Non-Gaussianity from Axion-Gauge Fields Interactions during Inflation

Emanuela Dimastrogiovanni [a,b], **Matteo Fasiello** [c], **Robert J. Hardwick** [c], **Hooshyar Assadullahi** [c], **Kazuya Koyama** [c], **David Wands** [c]

[a] CERCA & Department of Physics, Case Western Reserve University, Cleveland, OH, 44106, USA.

[b] Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, N2L 2Y5, Canada.

[c] Institute of Cosmology and Gravitation, University of Portsmouth, PO1 3FX, UK.

Abstract. We study the scalar-tensor-tensor non-Gaussian signal in an inflationary model comprising also an axion coupled with SU(2) gauge fields. In this set-up, metric fluctuations are sourced by the gauge fields already at the linear level providing an enhanced chiral gravitational waves spectrum. The same mechanism is at work in generating an amplitude for the three-point function that is parametrically larger than in standard single-field inflation.

Contents

1 Introduction

A period of accelerated expansion in the early universe, *inflation* has been hypothesized [1] in order to solve a number of puzzling initial conditions in the standard hot big-bang cosmology. Already in its simplest formulation, that of a scalar field minimally coupled to gravity, inflation can resolve such issues and provides a mechanism by which quantum fluctuations at early times are swept up by inflation to become the primordial seeds for structures to form in the universe.

The spectacular advances in observational cosmology in recent decades have refined the allowed range for viable inflationary models. A nearly scale-invariant spectrum of primordial adiabatic scalar fluctuations are required for agreement with observations with only small non-Gaussianities (e.g. for local non-Gaussianity the bispectrum amplitude is constrained to $f_{\rm nl}^{\rm local} \lesssim \mathcal{O}(10)$ [2]). These constraints notwithstanding, the inflationary paradigm can accommodate a rich particle content. An observational window on inflation is then automatically also a precious portal to high energy physics and a very special one at that; it provides access to beyond-standard-model energy scales that can be as high as 10^{14}GeV, well out of the reach of earth-bound particle colliders. Extra field content is not just an interesting possibility for inflation, it is also a natural one. To give just one example, in supersymmetric theories, unless supersymmetry is broken at scales much higher than the Hubble scale, $E \gg H$, the inflationary vacuum energy will break supersymmetry so that some of the resulting particles will have masses $m \sim H$. Even if such massive fields are long lost to us today, cosmological observables (e.g., the squeezed configuration of the bispectrum) can carry the imprint of their early dynamics so that one can engage in cosmological

archeology and search for such *fossils* [3]. Interestingly, information on the spin, mass and coupling of these particles can still be accessible today [4].

Given the plethora of inflationary setups still compatible with observational bounds, one may rely on future cosmological probes to identify the most compelling scenarios, as well as the requirement of a theoretically robust implementation of inflation. The latter includes navigating the perils of the so-called η-problem; in the absence of a sufficiently powerful symmetry, the inflaton potential will receive loop corrections of the form $V_0(\phi/M_P)^n$ making its mass too large ($m \sim H$) to sustain a sufficiently long expansion.

An approximately shift-symmetric potential can significantly ameliorate the η-problem as in the well-known case of natural inflation [5]. In this context (see [6] for a review on the subject), the axion potential receives non-perturbative contributions from the gauge sector resulting in a left-over discrete shift symmetry for the field $\phi \to \phi + 2\pi f$, with the dimensionful quantity f regulating the periodicity. Known string-theory constructions [7] suggest the constraint $f < M_P$; this hierarchy is further motivated by the fact that quantum gravity is expected to break all global symmetries. Given that observationally viable inflation via a single axion requires $f > M_P$, in order to operate in an under-control inflationary regime one may couple the axion to other[1] sectors so as to effectively lower f.

One such example is that of an axionic inflaton directly coupled to gauge fields via the least-irrelevant shift-symmetric operator $\phi F\tilde{F}$. There exists a vast literature [9] on what remains a very active subject, and includes the possibility of $F\tilde{F}$ being standard model gauge fields thus providing a natural reheating scenario (see e.g. [10]). In light of the axion-gauge coupling, an entire class of axion inflation models share intriguing potential signatures: a chiral gravitational wave signal, and in particular one that can grow at smaller scales (blue spectrum)[2]. It is worth pointing out that similar models have recently been employed in the context leptogenesis via axial-gravitational anomaly [13].

An intriguing specific realization of axion inflation is known as chromo-natural inflation (CNI) [14]: here the coupling is to SU(2) gauge fields[3], allowing for isotropic background solutions ([16, 17] provide a non-exhaustive list of works on the subject). Further studies [18] showed that the simplest realization of CNI is excluded by Planck data. This has lead to an extension of the model [19] (see also [20]) that retains all its original intriguing features; the tension with data is resolved by equipping the scalar sector with an additional field, now driving inflation. The extra field is not necessarily an axion and therefore its potential need not be shift-symmetric. Crucially, detection-level gravitational waves at CMB scales can be generated already at sub-Planckian values for the axion field-excursion, thereby reducing the effect of loop corrections on

[1] Another intriguing possibility is to have multiple axions [8].

[2] Other classes of inflationary scenarios endowed with non-standard gravitational waves production mechanisms include, e.g., scalar spectator fields with a small sound speed [11] and modified gravity models [12].

[3] See [15] for a very recent analysis pointing out one extra advantage that comes with the use of an $SU(2)$ as opposed to $U(1)$ model in the context of Schwinger pair creation and backreaction.

the inflationary potential. Interestingly, it has recently been shown how both CNI and its extension can be embedded in supegravity and string theory [21].

We consider in this paper the model introduced in [19]. Remarkably, a scan of the parameter space of this model reveals regions generating signatures detectable by both CMB probes and interferometers. The SU(2)-based enhancement of gravitational waves can lead to detectable $\langle EB \rangle$, $\langle TB \rangle$ signals for upcoming CMB probes and may be searched for using existing interferometers (and a cross-correlation thereof[4]) [22]. It has recently been shown that this setup supports large tensor non-Gaussianities [23] and we will show here that the same is true for the scalar-tensor-tensor correlation. It is interesting to point out that this specific observable, the scalar-tensor-tensor bispectrum, has received attention in the context of other models that also exhibit a non trivial chirality of the signal. A case in point is Chern-Simons gravity [24, 25]. The effect on CMB observables of parity-violation in the tensor sector was also the subject of studies in [26] and [27–30]. We refer the reader to the works above and references therein for further literature on the subject.

This paper is organized as follows: in *Section* 2 we review the model and its predictions at the level of the power spectra; in *Section* 3 we present the calculation of the scalar-tensor-tensor bispectrum and discuss our findings on its shape and amplitude with an eye on perturbativity bounds inherited also from the scalar-sector; conclusions are in *Section* 4. More details about the calculations can be found in the Appendices.

2 The model

As mentioned above, our model includes *spectator* fields (i.e. fields providing a sub-leading contribution to the total energy density during inflation), including an axion field χ and an SU(2) gauge field A^a_μ, in addition to the inflaton sector,

$$\mathcal{S} = \int d^4x\sqrt{-g}\left[\frac{M_{\rm Pl}^2}{2}R + \mathcal{L}_\phi - \frac{1}{2}\left(\partial\chi\right)^2 - U(\chi) - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \frac{\lambda\,\chi}{4f}F^a_{\mu\nu}\tilde{F}^{a\mu\nu}\right], \quad (2.1)$$

where $\mathcal{L}_\phi$ is the inflaton Lagrangian, $F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g^{abc}A^b_\mu A^c_\nu$ and the definition $\tilde{F}^{a\mu\nu} \equiv \epsilon^{\mu\nu\rho\sigma}F^a_{\rho\sigma}/(2\sqrt{-g})$ has been used.

The background equations of motion and the linear perturbation analysis were first presented in [19]. In this section, we review the main results and identify the model parameters that will appear in the bispectrum computation. The background for the gauge field can be chosen as $A^a_0 = 0$, $A^a_i = \delta^a_i a(t)Q(t)$. The scalars Q and χ have coupled equations of motion. Under minimal assumptions on the parameters and in a regime of slow-roll for the fields, the effective potential for Q is minimized by

$$Q = \left(\frac{-f\,U_\chi}{3g\lambda H}\right)^{1/3}. \quad (2.2)$$

[4]Cross-correlations among signals from different interferometers or from different constellations making up the same interferometer (as in the case of an advanced design of LISA or BBO discussed in [22]) would be necessary in order to detect circular polarization.

From the same equations of motion it also follows that

$$\frac{\lambda}{2fH}\dot{\chi} \simeq m_Q + \frac{1}{m_Q}\,, \tag{2.3}$$

where the parameter $m_Q \equiv g\,Q/H$ is to be interpreted as the mass, in units of Hubble, of the gauge field fluctuations. Einstein's equations lead to the following relation among slow-roll parameters:

$$\epsilon_H \equiv -\dot{H}/H^2 = \epsilon_\phi + \epsilon_\chi + \epsilon_B + \epsilon_E\,, \tag{2.4}$$

with

$$\epsilon_\phi \equiv \frac{\dot{\phi}^2}{(2H^2M_{\rm Pl}^2)}\,; \quad \epsilon_\chi \equiv \frac{\dot{\chi}^2}{(2H^2M_{\rm Pl}^2)}\,; \quad \epsilon_B \equiv \frac{g^2Q^4}{(HM_{\rm Pl})^2}\,; \quad \epsilon_E \equiv \frac{(HQ+\dot{Q})^2}{(HM_{\rm Pl})^2}\,. \tag{2.5}$$

The metric tensor fluctuations (h_{ij}) are linearly sourced by the tensor perturbations of the gauge field. The latter experience (near horizon crossing) a growth in one of their two polarizations that is controlled by m_Q; as a result, the corresponding helicity in the gravitational waves is enhanced. This non-zero chirality can be understood as a consequence of the parity-breaking nature of the gauge-field background. The expression for the sourced power spectrum is given by

$$\mathcal{P}_h^{\rm s} = \epsilon_B\,\frac{H^2}{\pi^2 M_{Pl}^2}\mathcal{F}^2\,, \tag{2.6}$$

where $\mathcal{F} = \mathcal{F}(m_Q)$ (a detailed derivation can be found in [19]). The transient instability of the gauge field tensor fluctuations can be understood as an energy transfer from the rolling axion.

The $SU(2)$-sensitive contribution to the power spectrum of gravitational waves (GW) can be larger than the one from vacuum fluctuations and within reach of upcoming experimental probes[5]; the model predicts chiral gravitational waves that would be observable for a sizable portion of its parameter space [22]. Our set-up serves as an explicit example of the fact that detectable GW may be generated even at a relatively low value for H, thus breaking the one-to-one $r \leftrightarrow H$ correspondence between the tensor-to-scalar ratio and the energy scale of inflation (see [34] for more about the lower bound on H in this context).

The tensor power spectrum in Eq. (2.6) is characterized by a broad (depending on model parameters) feature, a distinctive scale dependent "bump" that results from the background evolution of the axion-gauge field system, i.e. from the time-dependence, within the $\mathcal{F}$ function, of $m_Q(t) \equiv g\,Q(t)/H$ [19] (see also Appendix A of [22] for an analytical approximation of the power spectrum as a function of scale, valid around the peak value and for $3 \leq m_Q \leq 7$). From this feature originates the fact that there is

[5]These include, for example, the ground based CMB-S4 [31] and planned space missions for CMB polarization anisotropies such as LITEBIRD [32], as well as space-based interferometers such as LISA [33].

in this model ample room for a blue tensor spectral index, crucial for direct detection by interferometers. Having lifted the burden of driving inflation from the axion (in order to recover compatibility with data [18]), in the extended model one may enhance sourced gravitational waves on different scales by sampling the parameter space and acting on the coupling between the fields.

Moving on to the power spectrum of curvature fluctuations, this will depend on the precise form of $\mathcal{L}_\phi$ and, naturally, also on the specific dynamics one may postulate for the post-inflationary evolution of the spectator sector. The axion ($\delta\chi$) and gauge field scalar fluctuations are directly coupled with one another (and only gravitationally coupled to the inflaton fluctuations). We note that these modes will undergo a tachyonic instability [18] starting in the sub-horizon regime unless $m_Q \geq \sqrt{2}$. We will confine our analysis to such viable region of the parameter space.

Both the inflaton field and scalar fluctuations of the spectator sector contribute to curvature perturbations. The authors of [19] chose to be as agnostic as possible on the details of $\mathcal{L}_\phi$. It is nevertheless necessary to ensure that the field ϕ is the one driving inflation and that furthermore there exists a hierarchy among the slow-roll parameters with $\epsilon_\phi \simeq \epsilon_{\rm max}$. The latter condition ensures that the spectral index can satisfy existing observational constraints. Under such conditions the power spectrum of curvature perturbation is dominated by the inflaton contributions and is only mildly affected by the axion and gauge fields. However, for a more careful analysis, see *Section* 3.2.

3 STT bispectrum from Chern-Simons interactions

The $\chi F\tilde{F}$ interaction of Eq. (2.1) supports a transient growth in one of the $SU(2)$ tensor polarizations that propagates to the corresponding helicity in the GW power spectrum. This mechanism is also in place for higher-order correlation functions. The GW bispectrum for the theory in Eq. (2.1) has been calculated in [23], where it was shown that the $SU(2)$ contribution to tensor non-Gaussianity can be significantly larger than the one of standard single field inflation.

It is intuitively clear that, because the growth of the sourcing mode function occurs (only) near horizon crossing, the bispectrum shape will very much resemble the equilateral one, although there are some subtle differences with respect to the exact equilateral template. In an analogous fashion, one expects also mixed tensor-scalar correlators to receive the most sizable contributions from the gauge sector in equilateral configurations. The axion $\delta\chi$ is sourced by gauge tensor fluctuations via Chern-Simons interactions while in turn the curvature perturbation ζ receives contributions from $\delta\chi$.

Scalar non-Gaussianity is constrained on large CMB scales by, for example, $f_{\rm nl}^{\rm local} \lesssim \mathcal{O}(10)$ and $f_{\rm nl}^{\rm equil} \lesssim \mathcal{O}(50)$ (respectively for the local and equilateral shapes [2]). These bounds will soon improve thanks to upcoming large-scale structure observations and new CMB polarization data. The ongoing development of new interferometers with improved sensitivity to the stochastic background of primordial GW [33] will also help us place stronger constraints on tensor and mixed non-Gaussianity. Non-Gaussian ob-

servables are invaluable as a probe of the production mechanism of primordial GW and, more broadly, the inflationary particle content [3].

In what follows, we shall focus on the the (χ, t)-mediated $\langle h\, h\, \zeta\rangle$ bispectrum contribution, where we define t as the transverse and traceless part of the gauge field fluctuations, $\delta A_i^a \supset t_{ai}$. The presence and form of the Chern-Simons interaction suggest that this one observable is particularly sensitive to the effects of gauge fields. A typical contribution of this kind is represented[6] in the diagram of Fig. (1), which we evaluate in details in the remaining of this section.

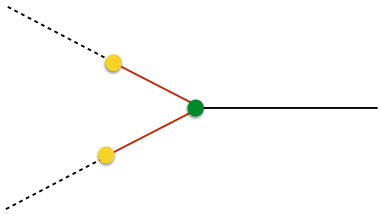

Figure 1. Black dotted lines represent the metric tensor h; red lines stand for the gauge field tensor perturbation t; the solid black line indicates the curvature perturbation $\zeta^{(\chi)}$. The green vertex arises from the Chern-Simons contribution to $\delta\chi\, t\, t$. The yellow vertex (to be understood according to the caveat stressed above) arises from the quadratic mixing term $h\, t$ in the Lagrangian.

We will henceforth work with comoving fields, $\Psi_{ij} \equiv a\, M_{Pl}\, h_{ij}/2$, and $u \equiv a\, \delta\chi$. The relation between the comoving curvature perturbation ζ and the axion field fluctuations, $\delta\chi$ is given, at leading order in $\dot\chi/\dot\phi$, by

$$\zeta^\chi = \frac{U_\chi}{V_\phi}\left(\frac{H}{\dot\phi}\right)\delta\chi\,, \tag{3.1}$$

where $U_\chi \equiv \partial_\chi U$ and $V_\phi \equiv \partial_\phi V$. The scalar-tensor-tensor bispectrum then reads

$$\langle h_{\mathbf{p}}(\tau)\, h_{\mathbf{q}}(\tau)\, \zeta^\chi_{\mathbf{k}}(\tau)\rangle = -\frac{4}{a^3(\tau) M_{Pl}^2}\left(\frac{U_\chi}{V_\phi}\right)\left(\frac{H}{\dot\phi}\right)\langle \Psi_{\mathbf{p}}(\tau)\, \Psi_{\mathbf{q}}(\tau)\, u_{\mathbf{k}}(\tau)\rangle\,. \tag{3.2}$$

3.1 Perturbative solutions

Tensor perturbations are expanded in Fourier space as

$$\mathcal{T}_{ij}(\mathbf{x}, t) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot\mathbf{x}} \sum_{\lambda=R,L} e_{ij}^\lambda(\hat k)\, \mathcal{T}_{\mathbf{k}}^\lambda(t)\,, \tag{3.3}$$

[6] It is important to note here that the diagram in Fig. 1 is meant as a pictorial reminder of the fields and interactions in play but should not be intended as the exact *in-in* formalism representation of the calculation. This is because the presence of the two-fields vertex in yellow requires a specific hierarchy between the interaction it represents and the rest of the quadratic action: $\delta\mathcal{L}^{(2)}_{\rm yellow} \ll \mathcal{L}^{(2)}_{\rm rest}$. Such inequality is not satisfied at all times therefore a consistent calculation entails either diagonalizing the system to avoid quadratic interactions or the use of Green's functions methods. We adopt the latter.

where $\mathcal{T}^{\lambda}_{\mathbf{k}}(t) = \mathcal{T}^{(1)\lambda}_{\mathbf{k}}(t) + \mathcal{T}^{(2)\lambda}_{\mathbf{k}}(t) + \dots$, and dots indicate higher-order terms in the perturbative expansion. In the equation just above $\mathcal{T}$ is a placeholder for Ψ_{ij} as well as for t_{ij}. This expansion will be convenient in light of the Green's function method, which we adopt throughout this manuscript[7]. Similarly, for the scalar field we have

$$u(\mathbf{x}, t) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot\mathbf{x}} u_{\mathbf{k}}(t)\,, \tag{3.4}$$

with $u_{\mathbf{k}}(t) = u^{(1)}_{\mathbf{k}}(t) + u^{(2)}_{\mathbf{k}}(t) + \dots$. To lowest order in the perturbative expansion, one finds

$$\langle \Psi_{\mathbf{p}}\, \Psi_{\mathbf{q}}\, u_{\mathbf{k}} \rangle = \langle \Psi^{(1)}_{\mathbf{p}}\, \Psi^{(1)}_{\mathbf{q}}\, u^{(2)}_{\mathbf{k}} \rangle + \langle \Psi^{(1)}_{\mathbf{p}}\, \Psi^{(2)}_{\mathbf{q}}\, u^{(1)}_{\mathbf{k}} \rangle + \langle \Psi^{(2)}_{\mathbf{p}}\, \Psi^{(1)}_{\mathbf{q}}\, u^{(1)}_{\mathbf{k}} \rangle\,. \tag{3.5}$$

We are interested in the non-Gaussianity arising from the gauge field effect on the metric tensor perturbations. The latter are linearly sourced by the $SU(2)$ tensor fluctuations, with one helicity acquiring a larger amplitude than the other, as reviewed in *Section* 2. We will focus here on the leading helicity mode, setting $\Psi = \Psi^R$ from now on.

Formally, the equation of motion for the metric tensor fluctuations reads

$$\mathcal{O}_{\Psi}\, \Psi_{\mathbf{k}} = \mathcal{S}^{\Psi}_{\mathbf{k}}\,, \tag{3.6}$$

where $\mathcal{O}_{\Psi}$ is the operator describing the homogeneous equation of motion, $\mathcal{O}_{\Psi}\Psi_{\mathbf{k}} = 0$, while $\mathcal{S}^{\Psi}$ acts as a source due to self-couplings as well as to interactions with other fields. The solution to (3.6) at i^{th} order takes the form

$$\Psi^{(i)}_{\mathbf{k}}(\tau) = \int_{\infty}^{\tau} d\tau'\, \mathcal{G}^{\Psi}_{\mathbf{k}}(\tau, \tau')\, \mathcal{S}^{(i)\Psi}_{\mathbf{k}}(\tau')\,, \tag{3.7}$$

where $\mathcal{G}^{\Psi}_{\mathbf{k}}$ is the Green's function of Ψ and $\mathcal{S}^{(i)\Psi}_{\mathbf{k}}$ the source term, expanded at the same order. The leading-order terms relevant for the diagram in Fig. 1, and included in the expansion (3.5), are

$$\Psi^{(1)}_{\mathbf{k}}(\tau) = \int d\tau'\, \mathcal{G}^{\Psi}_{\mathbf{k}}(\tau, \tau')\, \mathcal{S}^{[\Psi^{(1)}\cdot t^{(1)}]}_{\mathbf{k}}(\tau')\,, \tag{3.8}$$

$$\Psi^{(2)}_{\mathbf{k}}(\tau) = \int d\tau'\, \mathcal{G}^{\Psi}_{\mathbf{k}}(\tau, \tau')\, \mathcal{S}^{[t^{(2)}]}_{\mathbf{k}}(\tau')\,. \tag{3.9}$$

The quantity $\mathcal{S}^{[\Psi^{(1)}\cdot t^{(1)}]}_{\mathbf{k}}$ originates from the quadratic interactions between tensor modes in the metric and the fluctuations of the gauge fields. $\mathcal{S}^{[t^{(2)}]}_{\mathbf{k}}$ is the source term for Ψ due to the second order perturbation in the gauge field, specifically the one corresponding to the terms $\delta\chi\, t^2$ in the cubic Lagrangian. More explicitly, for the gauge field one has

$$\mathcal{O}_t\, t_{\mathbf{k}} = \mathcal{S}^{t}_{\mathbf{k}}\,, \tag{3.10}$$

[7] Alternatively, one may switch to a different basis to decouple, up to second order, $SU(2)$ fields from standard tensor modes and then employ the in-in formalism. The two approaches are equivalent.

and

$$t^{(i)}_{\mathbf{k}}(\tau) = \int d\tau' \, \mathcal{G}^{t}_{\mathbf{k}}(\tau,\tau') \, \mathcal{S}^{(i)t}_{\mathbf{k}}(\tau') \,, \tag{3.11}$$

where $\mathcal{G}^{t}_{\mathbf{k}}$ is the Green's function for t. The relevant contribution to $t^{(2)}$ is given by[8]

$$t^{(2)}_{\mathbf{k}}(\tau) = \int d\tau' \, \mathcal{G}^{t}_{\mathbf{k}}(\tau,\tau') \, \mathcal{S}^{[t^{(1)}\cdot u^{(1)}]}_{\mathbf{k}}(\tau') \,. \tag{3.12}$$

The linear order $t^{(1)}$ is sourced by $\Psi^{(1)}$. However, it can be shown that the homogeneous solution for t is a good approximation for the full 1st order solution up to late times [19]. This approximation works well because the transient growth experienced by one of the polarizations of gauge field fluctuations is inherently due to the coupling to the axion field, which is already manifest at the level of the homogeneous equation of motion for the gauge field.
Once free fields are quantized, the corresponding sourced fields inherit the same set of creation/annihilation operators. Indicating t^{λ}_{k} as the solution to the homogenous equation of motion for the $SU(2)$ tensor modes, one may write the field operator as

$$t^{(1)\lambda}_{\mathbf{k}}(\tau) = a^{\lambda}_{\mathbf{k}} \, t^{\lambda}_{k}(\tau) + a^{\lambda\dagger}_{-\mathbf{k}} \, [t^{\lambda}_{k}(\tau)]^{*} \,, \tag{3.13}$$

where $[a^{\lambda_1}_{\mathbf{k_1}}, a^{\lambda_2\dagger}_{-\mathbf{k_2}}] = \delta_{\lambda_1\lambda_2} \, \delta^{(3)}(\mathbf{k_1} + \mathbf{k_2})$.

Given the equation of motion for the scalar field, $\mathcal{O}_u \, u_{\mathbf{k}} = \mathcal{S}^{u}_{\mathbf{k}}$, one derives

$$u^{(i)}_{\mathbf{k}}(\tau) = \int d\tau' \, \mathcal{G}^{u}_{\mathbf{k}}(\tau,\tau') \, \mathcal{S}^{(i)u}_{\mathbf{k}}(\tau') \,, \tag{3.14}$$

with $\mathcal{G}^{u}_{\mathbf{k}}$ the Green's function for u. With $u^{(1)}$ the free-field, the expression for $u^{(2)}$ is given by

$$u^{(2)}_{\mathbf{k}}(\tau) = \int d\tau' \, \mathcal{G}^{u}_{\mathbf{k}}(\tau,\tau') \, \mathcal{S}^{[t^{(1)}\cdot t^{(1)}]}_{\mathbf{k}}(\tau') \,, \tag{3.15}$$

where $\mathcal{S}^{[t^{(1)}\cdot t^{(1)}]}_{\mathbf{k}}$ is, once again, obtained from the cubic Lagrangian in $\delta\chi \, t^2$.

Let us begin by focusing on $\langle \Psi^{(1)}_{\mathbf{p}} \, \Psi^{(1)}_{\mathbf{q}} \, u^{(2)}_{\mathbf{k}} \rangle$. To derive $u^{(2)}$, one expands the Chern-Simons interaction to third order

$$\mathcal{S}^{(3)}_{\rm CS} = \frac{\lambda}{2f} \int d^4x \left\{ - g \, a \, Q \, \delta\dot{\chi} \, (t_{ij})^2 + 2 \, \delta\chi \, \dot{t}_{ia} \, \epsilon^{0ijk} \, \partial_j t_{ka} \right\} \,. \tag{3.16}$$

[8] The one detailed in Eq. (3.12) is expected to be the leading contribution to the second-order tensor fluctuations of the gauge fields. There exist also metric tensor contributions to $t^{(2)}$, however the coupling λ/f is typically much stronger than gravitational interactions.

The corresponding equation of motion for u reads (more details on the derivation can be found in Appendix A)

$$u_{\mathbf{k}}^{''}+\left(a^2m_\chi^2+k^2-\frac{a^{''}}{a}\right)u_{\mathbf{k}}=\frac{\lambda}{2f}\int\frac{d^3k_2}{(2\pi)^3}\sum_{\lambda_1\lambda_2}e_{ij}^{\lambda_1}(\hat{k}_1)e_{ij}^{\lambda_2}(\hat{k}_2)$$
$$\times\Big\{aH\Big[g\sqrt{\epsilon_E}\,M_{Pl}t_{\mathbf{k_1}}^{\lambda_1}\,t_{\mathbf{k_2}}^{\lambda_2}$$
$$+\,m_Q\frac{d}{dt}\left(t_{\mathbf{k_1}}^{\lambda_1}\,t_{\mathbf{k_2}}^{\lambda_2}\right)\Big]-2\,k_2\,\frac{dt_{\mathbf{k_1}}^{\lambda_1}}{dt}\,t_{\mathbf{k_2}}^{\lambda_2}\Big\}\,. \tag{3.17}$$

The Green's function for u in the limit of negligible mass for the axion and in the regime $k\tau\to 0$, is given by

$$\mathcal{G}_{\mathbf{k}}^{u}(\tau,\tau')=\frac{\theta(\tau-\tau')}{k^3\,\tau\,\tau'}\left(k\,\tau'\cos k\tau'-\sin k\tau'\right)\,. \tag{3.18}$$

Notice that, in the massless limit for the axion, one has $\mathcal{G}_{\mathbf{k}}^{u}=\mathcal{G}_{\mathbf{k}}^{\Psi}$. The final expression for $u_{\mathbf{k}}^{(2)}$ is given by Eq. (3.15), with $\mathcal{S}_{\mathbf{k}}^{[t^{(1)}\cdot t^{(1)}]}$ equal to the right-hand side of Eq. (3.17). Combining Eqs. (3.8) and (3.15), after summing over all permutations, one obtains the final result

$$\langle\Psi_{\mathbf{p}}^{(1)R}\,\Psi_{\mathbf{q}}^{(1)R}\,u_{\mathbf{k}}^{(2)}\rangle\simeq(2\pi)^3\delta^{(3)}(\mathbf{p}+\mathbf{q}+\mathbf{k})\left(\frac{-\lambda}{2\,f}\right)e_{ij}^{R}(-\mathbf{p})e_{ij}^{R}(-\mathbf{q})\int d\tau^{'}\mathcal{G}_{\mathbf{p}}^{\Psi}(\tau,\tau^{'})\,D_p(\tau^{'})$$
$$\times\int d\tau^{''}\mathcal{G}_{\mathbf{q}}^{\Psi}(\tau,\tau^{''})\,D_q(\tau^{''})\int d\tau^{'''}\mathcal{G}_{\mathbf{k}}^{u}(\tau,\tau^{'''})\Big\{\mathcal{A}(p,q,\tau^{'},\tau^{''},\tau^{'''})$$
$$+\,\mathcal{B}(p,q,\tau^{'},\tau^{''},\tau^{'''})+\mathcal{C}(p,q,\tau^{'},\tau^{''},\tau^{'''})\Big\}\,, \tag{3.19}$$

where D_p is the differential operator defining the quadratic mixing between tensor modes of the metric and tensor perturbations of the gauge field, $D_p(\tau)\equiv\frac{2\sqrt{\epsilon_B}}{m_Q\tau}\partial_\tau+\frac{2\sqrt{\epsilon_B}}{\tau^2}(m_Q+p\,\tau)$ (see also Eq. (B.2)). Note that in writing Eq. (3.19) we defined

$$\mathcal{A}(p,q,\tau^{'},\tau^{''},\tau^{'''})\equiv\sqrt{\epsilon_E}\,g\;M_{Pl}\,\frac{2}{\tau^{'''}}\mathrm{Re}\Big\{t_p(\tau^{'})t_q^{*}(\tau^{''})t_p^{*}(\tau^{'''})t_q(\tau^{'''}) \tag{3.20}$$
$$+2\,t_p(\tau^{'})t_q(\tau^{''})t_p^{*}(\tau^{'''})t_q^{*}(\tau^{'''})\Big\}\,,$$

$$\mathcal{B}(p,q,\tau^{'},\tau^{''},\tau^{'''})\equiv-H\,m_Q\,\mathrm{Re}\Big\{t_p(\tau^{'})t_q^{*}(\tau^{''})\left[t_p^{'*}(\tau^{'''})t_q(\tau^{'''})+t_p^{*}(\tau^{'''})t_q^{'}(\tau^{'''})\right] \tag{3.21}$$
$$+2\,t_p(\tau^{'})t_q(\tau^{''})\left[t_p^{'*}(\tau^{'''})t_q^{*}(\tau^{'''})+t_p^{*}(\tau^{'''})t_q^{'*}(\tau^{'''})\right]\Big\}\,,$$

$$\mathcal{C}(p,q,\tau^{'},\tau^{''},\tau^{'''})\equiv-2\,H\,\tau^{'''}\mathrm{Re}\Big\{t_p(\tau^{'})t_q^{*}(\tau^{''})\,q\,t_p^{'*}(\tau^{'''})t_q(\tau^{'''}) \tag{3.22}$$
$$+t_p(\tau^{'})t_q^{*}(\tau^{''})\,p\,t_q^{'}(\tau^{'''})t_p^{*}(\tau^{'''})$$
$$+2\,t_p(\tau^{'})t_q(\tau^{''})\,q\,t_p^{'*}(\tau^{'''})t_q^{*}(\tau^{'''})$$
$$+2\,t_p(\tau^{'})t_q(\tau^{''})\,p\,t_q^{'*}(\tau^{'''})t_p^{*}(\tau^{'''})\Big\}\,,$$

where "Re" stands for real part and the index R on the mode function for the gauge fields has been dropped for simplicity.
Before proceeding any further, one ought to point out that the amplitudes stemming from the three contributions in Eq. (3.5) are all parametrically similar to one another (see Eq. B.8). However, the structure of the contributions with $\Psi^{(2)}$ in Eq. (3.5) is different from those with $u^{(2)}$ in that they entail a double time-integral rather than products of independent integrals (see Appendix B for the explicit expressions). Nevertheless, for the reasons outlined above, we expect a similar shape, i.e. with a peak in the equilateral configuration. In the remainder of the section we focus on the $u^{(2)}$ contribution, but we stress that the discussion on the final results applies to both contributions.

3.2 Amplitudes and shapes

We report in Fig. 2 the three contributions, (3.20) through (3.22), to the scalar-tensor-tensor bispectrum for a sample set of parameters. The sum of the three terms is also shown. As anticipated, the shape profile peaks in the equilateral configuration.
Let us now move on to the bispectrum amplitude. It is instructive to report here the three contributions labelled $\mathcal{A}, \mathcal{B}, \mathcal{C}$:

$$\langle h^{(1)}\, h^{(1)}\, \zeta^{(2)\chi}\rangle_{\mathcal{A},\mathcal{B},\mathcal{C}} \sim \left(\frac{H}{M_{Pl}}\right)^3 \left(\frac{U_\chi}{V_\phi}\right) \frac{\epsilon_B}{\sqrt{\epsilon_\phi}}\, \mathcal{I}_{\mathcal{A},\mathcal{B},\mathcal{C}} \times \left(\int G^{\Psi}\cdot t\right)^2 \left(\int G^{u}\cdot t\cdot t\right) , \tag{3.23}$$

where

$$\mathcal{I}_{\mathcal{A}} \equiv \frac{\lambda\, g\, \sqrt{\epsilon_E}\, M_{Pl}}{f} , \qquad \mathcal{I}_{\mathcal{B}} \equiv \frac{m_Q\, \lambda\, H}{f} , \qquad \mathcal{I}_{\mathcal{C}} \equiv \frac{\lambda\, H}{f} . \tag{3.24}$$

Before elaborating further on the magnitude of the bispectrum, we take a quick detour to discuss the power spectrum contributions arising from the same interactions, i.e. $\langle \zeta^{(2)\chi}\, \zeta^{(2)\chi}\rangle$. This is a one-loop correction to the tree-level scalar power spectrum. The consistency of the perturbative expansion rests on the fact that such contribution, as well as those at higher loops, is sub-leading with respect to the tree-level observable. This fact will also be reflected on the bispectrum. Using Eqs. (3.20)-(3.22), one finds that the amplitude of the bispectrum (3.23) can be expressed in the following form

$$\langle h^{(1)}\, h^{(1)}\, \zeta^{(2)\chi}\rangle_{\mathcal{A},\mathcal{B},\mathcal{C}} \sim \mathcal{P}_h^{\rm s} \left(\mathcal{P}_\zeta^{\rm tree}\cdot \Delta^2\right)^{1/2} \tag{3.25}$$

where we have defined $\mathcal{P}_\zeta^{\rm tree}$ as the scale-invariant tree-level power spectrum of curvature fluctuations and Δ^2 as the correction introduced by the loop, i.e. $\langle \zeta^{(2)\chi}\, \zeta^{(2)\chi}\rangle \simeq \mathcal{P}_\zeta^{\rm tree}\cdot\Delta^2$. In deriving (3.25), the expression for $\mathcal{P}_h^{\rm s}$ from Eq. (2.6) has been used. Parametrizing the scalar-tensor-tensor non-Gaussianity as $f_{\rm nl}\equiv \mathcal{B}_{hh\zeta}/\mathcal{P}_\zeta^2$, one obtains

$$f_{\rm nl}^{\rm SU(2)} \sim r^2\cdot \frac{\mathcal{P}_h^{\rm s}}{\mathcal{P}_h^{\rm tot}} \left(\frac{\Delta^2}{r^2\, \mathcal{P}_\zeta^{\rm tree}}\right)^{1/2} . \tag{3.26}$$

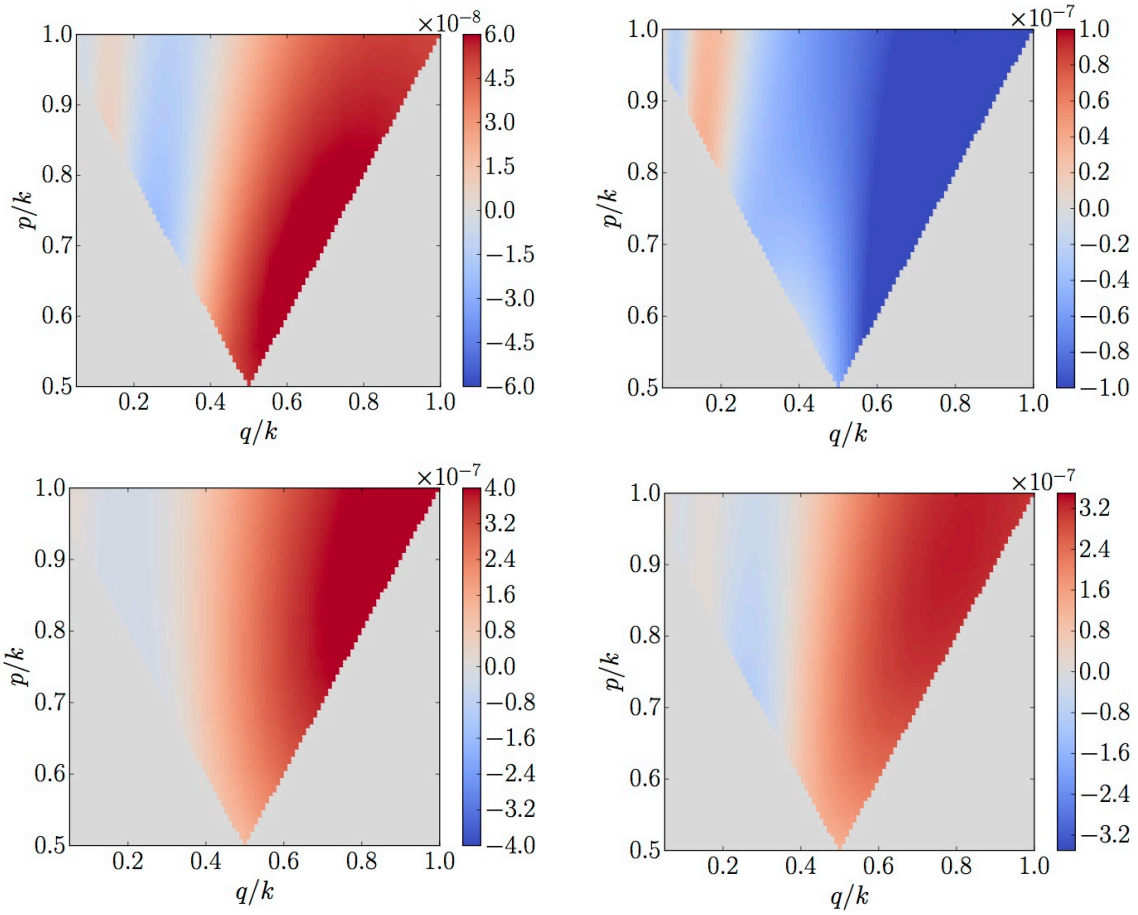

Figure 2. Plot of the contributions to $(pqk)^2\mathcal{B}_{hh\zeta}$ arising from the $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ terms in Eq. (3.23) (upper panels and lower left panel), and plot of the total $(pqk)^2\mathcal{B}_{hh\zeta}$ (lower right panel). The following set of parameters has been chosen in generating the plots above: $m_Q = 3.45$, $\epsilon_B = 3 \times 10^{-5}$, $\epsilon_\phi = 3 \times 10^{-3}$, $\epsilon_\chi = 3 \times 10^{-8}$, $f = 10^{-2}\, M_{Pl}$, $g = 10^{-2}$. We considered these particular values for the parameters for the sake of comparison with the existing literature. However, for a parameter region that has been filtered through the lenses of perturbativity bounds see *Section* 3.2 and *Appendix* C. Naturally, we expect that the shape, unlike the amplitude, does not depend on the choice of these parameters.

One may verify that, for $\Delta^2 < 10^{-5/3}$, the parameter space of the model supports a one-loop contribution to the scalar power spectrum that is subdominant with respect to the tree-level contribution. Under the same condition, the scalar non-Gaussianity arising from these interactions remains below the upper bounds from Planck for the amplitude associated with the equilateral template and a sizable $\mathcal{P}_h^{\rm s} \gtrsim \mathcal{P}_h^{\rm vacuum}$ is allowed (see Appendix C for more details).

It is useful at this stage to compare the result in (3.26) to the scalar-tensor-tensor non-Gaussianity in standard single field inflation. From [35], and using the above definition for $f_{\rm nl}$, one finds $f_{\rm nl}^{\rm standard} \sim r^2$. On the other hand, the parameter space of the model

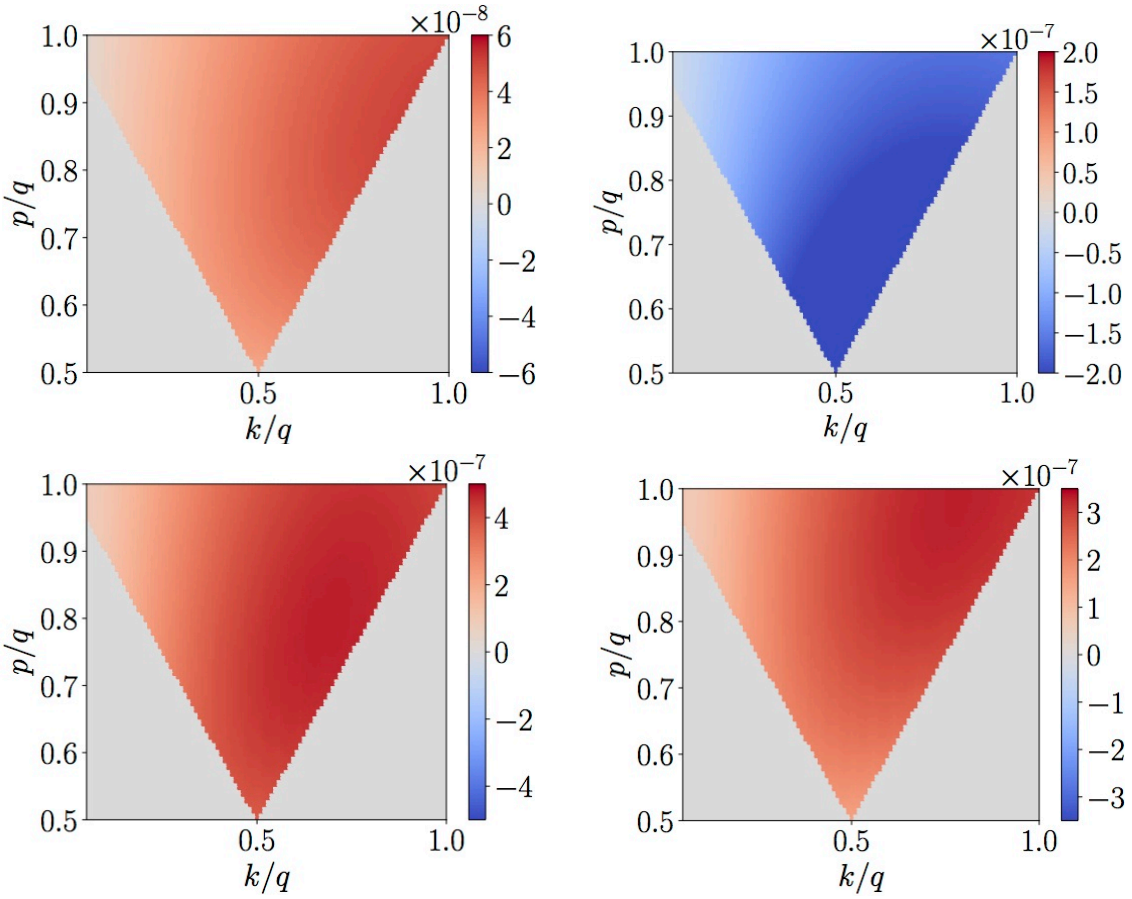

Figure 3. Same as Fig. 2, except for choosing as variables k/q and p/q (scalar fluctuation and one of the tensor fluctuation modes respectively).

we have been studying allows for a bispectrum as large as

$$f_{\rm nl}^{\rm gauge} \sim \frac{10^3}{r} f_{\rm nl}^{\rm standard} \,. \tag{3.27}$$

In deriving (3.27), which represents the maximum tensor-tensor-scalar amplitude the model allows, we have maximized the right-hand-side of Eq. (3.26) by setting $\mathcal{P}_h^s \simeq \mathcal{P}_h^{\rm tot}$, i.e. considering a tensor power spectrum dominated by the sourced contribution, and taking the upper limit allowed by the bounds on the scalar bispectrum[9], i.e. $\Delta^2 = 10^{-5/3}$. Using the results of Appendix C, one can verify that these conditions are easily satisfied by taking, for instance, $m_Q \simeq 2$ and $\epsilon_B \simeq$ (a few)$\epsilon_\chi = 7\times10^{-3}$.

The perturbativity bound notwithstanding, the scalar-tensor-tensor non-Gaussianity from SU(2) gauge fields shows a remarkable enhancement with respect to the standard

[9]Notice that Eq. (3.26) was been derived under the assumption that the scalar power spectrum is dominated by its tree-level contribution. This is entirely compatible with the chosen value of Δ^2 (see also Appendix A for more details).

result.

3.3 Tensor-scalar-scalar bispectrum

For the sake of completeness, we provide below an estimate of the SU(2) contribution to the tensor-scalar-scalar (tss) three-point function of the model. We leave a more thorough treatment to future work. There are two relevant diagrams contributing to the tss bispectrum that originate from the Chern-Simons interaction and from metric tensor-gauge field interactions:

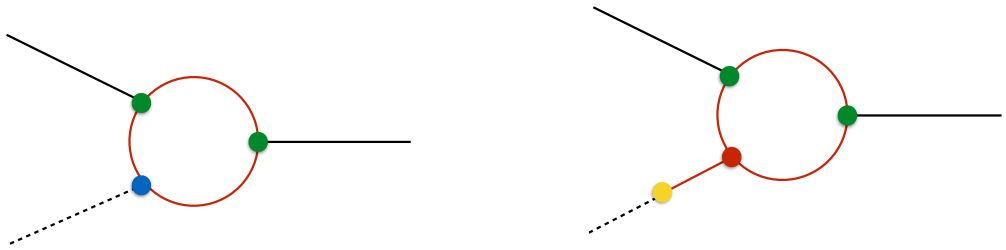

Figure 4. Black dotted lines stand for propagators of the metric tensor perturbation h; red lines represent the gauge field tensor perturbation t; the solid black line indicates the curvature perturbation, ζ^{χ}. The green vertex stems from the Chern-Simons contribution to the $\delta\chi\, t^2$ interaction, the yellow vertex from the quadratic Lagrangian in $h\,t$ (usual caveats apply), the blue and the red vertices from the cubic Lagrangian respectively in $h\,t^2$ and t^3.

Let us call (a) the diagram on the left and (b) the one of the right of Fig. 4. We report below the estimate for the corresponding amplitudes:

$$\mathcal{B}^{(a)}_{h\zeta\zeta} \approx \mathcal{P}^{1\text{loop}}_{\zeta}\, c^{(ii)} \left(\frac{H}{M_{Pl}}\right) (t\cdot t) \simeq \mathcal{P}^{1\text{loop}}_{\zeta}\, m_Q \left(\frac{H}{M_{Pl}}\right)^2 (t\cdot t) \simeq \mathcal{P}^{1\text{loop}}_{\zeta}\, \mathcal{P}^{\rm s}_h\, \frac{m_Q}{\epsilon_B}\,, \tag{3.28}$$

$$\begin{aligned}\mathcal{B}^{(b)}_{h\zeta\zeta} &\approx \mathcal{P}^{1\text{loop}}_{\zeta}\, \left(\mathcal{P}^{\rm s}_h\right)^{1/2} c^{(i)}\, (t\cdot t\cdot t) \simeq \mathcal{P}^{1\text{loop}}_{\zeta}\, \left(\mathcal{P}^{\rm s}_h\right)^{1/2} \frac{m_Q^2}{\sqrt{\epsilon_B}} \left(\frac{H}{M_{Pl}}\right) (t\cdot t\cdot t) \\ &\simeq \mathcal{P}^{1\text{loop}}_{\zeta}\, \mathcal{P}^{\rm s}_h\, m_Q\, \frac{m_Q}{\epsilon_B}\, (t\cdot t)\,,\end{aligned} \tag{3.29}$$

where the coefficients $c^{(i)}$ and $c^{(ii)}$ are the coupling constants characterizing the cubic Lagrangian for tensor fluctuations (see Eqs. (5)-(7) of [23]). In Eqs. (3.28)-(3.29) we also used the fact that the SU(2)-sourced tensor power spectrum, $P^{\rm s}_h$, is proportional to two (integrated) t mode-functions.

Introducing $f_{\rm nl} \equiv \mathcal{B}_{h\zeta\zeta}/\left(\mathcal{P}_\zeta\right)^2$, one finds

$$f^{(a)}_{\rm nl} \sim r \left(\frac{\mathcal{P}^{\rm s}_h}{\mathcal{P}^{\rm total}_h}\right) \frac{\Delta^2\, m_Q}{\epsilon_B}\,, \tag{3.30}$$

$$f^{(b)}_{\rm nl} \sim r \left(\frac{\mathcal{P}^{\rm s}_h}{\mathcal{P}^{\rm total}_h}\right) \frac{\Delta^2\, m_Q^2\, (t\cdot t)}{\epsilon_B}\,, \tag{3.31}$$

where $\mathcal{P}_\zeta^{\rm 1loop} \equiv \mathcal{P}_\zeta^{\rm tree}\,\Delta^2$. The amplitude above are to be compared with the result from standard single field inflation [35]

$$f_{\rm nl}^{\rm standard} \approx r\,. \tag{3.32}$$

Taking the limiting values $\Delta^2 = 10^{-5/3}$, $\mathcal{P}_h^{\rm s} = \mathcal{P}_h^{\rm total}$ and setting $(t\cdot t) \sim e^{3.6\,m_Q}$, the amplitudes in Eqs. (3.30)-(3.31) become

$$f_{\rm nl}^{(a)} \sim 10^{-2}\,r\cdot\frac{m_Q}{\epsilon_B} \approx 100\cdot r\,, \tag{3.33}$$

$$f_{\rm nl}^{(b)} \sim 10^{-2}\,r\cdot\frac{m_Q^2\,e^{3.6\,m_Q}}{\epsilon_B} \approx 10^{7}\cdot r\,, \tag{3.34}$$

where in the last step the sample values $m_Q = 3$, $\epsilon_B = 10^{-4}$ have been used to provide a concrete comparison with the standard case. We pause here to stress that, unlike for the scalar-tensor-tensor bispectrum, the results in this subsection are to be considered estimates and need to be confirmed by a full calculation. Since this observable is not the main focus of the paper, we leave a more thorough analysis to future work.

4 Conclusions

The model studied here belongs to an important class of theories characterized by a sourced gravitational waves signal in excess of tensor vacuum fluctuations. The analysis of the dynamics and the signatures of similar set-ups represents a cautionary tale against the temptation to immediately read off the inflationary energy scale directly from the value of the tensor-to-scalar ratio. The distinctive signatures of the SU(2)-equipped model [19] includes a blue or otherwise bumpy chiral gravitational waves power spectrum to a level accessible by upcoming observations [19, 22], along with enhanced tensor non-Gaussianity [23]. Mixed tensor-scalar non-Gaussianities are just as important. These provide additional predictive power, which is crucial to help constrain the model parameters.

In this paper we derive predictions for the scalar-tensor-tensor bispectrum and focus in particular on the effects of the axion-SU(2) fields coupling. We find that the three-point function is significantly enhanced with respect to its counterpart in the minimal inflationary scenario. Our focus has been on the impact on observables of a controlled growth in the gauge tensor modes near horizon crossing. Given that this dynamics is essentially localized at the horizon, the resulting shape function is expected to peak in the equilateral configuration. This is indeed the outcome of our analysis, as shown in Figs.(2-3).

The work presented here can be extended in a number of directions. It would be important to generate forecasts detailing the constraining power that upcoming experiments will have on mixed non-Gaussianity. A full analysis of the shape function also entails the comparison with existing templates in order to help distinguish this class of models from other scenarios.

Our results call for detailed studies of the scalar sector of the theory resulting from the choice of a specific inflaton Lagrangian, $\mathcal{L}_\phi$. Perhaps most importantly, for a

complete characterization of this and similar models it is essential to study the post-inflationary evolution of the axion and the gauge fields.

Acknowledgments

ED and MF are delighted to thank E. Komatsu for illuminating conversations and kind encouragement. ED would like to thank the Perimeter Institute for Theoretical Physics (Canada) for hospitality and support whilst this work was in progress. ED is supported in part by DOE grant DE-SC0009946. HA, MF, KK and DW are supported by STFC grant ST/N000668/1. RJH is supported by UK Science and Technology Facilities Council grant ST/N5044245. The work of KK has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No.646702 "CosTesGrav").

A Derivation of the second-order equation of motion for u

The equation of motion in real space, considering only the relevant source term, reads

$$u^{''} + \left(a^2 m_\chi^2 - \partial^2 - \frac{a^{''}}{a}\right) u = \frac{\lambda}{2f}\left[\frac{d}{dt}\left(g\, a\, Q\, t_{ij}^2\right) + 2\,\frac{dt_{ia}}{dt}\,\epsilon^{0ijk}\partial_j t_{ka}\right], \tag{A.1}$$

where $\partial^2 \equiv \delta^{ij}\partial_i\partial_j$, $m_\chi^2 \equiv d^2U/d\chi^2|_{\chi=\bar{\chi}}$ and $' \equiv d/d\tau$.

In momentum space one finds

$$u_{\mathbf{k}}^{''} + \left(a^2 m_\chi^2 + k^2 - \frac{a^{''}}{a}\right) u_{\mathbf{k}} = \frac{\lambda}{2f}\int \frac{d^3k_2}{(2\pi)^3}\sum_{\lambda_1\lambda_2}\left\{e_{ij}^{\lambda_1}(\hat{k}_1)e_{ij}^{\lambda_2}(\hat{k}_2)\left[\frac{d}{dt}\left(g\, a\, Q\, t_{\mathbf{k_1}}^{\lambda_1}\, t_{\mathbf{k_2}}^{\lambda_2}\right)\right]\right.$$
$$\left. + e_{ia}^{\lambda_1}(\hat{k}_1)e_{ka}^{\lambda_2}(\hat{k}_2)\, 2\,\epsilon^{0ijk}\, i\, k_{2j}\,\frac{dt_{\mathbf{k_1}}^{\lambda_1}}{dt}\, t_{\mathbf{k_2}}^{\lambda_2}\right\}, \tag{A.2}$$

where $\mathbf{k_1} \equiv \mathbf{k} - \mathbf{k_2}$. Using the relation $i\,\epsilon^{ijk}\, k_i\, e_{j\ell}^{\lambda} = \pm k e_{k\ell}^{\lambda}$, where $+$ is for $\lambda = L$ and $-$ is for $\lambda = R$, and the definitions for ϵ_E and m_Q, one arrives at Eq. (3.17).

B $\langle \Psi^{(2)}\Psi^{(1)}u^{(1)}\rangle$ computation

We present here our derivation of the contribution from $\langle \Psi_{\mathbf{q}}^{(2)}\Psi_{\mathbf{p}}^{(1)}u_{\mathbf{k}}^{(1)}\rangle$ to the scalar-tensor-tensor correlation:

$$\langle h_{\mathbf{q}}^{(2)}(\tau)h_{\mathbf{p}}^{(1)}(\tau)\zeta_{\mathbf{k}}^{(1)\chi}(\tau)\rangle = -\frac{4}{a^3(\tau)M_{Pl}^2}\left(\frac{U_\chi}{V_\phi}\right)\left(\frac{H}{\dot{\phi}}\right)\langle \Psi_{\mathbf{q}}^{(2)}(\tau)\Psi_{\mathbf{p}}^{(1)}(\tau)u_{\mathbf{k}}^{(1)}(\tau)\rangle\,, \tag{B.1}$$

where

$$\Psi_{\mathbf{p}}^{(1)}(\tau) = \int_{-\infty}^{0} d\tau^{'}\mathcal{G}_p^{(\Psi)}(\tau,\tau^{'})\, D_p(\tau^{'})t_p^{(1)}(\tau^{'})\,, \tag{B.2}$$

and

$$u_{\mathbf{k}}^{(1)}(\tau) = a_{\mathbf{k}} u_k(\tau) + a_{\mathbf{k}}^{\dagger} u_k^*(\tau)\,, \qquad\qquad u_k(\tau) = -\frac{1}{\sqrt{2\,k^3}\tau}(1+ik\tau)e^{-ik\tau}\,. \tag{B.3}$$

The metric fluctuation to second order reads

$$\Psi_{\mathbf{q}}^{(2)}(\tau) = \int d\tau^{'}\mathcal{G}_q^{\Psi}(\tau,\tau^{'})\,D_q(\tau^{'})t_p^{(2)}(\tau^{'})\,, \tag{B.4}$$

where

$$t_{\mathbf{q}}^{(2)}(\tau^{'}) = \int_{-\infty}^{\tau^{'}} d\tau^{''}\mathcal{G}_q^{t}(\tau^{'},\tau^{''})\mathcal{J}_{\mathbf{q}}^{(t)}(\tau^{''})\,, \tag{B.5}$$

and $\mathcal{G}_q^t$ is the Green's function for the gauge field tensor fluctuations. $\mathcal{J}_{\mathbf{q}}^{(t)}$ is the source term appearing in the second-order equation of motion for the gauge field, specifically the one due to the $\delta\chi\cdot t\cdot t$ Chern-Simons interaction:

$$t_{\mathbf{q}}^{''} + \left[k^2 + \frac{2}{\tau^2}\left(1 + m_Q^2 + k\tau\left(2m_Q + m_Q^{-1}\right)\right)\right] t_{\mathbf{q}} = \mathcal{J}_{\mathbf{q}}^{(t)}\,, \tag{B.6}$$

where $'$ indicates the derivative w.r.t. conformal time τ and

$$\mathcal{J}_{\mathbf{q}}^{(t)} \equiv -\frac{\lambda}{f}\epsilon_{\alpha\beta}^{R}(-\hat{q})\int\frac{d^3k_1}{(2\pi)^3}\sum_{\lambda=L,R}\epsilon_{\alpha\beta}^{\lambda}(\hat{k}_1)\left[\left(\frac{gQ}{H\tau}\pm k_1\right)t_{\mathbf{k}_1}^{\lambda}\,\delta\chi_{\mathbf{k}_2}^{'} + (q\pm k_1)\,t_{\mathbf{k}_1}^{\lambda'}\,\delta\chi_{\mathbf{k}_2}\right]. \tag{B.7}$$

Here $\mathbf{k}_2 = \mathbf{q} - \mathbf{k}_1$ and $\pm$ correspond, respectively, to $\lambda = L,\,R$. After performing the Wick contractions, one arrives at (for one permutation)

$$\begin{aligned}
\langle\hat{\Psi}_{\mathbf{q}}^{(2)}(\tau)\hat{\Psi}_{\mathbf{p}}^{(1)}(\tau)\delta\hat{\chi}_{\mathbf{k}}^{(1)}(\tau)\rangle \;=\; & (2\pi)^3\delta^{(3)}(\mathbf{q}+\mathbf{p}+\mathbf{k})\left(-\frac{\lambda}{f}\right) \\
& \times\int_{-\infty}^{\tau}d\tau_1\,\mathcal{G}_q^{\Psi}(\tau,\tau_1)D_q(\tau_1)\int_{-\infty}^{\tau_1}d\tau_2\,\mathcal{G}_q^{t}(\tau_1,\tau_2) \\
& \times a(\tau)\left[\frac{(p+q)^2-k^2}{4pq}\right]^2\int_{-\infty}^{\tau}d\tau_3\,\mathcal{G}_p^{\Psi}(\tau,\tau_3)D_p(\tau_3) \\
& \times\Big[\left(\frac{gQ}{H\tau_2}-p\right)\delta\chi_k^{'}(\tau_2)\delta\chi_k^{*}(\tau)t_p^{R}(\tau_2)t_p^{R*}(\tau_3) \\
& +\,(q-p)\,\delta\chi_k(\tau_2)\delta\chi_k^{*}(\tau)t_p^{R'}(\tau_2)t_p^{R*}(\tau_3)\Big]\,.
\end{aligned} \tag{B.8}$$

C Bounds from perturbativity and from scalar non-Gaussianity

We estimate here the one-loop power spectrum arising from the same interactions contributing to the tensor-tensor-scalar bispectrum analyzed in this paper:

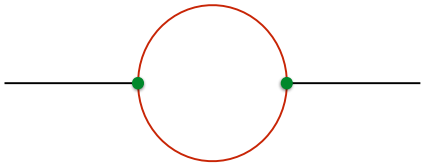

Figure 5. Diagrammatic representation of $P_\zeta^{1\text{loop}}$ arising from the Chern-Simons interaction. Red lines represent the gauge field tensor perturbation, t, solid black lines represents the curvature perturbation, ζ. The green vertex arises from the Chern-Simons contribution to $\delta\chi \cdot t \cdot t$.

One finds (schematically)

$$\mathcal{P}_\zeta^{1\text{loop}} \simeq \left(\frac{H}{\sqrt{\epsilon_\phi}\, M_{Pl}}\right)^2 \left(\frac{U_\chi}{V_\phi}\right)^2 (\mathcal{I}_\mathcal{A} + \mathcal{I}_\mathcal{B} + \mathcal{I}_\mathcal{C})^2 \left(\int G^u \cdot t \cdot t\right)^2 , \tag{C.1}$$

where $\mathcal{I}_{\mathcal{A},\mathcal{B},\mathcal{C}}$ were introduced in Eqs. (3.24) and we defined $\langle \zeta_{\mathbf{k}}^{(2)\chi}\, \zeta_{\mathbf{q}}^{(2)\chi} \rangle \simeq (2\pi)^3 \delta^{(3)}(\mathbf{k}+\mathbf{q})(k^3/2\pi^2)\mathcal{P}_\zeta^{1\text{loop}}(k)$. Let us now take a closer look at the parameters in the model to put (C.1) in a more explicit form. To this aim, we will make use of Eqs. (2.2)-(2.3), which we report below

$$Q = \left(\frac{-f\, U_\chi}{3g\lambda H}\right)^{1/3} , \qquad \frac{\lambda}{2fH}\dot{\chi} \simeq m_Q + \frac{1}{m_Q} . \tag{C.2}$$

We also remind the reader that $m_Q \equiv g\, Q/H$. In the slow-roll regime for Q, one finds $\epsilon_E \approx \epsilon_B/m_Q^2$. From (C.2) and from the definition of the slow-roll parameters (see Sec. 2) it follows that $(\lambda\, M_{Pl}/f)^2 = (m_Q + m_Q^{-1})^2/\epsilon_\chi$. One also finds $g = (m_Q^2\, H)/(\sqrt{\epsilon_B}\, M_{Pl})$. From the field equations, assuming a standard background equation of motion for the inflaton, one also has

$$\frac{U_\chi}{V_\phi} \approx \frac{\lambda\, M_{Pl}}{f}\sqrt{\frac{\epsilon_E \epsilon_B}{\epsilon_\phi}} . \tag{C.3}$$

Equipped with all of the above, one finds $\mathcal{P}_\zeta^{1\text{loop}} \approx \mathcal{P}_\zeta^{\text{tree}} \cdot \Delta^2$ where, schematically, we define $\Delta^2 = (\Delta_\mathcal{A} + \Delta_\mathcal{B} + \Delta_\mathcal{C})^2$, with

$$\Delta_i \approx 10^4 \cdot e^{3.6\, m_Q} \cdot \frac{\epsilon_B}{\epsilon_\chi \epsilon_\phi} \left(m_Q + m_Q^{-1}\right)^2 \left(\frac{H}{M_{Pl}}\right)^2 \times \begin{cases} 1, & i = \mathcal{A},\, B \\ m_Q^{-1}, & i = \mathcal{C} \end{cases} \tag{C.4}$$

It is straightforward to verify that the bound from scalar non-Gaussianity, i.e. from $\langle \zeta_{\mathbf{k}}^{(2)\chi}\, \zeta_{\mathbf{q}}^{(2)\chi}\, \zeta_{\mathbf{p}}^{(2)\chi} \rangle \propto (\mathcal{P}_\zeta^{1\text{loop}})^{3/2}$, is given by $\Delta^2 \lesssim 10^{-5/3}$. In deriving the latter bound we considered current Planck constraints on equilateral non-Gaussianity. The saturation of this bound, which is slightly more stringent than the perturbativity bound on the power spectrum, corresponds to the amplitude reported in Eq. (3.27).

One can easily verify that the perturbativity bound from the tensor power spectrum are weaker than those from scalars. Using Eqs. (B.4)-(B.5) and (B.7), and the aforementioned bounds from scalar non-Gaussianity, one arrives at

$$\frac{\langle h^{(2)}h^{(2)}\rangle}{\langle h^{(1)}h^{(1)}\rangle} \lesssim 10^{-1}\left(\frac{\epsilon_\phi\epsilon_\chi}{\epsilon_B^2}\right), \tag{C.5}$$

where $\langle h^{(2)}h^{(2)}\rangle \sim (1/M_{\rm Pl}^2)\langle\Psi^{(2)}\Psi^{(2)}\rangle$ and $\langle h^{(1)}h^{(1)}\rangle$ is the sourced tree-level tensor power spectrum from Eq. (2.6). Given the aforementioned hierarchy among the slow-roll parameters in (C.5), the sourced one-loop power spectrum is a few orders of magnitude smaller than the tree-level one.

References

[1] A. H. Guth, [Phys. Rev. D **23**, 347 (1981)]; D. H. Lyth and A. Riotto, Phys. Rept. **314**, 1 (1999) [hep-ph/9807278]; A. Riotto, ICTP Lect. Notes Ser. **14**, 317 (2003) [hep-ph/0210162]; W. H. Kinney, NATO Sci. Ser. II **123**, 189 (2003) [astro-ph/0301448]; D. Wands, Lect. Notes Phys. **738**, 275 (2008) [arXiv:0702187]; D. Baumann, [arXiv:0907.5424].

[2] P. A. R. Ade *et al.* [Planck Collaboration], Astron. Astrophys. **571**, A22 (2014) [arXiv:1303.5082]; P. A. R. Ade *et al.* [Planck Collaboration], Astron. Astrophys. **571**, A24 (2014) [arXiv:1303.5084]; P. A. R. Ade *et al.* [Planck Collaboration], Astron. Astrophys. **594**, A17 (2016) [1502.01592]; P. A. R. Ade *et al.* [Planck Collaboration], Astron. Astrophys. **594**, A20 (2016) [1502.02114].

[3] D. Jeong and M. Kamionkowski, Phys. Rev. Lett. **108**, 251301 (2012) [arXiv:1203.0302]; V. Assassi, D. Baumann and D. Green, JCAP **1211**, 047 (2012) [arXiv:1204.4207]; L. Dai, D. Jeong and M. Kamionkowski, Phys. Rev. D **87**, no. 10, 103006 (2013) [arXiv:1302.1868]; L. Dai, D. Jeong and M. Kamionkowski, Phys. Rev. D **88**, no. 4, 043507 (2013) [arXiv:1306.3985]; S. Brahma, E. Nelson and S. Shandera, Phys. Rev. D **89**, no. 2, 023507 (2014) [arXiv:1310.0471]; E. Dimastrogiovanni, M. Fasiello, D. Jeong and M. Kamionkowski, JCAP **1412**, 050 (2014) [arXiv:1407.8204]; E. Dimastrogiovanni, M. Fasiello and M. Kamionkowski, JCAP **1602**, 017 (2016) [arXiv:1504.05993]; R. Emami and H. Firouzjahi, JCAP **1510**, no. 10, 043 (2015) [arXiv:1506.00958]; X. Chen and Y. Wang, Phys. Rev. D **81**, 063511 (2010) [arXiv:0909.0496]; X. Chen and Y. Wang, JCAP **1004**, 027 (2010) [arXiv:0911.3380 [hep-th]]. [arXiv:0911.3380]; H. Lee, D. Baumann and G. L. Pimentel, JHEP **1612** (2016) 040 [arXiv:1607.03735]; M. Biagetti, E. Dimastrogiovanni and M. Fasiello, JCAP **1710**, no. 10, 038 (2017) [arXiv:1708.01587]; E. Dimastrogiovanni, M. Fasiello and G. Tasinato, [arXiv:1806.00850].

[4] N. Arkani-Hamed and J. Maldacena, [arXiv:1503.08043].

[5] K. Freese, J. A. Frieman and A. V. Olinto, [Phys. Rev. Lett. **65**, 3233 (1990)]; K. Freese and W. H. Kinney, Phys. Rev. D **70**, 083512 (2004) [hep-ph/0404012]; K. Freese and W. H. Kinney, JCAP **1503**, 044 (2015) [arXiv:1403.5277].

[6] E. Pajer and M. Peloso, Class. Quant. Grav. **30**, 214002 (2013) [arXiv:1305.3557].

[7] R. Kallosh, A. D. Linde, D. A. Linde and L. Susskind, Phys. Rev. D **52**, 912 (1995) [hep-th/9502069]; T. Banks, M. Dine, P. J. Fox and E. Gorbatov, JCAP **0306**, 001 (2003) [hep-th/0303252].

[8] J. E. Kim, H. P. Nilles and M. Peloso, JCAP **0501**, 005 (2005) [hep-ph/0409138]; S. Dimopoulos, S. Kachru, J. McGreevy and J. G. Wacker, JCAP **0808**, 003 (2008) [hep-th/0507205]; R. Kallosh, N. Sivanandam and M. Soroush, Phys. Rev. D **77**, 043501 (2008) [arXiv:0710.3429].

[9] M. M. Anber and L. Sorbo, Phys. Rev. D **81**, 043534 (2010) [arXiv:0908.4089]; N. Barnaby and M. Peloso, Phys. Rev. Lett. **106**, 181301 (2011) [arXiv:1011.1500]; L. Sorbo, JCAP **1106**, 003 (2011) [arXiv:1101.1525]; N. Barnaby, R. Namba and M. Peloso, JCAP **1104**, 009 (2011) [arXiv:1102.4333]; J. L. Cook and L. Sorbo, Phys. Rev. D **85**, 023534 (2012) [Erratum-ibid. D **86**, 069901 (2012)] [arXiv:1109.0022]; N. Barnaby, J. Moxon, R. Namba, M. Peloso, G. Shiu and P. Zhou, Phys. Rev. D **86**, 103508 (2012) [arXiv:1206.6117]; S. Mukohyama, R. Namba, M. Peloso and G. Shiu, JCAP **1408**, 036 (2014) [arXiv:1405.0346]; R. Z. Ferreira and M. S. Sloth, JHEP **1412**, 139 (2014) [arXiv:1409.5799]; O. Özsoy, K. Sinha and S. Watson, Phys. Rev. D **91**, no. 10, 103509 (2015) [arXiv:1410.0016]; N. Bartolo, S. Matarrese, M. Peloso and M. Shiraishi, JCAP **1501**, no. 01, 027 (2015) [arXiv:1411.2521]; P. Adshead, J. T. Giblin, T. R. Scully and E. I. Sfakianakis, JCAP **1512**, no. 12, 034 (2015) [arXiv:1502.06506]; N. Bartolo, S. Matarrese, M. Peloso and M. Shiraishi, JCAP **1507**, no. 07, 039 (2015) [arXiv:1505.02193]; R. Namba, M. Peloso, M. Shiraishi, L. Sorbo and C. Unal, JCAP **1601**, no. 01, 041 (2016) [arXiv:1509.07521]; R. Z. Ferreira, J. Ganc, J. Norea and M. S. Sloth, JCAP **1604**, no. 04, 039 (2016) [arXiv:1512.06116]; M. Peloso, L. Sorbo and C. Unal, [arXiv:1606.00459]; P. Adshead, J. T. Giblin, T. R. Scully and E. I. Sfakianakis, JCAP **1610**, 039 (2016) [arXiv:1606.08474]; J. Garcia-Bellido, M. Peloso and C. Unal, JCAP **1612**, no. 12, 031 (2016) [arXiv:1610.03763]; O. Özsoy, JCAP **1804**, no. 04, 062 (2018) [arXiv:1712.01991]; T. Fujita, I. Obata, T. Tanaka and S. Yokoyama, [arXiv:1801.02778].

[10] R. Z. Ferreira and A. Notari, JCAP **1709**, no. 09, 007 (2017) [arXiv:1706.00373]; R. Z. Ferreira and A. Notari, Phys. Rev. D **97**, no. 6, 063528 (2018) [arXiv:1711.07483];

[11] M. Biagetti, M. Fasiello and A. Riotto, Phys. Rev. D **88**, 103518 (2013) [arXiv:1305.7241]; M. Biagetti, E. Dimastrogiovanni, M. Fasiello and M. Peloso, JCAP **1504**, 011 (2015) [arXiv:1411.3029]; T. Fujita, J. Yokoyama and S. Yokoyama, PTEP **2015**, 043E01 (2015) [arXiv:1411.3658].

[12] S. Dubovsky, R. Flauger, A. Starobinsky and I. Tkachev, Phys. Rev. D **81**, 023523 (2010) [arXiv:0907.1658]; G. Cusin, R. Durrer, P. Guarato and M. Motta, JCAP **1505**, no. 05, 030 (2015) [arXiv:1412.5979]; M. Fasiello and R. H. Ribeiro, JCAP **1507**, no. 07, 027 (2015) [arXiv:1505.00404].

[13] R. R. Caldwell and C. Devulder, Phys. Rev. D **97**, no. 2, 023532 (2018) [arXiv:1706.03765];

[14] P. Adshead and M. Wyman, Phys. Rev. Lett. **108**, 261302 (2012) [arXiv:1202.2366]; P. Adshead and M. Wyman, Phys. Rev. D **86**, 043530 (2012) [arXiv:1203.2264]; E. Martinec, P. Adshead and M. Wyman, JHEP **1302**, 027 (2013) [arXiv:1206.2889]; E. Dimastrogiovanni, M. Fasiello and A. J. Tolley, JCAP **1302**, 046 (2013) [arXiv:1211.1396].

[15] K. D. Lozanov, A. Maleknejad and E. Komatsu, [arXiv:1805.09318].

[16] A. Maleknejad and M. M. Sheikh-Jabbari, Phys. Lett. B **723**, 224 (2013) [arXiv:1102.1513]; A. Maleknejad and M. M. Sheikh-Jabbari, Phys. Rev. D **84**, 043515 (2011) [arXiv:1102.1932]; A. Maleknejad and E. Erfani, JCAP **1403**, 016 (2014) [arXiv:1311.3361]; A. Bhattacharjee, A. Deshamukhya and S. Panda, Mod. Phys. Lett. A **30**, no. 11, 1550040 (2015) [arXiv:1406.5858]; I. Obata, T. Miura and J. Soda, Phys. Rev. D **92**, no. 6, 063516 (2015) [arXiv:1412.7620]; I. Obata *et al.*, Phys. Rev. D **93**, no. 12, 123502 (2016) [arXiv:1602.06024].

[17] J. T. Deskins, J. T. Giblin and R. R. Caldwell, Phys. Rev. D **88**, no. 6, 063530 (2013) [arXiv:1305.7226]; A. Maleknejad, Phys. Rev. D **90**, no. 2, 023542 (2014) [arXiv:1401.7628]; J. Bielefeld and R. R. Caldwell, Phys. Rev. D **91**, no. 12, 123501 (2015) [arXiv:1412.6104]; J. Bielefeld and R. R. Caldwell, Phys. Rev. D **91**, no. 12, 124004 (2015) [arXiv:1503.05222]; A. Maleknejad, JCAP **1612**, no. 12, 027 (2016) [arXiv:1604.06520]; R. R. Caldwell, C. Devulder and N. A. Maksimova, [arXiv:1604.08939]; R. R. Caldwell and C. Devulder, Phys. Rev. D **97**, no. 2, 023532 (2018) [arXiv:1706.03765].

[18] E. Dimastrogiovanni and M. Peloso, Phys. Rev. D **87**, no. 10, 103501 (2013) [arXiv:1212.5184];

P. Adshead, E. Martinec and M. Wyman, JHEP **1309**, 087 (2013) [arXiv:1305.2930];
R. Namba, E. Dimastrogiovanni and M. Peloso, JCAP **1311**, 045 (2013) [arXiv:1308.1366];
P. Adshead, E. Martinec and M. Wyman, Phys. Rev. D **88**, no. 2, 021302 (2013)
[arXiv:1301.2598].

[19] E. Dimastrogiovanni, M. Fasiello and T. Fujita, JCAP **1701**, no. 01, 019 (2017)
[arXiv:1608.04216].

[20] I. Obata *et al.*, Phys. Rev. D **93**, no. 12, 123502 (2016) [arXiv:1602.06024].

[21] G. Dall'Agata, [arXiv:1804.03104]; E. McDonough and S. Alexander, [arXiv:1806.05684].

[22] B. Thorne, T. Fujita, M. Hazumi, N. Katayama, E. Komatsu and M. Shiraishi, Phys. Rev. D
97, no. 4, 043506 (2018) [arXiv:1707.03240].

[23] A. Agrawal, T. Fujita and E. Komatsu, [arXiv:1707.03023]; A. Agrawal, T. Fujita and
E. Komatsu, [arXiv:1802.09284]; A. Agrawal, [arXiv:1804.01481].

[24] N. Bartolo and G. Orlando, JCAP **1707**, 034 (2017) [arXiv:1706.04627].

[25] N. Bartolo, G. Orlando and M. Shiraishi, [arXiv:1809.11170].

[26] M. Shiraishi, D. Nitta and S. Yokoyama, Prog. Theor. Phys. **126**, 937 (2011) [arXiv:1108.0175].

[27] M. Shiraishi, A. Ricciardone and S. Saga, JCAP **1311**, 051 (2013) [arXiv:1308.6769].

[28] N. Bartolo, S. Matarrese, M. Peloso and M. Shiraishi, JCAP **1501**, no. 01, 027 (2015)
[arXiv:1411.2521].

[29] N. Bartolo, S. Matarrese, M. Peloso and M. Shiraishi, JCAP **1507**, no. 07, 039 (2015) [arXiv:
1505.02193].

[30] R. Namba, M. Peloso, M. Shiraishi, L. Sorbo and C. Unal, JCAP **1601**, no. 01, 041 (2016)
[arXiv:1509.07521].

[31] K. N. Abazajian *et al.* [CMB-S4 Collaboration], [arXiv:1610.02743].

[32] M. Hazumi et al. (LiteBIRD), Proc. SPIE Int. Soc. Opt. Eng. 8442, 844219 (2012).

[33] N. Bartolo *et al.*, JCAP **1612** (2016) no.12, 026 [arXiv:1610.06481]; C. Caprini and
D. G. Figueroa, [arXiv:1801.04268].

[34] T. Fujita, R. Namba and Y. Tada, Phys. Lett. B **778**, 17 (2018) [arXiv:1705.01533].

[35] J. M. Maldacena, JHEP **0305**, 013 (2003) [astro-ph/0210603].