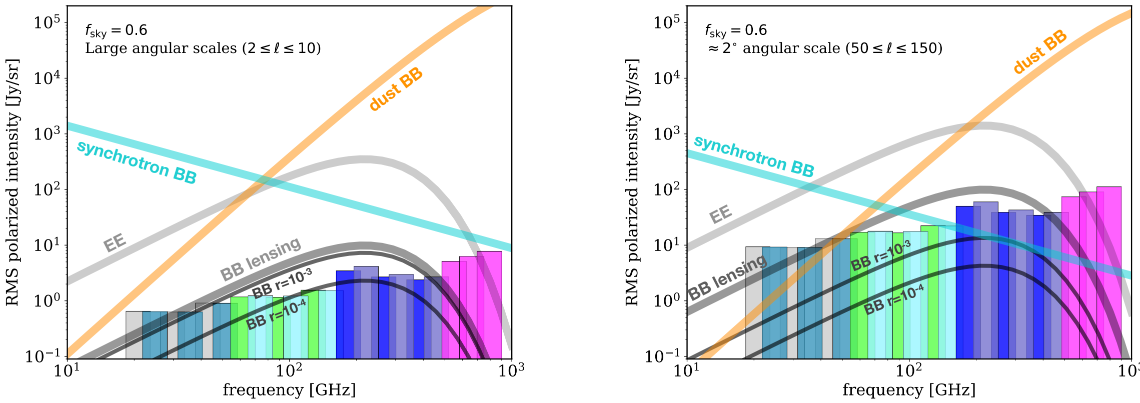

Figure 1: Polarization *BB* spectra of Galactic synchrotron and dust, compared to CMB polarization *EE* and *BB* spectra of different origins for two values of r and for two ranges of angular scales: large $\ell \leq 10$ corresponding to the reionization peak (left panel), and intermediate $50 \leq \ell \leq 150$ corresponding to the recombination peak (right panel). The location and sensitivity of the 21 PICO frequency channels is shown as vertical bands. (The color scheme is explained in Section 3.2.)

Diffuse Milky Way emissions dominate the sky's polarized intensity on the largest angular scales; see Figures **??** and **??**. Even though their levels decrease when considering smaller angular scales, they are still considerably brighter near the **IGW!** (**IGW!**) peak at $\ell = 80$ when averaging over 60% of the sky. Even in the cleanest, smaller patches of the sky, far from the galactic plane and thus relatively low in galactic emissions, their levels are expected to be substantial relative to the **IGW!** for $r \lesssim 0.01$, and dominate it for $r \lesssim 0.001$. Separating the cosmological and Galactic emissions signal, also called foreground separation, together with control of systematic uncertainties are *the* challenges facing any next decade experiment attempting to reach these levels of constraints on r. The foreground separation challenge would be easily surmountable if the Galactic emissions were already precisely characterized, or were known to have simple, fittable spectral emission laws. But neither is true. Until recently, the spectrum of Galactic synchrotron emission, arising from free electrons spiraling around Galactic magnetic fields, was modeled as a power law $I_{\rm sync} \propto \nu^{\alpha}$, with $\alpha \simeq -1$ (in brightness units). The spectrum of Galactic dust emission, arising from emission by Galactic dust grains, was modeled as $I_{\rm dust} \propto \nu^{\beta} B_{\nu}(T_{\rm dust})$, where $\beta \simeq 1.6$, $T_{\rm dust} \simeq 19\,{\rm K}$, and $B_{\nu}(T)$ is the Planck function; this is referred to as 'modified black body emission'. In principle, an experiment that had 6 frequency bands could determine the three emission parameters as well as the three amplitudes corresponding to that of dust, synchrotron, and the CMB. However, WMAP and Planck observations have shown that neither emission law is universal and that spectral parameters vary with the region of sky is this true? is there evidence from Planck; add references (thus that the values given above are valid only as averages across the sky). Also, while both emission laws are well-motivated phenomenological descriptions, the fundamental physics of emissions from grains of different materials, sizes and temperatures, and of electrons spiraling around magnetic fields implies that these laws are not expected to be exact, nor universal.
We know that we don't know enough about synchrotron and dust emission. We know even less about the polarization level of 'anomalous microwave emission', an excess of dust emission at frequencies between 10 and 100 GHz, and of infra-red sources. Depending on reasonable levels of polarization assumed their contributions to the total polarized signal may be appreciable or negligible (for $r \lesssim 0.001$) [**?**].

Faced with these uncertainties, but also with the opportunity provided by a platform that can host a broad range of frequencies – ground-based experiments are limited to several atmospheric windows and to frequencies of less than 300 GHz – PICO is designed with 21 frequency bands between 21 and 800 GHz; see Figure **??** and Table 3.2. This is the broadest frequency lever arm proposed by any imaging instrument to characterize and enable separation of Galactic emissions.

Foreground uncertainties, and the level of fidelity required in their characterization, also compel a transition in the way we assess and forecast the performance of a future experiment. We can no longer impose specific models upon the data; several publications have demonstrated that deviations between assumed models' parameters and the real sky could give rise to biases in r that are larger that future goals [**?**]. Al, make this more concrete? Rather, the data collected should provide information to constrain Galactic emissions with sufficient accuracy. For PICO we use the approach that has become the 'gold standard' in the community. In this approach we simulate sky maps that are constrained by available data, but otherwise have a mixtures of foreground properties, observe these maps just like a realistic experiment will do, and then apply foreground separation techniques to separate the Galactic and CMB emissions. We also provide forecasts using other techniques that use analytic calculations to estimate the efficacy of foreground separation, or others in which the simulated sky map is assumed to have specific Galactic emission models, which are then being fitted.

0.0.1 PICO Foreground Separation Methodology

Sky Maps For assessing the efficacy of foreground separation with PICO we used 8 different full sky models. All models were consistent with and constrained by available data and uncertainties from WMAP and *Planck*. The range of models included one test case that had a very simple realization of foregrounds, and others with varying degree of complexity including spatially varying spectral parameters and along the line of sight, anomalous microwave emission up to 2% polarized, dust polarization that rotates slightly as a function of frequency because of projection effects, or dust spectral energy distribution that departs from a simple modified blackbody. All foreground maps are generated at native resolution of 6.8 arcmin pixels [**?**]. They are generated using PySM and/or PSM codes [**?**]. Distinctly different realizations of the sky are allowed by current data, as demonstrated by Figure **??**. Karl, More details of the models are available at **?**].

For each of the 8 models we added CMB signals in both intensity and polarization matching a ΛCDM universe. The BB-lensing signal matched the level of 85% delensing forecasted for PICO. Each of these sky models had 100 different realization of the PICO CBE noise levels; 50 realizations had no **IGW!** signal and 50 others had a level of $r = 0.003$.

Foreground Separation The sky models were analyzed with a variety of techniques which are based on two broad categories: correlation methods, which exploit the fact that foreground emission is strongly correlated from frequency to frequency, but uncorrelated with the CMB, and parametric methods, which model the sky emission using specific (parametric) emission laws, and use spectral fits in independent pixels or sky regions to infer the amplitude and spectral parameters of each of the components in the sky. Correlation methods include SEVEM, and variants of the **ILC!** (**ILC!**) algorithm, such as the needlet-space **ILC!** (NILC) and a version generalised to multidmensional components (GNILC). Parametric methods include the Commander algorithm.

say something about the fisher methods, Raphael and Stephen

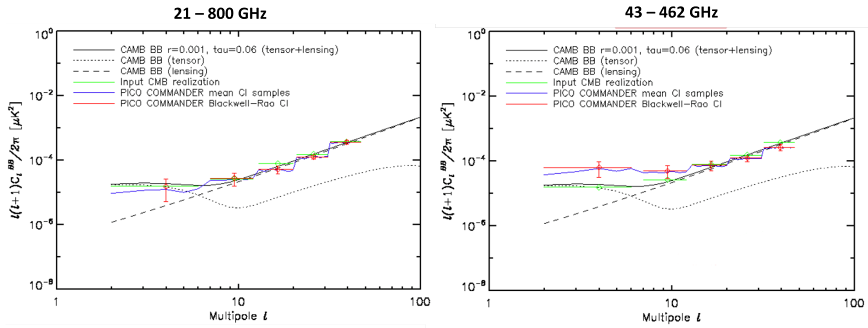

Figure 2: Foreground removal with all of PICO's 21 frequency bands (left panel) recovers the input CMB (green) without any bias (red) using the Commander algorithm on the *Planck* sky model (with 4 deg pixels, and 60% sky fraction). Running the same algorithm on the same sky without several of the lowest and highest bands (right panel) produces an output spectrum (red) that is biased relative to the input (green) at low ℓ multipoles. The bias would be interpreted as higher value of r relative to the model input (solid) with $r = 0.001$ (dots) and lensing (dash).

0.0.2 Results and Discussion

Our results validate the need for a broad frequency coverage with a strong lever arm on Galactic emissions outside of the primary CMB bands. Figure **??** shows that removing several of PICO's frequency bands, particularly those that monitor dust and synchrotron at high and low frequencies, respectively, significantly biases the extracted *BB* power spectrum, particularly at the lowest ℓ values.

There is also evidence that at levels of $r \simeq 0.001$ the combination of sensitivity and broad frequency coverage are efficacious in foreground removal. Figure **??** shows a result from the gold standard process described above for one of the sky models and with an input **IGW!** of $r = 0.003$. Residual foregrounds are below the cosmological signal over the important low ℓ range, where foregrounds are strongest. The residual spectra would likely be lower when analysis is carried out on only 50 or 40% of the sky, rather than the 60% used here.

There is other evidence that PICO can reach its stated target of $\sigma(r) = 0.0001$. Map-based simulations that were carried out for the forthcoming CMB-S4 experiment have shown that it can reach levels of $\sigma(r) = 0.0005$ in small, 3%-size, clean patches of the sky. The analysis only used frequencies up to 300 GHz. In principle, even smaller patches of 1-2% size are sufficient, and preferable, for attaining as low $\sigma(r)$ as possible. The PICO noise level per sky pixel is similar to that of CMB-S4, but PICO will have *full* sky coverage and thus access to *all* the clean patches available. Data from *Planck* indicate that there are ~ 10 check! patches as clean, or cleaner than those used for the CMB-S4 analysis, indicating that PICO's σr should be ~ 3 times more stringent. This scaling is very conservative because it only assumes CMB-S4's much narrower breadth of frequency coverage and its 7 bands check; it neglects PICOs much stronger rejection of foregrounds with 21 bands and up to 800 GHz.

few words about xforecast, and results from Raphael

While our results are encouraging, as they suggest that PICO's frequency coverage and sensitivity will be adequate for this level of r, more work should be invested to gain complete confidence. This work includes running numerous realizations of different sky models, and analyzing them with various techniques; optimizing sky masks; and using combination of techniques to handle large, intermediate, and small angular scale foregrounds differently.

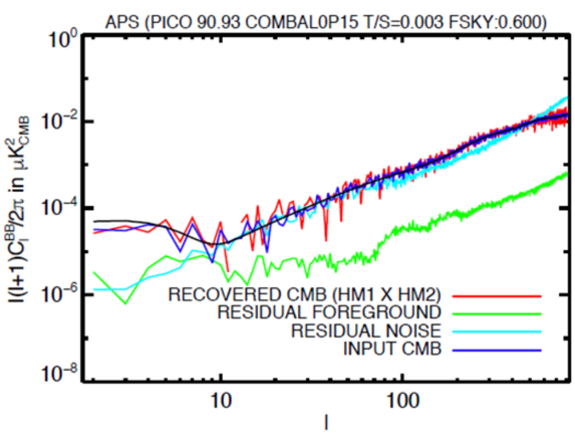

Figure 3: The power spectrum of residual *BB* foregrounds (green) has lower level than both the input CMB (blue) and the recovered CMB (red) which match well each other and the underlying cosmological model (black) after foreground separation with the NILC algorithm. This exercise assumed use of 60% of the sky.