

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 manuscript presents a wide-ranging, exploratory “screening” of ACT DR6 temperature maps (90/150/220 GHz) and the ACT–Planck NILC Compton- y product. It includes (i) a blind, threshold-based search for high-significance tSZ candidates in the NILC y -map (yielding a 200-object $\geq 5\sigma$ catalog), (ii) basic radial-profile morphology characterization of the brightest systems, (iii) simple multi-frequency temperature checks at candidate positions, (iv) several map-level statistical diagnostics (one-point y statistics; large excess kurtosis at 150 GHz; hemispherical power ratio tests; coarse grid-based f_{090} – f_{150} correlation “flags”), and (v) a set of internal null/consistency tests including split-map checks and polarization-based TB/EB-derived quantities. The paper’s breadth and transparency make it potentially useful as a DR6 “health check” and as a generator of follow-up targets, but many quantitative statements currently rest on methods that are either (a) too approximate (global thresholds in an inhomogeneous-noise map; single-pixel spectral ratios without beam matching), or (b) insufficiently documented to be reproducible (coherence/scale-cut methodology deferred to an internal report; incomplete power-spectrum estimation details on a cut sky). Several results labeled as anomalies (kurtosis significance, low-correlation cells, profile outliers, spectral outliers) need explicit null models, uncertainties, and a consistent multiple-testing treatment to separate robust findings from exploratory flags. Finally, there are reproducibility-critical presentation issues (notably unphysical coordinates in Table III and figure/caption mismatches) that must be corrected before publication.

Strengths

- Covers an unusually broad set of complementary diagnostics—tSZ candidate finding, y -noise inhomogeneity, one-point statistics, non-Gaussianity, isotropy tests, cross-frequency correlation/coherence, morphology outlier flagging, and internal null tests—giving readers a holistic view of ACT DR6 map-level behavior (Secs. III–XIII).
- Implements an independent blind search in the NILC Compton- y map and performs a basic sanity check by recovering known massive systems (e.g., Bullet Cluster) and characterizing bright-cluster profiles (Secs. III–IV, VIII).
- Clearly identifies concrete follow-up targets/regions (e.g., a compact feature with synchrotron-like spectrum but high y significance; low f_{090} – f_{150} correlation cells; morphology outliers), which is valuable for multi-wavelength validation (Secs. VI–VIII).

- The manuscript is generally candid that the analyses are exploratory and that foregrounds/noise/selection effects can dominate naïve statistics, which is the right framing for a screening paper (Secs. V, VII, XI–XII).
- Core definitions (Compton- y , tSZ temperature response, detection significance ν , Pearson ρ , coherence ρ_ℓ) are introduced and, once typeset cleanly, can provide a consistent language across the many tests (Secs. III, VII, IX).

Major issues

1. **tSZ candidate detection/catalog construction is not noise-aware and is insufficiently specified.** The catalog uses a single global threshold $\nu = (y_{\text{peak}} - \bar{y})/\sigma_y$ (Eq. (3), Sec. III.A) despite explicitly strong spatial noise inhomogeneity in the NILC y map (σ_y^{local} varies by ~ 3 ; Sec. IV.A/Fig. 6). This makes “ $\geq 5\sigma$ ” position-dependent and undermines interpretation of candidate counts, ν -distributions, morphology statistics of the “50 brightest,” and any inferred rarity (Secs. III.C–III.D, IV, VIII, XI–XII). In addition, the relationship between the reported $4194 > 5\sigma$ peaks at $2'$ smoothing (Sec. III.B) and the final 200-candidate $\geq 5\sigma$ catalog (Sec. III.C) is unclear (changes in smoothing scale, masking, peak merging, connected-component rules, edge handling, etc.). The text/captions also sometimes describe the smoothing-scale search as “matched filtering,” which it is not in the standard tSZ sense (Secs. III.B–III.C; Fig. 1).

Recommendation: Make the detection pipeline both reproducible and interpretable under inhomogeneous noise. Concretely: (i) define a local-noise significance ν_{local} using the σ_y^{local} map already computed in Sec. IV.A (same grid/resolution), and provide both ν_{global} and ν_{local} per object; (ii) explicitly list every step that maps the initial peak list to the final 200 (smoothing scale used for the published catalog; peak-finding method; connected-component definition; minimum area; deblending/merging radius; masking; edge cuts); (iii) quantify how candidate density depends on σ_y^{local} (or declination / distance to edges) and provide at least a coarse purity/completeness estimate via injections into representative low- and high-noise regions; (iv) rephrase “matched filter” as “Gaussian smoothing” unless a true matched filter (beam+template+noise PSD weighting) is implemented, and ensure Fig. 1/captions match the actual method.

2. **Multi-frequency spectral diagnostics (Secs. V.A–V.C, Table V) rely on single-pixel values in raw $f_{090}/f_{150}/f_{220}$ maps without beam matching, consistent filtering/transfer functions, aperture photometry/background subtraction, or split-based uncertainties.** Despite this, the manuscript draws quantitative-sounding conclusions (e.g., “only 1–2/20 show classical tSZ behavior,” most are foreground dominated) that are not supported by single-pixel, mismatched-resolution measurements in a CMB+foreground+anisotropic-noise field (Secs. V, XI–XIII).

Recommendation: Either strengthen Sec. V into a minimal but defensible photometric analysis or downscope the claims. Preferred: convolve all three frequency maps to a common beam (worst resolution), apply consistent filtering (or explicitly justify differences), and perform aperture photometry or matched-filter flux extraction at candidate positions with local annular background subtraction; use split maps to estimate per-band uncertainties and propagate them to spectral ratios (including covariances where relevant). Include bandpass/units details (μK_{CMB} ; effective band centers; expected tSZ null not exactly at nominal 220) and show how many candidates are consistent with a tSZ SED within errors. If this is beyond scope, recast the section as qualitative “foreground complexity at candidate positions,” remove/soften hard fractions like “2/20,” and add uncertainty estimates that demonstrate the classification is not robust.

3. **The highlighted compact feature at (RA, Dec) \approx (291.2°, -29.2°) (Sec. VI) is interpreted as synchrotron-like ($\alpha \approx -0.4$) yet very high-significance in Compton- y ($\approx 41\sigma$), but the measurements are under-specified and lack uncertainty quantification and basic external validation.** It is currently unclear (i) how α is estimated (single pixel vs aperture; whether $f220$ is included), (ii) whether the feature is stable in map splits, and (iii) whether it coincides with known radio/X-ray/optical/SZ sources. Given NILC’s sensitivity to foreground leakage/weighting, a radio source can plausibly bias a y reconstruction without being a true tSZ detection (Secs. VI.A–VI.C, XII.A, XIII).

Recommendation: Turn Sec. VI into a minimal, robust validation case study. Specify the photometric method (beam-homogenized aperture or matched-filter fluxes at $f090/f150/f220$ with background subtraction), fit α using all relevant bands, and report fluxes/temperatures and α with uncertainties (add error bars to Fig. 13). Repeat the measurement on independent splits (set0/set1 or equivalent) to demonstrate stability of both the y detection and the SED. Perform a basic cross-match (even a simple cone search) against appropriate radio catalogs (e.g., SUMSS/NVSS/PMN/AT20G depending on declination), major SZ catalogs (official ACT DR6; Planck; SPT), and X-ray catalogs (ROSAT/eROSITA if available), and report matches/non-matches explicitly. Based on these checks, adjust the language: either present it as a confirmed known system (cluster+radio AGN), or clearly label it as an exploratory “foreground/NILC-leakage flag” rather than a likely novel SZ+synchrotron object.

4. **Non-Gaussianity and “anomaly” claims (kurtosis, low-correlation cells, spectral outliers, morphology outliers) lack adequate null models, preprocessing detail, uncertainties, and consistent multiple-testing treatment.** The kurtosis result $\kappa \approx 47$ in f150 (Sec. VII.A) is compared to overly simplified Gaussian-noise simulations (“ $> 100\sigma$ ”), which is not an appropriate null for ACT temperature maps that contain CMB fluctuations, foregrounds, anisotropic/correlated noise, and filtering/transfer-function effects. Similarly, outlier classes—four low $f090$ – $f150$ corre-

lation cells (Sec. VII.C), “spectral ratio outliers” (Sec. V.C), and five morphology profile outliers (Sec. VIII)—are flagged without p -values under an explicit null, error bars on the metrics, or a look-elsewhere correction consistently applied (Secs. XI–XII). There is also an internal inconsistency in the multiple-testing discussion (11 vs 12 analyses; Secs. XI and XII.B).

Recommendation: For Sec. VII.A, document preprocessing step-by-step (masking/apodization; monopole/gradient removal; beam smoothing; any filtering; handling of masked pixels/edges; patch-selection scheme for 2000×2000 pixel patches), and calibrate κ against a realistic baseline (at minimum: lensed CMB + anisotropic noise consistent with splits; ideally also point sources/foreground residuals or apply a point-source mask and state it). Report κ with uncertainty (across patches) and show robustness to masking choices. For Secs. V.C, VII.C, VIII: define explicit null hypotheses and compute per-test p -values (via simulations or analytic approximations), include uncertainties on each statistic (spectral ratios, Pearson ρ per cell, morphology metrics), and then apply a clearly defined trials correction using a single, consistent N (list which analyses count toward it). Summarize per-test and global (post-trials) significance in a compact table in Sec. XI or XII.B, and reframe items that do not survive corrections as exploratory “flags.”

5. **Large-scale / harmonic-space methods are not sufficiently documented for reproducibility or for judging reliability on a partial CAR footprint.** This affects (i) the hemispherical power ratio test (Sec. VII.B) at very low multipoles (e.g., $2 \leq \ell \leq 50$) where ACT map filtering/transfer functions and cut-sky mode coupling can be important, and (ii) the cross-frequency coherence analysis and associated “scale-cut recommendations” (Sec. IX, Table IV), which are largely deferred to an internal report [23] without enough details to reproduce or assess (masks and f_{sky} , estimator type, beam/calibration treatment, noise debiasing, ℓ -binning, covariance estimation, null tests).

Recommendation: Add a concise but complete methods description in Secs. VII.B and IX. For hemispherical asymmetry: specify the exact sky regions/declination cuts, masks, apodization, pseudo- C_ℓ /MASTER (or alternative) estimator, mode-coupling correction, transfer-function/beam treatment, binning, and how uncertainties are obtained (splits vs sims). Justify the inclusion of $\ell \leq 50$ for these map products or restrict to a validated ℓ -range. For coherence (Sec. IX): describe preprocessing, the spectra used (cross-spectra between splits vs auto-spectra), noise debiasing, beam/calibration handling, binning, and uncertainties; state whether [23] is public, and if not, ensure Sec. IX alone is sufficient to reproduce Table IV. Rephrase “recommended scale cuts” as either diagnostic suggestions or analysis requirements, but make that distinction explicit in Sec. XII.C and the Conclusions.

6. **Several null/consistency-test claims are currently under-defined or potentially overstated, particularly those framed as constraints.** The cosmic birefringence bound $|\beta| < 0.01^\circ$ (Sec. X.B; Eq. (5)) is presented without the explicit TB/EB-to- β estimator equations, weighting, sign conventions, ℓ -range/binning, E/B purification and mode-coupling corrections, or inclusion of instrument polarization-angle systematics—details essential to assess whether such a tight bound is realistic for the dataset. Similarly, the “isocurvature limits” based on TB/EB consistency with zero (Sec. X.C) are not connected to a defined isocurvature parameter/model, and the rSZ- y sign-consistency discussion (Sec. X.E) is largely qualitative.

Recommendation: In Sec. X.B, provide the explicit small-angle rotation relations and the exact estimator used to infer β from C_ℓ^{EB} and/or C_ℓ^{TB} , including ℓ -weighting/bandpower combination, sign conventions for $Q/U \rightarrow E/B$, and treatment of masks/mode coupling (and whether E/B purification is used). State which systematics are included in the β uncertainty (especially polarization-angle calibration and leakage). In Sec. X.C, either reframe as a parity/null test (preferred given current content) or introduce a specific isocurvature model and show how TB/EB map to a parameter constraint. In Sec. X.E, quantify the rSZ- y check (e.g., pixel-wise or patch-wise correlation coefficients; dependence on y S/N; association with NILC weights/foreground templates) and provide at least one summary plot/table beyond anecdotal regions.

7. **Reproducibility-critical numerical/coordinate and figure/caption inconsistencies are present.** Most urgently, Table III lists declinations outside the physical range (e.g., $+166.3^\circ$, $+173.8^\circ$), which invalidates follow-up of the Sec. VIII morphology outliers. There are also smaller inconsistencies (e.g., Bullet Cluster ν quoted with different rounding across sections; unclear identification of the $\nu_{\text{max}} = 51.2\sigma$ object; ratio-definition flips between $T90/T150$ and $T150/T090$), and at least one figure/caption mismatch (Fig. 1 labeled as matched-filter optimization though Gaussian smoothing is used; mention of missing panels).

Recommendation: Perform a full audit of tables, coordinates, and repeated numeric claims (Secs. III–IV, VI, VIII, XI–XIII). Correct Table III positions and clearly state the coordinate system/epoch (e.g., ICRS/J2000 RA/Dec in degrees). Ensure the ν_{max} object is identified by catalog ID and coordinates in Secs. XII–XIII. Harmonize ratio definitions and sign/inequality interpretations across Sec. V and related figures. Fix figure completeness and ensure captions exactly match what is plotted and the method actually used (e.g., rename “matched filter” to “Gaussian smoothing” unless upgraded).

Minor issues

1. Connection of the 200 NILC y -map candidates to existing catalogs is discussed only qualitatively (e.g., Bullet Cluster recovery), limiting assessment of novelty, purity, and completeness (Secs. III.C, IV.C, XII–XIII).

Recommendation: Add a systematic cross-match to the official ACT DR6 cluster catalog [7] and (as feasible) Planck/SPT, reporting match fraction within a stated radius (e.g., $2'–3'$), typical offsets, and how matched/unmatched objects distribute in ν_{local} , σ_y^{local} , and proximity to footprint edges. If the full 200-object list is not in the main paper, provide it as a machine-readable supplement.

2. Morphology metrics and outlier selection in Sec. VIII (Fig. 17, Table III) are not defined with enough precision to reproduce: centering choice, radial binning, background subtraction, beam handling (deconvolved vs smoothed), and uncertainty estimation for concentration/FWHM/slope are unclear; z -score construction lacks a stated reference distribution/covariance.

Recommendation: In Sec. VIII, formally define each metric and the profile extraction pipeline, include uncertainty estimation (e.g., from noise/split realizations or bootstrap), and state how the multivariate reference distribution (mean/covariance) for z -scores is obtained. If morphology depends strongly on redshift/angular size, discuss this and/or control for it when flagging outliers.

3. Cross-frequency correlation grid analysis (Sec. VII.C, Fig. 16) lacks essential implementation details and uncertainty on ρ per cell, making it hard to assess whether the median $\rho \approx 0.69$ and the four $\rho < 0.12$ cells are unexpected given beams/noise/foregrounds.

Recommendation: State whether maps are beam-matched and similarly filtered before computing ρ , what masking and mean/gradient removal is applied per cell, how many pixels contribute per cell, and provide uncertainties (bootstrap, split-based, or simulations). Compare the observed distribution of ρ to a simple signal+noise model or simulations to justify calling specific cells “anomalous” versus “flagged for follow-up.”

4. Paper organization and formatting inconsistencies (section heading styles, early cross-references, terse table/figure introductions) reduce readability, especially given the many disparate diagnostics (Secs. I–XIII).

Recommendation: Standardize section/subsection formatting (consistent LaTeX `\section/\subsection` numbering), ensure cross-references resolve unambiguously, and add brief “what this figure/table shows and why it matters” text at first mention for the densest multi-panel figures/tables (notably those supporting Secs. V, VII, IX–X). Consider adding a short reproducibility appendix listing map versions, masks, smoothing kernels, and key parameter choices.

Very minor issues

1. Numerous typographical/typesetting issues (likely OCR artifacts) appear throughout: HTML escape sequences in inequalities (“>”), inconsistent spacing/formatting in equations and units (e.g., Eq. (5) rendering of $|\beta| < 0.01^\circ$), and occasional garbling in Eq. (4) coherence notation.

Recommendation: Proofread and clean the LaTeX source: replace HTML escapes with proper TeX, standardize spacing and units, ensure Eq. (4) is unambiguous (and state whether C_ℓ are cross-spectra between splits), and render Eq. (5) cleanly with consistent typography.

2. Figure accessibility and self-containment: some plots lack clear units/normalizations, sample sizes, or processing notes (masking/beam/filtering), and some color choices may be suboptimal for colorblind/grayscale viewing.

Recommendation: Make captions more self-contained (units, preprocessing, selection), increase label/font readability for print, and use colorblind-safe palettes and/or redundant line styles/markers where multiple series are compared.

3. Minor notation/style inconsistencies (e.g., “tSZ/TSZ”, “Compton-y/Compton-y”, “Fig.” vs “Figure”, frequency labels f_{090}/f_{150} vs f_{090}, f_{150} ; ratio definition flips in Sec. V).

Recommendation: Standardize notation and naming throughout text, captions, and tables, and keep a single convention for spectral ratio definitions (or prominently remind readers when using the inverse).

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 originally developed for full-sky CMB mapping, which has been described and validated in detail in prior studies of WMAP and Planck data using needlet ILC and constrained ILC methods.**
- *Reference(s):* 6, 12, 13
- • **A companion study measured the pairwise cross-correlation coefficient $\rho_\ell = C_\ell^{ab} / \sqrt{C_\ell^{aa} C_\ell^{bb}}$ across all 15 ACT DR6 channel pairs on a common mask with $f_{\text{sky}} \approx 0.46$, finding that same-band 150 GHz pairs maintain $\rho_\ell \gtrsim 0.99$ over $400 \lesssim \ell \lesssim 1500$, cross-frequency 90×150 pairs show $\rho_\ell \gtrsim 0.98$ over $500 \lesssim \ell \lesssim 1200$, and 220 GHz-related pairs have lower coherence limited to $\ell \lesssim 1000$ by dust foregrounds, thereby providing pair-specific scale-cut recommendations for downstream analyses.**
- *Reference(s):* 23

- • Using the PA5 f150 set0×set1 polarization cross-spectra over $10 \leq \ell \leq 100$, the authors constrain the cosmic birefringence rotation angle of the CMB polarization plane to $|\beta| < 0.01^\circ$ (1σ), with $\beta_{\text{EB}} \approx 0.0006^\circ \pm 0.001^\circ$ from EB and $\beta_{\text{TB}} \approx 0.006^\circ \pm 0.01^\circ$ from TB, consistent with zero rotation and compatible with existing Planck limits.
- *Reference(s)*: Planck Collaboration
- • The measured excess kurtosis of the ACT DR6 temperature field, $\kappa \approx 47$ (more than 100σ above Gaussian-noise simulations), is interpreted as arising from non-Gaussian contributions of unresolved extragalactic point sources, the thermal Sunyaev–Zel’dovich effect from low-mass clusters, and the cosmic infrared background, in line with theoretical and observational expectations for secondary CMB anisotropies and primordial non-Gaussianity constraints from previous work.
- *Reference(s)*: 17, 18, 19

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 core equations defining the Compton- y parameter, the non-relativistic tSZ spectral response, a simple detection significance, and several statistical diagnostics (kurtosis, correlation coefficients, power ratio). Most results are reported descriptively without detailed derivations, so the audit focuses on dimensional consistency, correctness of stated formulas, and consistency of definitions/notation across sections.

Checked items

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

- **Claim:** Defines the Compton- y parameter as $y = (\sigma_{\text{T}}/(m_e c^2)) \int n_e k_{\text{B}} T_e dl$.
- **Checks:** dimensional/units consistency, symbol definition consistency
- **Verdict:** PASS; confidence: high; impact: critical
- **Assumptions/inputs:** Line-of-sight integral over electron pressure $n_e k_{\text{B}} T_e$, Non-relativistic Thomson scattering regime
- **Notes:** Units are consistent: σ_{T} has area, $n_e k_{\text{B}} T_e dl$ gives energy per area, dividing by $m_e c^2$ gives dimensionless y . All symbols appearing are defined locally.

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

- **Claim:** States $\Delta T_{\text{tSZ}}/T_{\text{CMB}} = y[x \coth(x/2) - 4]$, with $x = h\nu/(k_{\text{B}} T_{\text{CMB}})$.

- **Checks:** algebraic form sanity-check, dimensionless consistency, limiting/sanity case
 - **Verdict:** PASS; confidence: high; impact: critical
 - **Assumptions/inputs:** Non-relativistic tSZ limit, Thermodynamic temperature convention for $\Delta T/T_{\text{CMB}}$
 - **Notes:** Expression is dimensionless and consistent with a frequency-dependent factor multiplying y . The paper's qualitative statements (decrement below ~ 217 GHz and increment above) match the sign behavior implied by the bracketed term.
3. ✓ **Detection significance definition** (Eq. (3), Sec. III.A, p.2)
- **Claim:** Defines $\nu \equiv (y_{\text{peak}} - \bar{y})/\sigma_y$ for thresholded detections in the y map.
 - **Checks:** definition consistency, normalization sanity-check
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** \bar{y} and σ_y computed over valid (masked) pixels, σ_y interpreted as global pixel RMS
 - **Notes:** Definition is internally consistent with the stated threshold $y > \bar{y} + 5\sigma_y$. The paper later interprets ν as ' σ ' significance; that is consistent only under an (unstated) approximate Gaussian-pixel assumption.
4. ✓ **Spectral ratio definition for candidate diagnostics** (Sec. V.A, p.4)
- **Claim:** Defines spectral ratio $R \equiv T_{90}/T_{150}$ using single-pixel temperatures at candidate positions.
 - **Checks:** symbol/definition consistency
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** T_{90} and T_{150} are in the same temperature units/convention, Single-pixel extraction is used as a diagnostic
 - **Notes:** The ratio definition is clear. Later sections use inverse ratios (T_{150}/T_{90}) for map visualization; not incorrect but requires careful interpretation of inequalities.
5. ✓ **Frequency difference and ratio maps** (Sec. V.C, p.5)
- **Claim:** Defines $\Delta T = T_{150} - T_{90}$ and uses T_{150}/T_{90} as a pixel-level spectral ratio.
 - **Checks:** algebraic correctness, notation consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** T_{150} and T_{90} share a common temperature convention, Division performed only on high-S/N pixels (as stated in figure caption context)

- **Notes:** Definitions are algebraically trivial and consistent. The paper correctly treats these as qualitative component tracers rather than a full component separation.
6. ✓ **Spectral index power-law convention** (Sec. VI.A, p.6)
- **Claim:** Uses $T_\nu \propto \nu^\alpha$ and reports α from the $f090/f150$ ratio.
 - **Checks:** definition sanity-check
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** Power-law approximation over 90–150 GHz, T_ν denotes a band-averaged temperature-like quantity
 - **Notes:** The convention $T_\nu \propto \nu^\alpha$ is self-consistent. The paper does not provide the explicit inversion formula $\alpha = \log(T_{90}/T_{150})/\log(\nu_{90}/\nu_{150})$; omission is acceptable but reduces auditability of sign conventions.
7. ✓ **Excess kurtosis definition** (Sec. VII.A, p.6)
- **Claim:** Defines excess kurtosis $\kappa \equiv \langle T^4 \rangle / \langle T^2 \rangle^2 - 3$.
 - **Checks:** formula correctness, normalization/sanity-check
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** Moments taken over pixels in a patch/window, T is mean-subtracted or mean negligible (not explicitly stated)
 - **Notes:** This is the standard excess kurtosis definition. If T is not mean-subtracted, the definition should technically use central moments; the paper does not state this, so interpretation depends on preprocessing (unclear but not an algebraic error).
8. △ **Split-map kurtosis 'confirmation' vs coadd kurtosis** (Sec. VII.A, p.6–7)
- **Claim:** States $\kappa \approx 47$ on the coadd and that split maps 'confirm' excess kurtosis with $\kappa \approx 1.66$ per split.
 - **Checks:** internal consistency/sanity-check
 - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
 - **Assumptions/inputs:** Coadd and split maps have different noise levels, κ computed in the same way across maps
 - **Notes:** A large reduction in standardized kurtosis from coadd to noisier splits can be mathematically plausible, but the paper provides no analytic explanation or mapping. Without clarifying whether κ is computed on identical preprocessing/masking and how noise affects κ , 'confirmation' is not verifiable.
9. ✓ **Hemispherical power ratio definition** (Sec. VII.B, p.7)
- **Claim:** Defines $R(\ell) = C_\ell^N / C_\ell^S$ using noise-bias-free split cross-power spectra $C_\ell^{\text{set0}} \times \text{set1}$.

- **Checks:** definition consistency, symbol consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** C_ℓ^N and C_ℓ^S are constructed consistently from cross-spectra, Same ℓ -binning and mask treatment in N and S regions
 - **Notes:** The ratio definition is coherent with the described methodology and avoids auto-spectrum noise bias (at least at the level described).
10. ✓ **Pearson cross-frequency correlation coefficient usage** (Sec. VII.C, p.7)
- **Claim:** Computes Pearson correlation $\rho(f090, f150)$ per grid cell and interprets low values as foreground/systematic indicators.
 - **Checks:** conceptual/statistical definition sanity-check
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** ρ is computed over pixel pairs within each cell, Maps are on common pixelization and comparable filtering
 - **Notes:** No explicit formula is printed, but the use of 'Pearson correlation ρ ' is standard and internally consistent with the later histogram/threshold discussion.
11. ✓ **Harmonic-space coherence coefficient** (Eq. (4), Sec. IX, p.8)
- **Claim:** Defines $\rho_\ell = C_\ell^{ab} / \sqrt{C_\ell^{aa} C_\ell^{bb}}$ for channel-pair coherence.
 - **Checks:** formula correctness, normalization/sanity-check, notation clarity
 - **Verdict:** PASS; confidence: medium; impact: moderate
 - **Assumptions/inputs:** C_ℓ^{ab} is a cross-spectrum; C_ℓ^{aa} , C_ℓ^{bb} are corresponding auto-spectra (or noise-bias-free estimates thereof)
 - **Notes:** The expression is mathematically consistent as a normalized cross-correlation in ℓ -space. However, the paper does not specify whether C_ℓ^{aa} and C_ℓ^{bb} are auto-spectra (noise-biased) or cross-spectra between splits (noise-bias-free), which affects interpretability but not algebra.
12. ✓ **Polarization amplitude definition** (Sec. X.D, p.9)
- **Claim:** Defines $P = \sqrt{Q^2 + U^2}$ and compares P across frequencies.
 - **Checks:** formula correctness
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** Q and U are in the same units and coordinate convention in each map
 - **Notes:** Definition is correct. (Note: P is Rice-distributed in noise; the paper uses it as a null consistency check, but no analytic bias correction is discussed—outside strict algebraic correctness.)
13. △ **Cosmic birefringence bound statement** (Eq. (5) and Sec. X.B, p.9)

- **Claim:** Reports $|\beta| < 0.01^\circ$ (1σ) from set0×set1 TB/EB spectra over $10 \leq \ell \leq 100$; quotes β_{EB} and β_{TB} estimates.
 - **Checks:** derivation completeness, definition/symbol consistency
 - **Verdict:** UNCERTAIN; confidence: low; impact: critical
 - **Assumptions/inputs:** A specific estimator relating observed TB/EB to β under a rotation model, Small-angle approximation and/or known unrotated spectra
 - **Notes:** The paper does not show the estimator (how β is computed from TB/EB), nor the assumptions about underlying spectra and sign conventions. As a result, the bound cannot be audited symbolically from the PDF alone.
14. ✘ **Bonferroni/look-elsewhere correction consistency** (Sec. XI, p.10 and Sec. XII.B, p.11)
- **Claim:** States the global significance should be adjusted by a Bonferroni correction dividing by 11 (Sec. XI) and elsewhere multiplying p -values by 12 (Sec. XII.B).
 - **Checks:** internal consistency
 - **Verdict:** FAIL; confidence: high; impact: moderate
 - **Assumptions/inputs:** A fixed number of independent (or at least counted) parallel analyses
 - **Notes:** The number of parallel tests is inconsistent (11 vs 12) and the described correction differs in presentation. This is a clear internal inconsistency in the statistical accounting (even though it does not affect algebraic formulas elsewhere).
15. ✘ **Coordinate validity in Table III** (Table III, Sec. VIII, p.8)
- **Claim:** Lists outlier clusters with RA/Dec and z -scores; several Dec values exceed $\pm 90^\circ$.
 - **Checks:** units/range consistency
 - **Verdict:** FAIL; confidence: high; impact: minor
 - **Assumptions/inputs:** Dec is standard declination in degrees
 - **Notes:** Declination values like $+166.3^\circ$ are outside the valid range for Dec, indicating a reporting/format error (e.g., wrong column, wrong units, or transcription problem).

Limitations

- Audit is limited to the equations and definitions explicitly present in the provided PDF text/images; many reported constraints (e.g., birefringence, isocurvature) are not derivable from the paper because the estimators/likelihood relations are not shown.

- Some equation text appears to contain OCR/transcription artifacts (e.g., Eq. (4) rendering), which can obscure notation-level checks; verdicts assume the intended standard forms when unambiguous from context.
- No external references were used to validate domain formulas; checks rely only on internal dimensional/algebraic consistency and the paper's own stated definitions.

Numerical results audit

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

17 candidate numerical checks were specified, but no checks were executed due to an execution error; consequently, all candidate checks are reported as UNCERTAIN based strictly on the provided inputs. Two internal-consistency risks are explicitly present in the inputs (trials factor 11 vs 12; $\kappa \approx 47$ vs $\kappa \approx 1.66$ per split), and several additional items are explicitly marked unverified because they require underlying map/figure data or external coordinates.

Checked items

1. \triangle C1 (Page 2, Section II.A (Frequency Maps))

- **Claim:** CAR projection at $0.5'$ pixel resolution corresponds to a map of $10,320 \times 43,200$ pixels.
- **Checks:** unit-consistent numeric recomputation (pixelization vs angular size)
- **Verdict:** UNCERTAIN
- **Notes:** Execution did not run; recomputation of implied angular extents was not performed.

2. \triangle C2 (Page 2, Table I)

- **Claim:** Beam FWHM ranges from $\sim 2.1'$ at 90 GHz to $\sim 1.0'$ at 220 GHz.
- **Checks:** min/max consistency between text and table
- **Verdict:** UNCERTAIN
- **Notes:** Execution did not run; table parsing and comparison to narrative values was not performed.

3. \triangle C3 (Page 2, Eq. (3) and Table II)

- **Claim:** Detection significance is $\nu \equiv (y_{\text{peak}} - \bar{y})/\sigma_y$; Table II reports (y, ν) for top candidates.
- **Checks:** cross-row constant check (implied linear relation)
- **Verdict:** UNCERTAIN
- **Notes:** Execution did not run; linear fit across Table II rows was not performed.

4. \triangle C4 (Page 2, Section III.C (Catalog))

- **Claim:** Significance distribution: 15 candidates above 30σ , 3 above 40σ , and a single object at 51.2σ .
 - **Checks:** count consistency from explicit table entries
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; counts from Table II ν values were not re-computed (and $> 30\sigma$ total requires full catalog beyond top-10 list).
5. \triangle **C5** (Page 3, Section III.D (Bullet Cluster Recovery))
- **Claim:** Candidate 2 has $\nu = 49.0\sigma$ and lies $3.4'$ from the Bullet Cluster position.
 - **Checks:** cross-reference consistency (table vs text)
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; verification of the repeated ν value was not performed. (Separation component is separately listed as unverified in inputs.)
6. \triangle **C6** (Page 3, Section IV.A (Overview and Noise Properties))
- **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 consistency
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; ratio recomputation was not performed.
7. \triangle **C7** (Page 4, Section IV.B (One-Point Distribution))
- **Claim:** At the 20σ level, 864 pixels exceed $y > 10^{-4}$; these pixels constitute $< 0.01\%$ of the total survey area.
 - **Checks:** percentage bound check using map pixel count
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; fraction vs full-grid bound was not recomputed. Inputs also note potential ambiguity between full CAR grid and masked footprint.
8. \triangle **C8** (Page 6, Section VII.A (Non-Gaussianity))
- **Claim:** Null simulations yield $\kappa_{\text{noise}} = -0.0001 \pm 0.0001$; measured kurtosis is $\kappa \approx 47$, corresponding to $> 100\sigma$ above expectation.
 - **Checks:** sigma significance recomputation
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; z-score recomputation was not performed.
9. \times **C9** (Page 7, Section VII.A (split-map verification))

- **Claim:** Split-map verification confirms excess kurtosis with $\kappa \approx 1.66$ per split.
 - **Checks:** internal plausibility / inconsistency flag (same metric different values)
 - **Verdict:** FAIL
 - **Notes:** Inputs specify $\kappa_{\text{full}} \approx 47$ and $\kappa_{\text{split}} \approx 1.66$, implying a ratio ≈ 28 (> 5 threshold noted in the check definition), which is flagged as requiring clarification.
10. \triangle **C10** (Page 7, Section VII.B (Hemispherical Power Asymmetry))
- **Claim:** Declination ranges: northern ($+4^\circ < \text{Dec} < +20^\circ$) and southern ($-21^\circ < \text{Dec} < -4^\circ$) with a $\pm 4^\circ$ equatorial buffer; ACT declination range provides only $\sim 16^\circ$ per hemisphere.
 - **Checks:** interval width recomputation
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; interval widths were not recomputed.
11. \triangle **C11** (Page 7, Section VII.B)
- **Claim:** Measure $\langle R \rangle = 0.93 \pm 0.07$; consistent with isotropy at the 1σ level.
 - **Checks:** difference-from-unity significance
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; z-score vs unity was not recomputed.
12. \triangle **C12** (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 ordering check
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; range sanity check was not performed.
13. \triangle **C13** (Page 7, Section VII.B)
- **Claim:** Mean temperature offsets of up to $88 \mu\text{K}$ between splits in certain declination bands.
 - **Checks:** unit/format sanity check
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; parse/positivity check was not performed.
14. \triangle **C14** (Page 7, Section VII.C and Table V row 'Cross-freq. correlation')
- **Claim:** Median cross-frequency correlation is $\rho_{\text{med}} = 0.69$; four cells show $\rho < 0.12$; Table V also lists 4 regions with $\rho < 0.12$.
 - **Checks:** repeated-constant consistency

- **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; consistency between narrative and Table V counts/threshold was not programmatically checked.
15. \triangle **C15** (Page 8, Section VIII and Table III)
- **Claim:** Five candidates are flagged as morphological outliers (z -score > 2); Table III lists five with z -scores **6.2, 4.2, 2.4, 2.9, 2.9**.
 - **Checks:** threshold count check from table
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; table parsing and threshold check was not performed.
16. \triangle **C16** (Page 8, Section IX and Table IV)
- **Claim:** Common mask has $f_{\text{sky}} \approx 0.46$; Table IV lists ℓ -ranges: 150 GHz family 400–1500; 90×90 400–1500; 90×150 500–1200; 220-related 500–1000.
 - **Checks:** range ordering and non-overlap sanity
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; ℓ -range ordering and endpoint consistency checks were not performed.
17. \triangle **C17** (Page 9, Section X.B (Cosmic Birefringence) and Eq. (5))
- **Claim:** Constraint $|\beta| < 0.01^\circ$ (1σ); EB gives $\beta_{\text{EB}} \approx 0.0006^\circ \pm 0.001^\circ$; TB gives $\beta_{\text{TB}} \approx 0.006^\circ \pm 0.01^\circ$.
 - **Checks:** inequality consistency (derived estimates within quoted bound)
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; bound/uncertainty comparisons were not re-computed.
18. \triangle **C18** (Page 9, Section X.C (Isocurvature Perturbation Limits))
- **Claim:** $\langle C^{\text{TB}} \rangle_{\ell} < 0.001, \langle \mu_{\text{K}} \rangle^2 \text{ and } \langle C^{\text{EB}} \rangle_{\ell} < 0.0001, \langle \mu_{\text{K}} \rangle^2 \text{ for } 10 \leq \ell \leq 100$; *both consistent with zero at $< 1\sigma$; Table V summarizes TB, EB $< 1\sigma$* .
 - **Checks:** order-of-magnitude and repeated-claim consistency
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; ratio check (0.001/0.0001) was not re-computed.
19. \triangle **C19** (Page 11, Table V and earlier sections)
- **Claim:** tSZ candidate recovery: **200** candidates, with $\nu_{\text{max}} = 51.2\sigma$ (also stated in abstract and Section III.C).

- **Checks:** repeated-constant consistency across sections
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; cross-section search for consistency was not performed.
20. \triangle **C20** (Page 11, Table V and Section V.B)
- **Claim:** Spectral diagnostics: raw maps 2/20 tSZ-consistent; srcfree maps 1/20 tSZ-consistent.
 - **Checks:** fraction-to-percentage recomputation and repeated-claim consistency
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; fraction recomputation (2/20, 1/20) was not performed.
21. \triangle **C21** (Page 1 (Abstract) and Page 11 (Table V))
- **Claim:** Compact source at $(291.2^\circ, -29.2^\circ)$: amplitude $758 \mu\text{K}$, spectral index $\alpha \approx -0.4$, Compton- y significance 41σ .
 - **Checks:** repeated-constant consistency across sections
 - **Verdict:** UNCERTAIN
 - **Notes:** Execution did not run; cross-location consistency check was not performed.
22. \times **C22** (Page 10, Section XI and Page 11, Section XII.B (look-elsewhere))
- **Claim:** Effective trials factor: 'eleven parallel analyses' (Section XI) vs 'twelve parallel analyses' (Section XII.B); also Bonferroni dividing by 11 vs multiplying by 12.
 - **Checks:** internal contradiction detection (repeated constants differ)
 - **Verdict:** FAIL
 - **Notes:** The provided inputs explicitly contain both 11 and 12 for the same concept (parallel analyses / trials factor), which is a direct inconsistency.

Limitations

- Audit uses only the provided PDF text; underlying ACT/NILC map data and any numerical arrays behind figures are unavailable, so many quantitative claims tied to maps/plots cannot be recomputed.
- No plot digitization is performed; figure-based quantities (e.g., candidate counts vs filter scale) are treated as unverified unless explicitly tabulated in text.
- Several checks (e.g., area fractions) depend on whether counts refer to the full CAR grid or the masked footprint; footprint pixel counts are not explicitly given, so such checks can only be conservative or must be marked unverified when ambiguity dominates.

- Execution error prevented running the planned numerical checks: Sandbox policy violation: from-import of 'typing' is not allowed.