

# Wavelet-Bayesian inference of cosmic strings embedded in the cosmic microwave background

J. D. McEwen<sup>1\*</sup>, S. M. Feeney<sup>2,3</sup>, H. V. Peiris<sup>4,5</sup>, Y. Wiaux<sup>6</sup>, C. Ringeval<sup>7</sup>  
and F. R. Bouchet<sup>8</sup>

<sup>1</sup>*Mullard Space Science Laboratory (MSSL), University College London, Surrey RH5 6NT, U.K.*

<sup>2</sup>*Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, U.K.*

<sup>3</sup>*Center for Computational Astrophysics, 160 5th Avenue, New York, NY 10010, USA*

<sup>4</sup>*Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.*

<sup>5</sup>*The Oskar Klein Centre for Cosmoparticle Physics, Stockholm University, Stockholm, Sweden*

<sup>6</sup>*Institute of Sensors, Signals, and Systems, Heriot-Watt University, Edinburgh EH14 4AS, U.K.*

<sup>7</sup>*Centre for Cosmology, Particle Physics and Phenomenology, Université Catholique de Louvain, Louvain-la-Neuve B-1348, Belgium*

<sup>8</sup>*Institut d'Astrophysique de Paris, Paris 75014, France*

Accepted —. Received —; in original form —

## ABSTRACT

Cosmic strings are a well-motivated extension to the standard cosmological model and could induce a subdominant component in the anisotropies of the cosmic microwave background (CMB), in addition to the standard inflationary component. The detection of strings, while observationally challenging, would provide a direct probe of physics at very high energy scales. We develop a framework for cosmic string inference from observations of the CMB made over the celestial sphere, performing a Bayesian analysis in wavelet space where the string-induced CMB component has distinct statistical properties to the standard inflationary component. Our wavelet-Bayesian framework provides a principled approach to compute the posterior distribution of the string tension  $G\mu$  and the Bayesian evidence ratio comparing the string model to the standard inflationary model. Furthermore, we present a technique to recover an estimate of any string-induced CMB map embedded in observational data. Using *Planck*-like simulations we demonstrate the application of our framework and evaluate its performance. The method is sensitive to  $G\mu \sim 5 \times 10^{-7}$  for Nambu-Goto string simulations that include an integrated Sachs-Wolfe (ISW) contribution only and do not include any recombination effects, before any parameters of the analysis are optimised. The sensitivity of the method compares favourably with other techniques applied to the same simulations.

**Key words:** cosmology: cosmic background radiation – cosmology: observations – methods: data analysis – methods: statistical.

## 1 INTRODUCTION

High-precision measurements of the anisotropies of the cosmic microwave background (CMB) strongly favour a standard cosmological model in which the large-scale structure of the Universe is seeded by nearly scale-invariant Gaussian density perturbations created during a phase of inflation (Hinshaw et al. 2013; Planck Collaboration XIII 2016). These measurements do, however, leave room for additional subdominant contributions to the CMB generated by processes beyond the standard inflationary paradigm. Cosmic strings represent a particularly well motivated extension to the standard model (for reviews see Brandenberger 1994; Vilenkin & Shellard 1994; Hindmarsh & Kibble 1995; Copeland & Kibble 2009). Arising in a range of attempts at Grand Unification, cos-

mic strings are linear topological defects produced when the Universe undergoes certain symmetry-breaking phase transitions. In an expanding Universe, the existence of causally separate regions prevents the symmetry from being broken in the same way throughout space, with a network of cosmic strings inevitably forming as a result (Kibble 1976). Such a string network cannot be solely responsible for producing the anisotropies of the CMB — cosmic strings cannot explain the acoustic peaks of the CMB power spectrum (Pen et al. 1997). However, cosmic strings could induce a subdominant contribution to the CMB through the Kaiser-Stebbins effect (Kaiser & Stebbins 1984), which induces a step-like (*i.e.*, highly non-Gaussian) temperature change between photons passing either side of a moving string. The magnitude of the contribution to the relative CMB temperature anisotropies from a *single, straight*

\* jason.mcewen@ucl.ac.uk

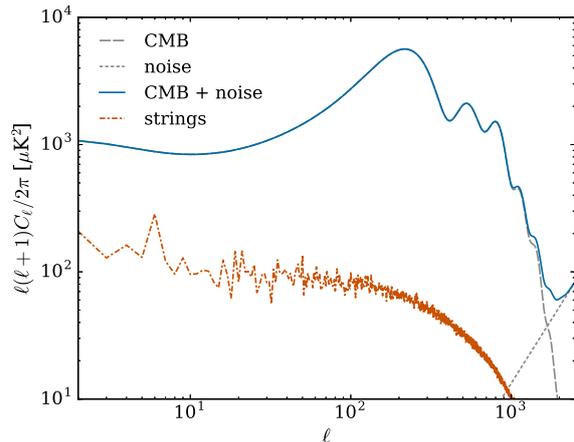
string is given by

$$\frac{\Delta T}{T_0} = 8\pi G\mu v\gamma_s, \quad (1)$$

where  $v$  is the transverse string velocity,  $\gamma_s$  is the corresponding relativistic gamma factor,  $G$  is the gravitational constant and  $\mu$  is the string tension (throughout we use natural units  $c = 1$ ). More generally, the situation is complicated by the existence of an evolving *network* of *wiggly* strings in an expanding universe. Calculating accurate observable effects of a network of cosmic strings is a rich and computationally demanding area of research (Albrecht & Turok 1989; Bennett & Bouchet 1989, 1990; Allen & Shellard 1990; Hindmarsh 1994; Bouchet et al. 1988; Vincent et al. 1998; Moore et al. 2002; Landriau & Shellard 2003; Ringeval et al. 2007; Fraisse et al. 2008; Landriau & Shellard 2011; Blanco-Pillado et al. 2011; Ringeval & Bouchet 2012), requiring the numerical evolution of the network in the presence of photons, matter and dark energy. Tools to simulate full-sky, high-resolution maps of string-induced CMB anisotropies incorporating all physical effects are not yet available. The current state-of-the-art methods produce matter-free simulations, which faithfully represent the small-scale structure imparted by the string network via the integrated Sachs-Wolfe (ISW) effect (Ringeval & Bouchet 2012) but do not include recombination effects. These simulations nevertheless remain computationally intensive, requiring hundreds of thousands of CPU hours to simulate a single full-sky map at *Planck* resolution.

The energy scale of the string-inducing phase transition  $\eta$  is directly related to the string tension  $\mu$  by  $\mu \sim \eta^2$ . Detecting the signatures of cosmic strings would therefore provide a direct probe of physics at extremely high energy scales. However, since any string signature must be subdominant, detecting strings is a significant observational challenge. The magnitude of the task is demonstrated in Fig. 1, in which we compare the power spectrum of a simulated string-induced CMB contribution (Ringeval & Bouchet 2012) (with amplitude close to current observational limits) and a standard inflationary component, as would be observed by *Planck*. Hereafter, we refer to string-induced CMB anisotropy maps with the shorthand “string maps”.

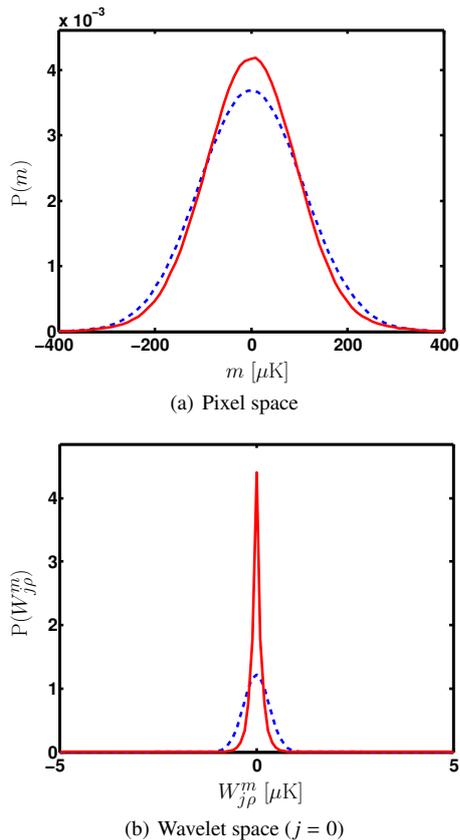
Various methods have been developed to search for string-induced contributions to the CMB, from power-spectrum constraints (Lizarraga et al. 2014a,b, 2016; Charnock et al. 2016), to higher-order statistics such as the bispectrum (Planck Collaboration XXV 2014; Regan & Hindmarsh 2015) and trispectrum (Fergusson et al. 2010), and tools such as edge detection (Lo & Wright 2005; Amsel et al. 2008; Stewart & Brandenberger 2009; Danos & Brandenberger 2010), Minkowski functionals (Gott et al. 1990; Ducout et al. 2013; Planck Collaboration XXV 2014), wavelets and curvelets (Starck et al. 2004; Hammond et al. 2009; Wiaux et al. 2010; Planck Collaboration XXV 2014; Hergt et al. 2016), level crossings (Sadegh Movahed & Khosravi 2011), and peak-peak correlations (Movahed et al. 2013). Current constraints on the string tension depend on the string model and simulation technique adopted. For Nambu-Goto strings, power spectrum analyses based on simulations computed by the unconnected segment model (USM; Albrecht et al. 1997, 1999; Pogosian & Vachaspati 1999) constrain the string tension to  $G\mu < 1.3 \times 10^{-7}$  (Planck Collaboration XXV 2014) using *Planck* temperature data and to  $G\mu < 1.1 \times 10^{-7}$  (Charnock et al. 2016) when *Planck* polarisation data are also included. Recombination effects have been considered by Regan & Hindmarsh (2015) but were found not to have a significant effect on the bispectrum. Beyond spectra, non-Gaussian analyses for Nambu-Goto strings—based on high-resolution sim-



**Figure 1.** Power spectra of the fiducial CMB, smoothed by the *Planck* beam and the HEALPIX pixel window function corresponding to  $N_{\text{side}} = 2048$  (grey, long dash), *Planck* instrumental noise (grey, short dash) and their sum (blue, solid). For comparison, a power spectrum estimated from a simulated full-sky cosmic string-induced CMB component (Ringeval & Bouchet 2012) is plotted, corresponding to  $G\mu = 5 \times 10^{-7}$  (orange, dot-dash). The string contribution is clearly subdominant, highlighting the challenge in constraining cosmic string models.

ulations of stringy CMB maps including only the ISW (integrated Sachs-Wolfe) contribution and no recombination effects (Ringeval & Bouchet 2012)—constrain the string tension to  $G\mu < 7.8 \times 10^{-7}$  (Planck Collaboration XXV 2014) using *Planck* temperature data. Considering only the ISW effect enables the production of high-resolution full-sky string maps, but these maps are necessarily conservative, and the resulting constraints are hence weaker. Furthermore, effects of recombination physics would increase the string anisotropy signal considerably (Planck Collaboration XXV 2014). While power spectrum statistics are inherently lossy, map-based analyses have the potential to better discriminate cosmic strings from other potential subdominant CMB signals.

As constraints on the amplitude of any string-induced component tighten, analysis techniques must become more sensitive to improve on the status quo. Wavelets are a particularly powerful tool for searching for cosmic strings due to their ability to simultaneously characterise signal structure in both scale and position. Furthermore, wavelets that are well-matched to the expected structure of string maps can be adopted, facilitating extraction of the string signal from the CMB and instrumental noise. Although string-induced CMB anisotropies are non-Gaussian, the statistical distribution of the pixels of a cosmic string map nevertheless remains close to Gaussian. In Fig. 2 histograms of simulated inflationary and cosmic string components are plotted, in both pixel (Fig. 2(a)) and wavelet (Fig. 2(b)) space. The shape of the distributions is reasonably similar in pixel space, whereas in wavelet space the distributions are markedly different. The distribution of the string component in wavelet space is highly peaked (*i.e.* leptokurtic) due to the sparse representation of the string component in wavelet space (*i.e.* due to the property that many of the wavelet coefficients of the string component are near zero). The inflationary CMB component, however, remains Gaussian distributed in wavelet space, since the wavelet transform is linear. The difference in the statistical properties of the string and inflationary CMB components in wavelet space can be exploited to isolate and estimate



**Figure 2.** Distributions of map  $m \in \{c, s\}$ , comprising either a CMB simulated inflationary component  $c$  (blue, dashed) or string component  $s$  (red, solid) CMB components, in both pixel space and scale-discretised wavelet space (for  $G\mu = 2 \times 10^{-6}$ ). The string-induced component is simulated by the method of Ringeval & Bouchet (2012). The shape of the distributions is reasonably similar in pixel space, whereas in wavelet space the distributions are markedly different.

the parameters of any string component. This is the approach taken in the current work.<sup>1</sup>

In this article we develop a hybrid wavelet-Bayesian approach to infer the presence and parameters of any cosmic string component in the CMB. We do not consider (insufficient) summary statistics like many alternative methods (*e.g.* the kurtosis), for which the origin of any non-Gaussian component cannot be rigorously determined. Instead, we learn and exploit the complex non-Gaussian structure of string-induced CMB contributions. We follow the approach of Hammond et al. (2009), generalising from the planar setting to the celestial sphere. In Hammond et al. (2009), techniques using planar wavelets are presented to learn the statistical structure of string-induced CMB contributions and to exploit this structure to recover an estimate of a planar map of the string component. We generalise these techniques to the full-sky setting using scale-discretised wavelets defined on the sphere (McEwen et al. 2015b; Leistedt et al. 2013; Wiaux et al. 2008), adopting direc-

<sup>1</sup> For very small scales, the underlying string distribution in pixel space becomes increasingly different to a Gaussian distribution in its tails, however these features can be washed out observationally by instrumental beams. In any case, for small instrumental beams that preserve these features the wavelet approach presented in the current work would characterise such structure, improving the sensitivity of the analysis.

tional wavelets with parameters selected to match the characteristic step-like temperature changes induced by strings in the CMB (Kaiser & Stebbins 1984). While Hammond et al. (2009) adopt a power spectrum approach to estimate the string tension, we recover the posterior distribution of the string tension in our wavelet formalism. Moreover, we also compute the Bayesian evidence to distinguish between the cosmic string model and the standard inflationary model. In summary, we present a principled and robust statistical framework based on Bayesian inference for parameter estimation and model selection, performing a Bayesian analysis in wavelet space where the inflationary and string induced CMB components have very different statistical properties. An overview of the string model and the recovery of any string-induced component is illustrated in Fig. 3, while an example of the scale-discretised wavelets considered is shown in Fig. 4.

We restrict our attention here to simulated observations modelling idealised *Planck* observations. An optimisation of the parameters of the method and the application to *Planck* data will be presented in a subsequent study. We employ the matter-free simulations of Ringeval & Bouchet (2012), which faithfully represent the small-scale string anisotropies produced by the ISW effect. As previously stated, these simulations come at a cost of hundreds of thousands of CPU hours per full-sky *Planck*-resolution map. Fortunately, our approach requires only two realisations of string maps: one to train our method and one to test it (see Fig. 6).

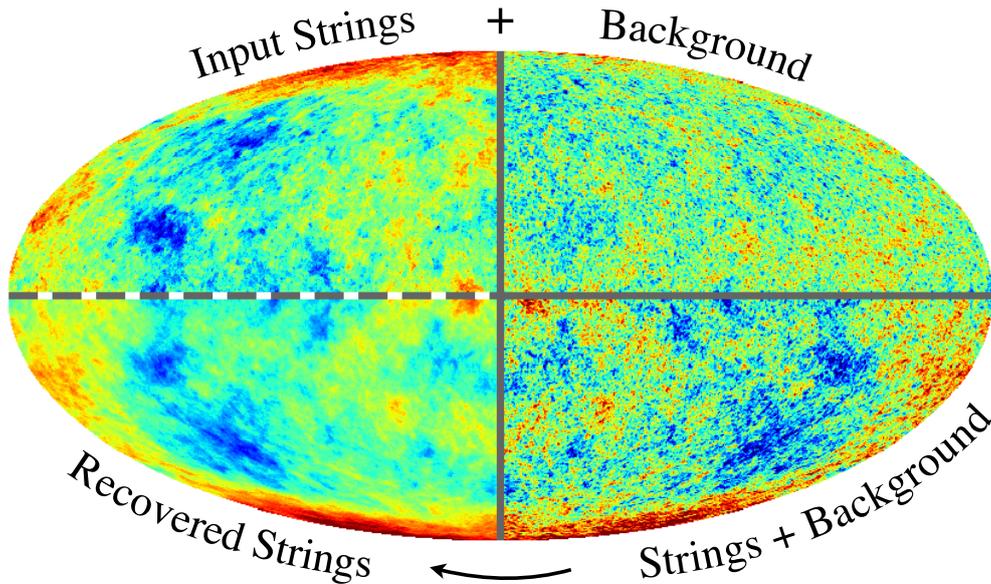
The outline of the paper is as follows. In Sec. 2, we review the wavelets used in this analysis. In Sec. 3 we describe in detail our hybrid wavelet-Bayesian framework for inferring the string tension and determining the Bayesian evidence for the string model relative to the standard inflationary model. We present in Sec. 4 the approach to recovering an estimate of the string-induced CMB component at the map level in the full-sky setting, which can be viewed as a Bayesian thresholding approach to denoising the observed CMB signal. In Sec. 5 we apply our framework to simulated observations and discuss the results. Concluding remarks are made in Sec. 6.

## 2 SCALE-DISCRETISED WAVELETS ON THE SPHERE

Wavelets on the sphere have found widespread use in analyses of the CMB (*e.g.* Vielva et al. 2004; McEwen et al. 2005, 2006b, 2008b; Vielva et al. 2006; McEwen et al. 2007b, 2008c; Feeney et al. 2011a,b; Feeney et al. 2012; Planck Collaboration XII 2014; Planck Collaboration XXIII 2014; Planck Collaboration XXV 2014) due to their ability to localise signal content in scale and space simultaneously (for a review see McEwen et al. 2007c).

Initial stable wavelet constructions on the sphere were based largely on continuous methodologies (*e.g.* Antoine & Vandergheynst 1998, 1999; Wiaux et al. 2005; Sanz et al. 2006; McEwen et al. 2006a), which do not support the exact synthesis of a sampled signal from its wavelet coefficients in practice. Consequently, cosmological analyses based on these constructions were limited to the analyses of wavelet coefficients; sampled signals on the sphere could not be accurately recovered from processed wavelet coefficients. Alternative discrete constructions based on the lifting scheme (Sweldens 1997) were developed (Schröder & Sweldens 1995; Barreiro et al. 2000; McEwen & Scaife 2008; McEwen et al. 2011), however these do not necessarily lead to a stable basis (Sweldens 1997).

More recently, a number of exact discrete wavelet frameworks on the sphere have been developed, with underlying contin-



**Figure 3.** Illustration of string map recovery. Under the string model, the CMB is comprised of a string-induced component (top left quadrant) and a Gaussian inflationary background component (top right quadrant), yielding the observed CMB map (bottom right quadrant). For visualisation purposes, in this simulation the string component is generated for a relatively large value of the string tension ( $G\mu = 2 \times 10^{-6}$ ), which controls the amplitude of the string map. We invert this process to recover an estimate of the input string component (bottom left quadrant). This is achieved by our hybrid wavelet-Bayesian analysis pipeline, which estimates the posterior distribution of the string tension, the Bayesian evidence for strings, and the mean-posterior string map. The same quadrant of each map is displayed in each panel, rotated and reflected as required.

uous representations and fast implementations that have been made available publicly, including needlets (Narcowich et al. 2006; Baldi et al. 2009; Marinucci et al. 2008), directional scale-discretised wavelets (Wiaux et al. 2008; Leistedt et al. 2013; McEwen et al. 2013), and the isotropic undecimated and pyramidal wavelet transforms (Starck et al. 2006). Each approach has also been extended to analyse spin functions on the sphere (Geller et al. 2008; Geller & Marinucci 2010, 2011; Geller et al. 2009; McEwen et al. 2015b, 2014; Starck et al. 2009) and functions defined on the three-dimensional ball formed by augmenting the sphere with the radial line (Durastanti et al. 2014; Leistedt & McEwen 2012; McEwen & Leistedt 2013; Leistedt et al. 2015; Lanusse et al. 2012). Ridgelet and curvelet wavelets on the sphere have also been constructed (Starck et al. 2006; McEwen 2017; Chan et al. 2016).

In this work we adopt directional scale-discretised wavelets (Wiaux et al. 2008; Leistedt et al. 2013; McEwen et al. 2013, 2015b, 2016), which are essentially the generalisation of needlets (Narcowich et al. 2006; Baldi et al. 2009; Marinucci et al. 2008) to directional wavelets (McEwen et al. 2016). Directional scale-discretised wavelets have recently been shown to satisfy quasi-exponential localisation and asymptotic uncorrelation properties similar to needlets (McEwen et al. 2016) and consequently have excellent spatial localisation properties.

In the remainder of this section we review directional scale-discretised wavelets concisely; for further details please see the related literature (Wiaux et al. 2008; Leistedt et al. 2013; McEwen et al. 2013, 2015b, 2016). The reader not interested in the details may safely skip the following subsections and simply note the notation used to denote wavelet coefficients specified in Eq. (3).

## 2.1 Wavelet transform and inversion

The scale-discretised wavelet transform of a function  $f \in L^2(\mathbb{S}^2)$  on the sphere  $\mathbb{S}^2$  is defined by the directional convolution of  $f$  with the wavelet  $\Psi^{(j)} \in L^2(\mathbb{S}^2)$ . In order to perform directional, spherical convolutions it is necessary to rotate functions on the sphere. The rotation operator  $\mathcal{R}_\rho$  is defined by

$$\Psi_{j\rho}(\omega) \equiv (\mathcal{R}_\rho \Psi_j)(\omega) \equiv \Psi_j(\mathcal{R}_\rho^{-1} \hat{\omega}), \quad (2)$$

where  $\mathcal{R}_\rho$  is the three-dimensional rotation matrix corresponding to  $\mathcal{R}_\rho$ . Spherical coordinates are denoted  $\omega = (\theta, \varphi) \in \mathbb{S}^2$  with colatitude  $\theta \in [0, \pi]$  and longitude  $\varphi \in [0, 2\pi)$ , where  $\hat{\omega}$  denotes the Cartesian vector corresponding to  $\omega$ . Rotations are specified by elements of the rotation group  $SO(3)$ , parameterised by the Euler angles  $\rho = (\alpha, \beta, \gamma) \in SO(3)$ , with  $\alpha \in [0, 2\pi)$ ,  $\beta \in [0, \pi]$  and  $\gamma \in [0, 2\pi)$ . The scale-discretised wavelet transform on the sphere then reads

$$\begin{aligned} W_{j\rho}^f &\equiv W_j^f(\rho) \equiv (f \otimes \Psi_j)(\rho) \equiv \langle f, \Psi_{j\rho} \rangle \\ &= \int_{\mathbb{S}^2} d\Omega(\omega) f(\omega) \Psi_{j\rho}^*(\omega), \end{aligned} \quad (3)$$

where  $j$  denotes the wavelet scale, which encodes the angular localisation of the wavelet,  $d\Omega(\omega) = \sin \theta d\theta d\varphi$  is the usual rotation-invariant measure on the sphere, and  $\cdot^*$  denotes complex conjugation. The inner product of functions on the sphere is denoted  $\langle \cdot, \cdot \rangle$ , while the operator  $\otimes$  denotes directional convolution on the sphere.

The wavelet transform of Eq. (3) thus probes directional structure in the signal of interest  $f$ , where  $\gamma$  can be viewed as the orientation about each point on the sphere  $(\theta, \varphi) = (\beta, \alpha)$ . Wavelet coefficient at scale  $j$  therefore live on the rotation group, *i.e.*  $W_j^f \in L^2(SO(3))$ . We adopt the shorthand notation  $W_{j\rho}^f$  to denote the wavelet coefficients of the signal  $f$  at scale  $j$  and position and orientation  $\rho$ , in order to simplify subsequent statistical calculations.

The wavelet coefficients do not encode the low-frequency content of the signal  $f$ ; a scaling function is introduced for this purpose. The scaling coefficients  $S^f \in L^2(\mathbb{S}^2)$  are given by the convolution of  $f$  with the axisymmetric scaling function  $\Phi \in L^2(\mathbb{S}^2)$  and read

$$\begin{aligned} S_\omega^f &\equiv S^f(\omega) \equiv (f \odot \Phi)(\omega) \equiv \langle f, \Phi_\omega \rangle \\ &= \int_{\mathbb{S}^2} d\Omega(\omega') f(\omega') \Phi_\omega^*(\omega'), \end{aligned} \quad (4)$$

where the rotated scaling function is defined by

$$\Phi_\omega(\omega') \equiv (\mathcal{R}_\omega \Phi)(\omega') \equiv \Phi(\mathcal{R}_\omega^{-1} \hat{\omega}'), \quad (5)$$

with  $\mathcal{R}_\omega = \mathcal{R}_{(\varphi, \theta, 0)}$ . The operator  $\odot$  denotes axisymmetric convolution on the sphere. Note that the scaling coefficients live on the sphere, and not the rotation group  $\text{SO}(3)$ , since directional structure of the low-frequency content of  $f$  is not typically of interest. We adopt the shorthand notation  $S_\omega^f$  to denote the scaling coefficients of the signal  $f$  at position  $\omega$ . In addition, we introduce the shorthand notation  $W^f = \mathcal{W}(f)$  to represent the overall wavelet analysis of  $f$ , *i.e.* including both wavelet and scaling coefficients.

Provided the wavelets and scaling function satisfy an admissibility condition (see Sec. 2.2), the original signal  $f$  can be synthesised exactly from its wavelet and scaling coefficients by

$$f(\omega) = \int_{\mathbb{S}^2} d\Omega(\omega') S_{\omega'}^f \Phi_{\omega'}(\omega) + \sum_{j=0}^J \int_{\text{SO}(3)} d\varrho(\rho) W_{j\rho}^f \Psi_{j\rho}(\omega), \quad (6)$$

where  $d\varrho(\rho) = \sin\beta d\alpha d\beta d\gamma$  is the usual invariant measure on  $\text{SO}(3)$ . We introduce the shorthand notation  $f = \mathcal{W}^{-1}(W^f)$  to represent the synthesis of a signal from its wavelet and scaling coefficients.

We adopt the same convention as [Wiaux et al. \(2008\)](#) and [McEwen et al. \(2016\)](#) for the wavelet scales  $j$ , with increasing  $j$  corresponding to larger angular scales, *i.e.* lower frequency content.<sup>2</sup> The maximum possible wavelet scale  $j$  is denoted by  $J_{\max}$  and is set to ensure the wavelets probe the entire scale (frequency) range (except zero) of the signal of interest, yielding  $J_{\max} = \lceil \log_\lambda(\ell_{\max}) \rceil$ , where  $\lambda$  is a dilation parameter (see [Wiaux et al. 2008](#); [Leistedt et al. 2013](#); [McEwen et al. 2015b, 2016](#)). The maximum wavelet scale considered in a given analysis  $J$  may be freely chosen, provided  $0 \leq J < J_{\max}$ . For  $J = J_{\max}$ , the wavelets probe the entire frequency content of the signal of interest except its mean, which is incorporated in the scaling coefficients.

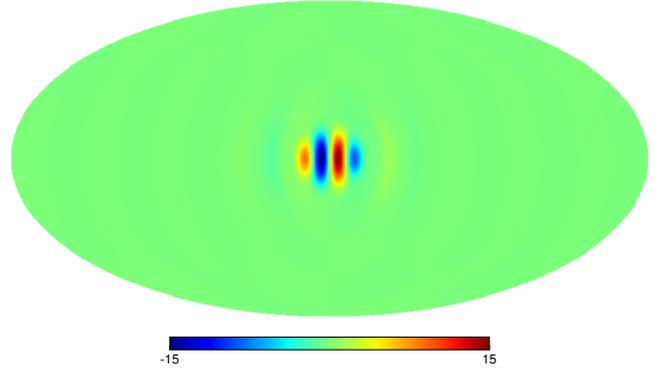
## 2.2 Wavelet construction

For the original signal to be synthesised perfectly from its wavelet and scaling coefficients through Eq. (6) the wavelets and scaling function must satisfy the following admissibility property:

$$\frac{4\pi}{2\ell+1} |\Phi_{\ell 0}|^2 + \frac{8\pi^2}{2\ell+1} \sum_{j=J_0}^J \sum_{m=-\ell}^{\ell} |(\Psi_j)_{\ell m}|^2 = 1, \quad \forall \ell, \quad (7)$$

where  $\Phi_{\ell 0} \delta_{m0} = \langle \Phi, Y_{\ell m} \rangle$  and  $(\Psi_j)_{\ell m} = \langle \Psi_j, Y_{\ell m} \rangle$  are the spherical harmonic coefficients of  $\Phi$  and  $\Psi_j$ , respectively, where  $\delta_{ij}$  for  $i, j \in \mathbb{Z}$  denotes the Kronecker delta. The spherical harmonic functions are denoted by  $Y_{\ell m} \in L^2(\mathbb{S}^2)$ , with  $\ell \in \mathbb{N}$  and  $m \in \mathbb{Z}$ ,  $|m| \leq \ell$ .

<sup>2</sup> Note that this differs from the convention adopted in [Leistedt et al. \(2013\)](#) and [McEwen et al. \(2015b\)](#) where increasing  $j$  corresponds to smaller angular scales and higher frequency content.



**Figure 4.** Directional scale-discretised wavelet with odd azimuthal symmetry for  $N = 4$ ,  $j = 7$ ,  $\ell_{\max} = 2048$ , and  $\lambda = 2$ . The wavelet is rotated from the North pole to the equator for visualisation purposes. The same wavelet parameters are assumed when analysing cosmic string maps, although lower  $j$  (*i.e.* smaller scales) are also considered. The wavelet is selected to match the step-like structure of contributions to the CMB due to cosmic strings, in order to yield a sparse representation of the string component in wavelet space.

Wavelets are defined in harmonic space in the separable form

$$(\Psi_j)_{\ell m} \equiv \sqrt{\frac{2\ell+1}{8\pi^2}} \kappa_j(\ell) \zeta_{\ell m}, \quad (8)$$

in order to control their angular and directional localisation separately, respectively through the kernel  $\kappa_j \in L^2(\mathbb{R}^+)$  and directionality component  $\zeta \in L^2(\mathbb{S}^2)$ , with harmonic coefficients  $\zeta_{\ell m} = \langle \zeta, Y_{\ell m} \rangle$ . Without loss of generality, the directionality component is normalised to impose

$$\sum_{m=-\ell}^{\ell} |\zeta_{\ell m}|^2 = 1, \quad \forall \ell. \quad (9)$$

An azimuthal band-limit  $N$  is imposed on the directionality component such that  $\zeta_{\ell m} = 0, \forall \ell, m$  with  $|m| \geq N$ , which controls the directional selectivity of the wavelet. Moreover, the wavelets are constructed to exhibit odd (even) azimuthal symmetry for  $N - 1$  odd (even). For further detail regarding the explicit construction of the wavelet kernel and directionality component see, *e.g.*, [McEwen et al. \(2016\)](#). An example of a scale-discretised wavelet on the sphere is plotted in Fig. 4

## 2.3 Computation

By appealing to sampling theorems on the sphere (*e.g.* [McEwen & Wiaux 2011](#)) and rotation group (*e.g.* [McEwen et al. 2015a](#)), the forward and inverse wavelet transforms of sampled signals can be computed exactly in theory for band-limited signals on the sphere, *i.e.* signals with spherical harmonic coefficients  $f_{\ell m} = 0, \forall \ell \geq \ell_{\max}$ , where  $f_{\ell m} = \langle f, Y_{\ell m} \rangle$ . The only error arising in numerical computations is that due to the finite representation of floating point numbers (indeed, numerical errors are found to be on the order of machine precision; see *e.g.* [McEwen et al. 2015b](#)). In practice, many real-world signals can be approximated accurately by band-limited signals. Furthermore, fast algorithms to compute the harmonic transforms associated with sampling theorems on the sphere and rotation group (*e.g.* [McEwen & Wiaux 2011](#); [McEwen et al. 2015a](#), respectively) can be exploited to render forward and inverse scale-discretised wavelet transforms computationally feasible for large cosmological data-sets (*e.g.* *Planck* maps).

The scale-discretised wavelet transform on the sphere is implemented in the s2LET code (Leistedt et al. 2013; McEwen et al. 2015b). The core algorithms of s2LET are implemented in C, while Matlab, Python and IDL interfaces are also provided. Consequently, s2LET is able to handle very large harmonic band-limits, corresponding to data-sets containing tens of millions of pixels. s2LET<sup>3</sup> is publicly available, and relies on the ssHT<sup>4</sup> code (McEwen & Wiaux 2011) to compute spherical harmonic transforms, the so3<sup>5</sup> code (McEwen et al. 2015a) to compute Wigner transforms and the FFTW<sup>6</sup> code to compute Fourier transforms. Note that it also supports the analysis of data on the sphere defined in the common HEALPIX<sup>7</sup> (Górski et al. 2005) format. s2LET provides the most recent and feature-rich implementation of scale-discretised wavelets, however, development on this project and s2LET was concurrent, and here we therefore use the previous s2dw<sup>8</sup> code (Wiaux et al. 2008) (which is functionally identical for the setting considered). s2dw is implemented in Fortran, and relies on the s2<sup>9</sup> code (McEwen et al. 2007a, 2008a) to handle data defined on the sphere and FFTW to perform Fourier transforms.

### 3 INFERENCE OF COSMIC STRING MODEL

While a cosmic string-induced component embedded in the CMB will not be Gaussian, the statistical distribution of the pixels of a cosmic string map nevertheless remains close to Gaussian. In Fig. 2 histograms of simulated inflationary and string-induced CMB components are plotted, in both pixel (Fig. 2(a)) and wavelet (Fig. 2(b)) space. In pixel space the distributions are similar. In wavelet space, however, while the distribution of the inflationary CMB component remains Gaussian (since the wavelet transform is linear), the distribution of the string-induced component is highly non-Gaussian. The latter distribution is peaked sharply about zero, illustrating the sparsifying nature of the wavelet transform for strings: the wavelet coefficients of the string-induced CMB component are sparsely distributed in wavelet space, while the coefficients of the inflationary CMB component are not.

We construct a hybrid wavelet-Bayesian framework to infer the presence of cosmic strings from CMB temperature observations. By constructing the statistical framework in wavelet space, where the inflationary and string-induced components have quite different statistical properties, we exploit the sparseness of the wavelet representation of the string signal to effectively determine the presence and parameters of any such component.

In this section we first describe the various models considered. We then define the statistical distributions of the inflationary and string-induced CMB components and noise, before presenting the framework for estimating the posterior distribution of the string tension  $G\mu$  and for estimating the Bayesian evidence in order to perform model selection.

<sup>3</sup> <http://www.s2let.org>

<sup>4</sup> <http://www.spinshht.org>

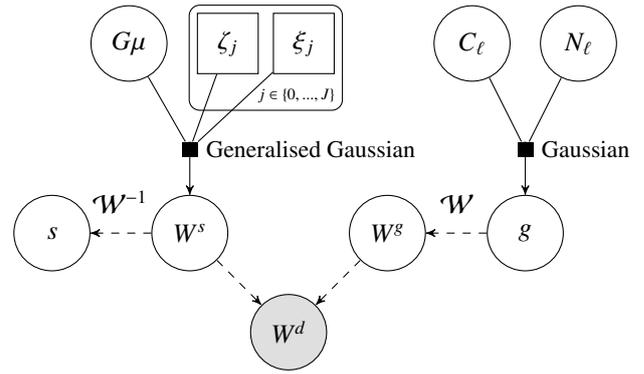
<sup>5</sup> <http://www.sohtree.org>

<sup>6</sup> <http://www.fftw.org>

<sup>7</sup> <http://healpix.jpl.nasa.gov>

<sup>8</sup> <http://www.s2dw.org>

<sup>9</sup> <http://www.jasonmcewen.org/codes.html>



**Figure 5.** Graphical Bayesian model  $M^s$  of the observed inflationary (Gaussian) and string induced (non-Gaussian) CMB components, represented in both wavelet and pixel spaces. Solid lines represent stochastic dependencies, while dashed lines represent deterministic dependencies. The string component is modelled by a generalised Gaussian distribution (GGD) in wavelet space, while the inflationary and noise components are modelled by Gaussian distributions in pixel space. The string component, Gaussian component and observed data are denoted by  $s$ ,  $g$ , and  $d$ , respectively, while wavelet coefficients are denoted by  $W$  with superscript representing the relevant signal. The forward and inverse wavelet transforms are represented by the shorthand notation  $\mathcal{W}$  and  $\mathcal{W}^{-1}$ , respectively. The GGDs modelling the wavelet coefficients of the string component are defined by the string tension  $G\mu$  and the scale and shape parameters  $\zeta_j$  and  $\xi_j$ , respectively. The Gaussian component is defined by the inflationary CMB and noise power spectra,  $C_\ell$  and  $N_\ell$ , respectively.

#### 3.1 Models

In the presence of a subdominant contribution due to cosmic strings, we model full-sky observational CMB data  $d$  as the sum of a string component  $s$ , an inflationary Gaussian component  $c$ , and noise  $n$ :

$$M^s : d = s + c + n. \quad (10)$$

All signals are assumed to be zero-mean since we study the perturbations of cosmological signals about their mean. We denote this string model by  $M^s$ . The alternative (standard) model is denoted by  $M^c$  and consists of an inflationary CMB component and noise only, absent of any string component:

$$M^c : d = c + n, \quad (11)$$

We work predominantly in wavelet space, where the inflationary and string-induced CMB components exhibit very different statistical distributions. Since the wavelet transform is linear, the models considered can be recast in wavelet space by, respectively,

$$M^s : W_{j\rho}^d = W_{j\rho}^s + W_{j\rho}^c + W_{j\rho}^n \quad (12)$$

and

$$M^c : W_{j\rho}^d = W_{j\rho}^c + W_{j\rho}^n, \quad (13)$$

where  $W^d = \mathcal{W}(d)$ ,  $W^s = \mathcal{W}(s)$ ,  $W^c = \mathcal{W}(c)$  and  $W^n = \mathcal{W}(n)$  are the wavelet coefficients of the observed CMB data, string component, inflationary component and noise, respectively. Here we denote the wavelet coefficients for each scale  $j$  and rotation  $\rho$  separately (hence the subscripts). Similar expressions hold for the scaling coefficients.

A graphical representation of the string model  $M^s$  in both wavelet and pixel space is shown in Fig. 5. The distributions modelling the string component, inflationary component and noise are

defined subsequently. While some of the variables used in the model shown in Fig. 5 have yet to be defined, it is nevertheless useful to present the general model now, which can then be used as a reference as the details of the model and distributions are specified in the subsequent subsections.

### 3.2 Statistical distributions

We determine the statistical distributions of the inflationary, noise and string-induced CMB components in wavelet space. The first two can be calculated analytically from an assumed power spectrum (since they are Gaussian), whereas the latter must be learnt from a string training simulation.

Since we determine these distributions in wavelet space, it is necessary in the following derivations to relate the wavelet coefficients of a signal to its spherical harmonic representation (e.g. McEwen et al. 2007a):

$$W_{j\rho}^d = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{n=-\ell}^{\ell} d_{\ell m}(\Psi_j)_{\ell n}^* D_{mn}^{\ell}(\rho), \quad (14)$$

where  $d_{\ell m} = \langle d, Y_{\ell m} \rangle$  and  $D_{mn}^{\ell} \in L^2(\text{SO}(3))$  are the Wigner D-functions. We adopt the shorthand notation  $\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{n=-\ell}^{\ell} = \sum_{\ell mn}$  henceforth and, assuming signals band-limited at  $\ell_{\max}$ , truncate sums over  $\ell$  to  $\ell_{\max}$ .

#### 3.2.1 CMB

The inflationary CMB component is assumed to be a homogeneous and isotropic Gaussian random field on the sphere defined by its power spectrum  $C_{\ell}$ :

$$\mathbb{E}[c_{\ell m} c_{\ell' m'}^*] = C_{\ell} \delta_{\ell\ell'} \delta_{mm'}, \quad (15)$$

where  $\mathbb{E}[\cdot]$  denotes expectation. The cosmological parameters defining the power spectrum  $C_{\ell}$  are assumed fixed at concordance values (Planck Collaboration XIII 2016) since the string contribution is subdominant. Since the wavelet transform is linear, the wavelet coefficients of the inflationary CMB component are also Gaussian. Their variance for scale  $j$  is

$$\begin{aligned} (\sigma_j^c)^2 &= \mathbb{E}[W_{j\rho}^c W_{j\rho}^{c*}] \\ &= \mathbb{E}\left[\sum_{\ell mn} D_{mn}^{\ell}(\rho) c_{\ell m}(\Psi_j)_{\ell n}^* \sum_{\ell' m' n'} D_{m' n'}^{\ell'}(\rho) c_{\ell' m'}^*(\Psi_j)_{\ell' n'}\right] \\ &= \sum_{\ell m} C_{\ell} |(\Psi_j)_{\ell m}|^2, \end{aligned} \quad (16)$$

where we have used Eq. (14) and the Wigner property (Varshalovich et al. 1989)

$$\sum_m D_{mn}^{\ell}(\rho) D_{m' n'}^{\ell*}(\rho) = \delta_{m m'} \quad (17)$$

(for an alternative proof of Eq. (16) see McEwen et al. 2007b). Consequently, the probability distribution of the wavelet coefficients of the inflationary CMB component on scale  $j$  read:

$$P_j^c(W_{j\rho}^c) = \frac{1}{\sqrt{2\pi(\sigma_j^c)^2}} \exp\left[-\frac{1}{2}\left(\frac{W_{j\rho}^c}{\sigma_j^c}\right)^2\right]. \quad (18)$$

We use  $P(\cdot)$  to denote generic probability distributions; however, when referring to a particular distribution we add appropriate superscripts and subscripts. Although this notation is not strictly necessary it improves the readability of the Bayesian analysis that follows. As typically considered in statistical wavelet analyses we assume wavelet coefficients are independent and do not include their

full covariance structure. We revisit the assumption of independence later and introduce measures to account for this approximation.

#### 3.2.2 Noise

Assuming Gaussian noise, we can include noise by simply modifying the Gaussian inflationary component to include the inflationary signal  $c$  and noise  $n$ :

$$g = c + n. \quad (19)$$

The resulting term  $g$  is Gaussian distributed since both  $c$  and  $n$  are Gaussian distributed. For modelling simplicity, we assume homogeneous and isotropic noise defined by power spectrum  $N_{\ell}$  such that

$$\mathbb{E}[g_{\ell m} g_{\ell' m'}^*] = (C_{\ell} + N_{\ell}) \delta_{\ell\ell'} \delta_{mm'}. \quad (20)$$

In practice a beam  $b_{\ell}$  and pixel window function  $p_{\ell}$  may also be incorporated, yielding

$$\mathbb{E}[g_{\ell m} g_{\ell' m'}^*] = (b_{\ell}^2 p_{\ell}^2 C_{\ell} + N_{\ell}) \delta_{\ell\ell'} \delta_{mm'}. \quad (21)$$

#### 3.2.3 Cosmic strings

Since the cosmic string-induced CMB component is not Gaussian and its map space distribution is not known *a priori*, it is not possible to analytically determine its distribution in wavelet space. Hence, we learn its distribution from a *training* simulated string map and test the distribution on a separate *testing* simulated string map.

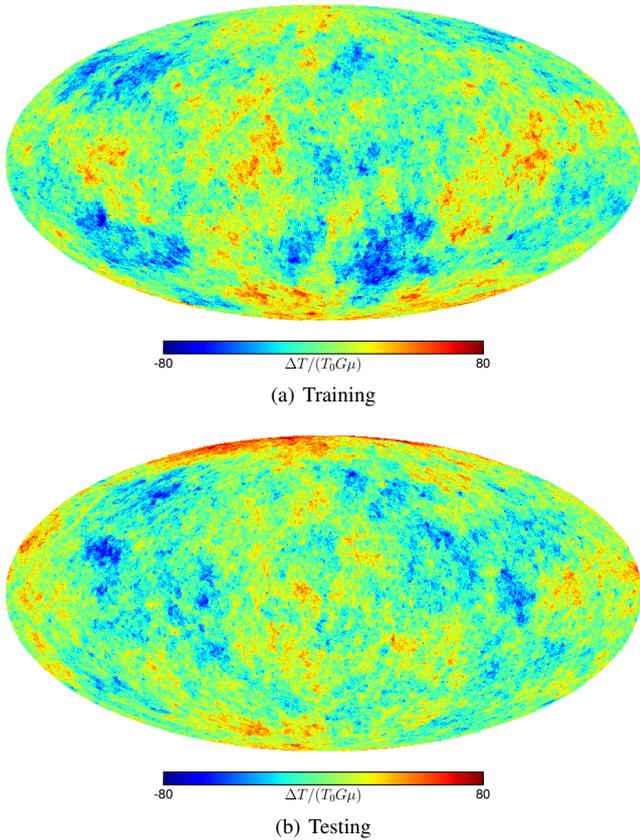
String maps simulated by the method of Ringeval & Bouchet (2012) are shown in Fig. 6. As previously stated, simulating these full-sky string maps at high resolution is extremely computationally demanding, requiring hundreds of thousands of CPU hours. Thankfully, we require only two simulated string maps: one for training, *i.e.* learning the statistical properties of string-induced CMB components; and one for testing our framework.

We adopt a generalised Gaussian distribution (GGD) to model the string-induced component in wavelet space and fit its parameters from the training simulation. The GGD of wavelet coefficients given a string tension  $G\mu$  is defined by

$$P_j^s(W_{j\rho}^s | G\mu) = \frac{\xi_j}{2G\mu\zeta_j\Gamma(\xi_j^{-1})} \exp\left(-\left|\frac{W_{j\rho}^s}{G\mu\zeta_j}\right|^{\xi_j}\right), \quad (22)$$

where  $\zeta_j$  and  $\xi_j$  are scale and shape parameters respectively (note that the overall scale of the distribution is dependent on the string tension and is given by  $G\mu\zeta_j$  at each scale) and  $\Gamma(\cdot)$  denotes the Gamma function. The GGD reduces to many common distributions for various shape parameters  $\xi$ . Gaussian and Laplacian distributions are recovered for  $\xi = 2$  and  $\xi = 1$  respectively, and in the limit  $\xi \rightarrow \infty$  the uniform distribution is recovered. The shape parameter can thus be considered as a measure of sparsity of the underlying signal. Note that GGDs have been used to model wavelet coefficients previously (e.g. Simoncelli & Adelson 1996). Due to statistical isotropy, the parameters of the GGD modelling the string contribution depend on wavelet scale  $j$  only and not the position or orientation of wavelet coefficients  $\rho$ . For small scales we expect the distribution of wavelet coefficients of the string map to be sparse in wavelet space, which we check by testing whether the shape of the distribution is leptokurtic, *i.e.* if  $\xi_j < 2$ .

We learn the shape and scale parameters of the GGD for the wavelet coefficients of a string-induced CMB component by the



**Figure 6.** Cosmic string-induced CMB anisotropies simulated by the method of Ringeval & Bouchet (2012).

method of moments approach outlined in Hammond et al. (2009). The variance and kurtosis of the GGD distributed wavelet coefficients of the string signal are given by, respectively,

$$(\sigma_j^s)^2 = \frac{(G\mu)^2 \zeta_j^2 \Gamma(3\xi_j^{-1})}{\Gamma(\xi_j^{-1})} \quad (23)$$

and

$$\kappa_j^s = \frac{\Gamma(5\xi_j^{-1})\Gamma(\xi_j^{-1})}{(\Gamma(3\xi_j^{-1}))^2}. \quad (24)$$

We compute the variance and kurtosis of the string training map and then solve these equations numerically to recover the scale and shape parameters of the GGD describing the wavelet coefficients of the string signal at each scale  $j$ . In practice, we train on the training string map with a beam and pixel windowed function applied.

The distributions of the cosmic string maps are shown in Fig. 7 and Fig. 8, while the estimated GGD shape parameters are listed in Table 1. The fitted GGD distribution of the training map matches the histogram of the testing map well for small scales (low  $j$ ), indicating that the learnt GGD accurately models the general statistical properties of cosmic string included CMB maps. As the scale becomes larger (higher  $j$ ) the match becomes less accurate due to cosmic variance. The distributions are also highly leptokurtic for small scales (low  $j$ ), *i.e.*  $\xi_j < 2$ , as apparent from the plots of the distributions (Fig. 7 and Fig. 8) and the fitted GGD shape parameters listed in Table 1: the string map is indeed sparse in wavelet space, as expected. As the scale increases the distribution becomes less leptokurtic, also as expected. We therefore consider wavelet

**Table 1.** GGD shape parameter  $\xi_j$  fitted to the testing string map. As expected the fitted GGD distributions are highly leptokurtic for small scales (low  $j$ ), with GGD shape parameter  $\xi_j < 2$ , due to the sparse representation of the string-induced CMB component in wavelet space.

Wavelet scale $j$	GGD shape $\xi_j$
0	0.94
1	1.08
2	1.40
3	1.69
4	1.68
5	1.79
6	1.91
7	1.76
8	1.84
9	1.80
10	2.57

coefficients up to and including scale  $j = 7$  only in the subsequent analysis, *i.e.* we set  $J = 7$ .

While we focus on inference in the current article, as an aside we note that once we have learnt the statistical properties of string maps, we can use the learnt distribution to simulate realisations of string maps for very low computational cost. However, in the current approach to training we do not learn the full covariance properties of the string components in wavelet space. We leave the development of a computationally efficient approach to simulating high-resolution, full-sky cosmic string-induced CMB maps to future work.

### 3.3 String tension estimation

In this section, we derive the posterior distribution for the string tension  $G\mu$  under the string model  $M^s$ . By Bayes theorem the string tension posterior  $P(G\mu | W^d)$  is related to the likelihood  $P(W^d | G\mu)$  by

$$P(G\mu | W^d) = \frac{P(W^d | G\mu) P(G\mu)}{P(W^d)} \propto P(W^d | G\mu) P(G\mu), \quad (25)$$

where  $P(G\mu)$  is the prior distribution for the string tension. For now we ignore the normalising denominator  $P(W^d)$  (the Bayesian evidence), which we return to in the following section. Recall that  $W^d$  are the wavelet coefficients of the observed CMB data.

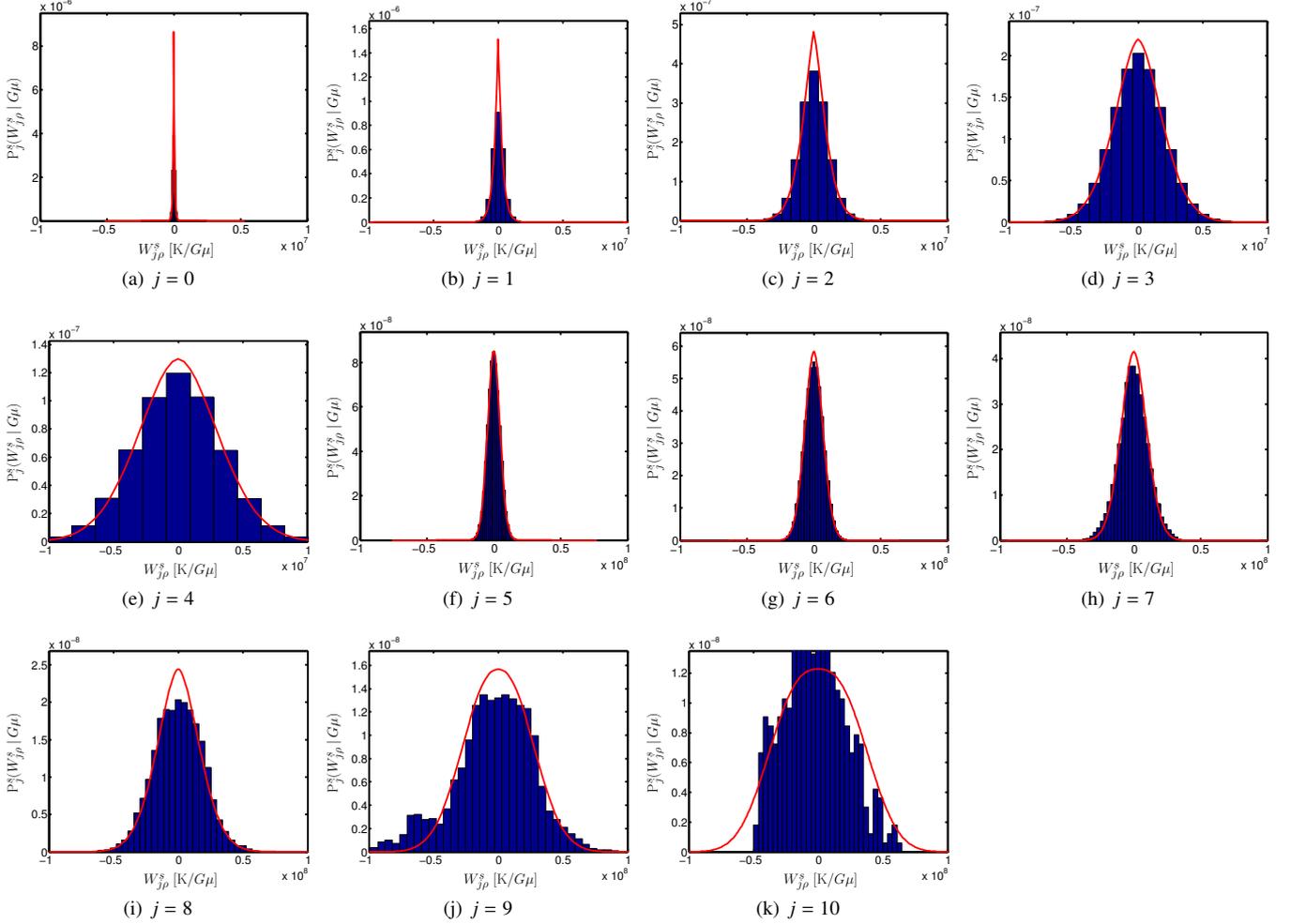
For each wavelet coefficient  $W_{j\rho}^d$  at scale  $j$  and position and orientation  $\rho$  the likelihood can be calculated by

$$P(W_{j\rho}^d | G\mu) = P(W_{j\rho}^s + W_{j\rho}^g | G\mu) \quad (26)$$

$$= \int_{\mathbb{R}} dW_{j\rho}^s P_j^s(W_{j\rho}^d - W_{j\rho}^s) P_j^s(W_{j\rho}^s | G\mu), \quad (27)$$

where  $W_{j\rho}^g$  are the wavelet coefficients of the Gaussian component  $g$ , which includes the inflationary CMB component and noise. The distributions comprising the integrand of Eq. (27) are precisely those described in Sec. 3.2, which we determine analytically or learn from a simulated string map. To compute the overall likelihood of the data, for speed of processing we assume each wavelet coefficient is independent, in which case the overall likelihood reads:

$$P(W^d | G\mu) = \prod_{j\rho} P(W_{j\rho}^d | G\mu). \quad (28)$$



**Figure 7.** Distribution of the cosmic string-induced CMB component in wavelet space, for each wavelet scale  $j$  (for parameters  $\ell_{\max} = 2048$ ,  $N = 4$ , and  $\lambda = 2$ ). The GGD distribution fitted to the training map is shown by the solid red curve, while the raw distribution of the testing map is shown by the solid blue histogram. The fitted GGD distribution of the training map matches the histogram of the testing map well for small scales (low  $j$ ), indicating that the learnt GGD accurately models the general statistical properties of cosmic string included CMB maps. As the scale becomes larger (higher  $j$ ) the match becomes less accurate due to cosmic variance. The distributions are also highly leptokurtic for small scales (low  $j$ ), indicating that the string map is indeed sparse in wavelet space, as expected. As the scale increases the distribution becomes less leptokurtic, also as expected. For these reasons we consider wavelet coefficients up to and including scale  $j = 7$  only in the subsequent analysis.

For numerical purposes we compute the log-likelihood, given by

$$\ln P(W^d | G\mu) = \sum_{j\rho} \ln P(W_{j\rho}^d | G\mu). \quad (29)$$

The assumption of independence of wavelet coefficients is approximate. Nevertheless, the covariance of wavelet coefficients decays rapidly with spatial separation (relative to the spatial size of the wavelet considered) and is zero for non-adjacent scales (*i.e.* for scales  $j$  and  $j'$  such that  $|j - j'| \geq 2$ ). We readdress the assumption of independence later and introduce measures to account for this approximation.

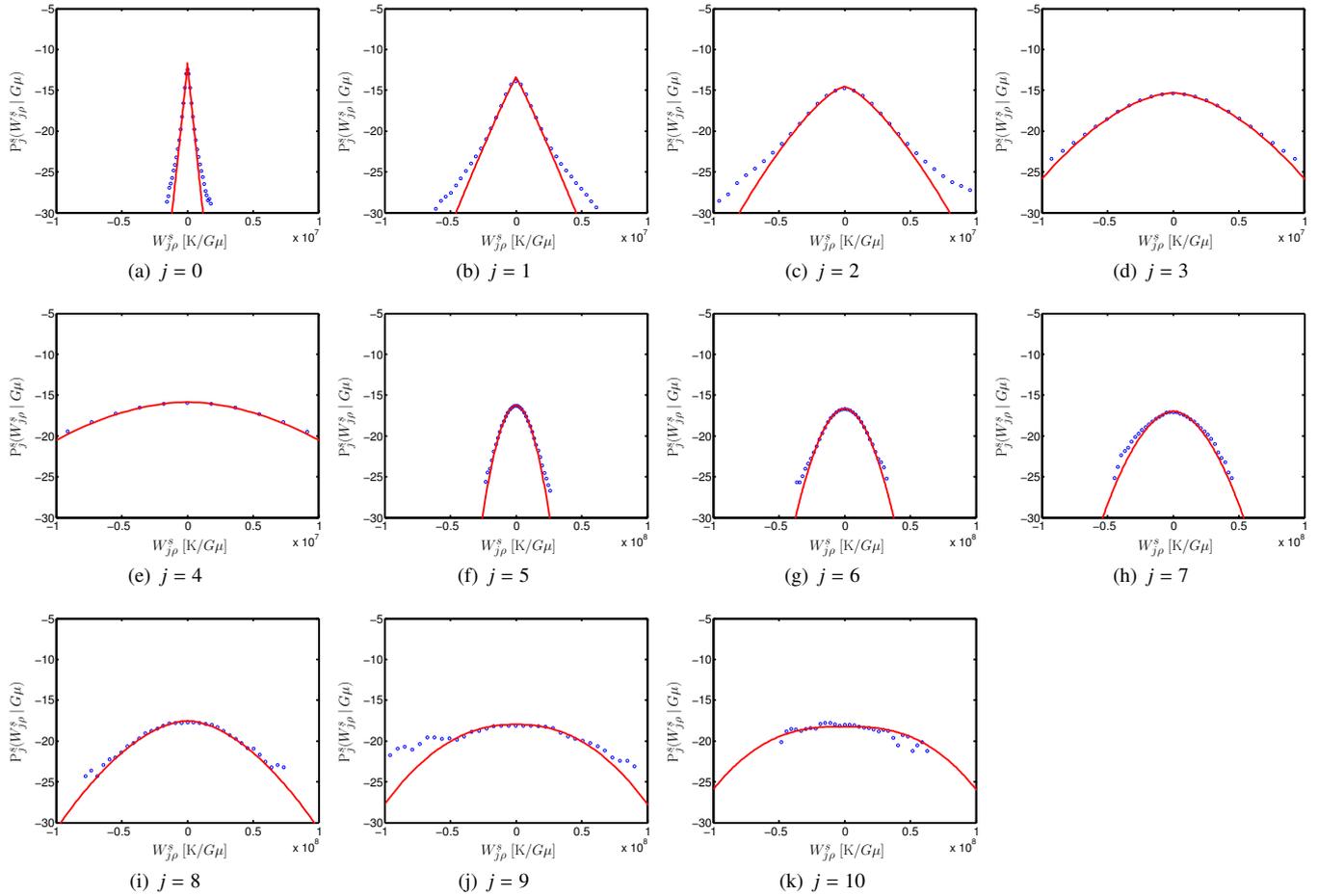
In practice, to compute the posterior distribution it is necessary to first evaluate the likelihood for each individual wavelet coefficient by Eq. (27), before combining these terms to compute the overall likelihood for the data by Eq. (28) or Eq. (29). In order to avoid recalculating integrals for identical (or similar) values of  $W_{j\rho}^d$ , we precompute look-up-tables (LUTs) for Eq. (27), storing the mapping from  $W_{j\rho}^d$  to  $P(W_{j\rho}^d | G\mu)$  for each  $j$ . When evaluating the likelihood of a given dataset we linearly interpolate the wavelet

coefficients onto the LUT grid. These LUTs are plotted in Fig. 9. Since the distributions  $P_j^g$  and  $P_j^s$  are properly normalised, the likelihood  $P(W_{j\rho}^d | G\mu)$  for a given  $G\mu$  is also a normalised probability distribution and should integrate to unity. To ensure the quadrature used to evaluate Eq. (27) is accurate, we check that the precomputed distributions  $P(W_{j\rho}^d | G\mu)$  integrate to unity (using the trapezium rule), which is indeed the case provided a sufficient number of samples is used to evaluate the integral.

### 3.4 String model comparison

To ascertain the overall evidence for cosmic strings we compare the Bayesian evidence of the string mode  $M^s$ , which includes string and inflationary induced CMB components, to the evidence of the standard inflationary model  $M^e$ . The Bayesian evidence of the string model is given by

$$E^s = P(W^d | M^s) = \int_{\mathbb{R}} d(G\mu) P(W^d | G\mu, M^s) P(G\mu | M^s), \quad (30)$$



**Figure 8.** Same as Fig. 7 but plotted on a  $\log_{10}$  scale.

where now we make the dependence on the model explicit. The Bayesian evidence of the CMB model is given by

$$E^c = P(W^d | M^c) = \prod_{j,\rho} P_j^s(W_{j\rho}^d). \quad (31)$$

For numerical purposes we compute the log-evidence, given by

$$\ln E^c = \ln P(W^d | M^c) = \sum_{j,\rho} \ln P_j^s(W_{j\rho}^d). \quad (32)$$

In the absence of any prior information favouring either model, the ratio of the model posterior probabilities is given by the ratio of the Bayesian evidences:

$$\frac{P(M^s | W^d)}{P(M^c | W^d)} = \frac{E^s}{E^c}. \quad (33)$$

We compute the ratio of evidences to determine the model favoured by the data. In practice we compute the difference in log-evidence (also called the Bayes factor):

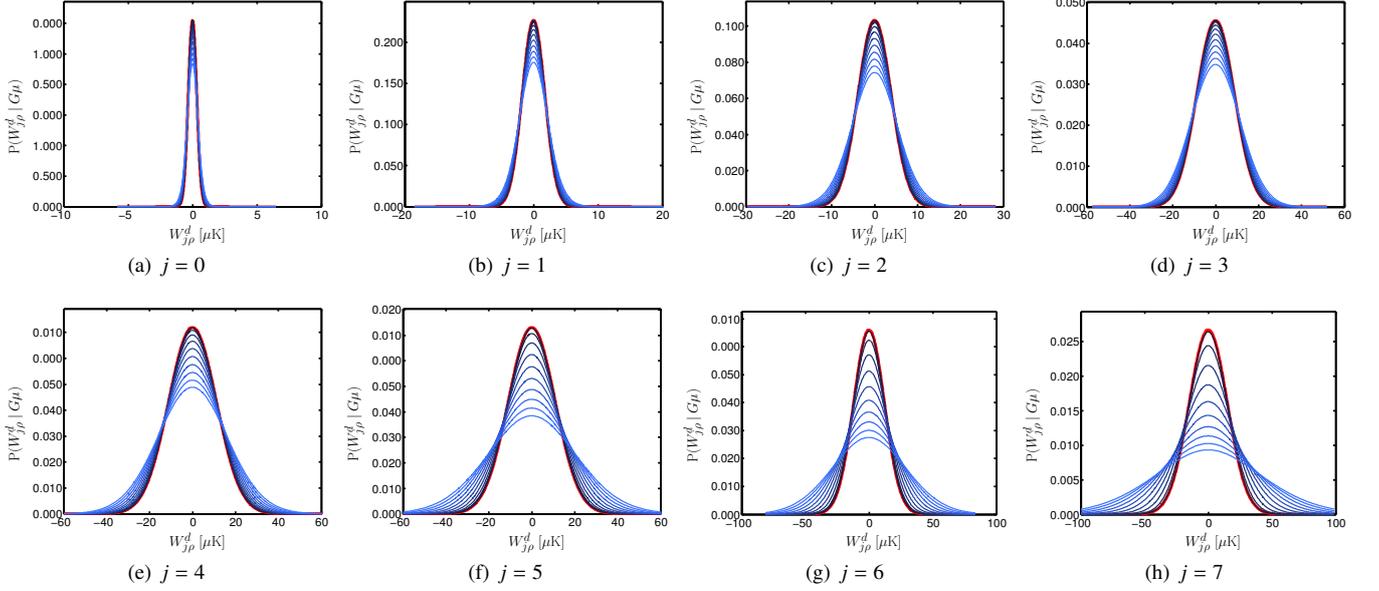
$$\Delta \ln E = \ln(E^s/E^c) = \ln E^s - \ln E^c. \quad (34)$$

The Jeffreys scale (Jeffreys 1961) is often used as a rule-of-thumb when comparing models via their Bayes factor. While we caution against using the Jeffreys scale as a strict test to classify models (since the boundaries of the scale are somewhat arbitrary), it can nevertheless be useful to gain some intuition for those not familiar with Bayesian model selection. The log-Bayes factor  $\Delta \ln E = \ln(E^{(1)}/E^{(2)})$  represents the degree by which model  $M^{(1)}$  is

favoured over model  $M^{(2)}$ , assuming the models are equally likely *a priori*. On the Jeffreys scale log-Bayes factors are given the following interpretation:  $0 \leq \Delta \ln E < 1$  is regarded as inconclusive;  $1 \leq \Delta \ln E < 2.5$  as significant;  $2.5 \leq \Delta \ln E < 5$  as strong; and  $\Delta \ln E \geq 5$  as conclusive (without loss of generality we have assumed  $E_1 \geq E_2$ ). For reference, a log-Bayes factor of 2.5 corresponds to odds of approximately 1 in 12, while a factor of 5 corresponds to odds of approximately 1 in 150.

#### 4 ESTIMATION OF COSMIC STRING MAPS

In addition to estimating the evidence for the cosmic string model and the posterior distribution of the string tension, we also recover a direct estimate of the string-induced CMB component itself. To estimate the string contribution at the map level we develop a Bayesian estimation approach in wavelet space, generalising the technique described in Hammond et al. (2009) from a planar region to the spherical full-sky setting. We first describe the string map estimation technique, before examining its properties as a Bayesian thresholding approach to denoise the inflationary CMB component from the observed data.



**Figure 9.** Look-up-tables (LUTs) to precompute the mapping  $W_{jp}^d$  to  $P(W_{jp}^d | G\mu)$  of Eq. (27) for each  $j$ . Precomputing LUTs avoids recalculating integrals for identical (or similar) values of  $W_{jp}^d$  when computing the posterior distribution of the string tension. The distributions  $P(W_{jp}^d | G\mu)$  are shown for different values of  $G\mu$  by the light blue lines. For comparison, the Gaussian CMB distribution  $P_j^c(W_{jp}^c)$  is shown by the heavy red line. As  $G\mu$  is reduced towards zero, the distributions  $P(W_{jp}^d | G\mu)$  approach  $P_j^c(W_{jp}^c)$  (the light blue curves darken and approach the heavy red curve). The LUTs are normalised probability distributions and integrate to unity. These plots are created using the testing string map and parameters  $\ell_{\max} = 2048$ ,  $N = 4$ ,  $J = 7$ , and  $\lambda = 2$ .

#### 4.1 String map estimation

Our inference of the wavelet coefficients of the underlying string map, and equivalently the string map itself, is encoded in the posterior probability distribution  $P(W_{jp}^s | W^d)$ . Various estimators can be considered to recover the wavelet coefficients of string map from their posterior distribution. We estimate the wavelet coefficients of the string map from the mean of the posterior distribution, which can be computed by

$$\bar{W}_{jp}^s = \int_{\mathbb{R}} dW_{jp}^s W_{jp}^s P(W_{jp}^s | W^d) \quad (35)$$

$$= \int_{\mathbb{R}} dW_{jp}^s W_{jp}^s \int_{\mathbb{R}} d(G\mu) P(W_{jp}^s | W^d, G\mu) P(G\mu | W^d) \quad (36)$$

$$= \int_{\mathbb{R}} d(G\mu) P(G\mu | d) \bar{W}_{jp}^s(G\mu), \quad (37)$$

where

$$\bar{W}_{jp}^s(G\mu) = \int_{\mathbb{R}} dW_{jp}^s W_{jp}^s P(W_{jp}^s | W_{jp}^d, G\mu) \quad (38)$$

$$= \frac{\int_{\mathbb{R}} dW_{jp}^s W_{jp}^s P(W_{jp}^d | W_{jp}^s, G\mu) P(W_{jp}^s | G\mu)}{P(W_{jp}^d | G\mu)} \quad (39)$$

$$= \frac{\int_{\mathbb{R}} dW_{jp}^s W_{jp}^s P_j^s(W_{jp}^d - W_{jp}^s | G\mu) P_j^s(W_{jp}^s | G\mu)}{P(W_{jp}^d | G\mu)}. \quad (40)$$

Note that we replace  $P(G\mu | W^d)$  with  $P(G\mu | d)$  in Eq. (37) (since there is a one-to-one relationship between a map and its wavelet coefficients) and appeal to Bayes theorem in Eq. (39). To summarise, for each  $G\mu$  we compute a denoised set of wavelet coefficients  $\bar{W}_{jp}^s(G\mu)$  by Eq. (40). We then combine these, taking the posterior distribution of the string tension  $P(G\mu | d)$  into account, to compute the overall denoised set of wavelet coefficients  $\bar{W}_{jp}^s$  by Eq. (37). The denominator of Eq. (40) is given by Eq. (27) for which LUTs have been precomputed already. Similarly, LUTs for the numerator

are precomputed for each  $j$ . Since we consider zero-mean signals, these LUTs should integrate to zero, which indeed they do provided a sufficient number of samples is used to evaluate the integrals. As we assume independence of the wavelet coefficients, wavelet coefficients are denoised pointwise.

Once we have recovered the denoised wavelet coefficients, a string map can be recovered through an inverse wavelet transform:

$$\bar{s} = \mathcal{W}^{-1}(\bar{W}^s). \quad (41)$$

Alternatively, string maps could also be estimated for each  $G\mu$  through an inverse wavelet transform:

$$\bar{s}(G\mu) = \mathcal{W}^{-1}(\bar{W}^s(G\mu)). \quad (42)$$

Since the wavelet transform is linear, the overall string map could then be recovered by

$$\bar{s} = \int_{\mathbb{R}} d(G\mu) P(G\mu | d) \bar{s}(G\mu). \quad (43)$$

While we use the string tension posterior distribution  $P(G\mu | d)$  estimated in wavelet space by the approach outlined in Sec. 3.3, one is free to substitute a posterior distribution estimated by alternative methods. The resulting recovered string maps could be considered as a pre-processed input to other map-based methods for estimating the string tension from the non-Gaussian structure of the string-induced CMB component, such as edge detection (*e.g.* Lo & Wright 2005; Amsel et al. 2008; Stewart & Brandenberger 2009; Danos & Brandenberger 2010). The enhanced string component and reduced background is likely to boost the effectiveness of subsequent string tension estimation.

An estimate of the variance of wavelet coefficients of the string component could also be performed in order to provide a measure of the accuracy of the recovered string-induced component. For this to be most useful it would be necessary to express the variance in

map space, which could be computed by an inverse wavelet transform with  $\Psi_{j\mu}$  and  $\Phi_{\omega}$  substituted by  $|\Psi_{j\mu}|^2$  and  $|\Phi_{\omega}|^2$ , respectively (for a related discussion see [Rogers et al. 2016a](#)). We leave this to future work.

## 4.2 Bayesian thresholding

The hybrid wavelet-Bayesian string map estimation technique outlined in Sec. 4.1 can be viewed as a Bayesian thresholding approach to denoise the observed data. The estimation of the wavelet coefficients of the string signal by Eq. (40) can be viewed as a mapping from the wavelet coefficients of the data  $W_{j\mu}^d$  to the estimated string signal  $\overline{W}_{j\mu}^s(G\mu)$  for a given  $G\mu$ . One then marginalises over the prior distribution for  $G\mu$  by Eq. (37). The thresholding mapping functions defined by Eq. (40) are plotted in Fig. 10.

As the wavelet scale  $j$  increases larger scale features are probed, for which the string distribution in wavelet space becomes less leptokurtic (as shown in Table 1), *i.e.* more Gaussian. Consequently, the thresholding functions become more linear as it becomes more difficult to distinguish the string and inflationary CMB distributions. For the small scales, corresponding to low  $j$ , the thresholding functions are less linear, with the energy in large coefficients more likely to be retained since these are more likely due to the string component, while the energy of small coefficients is more likely to be curtailed.

For each wavelet scale  $j$ , curves are plotted for different values of  $G\mu$ , with  $G\mu$  approaching zero as the shade of the curve darkens. As  $G\mu$  is reduced the amplitude of the string component is reduced relative to the inflationary component and the thresholding curves approach zero, as expected.

## 5 SIMULATIONS AND RESULTS

In this section, we demonstrate the application of our wavelet-Bayesian framework for cosmic string inference to simulated *Planck* observations. We do not optimise the parameters of the analysis and consider the standard dyadic wavelet scaling (*i.e.*  $\lambda = 2$ ). Alternative wavelet scalings, like that considered in [Rogers et al. \(2016a\)](#) and [Rogers et al. \(2016b\)](#), are likely to improve performance. The application to *Planck* data and the optimisation of the parameters of the method is left to future work. We first describe the CMB simulations performed, before presenting results from applying the framework outlined previously to these simulations for differing values of  $G\mu$ . We show results estimating the posterior distribution of the string tension, comparing the string model  $M^s$  with the standard inflationary model  $M^c$ , and recovering maps of the string-induced CMB component.

### 5.1 Simulations

Our simulations model idealised observations of combined string-induced and inflationary CMB skies by the *Planck* satellite's 143 GHz detectors, making heavy use of the *s2* code. As high-resolution, full-sky string simulations are computationally challenging to produce (see Sec. 3.2.3), we base all string simulations on the single testing string simulation (Fig. 6(b)), smoothing with a 7.3 arcmin Gaussian beam and rescaling by the appropriate  $T_0 G\mu$  (assuming the mean CMB temperature of  $T_0 = 2.725$  K; [Mather et al. 1999](#)). We do not touch the training string simulation (Fig. 6(a)) since this was used to fit the GGD distributions

modelling string-induced CMB components. We model the CMB and noise as pure Gaussian random fields and hence draw realisations directly from their combined power spectrum, using a band-limit of  $\ell_{\max} = 2500$ . We calculate the CMB power spectrum using *CAMB*<sup>10</sup> ([Lewis et al. 2000](#)), assuming the best-fit cosmology from *Planck*'s analysis of a compilation of CMB, lensing, baryon acoustic oscillation, supernova and expansion datasets ([Planck Collaboration XIII 2016](#)). To create the final power spectrum, we multiply the CMB power spectrum by the instrumental beam and *HEALPIX* window function, and then add white noise at 4.3  $\mu$ K per beam-sized pixel (the final sensitivity of *Planck*'s 143 GHz channel). The resulting power spectra are shown in Fig. 1. All maps are simulated at *HEALPIX* resolution  $N_{\text{side}} = 2048$ .

Examples of simulated CMB maps, with and without a string-induced component, are plotted in Fig. 11 for a string tension of  $G\mu = 5 \times 10^{-7}$ . It is not possible to determine the presence cosmic strings by eye.

### 5.2 String tension estimation

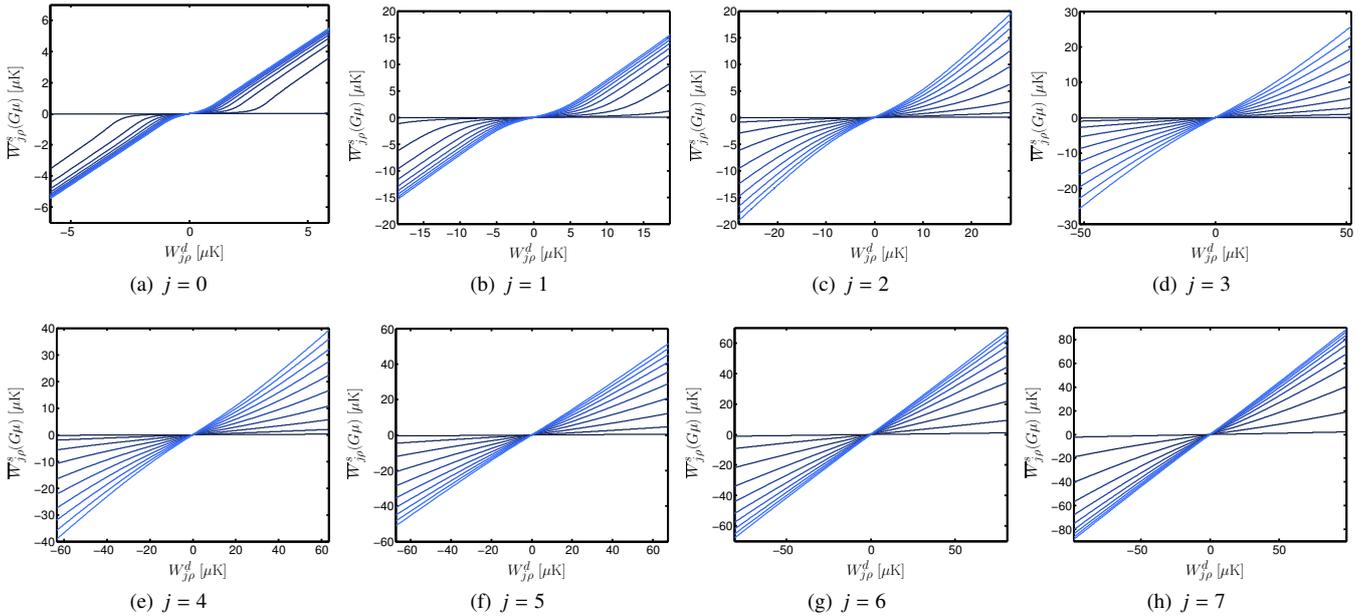
We perform the analysis outlined in Sec. 3.3 on a number of simulations with embedded string contributions of varying string tension  $G\mu$  to estimate the posterior distribution of the string tension.

In the framework presented in Sec. 3.3 we assume wavelet coefficients are independent. In practice, wavelet coefficients are not independent, but the covariance of wavelet coefficients does decay rapidly with spatial separation (relative to the spatial size of the wavelet considered) and is identically zero for non-adjacent scales (*i.e.* for scales  $j$  and  $j'$  such that  $|j - j'| \geq 2$ ). To better account for the covariance of wavelet coefficients, we fold into the analysis only those wavelet coefficients that are essentially uncorrelated. To achieve this, we compute a correlation length for each wavelet scale, which we define by the fifth zero crossing of the theoretical wavelet covariance when assuming an inflationary CMB power spectrum (for the derivation of the theoretical wavelet covariance see [McEwen et al. 2016](#)). We then downsample wavelet coefficients to the resolution defined by the correlation length and use the resulting downsampled maps of wavelet coefficients, for non-adjacent wavelet scales  $j$  only, when computing the full log-posterior by Eq. (29).

Here and subsequently we consider a dyadic wavelet scaling with  $\lambda = 2$ , as discussed previously. We consider a maximum wavelet scale of  $J = 7$ , as also discussed previously, since for these wavelet scales the GGD modelling the string component is highly leptokurtic (see Sec. 3.2.3). For the string tension  $G\mu$ , we assume a uniform prior over the domain ( $1 \times 10^{-10}$ ,  $4 \times 10^{-6}$ ), sampled with 200 uniformly spaced gridpoints. When constructing the LUTs, we evaluate tables sampled over a domain of 1000 uniformly spaced gridpoints for the wavelet coefficients of the data and use 9000 uniformly spaced gridpoints for the string wavelet coefficients when computing integrals (by the trapezium rule). The limits of the coefficient ranges are specified by the minimum and maximum values of the wavelet coefficients of the data. As discussed, we perform a number of tests to ensure the LUTs are evaluated accurately.

Since the string tension is a scaling parameter, an uninformative (Jeffreys) prior for the string tension would be a log-uniform prior. However, for this first work we instead choose to adopt a uniform prior so that the string tension posterior and likelihoods

<sup>10</sup> <http://camb.info>



**Figure 10.** Bayesian thresholding functions for each wavelet scale  $j$ . Each curve in a given panel shows a different value of  $G\mu$ , with  $G\mu$  approaching zero as the shade of the curve darkens. As  $G\mu$  is reduced the amplitude of the string contribution is reduced relative to the inflationary component and the thresholding curves approach zero, as expected. As the wavelet scale  $j$  increases, larger scale features are probed by the wavelet and the thresholding functions become more linear since the statistical distributions of the string and inflationary CMB components become more similar (see text for further discussion).

correspond, which can be useful for gaining further intuition regarding the effectiveness of the method. By using a uniform prior the impact of alternative priors can be approximately inferred by a kind of “posterior-by-eye” approach. In future, for applications to data, a log-uniform prior or indeed other priors can be considered.

The estimated posterior distributions are shown in Fig. 12 for a representative subset of the ground truth string tension values used in generating the simulated data. The full set of string tension values considered is shown in the first column of Table 2. An estimate of  $G\mu$  and the corresponding error, for each simulation, is recovered from the mean and the standard deviation of the posterior distribution and also shown Table 2. The ground truth string tension values used to embed the string-induced CMB component in the simulated data are recovered accurately above  $\sim 5 \times 10^{-7}$ . Below this approximate transition value the recovered estimates are biased high, likely due to unmodelled residual correlations, indicating the limit of the sensitivity of this unoptimised method.

### 5.3 String model comparison

For the same set of simulations we compute the Bayesian evidence ratio of the string model  $M^s$  and standard inflationary model  $M^c$ , performing the calculation outlined in Sec. 3.4. Again, we fold into the analysis only those wavelet coefficients that are essentially uncorrelated, following the approach outlined in Sec. 5.2.

The computed evidence ratios are shown in Table 2, where a positive value favours the string model  $M^s$  over the standard inflationary model  $M^c$ . For values of the string tension  $G\mu$  greater than  $\sim 5 \times 10^{-7}$  the string model is preferred. Interestingly, this is the same approximate transition value of  $G\mu$  as found for the estimation of the string tension in Sec. 5.2, further confirming the sensitivity of the (unoptimised) method for statistical inference.

**Table 2.** String tension values considered in simulations, with recovered estimates and corresponding Bayesian evidence ratio (positive evidence favours the string model  $M^s$ ). The ground truth string tension is recovered accurately above  $\sim 5 \times 10^{-7}$  but is biased high below this transition region. The Bayesian evidence ratio favours the string model also above  $\sim 5 \times 10^{-7}$  but favours the standard inflationary model below this transition, illustrating the sensitivity of the (unoptimised) method.

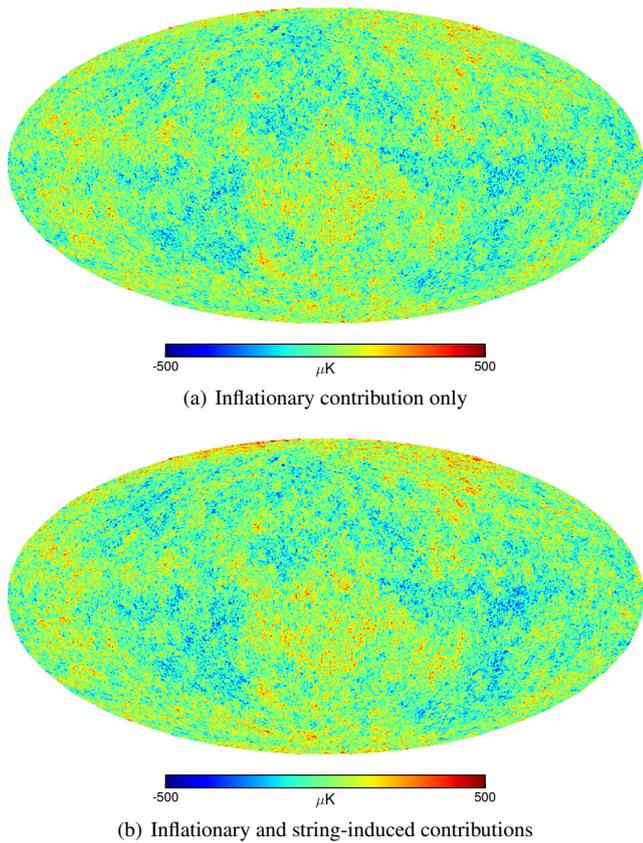
$G\mu$ truth / $10^{-7}$	$G\mu$ estimate / $10^{-7}$	Evidence ratio [ $\log_e$ ]
30.0	$29.58 \pm 0.45$	2 020
20.0	$19.60 \pm 0.47$	563
10.0	$9.90 \pm 0.58$	51.4
9.00	$8.97 \pm 0.61$	34.6
8.00	$8.06 \pm 0.65$	21.9
7.00	$7.18 \pm 0.69$	12.5
6.00	$6.36 \pm 0.73$	5.88
5.00	$5.63 \pm 0.75$	1.19
4.00	$5.06 \pm 0.75$	-1.86
3.00	$4.66 \pm 0.73$	-3.87

### 5.4 String map recovery

While we have examined the effectiveness of our wavelet-Bayesian method for statistical inference in the previous subsections, a significant advantage of our approach is the ability to also recover estimates of any embedded string-induced CMB component at the map level. For the same set of simulations we recover estimated string maps following the calculations outlined in Sec. 4.

Maps of the recovered string-induced CMB component are illustrated in Fig. 13. String maps are recovered well for large values of  $G\mu$ . As  $G\mu$  is reduced, the fidelity of the recovered maps is reduced as small scale features are washed out.

To assess the performance of the recovery of string maps quantitatively we plot in Fig. 14 the root-mean-squared (RMS) error and the signal-to-noise ratio (SNR) to quantify the error be-



**Figure 11.** Simulated CMB maps with and without a string-induced component. Panel (a) includes an inflationary CMB contribution only, whereas panel (b) includes an inflationary contribution and the string-induced contribution shown in Fig. 6(b), scaled to  $G\mu = 5 \times 10^{-7}$ . It is not possible to determine the presence of the string-induced component by eye.

tween the recovered string map and the ground truth string map. The SNR is defined as the ratio of the RMS value of the ground truth string map to the RMS error. We compute these error metrics for the simulations performed with varying values of the string tension  $G\mu$ . For comparison, results are also shown when not directly estimating the string-induced component. In this case, we simply consider the residuals between the observed data and the ground truth string maps. The RMS error is then simply given by the RMS of the inflationary CMB component and noise, hence the constant dashed blue curve in Fig. 14(a). As  $G\mu$  is reduced, though it is difficult to recover small scale string features (as shown in Fig. 13) the RMS error of the estimated string components is nevertheless reduced: by greater than  $200 \mu\text{K}$  for the lowest values of  $G\mu$  considered. From the SNR plotted in Fig. 14(b) it is clear that the relative improvement provided by the string estimation approach is reasonably constant over much of the domain considered.

Recall that, as discussed in Sec. 4.1, in order to recover the string-induced CMB component a posterior distribution for the string tension is required. Here, we adopt the posterior distributions recovered in Sec. 5.2 and shown in Fig. 12; however, if an alternative method provides a better estimate of the posterior distribution then the alternative posterior distribution can be substituted.

## 6 CONCLUSIONS

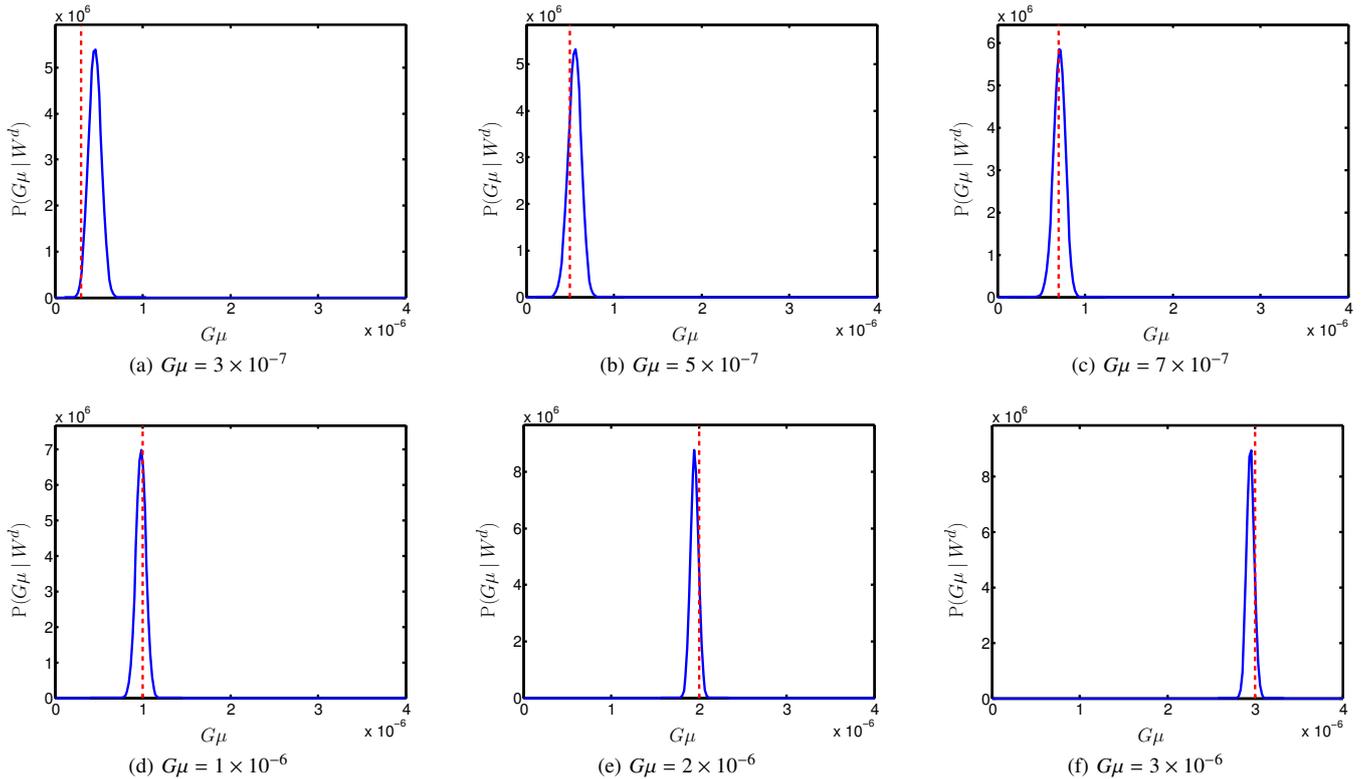
Cosmic strings are a well-motivated extension to the standard cosmological model and could induce a subdominant component in the anisotropies of the CMB. Detecting such a component would provide a direct probe of corresponding symmetry breaking phase transitions in the early Universe at very high energy scales. However, due to the weak nature of any string component its detection presents a significant observational challenge.

We present a hybrid wavelet-Bayesian framework for cosmic string inference, constructing a Bayesian analysis in wavelet space where the string-induced CMB component has very different statistical properties to the inflationary component. We learn and exploit the complex non-Gaussian structure of string-induced CMB contributions, rather than considering (insufficient) summary statistics like many alternative methods (*e.g.* the kurtosis), for which the origin of any non-Gaussian component cannot be rigorously determined. Our approach allows the full posterior distribution of the string tension to be estimated, from which a best estimate of the string tension and an associated error can be computed. The Bayesian evidence ratio comparing the string model, including an inflationary component and a subdominant string-induced component, and the standard inflationary model can also be computed. Moreover, it is also possible to recover an estimate of the string-induced component in the CMB at the map level.

We demonstrate the application of our wavelet-Bayesian framework and evaluate its performance using idealised simulations of CMB observations made by the *Planck* satellite, where a string component is embedded for a range of values of the string tension  $G\mu$ . For values of the string tension  $G\mu$  above  $\sim 5 \times 10^{-7}$ , we recover accurate estimates of its posterior distribution, which can be used to provide accurate point estimates of the string tension and associated error. The Bayesian evidence values computed also correctly favour the string model for values of the string tension  $G\mu$  above  $\sim 5 \times 10^{-7}$ , further highlighting the sensitivity of the method. The performance of our approach compares favourably with current constraints obtained using the same string simulations (that obtain the constraint  $G\mu < 7.8 \times 10^{-7}$ ; [Planck Collaboration XXV 2014](#)) and, moreover, is based on a principled statistical framework. A more robust and principled analysis is inevitably more conservative than less well-motivated alternatives but, nevertheless, we find our method generally compares favourably with other map-based techniques. While we consider slightly idealised *Planck* simulations we have not yet optimised the parameters of the analysis (alternative wavelet scalings, for example, are likely to improve performance; *cf.* [Rogers et al. 2016a,b](#)).

We find that the embedded string maps are recovered accurately for large values of the string tension  $G\mu$ . As  $G\mu$  is reduced, small-scale features in the recovered string maps are washed out, but the RMS error of the recovered maps is nevertheless reduced considerably. While maps of the string-induced CMB component are of interest in their own right, they can also be used as pre-processed inputs for alternative techniques to estimate the string tension from the non-Gaussian structure of the string-induced CMB component, such as computing the gradient. We leave post-processing of the recovered string maps for further work.

This is one of many areas to be considered in future work. First, more realistic *Planck* simulations will be considered, along with a mask to remove foreground emission (masking can be integrated in the wavelet analysis in a straightforward manner, following the approach of, *e.g.*, [McEwen et al. 2005, 2007b](#); [Leistedt et al. 2017](#)). Second, the parameters of the analysis will be optimised for



**Figure 12.** Posterior distributions of the string tension  $G\mu$  recovered from simulations. Each panel shows the recovered posterior distribution for a different ground truth value of  $G\mu$  as a solid blue curve; the ground truth value of  $G\mu$  is indicated by a vertical red dashed line. The ground truth value of  $G\mu$  is estimated accurately above  $\sim 5 \times 10^{-7}$ . For lower string tensions the posterior distribution is biased high, illustrating the sensitivity of the (unoptimised) method.

*Planck* observations (cf. Rogers et al. 2016a,b). Third, the steerability of scale-discretised wavelets will be exploited to provide more accurate inference when computing the posterior distribution of the string tension and the Bayesian evidence (cf. Planck Collaboration XXV 2014). Fourth, an estimate of the standard deviation of the recovered string map will be developed to characterise its accuracy, as outlined in Sec. 4.1. Fifth, techniques will be developed to better model the full covariance structure of signals in wavelet space.

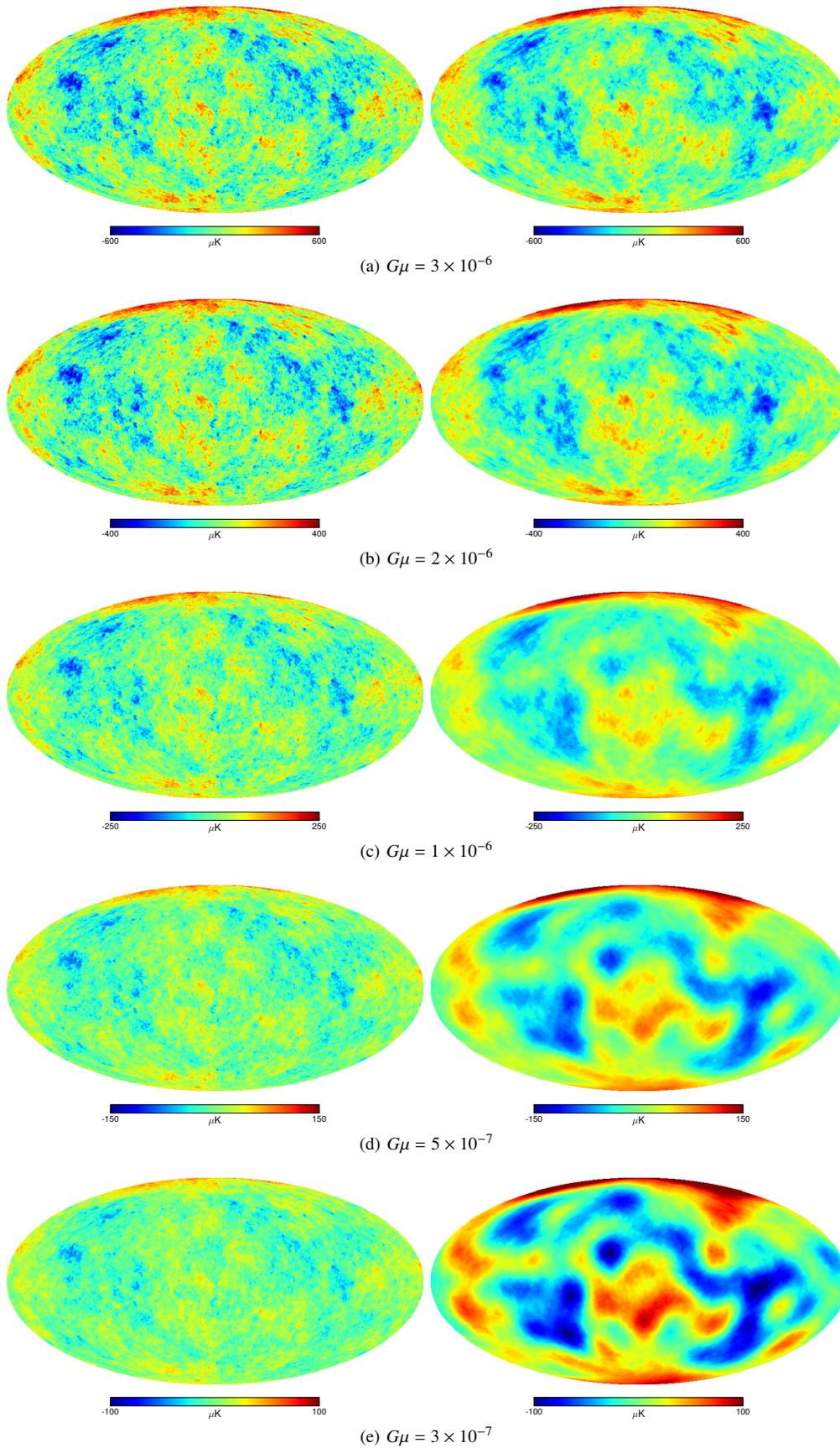
Our framework will in future be applied to observational data from *Planck* and other CMB experiments to provide constraints on the string tension that are based on a principled statistical analysis of the non-Gaussian structure of string-induced CMB contributions. While we focus in the current article on cosmic strings, the framework can also be adapted to other settings, such as other components embedded in the CMB.

## ACKNOWLEDGEMENTS

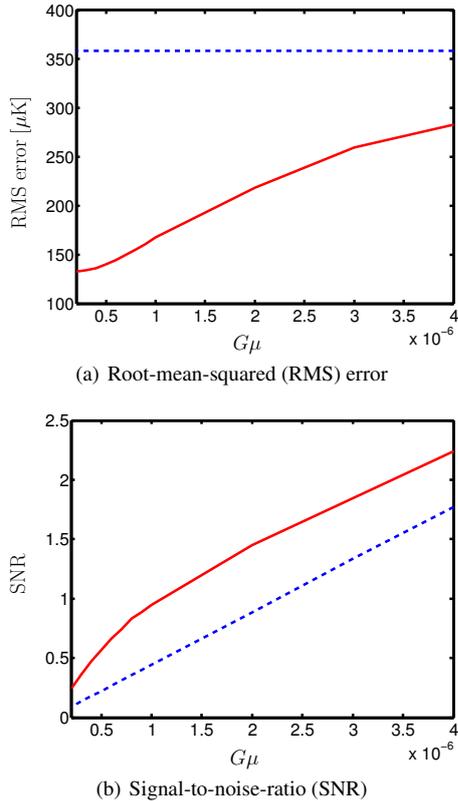
This work is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) by grant EP/M011852/1. SMF and HVP were partially supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) ERC GA n. 306478-CosmicDawn. SMF was additionally supported by the UK Science and Technology Facilities Council (STFC). We acknowledge use of the following public software packages: HEALPIX (Górski et al. 2005); CAMB (Lewis et al. 2000); s2DW (Wiaux et al. 2008); s2 (McEwen et al. 2007a, 2008a).

## REFERENCES

- Albrecht A., Battye R.A., Robinson J., 1997, Phys.Rev.Lett., 79, 4736, [astro-ph/9707129](#)
- Albrecht A., Battye R.A., Robinson J., 1999, Phys.Rev., D59, 023508, [astro-ph/9711121](#)
- Albrecht A., Turok N., 1989, Phys. Rev., D40, 973
- Allen B., Shellard P., 1990, Phys. Rev. Lett., 64, 119
- Amsel S., Berger J., Brandenberger R.H., 2008, J. Cosmol. Astropart. P., 4, 015, [0709.0982](#)
- Antoine J.P., Vandergheynst P., 1998, J. Math. Phys., 39, 8, 3987
- Antoine J.P., Vandergheynst P., 1999, Applied Comput. Harm. Anal., 7, 1
- Baldi P., Kerkycharian G., Marinucci D., Picard D., 2009, Ann. Stat., 37 No.3, 1150, [arXiv:math/0606599](#)
- Barreiro R.B., Hobson M.P., Lasenby A.N., Banday A.J., Górski K.M., Hinshaw G., 2000, Mon. Not. Roy. Astron. Soc., 318, 475, [astro-ph/0004202](#)
- Bennett D.P., Bouchet F.R., 1989, Phys. Rev. Lett., 63, 2776
- Bennett D.P., Bouchet F.R., 1990, Phys. Rev., D41, 2408
- Blanco-Pillado J.J., Olum K.D., Shlaer B., 2011, Phys.Rev., D83, 083514, [1101.5173](#)
- Bouchet F.R., Bennett D.P., Stebbins A., 1988, Nature, 335, 410
- Brandenberger R.H., 1994, International Journal of Modern Physics A, 9, 2117, [astro-ph/9310041](#)
- Chan J.Y.H., Leistedt B., Kitching T.D., McEwen J.D., 2016, IEEE Trans. Sig. Proc., 65, 1, 5, [arXiv:1511.05578](#)
- Charnock T., Avgoustidis A., Copeland E.J., Moss A., 2016, Phys. Rev. D., 93, 12, 123503, [1603.01275](#)
- Copeland E.J., Kibble T.W.B., 2009, Proceedings of the Royal So-



**Figure 13.** Recovered string-induced CMB maps for various ground truth values of the string tension  $G\mu$ . Ground truth maps are shown on the left and recovered maps on the right. String-induced CMB contributions are recovered well for large values of  $G\mu$ . As  $G\mu$  is reduced, the fidelity of the recovered maps is reduced as small scale features are washed out.



**Figure 14.** Error metrics quantifying the difference between the recovered string-induced CMB component and the ground truth map (solid red curve). For comparison, differences without estimating the string-induced component are also shown (dashed blue curve).

ciety of London Series A, 466, 623, [0911.1345](#)

Danos R.J., Brandenberger R.H., 2010, *International Journal of Modern Physics D*, 19, 183, [0811.2004](#)

Ducout A., Bouchet F.R., Colombi S., Pogosyan D., Prunet S., 2013, *Mon. Not. Roy. Astron. Soc.*, 429, 2104, [1209.1223](#)

Durastanti C., Fantaye Y., Hansen F., Marinucci D., Pesenson I.Z., 2014, *Phys. Rev. D*, 90, 10, 103532, [arXiv:1408.1095](#)

Feeney S.M., Johnson M.C., Mortlock D.J., Peiris H.V., 2011a, *Phys. Rev. Lett.*, 107, 071301, [1012.1995](#)

Feeney S.M., Johnson M.C., Mortlock D.J., Peiris H.V., 2011b, *Phys. Rev. D*, D84, 043507, [1012.3667](#)

Feeney S.M., Johnson M.C., Mortlock D.J., Peiris H.V., 2012, *Phys. Rev. Lett.*, 108, 24, 241301, [1203.1928](#)

Fergusson J.R., Regan D.M., Shellard E.P.S., 2010, *ArXiv e-prints*, [1012.6039](#)

Fraisse A.A., Ringeval C., Spergel D.N., Bouchet F.R., 2008, *Phys. Rev.*, D78, 043535, [0708.1162](#)

Geller D., Hansen F.K., Marinucci D., Kerkycharian G., Picard D., 2008, *Phys. Rev. D*, 78, 12, 123533, [arXiv:0811.2881](#)

Geller D., Lan X., Marinucci D., 2009, *Electron. J. Statist.*, 3, 1497, [0907.3369](#)

Geller D., Marinucci D., 2010, *J. Fourier Anal. and Appl.*, 16, 6, 840, [0811.2935](#)

Geller D., Marinucci D., 2011, *J. Math. Anal. Appl.*, 375, 2, 610, [arXiv:1006.3835](#)

Górski K.M., Hivon E., Banday A.J., Wandelt B.D., Hansen F.K., Reinecke M., Bartelmann M., 2005, *Astrophys. J.*, 622, 759, [astro-ph/0409513](#)

Gott III J.R., Park C., Juszkiewicz R., Bies W.E., Bennett D.P., Bouchet F.R., Stebbins A., 1990, *Astrophys. J.*, 352, 1

Hammond D.K., Wiaux Y., Vanderghaynst P., 2009, *Mon. Not. Roy. Astron. Soc.*, 398, 1317, [arXiv:0811.1267](#)

Hergt L., Amara A., Brandenberger R., Kacprzak T., Refregier A., 2016, *ArXiv e-prints*, [1608.00004](#)

Hindmarsh M., 1994, *Astrophys. J.*, 431, 534, [astro-ph/9307040](#)

Hindmarsh M.B., Kibble T.W.B., 1995, *Reports on Progress in Physics*, 58, 477, [hep-ph/9411342](#)

Hinshaw G., et al., 2013, *Astrophys. J. Supp.*, 208, 19, [1212.5226](#)

Jeffreys H., 1961, *Theory of probability*, Oxford University Press, Oxford, 3rd edition

Kaiser N., Stebbins A., 1984, *Nature*, 310, 391

Kibble T.W.B., 1976, *Journal of Physics A Mathematical General*, 9, 1387

Landriau M., Shellard E., 2003, *Phys.Rev.*, D67, 103512, [astro-ph/0208540](#)

Landriau M., Shellard E., 2011, *Phys.Rev.*, D83, 043516, [1004.2885](#)

Lanusse F., Rassat A., Starck J.L., 2012, *Astron. & Astrophys.*, 540, A92, [arXiv:1112.0561](#)

Leistedt B., McEwen J.D., 2012, *IEEE Trans. Sig. Proc.*, 60, 12, 6257, [arXiv:1205.0792](#)

Leistedt B., McEwen J.D., Büttner M., Peiris H.V., 2017, *Mon. Not. Roy. Astron. Soc.*, 466, 3, 3728, [arXiv:1605.01414](#)

Leistedt B., McEwen J.D., Kitching T.D., Peiris H.V., 2015, *Phys. Rev. D*, 92, 123010, [arXiv:1509.06750](#)

Leistedt B., McEwen J.D., Vanderghaynst P., Wiaux Y., 2013, *Astron. & Astrophys.*, 558, A128, 1, [arXiv:1211.1680](#)

Lewis A., Challinor A., Lasenby A., 2000, *Astrophys. J.*, 538, 473, [astro-ph/9911177](#)

Lizarraga J., Urrestilla J., Daverio D., Hindmarsh M., Kunz M., 2016, *ArXiv e-prints*, [1609.03386](#)

Lizarraga J., Urrestilla J., Daverio D., Hindmarsh M., Kunz M., Liddle A.R., 2014a, *Phys. Rev. Lett.*, 112, 17, 171301, [1403.4924](#)

Lizarraga J., Urrestilla J., Daverio D., Hindmarsh M., Kunz M., Liddle A.R., 2014b, *Phys. Rev. D*, 90, 10, 103504, [1408.4126](#)

Lo A.S., Wright E.L., 2005, *ArXiv*, [astro-ph/0503120](#)

Marinucci D., et al., 2008, *Mon. Not. Roy. Astron. Soc.*, 383, 539, [arXiv:0707.0844](#)

Mather J.C., Fixsen D.J., Shafer R.A., Mosier C., Wilkinson D.T., 1999, *Astrophys. J.*, 512, 511, [astro-ph/9810373](#)

McEwen J.D., 2017, *IEEE Trans. Image Proc.*, submitted, [arXiv:1510.01595](#)

McEwen J.D., Büttner M., Leistedt B., Peiris H.V., Vanderghaynst P., Wiaux Y., 2014, in *Proceedings IAU Symposium No. 306, 2014*, [arXiv:1412.1340](#)

McEwen J.D., Büttner M., Leistedt B., Peiris H.V., Wiaux Y., 2015a, *IEEE Sig. Proc. Lett.*, 22, 12, 2425, [arXiv:1508.03101](#)

McEwen J.D., Durastanti C., Wiaux Y., 2016, *Applied Comput. Harm. Anal.*, in press, [arXiv:1509.06767](#)

McEwen J.D., Hobson M.P., Lasenby A.N., 2006a, *ArXiv*, [astro-ph/0609159](#)

McEwen J.D., Hobson M.P., Lasenby A.N., 2008a, *IEEE Trans. Sig. Proc.*, 56, 8, 3813, [astro-ph/0612688](#)

McEwen J.D., Hobson M.P., Lasenby A.N., Mortlock D.J., 2005, *Mon. Not. Roy. Astron. Soc.*, 359, 1583, [astro-ph/0406604](#)

McEwen J.D., Hobson M.P., Lasenby A.N., Mortlock D.J., 2006b, *Mon. Not. Roy. Astron. Soc.*, 371, L50, [astro-ph/0604305](#)

McEwen J.D., Hobson M.P., Lasenby A.N., Mortlock D.J., 2008b,

- Mon. Not. Roy. Astron. Soc., 388, 2, 659, [arXiv:0803.2157](#)
- McEwen J.D., Hobson M.P., Mortlock D.J., Lasenby A.N., 2007a, IEEE Trans. Sig. Proc., 55, 2, 520, [astro-ph/0506308](#)
- McEwen J.D., Leistedt B., 2013, in *10th International Conference on Sampling Theory and Applications (SampTA)*, 329–333, [arXiv:1307.1307](#)
- McEwen J.D., Leistedt B., Büttner M., Peiris H.V., Wiaux Y., 2015b, IEEE Trans. Sig. Proc., submitted, [arXiv:1509.06749](#)
- McEwen J.D., Scaife A.M.M., 2008, Mon. Not. Roy. Astron. Soc., 389, 3, 1163, [arXiv:0803.2165](#)
- McEwen J.D., Vandergheynst P., Wiaux Y., 2013, in *SPIE Wavelets and Sparsity XV*, volume 8858, [arXiv:1308.5706](#)
- McEwen J.D., Vielva P., Hobson M.P., Martínez-González E., Lasenby A.N., 2007b, Mon. Not. Roy. Astron. Soc., 373, 1211, [astro-ph/0602398](#)
- McEwen J.D., Wiaux Y., 2011, IEEE Trans. Sig. Proc., 59, 12, 5876, [arXiv:1110.6298](#)
- McEwen J.D., Wiaux Y., Evers D.M., 2011, Astron. & Astrophys., 531, A98, [arXiv:1108.3900](#)
- McEwen J.D., Wiaux Y., Hobson M.P., Vandergheynst P., Lasenby A.N., 2008c, Mon. Not. Roy. Astron. Soc., 384, 4, 1289, [arXiv:0704.0626](#)
- McEwen J.D., et al., 2007c, J. Fourier Anal. and Appl., 13, 4, 495, [arXiv:0704.3158](#)
- Moore J.N., Shellard E.P.S., Martins C.J.A.P., 2002, Phys. Rev., D65, 2, 023503, [hep-ph/0107171](#)
- Movahed M.S., Javanmardi B., Sheth R.K., 2013, Mon. Not. Roy. Astron. Soc., 434, 3597, [1212.0964](#)
- Narcowich F.J., Petrushev P., Ward J.D., 2006, SIAM J. Math. Anal., 38, 2, 574
- Pen U.L., Seljak U., Turok N., 1997, Phys. Rev. Lett., 79, 1611, [astro-ph/9704165](#)
- Planck Collaboration XII, 2014, Astron. & Astrophys., 571, A12, [arXiv:1303.5072](#)
- Planck Collaboration XIII, 2016, Astron. & Astrophys., 594, A13, [arXiv:1502.01589](#)
- Planck Collaboration XXIII, 2014, Astron. & Astrophys., 571, A23, [arXiv:1303.5083](#)
- Planck Collaboration XXV, 2014, Astron. & Astrophys., 571, A25, [arXiv:1303.5085](#)
- Pogosian L., Vachaspati T., 1999, Phys.Rev., D60, 083504, [astro-ph/9903361](#)
- Regan D., Hindmarsh M., 2015, J. Cosmol. Astropart. P., 10, 030, [1508.02231](#)
- Ringeval C., Bouchet F.R., 2012, Phys. Rev. D., 86, 2, 023513, [1204.5041](#)
- Ringeval C., Sakellariadou M., Bouchet F., 2007, JCAP, 0702, 023, [astro-ph/0511646](#)
- Rogers K.K., Peiris H.V., Leistedt B., McEwen J.D., Pontzen A., 2016a, Mon. Not. Roy. Astron. Soc., 460, 3, 3014, [arXiv:1601.01322](#)
- Rogers K.K., Peiris H.V., Leistedt B., McEwen J.D., Pontzen A., 2016b, Mon. Not. Roy. Astron. Soc., 462, 3, 2310, [arXiv:1605.01417](#)
- Sadegh Movahed M., Khosravi S., 2011, J. Cosmol. Astropart. P., 3, 012, [1011.2640](#)
- Sanz J.L., Herranz D., López-Cañiego M., Argüeso F., 2006, in *EUSIPCO*, [astro-ph/0609351](#)
- Schröder P., Sweldens W., 1995, in *Computer Graphics Proceedings (SIGGRAPH '95)*, 161–172
- Simoncelli E.P., Adelson E.H., 1996, in *IEEE International Conference on Image Processing*, volume 1, IEEE, 379–382
- Starck J., Moudden Y., Bobin J., 2009, Astron. & Astrophys., 497, 931, [arXiv:0902.0574](#)
- Starck J.L., Aghanim N., Forni O., 2004, Astron. & Astrophys., 416, 9, [astro-ph/0311577](#)
- Starck J.L., Moudden Y., Abrial P., Nguyen M., 2006, Astron. & Astrophys., 446, 1191, [astro-ph/0509883](#)
- Stewart A., Brandenberger R., 2009, J. Cosmol. Astropart. P., 2, 009, [0809.0865](#)
- Sweldens W., 1997, SIAM J. Math. Anal., 29, 2, 511
- Varshalovich D.A., Moskalev A.N., Khersonskii V.K., 1989, *Quantum theory of angular momentum*, World Scientific, Singapore
- Vielva P., Martínez-González E., Barreiro R.B., Sanz J.L., Cayón L., 2004, Astrophys. J., 609, 22, [astro-ph/0310273](#)
- Vielva P., Martínez-González E., Tucci M., 2006, Mon. Not. Roy. Astron. Soc., 365, 891, [astro-ph/0408252](#)
- Vilenkin A., Shellard E.P.S., 1994, *Cosmic strings and other topological defects*, Cambridge monographs on mathematical physics, Cambridge Univ. Press, Cambridge
- Vincent G., Antunes N.D., Hindmarsh M., 1998, Phys. Rev. Lett., 80, 2277, [hep-ph/9708427](#)
- Wiaux Y., Jacques L., Vandergheynst P., 2005, Astrophys. J., 632, 15, [astro-ph/0502486](#)
- Wiaux Y., McEwen J.D., Vandergheynst P., Blanc O., 2008, Mon. Not. Roy. Astron. Soc., 388, 2, 770, [arXiv:0712.3519](#)
- Wiaux Y., Puy G., Vandergheynst P., 2010, Mon. Not. Roy. Astron. Soc., 402, 4, 2626, [arXiv:0908.4179](#)