Draft version July 16, 2019
Typeset using LaTeX default style in AASTeX62

Cleaning Foregrounds and Boosting Signals with Multi-line Intensity Mapping

Angus Beane,[1,2] Adam Lidz,[2] and others

[1] *Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[2] *Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA*

ABSTRACT

Intensity mapping of galactic emission lines stands to shed light on both galactic astrophysics and cosmology from $z \sim 1$ to $z \sim 10$ in the coming years. However, several observational challenges must first be met, including the problems of interloper and foreground contamination. Cross-correlations of two different emission lines can confront these challenges. By combining the cross-correlations between three or more lines, one can measure the large-scale bias encoded in the auto-spectrum of each line. We aim to understand the general case in more detail. Using the Fisher matrix formalism, we develop a framework for understanding the observational constraints possible with intensity mapping of an arbitrary number of lines and their associated cross-spectra. Results. We find that multi-line intensity mapping is a promising technique, and future intensity mapping efforts should take care to maximize the shared area between different surveys.

1. INTRODUCTION

Precision studies of the large-scale structure (LSS) of the universe provide important cosmological constraints. Traditional galaxy surveys are limited in their possible redshift extent to currently $z \lesssim 1$ due to the expense required in taking spectra of large numbers of high-redshift galaxies. Line-intensity mapping has emerged as a promising alternative, instead measuring the collective emission in a convenient line in fairly coarse pixels. By sacrificing resolution, intensity mapping allows one to measure large scales at high redshift.

Intensity mapping of the H I 21 cm line is the most well studied. During reionization, it is sourced by the intergalactic medium . Post-reionization, it is sourced by neutral hydrogen primarily in the interstellar medium . A detection of the 21 cm power spectrum and the baryon acoustic oscillation feature have been made at $z \sim 0.8$ (Chang et al. 2010).

2. TECHNICAL BACKGROUND

We first lay out the approach to modelling the average intensity of an arbitrary emission line, the spatial clustering of that signal, the cross-spectrum between two different lines, and the fisher matrix approach for estimating the bias factors from an arbitrary number of cross-spectra or cross-spectra and auto-spectra.

2.1. *Intrinsic Line Emission*

The average intensity of an arbitrary line is calculated according to the simple model of Lidz & Taylor (2016), which is based upon Lidz et al. (2011) and Pullen et al. (2013). The details are laid out in Appendix A, but in brief the model assumes a Schechter (1976) form for the star formation rate function and a linear relationship between a halo's luminosity and its star formation rate with constants of proportionality from Visbal & Loeb (2010).

We note that in this model the redshift evolution of each line is identical up to the luminosity-SFR constant of proportionality. In reality, we do not expect this to be the case. For example, the different CO lines should have varying intrinsic strengths since the typical phase of the ISM in each emitting galaxies should change as a function of redshift. This highlights one of the limitations of the line intensity model adopted here. Fully understanding the underlying galactic population will require more nuanced models, on which development is underway (**cite**).

To model the intrinsic spatial fluctuations of each line, we follow Lidz & Taylor (2016) and Cheng et al. (2016), though we adopt the notation of the former. We use the word "intrinsic" here not to mean the signal in the cosmological

Corresponding author: Angus Beane
abeane@sas.upenn.edu

rest frame, but rather the observed signal assuming the observed coordinates have been transformed to comoving coordinates with the correct redshift. This is not the case for interloping lines, for which their contribution to the signal is distorted because the assumed redshift and the emitted redshift are not the same.

The intrinsic emission from a line i is given by

$$P_{i,i}(k,\mu,z) = \langle I_i\rangle^2\langle b_i\rangle^2(1+\beta_i\mu^2)^2 D\left[\mu k\sigma_p(z)\right] P_{\delta,\delta}(k,z), \tag{1}$$

where $\mu = k_\parallel/k$ is the cosine of the angle between the wavevector $\boldsymbol{k}$ and the line of sight direction, $\langle I_i\rangle$ is the average specific intensity of line i, and $\langle b_i\rangle$ is the average luminosity-weighted bias of the emitting galaxies. The Kaiser effect (Kaiser 1987) is implemented through the $(1+\beta_i\mu^2)^2$ factor, where $\beta_i = f_\Omega/\langle b_i\rangle$ with $f_\Omega = \mathrm{d}\ln D/\mathrm{d}\ln a$ is the logarithmic derivative of the growth factor, well approximated by $f_\Omega \approx \Omega^{0.55}$ (Linder 2005). We assume a Lorentzian form for the finger-of-god suppression:

$$D(x) = \frac{1}{1+x^2}. \tag{2}$$

We approximate the pairwise velocity dispersion by $\sigma_p(z) = \sigma_v(z)/\sqrt{2}$ with $\sigma_v^2(z)$ being the variance of the line-of-sight component of the velocity field according to linear theory. Equation (3) neglects the shot noise component of the power spectrum, which is an acceptable omission for the large scales considered in this work. Our neglect to add the shot noise component is out of convenience considering the scales we will encounter, but this is not to imply that the shot noise component is uninteresting in its own right (**cite some?**) even in the cross-spectra (**cite Breysse 2019**). We use the Eisenstein & Hu (1998) model for the linear theory matter power spectrum, $P_{\delta,\delta}(k,z)$. All basic cosmological calculations are performed using the code `colossus` (Diemer 2018).

Throughout this work we will only consider the angle-averaged power spectrum, that is:

$$\begin{aligned} P_{i,i}(k,z) &= \langle I_i\rangle^2\langle b_i\rangle^2 C_i(k) P_{\delta,\delta}(k,z) \\ &\equiv B_i^2 C_i(k) P_{\delta,\delta}(k,z), \end{aligned} \tag{3}$$

where we have defined $C_i(k)$ to be the amplification of the power spectrum coming from the average over the angle-dependent terms in Equation (3), and B_i to be the product of $\langle b_i\rangle$ and $\langle I_i\rangle$. Henceforth all power spectra will be assumed to be angle averaged.

For the cross-spectrum between two lines i and j, we assume that

$$P_{i,j}(k,z) = r_{i,j}(k,z)\sqrt{P_{i,i}(k,z)P_{j,j}(k,z)}, \tag{4}$$

where $r_{i,j}$ is the cross-correlation coefficient between line i and j. We will only consider scales where $r_{i,j} = 1$.

2.2. *Multi-line Formalism*

Given a set of cross-spectra and their covariance matrix, we can, in principle, determine the optimal constraints that can be placed on both the average bias $\langle b_i\rangle$ and average intensity $\langle I_i\rangle$. However, such a procedure requires measuring the angular dependence of the power spectrum (**cite**), which may be fairly challenging (Chung 2019). Therefore, after fixing the value of b_i for each line i, we only consider the constraint that can be placed on each B_i. This is slightly inconsistent, since in principle we should first determine the constraint on $\langle b_i\rangle$ and $\langle I_i\rangle$ and then convert this into a constraint on B_i. But our approximation is reasonable because **of some reason**.

We can write down the Fisher matrix as

$$F_{i,j} = \int \frac{\mathrm{d}^3k}{(2\pi)^3} V_{\mathrm{surv}} \sum_l \sum_m \frac{\partial \hat{P}_l(\boldsymbol{k})}{\partial B_i} \left(\mathrm{Cov}[\hat{P}_l(\boldsymbol{k}), \hat{P}_m(\boldsymbol{k})]\right)^{-1} \frac{\partial \hat{P}_m(\boldsymbol{k})}{\partial B_j}, \tag{5}$$

where the sum over l and m each is over each *pair* of lines. Since we will be considering only angle-averaged power spectra, we can use the angle-averaged form of the Fisher matrix:

$$F_{i,j} = \int \frac{\mathrm{d}k\, k^2}{2\pi^2} V_{\mathrm{surv}} \sum_l \sum_m \frac{\partial \hat{P}_l(k)}{\partial B_i} \left(\mathrm{Cov}[\hat{P}_l(k), \hat{P}_m(k)]\right)^{-1} \frac{\partial \hat{P}_m(k)}{\partial B_j}. \tag{6}$$

In order to compute the covariance matrix, we use the result that (e.g. Visbal & Loeb 2010),

$$\begin{aligned} \mathrm{Var}[P_{i,j}] &= P_{i,j}^2 + P_{i,\mathrm{tot}}P_{j,\mathrm{tot}} \\ \mathrm{Cov}[P_{i,j}, P_{i,k}] &= P_{i,\mathrm{tot}}P_{j,k} + P_{i,j}P_{i,k}, \end{aligned} \tag{7}$$

where $P_{i,\rm tot} = P_{i,i} + N_i$. The computation of the noise N_i is given in the next subsection.

2.3. *Noise Components*

We describe the computation of each noise component. In principle the noise should consist of three components:

$$N_i = N_{i,\rm instrumental} + N_{i,\rm interlopers} + N_{i,\rm foreground}. \tag{8}$$

However, in this work we will ignore the foreground contribution. Our work should then be seen as a best-case scenario. **add more words**

2.3.1. *Instrumental Noise*

Add.

2.3.2. *Interloping and Extraloping Lines*

The noise for interloping and extraloping lines can be written as (Lidz & Taylor 2016; Cheng et al. 2016),

$$N_{i,\rm interlopers} = \sum_j \frac{1}{\alpha_\parallel(z_j)\alpha_\perp(z_j)^2} P_j\left(\frac{k_\parallel}{\alpha_\parallel(z_j)}, \frac{\boldsymbol{k}_\perp}{\alpha_\perp(z_j)}\right), \tag{9}$$

where the sum on j is over all interloping lines, z_j is the emitted redshift of line j, and where,

$$\alpha_\parallel = \frac{H(z_i)}{H(z_j)}\frac{1+z_j}{1+z_i} \qquad \alpha_\perp = \frac{D_A(z_j)}{D_A(z_i)}, \tag{10}$$

where z_i is the emitted redshift of line i and $D_A(z)$ is the comoving angular diameter distance to redshift z.

3. THREE FIELD APPROACH

We begin by considering the three field approach. We first start with this relatively simple case to gain some intuition for the S/N behavior when the cross-spectra of multiple lines are combined.

We are concerned with the problem of estimating the power spectrum of each line only from the three cross-spectra. To be more concrete, we model each auto-spectrum and cross-spectrum as,

$$\begin{aligned} P_{i,i}(k,z) &= B_i(z)^2 P_{\delta,\delta}(k,z) \\ P_{i,j}(k,z) &= B_i(z)B_j(z)P_{\delta,\delta}(k,z), \end{aligned} \tag{11}$$

where $B_i(z) \equiv b_i(z)\langle I_i\rangle(z)$. We will generally suppress the redshift and k arguments for brevity. We can model the auto-spectra and cross-spectra in this way only on sufficiently large scales, i.e. only when $k < k_{\rm max}$ for some value of $k_{\rm max}$. The value of $k_{\rm max}$ is set by two conditions. First, we must have that linear biasing is a good description for each field i so that B_i for each line is independent of k. Second, for each pair of lines (i,j) we must have that $|r_{i,j}(k)| = 1$ when $k < k_{\rm max}$.

Eventually it will be desirable to place constraints on $P_{\delta,\delta}(k)$. For now, however, we simply wish to understand how well future experiments will be able to determine each of B_i as a function of z. We can proceed simply by computing the relevant Fisher matrix:

$$F_{i,j} = \int \frac{{\rm d}^3k}{(2\pi)^3} V_{\rm surv} \sum_l \sum_m \frac{\partial \hat{P}_l(\boldsymbol{k})}{\partial B_i} \left({\rm Cov}[\hat{P}_l(\boldsymbol{k}), \hat{P}_m(\boldsymbol{k})]\right)^{-1} \frac{\partial \hat{P}_m(\boldsymbol{k})}{\partial B_j}, \tag{12}$$

where the sum over l and m each is over every *pair* of lines. Equation 12 can be readily modified to accomodate the case of more than three lines (by extending the summations to include these lines) and to accomodate the auto-spectra (by extending the sum over l and m to include them). We will consider these cases in Sections x and y, respectively.

For the variance and covariance of the power spectra, it can be shown that,

$$\begin{aligned} {\rm Var}[P_{i,j}] &= P_{i,j}^2 + P_{i,\rm tot}P_{j,\rm tot} \\ {\rm Cov}[P_{i,j}, P_{i,k}] &= P_{i,\rm tot}P_{j,k} + P_{i,j}P_{i,k}, \end{aligned} \tag{13}$$

where $P_{i,\rm tot} = P_{i,i} + N_i$. Note that N_i contains contributions from instrument noise, residual foreground contamination, and interloping/extraloping lines. While computing the Fisher matrix numerically is trivial, we will first work out the noise-dominated case which has a somewhat tractable analytic solution.

3.1. *White Noise-dominated Regime*

In the white noise-dominated regime, we have that each $N_i \gg P_{i,i}$ and $N_i(k) = N_i$. The relevant variance and covariance formulae then reduce to,

$$\begin{aligned} \mathrm{Var}[P_{i,j}] &= N_i N_j \\ \mathrm{Cov}[P_{i,j}, P_{i,k}] &= N_i P_{j,k}. \end{aligned} \tag{14}$$

Since the covariance formula is only first order in N_i, we will ignore contributions from terms like it and only consider terms involving $\mathrm{Var}[P_{i,j}]$. In such a case, the formula for the Fisher matrix simplifies greatly and we can write it down as:

$$F = V_k^2 \begin{pmatrix} \frac{B_2^2}{N_1 N_2} + \frac{B_3^2}{N_1 N_3} & \frac{B_1 B_2}{N_1 N_2} & \frac{B_1 B_3}{N_1 N_3} \\ \frac{B_2 B_1}{N_2 N_1} & \frac{B_3^2}{N_2 N_3} + \frac{B_1^2}{N_2 N_1} & \frac{B_2 B_3}{N_2 N_3} \\ \frac{B_3 B_1}{N_3 N_1} & \frac{B_3 B_2}{N_3 N_2} & \frac{B_1^2}{N_3 N_1} + \frac{B_2^2}{N_3 N_2} \end{pmatrix}, \tag{15}$$

where,

$$V_k^2 \equiv V_{\rm surv} \int \frac{\mathrm{d}k\, k^2}{2\pi^2} P_{\delta,\delta}(k)^2. \tag{16}$$

Inverting this matrix we find that the covariance matrix of the bias factors is,

$$C = F^{-1} = \frac{1}{4V_k^2} \begin{pmatrix} \frac{B_1^2 N_2 N_3 + B_2^2 N_3 N_1 + B_3^2 N_1 N_2}{B_2^2 B_3^2} & \frac{-B_1^2 N_2 N_3 - B_2^2 N_3 N_1 + B_3^2 N_1 N_2}{B_1 B_2 B_3^2} & \frac{-B_1^2 N_2 N_3 + B_2^2 N_3 N_1 - B_3^2 N_1 N_2}{B_1 B_2^2 B_3} \\ \frac{-B_1^2 N_2 N_3 - B_2^2 N_3 N_1 + B_3^2 N_1 N_2}{B_1 B_2 B_3^2} & \frac{B_1^2 N_2 N_3 + B_2^2 N_3 N_1 + B_3^2 N_1 N_2}{B_3^2 B_1^2} & \frac{B_1^2 N_2 N_3 - B_2^2 N_3 N_1 - B_3^2 N_1 N_2}{B_1^2 B_2 B_3} \\ \frac{B_1^2 N_2 N_3 - B_2^2 N_3 N_1 - B_3^2 N_1 N_2}{B_1^2 B_2 B_3} & \frac{B_1^2 N_2 N_3 - B_2^2 N_3 N_1 - B_3^2 N_1 N_2}{B_1^2 B_2 B_3} & \frac{B_1^2 N_2 N_3 + B_2^2 N_3 N_1 + B_3^2 N_1 N_2}{B_1^2 B_2^2} \end{pmatrix}. \tag{17}$$

Examining the diagonal components we have that:

$$\begin{aligned} C_{1,1} = \sigma_{B_1}^2 &= \frac{1}{4V_k^2} \left(B_1^2 \frac{N_2 N_3}{B_2^2 B_3^2} + N_1 (\frac{N_2}{B_2^2} + \frac{N_3}{B_3^2}) \right) \\ \implies \frac{\sigma_{B_1}^2}{B_1^2} &= \frac{1}{4V_k^2} \left(\frac{N_1 N_2}{B_1^2 B_2^2} + \frac{N_2 N_3}{B_2^2 B_3^2} + \frac{N_3 N_1}{B_3^2 B_1^2} \right). \end{aligned} \tag{18}$$

The symmetry of this result implies that,

$$\frac{B_1}{\sigma_{B_1}} = \frac{B_2}{\sigma_{B_2}} = \frac{B_3}{\sigma_{B_3}}. \tag{19}$$

Unfortunately the surprising result that the fractional error on each line is the same only holds in the three field case.

Now we wish to consider how the S/N on each B_i from the various cross-spectra is related to the S/N in the auto-spectrum case. We first write down the S/N from the auto-spectrum in the white-noise dominated regime:

$$\mathrm{S/N}(B_1, \text{auto-spectrum}) = \frac{B_1^2}{N_1} V_k \tag{20}$$

Now, let us compare the S/N of the two approaches in various cases. This will be simpler to write if we introduce $\tilde{B}_i \equiv B_i^2/N_i$. Then,

$$\frac{\mathrm{S/N}(B_1, \text{three-field})}{\mathrm{S/N}(B_1, \text{auto-spectrum})} = \frac{2}{\tilde{B}_1} \sqrt{\frac{\tilde{B}_1 \tilde{B}_2 \tilde{B}_3}{\tilde{B}_1 + \tilde{B}_2 + \tilde{B}_3}}. \tag{21}$$

There are several interesting cases to consider with this formula. For starters, let us assume that each field is measured to the same precision, i.e. $\tilde{B}_1 = \tilde{B}_2 = \tilde{B}_3 = \tilde{B}$. In this case, we have that

$$\frac{\mathrm{S/N}(B_1, \text{three-field})}{\mathrm{S/N}(B_1, \text{auto-spectrum})} = \frac{2}{\sqrt{3}} \approx 1.15. \tag{22}$$

So, the three-field approach gives a marginally better S/N than the auto-spectrum approach. This is encouraging, because it means that a tentative auto-spectrum measurement could be checked against the value derived from the

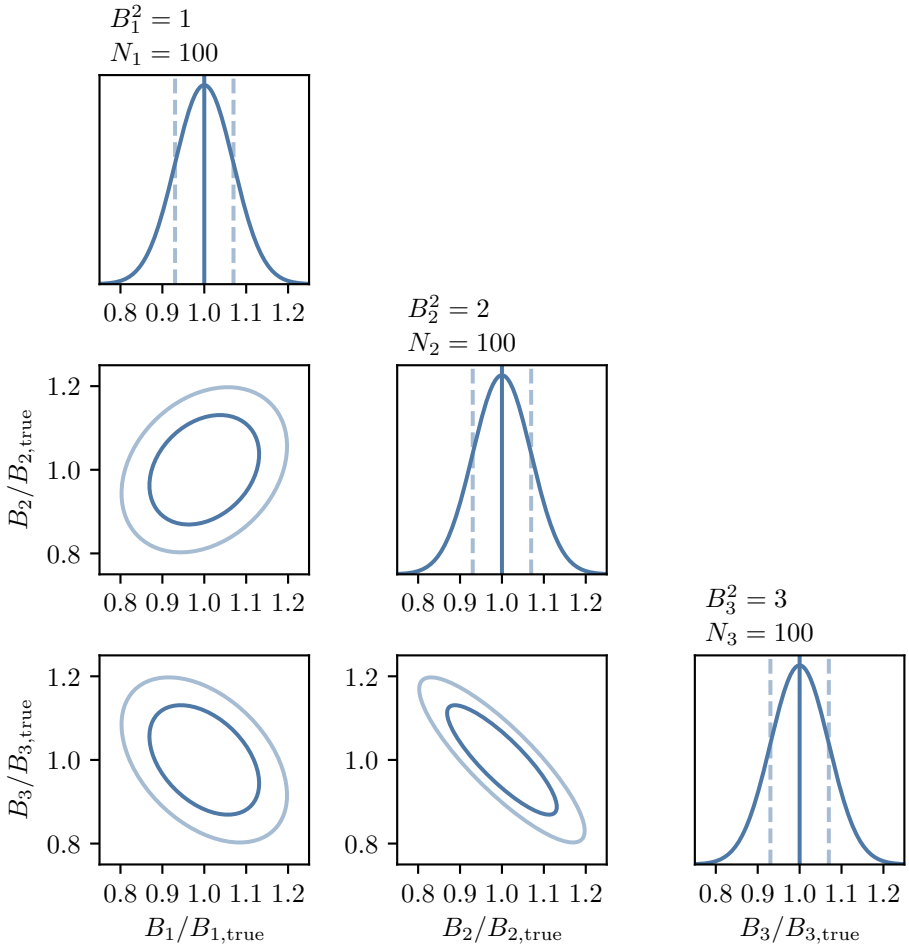

Figure 1. A toy corner plot showing the degeneracies between different lines with different intrinsic S/N ratios under the white-noise dominated assumption. We set $B_1^2 = 1$, $B_2^2 = 2$, and $B_3^2 = 3$ along with $N_1, N_2, N_3 = 100$ and $V_k^2 = 2 \times 10^6$. Since each line has the same value of N, the S/N of each line is driven by the value of B_i, with B_3 having the highest intrinsic S/N and B_1 the lowest. We plot each axis as a fraction of the true value for each of B_1, B_2, and B_3, and see that the marginalized posterior on each bias factor is the same (*diagonal panels*) as we expect from Equation 19. There is a strong negative covariance between B_2 and B_3, a mildly negative covariance between B_3 and B_1, and a mildly positive covariance between B_1 and B_2. An explanation for this behavior is given in the text.

three-field approach. Next, let us assume that the field 1 has a much stronger signal strength than 2 or 3 — i.e. $\tilde{B}_1 \gg \tilde{B}_2, \tilde{B}_3$. In this case we have that

$$\frac{\mathrm{S/N}(B_1, \text{three-field})}{\mathrm{S/N}(B_1, \text{auto-spectrum})} = 2\frac{\sqrt{\tilde{B}_2\tilde{B}_3}}{\tilde{B}_1}. \tag{23}$$

In this case we see that it is unlikely for the S/N on B_1 to be higher in the three-field case than in the auto-spectrum case. On the other hand, let us consider the case where one of the other lines is measured to much higher precision, e.g. $\tilde{B}_2 \gg \tilde{B}_1, \tilde{B}_3$. In this case, we have that

$$\frac{\mathrm{S/N}(B_1, \text{three-field})}{\mathrm{S/N}(B_1, \text{auto-spectrum})} = 2\sqrt{\frac{\tilde{B}_3}{\tilde{B}_1}}. \tag{24}$$

As long as field 1 and 3 are measured to comparable precision, a gain in the S/N of ~ 2 is possible with the three-field approach. Finally, let us consider the most interesting case in which fields 2 and 3 are measured to much higher

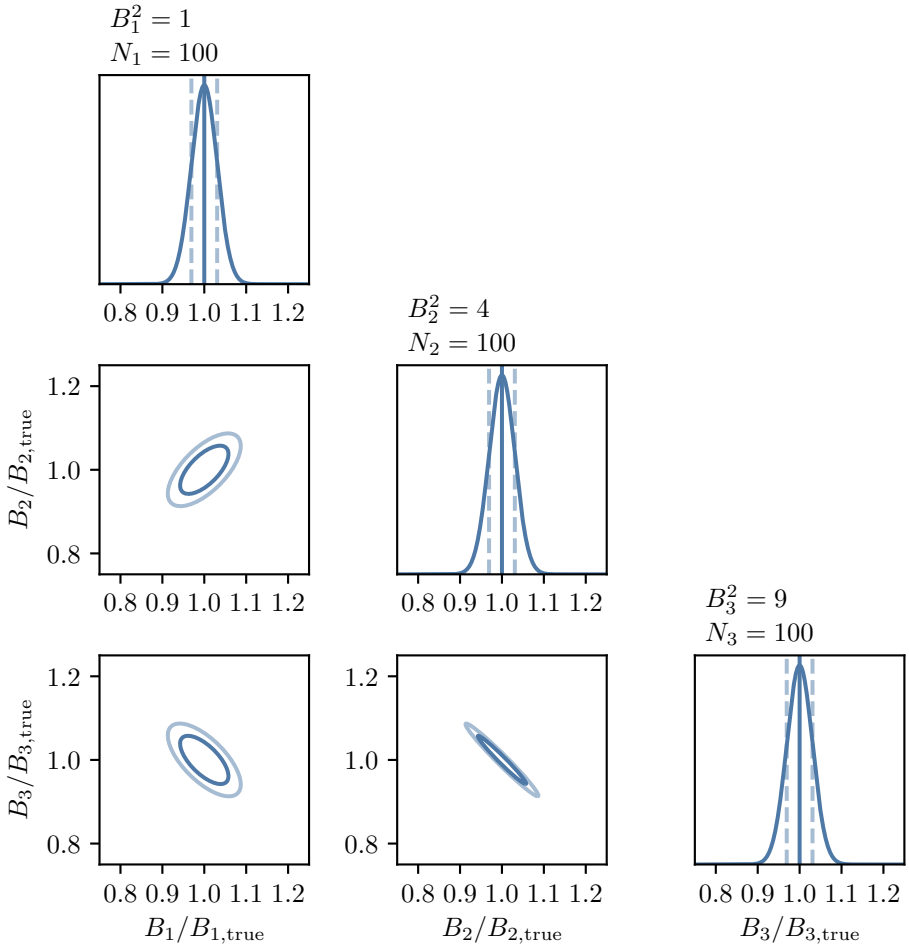

Figure 2. We perform the same procedure as in Figure 1, except now we slightly increase the intrinsic S/N of B_2 and B_3 by setting $B_2^2 = 4$ and $B_3^2 = 9$. The values for each N_i and V_k are the same as in Figure 1. We see that even though nothing about line 1 has changed (i.e. no additional measurements of line 1 have been made), the S/N on B_1 has improved because the fractional error on each line must be the same (Equation 19). The degeneracies between the different lines discussed in Figure 1 become stronger in this case.

precision than field 1 — i.e. $\tilde{B}_2, \tilde{B}_3 \gg \tilde{B}_1$. For simplicity let us also assume that $\tilde{B}_2 \approx \tilde{B}_3 = \tilde{B}$. In this case, we have that

$$\frac{\mathrm{S/N}(B_1, \text{three-field})}{\mathrm{S/N}(B_1, \text{auto-spectrum})} = \sqrt{2}\sqrt{\frac{\tilde{B}}{\tilde{B}_1}}. \tag{25}$$

As a result, if two lines are measured with strong precision, then we can strongly boost the S/N of a third, weakly measured line with the three-field approach.

As an illustration, we construct a corner plot in Figure 1 for the inferred bias values using toy values of $B_1^2 = 1$, $B_2^2 = 2$, and $B_3^2 = 3$ along with $N_1, N_2, N_3 = 100$ and $V_k^2 = 2 \times 10^6$. The units here are arbitrary. Since each line has the same value of N, the S/N of each line is driven by the value of B_i, with B_3 having the highest intrinsic S/N and B_1 the lowest. We make the same assumptions that go into Equation 15, i.e. that we are white-noise dominated in each line. By plotting each axis as a fraction of the true value for each of B_1, B_2, and B_3, we see that the marginalized posterior on each bias factor is the same (*diagonal panels*). This is equivalent to saying that the fractional error on each bias factor is the same for each line.

Moving on to the joint panels in Figure 1, we see that there is a strong negative covariance between B_2 and B_3, a mildly negative covariance between B_3 and B_1, and a mildly positive covariance between B_1 and B_2. We can

intuitively understand this behavior by considering that each cross-spectrum places a prior on the product B_1B_2, B_2B_3, and B_3B_1, with the width of that prior given by the S/N of that cross-spectrum.

In this toy scenario, the product B_2B_3 will have the highest S/N and therefore the tightest prior, and therefore to account for a larger B_3 one must have a smaller B_2 (and vice-versa), hence the strong negative correlation. We can make this argument for the relationship between B_2 and B_3 since B_2B_3 has the tightest prior, and therefore the products B_1B_2 and B_3B_1 are not relevant, since they have wider priors. The next tighest prior is on the product B_3B_1. Increasing the value of B_1 means that we must decrease the value of B_3 to keep B_3B_1 constant. Since B_2B_3 has a tighter prior, this means that we must also decrease the value of B_2. The prior on B_1B_2 is irrelevant, since it is wider than the other two.[1]

Finally, we turn to the product B_1B_2, which has the widest prior of the three products. If we increase B_1, then naively we would need to decrease B_2. However, increasing B_1 also means that we must decrease B_3 to keep B_3B_1 constant. Furthermore, decreasing B_3 means that we must increase B_2 to keep B_2B_3 constant. Since the priors on B_2B_3 and B_3B_1 are tighter than on B_1B_2, these constraints dominate, and as a result increasing B_1 means that we must increase B_2. Therefore, B_1 and B_2 are positively correlated.

In Figure 2 we perform the same procedure, except now we slightly increase the intrinsic S/N of B_2 and B_3 by setting $B_2^2 = 4$ and $B_3^2 = 9$. Notice that we have not changed the intrinsic S/N of field 1. Two things are apparent. First, the strength of the correlations that we discussed for Figure 1 are stronger, since the different priors have become more disparate in their relative tightness. Second, as a consequence of the fractional S/N being the same for each line, we see that the error on B_1 has become much smaller (and similarly for B_2 and B_3). However, this is particularly interesting in the case of B_1. No modification has been made to it's intrinsic S/N (or, in other words, no additional measurements have been made of line 1). Nonetheless, the S/N on the inferred B_1 has become higher.

4. RANDOM THOUGHTS

Like doing a matched filter??? Trying to understand intuitively why this works. If you have a super well measured field, then its like you're God who knows that the density field is. Its almost like using that density field as a matched filter?

APPENDIX

A. CO INTENSITY MODELLING

To model the statistical fluctuation and amplitude of the CO intensity maps we follow Lidz & Taylor (2016), and briefly summarize the procedure here. We begin by assuming a Schechter form for the star formation rate function,

$$\phi(\mathrm{SFR})\,\mathrm{dSFR} = \phi_* \left(\frac{\mathrm{SFR}}{\mathrm{SFR}_*}\right)^{\alpha} \exp\left[-\frac{\mathrm{SFR}}{\mathrm{SFR}_*}\right] \frac{\mathrm{dSFR}}{\mathrm{SFR}_*} \tag{A1}$$

where α is the faint-end slope, SFR_* is the characteristic star-formation rate, and ϕ_* is the characteristic number density. The average specific intensity in a line l can be related to the comoving emissivity by (Lidz et al. 2011; Pullen et al. 2013),

$$\langle I_l \rangle = \frac{\epsilon_l}{4\pi\nu_{\mathrm{rest},l}} \frac{c}{H(z)}, \tag{A2}$$

where $\nu_{\mathrm{rest},l}$ is the rest-frame emission frequency and ϵ_l is the comoving emissivity, each of line l. Assuming a linear luminosity-SFR relation with a constant of proportionality of L_0^l, along with the Schechter form for the SFR function (Equation A1), it follows that (Pullen et al. 2013),

$$\epsilon_l = \phi_* L_0^l \frac{\mathrm{SFR}_*}{1\,\mathrm{M}_\odot \mathrm{yr}^{-1}} \Gamma(2+\alpha). \tag{A3}$$

We adopt the value of L_0^{CO} for each line transition from Visbal & Loeb (2010).

REFERENCES

[1] Although in this case, to keep B_1B_2 constant while increasing B_1 we would need to decrease B_2 anyways.

Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463, doi: 10.1038/nature09187

Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165, doi: 10.3847/0004-637X/832/2/165

Chung, D. T. 2019, arXiv e-prints, arXiv:1905.00209. https://arxiv.org/abs/1905.00209

Diemer, B. 2018, ApJS, 239, 35, doi: 10.3847/1538-4365/aaee8c

Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605, doi: 10.1086/305424

Kaiser, N. 1987, MNRAS, 227, 1, doi: 10.1093/mnras/227.1.1

Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70, doi: 10.1088/0004-637X/741/2/70

Lidz, A., & Taylor, J. 2016, ApJ, 825, 143, doi: 10.3847/0004-637X/825/2/143

Linder, E. V. 2005, PhRvD, 72, 043529, doi: 10.1103/PhysRevD.72.043529

Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15, doi: 10.1088/0004-637X/768/1/15

Schechter, P. 1976, ApJ, 203, 297, doi: 10.1086/154079

Visbal, E., & Loeb, A. 2010, JCAP, 11, 016, doi: 10.1088/1475-7516/2010/11/016