Supplement S3: Observable Dataset

Complete GIFT Observable Dataset (95 entries) Brieuc de La Fournière
Independent researcher, Beaune, France May 2026 | v3.4

Contents

0.1 S3.1 Introduction

This supplement provides the complete GIFT observable dataset: 95 observables derived from a single G_2 metric (169 Chebyshev-Cholesky geometric parameters, zero freely adjustable physical parameters), organized by derivation type and cross-referenced with Lean formal verification. The dataset includes 2 BH remnant predictions from Pinčák et al. 2026 [5] (classified Type IV) and 6 new Type IV metric block eigenvalues.

The 95 observables decompose as **33(I) + 19(II) + 21(III) + 22(IV)** across four types of increasing derivation complexity. All predictions trace to a single compact G_2 manifold K_7 with (b_2, b_3) = (21, 77) and $E_8 \times E_8$ gauge structure, 169 metric parameters, zero continuously adjustable physical parameters. The pair (b_2, b_3) = (21, 77) does not appear among the catalogued compact G_2 manifolds; a complete geometric construction remains an open problem. The certified metric provides computational evidence for a new G_2 manifold.

The dataset serves as the canonical reference for all GIFT predictions, superseding scattered results across supplements S6–S25. Every entry traces back to a specific computation, a JSON artifact, and (where applicable) a Lean theorem.

0.2 S3.2 Type Classification

Observables are classified into four types by derivation directness. The type assignment reflects the number of physical identification steps between the G_2 metric and the final prediction:

Type	Count	Description	Derivation chain	Lean coverage
I	33	Direct algebraic from structural constants	topology → formula → ratio	33/33 (100%)
II	19	One physical identification step	formula + VEV or scale → dimensional quantity	0/19 (0%)
III	21	Multi-step dynamical chains	metric → mechanism → observable	14/21 (67%)
IV	22	Structural/internal quantities	metric → diagnostic	8/22 (36%)

Type I (33 observables): Direct from metric geometry and topological invariants (b_2, b_3, dim(G_2), dim(E_8), etc.). These are dimensionless ratios expressed as algebraic combinations of integers. Examples: $\sin^2\theta$_W = 3/13, Q_Koide = 2/3, m_τ/m_e = 3477. All 33 are Lean-certified. The two BH remnant topological predictions are classified as Type IV structural diagnostics (not part of the 33 Type I Lean-certified core).

Type II (19 observables): One physical identification step beyond Type I. Absolute masses from ratios × VEV (e.g., m_u = (m_u/m_d) × m_d_exp), CKM magnitudes from Wolfenstein parametrization, extended quark ratios. From the `gift_observables.csv` pipeline.

Type III (21 observables): Multi-step chains involving dynamical mechanisms: - wilson_line non-adiabatic eigenvalue splitting (3 observables: raw lepton ratios) - S11 RGE gauge coupling running (4 observables: $\sin^2\theta$_W, α_em^{-1}, α_s at M_Z, M_GUT) - spectral 7D spectral geometry (5 observables: Weyl exponent, KK states, fiber channels) - gauge_bundle gauge bundle data (4 observables: f_IJ conditioning, α_ratio, Yukawa rank) - instanton instanton volumes + combined wilson_line+instanton (5 observables: ΔV's, combined ratios)

Type IV (22 observables): Structural and internal consistency quantities with no direct experimental comparison. These include topology diagnostics (b_2, b_3, $\chi(K_7)$), Newton-Kantorovich certification values (h, distance, margin), Gram matrix conditioning (cond(G_K3), cond(G_77)), spectral counts, metric block eigenvalues (g_ss = 19/6, g_{T^2} = 7/6, g_K3 $\approx$ 64/77), and Pinčák et al. 2026 [5] topological diagnostics (N_QNM = 98, M_res, instanton suppression $\times$77). M_res and N_QNM are the two BH-remnant-related entries; they are classified here as structural diagnostics, not as Type I or Type III, because they require a physical scale identification (τ_0 = v_EW) not derivable from topology alone. These quantities verify internal consistency rather than predict measurements directly.

Derivation depth: The type assignment correlates with derivation chain length. Type I observables require 1 algebraic step; Type II requires 2 steps (algebra + scale identification); Type III requires 3–5 steps (algebra + dynamics + calibration); Type IV involves diagnostic computation chains of variable length.

0.3 S3.3 Methodology

0.3.1 Experimental Sources

All experimental reference values are drawn from four standard compilations:

Source	Coverage	Date
PDG 2024 [1]	Particle masses, gauge couplings, CKM elements	2025
NuFIT 6.0 [2]	Neutrino oscillation parameters (NO w/o SK)	Oct 2024
Planck PR4 (Tristram+ 2024) [3]	Cosmological parameters: h, σ_8, Ω_Λ	2024
CODATA 2022 [4]	Lepton mass ratios	2022

Selected NuFIT 6.0 values: $\sin^2\theta_{23}$ = 0.561, δ_CP = 177°. Selected Planck PR4 values: h = 0.6764, σ_8 = 0.807, Ω_Λ = 0.6847.

0.3.2 Structural Constants

The 20 structural constants from which all 95 observables derive:

#	Constant	Value	Origin
1	$\dim(E_8)$	248	E_8 Lie algebra
2	$\mathrm{rank}(E_8)$	8	Cartan subalgebra
3	$\dim(G_2)$	14	$\mathrm{Aut}(\mathbb{O})$
4	$\dim(K_7)$	7	$\mathrm{Im}(\mathbb{O})$
5	$b_2(K_7)$	21	Input; confirmed by spectral analysis (Paper B [B], §2.3). Any Mayer-Vietoris decomposition is conditional on building block identification (open problem).
6	$b_3(K_7)$	77	Input; confirmed by spectral analysis (Paper B [B], §2.3). Any Mayer-Vietoris decomposition is conditional on building block identification (open problem).
7	N_gen	3	Index theorem
8	p_2	2	$\dim(G_2)/\dim(K_7)$
9	Weyl	5	Triple identity
10	H*	99	b_2+b_3+1
11	$\dim(J_3(\mathbb{O}))$	27	Jordan algebra
12	$\dim(E_6)$	78	E_6 Lie algebra
13	$\dim(E_7)$	133	E_7 Lie algebra
14	$\dim(F_4)$	52	$\mathrm{Aut}(J_3(\mathbb{O}))$
15	$\mathrm{fund}(E_7)$	56	E_7 fundamental
16	$\lvert PSL(2,7)\rvert$	168	Fano automorphisms
17	D_bulk	11	4+7
18	α_sum	13	$\mathrm{rank}(E_8)$+Weyl
19	det(g)_den	32	$b_2+\dim(G_2)-$N_gen
20	det(g)_num	65	Weyl$\times\alpha$_sum

0.3.3 Consolidation Pipeline

The master dataset (`gift_observables.json`) is generated by `run_observable_dataset.py` (observable_dataset), which collects all individual JSON results from S6–S25 and performs 5 verification checks:

1. **Completeness**: All 95 entries present across 4 types
2. **Uniqueness**: No duplicate observable IDs
3. **Consistency**: GIFT values match source JSONs to machine precision
4. **Source tracing**: Every entry maps to a computation step (S6–S25)
5. **Format validation**: JSON, CSV, and LaTeX outputs agree

Each observable record contains: GIFT prediction (float64), GIFT fraction (exact symbolic), experimental value and uncertainty, experimental source, relative deviation (%), Lean theorem name (if certified), and provenance (computation step).

0.4 S3.4 Complete Observable Table

Table S3.1: Complete GIFT observable dataset (95 entries). Columns: name, GIFT prediction, experimental value, relative deviation (%), type (I/II/III/IV), source computation (S6–S25), Lean verification status.

[The full LaTeX table is provided in `data/gift_observables_latex.tex` and will be included at typesetting.]

0.4.1 Sector-by-Sector Performance

The 66 observables with experimental comparison (Types I+II+III) organize into 11 sectors with varying precision:

Best-performing sectors (mean dev $< 0.5\%$): - **Boson** (3 obs): mean 0.13%, m_H/m_W (0.02%), m_W/m_Z (0.06%), m_H/m_t (0.31%) - **Quark** (4 obs): mean 0.21%, m_s/m_d (0.00%), m_u/m_d (0.05%), m_c/m_s (0.12%), m_b/m_t (0.79%) - **CKM** (3 obs): mean 0.59%: A_Wolf (0.29%), $\sin^2\theta_{12}$ (0.36%), $\sin^2\theta_{23}$ (1.13%) - **Cosmology** (7 obs): mean 0.15%, n_s (0.004%), h (0.09%), Ω_b/Ω_m (0.16%), σ_8 (0.18%), Ω_DE (0.21%), Y_p (0.37%), Ω_DM/Ω_b (0.00%)

Good sectors (mean dev 0.5–3%): - **Electroweak** (3 obs): mean 0.11%, α^{-1} (0.002%), α_s (0.126%), $\sin^2\theta$_W (0.20%) - **Lepton** (3 obs): mean 0.04%: Q_Koide (0.001%), m_τ/m_e (0.004%), m_μ/m_e (0.12%) - **PMNS** (4 obs): mean 0.29%, θ_{12} (0.03%), θ_{23} (0.10%), $\sin^2\theta_{12}$ (0.23%), θ_{13} (0.37%)

Challenging sectors (mean dev $> 3\%$): - **Gauge running** (4 obs): mean 2.3%, $\sin^2\theta$_W(RGE) (2.78%), α_em^{-1}(RGE) (2.53%), α_s(RGE) (3.7%), M_GUT (0.00%) - **Instanton** (3 obs): mean 7.4%: ΔV(e-τ) (5.9%), ΔV(e-μ) (15.9%), κ(gauge) (4.7%)

B-test consistency (derived, not counted as independent observable): The MSSM B parameter B $= (\alpha_1{}^{-1} - \alpha_2{}^{-1})/(\alpha_2{}^{-1} - \alpha_3{}^{-1})$ evaluates to 1.4033 using GIFT couplings (0.23% from the theoretical 7/5), closer than the purely experimental B = 1.3948 (0.37%). The identity B = 7/5 is equivalent to α_em^{-1}(M_Z) $= 91\sqrt{2} = \dim(\Lambda^2\mathfrak{g}_2)\cdot\sqrt{2}$ and implies the holonomy sequence $\alpha_1{}^{-1}{:}\alpha_2{}^{-1}{:}\alpha_3{}^{-1}$ = $\dim(G_2){:}\dim(K_7){:}p_2$ = 14:7:2. See §5.4 of the main text.

0.4.2 Exact Matches

11 observables achieve deviations below 0.01% (effectively exact):

Observable	GIFT	Exp.	Dev.
α^{-1}	137.033	137.036	0.002%
Q_Koide	2/3	0.666661	0.001%
m_τ/m_e	3477	3477.15	0.004%
m_s/m_d	20	20.0	0.000%
n_s	0.9649	0.9649	0.004%
m_u	2.16	2.16	0.00%
|V_tb|	0.999	0.999	0.00%
m_c/m_d	234.3	234.0	0.004%
M_GUT	2×10^{16}	2×10^{16}	0.00%
α_GUT^{-1}	25.3	25.3	0.00%
rank(Y)	3	3	0.00%

Note: δ_CP canonical prediction is 197° (11.3% from NuFIT 6.0 central); a compactification factor 62/69 brings it to 177.014° (0.008%) but is not adopted, see §4.4.1. m_c/m_s deviation is 0.12%, above the 0.01% threshold.

0.4.3 Known Outliers

Two observables deviate by more than 5%, each with an identified physical origin:

δ_CP: The GIFT prediction is 197° (pure topological: 7 × 14 + 99), deviating 11.3% from NuFIT 6.0 central (177° ± 20°) but at the edge of 1σ. PSLQ analysis identifies a compactification factor 62/69 (§4.4.1 of main paper), documented as a structural observation but not adopted: the canonical prediction remains 197°. The experimental central value shifted from 197° (NuFIT 5.2) to 177° (NuFIT 6.0) and carries ±20° uncertainty. Falsifiable by DUNE (2028–2040).

α_s(RGE) = +3.7%: G_2-MSSM split-spectrum matching (MSSM above M_squark = 3165 GeV, SM + gauginos below ($b_3 = -5$)) gives α_s = 0.1224 (exp 0.1180, dev 3.7%). A naive degenerate-spectrum treatment would give 12.1%. The topological value α_s = $\sqrt{2}/12$ = 0.11785 (Type I) matches experiment to 0.126%.

ΔV(e-μ) = +15.9%: The instanton volume assignment ΔV(e-μ) = 3.271 versus target ln(16.82) = 2.823 reflects the optimal assignment problem for 57 associative cycles. The combined wilson_line+instanton pipeline with α = e^K (§S3.5) reduces this to 0.75%.

κ(gauge) = +4.7%: The gauge kinetic function conditioning cond(f_IJ) = 1.047 measures departure from exact universality. The deviation from 1.000 reflects K_7 geometry at the percent level.

0.4.4 Summary Statistics

Metric	Value
Total observables	95
Type breakdown	33(I) + 19(II) + 21(III) + 22(IV)
With experimental comparison	66
Exact matches (dev < 0.01%)	11
Within 1% of experiment	53
Within 5% of experiment	63
Lean-certified	55
Structural constants	20
Metric parameters	169

0.5 S3.5 The Combined wilson_line+instanton Pipeline

The lepton mass hierarchy (combined lepton ratio observables) uses two complementary geometric mechanisms. The complete derivation is in §6 of the main text; results are summarised here for cross-reference.

Raw wilson_line (K3 fiber eigenvalue splitting at c = 0.452): m_τ/m_e = 3403 (2.1%), m_τ/m_μ = 16.54 (1.7%), m_μ/m_e = 205.7 (0.5%). Raw instanton (associative 3-cycle volumes, 57 cycles): ΔV(e-τ) = 8.633 (5.9%), ΔV(e-μ) = 3.271 (15.9%). Combined with α = e^{K_0} = $\hat{V}$^{−3} = 0.002763 (Kähler potential of K_7, zero free parameters):

- m_τ/m_e = 3485 (exp 3477, **0.24%**)
- m_τ/m_μ = 16.69 (exp 16.82, **0.75%**)
- m_μ/m_e = 208.8 (exp 206.8, **0.97%**)

Certified in Lean: `AssociativeVolumes.lean` (14 conjuncts, 0 sorry). Full mechanism description: main §6.4.

0.6 S3.6 Sensitivity Cross-Reference

The complete sensitivity analysis is in §7 of the main text. Key results for cross-reference:

- **Effective DOF**: r_eff = 15.53 (SVD of 20×33 constant-usage Jacobian)

- **Overdetermination**: 33/15.53 = 2.13× (more constraints than free dimensions)
- **Cross-coupling**: 155/528 observable pairs share a structural constant; all form one connected component
- **P(coincidence, uniform)** = 10^{-346} (χ^2 + Fisher combined, main §7.5)
- **P(coincidence, algebraic)** = 10^{-133} (4.2M random formulas from same 20 constants, main §7.5)

Figures: `fig_sensitivity_heatmap.png`, `fig_constant_usage.png`, `fig_observable_correlations.png`, `fig_mc_per_observable.png`.

0.7 S3.7 Lean Cross-References

Of the 95 observables, **55 are formally verified** in Lean 4 across the following certificate files:

Lean File	**Axioms**	**Theorems**	**Conjuncts**	**Coverage**
`Foundations.lean`	0*	n/a	38	Type I: metric, torsion, topology
`Predictions.lean`	0*	n/a	55	Types I–III: couplings, masses, mixing
`Spectral.lean`	0*	n/a	41	Type II: KK spectrum, Weyl law
`MetricEigenvalues.lean`	0	n/a	15	Metric fractions
`SpectralInvariants.lean`	0	n/a	10	Heat kernel, $\zeta'(0)$, $b_1 = 0$
`CompactificationCorrection.lean`	0	n/a	6	δ_{CP} compactification factor
`TCSGaugeBreaking.lean`	0	14	10	Type III: gauge breaking chain
`GaugeBundleData.lean`	0	12	11	Type III: bundle universality
`AssociativeVolumes.lean`	0	19	14	Type III: instanton hierarchy
`ComputedWeylLaw.lean`	0	8	7	Type III: 7D spectral geometry

*4 published axioms. All substantive: standard theorems (Cheeger inequality) + geometric structure (TCS spectral bounds) + physical inputs (literature package: CGN+Joyce). KK_YM_EFT, K7_spectral_data, K7_analysis_data are theorems/defs. G_2 group structure proven by `native_decide`.

Total certificate: 213 conjuncts, 0 `sorry`, 134 .lean files (128 core + 6 generated + 12 test/support), 8378 build jobs (Lean 4.29.0).

0.7.1 Coverage by Type

Type	Certified	Total	Coverage
I	33	33	100%
II	0	19	0%
III	14	21	67%
IV	8	22	36%
Total	**55**	**95**	**58%**

Type II uncertified: The 19 Type II observables derive from the CSV pipeline (absolute masses from ratio × VEV, CKM magnitudes from Wolfenstein). These involve dimensional quantities and intermediate data that have not yet been axiomatized in Lean. The underlying Type I ratios from which they derive are all certified.

Type IV partial: 8 of 22 structural quantities are certified (Betti numbers, holonomy dimension, NK certification bounds). The remaining 14 involve spectral counts, conditioning numbers, metric eigenvalues, and BH remnant estimates that require floating-point Lean infrastructure not yet developed.

0.8 S3.8 Data Availability

The complete dataset is available in three formats:

- **JSON**: `gift_observables.json` (machine-readable, full metadata per observable)
- **CSV**: `gift_observables.csv` (tabular, 95 rows × 13 columns, for analysis)
- **LaTeX**: `gift_observables_latex.tex` (publication-ready table)

All artifacts are archived in the Zenodo data deposit (DOI: 10.5281/zenodo.19893371), which also contains the 9 source JSON files from S6–S25, the 10 computation scripts, and the 10 figures.

0.9 References

[1] Particle Data Group, Phys. Rev. D 110, 030001 (2024) [2] NuFIT 6.0, www.nu-fit.org (Oct 2024) [3] L. Tristram et al., A&A 682, A37 (2024) (Planck PR4) [4] CODATA 2022, physics.nist.gov/cuu/Constants

---

GIFT Framework: Supplement S3: Observable Dataset 95 observables, 4 types, 55 Lean-certified

References

References

[1] Particle Data Group, "Review of Particle Physics," *Phys. Rev. D* **110**, 030001 (2024).

[2] NuFIT 6.0, https://www.nu-fit.org (Oct 2024).

[3] L. Tristram et al., "Planck PR4 cosmological parameters," *Astron. Astrophys.* **682**, A37 (2024).

[4] CODATA 2022 recommended values, https://physics.nist.gov/cuu/Constants.

[5] R. Pinčák, A. Pigazzini, M. Pudlák, E. Bartoš, "Geometric origin of a stable black hole remnant from torsion in G_2-manifold geometry," *Gen. Rel. Grav.* **58**, 29 (2026).

GIFT Framework — Supplement S3: Observable Dataset — 95 observables, 4 types, 55 Lean-certified