The Challenge of Ruling Out Inflation via the Primordial Graviton Background

Sunny Vagnozzi1,2 and Abraham Loeb3

1 Department of Physics, University of Trento, Via Sommarive 14, I-38123 Povo (TN), Italy; sunny.vagnozzi@unitn.it
2 Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
3 Department of Astronomy, Harvard University, 60 Garden Street, Cambridge, MA 02138, USA

Received 2022 April 28; revised 2022 October 16; accepted 2022 October 17; published 2022 November 3

Abstract

Recent debates around the testability of the inflationary paradigm raise the question of how to model-independently discriminate it from competing scenarios. We argue that a detection of the cosmic graviton background (CGB), the relic radiation from gravitons decoupling around Planck time, would rule out the inflationary paradigm, as realistic inflationary models would dilute the CGB to an unobservable level. The CGB contribution to the effective number of relativistic species, \( \Delta N_{\text{eff,c}} \approx 0.054 \), is well within the reach of next-generation cosmological probes. We argue that detecting the high-frequency stochastic gravitational wave background associated to the CGB will be challenging but potentially feasible. We briefly discuss expectations within alternatives to inflation, focusing on bouncing cosmologies and emergent scenarios.

Unified Astronomy Thesaurus concepts: Cosmic inflation (319); Gravitational waves (678); Cosmology (343); Observational cosmology (1146); Alternatives to cosmic inflation (29)

1. Introduction

Inflation, a postulated stage of quasi-de Sitter expansion in the primordial universe, is widely regarded as the leading paradigm for the very early universe. Originally introduced to address various fine-tuning problems of the hot big bang (hBB) model, inflation provides a compelling mechanism for generating the density perturbations from which structure in the primordial universe, is widely regarded as the leading paradigm from

We now discuss the features of the CGB in the absence of inflation. We adopt the working assumption that above the Planck scale pointlike four-particle interactions involving two gravitons, whose rate at temperature \( T \) is of order \( \Gamma_g \sim T^3/M^4_{\text{Pl}} \), keep gravitons in thermal equilibrium in the primordial plasma (see also Zhao et al. 2009; Giovannini 2020). If we assume adiabatic evolution throughout the early stages of the primordial plasma, and therefore that the universe was radiation dominated up to then, the Hubble rate scales as \( H \sim T^2/M_{\text{Pl}} \). Comparing the two rates indicates that gravitons decouple at a temperature \( T_{g,\text{dec}} \sim M_{\text{Pl}} \) (or equivalently around Planck time \( t_{g,\text{dec}} \sim t_{\text{Pl}} \)); besides ruling out inflation, a detection of the CGB would thus provide an experimental test bed for theories attempting to unify quantum mechanics and gravity.

Being massless and thus decoupling while relativistic, primordial gravitons preserve the blackbody form of their spectrum following decoupling, with the effective CMB redshifted with the scale factor \( a \) as \( T_g \propto 1/a \). Since the entropy density \( s = 2\pi^2g_*(T)T^3/45 \) scales as \( s \propto a^{-3} \), where \( g_*(T) \) is the (temperature-dependent) effective number of entropy degrees of freedom (dof), we can relate the present-day temperatures of the CGB and CMB, \( T_{g,0} \) and \( T_{\gamma,0} \), respectively, as follows:

\[
\frac{T_{g,0}}{T_{\gamma,0}} = \left( \frac{g_*^c(T_0)}{(g_*^c(T_{\text{Pl}}) - 2)} \right)^{1/3},
\]

where \( g_*^c(T_0) \approx 3.91 \) is the present-day effective number of entropy dof excluding gravitons (accounting for photons and neutrinos), and \( g_*^c(T_{\text{Pl}}) \) is the effective number of entropy dof prior to graviton decoupling, including gravitons. Accounting
only for standard model (SM) dof up to the Planck scale, above the electroweak (EW) scale $g^*_s(T_{Pl}) - 2 \approx 106.75$. Precise measurements of the CMB frequency spectrum from COBE/FIRAS fix $T_{r,0} \approx 2.7 \text{ K}$ and therefore under these minimal assumptions the present-day CGB temperature is predicted to be $T_{g,0} \approx (3.91/106.75)^{1/3} T_{r,0} \approx 0.9 \text{ K}$, making the CGB about three times colder than the CMB.

Lacking a precise knowledge of the type of new physics lying beyond the TeV scale, the assumption of only considering SM dof up to the Planck scale is conservative, but likely somewhat unrealistic, as one might expect several additional dof to appear in the “desert” between the EW and Planck scales. If so, $g^*_s(T_{Pl})$ in the denominator of Equation (1) can only increase, decreasing $T_{g,0}$ with respect to the previous estimate $T_{g,0} \approx 0.9 \text{ K}$, which therefore should be viewed more as a conservative upper bound on $T_{g,0}$. However, the exact numbers are highly model dependent and depend on the specific new physics model. For instance, $T_{g,0} \approx 0.7 \text{ K}$ in a supersymmetric-like scenario where $g^*_s(T_{Pl})$ doubles, whereas $T_{g,0} \approx 0.4 \text{ K}$ in a hypothetical scenario where $g^*_s(T_{Pl})$ increases by an order of magnitude.

3. Can Inflation Be Ruled Out?

Our assumption of adiabatic evolution from $T_{Pl}$ down to present times breaks down whenever comoving entropy is generated, e.g., during reheating at the end of inflation. An inflationary phase alters the relation between $T_{g,0}$ and $T_{r,0}$ in Equation (1), as the latter would be determined by the dynamics of reheating, which however can at most produce out-of-equilibrium graviton excitations, unless the effective gravitational constant $G_{eff}$ was significantly higher at reheating. Since the scale factor increases exponentially during inflation, the CGB temperature itself is exponentially suppressed by a factor of $e^{-N}$, with $N$ the number of $e$ folds of inflation.

We can obtain an extremely conservative upper limit on $T_{g,0}$ in the presence of a phase of inflation (the tilde distinguishes the present-day graviton temperatures with and without inflation), using the facts that (a) solving the horizon and flatness problems requires $N \gtrsim 60$, and (b) reheating should occur at $T_{rh} \gtrsim 5 \text{ MeV}$ in order to not spoil Big Bang nucleosynthesis predictions (de Salas et al. 2015). From these requirements we find that $T_{g,0} \lesssim 50 \mu \text{ K}$, implying that inflation would dilute the CGB to an unobservable level. More generally, we find the following upper limit:

$$T_{g,0} \ll 0.25 \left( \frac{T_{rh}}{\text{GeV}} \right)^{-1} e^{60-N} \mu \text{ K}. \quad (2)$$

However, $T_{g,0} \lesssim 50 \mu \text{ K}$ is a very conservative upper limit, for two reasons. First, in most realistic models inflation typically proceeds for more than 60 $e$ folds, leading to further exponential suppression (see Equation (2)). Next, although reheating at scales as low as $T_{rh} \simeq O(\text{MeV})$ is observationally allowed, models realizing this in practice are very hard to construct (see, e.g., Kawasaki et al. 1999; Hannestad 2004; Khouy & Steinhardt 2011). It is far more likely that, if inflation did occur, reheating took place way above the EW scale, further tightening the upper bound on $T_{g,0}$.

One may try to evade these conclusions invoking models of incomplete inflation with a limited number of $e$ folds $46 \lesssim N \lesssim 60$; however, if inflation is indeed the solution to the flatness problem, such models are essentially ruled out by current stringent bounds on spatial curvature (Vagnozzi et al. 2021), as argued explicitly in Efstathiou & Gratton (2020). Even if $N < 60$, bringing $T_{g,0}$ to a detectable level still requires an extremely low reheating scale, typically harder to achieve within models of incomplete inflation.

A caveat to our previous results is our assumption of inflation occurring at sub-Planckian scales. Specifically, $T_{rh} > M_{Pl}$ is required for the CGB not to be washed out by inflation. However, on general grounds there are serious concerns about the consistency of trans-Planckian effects both during inflation and at reheating (e.g., Brandenberger & Martin 2013; Brandenberger & Kamali 2022). A specific concern is given by the trans-Planckian censorship conjecture, which sets tight limits on the maximum inflationary scale $\Lambda_{\text{inf}}$ and reheating temperature: $\Lambda_{\text{inf}}$, $T_{rh} < M_{Pl}$ (Bedroya & Vafa 2020; Bedroya et al. 2020; Kamali & Brandenberger 2020; Mizuno et al. 2020).

More importantly, the lack of detection of inflationary B modes indicates that $\Lambda_{\text{inf}}$ is at least 4 orders of magnitude below the Planck scale. For instantaneous reheating, the reheating temperature is obviously limited to $T_{rh} < \Lambda_{\text{inf}}$, as reheating to higher temperatures would violate (covariant) stress-energy conservation. For noninstantaneous reheating, $T_{rh}$ is of course even lower (see also Cook et al. 2015). Therefore, we deem it very safe to assume that $T_{rh} \ll M_{Pl}$, corroborating all our earlier findings. In summary, within realistic inflationary cosmologies one does not expect to be able to detect the relic thermal graviton background—contrary, a convincing detection thereof would rule out the inflationary paradigm.

4. Detectability of the CGB

We now investigate whether detecting the CGB is experimentally feasible, considering our benchmark $T_{g,0} \approx 0.9 \text{ K}$ case. The contribution of the CGB to the effective number of relativistic species $N_{\text{eff}}$ is given by:

$$\Delta N_{\text{eff},g} = \frac{8}{7} \left( \frac{11}{4} \right)^2 \frac{\rho_k}{\rho_y} = \frac{8}{7} \left( \frac{11}{4} \right)^2 \left( \frac{g^*_s(T_{rh})}{(g^*_s(T_{Pl}) - 2)} \right). \quad (3)$$

For $g^*_s(T_{Pl}) - 2 \approx 106.75$, we therefore find that $\Delta N_{\text{eff},g} \approx 0.054$, as expected for a species with 2 spin dof decoupling before the QCD phase transition.

A contribution to $N_{\text{eff}}$ of this size is a factor of 3 below the sensitivity of current probes. However, this number is well within the reach of a combination of next-generation CMB and LSS surveys. For instance, even after marginalizing over the total neutrino mass, Brinckmann et al. (2019) forecast a sensitivity of $\sigma_{N_{\text{eff}}} \approx 0.021$ combining CMB data from CMB-S4 and LiteBIRD with galaxy clustering and cosmic shear data from Euclid, whereas with a PICO-like experiment in place of CMB-S4+LiteBIRD the sensitivity improves to $\sigma_{N_{\text{eff}}} \approx 0.017$. Therefore, if the benchmark 0.9 K CGB were present, CMB-S4+LiteBIRD+Euclid would be able to detect it through its imprint on $N_{\text{eff}}$ at $\approx 2.5\sigma$, whereas PICO+Euclid would be able to do so at $\approx 3.2\sigma$.\footnote{Note, however, that a viable interpretation of the signal recently observed by various pulsar timing arrays (e.g., Arzoumanian et al. 2020) is in terms of inflationary GWs given a rather low reheating scale (Kuroyanagi et al. 2021; Oikonomou 2021; Vagnozzi 2021; Benetti et al. 2022; Odintsov et al. 2022; Oikonomou 2022).}
Figure 1. Strain of the CGB stochastic background of high-frequency GWs, alongside the sensitivities of various detector concepts discussed in the main text. The red line (“EMC”) refers to enhanced magnetic conversion, with the more transparent extension referring to potential future technological improvements discussed in the main text.

Should the CGB contribution to \( N_{\text{eff}} \) be detected, one may wonder how we know that the excess radiation density is associated to the CGB, rather than another dark radiation component. To remove this ambiguity, we consider the stochastic background of (high-frequency) gravitational waves (GWs) associated to the CGB. It is useful to think in terms of characteristic strain \( h_c \), i.e., the dimensionless strain that would be produced due to the passing stochastic GW background (SGWB) in the arms of an interferometer with arms of equal length \( L \) in the \( x \) and \( y \)-directions, \( h_c(\nu) \approx \Delta L/L \). The characteristic CGB strain \( h_c(\nu) \) is given by:

\[
h_c(\nu) = \frac{1}{\nu} \sqrt{\frac{3h_0^2}{2\pi^2}} \Omega_{\text{g}}(\nu) \approx 1.26 \times 10^{-27} \left( \frac{\nu}{\text{GHz}} \right)^{-1/2} \Omega_{\text{g}}(\nu),
\]

where \( h_0^2 \Omega_{\text{g}}(\nu) \) is the CGB spectral energy density in units of the present-day critical density;

\[
h_0^2 \Omega_{\text{g}}(\nu) = \frac{15}{\pi^4} h_0^2 \Omega_{\text{g},0} \left( \frac{T_{\text{g},0}}{T_{\text{c},0}} \right)^4 F(x_{\nu}),
\]

with \( h \) the reduced Hubble parameter, \( h_0^2 \Omega_{\text{g},0} \) the photon density parameter today, \( x_{\nu} \equiv h\nu/(k_B T_{\text{c},0}) \), and \( F(x_{\nu}) \equiv x_{\nu}^4/(e^{x_{\nu}} - 1) \). The CGB spectrum peaks at frequencies \( \nu \approx 75 \text{ GHz} \), making it a source of high-frequency GWs: Figure 1 shows the characteristic CGB strain alongside demonstrated or forecasted sensitivities of various detector concepts (see Aggarwal et al. 2021).

Aside from optically levitated sensors (Arvanitaki & Geraci 2013) and bulk acoustic wave devices (Goryachev & Tobar 2014), all probes in Figure 1 exploit the inverse Gertsenshtein effect (IGE), whereby GWs convert to photons within a strong magnetic field (Gertsenshtein 1962). While apart from small prototypes, dedicated instruments exploiting the IGE do not exist; Ito et al. (2020) and Ejlli et al. (2019) showed how constraints on high-frequency GWs can be obtained reinterpreting data from ongoing or planned axion experiments: in Figure 1 this includes ADMX, SQMS, IAXO, single-photon detectors (SPDs), JURA, OSQAR, and DMRadio8-100 (Domcke et al. 2022). The IGE can also be exploited in strongly magnetized astrophysical environments (Chen 1995; Domcke & Garcia-Cely 2021), recasting observations from radio telescopes such as EDGES and ARCADE. For more details on these detector concepts, see Aggarwal et al. (2021), Berlin et al. (2022), and Domcke et al. (2022).

Unfortunately, as is clear from Figure 1, all these detector concepts fall short of the CGB signal by several orders of magnitude. The only promising probe is enhanced magnetic conversion (EMC), a proposal to enhance the efficiency of IGE-based magnetic conversion detectors by seeding the conversion volume with locally generated auxiliary EM fields, e.g., EM Gaussian beams (GBs) oscillating at the frequency of the GW signal searched for (Li & Yang 2004; Baker et al. 2008). Until recently, EMC appeared to be well beyond technological reach, particularly due to the requirement of a GB geometric purity at the \( 10^{-24} \) level to reach strain levels of \( h_c \sim 10^{-30} \) at \( \nu \sim 100 \text{ GHz} \).

However, Ringwald et al. (2021) argued that reaching the above benchmark limit is feasible, exploiting state-of-the-art superconducting magnets utilized in near-future axion experiments to generate the required EM signal, then enhanced by a GB produced by an MW-scale 40 GHz gyrotron. While this still leaves us 2 orders of magnitude short of the CGB peak strain, realistic improvements in the development of gyrotrons, SPDs, and superconducting magnets, can bring the projected sensitivity down to \( h_c \sim 10^{-32} \), sufficient to detect the CGB in our benchmark scenario. We estimate that an increase in the gyrotron available power to \( \sim 100 \text{ MW} \) (which is realistically achievable) over a stable running time of \( \sim 1 \) month (which is much more challenging), alongside improvements in SPD dark count rates to \( \sim 10^{-5} \text{ s}^{-1} \), would result in a sensitivity to strains.
of order $h_i \sim 10^{-33}$, sufficient to detect our benchmark CGB. All quoted sensitivities can be further improved by increasing the reflector size, and the intensity and length of the magnets. Therefore, measuring strains as small as $h_i \sim 10^{-33}$ at $\nu \sim \mathcal{O}(100)\text{GHz}$, and detecting the benchmark CGB, might be feasible in the not-too-far-off future.\footnote{We recall that the quoted CGB strength assumes that the SM holds up to the Planck scale, and that the appearance of additional dof would lower the CGB temperature and associated SGWB strength. However, even in the extremely unrealistic scenario where the number of dof increases by an order of magnitude, the temperature of the CGB would only decrease by a factor of $\gtrsim 2$, making the SGWB signal only a factor of $\approx 5$ weaker (see the $T_{\gamma \gamma}$ dependence in Equations (4) and (5)).}

Another interesting potential detection channel proposed very recently by Brandenberger et al. (2022) proceeds through a parametric instability of the EM field in the presence of GWs. This would allow for conversion of high-frequency GWs to photons without the need for a strong background magnetic field. Sensitivity reach estimates for this probe, while not yet available, are worth further investigation in this context. An important issue concerns how to distinguish the CGB from competing SGWB sources. Possible examples could be the SGWB produced during preheating (Easther & Lim 2006) or during oscillon formation (Zhou et al. 2013); however, both these sources are important at lower frequencies, $\mathcal{O}(10^6-10^9)\text{Hz}$ (see Aggarwal et al. 2021), and hence should not confuse the CGB detection. The CGB SGWB can also be distinguished from the SGWB produced by out-of-equilibrium gravitational excitations at reheating (Ringwald et al. 2021); the latter would not be of the blackbody form, and its strength would be orders of magnitude below the CGB as long as the reheating temperature is $T_{rh} \ll M_P$, which as argued earlier can be safely assumed. This highlights the importance of detecting the CGB over a range of frequencies, given the clear prediction for its frequency dependence. Within the EMC experimental setup, this can be achieved by tuning the gyrotron frequency: the output frequencies available for typical gyrotrons fall within the $\sim 20-500\text{GHz}$ range, perfectly suited to probe the CGB spectrum around its peak frequency. A similar tuning procedure should also be possible for the GW-photon parametric instability probe.

A caveat to our findings is the assumption of a pure blackbody spectrum for primordial gravitons. This is likely to be an approximation at best, particularly at low frequencies, whose modes would have been superhorizon at Planck time. However, in the absence of detailed knowledge regarding the underlying theory of quantum gravity, this is among the most conservative assumptions we can make (note that the same approximation has been made in several earlier works discussing primordial gravitons, e.g., Zhao et al. 2009 and Giovannini 2020). Moreover, what is important for our results is the high-frequency tail of the CGB spectrum, where our assumption is likely to be far more realistic. Overall, it remains true that finding any trace of a GW background of the estimated amplitude at the estimated frequency will rule out the standard inflationary scenario.

5 Alternatives to Inflation

Our previous discussion raises the question of whether an unambiguous CGB detection would also spell trouble for alternative paradigms, where density perturbations are produced during a noninflationary phase. While the answer to this question is highly model dependent, we wish to provide a brief qualitative assessment limited to two well-motivated paradigms: bouncing cosmologies and emergent scenarios.

Within bouncing cosmologies, the challenge is to produce a thermal CGB in the first place. This is hard to achieve during the contracting phase, when the characteristic energy scale is typically $\Lambda_s \ll M_P$ (e.g., Brandenberger & Peter 2017). Another possibility is one where a relatively long bouncing phase with energy density around the Planck scale occurs between the initial contracting phase and the later hBB expansion (e.g., Cai 2014), in which case a thermal CGB would be generated and would survive the phase transition between the bouncing and expanding phases.

In emergent scenarios, the universe emerges from an initial high density state with matter in global thermal equilibrium, and where producing the CGB is far less unlikely. A particularly well-studied emergent scenario is the string gas proposal of Brandenberger & Vafa (1989), where the universe originates from a quasi-static Hagedorn phase of a string gas at temperature close to the Hagedorn temperature, before a T-dual symmetry breaking-driven phase transition connects to the hBB expansion. On general grounds, the energy density in the emergent phase is close to the Planck density, making it likely for gravitons to be in thermal equilibrium and therefore for a CGB to be generated.

However, the initial state in string gas cosmology is not a thermal state of particles but of strings, giving a different scaling of thermodynamical quantities. It is therefore unlikely that the string gas CGB takes the blackbody form, although it is in principle possible that its spectral energy density may be higher than our benchmark CGB, enhancing detection prospects. Fully exploring these points requires a dedicated study, going beyond the scope of our work.

6. Conclusions

Despite its enormous success, recent debates around the inflationary paradigm raise the question of how to model-independently discriminate it from competing scenarios for the production of primordial density perturbations. In this Letter, we have argued that a detection of the CGB, the leftover graviton radiation from the Planck era, would rule out the inflationary paradigm, as realistic inflationary models dilute the CGB to an unobservable level. Assuming the validity of the SM up to the Planck scale, the CGB contribution to the effective number of relativistic species $\Delta N_{\text{eff},g} \approx 0.054$ is well within the reach of next-generation cosmological probes, whereas detecting the associated stochastic background of high-frequency GWs in the $\nu \sim \mathcal{O}(100)\text{GHz}$ range is challenging but potentially feasible. We also argued that the CGB may be detectable within well-motivated alternatives to inflation such as bouncing and emergent scenarios. We hope that this work will spur further investigation into the possibility of model-independently confirming or ruling out the inflationary paradigm with upcoming observations (for similar endeavors see, e.g., Chen et al. 2019).

We are grateful to Robert Brandenberger, Massimo Giovannini, Will Kinney, Nick Rodd, and Luca Vissinelli for useful discussions and suggestions. S.V. is partially supported by the Isaac Newton Trust and the Kavli Foundation through a Newton-Kavli Fellowship, by a grant from the Foundation Blanceflor Boncompagni Ludovisi, née Bildt, and by a College Research Associateship at Homerton College, University of
Cambridge. A.L. is partially supported by the black hole Initiative at Harvard University, which is funded by grants from the John Templeton Foundation and the Gordon and Betty Moore Foundation.

ORCID iDs

Sunny Vagnozzi @ https://orcid.org/0000-0002-7614-6677
Abraham Loeb @ https://orcid.org/0000-0003-4330-287X

References

Aggarwal, N., et al. 2021, LRR, 24, 4
Albrecht, A., & Steinhardt, P. J. 1982, PhRvL, 48, 1220
Arvanitaki, A., & Geraci, A. A. 2013, PhRvL, 110, 071105
Arzoumanian, Z., et al. 2020, ApJL, 905, L34
Baker, R. M. L., Stephenson, G. V., & Li, F. 2008, in AIP Conf. Proc. 969, Space Technology and Applications Int. FORUM-STAIF 2008 (Melville, NY: AIP), 1045
Bedroya, A., Brandenberger, R., Loverde, M., & Vafa, C. 2020, PhRvD, 101, 103502
Bedroya, A., & Vafa, C. 2020, JHEP, 09, 123
Benetti, M., Graef, L. L., & Vagnozzi, S. 2022, PhRvD, 105, 043520
Berlin, A., Blas, D., Tito D’Agnolo, R., et al. 2022, PhRvD, 105, 116011
Brandenberger, R., Delgado, P. C. M., Ganz, A., & Lin, C. 2022, arXiv:2205.08767
Brandenberger, R., & Kamali, V. 2022, EPJC, 82, 818
Brandenberger, R., & Peter, P. 2017, FoPh, 47, 797
Brandenberger, R., & Vafa, C. 1989, NuPhB, 316, 391
Brandenberger, R. H., & Martin, J. 2013, CQGra, 30, 113001
Brinckmann, T., Hooper, D. C., Archidiacono, M., Lesgourgues, J., & Sprenger, T. 2019, JCAP, 01, 059
Cai, Y.-F. 2014, SCPMA, 57, 1414
Chen, P. 1995, PhRvL, 74, 634

5