Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **469,** 2982–2996 (2017)
Advance Access publication 2017 May 1

doi:10.1093/mnras/stx1029

A 3D model for carbon monoxide molecular line emission as a potential cosmic microwave background polarization contaminant

G. Puglisi,[1,2]★ G. Fabbian[1,2,3]★ and C. Baccigalupi[1,2]

[1]*SISSA – International School for Advanced Studies, Via Bonomea 265, I-34136 Trieste, Italy*
[2]*INFN – National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*
[3]*Institut d'Astrophysique Spatiale, CNRS (UMR 8617), Univ. Paris-Sud, Université Paris-Saclay, bât. 121, F-91405 Orsay, France*

Accepted 2017 April 26. Received 2017 April 7; in original form 2017 February 12

ABSTRACT

We present a model for simulating carbon monoxide (CO) rotational line emission in molecular clouds, taking account of their 3D spatial distribution in galaxies with different geometrical properties. The model implemented is based on recent results in the literature and has been designed for performing Monte Carlo (MC) simulations of this emission. We compare the simulations produced with this model and calibrate them, both on the map and the power spectrum levels, using the second release of data from the *Planck* satellite for the Galactic plane, where the signal-to-noise ratio is highest. We use the calibrated model to extrapolate the CO power spectrum at low Galactic latitudes where no high sensitivity observations are available yet. We then forecast the level of unresolved polarized emission from CO molecular clouds which could contaminate the power spectrum of cosmic microwave background polarization B modes away from the Galactic plane. Assuming realistic levels of the polarization fraction, we show that the level of contamination is equivalent to a cosmological signal with $r \lesssim 0.02$. The MC MOlecular Line Emission (MCMOLE3D) PYTHON package, which implements this model, is being made publicly available.

Key words: ISM: lines and bands – ISM: magnetic fields – ISM: molecules – cosmic background radiation – diffuse radiation – cosmology: observations.

1 INTRODUCTION

The carbon monoxide (CO) molecule is one of the most interesting molecules present in molecular clouds within our Galaxy. Although the most abundant molecule in Galactic molecular clouds is molecular hydrogen (H_2), it is inconvenient to use the emission from that as a tracer since it is difficult to detect because of having a low dipole moment and so being a very inefficient radiator. We therefore need to resort to alternative techniques for tracing molecular clouds using rotational or vibrational transitions of other molecules such as CO. Observations of CO emission are commonly used to infer the mass of molecular gas in the Milky Way by assuming a linear proportionality between the CO and H_2 densities via the CO-to-H_2 conversion factor, $X_{\rm CO}$. A commonly accepted value for $X_{\rm CO}$ is 2×10^{20} molecules $\cdot$ cm^{-2} (K km s^{-1})$^{-1}$, although this could vary with position in the Galactic plane, particularly in the outer Galaxy (Balser et al. 2011).

The most intense CO rotational transition lines are the $J = 1 \rightarrow 0, 2 \rightarrow 1$ and $3 \rightarrow 2$ transitions at sub-millimetre (sub-mm) wavelengths (115, 230 and 345 GHz, respectively). These can usually be observed in optically thick and thermalized regions of the interstellar medium. Traditionally, the observations of standard ^{12}CO emission are complemented by measurements of ^{13}CO lines. Being less abundant (few per cent), this isotopologue can be exploited for inferring the dust extinction in nearby clouds and hence providing a better constraint for measuring the H_2 abundance (Bally et al. 1987; Jackson et al. 2006). However, there is growing evidence that ^{13}CO regions could be associated with colder and denser environments, whereas ^{12}CO emission originates from a diffuse component of molecular gas (Roman-Duval et al. 2016).

The spatial distribution of the CO line emission reaches a peak in the inner Galaxy and is mostly concentrated close to or within the spiral arms, in a well-defined ring, the so-called *molecular ring* between about 4–7 kpc from the Galactic Centre. This property is not unique to the Milky Way but is quite common in barred spiral galaxies (see Regan et al. 2002 for further references). The emission in the direction orthogonal to the Galactic plane is confined within a Gaussian slab with roughly 90 pc full width half-maximum (FWHM) in the inner Galaxy getting broader towards the outer Galactic regions, reaching an FWHM of several hundred parsecs outside the solar circle. In the centre of the Galaxy, we can also identify a very dense CO emission zone, rich in neutral gas and individual stars, stretching out to about 700 light-years from the centre and known as the *Central Molecular Zone*.

★ E-mail: giuspugl@sissa.it (GP); gfabbian@ias.u-psud.fr (GF)

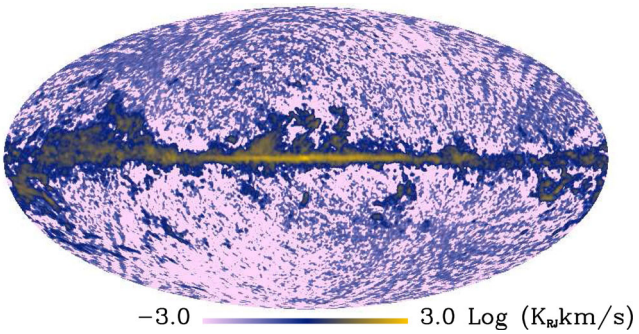

Figure 1. *Planck* CO 1 − 0 map (Planck Collaboration X 2016b). Note the predominance of instrumental noise in regions far from the Galactic plane.

Since the 1970s, many CO surveys of the Galactic plane have been carried out with ground-based telescopes, leading to accurate catalogues of molecular clouds (Dame, Hartmann & Thaddeus 2001; Mizuno & Fukui 2004). Usually these surveys have observed a strip of $|b| \lesssim 5$ deg around the Galactic plane. At higher Galactic latitudes (HGL, $|b| > 30$ deg), the low opacity regions of both gas and dust, together with a relatively low stellar background which is useful for spotting extinction regions, complicate the observation of CO lines making this very challenging. In fact, only ≈100 clouds have been detected so far in these regions.

The *Planck* satellite team recently released CO emission maps of the lowest rotational lines, $J = 1 - 0, 2 - 1$ and $3 - 2$ observed in the 100, 217 and 353 GHz frequency channels of the High Frequency Instrument (HFI; Planck Collaboration XIII 2014b; Planck Collaboration X 2016b). These were sensitive enough to map the CO emission even though the widths of these lines are orders of magnitude narrower than the bandwidth of the *Planck* frequency channels. These single-frequency maps have been processed with a dedicated foreground cleaning procedure so as to isolate this emission. The *Planck* maps were found to be broadly consistent with the data from other CO surveys (Dame et al. 2001; Heyer & Dame 2015), although they might still be affected by residual astrophysical emissions and instrumental systematics. In Fig. 1, we show the so-called Type 1 *Planck* map of the CO $J : 1 - 0$ line (Planck Collaboration XIII 2014b)[1] which will be used in the following.

Many current and future cosmic microwave background (CMB) polarization experiments[2] are designed to exploit the faint B-mode signal of CMB polarization as a cosmological probe, in particular to constrain the physics of large-scale structure formation or the inflationary mechanism in the early universe (Seljak & Zaldarriaga 1997; Hu & White 1997). One of the main challenges in the way of achieving these goals is the contamination of the primordial CMB signal by diffuse Galactic emission. In this respect, the synchrotron and thermal dust emission are known to be potentially the most dangerous contaminants, because they are intrinsically polarized. In fact, several analyses conducted on *Planck* and *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data from intermediate Galactic latitude and HGL at high (Planck Collaboration XXX 2016a) and low frequencies (Krachmalnicoff et al. 2015; Planck Collaboration XXV 2016c) showed that these emissions are dangerous at all microwave frequencies and locations on the sky (even if far from the galactic plane), confirming early studies using the *WMAP* satellite (Gold et al. 2011; Page et al. 2007; Baccigalupi 2003).

Appropriate observations and theoretical investigations and modelling of polarized foreground emission for all emissions at sub-mm frequencies are therefore crucial for the success of future experiments. As these will observe at frequencies overlapping with the CO lines, unresolved CO line emission could significantly contaminate these measurements as well.

CO lines are in fact expected to be polarized at the percent level or below (Goldreich & Kylafis 1981) because of interaction of the magnetic moment of the molecule with the Galactic magnetic field. This causes the so-called *Zeeman splitting* of the rotational quantum levels J into the magnetic sub-levels M which are intrinsically polarized. Moreover, if molecular clouds are somehow anisotropic (e.g. when in the presence of expanding or collapsing envelopes in star formation regions) or are asymmetric, population imbalances of the M levels can arise. This leads to different line intensities depending on the directions (parallel or perpendicular to the magnetic field) and to a net linearly polarized emission. Greaves et al. (1999) detected polarization in five star-forming regions near to the Galactic Centre while observing the CO lines $J = 2 - 1, 3 - 2$ and the $J = 2 - 1$ of the isotopologue ^{13}CO. The degree of polarization ranged from 0.5 to 2.5 per cent. Moreover, the deduced magnetic field direction was found to be consistent with previous measurements coming from dust polarimetry, showing that the polarized CO emission could become a sensitive tracer of small-scale Galactic magnetic fields.

The goal of this paper is to propose a statistical 3D parametric model of CO molecular cloud emission, in order to forecast the contamination of CMB signal by this, including in the polarization. Being able to perform statistical simulation of this emission is crucial for assessing the impact of foreground residual uncertainties on cosmological constraints coming from the CMB. In addition, the capability of modelling the Galactic foreground emission in its full complexity taking into account line-of-sight effects is becoming necessary in light of the latest experimental results and the expected level of sensitivity for the future experiments (Tassis & Pavlidou 2015; Planck Collaboration L 2017). In Section 2, we present the assumptions made for building the model and the simulation pipeline for its implementation. In Section 3, we describe the methodology for calibrating the CO simulations to match *Planck* observations. In Section 3.4, we show how the parameters describing molecular cloud distribution shape the angular power spectrum of CO emission. Finally, in Section 4, we forecast the expected level of polarized CO contaminations for the CMB B modes at HGL using our calibrated simulation of Section 3 to infer statistically the emission at HGL, where current observations are less reliable.

2 BUILDING A STATISTICAL 3D CO EMISSION MODEL

In order to build an accurate description of CO emission in the Galaxy, we collected the most up to date astrophysical data present in the literature concerning the distribution of molecular gas as a function of the Galactic radius (R) and the vertical scale of the Galactic disc (z) as well as of the molecular size and the mass function. The model has been implemented in a PYTHON package named Monte Carlo MOlecular Line Emission (MCMOLE3D),[3] which

[1] http://pla.esac.esa.int/pla

[2] For a complete list of the operating and planned probes, see e.g. lambda.gfsc.nasa.gov

[3] https://github.com/giuspugl/MCMole3D

is being made publicly available, and we present details of it in this section.[4] The model builds on and extends the method proposed by Ellsworth-Bowers et al. (2015) who conducted a series of analyses distributing statistically a relative large number of molecular cloud objects according to the axisymmetric distribution of H_2 observed in the Galaxy (Wolfire et al. 2003).

2.1 CO cloud spatial distribution

As mentioned in the Introduction, the CO emission is mostly concentrated around the molecular ring. We have considered and implemented two different spatial distributions of the molecular clouds: an axisymmetric *ring-shaped* one and one with four spiral arms, as shown in Figs 3(b) and (a), respectively. The first is a simplified model and is parametrized by $R_{\rm ring}$ and $\sigma_{\rm ring}$ which are the radius and the width of the molecular ring, respectively. On the other hand, the spiral arm distribution is in principle closer to the symmetry of our Galaxy and is therefore more directly related to observations. The distribution is described by two more parameters than for the axisymmetric case: the arm width and the spiral arm pitch angle. For the analysis conducted in the following sections, we fixed the value of the pitch angle to be $i \sim -13$ deg following the latest measurements of Davis et al. (2012) and Bobylev & Bajkova (2013) and fixed the arm half-width to be 340 pc (Vallée 2014).

Bronfman et al. (1988) found that the vertical profile of the CO emissivity can be optimally described by a Gaussian function of z centred on z_0 and having a half-width $z_{1/2}$ from the Galactic plane at $z = 0$. Both of the parameters z_0 and $z_{1/2}$ are in general functions of the Galactic radius R (see Heyer & Dame (2015) for recent measurements). Since we are interested in the overall distribution of molecular clouds mainly in regions close to the Galactic plane, where data are more reliable, we adopted this parametrization but neglected the effects of the mid-plane displacement z_0 and set it to a constant value $z_0 = 0$, following Delabrouille et al. (2013). The vertical profile is then parametrized just by $z_{1/2}$ and mimics the increase of the vertical thickness scatter that is observed when moving from the inner Galaxy towards the outer regions:

$$z_{1/2}(R) \propto \sigma_z(R) = \sigma_{z,0} \cosh\left(\frac{R}{h_R}\right), \tag{1}$$

where $\sigma_{z,0} = 0.1$ and $h_R = 9$ kpc corresponds to the radius where the vertical thickness starts increasing. The half-width $z_{1/2}$ is related to σ_z through the usual relation $z_{1/2} = \sqrt{2\ln 2}\sigma_z$. The final vertical profile is then:

$$z(R) = \frac{1}{\sqrt{2\pi}\sigma_z(R)} \exp\left[-\left(\frac{z}{\sqrt{2}\sigma_z(R)}\right)^2\right]. \tag{2}$$

2.2 CO cloud emission

The key ingredients for modelling the molecular cloud emission are the dimension of the cloud and its typical emissivity. We assume an exponential CO emissivity profile which is a function of the Galactic radius following (Heyer & Dame 2015; Roman-Duval et al. 2016):

$$\epsilon_0(R) = \epsilon_c \exp\left(R/R_{\rm em}\right), \tag{3}$$

where ϵ_c is the typical emissivity of a particular CO line observed towards the centre of the Galaxy and $R_{\rm em}$ the scalelength over which the emissivity profile changes. Clouds observed in the outer Galaxy are in fact dimmer.

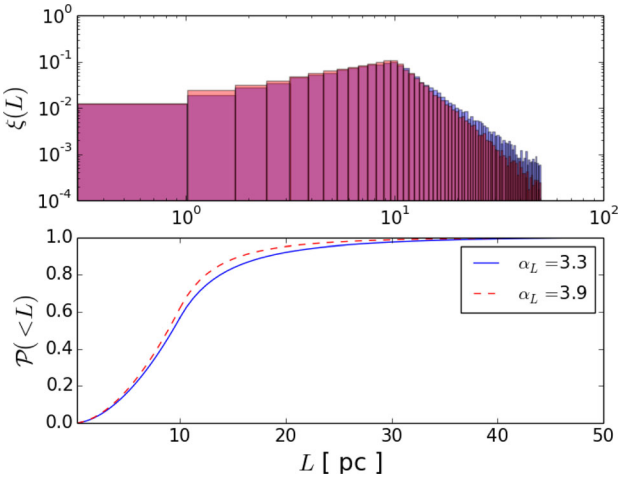

Figure 2. (Top) Histograms of dN/dL computed by assigning the size of each cloud with the probability function (bottom). The two spectral indices $\alpha_L \approx 3.3(3.9)$ refer, respectively, to clouds in the inner (outer) Galaxy.

We then assume the distribution of cloud size $\xi(L)$ defined by their typical size scale, L_0, the range of sizes $[L_{\rm min}, L_{\rm max}]$ and two power laws with spectral indices (Roman-Duval et al. 2010)

$$\xi(L) = \frac{dn}{dL} \propto \begin{cases} L^{0.8} & \text{if } L_{\rm min} < L < L_0, \\ L^{-\alpha_L} & \text{if } L_0 < L < L_{\rm max}, \end{cases} \tag{4}$$

with $\alpha_L = 3.3$, 3.9 for clouds inside or outside the solar circle, respectively. From the cloud size function $\xi(L)$, we derive the corresponding probability $\mathcal{P}(L)$ of having clouds with sizes smaller than L:

$$\mathcal{P}(<L) = \int_{L_{\rm min}}^{L} dL' \xi(L'). \tag{5}$$

The probability functions for different choices of the spectral index α_L are shown in Fig. 2. We then inverted equation (5) to get the cloud size associated with a given probability $L(p)$. The cloud sizes are drawn from a uniform distribution in [0, 1]. The histograms of the sizes generated following this probability function are shown in the top panel of Fig. 2 and are peaked around the most typical size L_0. In the analysis presented in the following, L_0 is considered as a free parameter.

Finally, we assume a spherical shape for each of the simulated molecular clouds once they are projected on the sky. However, we implemented different emissivity profiles that are function of the distance from the cloud centre, such as *Gaussian* or *cosine* profiles. These are particularly useful because, by construction, they give zero emissivity at the boundaries[5] and the maximum of the emissivity in the centre of the projected cloud on the sky. This not only mimics a decrease of the emission towards the outer regions of the cloud, where the density decreases, but also allows to minimize numerical artefacts when computing the angular power spectrum of the simulated maps (see Section 3). An abrupt top-hat transition at the boundary of each cloud would in fact cause ringing effects that could bias the estimate of the power spectrum.

[4] In the following, we will refer to this model as the MCMOLE3D model for the sake of clarity.

[5] For the Gaussian profile, we set σ in order to have the cloud boundaries at 6σ, i.e. where the Gaussian function is zero to numerical precision.

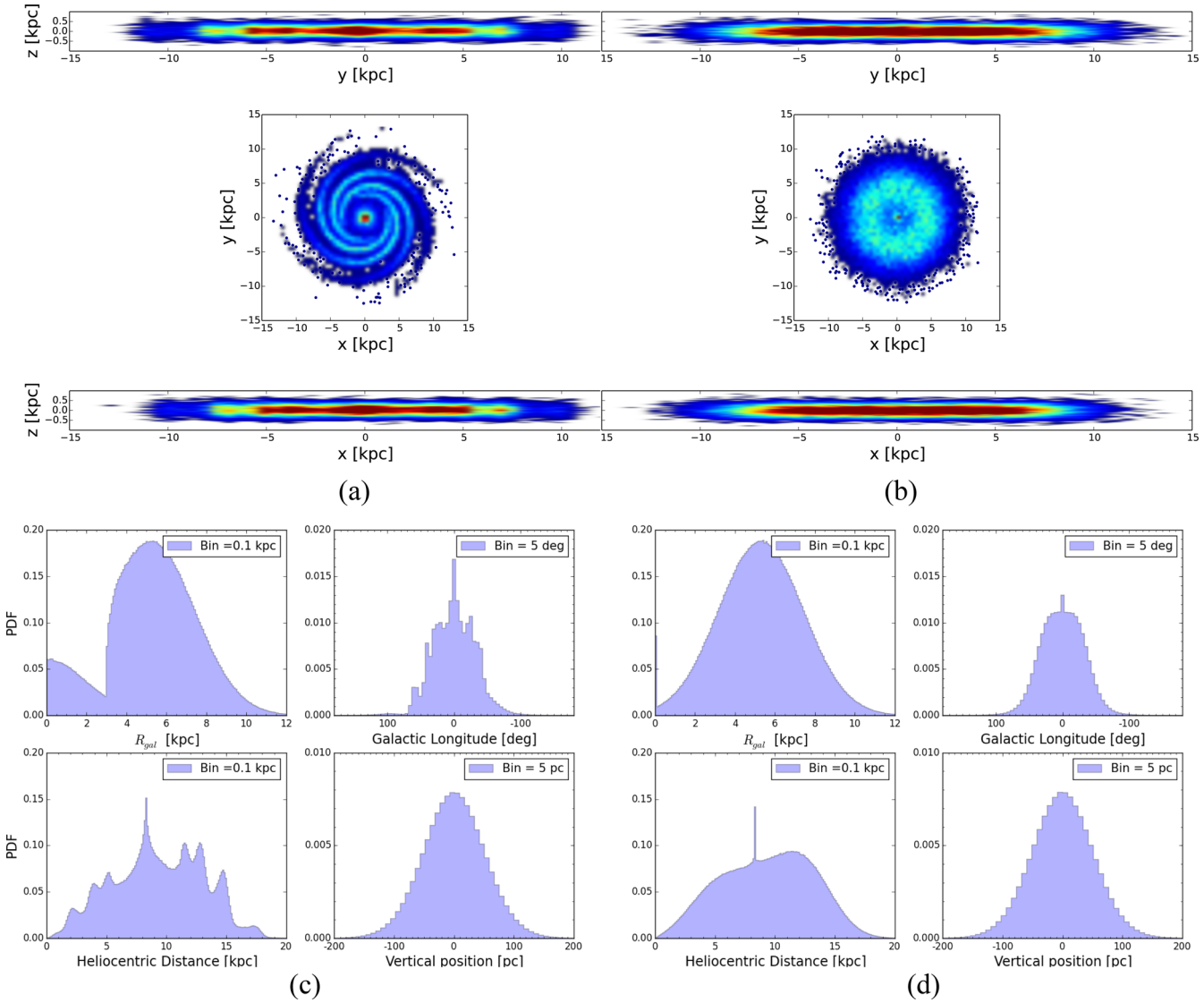

Figure 3. *Top panels*: density contour plots of an MC galaxy population with 40 000 objects distributed following the (a) LOGSPIRAL and (b) AXISYMMETRIC distributions. *Bottom panels*: probability density functions of 100 MC realizations of 40 000 molecular clouds following the (c) LOGSPIRAL and (d) AXISYMMETRIC geometry. The latter case is consistent with results in Ellsworth-Bowers et al. (2015).

2.3 Simulation procedure

The model outlined in the previous section enables statistical simulations of CO emission in our Galaxy to be performed for a given set of free parameters $\Theta^{\rm CO}$ that can be set by the user:

$$\Theta^{\rm CO} = \{N_{\rm clouds},\ \epsilon_c,\ R_{\rm em},\ R_{\rm ring},\ \sigma_{\rm ring},\ \times sigma_{z,0},\ h_R,\ L_{\rm min}, L_{\rm max}, L_0\}.$$

The values chosen for our analysis are listed in Table 1. For each realization of the model, we distribute by default 40 000 clouds within our Galaxy. This number is adopted for consistency with observations when observational cuts are applied (for further details see Ellsworth-Bowers et al. 2015). The product of each simulation is a map, similar to the one in Fig. 4, in the Hierarchical Equal Area Latitude Pixelization (HEALPIX, Górski et al. 2005)[6] pixelization scheme including all the simulated clouds as seen by an observer

Table 1. List of parameters used in MCMOLE3D simulations.

MCMOLE3D Default parameters	
$N_{\rm clouds}$	40 000
$R_{\rm ring}$ (kpc)	5.3
$L_{\rm min}$ (pc)	0.3
$L_{\rm max}$ (pc)	60
$\sigma_{z,0}$ (pc)	100
h_R (kpc)	9
$R_{\rm bar}{}^a$ (kpc)	3
i^a (deg)	−12
ϵ_c ($\rm K_{RJ}\,km\,s^{-1}$)	240
$R_{\rm em}$ (kpc)	6.6
L_0 (pc)	[5,50] Default: 20
$\sigma_{\rm ring}$ (kpc)	[1,5] Default: 2.5

[a]Only for LOGSPIRAL.

[6] http://healpix.sourceforge.net

placed in the Solar system. This map can be smoothed to match the resolution of a specific experiment and/or convolved with a realistic frequency bandwidth. When we compare with the *Planck* maps described in Section 3, we convolve the simulated maps to the beam resolution of the 100 GHz channel (∼10 arcmin).

The procedure implemented for each realization is the following:

(i) assign the $(R_{\rm gal}, \phi, z)$ Galactocentric positions. In particular:

(a) $R_{\rm gal}$ is extracted from a Gaussian distribution defined by the $R_{\rm ring}$ and $\sigma_{\rm ring}$ parameters. However, the $\sigma_{\rm ring}$ is large enough to give non-zero probability at $R_{\rm gal} \leq 0$. All of the negative values of $R_{\rm gal}$ are either automatically set to $R_{\rm gal} = 0$ (axisymmetric case), or re-computed extracting new positive values from a normal distribution centred at $R = 0$ and with the rms given by the scale of the Galactic bar (spiral-arm case). This choice allows us to circumvent not only the issue of negative values of $R_{\rm gal}$ due to a Gaussian distribution, but also to produce the high emissivity of the Central Molecular Zone (see Ellsworth-Bowers et al. 2015 for a similar approach).

(b) the z-coordinate is drawn randomly from the distribution in equation (2).

(c) the azimuth angle ϕ is computed from a uniform distribution ranging over $[0, 2\pi)$ in the case of the axial symmetry. Conversely, in the case of spiral arms, ϕ follows the *logarithmic spiral* polar equation

$$\phi(R) = A \log R + B,$$

where $A = (\tan i)^{-1}$ and $B = -\log R_{\rm bar}$ are, respectively, functions of the mean pitch angle and the starting radius of the spiral arm. In our case, we set $i = -12$ deg, $R_{\rm bar} = 3$ kpc;

(ii) assign cloud sizes given the probability function $\mathcal{P}(L)$ (equation 5);

(iii) assign emissivities to each cloud from the emissivity profile (see equation 3);

(iv) convert $(R_{\rm gal}, \phi, z)$ positions into the *heliocentric* coordinate frame $(\ell, b, d_\odot)$.

In Fig. 3, we show an example of the 3D distribution of the emission as well as the distribution of the location of the simulated clouds using both of the geometries implemented in the package.

2.4 Simulation results

In Fig. 4, we show two typical realizations of maps of CO emission for the AXISYMMETRIC and LOGSPIRAL geometries prior to any smoothing. As we are interested in the statistical properties of the CO emission, we report a few examples of the angular power spectrum $\mathcal{C}_\ell$ corresponding to different distributions of CO emission in Fig. 5. In the ones shown subsequently the spectra are $\mathcal{D}_\ell$ encoding a normalization factor $\mathcal{D}_\ell = \ell(\ell+1)\mathcal{C}_\ell/2\pi$.

We can observe two main features in the morphology of the power spectrum: a bump around $\ell \sim 100$ and a tail at higher ℓ. We interpret both of these features as the projection of the distribution of clouds from a reference frame off-centred (on the solar circle).

The bump reflects the angular scale (∼1 deg) related to the clouds which have the most likely size, parametrized by the typical size parameter, L_0, and which are close to the observer. On the other hand, the tail at $\ell \gtrsim 600$ (i.e. the arcminute scale) is related to the distant clouds which lie in the diametrically opposite position with respect to the observer. This is the reason why the effect is shifted to smaller angular scales. The L_0 and $\sigma_{\rm ring}$ parameters modify the power spectrum in two different ways. For a given typical size, if the width of the molecular ring zone $\sigma_{\rm ring}$ increases, the peak around $\ell \sim 100$ shifts towards lower multipoles, i.e. larger angular scales, and its amplitude increases proportionally to $\sigma_{\rm ring}$, see for instance the bottom right panel in Fig. 5. This can be interpreted as corresponding to the fact that the larger is $\sigma_{\rm ring}$, the more likely it is to have clouds closer to the observer at the solar circle with a typical size given by L_0. On the other hand, if we choose different values for the size parameter (left-hand panels in Fig. 5), the tail at small angular scales moves downwards and flattens as L_0 increases. Vice versa, if we keep L_0 constant (Fig. 5, bottom right panel), all of the tails have the same amplitude and an ℓ^2 dependency. In fact, if L_0 is small, the angular correlation of the simulated molecular clouds looks very similar to the one of *point sources* which has *Poissonian* behaviour. Conversely, if the typical size increases, the clouds become larger and they behave effectively as a coherent diffuse emission and less as point sources.

Far from the Galactic plane, the shape of the power spectrum is very different. In Fig. 6, we show an example of the average power spectrum of 100 MC realizations of CO emission at HGL, i.e. $|b| > 30$ deg, for both the AXISYMMETRIC and LOGSPIRAL geometries. For this run, we choose the so-called best-fitting values for the L_0 and $\sigma_{\rm ring}$ parameters discussed later in Section 3.4. In addition to the different shape depending on the assumed geometry, one can notice a significant amplitude difference with respect to the power spectrum at low latitudes. Moreover, this is in contrast with the trend observed in the Galactic plane, where the LOGSPIRAL geometry tends to predict a power spectrum of higher amplitude. In both cases, however, the model suppresses the emission in these areas, as shown in Fig. 4. In the LOGSPIRAL case, the probability of finding clouds in regions in between spiral arms is further suppressed and could explain this feature. The emission is dominated by clouds

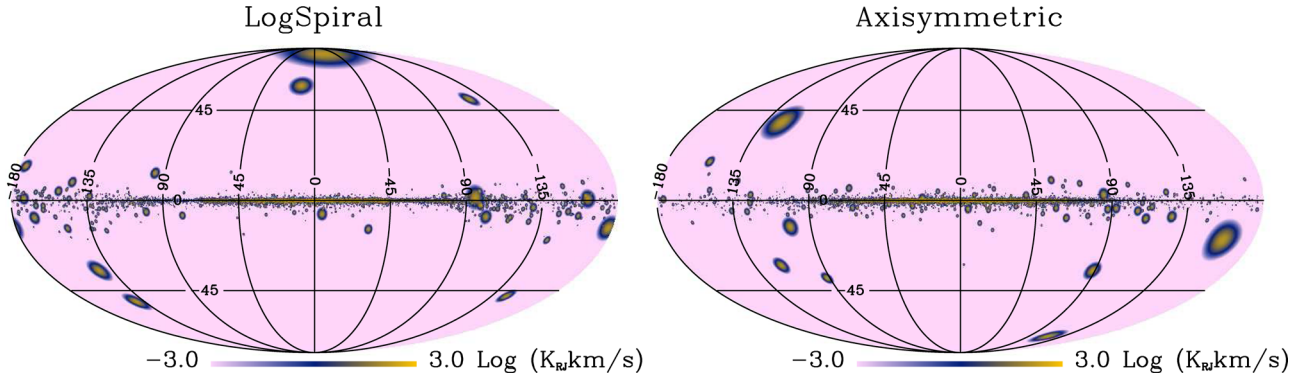

Figure 4. Two realizations of CO maps simulated with MCMOLE3D using the distribution parameters given by the values in Table 1.

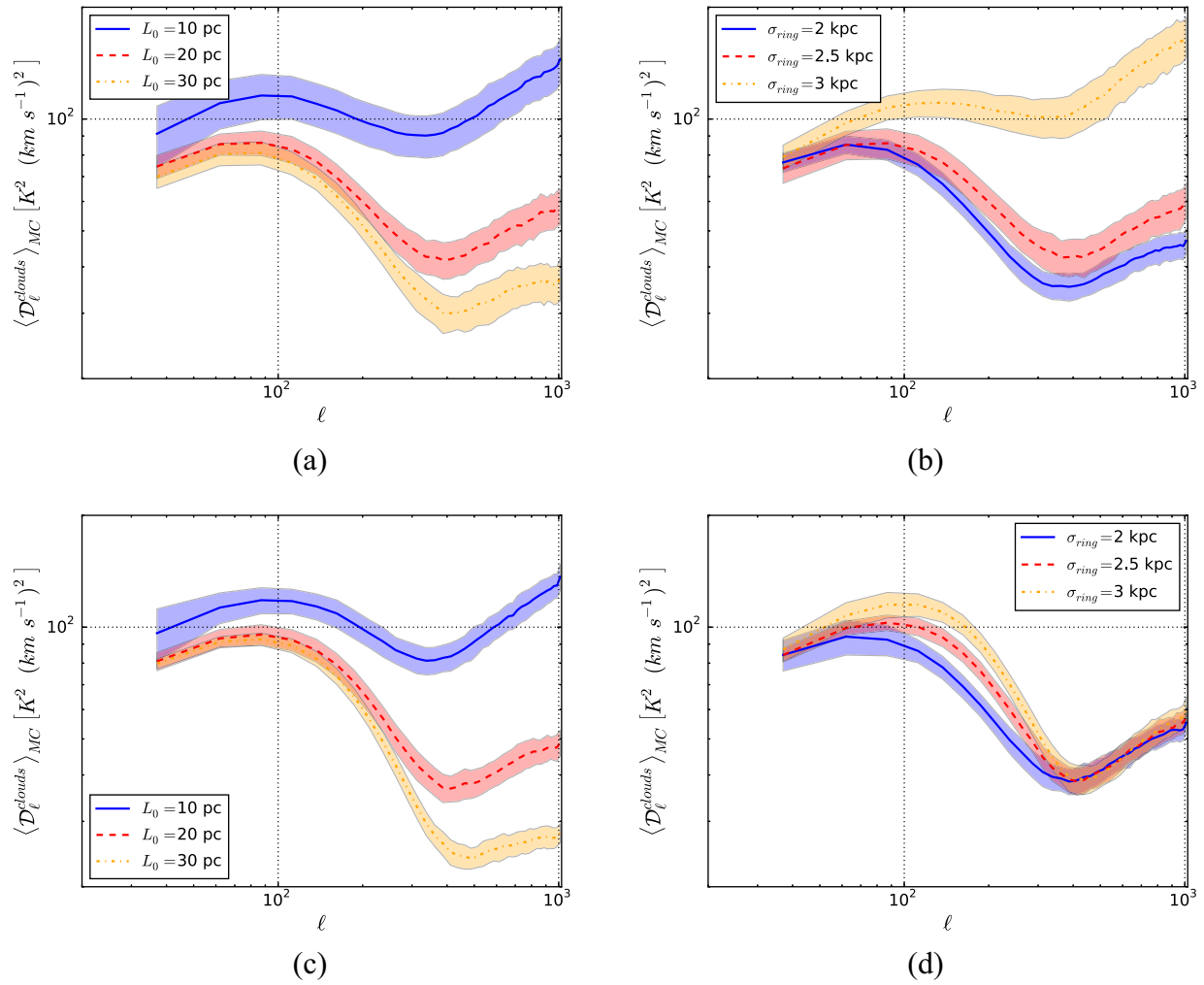

Figure 5. Angular power spectra of CO emission in the Galactic plane computed for 100 MC realizations of the MCMOLE3D model assuming different values of its free parameters. The mean value of the simulation is shown by solid, dashed and dot–dashed lines, while the shaded area represents the measured variance of the realizations. The top row shows the case of an AXISYMMETRIC geometry, while the bottom panel displays results for a LOGSPIRAL geometry. Results obtained by varying the L_0 ($\sigma_{\rm ring}$) parameters are shown on the left (right) column.

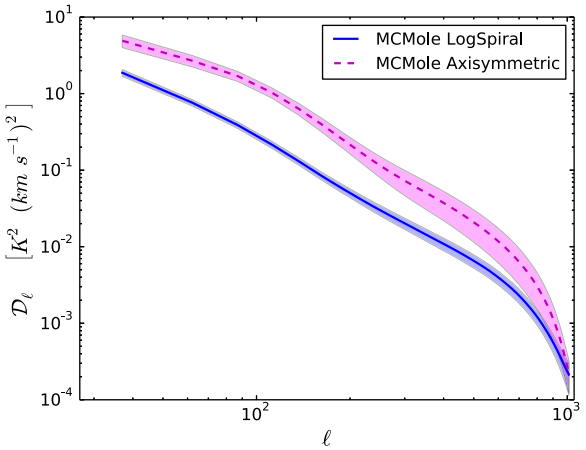

Figure 6. Examples of the power spectra of CO emission at HGL ($|b| > 30$ deg) for AXISYMMETRIC and LOGSPIRAL geometries. For both the geometries, we assumed the best-fitting values of the parameters describing the CO distribution presented in Section 3.4.

relatively close to the observer for both geometries, and so the angular correlation is mostly significant at large angular scales (of the order of a degree or more) and is damped rapidly at small angular scales.

3 COMPARISON WITH *PLANCK* DATA

3.1 Data set

The *Planck* collaboration released three different kinds of CO molecular line emission maps, described in Planck Collaboration XIII (2014b); Planck Collaboration X (2016b). We decided to focus our analysis on the so-called *Type 1* CO maps which have been extracted exploiting differences in the spectral transmission of a given CO emission line in all of the bolometer pairs relative to the same frequency channel. Despite being the noisiest set of maps, *Type 1* are in fact the cleanest maps in terms of contamination coming from other frequency channels and astrophysical emissions. In addition, they have been obtained at the native resolution of the *Planck* frequency channels, and so allow full control of the effective beam window function for each map.

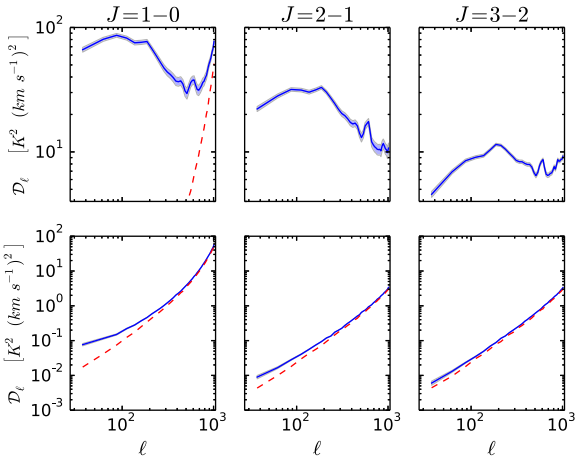

Figure 7. CO $1-0$ angular power spectrum (blue solid) estimated from the *Planck* map at low Galactic latitudes (top) and HGL (bottom). The shaded area shows the error bar due to the sample and noise variance. The expected noise level of the maps in the two regions is shown in (red dashed).

For this study, we considered in particular the CO $1-0$ line, which has been observed in the 100 GHz channel of the HFI instrument. This channel is in fact the most sensitive to the CO emission in terms of signal-to-noise ratio (SNR) and the $1-0$ line is also the one for which we have the most detailed external astrophysical observations. However, the *Planck* frequency bands were designed to observe the CMB and foreground emissions which gently vary with frequency and, thus, they do not have the spectral resolution required to resolve accurately the CO line emission. To be more quantitative, the *Planck* spectral response at 100 GHz is roughly 3 GHz, which corresponds to $\sim$8000 km s^{-1}, i.e. about 8 orders of magnitude larger than the CO rotational line width (which can be easily approximated as a Dirac delta). Therefore, the CO emission observed by *Planck* along each line of sight is integrated over the whole channel frequency band. Further details about the spectral response of the HFI instrument can be found in Planck Collaboration IX (2014a).

3.2 Observed co angular power spectrum

Since one of the goals of this paper is to understand the properties of diffuse CO line emission, we computed the angular power spectrum of the *Type 1* $1-0$ CO map to compare qualitatively the properties of our model with the single realization given by the emission in our Galaxy. We distinguish two regimes of comparison, low Galactic latitudes ($|b| \leq 30$ deg) and HGL ($|b| > 30$ deg). While at low Galactic latitudes, the signal is observed with high sensitivity, at high latitudes it is substantially affected by noise and by the fact that the emission in this region is faint due to its low density with respect to the Galactic disc.

In Fig. 7, we show the angular power spectra of the first three CO rotational line maps observed by *Planck* as well as the expected noise level at both HGL and low Galactic latitudes computed using a pure power spectrum estimator X^{2}PURE (Grain, Tristram & Stompor 2009). This is a pseudo-power spectrum method (Hivon et al. 2002) which corrects the so called E-to-B-modes leakage in the polarization field that arises in the presence of incomplete sky coverage (Smith & Zaldarriaga 2007; Bunn et al. 2003; Lewis, Challinor & Turok 2001). Although this feature is not strictly relevant for the analysis of this section, because we are considering the unpolarized component of the signal, it is important for the forecast presented in Section 4. We estimated the noise as the mean of 100 MC Gaussian simulations based on the diagonal pixel–pixel error covariance included in the *Planck* maps. One may notice how the noise has a level comparable to that of the CO power spectrum at HGL. However, we note that the released *Type 1* maps are obtained from the full mission data from *Planck*, and not from subsets of the data (e.g. using the so-called half-rings or half-mission splits). Thus, it was not possible to test whether the observed flattening of the power spectrum at large angular scale is due to additional noise correlation not modelled by the Gaussian uncorrelated model discussed above. We notice that, if these maps were present, we could have had an estimate of this correlation using the noise given by the difference between the map autospectra and the noise-bias-free signal obtained from the cross-spectra of the maps from data subsets. Since even for the $1-0$ line, the noise becomes dominant on scales $\ell \approx 20$ we decided to limit the comparison at low Galactic latitude where the SNR is very high.

We note that in the following we considered the error bars on the power spectrum as coming from the Gaussian part of the variance, i.e. following Hivon et al. (2002)

$$\Delta\tilde{C}_\ell = \sqrt{\frac{2}{\nu}}(C_\ell + N_\ell) \tag{6}$$

where ν is the number of degrees of freedom taking into account the finite number of modes going into the power spectrum calculation in each ℓ mode and the effective sky coverage. N_ℓ represents the noise power spectrum and the C_ℓ is the theoretical model describing the CO angular power spectrum with the tilde denoting measured quantities. Because we do not know the true CO theoretical power spectrum we assumed that $C_\ell + N_\ell = \tilde{C}_\ell$. The Gaussian approximation however underestimates the error bars. The CO field is in fact a highly non-Gaussian field with mean different from zero. As such, its variance should contain contributions coming from the expectation value of its 1 and 3 point function in the harmonic domain that are zero in the Gaussian approximation. These terms are difficult to compute and we considered the Gaussian approximation sufficient for the level of accuracy of this study.

As can be seen in Fig. 7, all of the power spectra of CO emission at low Galactic latitudes have a broad peak around the multipole $100 \div 300$, i.e. at the $\approx$1 deg angular scale. The signal power starts decreasing up to $\ell \sim 600$ and then grows again at higher ℓ due to the *Planck* instrumental noise contamination. Such a broad peak suggests that there is a correlated angular scale along the Galactic plane. This can be understood with a quick order of magnitude estimate. If we assume that most of the CO emission is localized at a distance of 4 kpc (in the molecular ring) and molecular clouds have a typical size of 30 pc, we find that each cloud subtends a $\sim$0.5 deg area in the sky. This corresponds to a correlated scale in the power spectrum at an ℓ of the order of a few hundred but the detail of this scale depends on the width of the molecular ring zone.

3.3 Galactic plane profile emission comparison

As a first test, we compared the profile of CO emission in the Galactic plane predicted by the model and the one observed in the data. Since we are mostly interested in a comparison as direct as possible with the *Planck* observed data, we convolved the MCMOLE3D maps with a Gaussian beam of 10 arcmin FWHM, corresponding to the nominal resolution of the 100 GHz channel of HFI, prior to any further processing.

In order to compare the data and the simulations, we constrained the total flux of the simulated CO maps with the one observed in

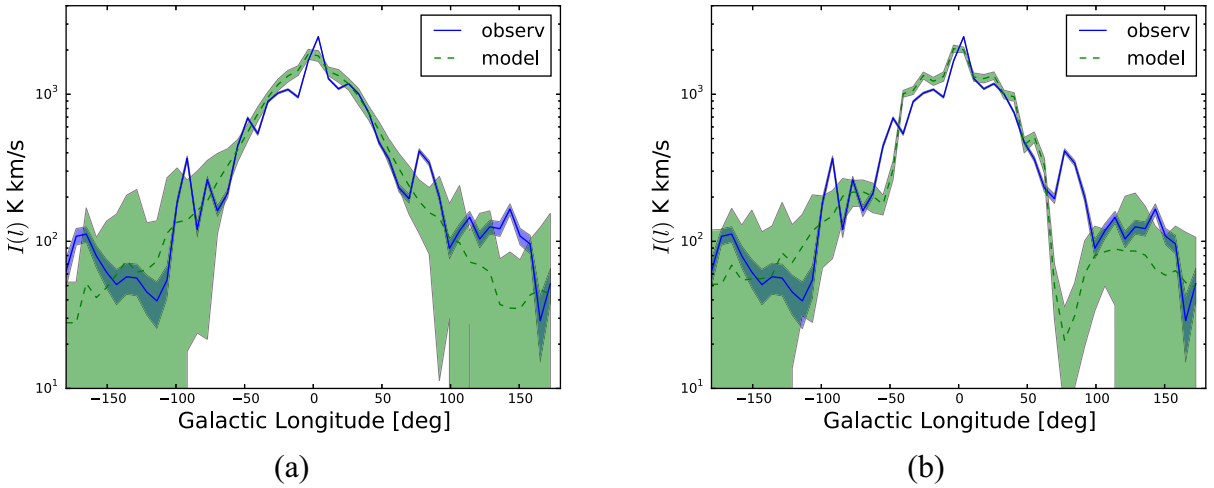

Figure 8. Comparison of 100 MC realization of MCMOLE3D simulated Galactic CO emission profiles (dashed green) with *Planck* observations (solid blue). The average profile integrals defined in equation (7) for the (a) AXISYMMETRIC and (b) LOGSPIRAL geometries are shown with dashed lines. The shaded area displays the standard deviation of all MCs in each longitude bin (green), or the noise level of *Planck* (solid blue) estimated from the *Type 1* null map.

the *Planck* data. This is necessary, otherwise the emission would be directly proportional to the number of clouds distributed in the simulated Galaxy. Such a procedure also breaks possible parameter degeneracies with respect to the amplitude of the simulated power spectra (see Section 3.4). Following Bronfman et al. (1988), we therefore computed the integrated flux of the emission along the two Galactic latitudes and longitudes (l, b) defined as

$$I^X(l) = \int dbI^X(l,b), \tag{7}$$

$$I^X_{\rm tot} = \int dldbI^X(l,b), \tag{8}$$

where X refers both to the *model* and to the *observed* CO map. We then rescaled the simulated maps, dividing by the factor f defined as:

$$f = \frac{I^{\rm observ}_{\rm tot}}{I^{\rm model}_{\rm tot}}. \tag{9}$$

We estimated the integrals in equations (7) and (8) by considering a narrow strip of Galactic latitudes within $[-2, 2]$ deg. We found that the value of f is essentially independent of the width of the Galactic latitude strip used to compute the integrals because most of the emission comes from a very thin layer along the Galactic plane of amplitude $|b| \lesssim 2$ deg.

In Fig. 8, we show the comparison between $I^{\rm observ}(l)$ and the $I^{\rm model}(l)$ as defined in equation (7) computed as the mean of 100 MC realizations of galaxies populated by molecular clouds for both the AXISYMMETRIC and LOGSPIRAL models as well as their typical standard deviation. In particular, we chose for these simulations the default parameters in Table 1. The emission profiles are quite consistent in the regions from which most of the CO emission comes, i.e. in the inner Galaxy, the I and the IV quadrants (longitude in $[-90, 90]$ deg[7]). On the contrary, the emission in the other two quadrants looks to be underestimated but within the scatter of the simulations. In fact, the observed emissions in both the II and III quadrants come mainly from the closer and more isolated system of clouds. These are actually more difficult to simulate because in that area (at Galactic longitudes $|l| > 100$ deg) the presence of noise starts to be non-negligible (see shaded blue in Fig. 8).

In addition, we note that the bump in the profile at $l \simeq 60$–70 deg, where we see a lack of power in both the AXISYMMETRIC and LOGSPIRAL cases, corresponds to the complex region of *Cygnus-X*, which contains the very well-known X-ray source Cyg-X1, massive protostars and one of the most massive molecular clouds known, $3 \times 10^6\,{\rm M}_\odot$, 1.4 kpc distant from the Sun. Given the assumptions made in Section 2, these large and closer clouds are not easy to simulate with MCMOLE3D especially where they are unlikely to be found, as in interspiral arm regions. Despite of this, one can notice an overall qualitative better agreement with observations for the LOGSPIRAL model than for the AXISYMMETRIC one. The latter reconstructs the global profile very well, but the former contains more peculiar features such us the central spike due to the Central Molecular Zone within the bar, or the complex of clouds at longitudes around ~ -140, -80 and 120 deg. We will perform a more detailed comparison of the two geometries in the following section and in Appendix A.

3.4 Constraining the MCMOLE3D model with *Planck* data

After comparing the CO profile emission, we checked whether the MCMOLE3D model is capable of reproducing the characteristic shape of the *Planck* CO angular power spectrum. Given the knowledge we have on the shape of the Milky Way, we decided to adopt the LOGSPIRAL geometry as a baseline for this comparison, and to fix the parameters for the specific geometry to the values describing the shape of our Galaxy (see Section 2.3). For sake of completeness, we reported the results of the same analysis adopting an AXISYMMETRIC geometry in Appendix A.

We left the typical cloud size L_0 and $\sigma_{\rm ring}$ (the width of the molecular ring) as free parameters of the model. While the former is directly linked to the observed angular size of the clouds, the role of the second one is not trivial, especially if we adopt the more realistic four spiral arms distribution. Intuitively, it changes the probability of observing more clouds closer to the observer and affects more the amplitude of the power on the larger angular scales.

[7] We stress that the definition of quadrants comes from the Galactic coordinates centred on the Sun. The I and IV quadrants are related to the inner Galaxy, while the II and the III ones look at its outer regions.

Table 2. Summary table of best-fitting parameters obtained using the two different *Planck* CO maps.

	L_0 (pc)	$\sigma_{\rm ring}$ (kpc)	$A_{\rm CO}$	$\tilde{\chi}^2$	dof	p-value	$\rho_{L\sigma}$
Type 1	14.50 ± 0.58	2.76 ± 0.19	0.69 ± 0.06	1.48	11	0.13	0.74
Type 2	11.59 ± 1.09	3.11 ± 0.32	1.37 ± 0.19	1.95	11	0.03	0.92

We defined a large interval, reported in Table 1, where L_0 and $\sigma_{\rm ring}$ are allowed to vary. Looking at the series of examples reported in Fig. 5, we can see that suitable parameter ranges which yield power spectra close to the *Planck* observations are $L_0 = 10 \div 30$ pc and $\sigma_{\rm ring} = 2 \div 3$ kpc. It is interesting to note that these are in agreement with estimates available in the literature (see e.g. Roman-Duval et al. 2010; Ellsworth-Bowers et al. 2015).

We then identified a set of values within the intervals just mentioned for which we computed the expected theoretical power spectrum of the specific model. Each theoretical model is defined as the mean of the angular power spectrum of 100 MC realizations of the model computed with X^{2}PURE. For each realization of CO distribution, we rescaled the total flux following the procedure outlined in the previous section before computing its power spectrum.

Once the expected angular power spectra for each point of the parameter domain had been computed, we built the hyper-surface $\mathcal{F}(\ell; \sigma_{\rm ring}, L_0)$ which for a given set of values $(\sigma_{\rm ring}, L_0)$ retrieved the model power spectrum, by interpolating it from its value at the closest grid points using splines. We checked that alternative interpolation methods did not impact significantly our results. We then computed the best-fitting parameters of the MCMOLE3D model by performing a χ^2 minimization with the *Planck* CO power spectrum data. For this procedure, we introduced a further global normalization parameter $A_{\rm CO}$ to take account of the *Planck* bandpass effects or other possible miscalibration of the model. These might come either from variations from the scaling laws employed in the model (that are thus not captured by the total flux normalization described earlier), or calibration differences between the *Planck* data and the surveys used to derive the scaling laws themselves. The bandpass effects tend to decrease the overall amplitude of the simulated signal because each line gets diluted over the width of the *Planck* frequency band.

Since the theoretical model has been estimated from MC simulations, we added linearly to the sample variance error of the *Planck* data an additional uncertainty budget corresponding to the uncertainty of the mean theoretical power spectrum estimated from MC. We note that when we compute the numerator of the f rescaling factor, we include not only the real flux coming from the CO lines but also an instrumental noise contribution. We therefore estimated the expected noise contribution to f by computing the integral of equation (8) on the *Planck* error map and found it to be equal to 10 per cent. We propagated this multiplicative uncertainty to the power spectrum level, rescaling the mean theoretical MC error bars by the square of this factor.

We limit the range of angular scales involved in the fit to $\ell \leq 400$ in order to avoid the regions that display an unusual bump at scales of around $\ell \approx 500$ that is not captured by any realization of our model (see Section 3.5). The best-fitting parameters are reported in Table 2

$$\begin{aligned} L_0 &= 14.50 \pm 0.58\,\text{pc}, \\ \sigma_{\rm ring} &= 2.76 \pm 0.19\,\text{kpc}, \\ A_{\rm CO} &= 0.69 \pm 0.06. \end{aligned} \tag{10}$$

The values are within the ranges expected from the literature. As can be seen in Fig. 9 the power spectrum corresponding to the model with the best-fitting parameters, describes the *Planck* data reasonably well. The minimum χ^2 obtained by the minimization process gives 1.48 that corresponds to a p-value of 13 per cent. We note, however, that all of the parameters are highly correlated. This is somewhat expected as the larger is $\sigma_{\rm ring}$, the closer the clouds get to the observer placed in the solar circle. This effect can be compensated by an overall decrease of the typical size of the molecular cloud as shown in Fig. 5(d).

Finally, we note that $A_{\rm CO} \lesssim 1$ suggests that, despite the rescaling procedure constraining quite well the overall power spectrum amplitude, the spatial distribution seems to be more complex than the one implemented in the model. This might partially be explained by the fact that we do not model explicitly any realistic bandpass effect of the *Planck* channel or the finite width of the CO line. Additional sources of signal overestimation could be residual contamination of ^{13}CO $1-0$ line or thermal dust in the map or variations of the emissivity profile in equation (3).

3.5 Consistency checks on other maps

The *Planck* collaboration released multiple CO maps extracted using different component separation procedures. We can test the stability of our results by using CO maps derived with these different approaches, in particular the so-called *Type 2* maps. These have been produced exploiting the intensity maps of several frequencies (*multichannel* approach) to separate the CO emission from the astrophysical and CMB signal (Planck Collaboration XIII 2014b). The maps are smoothed at a common resolution of 15 arcmin and have better SNR ratio than the *Type 1* ones. However, the CO is extracted by assuming several simplifications which may leak into contamination due to foreground residuals and systematics, as explained in Planck Collaboration X (2016b, section 5.5.3).

We repeated the procedure outlined in Sections 3.3 and 3.4 using the *Type 2* $1-0$ map. The values of the best-fitting parameters are summarized in Table 2 and we show in Fig. 9 the best-fitting model power spectrum together with the power spectrum of *Type 2* map data. We found that the values of $A_{\rm CO}$ obtained for *Type 2* are inconsistent with the one obtained for the *Type 1* maps. However, this discrepancy is consistent with the overall inter calibration difference between the two maps reported in Planck Collaboration X (2016b). Such differences are mainly related to a combination of bandpass uncertainties in the *Planck* observations and presence of a mixture of ^{12}CO and ^{13}CO (emitted at 110 GHz) lines for the *Type 1 maps*. While $\sigma_{\rm ring}$ is consistent between the two maps, the *Type 2* L_0 parameters are in slight tension at 2.7σ level. The overall correlation of the parameters is increased and the overall agreement between data and the MCMOLE3D mode is reduced although it remains acceptable. We cannot exclude however that this is a sign of additional systematic contamination in the *Type 2* maps.

The *Planck* collaboration provided maps of the $2-1$ line for both of the methods and we could use our model to constrain the relative amplitudes of the lines, while fixing the parameter of the cloud distribution. However, such analysis is challenging and might

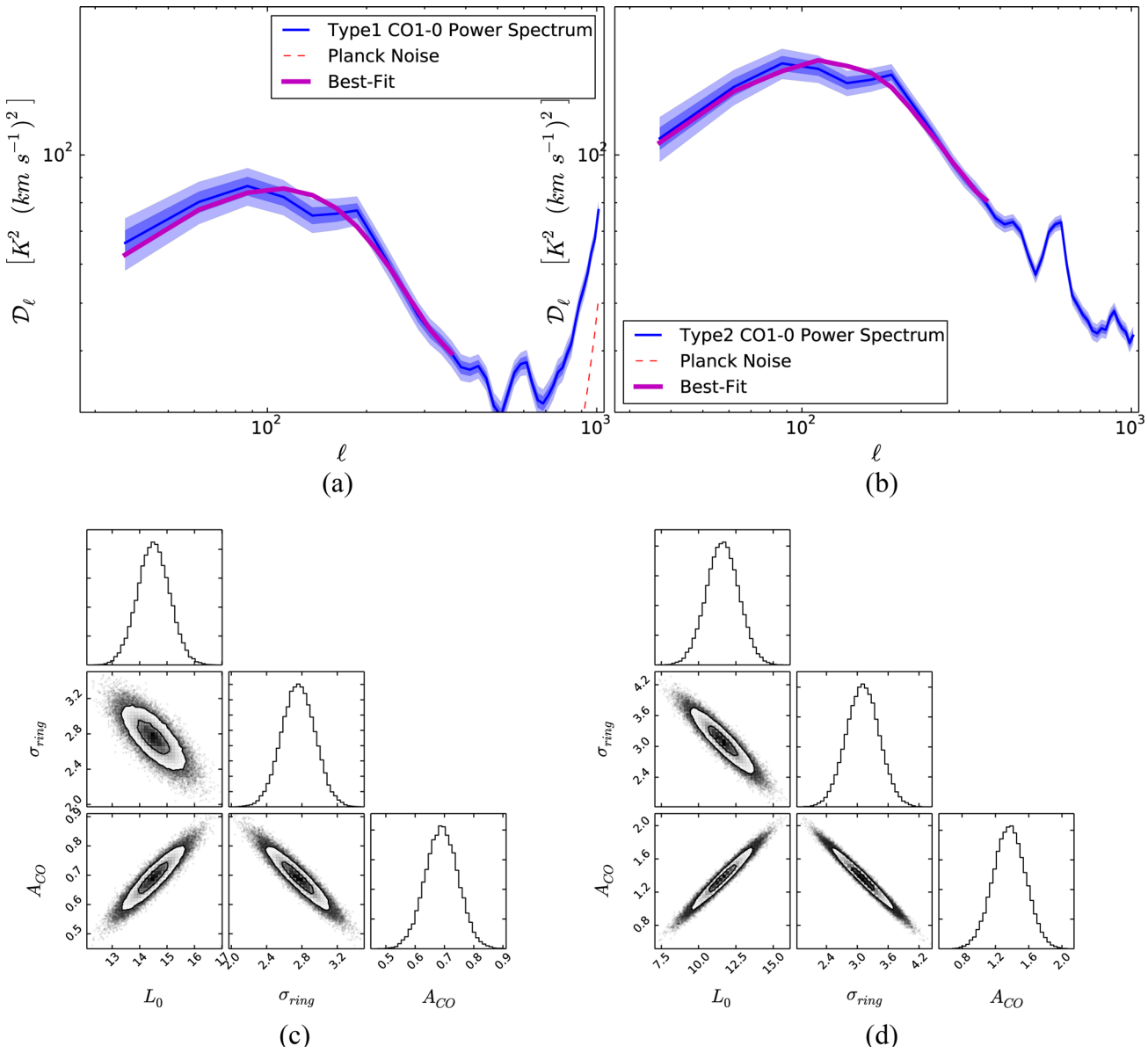

Figure 9. *Top panels*: angular power spectra (solid thin blue) of the *Planck Type 1* (left) and *Type 2* (right) maps. The shaded area correspond to the 1σ (dark blue) and 2σ (light blue) error bars including statistical and systematic uncertainties. The MCMOLE3D model CO angular power spectrum assuming the best-fitting parameters of equation (10) is shown in thick solid magenta. The *Planck* noise power spectrum is shown in red dashed, in the top right panel the noise level is about one order of magnitude smaller than the one in the top left panel. *Bottom panels*: best-fitting parameters of the MCMOLE3D model describing the *Planck* CO angular power spectrum and their correlations.

be biased by the presence of variations of local physical properties of the clouds (opacity and temperature) or by the red or blue shift of the CO line within the *Planck* bandpass induced by the motion of the clouds themselves (Planck Collaboration X 2016b). For these reasons, we decided to restrict our analysis only to the CO $1-0$ line, since it is the one for which the observational data are more robust.

We finally note that the observed angular power spectra of the *Planck* maps display an oscillatory behaviour at a scale of $\ell \geq 400$ with a clear peak at around $\ell \approx 500$. The fact that this feature is present in all of the lines and for all of the CO extraction methods means that it can reasonably be considered as a meaningful physical signature. Because a single cloud population produces an angular power spectrum with a characteristic peak scale, we speculate that this could be the signature of the presence of an additional cloud population with a different typical size or location. We however decided to leave the investigation of this feature for a future work.

3.6 Comparison with data at high Galactic latitudes

In Fig. 10, we compare the *Planck* CO $1-0$ power spectrum at HGL with the average power spectrum of 100 MC realizations of the MCMOLE3D model for the same region of the sky. We assumed for these runs the best-fitting values of the L_0 and $\sigma_{\rm ring}$ parameters reported in equation (10) and a LOGSPIRAL distribution. Because the *Planck* maps at these latitudes are dominated by noise, we subtracted our MC estimates of the noise bias data power spectrum so as to have a better estimate of the signal (blue circles).

As can be observed in Fig. 10, some discrepancy arises when comparing the power spectrum expected from the simulation of

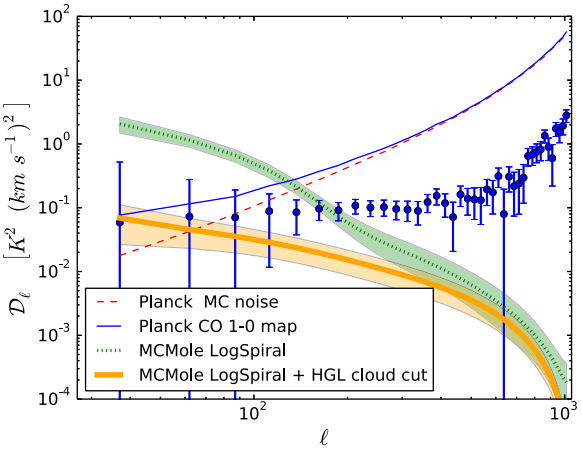

Figure 10. CO $1-0$ power spectrum at HGL of the LOGSPIRAL MCMOLE3D model (dotted green), using the parameters in equation (10). Thick solid orange—we show the power spectrum for the LOGSPIRAL geometry, the same parameters in equation (10) and with the HGL cut of clouds at $|b| > 30$ deg whose flux exceeds the *Planck* noise. The *Planck Type 1* CO power spectrum before and after noise bias subtraction is shown with the blue solid line and filled circles, respectively; the error bars account for both *Planck* data statistical uncertainties and systematics from the MCMOLE3D simulations. The noise bias is shown with the dashed red line.

LOGSPIRAL MCMOLE3D at HGL with the noise debiased *Planck* data. This is rather expected because the model has larger uncertainties at HGL than in the Galactic mid-plane (where the best-fitting parameters are constrained) given the lack of high sensitivity data. The discrepancy seems to point to an overestimation of the vertical profile parameters $\sigma_{z,0}$ and h_R (see equation 2) which gives a higher number of clouds close to the observer at high latitude. However, we also point out that, as explained in Section 3.2, the error bars in Fig. 10 might be underestimated especially at the largest angular scales where we are signal dominated. Therefore, discrepancies of order $\approx 3\sigma$ do not seem unlikely. Since we are mainly interested in using the model to forecast the impact of unresolved CO emission far from the Galactic plane ($|b| > 30$), we investigated whether removing the few HGL clouds in the simulation that appear close to the observer would improve the agreement with the *Planck* data. All of these clouds have, in fact, a flux exceeding the *Planck* CO map noise in the same sky area and they should have already been detected in real data. We will refer to this specific choice of cut as the HGL cut in the following. The power spectrum of the MCMOLE3D simulated maps obtained after the application of the HGL cut is shown in Fig. 10. We found that the model calibrated at low latitudes and after the application of the HGL cut agrees very well with the data on the angular scales where the signal slightly dominates, i.e. $\ell \lesssim 80$. We could not extend the comparison to smaller angular scales because the data become noise dominated and the residual increase of power observed on the power spectrum is dominated by a noise bias residual.

4 POLARIZATION FORECASTS

As noted in Section 1, CO lines are polarized and could contaminate sensitive CMB polarization measurements together with other polarized Galactic emission (synchrotron and the thermal dust) at sub-mm wavelengths. Future experiments will preferentially perform observations at intermediate and HGL, to minimize contamination from strong Galactic emissions close to the plane. Since CO data at HGL are not sensitive enough to perform accurate studies of this emission, we provide two complementary estimates of the possible contamination from its polarized counterpart to the CMB B-mode power spectrum in this sky region.

4.1 Data-based order of magnitude estimate

Starting from the measured *Planck* power spectrum at low Galactic latitudes, one can extrapolate a very conservative value of the CO power spectrum at higher latitudes. Assuming that all of the variance observed in the HGL region is distributed among the angular scales in the same way as in the Galactic plane, we can write

$$\mathcal{C}_\ell^{\rm high,CO} = \mathcal{C}_\ell^{\rm Gal}\frac{\rm var(high)}{\rm var(Gal)}. \tag{11}$$

This is a somewhat conservative assumption because we know that the bulk of the CO line emission is concentrated close to the Galactic disc and also because it assumes that the *Planck* noise at HGL is diffuse CO emission. The variance of the *Planck* CO map is $0.3\,{\rm K}^2({\rm km\,s}^{-1})^2$, at $|b| > 30$ deg, while for $|b| < 30$ deg we get a variance of $193.5\ {\rm K}^2\,({\rm km\,s}^{-1})^2$. Taking 1 per cent as the polarization fraction, $p_{\rm CO}$, of the CO emission and an equal power in E and B modes of polarized CO, we can convert $\mathcal{C}_\ell^{\rm COhigh}$ into its B-mode counterpart as $\mathcal{C}_\ell^{\rm COhigh,EE} = \mathcal{C}_\ell^{\rm COhigh,BB} = \mathcal{C}_\ell^{\rm high,CO}\,p_{\rm CO}^2/2$. We then apply the conversion factors of Planck Collaboration XIII (2014b) to convert the CO power spectrum into thermodynamic units (from ${\rm K_{RJ}kms}^{-1}$ to μK). We can compare $\mathcal{C}_\ell^{\rm COhigh,BB}$ to the amplitude of equivalent cosmological CMB inflationary B modes with tensor-to-scalar ratio $r = 1$ at $\ell = 80$. In terms of $\mathcal{D}_\ell^{\rm BB}$, this is equal to $\sim 6.67 \times 10^{-2}\mu{\rm K}^2$ for a fiducial *Planck* 2015 cosmology. We found that the amplitude of the extrapolated CO B-mode power spectrum is equal to a primordial B-mode signal having $r_{\rm CO} = 0.025$.

4.2 Simulation estimate

In order to verify and refine the estimate given in the previous section, we used the model presented in Section 2 to infer the level of contamination from unresolved polarized CO emission. For doing this, we first set the free parameter of the MCMOLE3D model to the best-fitting value derived in equation (10).

From the total unpolarized emission in each sky pixel of the simulation, $I^{\rm CO}$ we can then extract its linearly polarized part by taking into account the global properties of the Galactic magnetic field. Following Delabrouille et al. (2013) and Tassis & Pavlidou (2015), the Q and U Stokes parameters of each CO cloud can be related to the unpolarized emission as

$$Q(\hat{n})^{\rm CO} = p_{\rm CO}\, g_d(\hat{n}) I(\hat{n})^{\rm CO} \cos(2\psi(\hat{n})), \tag{12}$$

$$U(\hat{n})^{\rm CO} = p_{\rm CO}\, g_d(\hat{n}) I(\hat{n})^{\rm CO} \sin(2\psi(\hat{n})).. \tag{13}$$

where $p_{\rm CO}$ is the intrinsic polarization fraction of the CO lines, while g_d is the geometric depolarization factor which accounts for the induced depolarization of the light when integrated along the line of sight. The polarization angle ψ describes the orientation of the polarization vector and, for the specific case of Zeeman emission, it is related to the orientation of the component of the Galactic magnetic field orthogonal to the line of sight $B_\perp$. Following the findings of Greaves et al. (1999), we adopted a conservative choice of a constant $p_{\rm CO} = 1$ per cent for each molecular cloud of the simulation. Because the polarized emission in molecular clouds is correlated with the polarized dust emission (see e.g. Crutcher 2012), we used the g_d and ψ templates for the Galactic dust emission available in the public release of the Planck Sky Model suite[8] (Delabrouille

[8] http://www.apc.univ-paris7.fr/delabrou/PSM/psm.html

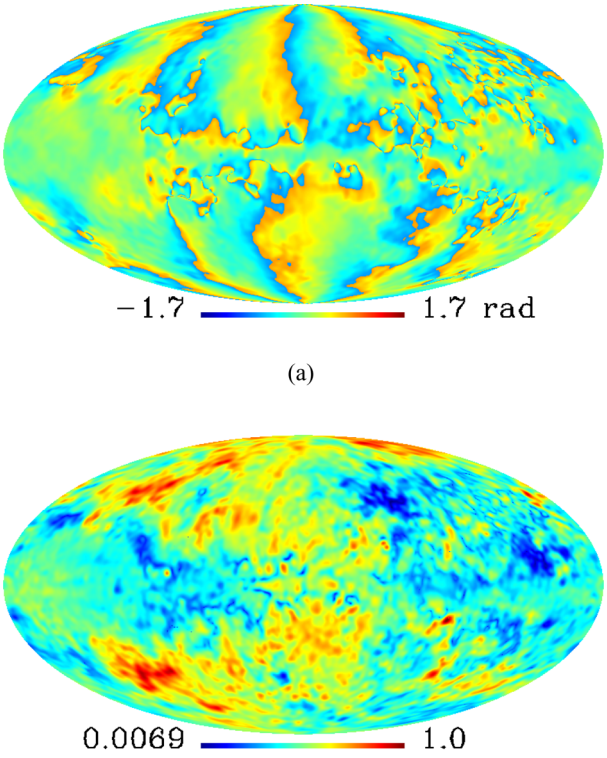

Figure 11. (a) Polarization angle and (b) geometrical depolarization maps used for simulating polarized CO emission in this work.

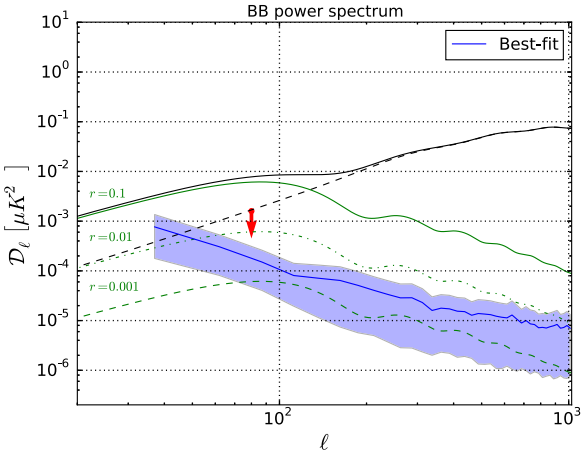

Figure 12. B-mode power spectra of polarized CO emission lines at HGL estimated using the best-fitting parameters of the LOGSPIRAL MCMOLE3D model (see equation 10). The expected *Planck* 2015 ΛCDM cosmological signal including the gravitational lensing contribution is shown in black. Potential contributions from inflationary B modes for tensor-to-scalar ratios of $r = 0.1$, 0.01 and 0.001 are shown with solid, dot–dashed and dashed green lines, respectively. The red arrow indicates the upper limits obtained in Section 4.1

et al. 2013). These have been derived from 3D simulations of the Galactic magnetic field (including both a coherent and a turbulent component) and data of the *WMAP* satellite and both are shown in Fig. 11.

Since we assumed a constant polarization fraction, the geometrical depolarization effectively induces a change in the polarization fraction as a function of Galactic latitudes decreasing it when moving away from the poles. This effect has already been confirmed by *Planck* observations (Planck Collaboration X 2016b) of thermal dust, whose polarization fraction increases at high latitudes.

In order to forecast the contamination of unresolved CO polarized emission alone, we apply the HGL cut as described in Section 3.6 to each realization of the model for consistency.

Once the $Q^{\rm CO}$ and $U^{\rm CO}$ maps have been produced, we computed the angular power spectrum using X^{2}PURE.

In Fig. 12, we show the mean and standard deviation of the B-mode polarization power spectrum extracted from 100 MC realizations of the CO emission following the procedure just outlined. Even though in Section 3.4 we showed that our model tends to slightly overestimate the normalization of the power spectrum, we decided not to apply the best-fitting amplitude $A_{\rm CO}$ to the amplitude of the B-mode power spectrum in order to provide the most conservative estimates of the signal.

As could be seen from Fig. 12, there is a significant dispersion compared to the results of the MC simulations at low Galactic latitudes (see Fig. 5). This simply reflects the fact that the observations, and hence our model, do not favour the presence of molecular clouds at HGL. Therefore their number can vary significantly between realizations. We repeated this test using the AXISYMMETRIC geometry and changing the parameter $\sigma_{\rm ring}$. The result is stable with respect to these assumptions. We found that the spatial scaling of the average E- and B-mode power spectrum can be approximated by a decreasing power law $\mathcal{D}_\ell \sim \ell^\alpha$, with $\alpha = -1.78$.

Our simulations suggest that the level of polarized CO emission from unresolved clouds, despite being significantly lower than synchrotron or thermal dust, can nevertheless significantly bias the primordial B-mode signal if not taken into account. The signal concentrates mainly on large angular scales and at $\ell \sim 80$, $\mathcal{D}_\ell = (1.1 \pm 0.8) \times 10^{-4}\ \mu{\rm K}^2$, where the uncertainty corresponds to the error in the mean spectra estimated from the 100 MC realizations. Therefore, the level of contamination is comparable to a primordial B-mode signal induced by tensor perturbations of amplitude $r_{\rm CO} = 0.003 \pm 0.002$, i.e. below the recent upper limit $r < 0.07$ reported by the BICEP2 Collaboration (BICEP2 Collaboration et al. 2016) but higher than the $r = 0.001$ target of upcoming experiments (Matsumura et al. 2014; Abazajian et al. 2016; CORE Collaboration et al. 2016). The contamination quickly becomes sub-dominant on small angular scales ($\ell \approx 1000$), where the B modes are mostly sourced by the gravitational lensing.

We finally note that these estimates are conservative since the assumed polarization fraction of 1 per cent of polarized is close to the high end of the polarization fractions observed in CO clouds.

5 CONCLUSIONS

In this work, we have developed a parametric model for CO molecular line emission which takes account of the CO clouds distribution within our Galaxy in 3D with different geometries, as well as the most recent observational findings concerning their sizes, locations and emissivity.

Despite most of the observations having so far been confined to the Galactic plane, we have built the model to simulate the emission over the full sky. The code implementing MCMOLE3D is being made publicly available.

We have compared the results of our simulations with *Planck* CO data on the map level and statistically (by matching angular power spectra). We found that:

(i) the parameters of the size function, L_0, and the width of the Galactic radial distributions $\sigma_{\rm ring}$ play a key role in shaping the power spectrum;
(ii) the choice of symmetries in the cloud distribution changes the profile of the integrated emission in the Galactic plane (equation 7), but not the power spectrum morphology;
(iii) our model is capable of reproducing fairly well the observations at low Galactic latitudes (see Fig. 8) and the power spectrum at high latitudes (Fig. 10).

We used our model to fit the *Planck* observed CO power spectrum and to estimate the most relevant parameters of the CO distribution, such as the typical size of clouds and the thickness of the molecular ring, finding results in agreement with values reported in the literature. The model which we have developed could easily be generalized and extended whenever new data become available. In particular, its accuracy at HGL would greatly benefit from better sub-mm measurements going beyond the *Planck* sensitivity, as well as from better information about the details of the CO polarization properties.

Finally, we used the best-fitting parameters obtained from comparing the MCMOLE3D model with *Planck* data to forecast the unresolved CO contamination of the CMB B-mode power spectrum at HGL. We conservatively assumed a polarization fraction of $p_{\rm CO} = 1$ per cent, which corresponds to the high end of those observed at low latitudes, since no polarized CO cloud has yet been observed far from the Galactic plane due to the weakness of this emission.

We found that this signal could mimic a B-mode signal with tensor-to-scalar ratio $0.001 \lesssim r \lesssim 0.025$. This level of contamination is indeed relevant for accurate measurements of CMB B modes. It should therefore be inspected further in light of the achievable sensitivities of upcoming and future CMB experiments together with the main diffuse polarized foreground (thermal dust and synchrotron). From the experimental point of view, trying to find dedicated instrumental solution for minimizing the impact of CO emission lines, appears to be particularly indicated in the light of these results.

ACKNOWLEDGEMENTS

We would like to thank Françoise Combe for many useful comments and suggestions for the development of this study, as well as Alessandro Bressan, Luigi Danese, Andrea Lapi, Akito Kusaka, Davide Poletti and Luca Pagano for several enlightening discussions. We thank John Miller for his careful reading of this work. We thank Guillaume Hurier for several clarifications about the *Planck* CO products. This work was supported by the RADIOFOREGROUNDS grant of the European Union's Horizon 2020 research and innovation programme (COMPET-05-2015, grant agreement number 687312) and by Italian National Institute of Nuclear Physics (INFN) INDARK project. GF acknowledges support of the CNES postdoctoral programme. Some of the results in this paper have been derived using the HEALPIX (Górski et al. 2005) package.

REFERENCES

Abazajian K. N. et al., 2016, preprint (arXiv:1610.02743)
Baccigalupi C., 2003, New Astron. Rev., 47, 1127
Bally J., Langer W. D., Stark A. A., Wilson R. W., 1987, ApJ, 312, L45
Balser D. S., Rood R. T., Bania T. M., Anderson L. D., 2011, ApJ, 738, 27
BICEP2 Collaboration et al., 2016, Phys. Rev. Lett., 116, 031302
Bobylev V. V., Bajkova A. T., 2013, Astron. Lett., 39, 759
Bronfman L., Cohen R. S., Alvarez H., May J., Thaddeus P., 1988, ApJ, 324, 248
Bunn E. F., Zaldarriaga M., Tegmark M., de Oliveira-Costa A., 2003, Phys. Rev. D, 67, 023501
CORE Collaboration et al., 2016, preprint (arXiv:1612.08270)
Crutcher R. M., 2012, ARA&A, 50, 29
Dame T. M., Hartmann D., Thaddeus P., 2001, ApJ, 547, 792
Davis B. L., Berrier J. C., Shields D. W., Kennefick J., Kennefick D., Seigar M. S., Lacy C. H. S., Puerari I., 2012, ApJS, 199, 33
Delabrouille J. et al., 2013, A&A, 553, A96
Ellsworth-Bowers T. P., Rosolowsky E., Glenn J., Ginsburg A., Evans II N. J., Battersby C., Shirley Y. L., Svoboda B., 2015, ApJ, 799, 29
Gold B. et al., 2011, ApJS, 192, 15
Goldreich P., Kylafis N. D., 1981, ApJ, 243, L75
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Grain J., Tristram M., Stompor R., 2009, Phys. Rev. D, 79, 123515
Greaves J. S., Holland W. S., Friberg P., Dent W. R. F., 1999, ApJ, 512, L139
Heyer M., Dame T., 2015, ARA&A, 53, 583
Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2
Hu W., White M. J., 1997, New Astron., 2, 323
Jackson J. M. et al., 2006, ApJS, 163, 145
Krachmalnicoff N., Baccigalupi C., Aumont J., Bersanelli M., Mennella A., 2015, A&A, 588, A65
Lewis A., Challinor A., Turok N., 2001, Phys. Rev. D., 65, 023505
Matsumura T. et al., 2014, J. Low Temp. Phys., 176, 733
Mizuno A., Fukui Y., 2004, in Clemens D., Shah R., Brainerd T., eds, ASP Conf. Ser. Vol. 317, Milky Way Surveys: The Structure and Evolution of our Galaxy. Astron. Soc. Pac., San Francisco, p. 59
Page L. et al., 2007, ApJS, 170, 335
Planck Collaboration IX, 2014a, A&A, 571, A9
Planck Collaboration XIII, 2014b, A&A, 571, A13
Planck Collaboration XXX, 2016a, A&A, 586, A133
Planck Collaboration X, 2016b, A&A, 594, A10
Planck Collaboration XXV, 2016c, A&A, 594, A25
Planck Collaboration L, 2017, A&A, 599, A51
Regan M. W., Sheth K., Teuben P. J., Vogel S. N., 2002, ApJ, 574, 126
Roman-Duval J., Jackson J. M., Heyer M., Rathborne J., Simon R., 2010, ApJ, 723, 492
Roman-Duval J., Heyer M., Brunt C. M., Clark P., Klessen R., Shetty R., 2016, ApJ, 818, 144
Seljak U., Zaldarriaga M., 1997, Phys. Rev. Lett., 78, 2054
Smith K. M., Zaldarriaga M., 2007, Phys. Rev. D, 76, 043001
Tassis K., Pavlidou V., 2015, MNRAS, 451, L90
Vallée J. P., 2014, ApJS, 215, 1
Wolfire M. G., McKee C. F., Hollenbach D., Tielens A. G. G. M., 2003, ApJ, 587, 278

APPENDIX A: BEST FITTING WITH AXISYMMETRIC GEOMETRY

In this appendix, we present the results of the analysis described in Section 3.4 to constraint the CO distribution using the MCMOLE3D model adopting an AXISYMMETRIC geometry instead of the LOGSPIRAL one. Following the procedure of Section 3.4, we construct a series of $\mathcal{F}(\ell; \sigma_{\rm ring}, L_0)$ hypersurfaces sampled on an ensemble of specific values of the L_0 and $\sigma_{\rm ring}$ parameters within the same ranges reported in Section 3.4.

In Fig. A1, we show the results of the fit of the axisymmetric MCMOLE3D model to the CO power spectrum of the *Planck Type*

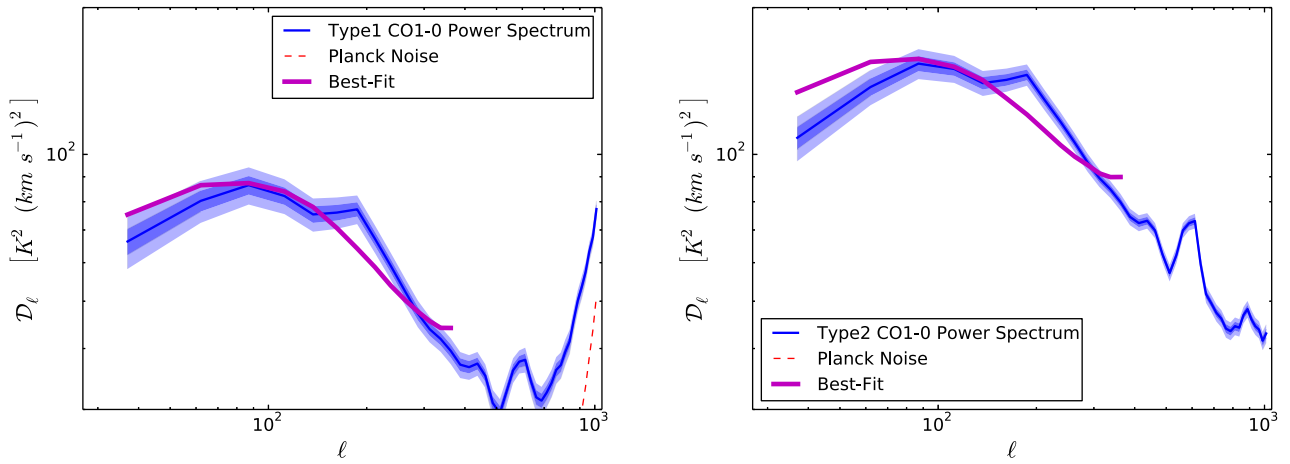

Figure A1. Angular power spectra of *Planck* CO $1-0$ line (blue) for *Type 1* (left) and *Type 2* (right) maps together with the results of the MCMOLE3D best-fitting model adopting an AXISYMMETRIC geometry (magenta).

Table A1. Best-fitting parameters for the AXISYMMETRIC MCMOLE3D model.

	L_0 (pc)	$\sigma_{\rm ring}$ (kpc)	$A_{\rm CO}$	$\tilde{\chi}^2$	dof	p-value	$\rho_{L\sigma}$
Type 1	19.47 ± 12.68	2.12 ± 0.23	1.00 ± 0.12	7.35	11	0.00	0.99
Type 2	16.24 ± 17.56	2.12 ± 0.30	2.25 ± 0.35	18.08	11	0.00	0.99

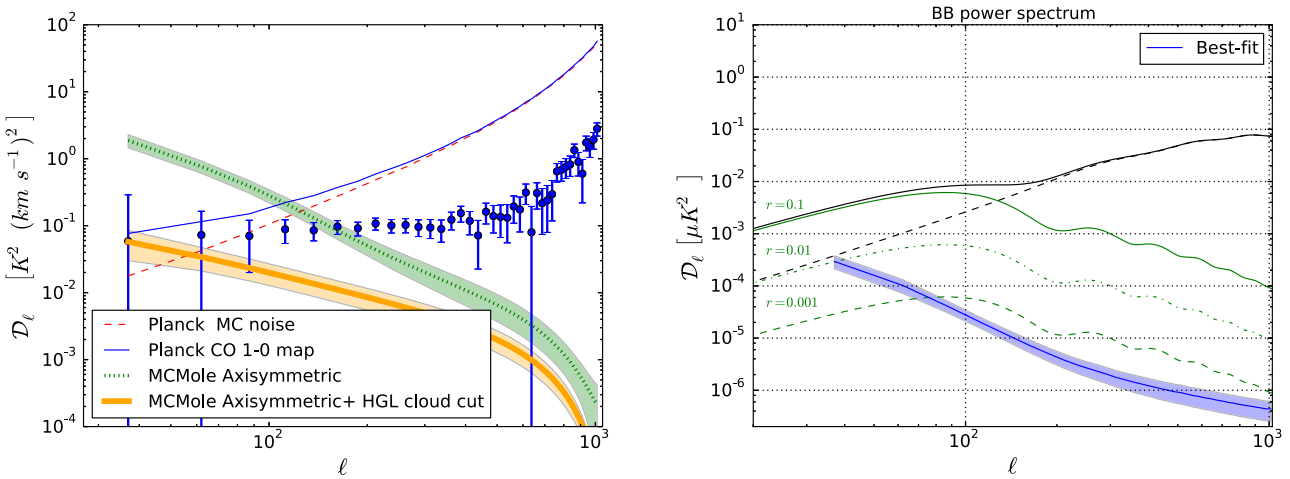

Figure A2. Left: CO $1-0$ power spectrum at HGL of the AXISYMMETRIC MCMOLE3D model (dotted green), for the best-fitting parameters reported in Table A1. Thick orange solid line shows the power spectrum for the AXISYMMETRIC geometry with the HGL cut applied. The *Planck Type 1* CO power spectrum before and after noise bias subtraction is shown with the blue solid line and dots, respectively. The *Planck* noise bias is shown with the dashed red line. Right: B-mode power spectra of polarized CO emission lines at HGL estimated using the best-fitting parameters of the AXISYMMETRIC MCMOLE3D model, see Fig. 12 for a comparison with the LOGSPIRAL geometry.

1 and *Type 2* CO maps in the Galactic plane. We summarize the best-fitting values of these parameters in Table A1. As it can be seen from the results of the χ^2 test in Table A1 the AXISYMMETRIC model does not fit the data satisfactorily. Moreover one of the parameters of the model, the typical cloud size L_0, is in practice unconstrained. For this reason, we decided to adopt the LOGSPIRAL geometry as a baseline choice for our forecast presented in Section 4. Nevertheless, we pushed the comparison between the two geometries in the HGL area for sake of completeness.

In Fig. A2, we show the comparison between the *Planck* data for *Type 1* maps and the MCMOLE3D axisymmetric best-fitting model after the application of the HGL cut described in the paper. The AXISYMMETRIC model describes the data similarly to the LOGSPIRAL model at the larger scales. The difference in the signal amplitude is in fact less than 30 per cent for angular scales $\ell \lesssim 100$ and the two models are compatible within the error bars. This seems to indicate that in this regime, the details of the CO distribution in the HGL region are mainly affected by the properties of the vertical profile rather than by the geometry of the distribution. Conversely, the difference between the two geometries becomes important at smaller angular scales reaching a level of $\approx$2 at $\ell \approx 1000$.

We finally performed a series of polarized simulations as in Section 4.2 to access the level of contamination to the CMB

B-modes power spectrum with the best-fitting AXISYMMETRIC model and found $r_{CO} \lesssim 0.001$. Moreover, the slope of the BB power spectrum in Fig. A2(b) is -2.2 similar to the one computed with the LOGSPIRAL geometry.

Because the LOGSPIRAL model describes the data both in the HGL and low galactic latitude area, we consider the upper limit derived with this setup more reliable and the reference estimate for the contamination to the cosmological signal due to the CO polarized emission.

This paper has been typeset from a TEX/LATEX file prepared by the author.