(\tau) = \alpha N^2 \tau^2 + 2N\Pi_x \tau + x_f - 2N\Pi_x - \alpha N^2, \quad (17)$$

$$\bar{y}(\tau) = -\beta N^2 \tau^2 + 2N\Pi_y \tau + y_f - 2N\Pi_y + \beta N^2. \quad (18)$$

We may use these solutions to perform the path integrals over x, y by writing a generic field evolution as a sum of such a solution plus a variation X, Y:

$$x(\tau) = \bar{x}(\tau) + X(\tau), \qquad y(\tau) = \bar{y}(\tau) + Y(\tau). \quad (19)$$

It is important that the variations are completely arbitrary, except that the total fields must obey the boundary conditions, implying that we must fix

$$\dot{X}(0) = 0, \quad X(1) = 0, \quad \dot{Y}(0) = 0, \quad Y(1) = 0. \quad (20)$$

The action can then be rewritten as

$$S = S_{\text{on-shell}} + \frac{1}{2}\int d\tau N\left(\frac{\dot{Y}^2 - \dot{X}^2}{4N^2}\right) + \left[\frac{\dot{\bar{y}}Y - \dot{\bar{x}}X}{4N}\right]_1^1$$
$$- \left[\frac{x\dot{X} - y\dot{Y}}{4N}\right]^0. \quad (21)$$

Due to (20) the surface terms vanish, and because the original action was quadratic we now obtain Gaussian integrals over X and Y,

$$\Psi = \int dN e^{iS_{\text{on-shell}}/\hbar} \cdot \int_{\dot{X}(0)=0}^{X(1)=0} \mathcal{D}X \int_{\dot{Y}(0)=0}^{Y(1)=0}$$
$$\times \mathcal{D}Y \exp\left(\frac{i}{2\hbar}\int d\tau \frac{\dot{Y}^2 - \dot{X}^2}{4N}\right). \quad (22)$$

The calculation of the fluctuation integrals with the required mixed Neumann-Dirichlet boundary conditions can be done explicitly, see the appendix of [8] for details, and the result is simply a numerical prefactor

$$\Psi[x_f, y_f, \gamma] = \frac{8\hbar}{\pi}\int dN \exp\left(\frac{i}{\hbar}S_{\text{on-shell}}[N, x_f, y_f, \gamma]\right). \quad (23)$$

In Appendix A we provide a quick, indirect way of verifying this result. Since this factor is just an overall constant, not affecting relative probabilities, we will drop it in the following. We are thus left with an ordinary integral over the lapse N, with the integrand arising from the explicit integration over the solutions (17), (18),

$$S_{\text{on-shell}} = \frac{N^3(\alpha^2 - \beta^2)}{6} + \frac{N^2(\alpha\Pi_x + \beta\Pi_y)}{2}$$
$$- \frac{N(\alpha x_f + \beta y_f)}{2} + \frac{y_f\Pi_y - x_f\Pi_x}{2}. \quad (24)$$

The saddle points of the integral are located at

$$\begin{cases} N_{\text{saddle}\pm}^{\alpha\neq\beta} = \frac{-(\alpha\Pi_x + \beta\Pi_y) \pm \sqrt{(\alpha\Pi_x + \beta\Pi_y)^2 + (\alpha^2 - \beta^2)(\alpha x_f + \beta y_f)}}{\alpha^2 - \beta^2}, \\ N_{\text{saddle}}^{\alpha=\beta} = \frac{1}{2}\frac{x_f + y_f}{\Pi_x + \Pi_y}. \end{cases} \quad (25)$$

Thus, for fixed boundary conditions, in particular fixed γ, we obtain two saddle points in a general potential, and a single saddle point for the exponential potential ($\alpha = \beta$). In terms of the original variables, the saddle points may be rewritten as

$$\begin{cases} N_{\text{saddle}\pm}^{\alpha\neq\beta} = \frac{1}{\alpha^2 - \beta^2}\left(-iV(\gamma) \pm \sqrt{(\alpha^2 - \beta^2)a_f^2\cosh(2\phi_f) - V(\gamma)^2}\right), \\ N_{\text{saddle}}^{\alpha=\beta} = -\frac{i}{2}a_f^2\exp\left(2\phi_f - 2\gamma\right). \end{cases} \quad (26)$$

As is evident, the saddle points (and thus also the semiclassical wave function) depend on the initial conditions via the parameter γ. Since γ fixes the initial momenta, and not the field values, it is not guaranteed that the saddle point geometries will be compact; i.e. it is not guaranteed that they close at $\tau = 0$. In fact, this will only occur for specific values of γ. We will adopt the prescription that we will simply sum over all values of γ that lead to at least one closed saddle points; i.e. we will take the total wave function to be given by

We should point out that this corresponds to a generalization of the path integral, where we now also sum over specific boundary conditions. In keeping with the spirit of the path integral, each possible history is weighted by $e^{iS/\hbar}$ only, and not by further factors, which is why it makes sense to weight the different boundary conditions

$$\Psi[x_f, y_f] = \sum_{\gamma \text{ s.t. } \exists\, \bar{a}(0)=0} \Psi[x_f, y_f, \gamma]. \quad (27)$$
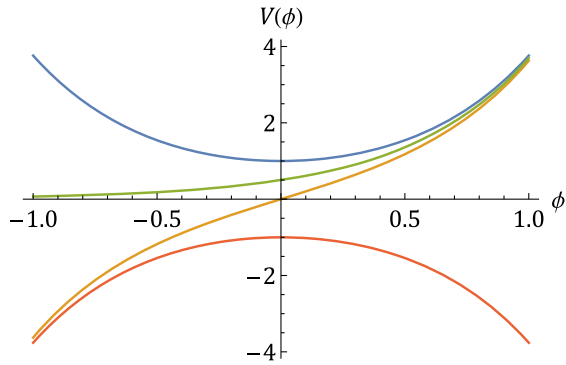
FIG. 1. Different scalar field potentials: $(\alpha, \beta) = (1, 0)$ hyperbolic cosine in blue, $(\alpha, \beta) = (0, 1)$ hyperbolic sine in orange, $(\alpha, \beta) = (1/2, 1/2)$ exponential in green and $(\alpha, \beta) = (-1, 0)$ negative hyperbolic cosine in red.

equally. We note that a closely related prescription has been implicitly used in earlier implementations of the no-boundary proposal, see e.g. [15]. There, Dirichlet boundary conditions were used, which specified both the initial values of the scale factor and the scalar field. In order to obtain closed saddle points, one then had to adapt the boundary condition on the scalar field depending on the final values of the fields. This may also be thought of as summing over appropriate initial conditions—we will return to discuss this important point in Sec. IV. But first, it is illuminating to study a few concrete examples to better see what consequences our definition entails. This will be the subject of the next section.

III. EXAMPLES

Despite its simplicity, the form of the potential $V(\phi) = \alpha \cosh(2\phi) + \beta \sinh(2\phi)$ that we are considering allows for a rich variety of qualitatively different behaviors, see Fig. 1. We will look at the case where the potential is positive, inflationary and has a minimum ($\beta = 0$, $\alpha > 0$). A second case is the opposite potential, with an AdS maximum at negative values of the potential ($\beta = 0$, $\alpha < 0$). Further, we can consider an exponential without

minimum ($\alpha = \beta$) and a potential that is unbounded both above and below ($\alpha = 0$). We will examine these cases in turn.

A. $\alpha = 1, \beta = 0$: $V(\phi) = \cosh(2\phi)$

The cosh potential is the most relevant example to study: it is inflationary, with a slow-roll region for very small values of $|\phi|$. Also, when the scalar field vanishes, the model reduces to gravity with a cosmological constant (plus scalar fluctuations), and hence we expect to recover the known no-boundary wave function in this limit. As we will see, this potential is the only one where a classical spacetime is predicted, which is a nontrivial result (a related discussion appeared in [26]). The model was also previously used to study large homogeneous quantum fluctuations in inflation [28].

Let us start by recalling the locations of the saddle points (25):

$$N_{\text{saddle}}^{\pm} = -i \cosh(2\gamma) \pm \sqrt{a_f^2 \cosh(2\phi_f) - \cosh^2(2\gamma)}. \quad (28)$$

The constraint equation (11), which is satisfied at the saddle points, directly implies that at the initial time $\tau = 0$ we have $x|_0 = 0$. Therefore, given that $a^4 = x^2 - y^2$, the initial geometry closes if and only if $y|_0 = 0$. From (18) we see that this is satisfied for

$$N_{\text{closed}} = \frac{y_f}{2\Pi_y} = \frac{a_f^2 \sinh(2\phi_f)}{2i \sinh(2\gamma)}. \quad (29)$$

Equating the saddle point value (28) with (29), we find that closed saddle point geometries must obey the equation

$$a_f^2 \sinh^2(2\phi_f) + 4 \sinh(2\gamma) \sinh(2\gamma - 2\phi_f) = 0. \quad (30)$$

We solve this equation for $\gamma \in \mathbb{C}$, and find three types of solutions valid in three different regions of the (a_f, ϕ_f) phase space:

$$\begin{cases} \gamma_{\mp}^{\text{I}} = \frac{\phi_f}{2} \mp \frac{i}{2} \arcsin(|\sinh(\phi_f)| \sqrt{a_f^2 \cosh^2(\phi_f) - 1}), \\[2mm] \gamma_{\mp}^{\text{II}} = \frac{1}{2} \cosh^{-1}\left(\sqrt{1 + \frac{\sinh^2(2\phi_f)}{4} (2 - a_f^2 \cosh(2\phi_f) \mp \sqrt{(2 - a_f^2 \cosh(2\phi_f))^2 - a_f^4})} \right), \\[2mm] \gamma_{\mp(n)}^{\text{III}} = \mp \frac{1}{2} \cosh[-1]\left(\frac{1}{2} \sinh(2\phi_f) \sqrt{a_f^2 \cosh(2\phi_f) - 2 \mp \sqrt{(a_f^2 \cosh(2\phi_f) - 2)^2 - a_f^4}} \right) + i\left(\frac{\pi}{4} + \frac{n \cdot \pi}{2} \right), \quad n \in \{0, 1\}. \end{cases} \quad (31)$$
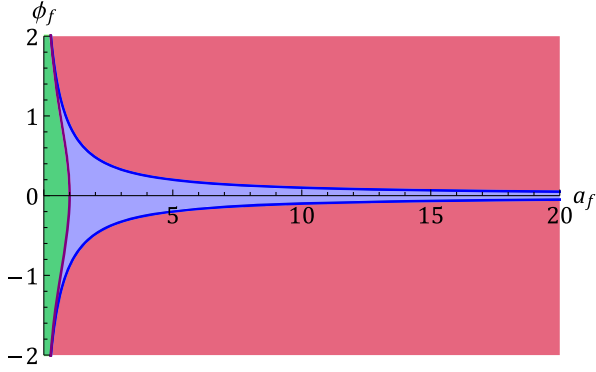
FIG. 2. Phase space regions (32). Region I is in blue, region II is in green and region III is in red. The purple curve is the condition $a_f^2 \cosh^2(\phi_f) = 1$, and the blue curve is the condition $a_f^2 \sinh^2(\phi_f) = 1$.

The above expressions are derived in Appendix B, and their regions of validity, represented in Fig. 2, are found to be

$$
\begin{cases}
\text{region I:} & a_f^2 \cosh^2(\phi_f) > 1 \text{ and } a_f^2 \sinh^2(\phi_f) < 1, \\
\text{region II:} & a_f^2 \cosh^2(\phi_f) < 1, \\
\text{region III:} & a_f^2 \sinh^2(\phi_f) > 1.
\end{cases}
\tag{32}
$$

We are now ready to compute the lapse integral. This is done by applying Picard-Lefschetz theory, as in [3]. For this, we must study the steepest descent contours (of the lapse integrand) emanating from the saddle points. On

these "thimbles" the lapse integral may then be defined as a sum of absolutely convergent integrals, which we can approximate to leading order in $\hbar$ by their saddle point values.

1. Region I: Complex saddle points

For both $\gamma_\mp^{\mathrm{I}}$ values, we find two saddle point values for the lapse N given by (28). All four saddle points are complex, but two of them are closed and the other two unclosed. This is because, for a given $\gamma_\mp^{\mathrm{I}}$ only one choice of sign in (28) may be made to match with the required value (29) for a closed geometry. But for each value of γ there are two saddle points, hence the second one must be unclosed. That the regularity condition may nevertheless be satisfied is due to the fact that the potential itself may vanish at $\tau = 0$. We explain this in somewhat more detail in Appendix C. Basically, this result follows from the fact that we have extended the fields to the complex plane, so that the potential is now a holomorphic (in fact entire) function of the fields. Picard's little theorem implies that such a function will generically vanish for some field value (s). Thus, the appearance of unclosed saddle points was in fact to be expected.

We illustrate the field evolutions at the saddle points with the numerical examples provided in Fig. 3. The saddle points appear in pairs, with opposite real parts of the lapse. This corresponds to geometries and scalar field evolutions that are complex conjugates of each other. As one can see very clearly in the figure, two of the saddle point geometries start out at zero size, while two others have an initial hypersurface with a nonzero, complex scale factor. In all
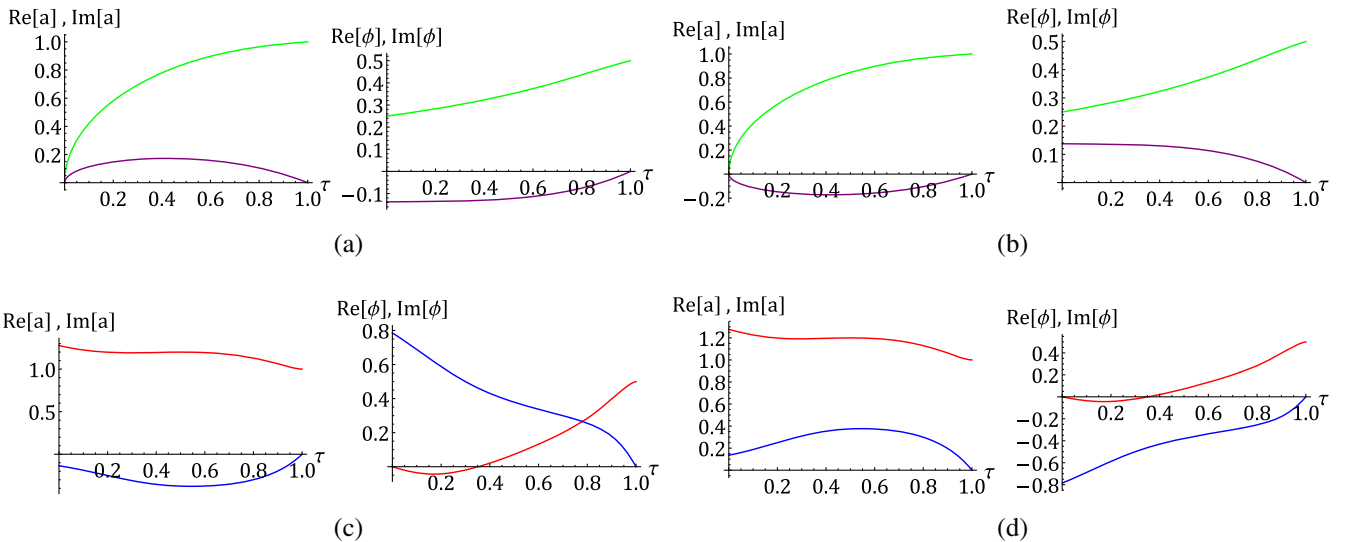

FIG. 3. Case $\alpha = 1$, $\beta = 0$, with N complex: $a_f = 1$, $\phi_f = 0.5$. In this case $\gamma_\mp = 0.25 \mp 0.137i$ and the saddle points are $N_+[\gamma_-] = 0.521{-}0.853i$, $N_-[\gamma_+] = -0.521 - 0.853i$, $N_-[\gamma_-] = -0.804 - 1.317i$, $N_+[\gamma_+] = 0.804{-}1.317i$. $N_+[\gamma_-]$ and $N_-[\gamma_+]$ are the closed saddle points. The graphs show the evolution of the scale factor and scalar field at the different saddle points. Real part in green for closed and in red for unclosed saddle points. Imaginary part in purple for closed and in blue for unclosed saddles. (a) Complex fields at $N_+[\gamma_-]$, (b) complex fields at $N_-[\gamma_+]$, (c) complex fields at $N_-[\gamma_-]$, and (d) complex fields at $N_+[\gamma_+]$.
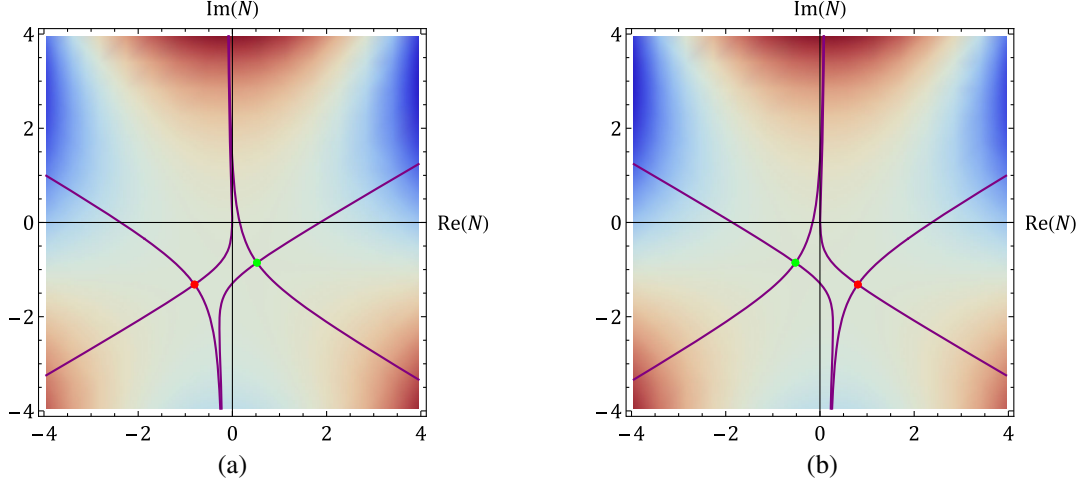
FIG. 4.   Case $\alpha = 1$, $\beta = 0$, with N complex: $a_f = 1$, $\phi_f = 0.5$. Density plot of the weighting $\mathrm{Re}[iS_{\mathrm{saddle}}^{\mathrm{on\text{-}shell}}]$ and flow lines in the complex N plane. Steepest descent (ascent) contours are those lines emanating from the saddle points and reaching blue (red) regions. The path integral is defined on sums of steepest descent contours. (a) $N_+[\gamma_-]$ (green dot) and $N_-[\gamma_-]$ (red dot) and (b) $N_-[\gamma_+]$ (green dot) and $N_+[\gamma_+]$ (red dot).

cases, the scalar starts from a complex field value at $\tau = 0$, for the unclosed saddles this is $\phi = \pm i\pi/4$, where the potential vanishes. For the closed saddle points, we have that

$$\phi(\tau = 0) = \lim_{\tau \to 0} \frac{1}{2}\mathrm{arctanh}\left(\frac{\bar{y}(\tau)}{\bar{x}(\tau)}\right) = \frac{1}{2}\mathrm{arctanh}\left(\frac{\Pi_y}{\Pi_x}\right) = \gamma. \tag{33}$$

Thus, for closed saddles, γ can simply be interpreted as the initial scalar field value. At $\tau = 1$, all field values are real, as required.

Figure 4 shows the steepest descent/ascent lines associated with the saddle points. We obtain one such figure for each value of γ. In each graph there is thus one closed and one unclosed saddle point. Their thimbles can be summed to a contour that runs parallel to the real N line—this is the most natural contour of integration, and most closely related to a Lorentzian contour. (It does not correspond to an integration over exactly Lorentzian metrics, because the initial momentum conditions imply that near $\tau = 0$ all metrics in the sum are necessarily complex.) Given that we are summing over the two relevant values of the initial conditions parameter γ, the resulting no-boundary wave function, to leading order in $\hbar$, is then given by

$$\Psi[x_f, y_f] = e^{\frac{i}{\hbar}S(N_+[\gamma_-])} + e^{\frac{i}{\hbar}S(N_-[\gamma_+])} + e^{\frac{i}{\hbar}S(N_+[\gamma_+])} + e^{\frac{i}{\hbar}S(N_-[\gamma_-])}. \tag{34}$$

The wave function is real, as the saddle points are pairwise complex conjugates of each other.

It remains to determine which saddle point(s) actually dominate, i.e. which saddles have the highest weighting. Recall that the saddle points are located at

$$N_\pm = -\Pi_x \pm \sqrt{x_f + \Pi_x^2}. \tag{35}$$

We can thus write the (on-shell) saddle point action as

$$S_{\mathrm{on\text{-}shell}} = \frac{N_\pm^3}{6} + \frac{N_\pm^2\Pi_x}{2} - \frac{N_\pm x_f}{2} + \frac{y_f\Pi_y - x_f\Pi_x}{2}, \tag{36}$$

$$= \mp \frac{\sqrt{x_f + \Pi_x^2}^3}{3} + \frac{\Pi_x^3}{3} + \frac{y_f\Pi_y}{2}. \tag{37}$$

Let us now compute the values of Π_x and Π_y for the different saddle points. The two possible solutions for γ are given from Eq. (31) by

$$\gamma_\mp = \frac{\phi_f}{2} \mp \frac{i}{2}\mathrm{arcsin}\left(|\sinh(\phi_f)|\sqrt{a_f^2\cosh^2(\phi_f) - 1}\right). \tag{38}$$

γ_- is closing the geometry of the saddle point N_+, while γ_+ is closing the geometry of the saddle point N_-. We then obtain

$$\begin{aligned}
\Pi_{x\mp} &= i\cosh(2\gamma_\mp) \\
&= \pm\sinh^2(\phi_f)\sqrt{a_f^2\cosh^2(\phi_f) - 1} \\
&\quad + i\cosh^2(\phi_f)\sqrt{1 - a_f^2\sinh^2(\phi_f)},
\end{aligned} \tag{39}$$

$$\begin{aligned}
\Pi_{y\mp} &= i\sinh(2\gamma_\mp) \\
&= \pm\frac{\sinh(2\phi_f)}{2}\sqrt{a_f^2\cosh^2(\phi_f) - 1} \\
&\quad + \frac{i\sinh(2\phi_f)}{2}\sqrt{1 - a_f^2\sinh^2(\phi_f)}.
\end{aligned} \tag{40}$$

Here we assume without loss of generality that $\phi_f > 0$, and we recall that we are looking for solutions in the region of the phase space where $a_f^2 \cosh^2(\phi_f) > 1$ and $a_f^2 \sinh^2(\phi_f) < 1$. Thus note that the imaginary parts of $\Pi_\pm$ do not depend on the choice of saddle point, only the real parts differ by a sign. The weighting of the saddle point is determined by the imaginary part of the action, $e^{-\text{Im}[S]}$. Hence from (37) we see that only the first term will be important in comparing weightings. If we define

$$\sqrt{x_f + \Pi_{x\mp}^2}^3 = \sqrt{A \pm iB}^3 \equiv (\rho e^{\pm i\theta})^{3/2} = \rho^{3/2} e^{\pm 3i\theta/2}, \quad (41)$$

then we find

$$\text{Im}\left[S_{\text{on-shell}}^{\text{closed}}[N_-[\gamma_+]]\right] - \text{Im}\left[S_{\text{on-shell}}^{\text{unclosed}}[N_+[\gamma_+]]\right] \quad (42)$$

$$= \text{Im}\left[S_{\text{on-shell}}^{\text{closed}}[N_+[\gamma_-]]\right] - \text{Im}\left[S_{\text{on-shell}}^{\text{unclosed}}[N_-[\gamma_-]]\right], \quad (43)$$

$$= -\frac{2}{3}\rho^{3/2}\sin\left(\frac{3\theta}{2}\right). \quad (44)$$

The saddle point that will contribute the most is the one with the smallest $\text{Im}[S]$ value. This depends on the sign of the sine, and hence on the value of the angle θ. With the definition (41) we have

$$B = 2\sinh^2(\phi_f)\cosh^2(\phi_f)\sqrt{a_f^2\cosh^2(\phi_f) - 1}\sqrt{1 - a_f^2\sinh^2(\phi_f)} > 0 \Rightarrow \theta \in (0, \pi), \quad (45)$$

$$A = a_f^2\cosh(2\phi_f) + \sinh^4(\phi_f)(a_f^2\cosh^2(\phi_f) - 1) - \cosh^4(\phi_f)(1 - a_f^2\sinh^2(\phi_f)). \quad (46)$$

Here B is always positive, while the sign of A depends on the boundary conditions. The sine in Eq. (44) will be positive as long as $\theta = \text{arccot}(\frac{A}{B}) < \frac{2\pi}{3}$, which implies $\cot(\theta) = \frac{A}{B} > -\frac{1}{\sqrt{3}}$, since the cotangent function is monotonically decreasing between 0 and π. Therefore, when $A/B > -1/\sqrt{3}$, the sine of $3\theta/2$ is positive and the two closed saddle points have the biggest amplitude, while when $A/B < -1/\sqrt{3}$, the sine is negative and the two unclosed saddle points have the largest amplitude. In Fig. 5 we plot the region in phase space (a_f, ϕ_f) where $A/B > -1/\sqrt{3}$, and we find that it englobes the region
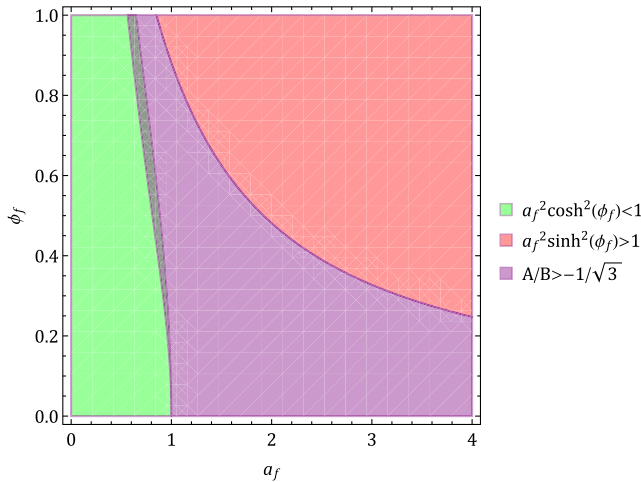

FIG. 5. The region where the closed saddle points dominate over the unclosed one is englobing the region where complex saddle points are defined. Note that our analysis assumed that we were outside the green and red regions, so from this picture nothing can be inferred about these regions.

where complex saddle point geometries are defined. Near the boundaries of this region, $a_f^2\cosh^2(\phi_f) = 1$ or $a_f^2\sinh^2(\phi_f) = 1$, the weightings approach each other, but in the interior the unclosed saddle points are exponentially suppressed compared to the closed saddle points. Thus we conclude that also in the case where a scalar field is included, the no-boundary wave function is dominated by compact, regular geometries.

Still, the interpretation of the wave function requires more refinements. As it stands, the wave function is a sum of four saddles, which are pairwise complex conjugates of each other. Thus the sum is dominated not by one, but by two saddles of equal weight. Taken at face value, this would mean that there would be a strong interference between two universes. However, it was shown in [29,30] that once perturbations are added, they lead to efficient decoherence of the two saddles as the Universe grows. Thus, once $a_f \gg 1$ we may focus on a single (dominant) saddle point, say $N_+[\gamma_-]$. For this saddle point, the action may be expanded in the scalar field (which is required to be small for these saddle points to exist), with the result that

$$S(N_+[\gamma_-]) = -\frac{1}{3}(a_f^2 - 1)^{3/2} - \frac{1}{2}(a_f^2 - 1)^{1/2}a_f^2\phi_f^2$$
$$- \frac{i}{3}\left(1 - \frac{3}{2}a_f^2\phi_f^2\right) + \mathcal{O}(\phi_f^4). \quad (47)$$

To quadratic order in the final value of the scalar field, the weighting is thus given by

$$\left|e^{\frac{i}{\hbar}S(N_+[\gamma_-])}\right| = e^{\frac{1}{3}\left(1 - \frac{3}{2}a_f^2\phi_f^2\right)}, \qquad a_f\phi_f \ll 1, \quad (48)$$

recovering the well-known result that the no-boundary wave function gives the highest probability to histories that evolve low on the scalar potential. The on-shell action (47) further implies that at large scale factor we have to leading order

$$\left| \frac{\partial \mathrm{Im}(S)}{\partial a_f} \Big/ \frac{\partial \mathrm{Re}(S)}{\partial a_f} \right| \sim \frac{\phi_f^2}{a_f} \ll 1, \quad \left| \frac{\partial \mathrm{Im}(S)}{\partial \phi_f} \Big/ \frac{\partial \mathrm{Re}(S)}{\partial \phi_f} \right| \sim \frac{1}{a_f} \ll 1.$$
$$(49)$$

Thus, at large scale factor, the amplitude of the wave function varies very slowly compared to the phase. This demonstrates that this branch of the wave function becomes of WKB form, which implies that in this region of phase space a classical spacetime, with a classical background evolution, is predicted [15,23]. Finally, we should note that the highest probability occurs for the limit where the scalar sits at the minimum of the $\cosh(2\phi)$ potential. In this limit, γ tends to zero and the unclosed saddles disappear. This limit in fact simply corresponds to the pure gravity case with a cosmological constant, except that now fluctuations of the scalar field are also included. These, however, provide a subdominant contribution to the wave function.

2. Region II: Imaginary saddle points

In region II, where roughly speaking the Universe is very small [$a_f^2 \cosh(2\phi_f) < 1$], the solutions for γ are purely real. Therefore the saddle point lapse values (28) will be

purely imaginary, as well as the initial momenta Π_x and Π_y. Then the classical solutions $\bar{x}$ and $\bar{y}$ are purely real, see (17), (18). If $\bar{a} = (\bar{x}^2 - \bar{y}^2)^{1/4}$ is also real, then we obtain a purely Euclidean geometry. Moreover the on-shell action evaluated on the saddle points is imaginary from (24). This means that the wave function of each saddle point geometry is a pure amplitude:

$$\psi \propto e^{\text{real number}}, \quad (50)$$

hence this does not lead to a WKB evolution. In this region of phase space, the Universe is still in the nucleation phase, roughly the equivalent to the tunneling phase of a particle under a barrier potential.

We will now look at the saddle point solutions in a little more detail, see Fig. 6 for a numerical example. For the closed saddle point solutions, $\bar{x}$ and $\bar{y}$ both start from 0 when $\tau = 0$. $\bar{x}$ describes a downward-facing parabola (because N^2 is negative), while $\bar{y}$ is a straight line. The gradient of $\bar{x}$ in $\tau = 0$ is always bigger than that of $\bar{y}$:

$$\left| \frac{\dot{\bar{y}}}{\dot{\bar{x}}} \right|_0 = |\tanh(2\gamma)| < 1, \quad (51)$$

and finally at the end point of the trajectory $\tau = 1$, the value of $\bar{x}$ is always above the value of $\bar{y}$ [since $\cosh(2\phi_f) > \sinh(2\phi_f)$]. Hence $\bar{x}$ is always larger than $\bar{y}$, and $\bar{a}$ will be real for the whole trajectory. So, the geometry of the closed saddle points is purely Euclidean.
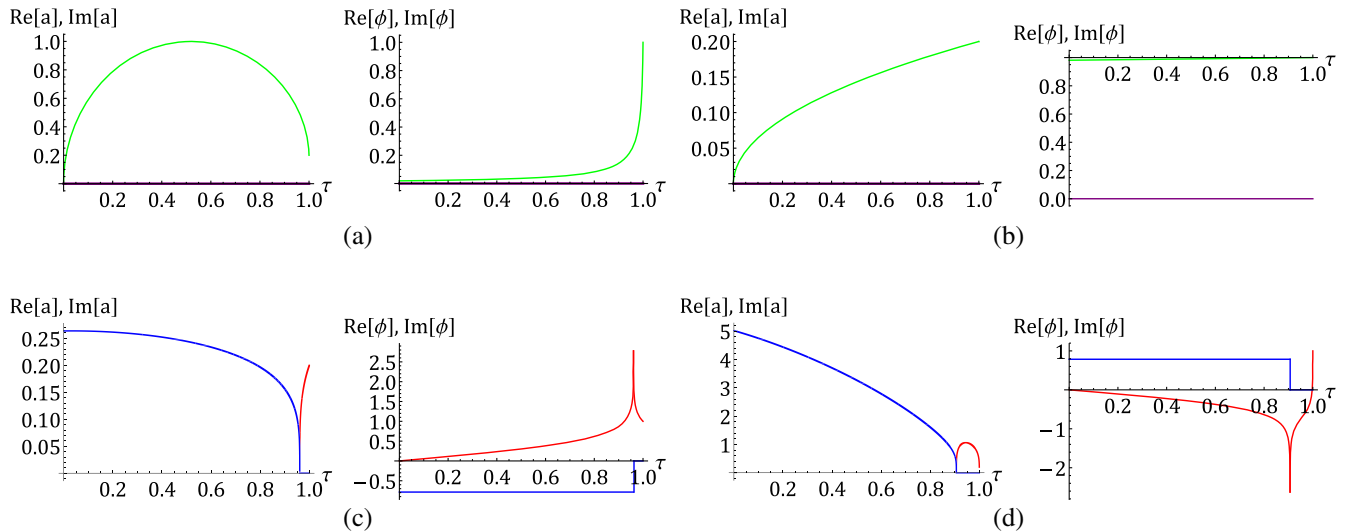

FIG. 6.   A numerical example of the field evolutions in region II. Here we have chosen $\alpha = 1, \beta = 0, a_f = 0.2, \phi_f = 1$. For this example $\gamma_- = 0.0189, \gamma_+ = 0.981, N_-[\gamma_-] = -1.92i, N_+[\gamma_-] = -0.0782i, N_+[\gamma_+] = -0.0208i,$ and $N_-[\gamma_+] = -7.24i$. We find two closed saddle point geometries—$N_+[\gamma_+]$ and $N_-[\gamma_-]$—and two unclosed geometries—$N_+[\gamma_-]$ and $N_-[\gamma_+]$. Real parts in green for closed saddle points, red for unclosed ones. Imaginary parts in purple for closed saddle points and blue for unclosed ones. Note that the imaginary part of the scalar typically jumps when the scale factor passes through zero, as the imaginary part of the scalar itself also changes there, given that it is obtained by taking a fourth root $a = (x^2 - y^2)^{1/4}$. (a) Closed euclidean geometry $N_-[\gamma_-]$, (b) closed euclidean geometry $N_+[\gamma_+]$, (c) unclosed euclidean geometry $N_+[\gamma_-]$, and (d) unclosed euclidean geometry $N_-[\gamma_+]$.

For the unclosed saddle points, only $\bar{x}$ starts at 0, so $\bar{x}^2 - \bar{y}^2$ will start with a negative value at $\tau = 0$. Because the end point value of $\bar{x}$ is always bigger than the end point of $\bar{y}$, this means that $\bar{x}^2 - \bar{y}^2$ will be positive at $\tau = 1$. Therefore by the mean value principle, there must be a $\tau_0 \in (0, 1)$ such that $\bar{a}(\tau_0) = 0$: these unclosed geometries exhibit a singular bounce in the course of their evolution. These solutions are thus unphysical, because fluctuations will blow up at this singularity, rendering the action infinite.

To examine which saddle points contribute to the path integral we must study the thimbles in the complexified N plane. A representative numerical example is shown in Fig. 7, with final values $a_f = 0.2$ and $\phi_f = 1$. In the figure the closed saddle points are marked with green dots. The interesting feature is that their locations relative to the unclosed saddles are reversed for the two values of γ. Hence, with the same contour of integration, we will necessarily get a contribution both from a closed and an unclosed saddle. If we stick to the same contour of integration as in region I, i.e. a line parallel to the real N axis, then for γ_- the unclosed saddle will contribute, while for γ_+ the relevant saddle will be the closed one,

$$\Psi[x_f, y_f] \approx e^{\frac{i}{\hbar} S_{\text{on-shell}}[N_+[\gamma_-]]} + e^{\frac{i}{\hbar} S_{\text{on-shell}}[N_+[\gamma_+]]}. \quad (52)$$

As explained above, the unclosed saddles contain a singular bounce and thus, once perturbations are added, their action will in fact diverge. Thus such saddle points essentially remove themselves from the wave function. We should note that, in [26], the point of view was taken that saddle points containing a singular bounce should be allowed, as one can deform the "path" taken in the complexified time plane as long as the end points remain unchanged, i.e. as long as the evolution starts at $\tau = 0$ and ends at $\int_0^1 N d\tau = N$. Such a deformed evolution could

then circumvent the singularity and render the field evolution regular. In view of Cauchy's theorem, nothing speaks against performing such deformations as long as the deformed path does not cross any singularities. Here, however, the original path actually does contain a singularity, at which all perturbations and matter configurations blow up—cf. for instance the scalar field evolution shown in Fig. 6(c). Thus Cauchy's theorem cannot be applied and we are led to disregard saddle point geometries that contain spacetime singularities. Note that we take the constant lapse version of the saddle point geometry, rather than one of the deformed ones, as fundamental, given that the path integral is evaluated precisely in the gauge where N is constant [31]. We acknowledge that this point of view is debatable, and that this is an issue that deserves a deeper study in the future. For now, let us add one further remark: the saddle point evolutions appear to be completely regular in terms of the x, y variables. One only sees the singularity after transforming back to the original a, ϕ variables. One could then argue that the x, y variables are better behaved and that these saddle points should therefore be retained. However, one should note that the transformation from the a, ϕ to the x, y variables then becomes ill defined. The fact that the a, ϕ fields have physical meaning—describing, respectively, the evolution of the volume of the Universe, and of the scalar field matter—provides in our view a strong argument for considering them as fundamental, and for discarding saddle points that are singular in these variables.

In light of the preceding discussion, with a single relevant saddle point remaining, the wave function may be approximated as

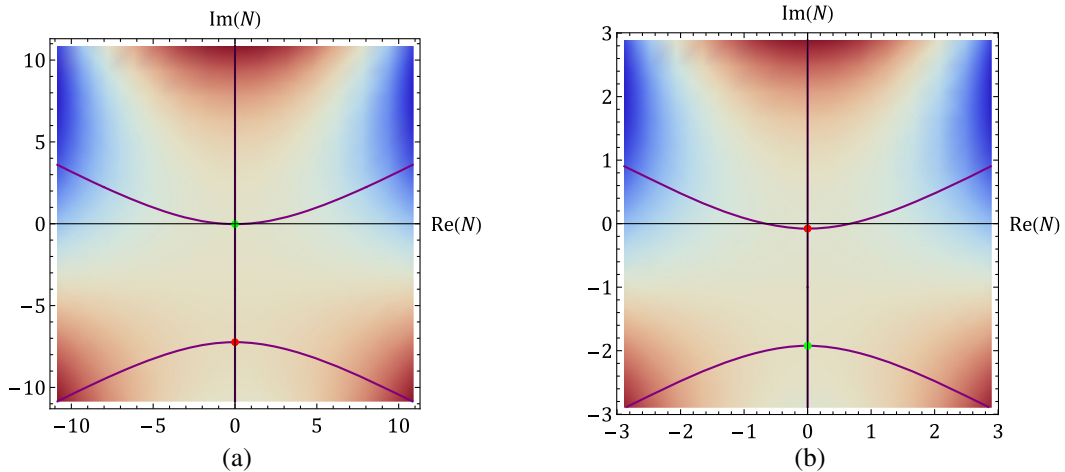$$\Psi \approx e^{\frac{i}{\hbar} S(N_+[\gamma_+])}, \quad (53)$$

FIG. 7. Case $\alpha = 1$, $\beta = 0$, $a_f = 0.2$, $\phi_f = 1$, with N imaginary. Density plot of the weighting $\text{Re}[iS_{\text{saddle}}^{\text{on-shell}}]$ and flow lines in the complex N plane. (a) $N_-[\gamma_-]$ (green dot) and $N_+[\gamma_-]$ (red dot) and (b) $N_+[\gamma_+]$ (green dot) and $N_-[\gamma_+]$ (red dot).

$$S(N_+[\gamma_+]) = -\frac{i}{3}[1 - (1 - a_f^2)^{3/2}]$$

$$+ \frac{i}{2}[1 - (1 - a_f^2)^{1/2}]a_f^2\phi_f^2 + \mathcal{O}(\phi_f^4). \quad (54)$$

The action is pure amplitude: it grows from zero at $a_f = 0$ until the boundary of the domain $a_f^2 \cosh(2\phi_f) < 1$ is reached. As the Universe grows further, a phase will develop, as described previously for region I.

It is noteworthy that of the two closed saddle points, it is the one whose geometry approximately corresponds to less than a hemisphere that is relevant to the path integral, rather than the one whose geometry is larger than half of an hemisphere, cf. also Fig. 6. This is in agreement with previous findings in the pure gravity case [10] and implies that when the scale factor a_f approaches zero, the dominant geometry is the vanishing geometry, rather than that of a full sphere, which would have signaled a dominant contribution from a nontrivial topology.

3. Region III: Real saddle points

In this region, $a_f^2 \sinh^2(2\phi) > 1$, the imaginary part of γ implies that $\cosh(2\gamma)$ and $\sinh(2\gamma)$ are purely imaginary. This means that the values of Π_x and Π_y as well as the classical solutions $\bar{x}$ and $\bar{y}$ are always purely real. Hence the saddle point values of the lapse (28) are also purely real, and the on-shell action (24) evaluated on these saddle points as well. Therefore whatever the relevant saddle points are, the wave function that we will get from the lapse integral will be of the following form:

$$\Psi[x_f, y_f] = \sum e^{i \cdot \text{real number}}, \quad (55)$$

which means that all the different saddle points have the same weighting, so this does not describe a true quantum state, but rather a collection of classical evolutions. However, we will now show that the classical solutions found in this region are not physical, because they exhibit a singular bounce in the course of their evolution. The $\bar{x}$ solution describes an upward facing parabola, while $\bar{y}$ is

still a straight line. For the closed saddle points, we start with $\bar{x}(0) = \bar{y}(0) = 0$, and the initial gradients are

$$\left|\frac{\dot{\bar{y}}}{\dot{\bar{x}}}\right|_0 = |\coth(\text{Re}[\gamma])| > 1, \quad (56)$$

therefore, since at the end point we still have $\bar{x}(1) > \bar{y}(1)$, there will be a τ_0 such that the x and y curve cross, and we find $a(\tau_0) = 0$. Closed geometries therefore contain a singular bounce, at which perturbations will blow up, and are therefore excluded as potential saddle points. For unclosed saddle points, we start from $x(0)^2 - y(0)^2 < 0$ and we go to $x(1)^2 - y(1)^2 > 0$, so again their must be a point $\tau_0 \in (0, 1)$ where the geometry vanishes, $a(\tau_0) = 0$, and we find another singular bounce.

In this region, we therefore do not find any physically acceptable solutions, but this is not a contradiction because this region is by definition not continuously connected to the pure gravity case $\phi_f \to 0$, so it just means that there exists a bound on the values of ϕ_f that lead to classical spacetime, and this bound is $a_f^2 \sinh^2(\phi_f) < 1$.

B. $\alpha = -1, \beta = 0$: $V(\phi) = -\cosh(2\phi)$

The negative $\cosh(2\phi)$ potential is interesting to study, as potentials of this kind appear rather naturally in string compactifications (though the potential would then be expected to be bounded below; i.e. it would turn upwards at large $|\phi|$). At the maximum of the potential, where the scalar field vanishes, we expect to recover the pure anti–de Sitter case, studied in [8,10]. In these works it was found that the AdS path integral is formally completely analogous to a no-boundary path integral, including the requirement for a regularity condition in the interior of the geometry.

Here, with a scalar included, we now have saddle points at

$$N_\pm^{\text{saddle}} = i\cosh(2\gamma) \pm \sqrt{-\cosh^2(2\gamma) - a_f^2\cosh(2\phi_f)}. \quad (57)$$

The γ values that lead to closed geometries are given by (see Appendix B)

$$\begin{cases} \gamma_{\mp}^{\text{I}} = \mp\frac{1}{2}\cosh^{-1}\left(\sqrt{1 + \frac{\sinh^2(2\phi_f)}{4}(2 + a_f^2\cosh(2\phi_f) - \sqrt{(2 + a_f^2\cosh(2\phi_f))^2 - a_f^4})}\right), \\ \gamma_{\mp}^{\text{II}} = \mp\frac{1}{2}\cosh^{-1}\left(\sqrt{1 + \frac{\sinh^2(2\phi_f)}{4}(2 + a_f^2\cosh(2\phi_f) + \sqrt{(2 + a_f^2\cosh(2\phi_f))^2 - a_f^4})}\right). \end{cases} \quad (58)$$

The solutions γ_-^{I} and γ_+^{II} are defined for $\phi_f > 0$, while γ_+^{I} and γ_-^{II} are defined for $\phi_f < 0$. The solutions I close the saddle point N_+, and the solutions II close the saddle point N_-. At each phase space point, i.e. for each choice of final

boundary conditions (a_f, ϕ_f), we get a total of four saddle point solutions, two closed ($N_+[\gamma^{\text{I}}]$ and $N_-[\gamma^{\text{II}}]$) and two unclosed ($N_+[\gamma^{\text{II}}]$ and $N_-[\gamma^{\text{I}}]$). It is straightforward to see that all γ values are real, and thus the saddle points (57) are
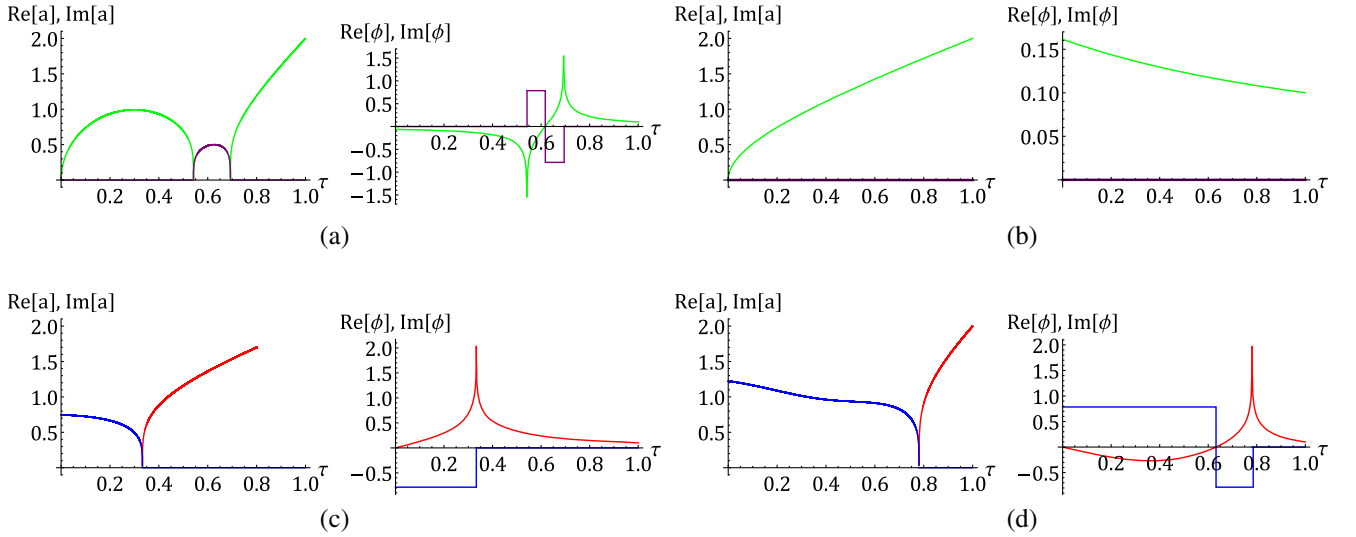
FIG. 8. Case $\alpha = -1, \beta = 0, a_f = 2, \phi_f = 0.1$, with N imaginary. The relevant numerical values are $\gamma^{\mathrm{I}}_- = -0.0615, N_+[\gamma^{\mathrm{I}}_-] = 3.26i, N_-[\gamma^{\mathrm{I}}_-] = -1.25i; \gamma^{\mathrm{II}}_+ = 0.162, N_-[\gamma^{\mathrm{II}}_+] = -1.23i, N_+[\gamma^{\mathrm{II}}_+] = 3.33i$. Geometry associated to the different saddle points. Same conventions as in Fig. 3. (a) Closed euclidean geometry $N_+[\gamma^{\mathrm{I}}_-]$, (b) closed euclidean geometry $N_-[\gamma^{\mathrm{II}}_+]$, (c) unclosed euclidean geometry $N_-[\gamma^{\mathrm{I}}_-]$, and (d) unclosed euclidean geometry $N_+[\gamma^{\mathrm{II}}_+]$.

pure imaginary. Hence we may deduce from (24) that the on-shell action will also be purely imaginary: e^{iS} is only an amplitude and does not contain any phase. Thus no classical spacetime will be implied. Examples of specific solutions are shown in Fig. 8. The two saddle points associated with γ^{I}, as well as the unclosed saddle associated to γ^{II} have a singular bounce within their evolution, they are therefore suppressed by the blowing-up fluctuations. Using similar techniques as for the $\alpha = 1$ case, one can show this is always the case. Therefore the only relevant solution is the closed saddle point geometry $N_-[\gamma^{\mathrm{II}}_+]$. Note that for this saddle point, the scalar field is purely real valued, and in fact runs up the potential as the scale factor grows. This is

characteristic of approximately anti–de Sitter space, where even potential maxima with negative second derivatives are stable, as long as they satisfy the Breitenlohner-Freedman bound [32], as our potential does (if one undoes the rescaling (4), then $V = -\cosh\left(\sqrt{\frac{2}{3}}\Phi\right)$ so that the effective mass is $m^2 = -\frac{2}{3} > m^2_{BF} = -\frac{9}{4}$).

The structure of the flow lines is again most easily grasped with the help of a numerical example, see Fig. 9. The closed saddle point is selected by Picard-Lefschetz theory only if the defining contour of integration contains the lower part of the imaginary lapse axis. In fact, given that we expect the wave function to be real valued [especially in
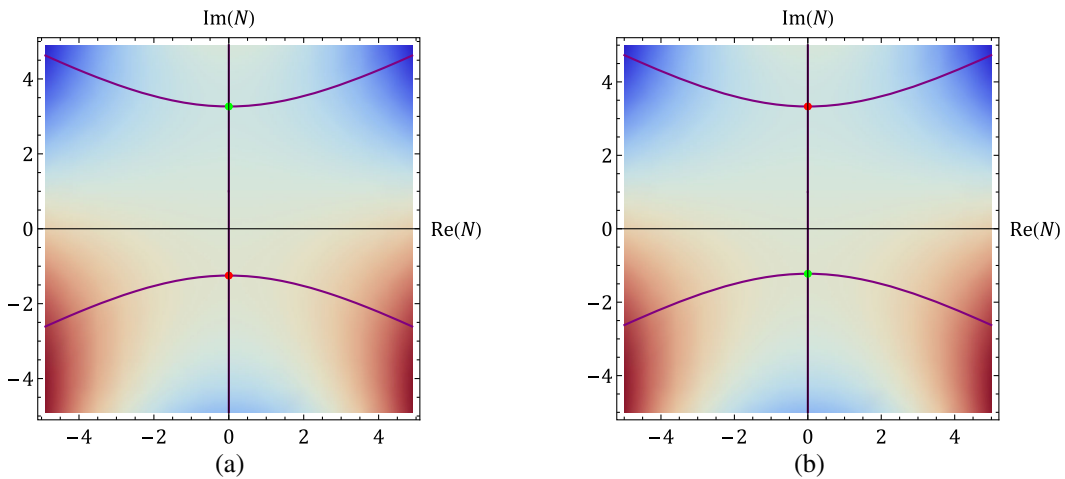

FIG. 9. Case $\alpha = -1, \beta = 0, a_f = 2, \phi_f = 1$. Density plot of the weighting $\mathrm{Re}[iS^{\mathrm{on\text{-}shell}}_{\mathrm{saddle}}]$ and flow lines in the complex N plane. Same conventions as in Fig. 4. (a) $N_+[\gamma^{\mathrm{I}}_-]$ (green dot) and $N_-[\gamma^{\mathrm{I}}_-]$ (red dot) and (b) $N_-[\gamma^{\mathrm{II}}_+]$ (green dot) and $N_+[\gamma^{\mathrm{II}}_+]$ (red dot).
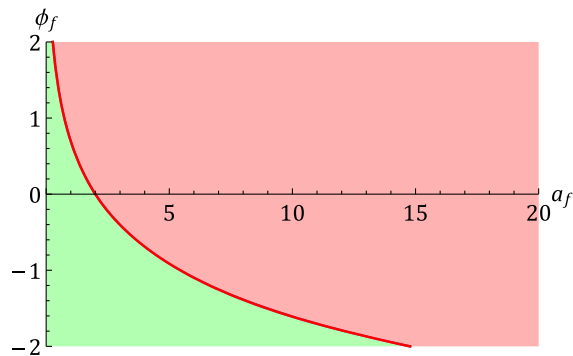
FIG. 10. Phase space regions delimited by the condition $a_f^2 \frac{\alpha}{2} e^{2\phi_f} = 1$ (red curve), with $\alpha = 1/2$. In the red region, the saddle point value for the lapse is purely real. In the green region, the saddle point value for the lapse is purely imaginary.

view of a potential conformal field theory (CFT) dual [33]], one should define the path integral as a sum of two contours: the first from negative imaginary infinity up to the saddle point in the upper half plane and then on to either the left or right, following the steepest descent contour; plus a contribution from a contour that is reflected across the imaginary lapse axis. These two contours will add up two complex conjugate contributions, just as was described for the pure gravity case in [8,10]. The wave function may then be estimated as

$$\Psi \approx e^{\frac{i}{\hbar} S(N_-[\gamma_+^{II}])}, \qquad (59)$$

$$S(N_-[\gamma_+^{II}]) = -\frac{i}{3}[(a_f^2 + 1)^{3/2} - 1]$$
$$- \frac{i}{2}[(a_f^2 + 1)^{1/2} - 1]a_f^2 \phi_f^2 + \mathcal{O}(\phi_f^4). \quad (60)$$

For completeness, let us mention that in applications to AdS/CFT one would let the scale factor run to infinity, $a_f \to \infty$, then add the appropriate counter terms to remove the volume divergence in the wave function.

We may conclude that the prescription to define the path integral with a momentum condition in the interior, rather than a Dirichlet condition, also works once a scalar field is included.

C. $\alpha = \beta$: $V(\phi) = \alpha \exp(2\phi)$

When $\alpha = \beta$, the potential becomes even simpler, and is given by a single exponential $V(\phi) \propto e^{2\phi}$. Classically, this potential allows for inflationary scaling solutions, with a constant equation of state [26]. However, because the potential is nowhere flat, it is not possible to achieve a de Sitter limit. The classical scaling symmetry, which consists of a rescaling of the metric accompanied by a shift in the scalar field, explains why the analysis is simplified for this potential. For instance, we are left with a unique expression for the saddle point, cf. (25),

$$N_{\text{saddle}} = -i \frac{a_f^2}{2} \exp(2\phi_f - 2\gamma). \qquad (61)$$

We find that the values of γ closing the saddle point geometry in this case are given by

$$\begin{cases} \gamma_{(n)}^{I} = \frac{1}{4} \ln \left(\frac{1 - a_f^2 \frac{\alpha}{2} e^{2\phi_f}}{e^{-4\phi_f}} \right) + in \cdot \frac{\pi}{2} & \text{for } n \in \{0, 1\}, \\ \gamma_{(n)}^{II} = \frac{1}{4} \ln \left(\frac{a_f^2 \frac{\alpha}{2} e^{2\phi_f} - 1}{e^{-4\phi_f}} \right) + i \left(\frac{\pi}{4} + n \cdot \frac{\pi}{2} \right) & \text{for } n \in \{0, 1\}, \end{cases} \qquad (62)$$

and their regions of validity, represented in Fig. 10 for $\alpha = 1/2$, are

$$\text{region I: } a_f^2 \frac{\alpha}{2} e^{2\phi_f} < 1, \qquad \text{region II: } a_f^2 \frac{\alpha}{2} e^{2\phi_f} > 1. \quad (63)$$

Note that in this case there is only one saddle point for each value of γ, and it is closed. This is a special case of Picard's little theorem, see Appendix C. Let us study the geometries of these solutions.

In region I the solution for γ implies that $\cosh(2\gamma)$ and $\sinh(2\gamma)$ are both real, so that the momenta Π_x and Π_y are purely imaginary. The saddle point values for the lapse (61) are thus also purely imaginary, as are the on-shell actions evaluated on these saddle point values. The wave function associated to these geometries is a pure amplitude, without phase. Thus no classical spacetime is predicted in this region.

In region II, the solution for γ instead leads to $\cosh(2\gamma)$ and $\sinh(2\gamma)$ being pure imaginary, so that the momenta Π_x and Π_y are real. The saddle point values, and the on-shell actions, are consequently real. In principle, we would then expect the wave function to represent a collection of classical solutions, all with equal weighting. However, as was also already noted in [26], the geometries all contain a singular bounce. One may see this as follows: the background solutions $\bar{x}$ and $\bar{y}$ both start at zero and are real valued, with $\bar{x}$ being a parabola with a local minimum, while $\bar{y}$ is a parabola with a local maximum. Given that at the final time $\bar{x}(1) > \bar{y}(1)$, there necessarily exists an intermediate time at which $|\bar{x}| = |\bar{y}|$, where the scale factor vanishes. At this time all matter fields and perturbations blow up and cause a divergence in the local curvature as well as the action. As explained in Sec. III A, we exclude such solutions and in fact find that for the pure exponential potential, no regular no-boundary solutions exist which would predict a classical evolution of the fields.

D. $\alpha = 0$, $\beta = 1$: $V(\phi) = \sinh(2\phi)$

Finally, we will look at the case $\alpha = 0$, $\beta = 1$, i.e. at the pure $\sinh(2\phi)$ potential. At large $\phi_f \gg 1$, we expect to recover the results of the $\cosh(2\phi)$ potential, where we did not find instantons that would predict a classical spacetime. However, at small ϕ we expect significant differences, as,
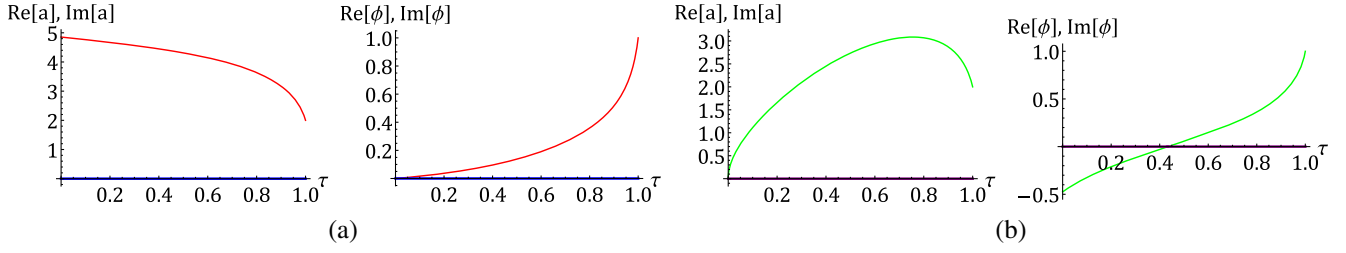
FIG. 11. Case $\alpha = 0$, $\beta = 1$, $a_f = 2$, $\phi_f = 1$, with N imaginary. The relevant numerical values are $\gamma_{(0)}^I = -0.475, N_+[\gamma_{(0)}^I] = 2.87i, N_-[\gamma_{(0)}^I] = -5.06i$. Geometry associated to the different saddle points. Same conventions as in Fig. 3. (a) Unclosed euclidean geometry $N_+[\gamma_{(0)}^I]$ and (b) closed euclidean geometry $N_-[\gamma_{(0)}^I]$.

just as for the exponential potential, there is no minimum where the scalar field can sit. Then at negative ϕ, the potential is unbounded below and quite steep, but not steep enough to allow for ekpyrotic solutions. Thus there is no region in this potential which obviously allows for useful/realistic instanton solutions. This expectation will be confirmed by the detailed analysis that follows.

The saddle point values (25) of the lapse here read

$$N_\pm^{\text{saddle}} = i \sinh(2\gamma) \pm \sqrt{-\sinh^2(2\gamma) - a_f^2 \sinh(2\phi_f)}. \quad (64)$$

Using once again similar techniques to those used in Appendix B, we find the following values of γ that can potentially close these saddle point geometries:

$$\begin{cases} \gamma_{(n)}^I = -\frac{1}{2}\cosh^{-1}\left(\frac{\cosh(2\phi)}{2}\sqrt{2 - a_f^2 \sinh(2\phi_f) + \sqrt{a_f^4 + (2 - a_f^2 \sinh(2\phi_f))^2}}\right) + i\frac{n\cdot\pi}{2} \quad \text{with} \quad n \in \{0, 1\}, \\ \gamma_{(n)}^{II} = \frac{1}{2}\cosh^{-1}\left(\sqrt{1 - \frac{\cosh^2(2\phi_f)}{4}(2 - a_f^2 \sinh(2\phi_f) - \sqrt{a_f^4 + (2 - a_f^2 \sinh(2\phi_f))^2})}\right) + i\left(\frac{\pi}{4} + \frac{n\cdot\pi}{2}\right), \quad n \in \{0, 1\}. \end{cases} \quad (65)$$

These expressions are both defined on the whole of phase space.

The solutions γ^I have $\text{Im}[\gamma^I] = n \cdot \pi/2$, therefore $\cosh(2\gamma^I) = (-1)^n \cosh(2\text{Re}[\gamma^I]) \in \mathbb{R}$ and $\sinh(2\gamma^I) = (-1)^n \sinh(2\text{Re}[\gamma^I]) \in \mathbb{R}$. Hence the momenta Π_x and Π_y

FIG. 12. Case $\alpha = 0$, $\beta = 1$, $a_f = 2$, $\phi_f = 1$. Density plot of the weighting $\text{Re}[iS_{\text{saddle}}^{\text{on-shell}}]$ and flow lines in the complex N plane. Same conventions as in Fig. 4. $N_-[\gamma_{(0)}^I]$ (green dot) and $N_+[\gamma_{(0)}^I]$ (red dot).

are purely imaginary and the saddle point values of the lapse also. As before, we impose that $\text{Im}(\Pi_x) > 0$ and this eliminates the values $\gamma_{(1)}^I$. The geometries are Euclidean, and the remaining ones are shown in Fig. 11. We once again obtain a closed and an unclosed geometry. The associated flow lines are shown in Fig. 12. One needs to define the integral on a contour that runs parallel to the real N line, but this time one needs to shift it into the lower half plane for convergence [this is because $\alpha = 0$, cf. Eq. (24)]. The relevant saddle point is then the closed saddle at $N_-[\gamma_{(0)}^I]$, and the wave function can be approximated as

$$\Psi \approx e^{\frac{i}{\hbar}S(N_-[\gamma_{(0)}^I])}, \quad (66)$$

$$S(N_-[\gamma_{(0)}^I]) = -\frac{i}{2}[3a_f^2(\sqrt{a_f^4 + 4} + 2)^{1/2} - (\sqrt{a_f^4 + 4} - 2)^{3/2}] + \frac{i}{2}(\sqrt{a_f^4 + 4} - 2)^{1/2}a_f^2\phi_f + \mathcal{O}(\phi_f^2). \quad (67)$$

The wave function is pure amplitude. This could describe the nucleation of the Universe at small scale factor, but there is no follow-up saddle point that could describe the

subsequent evolution of the Universe. Hence we conclude that these saddle points do not predict the emergence of classical spacetime. In fact, the wave function in this case resembles more the quasi-AdS case of Sec. III B. From this point of view, it is understandable that the scalar field starts at negative values (for the closed saddle), but then, to finally reach positive values of the potential, the Universe recollapses, which allows the scalar to run up to positive values, see again Fig. 11. It would be interesting to see whether the present solutions have an application in finite bulk extensions of AdS/CFT [34].

For the solutions γ^{II}, we have $\text{Im}[\gamma^{II}] = \pi/4 + n \cdot \pi/2$, so $\cosh(2\gamma^{II}) = i(-1)^n \sinh(2\text{Re}[\gamma^{II}]) \in i\mathbb{R}$ and $\sinh(2\gamma^{II}) = i(-1)^n \cosh(2\text{Re}[\gamma^{II}]) \in i\mathbb{R}$. The momenta Π_x and Π_y, as well as the saddle point values of the lapse are purely real. However, using analogous arguments to those used in previous paragraphs, one may straightforwardly show that all these saddle point geometries contain a singular bounce. For the same reasons as those mentioned previously, namely that such geometries lead to divergences of the curvature and the action once perturbations or additional matter fields are included, we deem such solutions to be unphysical and discard them. Thus we are unfortunately left without any complex saddle points, so that for this potential we conclude that the no-boundary wave function does not predict the emergence of classical evolution. In some sense, this is a very good example showing the predictive power of the no-boundary proposal: only in suitably flat regions of the potential does the no-boundary wave function exhibit classical (inflationary) spacetime evolution (while in sufficiently steep and negative regions of the potential the same could be achieved via ekpyrotic no-boundary instantons [35,36]). This is thus a good example showing that the classicality of the early Universe should be seen as an important clue as to its origin.

IV. DISCUSSION

In this paper we have been able to show, in the context of a minisuperspace model, that the no-boundary proposal can be defined as a sum over regular metrics and scalar field configurations. The regularity condition has the advantage that in the sum over metrics, one is considering metrics on which perturbations are damped, rather than enhanced. This is a great advantage compared to defining the no-boundary path integral as a sum over compact geometries, since in the latter case one is forced to include unstable configurations too. The price to pay is that a sum over regular metrics does not guarantee that the geometries are also closed. Off-shell geometries are certainly not closed in general, and in fact the best one can hope for is that the saddle points are closed. As we have demonstrated, even this does not occur in full generality: alongside saddle points that are both regular and compact, we found saddle points that are unclosed. Moreover, we argued that such unclosed saddle points cannot be avoided. However, they

were always subdominant, so that the semiclassical no-boundary wave function ended up being dominated by a closed, regular geometry with a regular scalar field configuration, as originally intended.

Let us add a few words on integration contours, because in defining the wave function, it is not enough to specify the boundary conditions, one also has to specify the contours of integration. The two issues are however connected: our boundary conditions contain a condition on the expansion rate of the Universe and, as we discussed, this expansion rate must be imaginary and with the correct sign in order to enforce the appropriate Wick rotation. The fact that the expansion rate starts out being imaginary implies that the metrics we are summing over cannot be purely Lorentzian. That being said, what we find is that in cases where the scalar potential takes positive values, the integration contour must be taken to be "as Lorentzian as possible." More specifically, we may take the integration contour to be the line of real values of the lapse N, and this contour is deformable into the two Lefschetz thimbles, cf. Fig. 4. By contrast, when the potential is negative, such as the case discussed in Sec. III B, then the contour of integration can be taken to be a sum of two contours symmetric about the imaginary lapse line. Thus we recover a phenomenon first discussed in [10], namely that negative potentials tend to require sums over (near-)Euclidean geometries, while positive potentials require sums over (near-)Lorentzian geometries.

Our construction highlights a puzzle inherent in the no-boundary proposal. This puzzle is not new, but it has come into renewed focus here. This work showed that imposing regularity alone is insufficient to fully fix the initial conditions. It only fixes them up to a complex number γ which, as we have discussed, can often be interpreted as the initial value of the scalar field. (Thus, the same problem arises when attempting to define the no-boundary integral over compact metrics.) This number must be fixed such that the path integral may contain saddle points that are closed (for generic values of γ the saddle points will all be unclosed). The puzzle is then that the required values of γ depend on the final conditions; i.e. as the Universe grows and the scalar field evolves, γ necessarily has to change. When the wave function is of WKB form, γ changes less and less as the Universe grows [15], but nevertheless it is not constant. This is as nonlocal/noncausal as it gets: at every moment in the history of the Universe, it suggests that the Universe must "recalculate" its entire wave function from the very beginning of space and time. The conclusion we can draw from this observation is that either quantum gravity really is this drastically nonlocal, or there is something wrong with the prescription that we have outlined.

We can conceive of one possible way in which the prescription above could be refined, such that the puzzle disappears: this is to require the scalar field to always take real values. A motivation for such a condition stems from the recent discussions of allowable metrics, where the criterion

of allowability is derived from demanding convergence of path integrals over real scalars and p-form fields, see the works of Kontsevich-Segal [37] and Witten [38], based on earlier work by Louko-Sorkin [39] (consequences for minisuperspace models started being explored in [40,41]). For positive potentials, the only place where γ turned out to be real and where the saddle point geometries are regular and lead to a classical spacetime, was at the minimum of the $\cosh(2\phi)$ potential. More generally, the scalar field value is both real and constant at extrema of the scalar potential. Thus, at such locations of the potential, the initial conditions for the no-boundary wave function remain unchanged over time. For the negative potential studied in Sec. III B, we also found real, yet evolving, values of the scalar field. However, the corresponding geometries were Euclidean, and in Euclidean space there is no notion of causality, so that the puzzle disappears in any event.

What would a restriction to extrema of the scalar potential imply for observations? First, we should note that the no-boundary wave function scales roughly as $e^{1/(\hbar V(\phi))}$, implying that low values of the potential come out as preferred. Thus, in a theory with many scalar fields (as expected in string theoretic models, for instance), the most likely configuration would be for all scalars to reside in the lowest minima of the potential. This is in fact a desirable feature, as it can help explain why all coupling constants in our Universe are found not to vary over time. However, with such a configuration, the Universe is also left empty. If one scalar instead starts out at a local maximum of the potential, then there is the possibility for this scalar to drive an inflationary phase during which density perturbations may also be generated, and structure may form in the Universe. Due to quantum fluctuations, the scalar field would then also eventually roll off the potential maximum, allowing inflation to end. Of course, to make this precise and see if this scenario is truly viable, one would have to understand the precise structure of the scalar potential, and in particular one would have to find out if there exists a local maximum in which sufficiently large density perturbations can be generated, alongside potential minima that are sufficiently stable. This is a substantial unsolved open problem. However, what this discussion illustrates is that, given a suitable dynamical theory, the no-boundary proposal remains an appealing, consistent (currently the only?) candidate for a theory of initial conditions.

APPENDIX A: FLUCTUATION INTEGRALS

The Wheeler-DeWitt equation in "position" space is obtained from the constraint (11)

$$\left[\hbar^2(\partial_x^2 - \partial_y^2) - \frac{1}{4}(1 - \alpha x - \beta y)\right]\Psi = 0. \quad \text{(A1)}$$

Assuming that the integrations of x and y have already been performed implies that the wave function can be written as an ordinary integral over the lapse, but with a currently unknown measure factor $m(N)$,

$$\Psi[x_f, y_f] = \int \mathrm{d}N m(N) e^{iS_0^{\text{on-shell}}[x_f, y_f, N]/\hbar}. \quad \text{(A2)}$$

We abbreviate the on-shell (in x and y) action (24) simply by S, and x_f, y_f by x, y. Then

$$\hbar^2(\partial_x^2 - \partial_y^2)\Psi = \int \mathrm{d}N m(N)\hbar^2 \left(\frac{iS_{,xx}}{\hbar} - \frac{(S_{,x})^2}{\hbar^2} - \frac{iS_{,yy}}{\hbar} + \frac{(S_{,y})^2}{\hbar^2}\right)e^{iS/\hbar}. \quad \text{(A3)}$$

Starting from (24), the explicit expressions are given by

$$S_{,x} = -\frac{\alpha N}{2} - \frac{\Pi_x}{2}, \qquad S_{,xx} = 0; \qquad S_{,y} = -\frac{\beta N}{2} - \frac{\Pi_y}{2}, \qquad S_{,yy} = 0, \quad \text{(A4)}$$

$$\Rightarrow (S_{,y})^2 - (S_{,x})^2 = \frac{1}{4}(N^2(\alpha^2 - \beta^2) - 2N(\alpha\Pi_x + \beta\Pi_y) + \Pi_y^2 - \Pi_x^2), \quad \text{(A5)}$$

$$S_{,N} = \frac{N^2(\alpha^2 - \beta^2)}{2} + N(\alpha\Pi_x + \beta\Pi_y) - \frac{(\alpha x + \beta y)}{2}, \quad \text{(A6)}$$

$$\Rightarrow \text{if } \Pi_y^2 - \Pi_x^2 = 1, (S_{,y})^2 - (S_{,x})^2 = -\frac{S_{,N}}{2} + \frac{(1 - \alpha x - \beta y)}{4}. \quad \text{(A7)}$$

Inserting into the WDW equation we obtain

$$0 = \int dN m(N) \cdot \left[-\frac{S_{,N}}{2} \right] e^{iS/\hbar} \Leftrightarrow 0 = \int dN \left(-\frac{i\hbar}{2} \right) m_{,N} e^{iS/\hbar} \Leftrightarrow m = \text{const.}, \tag{A8}$$

which verifies that, with mixed Neumann-Dirichlet boundary conditions, the fluctuation integrals lead to no additional N dependence in the integrand of the lapse integral.

APPENDIX B: SOLVING FOR γ

In this appendix, we will present the details on how to solve for the initial conditions parameter γ, under the condition that a closed saddle point geometry should exist. We will restrict the analysis to the $\cosh(2\phi)$ potential ($\alpha = \pm 1, \beta = 0$). For the other potentials, the analysis is analogous.

$\alpha = +1$ *and* $\beta = 0$

We need to solve the following equation for $\gamma \in \mathbb{C}$:

$$a_f^2 \sinh^2(2\phi_f) + 4 \sinh(2\gamma) \sinh(2\gamma - 2\phi_f) = 0. \tag{B1}$$

The real and imaginary parts of this equation must both hold separately. The imaginary part yields

$$\sin(4\text{Im}[\gamma]) \sinh(4\text{Re}[\gamma] - 2\phi_f) = 0, \tag{B2}$$

so we need either $\text{Im}[\gamma] = k \cdot \pi/4$ for $k \in \mathbb{Z}$, or $\text{Re}[\gamma] = \phi_f/2$. We now analyze the real part of Eq. (30) for these different cases.

(1) $\text{Re}[\gamma] = \phi_f/2$: then the real part of (30) becomes

$$\sin^2(2\text{Im}[\gamma]) = (a_f^2 \cosh^2(\phi_f) - 1) \sinh^2(\phi_f). \tag{B3}$$

Solutions to this equation only exist if the right-hand side is between 0 and 1, which translates into

$$a_f^2 \cosh^2(\phi_f) > 1 \quad \text{and} \quad a_f^2 \sinh^2(\phi_f) < 1. \tag{B4}$$

When these conditions are satisfied, Eq. (B3) possesses four families of solutions

$$\begin{cases} \gamma^{1,2(k)} = \frac{\phi_f}{2} \mp \frac{i}{2} (\arcsin(|\sinh(\phi_f)| \sqrt{a_f^2 \cosh^2(\phi_f) - 1}) + 2\pi \cdot k), \\ \gamma^{3,4(k)} = \frac{\phi_f}{2} \mp \frac{i}{2} (\pi - \arcsin(|\sinh(\phi_f)| \sqrt{a_f^2 \cosh^2(\phi_f) - 1}) + 2\pi \cdot k). \end{cases} \quad \forall \ k \in \mathbb{Z}. \tag{B5}$$

We still need to make sure that we obtain only saddle points that admit stable fluctuations, i.e. which imply the correct orientation for the Wick rotation. This means that we must get rid of spurious solutions for which $a\dot{a}|_0 \to -iN$ (rather than $a\dot{a}|_0 \to +iN$) when $\phi \to 0$; i.e. we only keep solutions where

$$\Pi_x = \frac{\dot{x}}{2N} \Big|_0 = \frac{a\dot{a}\cosh(2\phi) + a^2 \dot{\phi} \sinh(2\phi)}{N} \Big|_0 \xrightarrow{\phi \to 0} \frac{a\dot{a}}{N} \Big|_0 = +i. \tag{B6}$$

Writing Π_x in terms of γ, we obtain

$$\Pi_x = i \cosh(2\gamma) = i(\cosh(\phi_f) \cos(2\,\text{Im}[\gamma]) + i\,\sinh(\phi_f)\,\sin(2\,\text{Im}[\gamma])). \tag{B7}$$

Expanding this expression at zeroth order in ϕ_f, the second term vanishes and we find, by plugging in the solutions (B5) for γ,

$$\begin{cases} \Pi_x^{1,2(k)} = i \cos(\mp \arcsin(|\sinh(\phi_f)| \sqrt{a_f^2 \cosh^2(\phi_f) - 1}) + 2\pi \cdot k) \xrightarrow{\phi_f \to 0} + i, \\ \Pi_x^{3,4(k)} = i \cos(\mp \pi \pm \arcsin(|\sinh(\phi_f)| \sqrt{a_f^2 \cosh^2(\phi_f) - 1}) + 2\pi \cdot k) \xrightarrow{\phi_f \to 0} - i. \end{cases} \tag{B8}$$

We therefore exclude the last two families of solutions, $\gamma^{3,4(k)}$, and keep only the first two, $\gamma^{1,2(k)}$. Since γ only enters expressions through its hyperbolic sine or cosine, the value of k is unimportant and we can simply set it to 0. In the end we only have two values for γ:

$$2\gamma_{\mp} = \phi_f \mp i \arcsin(|\sinh \phi_f| \sqrt{a_f^2 \cosh^2(\phi_f) - 1}). \tag{B9}$$

The saddle point for the lapse computed from these γ values will be complex.

(2) $\mathrm{Im}[\gamma] = k \cdot \pi/2, k \in \mathbb{Z}$: $\sin(2\mathrm{Im}[\gamma]) = 0$ and $\cos(2\mathrm{Im}[\gamma]) = (-1)^k$. All expressions then only depend on whether k is odd or even, so we can restrict to $k = 0, 1$. The real part of the equation (30) yields

$$a_f^2 \sinh^2(2\phi_f) = 4\sinh(2\mathrm{Re}[\gamma])\sinh(2\phi_f - 2\mathrm{Re}[\gamma]). \tag{B10}$$

This admits a solution only if

$$a_f^2 \sinh^2(2\phi_f) < 4\left|\max_{\mathrm{Re}[\gamma]}(\sinh(2\mathrm{Re}[\gamma])\sinh(2\phi_f - 2\mathrm{Re}[\gamma]))\right| = 4\sinh^2(\phi_f),$$

$$\Leftrightarrow a_f^2 \cosh^2(\phi_f) < 1. \tag{B11}$$

When this condition is satisfied, we find two solutions for each k value:

$$\gamma_{\mp}^{(k)} = \frac{1}{2}\cosh^{-1}\left(\frac{1}{2}[4 + \sinh^2(2\phi_f)(2 - a_f^2\cosh(2\phi_f))\right.$$
$$\left.\mp \sinh^2(2\phi_f)\sqrt{(2 - a_f^2\cosh(2\phi_f))^2 - a_f^4}]^{1/2}\right) + ik \cdot \frac{\pi}{2}, \tag{B12}$$

which implies that

$$\Pi_{x\mp} = i\cosh(2\gamma_{\mp}) = \frac{i(-1)^k}{2}[4 + \sinh^2(2\phi_f)(2 - a_f^2\cosh(2\phi_f))$$
$$\mp \sinh^2(2\phi_f)\sqrt{(2 - a_f^2\cosh(2\phi_f))^2 - a_f^4}]^{1/2} \in i\mathbb{R}, \tag{B13}$$

$$\Pi_{y\mp} = i\sinh(2\gamma_{\mp}) = \frac{i(-1)^k}{2}|\sinh(2\phi_f)|[(2 - a_f^2\cosh(2\phi_f)) \mp \sqrt{(2 - a_f^2\cosh(2\phi_f))^2 - a_f^4}]^{1/2} \in i\mathbb{R}. \tag{B14}$$

We look at the limit of Π_x when $\phi_f \to 0$ to exclude spurious solutions, and we find

$$\Pi_{x\mp} \to i(-1)^k, \tag{B15}$$

thus we only keep $k = 0$ and we reject $k = 1$. Finally the γ values are

$$\gamma_{\mp} = \frac{1}{2}\cosh^{-1}\left(\sqrt{1 + \frac{\sinh^2(2\phi_f)}{4}((2 - a_f^2\cosh(2\phi_f)) \mp \sqrt{(2 - a_f^2\cosh(2\phi_f))^2 - a_f^4})}\right).$$

Because these γ values are real, the saddle points for the lapse will be purely imaginary.

(3) $\mathrm{Im}[\gamma] = \pi/4 + k \cdot \pi/2$: $\sin(2\mathrm{Im}[\gamma]) = (-1)^k$ and $\cos(2\mathrm{Im}[\gamma]) = 0$. All expressions will only depend on whether k is odd or even, so we can restrict to $k = 0, 1$. The real part of Eq. (30) becomes

$$a_f^2 \sinh^2(2\phi_f) - 4\cosh(2\mathrm{Re}[\gamma])\cosh(2\mathrm{Re}[\gamma] - 2\phi_f) = 0, \tag{B16}$$

which admits a solution only if

$$a_f^2 \sinh^2(2\phi_f) > 4\min_{\mathrm{Re}[\gamma]}(\cosh(2\mathrm{Re}[\gamma])\cosh(2\mathrm{Re}[\gamma] - 2\phi_f)) = 4\cosh^2(\phi_f),$$

$$\Leftrightarrow a_f^2 \sinh^2(\phi_f) > 1. \tag{B17}$$

This condition being fulfilled, we find the four following solutions:

$$\gamma_{\mp}^{(k)} = \mp \frac{1}{2} \cosh^{-1} \left(\frac{1}{2} \sinh(2\phi_f) \sqrt{a_f^2 \cosh(2\phi_f) - 2 \mp \sqrt{(a_f^2 \cosh^2(2\phi_f) - 2)^2 - a_f^4}} \right)$$
$$+ i \left(\frac{\pi}{4} + \frac{k \cdot \pi}{2} \right), \quad k \in \{0, 1\}. \tag{B18}$$

In this case it is not possible to take the limit $\phi_f \to 0$, as this would contradict the condition (B17). The saddle points for the lapse will be purely real.

$\alpha = -1$ *and* $\beta = 0$

The saddle point expressions are

$$N_{\pm}^{\text{saddle}} = i \cosh(2\gamma) \pm \sqrt{-\cosh^2(2\gamma) - a_f^2 \cosh(2\phi_f)}. \tag{B19}$$

The constraint equation automatically yields $x|_0 = 0$. Requiring the initial geometry to close hence implies

$$y|_0 = y_f - 2N\Pi_y = 0, \quad \Leftrightarrow \quad N_{\text{closed}} = \frac{a_f^2 \sinh(2\phi_f)}{2i \sinh(2\gamma)}. \tag{B20}$$

We thus get the following equation for a closed saddle point geometry that we will solve for γ:

$$a_f^2 \sinh^2(2\phi_f) - 4 \sinh(2\gamma) \sinh(2\gamma - 2\phi_f) = 0. \tag{B21}$$

Taking the imaginary part of this expression, we find

$$2a_f^2 \sin(4\,\text{Im}[\gamma]) \sinh(4\,\text{Re}[\gamma] - 2\phi_f) = 0. \tag{B22}$$

This is solved by $\text{Im}[\gamma] = k \cdot \pi/4$ for $k \in \mathbb{Z}$, or $\text{Re}[\gamma] = \phi_f/2$. For each of these cases we examine the real part of Eq. (B21):
(1) $\text{Re}[\gamma] = \phi_f/2$: then we find

$$4 \sin^2(2\,\text{Im}[\gamma]) = -a_f^2 \sinh^2(2\phi_f) - 4 \sinh^2(\phi_f). \tag{B23}$$

This equation does not admit any solution for $a_f > 0$ and $\phi_f \in \mathbb{R}$.
(2) $\text{Im}[\gamma] = k \cdot \pi/2$, $k \in \mathbb{Z}$: then $\cos(2\text{Im}[\gamma]) = (-1)^k$ and $\sin(2\text{Im}[\gamma]) = 0$. Again we only keep $k = \{0, 1\}$. The real part of the equation for γ is

$$a_f^2 \sinh^2(2\phi_f) = 4 \sinh(2\,\text{Re}[\gamma]) \sinh(2\,\text{Re}[\gamma] - 2\phi_f), \tag{B24}$$

which yields the following constraint:

$$a_f^2 \sinh^2(2\phi_f) > 4 \min_{\text{Re}[\gamma]} (\sinh(2\,\text{Re}[\gamma]) \sinh(2\,\text{Re}[\gamma] - 2\phi_f)). \tag{B25}$$

This minimum lies at $\text{Re}[\gamma] = \phi_f/2$ and hence implies $a_f^2 \sinh^2(2\phi_f) > -4\sinh^2(\phi_f)$. This condition is always satisfied, thus Eq. (B24) admits solutions for all values $a_f > 0$ and $\phi_f \in \mathbb{R}$, which are

$$\begin{cases} \text{Re}[\gamma_{\mp}^{\text{I}}] = \mp \cosh^{-1} \left(\sqrt{1 + \frac{\sinh^2(2\phi_f)}{4} \left(2 + a_f^2 \cosh(2\phi_f) - \sqrt{(2 + a_f^2 \cosh(2\phi_f))^2 - a_f^4} \right)} \right), \\ \text{Re}[\gamma_{\mp}^{\text{II}}] = \mp \cosh^{-1} \left(\sqrt{1 + \frac{\sinh^2(2\phi_f)}{4} \left(2 + a_f^2 \cosh(2\phi_f) + \sqrt{(2 + a_f^2 \cosh(2\phi_f))^2 - a_f^4} \right)} \right). \end{cases} \tag{B26}$$

We can further constrain the value of k by looking at the limit of Π_x when $\phi_f \to 0$. In that case, (B24) directly implies $\sinh^2(2\text{Re}[\gamma]) \to 0$, so $\gamma \to 0$. Then

$$\Pi_x = i \cosh(2\gamma) = i(-1)^k \cosh(2\operatorname{Re}[\gamma]) \underset{\gamma \to 0}{\to} i(-1)^k, \tag{B27}$$

so we recover the right limit $\Pi_x \to +i$ only for $k = 0$, so $\operatorname{Im}[\gamma] = 0$.

(3) $\operatorname{Im}[\gamma] = \pi/4 + k \cdot \pi/2$, $k \in \mathbb{Z}$: here we have $\cos(2\operatorname{Im}[\gamma]) = 0$ and $\sin(2\operatorname{Im}[\gamma]) = (-1)^k$, so

$$a_f^4 \sinh^2(2\phi_f) = -4a_f^2 \cosh(2\operatorname{Re}[\gamma]) \cosh(2\operatorname{Re}[\gamma] - 2\phi_f),$$

$$\Rightarrow a_f^2 \sinh^2(2\phi_f) < 4 \max_{\operatorname{Re}[\gamma]} (-\cosh(2\operatorname{Re}[\gamma]) \cosh(2\operatorname{Re}[\gamma] - 2\phi_f)). \tag{B28}$$

This maximum lies at $\operatorname{Re}[\gamma] = \phi_f/2$ and thus $a_f^2 \sinh^2(2\phi_f) < -4\cosh^2(\phi_f)$. This can never be satisfied for real values of a_f and ϕ_f.

APPENDIX C: PICARD'S LITTLE THEOREM

"If a function $f : \mathbb{C} \to \mathbb{C}$ is entire and nonconstant, then the set of values that $f(z)$ assumes is either the whole complex plane or the plane minus a single point." An entire function is a holomorphic function defined on the whole complex plane, and a function $f : \mathbb{C} \to \mathbb{C}$ is holomorphic if and only if the application $\mathbb{C} \to \mathbb{C} : z \to f'(z)$ is continuous on $\mathbb{C}$, and f is $\mathbb{C}$ differentiable; i.e. the corresponding function $F : \mathbb{R}^2 \to \mathbb{R}^2 : (\operatorname{Re}[z], \operatorname{Im}[z]) \to (\operatorname{Re}[f(z)], \operatorname{Im}[f(z)])$ is differentiable at all points $(\operatorname{Re}[z], \operatorname{Im}[z]) = (x, y)$, and for $f(x + iy) = u(x, y) + iv(x, y)$, the Cauchy-Riemann relations are satisfied:

$$\begin{cases} \frac{\partial u}{\partial x}(x, y) = \frac{\partial v}{\partial y}(x, y), \\ \frac{\partial u}{\partial y}(x, y) = -\frac{\partial v}{\partial x}(x, y). \end{cases} \tag{C1}$$

Based on this definition, one can show that the function

$$V : \mathbb{C} \to \mathbb{C} : z \to \alpha \cosh(z) + \beta \sinh(z) \tag{C2}$$

(i.e. our potential) is holomorphic. Generically, for a potential that is an entire function, this theorem means that the constraint equation at $\tau = 0$, $a|_0^2 V(\phi|_0) = 0$, is not only satisfied for closed geometries where $a|_0 = 0$ but also for situations where $V(\phi|_0) = 0$, which generically exist. There is one well-known exception given by the exponential potential, $V(\phi) = e^\phi$, which only vanishes at infinity when $\operatorname{Re}[\phi] < 0$. This is precisely the case we study when $\alpha = \beta$. That case put aside, this explains why for every γ, we find two regular saddle points; one that closes and for which $\phi|_0 = \gamma$, and one that remains unclosed, for which $\phi|_0 \neq \gamma$, but for which $\phi|_0$ will be such that $V(\phi|_0) = 0$.

Moreover this is not specific to this potential, but is a generic feature of all entire functions. Therefore the only assumption we are making is that this potential of the scalar field can be extended to the complex plane. This is however a strong assumption, which can be questioned, see e.g. the discussion in Sec. IV.

[1] S. W. Hawking, The boundary conditions of the universe, Pontif. Acad. Sci. Scr. Varia **48**, 563 (1982), https://ui.adsabs.harvard.edu/abs/1982ac...proc..563H/abstract.

[2] J. B. Hartle and S. W. Hawking, Wave function of the universe, Phys. Rev. D **28**, 2960 (1983); Adv. Ser. Astrophys. Cosmol. **3**, 174 (1987).

[3] J. Feldbrugge, J.-L. Lehners, and N. Turok, Lorentzian quantum cosmology, Phys. Rev. D **95**, 103508 (2017).

[4] J. J. Halliwell, J. B. Hartle, and T. Hertog, What is the no-boundary wave function of the universe?, Phys. Rev. D **99**, 043526 (2019).

[5] O. Janssen, J. J. Halliwell, and T. Hertog, No-boundary proposal in biaxial Bianchi IX minisuperspace, Phys. Rev. D **99**, 123531 (2019).

[6] A. Di Tucci, J.-L. Lehners, and L. Sberna, No-boundary prescriptions in Lorentzian quantum cosmology, Phys. Rev. D **100**, 123543 (2019).

[7] H. Matsui and T. Terada, Swampland constraints on no-boundary quantum cosmology, J. High Energy Phys. 10 (2020) 162.

[8] A. Di Tucci, M. P. Heller, and J.-L. Lehners, Lessons for quantum cosmology from anti-de Sitter black holes, Phys. Rev. D **102**, 086011 (2020).

[9] P. Caputa and S. Hirano, Airy function and 4d quantum gravity, J. High Energy Phys. 06 (2018) 106.

[10] J.-L. Lehners, Wave function of simple universes analytically continued from negative to positive potentials, Phys. Rev. D **104**, 063527 (2021).

[11] C. Jonas and J.-L. Lehners, No-boundary solutions are robust to quantum gravity corrections, Phys. Rev. D **102,** 123539 (2020).

[12] P. A. Cano, K. Fransen, and T. Hertog, Novel higher-curvature variations of R^2 inflation, Phys. Rev. D **103,** 103531 (2021).

[13] C. Jonas, J.-L. Lehners, and J. Quintin, Cosmological consequences of a principle of finite amplitudes, Phys. Rev. D **103,** 103525 (2021).

[14] J. J. Halliwell and S. W. Hawking, The origin of structure in the universe, Phys. Rev. D **31,** 1777 (1985).

[15] J. B. Hartle, S. Hawking, and T. Hertog, The classical universes of the no-boundary quantum state, Phys. Rev. D **77,** 123537 (2008).

[16] J. Feldbrugge, J.-L. Lehners, and N. Turok, No Smooth Beginning for Spacetime, Phys. Rev. Lett. **119,** 171301 (2017).

[17] J. Feldbrugge, J.-L. Lehners, and N. Turok, No rescue for the no boundary proposal: Pointers to the future of quantum cosmology, Phys. Rev. D **97,** 023509 (2018).

[18] A. Di Tucci and J.-L. Lehners, No-Boundary Proposal as a Path Integral with Robin Boundary Conditions, Phys. Rev. Lett. **122,** 201302 (2019).

[19] J. Louko, Canonizing the Hartle-hawking proposal, Phys. Lett. B **202,** 201 (1988).

[20] R. Bousso and S. W. Hawking, Lorentzian condition in quantum gravity, Phys. Rev. D **59,** 103501 (1999); **60,** 109903(E) (1999).

[21] D. Baumann, Inflation, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small* (World Scientific, Singapore, 2011), pp. 523–686, 10.1142/9789814327183_0010.

[22] J.-L. Lehners, Ekpyrotic and cyclic cosmology, Phys. Rep. **465,** 223 (2008).

[23] J.-L. Lehners, Classical inflationary and ekpyrotic universes in the no-boundary wavefunction, Phys. Rev. D **91,** 083525 (2015).

[24] J.-L. Lehners, New ekpyrotic quantum cosmology, Phys. Lett. B **750,** 242 (2015).

[25] R. Brandenberger and P. Peter, Bouncing cosmologies: Progress and problems, Found. Phys. **47,** 797 (2017).

[26] L. J. Garay, J. J. Halliwell, and G. A. Mena Marugan, Path integral quantum cosmology: A class of exactly soluble scalar field minisuperspace models with exponential potentials, Phys. Rev. D **43,** 2572 (1991).

[27] A. Vilenkin, Creation of universes from nothing, Phys. Lett. **117B,** 25 (1982).

[28] S. F. Bramberger, A. Di Tucci, and J.-L. Lehners, Homogeneous transitions during inflation: A description in quantum cosmology, Phys. Rev. D **101,** 063501 (2020).

[29] C. Kiefer, Continuous measurement of minisuperspace variables by higher multipoles, Classical Quantum Gravity **4,** 1369 (1987).

[30] J. J. Halliwell, Decoherence in quantum cosmology, Phys. Rev. D **39,** 2912 (1989).

[31] C. Teitelboim, Quantum mechanics of the gravitational field, Phys. Rev. D **25,** 3159 (1982).

[32] P. Breitenlohner and D. Z. Freedman, Positive energy in anti-De Sitter backgrounds and gauged extended super-gravity, Phys. Lett. **115B,** 197 (1982).

[33] J. M. Maldacena, The large N limit of superconformal field theories and supergravity, Adv. Theor. Math. Phys. **2,** 231 (1998).

[34] L. McGough, M. Mezei, and H. Verlinde, Moving the CFT into the bulk with $T\bar{T}$, J. High Energy Phys. 04 (2018) 010.

[35] L. Battarra and J.-L. Lehners, On the creation of the universe via ekpyrotic instantons, Phys. Lett. B **742,** 167 (2015).

[36] L. Battarra and J.-L. Lehners, On the no-boundary proposal for ekpyrotic and cyclic cosmologies, J. Cosmol. Astropart. Phys. 12 (2014) 023.

[37] M. Kontsevich and G. Segal, Wick rotation and the positivity of energy in quantum field theory, Quart. J. Math. Oxford Ser. **72,** 673 (2021).

[38] E. Witten, A note on complex spacetime metrics, arXiv:2111.06514.

[39] J. Louko and R. D. Sorkin, Complex actions in two-dimensional topology change, Classical Quantum Gravity **14,** 179 (1997).

[40] J.-L. Lehners, Allowable complex metrics in minisuperspace quantum cosmology, arXiv:2111.07816.

[41] K. Rajeev, Wave function of the universe as a sum over eventually inflating universes, arXiv:2112.04522.