Can small-scale baryon inhomogeneities resolve the Hubble tension? An investigation with ACT DR4

Leander Thiele,1 Yilun Guan,2 J. Colin Hill,3,4 Arthur Kosowsky,2 and David N. Spergel5,4

1Department of Physics, Princeton University, Princeton, NJ, USA 08544
2Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, USA 15260
3Department of Physics, Columbia University, 538 West 120th Street, New York, NY, USA 10027
4Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY, USA 10010
5Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ, USA 08544

Small-scale inhomogeneities in the baryon density around recombination have been proposed as a solution to the tension between local and global determinations of the Hubble constant. These baryon clumping models make distinct predictions for the cosmic microwave background anisotropy power spectra on small angular scales. We use recent data from the Atacama Cosmology Telescope to test these predictions. No evidence for baryon clumping is found, assuming a range of parameterizations for time-independent baryon density probability distribution functions. The inferred Hubble constant remains in significant tension with the SH0ES measurement.

I. INTRODUCTION

One cornerstone of the ΛCDM standard model of cosmology is the Hubble constant \( H_0 \). We may define it in two ways. In a model-agnostic way, \( H_0 \) measures the rate at which space is expanding at the present day. This definition enables us to determine \( H_0 \) locally by measuring the apparent speed of recession of galaxies in the Hubble flow. The other definition depends on the assumed cosmological model: it interprets \( c/H_0 \) as a measurement of the angular diameter distance to a given redshift, modulating the values of the background density parameters in the Friedmann equation. Measurements of \( H_0 \) according to this definition can be performed from the cosmic microwave background (CMB) power spectra or the late-time matter distribution. For both measurements, the baryonic acoustic oscillation (BAO) feature sets a characteristic length scale.

In the correct cosmological model the two definitions of the Hubble constant should be consistent, i.e., barring unknown systematics low- and high-redshift measurements should give consistent results. This, however, appears not to be the case, as is illustrated in Tab. I. Roughly, a > 4σ discrepancy between the low- and high-redshift inferences is observed.

Currently, the observed discrepancy is the most severe anomaly in the ΛCDM standard model, and hence the most promising avenue for the discovery of new physics in cosmology. Thus, it has spurred a wide range of attempts at an explanation (so many, in fact, that we now have dedicated reviews to keep track of them, e.g., [10]). Pre-recombination modifications include the Early Dark Energy models [11–16], which however are strongly constrained by large-scale structure data [17–19]. Additional light states may also be able to accommodate higher \( H_0 \), as long as their dynamics is non-trivial [20–22] (but see [23, 24]). Post-recombination modifications such as dynamical dark energy and modified gravity [25–30] or a fifth force [31] approach the tension from the opposite point of view, being not always physically well-motivated however, and often encountering problems with low-redshift expansion history data [32].

One attempted resolution of the \( H_0 \)-tension involves modifications of ΛCDM around the time of recombination that decrease the sound horizon \( r_s \). The peaks and troughs in the CMB power spectra are located at wave numbers scaling as \( (H_0 r_s)^{-1} \). Thus, a decrease in the sound horizon allows for higher values of \( H_0 \) as inferred

| dataset       | \( H_0 [\text{km/s/Mpc}] \) | type               |
|---------------|-----------------------------|--------------------|
| SH0ES [1]     | 73.2 ± 1.3                 | SNIa+Cepheids      |
| CSP [2]       | 69.6 ± 0.8 ± 1.7 DEB + TRGB|                   |
| MCP [3]       | 73.9 ± 3.0                 | VLBI geometric     |
| H0LiCOW [4]   | 73.3\(^{+1.1}_{-1.8}\)     | lensing time delay |
| TDCOSMO [5]   | 67.4\(^{+4.1}_{-3.2}\)     | lensing time delay |
| Planck [6]    | 67.4 ± 0.5                 | CMB                |
| ACT [7]       | 67.9 ± 1.5                 | CMB                |
| ACT+WMAP [7, 8]| 67.6 ± 1.1                 | CMB                |
| BOSS [9]      | 68.6 ± 1.1                 | BAO                |

Table I. Selection of recent \( H_0 \) measurements. See Sec. II for comment on the H0LiCOW measurement.
from the CMB. It should be noted that this picture is somewhat simplistic, since only changing $r_s$ introduces other tensions \cite{39}. The model considered in this work, however, has other secondary effects on the CMB power spectra that may be able to circumvent this constraint.

The mechanism to achieve the change in $r_s$ that we will focus on in this paper is the introduction of $O(1)$ inhomogeneities in the baryon density around the time of recombination. These inhomogeneities are required to be on comoving scales well below a Mpc so that they would not be directly visible in the currently available CMB data. Even in simple models of recombination physics it is clear that such small-scale baryon inhomogeneities will shift the surface of last scattering to higher redshift, thus decreasing the sound horizon. In fact, Peebles’s model \cite{41} contains a term $\dot{n}_e \propto -n_e^2$, which upon averaging gives faster recombination in inhomogeneous plasma on account of $(\langle n_e^2 \rangle > \langle n_e \rangle)^2$ \cite{41}.

One attractive proposal to achieve such a scenario is the introduction of primordial magnetic fields (PMFs) \cite{41,44}. This mechanism would have the advantage that it naturally sustains the required amount of small-scale inhomogeneity through a dynamical process, which in many other conceivable models would be entirely erased through Silk damping. Furthermore, theoretically well-motivated PMF configurations would lead to inhomogeneities of roughly the required amount on kpc scales. The necessary field strengths would be such that present-day large-scale magnetic fields could be explained with limited invocation of galactic dynamo effects. One open question that could be answered through MHD simulations is the effect of an inverse energy cascade from the turbulent small-scale magnetic fields and baryon inhomogeneities towards larger scales. If such an energy flow was appreciable, it would add additional constraints. We emphasize that this work deals with baryon inhomogeneities in an agnostic way and makes no reference to PMFs; however, it seems difficult to generate the required inhomogeneities without invoking PMFs. Furthermore, most other changes to recombination physics that would lead to appreciable shifts in $r_s$ are difficult to reconcile with atomic physics constraints \cite{45}.

While previous work – Ref. \cite{44}, hereafter JP20 – restricted CMB data to the results from the Planck satellite \cite{6}, this paper will include more recent data from the Atacama Cosmology Telescope (ACT), specifically their data release 4 (DR4) \cite{7,16}. Although the angular scales covered by Planck enable tight constraints on the location of the acoustic peaks (and thus on $H_0 r_s$), they are less constraining with regard to a secondary prediction of the inhomogeneous models, namely the change in amplitude in the Silk damping tail \cite{41}. While faster recombination decreases the time photons have to diffuse (thus suppressing Silk damping), helium recombination is also sped up which decreases the ionization fraction and therefore increases the photon mean free path (thus enhancing Silk damping). Generically, these two competing effects will not completely cancel; hence the inhomogeneous models have a characteristic signature on small angular scales that is not degenerate with the shift in peak positions. For this reason, the inclusion of small-scale ACT data provides valuable additional constraints on the proposed resolution of the $H_0$ tension. We may go even further and posit that the new ACT data covering scales that were not considered in JP20 provide an independent test of the secondary small-scale effects predicted by this model. If the ACT data do not show evidence of these effects, then the probability of this model resolving the Hubble tension is accordingly lessened, compared to a hypothetical additional data set that covers the same scales as Planck.

As the baryon inhomogeneities are on scales $10^3$ times below current resolution, their effect is completely described by the baryon density probability distribution function (PDF). While future MHD simulations assuming PMFs could provide constraints on the shape of this PDF, in this work we treat it as physical input. This has the advantage that we can treat the baryon inhomogeneities agnostically with regard to their origin, but we will have to make certain assumptions on the shape of the PDF in order to retain a predictive model. Future work will hopefully be able to give physically better motivated parameterizations of the PDF; however, we believe that the range of PDFs considered in this work provides strong motivation to believe that our results are in fact generic.

The rest of this paper is structured as follows. In Sec. II we review the three-zone model for the baryon density PDF from JP20 that we adopt in this work. Furthermore, we explain our choice of likelihoods and summarize the methods used. Sec. III contains the main results of this work, namely posteriors on $H_0$ and various parameters of the three-zone model for a variety of likelihoods. In Sec. IV we add local $H_0$ measurements to the CMB likelihoods in order to fully excite the baryon clumping mode and derive some intuition from the resulting best-fit models. We conclude in Sec. V.

II. SETUP

Following JP20, we parameterize the baryon density PDF as a three-zone model, where each zone is described by its density $\Delta_i = n_i^3/(n_0)$ and its volume fraction $f_i$ ($i = 1\ldots3$). We can think of the comoving coordinates being split into small regions, each of which realizes one
of the three zones.

We introduce the ‘clumping factor’ $b$ as

$$1 + b = \sum_i f_i \Delta_i^2 \quad (1)$$

and remind the reader of the constraints $\Sigma_i f_i = 1$ and $\Sigma_i f_i \Delta_i = 1$. The baryon clumping models are defined by four parameters, because there are three zones, each with density and volume fraction, and two constraints between these quantities. For physical convenience, we choose the 4-parameter model space to be spanned by $f_2$, $\Delta_1$, $\Delta_2$, and the clumping factor $b$ defined in Eq. (1).

In this paper, for ease of calculation and comparison with [JP20] we consider several subsets of these models, with either one or two of the four parameters being free and the others fixed to specific values, as listed in Tab. [II]. These likely bracket all of the reasonable physical cases, but a more extensive analysis could consider 4-parameter models instead.

While models M1 and M2 were chosen to facilitate direct comparison with [JP20] the more extended models ($f_2 = \ldots$) are constructed so as to cover a relatively wide range of possible baryon density PDFs while also being convenient computationally. They also constrain the variations of the baryon density to $O(1)$ excursions, consistent with the prediction of the PMF model.

We use two combinations of Boltzmann codes, recombination models, and Monte Carlo samplers to independently arrive at our results, namely CAMB [17] + RECFAST [48–50] + Cobaya [51, 52] and CLASS [53, 54] + HyRec [55] + MontePython [56, 57]. As previous work on our topic had relied on RECFAST which was calibrated for models close to $\Lambda$CDM and, certainly not for $O(1)$ deviations from the average baryon density, an independent check with the more sophisticated HyRec code is reassuring. All figures in this work were produced with the second toolchain. We run CLASS with the provided cl_permille.pre file instead of the default precision settings, and second Ref. [58] in recommending this as good practice, given the high precision of current and upcoming CMB data. Our convergence criterion is for the Gelman-Rubin $R_\alpha$ to achieve $\max_\alpha (R_\alpha - 1) < 0.05$ ($\alpha$ labels the parameters), and we generally achieve $R_{H_0} - 1 \lesssim 0.02$.

In both codes, we separately compute the recombination history $x_e^i(z)$ in each of the three zones and then perform the average to arrive at the large-scale ionization fraction

$$\langle x_e(z) \rangle = \sum_i f_i \Delta_i x_e^i(z). \quad (2)$$

We use the following likelihoods:

- **Planck 2018**: high-$\ell$ TTTEEE (lite), low-$\ell$ TT, low-$\ell$ EE [59];
- **ACT DR4**: high-$\ell$ TT, EE, TE [7, 46];
- **SMH**: SH0ES [60], MCP [3], H0LiCOW [4];
- **BAO**: SDSS DR7 [61] & BOSS DR9 [62], 6dFGS [63].

We found that working with the extended Planck likelihood (as opposed to the lite version) did not give appreciable changes in the posteriors in our parameter space.

The ACT likelihood is chosen such that covariance with Planck is negligible, with $\ell < 1800$ discarded in TT but no cuts in TE and EE [7].

The H0LiCOW measurement used in the SMH likelihood is not the most recent one; an updated lens model found substantially lower $H_0$ [5]. Similarly, we do not use the most recent SH0ES measurement [4] but rather the one used in [JP20] from Ref. [60]. The reason for this is that we use the SMH combination of likelihoods only for direct comparison to [JP20] (equal to their “H3”) as well as in contexts in which we simply want to excite the

1 Not all parameterizations are equally good, some will only allow narrow windows in the free parameters in which the constraint system has solutions at all, which complicates the sampling and interpretation.

2 the latter with the flag MODEL=FULL instead of the default RECFAST that ships with CLASS.

We performed a test with the default precision file on the $f_2 = 1/2$ model. In that case, there was a negligible $\sim 0.1\sigma$ bias on $H_0$, while the shape of the $b$-posterior was slightly but visibly altered.
baryon clumping mode to a considerable extent (i.e., not in posteriors).

We assume flat priors on $\omega_b$, $\omega_{\text{CDM}}$, $n_s$, $A_s$, $H_0$, $\tau_{\text{reio}}$, $A_{\text{Planck}}$ and the hard prior $0.9 < y_p < 1.1$ for the only ACT nuisance parameter (polarization efficiency). When they are used, the three-zone model parameters have the wide priors $0 < b < 10$ and $0 < \Delta_1 < 0.7$ (the upper bound on $\Delta_1$ is important as it removes the spurious second posterior mode that would otherwise be allowed by inverting the first and third zone). We assume a fixed neutrino mass sum of $0.06 \text{eV}$ (this helps us focus on the degeneracy between $H_0$ and baryon clumping, rather than having to worry about the $H_0 - \nu$ degeneracy in addition).

III. RESULTS: POSTERIORS

As a first step, we reproduce the results from [JP20], staying within their M1/M2 parameterization of the three-zone model and working solely with the $\text{Planck}+\text{SMH}$ likelihood (which we adopt only in order to reproduce previous results and assess the accuracy of $\text{RECFAST}$ in the baryon clumping models). The resulting posteriors are shown in Fig. 1. We observe that our posteriors on $H_0$ are almost identical with those found in [JP20] (which are plotted as dashed lines). There are some minor discrepancies in the posterior on the clumping parameter $b$, which we believe are most likely explained by differences between $\text{HyRec}$ and $\text{RECFAST}$.

In the remainder of this section, we will not include SMH (i.e. SH0ES, MCP, and the old $H_0\text{LiCOW}$ value) in the likelihoods. As we will see, the CMB(+BAO) likelihoods are not compatible with SMH even after allowing for baryon clumping, which in our view renders a formally combined likelihood problematic for interpretation.

In Fig. 2, we compare posteriors in the [JP20] parameterizations fitted to $\text{Planck}$ and $\text{Planck}+\text{ACT}$. For these posteriors, as well as for all the others plotted in this section, central values and 1-$\sigma$ uncertainties are listed in Tab. V in Appendix B. First, we observe that within $\Lambda\text{CDM}$ both likelihoods yield very similar posteriors on $H_0$ (this is to be expected, as ACT DR4 itself has only about 1/3 the

---

As mentioned in Sec. IV, we confirmed key results with the toolchain used in [JP20]. In that case, we find almost perfect agreement with their results.
statistical power of Planck, and moreover their ΛCDM constraints independently agree well \cite{2105.06343}). However, in the models allowing baryon clumping, M1 and M2, the situation is different. While the $H_0$ posterior’s lower end is virtually identical for both likelihoods, the inclusion of ACT significantly reduces the extent of the high-$H_0$ tail (due to this asymmetry, the posterior plot is more illuminating than the bare 1-$\sigma$ uncertainties). The effect is more pronounced for model M1 which yields a less severe tension with SH0ES than M2 does. Similarly, adding ACT reduces the 95% upper limits on the clumping parameter $b$ (c.f. Tab.\ V). Although M1 and M2 consistently lead to slightly higher central values of $H_0$, the shift is not nearly enough to get in reasonable agreement with the SH0ES measurement. In summary, at least in the JP20 parameterizations M1/M2, adding ACT to Planck increases the tension with SH0ES within the baryon clumping models.

We proceed by adding an additional degree of freedom to the baryon density PDF, treating $\Delta_1$ as a free parameter. The models considered here may be considered as spanning the entire range of which M1 and M2 sample only two discrete points, while also sampling $f_2$ (which was identical for M1 and M2) at three different points. Posteriors for the Planck and Planck+ACT likelihoods are shown in Fig.\ 3. We observe extremely similar posteriors in $H_0$ and $b$ for the three flavors of this extended parameterization. Again, no detection of the clumping effect (i.e., non-zero $b$) is made. Similarly to the posteriors shown in Fig.\ 2, the addition of ACT has the most pronounced effect in the high-$H_0$ tails, where it visibly reduces the leeway the baryon clumping models have in getting into somewhat decent agreement with SH0ES.

The extra degree of freedom introduced in these models, $\Delta_1$, is virtually unconstrained (for this reason, we do not include it in Tab.\ V). This may be a consequence of a poor choice of parameterization, however, it seems more likely that we are already close to covering the entire range of CMB spectra compatible with the data and allowed by the three-zone model. This view is supported by the fact that all three values of $f_2$ lead to extremely similar posteriors. If this is indeed the case, the $H_0$ posterior shown in Fig.\ 3 should be similar to those that could be obtained with any reasonable baryon density PDF. As we discuss in the conclusions, time dependence may prove to be an additional mode that could alter this picture.

We discuss a possible prior volume effect in the more extended $f_2 = \ldots$ models in Appendix\ A.

As a final result of this section, we show in Fig.\ 4 how adding $z \lesssim 1$ BAO data affects the posteriors in the extended models. As for $H_0$ the BAO likelihood has comparable statistical power to the CMB data, the observation that the low-$H_0$ tail is reduced by the addition of BAO is no surprise. On the other hand, the high-$H_0$ tail is almost invariant under adding BAO, which indicates that the CMB data alone are almost entirely responsible for constraining the baryon clumping modes of the model.
Planck

\[
\begin{array}{cccc}
0.3333 & 0.5000 & 2.2316 & 2.2561 \\
2.2511 & 2.2520 & 1.0000 & 1.0000 \\
3000 & 2500 & 1400 & 900 \\
1100 & 1000 & 0.9660 & 0.9742 \\
900 & 1.1822 & 1200 & 2.1086 \\
1200 & 2.1282 & 2.0940 & 2.0617 \\
& & 0.2865 & 0.3524 \\
& & 0.3801 & 0.1476 \\
& & 1.6784 & 3.3757 \\
\end{array}
\]

\[\Lambda_{\text{ionization fraction}} x_e (z)\]

\[
\begin{align*}
\text{In fact, the posterior on } b \text{ is slightly widened once BAO is included. Thus, we conclude that the BAO likelihood is not particularly useful in constraining the baryon clumping models, at least as long as the CMB-only likelihood does not allow appreciable increases in } H_0. \\

\text{IV. RESULTS: BEST-FIT MODELS}

In order to gain some further intuition, in this section we attempt to fully excite the baryon clumping mode and explore its consequences for the CMB spectra. In order to do this, we now add the previously discussed SMH likelihood to Planck+ACT. For this combined likelihood, we compute three best-fit models, listed in Tab. III (the model in the first column, Planck+ACT CDM, was not fitted to SMH and serves as the reference point). The models including baryon clumping, M1 and \( f_1 = 1/2 \), were chosen because they are the ones that had the highest centrval values for \( H_0 \) in the posteriors discussed in the previous section. We emphasize that it can be difficult to infer general trends from individual best-fit models, so the results in this section should be taken as starting points to develop intuition. The main results of this work have been presented in the previous section.

For each of the best-fit models, we compute the average ionization fraction \( x_e (z) \), shown in Fig. 5. In the figure, we also list \( z_{0.5} \) and \( \Delta z_{0.9 \rightarrow 0.1} \) as proxies for the location and width of the last scattering surface. First, we observe that without inclusion of baryon inhomogeneities the recombination history is essentially fixed. As expected, once the additional degrees of freedom in the baryon density PDF are introduced, recombination is pushed to higher redshift. If we write \( A \delta \log \Delta z_{0.9 \rightarrow 0.1} = \delta \log z_{0.5} \) (where \( \delta \) is with respect to the Planck+ACT CDM reference model), we find \( A = -1.60 \) and \( A = -3.53 \) for M1 and \( f_2 = 1/2 \) respectively. This is, to order of magnitude, consistent with \( A = -2 \), the expectation from the argument that the duration of recombination should scale as \( v_{\text{rec}}^{-1} \propto T^{-1/2} \propto z^{-1/2} \). However, it is interesting to note that both M1 and \( f_2 = 1/2 \) yield almost identical values for \( \Delta z_{0.9 \rightarrow 0.1} \).

\[\text{Figure 5. The ionization fraction as a function of redshift for the best-fit models listed in Tab. III}\]

5 The ACT DR4 data only contain band powers, which are related to the underlying CMB power spectra by \( C_{\ell} = W_{\ell} C_{\ell} \), where \( W_{\ell} \) is an m \( \times \) n matrix with m \( < \) n. Thus, it is not possible to uniquely infer the \( C_{\ell} \) from the band powers. In order to generate power spectra that can be directly compared to Planck and the theory models, we first find some solution \( C_{\ell} \) of the above equation (which is, of course, completely unphysical). Then, writing \( K_{\alpha} \) for the n \( \times \) (n + m) matrix whose columns are a basis for the kernel of \( W_{\ell} \), we can generate additional solutions \( C_{\ell}(a_{\alpha}) = C_{\ell} + K_{\alpha} a_{\alpha} \), where \( a_{\alpha} \) is an arbitrary vector. We now find the \( a_{\alpha} \) that minimizes \( ||\Lambda_{\alpha} C_{\ell}(a_{\alpha})|| \), where \( \Lambda_{\alpha} \) implements a discrete second derivative. We perform this procedure for both the ‘deep’ and ‘wide’ parts of the ACT data set, associating approximate error bars with the resulting \( C_{\ell} \) according to the normalized diagonal elements of the covariance matrix of the
Figure 6. Fractional changes in CMB power spectra for the best-fit models listed in Tab. III with respect to the Planck+ACT ΛCDM baseline. In the second panel (TE), because the data have zero crossings, we computed the normalizing factor $|C_\ell^\text{ref}|$ as a convolution of the absolute value of the reference TE power spectrum with a Gaussian filter of width $\sigma = 100$ (in $\ell$). This gives a useful normalization without artifacts from the roots. Note that for clarity the experimental data curves for ACT and Planck have been rescaled by the given factors. While the Planck data points are relatively close to what we use in our likelihood (except for binning and $\ell$-cuts to improve clarity, in particular for TE and EE), the ACT data points are for illustration purposes only and were generated as explained in the footnote.

Next, we consider the CMB power spectra arising from the listed best-fit models, shown in Fig. 5 where we compare them to Planck and ACT data. As already alluded to in Sec. III, the models with baryon inhomogeneities show significant deviations from ΛCDM in the high $\ell$ part of the spectra, particularly for TT and EE. These deviations are partly caused by shifts in other parameters, as is already evident from the orange curve depicting ΛCDM fit to Planck+ACT+SMH. However, in simple experiments where we put baryon inhomogeneities of magnitude comparable to the models shown here into otherwise unchanged ΛCDM cosmologies, similar trends occur (the same happens if $H_0$ is increased so as to neutralize the shift in peak positions). In particular, we always observe an increase in TT power at high-$\ell$ once the baryon inhomogeneities are introduced. Thus, we conclude that the effect of Silk damping is in fact reduced in the inhomogeneous models. This is contrary to previous statements in the literature, which presumably rested on numerical experiments made with much higher amounts of inhomogeneity.

It is interesting to disentangle how the suppression in the $\ell < 1000$ part of the EE spectrum is generated. The ratio of EE to TT for $\ell \gg 1$ can be approximated as

$$
\frac{C_{\ell}^{EE}}{C_{\ell}^{TT}} = \left( \frac{\eta_* \Delta \eta_*}{\sqrt{2} (1 + R)} \right)^2 B_\ell,
$$

where $\eta_*$ and $\Delta \eta_*$ are the location and thickness of the last scattering surface in conformal time respectively, $r$ is its distance from us, $R = 3 \rho_0/4 \rho_\gamma$, and $B_\ell$ is a geometric factor independent of cosmology. Comparing, for example, the best-fit models Planck+ACT ΛCDM and Planck+ACT+SMH $f_2 = 1/2$, we find that the differences in the various quantities entering Eq. (3) are all $\sim 2 \ldots 4\%$. On aggregate, however, the cancellation gives a $\sim 1\%$ effect.

The ACT data do not seem to show a preference for the general trends in the spectra introduced by the baryon inhomogeneities.

Finally, we present the changes in $\chi^2$ relative to the Planck+ACT ΛCDM baseline for the other three best-fit models, listed in Tab. IV. The model M1 increases $\chi^2$ by about equal amounts in Planck and ACT, while the model $f_2 = 1/2$ with one extra degree of freedom in the baryon density PDF is comparably more disfavoured by Planck than by ACT. ACT TT actually appears to favour the baryon clumping models, which holds true for both the entire $\ell$-range covered by ACT and $\ell > 1800$, the part used in our likelihoods. On the other hand, ACT TE and, to a lesser extent, EE, disfavour baryon clumping. This may be explained by the inability of those models to keep the last scattering surface’s thickness fixed (c.f. Fig. 5). Another possibility is simply that the steeper slopes in the polarization spectra translate into higher sensitivity to $H_0$. It should be mentioned that there have been hints of inconsistencies between ACT TE and other CMB data.

As no mechanism has been proposed that could explain these trends, and the tensions are still relatively mild, we do not see this work as the appropriate place to investigate them further.

$C_\ell$. Finally, we take the weighted average of the ‘deep’ and ‘wide’ $C_\ell$. 
The introduction of inhomogeneities in the baryon density around the time of recombination is an attractive proposal to resolve the Hubble tension. While we do not constrain ourselves to a particular mechanism that is supposed to generate the required small-scale power, PMFs appear to be the most natural candidate. A key prediction of the inhomogeneous models is, besides the geometric shift in acoustic peak positions, a change in the high-$\ell$ part of the CMB power spectra due to different amount of Silk damping. We have argued that the addition of small-scale ACT data to the previously considered Planck likelihood constitutes a powerful way to test this prediction. The higher statistical power of the thus enlarged likelihood also enables us to extend the parameter space; specifically, we are able to add an additional degree of freedom to the baryon density PDF.

Fitting the baryon clumping models to CMB-only likelihoods, we have shown that the inclusion of ACT narrows the posterior for the clumping parameter $b$ around zero compared to Planck alone. Likewise, although ACT does not appreciably alter the low-$H_0$ tails in the posteriors, it significantly shrinks the high-$H_0$ tails, indicating that the added small-scale information disfavors the introduction of baryon clumping. The CMB data still leave room for mild baryon clumping and some upward shift in $H_0$, thus somewhat relieving the Hubble tension, but values as high as the SH0ES 1-$\sigma$ lower bound are strongly disfavored.

Although the absence of theoretical predictions for the baryon density PDF restricts us to relatively simple parameterizations, we have argued that there is evidence that the posteriors on $H_0$ presented in this work are in fact generic for any time-independent baryon density PDF. An intriguing possibility that could evade the constraints presented would be the introduction of time dependence. Assuming the existence of PMFs, the decrease in magnetic field strength during recombination would lead to a homogenization of the baryon distribution (the relaxation time for kpc-scale baryon clumps is considerably smaller than recombination’s duration). It is possible that, in the large-scale average, this would lead to a faster onset of recombination while keeping the thickness of the last scattering surface approximately fixed, perhaps leading to a better fit with the polarization data.

Besides our primary results, we have also shown that while the previously used RECFAST recombination code is probably sufficiently accurate for the modifications considered in this work, there are some systematic differences to the more recent HyRec which we recommend should be used going forward. Furthermore, we have shown that, contrary to previous claims, baryon inhomogeneities of the amount necessary to appreciably shift $H_0$ do not generically lead to enhanced Silk damping.

We have found no evidence that the addition of small-scale ACT data to the likelihood favours the introduction of baryon inhomogeneities around recombination over $\Lambda$CDM. Looking forward, the increased small-scale CMB information from future ACT releases, SPT-3G [65], Simons Observatory [66], and CMB-S4 [67] will provide more constraining power. As long as no currently unknown systematics in ACT are discovered, however, smaller error bars will only serve to exclude the class of time-independent three-zone models considered in this paper with higher significance. Parallel to this improvement in data quality, theoretical advancements in constructing viable models of the baryon density PDF, for example through MHD simulations, will be essential in increasing the predictivity of the theory.

Modifications to CLASS used in this work are publicly available [here].

**ACKNOWLEDGMENTS**

We thank Karsten Jedamzik and Levon Pogosian for useful discussions during the final stages of this work. JCH and DNS thank the Simons Foundation for support. LT thanks Zack Li for help with the ACT DR4 likelihood.
[1] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scotic, Astrophys. J. Lett. 908, L6 (2021), arXiv:2012.08534 [astro-ph.CO].
[2] W. L. Freedman, B. F. Madore, T. Hoyt, I. S. Jang, R. Beaton, M. G. Lee, A. Monson, J. Neeley, and J. Rich, Astrophys. J. 891, 57 (2020) arXiv:2002.01550 [astro-ph.CO].
[3] M. J. Reid, J. A. Braatz, J. J. Condon, L. J. Greenhill, C. Henkel, and K. Y. Lo, Astrophys. J. 695, 287 (2009) arXiv:0811.4345 [astro-ph].
[4] K. C. Wong, S. H. Suyu, G. C. F. Chen, C. E. Rusu, M. Millon, D. Sluse, V. Bonvin, C. D. Fassnacht, S. Taubenberger, M. W. Auger, S. Birrer, J. H. H. Chan, F. Courbin, S. Hilbert, O. Tihhonova, T. Treu, A. Agnello, X. Ding, I. Jee, E. Komatsu, A. J. Shajib, A. Sonnenfeld, R. D. Blandford, L. V. E. Koopmans, P. J. Marshall, and G. Meylan, Mon. Not. Roy. Astron. Soc. 498, 1420 (2020) arXiv:1907.04869 [astro-ph.CO].
[5] S. Birrer, A. J. Shajib, A. Galan, M. Millon, T. Treu, A. Agnello, M. Auger, G. C. F. Chen, L. Christensen, T. Collett, F. Courbin, C. D. Fassnacht, L. V. E. Koopmans, P. J. Marshall, J. W. Park, C. E. Rusu, D. Sluse, C. Spiniello, S. H. Suyu, S. Wagner-Carena, K. C. Wong, M. Barnabè, A. S. Bolton, O. Czoske, X. Ding, J. A. Frieman, and L. Van de Vyvere, Astron. Astrophys. 643, A165 (2020) arXiv:2007.02941 [astro-ph.CO].
[6] Planck Collaboration, Astron. Astrophys. 641, A6 (2020) arXiv:1807.06209 [astro-ph.CO].
[7] ACT Collaboration, S. Aiolà, E. Calabrese, L. Maurin, S. Naess, B. L. Schmitt, et al., J. Cosmol. Astropart. Phys. 2020, 047 (2020) arXiv:2007.07288 [astro-ph.CO].
[8] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright, Astrophys. J. 695, 76 (2009) arXiv:0711.0993 [astro-ph].
[9] O. H. E. Philcox, M. M. Ivanov, M. Simonović, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 2020, 032 (2020) arXiv:2002.04035 [astro-ph.CO].
[10] L. Knox and M. Millea, Phys. Rev. D 101, 043533 (2020) arXiv:1908.03663 [astro-ph.CO].
[11] T. Karwal and M. Kamionkowski, Phys. Rev. D 94, 103523 (2016) arXiv:1608.01309 [astro-ph.CO].
[12] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, Phys. Rev. Lett. 122, 221301 (2019) arXiv:1811.04083 [astro-ph.CO].
[13] P. Agrawal, F.-Y. Cyr-Racine, D. Pinner, and L. Randall, arXiv e-prints , arXiv:1904.01016 (2019), arXiv:1904.01016 [astro-ph.CO].
[14] M.-X. Lin, G. Benevento, W. Hu, and M. Raveri, Phys. Rev. D 100, 063542 (2019) arXiv:1905.12618 [astro-ph.CO].
[15] J. Sakstein and M. Trodden, Phys. Rev. Lett. 124, 161301 (2020) arXiv:1911.11760 [astro-ph.CO].
[16] G. Ye and Y.-S. Piao, Phys. Rev. D 101, 083507 (2020) arXiv:2001.02451 [astro-ph.CO].
[17] J. C. Hill, E. McDonough, M. W. Toomey, and S. Alexander, Phys. Rev. D 102, 043507 (2020) arXiv:2003.07355 [astro-ph.CO].
[18] M. M. Ivanov, E. McDonough, J. C. Hill, M. Simonović, M. W. Toomey, S. Alexander, and M. Zaldarriaga, Phys. Rev. D 102, 103502 (2020) arXiv:2006.11235 [astro-ph.CO].
[19] G. D'Amico, L. Senatore, P. Zhang, and H. Zheng, arXiv e-prints , arXiv:2006.12420 (2020) arXiv:2006.12420 [astro-ph.CO].
[20] F. D'Eramo, R. Z. Ferreira, A. Notari, and J. L. Bernal, J. Cosmol. Astropart. Phys. 2018, 014 (2018) arXiv:1808.07430 [hep-ph].
[21] C. D. Kreisch, F.-Y. Cyr-Racine, and O. Doré, Phys. Rev. D 101, 123505 (2020) arXiv:1902.00534 [astro-ph.CO].
[22] A. Das and S. Ghosh, arXiv e-prints , arXiv:2011.12315 (2020) arXiv:2011.12315 [astro-ph.CO].
[23] T. Brinckmann, J. Hyeok Chang, and M. LoVerde, arXiv e-prints , arXiv:2012.11830 (2020) arXiv:2012.11830 [astro-ph.CO].
[24] S. R. Choudhary, S. Hannestad, and T. Tram, arXiv e-prints , arXiv:2012.07519 (2020) arXiv:2012.07519 [astro-ph.CO].
[25] Q.-G. Huang and K. Wang, European Physical Journal C 76, 506 (2016) arXiv:1606.05965 [astro-ph.CO].
[26] J. Kenk, M. Zimalačáregui, F. Montanari, and A. Barreira, J. Cosmol. Astropart. Phys. 2017, 020 (2017) arXiv:1707.02263 [astro-ph.CO].
[27] G.-B. Zhao, M. Raveri, L. Pogosian, Y. Wang, R. G. Critten, W. J. Handley, W. J. Percival, F. Beutler, J. Brinkmann, C.-H. Chung, A. J. Cuesta, D. J. Eisenstein, F.-S. Kitaura, K. Koyama, B. L’Huillier, R. C. Nichol, M. M. Fieria, S. Rodriguez-Torres, A. J. Ross, G. Rossi, A. G. Sánchez, A. Shaflieo, J. L. Tinker, R. Tejeiro, J. A. Vazquez, and H. Zhang, Nature Astronomy 1, 627 (2017) arXiv:1701.08165 [astro-ph.CO].
[28] R. C. Nunes, J. Cosmol. Astropart. Phys. 2018, 052 (2018) arXiv:1802.02281 [gr-qc].
[29] E. Di Valentino, E. V. Linder, and A. Melchiorri, Phys. Rev. D 97, 043528 (2018) arXiv:1710.02153 [astro-ph.CO].
[30] Y. Wang, L. Pogosian, G.-B. Zhao, and A. Zucca, Astrophys. J. Lett. 869, L8 (2018) arXiv:1807.03772 [astro-ph.CO].
[31] N. Khosravi, S. Baghram, N. Afshordi, and N. Altamirano, Phys. Rev. D 99, 103526 (2019) arXiv:1710.09366 [astro-ph.CO].
[32] M. Raveri, Phys. Rev. D 101, 083524 (2020) arXiv:1902.01366 [astro-ph.CO].
[33] A. Banihashemi, N. Khosravi, and A. H. Shirazi, Phys. Rev. D 101, 123521 (2020) arXiv:1808.02472 [astro-ph.CO].
[34] Y.-F. Cai, M. Khurshudyan, and E. N. Saridakis, Astrophys. J. **888**, 62 (2020), arXiv:1907.10813 [astro-ph.CO].

[35] E. Di Valentino, A. Melchiorri, O. Mena, and S. Vagnozzi, Phys. Rev. D **101**, 063502 (2020), arXiv:1910.09853 [astro-ph.CO].

[36] M. G. Damotti, B. De Simone, T. Schiavone, G. Montani, E. Rinaldi, and G. Lambiase, arXiv e-prints , arXiv:2103.02117 (2021), arXiv:2103.02117 [astro-ph.CO].

[37] H. Desmond, B. Jain, and J. Sakstein, Phys. Rev. D **100**, 043537 (2019), arXiv:1907.03778 [astro-ph.CO].

[38] G. Efstathiou, arXiv e-prints, arXiv:2103.08723 (2021), arXiv:2103.08723 [astro-ph.CO].

[39] K. Jedamzik, L. Pogosian, and G.-B. Zhao, arXiv e-prints, arXiv:2010.04158 (2020), arXiv:2010.04158 [astro-ph.CO].

[40] P. J. E. Peebles, *Principles of Physical Cosmology* (1976).

[41] K. Jedamzik and T. Abel, arXiv e-prints, arXiv:1108.2517 (2011), arXiv:1108.2517 [astro-ph.CO].

[42] R. Banerjee and K. Jedamzik, Phys. Rev. D **70**, 123003 (2004), arXiv:astro-ph/0410032 [astro-ph].

[43] K. Jedamzik and A. Saveliev, Phys. Rev. Lett. **123**, 021301 (2019), arXiv:1804.06115 [astro-ph.CO].

[44] K. Jedamzik and L. Pogosian, Phys. Rev. Lett. **125**, 181302 (2020), arXiv:2004.09487 [astro-ph.CO].

[45] M. Liu, Z. Huang, X. Luo, H. Miao, N. K. Singh, and L. Huang, Science China Physics, Mechanics, and Astronomy **63**, 290405 (2020), arXiv:1912.00190 [astro-ph.CO].

[46] ACT Collaboration, S. K. Choi, M. Hasselfield, S.-P. Ho, B. Koopman, M. Luntz, et al., J. Cosmol. Astropart. Phys. **2020**, 045 (2020), arXiv:2007.07289 [astro-ph.CO].

[47] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].

[48] S. Seager, D. D. Sasselov, and D. Scott, Astrophys. J. Lett. **523**, L1 (1999), arXiv:astro-ph/9909275 [astro-ph].

[49] S. Seager, D. D. Sasselov, and D. Scott, Astrophys. J. Suppl. **128**, 407 (2000), arXiv:astro-ph/9912182 [astro-ph].

[50] W. Y. Wong, A. Moss, and D. Scott, Mon. Not. Roy. Astron. Soc. **386**, 1023 (2008), arXiv:0711.1357 [astro-ph].

[51] J. Torrado and A. Lewis, arXiv e-prints , arXiv:2005.05290 (2020), arXiv:2005.05290 [astro-ph.IM].

[52] J. Torrado and A. Lewis, “Cobaya: Bayesian analysis in cosmology,” (2019), ascl:1910.019.

[53] J. Lesgourgues, arXiv e-prints, arXiv:1104.2932 (2011), arXiv:1104.2932 [astro-ph.CO].

[54] D. Blas, J. Lesgourgues, and T. Tran, J. Cosmol. Astropart. Phys. **2011**, 034 (2011), arXiv:1104.2933 [astro-ph.CO].

[55] Y. Ali-Haïmoud and C. M. Hirata, Phys. Rev. D **83**, 043513 (2011), arXiv:1011.3758 [astro-ph.CO].

[56] T. Brinckmann and J. Lesgourgues, (2018), arXiv:1804.07261 [astro-ph.CO].

[57] B. Audren, J. Lesgourgues, K. Benabed, and S. Prunet, JCAP **1302**, 001 (2013), arXiv:1210.7183 [astro-ph.CO].

[58] F. McCarthy, J. C. Hill, and M. S. Madhavacheril, arXiv e-prints, arXiv:2103.05582 (2021), arXiv:2103.05582 [astro-ph.CO].

[59] Planck Collaboration, Astron. Astrophys. **641**, A5 (2020), doi:10.1051/0004-6361/201935752.

[60] M. J. Reid, D. W. Pesce, and A. G. Riess, Astrophys. J. Lett. **886**, L27 (2019), arXiv:1908.05625 [astro-ph.GA].

[61] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera, Mon. Not. Roy. Astron. Soc. **449**, 835 (2015), arXiv:1409.3242 [astro-ph.CO].

[62] SDSS-III Collaboration, Mon. Not. Roy. Astron. Soc. **427**, 3435 (2012), arXiv:1203.6594 [astro-ph.CO].

[63] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, Mon. Not. Roy. Astron. Soc. **416**, 3017 (2011), arXiv:1106.3366 [astro-ph.CO].

[64] M. Zaldarriaga and D. D. Harari, Phys. Rev. D **52**, 3276 (1995), arXiv:astro-ph/9503108 [astro-ph].

[65] SPT-3G Collaboration, B. A. Benson, et al., in *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9153, edited by W. S. Holland and J. Zmuidzinas (2014) p. 91531P, arXiv:1407.2973 [astro-ph.IM].

[66] Simons Observatory Collaboration, J. Cosmol. Astropart. Phys. **2019**, 056 (2019), arXiv:1808.07445 [astro-ph.CO].

[67] CMB-S4 Collaboration, arXiv e-prints, arXiv:1907.04473 (2019), arXiv:1907.04473 [astro-ph.IM].
Figure 7. Comparison between the posterior in the unconstrained \( f_2 = 1/2 \) model, same as in Fig. 3, and the one with the constraint \( f_1 > 0.1 \). The latter should substantially reduce the prior volume effect arising from the degeneracy with ΛCDM at \( f_1 = 0 \). In the dashed posterior, the non-vanishing support at \( b = 0 \) is an artifact of MontePython’s posterior smoothing routine.

Figure 8. Likelihood profiles as a function of \( b \) for the \( f_2 = 1/2 \) Markov chain. The ‘mean’ and ‘best-fit’ lines are computed as described in the text. The orange line is the same marginalized likelihood that was shown in Fig. 3. The blue dashed line is a Gaussian fit with only the standard deviation varied, while the green dashed line is a 3-parameter Gaussian fit. The green solid line was rescaled so that the Gaussian fit attains its maximum at one. These fits should only be treated as visual guides.

Appendix A: Assessment of prior volume effects

The more extended parameterizations considered in this paper, \( f_2 = \ldots \), exhibit a pathology in that \( f_1 = 0 \) is equivalent to ΛCDM regardless of the value of \( \Delta_1 \). Thus, there may be a prior volume effect in these models that underestimates the posterior for non-vanishing baryon clumping (and thus increased \( H_0 \)). We assess this effect for the example of the \( f_2 = 1/2 \) model; for other choices of \( f_2 \) the conclusions are similar. Our methods are similar to those used in Ref. [18].

First, we consider changing the model such that a constraint \( f_1 > 0.1 \) is enforced, thereby avoiding ΛCDM entirely. The resulting posteriors are shown in Fig. 7. It is expected that the posteriors for \( b \) and \( H_0 \) move upward to some extent since the model, by construction, contains at least some baryon clumping now. The shift in the central value of \( H_0 \) is about \( \sigma/3 \) and we do not observe a substantial enlargement of the high-\( H_0 \) tail.

Second, we compute a frequentist mean likelihood profile in \( b \), which is the exponential of the average log-likelihood extracted from the Markov chain in 20 \( b \)-bins between 0 and 1, shown as the solid blue line in Fig. 8. We observe that this profile is somewhat broader than the posterior shown in Fig. 3. Furthermore, at the 95 \% upper limit on \( b \) from Tab. V the averaged likelihood is reduced by about \( \Delta \chi^2 \sim 2.4 \). This value would correspond to only an \( \sim 88 \% \) upper bound.

We also compute the more common likelihood profile in which we minimize the likelihood at a range of values of \( b \), shown as the solid green line in Fig. 8. In this case, we observe that the maximum is at non-zero \( b \), but then the likelihood profile drops off more steeply than the marginalized Bayesian posterior. For a discussion of the relative merits of the two ways to compute a likelihood profile, see Ref. [18].

We conclude that we observe evidence for a mild prior volume effect. However, the posteriors in Fig. 7 indicate that our main conclusions regarding \( H_0 \) are only marginally affected.
### Appendix B: Central values and uncertainties

| likelihood     | model  | $H_0$ [km/s/Mpc] | $100\omega_b$ | $\omega_{\text{cdm}}$ | $n_s$ | $10^9 A_s$ | $\tau_{\text{reio}}$ | $b$          |
|----------------|--------|------------------|----------------|------------------------|-------|------------|------------------------|-------------|
| **Planck**     | ACDM   | 67.24 ± 0.66     | 2.23(2)        | 0.120(1)               | 0.964(5) | 2.10(4)   | 0.054(9)               | –           |
|                | M1     | 68.45 ± 1.08     | 2.24(2)        | 0.122(2)               | 0.960(5) | 2.10(4)   | 0.053(8)               | <0.57       |
|                | M2     | 67.86 ± 0.83     | 2.24(2)        | 0.121(2)               | 0.964(5) | 2.10(4)   | 0.055(9)               | <0.32       |
| $f_2 = 1/3$    |        | 68.04 ± 0.97     | 2.24(2)        | 0.121(2)               | 0.963(6) | 2.10(4)   | 0.054(9)               | <0.42       |
| $f_2 = 1/2$    |        | 68.12 ± 1.04     | 2.24(2)        | 0.121(2)               | 0.963(6) | 2.10(4)   | 0.054(9)               | <0.53       |
| $f_2 = 2/3$    |        | 68.03 ± 0.96     | 2.24(2)        | 0.121(2)               | 0.964(6) | 2.10(4)   | 0.054(9)               | <0.52       |
| **Planck+ACT** | ACDM   | 67.26 ± 0.60     | 2.23(1)        | 0.120(1)               | 0.967(4) | 2.12(4)   | 0.055(8)               | –           |
|                | M1     | 68.18 ± 0.87     | 2.23(1)        | 0.121(2)               | 0.962(5) | 2.11(4)   | 0.053(8)               | <0.42       |
| $f_2 = 1/3$    |        | 67.74 ± 0.71     | 2.24(1)        | 0.121(1)               | 0.967(4) | 2.12(4)   | 0.054(8)               | <0.26       |
| $f_2 = 1/2$    |        | 67.99 ± 0.86     | 2.23(1)        | 0.121(2)               | 0.964(6) | 2.11(4)   | 0.054(8)               | <0.39       |
| $f_2 = 2/3$    |        | 68.06 ± 0.88     | 2.23(1)        | 0.121(2)               | 0.964(6) | 2.11(4)   | 0.054(8)               | <0.47       |
| **Planck+ACT+BAO** | M2     | 68.01 ± 0.89     | 2.24(2)        | 0.121(2)               | 0.965(5) | 2.11(4)   | 0.054(8)               | <0.49       |
| $f_2 = 1/3$    |        | 68.18 ± 0.77     | 2.24(2)        | 0.121(2)               | 0.964(6) | 2.11(4)   | 0.054(8)               | <0.38       |
| $f_2 = 1/2$    |        | 68.24 ± 0.83     | 2.24(1)        | 0.121(2)               | 0.964(6) | 2.11(4)   | 0.054(8)               | <0.46       |
| $f_2 = 2/3$    |        | 68.22 ± 0.79     | 2.24(1)        | 0.121(2)               | 0.966(5) | 2.11(4)   | 0.054(9)               | <0.50       |

Table V. Central values and 1-σ errors for the posteriors in this section. For conciseness, for $\Lambda$CDM parameters except $H_0$, the numbers in parentheses are the 1-σ uncertainties on the last digit. Since we do not observe a detection of the clumping effect, for the clumping parameter $b$ 95% upper limits are quoted.