New Astronomy 43 (2016) 26–36

Contents lists available at ScienceDirect

New Astronomy

journal homepage: www.elsevier.com/locate/newast

Cosmic bandits: Exploration versus exploitation in CMB B-mode experiments

Ely D. Kovetz [a,b,*], Marc Kamionkowski [b]

[a] *Theory Group, Department of Physics and Texas Cosmology Center, The University of Texas at Austin, TX 78712, USA*
[b] *Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

HIGHLIGHTS

- There is a tradeoff between exploration and exploitation in ground-based B-mode surveys.
- Machine-learning algorithms used elsewhere can be adopted for devising adaptive survey strategies.
- Efficient adaptive survey strategies can be used to steer clear of high-foreground sky patches.
- Using an adaptive survey strategy, sensitivity to primordial B-modes can be improved by a factor of 2 and 3.

ARTICLE INFO

Article history:
Received 24 January 2015
Revised 3 June 2015
Accepted 28 July 2015
Available online 3 August 2015

Communicated by J. Silk

Keywords:
(cosmology:) cosmic background radiation
cosmology: observations
methods: observational
methods: statistical

ABSTRACT

A preferred method to detect the curl-component, or B-mode, signature of inflationary gravitational waves (IGWs) in the cosmic microwave background (CMB) polarization, in the absence of foregrounds and lensing, is a prolonged integration over a single patch of sky of a few square degrees. In practice, however, foregrounds abound and the sensitivity to B modes can be improved considerably by finding the region of sky cleanest of foregrounds. The best strategy to detect B modes thus involves a tradeoff between exploration (to find lower-foreground patches) and exploitation (through prolonged integration). This problem is akin to the multi-armed bandit (MAB) problem in probability theory, wherein a gambler faces a series of slot machines with unknown winning odds and must develop a strategy to maximize his/her winnings with some finite number of pulls. While the optimal MAB strategy remains to be determined, a number of algorithms have been developed in an effort to maximize the winnings. Here, based on this resemblance, we tackle the search for IGW B modes with single frequency experiments in the presence of spatially varying foregrounds by developing adaptive survey strategies to optimize the sensitivity to IGW B modes. We demonstrate, using realistic foreground models and taking lensing-induced B modes into account, that adaptive experiments can substantially improve the upper bound on the tensor-to-scalar ratio (by factors of 2 and 3 in single frequency experiments, and possibly even more). Similar techniques can be applied to other surveys, including 21-cm measurements of signatures of the epoch of reionization, searches for a stochastic primordial gravitational wave background, deep-field imaging by the James Webb Space Telescope or various radio interferometers, and transient follow-up searches.

1. Introduction

Cosmology has become a science of surveys. Ever larger surveys are used to seek ever-more-subtle correlations to shed light on novel early-Universe phenomena or the physics of galaxy formation. The separation of the signals of interest from similar ones due to astrophysical foregrounds requires more sensitive measurements and clever algorithms. The issue of foregrounds can also be dealt with by restricting the survey to "clean" regions, where the foregrounds are absent or at least smaller in amplitude. But finding these clean regions requires a search which may then take time away from integration on a single patch of sky. Optimization of the sensitivity to a given signal may thus involve a tradeoff between *exploration* of several patches of sky, to find the cleanest one, and *exploitation*, deep integration on a single patch. What is the best strategy, under these circumstances, to optimize the sensitivity to the signal?

This question is somewhat analogous to the multi-armed bandit (MAB) problem, a well-known problem from probability theory and machine learning in computer science (Berry and Fristedt, 1985; Press, 2009; Robbins, 1952). In this problem, a gambler is faced with a set of slot machines with different reward probability distributions

* Corresponding author at: Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA. Tel.: +1 5125477817; fax: 410 516 7184.
E-mail address: elykovetz@gmail.com (E.D. Kovetz).

http://dx.doi.org/10.1016/j.newast.2015.07.010
1384-1076/

and has to maximize the total reward in a given number of plays, or actions. This is a classic learning problem, as repeated plays allow the gambler to *learn* the distributions of the different machines, with a tradeoff between exploration and exploitation governed by the total number of allowed plays. A popular manifestation of this problem, which has garnered growing attention in recent decades, is clinical trials (Press, 2009), where rewards—in the form of survival/fatality—are of particular importance. Theoretical study of the MAB problem has led to several theorems regarding the ultimate prospects of solution methods in the asymptotic limit of infinite number of plays (Lai and Robbins, 1985). In realistic scenarios, however, with only a finite number of plays, one must resort to heuristic approaches, and over the years several classes of these have been suggested in the literature and compared empirically to some extent (Gittins et al., 2011; Kuleshov and Precup, in press; Sutton and Barton, 1998).

In this paper we focus on the search for the curl, or the B-mode, signature of inflationary gravitational waves (IGWs) in the cosmic microwave background (CMB) polarization (Kamionkowski and Kosowsky, 1999; Kamionkowski et al., 1997; Seljak and Zaldarriaga, 1997). These B modes are the target of a number of ongoing and forthcoming CMB-polarization experiments (Ade, 2008; Austermann et al., 2012; Bischoff, 2012; Eimer et al., 2012; Essinger-Hileman et al., 2010; Johnson et al., 2003; Keating et al., 2003; Kermish et al., 2012; Montroy et al., 2006; Niemack et al., 2010; North et al., 2008; Reichborn-Kjennerud et al., 2010; Sheehy et al., 2011)[1]. The strategy of many of these experiments is to integrate deeply on a small patch of sky, as this optimizes the sensitivity to IGW B modes in an experiment with fixed detector sensitivity, or noise-equivalent temperature (NET), and duration (Jaffe et al., 2000). Realistically, though, these experiments will have to contend with foreground emission from Galactic dust and synchrotron radiation (Clark et al., 2012; Fantaye et al., 2011; Kogut et al., 2007; O'Dea et al., 2012; Stivoli et al., 2010; Verde et al., 2006). Since the amplitudes of these foregrounds may vary considerably from one region of the sky to another (Adam, 2014; Clark et al., 2012; Fantaye et al., 2011; O'Dea et al., 2012), the sensitivity to IGWs may be improved considerably by integrating on the cleanest patch. While measurements (mostly unpolarized) at other frequencies can be used to steer the experimentalist toward a clean region of the sky (Kovetz and Kamionkowski, in preparation), the *polarized* foregrounds in the electromagnetic and spatial frequencies of interest have only been measured to poor accuracy in the cleanest regions of sky (Adam, 2014). One can thus do an initial exploration of a broad region to find clean patches (Kovetz and Kamionkowski, in preparation), but that then takes time away from exploiting any particular region. An important challenge is thus to balance the tradeoff between exploration and exploitation in an optimal way, given the limits set by instrumental properties (including the total observation time of the experiment) and the expected distribution of foreground noise on the sky.

The purpose of this paper is to present a method inspired by heuristic solutions to the MAB problem to optimally perform the integration over sky patches so that noise from polarized foregrounds is minimized and the strongest possible upper bound can be placed on the amplitude of IGW B modes. We consider several fiducial experiments with instrumental properties representative of current and next-generation experiments, all operating at a single frequency of 150 GHz (a value common to many of the leading B-mode experiments) and focus on the dominant foreground source at this frequency, which is polarized emission from dust (PED) in the galaxy.

In order to forecast the variation of this foreground source across the sky, we use the FGPol (O'Dea et al., 2012) foreground templates for PED. We perform simulations of different survey (bandit) strategies on patches of sky within a low-noise region accessible from the South Pole, for which PED amplitudes are randomly drawn from the FGPol template and calculate the improvement (or degradation) in the upper bound on the tensor-to-scalar ratio r. While our analysis makes a number of simplifications (although we *do* include lensing-induced B modes (Hanson et al., 2013; Zaldarriaga and Seljak, 1998), an essential ingredient), our results demonstrate that the adaptive survey strategies we consider provide considerable advantage over prolonged integration on naively chosen patches.

While our focus here is on CMB polarization, the methods described in this work can also be applied to other observations in cosmology and astrophysics, such as 21-cm measurements (Furlanetto et al., 2006; Morales and Wyithe, 2010; Pritchard and Loeb, 2012), searches for a stochastic primordial gravitational wave background (Adams and Cornish, 2013; 2010), deep-field telescope imaging (Stiavelli et al., 2009; Windhorst et al., 2006), and transient searches (Djorgovski et al., 2011). We discuss potential issues pertaining to such applications, but leave their full study to future work.

The plan of the paper is as follows: in Section 2 we describe the PED templates used in our analysis, discuss the instrumental noise of our fiducial experiments and present the statistical tools for estimating the errors in measurements of the relevant power spectra. In Section 3, we describe how we construct and test adaptive survey strategies based on machine-learning heuristics and explain our prescription for simulating adaptive B-mode experiments. We present our results in Section 4 and discuss several assumptions and possible additional implementations in Section 5. We conclude in Section 6.

2. PED foreground

In order to remove the different foreground contributions, most experiments operate at several frequencies and use component separation (Fantaye et al., 2011; Stivoli et al., 2010) or template-based techniques (Fantaye et al., 2012) to extract as clean a signal as possible. In any such process, residuals remain at some level and will hinder the ability to detect the desired signal.

The major contributions of polarized foreground noise in the relevant frequency range (∼20–300 GHz) of CMB experiments are sourced by PED in the Galaxy and by synchrotron radiation. Synchrotron is more dominant at lower frequencies ($\lesssim$100 GHz), while PED overwhelms the CMB at higher frequencies ($\gtrsim$100 GHz). The well-known CMB foreground "sweet spot" is around 90 GHz, where the noise sources are similar in amplitude and both are comparably low.

For simplicity, we shall address fiducial experiments operating at a single frequency of 150 GHz, which is adopted by many of the sub-orbital polarization experiments (see Kovetz and Kamionkowski in preparation for a complementary discussion of multi-frequency approaches). Therefore, PED would be the major source for concern in terms of foregrounds and its subtraction would be difficult. To estimate the sky variation of the PED power spectrum, we use the FGPOL templates (Clark et al., 2012; O'Dea et al., 2012).[2] These templates are based on a three-dimensional bi-symmetric spiral (O'Dea et al., 2012) model of the Galactic magnetic field (including the turbulent component) and are normalized according to the results of WMAP (Kogut et al., 2007) so that the average dust polarization fraction outside the WMAP P06 polarization mask (Page, 2007) is 3.6%. Prior to the release of polarization results from the Planck experiment, the best constraint on the dust polarization fraction at higher frequencies came from the partial sky ($f_{\rm sky} = 17\%$) measurement at 353 GHz of the ARCHEOPS balloon-borne experiment (Benoit et al., 2004), which detected a polarization fraction around 4–5% in the Galactic plane

[1] In the past year, first detections of B-mode polarization from lensing of E-modes were announced (Ade, 2014a; Hanson et al., 2013; Zaldarriaga and Seljak, 1998), followed by a detection on degree scales (Ade, 2014), whose source remains under dispute (whether it is primordial or due to foregrounds) (Ade, 2014b; Flauger et al., 2014; Mortonson and Seljak, 2014).

[2] http://www3.imperial.ac.uk/people/c.contaldi/fgpol.