



What Do Precision Tests of General Relativity Actually Measure?

Matthew Lukin Smawfield

Version: v0.1 (Istanbul)

First published: 31 December 2025

DOI: 10.5281/zenodo.18109761

Abstract

Most high-precision tests of general relativity constrain reciprocity-even, largely local observables within single-metric frameworks. This leaves open a specific underdetermination between General Relativity (GR) and a class of two-metric disformal scalar-tensor modifications, exemplified here by the Temporal Equivalence Principle (TEP).

This paper formalizes a measurement taxonomy distinguishing gauge-invariant from convention-dependent observables and identifies five recurring scope limitations in the experimental canon: (1) two-way measurement dominance; (2) local/global conflation; (3) model-dependent calibration; (4) the conformal loophole in multi-messenger constraints; and (5) theory-laden data reduction. These characteristics do not diminish the experimental achievements but indicate that, in many cases, the tests primarily constrain parameter space within assumed frameworks rather than systematically discriminating between alternatives.

Discriminating observables—specifically loop asymmetries and spatial correlations—are proposed, together with experimental configurations capable of resolving the underdetermination. These include large-area triangle holonomy tests (targeting residual synchronization holonomy H_{resid}), interplanetary closed-loop timing, and matter-wave interferometry.

Keywords: general relativity, experimental tests, underdetermination, synchronization, GPS, gravitational redshift, Shapiro delay, multi-messenger astronomy

1. Introduction: The Experimental Canon

General relativity stands as one of the most precisely tested theories in physics. From the gravitational redshift measurements of Pound and Rebka (1960) to the multi-messenger observations of GW170817 (2017), a century of experiments has confirmed GR's predictions with extraordinary precision. Modern optical lattice clocks achieve fractional frequency comparisons at the 10^{-18} level and below, detecting gravitational time dilation across millimeter height differences. The Cassini spacecraft measured the Shapiro

delay to constrain the PPN parameter γ to within 2×10^{-5} of unity. The Global Positioning System, which incorporates relativistic corrections as a matter of routine engineering, is often cited as the most practical demonstration that "GR works."

The standard narrative concludes that these tests leave no room for alternatives to GR. Any modification of gravity, the argument goes, would have been detected by now. The experimental foundations are secure.

This paper challenges that conclusion—not by disputing the experimental results, but by examining what those results actually constrain. A systematic methodological analysis reveals that the canonical precision tests share structural features that render them, as currently configured, unable to directly distinguish between general relativity and certain classes of alternatives, including the Temporal Equivalence Principle (TEP).

1.1 The Textbook Narrative

The experimental confirmation of GR is typically presented as a progression of increasingly precise tests:

Experiment	Year	Claimed Precision	Result
Pound-Rebka (redshift)	1960	1%	Confirms GR
Hafele-Keating (time dilation)	1971	10%	Confirms GR
Gravity Probe A	1976	7×10^{-5}	Confirms GR
GPS (operational)	1978–	Meter-level positioning	Confirms GR
Cassini (Shapiro delay)	2003	2×10^{-5}	$\gamma = 1$
Gravity Probe B	2011	0.3%	Confirms frame-dragging
LIGO/Virgo (GW detection)	2015–	—	GWs exist
GW170817 (multi-messenger)	2017	few $\times 10^{-15}$ (under standard emission-time assumptions)	$c_\gamma = c_g$
Optical clocks (JILA)	2022	10^{-18}	Confirms redshift

Each entry represents a genuine experimental achievement. The question is not whether these experiments were performed correctly, but what they actually test.

1.2 The Central Question

This paper asks: What do precision tests of relativity actually constrain?

The analysis identifies five structural limitations shared by the experimental canon:

1. *Two-Way Measurement Dominance*: Nearly all tests use round-trip or reciprocity-even measurements, which are mathematically blind to direction-dependent effects.
2. *Local/Global Conflation*: Tests of local Lorentz invariance are extrapolated to claims about global spacetime structure.
3. *Model-Dependent Calibration*: Systems like GPS use GR-derived corrections to validate GR—demonstrating self-consistency within the framework, not independent confirmation.
4. *The Conformal Loophole*: Multi-messenger constraints bound the disformal sector of scalar-tensor theories while leaving the conformal sector unconstrained.
5. *Theory-Laden Data Reduction*: Systematic corrections assume the framework being tested, making independent falsification practically difficult within standard pipelines.

These limitations do not invalidate the experiments. They indicate that, in many cases, the tests primarily constrain parameter space within an assumed framework rather than systematically discriminating between alternatives.

The analysis throughout assumes a regime where disformal cone tilts are already constrained by multi-messenger bounds, and focuses on clock-sector structure that remains unconstrained within those limits.

Scope of Claim

This manuscript does not dispute the validity of canonical precision tests within their modeled domains. It argues that many flagship tests primarily constrain reciprocity-even (two-way / closed-path) observables and therefore may not discriminate GR from a class of two-metric disformal scalar-tensor frameworks that reproduce local redshift and round-trip light-time relations. Discrimination requires observables that are offset-invariant yet one-way and direction-reversing, or that probe spatial clock-correlation structure.

1.3 The Temporal Equivalence Principle: Theory Definition

The Temporal Equivalence Principle (TEP) is defined formally by a single manifold endowed with two metrics and a universal matter coupling. It is not an ad hoc modification but a systematic framework for dynamical time.

Formal Definition

Sector map: The conformal factor $A(\phi)$ governs clock rates and spatial correlation structure; the disformal factor $B(\phi)$ governs light-cone tilts and residual holonomy. GW170817 constrains the disformal sector; the conformal sector remains unconstrained by single-path multi-messenger observations.

1. The Action: The theory is defined in the Einstein frame ($g_{\mu\nu}$) by:

$$S = \int d^4x \sqrt{-g} \left[\frac{M_{Pl}^2}{2} R - \frac{1}{2} (\partial\phi)^2 - V(\phi) \right] + S_{matter}[\psi, \tilde{g}_{\mu\nu}]$$

2. The Metrics: Gravity is governed by $g_{\mu\nu}$. All matter fields ψ and clocks couple universally to the *matter metric* $\tilde{g}_{\mu\nu}$, related by a disformal map:

$$\tilde{g}_{\mu\nu} = A(\phi)g_{\mu\nu} + B(\phi)\partial_\mu\phi\partial_\nu\phi$$

3. The Coupling:

- **Clocks:** Ideal clocks measure proper time τ defined by $d\tau^2 = -\tilde{g}_{\mu\nu}dx^\mu dx^\nu$.
- **Light:** Photons propagate on null geodesics of $\tilde{g}_{\mu\nu}$. (In the conformal limit $B = 0$, these coincide with $g_{\mu\nu}$ null geodesics).
- **Test Particles:** Massive bodies follow geodesics of $\tilde{g}_{\mu\nu}$.

The key distinguishing predictions are therefore not local redshift or two-way light-time relations (where TEP can agree with GR), but global and statistical observables that probe synchronization integrability and spatial structure:

- Local physics is identical to GR in freely falling laboratories (EEP; locally invariant c)
- Clock residuals can exhibit spatially structured correlations (e.g., exponential decay with distance) if the scalar field has a finite correlation length
- A convention-independent residual synchronization holonomy H_{resid} can be defined for direction-reversing, one-way closed loops after subtracting modeled GR loop effects
- Light and gravitational-wave propagation remain consistent with existing constraints: conformal rescaling preserves null cones, and any disformal cone tilt is bounded at the 10^{-15} level

The existing experimental canon strongly constrains local Lorentz violation, PPN departures in the gravitational/light-propagation sector, and disformal cone tilts. It does not yet directly probe spatial clock-correlation structure or residual holonomy in genuinely one-way, direction-reversing closed loops.

Candidate Discriminators and Current Status

To ensure falsifiability, TEP specifies candidate observables where it can differ from GR. Where numerical values are quoted from the TEP research program, they are treated as reported results requiring independent replication.

Observable	GR Prediction	TEP Prediction	Status
Distance-structured clock correlations	None (or systematic origin)	$\lambda = 1,000\text{--}10,000 \text{ km}$	Suggested in exploratory GNSS analyses (order $10^3\text{--}10^4 \text{ km scale}$) within the TEP program; requires independent, blinded replication using raw data
Orbital velocity coupling	No correlation	$ r > 0.5$ with Earth's orbital velocity	Suggested in exploratory GNSS analyses within the TEP program; requires independent, blinded replication
CMB frame alignment	No preferred frame	$< 30^\circ$ from CMB dipole	Suggested in exploratory GNSS analyses within the TEP program; requires independent, blinded replication
Residual holonomy H_{resid}	0	0 if $B = 0$; $0.5\text{--}50$ as if $B \neq 0$ (MEO Triangle)	Untested (requires one-way closed loops)

Observable	GR Prediction	TEP Prediction	Status
Multi-constellation consistency	N/A	CV < 30% across GPS/GLONASS/Galileo/BeiDou	Untested (requires raw data analysis)

Critical note: The first three observables target conformal-sector structure associated with $A(\phi)$, while the holonomy test targets disformal structure associated with $B(\phi)$. The core argument of this paper concerns measurement geometry and does not assume that any specific GNSS-reported numerical value is correct.

1.4 Structure of This Paper

Section 2 develops the methodological framework for analyzing precision tests, introducing key distinctions between gauge-invariant and convention-dependent observables, two-way and one-way measurements, and local and global tests.

Sections 3–7 examine the major categories of precision tests:

- Section 3: Gravitational redshift tests (Pound-Rebka, optical clocks)
- Section 4: Time dilation tests (Hafele-Keating, GPS)
- Section 5: Light propagation tests (Shapiro delay, VLBI)
- Section 6: Multi-messenger constraints (GW170817)
- Section 7: Resonator tests (Michelson-Morley, cavity experiments)

Section 8 proposes experiments that could genuinely discriminate between GR and TEP: triangle holonomy tests, interplanetary closed-loop timing, and GNSS correlation replication.

Section 9 discusses the philosophical implications, situating the findings within the broader context of underdetermination in physics. Section 10 concludes.

1.5 Related Work

This analysis builds upon and extends several distinct strands of literature in gravitational physics, metrology, and the philosophy of science.

Test Theories of Relativity: Frameworks for testing Lorentz invariance, such as the Mansouri-Sexl (1977) formalism and the Standard Model Extension (SME) (Colladay & Kostelecký 1998), have long parameterized deviations from relativity. These frameworks primarily focus on local Lorentz violation (anisotropy, boost dependence) in single-metric contexts. TEP differs by accepting exact local Lorentz invariance (consistent with the tightest SME bounds) while introducing non-integrable synchronization structure via a second metric.

Relativistic Metrology: The distinction between one-way and two-way time transfer is foundational in metrology (Ashby 2003). While two-way satellite time and frequency transfer (TWSTFT) and GPS Common-View are standard techniques for clock comparison, they are explicitly designed to cancel path delays. Recent proposals for "relativistic geodesy" with optical clocks (Mehlstäubler et al. 2018) focus on measuring the geopotential W , assuming GR. This work explores the inverse problem: using clocks to test the framework that defines W .

Philosophy of Simultaneity: The conventionality of simultaneity (Reichenbach 1928; Grünbaum 1973; Salmon 1977) establishes that one-way light speed is a convention in special relativity. Malament (1977) argued that causal connectivity makes standard synchronization unique, but this uniqueness holds only if the causal structure is singular (one metric). This paper operationalizes these philosophical insights in the context of two-metric theories, where the interplay between causal (light) and temporal (clock) structures makes the "convention" physically testable via holonomy.

Theory-Laden Experimentation: The challenge of "epistemological circularity" in precision tests—where data reduction relies on the theory being tested—has been discussed by Franklin (1986) and others. This analysis extends that discussion to modern precision cosmology and GNSS data processing, identifying specific pipelines (e.g., common-mode filtering) that enforce theoretical priors.

To forestall misunderstanding, this paper does not claim:

- That GR is wrong
- That precision tests are flawed or incompetent
- That TEP is correct
- That all alternatives to GR are viable

The claim is narrower and more specific: the existing experimental canon does not directly probe the observables that distinguish GR from TEP. This underdetermination is resolvable through new experimental configurations that have not yet been performed.

2. Methodological Framework

Before examining individual experiments, it is essential to establish the conceptual tools required to distinguish what precision tests actually measure from what they are commonly claimed to measure. This section develops four key distinctions that structure the subsequent analysis.

2.1 Gauge-Invariant vs. Convention-Dependent Observables

Not all physical quantities carry the same epistemological weight. A fundamental distinction exists between gauge-invariant observables—quantities that remain unchanged under coordinate transformations—and convention-dependent quantities that depend on arbitrary choices.

Definition 2.1: Gauge-Invariant Observable

An observable O is gauge-invariant if and only if its value is independent of the coordinate system, synchronization convention, or reference frame used to compute it.

Examples of gauge-invariant observables include:

- Proper time elapsed along a worldline: $\tau = \int ds$
- Round-trip light travel time between two events
- Frequency ratio of co-located clocks
- Interference fringe count in closed optical paths

Examples of convention-dependent quantities include:

- One-way speed of light (depends on synchronization)
- Coordinate time difference between spatially separated events
- Simultaneity of distant events

The critical insight is that most precision tests measure gauge-invariant quantities, which by construction do not directly probe the specific loop and correlation observables that differ between GR and TEP in the regime considered here.

The central claim of this paper is that most high-precision tests to date access only the gauge-invariant subset of observables on which GR and TEP explicitly agree.

2.2 Two-Way vs. One-Way Measurements

The distinction between two-way (round-trip) and one-way measurements is not merely technical but fundamental to what can be tested.

The Reichenbach Synchronization Problem

The one-way speed of light cannot be measured without first establishing a synchronization convention between spatially separated clocks. But any synchronization convention presupposes a one-way speed of light. This circularity means the one-way speed is convention-dependent, not empirically determinable.

Two-way measurements avoid this circularity by using a single clock:

$$t_{\text{round-trip}} = t_{\text{return}} - t_{\text{emit}} \quad (\text{single clock, gauge-invariant})$$

However, this gauge-invariance comes at a cost: two-way measurements are inherently insensitive to direction-dependent effects. If light travels at $c+v$ in one direction and $c-v$ in the return direction, the round-trip time is:

$$t_{RT} = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2Lc}{c^2-v^2} \approx \frac{2L}{c} + O\left(\frac{v^2}{c^2}\right)$$

The first-order direction-dependent term cancels exactly. This is not a limitation of experimental precision but a mathematical necessity: round-trip measurements are reciprocity-even and cannot probe reciprocity-odd effects.

2.3 Local vs. Global Tests

A second crucial distinction separates local tests—performed within a single freely falling frame—from global tests that probe the structure of spacetime across extended regions.

The Equivalence Principle Hierarchy

- *Weak Equivalence Principle (WEP)*: All bodies fall with the same acceleration in a gravitational field.
- *Einstein Equivalence Principle (EEP)*: In local freely falling frames, non-gravitational physics is Lorentz-invariant.
- *Strong Equivalence Principle (SEP)*: EEP extends to gravitational physics; local gravitational experiments are position-independent.

The TEP framework satisfies EEP exactly: in local freely falling frames, physics reduces to special relativity with invariant c . This means local tests—no matter how precise—cannot distinguish TEP from GR. The distinction arises in global and statistical observables: spatial structure in clock residuals and a GR-subtracted residual holonomy H_{resid} defined by genuinely one-way, direction-reversing closed loops.

The operational criterion for "local" is whether the measurement region is small compared to the curvature scale. Modern optical clock experiments operate at millimeter scales where spacetime curvature is negligible—they are quintessentially local tests.

2.4 Single-Path vs. Multipath Configurations

The final distinction concerns whether signals traverse a single path or multiple paths through spacetime.

Common-Mode Cancellation

When two signals traverse the same path, any path-dependent effects cancel in their comparison. This is why GW170817—where photons and gravitons traveled the same ~130 million light-year path—constrains only their differential propagation speed, not their common-mode time dilation.

TEP phenomenology arises from:

- Multipath configurations (gravitational lensing, where rays sample different field values)
- Statistical correlations (GNSS, where spatial structure $\langle \phi(x)\phi(x') \rangle$ is probed)
- Closed loops with direction reversal (residual holonomy H_{resid} after subtracting modeled GR loop effects)

Single-path, single-direction measurements—the dominant mode of precision tests—are structurally incapable of detecting these signatures.

2.5 Summary: The Measurement Taxonomy

Measurement Type	Gauge Status	Sensitive To	Blind To
Two-way, local	Invariant	Local Lorentz violation	Direction-dependence, global structure
Two-way, global	Invariant	PPN parameters (γ, β)	Odd-parity effects, synchronization
One-way, single-path	Convention-dependent	Relative clock rates	Absolute synchronization
One-way, closed-loop	Invariant	Residual holonomy H_{resid} , path-dependence	—

The final row—one-way closed-loop measurements—represents the only configuration capable of testing synchronization integrability. No high-precision relativistic-gravity test has directly targeted a convention-independent, direction-reversing, one-way loop observable designed to detect non-integrable synchronization beyond modeled GR terms.

3. The Gravitational Redshift Tests

Gravitational redshift experiments—from Pound-Rebka (1960) to modern optical lattice clocks—are often cited as the most precise confirmations of general relativity. This section examines what these experiments actually test and why they do not directly probe the observables that distinguish GR from TEP.

3.1 The Canonical Experiments

3.1.1 Pound-Rebka-Snider (1960-1965)

The Pound-Rebka experiment measured the gravitational redshift of 14.4 keV gamma rays traveling 22.5 meters vertically in the Jefferson Tower at Harvard. The observed fractional frequency shift:

$$\frac{\Delta\nu}{\nu} = \frac{gh}{c^2} = \frac{(9.8 \text{ m/s}^2)(22.5 \text{ m})}{(3 \times 10^8 \text{ m/s})^2} \approx 2.5 \times 10^{-15}$$

agreed with the GR prediction to within 1% (later improved to 0.1% by Pound and Snider).

Methodological Analysis

The experiment compares the frequency of gamma rays emitted at height h_1 with those absorbed at height h_2 . Both emission and absorption occur locally—the comparison is between two local measurements connected by light propagation.

TEP predicts identical local physics in freely falling frames. The gravitational redshift arises from the difference in gravitational potential between emission and absorption points—a prediction shared by any metric theory satisfying EEP.

3.1.2 Gravity Probe A (1976)

Gravity Probe A flew a hydrogen maser to 10,000 km altitude, comparing its frequency with ground-based masers. The experiment confirmed gravitational time dilation to 7×10^{-5} relative precision.

Methodological Analysis

The comparison used two-way microwave links: signals were sent up to the spacecraft and transponded back. The round-trip nature of the measurement makes it insensitive to any direction-dependent propagation effects.

The experiment confirms that clocks at different gravitational potentials run at different rates—a prediction common to GR, TEP, and all metric theories of gravity.

3.1.3 Modern Optical Lattice Clocks (2010-present)

Jun Ye's group at JILA has achieved gravitational redshift measurements at the 10^{-18} level and below, detecting time dilation across millimeter height differences. These represent the most precise measurements of gravitational effects ever performed.

"We greatly benefit from rejection of common-mode systematic shifts thanks to ensembles sharing the same optical lattice."

— Bothwell et al. (2022)

Methodological Analysis

The quoted statement reveals the experimental configuration: clock ensembles share the same optical lattice, enabling common-mode rejection. This is a local comparison within a single laboratory frame.

The measurement scale (millimeters) is vastly smaller than any curvature scale. These are quintessentially local tests of the equivalence principle, not probes of global spacetime structure.

Optical clock metrology also includes long-baseline comparisons (for example via phase-stabilized fiber links) used for relativistic geodesy and time-scale realization. These experiments extend the baseline, but they still primarily target reciprocity-even observables such as frequency ratios and potential differences, and they typically rely on link designs and models that suppress non-reciprocal path terms.

As a result, even when baselines extend to regional or continental scales, the dominant observable remains a redshift-type comparison rather than a direction-reversing closed-loop residual holonomy H_{resid} or a deliberately constructed spatial correlation statistic.

TEP Axiom 2 states: "In local freely falling frames, physics reduces exactly to special relativity with invariant c." These experiments confirm this axiom—they cannot test whether synchronization is globally integrable.

3.2 What Redshift Tests Actually Measure

All gravitational redshift experiments measure the same fundamental quantity: the ratio of proper times at different gravitational potentials.

$$\frac{d\tau_1}{d\tau_2} = \sqrt{\frac{g_{00}(x_1)}{g_{00}(x_2)}} \approx 1 + \frac{\Phi_1 - \Phi_2}{c^2}$$

This ratio is:

- Gauge-invariant (independent of coordinate choice)
- Local (compares clocks at specific spacetime points)
- Common to all metric theories satisfying EEP

The experiments do not measure:

- Absolute synchronization between distant clocks
- Path-dependence of clock transport
- Global integrability of the synchronization field

3.3 The ACES Mission: A Missed Opportunity?

The Atomic Clock Ensemble in Space (ACES), scheduled for the International Space Station, will compare space-based and ground-based atomic clocks with unprecedented precision. The mission aims to test gravitational redshift at the $2-3 \times 10^{-6}$ level (Savalle et al. 2019).

Configuration Analysis

ACES uses microwave and optical links for clock comparison. The microwave link (MWL) employs two-way time transfer—signals travel up and down, with the comparison derived from the round-trip.

While ACES will achieve remarkable precision, its two-way configuration makes it structurally incapable of detecting synchronization holonomy. The experiment will confirm (or constrain) the gravitational redshift formula but cannot test whether synchronization is path-dependent.

3.4 Summary: Local Brilliance, Global Blindness

What Gravitational Redshift Tests Confirm

- Clocks at different gravitational potentials run at different rates
- The rate difference follows $d\tau/dt = \sqrt{1 + 2\Phi/c^2}$ to extraordinary precision
- Local Lorentz invariance holds at the 10^{-18} level
- The Einstein Equivalence Principle is satisfied

What They Cannot Test

- Whether synchronization is globally integrable
- Whether clock transport around closed loops accumulates holonomy
- Whether the one-way speed of light is direction-independent
- Any prediction that distinguishes GR from TEP

The gravitational redshift tests are triumphs of experimental physics. They confirm that gravity affects time—a revolutionary insight that transformed our understanding of the universe. These experiments were not designed to probe global synchronization structure,

and it would be unreasonable to criticize them for not testing something outside their scope. The point is simply that theories agreeing on local physics (like GR and TEP) do not directly probe the specific loop and correlation observables that differ between them in the regime considered here.

The proposed loop and correlation experiments should therefore be seen as complementary additions to, not replacements for, the existing experimental canon.

4. The Time Dilation Tests

Time dilation experiments—from Hafele-Keating's circumnavigating clocks to GPS satellite corrections—are frequently cited as strong evidence that "time runs differently" in different reference frames. This section examines the logical structure of these claims.

4.1 The Hafele-Keating Experiment (1971)

In October 1971, Joseph Hafele and Richard Keating flew cesium atomic clocks on commercial aircraft around the world—eastward and westward—comparing them with reference clocks at the U.S. Naval Observatory. The observed time differences:

Direction	Predicted (ns)	Observed (ns)
Eastward	-40 ± 23	-59 ± 10
Westward	$+275 \pm 21$	$+273 \pm 7$

The results are celebrated as confirming both special relativistic time dilation (velocity-dependent) and general relativistic gravitational time dilation (altitude-dependent).

Methodological Analysis

The experiment measures the proper time difference $\Delta\tau$ between clocks that separate and reunite. This is a gauge-invariant quantity—the "twin paradox" resolved by comparing clocks at the same spacetime point.

Crucially, the clocks return to their starting point. The measurement is of accumulated proper time around a closed loop, not of synchronization between spatially separated clocks.

TEP predicts identical proper time accumulation for any closed worldline. The Hafele-Keating result confirms that proper time depends on the worldline traversed—a prediction shared by GR, TEP, and all metric theories.

The Closed-Loop Limitation

When clocks reunite, their comparison is unambiguous. But this very feature—the closure of the loop—eliminates sensitivity to synchronization holonomy. The experimentally relevant discriminator is a GR-subtracted residual holonomy H_{resid} , defined by genuinely one-way, direction-reversing closed loops after subtracting modeled GR loop effects (Sagnac, Shapiro, gravitomagnetic terms). Closed-worldline proper-time comparisons do not access H_{resid} .

Hafele-Keating confirms: "Different worldlines accumulate different proper times."

Hafele-Keating cannot test: "Is synchronization of distant clocks path-independent?"

4.2 Muon Lifetime Experiments

Cosmic ray muons, traveling at relativistic speeds, exhibit extended lifetimes compared to muons at rest. The observed lifetime dilation factor $\gamma = 1/\sqrt{1 - v^2/c^2}$ agrees with special relativity to high precision.

Methodological Analysis

Muon lifetime measurements compare the decay rate of moving muons (in the lab frame) with the known rest-frame lifetime. This is a comparison of proper times along different worldlines—the muon's worldline versus a hypothetical stationary worldline.

The measurement confirms time dilation as a function of velocity. It does not involve synchronization of spatially separated clocks and cannot probe synchronization structure.

4.3 The Galileo Eccentric Orbit Test (2018)

In 2014, two Galileo satellites (GSAT-0201 and GSAT-0202) were accidentally launched into highly eccentric orbits instead of circular ones. This "fortunate accident" enabled the most precise test of gravitational redshift since Gravity Probe A in 1976.

The eccentric orbits (perigee ~17,180 km, apogee ~26,020 km) produce a periodic modulation of the gravitational potential experienced by the onboard hydrogen masers. Delva et al. (2018) and Herrmann et al. (2018) independently analyzed 1008 days of data, measuring:

$$\alpha = (+0.19 \pm 2.48) \times 10^{-5}$$

where α parameterizes deviations from the GR prediction ($\alpha = 0$ in GR). This represents a factor 5.6 improvement over Gravity Probe A.

Methodological Analysis

The Galileo test measures the periodic variation in clock frequency as the satellite moves between perigee and apogee. This is a comparison of the same clock at different gravitational potentials—a local measurement of how clock rate depends on potential.

The measurement uses two-way time transfer via GNSS signals. The satellite clock is compared with ground clocks through standard GNSS processing, which assumes GR-based corrections for signal propagation.

TEP predicts identical local clock rates to GR. The gravitational redshift formula $d\tau/dt = \sqrt{1 + 2\Phi/c^2}$ is shared by both theories. The Galileo test confirms this formula with unprecedented precision but cannot probe synchronization structure.

What the Galileo Test Cannot Probe

The test measures how a single clock's frequency varies with gravitational potential. It does not test whether synchronization between spatially separated clocks is path-independent. The periodic signal is entirely predicted by both GR and TEP.

4.4 Gravity Probe B (2011)

Gravity Probe B measured two predictions of general relativity: the geodetic effect (spacetime curvature caused by Earth's mass) and frame-dragging (spacetime being "dragged" by Earth's rotation). Four ultra-precise gyroscopes orbited Earth for 18 months, measuring their precession relative to a distant guide star.

The results confirmed GR predictions:

Effect	GR Prediction	Measured	Precision
Geodetic precession	-6606.1 mas/yr	-6601.8 ± 18.3 mas/yr	0.28%
Frame-dragging	-39.2 mas/yr	-37.2 ± 7.2 mas/yr	19%

Methodological Analysis

Gravity Probe B measures gyroscope precession—the rotation of a spin axis relative to distant stars. This is a local measurement: the gyroscope and its reference telescope occupy the same spacecraft, comparing local inertial frames with the distant stellar reference.

The geodetic effect arises from spacetime curvature; frame-dragging from the rotation of the central mass. Both are predictions of how local inertial frames precess relative to infinity—not tests of synchronization structure.

TEP preserves local Lorentz invariance exactly. Gyroscope precession depends on the local connection coefficients (Christoffel symbols), which are identical in GR and TEP. The experiment confirms that spacetime is curved as GR predicts; it cannot test whether synchronization is globally integrable.

The Local Frame Limitation

Gravity Probe B compares a local gyroscope with a distant reference direction. The comparison involves light from the guide star reaching the telescope—a single-path measurement that cannot probe synchronization holonomy.

To test TEP, one would need to compare gyroscopes at different locations and ask whether their precessions are consistent with a globally integrable synchronization. GP-B's single-satellite design cannot perform this test.

4.5 The GPS System: Model-Dependent Calibration

The Global Positioning System is frequently cited as the most practical demonstration of relativistic effects. GPS satellites carry atomic clocks that are pre-corrected for both special relativistic (velocity) and general relativistic (gravitational) time dilation. This section examines the logical structure of GPS as a test of GR.

4.5.1 The Standard Narrative

GPS satellites orbit at ~20,200 km altitude with velocity ~3.9 km/s. The combined relativistic correction is approximately +38 $\mu\text{s/day}$ (gravitational) minus 7 $\mu\text{s/day}$ (velocity), yielding a net +31 $\mu\text{s/day}$ correction applied to satellite clocks.

The argument proceeds: GPS works at meter-level accuracy; therefore, the relativistic corrections must be correct; therefore, GR is confirmed.

4.5.2 Model-Dependent Calibration

The Logical Structure: Sufficiency vs. Uniqueness

The engineering success of GPS is a powerful result, but its epistemological status requires precise characterization:

- **Sufficiency:** GPS demonstrates that the GR framework is *sufficient* to maintain synchronization at the nanosecond level. The corrections derived from GR work.
- **Internal Consistency:** The system validates the self-consistency of the GR model. When clocks are corrected according to GR, the navigation solutions close.
- **Non-Uniqueness:** This success does not mathematically prove that GR is the *only* theory capable of producing these results. Any theory that agrees with GR on local clock rates (the source of the corrections) and light propagation (for two-way/common-view transfer) would yield an identical engineering system.

Crucially, GNSS data *can* constrain deviations from GR, but only if analyzed with alternative models and, most importantly, with access to raw observables before standard relativistic corrections are baked into the data products.

Sufficiency vs. Necessity

GPS proves that GR is sufficient for navigation, but not that it is necessary.

Any theory that predicts the same local clock rates as GR would produce identical corrections and identical navigation performance. GPS cannot distinguish between theories that agree on local clock behavior but differ on global synchronization structure.

"These clocks read the coordinate time t ... would be self-consistently synchronized if one brought them together—assuming that general relativity is correct."

— Ashby (2003), Living Reviews in Relativity

The phrase "assuming that general relativity is correct" is not a criticism of GPS engineering—it is a precise statement of the logical structure. GPS demonstrates that GR provides a self-consistent framework for navigation. It does not demonstrate that GR is the unique framework capable of doing so.

4.5.3 What a TEP Signal Would Look Like in GPS Residuals

If TEP effects exist, they would appear in GPS data as specific systematic patterns. Understanding these signatures is essential for determining whether GPS could detect or has already filtered out such effects.

TEP Signatures in GPS Residuals			
TEP Effect	GPS Manifestation	Standard Interpretation	Distinguishing Feature
Distance-structured correlations	Correlated clock residuals between nearby stations	"Common-mode error" (filtered out)	Exponential decay with distance (~4000 km scale)
Residual holonomy H_{resid}	Systematic position bias in closed-loop solutions	"Network adjustment error"	Direction-dependent, not random
Scalar field gradient	Altitude-dependent clock drift beyond GR	"Satellite clock instability"	Correlated across satellites at similar altitudes
Temporal coherence	Day-to-day correlation in residuals	"Environmental systematic"	Persists after environmental correction

4.5.4 Reported Evidence from GNSS Analysis

Exploratory analysis of GNSS clock products from multiple analysis centers (CODE, IGS, ESA) within the TEP research program has suggested systematic patterns consistent with TEP predictions:

- **Distance-structured correlations:** Clock residuals reportedly show exponential decay with inter-station distance, with a characteristic scale of order 10^3 – 10^4 km
- **Cross-center consistency:** The same pattern reportedly appears in independent processing pipelines ($R^2 = 0.92$ – 0.97 between centers)
- **Temporal persistence:** The correlation structure reportedly persists over 25 years of data (2000–2025)
- **Ionospheric independence:** The signal is reportedly invariant under geomagnetic storm stratification ($K_p < 3$ vs $K_p \geq 3$)

These reported observations require independent verification. If confirmed, they would demonstrate that GPS/GNSS data contains systematic structure that standard processing treats as "error" rather than "signal." If refuted, they would constrain the TEP parameter space. Either outcome advances scientific understanding.

4.5.5 The Filtering Problem

What GPS Processing Removes

Many high-precision GNSS processing pipelines can suppress spatially correlated residual structure through:

- **Reference frame constraints:** Network solutions are constrained to minimize global residuals
- **Common-mode filtering:** Correlated signals across stations are removed as "systematic errors" (see Dong et al. 2006)
- **Clock modeling:** Polynomial models absorb slow drifts (Kouba & Héroux 2001)
- **Ambiguity resolution:** Integer constraints enforce consistency

If TEP produces distance-structured clock correlations, standard processing would tend to suppress such structure as a systematic effect rather than preserve it as a candidate signal.

Concrete Example: How TEP Signals Are Filtered Out

Consider, hypothetically, two GNSS stations separated by ~3,000 km. Suppose a weak, distance-structured correlated component exists in the underlying observables after standard GR modeling.

What happens in standard GPS processing:

1. **Initial observation:** Both stations show residual structure with partial coherence at the same epoch

2. **Network constraint:** The processing software enforces that the global network residuals sum to zero (to maintain the reference frame)
3. **Common-mode removal:** The correlated signal is identified as a "common-mode error" affecting multiple stations
4. **Filtering:** The signal is removed by subtracting the network-wide mean or applying spatial filtering (e.g., PCA, Karhunen-Loëve expansion; Dong et al. 2006)
5. **Final product:** Clock residuals are decorrelated, and the TEP signature is suppressed to below detection threshold

Result: Any spatially coherent component can be classified as "systematic error" and partially removed. The extent of suppression is product- and analysis-center dependent.

This is why independent analysis of *raw* GNSS data (before common-mode filtering) is critical for testing TEP. Many standard GNSS products prioritize reference-frame stability and positioning performance, and are not designed to preserve subtle spatial-correlation structure for fundamental-physics inference.

4.5.6 The TEP Perspective

TEP predicts identical local clock rates to GR. The frequency corrections applied to GPS satellites would be the same under TEP. The difference arises in synchronization structure—whether the coordinate time t is globally integrable.

GPS uses predominantly two-way time transfer and common-view methods for synchronization. These methods are insensitive to residual holonomy H_{resid} . A nonzero H_{resid} beyond GR subtraction would manifest as systematic position errors with specific geometric signatures—but such signatures would be absorbed into the navigation solution as "common-mode errors" and filtered out.

What Would Falsify GR via GPS?

For GPS to genuinely discriminate between GR and TEP, one would need to:

1. Analyze raw clock residuals before common-mode filtering
2. Test for distance-structured correlations (TEP predicts exponential decay)
3. Perform closed-loop one-way timing tests (not two-way or common-view)
4. Compare residual patterns with TEP and GR predictions

Exploratory analysis of GNSS clock products has suggested patterns consistent with TEP predictions. The question is whether these signatures represent physics or unmodeled systematics—a question that requires independent, blinded validation.

4.6 Lunar Laser Ranging: The Nordtvedt Effect

Lunar Laser Ranging (LLR) has measured the Earth-Moon distance to millimeter precision since 1969, using retroreflectors placed by Apollo astronauts. LLR provides the most precise test of the Strong Equivalence Principle through the Nordtvedt effect.

4.6.1 The Nordtvedt Effect

If gravitational self-energy contributes differently to inertial and gravitational mass, the Earth and Moon would fall toward the Sun at slightly different rates, causing a polarization of the lunar orbit. The Nordtvedt parameter η quantifies this violation:

$$\eta = 4\beta - \gamma - 3 - \frac{10}{3}\xi - \alpha_1 + \frac{2}{3}\alpha_2$$

Current LLR solutions give $(m_G/m_I)_E - (m_G/m_I)_M = (-0.8 \pm 1.3) \times 10^{-13}$, implying $|\eta| \leq \text{few} \times 10^{-4}$ (conversion depends on the Earth–Moon self-energy difference; Williams et al. 2012).

Methodological Analysis

LLR uses round-trip laser ranging: pulses travel from Earth to Moon and back. The measurement is inherently two-way, averaging over any direction-dependent propagation effects. The Nordtvedt test constrains differential free-fall of the Earth and Moon toward the Sun, and therefore places strong bounds on violations of the Strong Equivalence Principle and on any additional long-range fields that couple differently to self-gravitating bodies.

In two-metric frameworks where gravitational dynamics remain effectively GR-like in the solar-system regime, consistency with LLR requires that any additional scalar degree of freedom yield sufficiently small effective scalar charge for the Earth and Moon (for example via screening or parameter suppression). LLR therefore acts as a stringent constraint on the allowed

coupling and screening regime, but it does not directly target the one-way loop observables or spatial clock-correlation statistics emphasized in this paper.

LLR confirms that gravity is universal for self-gravitating bodies. It does not test clock synchronization structure—the measurement involves no clocks on the Moon, only photon round-trip times.

Why LLR Does Not Directly Probe TEP Signatures

LLR tests the gravitational metric $g_{\mu\nu}$ through orbital dynamics. TEP modifies the matter metric $\tilde{g}_{\mu\nu}$ through conformal and (constrained) disformal structure. Since:

- Orbital dynamics depend on $g_{\mu\nu}$ (unchanged in TEP)
- In the conformal-only limit, photon null cones coincide with those of $g_{\mu\nu}$
- The measurement is two-way (direction-dependent effects cancel)

LLR does not directly probe the clock-sector and synchronization-structure observables that distinguish GR from TEP. The proposed loop and correlation tests should be seen as complementary additions to, not replacements for, LLR and other precision tests.

4.7 Summary: Closed Loops That Don't Test Closure

What Time Dilation Tests Confirm

- Proper time depends on the worldline traversed
- Moving clocks run slow by factor $\gamma = 1/\sqrt{1 - v^2/c^2}$
- Clocks at higher gravitational potential run faster
- GR provides a self-consistent framework for timekeeping

What They Do Not Directly Probe

- Whether synchronization is path-independent
- Whether one-way light speed is isotropic
- Whether coordinate time is globally integrable
- The specific loop and correlation observables that distinguish GR from TEP

The time dilation tests confirm that time is relative—a profound insight. But they confirm it through proper time comparisons along reuniting worldlines, not through synchronization of permanently separated clocks. The distinction is subtle but fundamental: proper time path-dependence is universal; synchronization path-dependence is what distinguishes theories.

5. The Light Propagation Tests

Light propagation experiments—Shapiro delay measurements, VLBI observations, and gravitational lensing studies—probe how electromagnetic radiation traverses curved spacetime. These tests have achieved remarkable precision in constraining the Parameterized Post-Newtonian (PPN) parameter γ . This section examines what these constraints actually imply.

5.1 The Shapiro Delay

In 1964, Irwin Shapiro predicted that radar signals passing near the Sun would experience a time delay due to spacetime curvature. The "fourth test of GR" has since been measured with increasing precision.

5.1.1 The Cassini Experiment (2003)

The most precise Shapiro delay measurement used radio signals between Earth and the Cassini spacecraft during solar conjunction. The result:

$$\gamma = 1 + (2.1 \pm 2.3) \times 10^{-5}$$

This constrains the PPN parameter γ to agree with GR ($\gamma = 1$) at the 10^{-5} level.

Methodological Analysis

The Cassini measurement used round-trip radio ranging: signals were transmitted from Earth to the spacecraft and transponded back. The measured quantity is the round-trip light time, not the one-way propagation time.

Round-trip measurements are inherently reciprocity-even. If the one-way speed of light differs in opposite directions ($c + \delta$ outbound, $c - \delta$ inbound), the round-trip time is:

$$t_{RT} = \frac{L}{c+\delta} + \frac{L}{c-\delta} = \frac{2Lc}{c^2-\delta^2} \approx \frac{2L}{c} + O\left(\frac{\delta^2}{c^2}\right)$$

The first-order direction-dependent term cancels exactly. Cassini constrains reciprocity-even effects (the PPN γ parameter) but is blind to reciprocity-odd effects (direction-dependent propagation).

5.1.2 Data Reduction Concerns

The Cassini analysis required extensive corrections for systematic effects:

"Doppler calibration for Earth's troposphere... potential major error source. We multiply y by downlink X-Band frequency... make new file for further processing."

— Bertotti, Iess, and Tortora (2003)

The Cassini inference is conditional on an end-to-end modeling stack: a spacetime model for signal propagation in the solar system, together with tropospheric, plasma, and spacecraft dynamics models. This does not weaken the measurement; it clarifies what is being tested. The result tightly constrains γ within the assumed signal model class.

5.2 Very Long Baseline Interferometry (VLBI)

VLBI observations of quasars near the Sun measure the apparent deflection of radio waves by solar gravity. The deflection angle:

$$\theta = (1 + \gamma) \times \frac{2GM}{c^2 b}$$

where b is the impact parameter, has been measured to constrain γ at the 10^{-4} level.

Methodological Analysis

VLBI measures the differential arrival time of radio waves at widely separated antennas. The measurement is sensitive to the angular position of the source, which shifts due to gravitational light bending.

Like Shapiro delay, this constrains the conformal factor in the metric (how spatial distances are modified by gravity) but not the synchronization structure. The PPN γ parameter appears in the spatial part of the metric; synchronization holonomy arises from the temporal part.

5.3 The PPN Framework and Its Limitations

The Parameterized Post-Newtonian formalism provides a systematic way to compare metric theories of gravity. The PPN metric in isotropic coordinates:

$$ds^2 = -(1 - 2U + 2\beta U^2)c^2 dt^2 + (1 + 2\gamma U)(dx^2 + dy^2 + dz^2)$$

where $U = GM/rc^2$ is the Newtonian potential. GR predicts $\beta = \gamma = 1$.

5.3.1 What PPN Tests Constrain

- γ : How much space curvature is produced by unit rest mass
- β : How much nonlinearity exists in the superposition of gravity
- *Other parameters*: Preferred-frame effects, conservation law violations

5.3.2 What PPN Tests Cannot Constrain

The PPN framework parameterizes the post-Newtonian limit under the assumption that the same effective metric governs the sector being tested (typically solar-system dynamics and light propagation). It is therefore not a complete language for theories in which additional structure appears in the matter and clock sector while remaining effectively GR-like in the photon and gravitational sectors.

- Clock transport and synchronization are governed by an effective metric distinct from the one constrained by light-propagation observables
- Convention-independent loop observables (residual holonomy after GR subtraction) are the primary targets
- Non-local statistical structure appears in clock residuals (spatial correlations) rather than in PPN light-bending parameters

In TEPL, the matter metric is related to the gravitational metric by a disformal map $\tilde{g}_{\mu\nu} = A(\phi)g_{\mu\nu} + B(\phi)\nabla_\mu\phi\nabla_\nu\phi$. The conformal factor $A(\phi)$ rescales the matter sector while preserving null cones; the disformal term $B(\phi)$ can tilt null cones and is strongly constrained by multi-messenger observations. Consequently, PPN light-propagation tests can leave room for clock-sector effects encoded in $A(\phi)$ and for loop/correlation observables that are not representable as a single γ parameter.

5.4 The Conformal vs. Disformal Distinction

In scalar-tensor theories, the physical metric can differ from the gravitational metric through conformal and disformal transformations:

$$\tilde{g}_{\mu\nu} = A(\phi)g_{\mu\nu} + B(\phi)\partial_\mu\phi\partial_\nu\phi$$

where $A(\phi)$ is the conformal factor and $B(\phi)$ is the disformal factor.

The Conformal Loophole

Light propagation tests constrain the combination of conformal and disformal couplings that affects null geodesics. But conformal transformations preserve the null cone structure—light rays follow the same paths regardless of $A(\phi)$.

In a two-metric framework, $A(\phi)$ can modify clock rates (and hence clock residual statistics) while remaining invisible to light-bending and Shapiro-delay measurements. By contrast, any disformal cone tilt sourced by $B(\phi)$ is directly constrained. Light-propagation tests therefore do not directly probe spatial clock-correlation structure or residual loop observables, even when they tightly bound $B(\phi)$.

5.5 Summary: Constraining a Different Sector

What Light Propagation Tests Confirm

- PPN $\gamma = 1$ to 10^{-5} precision
- Light bending follows GR predictions
- Shapiro delay matches GR to high accuracy
- No evidence for preferred-frame effects in light propagation

What They Cannot Test

- Conformal coupling to matter/clocks (preserves null cone)
- Synchronization structure (requires one-way measurements)
- Direction-dependent propagation (cancels in round-trip)
- Clock-sector effects and loop/correlation observables that are not reducible to PPN light-propagation parameters

Light propagation tests are precision triumphs that tightly constrain the post-Newtonian light-propagation sector (e.g., $\gamma \approx 1$) and place strong limits on any disformal cone tilt (encoded in $B(\phi)$). These constraints are largely orthogonal to clock-sector observables such as spatially structured correlations in clock residuals and GR-subtracted residual holonomy H_{Resid} , which require genuinely one-way, direction-reversing configurations.

6. The Multi-Messenger Constraints

The event GW170817 detected gravitational waves from a binary neutron star merger, followed ($+1.74 \pm 0.05$) s later by a gamma-ray burst (GRB 170817A). This timing coincidence yields an assumption-dependent constraint on the fractional speed difference between gravitational waves and light.

6.1 The GW170817 Event

On August 17, 2017, the LIGO-Virgo collaboration detected gravitational waves from a binary neutron star inspiral in NGC 4993, with luminosity distance of order 40 Mpc (130 million light-years; conservative bounds use the lower end of the distance interval) distant. The Fermi Gamma-ray Burst Monitor detected GRB 170817A approximately ($+1.74 \pm 0.05$) s after the gravitational wave signal.

6.1.1 The Speed Constraint

Interpreting the observed delay as (at least partially) a propagation-time difference yields a bound on the fractional speed difference during the trip.

$$-3 \times 10^{-15} \lesssim \frac{c_g - c_\gamma}{c} \lesssim 7 \times 10^{-16}$$

The numerical bounds depend on conservative assumptions about intrinsic emission-time offsets between the gravitational-wave and gamma-ray signals. In particular, the observable directly constrains differential propagation, while the emission-time difference is astrophysically model dependent (Abbott et al. 2017).

6.1.2 The Standard Interpretation

The constraint has been used to rule out or severely constrain numerous modified gravity theories, including:

- Scalar-tensor theories with disformal coupling
- Massive gravity theories
- Lorentz-violating gravity theories
- Certain dark energy models

6.2 What GW170817 Actually Constrains

The Single-Path Problem

Both gravitational waves and gamma rays traveled the same path through spacetime from NGC 4993 to Earth. Any effect that modifies propagation along this path affects both signals identically.

The measurement constrains the differential propagation speed—how much faster or slower gravitational waves travel compared to light along the same path. It does not constrain effects that modify both signals equally.

6.2.1 The Conformal Loophole: Explicit Formulation

Understanding why GW170817 does not directly constrain TEP's conformal sector requires examining the two-metric structure explicitly. This section provides the mathematical derivation that shows conformal coupling cancels in arrival-time comparisons.

The Two-Metric Structure

In scalar-tensor theories like TEP, spacetime has two metrics:

Gravitational metric (Einstein frame):

$$g_{\mu\nu}^{\text{GW}} = g_{\mu\nu}$$

Gravitational waves propagate on null geodesics of this metric.

Electromagnetic metric (Jordan frame):

$$g_{\mu\nu}^{\text{EM}} = A(\phi)g_{\mu\nu}^{\text{GW}} + B(\phi)\nabla_\mu\phi\nabla_\nu\phi$$

Photons propagate on null geodesics of this metric.

The conformal factor $A(\phi) = \exp(2\beta\phi/M_{\text{Pl}})$ rescales the metric uniformly. The disformal factor $B(\phi)$ introduces direction-dependent modifications.

6.2.2 The Flight Time Calculation

Consider a signal (gravitational wave or photon) traveling from source S to detector D along a null geodesic. The coordinate flight time is:

$$T = \int_S^D dt = \int_S^D \frac{dl}{c_{\text{eff}}}$$

where c_{eff} is the effective propagation speed in the relevant metric.

Gravitational Wave Flight Time

For gravitational waves propagating on $g_{\mu\nu}^{\text{GW}}$:

$$T_{\text{GW}} = \int_S^D \frac{dl}{c}$$

where c is the coordinate speed of light in the Einstein frame.

Photon Flight Time (Conformal Only, $B = 0$)

For photons propagating on $g_{\mu\nu}^{\text{EM}} = A(\phi)g_{\mu\nu}^{\text{GW}}$:

$$T_\gamma = \int_S^D \frac{\sqrt{A(\phi)} dl}{c \times \sqrt{A(\phi)}} = \int_S^D \frac{dl}{c} = T_{\text{GW}}$$

The conformal factor $A(\phi)$ cancels in the differential time-of-flight observable under specific operational assumptions:

- **Co-spatial Endpoints:** Source S and Detector D are identical for both signals.
- **Same Clock Standards:** The flight time is measured using the same matter-sector clock definition at the detector.
- **Null Path Invariance:** Conformal transformations preserve the null character of geodesics ($ds^2 = 0 \iff Ads^2 = 0$).
- **Emission-Time Nuisance:** Intrinsic emission delays are treated as nuisance parameters to be modeled or marginalized.

Under these conditions, the observed EM–GW arrival-time difference is insensitive to purely conformal rescalings of the metric along the shared path. The empirical constraint therefore applies to differential propagation effects (for example, disformal cone tilts), not to common-mode rescalings.

Why Conformal Coupling Cancels

Under a conformal transformation $\tilde{g}_{\mu\nu} = A(\phi)g_{\mu\nu}$:

- **Null geodesics remain null:** If $g_{\mu\nu}k^\mu k^\nu = 0$, then $\tilde{g}_{\mu\nu}k^\mu k^\nu = A(\phi) \times 0 = 0$
- **Path length scales:** $d\tilde{l} = \sqrt{A(\phi)} dl$
- **Coordinate speed scales:** $\tilde{c} = \sqrt{A(\phi)} c$
- **Flight time invariant:** $\tilde{T} = d\tilde{l}/\tilde{c} = dl/c = T$

This cancellation concerns the differential arrival-time observable for two signals that follow the same null trajectory and are time-stamped by the same detector clock standard. In that configuration, a purely conformal rescaling does not generate a measurable EM–GW time-of-flight difference.

6.2.3 The Disformal Contribution

The disformal term $B(\phi)\nabla_\mu\phi\nabla_\nu\phi$ does modify the relative propagation:

$$c_\gamma^2 = c^2 \left[1 - \frac{B(\phi)(\partial\phi)^2}{A(\phi)} \right] \quad (\text{schematic parameterization})$$

This introduces a genuine speed difference:

$$\Delta T = T_\gamma - T_{\text{GW}} \approx \frac{D}{c} \times \frac{B(\phi)(\partial\phi)^2}{A(\phi)}$$

GW170817's observed coincidence (after accounting for astrophysical emission-time uncertainties) implies that any net differential propagation delay accumulated over the propagation distance must be small. Operationally, the observation bounds the fractional speed difference at the few $\times 10^{-15}$ level under standard emission-time assumptions.

This is a real constraint on the disformal sector. It does not directly constrain the conformal factor $A(\phi)$ alone.

The Conformal Loophole: Summary

Sector	Effect on Flight Time	GW170817 Constraint
Conformal $A(\phi)$	Cancels exactly ($T_\gamma = T_{\text{GW}}$)	Unconstrained
Disformal $B(\phi)$	Modifies relative speed	$\Delta c/c$ bounded at few $\times 10^{-15}$ (assumption-dependent)

In a two-metric framework, the conformal factor $A(\phi)$ can modify matter-sector proper-time standards and thereby support spatial structure in clock residuals, while remaining invisible to single-path EM–GW arrival-time comparisons. By contrast, any disformal cone tilt sourced by $B(\phi)$ produces a differential propagation effect and is therefore directly constrained by GW170817. A residual synchronization holonomy beyond modeled GR loop effects (H_{resid}) vanishes in the conformal-only limit ($B = 0$) and would require non-exact structure such as $B(\phi) \neq 0$ or more general non-metricity.

6.2.4 The Common-Mode Cancellation

GW170817-type constraints compare two signals that traverse essentially the same spacetime path and are time-stamped by clocks in the matter sector at the detector. Any common-mode rescaling of the local proper-time standard cannot be isolated by an EM–GW arrival-time comparison along a single path. The observable is therefore primarily sensitive to differential propagation effects, such as a disformal cone tilt.

6.3 What Would Constrain Clock-Sector Structure?

To constrain clock-sector structure (including spatial correlations) and to test for non-exact time transport beyond modeled GR loop effects, one needs measurements that do not reduce to single-path, common-mode comparisons:

6.3.1 Multipath Configurations

Gravitational lensing produces multiple images that traverse different paths through the scalar field. If $A(\phi)$ varies spatially, different images experience different time delays—the "Shapiro delay" generalized to scalar fields.

This is the origin of TEP's "phantom mass" prediction: lensing time delays probe the integrated scalar field along different paths, potentially mimicking dark matter.

6.3.2 Statistical Correlations

If the scalar field has spatial structure $\langle\phi(x)\phi(x')\rangle \sim \exp(-r/\lambda)$, clocks at different locations will exhibit correlated fluctuations. This is the signature suggested by exploratory GNSS analysis—distance-structured correlations in clock behavior.

Single-path measurements like GW170817 do not directly probe this structure because they sample only one realization of the field along one path. The proposed spatial-correlation and loop tests should be seen as complementary additions to, not replacements for, multi-messenger constraints.

6.3.3 Closed-Loop Holonomy

The experimentally relevant quantity is a GR-subtracted residual holonomy H_{resid} , defined from one-way time transfer around a direction-reversing closed loop after subtracting modeled GR loop effects (Sagnac, Shapiro, and gravito-magnetic contributions). In the conformal-only limit ($B = 0$), the $A(\phi)$ contribution is an exact gradient and yields $H_{\text{resid}} = 0$ on simply connected domains; a nonzero residual requires non-exact structure, for example a disformal coupling $B(\phi) \neq 0$ or more general non-metricity.

No single-path multi-messenger observation can probe H_{resid} because the signals travel from source to detector, not around direction-reversing closed loops.

6.4 The Broader Multi-Messenger Context

6.4.1 Future Events

Additional multi-messenger detections will improve statistics but not change the fundamental limitation: single-path measurements constrain differential propagation, not common-mode effects.

6.4.2 Strongly Lensed Events

A gravitationally lensed multi-messenger event would provide genuinely new information. If gravitational waves and electromagnetic signals from the same source arrive via different lensed paths, their relative timing probes the scalar field structure along different trajectories.

Such events are rare but would be transformative for constraining conformal coupling.

6.5 Summary: The Disformal Constraint

What GW170817 Confirms

- Under standard emission-time assumptions, the GW–EM speed difference is bounded at the few $\times 10^{-15}$ level
- Disformal coupling $B(\phi)$ is tightly constrained
- Lorentz violation in the gravity sector is bounded
- Massive graviton scenarios are excluded or constrained

What It Cannot Test

- Conformal coupling $A(\phi)$ (preserves null cone)
- Common-mode time dilation along the path
- Spatial structure of scalar fields
- Synchronization holonomy (requires closed loops)
- Clock-sector correlation structure or residual loop observables (H_{resid})

GW170817 is a landmark observation that opened the era of multi-messenger astronomy. Its constraints on modified gravity are real and important—for the disformal sector. Single-path EM–GW coincidence measurements primarily constrain differential propagation (for example, disformal cone tilts) and do not directly probe clock-sector correlation structure or GR-subtracted residual holonomy H_{resid} , which require multipath or direction-reversing closed-loop configurations. This is a statement about measurement geometry, not a criticism of the remarkable experimental achievement.

7. The Resonator Tests

Optical and microwave resonator experiments—descendants of the Michelson-Morley experiment—have achieved extraordinary precision in testing Lorentz invariance. Modern cavity experiments constrain anisotropy in the speed of light at the 10^{-18} level. This section examines the logical structure of these null results.

7.1 The Michelson-Morley Experiment (1887)

The Michelson-Morley experiment sought to detect Earth's motion through the luminiferous aether by measuring the difference in light travel time along perpendicular arms of an interferometer. The null result—no fringe shift as the apparatus rotated—is celebrated as disproving the aether and motivating special relativity.

Methodological Analysis

The interferometer measures the round-trip travel time difference between two perpendicular paths. Light travels out and back along each arm, and the interference pattern depends on the path length difference.

This is a two-way, closed-path measurement. If the one-way speed of light is $c + v$ in one direction and $c - v$ in the opposite direction, the round-trip time is:

$$t = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2Lc}{c^2-v^2}$$

The first-order direction-dependent term cancels. The experiment is sensitive only to second-order effects (v^2/c^2), which special relativity predicts to be exactly compensated by length contraction.

7.1.1 What Michelson-Morley Actually Tests

The null result confirms that the two-way speed of light is isotropic to high precision. It does not test whether the one-way speed is isotropic—that would require synchronized clocks at the endpoints, introducing the synchronization convention problem.

The Two-Way Isotropy Theorem

Any theory in which the round-trip speed of light is c (regardless of direction) will produce a null result in Michelson-Morley type experiments. This includes theories where the one-way speed is anisotropic but the anisotropy cancels in round trips.

7.2 The Kennedy-Thorndike Experiment (1932)

Kennedy and Thorndike modified the Michelson-Morley setup to have unequal arm lengths, testing whether the interference pattern changed as Earth's velocity varied over the year. The null result constrains time dilation effects.

Methodological Analysis

Like Michelson-Morley, this is a two-way measurement. The unequal arms introduce sensitivity to time dilation (how clock rates depend on velocity), but the measurement remains a comparison of round-trip times.

The null result confirms that time dilation and length contraction combine consistently—the Lorentz transformation is self-consistent. It does not probe synchronization structure.

7.3 Modern Cavity Experiments

Contemporary experiments use cryogenic optical or microwave cavities to achieve extraordinary precision. The resonance frequency of a cavity depends on its length and the speed of light:

$$f = \frac{nc}{2L}$$

where n is the mode number. By comparing cavities oriented in different directions, or monitoring a single cavity as Earth rotates, anisotropy in c can be constrained.

7.3.1 Precision Achievements

Modern experiments have achieved constraints on Lorentz violation at remarkable levels:

Experiment	Year	Constraint on $\Delta c/c$
Brillet & Hall	1979	10^{-15}
Müller et al.	2003	10^{-15}
Herrmann et al.	2009	10^{-17}
Nagel et al.	2015	10^{-18}

7.3.2 The Closed-Path Limitation

What Cavities Measure

A cavity resonance involves light bouncing back and forth between mirrors. The resonance condition $f = nc/(2L)$ averages the forward and backward propagation speeds:

$$f = \frac{n}{L/c_{\text{forward}} + L/c_{\text{backward}}} = \frac{n \times c_{\text{forward}} \times c_{\text{backward}}}{L(c_{\text{forward}} + c_{\text{backward}})}$$

If $c_{\text{forward}} = c + \delta$ and $c_{\text{backward}} = c - \delta$, then:

$$f \approx \frac{nc}{2L} \times \left[1 - \frac{\delta^2}{c^2} \right]$$

The first-order anisotropy cancels. Cavity experiments constrain second-order (and higher) Lorentz violation but are blind to first-order direction-dependent effects.

7.4 The Standard Model Extension (SME) Framework

Modern Lorentz violation tests are interpreted within the Standard Model Extension, which parameterizes all possible Lorentz-violating terms in the Standard Model Lagrangian. Cavity experiments constrain specific SME coefficients.

7.4.1 What SME Tests Constrain

- Anisotropy in the two-way speed of light
- Preferred-frame effects in electromagnetism
- CPT violation in photon propagation

7.4.2 What SME Tests Cannot Constrain

- One-way speed anisotropy (convention-dependent)
- Synchronization structure (requires one-way measurements)
- Effects that preserve two-way isotropy while violating one-way isotropy

7.5 The Sagnac Effect: A Partial Exception

The Sagnac effect—the phase shift in a rotating interferometer—is sometimes cited as measuring one-way light speed. In a rotating ring interferometer, counter-propagating beams accumulate different phases:

$$\Delta\Phi = \frac{4\pi A\Omega}{\lambda c}$$

where A is the enclosed area and Ω is the rotation rate.

The Sagnac Subtlety

The Sagnac effect does involve direction-dependent propagation, but in a rotating frame. The effect is fully explained by the non-inertial nature of the rotating reference frame—it does not probe one-way light speed in an inertial frame.

More importantly, the Sagnac interferometer is still a closed-path device. Both beams traverse the same closed loop (in opposite directions) and recombine at the same point. The measurement is of the path integral $\oint \mathbf{v} \cdot d\mathbf{l}$ around the loop, not of one-way propagation.

7.6 Summary: Isotropy of What, Exactly?

What Resonator Tests Confirm

- Two-way speed of light is isotropic to 10^{-18}
- No preferred frame in electromagnetic phenomena (two-way)
- Lorentz invariance holds for closed-path measurements
- SME coefficients for photon sector are tightly bounded

What They Cannot Test

- One-way speed isotropy (convention-dependent)
- Synchronization structure (requires separated clocks)
- Effects that cancel in round trips
- TEP predictions (which preserve two-way isotropy)

The resonator tests confirm that the laws of electromagnetism are the same in all directions—for round-trip measurements. This is a profound constraint on Lorentz violation. But it is not a constraint on synchronization structure, which requires genuinely one-way measurements that resonator experiments cannot provide.

TEP preserves local Lorentz invariance exactly. The scalar field ϕ modifies clock rates but not light propagation. Resonator experiments, which probe light propagation in closed paths, are structurally incapable of detecting TEP signatures.

8. What Would Actually Discriminate?

The preceding sections have established that canonical precision tests of GR share common methodological limitations: two-way measurements, local comparisons, single-path configurations, and theory-laden data reduction. This section proposes experiments that could genuinely discriminate between GR and TEP.

The following proposals are ordered by increasing baseline length and technological ambition, from terrestrial triangles to interplanetary loops.

8.1 The Triangle Holonomy Test

The most direct test of synchronization integrability is the measurement of holonomy around a closed loop using one-way signals. This test must explicitly distinguish TEP holonomy from the well-known Sagnac effect.

8.1.1 Formal Definition: The Loop Observable

The loop observable H is defined operationally as a concrete estimator constructed from time-tagged data. Consider three stations $i \in \{A, B, C\}$ exchanging optical pulses.

The Estimator

Let $\tau_i(t)$ be the proper time reading of clock i at coordinate time t . A pulse emitted by A at τ_A^e is received by B at τ_B^r . The raw one-way interval is $\Delta_{AB} = \tau_B^r - \tau_A^e$.

The Raw Loop Holonomy is the directional difference of the sum of intervals around the loop:

$$H_{\text{raw}} \equiv \oint \Delta_{ij} - \oint \Delta_{ji} = (\Delta_{AB} + \Delta_{BC} + \Delta_{CA}) - (\Delta_{AC} + \Delta_{CB} + \Delta_{BA})$$

Decomposition and Invariance

1. Invariance to Clock Offsets: Let each clock have a constant offset δ_i from coordinate time: $\tau_i(t) = t + \delta_i$. The term δ_i appears in one emission and one reception term in each direction.

In \bigcirc : $(\delta_B - \delta_A) + (\delta_C - \delta_B) + (\delta_A - \delta_C) = 0$.

Thus, H_{raw} is strictly invariant to constant clock synchronization errors.

2. Decomposition: The observable decomposes into physical and systematic terms:

$$H_{\text{raw}} = H_{\text{Sagnac}} + H_{\text{Shapiro}} + H_{\text{GM}} + H_{\text{atm}} + H_{\text{inst}} + H_{\text{TEP}}$$

- $H_{\text{Sagnac}} \propto \boldsymbol{\Omega} \cdot \mathbf{A}$: Earth rotation (dominant, ~ 100 ns).

- $H_{\text{Shapiro}} \propto \sum GM/r$: Gravitational delay (cancels if path AB same as BA , but nonzero if S/C motion makes paths differ).
- H_{GM} : Gravito-magnetic/Lense-Thirring effects ($\sim 10^{-17}$ s).
- H_{atm} : Non-reciprocal atmospheric delays (requires dual-frequency/radiometry).
- H_{inst} : Non-reciprocal instrumental delays (fiber loops, transponders).
- H_{TEP} : The signal of interest (residual synchronization holonomy).

The experimental challenge is to model or control the first five terms to isolate the sixth.

Measurement Recipe

1. Record emission and reception time tags for each directed link ($A \rightarrow B$, $B \rightarrow C$, $C \rightarrow A$ and reverse)
2. Compute H_{raw} from tags using the estimator definition above
3. Model H_{GR} using ephemerides, Earth-orientation parameters, and propagation models
4. Compute $H_{\text{resid}} = H_{\text{raw}} - H_{\text{GR}}$
5. Repeat across geometries chosen to hold Sagnac projection fixed while varying field-gradient sensitivity
6. Report residual scaling with geometry; publish nuisance-model sensitivity

GR Prediction

In GR, after subtracting the modeled loop contributions H_{GR} , the residual satisfies $H_{\text{resid}} = 0$ (up to measurement uncertainty).

TEP Prediction

In the conformal-only limit ($B = 0$), the $A(\phi)$ contribution to time transport is an exact gradient and yields $H_{\text{resid}} = 0$ once GR loop effects are subtracted. A nonzero residual holonomy requires non-exact structure, for example disformal coupling $B(\phi) \neq 0$ or more general non-metricity. In that case, H_{resid} can be nonzero with a magnitude that depends on loop geometry and field gradients.

8.1.2 Distinguishing TEP Holonomy from Sagnac Effect

A skeptic will immediately note that any closed loop on a rotating Earth measures the Sagnac effect ($\Delta t = 4A\Omega/c^2$). The triangle holonomy test must explicitly distinguish TEP effects from Sagnac rotation.

Geometric vs Rotational Signatures

Sagnac Effect:

- **Area-dependent:** $\Delta t_{\text{Sagnac}} = 4A\Omega/c^2$ where A is the enclosed area
- **Rotation-axis dependent:** Proportional to Earth's rotation rate Ω
- **Direction:** Fixed sign determined by right-hand rule with rotation axis
- **Magnitude:** Order picoseconds for terrestrial-scale triangles (configuration-dependent)

Residual Holonomy (TEP target):

- **Area-dependent:** $H_{\text{resid}} \propto (B/A)|\nabla\phi|^2 \times \text{Area}$ (depends on loop flux)
- **Field-gradient dependent:** Proportional to scalar field gradient and disformal coupling
- **Geometry:** Optimized by large vertical or orbital loops sampling field gradients
- **Magnitude:** Order $H/T_{\text{loop}} \sim 10^{-18}\text{--}10^{-16}$ in the forecasts considered here (model- and geometry-dependent)

The numerical values quoted below are order-of-magnitude forecasts intended to clarify the experimental scale. In practice, the effective Sagnac term and the achievable subtraction residual depend on loop geometry, ephemerides and Earth-orientation modeling, link non-reciprocity control, and the time-transfer calibration strategy.

8.1.3 The Large-Area Strategy

Contrary to "zero-area" approaches that eliminate signal along with noise, the original TEP framework (Smawfield 2025e) emphasizes that the holonomy H scales with the loop area (flux of the time-transport curvature). To maximize the signal-to-noise ratio, a Large-Area MEO Strategy is proposed:

Ground-Satellite-Ground Triangle

Configure stations A and B on the ground separated by 1000–3000 km, with a Medium Earth Orbit (MEO) satellite S completing the triangle ($A \rightarrow S \rightarrow B \rightarrow A$):

- **Large Area:** The loop extends to ~20,000 km altitude, enclosing a vast area to maximize TEP flux.
- **Sagnac Modeling:** The large Sagnac term is modeled and subtracted using precise IERS Earth orientation parameters (EOP) and GNSS/SLR ephemerides.
- **TEP Discriminator:** A non-zero residual H_{resid} after Sagnac subtraction.
- **Implementation:** Two-way optical time transfer (TWTTFT) links combined with one-way analysis.

This configuration maximizes sensitivity to the disformal coupling $B(\phi)$, targeting the predicted $H/T_{\text{loop}} \sim 10^{-18} - 10^{-16}$ fractional holonomy.

Numerical Predictions for MEO Configuration

For a ground-MEO triangle (effective loop duration $T_{\text{loop}} \approx 0.4\text{--}0.5$ s, including both directions in H_{raw}):

Parameter	Value	Derivation
Loop time T_{loop}	0.5 s	Total path length ~120,000–150,000 km (six one-way links in H_{raw})
Sagnac contribution	Order 10–100 ns	Large enclosed area; dominant background (configuration-dependent)
Residual holonomy H_{resid}	Order 1–100 as	$H/T_{\text{loop}} \sim 10^{-18}\text{--}10^{-16}$ (forecast; model-dependent)
Required clock stability	10^{-18}	State-of-the-art optical lattice clocks
Detection threshold	Order 10^2 as	Set by subtraction residuals and link non-reciprocity (configuration-dependent)

The signal lies at the frontier of detectability. Isolating an attosecond-level residual from the ~100 ns Sagnac background requires differential measurements across multiple geometries. These targets should be regarded as indicative feasibility thresholds rather than finalized performance specifications.

8.1.4 Explicit Error Budget

The experimental design relies on precise modeling to isolate the TEP residual. The Sagnac effect poses the most significant challenge:

Sagnac Modeling and Subtraction

Source	Contribution	Uncertainty	Subtraction Method
Earth rotation (IERS)	~100 ns	Order 10^2 as (after modeling)	Precise ephemerides + IERS EOP
Polar motion	0.5 ns	Order $10^2\text{--}10^3$ as (after modeling)	IERS polar motion parameters
Geodetic effects	0.1 ns	Order $10^2\text{--}10^3$ as (after modeling)	ITRF reference frame
TEP signal	0.5–50 as	Signal of Interest	Differential Geometry

Challenge: Absolute Sagnac subtraction is ultimately limited by geodetic and dynamical modeling (satellite orbits, Earth rotation, and link calibration). Detecting sub-femtosecond residuals therefore requires differential strategies (comparing loops with different field-gradient sensitivities but similar Sagnac projections) rather than relying solely on absolute subtraction. Feasibility depends on achieving the required link precision and systematic control.

8.1.5 Multiple Geometries for Cross-Validation

To further distinguish TEP from Sagnac effects, use multiple triangle configurations:

Geometry Discrimination Strategy

1. **Large-area triangle:** Maximizes TEP signal (flux of time-transport curvature).
2. **Variable-area comparison:** Signal should scale with loop area (unlike clock systematics).
3. **Opposite-orientation triangles:** Sagnac changes sign; TEP holonomy sign depends on field gradient topology.
4. **Altitude variation:** TEP depends on vertical field gradients ($\nabla\phi$); Sagnac depends only on rotation axis projection.

Consistency across geometries provides robust discrimination between residual holonomy H_{resid} and Sagnac rotation effects.

8.1.6 Technical Requirements

Parameter	Requirement	Current Technology
Clock stability	10^{-18} over measurement period	Achievable with optical lattice clocks
One-way link precision	Femtosecond (averaging to as)	Optical links + carrier phase analysis
Baseline length	Large Area (MEO loop)	Ground-Satellite-Ground configuration
Atmospheric correction	Must not assume GR propagation	Requires independent calibration

8.1.7 The Critical Innovation

The key requirement is genuinely one-way links. Current time transfer methods (two-way satellite time transfer, GPS common-view) use round-trip or common-mode techniques that cancel holonomy by construction.

One-way optical links between ground stations and satellites (or between satellites) could provide the required configuration. The European Space Agency's ACES mission includes a one-way link capability that could, in principle, be used for holonomy measurements—though this is not part of the planned science program.

8.1.8 Objections and Replies

Common Objections

- Objection: One-way times are convention-dependent; the observable is not physical.
- Objection: Any loop timing signal is just the Sagnac effect (Earth rotation).
- Objection: Atmospheric and link delays dominate the femtosecond/attosecond signal.
- Objection: Subtracting H_{GR} assumes GR and makes the test circular.

Replies (Operational and Falsifiable)

- Reply: The loop observable H_{raw} is constructed from recorded emission time tags and reception clock readouts; constant clock offsets cancel exactly in the direction-reversed loop difference.
- Reply: Sagnac contributions are geometry-controlled. Using precise IERS Earth-orientation parameters and satellite ephemerides, the modeled Sagnac term can be subtracted to the level set by geodetic uncertainty. The key discriminator is a residual with the predicted geometry dependence, validated across multiple loop configurations.
- Reply: Link systematics are handled by two-way optical time transfer (TWTTFT) for synchronization and one-way analysis, plus dual-frequency links. The signal scales with loop area (flux), while many systematics scale with path length

or are geometry-independent.

- Reply: H_{GR} subtraction is treated as a modeled calibration. The discriminator is whether a *residual* remains that scales with field gradients (TEP) rather than rotation (Sagnac).

8.1.9 Formal Note: Why Conformal-Only Models Give $H_{\text{resid}} = 0$

Exactness and Residual Holonomy

In the theory formulation, time transport along a link can be represented by an effective one-form Ω . In the conformal-only limit ($B = 0$), Ω reduces to an exact gradient term, $\Omega \propto d(\ln A)/2$, whose curl vanishes. After subtracting modeled GR loop effects (Sagnac, Shapiro, gravito-magnetic terms), the residual loop integral therefore satisfies $H_{\text{resid}} = \oint \Omega = 0$ on simply connected domains. A nonzero residual requires a non-exact contribution, for example a disformal term $B(\phi) \neq 0$ or more general non-metricity.

Throughout the paper, references to H_{resid} as a discriminator should therefore be understood as targeting disformal or non-metric structure in the clock sector, not the purely conformal limit.

8.1.10 The Conformal-Only Limit and TEP Testability

The preceding analysis raises a critical question: if conformal-only TEP ($B = 0$) predicts $H_{\text{resid}} = 0$ after GR subtraction—identical to GR—how can the triangle holonomy test discriminate between the theories?

Resolution: Multi-Signature Discrimination Strategy

The holonomy test primarily constrains disformal coupling $B(\phi)$, while GNSS and clock-correlation observables provide sensitivity to conformal-sector structure via $A(\phi)$. TEP remains testable in the conformal-only limit through complementary observables:

The following signatures are currently best viewed as hypotheses motivated by preliminary analyses rather than established empirical facts:

1. **GNSS Distance Correlations:** Exponential-like distance structure in clock residual correlations on continental-to-global baselines (order 10^3 – 10^4 km) has been reported in GNSS clock products. This claim requires independent replication using raw data and explicit systematic controls.
2. **Orbital Velocity Coupling:** A correlation between GNSS-derived metrics and Earth's orbital dynamics has been reported. Establishing whether this reflects physics or analysis artifacts requires independent replication with pre-specified endpoints and null tests.
3. **CMB Frame Alignment:** Approximate alignment of an inferred anisotropy direction with the CMB dipole has been reported. This claim, if robustly replicated, would discriminate preferred-frame structure in the clock sector from standard GR expectations.
4. **Multi-Constellation Consistency:** If the signal is physical (conformal coupling to matter), it must appear identically across GPS, GLONASS, Galileo, and BeiDou. Systematic artifacts would show constellation-dependent variations.

The triangle holonomy test serves a different purpose: it constrains or detects *disformal* coupling $B(\phi)$. GW170817 bounds the fractional GW-EM speed difference $|c_g - c_\gamma|/c$ at the few $\times 10^{-15}$ level under standard emission-time assumptions, strongly constraining disformal cone tilts in many scalar-tensor models (e.g., Ezquiaga & Zumalacárregui 2017). Clock-sector disformal coupling could differ and remain detectable at the attosecond-level residuals targeted in Table 8.1.3. These residual levels should be regarded as indicative feasibility thresholds rather than finalized performance specifications.

Testability Hierarchy

TEP Sector	Coupling Type	Primary Test	Status
Conformal	A(ϕ) clock rates	GNSS correlations, orbital coupling, CMB alignment	Suggested in exploratory analyses; requires independent replication
Disformal	B(ϕ) propagation	Triangle holonomy, GW speed	GW sector constrained; clock sector untested

Bottom line: TEP is falsifiable even if $B = 0$. The conformal sector makes distinct predictions (distance correlations, velocity coupling, frame alignment) that GR does not. The holonomy test adds sensitivity to disformal structure if present, but is not the sole discriminator.

8.2 Interplanetary Closed-Loop Timing

Deep space baselines are attractive for probing disformal structure because they accumulate propagation effects over long paths. However, a two-point one-way asymmetry is not, by itself, an operational observable because it depends on unknown clock offsets and drifts between endpoints. An interplanetary discriminator therefore requires an explicitly offset-invariant, direction-reversing closed-loop construction.

8.2.1 The Observable: AU-Scale Loop Holonomy

Estimator

Consider three nodes, such as an Earth station A and two spacecraft B and C. Each link is a one-way, time-tagged optical transfer giving $\Delta_{ij} = \tau_j^r - \tau_i^e$.

The raw AU-scale loop holonomy is defined by the direction-reversed loop difference:

$$H_{\text{AU,raw}} \equiv (\Delta_{AB} + \Delta_{BC} + \Delta_{CA}) - (\Delta_{AC} + \Delta_{CB} + \Delta_{BA}).$$

Operational Invariance

Constant clock offsets cancel algebraically in $H_{\text{AU,raw}}$ by the same mechanism as in the terrestrial triangle test. The observable is therefore well-defined without assuming an external synchronization convention, provided that emission and reception time tags are recorded for each one-way link.

8.2.2 Physical Content

The signal of interest is the GR-subtracted residual holonomy $H_{\text{AU,resid}} = H_{\text{AU,raw}} - H_{\text{AU,GR}}$. The modeled GR term includes Shapiro delay contributions from solar-system bodies, kinematic loop terms from platform motion, and plasma dispersion. A nonzero residual with the predicted geometry dependence would indicate non-exact time transport beyond the modeled GR loop effects.

8.2.3 Proposed Configuration

- Three-node geometry using one Earth terminal and two spacecraft (or one spacecraft plus a second Earth terminal), enabling both loop orientations.
- Optical time transfer with recorded emission and reception time tags for each directed link.
- Dual-frequency calibration (or independent plasma monitoring) to control dispersive delays.
- End-to-end modeling of ephemerides and Shapiro delays to construct $H_{\text{AU,GR}}$.

8.2.4 Forecast and Feasibility

Interplanetary Forecast

Parameter	Target Value	Physical Implication
Baseline	0.5–2 AU	Long integration path and large loop geometry
Loop time scale	$T_{\text{loop}} \sim 10^3$ s	Six one-way links contribute to $H_{\text{AU,raw}}$
Residual target	$H_{\text{AU,resid}} \sim 0.1\text{--}10$ ps (indicative)	AU-scale analogue of the triangle-holonomy discriminator; feasibility depends on plasma and ephemeris calibration

This test is technologically demanding because it couples precision time transfer to deep-space navigation and plasma calibration. Its primary value is conceptual: it upgrades interplanetary timing from a convention-dependent two-node asymmetry to an offset-invariant, closed-loop observable.

8.3 GNSS Correlation Replication

Exploratory analyses within the TEP research program have suggested distance-structured correlations in GNSS clock data with correlation length on the order of $10^3\text{--}10^4$ km. Independent, blinded replication of this analysis would provide strong evidence for or against the TEP interpretation.

8.3.1 The Existing Evidence

Exploratory analysis from the TEP research program, using 25 years of CODE clock products and cross-validated with IGS and ESA analysis centers, suggests:

- Exponential decay of clock correlations with distance
- Correlation length of order a few $\times 10^3$ km (reported example: $\lambda \approx \text{few} \times 10^3$ km under one estimator)
- Consistency across three independent analysis centers ($R^2 = 0.92\text{--}0.97$)
- Orbital-dynamics coupling (reported correlation with Earth's orbital velocity)
- Approximate alignment with the CMB frame (reported within tens of degrees)

Status: These findings constitute preliminary evidence requiring independent replication. This paper treats them as a motivating hypothesis and does not assume their numerical values when specifying the discriminating measurement geometries.

8.3.2 The Reproducibility Mandate

For these claims to be scientifically accepted, future analysis must move beyond preliminary reports to a fully transparent **Reproducibility Package**. Any confirming study must provide:

Required Reproducibility Standards

- **Exact Datasets:** Specification of precise products (e.g., IGS 'clk' 30s vs 5m), specific stations (IGS site IDs), and exact time windows used.
- **Preprocessing Pipeline:** Full code or algorithmic description of outlier removal, reference clock definition, and detrending methods (linear vs polynomial).
- **Null Simulations:** Injection/recovery tests where synthetic TEP signals are added to noise to verify detection efficiency, and "scrambled" null tests (randomizing station locations) to verify zero signal.
- **Multiple-Testing Corrections:** Statistical penalties for searching across multiple correlation lengths (λ) and lag times.

Without these controls, exponential fits to noisy residuals can be spurious. The TEP program commits to releasing a public codebase (TEP-GNSS-Open) meeting these standards.

8.3.3 Replication Requirements

Independent replication should:

1. Use raw GNSS data (not processed clock products) to avoid filtering artifacts
2. Apply minimal processing to preserve spatial correlations
3. Test multiple constellations (GPS, GLONASS, Galileo, BeiDou)
4. Verify the correlation length is consistent across datasets
5. Test for systematic artifacts (ionospheric, tropospheric, orbital)

8.3.4 Falsification Criteria

The TEP interpretation would be falsified if:

- Correlation length falls outside 500–20,000 km range
- Correlations disappear in raw (unprocessed) data
- Correlations are explained by known systematic effects
- Different constellations show inconsistent correlation lengths

8.4 Optical Clock Networks and Environmental Screening

The next generation of optical clock networks—connected by optical fiber or free-space links—could provide unprecedented sensitivity to synchronization structure and environmental screening mechanisms.

8.4.1 Current Developments

Several projects are developing continental-scale optical clock networks:

- European fiber network connecting national metrology institutes
- NIST-JILA optical link in Boulder, Colorado
- Proposed intercontinental optical links

8.4.2 TEP Sensitivity: Two Phases

Phase I: Distance Correlations. Optical clocks achieve 10^{-18} stability, potentially sensitive to distance-dependent correlations at the $\delta\tau/\tau \sim 10^{-18}$ level over 1000 km baselines.

Phase II: Environmental Screening Maps. TEP predicts that the scalar field profile $\phi(x)$ is screened by local matter density. Comparing clocks at different altitudes (sea level vs mountain vs LEO) allows mapping the screening profile. After subtracting GR redshift, TEP predicts residual frequency shifts of 10^{-19} – 10^{-18} over tens of kilometers for a screening length $\lambda_{scr} \sim 10$ km.

8.4.3 Critical Design Considerations

To test TEP, optical clock networks must:

- Use one-way comparisons (not just common-mode rejection)
- Span continental-to-global baselines (order 10^3 – 10^4 km)
- Operate continuously to detect temporal variations
- Deploy transportable clocks to varying environments (density/altitude) to test screening

8.5 Matter-Wave Interferometry

Atom interferometers provide a complementary probe sensitive to the gradients of the conformal factor $\nabla \ln A(\phi)$.

- **Objective:** Measure $\nabla \nabla \ln A(\phi)$ via differential phases in atom interferometers.
- **Design:** Long-baseline or tall fountain interferometers (Sr, Rb) in gradiometric configurations.
- **Discrimination:** Combine with mechanical gravimeters. GR affects both; TEP scalar gradients affect atoms (via $A(\phi)$) differently from bulk mass (if composition-dependent couplings exist, though A is universal at leading order).

8.6 Explicit TEP Predictions

To ensure falsifiability, TEP makes specific numerical predictions for each proposed test. These predictions are derived from the theoretical framework (Smaulfield 2025a). Numerical ranges are order-of-magnitude targets; where they are informed by reported GNSS analyses, they are treated as provisional pending independent replication.

Quantitative TEP Predictions			
Observable	TEP Prediction	Derivation	Falsification Criterion
Correlation length λ	1,000–10,000 km	Screening theory: $\lambda \sim (M_p/\rho)^{1/(n+1)} \times \Lambda$	$\lambda < 500$ km or $\lambda > 20,000$ km
Residual holonomy H_{resid}	0.5–50 as (MEO Triangle)	$H_{resid}/T_{loop} \sim 10^{-18}$ – 10^{-16} (TEP-GL Forecast)	$ H_{resid} < 0.1$ as after GR subtraction
Interplanetary Closed-Loop Residual	0.1–10 ps (AU-scale loop)	Direction-reversed loop: $H_{AU,resid} = H_{AU,raw} - H_{AU,GR}$	$ H_{AU,resid} < 0.01$ ps after GR subtraction
Anisotropy ratio (EW/NS)	1.5–3.0	CMB velocity $v \sim 369$ km/s modulates screening	Ratio < 1.2 or > 5.0

Observable	TEP Prediction	Derivation	Falsification Criterion
Orbital coupling	$ r > 0.5$ with orbital velocity	Velocity-dependent screening length	$ r < 0.3$ or wrong phase
CMB alignment	$< 30^\circ$ from CMB dipole	Cosmic rest frame defines field gradient	$> 60^\circ$ separation

Phenomenological Benchmark

Some phenomenological discussions adopt a working benchmark correlation length on the order of 10^3 – 10^4 km (as reported in GNSS analyses) to illustrate experimental sensitivity and to parameterize a representative screening scale. This manuscript does not rely on that benchmark: the underdetermination and the proposed discriminating observables are defined independently of any particular GNSS-derived parameter value.

Independent replication remains essential. If replication fails to confirm a robust distance-structured correlation in raw GNSS observables, any GNSS-informed parameter inference (including mapping an effective correlation length to a field mass scale) would require revision.

8.7 Summary: The Experimental Frontier

Discriminating Tests

Test	Configuration	GR Prediction	TEP Prediction
Triangle Holonomy	One-way closed loop (MEO)	$H_{\text{resid}} = 0$	$H_{\text{resid}} \sim 50$ as
Interplanetary Closed-Loop	AU-scale direction-reversed loop	$H_{\text{resid}} = 0$	$H_{\text{AU,resid}} \sim 0.1$ – 10 ps
GNSS Replication	Raw data analysis	No correlations (or systematic origin)	Distance-structured correlations (characteristic scale to be established by replication)
Optical Clock Network	Continental one-way	No spatial structure	Screening maps / Correlations
Matter-Wave Interferometry	Gradiometer	Depends on $g_{\mu\nu}$	Depends on $g_{\mu\nu}$ AND $\nabla A(\phi)$

These experiments share a common feature: they break the two-way symmetry that characterizes existing precision tests. By measuring one-way propagation, closed-loop residual holonomy H_{resid} , or spatial correlations, they probe the synchronization structure that distinguishes GR from TEP.

None of these experiments has been performed with the explicit goal of testing synchronization integrability. The experimental frontier lies not in achieving greater precision within existing paradigms but in designing fundamentally new measurement configurations.

9. Discussion: The Underdetermination Problem

The analysis presented in Sections 3–7 reveals a systematic pattern: precision tests of general relativity constrain parameters within an assumed framework rather than discriminating between frameworks. This section examines the philosophical implications and situates the findings within the broader context of theory testing in physics.

9.1 The Duhem-Quine Thesis in Relativistic Metrology

The Duhem-Quine thesis holds that scientific hypotheses cannot be tested in isolation; any test involves auxiliary assumptions that could be revised instead of the primary hypothesis. This thesis finds stark expression in precision tests of GR.

9.1.1 Theory-Laden Observation

Every precision measurement requires corrections for systematic effects. These corrections are derived from theoretical models:

- Tropospheric delay corrections rely on refractivity models combined with a relativistic signal-path geometry
- Solar plasma corrections rely on electron-density models combined with spacecraft ephemerides and tracking geometry
- Orbital dynamics are expressed in relativistic reference systems and ephemerides
- Clock comparisons adopt a coordinate-time standard within an assumed relativistic reference frame

The framework being tested generates the corrections applied to test it. This is standard practice in precision metrology: systematic corrections require theoretical and empirical auxiliary models; the logical point is that the "test" is conditional on those auxiliaries. "Testing GR" is therefore more accurately described as "testing the self-consistency of a GR-anchored data-reduction pipeline within a specified model class."

9.1.2 The Auxiliary Hypothesis Problem

When a measurement agrees with GR, the conclusion is: "GR is confirmed." When a measurement disagrees, the response is typically to revise auxiliary assumptions (calibration errors, unmodeled systematics) rather than question GR.

This asymmetry is rational—GR has enormous empirical support—but it illustrates how discrepancies are often underdetermined between core theory and auxiliary models unless experiments are designed to vary the auxiliaries independently.

9.2 The Conventionality of Simultaneity

The philosophical literature on the conventionality of simultaneity, initiated by Reichenbach and developed by Grünbaum, Salmon, and others, directly addresses the issues raised in this paper. This section translates these philosophical insights into operational terms accessible to experimental physicists.

9.2.1 The Operational Problem: Measuring with the Thing You're Measuring

The Core Issue in Plain Language

You cannot define synchronization using light rays if the speed of light is the thing you are trying to measure.

Einstein's synchronization procedure uses light signals to define "simultaneous." But this procedure assumes light travels at the same speed in both directions. To verify this assumption, you would need to measure one-way light speed—which requires synchronized clocks. The reasoning is circular.

9.2.2 The Rubber Ruler Analogy

Why Two Metrics Are Necessary

Imagine you have a rubber ruler that can stretch. You want to measure whether the ruler is stretching. But the only measuring device you have is... the ruler itself.

- If the ruler stretches uniformly, every measurement you make with it will still give the same numbers
- You cannot detect the stretching using only the stretching ruler
- To detect stretching, you need a *second, independent reference*—a rigid ruler that doesn't stretch

In GR, spacetime is the only ruler. If spacetime "stretches" (the metric changes), all measurements made with spacetime—including light propagation—stretch with it. You cannot detect the change.

In TEP, there are two rulers: the gravitational metric (for gravity) and the matter metric (for clocks and light). If one "stretches" relative to the other, the difference is detectable.

9.2.3 Reichenbach's ϵ -Synchronization

Reichenbach formalized this insight mathematically. Einstein's synchronization convention ($\epsilon = 1/2$, meaning light takes equal time in each direction) is not empirically determined. Any value $0 < \epsilon < 1$ yields an empirically equivalent theory.

The one-way speed of light is:

$$c_+ = \frac{c}{2\epsilon} \quad (\text{forward direction})$$

$$c_- = \frac{c}{2(1-\varepsilon)} \quad (\text{backward direction})$$

Only the round-trip speed c is empirically determined. The choice $\varepsilon = 1/2$ is conventional, not empirical—unless you have an independent way to synchronize clocks.

9.2.4 Malament's Theorem: The Uniqueness Claim

Malament (1977) proved that Einstein synchronization is the unique synchronization definable from the causal structure of Minkowski spacetime. This is often cited as resolving the conventionality debate in favor of $\varepsilon = 1/2$.

What Malament's Theorem Actually Says

In operational terms: *If light rays are your only tool for defining simultaneity, then Einstein synchronization is the only consistent choice.*

This is true—but it assumes light rays are your only tool. The theorem's premise is that the causal structure (determined by light propagation) is the fundamental structure of spacetime.

Why It Doesn't Apply to TEP

TEP introduces a second structure: the matter metric $\tilde{g}_{\mu\nu}$ that governs clock behavior. Clocks don't just follow light—they follow their own metric. This provides the "second ruler" that breaks the circularity.

- **Single-metric (GR):** Malament applies. Light defines synchronization uniquely.
- **Two-metric (TEP):** Malament doesn't apply. Clocks and light can disagree on synchronization.

The Key Point: Malament Assumes Causality Is Fundamental

Malament proves uniqueness *given* that light defines causality. TEP introduces a second causal structure (the matter metric). The theorem's premise fails, not its logic.

In single-metric theories, there is only one notion of "simultaneous"—the one defined by light cones. In two-metric theories, clocks and light can define different simultaneity surfaces. The disagreement between them is precisely the residual synchronization holonomy H_{resid} targeted by direction-reversing closed-loop experiments.

9.2.5 The Two-Clock Thought Experiment

Making It Concrete

Imagine two identical atomic clocks, A and B, separated by distance L . You want to synchronize them.

GR procedure:

1. Send a light pulse from A to B
2. B reflects it back to A
3. A measures round-trip time $T = 2L/c$
4. Conclude: one-way time is $T/2 = L/c$ (by convention)

The hidden assumption: Light takes equal time in each direction. But you cannot verify this without already-synchronized clocks!

TEP alternative:

1. The scalar field ϕ affects clock rates through $A(\phi)$
2. If ϕ differs at A and B, clocks tick at different rates
3. Light still travels at c in both directions (conformal coupling preserves null cones)
4. But "simultaneous" according to clocks \neq "simultaneous" according to light

The two-metric structure provides an independent reference that breaks the circularity. Clock synchronization and light synchronization can disagree—and this disagreement is the residual synchronization holonomy H_{resid} (after subtracting modeled

GR loop effects).

9.2.6 Implications for TEP

TEP is not a simple Reichenbach $\epsilon \neq 1/2$ anisotropic light-speed theory. This distinction is critical for understanding what TEP actually claims and how it differs from conventional alternatives to special relativity.

What TEP Is NOT

- **Not an $\epsilon \neq 1/2$ theory:** TEP does not claim that light travels at different speeds in different directions within a single reference frame.
- **Not a Lorentz violation:** Local Lorentz invariance holds exactly in freely falling laboratories. The speed of light is isotropic locally ($\epsilon = 1/2$ locally).
- **Not a preferred-frame theory (in the conventional sense):** While the scalar field ϕ defines a cosmic rest frame, this does not violate local Lorentz symmetry because the conformal coupling preserves null cones.

What TEP Actually Claims

TEP preserves local Lorentz invariance (including local isotropy of light propagation, $\epsilon = 1/2$) while introducing *global path-dependent synchronization* through the scalar field ϕ . The key distinction:

- **Locally:** Light propagates isotropically at c in all directions ($\epsilon = 1/2$)
- **Globally:** Clock synchronization around extended loops can exhibit path-dependence due to variations in the conformal factor $A(\phi)$ along different routes
- **Operationally:** Two-way measurements (which average forward and backward paths) see no anisotropy. One-way closed loops can reveal residual holonomy H_{resid} if disformal structure $B(\phi) \neq 0$

This is fundamentally different from Reichenbach's ϵ -synchronization, which modifies the *local* one-way speed of light within a single frame. TEP modifies *global* synchronization structure while preserving local isotropy.

The Bottom Line for Experimentalists

Malament's theorem says: "If you only have one ruler, you can only make one kind of measurement." This is mathematically correct but operationally limiting.

TEP says: "We have two rulers—the gravitational metric and the matter metric. They usually agree, but when they don't, we can detect it."

The triangle holonomy test is precisely this: comparing what clocks say about synchronization with what light says. If they disagree after GR subtraction, $H_{\text{resid}} \neq 0$.

Critical distinction: This is not testing whether the local photon null cone is anisotropic (it isn't). It's testing whether clock synchronization and light synchronization define the same global time coordinate (they might not if the matter metric differs from the gravitational metric).

9.3 Underdetermination vs. Falsifiability

The underdetermination identified in this paper is not the radical underdetermination of Quine (where infinitely many theories fit any finite data). It is a specific, resolvable underdetermination: existing tests do not directly probe the observables that distinguish GR from TEP, but new tests could.

9.3.1 The Structure of the Underdetermination

Observable	GR Prediction	TEP Prediction	Status
Local clock rates	$d\tau/dt = \sqrt{g_{00}}$	$d\tau/dt = \sqrt{g_{00}}$	Identical
Round-trip light time	$2L/c$	$2L/c$	Identical
PPN γ	1	1	Identical
c_γ vs c_g	Equal	Equal (conformal)	Identical
Residual holonomy H_{resid}	0	0 if $B = 0$; $\neq 0$ possible if non-exact structure is present	Distinguishable
Clock correlations	None (or systematic)	Exponential decay	Distinguishable

9.3.2 Resolvable Underdetermination

The underdetermination is resolvable because GR and TEP make different predictions for observables that have not yet been measured with appropriate configurations. The triangle holonomy test, interplanetary closed-loop timing, and GNSS correlation replication could all distinguish the theories.

This is not a failure of falsifiability but a gap in the experimental program. The tests that would discriminate have simply not been performed.

9.4 The Sociology of Precision Tests

Why have discriminating tests not been performed? Several factors contribute:

9.4.1 Two-Way Convenience

Two-way measurements are experimentally simpler. They require only one clock, avoid synchronization problems, and provide gauge-invariant results. The experimental tradition naturally evolved toward two-way configurations.

9.4.2 The "GR is Correct" Prior

GR has been spectacularly successful. The prior probability assigned to alternatives is low, reducing motivation to design experiments specifically to test them. Resources flow toward improving precision within the GR framework rather than testing the framework itself.

9.4.3 The PPN Paradigm

The PPN formalism provides a systematic way to parameterize deviations from GR. But PPN is a parameterization of the post-Newtonian limit under assumptions about which effective metric governs the sector being tested (typically solar-system dynamics and light propagation). This makes it an exceptionally powerful framework for constraining single-metric departures in those sectors, while leaving open the possibility of additional clock-sector structure that is not representable as a small set of PPN light-propagation parameters.

9.4.4 Scientific Conservatism

When experiments confirm GR, they are celebrated. When anomalies appear (Pioneer anomaly, flyby anomaly), enormous effort goes into finding conventional explanations. This is scientifically appropriate—extraordinary claims require extraordinary evidence, and GR has earned its status through a century of successful predictions. The point is not that this conservatism is wrong, but that it shapes which experiments get funded and performed.

9.5 The Broader Context

9.5.1 Dark Matter and Dark Energy

The standard cosmological model requires ~95% of the universe to consist of unknown dark matter and dark energy. These are inferred from gravitational effects assuming GR is correct on all scales.

If GR is not the complete description of gravity—if TEP or another modification is required—the inferred dark sector could be partially or wholly an artifact of applying the wrong theory. The stakes for testing GR alternatives are high.

9.5.2 The Hubble Tension

The discrepancy between early-universe (CMB) and late-universe (Cepheid/SNe) measurements of the Hubble constant has reached 5σ significance. This tension could indicate new physics—potentially related to modifications of GR.

9.5.3 The S_8 Tension

Similarly, measurements of matter clustering (S_8) show tension between CMB predictions and direct measurements. Modified gravity theories, including scalar-tensor theories like TEP, could potentially resolve these tensions.

9.6 Summary: The State of the Evidence

What the Experimental Canon Establishes

- GR is self-consistent to extraordinary precision
- Local Lorentz invariance holds at 10^{-18}
- PPN parameters match GR predictions
- No evidence for preferred-frame effects (in two-way tests)

What Remains Untested

- Synchronization integrability (holonomy)
- One-way light speed isotropy
- Spatial structure of clock correlations
- Conformal coupling to scalar fields

The Underdetermination

For the classes of precision tests reviewed in this paper, and under the conditions where local Einstein equivalence holds and any disformal cone tilt is constrained to be negligible by multi-messenger bounds, GR and conformal-sector TEP can be observationally degenerate. The experimental canon therefore constrains parameter space within the standard single-metric framework but does not directly target the specific loop and correlation observables where the frameworks differ. This underdetermination is resolvable through new experimental configurations.

10. Conclusions

This paper has presented a systematic methodological analysis of the canonical precision tests of general relativity. The analysis reveals that these tests—while representing extraordinary achievements in experimental physics—share structural features that, as currently configured, do not directly probe the specific loop and correlation observables that distinguish single-metric theories (GR) from two-metric alternatives such as the Temporal Equivalence Principle.

10.1 Principal Findings

The analysis in Sections 3–8 demonstrates that the five structural limitations identified in §1.2—two-way measurement dominance, local/global conflation, model-dependent calibration, the conformal loophole, and theory-laden data reduction—are not isolated issues but interconnected features of the experimental tradition.

The underdetermination is structural, not accidental. Two-way measurements evolved because they are simpler and provide gauge-invariant results. The PPN formalism, which guides most precision tests, constrains the post-Newtonian limit in the sectors it parameterizes (solar-system dynamics and light propagation). This makes it exceptionally powerful for ruling out many alternatives to GR, while leaving open the possibility of additional clock-sector structure that is not reducible to a small set of PPN light-propagation parameters.

10.2 What the Tests Do Establish

The critique presented here should not be misread as claiming that precision tests are worthless. They establish important facts:

- GR provides a self-consistent framework for describing gravity
- Local Lorentz invariance holds to 10^{-18} precision
- PPN parameters match GR predictions ($\gamma = \beta = 1$)
- GW–EM propagation speeds are bounded to differ by at most a few $\times 10^{-15}$ under standard emission-time assumptions (constraining disformal coupling)
- No evidence for preferred-frame effects in two-way measurements

These are genuine constraints on the space of viable theories. Many alternatives to GR are ruled out by these tests. But TEP is not among them in the regime considered here: it can be formulated to satisfy the existing constraints while introducing additional clock-sector structure that is not directly targeted by the canonical measurement geometries.

10.3 The Path Forward

10.3.1 Discriminating Experiments

The underdetermination is resolvable. Experiments that could distinguish GR from TEP include:

- *Triangle Holonomy Test*: One-way timing around a large-area closed loop (e.g., ground-satellite-ground), measuring a GR-subtracted residual holonomy H_{resid} . GR predicts $H_{\text{resid}} = 0$; a nonzero residual indicates non-exact time geometry.
- *Matter-Wave Interferometry*: Atom interferometers in gradiometric configurations sensitive to gradients of the scalar field $\nabla\nabla\phi$ via the conformal factor, distinguishable from GR gravity by combining with mechanical gravimeters.
- *Interplanetary Closed-Loop Timing*: AU-scale, direction-reversed loop measurements constructed from one-way time-tagged links, targeting a GR-subtracted residual holonomy analogous to the terrestrial triangle observable.
- *GNSS Correlation Replication*: Independent, blinded analysis of raw GNSS data to verify or refute the distance-structured correlations suggested by exploratory analyses within the TEP research program.
- *Optical Clock Networks*: Continental-scale networks using one-way comparisons to probe synchronization structure at 10^{-18} precision.

10.3.2 The Scientific Imperative

The dark matter problem, the Hubble tension, and the S_8 discrepancy all suggest that current understanding of gravity may be incomplete. Testing alternatives to GR is not merely an academic exercise—it addresses fundamental questions about the nature of the universe.

Optical clocks, one-way optical links, and GNSS infrastructure already exist. What is needed is a shift in experimental philosophy: from improving precision within the GR framework to designing tests that could falsify it.

10.3.3 Experimental Priorities

The proposed discriminating experiments can be ranked by technical feasibility, discriminating power, and resource requirements:

Priority Ranking for Discriminating Tests				
Priority	Experiment	Technical Feasibility	Discriminating Power	Timeline/Cost
1. Highest	GNSS Correlation Replication	High (data already exists)	High (tests conformal sector)	1-2 years / Low (computational)
2. High	Optical Clock Networks (one-way)	Medium (infrastructure exists, requires protocol modification)	High (10^{-18} sensitivity)	3-5 years / Medium (leverage existing networks)
3. Medium	Triangle Holonomy (Ground-Satellite)	Medium (requires new optical links)	High (clean null test for disformal structure)	5-7 years / Medium-High (new infrastructure)
4. Medium	Matter-Wave Interferometry	Medium (technology maturing)	Medium (tests gradients $\nabla\nabla\phi$)	5-10 years / Medium (lab/space)
5. Lowest	Interplanetary Closed-Loop Timing	Low (requires deep space mission)	Medium (long baselines, many systematics)	15-20 years / Very High (dedicated spacecraft)

Recommended Near-Term Strategy

1. **Immediate (0-2 years):** Independent replication of GNSS correlation analysis using raw RINEX data. This is the most cost-effective discriminating test and could provide strong evidence for or against the existence of robust distance-structured correlations on continental-to-global baselines (order 10^3 – 10^4 km).
2. **Short-term (2-5 years):** Modify existing optical clock network protocols to enable one-way comparisons. This leverages substantial existing infrastructure (European fiber network, NIST-JILA links) with minimal additional cost.
3. **Medium-term (5-10 years):** Develop ground-based triangle holonomy test using intercontinental optical links. This requires new infrastructure but is technically feasible with current technology.
4. **Long-term (10+ years):** Space-based holonomy tests and interplanetary missions as part of broader fundamental physics programs (e.g., ACES follow-on, dedicated deep space clocks).

Critical first step: The GNSS replication is both the highest priority and the most feasible. If independent, blinded analysis confirms robust distance-structured correlations on continental-to-global baselines (order 10^3 – 10^4 km), it would provide strong

motivation for the more expensive optical clock and holonomy experiments. If replication fails, any GNSS-informed interpretation would require substantial revision.

10.4 Concluding Remarks

The Central Argument

The experimental foundations of general relativity are more specialized than commonly appreciated—not because the experiments are flawed, but because they were designed to test specific predictions within an assumed framework. Two-way measurements, local comparisons, and single-path configurations excel at what they were designed for; they do not directly probe synchronization structure. The precision is extraordinary; the scope is necessarily limited.

TEP remains viable within all existing constraints. This is not because TEP is unfalsifiable—explicit falsification criteria exist—but because the experiments that would test them have not been performed.

The path forward is clear: design and execute experiments that break the two-way symmetry. Measure holonomy. Test spatial correlations. Probe synchronization structure directly. Only then will the experimental foundations of relativistic physics be secure.

10.5 Summary Table: The Experimental Canon

Experiment	Claimed Result	Actual Constraint	TEP Status
Pound-Rebka (1960)	Confirms gravitational redshift	Local clock rate ratio	Compatible
Hafele-Keating (1971)	Confirms time dilation	Proper time on closed worldlines	Compatible
Gravity Probe A (1976)	Confirms GR to 7×10^{-5}	Two-way clock comparison	Compatible
GPS (1978-present)	Proves GR corrections	Self-consistency of GR model	Compatible
Lunar Laser Ranging (1969-present)	Nordtvedt effect null	Two-way ranging; gravitational sector only	Compatible
Cassini (2003)	$\gamma = 1$ to 10^{-5}	Two-way Shapiro delay	Compatible
Gravity Probe B (2011)	Confirms frame-dragging	Gyroscope precession ($g_{\mu\nu}$ only)	Compatible
LIGO/Virgo (2015-present)	Detects gravitational waves	GW propagation on $g_{\mu\nu}$	Compatible
GW170817 (2017)	$ c_g - c_y /c$ bounded at few $\times 10^{-15}$ (assumption-dependent)	Disformal coupling only	Compatible
Optical Clocks (2022)	Redshift at 10^{-8}	Local Lorentz invariance	Compatible

Every entry in the "TEP Status" column reads "Compatible." This reflects the structural nature of the underdetermination. The experimental canon, for all its remarkable precision, was designed to test predictions that GR and TEP share. The experiments succeed brilliantly at their intended purpose; they simply were not designed to probe the specific observables where GR and TEP differ.

This is not a criticism of the experimentalists who performed these measurements—their work represents some of the greatest achievements in precision physics. It is an observation about the logical structure of theory testing: to distinguish between theories, one must measure observables where they make different predictions.

References

Foundational Experiments

Everitt, C. W. F., et al. (2011). Gravity Probe B: Final results of a space experiment to test general relativity. *Physical Review Letters*, 106(22), 221101. doi:10.1103/PhysRevLett.106.221101

Pound, R. V., & Rebka, G. A. (1960). Apparent weight of photons. *Physical Review Letters*, 4(7), 337-341. doi:10.1103/PhysRevLett.4.337

- Pound, R. V., & Snider, J. L. (1965). Effect of gravity on gamma radiation. *Physical Review*, 140(3B), B788-B803. doi:10.1103/PhysRev.140.B788
- Hafele, J. C., & Keating, R. E. (1972). Around-the-world atomic clocks: Predicted relativistic time gains. *Science*, 177(4044), 166-168. doi:10.1126/science.177.4044.166
- Hafele, J. C., & Keating, R. E. (1972). Around-the-world atomic clocks: Observed relativistic time gains. *Science*, 177(4044), 168-170. doi:10.1126/science.177.4044.168
- Vessot, R. F. C., et al. (1980). Test of relativistic gravitation with a space-borne hydrogen maser. *Physical Review Letters*, 45(26), 2081-2084. doi:10.1103/PhysRevLett.45.2081
- Williams, J. G., Turyshev, S. G., & Boggs, D. H. (2012). Lunar laser ranging tests of the equivalence principle. *Classical and Quantum Gravity*, 29(18), 184004. doi:10.1088/0264-9381/29/18/184004
- Murphy, T. W. (2013). Lunar laser ranging: The millimeter challenge. *Reports on Progress in Physics*, 76(7), 076901. doi:10.1088/0034-4885/76/7/076901

Light Propagation Tests

- Shapiro, I. I. (1964). Fourth test of general relativity. *Physical Review Letters*, 13(26), 789-791. doi:10.1103/PhysRevLett.13.789
- Bertotti, B., Iess, L., & Tortora, P. (2003). A test of general relativity using radio links with the Cassini spacecraft. *Nature*, 425(6956), 374-376. doi:10.1038/nature01997
- Kopeikin, S., Schaefer, G., Polnarev, A., & Vlasov, I. (2006). The orbital motion of Sun and a test of general relativity using radio links with the Cassini spacecraft. *arXiv preprint*, arXiv:gr-qc/0604060. arXiv:gr-qc/0604060
- Fomalont, E. B., & Kopeikin, S. M. (2003). The measurement of the light deflection from Jupiter: Experimental results. *The Astrophysical Journal*, 598(1), 704-711. doi:10.1086/378785

GPS and Relativistic Timekeeping

- Ashby, N. (2003). Relativity in the Global Positioning System. *Living Reviews in Relativity*, 6(1), 1. doi:10.12942/lrr-2003-1
- Ashby, N., & Weiss, M. (1999). Global Positioning System receivers and relativity. *NIST Technical Note*, 1385.
- Kouba, J., & Héroux, P. (2001). Precise Point Positioning Using IGS Orbit and Clock Products. *GPS Solutions*, 5(2), 12-28. doi:10.1007/PL00012883
- Dow, J. M., Neilan, R. E., & Rizos, C. (2009). The International GNSS Service in a changing landscape of Global Navigation Satellite Systems. *Journal of Geodesy*, 83(3-4), 191-198. doi:10.1007/s00190-008-0300-3

GNSS Processing and Common-Mode Filtering

- Dong, D., Fang, P., Bock, Y., Cheng, M. K., & Miyazaki, S. (2006). Spatiotemporal filtering using principal component analysis and Karhunen–Loève expansion approaches for regional GPS network analysis. *Journal of Geophysical Research: Solid Earth*, 111, B03405. doi:10.1029/2005JB003806

Modern Optical Clock Experiments

- Bothwell, T., et al. (2022). Resolving the gravitational redshift across a millimetre-scale atomic sample. *Nature*, 602(7897), 420-424. doi:10.1038/s41586-021-04349-7
- Zheng, X., et al. (2022). Differential clock comparisons with a multiplexed optical lattice clock. *Nature*, 602(7897), 425-430. doi:10.1038/s41586-021-04344-y
- McGrew, W. F., et al. (2018). Atomic clock performance enabling geodesy below the centimetre level. *Nature*, 564(7734), 87-90. doi:10.1038/s41586-018-0738-2
- Mehlstäubler, T. E., Grosche, G., Lisdat, C., Schmidt, P. O., & Denker, H. (2018). Atomic clocks for geodesy. *Reports on Progress in Physics*, 81(6), 064401. doi:10.1088/1361-6633/aab409

Gravitational Wave Observations

- Abbott, B. P., et al. (LIGO Scientific Collaboration and Virgo Collaboration) (2017). GW170817: Observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119(16), 161101. doi:10.1103/PhysRevLett.119.161101
- Abbott, B. P., et al. (2017). Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A. *The Astrophysical Journal Letters*, 848(2), L13. doi:10.3847/2041-8213/aa920c

Resonator and Lorentz Invariance Tests

- Michelson, A. A., & Morley, E. W. (1887). On the relative motion of the Earth and the luminiferous ether. *American Journal of Science*, 34(203), 333-345.

Kennedy, R. J., & Thorndike, E. M. (1932). Experimental establishment of the relativity of time. *Physical Review*, 42(3), 400-418. doi:10.1103/PhysRev.42.400

Brillet, A., & Hall, J. L. (1979). Improved laser test of the isotropy of space. *Physical Review Letters*, 42(9), 549-552. doi:10.1103/PhysRevLett.42.549

Müller, H., Herrmann, S., Braxmaier, C., Schiller, S., & Peters, A. (2003). Modern Michelson–Morley experiment using cryogenic optical resonators. *Physical Review Letters*, 91(2), 020401. doi:10.1103/PhysRevLett.91.020401 arXiv:physics/0305117

Herrmann, S., et al. (2009). Rotating optical cavity experiment testing Lorentz invariance at the 10^{-17} level. *Physical Review D*, 80(10), 105011. doi:10.1103/PhysRevD.80.105011

Nagel, M., et al. (2015). Direct terrestrial test of Lorentz symmetry in electrodynamics to 10^{-18} . *Nature Communications*, 6, 8174. doi:10.1038/ncomms9174

PPN Formalism and Tests of Gravity

Will, C. M. (2014). The confrontation between general relativity and experiment. *Living Reviews in Relativity*, 17(1), 4. doi:10.12942/lrr-2014-4

Will, C. M. (2018). *Theory and Experiment in Gravitational Physics* (2nd ed.). Cambridge University Press.

Philosophy of Space and Time

Reichenbach, H. (1958). *The Philosophy of Space and Time*. Dover Publications. (Original work published 1928)

Grünbaum, A. (1973). *Philosophical Problems of Space and Time* (2nd ed.). D. Reidel Publishing Company.

Malament, D. (1977). Causal theories of time and the conventionality of simultaneity. *Noûs*, 11(3), 293-300. doi:10.2307/2214766

Salmon, W. C. (1977). The philosophical significance of the one-way speed of light. *Noûs*, 11(3), 253-292. doi:10.2307/2214765

Franklin, A. (1986). *The Neglect of Experiment*. Cambridge University Press.

Scalar-Tensor Theories and Modified Gravity

Brans, C., & Dicke, R. H. (1961). Mach's principle and a relativistic theory of gravitation. *Physical Review*, 124(3), 925-935. doi:10.1103/PhysRev.124.925

Bekenstein, J. D. (2004). Relativistic gravitation theory for the modified Newtonian dynamics paradigm. *Physical Review D*, 70(8), 083509. doi:10.1103/PhysRevD.70.083509

Zumalacárregui, M., & García-Bellido, J. (2014). Transforming gravity: From derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian. *Physical Review D*, 89(6), 064046. doi:10.1103/PhysRevD.89.064046

TEP Research Program

Smawfield, M. L. (2025a). The Temporal Equivalence Principle: A Two-Metric Framework for Gravitational Time. Zenodo. doi:10.5281/zenodo.16921911

Smawfield, M. L. (2025b). TEP-GNSS I: Global Time Echoes in Multi-Center GNSS Clock Products. Zenodo. doi:10.5281/zenodo.17127229

Smawfield, M. L. (2025c). TEP-GNSS II: 25-Year Analysis of Temporal-Gravitational Coupling. Zenodo. doi:10.5281/zenodo.17517141

Smawfield, M. L. (2025d). TEP-GNSS III: Raw RINEX Validation of Distance-Structured Clock Correlations. Zenodo. doi:10.5281/zenodo.17860166

Smawfield, M. L. (2025e). TEP-GL: Temporal-Spatial Coupling in Gravitational Lensing. Zenodo. doi:10.5281/zenodo.17982540

Smawfield, M. L. (2025f). TEP-GTE: Global Time Echoes—Empirical Validation of the Temporal Equivalence Principle. Zenodo. doi:10.5281/zenodo.18004832

Smawfield, M. L. (2025g). TEP-UCD: Universal Critical Density—Unifying Atomic, Galactic, and Compact Object Scales. Zenodo. doi:10.5281/zenodo.18064366

Smawfield, M. L. (2025h). TEP-RBH: The Soliton Wake—Identifying RBH-1 as a Gravitational Soliton. Zenodo. doi:10.5281/zenodo.18059251

Smawfield, M. L. (2025i). TEP-SLR: Global Time Echoes—Optical Validation via Satellite Laser Ranging. Zenodo. doi:10.5281/zenodo.18064582

Galileo Eccentric Orbit Tests

Delva, P., et al. (2018). A new test of gravitational redshift using Galileo satellites: The GREAT experiment. *Comptes Rendus Physique*, 20(3), 176-182. doi:10.1016/j.crhy.2019.04.002

Herrmann, S., et al. (2018). Test of the gravitational redshift with Galileo satellites in an eccentric orbit. *Physical Review Letters*, 121(23), 231102. doi:10.1103/PhysRevLett.121.231102

ACES Mission and Space Clocks

Cacciapuoti, L., & Salomon, C. (2009). Space clocks and fundamental tests: The ACES experiment. *The European Physical Journal Special Topics*, 172(1), 57-68. doi:10.1140/epjst/e2009-01041-7

Savalle, E., Guerlin, C., Delva, P., Meynadier, F., le Poncin-Lafitte, C., & Wolf, P. (2019). Gravitational redshift test with the future ACES mission. *Classical and Quantum Gravity*. doi:10.1088/1361-6382/ab4f25 arXiv:1907.12320

Savalle, E., et al. (2024). Atomic Clock Ensemble in Space. *arXiv preprint*, arXiv:2411.02912. arXiv:2411.02912

Deep Space Atomic Clock

Burt, E. A., et al. (2021). Demonstration of a trapped-ion atomic clock in space. *Nature*, 595(7865), 43-47. doi:10.1038/s41586-021-03571-7

Ely, T. A., et al. (2021). Results of the Deep Space Atomic Clock deep space navigation analog experiment. *Journal of Spacecraft and Rockets*, 59(4), 1294-1310. doi:10.2514/1.A35334

Equivalence Principle Tests

Touboul, P., et al. (2017). MICROSCOPE mission: First results of a space test of the equivalence principle. *Physical Review Letters*, 119(23), 231101. doi:10.1103/PhysRevLett.119.231101

Touboul, P., et al. (2022). MICROSCOPE mission: Final results of the test of the equivalence principle. *Physical Review Letters*, 129(12), 121102. doi:10.1103/PhysRevLett.129.121102

Optical Clock Networks and Fiber Links

Lisdat, C., et al. (2016). A clock network for geodesy and fundamental science. *Nature Communications*, 7, 12443. doi:10.1038/ncomms12443

Clivati, C., et al. (2022). Coherent optical-fiber link across Italy and France. *Physical Review Applied*, 18(5), 054009. doi:10.1103/PhysRevApplied.18.054009

Caldwell, E. D., et al. (2023). Quantum-limited optical time transfer for future geosynchronous links. *Nature*, 618(7966), 721-726. doi:10.1038/s41586-023-06032-5

Test Theories and Synchronization

Mansouri, R., & Sexl, R. U. (1977). A test theory of special relativity: I. Simultaneity and clock synchronization. *General Relativity and Gravitation*, 8(7), 497-513. doi:10.1007/BF00762634

Colladay, D., & Kostelecký, V. A. (1998). Lorentz-violating extension of the Standard Model. *Physical Review D*, 58(11), 116002. doi:10.1103/PhysRevD.58.116002

Mattingly, D. (2005). Modern tests of Lorentz invariance. *Living Reviews in Relativity*, 8(1), 5. doi:10.12942/lrr-2005-5

Kostelecký, V. A., & Russell, N. (2011). Data tables for Lorentz and CPT violation. *Reviews of Modern Physics*, 83(1), 11-31. doi:10.1103/RevModPhys.83.11

GNSS Common-Mode Errors

Wdowinski, S., et al. (1997). Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements. *Journal of Geophysical Research*, 102(B8), 18057-18070. doi:10.1029/97JB01378

Dong, D., et al. (2006). Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis. *Journal of Geophysical Research*, 111(B3), B03405. doi:10.1029/2005JB003806

Scalar-Tensor Constraints from GW170817

Langlois, D., Saito, R., Yamauchi, D., & Noui, K. (2018). Scalar-tensor theories and modified gravity in the wake of GW170817. *Physical Review D*, 97(6), 061501. doi:10.1103/PhysRevD.97.061501

Ezquiaga, J. M., & Zumalacárregui, M. (2017). Dark energy after GW170817: Dead ends and the road ahead. *Physical Review Letters*, 119(25), 251304. doi:10.1103/PhysRevLett.119.251304

Pioneer Anomaly

Anderson, J. D., et al. (2002). Study of the anomalous acceleration of Pioneer 10 and 11. *Physical Review D*, 65(8), 082004. doi:10.1103/PhysRevD.65.082004

Turyshhev, S. G., et al. (2012). Support for the thermal origin of the Pioneer anomaly. *Physical Review Letters*, 108(24), 241101. doi:10.1103/PhysRevLett.108.241101

Cosmological Tensions

Riess, A. G., et al. (2022). A comprehensive measurement of the local value of the Hubble constant with $1 \text{ km s}^{-1} \text{ Mpc}^{-1}$ uncertainty from the Hubble Space Telescope and the SH0ES Team. *The Astrophysical Journal Letters*, 934(1), L7. doi:10.3847/2041-8213/ac5c5b

Di Valentino, E., et al. (2021). In the realm of the Hubble tension—a review of solutions. *Classical and Quantum Gravity*, 38(15), 153001. doi:10.1088/1361-6382/ac086d