Paper Feedback Report (v2)

Prepared by ProofInk.AI (ProofInk.AI can make mistakes. Always double-check responses)

Generated: 8/2/2026, 5:01:20 PM

Paper: Information geometry for approximate Bayesian computation

Feedback Mode: Deep Analysis

Review Version: v2

Processing Time: 403.01 seconds

3/5

Overall Score

Adjusted Confidence: 74%

> **Likely outcome (approximate):** Likely major to likely minor revisions — an uncertain interpretation of the score, not a guarantee.

> *This feedback report is intended to help strengthen your submission before it reaches human reviewers. It highlights potential concerns and opportunities for improvement — not to discourage, but to give you the best chance of success. Not all feedback points carry equal weight. Use your expertise to prioritize.*

About This Paper

This paper develops a theoretical, information-geometric framework for analyzing Approximate Bayesian Computation (ABC), using relative entropy between the true and ABC-simulated posterior distributions to characterize the effect of the threshold parameter ε (allowed to be a vector), the distance metric, and the size of the data n. The authors derive leading-order asymptotic expansions of the relative entropy and of estimator bias as $|\varepsilon|\to 0$, specialize these results to the exponential family, and compare mean rejection rates for ball versus ellipse acceptance regions. Two worked examples (exponential and normal-gamma models) and an accompanying simulation study illustrate that allowing direction-dependent thresholds (an ellipsoidal acceptance region) can reduce rejection rates relative to the standard ball-shaped tolerance region for the same relative-entropy tolerance.

> **Highlights:** The paper introduces a genuinely novel information-geometric framework for analyzing ABC threshold selection, using relative entropy to rigorously characterize approximation error as a function of ε, distance metric, and sample size. The extension to vector-valued thresholds enabling a principled ellipsoid-versus-ball comparison (Sections 5-7) is a creative and useful contribution, and the exponential-family specialization adds welcome analytical tractability to a field dominated by simulation-based heuristics. The writing is clear despite technical density, and the worked examples plus simulation study offer readers concrete intuition for otherwise abstract asymptotic results.

Executive Summary

This is an ambitious and largely novel theoretical contribution that brings information-geometric tools to ABC threshold analysis, with core derivations (Lemma A.1, Proposition 2.2, Lemma 5.1, Lemma 3.4) that are internally sound and numerically verifiable where checkable. However, the paper would benefit from more careful alignment between the strength of claims made in the abstract and conclusion and what the asymptotic ($\varepsilon \to 0$) and single-setting empirical results actually demonstrate, most notably the abstract's claim of bounds 'for positive tolerances and finite sample sizes' and the conclusion's leap from rejection-rate to convergence-rate comparisons. An explicit limitations discussion — covering the sufficient-statistics assumption, tractable-likelihood examples, and the '$n|\varepsilon|$ small' validity regime — along with tighter justification of a few proof steps (Proposition 2.2, Condition 3.3) and clarification of the RB/RE ratio definition, would substantially strengthen the paper's rigor and credibility.

At a Glance

3/5 — The paper has a clear motivation, precise formal aims, and some strong theoretical structure, but the overall score is held to the middle by limited practical scope, a heavy reliance on a strong sufficiency assumption, and empirical support that does not yet fully match the breadth of the claims.

Criterion	Score	Bar
Hypothesis	4/5	
Problem Statement	4/5	
Methodology	3/5	

Results	3/5	
Novelty	4/5	
Contributions	3/5	
Limitations	2/5	
References	4/5	
Writing Quality	4/5	
Discussion	3/5	

Key Strengths

- The paper's aims are explicit, well structured, and easy to evaluate, which gives the work a strong sense of direction.
- The theoretical development is careful and internally consistent, with useful self-checks and rigorous symmetry arguments supporting the expansions.
- The examples reveal a nontrivial insight: certain leading-order bias terms can be largely independent of sample size, which is a valuable takeaway for ABC practice.

Top Priorities for Improvement

1. Broaden the practical framing of the main results by discussing more explicitly how the sufficiency assumption constrains real ABC workflows and where the approach can still be useful without it. This would make the contribution feel more applicable beyond the idealized setup.
2. Strengthen the empirical validation of the ellipsoidal-threshold claim by testing additional models, parameter regimes, or observables rather than relying mainly on one normal example. More varied evidence would better support the generality of the reported acceptance-rate advantage.
3. Interpret the asymptotic comparisons more carefully relative to the simulated settings, especially where the product $n|\varepsilon|$ is not especially small. A brief discussion of where the approximation is expected to be accurate would improve alignment between theory and experiments.
4. Replace or supplement the purely illustrative examples with a case closer to a genuine ABC use case. This would better demonstrate practical relevance and make the numerical section more convincing.
5. Tighten the discussion of limitations and scope in the introduction and conclusion so the paper's claims are easier to calibrate. Clearer boundaries would help readers understand exactly when the method should be expected to perform well.

Revision Progress Dashboard

Comparing with v1 (quick review, 2026-08-02)

Limited progress -- 11 of 15 issues remain unresolved.

Overall Score: 4 -> 3 (-1)

Per-Criterion Changes:

- Hypothesis: ? -> 4
- Problem Statement: ? -> 4
- Methodology: ? -> 3
- Results: ? -> 3
- Novelty: ? -> 4
- Contributions: ? -> 3
- Limitations: ? -> 2
- References: ? -> 4
- Writing Quality: ? -> 4
- Discussion: ? -> 3

Issue Summary: 0 resolved, 4 partial, 11 unresolved, 2 new

New Issues:

- [NEW] CORRECTED: The paper's applicability is narrowed by its reliance on sufficient statistics, and this limitation would benefit from fuller discussion. The introduction acknowledges that 'One would like the statistics to be sufficient statistics, but that is usually hard to come up with in practice' (Page 2), while the formal setup then assumes 'we are working with sufficient statistics' (Page 5). The assumption is explicit, but its practical consequences could be discussed more directly.
- [NEW] CONFIRMED: The generality of the practical claim for ellipsoidal thresholds is broader than the empirical support shown here. The introduction states that 'by exploiting the asymmetric effect of each direction, one is able to enlarge the acceptance region of ABC' and 'the mean rejection rate is typically smaller in the case of ε ∈ Rq+' (Page 3), but the simulation study is a single normal example with one true parameter setting, '(μ, σ2) = (0, 1)' (Page 24). The claim would be stronger with additional models or parameter regimes.

Score Timeline:

- v1: 4/5 (2026-08-02)
- v2: 3/5 (2026-08-02)

Priority Checklist for Next Draft:

1. The new review still characterizes the results as asymptotic and only accurate in the limit, so it does

not supply non-asymptotic finite-sample error bounds. The abstract/body mismatch therefore remains.

2. The examples are still described as analytically convenient but not representative because ABC is unnecessary when the likelihood is tractable. This does not address the limited practical scope.
3. The new review notes that the quantitative match to the asymptotic formula should be interpreted cautiously, but it does not provide a quantitative correction or calibration of the overestimation. The original concern remains open.
4. The new review still indicates that the paper does not actually provide explicit finite-sample, non-asymptotic bounds. The abstract claim remains insufficiently supported by the main results.
5. The examples are still acknowledged to be cases where ABC is not needed, so they remain pedagogical rather than compelling demonstrations of genuine ABC utility. The limitation persists.
6. The new review does not offer a quantitative error characterization for the rejection-rate approximation, only a cautionary note about interpreting the asymptotics. The approximation gap remains unquantified.
7. The paper still lacks an example with an intractable likelihood, and the new review confirms the examples are chosen for analytic convenience. So the gap in practical ABC relevance remains.
8. No quantitative characterization of the approximation error is added, only a warning that the asymptotic formulas should be interpreted cautiously. Confidence for non-infinitesimal tolerances is still not strengthened in a measurable way.

— Detailed analysis follows below —

Strengths

- CONFIRMED: The paper states a clear and testable set of aims up front: 'The contribution of the paper is fourfold' (Pages 2-3), followed by specific tasks including computing the leading-order relative entropy term, allowing vector thresholds, analyzing bias, and studying rejection rates. This makes it straightforward to assess whether the paper delivers on its promises.
- CONFIRMED: The practical motivation is well framed. The introduction identifies concrete ABC pain points—'known issues on sensitivity in performance with respect not only to the distance metric ... but also on the choice of the threshold parameter ε' and the fact that ABC 'typically scales badly with the size of the data' (Page 2)—and ties them to a clearly stated theoretical gap: 'theoretical developments are still in their infancy' (Page 2).

- CONFIRMED: The mathematical problem statement is precise. The paper formulates threshold choice as 'maximize ε-dependent acceptance region subject to $H(P|P\varepsilon)(\tau*) \le$ tol' (1.1) and then derives the tractable asymptotic surrogate (1.2) on Page 4.

- ADDED: The theory includes meaningful internal self-checks. Remark 2.3 states that the ellipse formula 'immediately reduces to the formula for the ball, when $\varepsilon1 = \cdots = \varepsilon q = \varepsilon$' (Page 9), and Corollary 3.5 adds 'if also q = 1 then we obtain the statement of Lemma 3.1' (Page 12). These reductions strengthen confidence in the formal development.

- ADDED: Lemma A.1 provides a careful symmetry argument supporting the leading-order expansion. The appendix shows that '$\nabla\varepsilon f\varepsilon(\theta|\tau*)|\varepsilon=0 = 0$' and that mixed second derivatives vanish, 'for all i, j ... with i̸= j, $\partial 2\varepsilon i\varepsilon jf\varepsilon(\theta|\tau*)|\varepsilon=0 = 0$' (Pages 27-29). This is a rigorous and elegant justification for why odd-order and mixed terms drop out.

- ADDED: The examples yield a genuinely useful and somewhat surprising insight: the leading-order bias can be independent of n for selected observables. In the exponential example, 'the bias does not depend on the size of the data set n' (Page 19), and in the normal example Corollary 6.4 says that, when the prior mean and observed mean are small, 'C1, up to leading order, does not depend on n' (Page 22).

Weaknesses

- CORRECTED: The paper's applicability is narrowed by its reliance on sufficient statistics, and this limitation would benefit from fuller discussion. The introduction acknowledges that 'One would like the statistics to be sufficient statistics, but that is usually hard to come up with in practice' (Page 2), while the formal setup then assumes 'we are working with sufficient statistics' (Page 5). The assumption is explicit, but its practical consequences could be discussed more directly.

- CONFIRMED: The generality of the practical claim for ellipsoidal thresholds is broader than the empirical support shown here. The introduction states that 'by exploiting the asymmetric effect of each direction, one is able to enlarge the acceptance region of ABC' and 'the mean rejection rate is typically smaller in the case of $\varepsilon \in Rq+$' (Page 3), but the simulation study is a single normal example with one true parameter setting, '$(\mu, \sigma 2) = (0, 1)$' (Page 24). The claim would be stronger with additional models or parameter regimes.

- CORRECTED: The asymptotic regime used by the theory deserves more caution in interpreting the numerical comparisons. Section 3 states that 'the expansion of relative entropy in ε is typically correct if the product $n|\varepsilon|$ is small' (Page 9), and Section 7 reiterates that the approximation is 'accurate only in the limit as $n|\varepsilon| \downarrow 0$' (Page 23). Some reported settings involve moderate products—for example, Table 1 gives n = 1000 and ε = 0.022 for (tol, n) = (1, 1000) (Page 24)—so the simulations are still informative, but the quantitative match to asymptotic formulas should be interpreted cautiously.

- CONFIRMED: The illustrative examples are analytically convenient but not representative of the main use-cases where ABC is needed. Section 6 explicitly notes that 'in these cases, we have explicit information for the likelihood and thus ABC methods are not necessary for inference' (Page 17). This is reasonable for exposition, but it limits the paper's practical demonstration.

Gaps and Inconsistencies

- CONFIRMED: A self-contained derivation of the central bias expansion is not found in the retrieved text. On Page 14, the paper states, 'Algebraic computations similar to those in the case of a ball in [3] ... shows that bias(...) = Σ εi2 Ci(τ∗) + O(|ε|3),' but an explicit proposition or proof is not provided in the retrieved pages. Since Sections 6.1 and 6.2 use this formula, a fuller derivation would improve traceability.
- ADDED: Lemma 3.1's proof would benefit from an explicit verification of the displayed fourth-order algebra. On Page 11 the proof writes terms including '2/(5!3)(nε)4 Eη4(θ)' and later '+ 2/5!(nε)4 Eη4(θ)' before concluding with '(nε)4 1/72 Var(η2(θ)) + O(|nε|6)'. As written, the η4-coefficient cancellation/simplification is not transparent from the displayed steps and merits checking or a short intermediate line.
- ADDED: The rejection-rate comparison in Lemma 5.1 is not fully pinned down as an apples-to-apples comparison between ball and ellipse. Page 15 defines 'pB = P(τ ∈ Bε(τ∗))' for a scalar ε and 'pE = P(τ ∈ Dε(τ∗))' for vector ε ∈ Rq+, then gives formula (5.1) mixing 'ε' in the numerator with 'εi' in the denominator. Section 7 later compares thresholds chosen under 'H(P|Pε)(τ∗) ≤ tol' and, for the ellipse only, solves 'maximize (πε1ε2) subject to leading order term in (6.7) ≤ tol' (Page 23). Stating explicitly how the ball radius ε is matched to the ellipse components εi would clarify the interpretation of U(τ∗).
- CORRECTED: There are small cross-table discrepancies in the reported observed statistics that complicate exact row-wise comparison, even if they may be transcription issues. For example, for (tol, n) = (1, 100), Table 1 reports 'τ∗ = (−0.181, 0.866)' (Page 24) while Table 2 reports '(−0.081, 0.866)' (Page 25); for (0.05, 1000), Table 1 reports '(−0.027, 1.013)' (Page 24) while Table 2 reports '(−0.027, 1.033)' (Page 25).
- ADDED: Table 1 contains an internal tolerance/threshold inconsistency that likely reflects a typo and affects interpretation of the rejection-rate trend. On Page 24, the row '(0.5, 600)' reports 'ε = 0.027' with 'R̂B = 5805', whereas the looser row '(1, 600)' reports the much smaller 'ε = 0.003' with 'R̂B = 4172'. Since larger tolerance would ordinarily not correspond to a dramatically smaller threshold, this row should be checked.
- CORRECTED: The introductory statement about lower rejection rates would benefit from qualification to match the later theory. The introduction says, 'We find that the mean rejection rate is typically smaller in the case of ε ∈ Rq+' (Page 3), whereas Section 5 gives a conditional statement: 'if

> $\varepsilon_i/\varepsilon \gg 1$ ... one can have benefits using the ellipse' (Page 16). Aligning the introduction more closely with the conditional form of the result would improve precision.

Criterion-by-Criterion Analysis

Hypothesis - Score: 4/5 (confidence: 0.55)

Strengths:

- The paper's four-fold research goals are stated with precision in the introduction, e.g., 'The contribution of the paper is fourfold... Firstly, in Section 2 we compute the leading order term of the relative entropy...' (Page 2-3).
- A specific, checkable claim is advanced: 'We find that the mean rejection rate is typically smaller in the case of $\varepsilon \in R_q^+$' (Page 3), which is later tested numerically.

Weaknesses:

- As a mathematical/methodological paper, there is no classical falsifiable empirical hypothesis with clearly defined independent/dependent variables and a predicted directional relationship stated up front
- the 'hypothesis' is really a set of derivations to be carried out.
- The claim that ellipsoidal thresholds are advantageous is stated somewhat generally (Page 3: 'by exploiting the asymmetric effect of each direction, one is able to enlarge the acceptance region of ABC') but is only demonstrated in one bivariate example, so the generality of the implicit hypothesis is not fully established.

Suggestions:

- State explicit, numbered conjectures/claims (e.g., 'Claim 1: for fixed tolerance, ellipsoidal thresholds yield lower rejection rate than balls whenever weight variances differ across directions') that can be directly matched to the propositions and simulation results.
- Discuss the conditions under which the reverse could occur (i.e., when ball and ellipse perform similarly or when ellipse could underperform) to sharpen the testable content of the central claim.

Problem Statement - Score: 4/5 (confidence: 0.75)

Strengths:

- The practical challenges of ABC (choice of distance metric, threshold ε, and scaling with data size n) are clearly laid out with citations, e.g., 'known issues on sensitivity in performance with respect not only to the distance metric... but also on the choice of the threshold parameter ε' (Page 1-2).
- The gap in the literature is explicitly identified: 'Despite significant developments in methodological approaches, theoretical developments are still in their infancy' (Page 2), with specific prior work [3,6] contrasted against the paper's different focus.
- The introduction gives a precise mathematical formulation of the underlying optimization problem (1.1)-(1.2), sharpening what the 'problem' actually is (Page 4).

Weaknesses:

- The paper motivates the sufficient-statistics assumption as standard practice but simultaneously acknowledges its impracticality ('One would like the statistics to be sufficient statistics, but that is usually hard to come up with in practice', Page 2), creating some tension with the problem framing that is not fully resolved.

Suggestions:

- Add a paragraph early on clarifying how the theoretical results might extend or degrade when only approximately sufficient statistics are used in practice, since this is the realistic ABC setting.

Methodology - Score: 3/5 (confidence: 0.65)

Strengths:

- Assumptions are stated explicitly as formal conditions (Condition 2.1, 3.2, 3.3), which aids reproducibility of the theoretical derivations (Pages 7, 12).
- The proof techniques (Taylor expansion of relative entropy, Laplace asymptotics for exponential family expectations) are standard, transparent, and traceable step-by-step (Pages 8, 13-14).
- The simulation setup describes concrete hyperparameters, sample sizes, and optimization tools (R's nloptr, Augmented Lagrangian algorithm) enabling replication (Page 23-24).

Weaknesses:

- The core theory requires 'n|ε| small' or 'nε small' for validity (e.g., Page 9: 'the expansion of relative entropy in ε is typically correct if the product n|ε| is small'), yet several simulated settings (e.g., Table 1, tol=1, n=100, ε≈0.096, giving nε≈9.6) appear to fall well outside a 'small' regime, raising questions about whether the reported computational-gain comparisons are within the theory's valid range.

- Both worked examples (exponential rate, normal mean/variance) explicitly use models whose likelihoods are fully known and for which 'ABC methods are not necessary for inference' (Page 17), which limits the demonstrated applicability of the theory to genuine likelihood-free settings that motivate ABC in the first place.
- The number of accepted particles is fixed at K=1000 (Page 24) without justification (e.g., no discussion of sensitivity of conclusions to K, or power/precision analysis for the Monte Carlo estimates of rejection counts).
- It is unclear whether the reported rejection-rate statistics (ˆRB, ˆRE) and the computational-gain ratios in Table 3 / Figure 2 are based on a single simulation run or averaged over multiple replications, and no measure of Monte Carlo variability (e.g., standard errors across replications) is reported for these quantities.

Suggestions:

- Report the value of n|ε| (or nε) alongside each simulation setting in Tables 1-2 so readers can judge how close the empirical study is to the theoretically valid asymptotic regime.
- Include at least one example with an intractable likelihood (the actual use case for ABC) to demonstrate that the weight-function computations are feasible via simulation, as suggested in the conclusion (Page 26).
- Report replication-based standard errors or confidence intervals for ˆRB, ˆRE and the computational gain ratio to substantiate claims of consistent (rather than noise-driven) improvement.
- Provide a brief justification or sensitivity check for the fixed K=1000 accepted-particle choice.

Results - Score: 3/5 (confidence: 0.60)

Strengths:

- The theoretical results are internally consistent: the ellipse formula for relative entropy correctly reduces to the ball formula when ε1=...=εq=ε (Remark 2.3, Page 9), and Corollary 3.5 correctly reduces to Lemma 3.1 when q=1 (Page 12).
- The simulation results in Tables 1-3 qualitatively confirm the theoretical prediction that ellipse acceptance regions reduce rejection counts relative to balls at matched tolerance (e.g., ˆRE/ˆRB consistently below 1 in Table 3, Page 25).
- The paper is candid that the asymptotic approximation 'overestimates the ratio of mean rejection rates' relative to the empirical values (Page 24-25), an honest reporting of a discrepancy rather than an overstated fit.

Weaknesses:

- The observed overestimation by the asymptotic formula Ũ compared to the empirical ˆRE/ˆRB (Table 3, e.g., 0.986 vs 1.014 or more starkly 0.648 vs 0.537 at (tol=0.25, n=300)) is acknowledged but not quantitatively analyzed (e.g., no discussion of the systematic direction or magnitude of this bias across settings).
- Results are presented for only a single true parameter setting $(\mu,\sigma^2)=(0,1)$ and one set of hyperparameters (Page 24), so it is unclear how sensitive the reported computational gains are to the location of τ^* or to prior specification, despite the paper itself noting elsewhere that the size of τ^* matters (Section 6.1, Page 18-19).
- Figure 2's observation that 'computational gains are more evident as n increases, but seem to stabilize for large n' (Page 26) is presented as an empirical pattern but not explained theoretically, leaving a gap between the asymptotic scaling results (Section 3) and this observed n-dependence of the gain itself.

Suggestions:

- Add a systematic comparison across multiple true parameter values and prior settings to establish the robustness of the reported computational gains.
- Discuss quantitatively why the asymptotic formula systematically overestimates or underestimates the empirical ratio, perhaps tying it to the O(|ε|6) or O((n|ε|)3) remainder terms.

Novelty - Score: 4/5 (confidence: 0.60)

Strengths:

- Casting ABC threshold selection as a relative-entropy-constrained optimization problem, with ε allowed to be a vector, is presented as a genuinely new geometric viewpoint distinguishing this work from prior threshold/bias analyses (Page 2: 'The focus of our paper is different... studying the relative entropy between the underlying and the simulated posterior distribution').
- The explicit connection and generalization of the bias formula in [3] from the ball case to the more general ellipse case is clearly and honestly flagged (Page 14: 'the computations in [3] are for a ball as an acceptance region..., whereas here we present the result for the more general case of an ellipse').

Weaknesses:

- The core technical machinery (Taylor expansion of relative entropy, Laplace asymptotics for exponential families) uses standard tools, so novelty rests primarily on the specific combination and application to ABC rather than new mathematical techniques.
- The paper does not explicitly discuss whether other information-theoretic ABC analyses (e.g., relative

entropy applied to different divergence measures, KL between simulated/observed statistics rather than posteriors) have been considered elsewhere, which would help position the novelty more precisely relative to the broader information-theoretic UQ literature it cites (Page 2, refs [1,2,11,15,16]).

Suggestions:

- Include a short related-work paragraph explicitly contrasting this relative-entropy-on-posteriors approach with any other information-theoretic ABC treatments, if they exist, to sharpen the novelty claim.

Contributions - Score: 3/5 (confidence: 0.60)

Strengths:

- The derived weight functions and closed-form expansions (Propositions 2.2, Lemma 3.4) provide a concrete, actionable mathematical tool for practitioners to potentially tune ε_i per direction (Page 4).
- The comparison of mean rejection rates between ball and ellipse acceptance regions (Section 5, Lemma 5.1) offers a practically interpretable metric (computational gain) connecting theory to algorithmic efficiency.

Weaknesses:

- Both illustrative examples use models where the likelihood is fully tractable and 'ABC methods are not necessary for inference' (Page 17), which limits the paper's demonstrated practical impact for genuine likelihood-free applications that are ABC's core use case.
- For non-exponential-family models (which include most realistic ABC applications), the paper states 'closed form computation may not be possible and one would have to result to simulation' (Page 26) and explicitly defers development of the needed online/Monte Carlo weight-estimation methods to future work, meaning the practical contribution for typical ABC users is not yet realized.
- The mean rejection rate framework depends on ε being small enough for the asymptotic approximations, which as discussed under Methodology, is not clearly satisfied in some of the reported simulation settings, tempering how much practical guidance can currently be drawn.

Suggestions:

- Clarify in the conclusion the specific class of real-world ABC problems for which the results are immediately actionable versus those requiring future methodological development.
- Consider a proof-of-concept experiment using simulation-based estimation of the weight functions on a model with an intractable likelihood, even if only approximately, to demonstrate practical feasibility.

Limitations - Score: 2/5 (confidence: 0.60)

Strengths:

- The conclusion acknowledges that closed-form weight function computation may not generalize beyond the presented examples: 'For more general models, closed form computation may not be possible and one would have to result to simulation' (Page 26).

- The paper is honest that the theoretical formula for the mean rejection ratio is 'an approximation formula valid only in the limit and thus care is needed in its application and interpretation' (Page 22).

Weaknesses:

- There is no dedicated limitations discussion addressing the strong reliance on sufficient statistics being available (Page 5-6), despite the introduction itself noting this is 'usually hard to come up with in practice' (Page 2).
- The restriction to diagonal-matrix (axis-aligned ellipse/ball) acceptance regions is not clearly flagged as a limitation, even though the paper notes 'Other choices for the acceptance region... are of course possible' (Page 6) without discussing what is lost by not considering general positive-definite $A(\varepsilon)$.
- The potential mismatch between the small-$|\varepsilon|$ asymptotic regime required for the theory and the ε values actually used in the simulation study (as discussed under Methodology) is not explicitly flagged as a limitation of the empirical validation.
- Condition 3.3's requirement of a unique, non-degenerate interior posterior mode (Page 12) is a non-trivial restriction (excluding multimodal posteriors) that is not discussed as a boundary of applicability.

Suggestions:

- Add an explicit 'Limitations' subsection discussing the sufficient-statistic assumption, the diagonal-metric restriction, the small-ε validity regime, and the unimodality assumption (Condition 3.3).
- Discuss how results might change qualitatively when sufficient statistics are unavailable and only approximately sufficient summary statistics are used, since this is the norm in applied ABC.

References - Score: 4/5 (confidence: 0.70)

Strengths:

- The bibliography covers foundational ABC methodology papers (e.g., [4,12,13,20,21,23]) as well as relevant information-theoretic UQ papers by the author's collaborators (e.g., [1,2,11,15,16]), appropriately supporting the paper's claims.
- Citation style is consistent (numbered bracket style) throughout the text and reference list (Pages 1-2, 31-32), with no evident mixing of citation formats.

Weaknesses:

- The reference list appears to stop around 2017 ([15], [16]) even though the arXiv version is dated 2019 (Page 1: 'arXiv:1812.02127v2... 12 Aug 2019'), so more recent (2017-2019) developments in ABC theory, if any, may not be reflected.

- Reference [18] (Sottoriva and Tavaré) is listed but not clearly tied to an in-text citation location in the excerpted content, making its relevance to the specific arguments harder to verify from the retrieved text.

Suggestions:

- Update the literature review to include any ABC theory papers published between the initial 2016 drafting and the 2019 revision to ensure the 'theoretical developments are still in their infancy' claim (Page 2) reflects the most current state of the field.

Writing Quality - Score: 4/5 (confidence: 0.65)

Strengths:

- The introduction provides an explicit roadmap of the paper's organization matching the actual section structure (Page 5: 'The paper is organized as follows...'), aiding navigability.
- Key results are consistently restated in intuitive language after formal statements (e.g., Remark 2.3, Page 9
- discussion after Lemma 3.4, Page 12), helping readers connect formulas to concepts.

Weaknesses:

- Some derivations are dense with repeated near-identical intermediate expressions (e.g., the multi-line proof of Proposition 2.2, Page 8) that could benefit from more explanatory prose bridging the equations.
- Minor notational overload occurs, e.g., τ is used both as vector of statistics and appears with sub/superscripts across sections in ways that require careful re-reading (e.g., τε notation introduced without much explanation, Page 11).

Suggestions:

- Add brief prose annotations between major algebraic steps in the longer proofs to help readers follow the logical chain without re-deriving each line.
- Consider a notation table summarizing key symbols (τ*, τε, ε, q, n, wi) for reader reference given the density of overlapping symbols.

Discussion - Score: 3/5 (confidence: 0.55)

Strengths:

- The conclusion connects back to the paper's main contributions and appropriately flags an unresolved question raised by the simulation results: 'It would be of interest to theoretically address this aspect of the algorithm' regarding the stabilizing computational gain as n increases (Page 26).
- The discussion is honest that closed-form results in the examples were chosen 'to focus on the more theoretical aspects of the algorithm' rather than claiming broad practical readiness (Page 26).

Weaknesses:

- The discussion does not address the discrepancy between the asymptotic ratio Ũ and empirical ˆRE/ˆRB beyond a brief mention (Page 24-25), missing an opportunity to interpret what this gap implies about the practical reliability of the theoretical formulas.
- The discussion section does not revisit the tension between the sufficient-statistics assumption underpinning all the theory and the practical unavailability of sufficient statistics noted in the introduction (Page 2), leaving this important caveat unaddressed at the conclusion.
- There is limited discussion of alternative explanations for the observed stabilization of computational gain with increasing n (Page 26) beyond stating it would be 'of interest' to study theoretically, without offering even a qualitative hypothesis.

Suggestions:

- Expand the conclusion to explicitly discuss the practical implications and caveats of the sufficient-statistics assumption for real ABC applications.
- Offer a qualitative explanation (even if not fully rigorous) for the observed stabilization of computational gains with n, connecting it to the specific weight function scalings derived in Section 3.

Improvement Suggestions

1. [Hypothesis] State explicit, numbered conjectures/claims (e.g., 'Claim 1: for fixed tolerance, ellipsoidal thresholds yield lower rejection rate than balls whenever weight variances differ across directions') that can be directly matched to the propositions and simulation results.
2. [Hypothesis] Discuss the conditions under which the reverse could occur (i.e., when ball and ellipse perform similarly or when ellipse could underperform) to sharpen the testable content of the central claim.

3. [Problem Statement] Add a paragraph early on clarifying how the theoretical results might extend or degrade when only approximately sufficient statistics are used in practice, since this is the realistic ABC setting.
4. [Methodology] Report the value of n|ε| (or nε) alongside each simulation setting in Tables 1-2 so readers can judge how close the empirical study is to the theoretically valid asymptotic regime.
5. [Methodology] Include at least one example with an intractable likelihood (the actual use case for ABC) to demonstrate that the weight-function computations are feasible via simulation, as suggested in the conclusion (Page 26).
6. [Methodology] Report replication-based standard errors or confidence intervals for ˆRB, ˆRE and the computational gain ratio to substantiate claims of consistent (rather than noise-driven) improvement.
7. [Methodology] Provide a brief justification or sensitivity check for the fixed K=1000 accepted-particle choice.
8. [Results] Add a systematic comparison across multiple true parameter values and prior settings to establish the robustness of the reported computational gains.
9. [Results] Discuss quantitatively why the asymptotic formula systematically overestimates or underestimates the empirical ratio, perhaps tying it to the O(|ε|6) or O((n|ε|)3) remainder terms.
10. [Novelty] Include a short related-work paragraph explicitly contrasting this relative-entropy-on-posteriors approach with any other information-theoretic ABC treatments, if they exist, to sharpen the novelty claim.

Priority Revisions

- Add an explicit limitations section addressing the sufficient-statistics assumption, restriction to tractable-likelihood examples, and axis-aligned region geometry, clarifying how findings might transfer to realistic intractable-likelihood ABC settings
- Revise abstract and conclusion language to accurately reflect that results are asymptotic (ε→0) rather than finite-sample bounds, and to distinguish rejection-rate findings from convergence-rate claims
- Clarify the definition and consistent usage of RB/RE as rejection counts versus the ratio U(τ*), and expand discussion of why the asymptotic ratio Ũ overestimates the empirical ratios in Table 3
- State the regularity/normalization conditions (positivity, differentiation-under-integral, dominated remainder control) needed for the O(|ε|^6) remainder bounds, and clarify the practical relevance of the 'n|ε| small' regime
- Where feasible, extend empirical illustration beyond the single parameter setting and hyperparameter configuration to demonstrate robustness of the ellipse-vs-ball comparison

Cross-Validation Summary

After reconciling both passes, the overall 3/5 assessment remains appropriate. The paper has a strong and carefully organized theoretical core with good internal reduction checks, but confidence is limited by one underived bias formula, a proof step in Lemma 3.1 that deserves algebraic verification, and several table-level inconsistencies or ambiguities in the rejection-rate comparison. Some initial criticisms were also recalibrated to better fit a theoretical paper and the paper's explicitly asymptotic scope.

Validated Strengths

- CONFIRMED: The paper states a clear and testable set of aims up front: 'The contribution of the paper is fourfold' (Pages 2-3), followed by specific tasks including computing the leading-order relative entropy term, allowing vector thresholds, analyzing bias, and studying rejection rates. This makes it straightforward to assess whether the paper delivers on its promises.
- CONFIRMED: The practical motivation is well framed. The introduction identifies concrete ABC pain points —'known issues on sensitivity in performance with respect not only to the distance metric ... but also on the choice of the threshold parameter ε' and the fact that ABC 'typically scales badly with the size of the data' (Page 2)—and ties them to a clearly stated theoretical gap: 'theoretical developments are still in their infancy' (Page 2).
- CONFIRMED: The mathematical problem statement is precise. The paper formulates threshold choice as 'maximize ε-dependent acceptance region subject to H(P|Pε)(τ∗) ≤ tol' (1.1) and then derives the tractable asymptotic surrogate (1.2) on Page 4.
- ADDED: The theory includes meaningful internal self-checks. Remark 2.3 states that the ellipse formula 'immediately reduces to the formula for the ball, when ε1 = · · · = εq = ε' (Page 9), and Corollary 3.5 adds 'if also q = 1 then we obtain the statement of Lemma 3.1' (Page 12). These reductions strengthen confidence in the formal development.
- ADDED: Lemma A.1 provides a careful symmetry argument supporting the leading-order expansion. The appendix shows that '∇εfε(θ|τ∗)|ε=0 = 0' and that mixed second derivatives vanish, 'for all i, j ... with i̸= j, ∂2εiεjfε(θ|τ∗)|ε=0 = 0' (Pages 27-29). This is a rigorous and elegant justification for why odd-order and mixed terms drop out.
- ADDED: The examples yield a genuinely useful and somewhat surprising insight: the leading-order bias can be independent of n for selected observables. In the exponential example, 'the bias does not depend on the size of the data set n' (Page 19), and in the normal example Corollary 6.4 says that, when the prior mean and observed mean are small, 'C1, up to leading order, does not depend on n' (Page 22).

Validated Weaknesses

- CORRECTED: The paper's applicability is narrowed by its reliance on sufficient statistics, and this limitation would benefit from fuller discussion. The introduction acknowledges that 'One would like the statistics to be sufficient statistics, but that is usually hard to come up with in practice' (Page 2), while the formal setup then assumes 'we are working with sufficient statistics' (Page 5). The assumption is explicit, but its practical consequences could be discussed more directly.

- CONFIRMED: The generality of the practical claim for ellipsoidal thresholds is broader than the empirical support shown here. The introduction states that 'by exploiting the asymmetric effect of each direction, one is able to enlarge the acceptance region of ABC' and 'the mean rejection rate is typically smaller in the case of ε ∈ Rq+' (Page 3), but the simulation study is a single normal example with one true parameter setting, '(μ, σ2) = (0, 1)' (Page 24). The claim would be stronger with additional models or parameter regimes.
- CORRECTED: The asymptotic regime used by the theory deserves more caution in interpreting the numerical comparisons. Section 3 states that 'the expansion of relative entropy in ε is typically correct if the product n|ε| is small' (Page 9), and Section 7 reiterates that the approximation is 'accurate only in the limit as n|ε| ↓ 0' (Page 23). Some reported settings involve moderate products—for example, Table 1 gives n = 1000 and ε = 0.022 for (tol, n) = (1, 1000) (Page 24)—so the simulations are still informative, but the quantitative match to asymptotic formulas should be interpreted cautiously.
- CONFIRMED: The illustrative examples are analytically convenient but not representative of the main use-cases where ABC is needed. Section 6 explicitly notes that 'in these cases, we have explicit information for the likelihood and thus ABC methods are not necessary for inference' (Page 17). This is reasonable for exposition, but it limits the paper's practical demonstration.

Validated Gaps

- CONFIRMED: A self-contained derivation of the central bias expansion is not found in the retrieved text. On Page 14, the paper states, 'Algebraic computations similar to those in the case of a ball in [3] ... shows that bias(...) = Σ εi2 Ci(τ∗) + O(|ε|3),' but an explicit proposition or proof is not provided in the retrieved pages. Since Sections 6.1 and 6.2 use this formula, a fuller derivation would improve traceability.
- ADDED: Lemma 3.1's proof would benefit from an explicit verification of the displayed fourth-order algebra. On Page 11 the proof writes terms including '2/(5!3)(nε)4 Eη4(θ)' and later '+ 2/5!(nε)4 Eη4(θ)' before concluding with '(nε)4 1/72 Var(η2(θ)) + O(|nε|6)'. As written, the η4-coefficient cancellation/simplification is not transparent from the displayed steps and merits checking or a short intermediate line.
- ADDED: The rejection-rate comparison in Lemma 5.1 is not fully pinned down as an apples-to-apples comparison between ball and ellipse. Page 15 defines 'pB = P(τ ∈ Bε(τ∗))' for a scalar ε and 'pE = P(τ ∈ Dε(τ∗))' for vector ε ∈ Rq+, then gives formula (5.1) mixing 'ε' in the numerator with 'εi' in the denominator. Section 7 later compares thresholds chosen under 'H(P|Pε)(τ∗) ≤ tol' and, for the ellipse only, solves 'maximize (πε1ε2) subject to leading order term in (6.7) ≤ tol' (Page 23). Stating explicitly how the ball radius ε is matched to the ellipse components εi would clarify the interpretation of U(τ∗).
- CORRECTED: There are small cross-table discrepancies in the reported observed statistics that complicate

exact row-wise comparison, even if they may be transcription issues. For example, for (tol, n) = (1, 100), Table 1 reports '$\tau* = (-0.181, 0.866)$' (Page 24) while Table 2 reports '$(-0.081, 0.866)$' (Page 25); for (0.05, 1000), Table 1 reports '$(-0.027, 1.013)$' (Page 24) while Table 2 reports '$(-0.027, 1.033)$' (Page 25).

- ADDED: Table 1 contains an internal tolerance/threshold inconsistency that likely reflects a typo and affects interpretation of the rejection-rate trend. On Page 24, the row '(0.5, 600)' reports '$\varepsilon = 0.027$' with '$\hat{R}B = 5805$', whereas the looser row '(1, 600)' reports the much smaller '$\varepsilon = 0.003$' with '$\hat{R}B = 4172$'. Since larger tolerance would ordinarily not correspond to a dramatically smaller threshold, this row should be checked.
- CORRECTED: The introductory statement about lower rejection rates would benefit from qualification to match the later theory. The introduction says, 'We find that the mean rejection rate is typically smaller in the case of $\varepsilon \in Rq+$' (Page 3), whereas Section 5 gives a conditional statement: 'if $\varepsilon i/\varepsilon \gg 1$ ... one can have benefits using the ellipse' (Page 16). Aligning the introduction more closely with the conditional form of the result would improve precision.

Corrections Applied

- The review no longer treats the rejection-rate identity on Page 15 as a definitional error; it is now framed as an asymptotic approximation consistent with the paper's stated 'small $|\varepsilon|$ regime.'
- The concern about the small-parameter validity regime was softened. Rather than asserting that particular simulations definitively fall outside the theory, the final review now notes that Pages 9 and 23 themselves restrict quantitative accuracy to cases where '$n|\varepsilon|$' is small and therefore some rows should be interpreted cautiously.
- The earlier empirical-hypothesis criticism was removed because it applied an inappropriate standard to a theory paper.
- The cross-table $\tau*$ discrepancies were retained but reframed as likely transcription issues that complicate row-wise comparison, rather than strong evidence that the simulations are not controlled.
- The abstract/finite-sample 'error bounds' criticism was softened from an outright mismatch claim to a wording/clarity issue.

New Findings Added

- Lemma 3.1's displayed fourth-order algebra on Page 11 needs verification because the $\eta 4$ terms shown in the proof do not transparently collapse to the final '$1/72\ \mathrm{Var}(\eta 2(\theta))$' expression.
- Table 1 contains an internal inconsistency at $n = 600$: '(0.5, 600)' reports $\varepsilon = 0.027$ while the looser '(1, 600)' row reports $\varepsilon = 0.003$ (Page 24), which is difficult to reconcile and likely a typo.
- Lemma 5.1's ball-versus-ellipse comparison is conceptually under-specified because formula (5.1) on Page 15 mixes a scalar ε for the ball with vector εi for the ellipse, while Section 7 compares thresholds chosen under different optimization setups (Page 23).

- The final review now credits the paper's internal consistency checks more explicitly, especially the reduction of the ellipse formulas to the ball case in Remark 2.3 (Page 9) and Corollary 3.5 (Page 12).
- The final review also adds credit for the paper's concrete n-independent leading-bias results in the examples on Pages 19 and 22.

Discarded Findings

- DISCARDED: The initial complaint that the paper lacks a 'falsifiable empirical hypothesis' was removed. For a theoretical/methodological paper, the more appropriate standard is whether it makes concrete, checkable claims, and it does—for example, 'We find that the mean rejection rate is typically smaller in the case of ε ∈ Rq+' (Page 3), which is then examined numerically in Tables 1-3.
- DISCARDED: The initial claim that 'U(τ∗) = E[RE|τ∗]/E[RB|τ∗] = pB/pE' is a definitional inconsistency was removed. Page 15 explicitly frames the analysis in 'the small |ε| regime,' so the issue is better understood as an asymptotic approximation than as a strict contradiction with the earlier geometric-distribution wording.
- DISCARDED: The earlier criticism that the abstract's 'error bounds on performance for positive tolerances and finite sample sizes' (Page 1) is flatly inconsistent with later O(|ε|6) and O(|ε|3) terms was too strong. Those remainder terms are asymptotic finite-ε error terms; the fairer point is simply that the paper could describe them more explicitly as asymptotic rather than explicit non-asymptotic bounds.

Revision Status vs. Previous Review

Status	Finding
[UNRESOLVED]	The new review still characterizes the results as asymptotic and only accurate in the limit, so it does not supply non-asymptotic finite-sample error bounds. The abstract/body mismatch therefore remains.
[UNRESOLVED]	The examples are still described as analytically convenient but not representative because ABC is unnecessary when the likelihood is tractable. This does not address the limited practical scope.
[UNRESOLVED]	The new review notes that the quantitative match to the asymptotic formula should be interpreted cautiously, but it does not provide a quantitative correction or calibration of the overestimation. The original concern remains open.
[PARTIAL]	The new review acknowledges the sufficient-statistics assumption and flags its practical consequences more explicitly. However, it does not remove the underlying restriction that the main results require a unique interior non-degenerate maximum.

[PARTIAL]	The revised text now cautions that the expansion is accurate only when n|ε| is small and that numerical comparisons should be interpreted cautiously. This clarifies the regime, but it does not fully resolve the tension with practical finite-n ABC settings.
[UNRESOLVED]	The new review still indicates that the paper does not actually provide explicit finite-sample, non-asymptotic bounds. The abstract claim remains insufficiently supported by the main results.
[UNRESOLVED]	The examples are still acknowledged to be cases where ABC is not needed, so they remain pedagogical rather than compelling demonstrations of genuine ABC utility. The limitation persists.
[UNRESOLVED]	The new review does not offer a quantitative error characterization for the rejection-rate approximation, only a cautionary note about interpreting the asymptotics. The approximation gap remains unquantified.
[PARTIAL]	The new review makes the sufficient-statistics assumption explicit and comments on its practical implications, which improves transparency. But it still does not discuss which model classes violate the unique-interior-maximum condition or how broadly the scaling results apply.
[PARTIAL]	The revised review now explicitly states when the asymptotics are accurate and cautions against over-interpreting moderate n|ε| values. That helps distinguish the regimes, but the practical validity range for routine finite-sample ABC is still not sharply delineated.
[UNRESOLVED]	The paper still lacks an example with an intractable likelihood, and the new review confirms the examples are chosen for analytic convenience. So the gap in practical ABC relevance remains.
[UNRESOLVED]	No quantitative characterization of the approximation error is added, only a warning that the asymptotic formulas should be interpreted cautiously. Confidence for non-infinitesimal tolerances is still not strengthened in a measurable way.
[UNRESOLVED]	The new review does not discuss model classes that violate the unique-interior-max condition, so the scope of the exponential-family scaling results remains unclear.
[UNRESOLVED]	The new review continues to describe the results as asymptotic and only accurate in the limit, without providing explicit non-asymptotic finite-sample bounds. The abstract's stronger wording remains unsupported.
[UNRESOLVED]	The revision still does not derive explicit constants for the big-O terms or otherwise furnish genuine non-asymptotic performance bounds. The claim is therefore not

directly substantiated.

Generated by **ProofInk.AI** on August 2, 2026.

Note: Mathematical expressions appear as plain text in Word exports. For fully rendered equations, use the PDF or on-screen view.