ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Swiss Federal Institute of Technology, EPFL
Signal Processing Laboratory, LTS2

Prof. PIERRE VANDERGHEYNST
Dr. YVES WIAUX

Master project

Sep. 2008 - Jan. 2009

Cosmic string detection from interferometric data of the microwave background radiation

GILLES PUY

February 23, 2009

Contents

1 Introduction **1**

2 Radio interferometry **3**
2.1 Visibilities . . . 3
2.2 Interferometric inverse problem . . . 3
2.2.1 General case . . . 3
2.2.2 Cosmic string case . . . 4

3 Bayesian perspective **6**
3.1 Statistical model of the Gaussian CMB signal . . . 6
3.2 Statistical model of the string signal . . . 6
3.3 Estimation of the string tension . . . 9
3.4 Wavelet domain Bayesian reconstruction scheme . . . 9

4 Compressed sensing **11**
4.1 Introduction to the theory . . . 11
4.2 Compressed sensing reconstruction of a string signal . . . 12

5 CLEAN **14**
5.1 The CLEAN algorithm . . . 14
5.2 CLEAN reconstruction of a string signal . . . 14

6 Simulations and results **16**
6.1 Experimental set up . . . 16
6.2 Performance of the estimation of the string tension . . . 18
6.3 Performance of the reconstruction algorithms . . . 18

7 Conclusion and perspectives **22**

8 Acknowledgements **22**

References **23**

A Simulation of Gaussian CMB maps **24**

1 Introduction

Figure 1: The Arcminute Microkelvin Imager (Large Array)

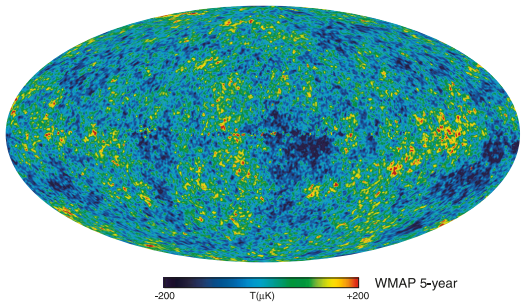

Figure 2: WMAP data

In the hot early universe, standard matter (electrons and protons) and radiation (photon) were in equilibrium. When the temperature of the universe was low enough (400 000 years after the Big Bang), electrons and protons recombined to form neutral hydrogen atoms and the universe became transparent to radiation. The cosmic microwave background (CMB) is the remnant of this radiation, in which we are still bathed today and that cosmologists have been studying for nearly half a century now. According to the concordance cosmological model the cosmic structures and the CMB originate from Gaussian adiabatic perturbations seeded in the early phase of inflation, a period of exponential expansion in the very early universe. The corresponding component of the CMB is thus also Gaussian. Its anisotropies on the celestial sphere are divided into the so-called primary anisotropies, imprinted at the time of hydrogen recombination itself, and secondary anisotropies, imprinted during the subsequent evolution of the universe.

Some models motivated by theories for the unification of the fundamental interactions suggest the formation of topological defects such as cosmic strings, originating from phase transitions at the end of inflation. These defects would have imprinted a non-Gaussian component in the CMB superimposed to the standard Gaussian component. The overall string network is parameterized by a string tension μ (a mass per unit length of string) which sets the overall amplitude of the string contribution. Current observations of the anisotropies of the temperature of the CMB radiation from WMAP (Wilkinson Microwave Anisotropy Probe) data[1] still allow the presence of cosmic strings with a string tension below:

$$\frac{G\mu}{c^2} \leq 2 \times 10^{-7},$$

where $G = 6.67\text{e} - 11$ m^{3}kg^{-1}s^{-2} stands for the gravitational constant and $c = 3.00\text{e} + 08$ m/s the speed of light.

Beyond standard radiometers, radio-interferometers are also used to observe the CMB signal. For example, the Arcminute Microkelvin Imager (AMI) in Cambridge[2] is currently mapping the CMB at radio frequencies between 13.5 Ghz and 18 Ghz on field of view around 1 degree, and with a resolution between 30 arcseconds and several arcminutes. Radio interferometry is a powerful technique used in astronomy to achieve high angular resolution. Radio telescope arrays synthesize a very large effective aperture by using only small telescopes. Considering small field of views on the celestial sphere, one may see the signal probed as a planar image on the plane perpendicular to the line of sight. In this context, the synthesized aperture is the same as the one of a unique telescope whose size would be the maximum projected distance between two telescopes of the array on the plane of the image. A radio interferometer directly acquires measurements in the Fourier plane of this image. The aim of the project is to develop and test an algorithm for the detection of cosmic strings from such

[1] http://lambda.gsfc.nasa.gov/

[2] http://www.mrao.cam.ac.uk/telescopes/ami/

interferometric data of the CMB.

In Section 2, we present the inverse problem posed to reconstruct a string signal from radio-interferometric data. Then we describe the statistical models of the string and Gaussian CMB signals in Section 3. In Section 4, we introduce the theory of compressed sensing and describe two algorithms to reconstruct a string signal that take into account prior information on the signal. Then we present in Section 5 a very common algorithm called CLEAN and explain how to apply it to our problem. This algorithm is used in radio-astronomy to solve the general interferometric inverse problem. Finally, we study the performance of these three algorithms in Section 6 and conclude in Section 7.

2 Radio interferometry

In this section, we describe the principles of radio-interferometry and formulate the general inverse problem posed for the reconstruction of an image from radio-interferometric data. After presenting briefly the string and Gaussian CMB signals, the inverse problem is then posed for these specific signals.

2.1 Visibilities

Let us consider n radio telescopes that constitute an interferometric array. All these radio telescopes, identified by an index k ($1 \leq k \leq N$), point in the same direction and receive a monochromatic electric field $E(\vec{r}_k, t)$ of wavelength λ (t denotes the time variable, and $\vec{r}_k \in \mathbb{R}^3$ the three-dimensional position of the k$^{\text{th}}$ telescope). The vector $\vec{R}_{k\,k'} = \vec{r}_k - \vec{r}_{k'}$ defines the relative position between the two telescopes k and k', and is called the baseline. The field of view of each telescope is limited by an illumination function $A(\theta, \phi)$ (where $(\theta, \phi) \in [0, 2\pi) \times [0, \pi]$ are the celestial coordinates). At each instant of observation, each pair of telescopes k and k' measures a complex visibility $y_{k\,k'}$ defined as the correlation between the incoming electric fields:

$$y_{kk'} = \langle E(\vec{r}_k, t) E(\vec{r}_{k'}, t)^* \rangle_{\Delta t}\,, \tag{1}$$

where $\langle \cdot \rangle_{\Delta t}$ represents an average over a time long compared with the period of the wave detected. This correlation measures a flux and the visibility is in unit of Jansky (1 Jy = 10^{-26} W m^{-2} Hz^{-1}). By combining all pairs of telescopes, it is thus possible to measure $\binom{n}{2} = n(n-1)/2$ visibilities at each instant of observation.

We now assume that the field of view of the telescopes is small enough to consider that the observed patch of the sphere may be approximated by a planar patch P. The illumination function may now be defined on this plane as a function of a two dimensional vector $\vec{p} \in \mathbb{R}^2$ whose origin is set at the pointing direction of the array. The projected baseline, $\vec{R}^{\perp}_{kk'}$, is defined as the projection of the baseline $\vec{R}_{kk'}$ onto the plane P. In these conditions, the van Cittert-Zernike theorem states that the measured visibility $y_{kk'}$, defined by expression (1), is the Fourier transform of the intensity of E, $I(\vec{p})$ with $\vec{p} \in \mathbb{R}^2$, multiplied by the illumination function $A(\vec{p})$ at the spatial frequency $\vec{u}_{k_1 k_2}$ defined as the projected baseline divided by the wavelength of observation:

$$y_{kk'} = \int_{\mathbb{R}^2} A(\vec{p}) I(\vec{p}) \exp^{-2i\pi \vec{p}\cdot\vec{u}_{kk'}}\, d^2\vec{p}, \tag{2}$$

with

$$\vec{u}_{kk'} = \frac{\vec{R}^{\perp}_{kk'}}{\lambda}. \tag{3}$$

Thanks to the Earth rotation, the projected baseline $\vec{R}^{\perp}_{kk'}$ is continuously changing and traces arcs of ellipses on the Fourier plane. The combination of all arcs for all pair of telescopes results in a non-uniform and non-complete Fourier coverage with $m/2$ visibilities characterizing the interferometer. For illustration, Figure 3 shows an example of the Fourier sampling in the so-called (u, v) plan.

2.2 Interferometric inverse problem

2.2.1 General case

Assuming that the signals $I(\vec{p})$ and $A(\vec{p})$ are band-limited, the Nyquist-Shannon theorem requires to uniformly sample these signals at a resolution at least twice their bandwidth to fully describe them. We can thus defined the points $\vec{p}_i \in \mathbb{R}^2$, with $1 \leq i \leq N$, of the uniform grid of size $N^{1/2} \times N^{1/2}$ where these signals are sampled. Their sampled version may be denoted as $x \in \mathbb{R}^N = \{x_i = I(\vec{p}_i)\}_{1 \leq i \leq N}$, and $a \in \mathbb{R}^N = \{a_i = A(\vec{p}_i)\}_{1 \leq i \leq N}$. These sampled signals are also described by their Fourier coefficients, $\hat{x} \in \mathbb{C}^N$ and $\hat{a} \in \mathbb{C}^N$, computed on a discrete uniform grid of $N^{1/2} \times N^{1/2}$ spatial frequencies $\{\vec{u}_i\}_{1 \leq i \leq N}$: $\hat{x} = \{\hat{I}(\vec{u}_i)\}_{1 \leq i \leq N}$ and $\hat{a} = \{\hat{A}(\vec{u}_i)\}_{1 \leq i \leq N}$. The signals $I(\vec{p})$ and $A(\vec{p})$ being real, we also have the symmetry $\hat{I}(\vec{u}_i) = \hat{I}^*(-\vec{u}_i)$ and $\hat{A}(\vec{u}_i) = \hat{A}^*(-\vec{u}_i)$.

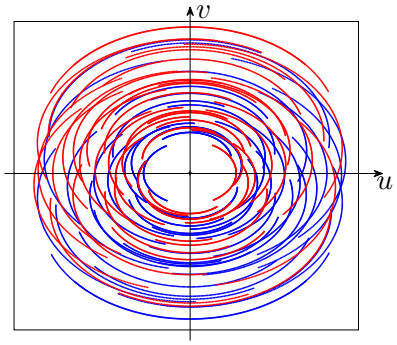

Figure 3: An example of the (u, v) sampling over arcs of ellipses.

During an observation period, $m/2$ complex visibilities are acquired by the interferometric array at different frequencies $\vec{u}_b \in \mathbb{R}^2$ with $1 \leq b \leq m/2$. In the general case, the frequencies probed $\{\vec{u}_b\}_{1\leq b\leq m/2}$ do not belong to the set of discrete frequencies $\{\vec{u}_i\}_{1\leq i\leq N}$ describing the signals x and a, and a gridding operation is usually performed on the visibilities before doing any reconstruction. This operation allows the use of the standard fast Fourier transform. However, it also induces some artifacts in the reconstructed image and leads to a loss of information. This gridding operation might be avoided by the use of some fast algorithms which have been designed to compute the Fourier transform of a signal on non-equispaced frequencies (NFFT[3]). For simplicity, we will consider here that the probed frequencies $\{\vec{u}_b\}_{1\leq b\leq m/2}$ belong to the set of discrete frequencies $\{\vec{u}_i\}_{1\leq i\leq N}$. Furthermore, as the signals probed are real, we can also consider that the set of frequencies $\{\vec{u}_b\}_{1\leq b\leq m/2}$ is limited to one-half of the Fourier plane. The $m/2$ measured complex visibilities $\{\widehat{AI}(\vec{u}_b)\}_{1\leq b\leq m/2}$ might be affected by independent noise: $n \in \mathbb{C}^{m/2}$. The symmetric visibilities can be deduced afterwards and the resulting m complex visibilities can be regrouped in a vector denoted $y \in \mathbb{C}^m$ (the noise vector becomes thus $n \in \mathbb{C}^m$ but with only $m/2$ independent values). In this setting, the number of independent complex visibilities $m/2$ is smaller than the number of Fourier coefficients $N/2$ needed to characterize exactly the signals x and an ill-posed inverse problem is defined to reconstruct the image:

$$y = \Phi x + n, \tag{4}$$

where

$$\Phi = MFD. \tag{5}$$

In the relation (5), $D = \{D_{ij} = a_i\delta_{ij}\}_{1\leq i,j\leq N} \in \mathbb{R}^{N\times N}$ is a matrix representing the illumination function, $F \in \mathbb{C}^{N\times N}$ is the matrix notation of the discrete Fourier transform, and $M \in \mathbb{R}^{m\times N}$ is a binary matrix which selects the m complex visibilities probed by the interferometer. This matrix contains only one non-zero value per line. The resulting matrix Φ is called the sensing matrix.

2.2.2 Cosmic string case

As explained in the introduction (see Section 1), some theories predict the existence of two components in the cosmic microwave background. In this context, the overall CMB signal may be seen as a linear superposition of a string signal $X^{stg}(\vec{p})$ and a pure Gaussian component $G(\vec{p})$, seen as noise. The amplitude of the string signal X^{stg} is controlled by the dimensionless string tension ρ: $\rho X^{stg}(\vec{p}) + G(\vec{p})$. In the following, we will denote $X^{\rho}(\vec{p})$ the string signal multiplied by the string tension. The discrete versions of these signals are denoted: $x^{\rho} \in \mathbb{R}^N = \{x_i^{\rho} = X^{\rho}(\vec{p}_i)\}_{1\leq i\leq N}$ for the string signal, and $g \in \mathbb{R}^N = \{g_i = G(\vec{p}_i)\}_{1\leq i\leq N}$ for the Gaussian CMB signal. For illustration, Figure 4 represents simulated maps of both signals (on a field of view of $1.8° \times 1.8°$ and at resolution of 0.42 arcminute) and their linear superposition for a string tension of $2.0e-07$ as well as the corresponding magnitude of gradient. The simulation of the string induced CMB map is provided by Fraisse et al. ([8]). The simulation of the Gaussian CMB map is performed thanks to the technique described in article [12] and recalled in Appendix A. We can notice that the strings are not detectable by eye in

[3]http://www-user.tu-chemnitz.de/ potts/nfft/

the noisy temperature map (top right) while a part of it appears in the magnitude of the gradient (bottom right). The gradient operator actually enhance the high frequency features such as the strings.

In the context of radio-interferometry, it is possible to re-write the general inverse problem (4) for this particular case. It becomes:

$$y = \Phi x^{\rho} + \Phi g, \tag{6}$$

in the absence of any instrumental noise. As the signal g is seen as noise here, the term Φg will be further denoted as $n^g \in \mathbb{C}^m = \{n_i^g\}_{1 \leq i \leq m}$ and the above equation becomes:

$$y = \Phi x^{\rho} + n^g. \tag{7}$$

The signals at our disposition are in unit of temperature, so to compute the visibilities in Jy, we have to convert our map in flux density as described in the article [14]. The goal is now to reconstruct the string signal x^{ρ} from the $m/2$ independent measured complex visibilities.

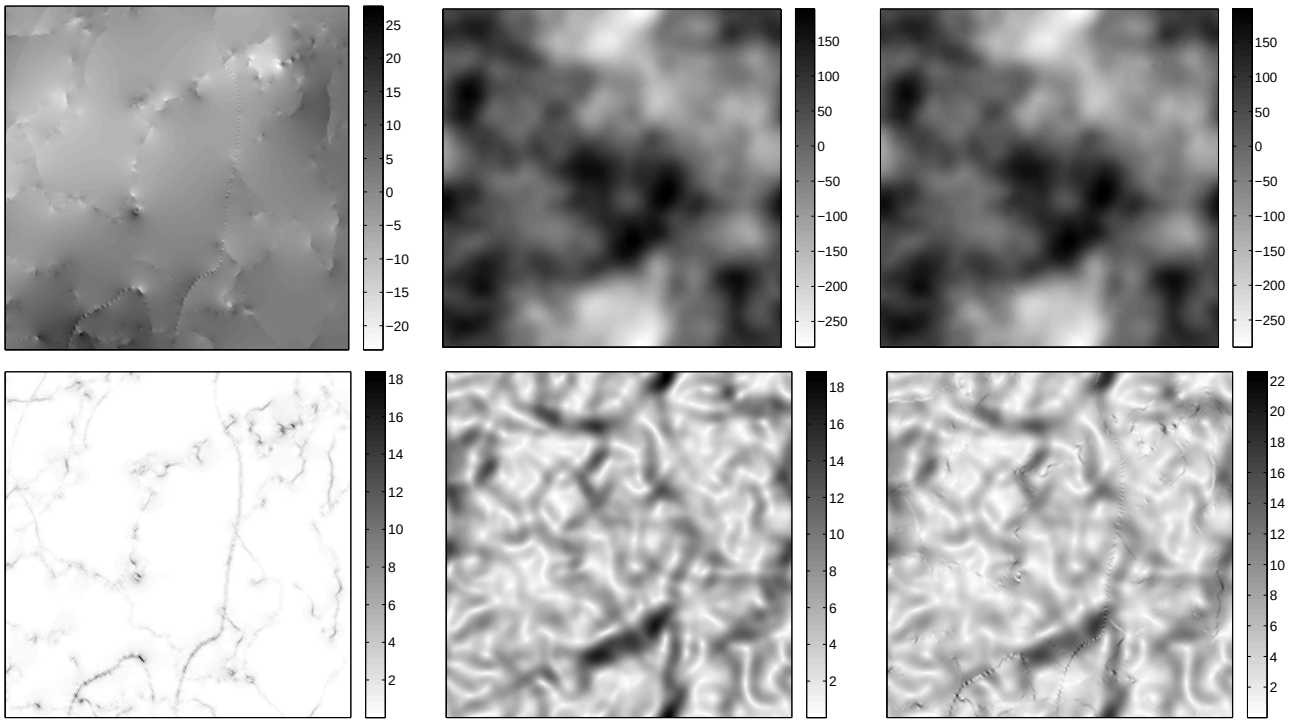

Figure 4: Top panels: simulated string induced CMB map in μK for a string tension of $2.0e-07$ on a field of view of $1.8^{\circ} \times 1.8^{\circ}$ (left), simulated Gaussian CMB map in μK on the same field of view with the primary CMB anisotropies only (center), combination of the string induced CMB map and the Gaussian CMB map (right). Bottom panels: corresponding maps of the magnitude of the gradient.

3 Bayesian perspective

In this section, we present the statistical model of the Gaussian CMB signal and of the string signal. We then present a method to estimate the string tension from the interferometric data acquired and show finally how to address the inverse problem posed in a Bayesian perspective.

3.1 Statistical model of the Gaussian CMB signal

The Gaussian noise component g can be modelled as a result of a random Gaussian process with a fixed discrete angular power spectrum $\mathcal{C}_l^g$ set by the standard cosmological model (see Appendix A). In this application, we only consider the primary CMB anisotropies of the power spectrum (see Fig. 6). Let us denote l the angular frequency which is equal to $2\pi\|\vec{u}\|_2$ (where $\|\cdot\|_2$ is the ℓ_2 norm and $\vec{u}$ is the spatial frequency). In this model, the Fourier coefficients $\hat{g}(\vec{u})$ of g can be seen as independent random Gaussian variables with mean 0 and variance:

$$\mathcal{C}_l^g = \langle \hat{g}(\vec{u})\hat{g}^*(\vec{u'})\rangle = \Omega_{pix} N \, \mathcal{C}_l^G \delta_{ll'}, \tag{8}$$

where Ω_{pix} is the pixel area of the discrete map in the real space and $\mathcal{C}_l^G$ is the continuous angular power spectrum of the signal $G(\vec{p})$:

$$\langle \hat{G}(\vec{u})\hat{G}^*(\vec{u'})\rangle = (2\pi)^2 \mathcal{C}_l^G \delta(l-l'). \tag{9}$$

When this signal is observed by an interferometric array as the one defined in the previous section, it is first multiplied by the illumination function A. This multiplication induces a convolution in the Fourier domain. Consequently, the resulting Fourier coefficients $\widehat{Ag}(\vec{u})$ (i.e. the coefficients $\{n_b^g\}_{1\leq b\leq m}$ of the vector n^g) can still be considered as zero-mean random Gaussian variables but not any more independent. However, the full width at half maximum (FWHM) of the illumination function is large enough (40 arcminute for our simulations) to consider that its FWHM in the Fourier domain is very small (0.022 arcminute^{-1}). The correlations induced are thus very local and the variables may reasonably be considered independent. In this context, it is possible to show that the variance of the zero-mean Gaussian discrete variable n_b^g at the frequency $\vec{u}_b$ can be computed as follows:

$$\sigma_b^2 = \mathcal{C}_l^{ag} = \mathcal{C}_l^g \star |\hat{a}(l)|^2 = \int dl \, \mathcal{C}_l^g \, |\hat{a}(l_b - l)|^2. \tag{10}$$

The total probability distribution of the entire noise vector $n^g \in \mathbb{C}^m$ signal is:

$$\pi(n^g) \propto \exp\left[-\sum_{b=1}^{m/2} \frac{|n_b^g|^2}{\sigma_b^2} \right], \tag{11}$$

as only the first $m/2$ variables are independent (the others being completely fixed by the first $m/2$ variables).

3.2 Statistical model of the string signal

The string signal can be understood as a result of a statistical process well modelled by Generalized Gaussian Distribution (GGD) in a wavelet space ([9]). We consider here a redundant steerable wavelet basis Ψ with 6 scales $j\,(1 \leq j \leq 6)$ including a low pass and a high pass axisymmetric filters and four intermediate scales defining steerable wavelets with 6 orientations $q\,(1 \leq q \leq 6)$ ([13]). We denote the wavelet vector $\alpha \in \mathbb{R}^T = \{\alpha_i\}_{1\leq i\leq T}$: $x^\rho = \Psi\alpha$. If we assume that the string tension ρ is known, the GGD prior distribution π_j of a wavelet coefficient α_i only depends on the scale and can be modelled as follows:

$$\pi_j(\alpha_i|\rho) \propto \exp\left[-\left| \frac{\alpha_i}{\rho u_j} \right|^{v_j} \right]. \tag{12}$$

The parameters v_j are called shape parameters. A shape parameter equals to 2 identifies a Gaussian distribution, a one equals to 1 identifies a Laplacian distribution. It can be understood as a measure of the compressibility of the signal. Values close to 0 yields very peaked probability distributions with heavy tails, i.e. very compressible

signals, whereas values close to 2 represents a non-sparse distributions. The parameters u_j are called scale parameters and are linked directly to the standard deviation of the GGD distribution. These parameters are estimated from the wavelet decompositions of 16 training simulations of a string signal similar to the one presented in Section 2.2.2, in the unit of flux density (Jy) and for a string tension of 1. To do so, we use a moment fitting method. The kurtosis k_j (the ratio of the fourth central moment to the square of the variance) and the variance σ_j^2 within each scale are estimated assuming statistical isotropy. As the theoretical expressions of the kurtosis and the variance of a GGD depend only on u_j and v_j, it is possible to get an estimation of these last two parameters from the estimated kurtosis and variance by solving numerically:

$$k_j = \frac{\Gamma(5/v_j)\Gamma(1/v_j)}{\Gamma(3/v_j)^2} \tag{13}$$

$$\sigma_j^2 = \rho^2 \frac{\Gamma(3/v_j)}{\Gamma(1/v_j)} u_j^2, \tag{14}$$

where Γ stands for the Gamma function. The values obtained are: $\{v_1 = 0.43, v_2 = 0.39, v_3 = 0.47, v_4 = 0.58, v_5 = 0.76, v_6 = 1.86\}$ and $\{u_1 = 8.9e-03, u_2 = 2.8e-03, u_3 = 2.2e-02, u_4 = 1.5e-01, u_5 = 9.5e-01, u_6 = 5.7e+01\}$. For illustration, Figure 5 shows the GGD's corresponding to these estimated shape and scale parameters, superimposed on the histograms obtained from the 16 simulations.

If we now assume that the wavelet coefficients are independent, the total probability distribution of the entire signal is simply the product of the probability distributions of each wavelet coefficient:

$$\pi(\alpha|\rho) \propto \exp\left[-\sum_{i=1}^{T}\left|\frac{\alpha_i}{\rho u_j}\right|^{v_j}\right]. \tag{15}$$

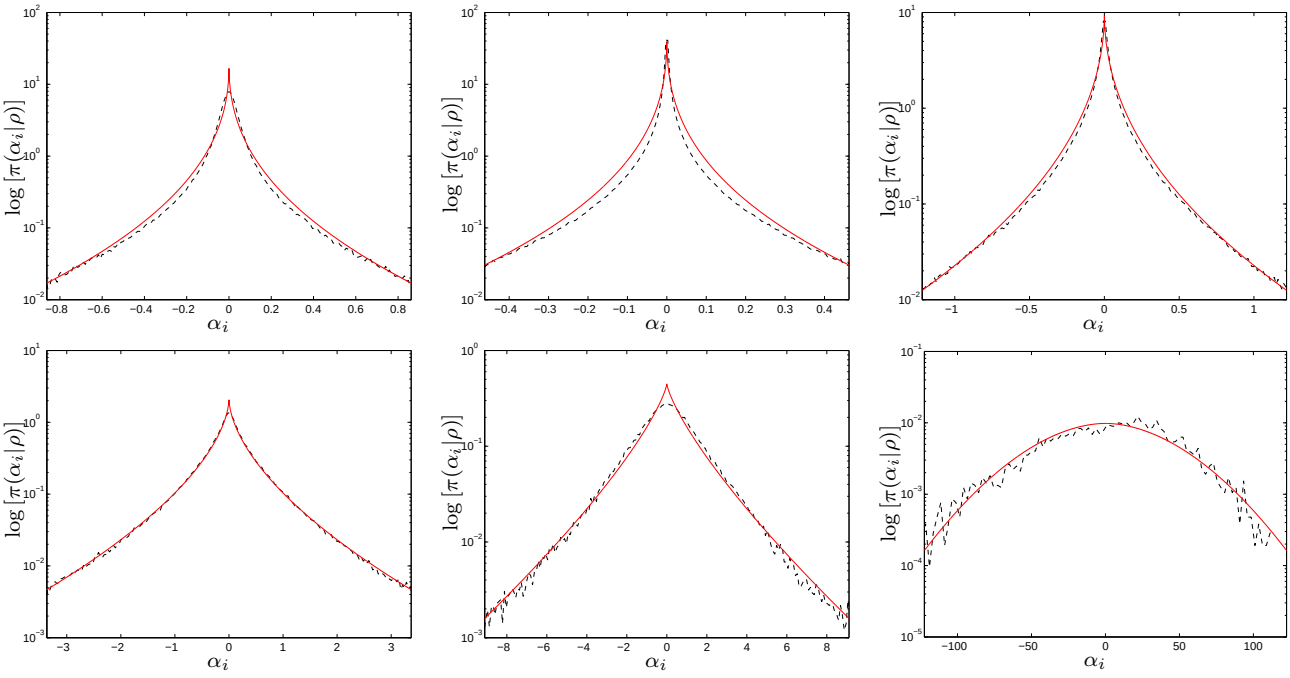

Figure 5: Logarithm of the generalized Gaussian distributions $\pi(\alpha_i|\rho)$ (red solid line) with the corresponding histograms of the coefficients α_i obtained from the 16 training string simulations (black dashed line). Top panels: high pass axisymmetric filter (left), first scale of steerable wavelets (center), second first scale of steerable wavelets (right). Bottom panels: third scale of steerable wavelets (right), fourth scale of steerable wavelets (center), low pass axisymmetric filter (right).

As for the Gaussian CMB signal, the discrete angular power spectrum of the string signal $\mathcal{C}_l^{x^\rho}$ is also known. An analytical expression of the continuous power spectrum $\mathcal{C}_l^{X^\rho}$ was provided by Fraisse et al. ([8])

$$\forall l > 1000, \quad l(l+1)\,\mathcal{C}_l^{X^\rho} = \rho^2 \frac{14}{1000^{-p}} l^{-p} \text{ with } p = 0.889. \tag{16}$$

The discrete angular power spectrum is deduced by replacing $\mathcal{C}_l^G$ by $\mathcal{C}_l^{X^\rho}$ in equation (8):

$$\forall l > 1000, \quad \mathcal{C}_l^{x^\rho} = \Omega_{pix} N\, \mathcal{C}_l^{X^\rho} \delta_{ll'}. \tag{17}$$

When this signal is multiplied by the illumination function A, the new power spectrum can be computed by using relation (10) thus obtaining:

$$\mathcal{C}_l^{ax^\rho} = \mathcal{C}_l^{x^\rho} \star |\hat{a}(l)|^2 = \int dl'\, \mathcal{C}_{l'}^{x^\rho}\, |\hat{a}(l-l')|^2. \tag{18}$$

For illustration, Figure 6 shows the continuous angular power spectrum of the string signal for a string tension of $2.0e-07$.

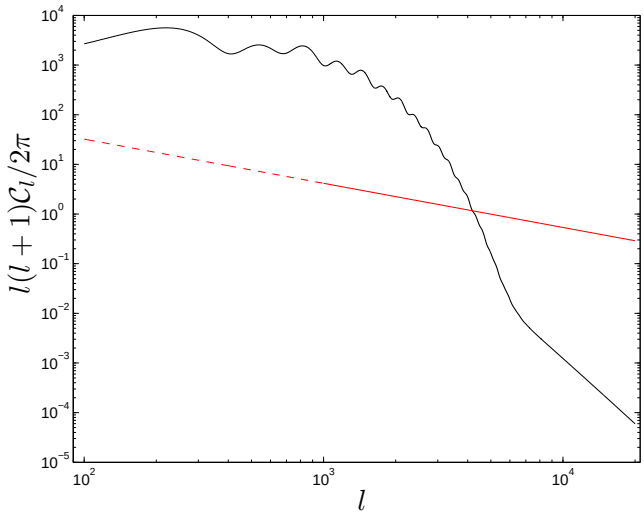

Figure 6: Continuous angular power spectrum in μK^2 of the primary anisotropies of the CMB (black line) and of the string signal for a string tension of $2e-07$ (red line). The angular frequencies range from $l = 1.0e+02$ and $l = 2.0e+04$ and the data are represented on $\log_{10} - \log_{10}$ axes.

3.3 Estimation of the string tension

The estimation of the string tension is done by fitting the power spectrum C_l^y of the acquired data y to the power spectrums of the Gaussian CMB signal $\mathcal{C}_l^{ag}$ of the string signal $\mathcal{C}_l^{ax^\rho}$ at the available angular frequencies l (see article [9]). The model used in this article assumes that both signals arise from two statistically independent isotropic Gaussian random processes. Under this model, the observed complex visibilities y_b have a power spectrum:

$$\mathcal{C}_{l_b}^y = \left\langle |y_b|^2 \right\rangle = \mathcal{C}_{l_b}^{ax^\rho} + \mathcal{C}_{l_b}^{ag}, \tag{19}$$
$$\forall l_b > 1000 \text{ and } l_b \in \{2\pi \|\vec{u}_b\|_2\}_{1 \leq b \leq m/2}.$$

The probability distribution $\pi(y|\rho)$ of y given ρ reads as:

$$\pi(y|\rho) = \frac{1}{\pi \prod_{b=1}^{m/2} \mathcal{C}_{l_b}^y} \exp\left[-\sum_{b=1}^{m/2} \frac{|y_b|^2}{\mathcal{C}_{l_b}^y} \right]. \tag{20}$$

An estimation of the string tension may be obtained by computing the maximum *a posteriori* of the probability $\pi(\rho|y)$:

$$\begin{aligned} \tilde{\rho} = \arg\max_{\rho' \in \mathbb{R}} \pi(\rho'|y) &= \arg\max_{\rho' \in \mathbb{R}} \frac{\pi(\rho')\pi(y|\rho')}{\pi(y)} \text{ with } \pi(y) = C^{st} \\ &= \arg\max_{\rho' \in \mathbb{R}} \pi(\rho')\pi(y|\rho'), \end{aligned} \tag{21}$$

where the prior $\pi(\rho')$ is considered uniform within an interval $[0, \rho_{max}]$ with a upper bound $\rho_{max} > 2.0e-07$. Finally, an estimation of ρ is obtained by solving numerically:

$$\tilde{\rho} = \arg \min_{\rho' \in [0, \rho_{max}]} \left[\sum_{b=1}^{m/2} \frac{|y_b|^2}{\mathcal{C}_{l_b}^y} + \sum_{b=1}^{m/2} \log(\pi \mathcal{C}_{l_b}^y) \right]. \tag{22}$$

3.4 Wavelet domain Bayesian reconstruction scheme

As we have statistical models for both signals, we may design now a Bayesian method to reconstruct the string signal from the observed visibilities. The prior distribution of our string signal is modelled in the wavelet domain, we have thus to re-write the inverse problem (7) as follows, to use this information:

$$y = \Phi\Psi\alpha + n^g. \tag{23}$$

The idea is now to recover the vector α from the visibilities y. The original signal might be obtained afterwards by applying Ψ to the estimated vector $\tilde{\alpha}$: $\tilde{x}^\rho = \Psi\tilde{\alpha}$.

If we assume that the string tension ρ is available or has been estimated thanks to the method described before, we may find an estimation $\tilde{\alpha}$ of the vector α by computing the maximum *a posteriori* (MAP) of the probability $\pi(\alpha^\rho|y, \rho)$:

$$\begin{aligned} \tilde{\alpha}^\rho &= \arg\max_{\alpha' \in \mathbb{R}^T} \pi(\alpha'|y, \rho) \\ &= \arg\max_{\alpha' \in \mathbb{R}^T} \frac{\pi(y|\alpha', \rho)\pi(\alpha'|\rho)}{\pi(y)} \\ &= \arg\max_{\alpha' \in \mathbb{R}^T} \pi(y|\alpha', \rho)\pi(\alpha'|\rho) \end{aligned} \tag{24}$$

where $\pi(\alpha'|\rho)$ is given by equation (15) and $\pi(y|\alpha', \rho)$ is the probability of the noise $n^g = y - \Phi\Psi\alpha'$:

$$\pi(y|\alpha', \rho) = \pi(n^g) \propto \exp\left[-\sum_{b=1}^{m/2} \frac{|y_b - (\Phi\Psi\alpha')_b|^2}{\sigma_b^2} \right]. \tag{25}$$

Combining all these equations yields:

$$\tilde{\alpha} = \arg \min_{\alpha' \in \mathbb{R}^T} \left[\sum_{b=1}^{m/2} \frac{|y_b - (\Phi\Psi\alpha')_b|^2}{\sigma_b^2} + \sum_{i=1}^{T} \left| \frac{\alpha'_i}{\rho u_j} \right|^{v_j} \right]. \tag{26}$$

In the equation above, some shape parameters v_j are smaller than 1. The problem to solve is thus non-convex and finding the MAP estimation $\tilde{\alpha}$ is not an easy task. In the next section we show how to address this problem after reformulating it in the flexible framework of compressed sensing.

4 Compressed sensing

In this section, we introduce briefly the theory of compressed sensing and its main results. Using these results, we then describe two possible methods which try to reconstruct a string signal from radio-interferometric data. The first one is a very general method that uses only the fact that our signal is compressible in a wavelet basis. The second method is specific to the string signal and uses the statistical prior detailed in the previous section.

4.1 Introduction to the theory

The theory of compressed sensing shows that it is possible to reconstruct a signal accurately, sometimes even exactly, from a far smaller number of samples than required by the Nyquist-Shannon theorem ([1], [2], [3], [7]). This theory applies to sparse or compressible signals. Let us expand a signal $x \in \mathbb{R}^N$ in a basis $\Psi \in \mathbb{R}^{N \times T} = \{\psi_1 \psi_2 \dots \psi_T\}$:

$$x = \sum_{i=1}^{T} \alpha_i \psi_i. \tag{27}$$

The coefficients α_i are the scalar product of the vector x with the vectors ψ_i: $\alpha_i = \langle x, \psi_i \rangle$. In a matrix notation, we have $x = \Psi\alpha$ with $\alpha \in \mathbb{R}^T = \{\alpha_i\}_{1 \leq i \leq T}$. The signal x is said to be sparse in the basis Ψ if it decomposition α contains only a small number $K \ll N$ of non-zero values. It is compressible if it contains a small number of significant coefficients $K \ll N$.

This signal x is probed by collecting partial information through $m \approx K$ linear measurements in a sensing basis $\Phi \in \mathbb{R}^m \times R^N$ possibly affected by independent and identically distributed Gaussian noise $n \in \mathbb{R}^m$. We can consider, without any loss of generality, a noise with a mean 0 and a variance 1. We regroup the measurements in a vector $y \in \mathbb{R}^m$:

$$y = \Phi x + n. \tag{28}$$

The number m of measurements being far smaller than the number N of samples needed to fully describe the signal x, the above problem is ill-posed. However, as the signal x has a sparse representation in the basis Ψ, it depends on much smaller number of unknowns. The ill-posed problem is regularized by this assumption of compressibility. In this context, the theory of compressed sensing states that if the sensing matrix $\Theta = \Phi\Psi \in \mathbb{R}^m \times \mathbb{R}^T$ respects the restricted isometry property (RIP), then it is possible possible to have a very accurate approximation of the compressible signal x by solving the so-called Basis Pursuit denoise (BPDN) problem. Let us first re-write the problem above to introduce the vector α and the matrix Θ:

$$y = \Phi\Psi\alpha + n = \Theta\alpha + n. \tag{29}$$

The BPDN consists in the minimization of the ℓ_1 norm of the vector α under a constraint on the ℓ_2 norm of the residual noise:

$$\min_{\alpha' \in \mathbb{R}^T} \|\alpha'\|_1 \text{ subject to } \|y - \Theta\alpha'\|_2 \leq \epsilon. \tag{30}$$

$\|y - \Theta\alpha'\|_2^2 = \sum_{b=1}^{m} (y - \Theta\alpha')_b^2$ is an estimate of the noise n and follows thus a chi-square distribution with m degrees of freedom. ϵ^2 is thus equals to some suitable percentile of this distribution. The solution of (30) may be obtained through convex optimization algorithms.

The matrix Θ respects the restricted isometry property of order K if, for any $k \leq K$ sparse vector α_k in the basis Ψ, there exists a constant $\delta_K < 1$ such that:

$$(1 - \delta_K)\|\alpha_k\|_2^2 \leq \|\Theta\alpha_k\|_2^2 \leq (1 + \delta_K)\|\alpha_k\|_2^2. \tag{31}$$

This property says that the matrix Θ should preserve the ℓ_2 norm of any K sparse signals and, as $\delta_K < 1$, means that no K sparse vector can be in the null space of Θ. This is useful as otherwise there would be no hope of reconstructing these vectors. The theory of compressed sensing shows that if the matrix Θ satisfies the RIP of order $2K$ with $\delta_{2K} < \sqrt{2} - 1$, then the solution to (30) provides an accurate reconstruction of the signal α.

In the absence of noise, $\epsilon = 0$, the problem (30) as a unique solution and the reconstruction is exact if the signal is exactly sparse. In the presence of noise and for compressible signal, we have:

$$\|\alpha - \tilde{\alpha}\|_2 \leq C_{1,K}\epsilon + C_{2,K}\frac{\|\alpha - \alpha_K\|_1}{\sqrt{K}}, \tag{32}$$

where α_K is the best approximation of α obtained by keeping the K largest coefficients of the vector α. This relation provides a strong stability result on the quality of the reconstructions. It has been emphasized that the constant $C_{1,K}$ and $C_{2,K}$ are rather small. For example, for $\delta_{2K} = 0.2$ then $C_{1,K} = 8.5$ and $C_{2,K} = 4.2$.

We can also recall that $\forall\epsilon$, there exists a parameter $\tau > 0$ such that the BPDN problem (30) is strictly equivalent to:

$$\min_{\alpha' \in \mathbb{R}^T} \frac{1}{2}\|y - \Phi\Psi\alpha'\|_2 + \tau\|\alpha\|_1, \tag{33}$$

In a Bayesian perspective, the problem (33) is equivalent to finding the maximum *a posteriori* of a signal with a Laplacian prior distribution in presence of independent and identically distributed Gaussian noise.

Finally, other minimization problems have been proposed for the recovery of α ([4], [6]). We can for example substitute the ℓ_1 norm in the BPDN problem by a ℓ_p norm. In a Bayesian perspective, it corresponds to suppose that the signal α is described by a generalized Gaussian distribution. This description is of a particular interest for us because one would have already noticed the similarity between the general inverse problem (29) for the compressed sensing and ours (23), as well as the GGD distribution of our string signal and the possibly to use a ℓ_p norm in (30).

4.2 Compressed sensing reconstruction of a string signal

We have seen that the string signal is compressible in a wavelet space (see section 3.2 page 6). As only partial information of it is acquired through $m/2$ linear measurements, the first idea is to try to reconstruct the signal by solving the BPDN problem (30). However, we must recall that the noise in the BPDN problem (30) was assumed identically distributed. It is not the case for the Gaussian CMB signal and the constraint $\|y - \Phi\Psi\alpha'\|_2 \leq \epsilon$ has not any more any statistical meaning, i.e. it does not follows any more a chi-square distribution. For the Gaussian CMB noise, we can consider the sum $\sum_{b=1}^{m} |y_b - (\Phi\Psi\alpha')_b|^2/\sigma_b^2$. Because of the symmetry in the Fourier domain, we have:

$$\begin{aligned}\sum_{b=1}^{m} \frac{|y_b - (\Phi\Psi\alpha')_b|^2}{\sigma_b^2} &= 2\sum_{b=1}^{m/2} \frac{|y_b - (\Phi\Psi\alpha')_b|^2}{\sigma_b^2} \\ &= \sum_{b=1}^{m/2} \frac{(n_b^{re})^2}{\sigma_b^2/2} + \frac{(n_b^{im})^2}{\sigma_b^2/2}.\end{aligned} \tag{34}$$

where n_b^{re} is the real part of $y_b - (\Phi\Psi\alpha')_b$ and n_b^{im} its imaginary part. $(n_b^{re})^2/(\sigma_b^2/2)$ and $(n_b^{im})^2/(\sigma_b^2/2)$ are independent and identically real Gaussian variables of mean 0 and variance 1. The previous sum follows thus a chi-square distribution with m degrees of freedom. The level of the noise residual ϵ^2 is chosen equals to some $100\,\theta^{\text{th}}$ percentile of this chi-square distribution, i.e. $p(\chi^2 \leq \epsilon^2) = \theta$. Finally, the problem to solve reads as:

$$\min_{\alpha' \in \mathbb{R}^T} \|\alpha'\|_1 \text{ subject to } \sum_{b=1}^{m} \frac{|n_b'|^2}{\sigma_b^2} \leq \epsilon^2. \tag{35}$$

As already explained in the previous section, the ℓ_1 norm in the above problem can be linked to the fact that the signal α to recover is well described by a Laplacian distribution. However, we know in our case that the signal α is well described by GGD's. The idea is know to replace the l_1 by another one to take into account this *a priori* information.

If a signal is well modeled by a GGD with a shape parameter p, we can replace the ℓ_1 norm in the BPDN problem by a ℓ_p norm. The string signal is well modeled by the product of 6 different GGD's in the wavelet domain, we can thus define the corresponding "s norm" and use it to recover the signal. We define the "s norm" as follows:

$$\|\alpha\|_s = \sum_{i=1}^{T} \left| \frac{\alpha'_i}{\rho u_j} \right|^{v_j}. \tag{36}$$

And by replacing the ℓ_1 norm in the BPDN problem by this s norm and the problem to solve becomes:

$$\min_{\alpha' \in \mathbb{R}^T} \|\alpha'\|_s \text{ subject to } \sum_{b=1}^{m} \frac{|n'_b|^2}{\sigma_b^2} \leq \epsilon^2, \tag{37}$$

and will be called statistical basis pursuit (SBP) problem. As for the MAP problem (26), the SBP problem is not convex. However there exists some techniques to converge to the solution. Candès et al. present a simple iterative algorithm to converge to the ℓ_0 norm in article [4]. The algorithm presented here is a modification of the former to converge to the s-norm (36):

1. Set the weights $\omega_i^{(0)} = 1,\ i = 1, \ldots, T$.
2. Solve the weighted BPDN problem:

$$\alpha^{(l)} = \arg\min \sum_{i=1}^{T} \omega_i^{(l)} |\alpha_i| \text{ subject to } \sum_{b=1}^{m} \frac{|n'_b|^2}{\sigma_b^2} \leq \epsilon^2. \tag{38}$$

3. Update the weights:

$$\omega_i^{(l+1)} = \frac{1}{(\rho u_j)^{v_j} (|\alpha_i^{(l)}|^{(1-v_j)} + \xi)},\ i = 1, \ldots, T. \tag{39}$$

4. Terminate on converge or after a given number of iterations. Otherwise, increment l and go to step 2.

The value ξ is a regularization parameter and is here to avoid numerical errors if $\alpha_i^{(l)}$ is equal to zero. It should be negligible and we choose it equal to a small percentage of the standard deviation of the GGD at the considered scale j (14). No proof of convergence of this algorithm and stability results of the reconstructed signal are yet provided. Its performances are just assessed on the basis of simulations (see Section 6). However, we can easily check that if the algorithms converge (i.e. $\alpha^{(l+1)} = \alpha^{(l)}$) and the value of ξ is negligible, then it converges to the solution of (37):

$$\begin{aligned}
\sum_{i=1}^{T} \omega_i^{(l+1)} |\alpha_i^{(l+1)}| &= \sum_{i=1}^{T} \frac{|\alpha_i^{(l+1)}|}{(\rho u_j)^{v_j} (|\alpha_i^{(l)}|^{(1-v_j)} + \xi)} \\
&\approx \sum_{i=1}^{T} \frac{|\alpha_i^{(l)}|}{(\rho u_j)^{v_j} (|\alpha_i^{(l)}|^{(1-v_j)})} \\
&= \sum_{i=1}^{T} \frac{|\alpha_i^{(l)}|^{v_j}}{(\rho u_j)^{v_j}} \\
&= \sum_{i=1}^{T} \left| \frac{\alpha_i^{(l)}}{\rho u_j} \right|^{v_j} \\
&= \|\alpha^{(l)}\|_s.
\end{aligned} \tag{40}$$

Finally, we can also notice that this algorithm needs the knowledge of the string tension ρ. Its value is estimated prior to any reconstruction thanks to the techniques detailed in Section 3.3.

5 CLEAN

In this section, we describe a very common algorithm used in radio astronomy to reconstruct an image from radio-interferometric data. This algorithm is called CLEAN. We then explain how to apply it to reconstruct our string signal.

5.1 The CLEAN algorithm

CLEAN is an iterative algorithm that aims at solving the general inverse problem (4) ([10]). This algorithm tries to reconstruct the initial image x multiplied by the illumination function: $\bar{x} = Ax \in \mathbb{R}^N$. The measured visibilities are first put on the uniform discrete Fourier grid of size $N^{1/2} \times N^{1/2}$ (a zero value is affected to the Fourier coefficients at all non-probed frequencies), and a discrete inverse Fourier transform is performed. The resulting image is called the dirty image $\bar{x}^{(d)} \in \mathbb{R}^N$. In our case, this operation is actually equivalent to apply the adjoint operator $\bar{\Phi}^\dagger$, of $\bar{\Phi} = MF$, to the visibilities y: $\bar{x}^{(d)} = \bar{\Phi}^\dagger y$. The inverse Fourier transform of the binary mask which selects the frequencies probed is called the dirty beam: $d \in \mathbb{R}^N = \bar{\Phi}^\dagger 1_m$, where $1_m = \{c_i = 1\}_{1 \leq i \leq m}$. We can also compute the expression of the noise in real space: $\bar{n}^{(d)} = \bar{\Phi}^\dagger n$. The general inverse problem (4) becomes in real space:

$$\bar{x}^{(d)} = d \star \bar{x} + \bar{n}^{(d)}. \tag{41}$$

The problem above being ill-posed, an assumption on the signal $\bar{x}$ must be made to regularize it. The assumption here is that the initial signal $\bar{x}$ can be expressed as a sum of Dirac spikes. The convolved image $d \star \bar{x}$ is thus constituted of a sum of dirty beams d centered at each location of a Dirac spikes. Based of this interpretation, the CLEAN algorithm is actually very simple:

1. Initialize the residual map $\bar{x}^{(0)} \in \mathbb{R}^N$ to $\bar{x}^{(d)}/c$ (c is the central value of the dirty beam) and the reconstructed map $\tilde{x} \in \mathbb{R}^N$ to the null vector.

2. At the iteration l, find the index i_{max} where the residual image $\bar{x}^{(l)} = \{\bar{x}_i^{(l)}\}_{1 \leq i \leq N}$ has its maximum absolute value and create a vector $\beta \in \mathbb{R}^N \{\beta_i = \bar{x}_i^{(l)} \delta_{i i_{max}}\}_{1 \leq i \leq N}$.

3. Compute the residual map at the next iteration $(l+1)$: $\bar{x}^{(l+1)} = \bar{x}^{(l)} - \gamma \bar{\Phi}^\dagger \Phi \beta / c$. γ is called the loop-gain factor and its value (smaller than 1) is usually equals to a few tenths. This factor enhances a lot the quality of the reconstruction even if it also increases the computational cost.

4. Add the vector β to the reconstructed map: $\tilde{x} = \tilde{x} + \gamma \beta$

5. Terminate if the noise level in the residual image is low enough or after a certain number of iteration. Otherwise go to step 2.

With this algorithm, the only assumption is that the original signal can be expressed as a sum of Dirac spikes. The sparsity or compressibility of the original signal is not explicitly imposed but implicitly assumed: the original signal should have its energy concentrated at specific location in real space. We can also acknowledge that some multi-scale versions of CLEAN are under development ([5]). The idea here is that the signal can be represented as a sum of wavelets. The algorithms is the same but the matrix Φ is replaced by $\Phi\Psi$ where Ψ is a wavelet basis. Some other techniques exists to solve the interferometric inverse problem (MEM, WIPE). They regularize the problem by the introduction of a smoothness prior and thus do not use the fact that a lot of signals in the Nature have a sparse or compressible representation like our string signal.

5.2 CLEAN reconstruction of a string signal

CLEAN can be directly be applied to our problem. Let us re-write the problem (41) to introduce the string and the Gaussian CMB signals:

$$\bar{x}^{(d)} = d \star \bar{x}^\rho + d \star \bar{g}. \tag{42}$$

The string signal $\bar{x}^{\rho}$ is completely buried in the Gaussian CMB signal (see Section 2.2.2). Before trying to reconstruct the string signal from the dirty image $\bar{x}^{(d)}$, a pre-filtering must be performed. One very simple and common solution is to do a matched filtering: the visibilities are divided by the variance of the Gaussian CMB noise. To do so, a diagonal matrix $W \in \mathbb{R}^{m \times m} = \{1/\sigma_b\}_{1 \leq b \leq m}$, called a whitening operator, is introduced in the operator $\bar{\Phi}$. This filter maximizes the signal-to-noise ratio in real space. We denote this new operator $\tilde{\Phi}$: $\tilde{\Phi} = W\bar{\Phi}$. CLEAN can now be applied by replacing $\bar{\Phi}$ by $\tilde{\Phi}$ in the above description. In a statistical sense, the stopping criterion used is set in terms of (34). For a given candidate reconstruction $\tilde{x}$, the noise residual $y - \bar{\Phi}\tilde{x}$ is computed and the algorithm is stopped if the level of the noise is low enough.

The goal is now to compare the performances of CLEAN, which is a very standard algorithm in radio-astronomy and does not use any *a priori* information for the reconstruction of the signal, to the ones of BPDN which explicitly impose the compressibility of the signal, and the ones of SBP which is adapted to the statistical model of the string signal.

6 Simulations and results

In this section we discuss the performance of the three algorithms proposed to reconstruct a string signal from radio-interferometric data. We describe the experimental set up and the criteria measured to evaluate the performance and interpret the results.

6.1 Experimental set up

The performance of the three methods considered (BPDN (35), SBP (37), CLEAN) are studied at different string tensions and for different distributions of the visibilities. The string tensions are equi-spaced on the logarithmic scale and read as follows: $\{\rho_1 = 1.0e-09, \rho_2 = 3.2e-09, \rho_3 = 1.0e-08, \rho_4 = 3.2e-08, \rho_5 = 1.0e-07\}$. The distribution of the visibilities are very dependent of the interferometric array. In order to draw general conclusions, these distributions are simulated by uniformly selecting random visibilities in one half of the Fourier plane (the others being obtained by symmetry). We consider here 5 different distributions by selecting $m/2$ visibilities in one half of the Fourier plane, so that m/N is equal to 5, 10, 15, 20, and 25 percent. For one given distribution and one given string tension, the performance are evaluated over 30 simulations which consist of the superposition of one unique string induced CMB map with 30 simulations of the Gaussian CMB signal. Three different criteria are used to assess the performance. The first one is the signal-to-noise ratio between the magnitude of gradient of the original signal x^ρ multiplied by the illumination function and the magnitude of gradient the reconstructed images $\tilde{x}$ obtained with CLEAN, BPDN and SBP (for the last two algorithms the reconstructed signals are first re-multiplied by the illumination function):

$$\text{SNR}^{(|\vec{\nabla}x^\rho|,|\vec{\nabla}\tilde{x}|)} = 20\log_{10}\frac{\sigma^{|\vec{\nabla}x^\rho|}}{\sigma^{(|\vec{\nabla}x^\rho|-|\vec{\nabla}\tilde{x}|)}}, \tag{43}$$

where $|\vec{\nabla}\cdot|$ denotes the magnitude of gradient and $\sigma^{(\cdot)}$ is the standard deviation of the considered signal. We also measure the correlation coefficient of these signals:

$$r^{(|\vec{\nabla}x^\rho|,|\vec{\nabla}\tilde{x}|)} = \frac{\text{cov}^{(|\vec{\nabla}x^\rho|,|\vec{\nabla}\tilde{x}|)}}{\sigma^{|\vec{\nabla}x^\rho|}\sigma^{|\vec{\nabla}\tilde{x}|}}, \tag{44}$$

where $\text{cov}^{(|\vec{\nabla}x^\rho|,|\vec{\nabla}\tilde{x}|)}$ stands for the covariance between $|\vec{\nabla}x^\rho|$ and $|\vec{\nabla}\tilde{x}|$. Finally, we measure the kurtoses of the magnitudes of gradient of the reconstructed maps and compare them to the kurtosis of the magnitude gradient of the original signal.

The illumination function $A(\vec{p})$ is modelled as a Gaussian window with a FWHM of 40 arcminute. For BPDN and SBP, the constraint on the noise ϵ^2 (see (34)) is chosen equal to the 95^{th} percentile of the chi-square distribution with m degrees of freedom. This constraint is also chosen to be the stopping criterion of CLEAN. The loop gain factor for CLEAN is taken equal to 0.1. The BPDN problem is solved by using the SPGL1[4] toolbox.

To solve the SBP problem, a value of ξ has to be set. We take it equals to 1% of the standard deviation in the scale j considered. This value is very critical and controls the converge or not of the algorithm. These problems of convergence lead us to use a value of v_6 equal to 1 instead of 1.86. This restriction should induce only negligible artifacts because the coefficients of this scale (the last one, i.e. the low pass filter) do not participate to the identification of the string network as the string signal is buried in the noise at large scale ([9]).

For illustration, Figure 7 shows the magnitude of gradient of the simulated string signal multiplied by the illumination function. Figure 8 shows the magnitudes of gradient of the reconstructed maps obtained with CLEAN and SBP for two different Fourier coverages (5% and 10%), as well as the magnitudes of gradient of the dirty maps renormalized by the central value of the dirty beam.

[4] http://www.cs.ubc.ca/labs/scl/spgl1/

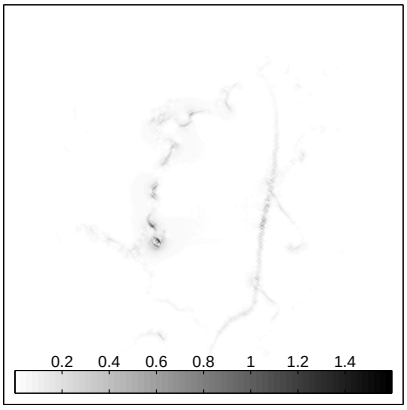

Figure 7: Magnitude of the gradient of the original simulated string signal in μK multiplied by the illumination function.

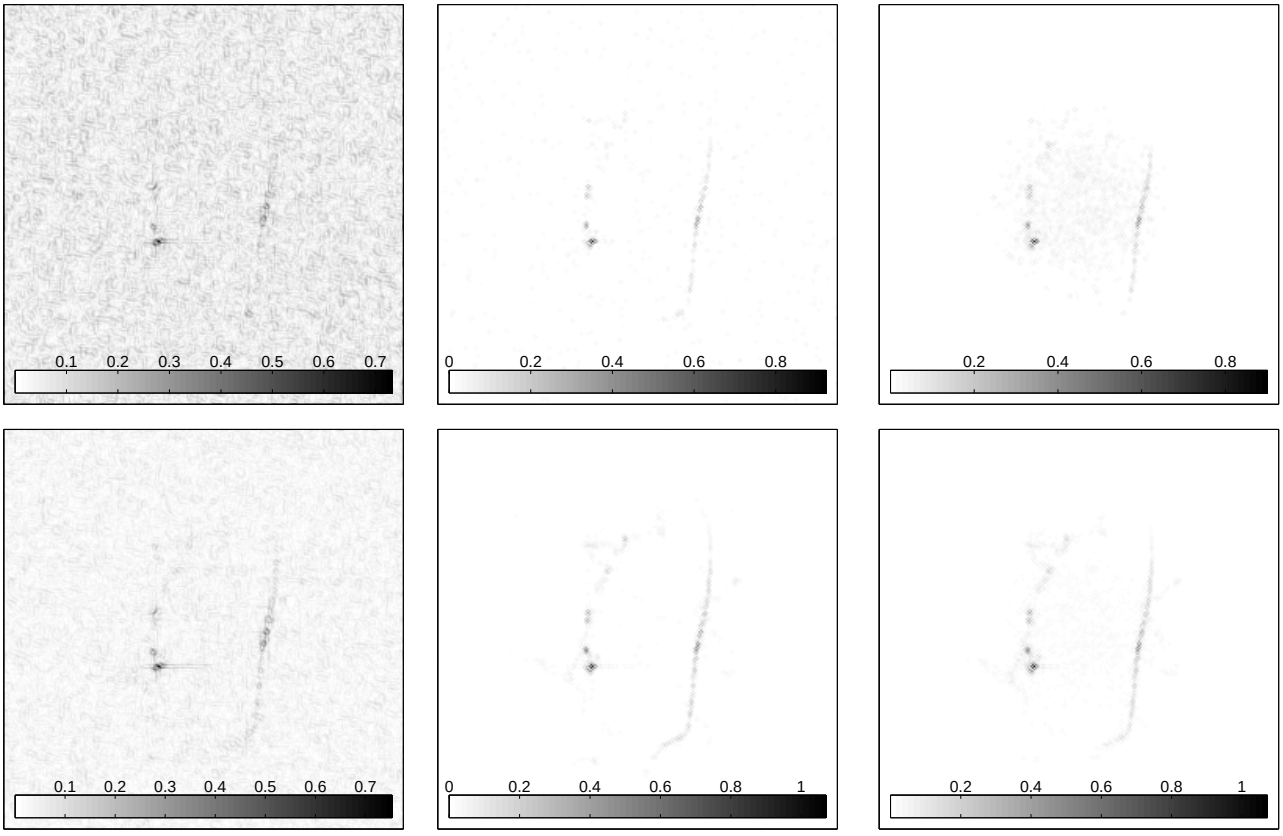

Figure 8: Magnitude of gradient of the dirty map (μK) renormalized by the center value of the dirty beam (left), the CLEAN reconstruction (μK) (center), the SBP reconstruction (μK) re-multiplied by the illumination function (right) for a Fourier coverage of 5% (top panels) and 10% (bottom panel).

6.2 Performance of the estimation of the string tension

The string tension is estimated prior to any SBP reconstruction thanks to the method described in Section 3.3 (page 9). For illustration, Figure 9 shows the posterior probability distributions of $\pi(\rho|y)$ at two string tensions ($1.0e-08$ and $3.2e-08$) for three different Fourier coverages (5, 15, and 25 percent). The accuracy of this method was already outlined in article [9]. Here, the results presented in Figure 9 show that the estimated string tensions are quite close to the true string tension ($7.9e-08$ for $1.0e-08$ and $2.5e-08$ for $3.2e-08$) even if at least 75% percent of the Fourier coefficients are not available to estimate the power spectrum of the superposition of the string and the Gaussian CMB signals.

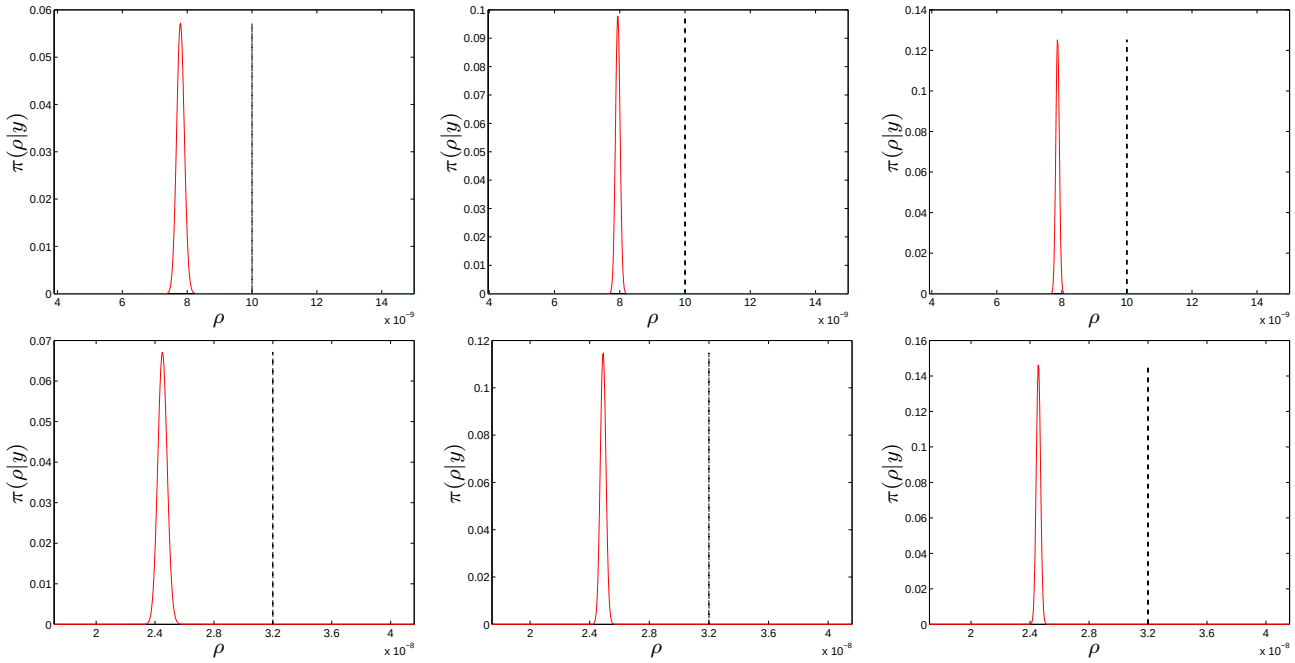

Figure 9: Posterior probability distribution of the string tension ρ given the visibilities y (red solid line). Top panels: true string tension of $1.0e-08$ (black dashed line) and a Fourier coverage equals to 5% (left), 15% (center), 25% (right). Bottom panels: true string tension of of $3.2e-08$ (black dashed line) and a Fourier coverage equals to 5% (left), 15% (center), 25% (right).

6.3 Performance of the reconstruction algorithms

The mean signal-to-noise ratio, correlation coefficient and kurtosis of the gradient magnitude of the reconstructed maps as well as the 1σ error bars are presented respectively Figure 10, Figure 11, and Figure 12. These values are presented as a function of the string tension for the 5 different distributions considered.

At $1.0e-09$, which is the limit of eye detectibility of the strings in the gradient magnitude of the reconstructed maps, the SNR's are roughly the same whatever the method used. However, it also means that there exist some techniques which are able to reconstruct a string signal from radio-interferometric data at a string tension lower than the experimental limit of $2.0e-07$ in the presence of the primary anisotropies of the CMB only. At higher string tensions, BP and CLEAN give SNR's close to each other but the SNR's obtained with SBP are much higher even if the difference decreases when the number of visibilities increases.

For the Fourier coverage of 5% and 10%, the correlation coefficient shows that the reconstructions obtained with BP and SBP are much better than those obtained with CLEAN. The difference decreases when the Fourier coverage increases but the mean value is always higher for SBP.

The huge error bars of the kurtosis at $1.0e-09$ shows that we are at the limit of the detectibility of the strings. At higher string tension, we notice that CLEAN always fails to recover the original kurtosis. It can be explained by the fact that CLEAN reconstructs the string map by a sum of Dirac spikes and the kurtoses are thus higher than for BP and SBP which reconstruct the string map by a sum of steerable wavelets. Whatever

the string tension and the Fourier coverage, the kurtoses of SBP are the closest of the original one. The errors bars of SBP are much smaller than for BP which shows the stability of SBP.

Finally, we should also outline the fact that no value of the SNR, correlation coefficient, and kurtosis are yet available for CLEAN at $1.0e-07$. This is due to the fact CLEAN needs 1 day to converge to solution at this string tension whereas only 7 hours are sufficient for SBP.

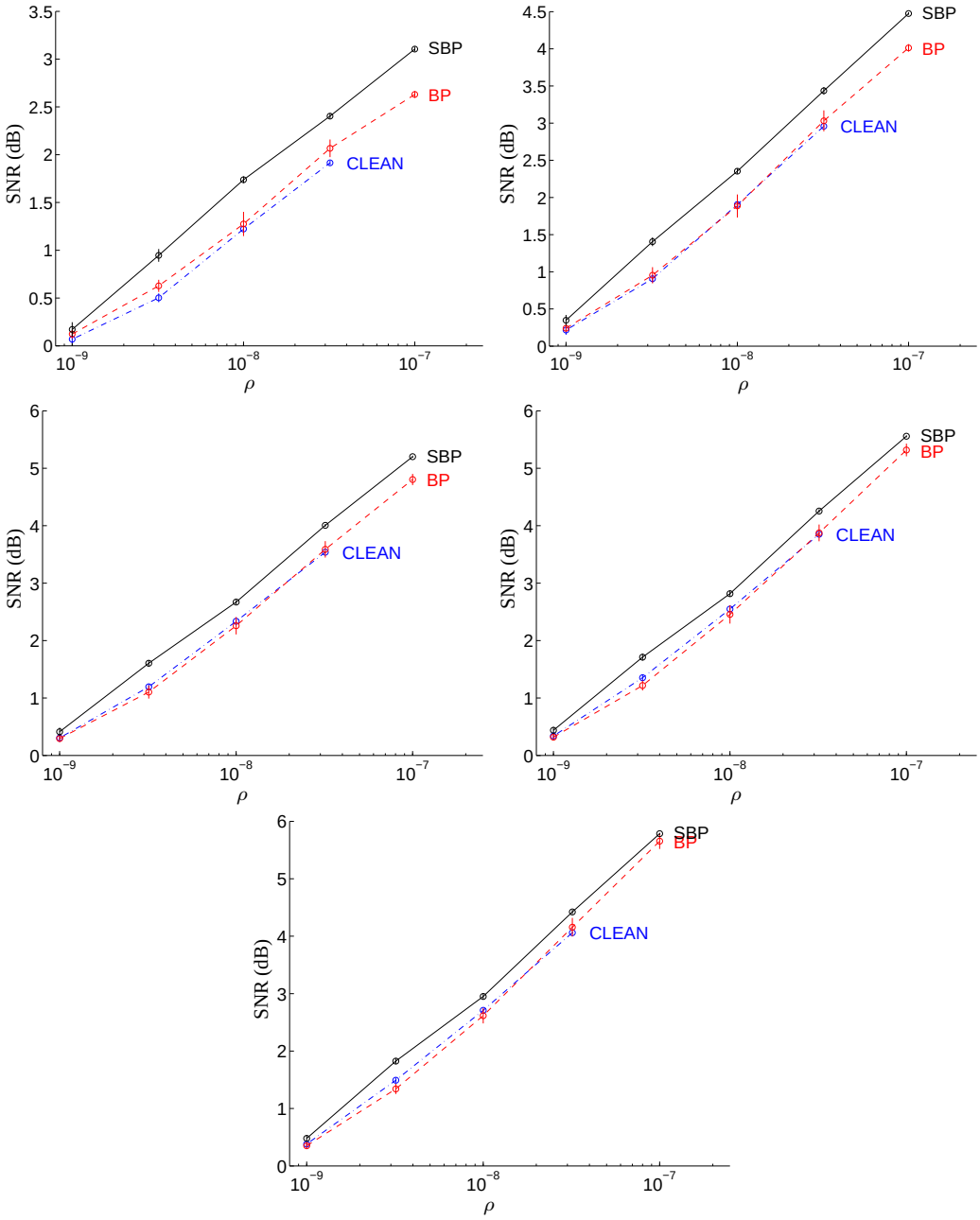

Figure 10: Graph of the mean signal-to-noise ratio in decidels (dB) of the magnitude of gradient at different string tension in the range $[1.0e-09, 1.0e-07]$ for different Fourier coverages: 5% (top left panel), 10% (top right panel), 15% (middle left panel), 20% (middle right panel), 25% (bottom panel). The vertical lines on the curves represent the 1σ error bars estimated over 30 simulations.

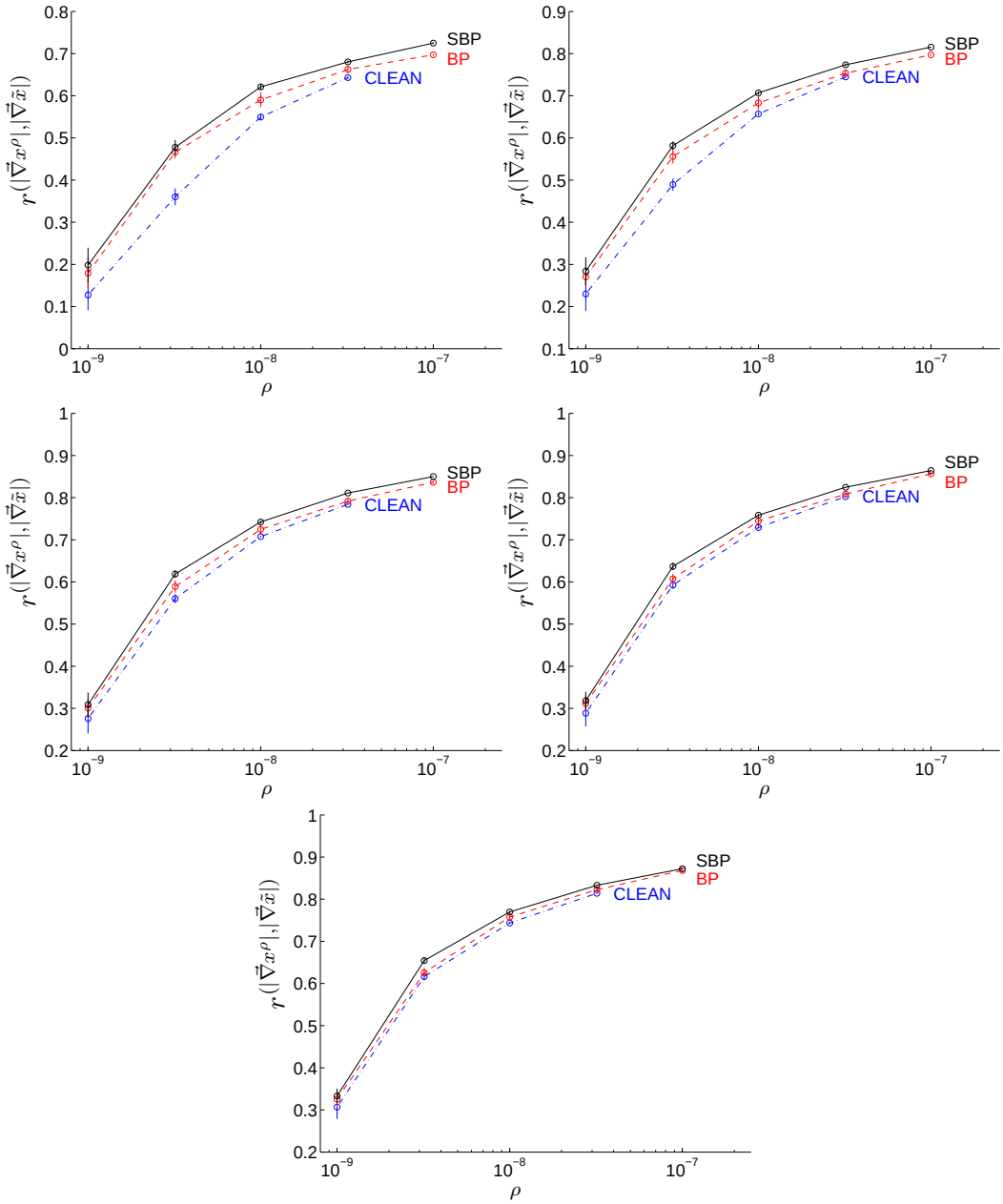

Figure 11: Graph of the mean correlation coefficient of the magnitude of gradient at different string tension in the range $[1.0e-09, 1.0e-07]$ for different Fourier coverages: 5% (top left panel), 10% (top right panel), 15% (middle left panel), 20% (middle right panel), 25% (bottom panel). The vertical lines on the curves represent the 1σ error bars estimated over 30 simulations.

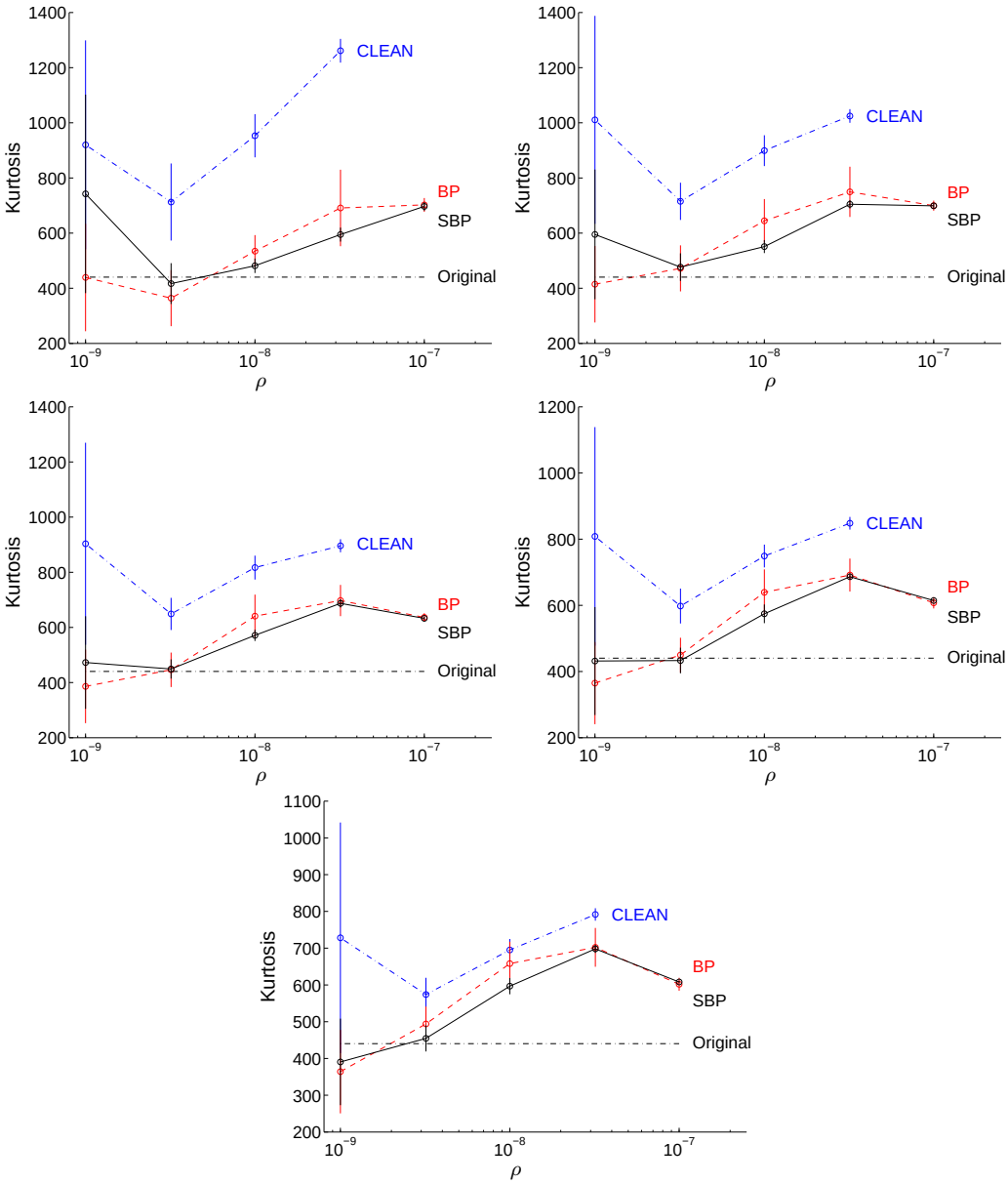

Figure 12: Graph of the mean kurtosis of the magnitude of gradient at different string tension in the range $[1.0e-09, 1.0e-07]$ for different Fourier coverages: 5% (top left panel), 10% (top right panel), 15% (middle left panel), 20% (middle right panel), 25% (bottom panel). The vertical lines on the curves represent the 1σ error bars estimated over 30 simulations.

7 Conclusion and perspectives

A radio-interferometer probes an astrophysical signal through much smaller Fourier measurements than required by the Nyquist-Shannon theorem. The number of measurements being far smaller than the number of unknowns, an ill-posed inverse problem is defined to reconstruct the original signal. If this signal is compressible in a certain basis, the theory of compressed sensing shows that these measurements may be sufficient to have an accurate reconstruction. In this context, the sensing matrix must respect the restricted isometry property and a very accurate reconstruction may be obtained by solving the basis pursuit denoise (BPDN) problem. This problem consists in the minimization of the ℓ_1 norm of the signal to recover under a constraint on the ℓ_2 norm of the residual noise. The string signal being compressible in a steerable wavelet basis ([9]), a first possibility to reconstruct it is to solve this BPDN problem.

However, this method does not use all the prior information available on the string signal. We might want to use the fact that it is well modelled by GGD's in the steerable wavelet basis. Here we used the flexibility of the compressed sensing to introduce this prior information by changing the ℓ_1 norm of the BPDN problem by a "s norm" directly linked to the GGD's. We called this method statistical basis pursuit (SBP).

The third and last method that we tested is a very common algorithm used in radio-astronomy which does not use explicitly any prior information. This algorithm is called CLEAN. We implemented it to assess the performance of BPDN and SBP compared to a standard algorithm.

The simulations show that the performance of BPDN is quite similar to that of CLEAN in term of SNR, correlation coefficient and kurtosis. However, SBP is always better than the two others whatever the evaluation criterium considered. As already emphasized in [15], this work shows that the introduction of prior information in the reconstruction techniques improves the quality of the recovered maps.

Another important information is that it is possible to recover a string signal at string tensions much below the current experimental limit of $2.0\text{e}-07$ when we only consider the primary anisotropies of the CMB. This result leaves strong hope for string detection from current and forthcoming interferometric data.

In this perspective, further evolution of the algorithms developed may be envisaged. Firstly, the convergence of the SBP algorithm was found to be sensitive to the way the regularization parameter is set in the reweighted algorithm. A smaller value of this parameter at 0.1% of the variance of the GGD was tried. It allows us to take the right value for the shape parameter of the largest scale. It increases the performance a lot ($+1$dB compared to CLEAN at a string tension of $3.2\text{e}-08$) but the time of computation quadruples.

Secondly, some other techniques have also to be tested and compared to the results presented in this document. For example, the gradient of the string signal is compressible and well modelled by a generalized Gaussian distribution with a shape parameter of 0.4. We can thus think of minimizing the TV norm of the signal or even the TV norm to the power 0.4. Some preliminary results also show that minimizing the ℓ_0 norm of the signal in real space gives good results. This solution has to be tested.

The secondary anisotropies of CMB and the introduction of instrumental noise will also have to be considered for a complete study of the performance in realistic conditions.

Finally, in regard of the forthcoming radio-telescopes with wide field of views such as the future Square Kilometer Array (SKA), we may think to extend these methods developed on the plane to the sphere ([11]).

8 Acknowledgements

I would like to thank Dr. Yves Wiaux, who supervised this work, for his guidance and assistance throughout the project. I also wish to thank Dr. Laurent Jacques for answering my questions about the theory of compressed sensing, and Dr. David Kenric Hammond for providing me the code which estimates the string tension as well as answering my questions about the wavelet decomposition.

References

[1] E. J. Candès. Compressive sampling. *Int. Congress of Mathematics, Madrid, Spain*, 2006.

[2] E. J. Candès, J. K. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Info. Theory*, 52(2):489–509, February 2006.

[3] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[4] E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted l1 minimization, 2007, preprint (arXiv:0711.1612).

[5] T. J. Cornwell. Multi-Scale CLEAN deconvolution of radio synthesis images. *IEEE Special Issue on Signal Processing*, submitted, 2008, preprint (arXiv:0806.2228).

[6] M. E. Davies and R. Gribonval. Restricted Isometry Constants where ℓ_p sparse recovery can fail for $0 < p \leq 1$. Research report, 2008, (http://www.dsp.ece.rice.edu/cs/).

[7] D. L. Donoho and Y. Tsaig. Compressed sensing. *IEEE Trans. Inform. Theory*, 52:1289–1306, 2006.

[8] A. A. Fraisse, C. Ringeval, D. N. Spergel, and F. R. Bouchet. Small-angle cmb temperature anisotropies induced by cosmic strings. *Physical Review D*, 78:043535, 2008.

[9] D. K. Hammond, Y. Wiaux, and P. Vandergheynst. Wavelet domain Bayesian denoising of string signal in the cosmic microwave background. *Mon. Not. Roy. Astron. Soc.*, submitted, 2008, preprint (arXiv:0811.1267 [astro-ph]).

[10] J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *Astronomy and Astrophysics Supplement*, 15:417–+, June 1974.

[11] J. D. McEwen and A. M. M. Scaife. Simulating full-sky interferometric observations. *Mon. Not. Roy. Astron. Soc.*, 389:1163, 2008.

[12] G. Rocha, M. P. Hobson, S. Smith, P. Ferreira, and A. Challinor. Simulation of non-Gaussian CMB maps. *Mon. Not. Roy. Astron. Soc.*, 357:1–11, 2005.

[13] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *International Conference on Image Processing*, volume 3, pages 444–447, 23-26 Oct. 1995, Washington, DC, USA, 1995.

[14] Martin White, John E. Carlstrom, Mark Dragovan, and William L. Holzapfel. Interferometric observation of cosmic microwave background anisotropies. *The Astrophysical Journal*, 514:12, 1999.

[15] Y. Wiaux, L. Jacques, G. Puy, A. M. M. Scaife, and P. Vandergheynst. Compressed sensing imaging techniques for radio interferometry. *Mon. Not. Roy. Astron. Soc.*, submitted, 2008, preprint (arXiv:0812.4933 [astro-ph]).

A Simulation of Gaussian CMB maps

A method for simulation of statistically-isotropic non-Gaussian CMB maps was developped by Rocha et al. and is presented in the article [12]. In this section, we only focus on the flat-sky approximation (article [12], Appendix C) applied for the simulation of Gaussian CMB maps.

Let $s_p = s(x_p), p \in [0, N_{pix}]$ be a pixelised map of Gaussian white noise. Each pixel of this map are drawn from the standard normal distribution (with a mean of zero and a variance of one). The Fourier transform $\hat{s}$ can be approximated as follows in the discrete case:

$$\hat{s}(l) = \int d^2x s(x)\, e^{-il\cdot x} \approx \Omega_{pix} \sum_p s(x_p)\, e^{-il\cdot x_p}, \tag{45}$$

where Ω_{pix} is the pixel-area. The power spectrum of these coefficients is given by:

$$\langle \hat{s}(l)\hat{s}^*(l')\rangle = \Omega_{pix}^2 N_{pix} \delta_{ll'}, \tag{46}$$

In the continuum limit, we have:

$$\delta_{ll'} = \frac{1}{N_{pix}} \sum_p e^{i(l-l')\cdot x_p} \approx \frac{1}{N_{pix}\Omega_{pix}} \int d^2x e^{i(l-l')\cdot x} = \frac{(2\pi)^2}{N_{pix}\Omega_{pix}} \delta(l-l'), \tag{47}$$

and equation (46) becomes thus

$$\langle \hat{s}(l)\hat{s}^*(l')\rangle = (2\pi)^2 \Omega_{pix}^2 \delta(l-l'). \tag{48}$$

In order to obtain a final map with a particular power spectrum $(2\pi)^2\mathcal{C}_l$, we rescale all the Fourier coefficients as follows:

$$\hat{\bar{s}}(l) = \sqrt{\frac{\mathcal{C}_l}{\Omega_{pix}}}\hat{s}(l), \tag{49}$$

so that, in the continuum limit:

$$\langle \hat{\bar{s}}(l)\hat{\bar{s}}^*(l')\rangle = (2\pi)^2 \mathcal{C}_l \delta(l-l'). \tag{50}$$

The final Gaussian CMB map $\bar{s}_p$ with the given power spectrum is obtain by inverting the Fourier transform:

$$\bar{s}_p = \frac{2\pi}{N_{pix}\Omega_{pix}} \sum_l \hat{\bar{s}}(l)\, e^{il\cdot x_p} \approx \int \frac{d^2x}{4\pi^2} \hat{\bar{s}}(l)\, e^{-il\cdot x}. \tag{51}$$