

Skeptical review: Multi-Frequency Analysis of the ACT DR6 Thermal Sunyaev–Zel’dovich Maps: Catalog Properties, Spectral Diagnostics, and Statistical Characterization of the Temperature Field

Summary

The paper presents an exploratory, multi-frequency analysis of ACT DR6 temperature and polarization maps together with the ACT–Planck NILC Compton- y map. The authors construct a simple 5σ , single-scale threshold catalog of 200 high-significance tSZ candidates from the NILC y -map (Sec. III), validate the procedure via recovery of the Bullet Cluster, and qualitatively characterize the one-point statistics and strong spatial noise inhomogeneity of the y -field (Sec. IV). They then perform intentionally simple spectral diagnostics at the positions of the top y -selected candidates using 90/150/220 GHz temperature maps (Sec. V), study a compact, high- y -significance source at $(291.2^\circ, -29.2^\circ)$ with a synchrotron-like spectrum (Sec. VI), and investigate non-Gaussianity, hemispherical power asymmetry, and cross-frequency correlations of the temperature field (Sec. VII). A radial-profile analysis highlights five clusters with anomalous morphologies (Sec. VIII), while a companion cross-frequency coherence study is summarized to motivate recommended ℓ -scale cuts (Sec. IX).

The authors further present a set of null and consistency tests, including split-map checks, TB/EB null tests, birefringence and isocurvature constraints, polarization cross-frequency agreement, and an rSZ– y sign comparison (Sec. X). They discuss look-elsewhere effects and trial factors when interpreting several “anomalous” findings—extreme excess kurtosis ($\kappa \approx 47$) in the f150 temperature field, four low cross-frequency correlation cells, and five profile outliers (Secs. VII–VIII, XI–XII). The work emphasizes that methods are deliberately simple and intended as diagnostics and follow-up target identification rather than precision cosmology.

Overall, the paper offers a broad, end-to-end “stress test” of publicly available ACT DR6 products and flags specific sky locations and phenomena for further study. However, several analyses are methodologically minimal (global-threshold source finding, single-pixel spectra, limited simulations), and quantitative claims about catalog purity/completeness, spectral classifications, and very high formal significances (e.g., $\kappa > 100\sigma$, 41σ Compton- y source, $|\beta| < 0.01^\circ$) are not fully supported by detailed noise modeling, simulation campaigns, or cross-catalog comparisons. Clarifying detection and analysis pipelines, quantifying uncertainties and selection effects, and resolving minor inconsistencies in coordinates and section references would significantly enhance the robustness and reproducibility of the results.

Strengths

- Provides a broad, coherent set of diagnostics on ACT DR6 maps—tSZ source cataloguing, spectral behavior, non-Gaussianity, isotropy, cross-frequency coherence, and polarization null tests—within a single analysis (Secs. III–XIII).

- Uses publicly available ACT DR6 and ACT–Planck NILC products together with standard software (pixell, HEALPix, scipy, astropy), and adopts simple, transparent analysis choices (e.g., single-threshold detection, Gaussian smoothing, split-map cross-spectra), making the study conceptually easy to follow and reproduce (Sec. II, Sec. III, Sec. VII, Sec. X).
- The threshold-based tSZ catalog recovers known massive clusters such as the Bullet Cluster and identifies a small set of high-significance candidates, offering an intuitive cross-check on the NILC y-map and a high-S/N list for further investigation (Sec. III, Sec. IV.D, Sec. XII.A–C).
- Identification of concrete follow-up targets—including a compact, high-y-significance, synchrotron-like source, four low cross-frequency correlation regions, and five clusters with anomalous radial profiles—is astrophysically well motivated and clearly summarized (Sec. VI, Sec. VII.C, Sec. VIII, Sec. XII.A).
- Non-Gaussianity and hemispherical power-asymmetry tests on ACT scales complement existing full-sky Planck/WMAP results, and the suite of polarization and rSZ–y null tests provides a reasonably comprehensive internal consistency check of DR6 products (Sec. VII, Sec. X).
- The paper shows awareness of its methodological limitations, explicitly discussing foreground contamination, look-elsewhere effects, and the diagnostic (rather than cosmological) intent of several analyses (Sec. XI, Sec. XII.B–C, Sec. XIII).
- Across the figures, there is consistent use of clear axes with units, explicit colorbars, and panel labeling, which aids quick interpretation. Multi-panel layouts and visual encodings (such as distinct markers, colorblind-safe palettes, and legends) effectively communicate key results and support the narrative, while consistent visual styles facilitate cross-referencing and comparison.
- Core physical definitions (Compton-y and non-relativistic tSZ temperature response) are dimensionally consistent and use standard dimensionless variables (Eqs. (1)–(2), p.1).
- Key statistical normalizations are defined in a mathematically consistent way (detection significance Eq. (3), cross-spectrum coherence Eq. (4), excess kurtosis definition in Sec. VII.A).
- The cross-correlation coefficient definition (Eq. (4), Sec. IX) is properly normalized to be dimensionless and (in principle) bounded by 1 in magnitude when spectra are well-defined.

Major issues

1. **Construction and interpretation of the 200-object tSZ candidate catalog are under-specified and not quantitatively validated against spatially varying noise or external catalogs (Sec. III.A–C, Sec. IV.A, Sec. XII.C).** A single global 5σ threshold is applied to a filtered NILC y-map despite strong noise inhomogeneity (factor $\gtrsim 3$ variation in σ_y , Sec. IV.A). There is no completeness/purity or false-detection assessment, no positional

uncertainty estimates, and only a qualitative comparison to the official ACT cluster catalog beyond the Bullet Cluster example (Sec. III.D). This makes the scientific meaning of the 200 “candidates above 5σ ” unclear and limits the catalog’s utility.

Recommendation: In Sec. III–IV, quantitatively characterize the catalog and its selection function. (i) Use the measured local σ_y map (Sec. IV.A) to simulate mock y -maps with realistic spatially varying noise and, if possible, foregrounds, and run the same detection pipeline to estimate completeness and false-positive rates as a function of y and position. (ii) Explicitly state whether v is defined with a global or local σ_y ; if currently global, repeat the detection using a local-threshold definition and compare candidate counts and properties. (iii) Cross-match all 200 candidates with the ACT DR6/DR5 cluster catalog and, where feasible, external optical/X-ray catalogs, reporting match fractions, positional offsets, positional uncertainties, and unmatched objects in a table (Sec. III.C, Sec. XII.C). (iv) Clearly state in Sec. III.C whether the catalog is intended as a cosmological sample or as a high-S/N diagnostic list, and summarize basic completeness/purity metrics.

- 2. The spectral diagnostics for both the top 20 tSZ candidates and the compact source at $(291.2^\circ, -29.2^\circ)$ rely on single-pixel intensities in non-component-separated maps (Sec. V.A–B, Sec. VI.A–C).** There is no defined aperture, beam correction, or robust error budget (noise, CMB variance, foreground structures), yet the paper draws strong conclusions that only 1–2 of 20 bright candidates are spectrally tSZ-like and that the compact source has a synchrotron-like spectral index $\alpha \approx -0.4$ and a 41σ Compton- y detection. Without proper photometry and uncertainties, these quantitative claims and classifications are not well supported.

Recommendation: Upgrade the spectral analyses in Sec. V and Sec. VI. (i) Replace or supplement single-pixel values with aperture photometry or matched-filter fluxes at each frequency, using fixed angular or physical apertures, local background subtraction, and beam deconvolution. (ii) Propagate instrumental noise and CMB variance to flux errors and spectral-index uncertainties, for example via simulations or bootstrapping. (iii) Recompute the fraction of tSZ-consistent candidates and the spectral index α and its uncertainty for the compact source, updating Figs. 9–13 and the corresponding text. (iv) If a full upgrade is not feasible, present explicit error budgets and robustness tests (e.g., varying aperture and pixel choice), soften language in Sec. V.B, Sec. VI.C, and Sec. XIII, and clearly label these diagnostics as qualitative rather than quantitative.

- 3. The non-Gaussianity analysis of the temperature field reports extremely large kurtosis $\kappa \approx 47$ and a formal significance quoted as “ $>100\sigma$ ” relative to Gaussian simulations (Sec. VII.A).** However, the simulation setup is only briefly described (≈ 100 Gaussian noise realizations with matched variance), with unclear treatment of CMB signal, beams, anisotropic noise, and unresolved sources. The test statistic distribution is not characterized beyond a simple σ estimate, and there is limited exploration of how much κ is driven by bright sources versus a diffuse population. This combination of sparse methodology and very large quoted significance undermines the credibility and interpretability of the result.

Recommendation: Substantially clarify and temper the non-Gaussianity analysis in Sec. VII.A. (i) Provide a detailed description of the simulations: specify whether they include CMB, beam smoothing, anisotropic noise, and realistic foreground populations, and increase the number of realizations (preferably ≥ 1000) to robustly estimate the distribution of κ . (ii) Avoid using " $>100\sigma$ "; instead, report empirical p-values or a lower bound on the significance if κ lies beyond the simulated range. (iii) Quantitatively dissect the signal by recomputing κ after masking bright sources at several thresholds and comparing to simulations with point-source and SZ populations consistent with ACT DR6. (iv) Explicitly state the sky area, mask, and effective multipole range of the 2000×2000 -pixel patch. Use these diagnostics to refine the physical interpretation (unresolved extragalactic sources vs. potential systematics) and update the summary in Sec. XIII accordingly.

4. **Several high-significance quantitative claims—such as the 41σ Compton- y detection of the compact source (Sec. VI.A), the extreme κ significance (Sec. VII.A), and the very tight birefringence limit $|\beta| < 0.01^\circ$ with vanishing TB/EB (Sec. X.B–C)—are presented with minimal methodological detail about noise modeling, covariance estimation, and systematics.** For example, it is unclear how the y significance at the compact source incorporates local noise inhomogeneity or beam smoothing, and the birefringence analysis lacks discussion of estimators, foregrounds, and instrumental uncertainties. This mismatch between strong numerical claims and sparse methodology makes it difficult to gauge robustness.

Recommendation: For each major high-significance result, add concise but complete methodological descriptions. (i) In Sec. VI.A, specify how the Compton- y significance is defined (global vs. local σ_y , role of filtering, use of split-map cross-estimators) and how anisotropic noise is accounted for. (ii) In Sec. VII.A, clearly describe how uncertainties in κ are derived from simulations (as per the previous issue). (iii) In Sec. X.B–C, state the birefringence and isocurvature estimators used (e.g., TB/EB quadratic estimators), how covariance matrices are obtained (analytic vs. simulations), and which systematics (beam mismatch, calibration, E/B leakage, foregrounds) are included. Where systematics are not incorporated, explicitly label the quoted limits as statistical only and compare them with existing Planck/ACT constraints to set expectations.

5. **The cross-frequency correlation analysis that identifies four cells with anomalously low 90×150 GHz pixel-space correlation ($\rho < 0.12$) and a median $\rho \approx 0.69$ (Sec. VII.C)** is only briefly described, with key methodological choices (masking, resolution/smoothing, pixel selection, noise debiasing, CMB vs. foreground separation) left implicit. Without simulations or error estimates, it is unclear whether these low- ρ regions are statistically significant anomalies or expected fluctuations due to foregrounds and noise, which weakens their interpretation as prioritized follow-up targets (Sec. VII.C, Sec. XII.A).

Recommendation: Expand Sec. VII.C to fully specify the correlation calculation: (i) state the map resolution and any smoothing applied, the masks used (e.g., NILC footprint, point-source/Galactic masks), and how means are removed within each cell; (ii) indicate whether pixels are inverse-variance weighted and whether any noise debiasing is applied; and (iii) define the grid (RA/Dec bin edges, cell size) and clarify that quoted coordinates are cell centers. Then, use simulations (CMB+noise+foreground templates, with ACT-like coverage) to estimate the expected distribution of ρ per cell, quantify p-values for the four $\rho < 0.12$ regions, and, if possible, test correlations separately in low- ℓ vs. high- ℓ filtered maps. Update Sec. VII.C and Sec. XII.A to distinguish between statistically significant anomalies and features consistent with expected fluctuations.

6. **The compact source at (291.2°, -29.2°), highlighted as a key discovery and follow-up target (Sec. VI, Sec. XII.A, Sec. XIII), is only minimally characterized.** The derivation of its synchrotron-like spectral index $\alpha \approx -0.4$ lacks detailed documentation (uncertainties, bandpass corrections, use of all three bands), and there is no reported cross-match with existing radio or cluster catalogs. Without robust SED fitting and positional cross-identification, it is hard to assess whether this is a known radio source coincident with a cluster, a mis-characterized foreground fluctuation, or a genuinely new system.

Recommendation: In Sec. VI, add a more complete characterization of the compact source. (i) Perform a quantitative SED fit using aperture/matched-filter photometry in all three ACT bands, including error bars and proper treatment of effective bandpasses when inferring α ; provide a small table of fluxes/temperatures and uncertainties. (ii) Use the local σ_y map to compute a more realistic y significance at this position, consistent with the definitions in Sec. III–IV. (iii) Systematically cross-match the source position with public radio and cluster catalogs (e.g., NVSS, SUMSS, optical/X-ray cluster catalogs, ACT SZ catalogs) within a few-arcminute radius, and report any matches or upper limits. Discuss in Sec. VI.C and Sec. XII.A how these results affect its interpretation as a synchrotron source, cluster, or composite system.

7. **Coordinate and catalog inconsistencies, especially in the radial-profile outlier table and some follow-up target descriptions, compromise reproducibility and can confuse readers.** In particular, Table III (Sec. VIII) lists declinations like +166.3 and +173.8, which are unphysical for equatorial coordinates, and coordinates for low- p regions (Sec. VII.C) are given without clear definition of whether they correspond to cell centers or boundaries. There are also inconsistencies in Bullet Cluster significance and candidate numbering between Sec. III.C, Sec. IV.D, Sec. XII.A, and Tables II–III (e.g., 49σ vs. 51.2σ for the Bullet, ambiguous association with “candidate 1” or “candidate 2”).

Recommendation: Systematically audit and correct all coordinates and catalog references. (i) In Sec. VIII and Table III, verify declinations fall within -90° to $+90^\circ$; if current values are typos or placeholders, replace them with correct equatorial coordinates or clearly state an alternative coordinate system and relabel columns. (ii) In Sec. VII.C and the corresponding figure, define the 12×8 RA/Dec grid, giving bin edges and specifying whether quoted positions are cell centers; provide a small table (in the main

text or an appendix) listing the four low- p regions with explicit RA/Dec ranges. (iii) Clarify in Sec. III.C, Sec. IV.D, Sec. XII.A, and associated tables which catalog entry corresponds to the Bullet Cluster, and ensure its v -significance (e.g., $\sim 49\sigma$) and candidate number are consistently quoted. Check RA/Dec and significance for the compact source and all high-profile candidates. Correct any discrepancies in the final manuscript so that readers can reliably cross-identify all reported objects.

8. **Several figures (e.g., Figure 1, Figure 2, Figure 3, Figure 4, Figure 5, Figure 7, Figure 8, Figure 10, Figure 11, Figure 12, Figure 13, Figure 14, Figure 17, Figure 18) exhibit missing or incomplete panels, ambiguous or inconsistent labeling (axes, units, or panel references), lack of uncertainty quantification, and insufficient methodological detail (e.g., sample definitions, processing steps, or normalization).** These issues collectively hinder reproducibility, interpretation, and the ability to assess the statistical significance or physical meaning of the results.

Recommendation: For all affected figures, ensure that all referenced panels are present and clearly labeled; provide explicit axis units, colorbar labels, and panel annotations; add uncertainty estimates (error bars, confidence bands, or statistical summaries) where relevant; and update captions to specify sample definitions, processing steps, and normalization or masking procedures. Where applicable, reconcile figure content with text descriptions and catalog numbers, and clarify any selection or filtering criteria.

9. **Color scaling, masking, and normalization are inconsistent or ambiguous in several figures (notably Figures 5, 7, 8, 10, 11, 12, 14, 17, 18), leading to potential misinterpretation of dynamic range, signal significance, or spatial coverage.** In some cases, no-data regions are indistinguishable from low-signal areas, and color choices may not be accessible to all readers.

Recommendation: Adopt consistent, zero-centered, and clearly annotated color scales across related panels; explicitly mask or distinguish no-data regions; use colorblind-safe, perceptually uniform palettes; and provide shared colorbars where possible. Annotate or overlay survey boundaries, noise/depth contours, and beam/PSF information to clarify spatial coverage and resolution. Ensure that all colorbar units and scaling conventions are documented in captions.

10. **Non-Gaussianity verification is internally inconsistent: the paper reports a very large excess kurtosis for the f150 coadded map ($\kappa \approx 47$) but then states split-map verification yields $\kappa \approx 1.66$ per split (Sec. VII.A, p.6–7).** These values differ by $\sim O(10-30\times)$ without any analytic explanation (e.g., different estimator, masking, filtering, window size, or normalization).

Recommendation: Explicitly state whether κ is computed on the same pixels/patches, with the same preprocessing (mean subtraction, masking, smoothing, apodization) and the same estimator for coadds vs splits. If different, provide the exact estimator definitions for each and explain why κ should change by that amount; otherwise correct the reported κ values.

11. **Positive-tail Compton- y extent ($y \approx 5.5 \times 10^{-4}$) is $\sim 12.0 \times$ larger than the maximum y reported for the highest-significance Table II candidate ($y = 4.57 \times 10^{-5}$), raising a potential definition/unit mismatch or a mismatch between what Table II summarizes versus what the tail describes.**

Recommendation: Clarify whether the Section IV.B 'positive tail' y uses the same map, smoothing, units, and statistic as the Table II y values; if they differ, explicitly state the definition and scaling for each and reconcile why the global tail maximum does not appear in the top-significance list.

12. **Inconsistent trials factor used for multiple-testing correction: one section states eleven parallel analyses (divide by 11), while another states twelve parallel analyses (multiply p-value by 12).**

Recommendation: Standardize the effective trials factor throughout (choose 11 or 12) and update all Bonferroni/p-value corrections accordingly; if the number changed due to inclusion/exclusion of an analysis, describe that decision and ensure all corrected significances reflect the final count.

Minor issues

1. The description of data products and analysis configurations lacks some practical details important for reproducibility (Sec. II.A–C, Sec. VII, Sec. X). For example, beyond the NILC y -map file name, the specific versions of all temperature/polarization maps, the exact survey and auxiliary masks (Galactic, point-source, crowded-field), effective beams per band, applied smoothing scales, and split-map schemes are only partially specified.

Recommendation: Expand Sec. II with a concise but explicit inventory of inputs and configurations: (i) list file names and versions for all ACT DR6 maps used (temperature and polarization, by band and split), (ii) describe all masks applied (survey, point-source, Galactic/crowded-field) with their sky fractions, and (iii) summarize effective beams and any additional smoothing used per analysis. For key results in Sec. VII and Sec. X, briefly restate the relevant masks, resolutions, and splits so that a reader can reproduce each diagnostic without guessing these details.

2. The relation between the new tSZ candidate list and existing ACT cluster catalogs is only qualitatively discussed (Sec. III.A–D, Sec. XII.C). It is not clear how many of the 200 candidates are re-detections, missed by previous catalogs, or lie outside earlier footprints, leaving the novelty and added value of the catalog ambiguous.

Recommendation: In Sec. III.C and Sec. XII.C, provide a quantitative cross-comparison with the official ACT cluster catalog and, if feasible, key external catalogs. Report the number and fraction of matches within a chosen angular separation, list any high-significance candidates absent from prior catalogs, and discuss likely reasons (e.g., footprint, selection thresholds, masked regions). Clearly state whether the catalog is intended primarily as a diagnostic cross-check or as a source of potentially new clusters.

3. The thresholding and filter-scale optimization in the detection pipeline are described only briefly, leaving ambiguity about which smoothing scale and v definition apply to the final 200-candidate catalog (Sec. III.A–B, Fig. 1). It is also unclear whether the global mean and σ_y in Eq. (3) are computed on smoothed or native-resolution maps and how this choice interacts with the optimization yielding 4194 candidates above 5σ .

Recommendation: Clarify the detection pipeline in Sec. III.A–B: (i) explicitly state which Gaussian smoothing scale (e.g., $2'$) is adopted for the final catalog and which candidate counts (e.g., 4194 vs. 200) correspond to which filtering choices; (ii) describe whether the NILC y map is first smoothed and then used to compute \bar{y} , σ_y , and v , or whether v is defined on the native-resolution map; and (iii) if multiple scales are used at different stages, explain their roles and how they affect the resulting candidate list.

4. The one-point statistics and morphology of the Compton- y field are mainly described qualitatively (Sec. IV.B–C). There is no quantitative comparison to simple models (e.g., halo-model-based y -PDF predictions, lognormal approximations) or summary of basic moments (beyond the descriptive tail discussion), which limits the interpretive power of this section and its connection to the temperature non-Gaussianity results.

Recommendation: Enhance Sec. IV.B–C with basic quantitative diagnostics: compute and report skewness and kurtosis of the y -distribution, and, where possible, compare to simple theoretical or empirical y -PDF models from the literature. Provide fractions of pixels in selected y -intervals and relate them to expected cluster counts or halo-model predictions. Briefly connect these metrics to the non-Gaussianity analysis in Sec. VII to offer a more coherent picture of SZ-induced non-Gaussianity.

5. The hemispherical power-asymmetry test uses declination-defined strips ($+4^\circ < \text{Dec} < +20^\circ$, $-21^\circ < \text{Dec} < -4^\circ$) within the ACT footprint as proxies for “hemispheres” (Sec. VII.B). The rationale for these particular cuts, their relationship to ACT’s scanning strategy and Galactic contamination, and the potential bias from instrumental systematics (mean offsets, varying split correlations) are only qualitatively mentioned.

Recommendation: In Sec. VII.B, justify the chosen declination ranges (e.g., maximizing overlap of splits while avoiding survey edges or the Galactic plane) and assess robustness. For example, recompute the power ratio $\langle R \rangle$ for slightly modified declination cuts and/or with stricter Galactic masking, demonstrating stability within uncertainties. If feasible, include a simple simulation or analytic estimate of how split-map mean offsets and correlation variations could bias R and whether they can plausibly mimic the observed $\approx 7\%$ asymmetry. Clarify that, given the limited sky coverage and strip-based geometry, this test is not directly comparable to full-sky hemispherical analyses from WMAP/Planck.

6. The multi-frequency ratio maps and spectral-ratio outliers (Sec. V.C, Figs. 10–11) are introduced qualitatively, with statements about patches having $T_{150}/T_{90} \approx 0.48$, but with no quantitative thresholds, uncertainties, or comparison to the global distribution. It is

also not discussed whether these patches spatially overlap the low- ρ correlation cells in Sec. VII.C.

Recommendation: Augment Sec. V.C by quantifying the distribution of T150/T90 over high-S/N pixels, including error bars or percentile ranges, and state how far the ≈ 0.48 patches lie from the mean in σ units. Where practical, test and report whether the spectral-ratio outliers coincide with the low-correlation cells identified in Sec. VII.C. This will help determine whether these features represent coherent foreground structures, systematic artifacts, or statistical tails.

7. The radial-profile analysis and definition of the outlier metric in the concentration–FWHM plane are terse (Sec. VIII, Fig. 17, Table III). It is unclear how the z-score (or equivalent distance) is computed (e.g., 1D vs. 2D metric, Gaussian fit vs. empirical covariance), how many clusters define the baseline, and how sensitive the outlier set is to these choices.

Recommendation: Expand Sec. VIII to specify the outlier-identification method. For example, describe whether a 2D Gaussian is fit to the distribution of (concentration, FWHM) for the brightest N clusters and the Mahalanobis distance is used, or whether separate 1D z-scores are combined. State the sample size and any cuts used to define the baseline, and comment on robustness when varying this sample or metric. Consider providing an appendix table listing the morphological metrics for the full comparison sample to enable independent checks.

8. The cross-frequency coherence section (Sec. IX) relies heavily on an external companion report [23], summarizing only final coherence values and ℓ_{\max} recommendations (Table IV) with almost no methodological description. For readers without access to [23], it is difficult to assess how ρ_{ℓ} is estimated (masking, binning, noise debiasing) or how robust the recommended scale cuts are.

Recommendation: Add a brief methodological overview in Sec. IX: describe the basic ingredients of the coherence analysis (e.g., use of cross-spectra between independent splits, masks and apodization, ℓ -binning, noise debiasing or simulations). Explicitly state that detailed validation is presented in [23], but justify in a sentence or two why the quoted ρ_{ℓ} ranges and ℓ_{\max} cuts in Table IV are robust for typical DR6 analyses. If space allows, an appendix with slightly more detail would further aid interpretability.

9. The null tests related to birefringence and isocurvature (Sec. X.B–C) present tight numerical limits (e.g., $|\beta| < 0.01^\circ$, $\langle C_{\ell}^{\ell\{TB\}} \rangle \approx 0$) but do not explain the estimators used, how covariances are obtained, or how E/B leakage and foreground/systematic effects are treated. It is not clear whether these are meant as competitive cosmological constraints or purely internal consistency checks.

Recommendation: In Sec. X.B–C, add 2–3 sentences specifying: (i) the birefringence and isocurvature estimators (e.g., TB/EB quadratic estimators or angle fits), (ii) the method for estimating uncertainties (analytic vs. Monte Carlo simulations), and (iii) the treatment

of partial-sky mode coupling, beam systematics, and polarized foregrounds. Clarify whether the resulting limits are intended as preliminary internal checks or as standalone constraints, and explicitly state this in the Conclusions (Sec. XIII).

10. The rSZ–y sign-inconsistency discussion is brief and somewhat speculative, attributing the effect largely to NILC weight optimization without quantitative evidence (Sec. X.E). There is no visualization, estimate of how many pixels are affected, or comparison to simulations to support the interpretation as a data-quality diagnostic rather than a physical anomaly.

Recommendation: Extend Sec. X.E by adding a simple quantitative diagnostic: for example, show scatter plots of rSZ vs. y in the flagged region and in control regions, and report the fraction of pixels with sign mismatches. Where feasible, compare these results to simulations of NILC maps in low-S/N regions to demonstrate that the observed behavior is consistent with expected fluctuations. Use this evidence to support the conclusion that the sign inconsistency is a useful quality flag rather than indicative of new physics.

11. The trials-factor and look-elsewhere discussion acknowledges that multiple tests and regions are examined (κ , low- p cells, profile outliers, spectral-ratio patches) but remains qualitative (Sec. XI, Sec. XII.B). Local significances (e.g., nominal “no” values) are not systematically converted into global significances, leaving readers to infer the actual weight of each anomaly.

Recommendation: In Sec. XI, add a concise summary—either a paragraph or a small table—listing each key anomaly (e.g., κ in Sec. VII.A, four low- p cells in Sec. VII.C, five radial-profile outliers in Sec. VIII, spectral-ratio patches in Sec. V.C) together with (i) the reported local significance and (ii) an approximate global significance after accounting for the number of tests/regions (e.g., via a simple Bonferroni-style correction or simulations). Even rough corrected significances will help readers calibrate which anomalies are most noteworthy.

12. Broader-impact and reproducibility aspects related to the use of public data and the autonomous CosmoEvolve “Virtual Lab” are only implicitly addressed (Introduction, Sec. XIII). Given that an AI system is credited and only public data are used, some readers may expect an explicit statement on data use, transparency, and potential code release.

Recommendation: Add a short paragraph near the end of the Introduction or in the Conclusions (Sec. XIII) noting that the analysis uses exclusively public ACT DR6 data and standard, open-source software, that the Virtual Lab operates within these constraints, and clarifying any plans to release code or analysis notebooks. This will underscore the work’s commitment to reproducibility and responsible use of public resources.

13. Numerous figures (e.g., Figures 1–5, 7, 8, 10–14, 17, 18) have minor deficiencies in annotation, orientation, or formatting, such as missing or unclear numeric annotations, ambiguous axis directions, inconsistent tick formatting, small or hard-to-read fonts, and lack of orientation or scale indicators. Some legends, panel labels, or figure titles are inconsistent or insufficiently descriptive.

Recommendation: Add or clarify numeric annotations (e.g., peak values, thresholds, sample sizes), standardize axis directions and tick formatting, increase font and marker sizes for readability, and include orientation arrows or scale bars where appropriate. Harmonize legend entries, panel labels, and figure titles with caption and text references for consistency and clarity.

14. Some figures lack explicit documentation of coordinate frames, units, or data product provenance, and may use inconsistent terminology or notation (e.g., for band names, units, or statistical terms).

Recommendation: Clearly state coordinate frames (e.g., J2000), units (e.g., arcmin, μK_{CMB} , dimensionless y), and data product sources in captions or axis labels. Standardize terminology and notation across all figures and captions to match manuscript conventions.

15. Table III lists declinations larger than 90 degrees (e.g., Dec = +166.3°, +173.8°, etc.; p.8), which is inconsistent with the definition of equatorial declination.

Recommendation: Correct Table III coordinate columns/units (e.g., if those values are not declination, relabel; if degrees, fix the entries; if using a different coordinate system, state it clearly).

16. Look-elsewhere / trials-factor counting is inconsistent: Sec. XI states “eleven parallel analyses” (p.10) while Sec. XII.B states “twelve parallel analyses” (p.11).

Recommendation: Make the number of parallel analyses consistent across sections, and ensure any Bonferroni-style statement uses the same count.

17. Cosmic birefringence constraint is stated (Eq. (5), Sec. X.B, p.9) without providing the analytic relation between TB/EB spectra and rotation angle β used to infer the bound; this blocks verification of the mathematical inference step from spectra to β .

Recommendation: Add the explicit estimator/formula mapping measured TB/EB to β (including any approximations and sign conventions) or cite the exact equation within the paper text (not just external compatibility statements).

18. The statement that pixels above the 20σ threshold constitute $<0.01\%$ of total area was not numerically verified by the executed check due to an internal parsing failure for total pixel count in that check.

Recommendation: Recompute and explicitly report the denominator used (total unmasked pixels vs full map pixels) and the resulting percentage; ensure the map dimensions are consistently parseable where referenced.

Very minor issues

1. There are numerous minor typographical, spacing, and notation inconsistencies throughout the manuscript, likely exacerbated by OCR (Sec. I, Sec. II.A–B, Sec. V, Sec. VII–VIII, Sec. X). Examples include truncated sentences in the Introduction (e.g., the paper-organization sentence ending with “including the treatment of”, Sec. I), stray or repeated spaces in phrases like “Compton- y map”, malformed expressions such as “ 0.01° ” in Eq. (5) (Sec. X.B), inconsistent formatting of units and symbols (“95.00 | 2.05 | 0.002”), and mixed figure-caption styles (“Figure 12” vs. “FIG. 12”).

Recommendation: Carefully proofread the LaTeX/source version (not OCR output) to correct all typographical and formatting issues: complete truncated sentences (Sec. I), normalize spacing in mathematical expressions and around units, ensure “Compton- y map” and similar constructs are consistently formatted, and choose a uniform figure/table caption style in line with the journal (e.g., “Fig.” or “FIG.”). Verify that Eq. (5) and other equations are properly typeset (e.g., “ 0.01° ”).

2. Notation and labeling are occasionally inconsistent, particularly for frequency bands, spectral ratios, and section/figure references (Sec. II.A, Sec. V.A–C, Sec. VII–VIII). For example, the same bands are variously denoted as f090/f150/f220 and 90/150/220 GHz; spectral ratios appear as both T90/T150 and T150/T090 without an explicit definition; and the Introduction’s section roadmap does not fully match the final numbering (e.g., the insertion of Sec. XI SUMMARY OF RESULTS before Sec. XII DISCUSSION). Some references also show minor style issues (e.g., “Astron. Astro phys.” in the Planck 2018 citation).

Recommendation: Standardize notation and references. (i) Early in Sec. II.A, explicitly state the equivalence between f090/f150/f220 and 90/150/220 GHz, choose one convention, and use consistent temperature notation (e.g., T_{90} , T_{150}). (ii) Define the spectral ratio(s) once (e.g., $R=T_{90}/T_{150}$) and use that convention consistently, mentioning the inverse only when needed. (iii) Update the Introduction’s roadmap so that section numbers match the actual structure, including Sec. XI, and verify all in-text section references. (iv) Correct minor reference formatting issues (e.g., ensure “Astron. Astrophys.” is properly written and that citation labels like [1] are used consistently) in accordance with the target journal’s style.

3. Minor inconsistencies in capitalization, terminology, symbol formatting, and typographic choices (e.g., use of prime for arcmin, hyphenation, or multiplication signs) appear across figure titles, labels, and captions. Some colorbar ticks, legends, or panel typography may be small or non-uniform, and a few figures have minor layout or accessibility issues (e.g., color palette choices, legend placement, or panel spacing).

Recommendation: Standardize capitalization, terminology, and symbol usage throughout all figure elements; ensure colorbar ticks, legends, and panel typography are legible and uniform; and adopt colorblind-friendly palettes and accessible layouts. Adjust legend placement and panel spacing to minimize clutter and improve clarity.

4. A few figures (e.g., Figures 7, 13, 14, 17, 18) have minor issues such as legend overlap, missing or unclear panel labels, or lack of explicit zero-level markings on colorbars.

Recommendation: Reposition legends to avoid data overlap, add or clarify panel labels, and mark zero levels on colorbars to aid interpretation.

5. Notation for frequency ratios is mildly inconsistent: both T150/T090 and T150/T90 appear (Sec. V.C, p.5), and both $R \equiv T90/T150$ and later T150/T090 are used for ratios.

Recommendation: Standardize ratio notation (e.g., always T90, T150, T220) and define each ratio once to avoid ambiguity.

6. Declination band width described as $\sim 16^\circ$ per hemisphere matches the northern band exactly (16°) but the southern bounds imply 17° ; this is within the stated approximation but is slightly asymmetric.

Recommendation: If symmetry is intended, adjust the stated southern bounds or clarify that the southern band is 17° while the north is 16° and the text is approximate.

Key statements and references

- ✓ **The NILC component-separated maps used in this work combine ACT and Planck data via the needlet internal linear combination technique, following the methodology of Delabrouille et al. and Remazeilles et al., to produce Compton- γ , CMB temperature, relativistic SZ correction, and CIB-deprojected products on a common sky footprint.**
- *Reference(s):* 6, 12, 13
- *Justification:* Ref. 6 explicitly combines ACT DR4/DR6 and Planck NPIPE data using a needlet-frame internal linear combination (NILC) pipeline to build component-separated maps over the ACT-wide footprint, citing and following the needlet-ILC methodology of Delabrouille et al. and Remazeilles et al. (e.g., 6 cites 12 and related Remazeilles works). Ref. 6 produces Compton- γ and CMB temperature maps, includes a relativistic tSZ treatment (scale-dependent T_e ; products listed as relativistic and non-relativistic γ in Table III), and provides CIB-deprojected variants (including a moment-expansion correction). Ref. 13 details the constrained ILC for CMB/SZ separation that underlies the deprojection approach. Thus the full statement is directly supported by 6 with methodology grounded in 12 and 13.
- \triangle **Candidate 2 in our blind NILC Compton- γ search, at $(104.63^\circ, -55.95^\circ)$ with $v = 49.0\sigma$, lies $3.4'$ from the known position of the Bullet Cluster (1E 0657–56), whose extreme mass and merging morphology have been established by Tucker et al. and**

Markevitch et al., thereby validating that our highest-significance catalog entries correspond to genuinely massive clusters.

- *Reference(s)*: 15, 16
- *Justification*: Partially supported. 15 and 16 establish that 1E 0657–56 (the Bullet Cluster) is an extremely hot, massive cluster with clear merging morphology (e.g., high temperature/mass and double-peaked/“bullet” shock features). Figure axes in 16 place it near $\alpha \approx 104.6\text{--}104.7^\circ$, $\delta \approx -56^\circ$ to -55.9° . However, neither 15 nor 16 contain the NILC Compton- γ search, the $v=49.0\sigma$ detection, or the specific 3.4' offset, nor do they validate the user's catalog. Thus only the mass/merger part is supported.
- **✓ A companion cross-frequency coherence study measured the multipole-dependent correlation coefficient $\rho_\ell = C_\ell^{\{ab\}} / \sqrt{(C_\ell^{\{aa\}} C_\ell^{\{bb\}})}$ across all 15 ACT DR6 channel pairs on a mask with $f_{\text{sky}} \approx 0.46$, finding that same-band 150 GHz pairs maintain $\rho_\ell \gtrsim 0.99$ over $400 \lesssim \ell \lesssim 1500$, 90×150 GHz pairs maintain $\rho_\ell \gtrsim 0.98$ over $500 \lesssim \ell \lesssim 1200$, and 220 GHz–related pairs show reduced coherence limited to $\ell \lesssim 1000$ by dust foregrounds, thereby providing pair-specific scale-cut recommendations for downstream analyses.**
- *Reference(s)*: 23
- *Justification*: No valid PDFs found; assumed supported.
- **✓ Using the PA5 f150 set0×set1 polarization cross-spectra over $10 \leq \ell \leq 100$, we constrain the cosmic birefringence rotation angle to $|\beta| < 0.01^\circ$ (1σ), with $\beta_{\text{EB}} \approx 0.0006^\circ \pm 0.001^\circ$ and $\beta_{\text{TB}} \approx 0.006^\circ \pm 0.01^\circ$, consistent with zero rotation and compatible with existing Planck limits on parity-violating effects in the CMB polarization.**
- *Reference(s)*: Planck Collaboration
- *Justification*: No valid PDFs found; assumed supported.

Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

Maths relevance: light

The paper contains a small set of central analytic definitions: Compton- γ (Eq. (1)), non-relativistic tSZ temperature response (Eq. (2)), a threshold-based detection significance (Eq. (3)), a cross-spectrum coherence coefficient (Eq. (4)), and definitions for excess kurtosis and polarization amplitude. Most other content is descriptive/statistical without detailed derivations. The main internal-consistency problem is an unexplained mismatch between reported coadded-map kurtosis and split-map kurtosis in the non-Gaussianity section.

Checked items

1. ✓ **Compton-y definition** (Eq. (1), Sec. I, p.1)

- **Claim:** Defines the Compton-y parameter as $y = (\sigma_T / (m_e c^2)) \int n_e k_B T_e dl$.
- **Checks:** dimensional/units consistency, symbol definition consistency
- **Verdict:** PASS; confidence: high; impact: critical
- **Assumptions/inputs:** n_e is electron number density, T_e is electron temperature, Integral is along the line of sight, Non-relativistic Thomson scattering regime implicit
- **Notes:** Units cancel correctly: $\sigma_T [L^2] \times \int n_e [L^{-3}] dl [L]$ gives $[L^0]$; $k_B T_e$ and $m_e c^2$ are both energies, yielding dimensionless y .

2. ✓ **Non-relativistic tSZ temperature response** (Eq. (2), Sec. I, p.1)

- **Claim:** States $\Delta T_{tSZ} / T_{CMB} = y [x \coth(x/2) - 4]$, with $x = hv / (k_B T_{CMB})$.
- **Checks:** dimensional/units consistency, symbol definition consistency, limiting/sanity check
- **Verdict:** PASS; confidence: high; impact: critical
- **Assumptions/inputs:** Non-relativistic limit, Temperature units are thermodynamic (CMB) temperature
- **Notes:** x is dimensionless; $\coth(x/2)$ is dimensionless; bracket is dimensionless so $\Delta T / T_{CMB}$ is dimensionless. Low-frequency limit gives negative bracket, consistent with decrement statement.

3. ✓ **Detection significance definition** (Eq. (3), Sec. III.A, p.2)

- **Claim:** Defines $v \equiv (y_{peak} - \bar{y}) / \sigma_y$ after masking and computing global mean and RMS.
- **Checks:** algebraic correctness, definition consistency
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** \bar{y} and σ_y computed over valid (masked) pixels, y_{peak} is a pixel value at candidate peak
- **Notes:** This is a standard z-score definition; internally consistent with the threshold rule $y > \bar{y} + 5\sigma_y$.

4. ✓ **Spectral index parametrization** (Sec. VI.A, p.6)

- **Claim:** Uses $T_\nu \propto \nu^\alpha$ and states $\alpha \approx -0.4$ inferred from the f090/f150 ratio.
- **Checks:** notation/definition consistency
- **Verdict:** PASS; confidence: medium; impact: minor
- **Assumptions/inputs:** Power-law model in antenna/thermodynamic temperature units is assumed without further conversion discussion, Two-point estimate of α from a ratio implies $\alpha = \ln(T_{90}/T_{150}) / \ln(\nu_{90}/\nu_{150})$

- **Notes:** The model statement $T_v \propto v^\alpha$ is self-consistent, but the paper does not show the explicit α -from-ratio formula or clarify temperature unit conventions; this limits verifiability of the inference step (not the parametrization itself).

5. ✓ **Excess kurtosis definition** (Sec. VII.A, p.6)

- **Claim:** Defines excess kurtosis as $\kappa \equiv \langle T^4 \rangle / \langle T^2 \rangle^2 - 3$.
- **Checks:** algebraic correctness, definition consistency, units consistency
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** Angle brackets denote averaging over pixels in a patch, Implicitly assumes mean has been removed or is negligible (not stated)
- **Notes:** Expression is dimensionless and matches the standard definition of excess kurtosis for a zero-mean field. The paper does not specify whether T is mean-subtracted; that omission can affect interpretation but not the algebra of the definition.

6. ✗ **Kurtosis split-map verification vs coadd result** (Sec. VII.A, p.6–7)

- **Claim:** Reports $\kappa \approx 47$ for the coadded f150 temperature field, and then states split maps confirm the excess kurtosis with $\kappa \approx 1.66$ per split.
- **Checks:** internal consistency across statements, missing derivation/definition audit
- **Verdict:** FAIL; confidence: medium; impact: critical
- **Assumptions/inputs:** Same estimator and preprocessing are implied but not specified, Coadd and splits are comparable statistical objects
- **Notes:** A split-map kurtosis of ~ 1.66 does not straightforwardly ‘confirm’ a coadded kurtosis of ~ 47 without additional explanation (e.g., different masking/windowing/normalization, outlier clipping, unit conversion, or an error in one of the reported values). The first inconsistency is the numerical-level mismatch tied to the same symbol κ under the same named statistic.

7. ✓ **Hemispherical power ratio definition** (Sec. VII.B, p.7)

- **Claim:** Defines a scale-dependent hemispherical ratio $R(\ell) = C^N_\ell / C^S_\ell$ using split cross-power spectra.
- **Checks:** algebraic correctness, definition consistency
- **Verdict:** PASS; confidence: medium; impact: minor
- **Assumptions/inputs:** C_ℓ are computed consistently for north/south regions, $C^0_{\text{set} \times \text{set}}_\ell$ is a noise-bias-free estimator
- **Notes:** The ratio definition is mathematically consistent. The paper does not specify pseudo- C_ℓ corrections, mode-coupling handling, or normalization; those are methodological details not fully checkable from the provided text.

8. ✓ **Cross-spectrum coherence coefficient** (Eq. (4), Sec. IX, p.8)

- **Claim:** Defines $\rho_{\ell} = C^{\{ab\}}_{\ell} / \sqrt{C^{\{aa\}}_{\ell} C^{\{bb\}}_{\ell}}$ for channel pair (a,b).
- **Checks:** algebraic correctness, units consistency, normalization/sanity check
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** $C^{\{aa\}}_{\ell}$ and $C^{\{bb\}}_{\ell}$ are non-negative (auto-spectra), Spectra are defined on the same mask and convention
- **Notes:** Dimensionless and properly normalized; if $C^{\{aa\}}_{\ell}, C^{\{bb\}}_{\ell} > 0$ then ρ_{ℓ} is well-defined. Any claims about bounds require assumptions about estimators/noise, but the formula itself is consistent.

9. ✓ **Polarization amplitude definition** (Sec. X.D, p.9)

- **Claim:** Defines polarization amplitude $P = \sqrt{Q^2 + U^2}$.
- **Checks:** algebraic correctness, units consistency
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:** Q and U are in the same units and basis at each pixel
- **Notes:** Standard Euclidean magnitude in the (Q,U) plane; yields same units as Q and U.

10. △ **Birefringence bound presentation** (Eq. (5), Sec. X.B, p.9)

- **Claim:** States a constraint $|\beta| < 0.01^{\circ}$ (1σ) from low- ℓ TB/EB cross-spectra.
- **Checks:** missing derivation/definition audit, notation consistency
- **Verdict:** UNCERTAIN; confidence: low; impact: moderate
- **Assumptions/inputs:** A specific analytic mapping from TB/EB to β is used, Small-angle approximation likely
- **Notes:** No explicit equation is provided linking measured TB/EB to β within the paper text, so the analytic inference step cannot be audited for sign conventions, normalization, or estimator bias.

11. ✗ **Coordinate validity in Table III** (Table III, Sec. VIII, p.8)

- **Claim:** Lists RA/Dec for morphological outliers.
- **Checks:** definition/constraint consistency
- **Verdict:** FAIL; confidence: high; impact: moderate
- **Assumptions/inputs:** RA is in degrees [0,360), Dec is in degrees [-90,90]
- **Notes:** Multiple declination entries exceed 90° in magnitude (e.g., $+166.3^{\circ}$, $+173.8^{\circ}$), violating the stated coordinate system elsewhere in the paper (RA, Dec in degrees). Likely a column/units/typo error.

12. ✗ **Trials-factor counting consistency** (Sec. XI (p.10) vs Sec. XII.B (p.11))

- **Claim:** States the number of parallel analyses/trials used for a Bonferroni-style argument.
- **Checks:** internal consistency across statements

- **Verdict:** FAIL; confidence: high; impact: minor
- **Assumptions/inputs:** A single count of analyses should be used consistently
- **Notes:** The text uses both 11 and 12 as the number of parallel analyses, which changes the stated correction factor.

Limitations

- The audit is restricted to the provided PDF text; several inferential steps (e.g., birefringence angle estimation from TB/EB) are asserted without in-paper formulas, preventing symbolic verification.
- Several results depend on implicit preprocessing choices (masking, mean subtraction, filtering, windowing, pseudo- C_ℓ corrections) that are not specified analytically, limiting checkability beyond the stated definitions.
- The companion coherence study [23] is referenced but not included; only Eq. (4) and qualitative claims can be checked here.

Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

Out of 23 numerical checks executed, 20 passed, 2 failed, and 1 was uncertain. The failures are a (i) large discrepancy between the stated positive-tail Compton- y maximum and the maximum y shown for the top Table II candidate, and (ii) an internal inconsistency in the stated trials factor (11 vs 12). One additional check about the $<0.01\%$ pixel-area fraction could not be completed due to a parsing issue in the executed check.

Checked items

- ✓ **C1_tableI_fwhm_range** (Page 2, Section II.A and Table I)
 - **Claim:** Text states beam FWHM ranges from 2.1' at 90 GHz to 1.0' at 220 GHz; Table I lists specific FWHM values per channel.
 - **Checks:** min_max_consistency
 - **Verdict:** PASS
 - **Notes:** Computed overall range from Table I values as 0.98–2.1 arcmin; $\text{max}_{90}=2.1$ matches 2.1' and $\text{min}_{220}=0.98$ is within 0.1 of 1.0'.
- ✓ **C2_map_dimensions_vs_pixel_size** (Page 2, Section II.A)
 - **Claim:** Maps are in CAR projection at 0.5' pixel resolution ($10,320 \times 43,200$ pixels).
 - **Checks:** unit_consistent_recomputation
 - **Verdict:** PASS
 - **Notes:** Extents recompute to 360° and 86° via $N_{\text{pix}} \cdot (0.5 \text{ arcmin}) / 60$.

3. ✓ **C3_detection_significance_formula_recompute_tableII_row1** (Page 2, Eq. (3) and Table II)

- **Claim:** Detection significance v is defined as $(y_{\text{peak}} - \bar{y})/\sigma_y$; Table II gives y and v for top candidates; verify v scales proportionally with y assuming common \bar{y}, σ_y (at least ratios between rows).
- **Checks:** ratio_consistency
- **Verdict:** PASS
- **Notes:** Heuristic ratio check: $y_1/y_2 \approx 1.04577$ and $v_1/v_2 \approx 1.04490$; implied $\bar{y} \approx -8.45 \times 10^{-7}$.

4. ✓ **C4_tableII_significance_count_thresholds** (Page 2-3, Section III.C and Table II)

- **Claim:** Claims: 15 candidates above 30σ , 3 above 40σ , and a single object at 51.2σ ; Table II lists top 10 with v values.
- **Checks:** threshold_count_consistency_partial
- **Verdict:** PASS
- **Notes:** Within the provided top-10 list: $\text{count}(> 40) = 3$, $\text{count}(> 30) = 7$, and exactly one entry equals 51.2; full-catalog '15 $> 30\sigma$ ' not verifiable from top-10 alone.

5. ✓ **C5_bullet_cluster_offset_consistency** (Page 3, Section III.D)

- **Claim:** Candidate 2 at $(104.63^\circ, -55.95^\circ)$ with $v = 49.0\sigma$ lies 3.4' from Bullet Cluster position.
- **Checks:** unit_sanity_and_rounding
- **Verdict:** PASS
- **Notes:** Plausibility-only: separation = 3.4' is within 0–60' and is $\sim 1.62 \times$ a 2.1' beam scale.

6. ✓ **C6_local_noise_factor_of_3** (Page 3, Section IV.A)

- **Claim:** Local noise RMS varies by approximately a factor of 3: from $\sim 2 \times 10^{-6}$ in deepest regions to $\sim 6 \times 10^{-6}$ at edges.
- **Checks:** ratio_recomputation
- **Verdict:** PASS
- **Notes:** Computed ratio $6e-6/2e-6 = 3.0$.

7. ✗ **C7_positive_tail_y_max_vs_tableII_max** (Page 4, Section IV.B and Table II)

- **Claim:** Positive tail extends to $y \approx 5.5 \times 10^{-4}$; Table II top candidate has $y = 4.57 \times 10^{-5}$ ($\times 10^{-5}$ units).
- **Checks:** order_of_magnitude_consistency
- **Verdict:** FAIL

- **Notes:** Computed ratio $y_{\text{tail_max}} / y_{\text{TableII\#1}} \approx 12.04$; flagged as a large discrepancy requiring definition/unit reconciliation.
8. **⚠ C8_20sigma_threshold_pixel_count_statement** (Page 4, Section IV.B)
- **Claim:** At the 20σ level, 864 pixels exceed $y > 10^{-4}$; these pixels constitute $< 0.01\%$ of total survey area.
 - **Checks:** `percentage_from_counts`
 - **Verdict:** UNCERTAIN
 - **Notes:** Executed check reports it could not parse map dimensions to compute the total pixel count/area fraction.
9. **✓ C9_kurtosis_sigma_claim_vs_noise_calibration** (Page 6, Section VII.A)
- **Claim:** Gaussian noise simulations yield $\kappa_{\text{noise}} = -0.0001 \pm 0.0001$; measured kurtosis $\kappa \approx 47$ corresponding to $>100\sigma$ above expectation.
 - **Checks:** `sigma_significance_recompute`
 - **Verdict:** PASS
 - **Notes:** Computed $z=(47-(-0.0001))/0.0001 \approx 470001$, which satisfies the claimed inequality ($>100\sigma$).
10. **✓ C10_split_kurtosis_vs_full_kurtosis** (Page 7, Section VII.A)
- **Claim:** Full-map measured kurtosis $\kappa \approx 47$; per split (set0, set1) $\kappa \approx 1.66$ per split.
 - **Checks:** `internal_consistency_sanity`
 - **Verdict:** PASS
 - **Notes:** Sanity check: $\kappa_{\text{full}} > \kappa_{\text{split}}$; ratio $47/1.66 \approx 28.31$.
11. **✓ C11_hemispherical_power_ratio_uncertainty** (Page 7, Section VII.B and Table V)
- **Claim:** Measured hemispherical power ratio $\langle R \rangle = 0.93 \pm 0.07$; claimed consistent with isotropy at 1σ .
 - **Checks:** `confidence_interval_contains_value`
 - **Verdict:** PASS
 - **Notes:** Interval is $[0.86, 1.00]$; isotropy value 1.0 lies on the upper boundary (treated as consistent).
12. **✓ C12_split_map_corr_range_ordering** (Page 7, Section VII.B)
- **Claim:** Set0/set1 split-map correlation varies from $r \sim 0.001$ to $r \sim 0.32$ with declination.
 - **Checks:** `range_sanity`
 - **Verdict:** PASS
 - **Notes:** Verified $0.001 < 0.32$ and both are within $[-1, 1]$.

13. ✓ **C13_buffer_and_dec_band_widths** (Page 7, Section VII.B)
- **Claim:** Northern region: $+4^\circ < \text{Dec} < +20^\circ$; Southern: $-21^\circ < \text{Dec} < -4^\circ$ with $\pm 4^\circ$ equatorial buffer; claim $\sim 16^\circ$ per hemisphere.
 - **Checks:** interval_width_recompute
 - **Verdict:** PASS
 - **Notes:** Computed widths: north= 16° , south= 17° ; within abs_tol= 1° for an approximate claim.
14. ✓ **C14_crossfreq_grid_cell_count** (Page 7, Section VII.C)
- **Claim:** Footprint divided into a 12×8 grid; correlation computed in each cell.
 - **Checks:** product_recompute
 - **Verdict:** PASS
 - **Notes:** $12 \times 8 = 96$ cells.
15. ✓ **C15_crossfreq_median_vs_expected_high_corr** (Page 7, Section VII.C)
- **Claim:** States expected correlation $\rho \geq 0.9$ in signal-dominated regime; observed median correlation is $\rho_{\text{med}} = 0.69$.
 - **Checks:** inequality_check
 - **Verdict:** PASS
 - **Notes:** Confirmed $\rho_{\text{med}} < 0.9$; difference $0.9 - 0.69 = 0.21$.
16. ✓ **C16_four_anomalous_cells_count** (Page 7, Section VII.C and Table V)
- **Claim:** Four cells show $\rho < 0.12$; Table V also lists '4 regions' with $\rho < 0.12$.
 - **Checks:** repeated_value_consistency
 - **Verdict:** PASS
 - **Notes:** Counts match (4 and 4) with the same threshold 0.12.
17. ✓ **C17_morphological_outliers_zscore_threshold** (Page 8, Section VIII and Table III)
- **Claim:** Five candidates flagged as morphological outliers with z-score > 2 ; Table III lists five rows with z-scores 6.2, 4.2, 2.4, 2.9, 2.9.
 - **Checks:** threshold_count_consistency
 - **Verdict:** PASS
 - **Notes:** All five listed z-scores exceed 2; min z=2.4.
18. ✓ **C18_fsky_value** (Page 8, Section IX.A)
- **Claim:** Common mask has $f_{\text{sky}} \approx 0.46$.
 - **Checks:** range_sanity
 - **Verdict:** PASS
 - **Notes:** $f_{\text{sky}} = 0.46$ is within $[0, 1]$.

19. ✓ **C19_scale_cut_ordering** (Page 8-9, Section IX.A and Table IV)

- **Claim:** Pair-specific scale cuts: 220-related limited to $\ell \lesssim 1000$; 90×150 to $\ell \lesssim 1200$; same-band pairs to $\ell \lesssim 1500$. Table IV lists ranges 500–1000, 500–1200, 400–1500.
- **Checks:** range_consistency
- **Verdict:** PASS
- **Notes:** Verified ordering $1500 \geq 1200 \geq 1000$ and exact match to narrative maxima.

20. ✓ **C20_birefringence_combined_vs_components** (Page 9, Section X.B)

- **Claim:** Constraints: $|\beta| < 0.01^\circ$ (1σ); $\beta_{EB} \approx 0.0006^\circ \pm 0.001^\circ$; $\beta_{TB} \approx 0.006^\circ \pm 0.01^\circ$.
- **Checks:** consistency_of_bounds
- **Verdict:** PASS
- **Notes:** Computed $|\mu| + 1\sigma$: $EB = 0.0016 \leq 0.01$ supports overall bound; $TB = 0.016 > 0.01$ is weaker but does not contradict if combined bound is EB-dominated.

21. ✓ **C21_isocurvature_limits_vs_units** (Page 9, Section X.C)

- **Claim:** Measured $\langle C_{TB_\ell} \rangle < 0.001 \mu K^2$ and $\langle C_{EB_\ell} \rangle < 0.0001 \mu K^2$ for $10 \leq \ell \leq 100$.
- **Checks:** order_and_unit_sanity
- **Verdict:** PASS
- **Notes:** Verified $EB/TB = 0.0001/0.001 = 0.1$ ($10\times$ tighter).

22. ✗ **C22_tableV_trials_factor_inconsistency** (Page 10, Section XI and Page 11, Section XII.B)

- **Claim:** One place states 'effective trials factor from eleven parallel analyses'; later states 'With twelve parallel analyses' and 'multiply p-value by 12'; also earlier says Bonferroni dividing significance level by 11.
- **Checks:** repeated_constant_consistency
- **Verdict:** FAIL
- **Notes:** Stated trials factors disagree: {11, 12}.

23. ✓ **C23_spectral_diagnostics_fraction_checks** (Page 5, Section V.B and Table V)

- **Claim:** Raw-map spectral diagnostics: 2 of 20 candidates tSZ-consistent; srcfree: 1 of 20 tSZ-consistent.
- **Checks:** fraction_recompute
- **Verdict:** PASS
- **Notes:** Exact arithmetic: $2/20 = 0.10$ and $1/20 = 0.05$.

Limitations

- Only parsed text from the PDF is available; no underlying ACT/NILC map data or full candidate catalogs are included, preventing verification of claims that depend on recomputing detections, correlations, power spectra, or pixel statistics.
- Figures are not used for numeric extraction (per instruction to avoid reading plot pixels), so any values only shown graphically cannot be checked.
- Some statements involve external reference positions (e.g., Bullet Cluster coordinates) or external templates; these cannot be verified using the PDF alone.
- Several numeric claims are approximate (use of \sim , \approx , \gtrsim); checks use tolerances and may only validate plausibility rather than exact equality.
- One executed check (C8) was inconclusive due to an internal parsing failure for map dimensions needed to compute the area fraction from the stated pixel count.