FQM Paper 2 — v2.3 Supplementary Source Audit

Audit of the two uploaded archives; source scripts and JSON outputs preserved unchanged.

Scope

The audit covers the complete v2 supplement (29 entries) and the Levels 2–3 falsification archive (9 entries). It checks manuscript/script alignment, input provenance, computational reproducibility, and whether the statistical language matches the supplied code.

Key findings

Area	Audit result
Test 10 — GRB	Synthetic pilot only; 24 trials per condition and B=10 bootstrap resamples per trial. Not a real GRB photon analysis; power
Test 11 — planetary DSI	Separate Solar-System eight-planet semi-major-axis test and synthetic log-periodic power study. Re-executed; JSON exa
Exoplanet catalogue	Distinct script reports 879 systems / 1003 ratios, but OEC snapshot, revision, date and hash are absent. Cannot regenera
Look-elsewhere calibration	Only 40 null simulations and a downsampled grid (`grid[::4]`) versus 371 observed centres. Global threshold is coarse.
Test 12 — relaxation	Re-executed; JSON exact match. Model/identifiability result, not a physical relaxation measurement.
Test 13 — pairing	Mean-field `z_fib_vs_random` is Δ divided by random SD, not an inferential z-test. DMRG random controls n=6 and n=3; T
Model conventions	Mean-field pairing kernel uses positive U_pair=+2; DMRG attractive Hubbard uses U=-4. Manuscript v2.3 distinguishes th
Execution	Tests 10 and 13 exceeded the available execution window. ngspice 42/scikit-rf, Astropy, and TeNPy environments were n

Interpretation and required upgrades

• Do not describe the synthetic Test 10 rates as observational constraints. Increase bootstrap resamples to at least 2,000 and independent trials to at least 1,000 per condition for confirmatory calibration.

• Archive the exact OEC snapshot and checksum; rerun the full 371-centre scan with at least 10,000 null simulations.

• For Test 13, pin TeNPy/BLAS versions, archive convergence logs and all random controls, increase sample counts, and add finite-size scaling uncertainties.

• The supplied Stage 6 ngspice 47 workflow is independent of the older ngspice 42/scikit-rf experiment and does not replace it.

• No physical VNA measurement is included. The LC predictions remain model predictions until a calibrated bench experiment is performed.

Re-execution status

Test 11 and Test 12: exact JSON match after re-execution. Test 10 and Test 13: not completed within the available execution window. Exoplanet catalogue, ngspice/scikit-rf experiment A, and DMRG experiment C: not re-executed because required data or matching dependencies were absent.

The original supplied files and outputs are retained unchanged. Corrections in the v2.3 manuscript concern interpretation, statistical terminology, and provenance disclosure.