

Skeptical review: Decisive Cosmological Evidence for the Normal Neutrino Mass Hierarchy from DESI Data Release 2

Summary

The manuscript performs a Bayesian model comparison between the Normal (NH) and Inverted (IH) neutrino mass hierarchies by combining (i) cosmological constraints on the summed neutrino mass Σm_ν , extracted from DESI DR2 (+ Planck CamSpec) chains and (ii) oscillation information from NuFIT 6.0. The DESI cosmological information is compressed to a 1D distribution in Σm_ν , approximated as a truncated Gaussian (Sec. 2.1, Sec. 3.1), and combined with an oscillation likelihood and two prior families on the neutrino masses: an exchangeable hierarchical log-normal prior (SJPV) and a Fisher-information/reference prior (HS) (Sec. 2.3). The resulting Bayes factors $K = P(D|\text{NH})/P(D|\text{IH})$ are reported to be very large in ΛCDM (and still strong in $w_0w_a\text{CDM}$), and the paper further propagates the inferred mass posteriors to predictions for $m_{\beta\beta}$ and $0\nu\beta\beta$ half-lives (Sec. 3.5–3.6), plus a historical reconstruction of K from earlier cosmological Σm_ν limits (Sec. 3.3, Fig. 4). The topic is timely and the overall Bayesian framing is appropriate, but several core steps are currently ambiguous or conceptually fragile: most importantly, the use of a marginalized DESI posterior as an effective likelihood (risking prior double counting), the sensitivity of K to tail modeling near the IH threshold, incomplete specification of priors and numerical evidence computation, and limited discussion of cosmological-model/systematic dependence given the strength of the paper’s “decisive” claims. Addressing these points is necessary for the conclusions to be both robust and reproducible.

Strengths

- Addresses a high-impact question (cosmological discrimination of NH vs IH) using newly available DESI DR2 constraints (Sec. 1, Sec. 3).
- Uses a clear Bayesian model-comparison target (evidence and Bayes factor) and attempts prior-robustness checks by considering two qualitatively different prior families (Sec. 2.2–2.3, Sec. 3.4).
- Provides phenomenology-facing outputs (individual mass posteriors; $m_{\beta\beta}$ implications for $0\nu\beta\beta$) that are useful to experimental audiences (Sec. 3.5–3.6).
- Includes some robustness exploration across dataset choices and an extension beyond ΛCDM ($w_0w_a\text{CDM}$), and provides a useful historical narrative of how constraints have evolved (Sec. 3.1–3.4, Sec. 3.7).
- Figures are generally readable and the logical flow from Σm_ν constraints to hierarchy preference is easy to follow at a high level.

Major issues

1. **Using a marginalized DESI Σm_ν posterior as a cosmological likelihood risks prior double counting and conceptual inconsistency (Sec. 2.1, Sec. 3.1).** The manuscript states it extracts the 1D marginalized posterior for Σm_ν from DESI DR2 MCMC chains and then uses it as $P(D_{\text{cosmo}}|\Sigma m_\nu)$ in the evidence integral. But a chain-derived marginalized posterior is proportional to $L(\Sigma m_\nu, \text{other}) \times \pi_{\text{DESI}}(\Sigma m_\nu, \text{other})$ marginalized over other parameters using DESI priors, not a likelihood in Σm_ν . Treating it as a likelihood can silently import DESI priors and volume effects into the second-stage evidence, undermining the interpretation of differences between SJPV vs HS and potentially biasing K .

Recommendation: In Sec. 2.1 (and where Eq. (1)–(2) are defined), explicitly distinguish likelihood vs posterior and state the DESI priors relevant to Σm_ν and the parameters being marginalized. Then do one of the following: (i) reconstruct an approximate likelihood in Σm_ν (e.g., from a profile likelihood if available, or by deconvolving a known Σm_ν prior and carefully discussing what remains after marginalization), or (ii) if you intentionally use a posterior-based surrogate, rewrite the evidence expression to avoid double counting and clearly state the approximation and its limitations. Provide a quantitative comparison of K obtained from (a) a likelihood-based surrogate (profile-likelihood or equivalent) versus (b) your posterior-derived surrogate to demonstrate stability.

2. **The DESI cosmological information is overly compressed to a 1D truncated Gaussian in Σm_ν without quantitative validation in the IH-critical tail and without accounting for degeneracies (Sec. 2.1, Sec. 3.1–3.2, Sec. 3.7).** The Bayes factor is dominated by the likelihood/posterior tail near $\Sigma m_\nu \approx 0.09\text{--}0.12$ eV (around the IH minimum). Small mismodeling of the tail or ignoring correlations with parameters that broaden Σm_ν (especially in $w_0 w_a$ CDM) can change $\ln K$ materially. The negative untruncated mean ($\mu_0 < 0$) further emphasizes that the effective distribution is boundary-dominated.

Recommendation: In Sec. 2.1 and Sec. 3.1, validate the 1D surrogate specifically in the range $0.06\text{--}0.12$ eV: provide goodness-of-fit diagnostics in the tail (not only bulk), and compare to at least one alternative surrogate (e.g., KDE/interpolated density used directly, skew-normal/mixture model). Report how $\ln K$ changes across these alternatives for both Λ CDM and $w_0 w_a$ CDM (Sec. 3.7). Where feasible, include at least a low-dimensional approximation that retains the main Σm_ν degeneracies (e.g., Σm_ν with Ω_m and w_0) or reweight the original chains under NH/IH priors to avoid the 1D collapse.

3. **Oscillation information (NuFIT 6.0) and the hierarchy preference $\Delta\chi^2$ are incorporated ambiguously and may be double-counted or mis-applied (Sec. 2.1–2.2, Sec. 3.2).** The manuscript mentions a NuFIT $\Delta\chi^2 \approx 6.1$ preference for NH

and applies it via a multiplicative factor $\exp(\Delta\chi^2/2)$ to produce K_{full} , but it is unclear whether $P(D_{\text{osc}}|\theta)$ is also integrated over oscillation parameters in the evidence integral. The sign convention and the mapping from NuFIT’s global-fit $\Delta\chi^2$ to a Bayes factor between discrete hierarchies are not defined.

Recommendation: In Sec. 2.1–2.2 and Sec. 3.2: (i) define $\Delta\chi^2$ explicitly (e.g., $\Delta\chi^2 \equiv \chi_{\text{IH}}^2 - \chi_{\text{NH}}^2$) and therefore the likelihood ratio mapping; (ii) state precisely what enters the evidence integral for each hierarchy (full $P(D_{\text{osc}}|\theta)$ surface vs approximations vs none); and (iii) if you keep the $K_{\text{base}}/K_{\text{full}}$ decomposition, make it mathematically explicit and prove/justify that the $\exp(\Delta\chi^2/2)$ factor does not double-count any oscillation information already included. Ideally, implement a single end-to-end evidence calculation that multiplies $P(D_{\text{cosmo}}|\Sigma m_\nu(\theta)) \times P(D_{\text{osc}}|\theta)$ and integrates over θ for NH and IH, and tabulate the resulting $\ln K$ alongside the decomposed values.

4. **Prior definitions (SJPV and HS) are insufficiently specified to reproduce results and to interpret prior dependence (Sec. 2.3, Sec. 3.4).** The SJPV hierarchical log-normal prior needs explicit generative form, hyperpriors, bounds, and the exact procedure for mapping exchangeable draws to ordered eigenmasses under NH/IH (label switching/order constraints). The HS “reference” prior is given in a closed form (e.g., proportional to a polynomial in masses) without a transparent derivation, definition of the base measure, parameterization choices, or normalization domain; evidence values can depend strongly on those choices despite “reference” wording.

Recommendation: Expand Sec. 2.3 (preferably with an Appendix) to fully define both priors: for SJPV, write the full hierarchical model (including hyperparameter priors and explicit bounds), clarify whether masses are treated as unordered then ordered, and show how NH vs IH mapping is done. For HS, provide a step-by-step derivation from the chosen observable set/Fisher information to the stated mass-space density, including the Jacobian, the measure with respect to which the density is defined, and the normalization domain (upper mass bounds). In Sec. 3.4, add sensitivity checks showing how $\ln K$ changes under reasonable variations of hyperprior bounds and HS normalization bounds.

5. **Evidence computation methodology and numerical uncertainty are not documented at the level required for a paper that claims very large Bayes factors (Sec. 2.2, Sec. 3.2–3.4).** The manuscript reports K values with high apparent precision (e.g., $K = 10231.4$), but does not specify the integral dimensionality, parameterization (three masses vs m_{lightest} +splittings vs hyperparameters), numerical method (grid/MC/nested sampling/importance sampling), convergence checks, or uncertainty on $\ln K$. It is also unclear whether the “likelihood” surrogate is treated as normalized or up to a constant, which matters for evidences.

Recommendation: In Sec. 2.2 and/or an Appendix, specify: (i) the exact integration variables for each hierarchy/prior; (ii) the numerical integration/sampling method; (iii) sample sizes / grid resolution; (iv) convergence diagnostics; and (v) an estimated numerical error bar on $\ln Z$ and $\ln K$ (report $\ln K \pm \sigma$). Reduce reported significant figures accordingly. Also state explicitly whether your $P(D_{\text{cosmo}}|\Sigma m_\nu)$ surrogate is normalized and how any normalization constants are handled consistently across hierarchies.

6. **Cosmological-model and dataset/systematic dependence is not explored or discussed in proportion to the strength of the “decisive” claim (Sec. 1, Sec. 3.1–3.3, Sec. 3.7, Conclusion).** Only Λ CDM and w_0w_a CDM are analyzed directly. However, Σm_ν constraints can weaken in standard extensions (e.g., non-flat Λ CDM, Λ CDM + N_{eff} , early dark energy, modified gravity), and K is highly sensitive to any broadening of the Σm_ν tail above the IH threshold. Additionally, the paper should more clearly separate “DESI+Planck (CamSpec) under a specific pipeline” from a general statement about cosmology.

Recommendation: In Sec. 3.7 and the Conclusion, add a structured robustness discussion: (i) explicitly list which cosmological extensions are expected to most impact Σm_ν and why; (ii) either run at least one or two additional representative extensions (e.g., Λ CDM + N_{eff} and non-flat Λ CDM) and report $\ln K$, or (if infeasible) use published constraints in those models to bound how much $\ln K$ could plausibly decrease; and (iii) provide a table of $\ln K$ across CMB likelihood choices (CamSpec vs alternatives) and dataset combinations (DESI-only, DESI+Planck, DESI+Planck+SN) using a consistent methodology. Soften global statements (“definitive”/“decisive”) to be explicitly conditional on the assumed cosmological model class and dataset combination.

7. **$0\nu\beta\beta$ implications are presented too definitively without propagating key nuclear/theory uncertainties and mechanism dependence (Sec. 3.6, Conclusion).** Translating $m_{\beta\beta}$ to half-life requires isotope choice, phase-space factors, and nuclear matrix elements (NMEs), which carry sizable spreads; moreover, the link assumes standard light Majorana neutrino exchange dominance.

Recommendation: In Sec. 3.6 and the Conclusion, specify the isotope(s), phase-space factors, and NME set(s) used, and propagate NME uncertainties into a band of predicted $T_{1/2}$ for a given $m_{\beta\beta}$ (e.g., show ranges across multiple standard NME calculations). State explicitly that results are conditional on the light-Majorana exchange mechanism and that alternative mechanisms can decouple $m_{\beta\beta}$ from the rate. Rephrase detection-probability statements to make these conditions explicit.

8. **Scholarly grounding and references need strengthening, and the framing risks overstating novelty/authority (Sec. 1, Sec. 4/References, Conclusion).** Key primary references for DESI DR2 cosmology outputs, Planck likelihoods, NuFIT

6.0, prior Bayesian hierarchy studies, and $0\nu\beta\beta$ experimental sensitivities should be clearly and reliably cited. Any non-standard, placeholder, or hard-to-verify references undermine credibility, especially for a “decisive evidence” claim.

Recommendation: Revise Sec. 4/References to ensure all inputs are traceable to primary, citable sources (DESI DR2 cosmology/likelihood papers, Planck likelihood documentation, NuFIT 6.0 global-fit papers/data release, and key prior work on cosmological hierarchy inference). Update the Introduction/Conclusion phrasing to position the contribution relative to that literature (what is new: DESI DR2-specific update, prior comparison, etc.), and remove/replace any references that are not verifiable or not appropriate for the venue.

Minor issues

1. Equation and notation clarity: Eq. (1) is rendered ambiguously (integral/evidence notation and the symbol Z), and mass-index notation alternates between (m_1, m_2, m_3) and (m_L, m_M, m_H) without an explicit mapping under NH vs IH (Sec. 2.2–2.4).

Recommendation: Rewrite Eq. (1) unambiguously as $Z_M \equiv P(D|M) = \int P(D|\theta, M)P(\theta|M) d\theta$, and use consistent symbols thereafter. Add a small mapping table (Sec. 2.3 or 2.4): under NH, $(m_L, m_M, m_H) = (m_1, m_2, m_3)$; under IH, $(m_L, m_M, m_H) = (m_3, m_1, m_2)$, and state which basis each prior is defined on.

2. Construction of the truncated-Gaussian fit is under-described (Sec. 2.1, Sec. 3.1): KDE/binning choice, fit criterion, and how the 95% upper limit 0.0642 eV is computed (from chains vs from the fit vs from DESI publication) are not clearly distinguished.

Recommendation: Add a concise, reproducible description: chain selection, burn-in/thinning, method to estimate the 1D density (KDE bandwidth or binning), fitting procedure for (μ_0, σ) (e.g., MLE for truncated normal), and whether the reported UL is computed directly from the chain or from the parametric fit. Explicitly write the truncated-normal formula and normalization constant.

3. Historical reconstruction of Bayes factors lacks a reproducible mapping from published Σm_ν upper limits to likelihood surrogates (Sec. 3.3, Fig. 4). It is unclear which experiments/datasets are used at each epoch, what cosmological model each limit assumes, and whether oscillation inputs are time-dependent or fixed to modern NuFIT.

Recommendation: Add a table (main text or Appendix) listing each epoch: reference, dataset/model, quoted Σm_ν constraint, the assumed likelihood surrogate (and how σ is inferred from a 95% UL), and whether oscillation information is contemporaneous or held fixed. This will make Fig. 4 auditable.

4. Interpretation of K relies heavily on Jeffreys-scale labels (“strong/decisive”) without sufficient caveats about prior/model dependence (Sec. 2.2, Sec. 3.2–3.4, Conclusion).

Recommendation: Keep reporting numerical $\ln K$ as the main result; add a short note that Jeffreys-scale categories are heuristic. Consider translating headline $\ln K$ values into posterior model probabilities for a few plausible prior odds to help readers interpret the magnitude without categorical language.

5. $m_{\beta\beta}$ Monte Carlo propagation is not described in enough detail to reproduce (Sec. 2.4, Sec. 3.6): sample counts, correlation handling among oscillation parameters, convergence checks, and whether IH $m_{\beta\beta}$ curves are conditional on IH or hierarchy-marginalized are unclear.

Recommendation: State the sampling procedure (N samples, how correlated NuFIT parameters are drawn), show a simple convergence test (stability of key quantiles), and clarify in captions/text whether plotted distributions are conditional on each hierarchy or weighted by the hierarchy posterior odds.

6. Figures and captions: several plots mix likelihood/posterior language and do not always specify normalization, scale ($\ln K$ vs $\log_{10} K$), or credible-interval definitions; some legends/labels are hard to parse (Sec. 3; Figs. 1, 3–6, 8–9).

Recommendation: Standardize figure labeling: explicitly state whether curves are posteriors or likelihood surrogates, include axis units and log base, and define the interval/UL construction used. Add a compact table in Sec. 3.2 summarizing $\ln K_{\text{base}}$ and $\ln K_{\text{full}}$ for each dataset/prior/model shown in figures.

Very minor issues

1. Terminology/typos and consistency: inconsistent formatting of model names (“ $w_0w_a\text{CDM}$ ” vs $w_0w_a\text{CDM}$), interval terminology (95% C.L./UL/credible), and occasional stray acronyms/spacing issues (Sec. 2–4).

Recommendation: Proofread for consistent notation and terminology throughout; standardize model names and interval language; fix minor typographical issues in equations, captions, and references.

2. Truncated Gaussian parameter interpretation: μ_0 and σ are described as “central value” and standard deviation even though $\mu_0 < 0$ and truncation changes the mean/median (Sec. 2.1, Sec. 3.1).

Recommendation: Clarify that (μ_0, σ) refer to the underlying untruncated normal parameters, and (optionally) report the truncated distribution mean/median for intuition.

3. Majorana phases in Eq. (3) are not explicit though the text says phases are sampled (Sec. 2.4, Eq. (3)).

Recommendation: State explicitly whether the PMNS elements U_{ei} in Eq. (3) include the Majorana phases, or write the phases as separate factors.

Key statements and references

- **✘** The constraints on the neutrino mass-squared differences (Δm_{21}^2 and $\Delta m_{3\ell}^2$) and the leptonic mixing angles used in this work are taken from the NuFIT 6.0 global analysis of neutrino oscillation data, which also provides a slight preference for the Normal Hierarchy corresponding to $\Delta\chi^2 = 6.1$ between NH and IH, and this $\Delta\chi^2$ contribution is explicitly included in the final Bayes factor calculation.
- *Reference(s)*: [1]
- *Justification*: In [1], the authors reference NuFIT-6.0 and NuFIT-6.1 but explicitly use JUNO and NuFIT-6.1 values for Δm_{21}^2 , Δm_{31}^2 , and mixing angles in their numerical examples. They assume normal ordering rather than incorporating any $\Delta\chi^2$ preference between orderings, and there is no Bayesian analysis or Bayes factor calculation described, nor any mention of including a $\Delta\chi^2 = 6.1$ preference. Thus, the specific claim about using NuFIT 6.0 constraints and explicitly including the $\Delta\chi^2$ preference in a final Bayes factor is not supported.
- **✘** The Bayesian model comparison between the Normal and Inverted Hierarchies is performed by computing the marginal likelihoods $P(D|M)$ using a framework in which the total likelihood factorizes as $P(D|\theta, M) = P(D_{\text{cosmo}}|\Sigma m_\nu(\theta)) \times P(D_{\text{osc}}|\theta)$, following standard cosmological-neutrino analyses, and the relative evidence is quantified by the Bayes factor $K = P(D|\text{NH})/P(D|\text{IH})$, interpreted on the Jeffreys scale where $K > 10$ and $K > 100$ correspond to “strong” and “decisive” evidence, respectively.
- *Reference(s)*: [3], [4], [8]
- *Justification*: Neither [3] nor [4] performs a Bayesian model comparison between Normal and Inverted Hierarchies via marginal likelihoods $P(D|M)$ or computes a Bayes factor. [4] estimates parameter bounds under different hierarchies using MCMC posteriors and profile likelihoods, imposing oscillation limits as priors rather than a separate oscillation likelihood factor, and does not evaluate evidences or use the Jeffreys scale. [3] provides cosmological constraints on Σm_ν and N_{eff} without hierarchy model selection or Bayes factors.
- **✘** To contextualize the impact of DESI DR2, the authors reconstruct the historical evolution of the Bayes factor from 2002 onward by re-computing K using a sequence of published cosmological upper limits on Σm_ν (e.g., 2dFGRS with $\Sigma m_\nu < 1.80$ eV and Planck 2015 with $\Sigma m_\nu < 0.18$ eV), demonstrating that the evidence for the Normal Hierarchy grows monotonically as these external cosmological constraints tighten, culminating in a decisive preference once the DESI DR2 limit $\Sigma m_\nu < 0.0642$ eV falls below the Inverted Hierarchy minimum mass (~ 0.099 eV).
- *Reference(s)*: [15]

- *Justification:* The attached paper [15] analyzes DESI DR2 via Bayesian model comparison and tension quantification across cosmological models (e.g., Λ CDM, w_0w_a CDM) and datasets. It does not reconstruct a historical Bayes factor for neutrino mass hierarchy, does not recompute K from published Σm_ν upper limits, and does not discuss Normal vs Inverted Hierarchy or a DESI DR2 Σm_ν limit of 0.0642 eV. Hence the claimed monotonic growth of evidence for the Normal Hierarchy culminating in a decisive preference is not presented in [15].

Mathematical consistency audit

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

Maths relevance: light

The paper’s analytic content centers on (i) defining Bayesian evidence and Bayes factors for NH vs IH, (ii) factorizing the likelihood into a cosmological component depending on Σm_ν and an oscillation component, (iii) specifying two priors over neutrino masses (SJPV hierarchical and HS reference), and (iv) defining the effective Majorana mass $m_{\beta\beta}$. The main internal-consistency risks are definitional: whether the extracted DESI 1D Σm_ν distribution is a likelihood or a posterior, ambiguity in the written evidence integral (Eq. (1)), and insufficient derivation detail for the HS prior and for the $\Delta\chi^2$ -to-likelihood-ratio conversion.

Checked items

1. **⚠ Evidence (marginal likelihood) definition** (Eq. (1), Sec. 2.2, p.3)
 - **Claim:** Defines the model evidence $P(D|M)$ as an integral of likelihood times prior over parameters θ .
 - **Checks:** notation consistency, definition correctness
 - **Verdict:** UNCERTAIN; confidence: medium; impact: critical
 - **Assumptions/inputs:** D denotes combined cosmological and oscillation data, θ denotes neutrino mass parameters under a fixed hierarchy M
 - **Notes:** As rendered, Eq. (1) appears to miss an integral sign and places 'Z' as a multiplicative factor rather than an equality to the integral. If the intended equation is ' $Z \equiv P(D|M) = \int \dots d\theta$ ', it is fine; but the text alone is ambiguous and blocks strict verification.
2. **✓ Likelihood factorization** (Sec. 2.2, p.3 (unnumbered equation after Eq. (1)))
 - **Claim:** Uses $P(D|\theta, M) = P(D_{\text{cosmo}}|\Sigma m_\nu(\theta)) \times P(D_{\text{osc}}|\theta)$.
 - **Checks:** symbol consistency, assumption clarity
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** Conditional independence of cosmological and oscillation datasets given θ (and M), Σm_ν is a deterministic function of θ

- **Notes:** Factorization is internally consistent with the narrative: cosmology depends only on Σm_ν , while oscillations depend on mass splittings and mixing parameters in θ .
3. ✓ **Bayes factor definition** (Eq. (2), Sec. 2.2, p.4)
- **Claim:** Defines $K = P(D|\text{NH})/P(D|\text{IH})$.
 - **Checks:** definition correctness, notation consistency
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** Same dataset D used for both evidences, NH and IH treated as competing models
 - **Notes:** Standard and consistent with later text ($K_{\text{base}}, K_{\text{full}}$).
4. ✗ **Use of DESI chain-derived Σm_ν distribution as likelihood** (Sec. 2.1, p.3; Sec. 3.1, p.5)
- **Claim:** Extracts the 1D marginalized posterior for Σm_ν from DESI DR2 MCMC chains and models it as a truncated Gaussian likelihood $P(D_{\text{cosmo}}|\Sigma m_\nu)$.
 - **Checks:** definition/measure consistency, Bayesian coherence
 - **Verdict:** FAIL; confidence: medium; impact: critical
 - **Assumptions/inputs:** The chain-derived 1D distribution is proportional to the likelihood in Σm_ν (or can be treated as such), Any DESI prior on Σm_ν is negligible/flat or appropriately accounted for
 - **Notes:** Within the paper text, no step is provided to justify replacing a marginalized posterior with a likelihood, nor is the DESI prior on Σm_ν specified and removed. This is an internal mathematical inconsistency relative to Eq. (1), where $P(D_{\text{cosmo}}|\Sigma m_\nu)$ must be a likelihood factor rather than a posterior density.
5. ✓ **Truncated Gaussian on $\Sigma m_\nu \geq 0$** (Sec. 2.1, p.3; Sec. 3.1, p.5)
- **Claim:** Approximates the Σm_ν distribution by a Gaussian with parameters (μ_0, σ) truncated to $\Sigma m_\nu \geq 0$.
 - **Checks:** domain/constraint consistency, notation clarity
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** Σm_ν is physically nonnegative, Truncation is implemented as conditional normalization on $[0, \infty)$
 - **Notes:** Truncating at $\Sigma m_\nu \geq 0$ is mathematically consistent even when $\mu_0 < 0$; however, calling μ_0 a 'central value' of the truncated distribution is imprecise (it is the underlying normal's location parameter).
6. △ **HS prior functional form** (Sec. 2.3, p.4)

- **Claim:** States the HS reference prior on (m_L, m_M, m_H) is proportional to $m_L m_M + m_L m_H + m_M m_H$ as a Jacobian-derived reference prior.
- **Checks:** derivation completeness, symbol/definition consistency
- **Verdict:** UNCERTAIN; confidence: low; impact: moderate
- **Assumptions/inputs:** Reference prior is derived from Fisher information of $(\Delta m_{21}^2, \Delta m_{3\ell}^2)$, A specific parameter transformation is used, with ordering constraints
- **Notes:** The claimed polynomial form is not derivable from the provided text alone: the mapping and Jacobian (and any additional parameter needed beyond two Δm^2 's) are not shown. Without these, the prior's correctness cannot be audited.

7. ✓ SJPV prior description vs parameterization (Sec. 2.3, p.4)

- **Claim:** Assumes exchangeable (m_1, m_2, m_3) drawn from a common log-normal distribution, inducing a 'geometric volume penalty' against IH.
- **Checks:** conceptual consistency, notation consistency
- **Verdict:** PASS; confidence: medium; impact: minor
- **Assumptions/inputs:** Exchangeability in the prior over labeled eigenstates, Hierarchy imposed as a constraint defining NH/IH subspaces
- **Notes:** While no equations are given, the qualitative claim is internally consistent: imposing ordering/near-degeneracy constraints typically reduces allowed prior volume under an exchangeable prior.

8. ✓ $m_{\beta\beta}$ definition (Eq. (3), Sec. 2.4, p.5)

- **Claim:** Defines $m_{\beta\beta} = \left| \sum_{i=1}^3 U_{ei}^2 m_i \right|$.
- **Checks:** definition correctness, notation consistency
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:** U_{ei} are (possibly complex) PMNS elements in a convention compatible with $0\nu\beta\beta$, m_i are neutrino mass eigenvalues
- **Notes:** Expression is internally consistent with later description of sampling CP phases; however, the equation does not explicitly display Majorana phases, so the reader must infer they are included in U or applied separately.

9. △ Inclusion of oscillation $\Delta\chi^2$ as multiplicative factor (Sec. 3.2, p.6)

- **Claim:** Computes $K_{\text{full}} = K_{\text{base}} \times \exp(\Delta\chi^2/2)$ using $\Delta\chi^2 = 6.1$ from NuFIT.
- **Checks:** algebraic transformation, assumption clarity
- **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
- **Assumptions/inputs:** $\Delta\chi^2$ corresponds to $-2\log$ likelihood ratio between IH and NH (with a defined sign), Oscillation preference can be separated multiplicatively from the cosmology-driven K_{base}

- **Notes:** Multiplying by $\exp(\Delta\chi^2/2)$ is only correct for a specific $\Delta\chi^2$ sign convention (typically $\Delta\chi^2 = \chi_{\text{IH}}^2 - \chi_{\text{NH}}^2$). The paper does not define the convention, so the factor could be inverted.

10. ✓ **Parameter dependence of cosmological likelihood** (Sec. 2.2–2.4, p.3–5)

- **Claim:** Uses $\Sigma m_\nu(\theta)$ to connect neutrino mass parameters θ to the cosmological likelihood term.
- **Checks:** mapping consistency
- **Verdict:** PASS; confidence: medium; impact: minor
- **Assumptions/inputs:** Σm_ν is computed as a sum of mass eigenstates consistent with the hierarchy constraints, θ includes enough parameters to define the spectrum
- **Notes:** Although $\Sigma m_\nu(\theta)$ is not explicitly written, the mapping is conceptually straightforward and consistent with the rest of the paper’s notation.

Limitations

- Only three explicit equations are provided; many central computations (evidence integrals, prior normalizations, Jacobians, mapping from oscillation parameters to masses) are described qualitatively without derivations, limiting the scope of symbolic verification.
- The audit cannot verify any claims requiring omitted intermediate steps (e.g., the HS prior derivation from Fisher information, or the exact construction of K_{base} from the truncated Gaussian and priors).
- The paper does not specify the prior used in the DESI MCMC analysis for Σm_ν , preventing a purely internal check that the extracted 1D distribution is a likelihood rather than a posterior.

Numerical results audit

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

Out of 12 candidate checks, 11 PASS and 1 is UNCERTAIN. Recomputed Bayes-factor scaling with $\exp(\Delta\chi^2/2)$ matches reported K_{full} values closely for both SJPV and HS priors. Truncated-normal 95% upper limits computed from (μ_0, σ) parameters agree with the reported baseline UL and show the Feldman–Cousins parameterization yields a very similar UL under the same truncation convention. Mass-sum minima for NH and IH are consistent with the stated approximate component masses. The inequality $\text{UL}(\Sigma m_\nu) < \text{IH}$ minimum holds with a **0.0348** eV margin. The only unresolved item is a qualitative claim about being "almost entirely below" a 5 meV sensitivity threshold, which cannot be established from endpoints alone.

Checked items

1. ✓ **C1_Kfull_from_Kbase_SJPV** (Page 6, Section 3.2 (Bayes factors paragraph))
 - **Claim:** Under the SJPV prior, $K_{\text{base}} = 484.5$. Including $\Delta\chi^2 = 6.1$ multiplicatively ($K_{\text{full}} = K_{\text{base}} \times \exp(6.1/2)$) gives $K_{\text{full}} = 10231.4$.
 - **Checks:** recompute_formula
 - **Verdict:** PASS
 - **Notes:** Computed $K_{\text{full,calc}} = 10230.3844$ vs reported 10231.4; relative agreement is very tight, with a rounding-level absolute difference.
2. ✓ **C2_Kfull_from_Kbase_HS** (Page 6, Section 3.2 (HS prior paragraph))
 - **Claim:** Under the HS prior, $K_{\text{base}} = 22.0$. Including the oscillation preference yields $K_{\text{full}} = 464.6$.
 - **Checks:** recompute_formula
 - **Verdict:** PASS
 - **Notes:** Computed $K_{\text{full,calc}} = 464.5376$ vs reported 464.6; consistent with one-decimal rounding.
3. ✓ **C3_truncated_normal_UL_from_mu_sigma** (Page 5, Section 3.1 and Figure 1 caption)
 - **Claim:** Baseline Λ CDM posterior for Σm_ν modeled as truncated Gaussian with $\mu_0 = -0.009$ eV and $\sigma = 0.036$ eV, yielding 95% upper limit $\Sigma m_\nu < 0.0642$ eV.
 - **Checks:** recompute_quantile_truncated_distribution
 - **Verdict:** PASS
 - **Notes:** Computed truncated-normal $\text{UL}_{95} = 0.0648869$ eV vs reported 0.0642 eV; within tolerance given possible convention details.
4. ✓ **C4_truncated_normal_UL_from_FC_mu_sigma** (Page 5, Section 3.1 and Figure 1 caption)
 - **Claim:** Feldman-Cousins profile likelihood parameters $\mu_0 = -0.036$ eV and $\sigma = 0.043$ eV are shown; check implied 95% upper limit under truncation at $\Sigma m_\nu \geq 0$ for consistency with being 'close agreement' to baseline.
 - **Checks:** recompute_quantile_truncated_distribution
 - **Verdict:** PASS
 - **Notes:** Computed truncated-normal UL_{95} (FC params) = 0.0639334 eV vs baseline 0.0642 eV; supports the stated qualitative 'close agreement'.
5. ✓ **C5_NH_mass_sum_vs_reported_min** (Page 10, Section 3.5 and Figure 6 caption)

- **Claim:** For NH, $m_2 \approx 0.008$ eV and $m_3 \approx 0.050$ eV, with m_1 nearly massless (< 0.01 eV), forcing Σm_ν near theoretical minimum ~ 0.059 eV.
 - **Checks:** sum_components_vs_total
 - **Verdict:** PASS
 - **Notes:** $m_2 + m_3 = 0.058$ eV vs reported ~ 0.059 eV; consistent with rounding/approximation.
6. ✓ **C6_IH_mass_sum_vs_reported_min** (Page 12, Figure 7 caption and Page 13, Figure 8 caption)
- **Claim:** IH requires $m_1 \approx m_2 \approx 0.05$ eV with $m_3 \approx 0$, implying Σm_ν minimum ~ 0.099 eV.
 - **Checks:** sum_components_vs_total
 - **Verdict:** PASS
 - **Notes:** $0.05 + 0.05 + 0 = 0.10$ eV vs reported ~ 0.099 eV; consistent given rounded inputs.
7. ✓ **C7_UL_below_IH_min_tension_margin** (Page 1 abstract; Page 5 Section 3.1; Page 13 Figure 8 caption)
- **Claim:** DESI DR2 gives 95% UL $\Sigma m_\nu < 0.0642$ eV while IH minimum is ~ 0.099 eV (or ≥ 0.099 eV), implying UL is below IH minimum.
 - **Checks:** inequality_consistency
 - **Verdict:** PASS
 - **Notes:** Computed gap $IH_{\min} - UL = 0.0348$ eV (> 0); ratio $UL/IH_{\min} = 0.6485$.
8. ✓ **C8_w0wa_UL_relaxation_factor** (Page 13, Section 3.7)
- **Claim:** In w_0w_a CDM, 95% UL on Σm_ν relaxes to **0.163** eV from baseline **0.0642** eV.
 - **Checks:** ratio_change
 - **Verdict:** PASS
 - **Notes:** Computed relaxation factor = **2.53894**; absolute increase = **0.0988** eV (153.9% increase).
9. ✓ **C9_HS_Kfull_drop_factor_w0wa_vs_LCDM** (Page 13, Section 3.7)
- **Claim:** HS prior K_{full} drops from **464.6** (Λ CDM baseline) to **42.6** (w_0w_a CDM).
 - **Checks:** ratio_change
 - **Verdict:** PASS
 - **Notes:** Computed ratio = **0.09169**; percent change = **-90.83%** (drop of **422.0** in absolute terms).

10. ✓ **C10_mbb_NH_CI_contains_median** (Page 11, Section 3.6; Figure 9 caption)
- **Claim:** NH $m_{\beta\beta}$ 95% credible interval is [0.95, 11.55] meV with median 3.28 meV.
 - **Checks:** interval_contains_point
 - **Verdict:** PASS
 - **Notes:** Median lies within the stated 95% CI.
11. ✓ **C11_mbb_IH_CI_contains_median** (Page 12, Section 3.6; Figure 9 caption)
- **Claim:** IH median $m_{\beta\beta}$ is 37.03 meV with 95% CI [18.36, 49.51] meV.
 - **Checks:** interval_contains_point
 - **Verdict:** PASS
 - **Notes:** Median lies within the stated 95% CI.
12. △ **C12_compare_experimental_sensitivities_vs_CI_NH** (Page 11-12, Section 3.6)
- **Claim:** NH 95% CI [0.95, 11.55] meV is 'far below' ~ 20 meV and 'almost entirely below' ~ 5 meV.
 - **Checks:** threshold_comparison
 - **Verdict:** UNCERTAIN
 - **Notes:** Endpoints confirm $CI_{\text{high}}(11.55) < 20$ meV, but $CI_{\text{high}}(11.55)$ is not < 5 meV. The phrase 'almost entirely below 5 meV' cannot be verified without distributional information; an endpoint-based crude fraction above 5 meV using only endpoints is ~ 0.6179 .

Limitations

- Checks are limited to explicit numbers present in the provided PDF text; no external datasets, MCMC chains, or plot-value extraction are used.
- Some claims use approximate symbols (\approx , \sim) and qualitative phrasing (e.g., 'close agreement', 'almost entirely'); only coarse numerical sanity checks can be performed without full distributions.
- Upper limits and credible intervals may depend on one-sided vs two-sided conventions and on truncation; the PDF does not fully specify these conventions, so tolerances are provided.