EDGES-low GHA dependent data: cross-checks and null tests

Peter H. Sims
McGill Space Institute, McGill University

30th March 2021

Cosmic Dawn (CD), when the first stars and proto-galaxies began to form, is commonly expected to be accompanied by an absorption signature at radio frequencies. This feature arises as Lyman-α photons emitted by these first luminous objects couple the 21 cm excitation temperature of intergalactic hydrogen gas to its kinetic temperature, driving it into absorption relative to the CMB. In fiducial models of CD, significant spatial fuctuations in the 21 cm brightness temperature at CD, sourced from inhomogeneities in the density and Lyman-α fields, are limited to sub-degree angular scales; on larger scales, the sky-averaged 21 cm signal is well approximated as isotropic. The directivity pattern on the EDGES-low antenna has significant weight on angular scales of 10s of degrees. Correspondingly, an observable consequence of the isotropy of the signal on these scales is that one should expect the CD absorption trough to be independent of the Galactic Hour Angle (GHA) in the EDGES-low data. In contrast, both GHA-independent and GHA-dependent multiplicative systematic effects, arising due to imperfect calibration or instrumental modelling coupled to anisotropic foreground emission, are possible. As such, GHA-independence of estimates of the CD absorption trough provides a useful cross-check against contamination by unmodelled GHA-dependent systematic effects. In this memo, I examine the Galactic hour angle dependence of EDGES low-band data in relation to this cross-check.

1 Data

The data sets analysed in this memo derive from observations with EDGES low between 2016 day 252 and 2017 day 94. Two versions of the data, derived from independent data-processing pipelines (including calibration, RFI excision and GHA and frequency averaging), are considered:

- *Data processed by A. Rogers* - (i) 'spegha' tables of EDGES-low data (temperature spectra), with flags, as a function of UTC time. One table per 4 hour GHA block, for blocks centered on GHA = 0, 4, 8, 12, 16 and 20 hours. (ii) '4hour_bin' tables of 'spegha' data averaged over GHA. (iii) 'spegha4hrtable' table of estimated 21 cm absorption trough amplitude, SNR and sky temperature at 78 MHz as a function of GHA bin. Hereafter, I refer to these as the p1 data sets.

- *Data processed by R. Monsalve* - tables of EDGES-low data binned in GHA blocks of 4 and 6 hours, centered on GHA = 0, 4, 8, 12, 16 and 20 hours and 0, 6, 12 and 18 hours, respectively, and their respective weights. Hereafter, I refer to these as the p2 data sets.

Here, I restrict my attention to the three 4-hour GHA block data sets centered on GHA = 8, 12 and 16 hours, matching the GHA range of the data used in the primary analysis in B18. The GHA-average of the data in these blocks approximately matches that used to derive the quantities estimated in Extended Data Table 2 of Bowman et al. (2018) (hereafter, B18 EDT2).

2 Methodology

To provide some intuition for the impact of large scale isotropy of the 21 cm signal on recovered parameter estimates as a function of GHA, I start by constructing and analysing mock-data sets with a priori known foreground, 21 cm signal and noise components. The approach used to construct these data sets is described in Section 2.1. The models used to fit both the mock-data sets and the data are described in Section 2.2.

2.1 Data model & mock-data sets

I produce a mock-data set for each of the three 4-hour GHA block data sets centered on GHA = 8, 12 and 16 hours, each described by a data model of the form,

$$\boldsymbol{T}_{\rm ant}(\mathrm{GHA}) = \boldsymbol{T}_{21} + \boldsymbol{T}_{\mathbf{Fg}}(\mathrm{GHA}) + \boldsymbol{n}(\mathrm{GHA}) \ . \tag{1}$$

Here, $\boldsymbol{T}_{21}$ is a vectorised 21 cm absorption trough spectrum parametrised as a flattened Gaussian of the form,

$$T_{21}(\nu) = -A \left(\frac{1 - e^{-\tau e^{B}}}{1 - e^{-\tau}} \right) \ , \tag{2}$$

where,

$$B = \frac{4(\nu - \nu_0)^2}{w^2} \log \left[-\frac{1}{\tau} \log \left(\frac{1 + e^{-\tau}}{2} \right) \right] \ , \tag{3}$$

and A, ν_0, w and $\tau = 7$ describe the amplitude and central frequency, width and flattening of the absorption trough, respectively. $\boldsymbol{T}_{\mathbf{Fg}}(\mathrm{GHA})$ is a vectorised model for the foreground component of the data parametrised as a 'lin-log' polynomial of the form,

$$T_{\rm fg}(\mathrm{GHA}, \nu) = \sum_{i=0}^{n-1} a_i(\mathrm{GHA}) \left(\frac{\nu}{\nu_0} \right)^{i-2.5} \ , \tag{4}$$

which is proportional to the foreground model used for calculating the values quoted in B18 EDT2. $\boldsymbol{n}(\mathrm{GHA})$ is a vectorised model for the noise on the data, which, here, I approximate as Gaussian and white, mirroring the noise model used in B18 EDT2.

When constructing the mock data sets, I assume a 21 cm signal that is fixed between data sets, in the manner expected for a signal that is isotropic on the scales measured by EDGES. In each GHA data set, the signal is described by Equation 2 with parameters: $A = 0.53$ K, $\nu_0 = 78.1$ MHz, $w = 18.7$ MHz and $\tau = 7$, selected to match the 21 cm signal parameters inferred in B18. I construct the foreground model according to Equation 4, with $n = 6$ and with foreground coefficients, $\hat{a}_i$, given by the elements of the vector of maximum likelihood coefficients, $\hat{\boldsymbol{a}}(\mathrm{GHA}) = (\mathbf{Q}^T\mathbf{N}^{-1}\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{N}^{-1}\boldsymbol{d}(\mathrm{GHA})$, of the fit of Equation 4 to the data as a function of GHA. Here, $\mathbf{Q}$ is the design matrix of lin-log polynomials in Equation 4 and $\mathbf{N}$ is the covariance matrix of $\boldsymbol{n}(\mathrm{GHA})$. Finally, for simplicity, I assume a fixed RMS noise level of ~ 45 mK in each of the GHA data sets. In practice, there is an $\sim 20\%$ increase in the foreground brightness temperature averaged over 4 h GHA blocks between the GHA block centered on 8 and 16 hours; however, since the expected signal is an order of magnitude larger than the assumed noise, with or without this correction, the impact of this assumption will be small[1].

[1]The impact of deviations from this assumption of uniformity of the noise in the GHA-averaged EDGES-low data has been shown to be of greater significance (Sims & Pober 2020); however, for direct comparison with B18 EDT2, here, I restrict consideration to models with uniform noise.

Parameter	Distribution	Prior
A	uniform	$U(0, 4)$ K
ν_0	uniform	$U(60, 90)$ MHz
w	uniform	$U(5, 40)$ MHz
τ	uniform	$U(0, 40)$

Table 1: Default priors on flattened Gaussian 21 cm signal model.

2.2 Generalised GHA-dependence cross-check

Under the assumption that Equation 1 accurately describes the EDGES-low data, one should expect to recover consistent estimates of the CD absorption trough as a function of GHA. In contrast, inconsistency between 21 cm signal estimates as a function of GHA implies the presence of a GHA-dependent bias in the signal estimates, resulting from structure in the data that is not modelled by $T_{\rm fg}(\mathrm{GHA}, \nu)$.

In Section 3.1, I validate this assertion using fits to the mock data described in Section 2.1, before examining the GHA-dependence of the EDGES-low data in Section 3.3. In both cases, I assume that the probability of the data given Equation 1 is described by a Gaussian likelihood function of the form,

$$\mathcal{L}(\mathbf{\Theta}) \propto \frac{1}{\sqrt{\det(\mathbf{N})}} \exp\left[-\frac{1}{2}\left(\boldsymbol{d} - \boldsymbol{T}_{\rm ant}\right)^{\dagger} \mathbf{N}^{-1} \left(\boldsymbol{d} - \boldsymbol{T}_{\rm ant}\right)\right], \tag{5}$$

where $\boldsymbol{T}_{\rm ant}$ is the model for the signal and I model the noise on the spectrum as an uncorrelated Gaussian random field, with covariance matrix $\mathbf{N}$. The elements of the covariance matrix are given by $\mathbf{N}_{ij} = \left\langle n_i n_j^* \right\rangle = \delta_{ij}\sigma_j^2$, where $\langle ... \rangle$ represents the expectation value and $\sigma_j = 45$ mK denotes the noise expectation in element j of the spectrum. The priors that I place on the parameters of the flattened Gaussian 21 cm signal model when fitting the data are as listed in Table 1. These prior distributions are the used for all fits including the flattened Gaussian 21 cm signal model except where otherwise stated in the text.

3 Results

3.1 Parameter recovery from mock-data sets

In Section 3.1.1, I examine the effect of adding unnecessary complexity to a foreground model that is already sufficiently complex to describe the data. In Section 3.1.2, I examine the prior dependence of recovered 21 cm parameter estimates when they are correlated in the manner seen in the EDGES low-band parameter estimates in B18 Extended Data Figure 10.

3.1.1 Foreground model complexity

When fitting Equation 5, I consider both $n = 6$ and $n = 7$ variants of Equation 4 for the purposes of constraining the 21-cm signal. This enables validation of the expectation that, for a sky temperature that is accurately described by Equation 1 with an $n = 6$ foreground, if the $\mathbf{Q}_{\rm n}$ basis for a more complex foreground model includes $\mathbf{Q}_6$, one should expect to recover 21 cm signal estimates consistent with those obtained with $\mathbf{Q}_6$, within their respective uncertainties.

The posterior probability distributions of the parameters of the flattened Gaussian 21 cm signal model and a 6 term lin-log foreground model, jointly estimated from the $6 \geq \mathrm{GHA} > 10$

(grey), $10 \geq \mathrm{GHA} > 14$ (red), $14 \geq \mathrm{GHA} > 18$ (blue) EDGES low band data sets, are shown in Figure 1 and Figure 2, respectively. The posterior probability distributions of the parameters of the corresponding fit of the 21 cm signal and a 7 term lin-log foreground model are shown in Figure 3 and Figure 4, respectively. The recovered parameters are consistent with one another, as a function of GHA, and with the input 21 cm signal parameters in the mock data set. Furthermore, the highest order term in the $n = 7$ data set is consistent with zero and simply increases the correlation between, and uncertainties on, the 21 cm signal parameters.

The unnecessary increase in complexity of the foreground model between $n = 6$ and $n = 7$ is also evident if one compares the Bayesian evidence of the models,

$$\mathcal{Z} = \int \mathcal{L}(\Theta)\pi(\Theta)\mathrm{d}^n\Theta \,, \tag{6}$$

where $\pi(\Theta)$ are the priors on the model parameters Θ. For uniform priors on the parameters of the model over the ranges illustrated in Figure 1 and Figure 2, and assuming that the $n = 6$ and $n = 7$ foreground models are a priori equally likely, the log-Bayes factor between the models is $\ln(B_{6,7}) = \ln(\mathcal{Z}_6) - \ln(\mathcal{Z}_7) = 9.6$, corresponding to decisive evidence in favour of the model including the $n = 6$, rather than the $n = 7$, foreground model as a description of the mock data set (e.g. Kass & Raftery 1995).

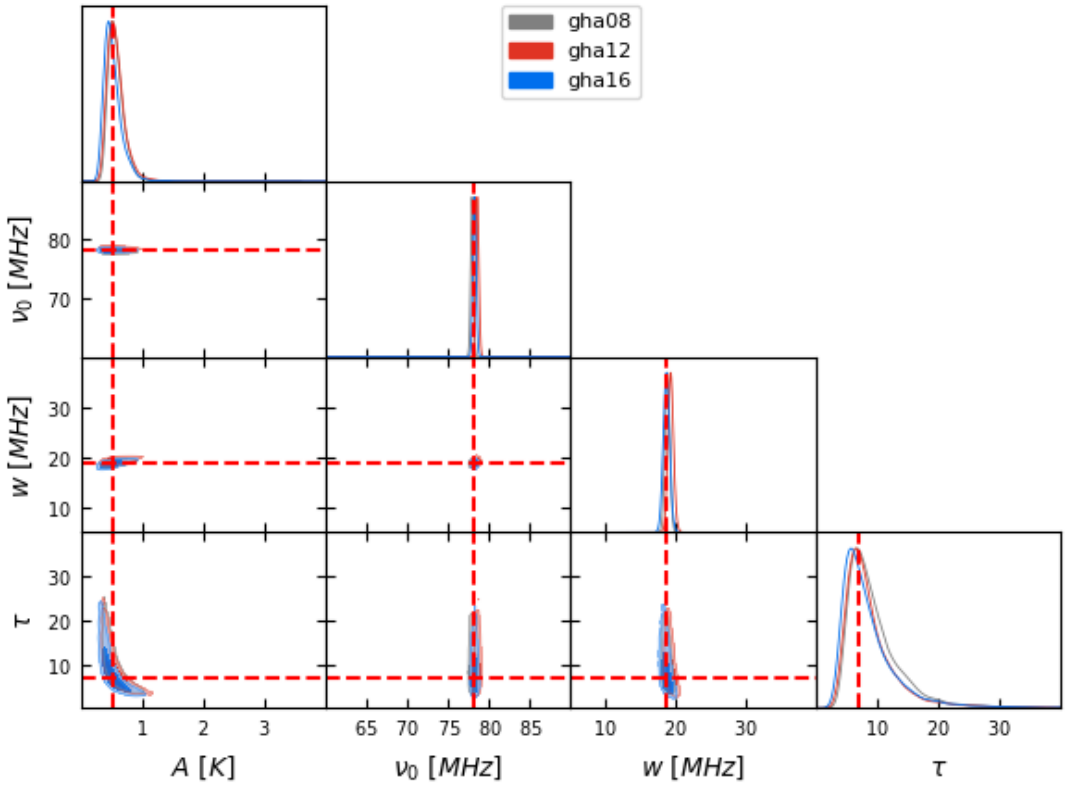

Figure 1: Posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 term lin-log foreground model (Equation 4), from three EDGES low band mock-data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red), $14 \geq \mathrm{GHA} > 18$ (blue).

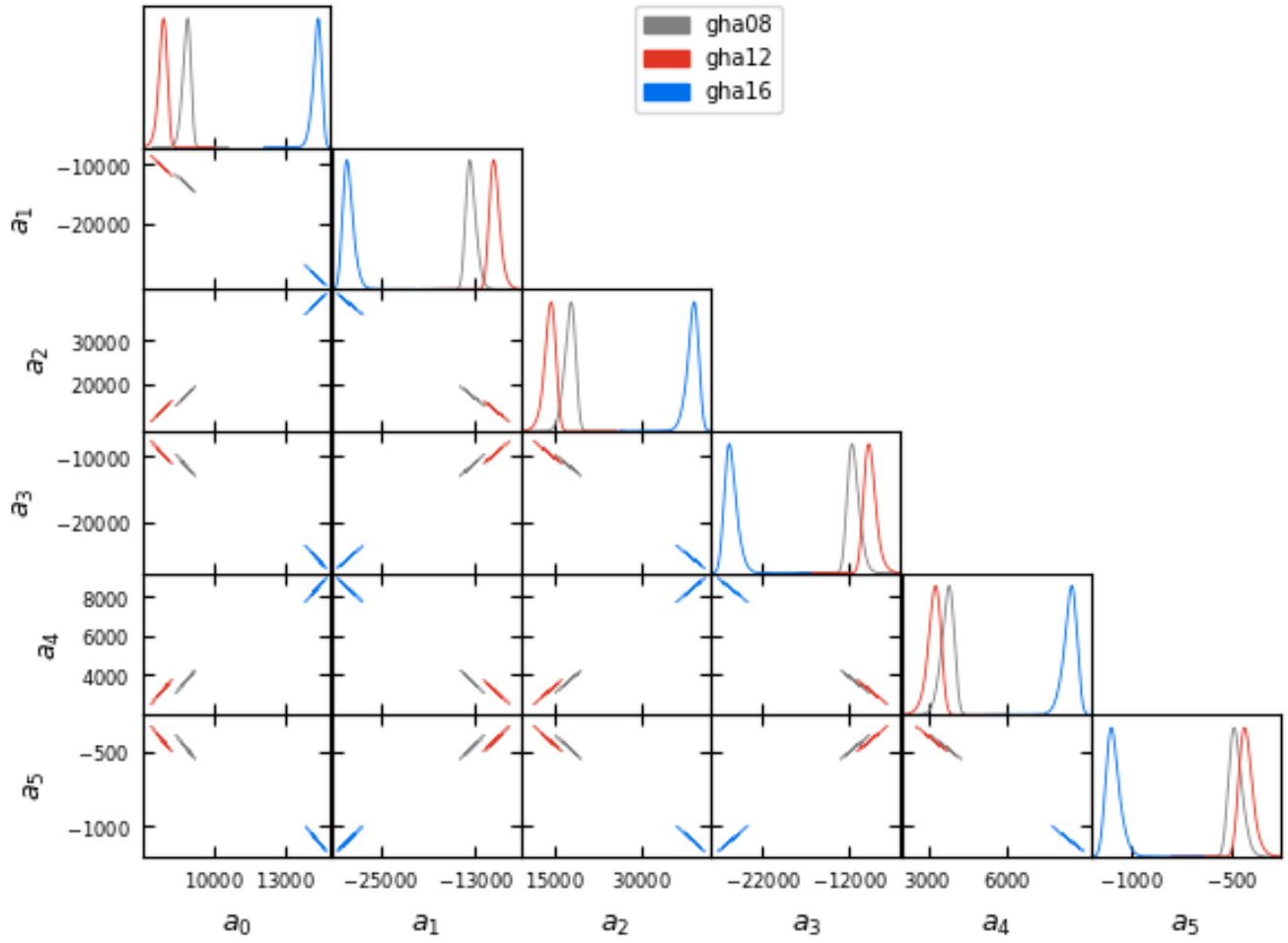

Figure 2: Posterior probability distributions of the parameters of a 6 term lin-log foreground model (Equation 4), when jointly estimated with a flattened Gaussian 21 cm signal model, from three EDGES low band mock-data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red), $14 \geq \mathrm{GHA} > 18$ (blue).

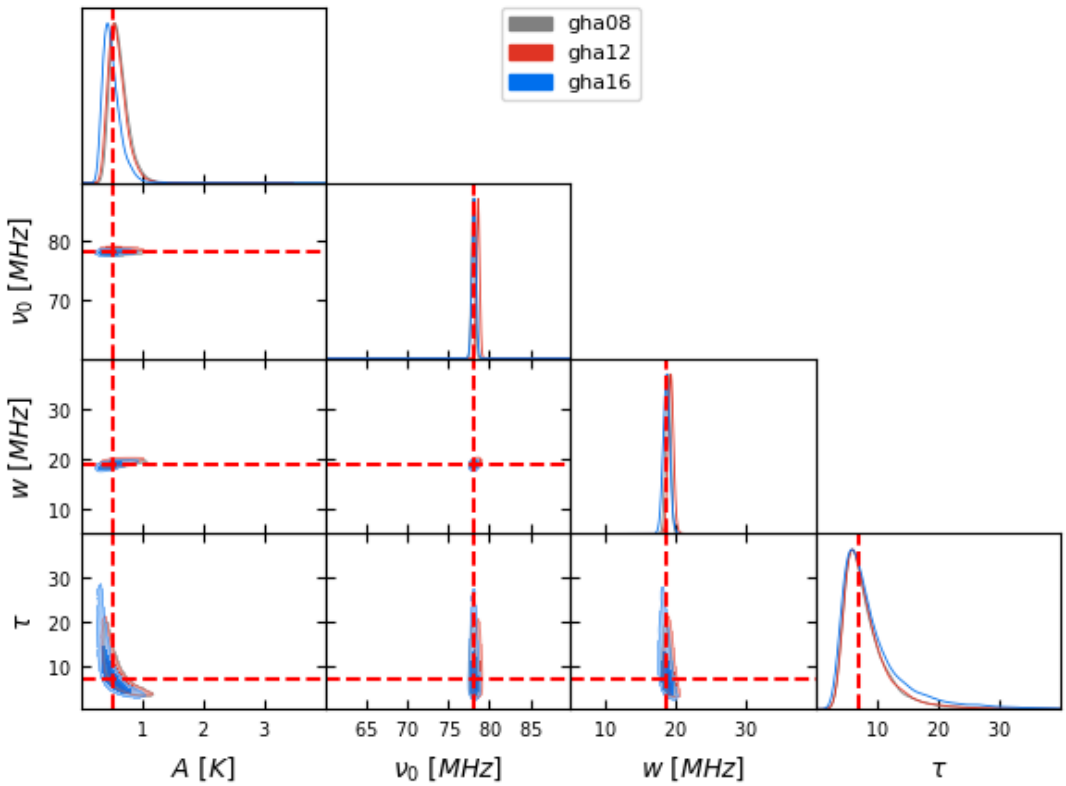

Figure 3: Posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 term lin-log foreground model (Equation 4), from three EDGES low band mock-data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red), $14 \geq \mathrm{GHA} > 18$ (blue).

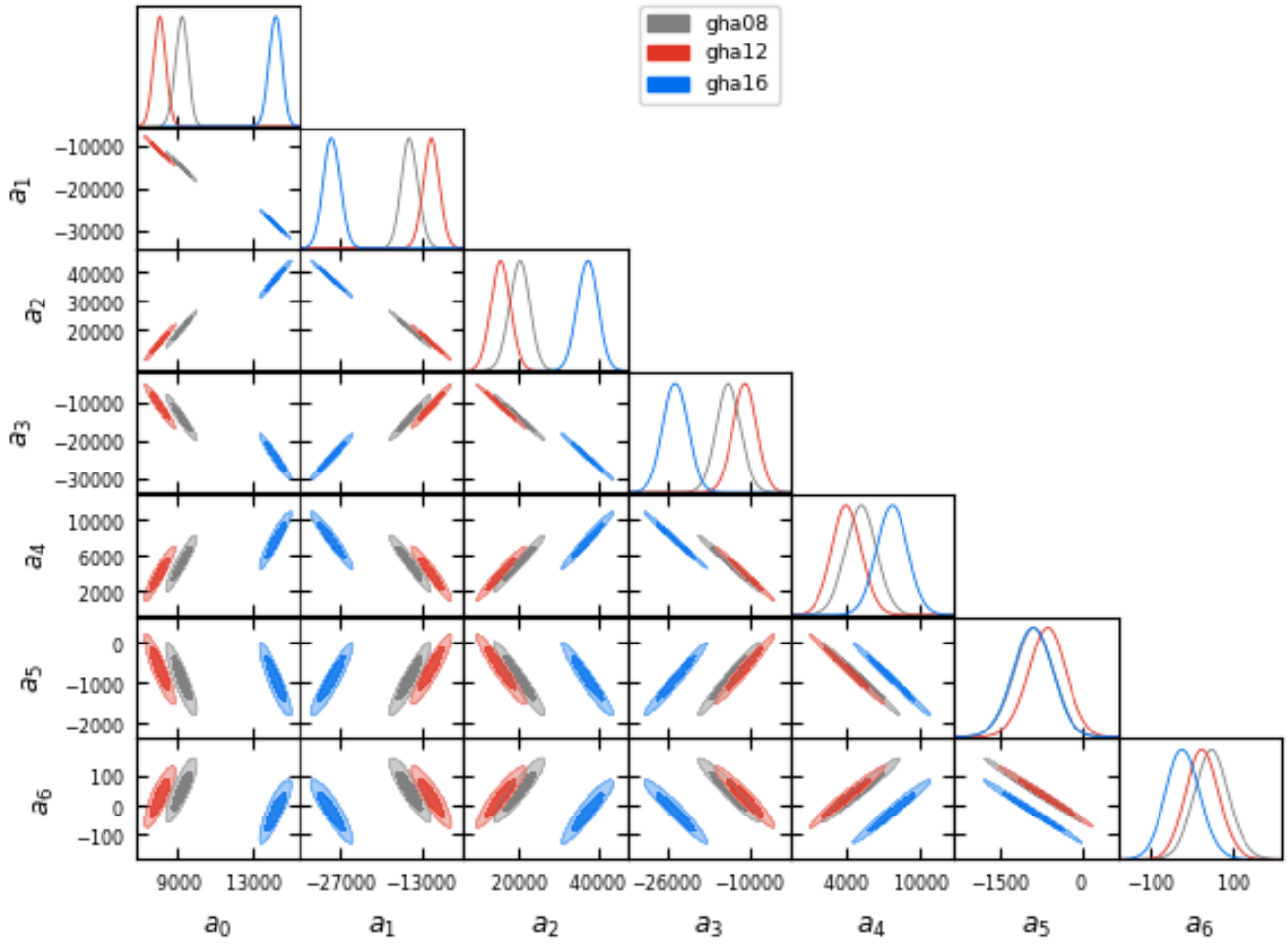

Figure 4: Posterior probability distributions of the parameters of a 6 term lin-log foreground model (Equation 4), when jointly estimated with a flattened Gaussian 21 cm signal model, from three EDGES low band mock-data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red), $14 \geq \mathrm{GHA} > 18$ (blue).

3.1.2 Prior-dependence of 21 cm parameters

As seen in Figures 1 and 3, when fitting the mock data sets, the amplitude and flattening of a flattened Gaussian model for the 21 cm signal are degenerate. This degeneracy implies that the recovered estimates of A and τ will be sensitive to their respective priors. This is illustrated in Figure 5 where the posteriors for the 21 cm signal estimates are plotted for particular choices of restrictive priors on τ: $0 < \tau < 4.5$ in grey, $4.5 < \tau < 12$ in red[2], $20 < \tau < 40$ in blue. A degeneracy of this sort requires a careful choice of priors on the degenerate parameters. Ideally, this choice would be motivated by the physics of the problem, but in the absence of theoretical priors, broad uninformative priors that avoid artificially restricting the recovered estimates of A through a restrictive choice of prior on τ, and vice versa, are a conservative choice.

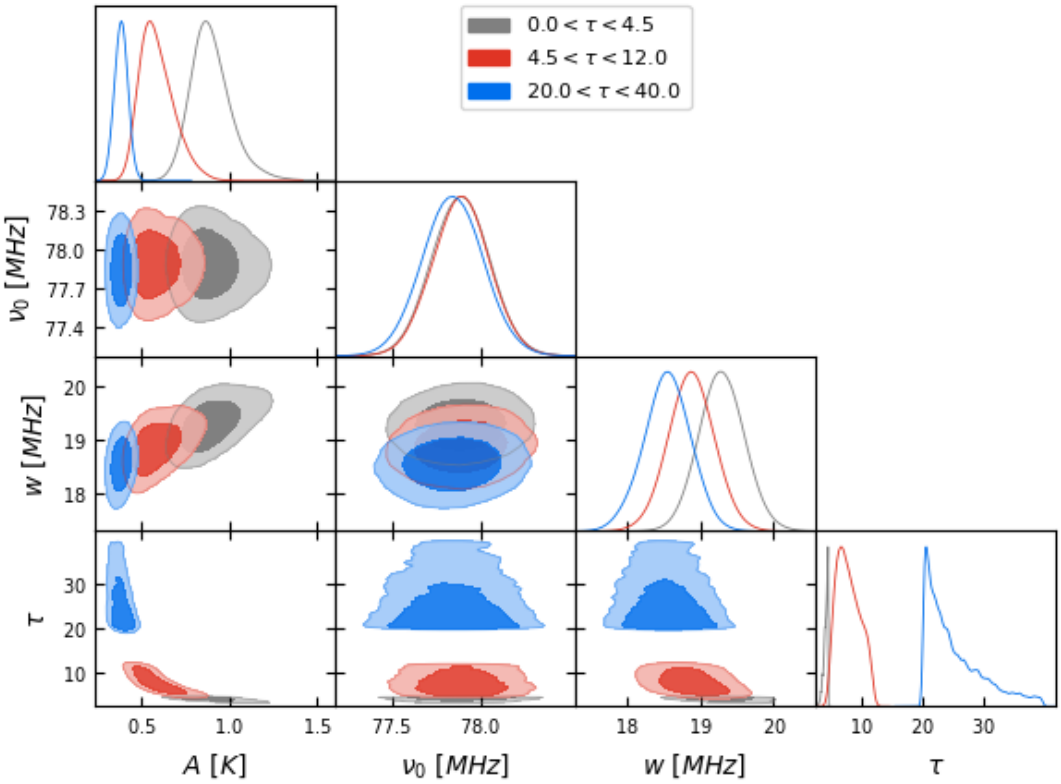

Figure 5: Posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 term lin-log foreground model (Equation 4), from a single EDGES low band mock-data set, for three choices of uniform priors on τ: $0 < \tau < 4.5$ (grey), $4.5 < \tau < 12$ (red), $20 < \tau < 40$ (blue). The degeneracy between A and τ, seen in e.g. Figure 1, means the posterior for A is a function of the prior on τ (and vice versa).

[2]The intermediate prior in τ, in red, has been selected to match the range considered in EDGES MIT-series memo 272: `https://www.haystack.mit.edu/wp-content/uploads/2020/07/memo_EDGES_272.pdf`

3.2 Signal-recovery null-test

3.2.1 Foreground-only mock-data set

In this section I start by repeating the analysis described in Section 3.1.1, but now using mock-data sets consisting of only foreground components. I then fit each mock-data set using: (i) a model consisting exclusively of a lin-log foreground model (Equation 4), with either $n = 6$ or $n = 7$ or (ii) a consisting of one of the foreground models described in i) plus a flattened Gaussian 21 cm signal model (Equation 2).

Comparing the Bayesian evidence of the models when fit to the foreground-only mock-data set, I find the log-Bayes factor of 3.8 and 4.4, respectively, between the model including an $n = 6$ or $n = 7$ lin-log foreground model, versus one including both an with $n = 6$ or $n = 7$ lin-log foreground and flattened Gaussian 21 cm signal model. In both case, this corresponds to strong evidence in favour of the foreground-only model (i.e. no detection of the 21 cm signal) as expected.

3.2.2 Foreground plus B18 21 cm signal mock-data set

Next, I repeat the test described in Section 3.2.1 on the mock-data sets analysed in Section 3.1.1, which include the absorption trough recovered in B18. In this case the Bayesian evidence decisively favours the models including the flattened Gaussian 21 cm signal over those without it, with a log-Bayes factor of greater than 10 in each case.

3.2.3 GHA = 8, 12 and 16 hour AR EDGES low band data sets

Repeating the test described in Section 3.2.2 but now on each of the GHA = 8, 12 and 16 hour AR EDGES low band data sets, I again find that the Bayesian evidence decisively favours the models including the flattened Gaussian 21 cm signal over those including only a foreground model, with a log-Bayes factor of greater than 10 in each case.

3.3 EDGES-low GHA parameter dependence

3.3.1 21 cm signal amplitude estimates when fixing the shape of the signal and fitting to a sub-band of the data

In this Section, I analyse GHA binned EDGES low data similar (up to pre-processing differences; see Section 3.3.4) to that which was analysed and reported on in a test of foreground-temperature independence of the recovered 21 cm absorption trough in B18 EDT2. In B18, this test is carried out by fitting for the maximum likelihood amplitude of a flattened Gaussian absorption trough, whose shape parameters are fixed to match the shape of the 21 cm signal recovered from the GHA-averaged data, jointly with a 6 term lin-log model for the foregrounds (Equation 4), over a $65 < \nu < 95$ MHz sub-band of the EDGES low data. In Table 2, I list the maximum likelihood values for the 21 cm absorption trough, in 4-hour GHA bins centered on 8, 12 and 16 hours, reported (i) in B18 EDT2 and (ii) from analysis of the data set under consideration in this memo. Table 2 also includes 1-sigma uncertainties associated with the posterior probability distribution for the amplitude of the 21 cm signal derived via an analogous analysis to that described in Section 3.1, but with the additional restrictions described above on the shape of the 21 cm signal and frequency range analysed. The mean of the posteriors for the parameters are consistent with the estimates in B18 EDT2 to within their 1-sigma uncertainties in of the p1 data sets and in all but the 16 h GHA p2 data set, which is consistent at 2-sigma. The amplitudes recovered from

the p1 data sets are also systematically higher than the p2 data sets (at the $\sim$ 1-sigma level in the GHA=8 and 12 h data sets and at $\sim$ 3-sigma in the GHA=16 h data set), perhaps due to differences in processing of the raw data.

			Sub-band estimates			
			p1 data sets		p2 data sets	
Parameter	GHA	EDT2 (K)	Recovered mean (K)	1-sigma uncertainty (K)	Recovered mean (K)	1-sigma uncertainty (K)
A	08	0.44	0.49	0.08	0.37	0.08
	12	0.57	0.59	0.08	0.49	0.08
	16	0.59	0.67	0.08	0.44	0.08

Table 2: Mean and 1-sigma uncertainties on the amplitude of a flattened Gaussian 21 cm signal model with fixed central frequency, width and flattening factor ($\nu_0 = 78.1$ MHz, $w = 18.7$ MHz, $\tau = 7$), when jointly estimated with a 6 term lin-log foreground model (Equation 4), from the GHA = 8, 12 and 16 hour EDGES low band data sets, over the restricted frequency range $65 < \nu < 95$ MHz.

3.3.2 GHA dependence of the amplitude and shape of the 21 cm signal fitting across the full EDGES-low band

As discussed in the introduction to this memo, a predicted consequence of the isotropy of the 21 cm signal on large spatial scales is that one should expect the CD absorption trough to be independent of GHA in EDGES-low data. In contrast, both GHA-independent and GHA-dependent multiplicative systematic effects, arising due to imperfect calibration or instrumental modelling coupled to anisotropic foreground emission, are possible.

In this section, I build on and generalise the GHA dependence of the amplitude of the 21 cm signal test described in Section 3.3.1, in order to assess the level of consistency in recovered estimates of the shape and amplitude of the 21 cm signal as a function of GHA. For the p1 data sets, this is done over the $65 < \nu < 95$ MHz sub-band, in which the processed data has non-zero weight. For the p2 data sets, this is done both over the $65 < \nu < 95$ MHz sub-band, for direct comparison with the p1 data sets and over the full $50 < \nu < 100$ MHz EDGES-low spectral band in which it has non-zero weight (the same band used to derive the 21 cm signal parameters in the analysis of GHA-averaged data in B18).

Figures 6 and 7 show the posterior probability distributions for the flattened Gaussian 21 cm signal parameters when jointly fitting the data with an $n = 6$ and $n = 7$ lin-log foreground model, respectively. In both figures, the posteriors derived from the p1 and p2 data sets are shown in the top and bottom subplots, respectively. In all cases, the significant correlation between the foreground and 21 cm signal model when fitting to data restricted to the $65 < \nu < 95$ MHz sub-band, results in complex multi-modal posterior probability distributions. Differences are also apparent in the posteriors recovered from the p1 and p2 data sets. The corresponding posteriors for the foreground models parameters are shown in Figures 20 – 23 in Appendix C.

The complexity of the posteriors and differences between the constraints derived from the p1 and p2 data sets make assessing GHA-independence of the signal when fitting to data restricted to the $65 < \nu < 95$ MHz sub-band challenging. Additional data has the potential to break the degeneracy between models[3]; thus, next, I turn to analysis of the full-band p2 data sets.

[3]Also, see Appendix B for an overview of the impact of having greater a priori knowledge of the signal, relevant if, for example, such constraints were available from external data sets.

Figure 8 shows the posterior probability distributions for the flattened Gaussian 21 cm signal parameters when jointly fitting the p2 data sets with an $n = 6$ (top) and $n = 7$ (bottom) lin-log foreground model, respectively, over the full $50 < \nu < 100$ MHz EDGES low spectral band. The corresponding posteriors for the parameters of the $n = 6$ and $n = 7$ lin-log foreground models are shown in Figures 24 and 25, respectively.

For a uniform prior $U(0, 40)$ on the flattening factor, τ, of the flattened Gaussian 21 cm, used up to now in this memo, the posterior probability distribution for the flattening factor, τ, of the flattened Gaussian 21 cm signal is prior limited when analysing this data set (see Appendix A). Thus, when fitting the full $50 < \nu < 100$ MHz EDGES low spectral band, I broaden the uniform prior on τ to $U(0, 1000)$.

Comparing the Bayesian evidence for the two models incorporating $n = 6$ and $n = 7$ lin-log foreground components, yields a log-Bayes factors greater than 5 between models for all of the GHA data sets analysed. This corresponds to strong evidence in preference of the model incorporating the $n = 6$, over $n = 7$, lin-log foreground components. This is also reflected in the strong consistency between the 21 cm signal posteriors that are recovered between joint fits with the $n = 6$ or $n = 7$ lin-log foreground models in Figure 8, which indicates that the added complexity of the $n = 7$ lin-log foreground model is redundant when fitting these data sets.

Comparing Figure 8 to Figures 6 and 7 it is evident that the additional data available when fitting to the full EDGES low band plays an important role in reducing degeneracy between regions of the parameter space of the models being fit, with uni-modal posterior solutions recovered when fitting over the full spectral band. Relatively consistent posteriors are recovered for the GHA=08 and 12 hours data sets; however, the solutions are inconsistent with those recovered at GHA=16 hours where a deeper feature at a lower frequency is preferred. The recovered estimates in each of the data sets are not in agreement with the estimates recovered from the GHA-averaged data in B18. This inconsistency may be due to updates to the pre-processing (e.g. calibration, RFI excision, LST and frequency averaging) of the data, relative to that used for the data analysed in B18. The specifics of this updated pre-processing and its impact on the recovered signal requires additional consideration but is not discussed further here.

The mean 21 cm signal parameter and their 1-sigma uncertainties corresponding to the posteriors in Figure 8 are listed in Table 3 for the analysis using the $n = 6$ lin-log foreground model. The parameters with the $n = 7$ lin-log foreground model aren't listed but are consistent with the values quoted for the $n = 6$ lin-log foreground model. In the GHA=8 and 12 hour data sets, the amplitude of the recovered signals exceed, by 2 and 3-sigma, respectively, the standard cosmological expectations for the absorption depth, which predicts a maximum depth[4] of ~ 200 mK. Additionally, the signal is highly flattened, with flattening factors more than an order of magnitude larger than found in B18 being preferred. This may be suggestive of sharp residual systematic structure, such as RFI, in the GHA-dependent data that may have been excised in the GHA averaged data set analysed in B18.

In the GHA=16 hour data set, the amplitude of the recovered signals exceed, the standard cosmological expectations for the absorption depth by ~ 10-sigma. The signal is smooth and narrow, with a flattening parameter of 1.9 ± 0.8 MHz and width of 5.7 ± 0.3 MHz.

Figures 9 and 10 shows the maximum a posteriori (MAP) flattened Gaussian signal components $(d - m_n)$ residuals, with $m_n = T_{\rm Fg,n} + T_{21}$ and $T_{\rm Fg,n} + T_{21}$ a n-term linlog polynomial foreground model with $n = 6$ (left) and $n = 7$ (right), in fits to GHA = 8, 12 and 16 hour

[4]This model assumes adiabatic cooling of the hydrogen kinetic temperature between $z \approx 150$ and CD, no reheating of the gas prior to the Wouthuysen–Field coupling of the spin and kinetic temperatures of the gas and a background radiation temperature dominated by the CMB at CD (e.g. Barkana 2018).

		Full band estimates 6 term lin-log	
Parameter	GHA	Recovered mean (K)	1-sigma uncertainty (K)
A	08	0.31	0.05
	12	0.31	0.04
	16	0.51	0.03
ν_0	08	65.2	0.2
	12	65.7	0.2
	16	60.0	0.2
w	08	17.3	0.3
	12	19.4	0.4
	16	5.7	0.3
τ	08	165	197
	12	278	254
	16	1.9	0.8

Table 3: Mean and 1-sigma uncertainties on the parameters of a flattened Gaussian 21 cm signal model (Equation 2), when jointly estimated with either a 6 term lin-log foreground model (Equation 4), from the GHA = 8, 12 and 16 hour p2 EDGES low band data sets.

GHA	RMS$(d - m_6)$ [mK]	RMS$(d - m_7)$ [mK]
08	44.6	43.3
12	48.3	48.3
16	68.8	68.8

Table 4: RMS residuals of maximum a posteriori joint fits, to GHA = 8, 12 and 16 hour AR EDGES low band data sets, of a flattened Gaussian absorption trough and a 6 or 7 term linlog polynomial foreground model: RMS$(d - m_n)$, with $m_n = T_{\mathrm{Fg,n}} + T_{21}$ and $n = 6$ and $n = 7$, respectively.

p2 EDGES low band data sets. The RMS of the residuals of the fits are listed in Table 4. Structure inconsistent with noise is evident at the low end of the band (in the spectral range $50 < \nu \lesssim 55\,\mathrm{MHz}$) in all of the data sets analysed with both of the foreground models considered. This structure is not evident in fits to the publicly available EDGES low-band data, suggesting differences in processing of the raw data as a possible cause.

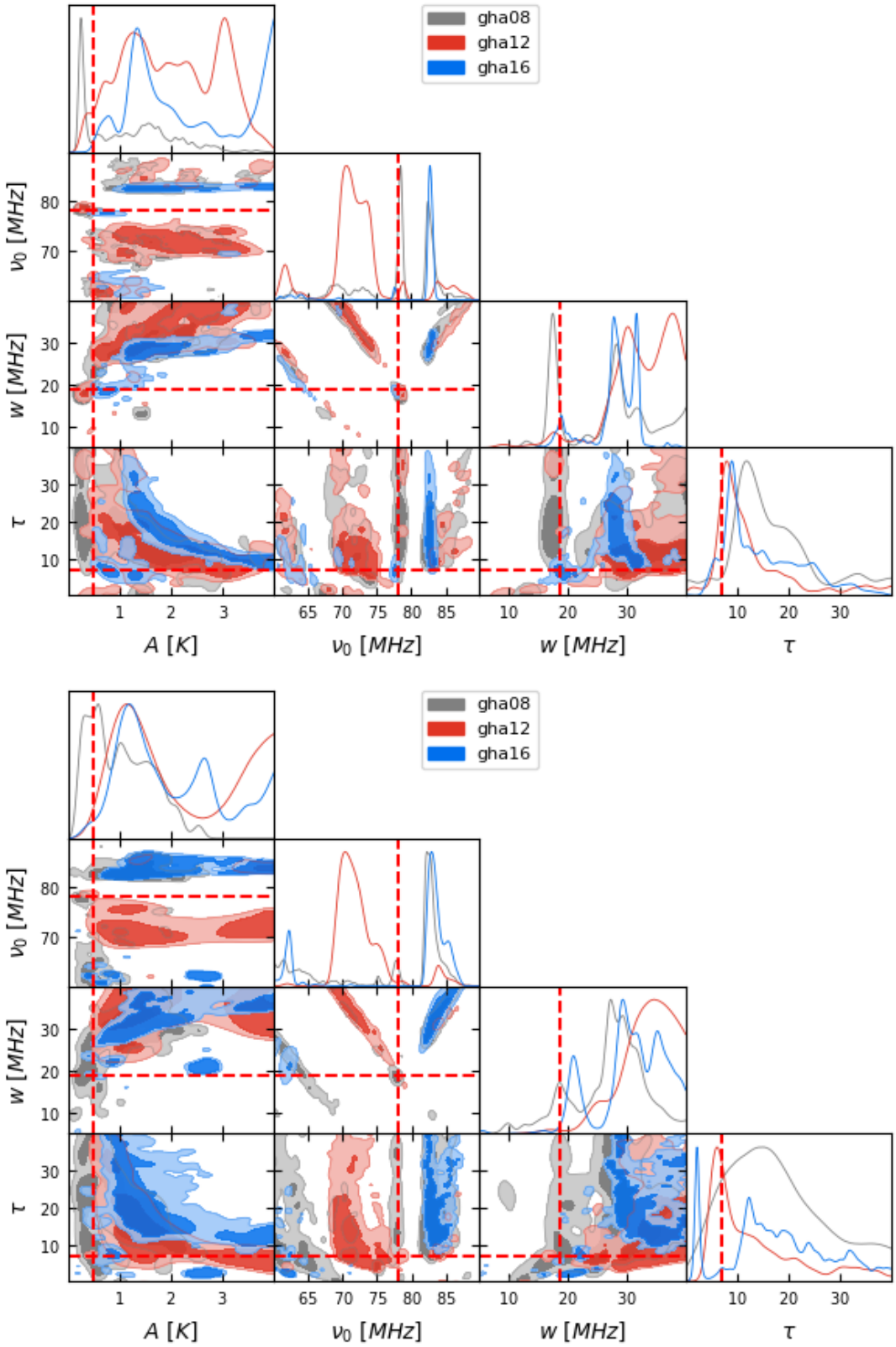

Figure 6: Posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 term lin-log foreground model (Equation 4), from EDGES low band data sets derived from the raw data using the p1 and p2 (see Section 1) processing pipelines (top and bottom, respectively). In each case, the three data sets analysed have Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red) and $14 \geq \mathrm{GHA} > 18$ (blue) and are fit over the restricted frequency range $65 < \nu < 95$ MHz.

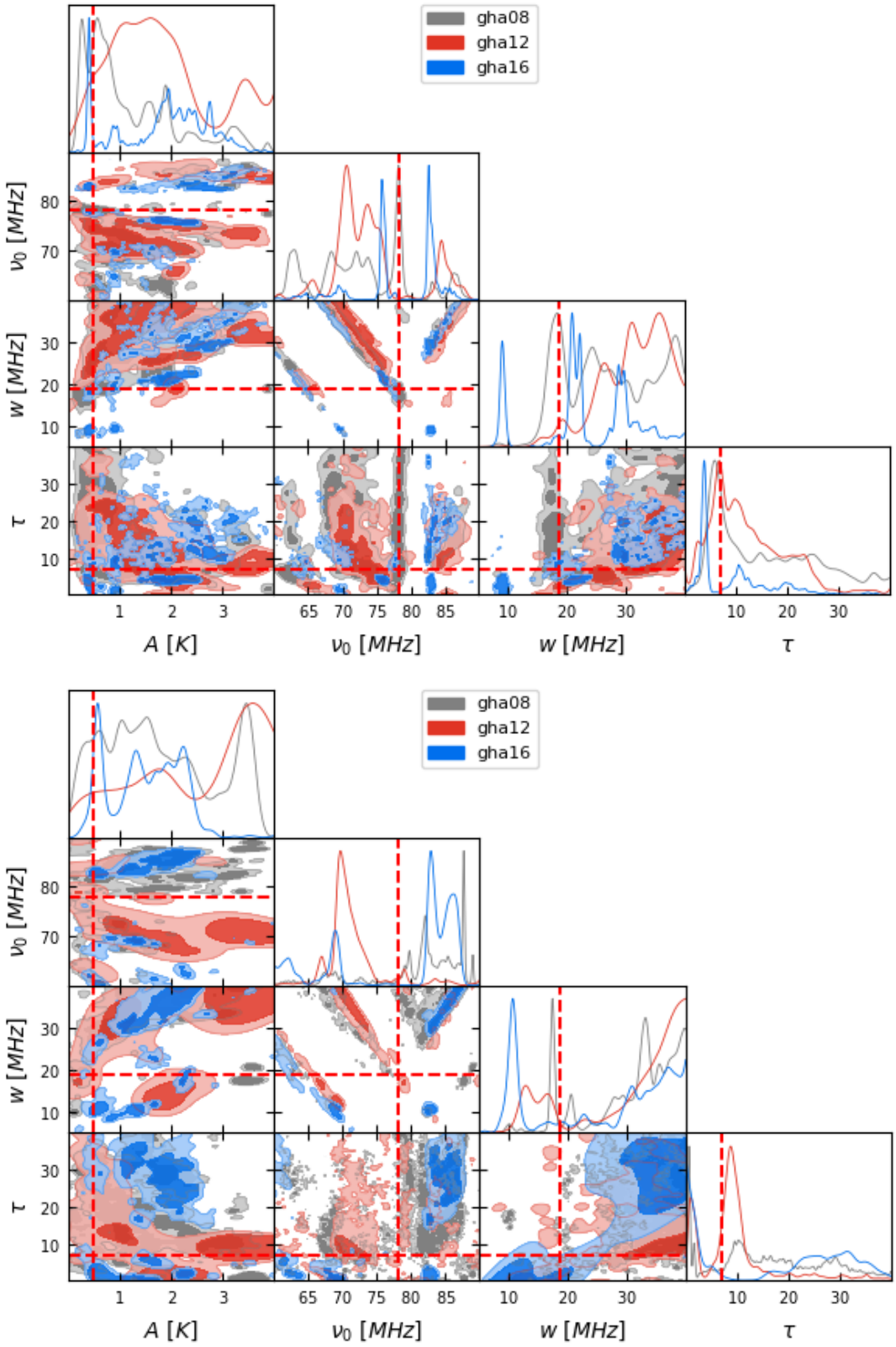

Figure 7: As in Figure 6 but for a 7 term lin-log foreground model.

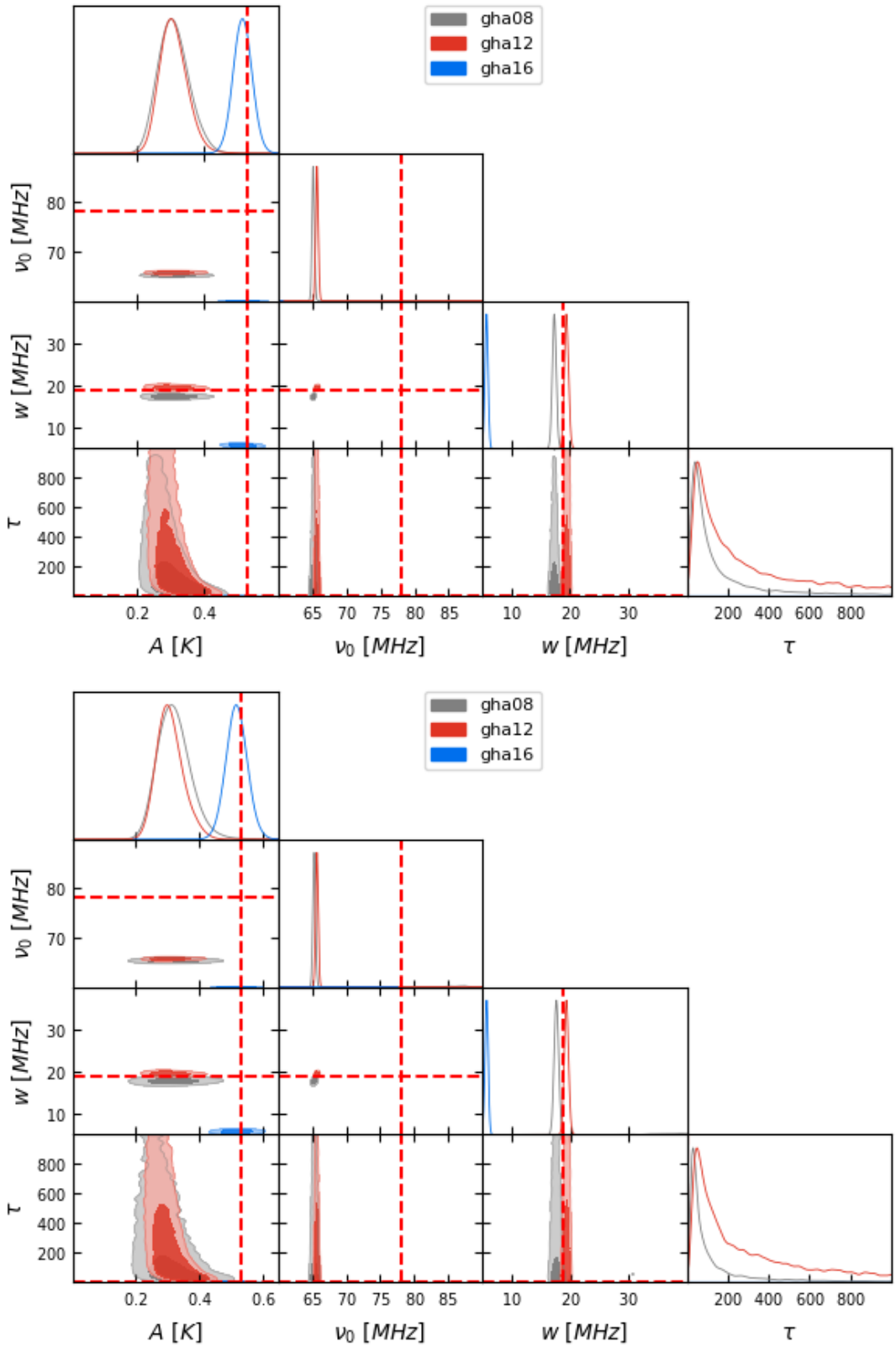

Figure 8: Posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 (top) or 7 (bottom) term lin-log foreground model (Equation 4), from the p2 data sets over the full-band frequency range $50 < \nu < 100$ MHz. In each case, the three data sets analysed have Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red) and $14 \geq \mathrm{GHA} > 18$ (blue).

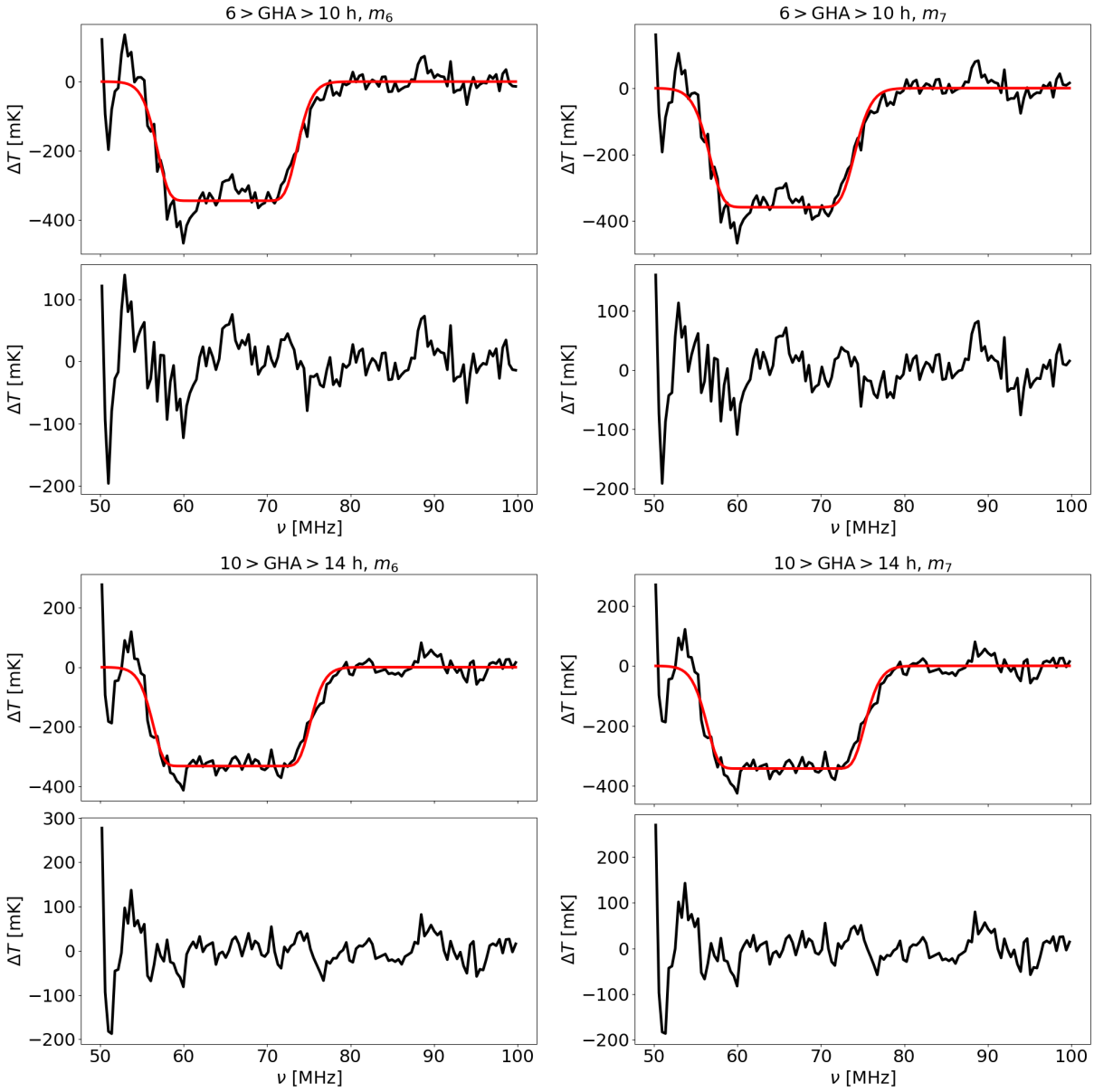

Figure 9: MAP flattened Gaussian signal components overlaid on $(d - T_{\rm Fg})$ residuals (top subfigures) and $(d - m_n)$ residuals, with $m_n = T_{\rm Fg,n} + T_{21}$ and $T_{\rm Fg,n} + T_{21}$ a n-term linlog polynomial foreground model with $n = 6$ (left) and $n = 7$ (right), in fits to GHA = 8 (top row) and 12 (bottom row) p2 EDGES low band data sets. The RMS of the residuals are listed in Table 4.

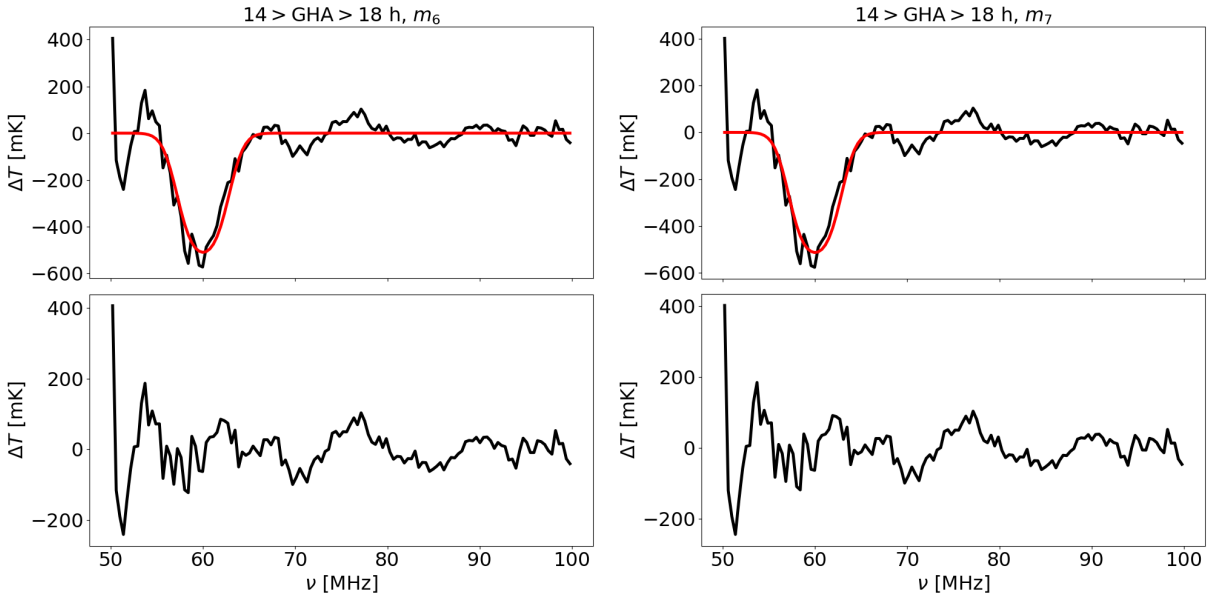

Figure 10: As in Figure 9 but for a fit to the GHA = 16 hour p2 EDGES low band data set. The RMS of the residuals are listed in Table 4.

3.3.3 GHA dependence of the amplitude and shape of the 21 cm signal fitting across alternate sub-bands

Fitting to the full-band data provides constraints, on the 21 cm signal, that are most directly comparable with those obtained in B18. However, in the absence of systematic-free data at the low end of the band, in the data sets under consideration here, we also perform analyses of sub-bands of the data that excludes this systematics-contaminated region. When selecting the specific sub-band to analyse a compromise must be found between between systematic suppression and use of a band that is larger than the $65 < \nu < 95$ MHz sub-band, used for the analysis shown in Figures 6 and 7, in order to reduce model degeneracy.

Figures 11 and 12 show the posterior probability distributions for the flattened Gaussian 21 cm signal parameters when jointly fitting $55 < \nu < 100$ MHz and $60 < \nu < 100$ MHz sub-band of the p2 data sets, respectively, in each case with $n = 6$ (top sub-figures) and $n = 7$ (bottom sub-figures) lin-log foreground models.

Figures 13 – 14 shows the MAP flattened Gaussian signal components $(d - m_n)$ residuals, with $m_n = T_{\rm Fg,n} + T_{21}$ and $T_{\rm Fg,n} + T_{21}$ a n-term linlog polynomial foreground model with $n = 6$ (left) and $n = 7$ (right), in fits to $55 < \nu < 100$ MHz sub-bands of GHA = 8, 12 and 16 hour p2 EDGES low band data sets. Figures 15 – 16 shows the corresponding MAP flattened Gaussian signals in fits to $60 < \nu < 100$ MHz sub-bands of the data sets. Any residual systematic effects that are present are less visually apparent in the residuals of the MAP fits to either the $55 < \nu < 100$ MHz or $60 < \nu < 100$ MHz sub-bands of the data sets. There is a reasonable level of consistency between the MAP signal estimates in the two sub-bands; however, the MAP signals are not consistent between the foreground models or as a function of GHA. Similar conclusions can be drawn from the posteriors for the parameters of the signals in the various cases considered. In addition, the posteriors on the parameters also illustrate the increasing degeneracy between probable signal models when fitting to the narrower sub-band of the data with the higher complexity foreground model.

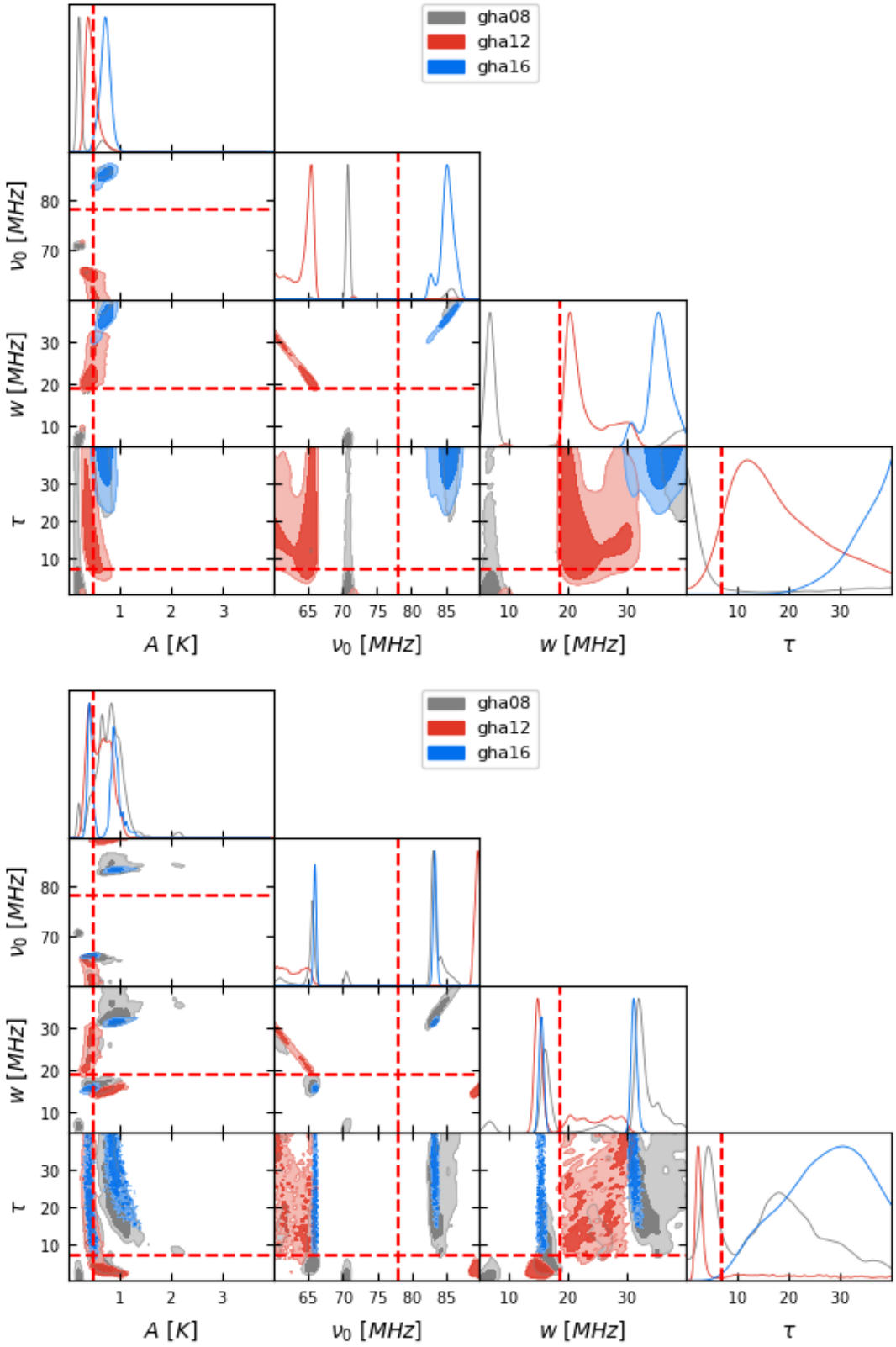

Figure 11: Posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 (top) or 7 (bottom) term lin-log foreground model (Equation 4), from the p2 data sets over the sub-band frequency range $55 < \nu < 100$ MHz. In each case, the three data sets analysed have Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red) and $14 \geq \mathrm{GHA} > 18$ (blue).

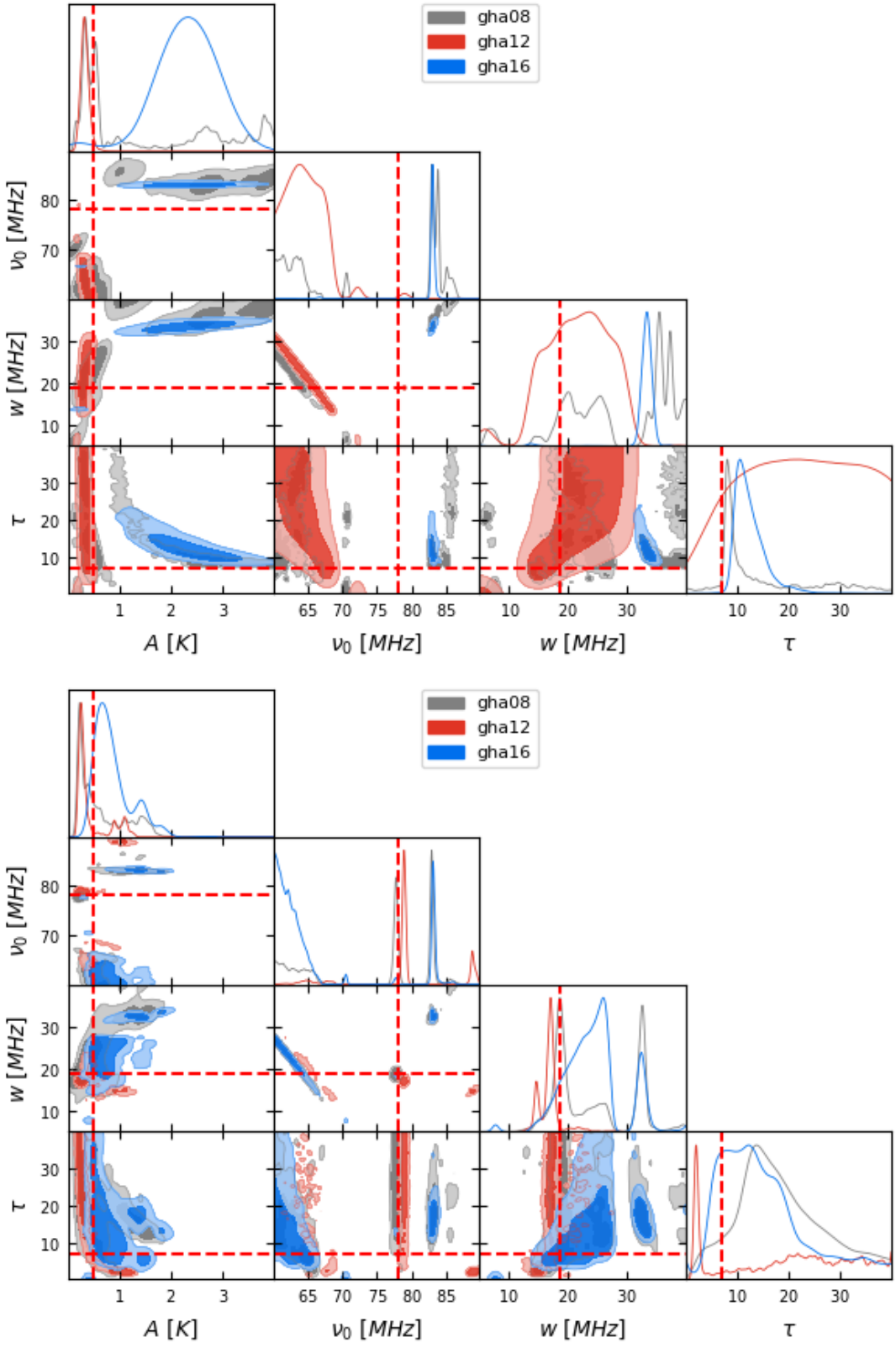

Figure 12: As in Figure 11 but when fitting the data sets over the sub-band frequency range $60 < \nu < 100$ MHz.

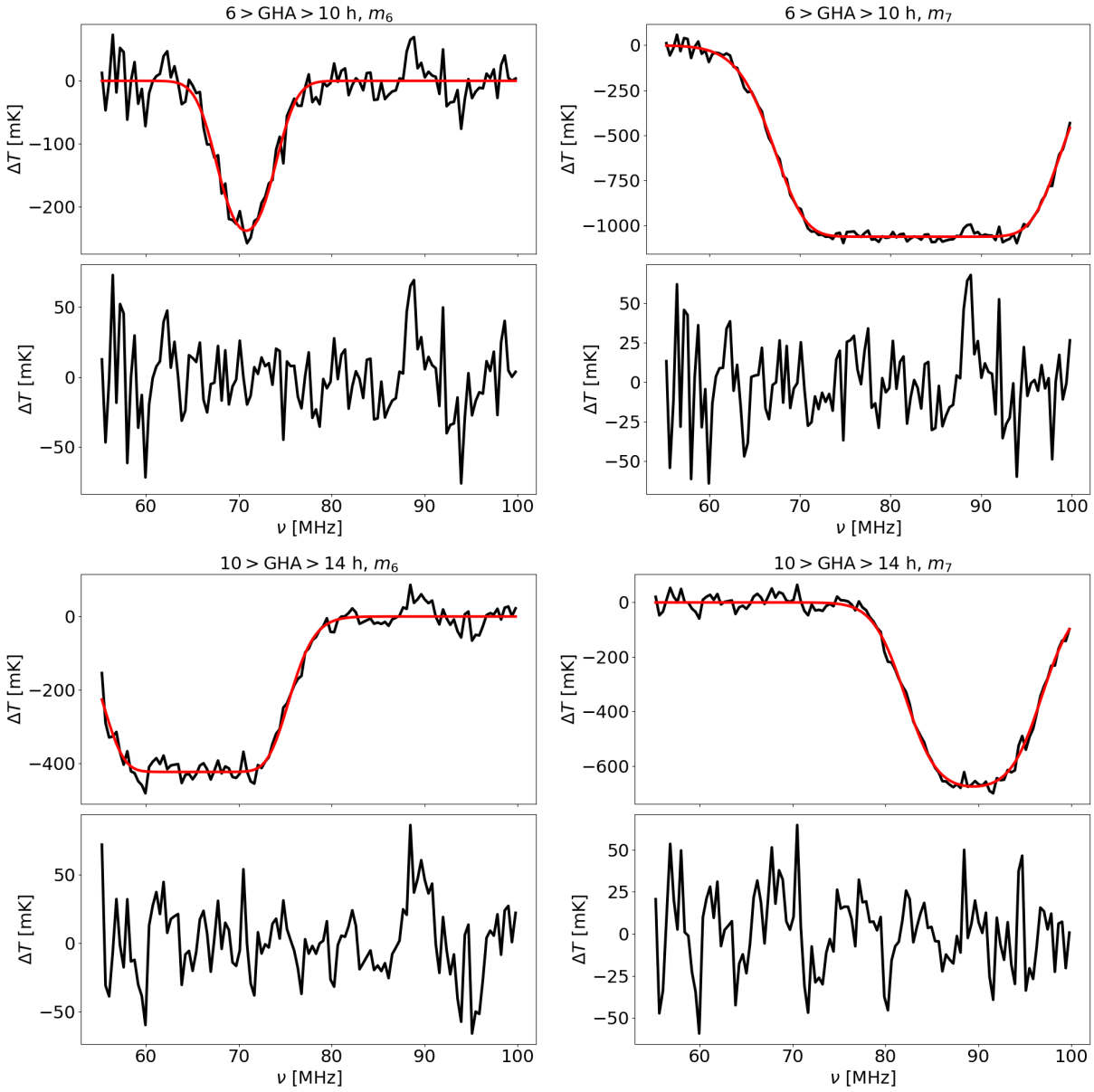

Figure 13: MAP flattened Gaussian signal components overlaid on $(d - T_{\rm Fg})$ residuals (top subfigures) and $(d - m_n)$ residuals, with $m_n = T_{\rm Fg,n} + T_{21}$ and $T_{\rm Fg,n} + T_{21}$ a n-term linlog polynomial foreground model with $n = 6$ (left) and $n = 7$ (right), in fits to GHA = 8 (top row) and 12 (bottom row) p2 EDGES low band data sets over the sub-band frequency range $55 < \nu < 100$ MHz.

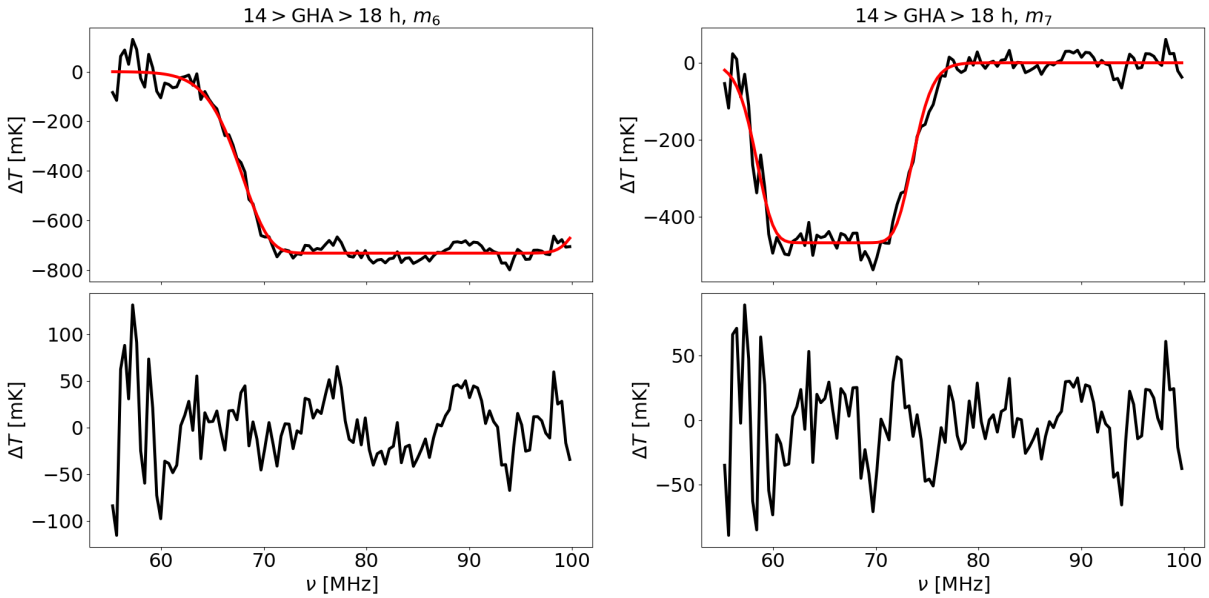

Figure 14: As in Figure 13 but for a fit to the GHA = 16 hour p2 EDGES low band data set.

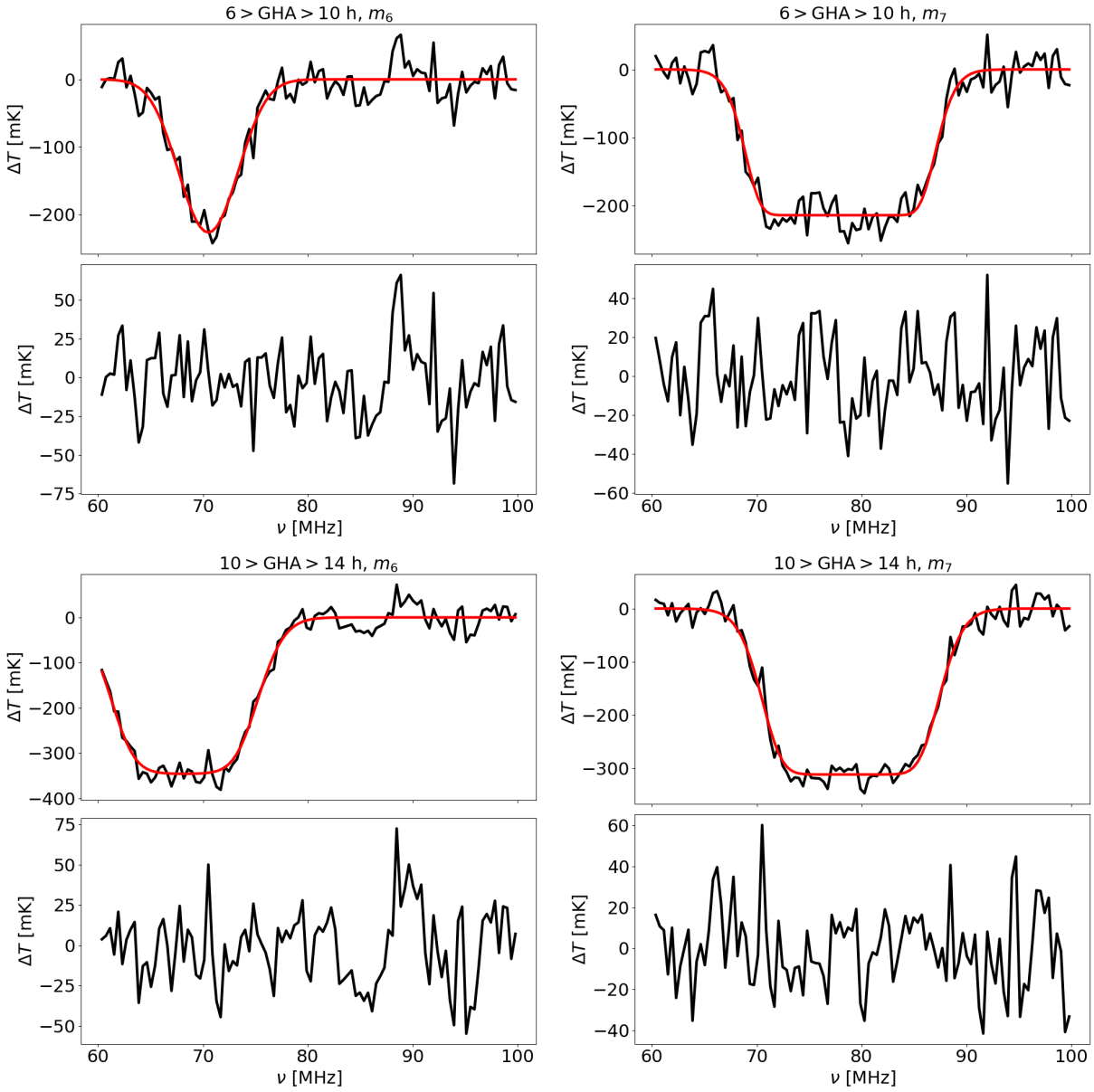

Figure 15: MAP flattened Gaussian signal components overlaid on $(d - T_{\rm Fg})$ residuals (top subfigures) and $(d - m_n)$ residuals, with $m_n = T_{\rm Fg,n} + T_{21}$ and $T_{\rm Fg,n} + T_{21}$ a n-term linlog polynomial foreground model with $n = 6$ (left) and $n = 7$ (right), in fits to GHA = 8 (top row) and 12 (bottom row) p2 EDGES low band data sets over the sub-band frequency range $60 < \nu < 100$ MHz.

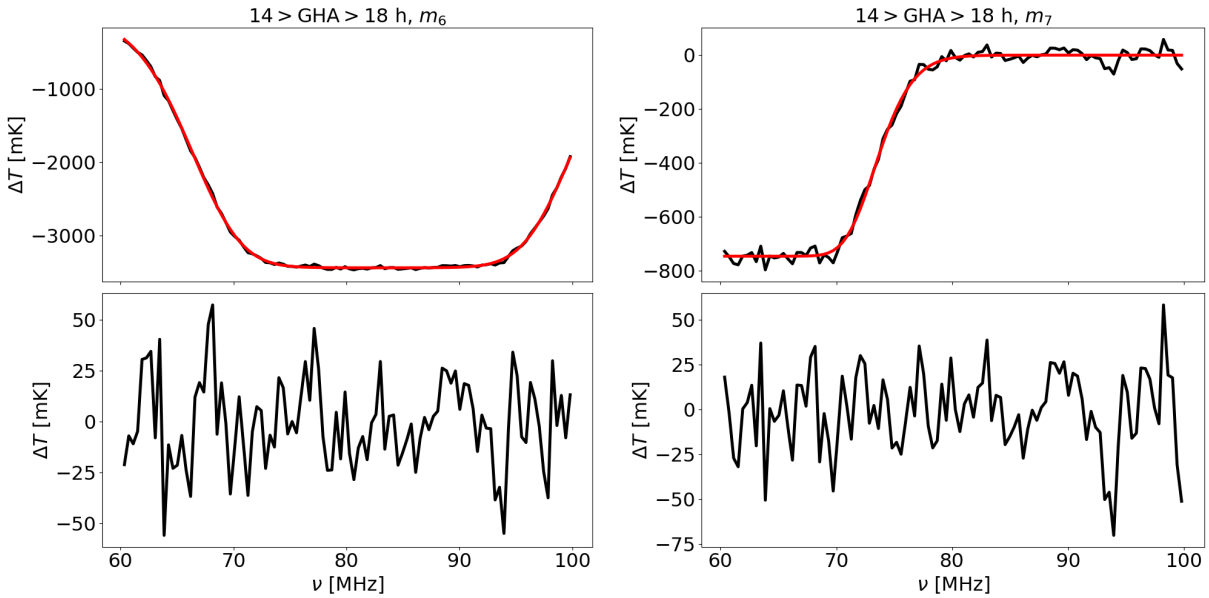

Figure 16: As in Figure 15 but for a fit to the GHA = 16 hour p2 EDGES low band data set.

3.3.4 Limitations of data model

The components of the data model used in this memo are described in Section 2.1. They include a flattened Gaussian model for the 21 cm signal, a lin-log polynomial foreground model with either $n = 6$ or $n = 7$ terms and a model for the data covariance matrix that assumes that the noise on the data is Gaussian and white. It is evident from the foreground posteriors, which typically have oscillating positive and negative coefficients (see Appendix C), that the nominal model for the foreground used here must also be modelling non-foreground structure in the data. This additional structure could be due, for example, to imperfections in data calibration, chromaticity correction or RFI excision. Better understanding of these systematics would be valuable. Investigations to understand sources of systematics, in general, is ongoing throughout the EDGES collaboration.

References

Barkana R., 2018, Natur, 555, 71

Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Natur, 555, 67

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 791

Sims P. H., Pober J. C., 2020, MNRAS, 492, 22. doi:10.1093/mnras/stz3388

A Full-band astro parameter posteriors – restricted τ prior

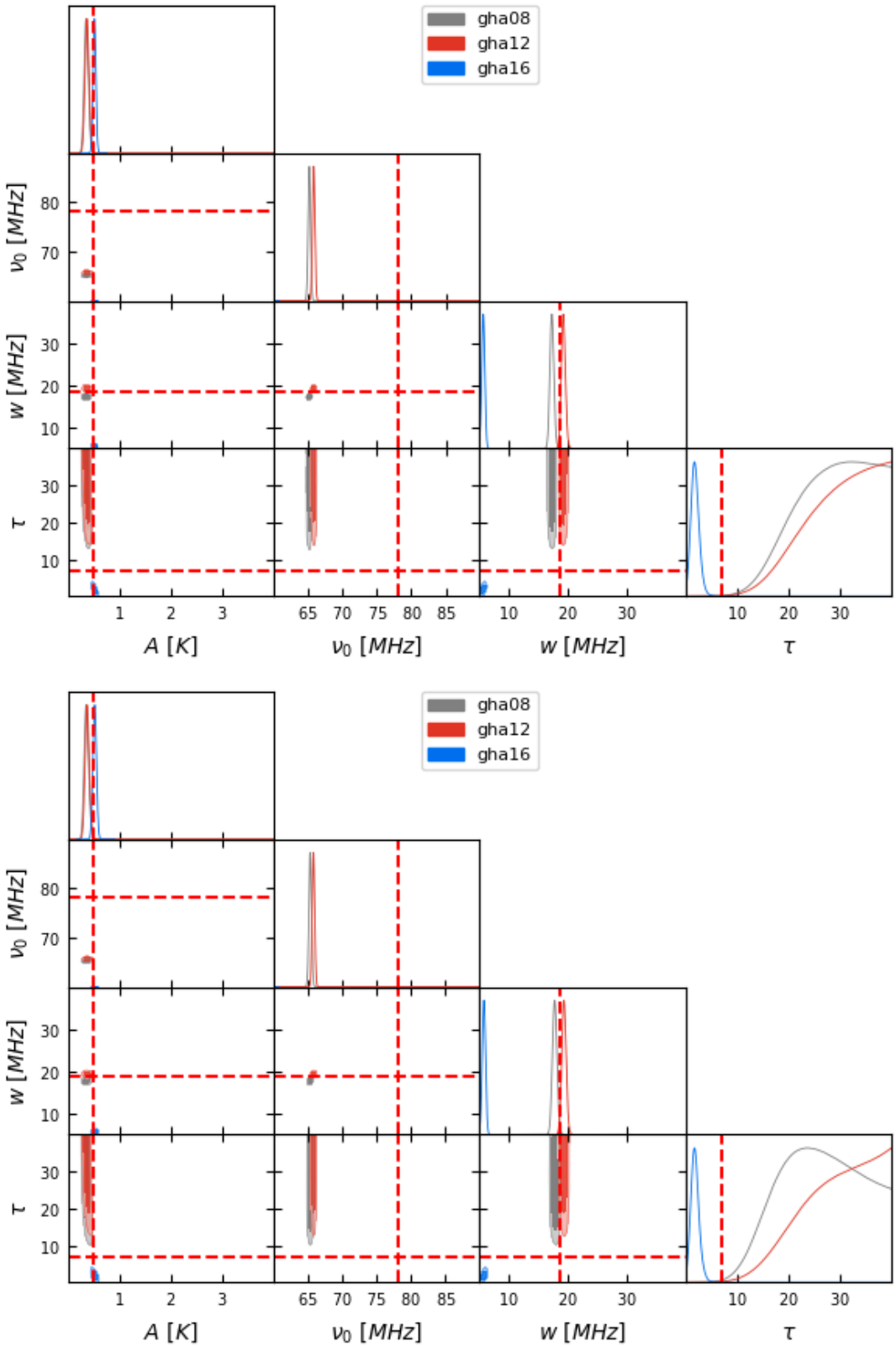

Figure 17: As in Figure 8 but for a narrower $U(0, 40)$ uniform prior on the flattening factor, τ, of the flattened Gaussian 21 cm model.

B Fixing individual 21 cm signal parameters to match estimates recovered in B18

Given the degeneracy between 21 cm signal models when fitting to data restricted to the $65 < \nu < 95$ MHz sub-band, one could ask, hypothetically, whether a given level of a priori knowledge of the signal would be sufficient to eliminate that degeneracy[5]. The impact of imposing a restrictive prior on a single parameter of the 21 cm signal is illustrated in Figures 18 and 19. Restricting the central frequency or width of the 21 cm signal model to their respective means recovered in B18 results in stronger constraints on the the non-restricted parameters and greater consistency with the values of those parameters found in B18. The impact of fixing the amplitude or τ is less strong and has a greater GHA-dependence.

[5]Physical motivation for such a procedure could arise if strong external constraints the signal were available; however, given current constraints, fixing parameters in this manner is less strongly motivated.

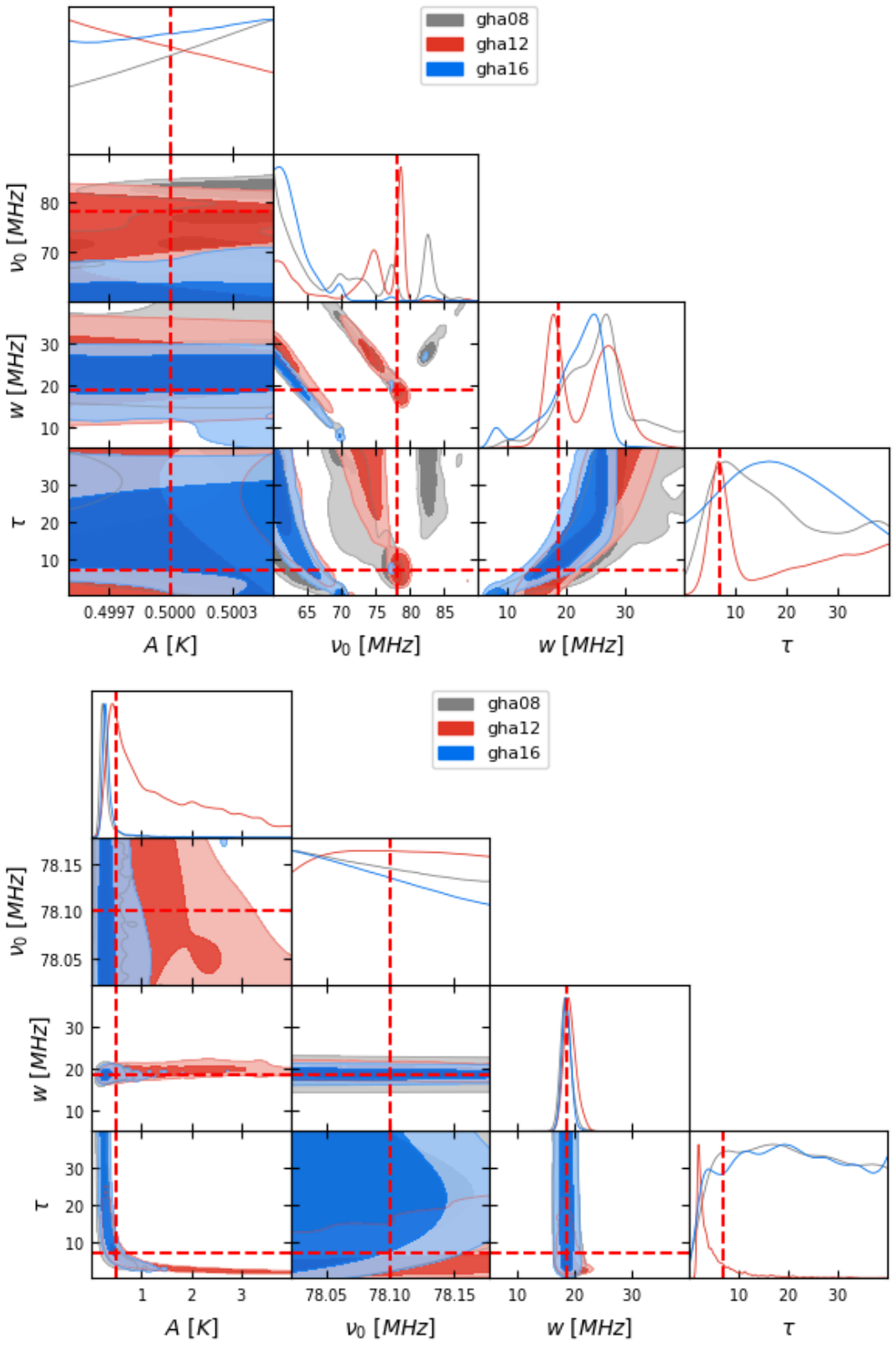

Figure 18: The impact of a restrictive prior on a single parameter of the 21 cm signal. [Top] The posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 term lin-log foreground model (Equation 4), from the p2 data sets over the sub-band frequency range $65 < \nu < 95$ MHz, with a uniform prior on the signal amplitude between $A = 0.5 \pm 0.1\%$ K. [bottom] As top, but assuming a uniform prior on the central frequency of the 21cm signal between $\nu_0 = 78.1 \pm 0.1\%$ MHz.

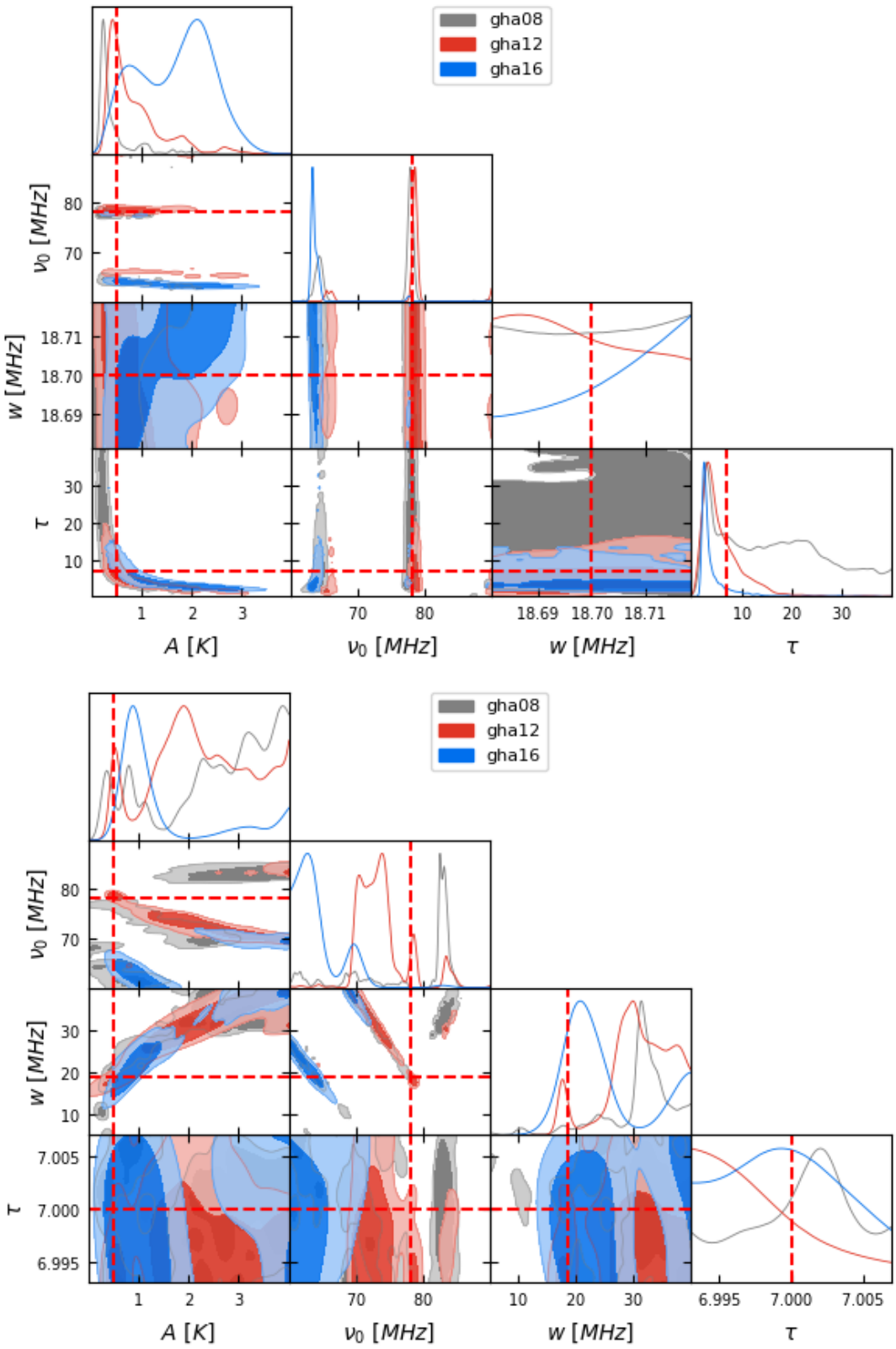

Figure 19: The impact of a restrictive prior on a single parameter of the 21 cm signal. [Top] The posterior probability distributions of the parameters of a flattened Gaussian 21 cm signal model, when jointly estimated with a 6 term lin-log foreground model (Equation 4), from the p2 data sets over the sub-band frequency range $65 < \nu < 95$ MHz, with a uniform prior on the signal amplitude between $w = 18.7 \pm 0.1\%$ MHz. [bottom] As top, but assuming a uniform prior on the central frequency of the 21cm signal between $\tau = 7.0 \pm 0.1\%$.

C Foreground parameter posteriors

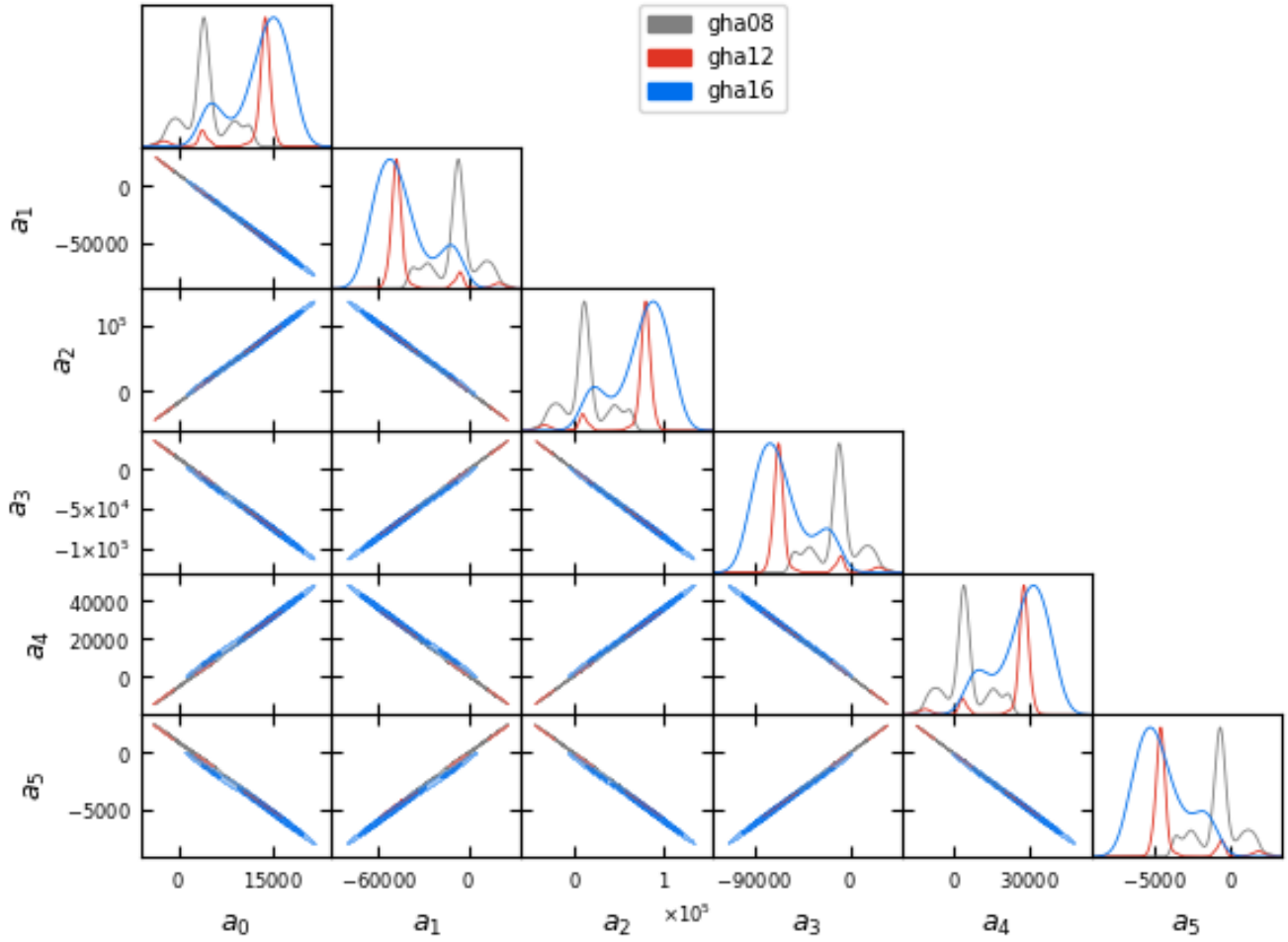

Figure 20: Posterior probability distributions of the parameters of a 6 term lin-log foreground model (Equation 4), when jointly estimated with a flattened Gaussian 21 cm signal model, from three p1 EDGES low band data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red) and $14 \geq \mathrm{GHA} > 18$ (blue), over the restricted frequency range $65 < \nu < 95$ MHz.

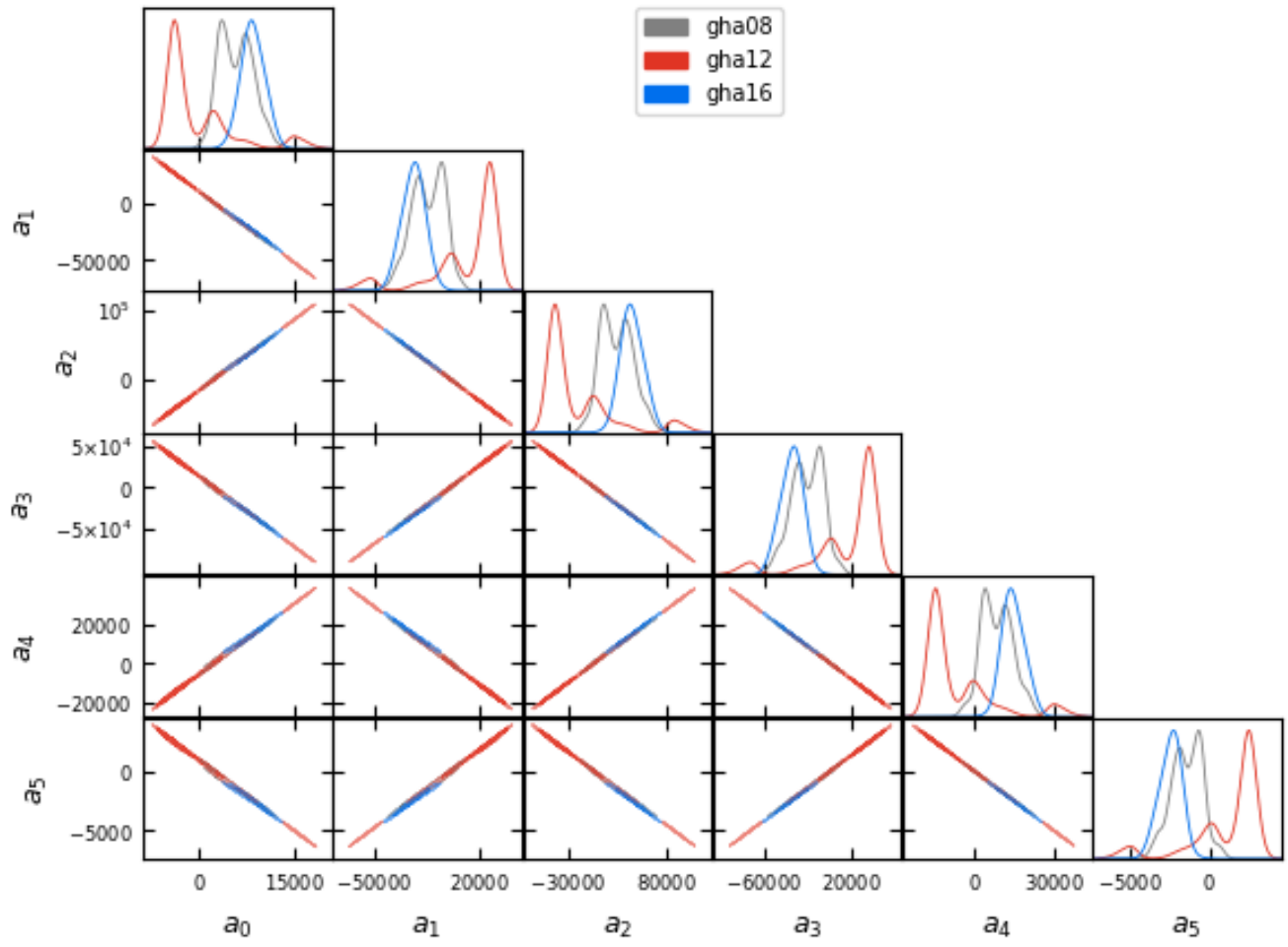

Figure 21: As in Figure 20 but for data set produced by processing the raw data with the Monsalve-pipeline.

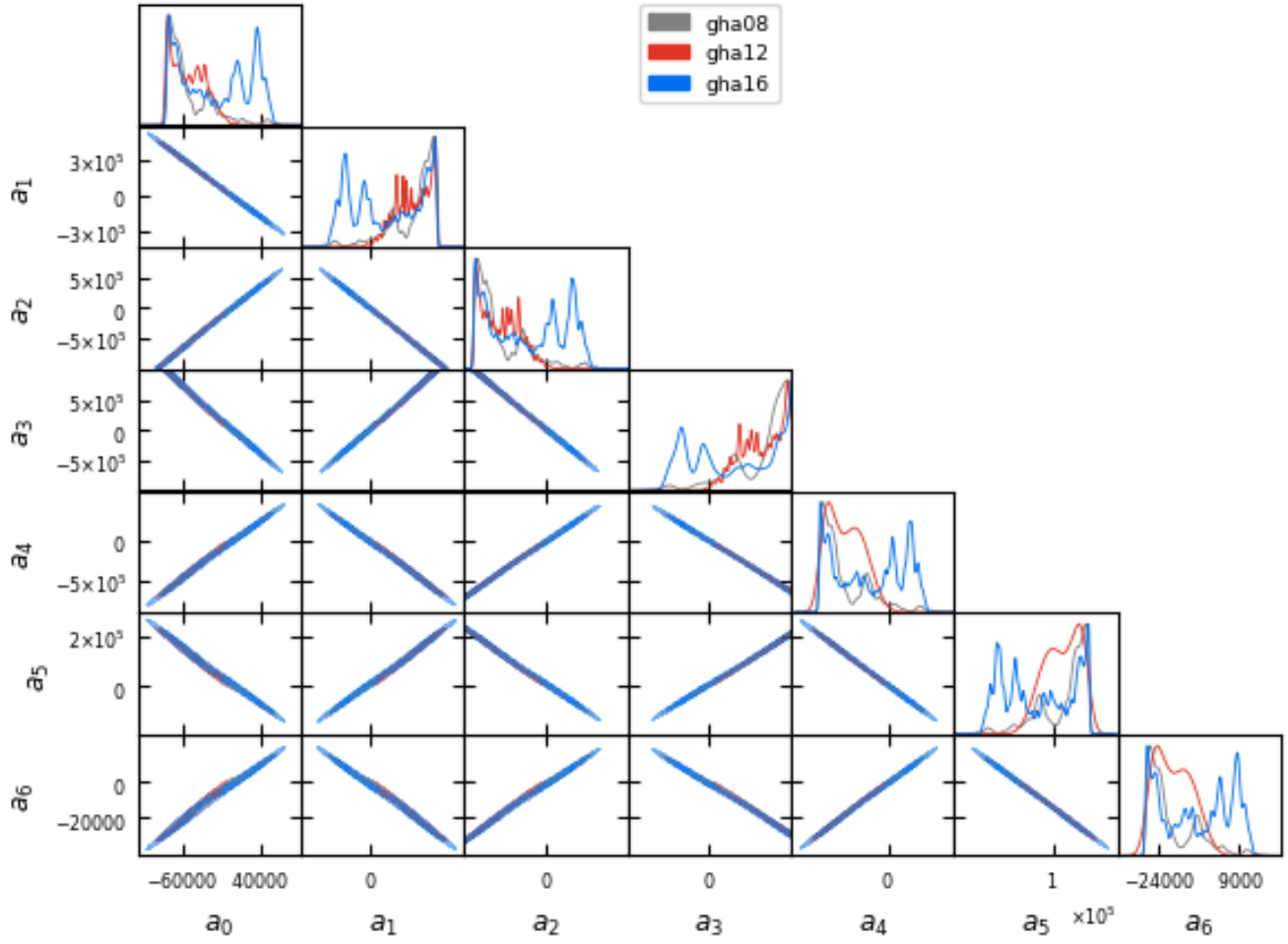

Figure 22: Posterior probability distributions of the parameters of a 7 term lin-log foreground model (Equation 4), when jointly estimated with a flattened Gaussian 21 cm signal model, from three p1 EDGES low band data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red) and $14 \geq \mathrm{GHA} > 18$ (blue), over the restricted frequency range $65 < \nu < 95$ MHz.

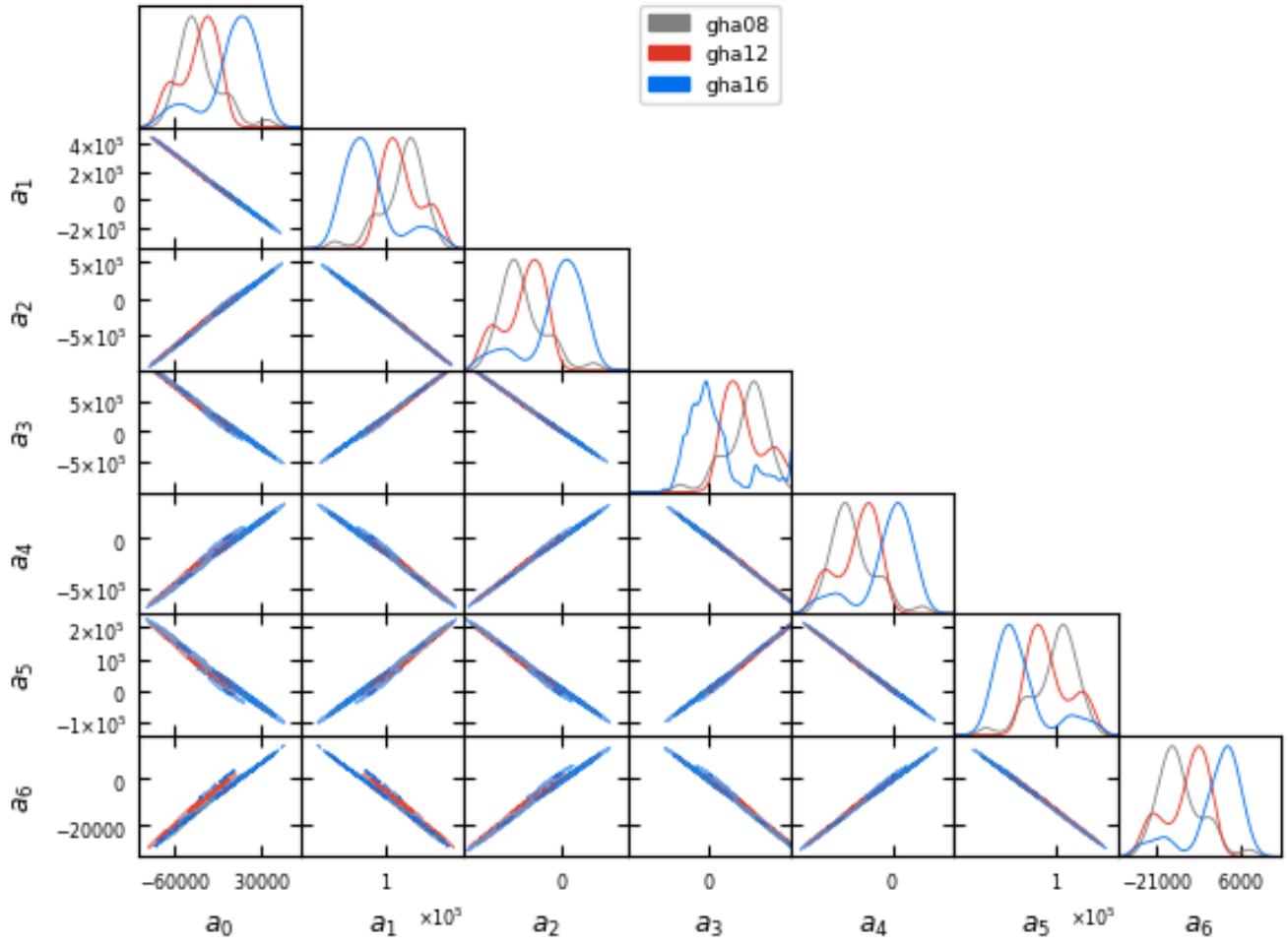

Figure 23: As in Figure 22 but for data set produced by processing the raw data with the Monsalve-pipeline.

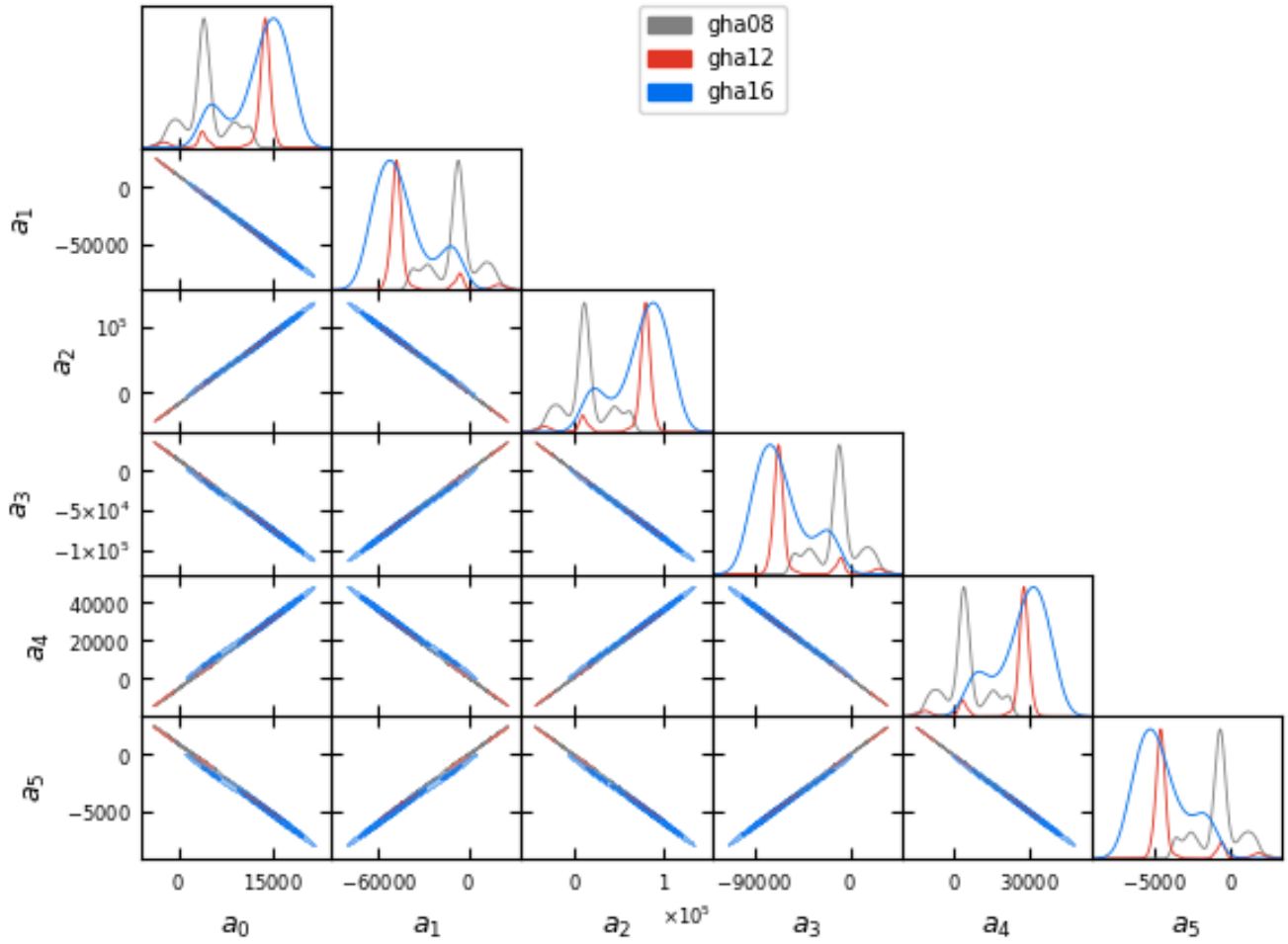

Figure 24: Posterior probability distributions of the parameters of a 6 term lin-log foreground model (Equation 4), when jointly estimated with a flattened Gaussian 21 cm signal model, from three Monsalve EDGES low band data sets with Galactic hour angle ranges: $6 \geq \mathrm{GHA} > 10$ (grey), $10 \geq \mathrm{GHA} > 14$ (red) and $14 \geq \mathrm{GHA} > 18$ (blue), over the full EDGES-low frequency range $50 < \nu < 100$ MHz.

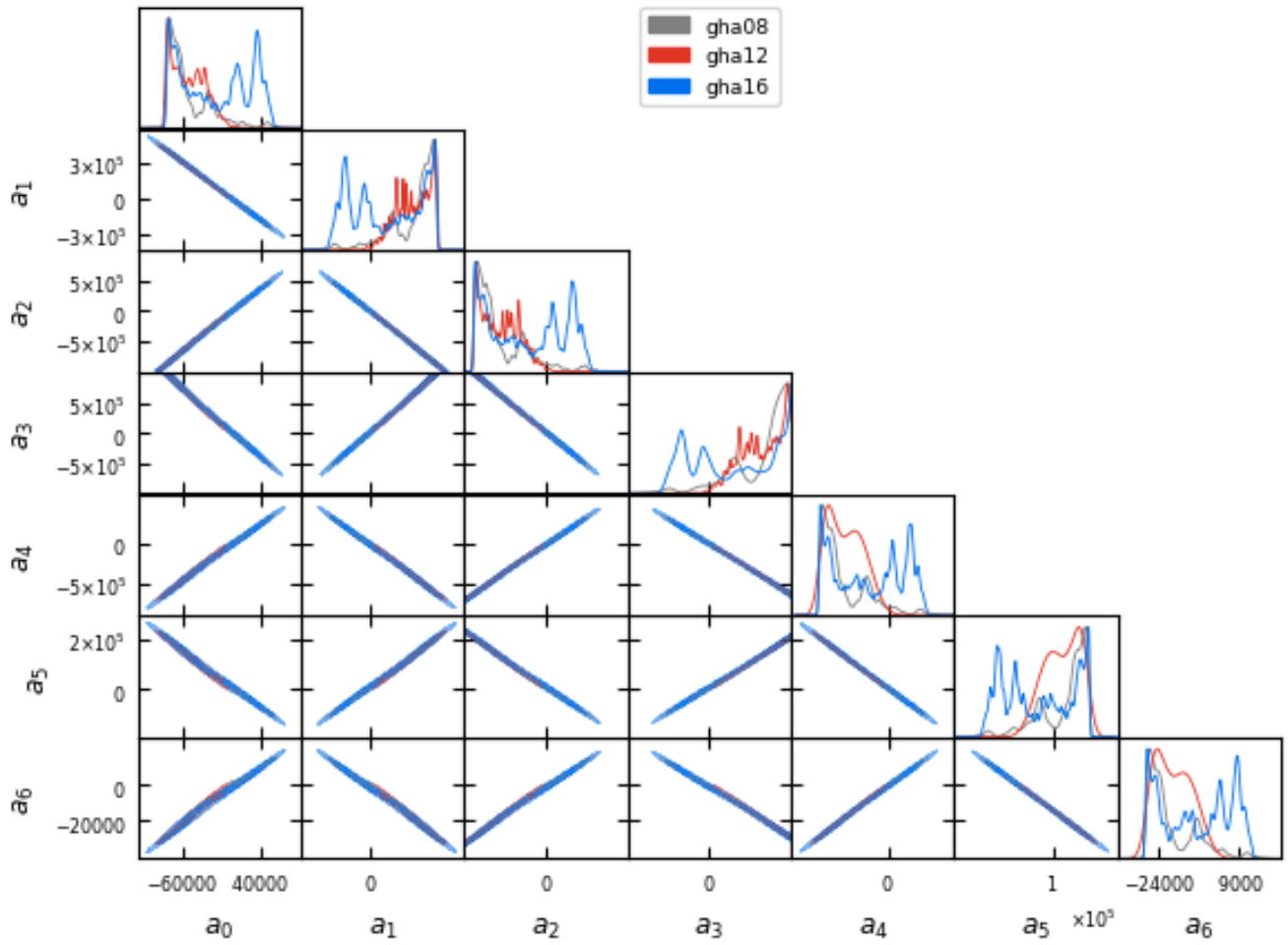

Figure 25: As in Figure 24 but for the parameters of a 7 term lin-log foreground model.