One Sky Alien Array

An archive-first test of intermittent signal coverage and a sensitivity-bounded receiving network

Andrew T. Doemling

AI research assistance from ChatGPT (OpenAI), Claude (Anthropic), Gemini (Google DeepMind), and Grok (xAI)

Project codename Lost Voices

6 October 2026

Theoretical research proposal · Final draft · Not peer reviewed

Authorship note: the AI systems named above are research tools, not human authors, and cannot carry authorship accountability. The human author is responsible for every number and claim in this document.

Abstract

Intermittent radio technosignatures require a search-completeness model that includes observing times and the full candidate-selection pipeline. A geometric opportunity to intercept one burst is not the same as recovery by a search that requires a signal in all three on-source scans. This distinction prevents a direct conversion of published continuous-transmitter limits into intermittent-beacon prevalence limits. One Sky Alien Array proposes an archive-first test: reconstruct frequency-qualified observing windows in a public Breakthrough Listen release, measure end-to-end recovery with blinded injections, and compare full-scan and short-window searches under both all-ON and any-ON acceptance at matched review cost. A hypothetical homogeneous survey illustrates how temporal coverage enters prevalence inference without claiming a reanalysis of existing survey data. A second, conditional track considers a receive-only network of independent small stations. Under optimistic assumptions a 3 m dish requires about 3.2×10^{14} W of equivalent isotropic radiated power in a 1 Hz line at 10 pc for 300 s of signal integration. This restricts the accessible signals to very powerful emission toward Earth; it does not establish transmitter purpose or duty cycle. Reproducible geometric examples show narrow scheduling resonances, the cost of paired corroboration, and the role of longitude for a specified target geometry. A 50% daily phase-coverage goal requires at least 144 paired five-minute visits per target, before visibility, weather, and recovery losses. Differential Doppler, measured interference vetoes, and blinded evaluation determine whether those opportunities become useful evidence. No completed archive benchmark, instrument advantage, extraterrestrial detection, or contact is claimed.

Plain-language summary

A radio telescope can miss a signal because it is too faint, because it appears while the telescope is looking elsewhere, or because its pattern fails the search's rules. Those possibilities must be measured separately. Many nearby-star searches were designed to find a tone that persists across several observations. Their published limits cannot simply be reused for a brief burst that appears in only one of those observations.

One Sky Alien Array begins with existing public data. It would measure which possible signal timings were observed, test which synthetic signals survive the complete search, and publish both detections and blind spots. A later network of small receiving stations would be considered only if these tests justify it. Small dishes would require exceptionally powerful signals at interstellar distances, and using two dishes to confirm a burst consumes time that could otherwise cover another possible arrival. The proposal measures that tradeoff. Its first useful result could be a reproducible account of what has been searched, even if no unexplained signal is found and no network is built.

Scope and claims

One Sky Alien Array is an unbuilt concept. Neither the archive-analysis tool nor the receiving network has been built or tested. No new observations, detections, experimental results, contact, or demonstrated performance

are reported here. Numerical examples are calculations under stated assumptions. Sensitivity figures are optimistic planning values to be replaced by measurement; phase-coverage results are geometry under ideal detection, not observed recovery. Section 2 distinguishes published survey results from a hypothetical any-ON calculation and specifies the completeness needed for a valid future prevalence inference. The acceptance thresholds in Section 4 are proposed design gates for a future preregistered experiment. The project claims no invention of volunteer computing, distributed SETI, geographically separated verification, satellite screening, or machine-learning technosignature searches, all of which have substantial prior art reviewed in Section 3. Its intended contribution is a frequency-qualified temporal audit, a recovery-calibrated archive benchmark, and a testable resource-accounting framework for a sensitivity-limited receiving network.

1 The question this paper asks

1 1 Sensitivity and the intermittency hypothesis

A signal must be received, preserved, and investigated independently before its origin can be established. Increasing collecting area helps with faint signals, but it does not solve a temporal mismatch: a telescope pointed at the right field during the wrong interval receives no burst.

Table 2 shows that an illustrative 3 m dish with a good receiver needs about 3.2×10^{14} W of equivalent isotropic radiated power within a 1 Hz line at 10 pc for 300 s of signal integration. The Arecibo planetary radar provides a comparison of about 2×10^{13} W [23]. Ordinary terrestrial leakage would be far below this small-dish threshold at interstellar distances. A detectable signal would therefore have to deliver unusually large power toward the receiver. That fact does not identify its purpose, establish deliberate Earth targeting, or imply that it is intermittent.

Intermittency is nevertheless a useful hypothesis to test. Benford, Benford, and Benford describe cost-optimized beacon designs that favor powerful directional pulses rather than continuous isotropic illumination [1]. Gray examines modulation associated with a transmitter on a rotating planet [2]. These conditional models motivate a signal family; they are not evidence about extraterrestrial technology. A small-dish proposal should quantify that family's temporal coverage while retaining continuous signals and alternative behaviors as controls. Larger partner telescopes relax the power limit and widen the accessible family.

1 2 What survey nondetections can establish

The nearby-star surveys discussed here used three five-minute on-source scans interleaved with off-source comparisons, a cadence designed to distinguish persistent celestial narrowband signals from interference [23,24]. In Price et al., a surviving candidate had to appear in all three ON scans. A five-minute isolated burst can overlap an ON scan yet fail that selection by construction. Thus a published null in the accepted candidate list is not a measured null for a new any-ON burst search.

The correct question is how often a specified transmitter would survive the complete search, including the actual observing times, frequency coverage, sensitivity, and selection rules. Margot et al. already incorporate efficiency and duty cycle into prevalence inference [7]. Section 2 uses that principle, separates geometric opportunity from calibrated recovery, and shows why a release-specific audit and benchmark are necessary before interpreting intermittent-transmitter prevalence.

1 3 Two tracks, one evidence record

A complementary starting point is telescope data that has already been recorded. Doemling proposes an opt-in downloadable tool, accessible through an AI assistant or skill, that lets participants analyze public observations and contribute results to a shared candidate database. This paper develops that concept into a bounded scientific workflow whose first product is a temporal-coverage audit of one documented public archive release, followed by a blinded benchmark of an intermittency-aware search against strong existing pipelines. The audit links the two tracks: it tests whether the hypothetical coverage gaps in Section 2 occur in the selected release and whether a prospective observing experiment could address them.

Intermittency has plausible physical motivations beyond beacon economics. A narrow transmitting beam concentrates energy but may sweep past Earth briefly; propagation and receiver-side visibility can also create intermittent reception. A search can be beautifully optimized for the wrong assumptions, and continuous narrowband carriers remain mandatory controls. The multidimensional "cosmic haystack" analysis shows why an unqualified claim that a null survey excludes extraterrestrial technology would be unjustified [3]. A useful project states the signal family to which it is sensitive, measures the completeness of that search, and publishes its blind spots.

The proposition developed here is specific: make geographically separated observations overlap when independent verification is valuable, but make their repeat visits sample different possible transmission phases. Those objectives compete for telescope time. Two stations looking together spend two station-hours on one hour of sky exposure; two stations looking separately can cover more sky or more time but cannot directly corroborate the same short burst. Section 7 shows that geography partly resolves the competition and exactly where it does not. The scientific task is to measure the tradeoff, not to assume that adding stations must help.

2 Temporal opportunity and prevalence inference

2 1 Geometry and the actual selection function

Define a periodic signal by period P, burst duration B with $0 < B \le P$, and unknown start phase φ, uniform over one cycle. For an on-source scan of duration V starting at a, require overlap of at least τ, with $0 < \tau \le \min(B, V)$. The admissible burst-start interval is $[a + \tau - B, a + V - \tau]$, of width $W = V + B - 2\tau$. Taking the union of these intervals modulo P gives the any-ON geometric opportunity:

$$G(P, B; \tau) = \frac{1}{P}\int_0^P I\,(\varphi; \mathcal{W}, B, \tau)\, d\varphi. \quad (1)$$

Here I is one when at least one scan meets the overlap requirement. For the examples below, V = 5 min and $\tau = \min(B/2, 2.5\ \text{min})$. Thus bursts of five minutes or less admit a five-minute phase window per scan; long bursts need only 2.5 min captured, not half of their duration. With three scans beginning at 0, 10, and 20 min, short bursts give G = 15 min/P once the windows are distinct. A 60 min burst merges those windows into an 80 min interval. At P = 24 h these opportunities are 1.04% and 5.56%, respectively. They describe a hypothetical any-ON rule, not the published all-three-ON selection.

Geometry alone does not measure detection probability. For target i and signal parameters θ, define $C_i(\theta)$ as the probability, averaged over unknown phase and the declared signal population, that the full observation and analysis process returns an accepted event. θ includes frequency, power, period, duration, linewidth, and drift behavior. C_i must include usable bands, repeated epochs, signal capture, interference masks, detection thresholds, and all candidate filters. Injection recovery through the complete pipeline estimates it. Under independent target trials and a homogeneous host fraction f,

$$P_0(f \mid \theta) = \prod_{i=1}^{N}[1 - fC_i(\theta)]. \quad (2)$$

Each physical target appears once; repeated observations contribute inside its completeness. Shared failures require an explicit dependence model rather than the product approximation. A 95% upper bound solves $P_0 = 0.05$ within $0 \le f \le 1$. If no solution lies in that interval, the chosen signal family has no upper bound below 100% at that confidence level. This does not mean that its likelihood contains no information.

2 2 Published survey facts and a separate toy example

Price et al. report observations of 1,327 distinct stars across heterogeneous receiver samples: 882 L-band targets, 1,005 S-band targets, and 189 Parkes targets, with different band coverage, sensitivity, and repeat observations [24]. These overlapping samples cannot be pooled into 1,327 equivalent trials at a common L-band threshold. The quoted 0.45% continuous-transmitter limit belongs to the L-band result. Margot et al. explicitly correct the 95% limits for assumed detection efficiency p = 0.5 to 0.68%, 0.59%, and 3.1% for the respective samples [7, Section V.3]. For example, the homogeneous binomial expression with N = 882 gives 0.678%.

The all-three-ON requirement is equally important. A short isolated burst that satisfies Equation (1) need not pass that requirement. Neither the geometry nor the published null establishes the result of a search that accepts any one ON scan. New limits for intermittent signals require a defined reprocessing search and calibrated C_i for its actual target sample. There is no empirical intermittent-prevalence reanalysis in this paper.

To illustrate the statistical effect of sparse coverage, consider an explicitly hypothetical survey of 1,000 equivalent targets in one common frequency band, with the three scans above. Assume any-ON acceptance and a duration-independent conditional recovery p = 0.5 whenever the geometric overlap condition is met. The latter is a simplifying assumption, not a measured pipeline efficiency. Only under these assumptions does C = pG, giving

$$P_0(f) = (1 - fpG)^{1000}, \qquad f_{95} = \frac{1-0.05^{1/1000}}{pG}. \text{ (2a)}$$

The corresponding even-odds prevalence is

$$f_{50} = \frac{1-0.5^{1/1000}}{pG}. \text{ (2b)}$$

Table 1 Hypothetical any ON survey with 1000 equivalent targets

Period	G for B ≤ 5 min	f_{95}	f_{50}	G for B = 60 min	f_{95}
Continuous	100%	0.598%	0.139%	100%	0.598%
24 h	1.042%	57.4%	13.3%	5.556%	10.8%
2 days	0.521%	>100%	26.6%	2.778%	21.5%
7 days	0.149%	>100%	93.1%	0.794%	75.4%
27 days	0.0386%	>100%	>100%	0.206%	>100%

Note: Calculated illustration, not a limit on a published survey. N = 1,000, p = 0.5, V = 5 min, scans start at 0, 10, and 20 min, and τ = min(B/2, 2.5 min). ">100%" means the equation's crossing lies beyond physically allowed prevalence. For f_{95}, this means no nontrivial 95% upper bound; for f_{50}, even a transmitter at every target leaves a null more likely than a detection. The assumed common sensitivity is held fixed solely for the toy calculation. Real thresholds depend on captured duration and must enter C_i.

For short bursts the hypothetical 95% bound reaches 100% at about 1.74 days, where the probability of a null at f = 1 is 5%. A null becomes more likely than a detection only beyond about 7.52 days, where that probability exceeds 50%. At seven days it is about 47.5%. Those are different statistical statements. The toy example demonstrates the importance of completeness without attributing an unmeasured result to Price et al.

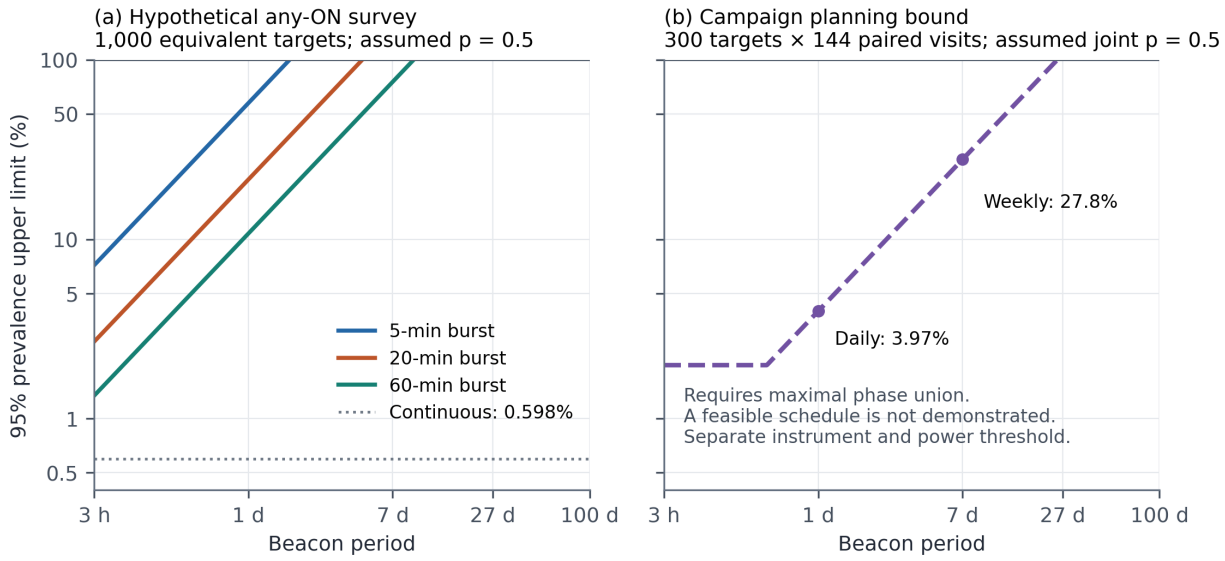


Figure 1 Hypothetical prevalence limits and an optimistic campaign bound. Solid curves use Equation (2a) for 1,000 equivalent targets, p = 0.5, and any-ON geometry for three five-minute scans. The continuous reference is 0.598%. Values above 100% yield no upper bound below 100% at 95% confidence. The separate campaign panel uses 300 targets and the exposure ceiling G ≤ min(1, 144W/P), with W = 5 min and an assumed joint recovery p = 0.5. It is an optimistic lower envelope for the achievable prevalence bound, not an evaluated schedule or a performance forecast. The campaign and toy survey need not have the same sensitivity; no direct instrument comparison is implied.

2 3 A geometric intermittent transmitter figure of merit

Enriquez et al. define the Continuous Waveform Transmitter Figure of Merit, $CWTFM = \zeta_{AO} EIRP_{\min} / (N_{stars} \nu_{rel})$, normalized so that a survey reaching 10^{13} W over a fractional bandwidth of 0.5 on 1,000 stars scores 1 [23]. Lower is better. Their reported score is 0.85; Price et al. report an improvement factor of 6.95, corresponding to about 0.12 [24]. For a fixed common band, common sensitivity convention, and specified burst family, a useful geometric scheduling heuristic is

$$ITFM(P, B; \tau) = \zeta_{AO} \frac{EIRP_{\min}}{\nu_{rel} \sum_i G_i(P, B; \tau)}. \quad (3)$$

It reduces to CWTFM when every $G_i = 1$. A thousand targets each with G = 0.01 contribute ten target-equivalents of geometric opportunity. That is not ten calibrated detection trials: the recovery and acceptance losses still belong in C_i. Different bands and duration-dependent thresholds must not be hidden in one score. The archive audit reports this heuristic by frequency-qualified sample alongside the full coverage surface; prevalence inference uses Equation (2). Margot et al.'s existing efficiency and duty-cycle treatment [7] is the methodological foundation, rather than a novelty claimed here.

2 4 What follows from the calculation

The calculation motivates a measurement of temporal coverage and selection, not a claim that surveys have excluded nothing or that intermittent transmitters are common. It cannot establish a transmitter's purpose or preferred schedule. Published continuous-carrier searches remain informative within their stated assumptions and corrections. The contribution proposed here is to audit actual per-target, frequency-qualified windows and measure end-to-end recovery for a declared intermittent family, then use those quantities to interpret either detections or nondetections.

3 What earlier work already establishes

Both distributed observing and distributed computing have substantial prior art. Project Argus described a worldwide network of small privately operated radio telescopes, including independent verification and the challenge of intermittent signals, and bought sky coverage with wide beams at the price of range [4]. One Sky

treats that history as a foundation and claims no invention of citizen SETI, geographically separated verification, or simultaneous observations.

Modern dual-site work makes that distinction concrete. Johnson and colleagues conducted a simultaneous technosignature search using international LOFAR stations in Ireland and Sweden at 110–190 MHz, demonstrating a noninterferometric route to using geographically separated stations for interference rejection and underscoring why frequency comparisons require corrections for the stations' different motions [5]. Section 6 makes that correction explicit.

Intermittency is an established research subject. Gray examined possible planetary-day modulation and the value of much longer temporal coverage [2]. Suresh and colleagues developed BLIPSS, a periodic-signal search applied to 4–8 GHz Galactic-center data for periods of 11–100 s and duty cycles of 10–50% [6]. Those examples show that a technosignature search need not be restricted to an uninterrupted spectral line; they do not establish which temporal behavior, if any, real transmitters favor.

The relevant methodological standard is recovery, not pipeline activity. Margot and colleagues used end-to-end signal injections and treated signal duty cycle in interpreting a radio technosignature survey [7]. The SETI@home analysis used injected synthetic signals and accumulated evidence across a long observation history [8]. A list of millions of detected features is not a measurement of sensitivity unless the experiment also determines which signals the process misses.

SETI@home is the direct predecessor for downloading radio data to volunteer computers [8]. Its history carries a lesson this proposal takes seriously: the project ended active distribution in 2020 not because volunteer computation failed, but because the scientific bottleneck had moved to analyzing, de-duplicating, and interference-screening the accumulated results. BOINC provides established mechanisms for replicated task validation [9]. Replication checks computation; it does not turn repeated processing of one observation into independent astronomical evidence.

Modern archive searches provide strong algorithmic comparisons. Breakthrough Listen's nearby-star surveys established the ABACAD cadence and the narrowband drift search [23,24]. Ma and colleagues used deep learning to inspect observations of 820 nearby stars; their training favored signals present across all three ON scans, leaving shorter or intermittent signals outside that study's assessment, and the reported signals of interest were not redetected in follow-up [10,11]. Painter and colleagues' Pickles pipeline applied comparatively simple statistics to a large Breakthrough Listen archive sample and recovered signals found by both conventional and machine-learning searches; its stated limitations include losses from intermittent signals and dense interference [12]. These limitations motivate a benchmark; they do not establish that missed extraterrestrial signals are present.

Interference from satellite constellations is now a research field of its own. Hellbourg's primer summarizes the protection problem in the era of mega-constellations [29]; Nhan and colleagues describe the NRAO–SpaceX Operational Data Sharing system and Telescope Boresight Avoidance, with trials at the Very Large Array in the 1990–1995 MHz and 10.7–12.7 GHz bands [30]; and the ITU-R RA.769 protection criteria remain the formal benchmark for a station's own noise budget [31]. Section 6 inherits that stack and adds a veto whose damage to real injections is measured.

The most instructive caution is scientifically productive. The Breakthrough Listen candidate blc1 initially displayed attractive properties, yet follow-up analysis identified a pattern consistent with local interference and intermodulation [13]. The lesson is not to abandon unusual signals. It is to preserve enough information to discover what made them unusual. One Sky's design must reward correct rejection as well as genuine recovery, without making the rejection rules so rigid that a real intermittent transmitter is automatically discarded.

4 Hypotheses and preregistered endpoints

4 1 The archive hypothesis

The archive track asks whether a declared intermittency-aware search improves held-out injection recovery relative to strong frozen baselines at matched searched exposure, review workload, and computing budget.

Integration duration and ON-scan acceptance must be varied separately in the 2 × 2 comparison of Section 5 5. An optional AI ranking arm must be evaluated separately, so an improvement from shorter search windows cannot be credited to AI without an ablation. A second engineering test measures validated throughput after transfer, replication, and central recomputation. Faster unvalidated output is not the endpoint. Archive results report recovery by signal family, surviving background candidates per unique searched GHz-hour, reviewer time, device-hours, and transferred bytes. Reprocessing the same file does not add new sky exposure.

4 2 The observing hypothesis

The observing hypothesis is a resource-matched comparison. In a preregistered population of synthetic intermittent beacons, phase-diverse paired revisits recover a higher fraction of independently corroborated source realizations than the strongest preregistered practical baseline, under equal aggregate station-hours, frequency coverage, processing budget, and network-wide false-alert budget. The primary population must contain signals physically detectable by at least two participating stations; results for weaker signals form a separate, mandatory analysis rather than disappearing from the denominator.

The injection population, locked before any held-out score, spans periods of 10 minutes, 1 hour, 6 hours, 23.93 hours, 24 hours, 2 days, 7 days, and 27 days; duty cycles of 0.001, 0.01, 0.05, 0.2, and 0.5; drift that is constant, linear between −4 and +4 Hz/s, or weakly quadratic; uniformly randomized initial phase; and amplitudes bracketing the two-station detection boundary, with a known fraction within 3 dB of threshold. The first registration is restricted to unresolved narrowband carriers; pulsed and broadband classes are a later registration, not a silent expansion. Aperiodic bursts, continuous carriers, periods outside the grid, and nonlinear drift form separate challenge populations.

The primary endpoint is the fraction of injections recovered as corroborated events, defined in Section 6 6. The proposed practical gate is an absolute recovery improvement of at least 10 percentage points over the strongest practical baseline, with a one-sided 95% block-bootstrap lower confidence bound above that margin and the background-alert bound within budget. This is a demanding design requirement, not an expected result. Figure 3 does not establish it: its single-site greedy schedule gains only about 8.01 points over fixed transit at exact sidereal resonance, and randomized scheduling performs comparably to the greedy policy. A feasibility pilot must establish a defensible primary population and sample size before committing to a held-out trial.

For that registration, define the primary day-scale stratum explicitly: periods at the sidereal day and 24 h plus neighborhoods whose widths are fixed from the campaign duration before testing; duty cycles 0.001, 0.01, and 0.05; stated weights for each period and duty-cycle cell; and a frozen amplitude distribution around the measured two-station threshold. Record the exact neighborhood widths, amplitude law, and denominator after commissioning and before any held-out labels are opened. Until those choices are registered, the confirmatory experiment is not fully specified. A relative gain R_proposed/R_baseline − 1 is secondary, with the absolute difference reported when baseline recovery is zero. Report all individual cells and broad-prior results to expose dependence on those weights.

The operational background-alert target is an upper 95% bound no greater than one surviving background candidate per 1,000 aggregate usable station-hours. A smaller or statistically inconclusive gain does not pass the practical gate; it does not prove zero scientific effect. Development data may justify changing the design before registration, after which the Open Science Framework record fixes software hashes, signal populations, thresholds, satellite rules, monitor windows, analysis, and stopping criteria [32]. A failed gate is publishable.

4 3 An observing window rather than a guess about aliens

Here phase means position within a hypothetical beacon's temporal cycle, not radio-frequency wave phase. Equation (1) applied to a planned schedule is a design quantity: it exposes temporal holes and repeated observations of the same possible phase. A frequently used shortcut, $1 - (1 - \delta)^m$ for m visits and duty cycle δ, assumes independent random transmission phases. Regular revisits violate that independence in a specific way that Section 7 quantifies: a nightly schedule samples phases on a lattice, and a lattice is excellent for most periods and catastrophic for some. No schedule can be best for every unknown population. The period

distribution used for scheduling is an engineering prior, and the final evaluation must include populations that disagree with it.

5 The archive experiment

5 1 What the tool is for

SETI@home demonstrated that volunteers will donate computation; it also demonstrated that computation was not the scarce resource. The scarce resources in a technosignature search are reproducibility, candidate review, and an honest record of what has and has not been searched. The proposed tool is therefore designed first as a reproducible analysis and review environment with a shared evidence registry, and only second as a way to distribute computation. Whether distribution adds useful validated throughput after transfer, replication, and review costs is an engineering question the pilot answers; the registry and review workflow remain valuable even if the answer is no. Runtime, storage location, transfer cost, and expert-review capacity are measured on the selected workload.

5 2 The temporal coverage audit

Before any new search is run, the archive's own metadata can address the geometric question illustrated in Section 2: how much of a hypothetical beacon's cycle has already been observed, target by target? Begin with one documented Breakthrough Listen release and a declared target list [14,15]. For each target, reconstruct actual on-source start and stop times, pointing, cadence links, frequency coverage, and duplicated or overlapping records. Preserve the distinction between missing metadata and a confirmed absence of observing. Where only catalogue times are available, label the result nominal geometric opportunity; a later pass uses usable-band and quality information to refine it.

For each declared period P, burst duration B, and required overlap τ, compute the union of recoverable burst-start intervals modulo P as in Equation (1), merging overlaps so that three nearby scans do not count a shared phase three times. Observations in disjoint frequency ranges do not form one continuous spectral search; publish results by searched band. The first release contains per-target coverage versus period and burst duration, the sample distribution, the intermittent figure of merit of Equation (3) on the locked grid, frequency-qualified observing windows, exclusions, and independently testable code. Publish both a broad, explicitly weighted period summary and the deepest coverage notches: a plot weighted toward resonances misrepresents average performance, and a smooth average conceals blind spots. If the audit finds that day-scale, low-duty-cycle coverage is already high on the chosen release, the motivating hole is not in that dataset and the network is not justified by it. That stop condition is a result.

5 3 Public measurements and bounded work units

The scientific inputs are numerical measurements, usually time–frequency power arrays in filterbank or HDF5 products; selected observations also retain voltage data. An observation manifest records the official archive identifier and URL, source checksum, acquisition time system, sky pointing, frequency reference frame, resolutions, polarization products, calibration, quality flags, and ON/OFF cadence relationships. Public accessibility and permission to redistribute are checked separately for each selected release; a software license does not establish the license of every data file.

The feasibility pilot must name the exact release, product, and at least one complete usable cadence in a checksum-bearing manifest before a benchmark is promised. The Price et al. high-spectral-resolution product has approximately 18.25 s time samples and 2.79 Hz channels [24]; Painter et al. describe comparable approximately 18 s fine-frequency data [12]. A genuine 10 s narrowband window cannot be reconstructed from those power samples. A faster product with coarser spectral channels is a different signal domain, and interpolation does not restore discarded time resolution. Raw-voltage reprocessing is an optional, separately justified path only where the required recordings and rights are available. The proposed archive trial remains conditional on that product-level check.

Raw archives are cut, near the archive, into immutable frequency-slice bundles of at most 128 MiB that keep the full ON/OFF time axis, calibration bits, and quality mask, with drift-dependent padding and explicit ownership of overlapping boundaries so that a track is neither lost nor counted twice. Each bundle is hashed. Dynamic remote subsetting of HDF5 is out of scope because it makes a volunteer workload unreproducible. The bundles are a method, not a result.

5 4 Numerical processing and the AI interface

Pinned numerical engines load data, apply documented masks, search drift and time-window patterns, and produce measurements. turboSETI supplies the conventional narrowband drift-search baseline [16]; Pickles supplies an independently defined statistical search on the same input observations [12]; blimpy provides the format-handling layer [18]. The intermittency-aware arm is a registered third path restricted to the locked period grid; it does not get to redefine a hit after the score.

A learned spectrogram model may rank candidates or flag interference morphology, using frozen weights and an independently assessed operating point, initially supplementing the baseline streams with a prespecified audit sample of low-ranked cases. A language-model interface can explain computed results, retrieve documentation, and request approved jobs through a narrow API. It cannot supply missing measurements, revise thresholds after seeing test results, or turn a model score into a probability of alien origin. The numerical worker remains usable without a language-model subscription, and sending any data to an external AI provider requires a separate explicit opt-in. Signed declarative manifests select reviewed installed algorithms, input hashes, bounded parameters, expiry, and expected outputs; they do not carry arbitrary shell commands or downloaded executable code.

5 5 A duration bank and a separate acceptance test

For each frequency-drift template, integrate the dedrifted power time series over a declared bank of integer-bin windows and normalize against a locally estimated noise distribution. On an approximately 18.25 s product, an illustrative bank is 1, 2, 3, 8, and 16 bins, or approximately 18.25, 36.5, 54.75, 146, and 292 s, together with the full available scan. The actual manifest fixes the bin size, complete-sample counts, window stride, edge treatment, masking, and trials. Inject sub-bin burst phases and account for intra-bin drift smearing. Shorter windows require a suitable higher-time-resolution product.

The ideal motivation is straightforward. For a constant-amplitude burst of duration b fully captured in a matched window w = b with stationary independent noise, relative to integrating the same signal over a full scan T,

$$\frac{(S/N)_{w=b}}{(S/N)_T} = \sqrt{\frac{T}{b}}. \quad (4)$$

For T = 300 s the ideal ratios are 3.87 for a 20 s burst, 2.24 for 60 s, and 1.41 for 150 s. These are pedagogical values, not realized gains in the cited archive. Finite bins, partial capture, channelization, drift, correlated noise, and threshold recalibration change them. The full-scan statistic remains in the bank, but adding trials can still reduce continuous-signal recovery at a matched background budget. The effective trials count must be measured or conservatively calibrated; it is not fixed by the number of nominal window lengths.

The benchmark separates two changes in a factorial design: full-scan integration with all-ON acceptance; full-scan integration with any-ON acceptance; short-window integration with all-ON acceptance; and short-window integration with any-ON acceptance. The same input observations, blinded injection realizations, frequency and drift domain, and background-calibrated reviewer budget apply to all four cells. All-ON means a qualifying detection in each of the three ON scans; any-ON means one or more qualifying ON detections. Both apply a frozen, sensitivity-aware OFF-source association rule. Preserve the unmodified, pinned turboSETI and Pickles pipelines as additional named baselines, since modifying their acceptance changes the algorithm being assessed.

An isolated short burst can benefit automatically from relaxing all-ON acceptance. The factorial comparison reveals whether shorter integration adds recovery beyond that change and what each change costs in interference and review time. The primary archive result is this recovery-versus-cost surface. An optional

learned-ranking ablation follows it; no gain from changed integration or acceptance is credited to AI. Boxcar searches for fast radio transients provide prior art for duration banks [28], while BLIPSS addresses periodic folding in a different temporal domain [6].

5 6 Consent, reliable execution, and untrusted submissions

Participants explicitly choose CPU and optional GPU use, memory, disk cache, active hours, idle and battery behavior, bandwidth, daily transfer, and per-job time limits. A cautious pilot could begin CPU-only, idle and plugged-in, with one core, 2 GiB RAM, a 2 GiB cache, and at most 1 GiB of daily transfer. Local settings override the scheduler; pause, cache deletion, and uninstall must be straightforward. Workers run unprivileged with access limited to dedicated input and output directories. Reviewed signed releases, input-hash checks, strict schemas, parser and decompression limits, and bounded network destinations reduce risk. Checkpoints, expiring leases, capped retries, idempotent uploads, and server backpressure prevent failed jobs from producing unbounded work.

Volunteer submissions remain untrusted. Replication across unrelated hosts, hidden controls, and random audits can reveal errors, while important candidates are recomputed on a trusted reference system using original data; numerical agreement tolerances must be specified for different hardware [9]. Account reputation or a majority of matching answers cannot exclude coordinated poisoning. Public records use pseudonymous contributions, and the service retains no unnecessary personal files or device access.

5 7 A shared database of evidence and coverage

The database stores completed searches and null results as well as candidates. Its minimum scientific record joins the observation and slice identifiers; source hashes; pipeline, model, calibration, and mask versions; parameters and seeds; usable searched coverage; candidate time, frequency, drift, duration, and statistic definitions; bounded diagnostic cutouts; replication outcomes; and review history. Exact analysis identifiers deduplicate repeated computation. Rejections and corrected interpretations remain visible in an append-only audit trail. Unsearched regions, failed jobs, masked bands, and analyzed nondetections have distinct states; otherwise the database would exaggerate search completeness. Candidate states distinguish an unverified submission, reproduced numerical result, explained interference, unexplained signal of interest, and independently observed candidate. Crowd labels, model labels, and expert decisions remain separate, and public submissions do not automatically become training truth.

6 A physically credible receiving system

6 1 Independent stations and honest sensitivity

The initial architecture uses independent receiving stations rather than a phase-coherent interferometer. A station enters the corroboration sample only with a characterized antenna and receiver including its sidelobes, a measured system temperature, a stable time and frequency reference, known pointing, a logged local-interference record, and the same pinned pipeline as the archive. Stations share observation schedules and candidate records and, where possible, preserve calibrated voltage or spectral data around an event for independent reanalysis. Stations must meet the commissioning requirements in Sections 6 4, 6 5, and 8 2; all remaining thresholds must be locked before testing. Stations that fail those requirements do not enter the corroboration sample.

This architecture allows candidate-level coincidence. It does not add dish areas, produce high-resolution localization, or lower the flux threshold by the square root of the number of stations. Coherent combination requires phase calibration, precise delay models, compatible signal paths, and an actual beamforming or correlation system; neither coherent nor incoherent combination is assumed. Internet links, including satellite-delivered internet, can transport telescope data; a communications connection does not itself provide calibrated sky measurements or collecting area.

For a circular dish with diameter D and aperture efficiency η_a, effective collecting area is $A_{eff} = \frac{\eta_a \pi D^2}{4}$. The system-equivalent flux density relates the system temperature T_{sys} to collecting area [17]; lower SEFD means better sensitivity:

$$SEFD = \frac{2k_B T_{sys}}{A_{eff}} \times 10^{26}\ Jy \quad (5)$$

For a detected narrowband signal whose integrated received flux is F, an ideal radiometer approximation gives the limiting flux, where q is the detection signal-to-noise requirement, W_{eff} the matched noise bandwidth, n_{pol} the number of independent polarization power streams, t_{on} the time during which the signal is present, and ℓ the fraction of signal power captured by the complete system:

$$F_{\min} = \frac{q\ SEFD \times 10^{-26}}{\ell} \sqrt{\frac{W_{eff}}{n_{pol} t_{on}}}\ W\ m^{-2} \quad (6)$$

This is a planning approximation, not a substitute for injecting signals through the actual acquisition and search pipeline. Pointing errors, receiver drift, propagation, digitization, polarization mismatch, and processing losses reduce ℓ. A 300 s observation does not provide 300 s of signal integration if the beacon appears for only 20 s.

A small dish also has a wide beam. Near 1.42 GHz the half-power beamwidth is about 4.8° for a 3 m dish, 12° for a 1.2 m dish, and 0.15° for a 100 m dish; the coefficient depends on illumination and these are estimates to be replaced by a measured beam pattern. A 5° field contains many stars at all distances, so a small station "pointing at a star" is pointing at a field, and ON/OFF nodding must move the antenna many degrees to leave it. Interference entering through the broad main beam and sidelobes is harder to separate by pointing alone, and the beam is also why low-Earth-orbit satellite passes matter so much in Section 6 4. This is a further reason that the celestial-origin test of Section 6 3 and multi-site coincidence carry more weight for small stations than for large ones, and a reason that archive-tuned ON/OFF logic cannot be transferred to small dishes without re-validation.

6 2 An illustrative link budget

Table 2 Illustrative single dish carrier thresholds at 10 pc

Diameter m	Efficiency	System K	SEFD Jy	Beam deg	EIRP at 10 pc W	Relative to Arecibo radar
1.2	0.60	100	406,921	12.1	1.99×10^{15}	99×
3	0.60	100	65,107	4.8	3.18×10^{14}	16×
6	0.60	100	16,277	2.4	7.95×10^{13}	4×
25	0.65	35	302.9	0.58	1.48×10^{12}	0.07×
100	0.70	25	12.56	0.15	6.13×10^{10}	0.003×

Note: Calculated examples, not measured One Sky or named observatory performance. A 1 Hz matched signal, 300 s on-signal time, two independent polarization power streams, q = 10, and full signal capture are assumed in every row. The beam is 1.2 λ/D at 1.42 GHz. The last column divides by an Arecibo planetary radar EIRP of 2×10^{13} W [23]. The 25 m and 100 m rows use separately assumed specifications to illustrate the scale of larger-aperture collaboration, not the performance of a named observatory; the Green Bank Telescope at L-band has a published SEFD of about 10 Jy [23], so the 3 m example is about 38 dB less sensitive.

The value q = 10 is a nominal threshold relative to the noise-only radiometer RMS, not a measured recovery probability or a calibrated global false-alarm significance. The 1 Hz quantity is matched signal bandwidth, not the total searched band. Equivalent isotropic radiated power is the transmitter's effective power toward the receiver expressed as the power an isotropic radiator would need; the tabulated EIRP is integrated power in the 1 Hz-contained line and is not the transmitter's electrical consumption. At distance d, free-space propagation gives:

$$EIRP_{\min} = 4\pi d^2 F_{\min} \quad (7)$$

For the illustrative 3 m dish, the threshold is about 3.18×10^{14} W at 10 pc, 3.18×10^{16} W at 100 pc, and 3.18×10^{18} W at 1,000 pc. These values delimit the question a small station can ask: whether a sufficiently powerful signal is directed toward Earth, in the observed band, during the usable observing window. Sensitivity improves only as the square root of on-signal integration time in this model: increasing captured duration from 300 to 1,200 s halves the limiting EIRP; reducing capture to $\ell = 0.7$ raises the threshold by about 43%; an equal-duration ON-minus-OFF measurement adds approximately $\sqrt{2}$ to the noise; a 150 s burst raises the 300 s threshold by $\sqrt{2}$. A station substantially less sensitive than its partner becomes the bottleneck for mandatory coincidence: the 1.2 m and 3 m examples differ in SEFD by a factor of 6.25, and the weak station's nondetection then provides little evidence against a signal. Stations should be paired by measured capability, and the pipeline should predict whether a partner should have detected a candidate before applying a veto.

6 3 A calibrated differential-Doppler test

A celestial signal received at separated stations does not occupy identical topocentric frequencies. Earth's rotation and orbital motion, transmitter motion, and receiver-frequency errors influence the observed frequency track, and the correct comparison is agreement with a shared sky-direction and motion model within calibrated uncertainty [5]. Orbital motion is common to all stations and cancels in a between-station comparison; rotation does not. For a source at declination δ observed from station i at geodetic latitude ϕ_i and hour angle H_i, the rotational line-of-sight velocity is $v_i = \omega_\oplus R_\oplus \cos\phi_i \cos\delta \sin H_i$, so the predicted between-station frequency offset for a transmitted frequency f is:

$$\Delta f_{12} = \frac{f\,\omega_\oplus R_\oplus \cos\delta}{c}(\cos\phi_1 \sin H_1 - \cos\phi_2 \sin H_2) \quad (8)$$

At 1.42 GHz the prefactor $f\omega_\oplus R_\oplus / c$ is about 2.2 kHz. Two stations at latitude 40° separated by 15° of longitude, observing an equatorial source symmetrically about the pair's midpoint, differ by about 440 Hz, and the differential drift rate between them, the time derivative of that offset, has an amplitude of about 0.032 Hz/s. Both are computable from station coordinates and time to far better than 1 Hz; a GPS-disciplined oscillator at 10^{-11} fractional stability contributes about 0.014 Hz. Operational software must handle clock and local-oscillator offsets, geometric delays, time standards, Earth orientation, and higher-order corrections consistently [26,27].

A several-degree search beam dominates the prediction uncertainty. Moving the assumed direction by 2.4°, the half-power half-width of a 3 m dish, changes the predicted offset by less than 1 Hz for an equatorial source observed symmetrically about the midpoint, where the first-order terms cancel, and by up to about 14 Hz for a source at declination 45°, where the offset itself is about 311 Hz. Direction uncertainty therefore exceeds the oscillator term by three orders of magnitude yet remains a few percent of the predicted offset. A terrestrial transmitter fixed to the ground predicts zero differential offset; a geostationary satellite almost none; a low-orbit satellite a large and different offset predictable from orbital elements. The test separates a ground-fixed emitter from a celestial source with margin to spare; it does not localize the source within the beam. Search or marginalize over the permitted sky directions and include those trials in background calibration. Agreement with the sky model is evidence of consistency, not proof of extraterrestrial origin; small baselines or unfavorable geometry offer little discrimination, and calibrated injections must establish which cases the test can reject.

A linear drift search is a tractable starting point. Near 1.42 GHz, a drift of 1 Hz/s corresponds to a line-of-sight acceleration of about 0.21 m/s^2; a ±10 Hz/s search is a bounded engineering choice informed by the distribution of plausible transmitter accelerations [25]. With 1 Hz channels over 300 s, adjacent drift trials spaced by 0.0033 Hz/s differ by about one channel at the endpoint; a signal drifting 10 Hz/s smears across about ten bins within the 1 s transform, and shifting completed spectra cannot reverse that loss, so high-drift trials require raw-voltage dechirping or coarser resolution with a recomputed limit. Even a 1 MHz band contains roughly one million 1 Hz bins; approximately 6,000 drift trials imply around six billion nominal templates per scan, so a single-template "10 sigma" label cannot establish a network false-alert probability and empirical end-to-end background calibration is essential. Two polarization streams at 1 million complex samples per second with 8-bit I and Q produce about 4 MB/s, or 14.4 GB per hour; a 300 s raw segment is about

1.2 GB. Local ring buffers, candidate-triggered retention, and documented reduced products are practical necessities, and the ring-buffer duration and trigger latency must be measured.

6 4 Satellite ephemeris as a measured veto

Public two-line elements and general-perturbation element sets are the starting catalogue. CelesTrak publishes grouped sets without an account, including supplemental elements that describe actively maneuvering low-Earth-orbit spacecraft better than a stale classic set; Space-Track is the official catalogue and the source for history, decay, and objects CelesTrak does not carry [33,34]. Propagation uses an SGP4 implementation checked against the published test vectors before any candidate is scored [35]; a propagator that fails those vectors does not enter the pipeline.

Position coincidence inside a 4.8° beam can be a weak veto. Depending on sky direction and epoch, catalogued objects may occupy a substantial fraction of observing time; flagging every such coincidence can discard real sky signals. The loss must be measured rather than assumed. Three numbers are therefore registered, computed on commissioning data before any science score. Mask fraction: the fraction of station-time in which at least one catalogued object lies inside the half-power beam; if this exceeds 30% at 1.42 GHz, the station does not enter the corroboration sample at that frequency. False-veto rate: the fraction of injected celestial beacons removed by the satellite rule alone; a screen that deletes more than 5% of injections is rejected and the rule tightened before the trial. Residual leakage: the fraction of known satellite passes that produce a narrowband hit the screen fails to flag, reported but not required to be zero.

The discriminating observable is angular rate. A celestial source moves at sidereal rate; a low-Earth-orbit object crosses a 4.8° beam in seconds to a few tens of seconds. Events whose amplitude envelope matches the predicted crossing are vetoed. Events that persist across a sidereal dwell, match a second station that does not share the same pass, satisfy Equation (8), and are absent from the local monitor are retained.

6 5 Operator coordination, local monitors, and independent transmit evidence

Steered satellite downlinks and unintentional emission from satellite electronics are different problems. The National Radio Astronomy Observatory and SpaceX operate an Operational Data Sharing system in which a participating telescope publishes where it is pointing, at what frequency, and over what bandwidth, at a prearranged lead time; the constellation applies Telescope Boresight Avoidance, taking no action outside an agreed outer region, steering downlink beams away from the site, typically by about 180 km, inside it, and briefly disabling beamforming inside an inner region [30]. Trials at the Very Large Array have shown reduced downlink interference in the 1990–1995 MHz and 10.7–12.7 GHz bands. This proposal uses that system where it is actually offered and does not invent membership it does not have. A station accepted into an operator feed publishes its schedule and logs every interval in which avoidance was requested and confirmed. Avoidance is a reduction in steered downlink power, not a veto: candidates in an avoidance window still pass the ephemeris screen, the local monitor, and the second site. The published system does not address unintentional electromagnetic radiation from onboard electronics, so out-of-band energy in the protected 1400–1427 MHz band remains in scope for the local monitor and the ephemeris screen. A quiet downlink beam is not a quiet satellite. User terminals near a station are a separate local interferer logged at site selection. Aggregate constellation noise is a limit this project does not solve; the single-interferer criteria of ITU-R RA.769 are used only as the commissioning benchmark for a station's own noise budget [31], and no claim is made that the band is clean [29].

Each station runs a broadband reference sensor, an omnidirectional antenna and a software-defined radio, in parallel with the primary feed. A veto requires coincidence inside one native integration of the candidate and a power excess above a locked multiple of the sensor's own median. A monitor that vetoes more than the locked fraction of celestial injections is a broken monitor, and the station drops out until it is fixed. SatNOGS, the Libre Space Foundation ground-station network, is the independent check on whether a catalogued object is known to transmit [36]: a candidate matching a SatNOGS detection of the same object at a consistent frequency and pass is labelled satellite and leaves the list. Absence from SatNOGS is not evidence of a celestial source; the catalogue is incomplete and a non-detection there clears nothing.

6 6 The corroboration gate

A candidate leaves the network as a corroborated event only if all of the following hold, each locked before the trial. It is present in the primary beam above threshold and absent from the local reference sensor inside the locked window. It is not matched to a catalogue beam-crossing under Section 6 4 and not matched to a SatNOGS transmit record for that pass. A second station, with shared satellite illumination and correlated interference explicitly assessed rather than assumed absent, records a hit at the frequency predicted by Equation (8) and the locked drift model, or a consistent independent redetection occurs within the common follow-up budget. For a later redetection, its phase window must come from a blinded period-search or prediction procedure, with period-template trials and association tolerances registered in advance. The injected true period is never supplied to the corroboration gate. Simultaneous two-site agreement can corroborate a single burst without establishing a recurrence period; a single-site pulse without independent confirmation remains outside the primary corroborated endpoint. The alert sits inside the network-wide false-alert budget, counted as background alerts per 1,000 usable station-hours. Anything short of this is a hit, published in an appendix, not a candidate. This is the operational content of the 2026 International Academy of Astronautics requirement that authentication use independent facilities and different instrumentation, applied before any public claim rather than after one [21].

7 Why the revisit schedule can change the result

7 1 An exact aliasing example

Consider a periodic beacon with a 5 min transmission once every 24 h and a start phase uniformly distributed over the cycle. A pair of stations makes eight simultaneous 5 min visits and requires at least 2.5 min of burst overlap in one visit. Assume adequate signal strength at both stations, ideal recovery, no outages, and visibility at every planned visit. One visit admits a 5 min interval of possible burst-start phases. Eight repeats at the same clock phase cover that interval again, leaving 5/1,440, or 0.347%, geometric phase coverage. Eight nonoverlapping phase windows cover 40/1,440, or 2.778%. Both schedules cost 80 aggregate station-minutes.

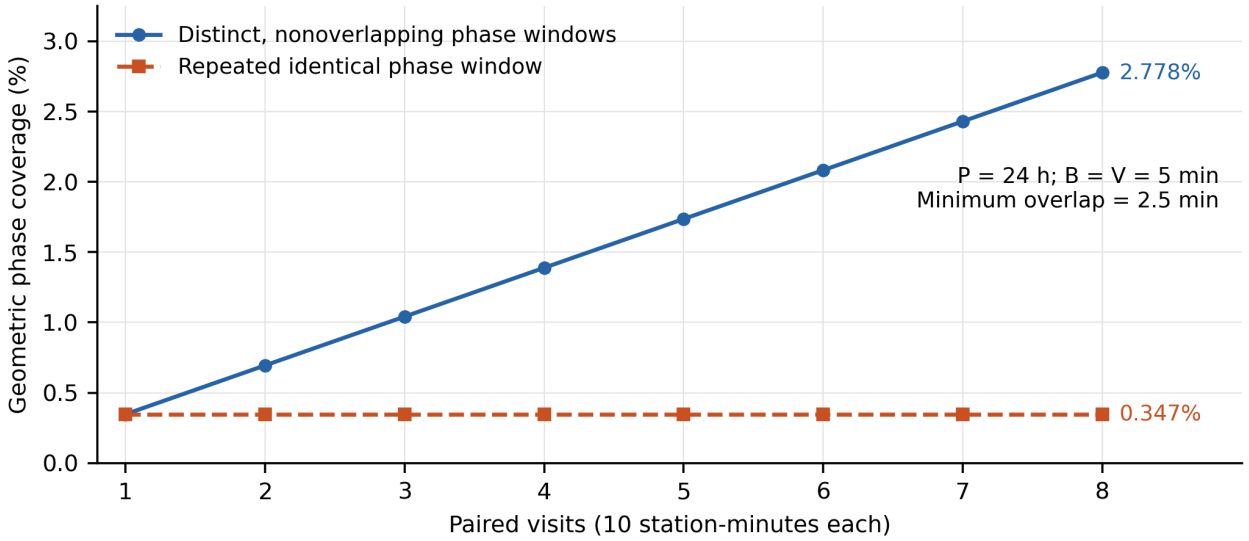


Figure 2 A constructed aliasing example. Eight 5 min paired visits cost 80 station-minutes. For a 24 h beacon period, 5 min bursts, and at least 2.5 min required overlap, distinct phase windows cover eight times as many possible start phases as identical repeats. Exact geometry under ideal detection.

7 2 Fixed schedules are the default, and they alias

The identical-clock-time schedule in Figure 2 is not a straw man. A star transits at the same local sidereal time every day, so an observer who observes at transit, or at any fixed hour angle, samples the same phase of any beacon whose period is one sidereal day or an integer fraction of it; an observer who observes at the same civil

time each night does the same for a 24 h period. Gray's planetary-day analysis is a special case of this aliasing [2].

Figure 3 replaces an assumed scheduler advantage with a reproducible comparison. A station at latitude 40° observes an equatorial source above 20° elevation; the spherical-Earth model gives a visibility interval of about 8.441 h per 23.93447 h sidereal day. This is a relative-time sidereal-cycle geometry, without dated observations or daylight, Sun-avoidance, or weather constraints. A real schedule must supply dates, station and target coordinates, and operating exclusions. Each of 30 cycles contains one 5 min visit, every complete visit remains inside its daily visibility window, and each policy costs 150 station-minutes. Four policies are compared. The fixed policy always observes at transit. The linear sweep moves the visit start from the first feasible start on cycle one to the last feasible start on cycle thirty. The random policy draws each night's start uniformly from the feasible range with a fixed seed. The phase-diverse policy is a greedy scheduler: each night it chooses, from a 5 min grid of feasible starts, the start that most increases geometric coverage averaged over a design prior of 160 periods spaced log-uniformly between 1 and 48 h. Ties within 10^{-12} in mean coverage choose the earliest candidate. Broad-prior results use a separate 160,000-point log-midpoint evaluation grid, supplemented by exact resonance and local zoom calculations, rather than interpolation across narrow notches.

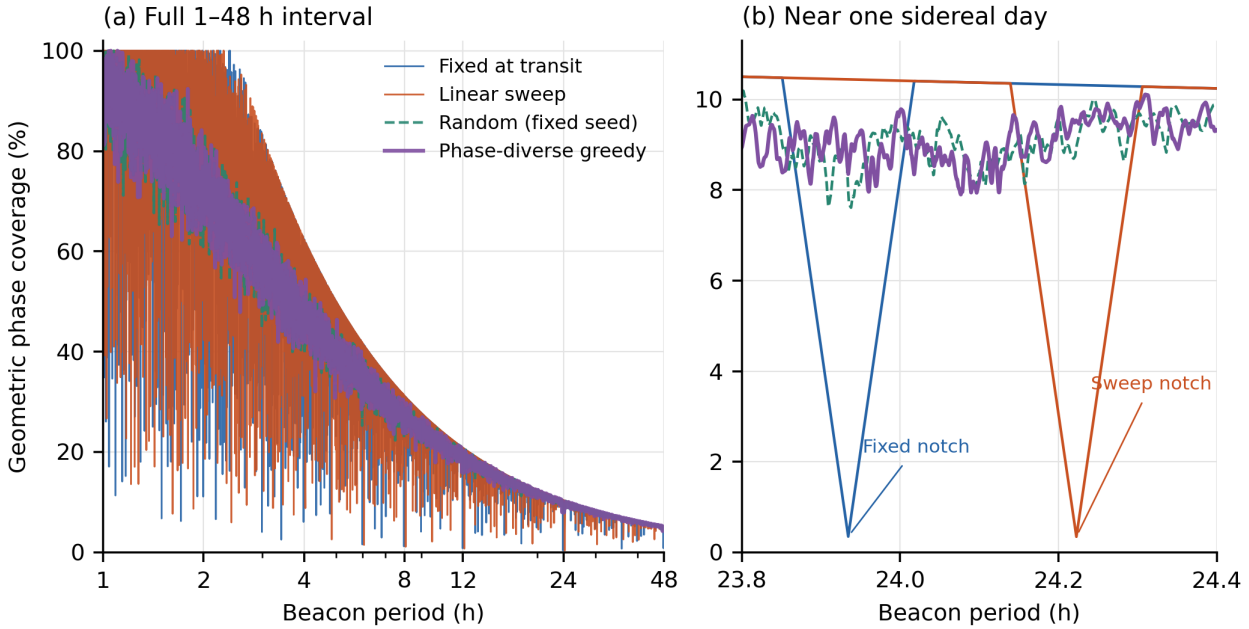


Figure 3 Thirty nightly 5 min visits from one station, four scheduling policies. (a) Periods from 1 to 48 h on a logarithmic axis. The fixed-transit and linear-sweep lattices have deep notches at resonant periods; the random and greedy phase-diverse policies avoid the deep lattice notches on the evaluated grid and sit about two percentage points lower on average. (b) The neighborhood of one sidereal day. The fixed policy's notch is at 23.93 h and the linear sweep's at 24.22 h; the greedy policy varies between about 8% and 10% without the deep fixed-cadence notch in this evaluated range. Exact geometry under ideal detection. Code in Appendix A.

Three results follow. First, averaged over the log-uniform prior, the policies are close: fixed and linear-sweep schedules each cover about 40.2% of possible phases, while random and greedy policies each cover about 38.2%. A regular spacing can be efficient over a broad prior even while failing at selected periods. Second, at exactly one sidereal day the fixed policy covers 0.348% while the greedy policy covers 8.356%, an absolute difference of 8.008 percentage points. The linear sweep is itself a uniform lattice with spacing D + (L − V)/(N − 1); its deepest notch is near 24.22267 h, with coverage 0.344%. Over the dense log-uniform evaluation grid, fixed and sweep policies each fall below half the greedy coverage for about 4.3% of periods. Third, random and greedy policies avoid those deep lattice notches in this evaluated example. The greedy policy ranges from 7.90% to 10.11% in the displayed 23.8–24.4 h neighborhood. These results support testing phase diversity but do not establish that a greedy scheduler improves on a randomized practical baseline. A nonlinear deterministic schedule also need not have the linear sweep's alias. Any claimed coverage floor is restricted to the evaluated prior, campaign, and grid and must be stress-tested.

If the true period is P = D + ε for daily spacing D, successive visits shift by approximately −ε in beacon phase per cycle, and the accumulated shift over N visits is about (N − 1)|ε| before wrapping. A period merely near a sidereal day therefore does not inherit a permanent single-site ceiling; the accessible phase changes over a long enough campaign. Finite duration, visibility, orbital timing corrections, and schedule variation must all enter a realistic audit.

7 3 Longitude and the cost of corroboration

Target visibility sets an upper envelope even with unlimited visits. For latitude φ, declination δ, and minimum elevation e, the spherical model gives $\cos H_0 = (\sin e - \sin \varphi \sin \delta)/(\cos \varphi \cos \delta)$, and the daily visible duration is $L = D\, H_0/\pi$. For a continuous visibility interval of length L, allowing a complete visit anywhere within it admits a burst-start interval of length L + B − 2τ, which for B = V = 5 min and τ = 2.5 min is simply L. At exact sidereal resonance the reachable fraction of phases from one site is therefore:

$$\rho = \min\left(1, \frac{L + B - 2\tau}{P}\right) \quad (9)$$

which for this equatorial source, latitude 40° station, and 20° elevation cut is 35.3% at exact sidereal resonance P = D and about 71% at P = D/2. These are geometry-specific ceilings. Circumpolar targets may remain visible continuously; periods away from exact resonance can sweep through different accessible phases during a sufficiently long campaign. Longitude diversity raises the ceiling in this particular resonant geometry. Two sites whose transit times are 6 h apart reach a union of about 60.3% of phases, and three sites spaced D/3 apart reach all of them. But their simultaneous-visibility intersections are small: 10.2% for the pair and 5.8% for the union of phases visible to at least two of the three. Detection somewhere in a network and instantaneous independent corroboration are different endpoints. Three designated pairs with 1 h internal transit separation and pair centers D/3 apart offer about 93.3% simultaneous paired-phase reach alongside 100% any-site reach; internal separations of 2 or 3 h reduce the paired envelope further.

The envelopes above assume unlimited visits. Figure 4 adds the budget. Six stations each make one 5 min visit per night for up to 120 nights, in three layouts: three pairs at a single longitude, three regional pairs with 1 h internal separation and centers D/3 apart, and six independent stations spaced D/6 apart. Pairs observe together, so a paired visit covers one phase window for two station-visits; the independent stations each cover one window per visit and corroborate nothing. At one sidereal day the two paired layouts coincide initially because both spend two station-visits per admitted phase window. Independent stations supply twice their any-site coverage in this regime. At 180 station-visits the paired layouts cover 31.3% of phases, while the independent layout covers 62.7%. With greater exposure the single-longitude pairs approach their 35.3% envelope, regional pairs approach 93.3%, and independent stations reach 100% any-site coverage without simultaneous corroboration. Tiling is an ideal allocation, not a multi-target operational schedule.

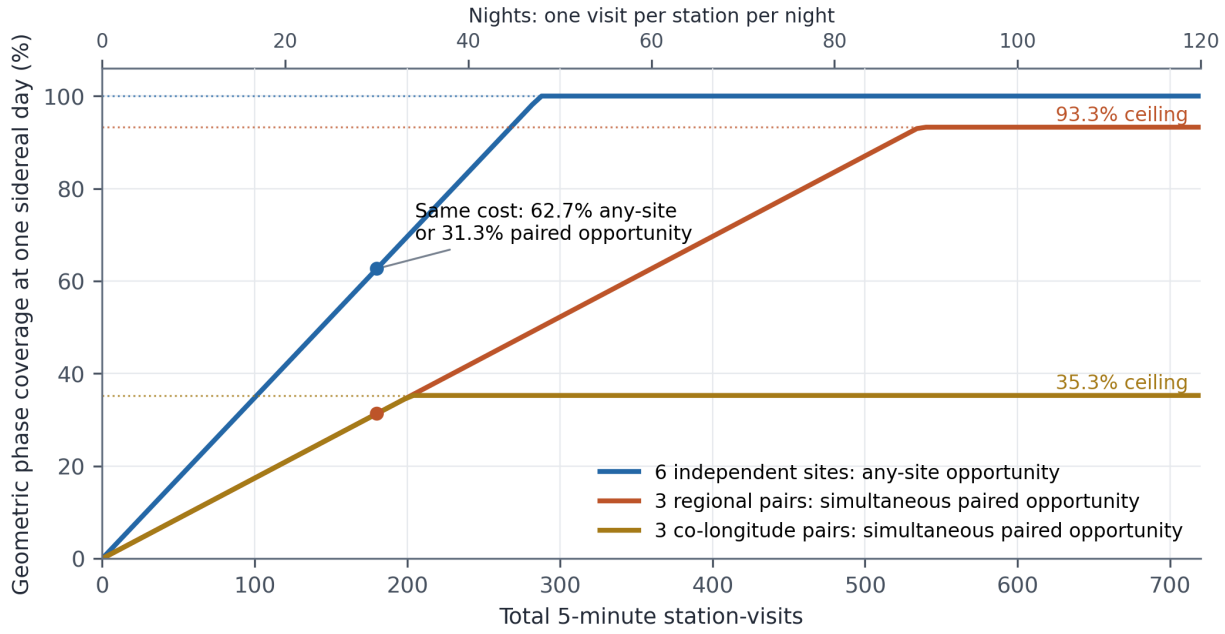


Figure 4 Coverage and corroboration at exact sidereal resonance. Six stations each have one five-minute visit per night; curves use ideal tiled allocations and the target geometry of Section 7 2. Dotted lines mark unlimited-visit envelopes. The last phase tile is aligned to its visibility boundary, and redundant visits after saturation remain in the budget. The paired layouts initially coincide, while independent stations provide twice their any-site phase coverage for the same station-visits. At 180 station-visits the values are 31.3% and 62.7%. Single-longitude pairs approach 35.3% and regional pairs 93.3%; independent stations reach 100% any-site coverage with no simultaneous corroboration. These are distinct endpoints. Code is in Appendix A.

The cost of phase completeness follows from the same geometry. With admitted window $W = V + B - 2\tau$ per visit, reaching geometric coverage C of a period P requires at least

$$N_{req} = \text{ceil}\left(\frac{CP}{V+B-2\tau}\right) \quad (10)$$

visits per target. Attaining this integer lower bound requires adequate visibility and sufficiently little phase overlap. For five-minute windows, 50% coverage of a 24 h cycle requires at least 144 visits and 90% requires 260. At 12 h the corresponding counts are 72 and 130; at 48 h they are 288 and 519. Mandatory pairing doubles station-visits. For a hypothetical 300-target campaign, 144 paired visits per target consume at least 7,200 station-hours of on-source exposure. Ten stations each delivering eight usable on-source hours per night would require at least 90 fully utilized nights. Visibility, shared target access, slews, calibration, weather, masks, and allocation conflicts can only increase that elapsed time; no feasible multi-target campaign has yet been constructed.

For 144 visits the union obeys $G(P) \leq \min(1, 144W/P)$. This is an exposure ceiling, not an achieved schedule, and a single fixed schedule need not attain it at every period. If a 300-target campaign attained that ceiling and had measured joint conditional recovery $p = 0.5$ at a specified common signal threshold, the homogeneous model would give an optimistic 95% bound of 3.97% for daily five-minute bursts and 27.8% for weekly bursts. Figure 1 labels this conditional planning curve explicitly. Real limits use per-target C_i; a geometry-only overlap does not preserve the full 300 s sensitivity, and the hypothetical three-scan example is not an empirical comparator.

A regional-pair network spread in longitude is therefore one architecture worth testing: pairs or small clusters within an hour or two of longitude provide coincidence and the differential-Doppler test, and the clusters are distributed around the globe to broaden phase reach for the declared targets and periods. It does not guarantee complete coverage for every source. Which clusters observe which targets on which nights is a scheduling problem over the whole ring, and Equation (1) is its objective function. A network may also corroborate a repeating source through a later independent observation at a predicted phase instead of simultaneous coincidence; that endpoint is reported separately and tested within a shared follow-up budget.

7 4 A reproducible scheduling policy and its limits

The operational policy uses a declared distribution of possible periods, not one guessed extraterrestrial day. An initial scheduler evaluates periods from 1 to 48 h on a logarithmic grid while retaining out-of-range and aperiodic signals as stress tests. At each decision it selects a feasible paired observation that most increases sensitivity-weighted coverage of previously untested phases per additional station-hour, including measured sensitivity, shared usable bandwidth, source elevation, interference occupancy, mask fraction, slew overhead, and buffer availability, and penalizing uncertain station performance rather than assuming ideal behavior. A simple frozen greedy scheduler, of the kind used for Figure 3, is a better first experiment than an opaque optimization system. Save target lists, feasible slots, period weights, scoring rules, tie breaking, seeds, and resource constraints; keep optimization periods separate from a denser evaluation grid; report broad-prior performance and preregistered near-resonance slices with widths declared before testing.

If signals are weak, revisiting the same phase may improve recovery within that phase. For a continuous transmitter, phase diversity provides no advantage. A candidate may justify immediate follow-up, but follow-up consumes the same finite budget as planned observations; reserve an equal allowance for every policy or debit the extra resources explicitly. The strategy is wrong, and this paper will say so, if under the frozen rules the metadata audit finds day-scale, low-duty-cycle coverage already high on the chosen release; if a registered observing trial does not meet its practical recovery gate; if its gain appears only where the baseline was already nearly complete; if the satellite screen deletes more injections or masks more station-time than the locked ceilings; or if a commissioned station cannot meet the noise, mask, or monitor ceilings. In every one of those cases the deliverable is the measurement, not a network.

8 Experiments that could prove the strategy wrong

8 1 Test archive recovery before distributing it

The first benchmark runs on a reference workstation using a predeclared sample of complete ON/OFF cadences from independent observing sessions, including representative quiet and interference-contaminated regions rather than only attractive candidate images. Verify timestamps, units, frequency orientation, missing samples, and calibration before deriving searched exposure. Run the four duration-by-acceptance cells of Section 5 5 and unmodified pinned turboSETI and Pickles on the same inputs, with any learned-ranking ablation kept separate; document each engine's supported signal domain and the losses of its complete filtering chain; compare operating points at matched background-candidate and reviewer budgets with device-hours and data movement included; measure recovery over signal strength, drift, linewidth, duration, occupancy, band edges, and interference overlap. Performance on short windows must not conceal losses for continuous signals, which remain mandatory controls.

Use setigen or another validated generator to insert blinded synthetic signals into real backgrounds, with the Voyager tutorial file as a positive control [19,22]. Split development, threshold calibration, and held-out observations by independent session or target, keeping adjacent cutouts and repeated interference episodes in the same partition; keep held-out labels and seeds unavailable to tuning; include out-of-distribution signal families; disclose the portion of the instrument tested by the injection stage. Power-domain tests cannot validate voltage-domain acquisition claims. Publish recovery curves with block-resampled uncertainty, surviving background candidates per unique searched GHz-hour, review minutes, peak memory, device-hours, transfer bytes, and exclusions. A nominal gain that disappears at matched review burden or depends on leaked test information fails the archive claim.

8 2 Establish the receiver before testing the search

Commissioning begins with measurable checks: receiver noise and gain stability, pointing, antenna response including sidelobes, time stamps, oscillator drift, interference occupancy, mask fraction, false-veto rate, monitor dead time, and the dependence of usable bandwidth on weather, elevation, and local activity. Known astronomical continuum or spectral sources test pointing and bandpass where sensitivity allows; detecting the Galactic hydrogen line demonstrates a useful receiver, not the ability to detect a beacon. Digital injections enter

early enough in the signal path to test the parts of the system being claimed; a synthetic track added after channelization does not test digitizer behavior or earlier excision losses. Bench tests with cabled or shielded injection probe those stages without radiating an unauthorized signal, and injection provenance is recorded so synthetic events cannot be mistaken for detections. Development, calibration, and evaluation data are separated by site and time, not merely by random time-frequency pixels.

8 3 Compare policies under identical constraints

At minimum, compare four policies: the proposed phase-diverse paired revisits; fixed-cadence paired visits; randomized feasible paired visits; and a coverage-first policy that observes independently and spends an identical reserved allowance on confirmation. Add a strongest-available-station policy as a sensitivity benchmark; it is an ablation, not a cost-matched demonstration of network superiority. Resource accounting includes observation, calibration, slews, local computation, network transfer, and follow-up; the primary matched budget is aggregate station-hours. The Figure 3 single-station geometry has similar broad-prior mean coverage across its four policies, but it does not predict equal recovery for these network policies; sensitivity, paired-confirmation costs, interference, and feasible allocation must be measured. Preregistered resonance neighborhoods test whether any benefit is concentrated there. Report recovery separately for the broad benchmark distribution and for preregistered resonance neighborhoods, with widths and weights fixed before evaluation. Test all policies on the same injected source realizations; when counterfactual raw data do not exist for an alternative pointing, do not pretend that replay supplies them. Begin with synchronized recordings of a common target or a validated interference library, then perform a prospective randomized schedule comparison on new sky observations. Population weights are benchmark weights, not a model of the true prevalence of civilizations; report the full recovery surface as well as the weighted summary.

8 4 Measure false alerts in a difficult world

A network must survive correlated interference. Satellites can illuminate separated sites; shared software can manufacture matching artifacts; reference clocks, processing conventions, and common masks can generate correlated failures. Tests therefore include same-time multi-site interference, local intermodulation, clock offsets, gain changes, outages, duplicated files, and intentionally incorrect metadata. Off-target observations, local monitors, the differential-Doppler test, the ephemeris screen, time-shift controls, and additional sites offer complementary checks; time shifts erase real shared interference as well as sky signals, so same-epoch controls are essential. The operational alert-rate denominator is aggregate usable station-hours searched under the specified band and pipeline configuration, with background alerts per wall-clock network-hour also reported. Count every surviving alert against the budget conservatively; separately label confirmed instrumental or terrestrial causes and unresolved cases. Zero observed background alerts does not establish a zero rate: under a Poisson model, zero events over exposure E gives an approximate one-sided 95% upper rate of 3/E, so demonstrating a rate below one per 1,000 station-hours requires roughly 3,000 representative station-hours with zero alerts before correlations and domain shift are considered.

8 5 Freeze the endpoint and publish the failure modes

Recovery differences use paired comparisons across the same source realizations, with uncertainty resampled over independent observing blocks or site-day clusters; thousands of nearly identical injections into one interference episode are not thousands of independent trials. A pilot estimates the variance needed for a sample-size calculation; the stopping rule is set before the held-out result is opened. After the feasibility pilot and registration, the observing design passes its proposed gate only if the one-sided 95% lower bound on improvement exceeds 10 percentage points in the locked primary stratum and the background-alert bound stays within budget; a failed operational gate blocks acceptance without proving that the scientific gain is negative. Inconclusive results warrant a revised experiment, not a discovery announcement. The strongest deliverable is an auditable release: observing-window files, calibration summaries, injection specifications and seeds, blinded and unblinded recovery labels, frozen pipeline and scheduler versions, resource logs, and the analysis needed to reproduce every figure, with pseudonymous station identities and controlled access to

precise geometry. Publication of a negative result would still establish which kinds of networking are not worth building.

9 From a promising signal to possible communication

A scientifically consequential signal needs coordinated observation, a defensible account of alternative explanations, and enough raw information for investigators outside the original team to evaluate it. The first stage is a calibrated pipeline hit. The next is a signal of interest after local interference, harmonics, intermodulation, pointing, and receiver faults have been investigated. A stronger stage is independent confirmation with a different site or instrument, with timing and frequency behavior consistent with the same celestial direction under Equation (8). An accepted extraterrestrial interpretation would require sustained scrutiny of natural and human explanations. None of these stages is interchangeable with a message, and an anomaly-detection model cannot supply the missing inference by labeling something “unusual.”

Repeatability need not mean permanent transmission. A genuinely intermittent source may return at predicted phases, and a network that has recorded which phases it has and has not observed is in a position to predict and test the next arrival. Once evidence warrants it, larger observatories and independent teams can improve sensitivity, localization, spectral resolution, polarization measurements, and temporal coverage; a small dish’s several-degree beam does not isolate an individual star. Information recovery comes next if the signal contains modulation, beginning with measurable structure and testing multiple decoding hypotheses on withheld data; an interpretation that explains only the segment used to invent it is weak evidence of a message. The return direction would require its own link budget [20], and detecting a powerful incoming beacon does not establish that its sender is listening.

This proposal is receive-only. No stage builds, schedules, or authorizes a transmission. The International Academy of Astronautics declaration dated 1 June 2026 replaces the 2010 declaration; it asks the discoverer to authenticate with independent observations by more than one organization using different instrumentation before any public claim, to preserve records, to protect frequencies through the ordinary International Telecommunication Union path, and to treat any reply as a matter for broad international consultation rather than the discoverer’s decision [21]. Its scope explicitly excludes messaging in advance of a confirmed detection, and it is professional guidance rather than binding law or a transmission authorization. This project stops at a documented candidate and a request for independent observing time. Even an authorized reply would not produce a rapid conversation: at 10 pc, light takes about 32.6 years each way, so the earliest physically possible return would arrive about 65 years after our transmission. Contact may first mean learning that another technological intelligence exists; dialogue could be a project spanning generations.

10 A staged project with stop conditions

Table 3 Stages and stop conditions

Stage	Work and evidence	Stop condition
1 Metadata and feasibility	Select a documented release and exact product; reconstruct target windows and band masks; verify one complete cadence; publish geometric coverage	Inadequate data or rights, or no relevant temporal gap in the selected release
2 Archive registration	Freeze duration-by-acceptance cells, complete baselines, injection population, splits, review cost, and software hashes	Incomplete registration; no held-out score is run
3 Archive benchmark	Measure blinded recovery at fixed archival timestamps on a reference workstation; test a bounded bundle prototype	No useful gain at matched review cost, invalid controls, or unreproducible results
4 Participation pilot	Up to ten consenting volunteers for two weeks; replicated jobs and central candidate recomputation	No validated throughput benefit, or a security or privacy incident

Stage	Work and evidence	Stop condition
5 Commissioning and observing pilot	Calibrate one station and a partner; evaluate backgrounds, feasible schedules, variance, and a plausible primary effect; register the prospective trial	Noise or veto ceilings fail, or the proposed recovery gate is infeasible
6 Network trial	Prospective randomized, blinded, resource-matched scheduling comparison within the measured sensitivity range; test an additional longitude cluster where justified	Not run without commissioning, a feasible schedule, and locked endpoints; stop at the registered rule

Stage 3 can end as a pipeline paper, and Stage 4 can end with computation staying central and the volunteer role becoming review. Archive recovery and observing-policy recovery are separate gates: fixed archive timestamps cannot validate a schedule that was never observed. Station construction requires its own feasibility and review decision. Stage 6 publishes the intermittent figure of merit and the alert-rate denominator whether or not a candidate appears. Advancement at every stage requires reproducible candidate and null records, enforceable participant limits, usable metadata, viable retention, and independent scientific review of the methods. Longitude expansion is planned against Equation (10): a campaign that cannot afford hundreds of visits per target for day-scale periods should narrow its period prior, accept partial phase coverage explicitly in its published limits, or concentrate on shorter periods where completeness is affordable.

Null sky observations produce a valuable result. For each declared signal family, publish the observing windows, directional coverage, frequency range, limiting EIRP, duty-cycle assumptions, and injection-calibrated recovery, and place the result on the scale of Equations (2) and (3) so that it can be compared with the published literature and combined with compatible evidence. Repeated visits to one star change its coverage; they are not extra independent civilizations.

11 Operational technology and scope

Table 4 Operational stack

Function	Tool	Role and limit
Archive	Breakthrough Listen open data; blimpy	One cited release; filterbank and HDF5 read path; no new observations in Stages 1 to 4
Drift baseline	turboSETI	Conventional narrowband search; known to miss some morphologies and to lose efficiency at high drift
Comparison filter	Pickles	Independent statistical search on the same observations; preserve its complete selection rules
Intermittency arm	Matched-duration filter bank, Section 5 5	Registered third path on the locked grid; background calibrated empirically
Injections	setigen	Locked population only; Voyager tutorial is the positive control
Ephemeris	CelesTrak GP and TLE; Space-Track; SGP4	Public elements, official catalogue, propagator checked on test vectors
Downlink avoidance	NRAO–SpaceX ODS; Telescope Boresight Avoidance	Used only if a station is accepted; does not quiet unintentional emission
Transmit cross-check	SatNOGS	Known transmissions of catalogued objects; a non-detection clears nothing
Local veto	Broadband reference sensor and SDR	Coincidence veto; a monitor that eats injections is failed

Function	Tool	Role and limit
Celestial-origin test	Differential Doppler, Equation (8)	Separates ground-fixed from celestial; does not localize within the beam
Registration	Open Science Framework	Rules frozen before the held-out split
Volunteer compute	Hashed 128 MiB slices; BOINC-style replication	Reproducible work units; not a scientific result

Language-model tools may be used to draft registration text, script metadata audits, bookkeep injection campaigns, explain computed results, and, as in this paper, perform and check calculations. They do not select candidates, set thresholds, or adjudicate a hit. Declined: optical technosignatures, transmission or active radar, any claim of all-sky coverage, dynamic remote HDF5 subsetting, any score on a data split not named in the registration, and the hope that a megaconstellation avoidance agreement makes the hydrogen band quiet.

12 Limits and interpretation

This design searches a bounded radio-signal family. It cannot establish that extraterrestrial civilizations exist, favor radio, choose the observed frequencies, or direct energy toward Earth. Beacon economics motivates one possible intermittent family without assigning it a known probability. A sufficiently high EIRP does not reveal intent, and the small-dish threshold excludes much of the conceivable signal space.

Section 2 separates published results from a hypothetical homogeneous calculation. Its toy $p = 0.5$ is an assumption. Published all-ON selections, heterogeneous target samples, and duration-dependent sensitivity prevent the toy any-ON model from serving as an empirical reanalysis. Prevalence bounds require calibrated target-level completeness for the actual data and pipeline. The geometric figure of merit is a scheduling summary, not a substitute for that completeness.

The scheduling results are exact phase unions for specified idealized visits, with numerically evaluated period grids and a finite candidate grid for the greedy policy. Figure 3 concerns one station, target geometry, and campaign; Figure 4 compares ideal tiled layouts at exact sidereal resonance. Their finite-grid minima are not universal guarantees. The 35.3% single-site ceiling belongs to the specified equatorial target and elevation cut at exact resonance. The 144-visit campaign curve is an optimistic exposure bound, not an implemented schedule. None includes measured signal recovery, weather, or a complete multi-target allocation.

A sparse network is vulnerable to weak signals, uncertain calibration, shared interference, and incomplete localization. Mandatory two-site detection reduces completeness near the weaker station's threshold. Wider searches increase both computational cost and statistical trials; narrower searches create blind spots. Link budgets must be replaced by measured performance. Satellite vetoes trade retained exposure against interference, and aggregate constellation noise is a limit the project does not solve. Archive reanalysis cannot reconstruct missing observations or recover discarded signal information. Model scores do not measure the probability of extraterrestrial origin. Public availability does not establish redistribution rights or guarantee independent follow-up time.

13 Conclusion

Intermittent-signal searches require an honest account of both when a telescope observed and which signals its complete pipeline would accept. Published continuous-carrier limits cannot be converted to short-burst prevalence limits by dividing by a geometric coverage fraction alone. The first One Sky Alien Array deliverable is therefore an audit and experiment: reconstruct frequency-qualified target windows, measure blinded end-to-end recovery, separate integration duration from ON-scan acceptance, and compare useful recovery at matched review cost.

A later receive-only network would address a narrower, much higher-power signal family. Its purpose and architecture remain conditional on measured sensitivity and a feasible experiment. The geometric examples establish resource tradeoffs: fixed schedules can repeat the same phase near resonances, longitude can broaden

reach in a specified visibility geometry, and paired confirmation costs discovery exposure from the first visit. They do not establish a universal single-site ceiling, a guaranteed scheduler gain, or a completed campaign's prevalence limit.

A favorable result would justify a measured way to use sensitivity-limited stations and a defensible interpretation of their observing record. A negative result would identify where reprocessing, larger apertures, different scheduling, or central computation is more useful. Both outcomes advance the question this proposal can actually answer: how completely have we searched a stated signal family, and what additional observation or analysis would improve that completeness?

Author contributions

Andrew T. Doemling conceived One Sky Alien Array, including the opt-in AI-assisted distributed telescope-archive analysis tool and the shared candidate database, set the research direction, and is responsible for the content of this document. The analysis, calculations, figures, verification of references against primary sources, reproduction program, and text were developed together with the AI research assistants ChatGPT (OpenAI), Claude (Anthropic), Gemini (Google DeepMind), and Grok (xAI). These systems are artificial intelligence tools; they cannot be held accountable for the work in the way that journal and preprint authorship policies require of human authors, and their contributions are acknowledged as assistance rather than human authorship. This attribution concerns the concept and analysis developed in this paper; it does not claim invention of volunteer computing, distributed SETI, satellite screening, or machine-learning technosignature searches, and no implementation, experimental validation, or discovery is asserted.

Acknowledgments and research status

AI research assistance from ChatGPT (OpenAI), Claude (Anthropic), Gemini (Google DeepMind), and Grok (xAI) was used for literature discovery and synthesis, calculations, figures, verification of references, research-design development, and drafting and revision. The document has not undergone independent human peer review. Authorship does not imply that the proposed system has been built or that its methods have been experimentally validated. No new telescope observations, completed archive benchmark, candidate discovery, restricted-data access, or extraterrestrial contact is reported. Dataset selection, release-specific usage rights, software versions, reproducibility, and independent expert review remain necessary before implementation or scientific submission. The single program in Appendix A reproduces the paper's model calculations, tables, and figures. Literature measurements and operational design thresholds are attributed or identified separately; the program does not validate them.

References

[1] Benford, G., Benford, J., and Benford, D. (2010 version). Searching for Cost Optimized Interstellar Beacons. arXiv:0810.3966, first submitted 2008. arXiv 0810.3966

[2] Gray, R. H. (2020). Intermittent Signals and Planetary Days in SETI. International Journal of Astrobiology, 19(4), 299–307. DOI 10.1017/S1473550420000038; open manuscript arXiv 2109.06175

[3] Wright, J. T., Kanodia, S., and Lubar, E. G. (2018). How Much SETI Has Been Done? Finding Needles in the n-Dimensional Cosmic Haystack. The Astronomical Journal, 156, 260. DOI 10.3847/1538-3881/aae099; open manuscript arXiv 1809.07252

[4] Shuch, H. P. (1996). Project Argus and the Challenge of Real-Time All-Sky SETI. Presented at the Fifth International BioAstronomy Conference, Capri, 5 July 1996. Author's primary text

[5] Johnson, O. A., et al. (2023). A Simultaneous Dual-site Technosignature Search Using International LOFAR Stations. The Astronomical Journal, 166, 193. DOI 10.3847/1538-3881/acf9f5; open manuscript arXiv 2310.15704

[6] Suresh, A., et al. (2023). A 4–8 GHz Galactic Center Search for Periodic Technosignatures. The Astronomical Journal, 165, 255. DOI 10.3847/1538-3881/acccf0; open manuscript arXiv 2305.18527

[7] Margot, J.-L., et al. (2023). A Search for Technosignatures Around 11,680 Stars with the Green Bank Telescope at 1.15–1.73 GHz. The Astronomical Journal, 166, 206. DOI 10.3847/1538-3881/acfda4; open manuscript arXiv 2308.02712

[8] Anderson, D. P., Korpela, E. J., Werthimer, D., Cobb, J., and Allen, B. (2025). SETI@home: Data Analysis and Findings. The Astronomical Journal, 170, 111. DOI 10.3847/1538-3881/ade5ab; open manuscript arXiv 2506.14737

[9] BOINC Project. Job replication. Official application-validation documentation, accessed 6 October 2026. BOINC validation documentation

[10] Ma, P. X., et al. (2023). A deep-learning search for technosignatures of 820 nearby stars. Nature Astronomy. DOI 10.1038/s41550-022-01872-z; open manuscript arXiv 2301.12670

[11] Berkeley SETI Research Center. A deep learning search for technosignatures in Breakthrough Listen data. Official study summary and follow-up caveat, accessed 6 October 2026. Study and follow-up summary

[12] Painter, C., et al. (2025 revised manuscript). A Novel Technosignature Search in the Breakthrough Listen Green Bank Telescope Archive. arXiv:2412.05786v2, revised 5 August 2025. arXiv 2412.05786v2

[13] Sheikh, S. Z., et al. (2021). Analysis of the Breakthrough Listen signal of interest blc1 with a technosignature verification framework. Nature Astronomy. Journal article; open manuscript arXiv 2111.06350

[14] Lebofsky, M., et al. (2019; manuscript revised 2020). The Breakthrough Listen Search for Intelligent Life: Public Data, Formats, Reduction and Archiving. Publications of the Astronomical Society of the Pacific. DOI 10.1088/1538-3873/ab3e82; open manuscript arXiv 1906.07391

[15] Berkeley SETI Research Center. Breakthrough Listen Open Data Archive. Official catalogue, accessed 6 October 2026. Official archive

[16] Berkeley SETI Research Center. turboSETI radio technosignature search software. Official repository, accessed 6 October 2026. Official software repository

[17] Condon, J. J., and Ransom, S. M. Essential Radio Astronomy, Chapter 3. National Radio Astronomy Observatory. Radiometer equation and radio receiver sensitivity. NRAO documentation

[18] Berkeley SETI Research Center. blimpy data-format reading and analysis tools. Official repository, accessed 6 October 2026. Official software repository

[19] Brzycki, B., et al. (2022). Setigen: Simulating Radio Technosignatures for SETI. The Astronomical Journal. DOI 10.3847/1538-3881/ac5e3d; open manuscript arXiv 2203.09668

[20] Messerschmitt, D. G. (2013). End-to-end Interstellar Communication System Design for Power Efficiency. arXiv:1305.4684. arXiv 1305.4684

[21] International Academy of Astronautics SETI Permanent Committee (1 June 2026). Declaration of Principles Concerning the Conduct of the Search for Extraterrestrial Intelligence (SETI), 2026 Update. Official declaration; announcement IAA press release 26PR-01

[22] Berkeley SETI Research Center. Breakthrough Listen data introduction and Voyager tutorial. Official project resources, accessed 6 October 2026. Official data introduction; tutorial Voyager tutorial

[23] Enriquez, J. E., et al. (2017). The Breakthrough Listen Search for Intelligent Life: 1.1–1.9 GHz Observations of 692 Nearby Stars. The Astrophysical Journal, 849, 104. DOI 10.3847/1538-4357/aa8d1b; open manuscript arXiv 1709.03491

[24] Price, D. C., et al. (2020). The Breakthrough Listen Search for Intelligent Life: Observations of 1327 Nearby Stars Over 1.10–3.45 GHz. The Astronomical Journal, 159, 86. DOI 10.3847/1538-3881/ab65f1; open manuscript arXiv 1906.07750

[25] Sheikh, S. Z., Wright, J. T., Siemion, A., and Enriquez, J. E. (2019). Choosing a Maximum Drift Rate in a SETI Search: Astrophysical Considerations. The Astrophysical Journal, 884, 14. DOI 10.3847/1538-4357/ab3fa8; open manuscript arXiv 1910.01148

[26] Eastman, J., Siverd, R., and Gaudi, B. S. (2010). Achieving Better Than 1 Minute Accuracy in the Heliocentric and Barycentric Julian Dates. Publications of the Astronomical Society of the Pacific, 122, 935–946. DOI 10.1086/655938; open manuscript arXiv 1005.4415

[27] Astropy Project. Working with Velocities in Astropy Coordinates, section on precision of radial velocity correction. Official documentation, accessed 6 October 2026. Astropy precision documentation

[28] Cordes, J. M., and McLaughlin, M. A. (2003). Searches for Fast Radio Transients. The Astrophysical Journal, 596, 1142–1154. DOI 10.1086/378231; open manuscript arXiv astro-ph/0304364

[29] Hellbourg, G. (2025). Quiet Skies Report: A Primer on Protecting Radio Astronomy in the Age of Satellite Mega-Constellations. arXiv:2512.00941. arXiv 2512.00941

[30] Nhan, B. D., De Pree, C. G., Beasley, A., et al. (2025). ODS: A self-reporting system for radio telescopes to coexist with adaptive satellite constellations. arXiv:2502.15068. arXiv 2502.15068

[31] International Telecommunication Union, Radiocommunication Sector. Recommendation ITU-R RA.769, Protection criteria used for radio astronomical measurements. Used here only as a station noise benchmark. ITU-R RA.769

[32] Center for Open Science. Open Science Framework preregistration. Accessed 6 October 2026. OSF

[33] CelesTrak. General perturbations element sets and supplemental element sets. Operational data source, accessed 6 October 2026. CelesTrak

[34] Space-Track.org, 18th Space Defense Squadron. Official satellite catalogue. Operational data source, accessed 6 October 2026. Space-Track

[35] Vallado, D. A., Crawford, P., Hujsak, R., and Kelso, T. S. (2006). Revisiting Spacetrack Report #3. AIAA/AAS Astrodynamics Specialist Conference, AIAA 2006-6753. Source of the SGP4 reference implementation and test vectors. DOI 10.2514/6.2006-6753

[36] Libre Space Foundation. SatNOGS network and observation database. Operational data source, accessed 6 October 2026. SatNOGS

Appendix A Numerical reproduction program

The following standalone Python program reproduces the hypothetical prevalence table, link budgets, phase-union figures, scheduling summaries, and conditional campaign bounds in this paper. Save the complete listing as reproduce_final.py and run it with Python 3, NumPy, and Matplotlib in a writable directory. The program writes four figure files and a JSON results record. It uses no telescope data and makes no network requests. A fixed random seed, an explicit greedy tie rule, analytic checks, and independent interval-union tests support reproducibility. The dense log-period grid is numerical quadrature, not a proof of a continuous-period minimum. All results remain conditional on the assumptions in the main text; reproduction is not experimental validation.

```
"""One Sky Alien Array: independent numerical reproduction of final Figures 1-4.

Run: python reproduce_final.py --out results
Requires Python 3, NumPy, and Matplotlib. No network or survey data are used.
All outputs are calculations under the stated assumptions, not observations.
The survey example uses 1,000 hypothetical homogeneous any-ON targets.
The campaign curve is an optimistic exposure bound, not a realized schedule.
"""
import os
os.environ.setdefault("MPLCONFIGDIR", "/tmp/onesky-matplotlib")
os.environ.setdefault("XDG_CACHE_HOME", "/tmp/onesky-cache")

# Figures 1 and 4: toy prevalence and paired-cost geometry
from pathlib import Path
import json
import math
import numpy as np
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt

SIDEREAL_H = 23.9344696
VISIT_H = 5.0 / 60.0
TOY_N = 1000
TOY_P = 0.5
COLORS = {'blue': '#2467A6', 'orange': '#BD5429', 'green': '#168073',
          'purple': '#7151A2', 'gold': '#A57A17', 'ink': '#222936'}

def circular_union(intervals, period):
    """Measure a union of real intervals modulo period, divided by period."""
    if period <= 0:
        raise ValueError('Period must be positive')
    pieces = []
    for lower, upper in intervals:
        width = upper - lower
        if width < 0:
            raise ValueError('Interval has negative length')
        if width >= period:
            return 1.0
        if width == 0:
            continue
        left = lower % period
        right = left + width
        if right <= period:
            pieces.append((left, right))
        else:
            pieces.extend(((left, period), (0.0, right - period)))
    if not pieces:
        return 0.0
    pieces.sort()
    total = 0.0
    lo, hi = pieces[0]
    for a, b in pieces[1:]:
        if a <= hi + 1e-13:
            hi = max(hi, b)
        else:
            total += hi - lo
            lo, hi = a, b
    return min(1.0, (total + hi - lo) / period)
```

```
def phase_coverage(starts, period, burst=VISIT_H, visit=VISIT_H,
                   overlap=VISIT_H / 2):
    """Opportunity for >=overlap of a single burst in any individual visit."""
    if not (0 < overlap <= min(burst, visit)) or burst > period:
        raise ValueError('Require 0 < overlap <= min(burst, visit) and burst <= P')
    return circular_union([(a + overlap - burst, a + visit - overlap)
                           for a in starts], period)

def toy_cadence_coverage(period, burst_minutes):
    burst = burst_minutes / 60
    overlap = min(burst / 2, 2.5 / 60)
    return phase_coverage([0.0, 10 / 60, 20 / 60], period,
                          burst=burst, overlap=overlap)

def homogeneous_root(q, n=TOY_N, null_probability=0.05):
    """Unclipped prevalence solving (1 - f*q)**n = null_probability.
    A result >1 means no solution within the physical prevalence interval.
    It is not an upper bound above 100% and is not a posterior probability.
    """
    if not (0 < null_probability < 1) or n < 1 or not (0 <= q <= 1):
        raise ValueError('Invalid model inputs')
    return math.inf if q == 0 else -math.expm1(math.log(null_probability) / n) / q

def prevalence_limit(completenesses, alpha=0.05):
    """Invert the independent-target zero-detection likelihood.
    q_i includes all visits, frequency, sensitivity and acceptance rules.
    None means no upper bound below one.
    """
    q = np.asarray(completenesses, dtype=float)
    if np.any((q < 0) | (q > 1)) or not (0 < alpha < 1):
        raise ValueError('Invalid completeness or alpha')
    target = math.log(alpha)
    def log_null(f):
        with np.errstate(divide='ignore'):
            return float(np.log1p(-f * q).sum())
    if log_null(1.0) > target:
        return None
    lo, hi = 0.0, 1.0
    for _ in range(80):
        mid = (lo + hi) / 2
        if log_null(mid) > target:
            lo = mid
        else:
            hi = mid
    return (lo + hi) / 2

def campaign_coverage_ceiling(period, visits=144):
    """Exposure-only upper bound; no feasible campaign is asserted."""
    return min(1.0, visits * VISIT_H / period)

def visibility_hours(latitude=40, declination=0, minimum_elevation=20):
    phi, dec, elev = map(math.radians, (latitude, declination, minimum_elevation))
    x = (math.sin(elev) - math.sin(phi) * math.sin(dec)) / (math.cos(phi) * math.cos(dec))
    return SIDEREAL_H * math.acos(x) / math.pi

def tiled_local_starts(number, duration):
    """Tile a visibility interval; boundary-align the final tile, then repeat."""
    if duration < VISIT_H:
        return []
    return [min(j * VISIT_H, duration - VISIT_H) for j in range(number)]

def layout_coverage(nights, layout):
    """At P=D, tile visible phase with one visit/station/night, six stations.
    Co-longitude pairs share a region; regional pairs tile their intersections.
    Each shared paired opportunity costs two station-visits.
    """
    visible = visibility_hours()
    intervals = []
    if layout == 'one_longitude_pairs':
        local = tiled_local_starts(3 * nights, visible)
        intervals = [(-visible / 2 + x - VISIT_H / 2,
                      -visible / 2 + x + VISIT_H / 2) for x in local]
    elif layout == 'regional_pairs':
        common = visible - 1.0
        for pair in range(3):
```

```
            center = pair * SIDEREAL_H / 3
            for x in tiled_local_starts(nights, common):
                a = center - common / 2 + x
                intervals.append((a - VISIT_H / 2, a + VISIT_H / 2))
    elif layout == 'independent_sites':
        for station in range(6):
            center = station * SIDEREAL_H / 6
            for x in tiled_local_starts(nights, visible):
                a = center - visible / 2 + x
                intervals.append((a - VISIT_H / 2, a + VISIT_H / 2))
    else:
        raise ValueError('Unknown layout')
    return circular_union(intervals, SIDEREAL_H)

def style():
    plt.rcParams.update({'font.family': 'DejaVu Sans', 'font.size': 9,
        'axes.spines.top': False, 'axes.spines.right': False,
        'axes.labelcolor': COLORS['ink'], 'axes.edgecolor': '#647080',
        'xtick.color': '#47515F', 'ytick.color': '#47515F',
        'axes.grid': True, 'grid.color': '#E6E9ED', 'grid.linewidth': 0.55,
        'legend.frameon': False, 'savefig.facecolor': 'white'})

def save_figure(fig, directory, stem):
    for suffix in ('png', 'svg', 'pdf'):
        fig.savefig(directory / (stem + '.' + suffix), dpi=300,
                    bbox_inches='tight', pad_inches=0.12)
    plt.close(fig)

def make_coverage_outputs(directory):
    directory = Path(directory)
    directory.mkdir(parents=True, exist_ok=True)
    style()
    rows = []
    for days in (1, 2, 7, 27):
        row = {'period_days': days, 'families': {}}
        for burst in (5, 20, 60):
            g = toy_cadence_coverage(days * 24, burst)
            f95 = homogeneous_root(TOY_P * g)
            f50 = homogeneous_root(TOY_P * g, null_probability=0.5)
            row['families'][str(burst)] = {'coverage': g,
                'f95': f95 if f95 <= 1 else None, 'unclipped_f95_root': f95,
                'even_odds_f': f50 if f50 <= 1 else None,
                'unclipped_even_odds_root': f50,
                'null_probability_at_f1': math.exp(TOY_N * math.log1p(-TOY_P * g))}
        rows.append(row)
    short_daily_g = toy_cadence_coverage(24, 5)
    continuous = homogeneous_root(TOY_P)
    result = {'conventions': {'toy_N': TOY_N, 'conditional_recovery_p': TOY_P,
        'on_scan_starts_minutes': [0, 10, 20], 'scan_minutes': 5,
        'overlap_rule_minutes': 'min(B/2, 2.5)',
        'model_status': 'hypothetical homogeneous any-ON illustration'},
        'continuous_f95': continuous,
        'continuous_even_odds_f': homogeneous_root(TOY_P, null_probability=0.5),
        'toy_rows': rows,
        'short_burst_crossover_days': {
            'f95_reaches_one': TOY_P * short_daily_g / (-math.expm1(math.log(.05) / TOY_N)),
            'null_probability_at_f1_reaches_half': TOY_P * short_daily_g / (
                -math.expm1(math.log(.5) / TOY_N))},
        'campaign': {'targets': 300, 'paired_visits_per_target': 144,
            'assumed_joint_recovery': .5, 'station_exposure_hours': 7200,
            'status': 'exposure ceiling; f95 values are optimistic lower envelopes',
            'rows': []}}
    for days in (1, 2, 7, 27):
        g = campaign_coverage_ceiling(days * 24)
        result['campaign']['rows'].append({'period_days': days, 'coverage_ceiling': g,
            'optimistic_f95_root': homogeneous_root(.5 * g, n=300)})
    days = np.geomspace(.125, 100, 900)
    fig, axes = plt.subplots(1, 2, figsize=(8.4, 3.65), sharey=True)
    for burst, color in ((5, 'blue'), (20, 'orange'), (60, 'green')):
        roots = [100 * homogeneous_root(TOY_P * toy_cadence_coverage(d * 24,
            burst)) for d in days]
        axes[0].plot(days, roots, color=COLORS[color], lw=1.8, label=f'{burst}-min burst')
    axes[0].axhline(100 * continuous, color='#707A87', ls=':', lw=1.2,
                    label='Continuous: 0.598%')
    axes[0].set_title(('(a) Hypothetical any-ON survey\n'
```

```
    '1,000 equivalent targets; assumed p = 0.5'), loc='left', fontsize=9)
axes[0].set_ylabel('95% prevalence upper limit (%)')
axes[0].legend(loc='lower right', bbox_to_anchor=(1.0, .12), fontsize=8)
campaign_roots = [100 * homogeneous_root(.5 * campaign_coverage_ceiling(d * 24),
    n=300) for d in days]
axes[1].plot(days, campaign_roots, color=COLORS['purple'], ls='--', lw=2,
            label='Optimistic lower envelope for f95')
axes[1].scatter([1, 7], [3.9744327778, 27.8210294445], s=20,
                color=COLORS['purple'], zorder=3)
axes[1].annotate('Daily: 3.97%', (1, 3.9744327778), xytext=(1.25, 2.6), fontsize=8)
axes[1].annotate('Weekly: 27.8%', (7, 27.8210294445), xytext=(9.3, 15), fontsize=8)
axes[1].set_title(('(b) Campaign planning bound\n'
    '300 targets × 144 paired visits; assumed joint p = 0.5'), loc='left', fontsize=9)
axes[1].text(.04, .11, ('Requires maximal phase union.\n'
    'A feasible schedule is not demonstrated.\n'
    'Separate instrument and power threshold.'),
            transform=axes[1].transAxes, fontsize=8, color='#46505E')
for ax in axes:
    ax.set_xscale('log'); ax.set_yscale('log')
    ax.set_xlim(.125, 100); ax.set_ylim(.4, 100)
    ax.set_xticks([.125, 1, 7, 27, 100], ['3 h', '1 d', '7 d', '27 d', '100 d'])
    ax.set_yticks([.5, 1, 5, 10, 50, 100], ['0.5', '1', '5', '10', '50', '100'])
    ax.minorticks_off()
    ax.set_xlabel('Beacon period')
    ax.axhline(100, color='#596472', lw=.8)
fig.subplots_adjust(wspace=.15, bottom=.17, top=.89)
save_figure(fig, directory, 'figure_1_toy_prevalence')
names = ['one_longitude_pairs', 'regional_pairs', 'independent_sites']
budgets = np.arange(0, 721, 6)
coverage = {name: [layout_coverage(int(b // 6), name) for b in budgets] for name in names}
visible = visibility_hours()
ceilings = {'one_longitude_pairs': visible / SIDEREAL_H,
    'regional_pairs': 3 * (visible - 1) / SIDEREAL_H, 'independent_sites': 1.0}
result['layouts'] = {'sidereal_period_hours': SIDEREAL_H, 'visibility_hours': visible,
    'continuous_envelopes': ceilings, 'coverage_at_station_budgets': {
        str(b): {name: layout_coverage(b // 6, name) for name in names}
        for b in (180, 204, 288, 360, 534, 540, 720)},
    'construction': ('full visits; final tile aligned to visibility boundary; '
                     'repeated after saturation')}
fig, ax = plt.subplots(figsize=(7.25, 4.15))
labels = [('independent_sites', 'blue', '6 independent sites: any-site opportunity'),
    ('regional_pairs', 'orange', '3 regional pairs: simultaneous paired opportunity'),
    ('one_longitude_pairs', 'gold', '3 co-longitude pairs: simultaneous paired opportunity')]
for name, color, label in labels:
    ax.plot(budgets, 100 * np.asarray(coverage[name]), color=COLORS[color], lw=2,
        label=label)
    ax.axhline(100 * ceilings[name], color=COLORS[color], ls=':', lw=.8, alpha=.75)
ax.scatter([180, 180], [62.6711, 31.3356], s=22, color=[COLORS['blue'],
    COLORS['orange']], zorder=4)
ax.annotate('Same cost: 62.7% any-site\nor 31.3% paired opportunity',
    (180, 62.6711), xytext=(205, 69), fontsize=8,
    arrowprops={'arrowstyle': '-', 'color': '#788290', 'lw': .8})
ax.text(712, 35.2679 + 1.4, '35.3% ceiling', ha='right', color=COLORS['gold'], fontsize=8)
ax.text(712, 93.2695 + 1.4, '93.3% ceiling', ha='right', color=COLORS['orange'], fontsize=8)
ax.set_xlim(0, 720); ax.set_ylim(0, 106)
ax.set_xlabel('Total 5-minute station-visits')
ax.set_ylabel('Geometric phase coverage at one sidereal day (%)')
ax.legend(loc='lower right', fontsize=8)
top = ax.secondary_xaxis('top', functions=(lambda b: b / 6, lambda n: n * 6))
top.set_xlabel('Nights: one visit per station per night', fontsize=8)
top.tick_params(labelsize=8)
fig.subplots_adjust(top=.83, bottom=.16, left=.12, right=.98)
save_figure(fig, directory, 'figure_4_layout_cost')
result['visibility_network_examples'] = {
    'single_site_at_P_D': min(1., visible / SIDEREAL_H),
    'single_site_at_P_D_over_2': min(1., 2 * visible / SIDEREAL_H),
    'two_sites_6h_apart_any_site': circular_union(
        [(-visible / 2, visible / 2), (6 - visible / 2, 6 + visible / 2)], SIDEREAL_H),
    'two_sites_6h_apart_simultaneous': max(0., visible - 6) / SIDEREAL_H,
    'three_sites_D_over_3_apart_any_site': circular_union(
        [(j * SIDEREAL_H / 3 - visible / 2, j * SIDEREAL_H / 3 + visible / 2)
         for j in range(3)], SIDEREAL_H),
    'three_sites_D_over_3_apart_at_least_two_visible':
        3 * max(0., visible - SIDEREAL_H / 3) / SIDEREAL_H,
    'regional_pairs_envelope_by_internal_separation_h': {
```

```
            str(h): min(1., 3 * max(0., visible - h) / SIDEREAL_H) for h in (1, 2, 3)}}
    result['minimum_integer_visits'] = {str(h): {str(c): math.ceil(c * h / VISIT_H)
        for c in (.5, .9)} for h in (12, 24, 48)}
    (directory / 'coverage_results.json').write_text(json.dumps(result, indent=2) + '\n')
    return result

def run_coverage_tests():
    assert abs(circular_union([(0.9, 1.1)], 1) - .2) < 1e-12
    assert abs(circular_union([(0, .3), (.2, .4)], 1) - .4) < 1e-12
    assert circular_union([], 1) == 0
    assert circular_union([(0, 2)], 1) == 1
    for d in (1, 2, 7, 27):
        for b, numerator_min in ((5, 15), (20, 40), (60, 80)):
            assert abs(toy_cadence_coverage(24 * d, b) - numerator_min / (1440 * d)) < 1e-12
    assert abs(prevalence_limit([.5] * 1000) - homogeneous_root(.5)) < 1e-12
    assert prevalence_limit([1e-6] * 1000) is None
    assert abs(prevalence_limit([1.0]) - .95) < 1e-12
    for n in range(121):
        assert (abs(layout_coverage(n, 'one_longitude_pairs') - min(3 * n * VISIT_H,
            visibility_hours()) / SIDEREAL_H) < 1e-10)
        assert (abs(layout_coverage(n, 'regional_pairs') - 3 * min(n * VISIT_H,
            visibility_hours() - 1) / SIDEREAL_H) < 1e-10)
        assert (abs(layout_coverage(n, 'independent_sites') - min(6 * n * VISIT_H
            / SIDEREAL_H, 1)) < 1e-10)
    return True

# Figures 2 and 3: aliasing and scheduling, plus physical planning values
import csv
from matplotlib.ticker import NullFormatter

SIDEREAL_H = 23.9344696
VISIT_H = BURST_H = 5 / 60
SEED = 20261006
SCHEDULE_COLORS = ("#2563A6", "#C95020", "#278570", "#8050A1")

def schedule_phase_coverage(starts, periods, visit=VISIT_H, burst=BURST_H, tau=None):
    """Any-scan coverage from circular gaps, max(delta_phase-width,0).
    Admitted interval: [s+tau-B,s+V-tau]; tau=min(B/2,V/2) by default.
    """
    periods = np.atleast_1d(np.asarray(periods, dtype=float))
    times = np.asarray(starts, dtype=float)
    tau = min(burst / 2, visit / 2) if tau is None else tau
    if (np.any(periods <= 0) or visit <= 0 or burst <= 0
            or not 0 < tau <= min(visit, burst)):
        raise ValueError("Require P,V,B>0 and 0<tau<=min(V,B).")
    if times.size == 0:
        return np.zeros(periods.shape)
    width = visit + burst - 2 * tau
    phases = np.sort(np.remainder(times[None, :], periods[:, None]), axis=1)
    gaps = np.diff(np.column_stack((phases, phases[:, 0] + periods)), axis=1)
    return np.clip(1 - np.maximum(gaps - width, 0).sum(axis=1) / periods, 0, 1)

def schedule_visibility_hours(latitude=40., declination=0., elevation=20.):
    phi, dec, el = np.deg2rad([latitude, declination, elevation])
    q = (np.sin(el) - np.sin(phi) * np.sin(dec)) / (np.cos(phi) * np.cos(dec))
    return float(SIDEREAL_H * np.arccos(np.clip(q, -1, 1)) / np.pi)

def schedule_policies():
    """30 complete scans, one per visibility cycle, with fixed resources."""
    n = 30
    visible = schedule_visibility_hours()
    transits = 12 + np.arange(n) * SIDEREAL_H
    lo, hi = transits - visible / 2, transits + visible / 2 - VISIT_H
    design = np.geomspace(1., 48., 160)
    policies = {"fixed": transits - VISIT_H / 2,
                "sweep": np.linspace(lo[0], hi[-1], n),
                "random": np.random.default_rng(SEED).uniform(lo, hi)}
    chosen = []
    for a, b in zip(lo, hi):
        candidates = a + np.arange(math.floor((b - a) / VISIT_H + 1e-9) + 1) * VISIT_H
        scores = np.array([schedule_phase_coverage(chosen + [t],
            design).mean() for t in candidates])
        earliest_best = np.flatnonzero(scores >= scores.max() - 1e-12)[0]
        chosen.append(float(candidates[earliest_best]))
    policies["greedy"] = np.asarray(chosen)
```

```
    for values in policies.values():
        assert np.all(values >= lo - 1e-9) and np.all(values <= hi + 1e-9)
    return policies, design

def schedule_save_figure(fig, output, name):
    for suffix in ("png", "svg"):
        fig.savefig(output / f"{name}.{suffix}", dpi=320, facecolor="white")
    plt.close(fig)

def figures(output):
    plt.rcParams.update({"font.family": "DejaVu Sans", "font.size": 8.5,
        "axes.spines.top": False, "axes.spines.right": False,
        "axes.grid": True, "grid.color": "#E4E4E4", "grid.linewidth": .5,
        "axes.axisbelow": True, "legend.frameon": False,
        "svg.hashsalt": "onesky-final"})
    counts = np.arange(1, 9)
    repeated = 12 + np.arange(8) * 24
    distinct = repeated + np.arange(8) * 3
    same = [100 * schedule_phase_coverage(repeated[:n], [24])[0] for n in counts]
    spread = [100 * schedule_phase_coverage(distinct[:n], [24])[0] for n in counts]
    fig, ax = plt.subplots(figsize=(6.4, 2.8), layout="constrained")
    ax.plot(counts, spread, "o-", color=SCHEDULE_COLORS[0], markersize=4,
            label="Distinct, nonoverlapping phase windows")
    ax.plot(counts, same, "s--", color=SCHEDULE_COLORS[1], markersize=4,
            label="Repeated identical phase window")
    for value, col in ((spread[-1], SCHEDULE_COLORS[0]), (same[-1], SCHEDULE_COLORS[1])):
        ax.text(8.12, value, f"{value:.3f}%", va="center", color=col, fontsize=8)
    ax.set(xlim=(.8, 9), ylim=(0, 3.25), xticks=counts,
           xlabel="Paired visits (10 station-minutes each)",
           ylabel="Geometric phase coverage (%)")
    ax.legend(loc="upper left", fontsize=8)
    ax.text(.98, .56, "P = 24 h; B = V = 5 min\nMinimum overlap = 2.5 min",
            transform=ax.transAxes, ha="right", fontsize=8)
    schedule_save_figure(fig, output, "figure2_eight_visits")
    policies, design = schedule_policies()
    alias = SIDEREAL_H + (schedule_visibility_hours() - VISIT_H) / 29
    # Direct quadrature at 160,000 log-bin midpoints; no interpolation and no
    # design-grid points. Extra resonance points are for plotting only.
    evaluation = np.exp((np.arange(160000) + .5) * math.log(48) / 160000)
    dense = np.geomspace(1, 48, 4000)
    neighborhoods = [np.linspace(center * .985, center * 1.015, 400)
                     for k in (1, 2, 3, 4) for center in (SIDEREAL_H / k, 24 / k)]
    broad = np.unique(np.concatenate([dense, [SIDEREAL_H, 24, alias]] + neighborhoods))
    zoom = np.unique(np.r_[np.linspace(23.8, 24.4, 1801), SIDEREAL_H, alias])
    exact = {key: schedule_phase_coverage(times, evaluation) for key, times in policies.items()}
    curves = {key: schedule_phase_coverage(times, broad) for key, times in policies.items()}
    close = {key: schedule_phase_coverage(times, zoom) for key, times in policies.items()}
    styles = {"fixed": ("Fixed at transit", SCHEDULE_COLORS[0], "-", .65),
              "sweep": ("Linear sweep", SCHEDULE_COLORS[1], "-", .65),
              "random": ("Random (fixed seed)", SCHEDULE_COLORS[2], "--", 1.0),
              "greedy": ("Phase-diverse greedy", SCHEDULE_COLORS[3], "-", 1.45)}
    fig, axes = plt.subplots(1, 2, figsize=(6.5, 3.3), layout="constrained",
                             gridspec_kw={"width_ratios": [1.18, 1]})
    for key, (label, color, ls, width) in styles.items():
        axes[0].plot(broad, 100 * curves[key], label=label, color=color,
                     ls=ls, lw=width, alpha=.9)
        axes[1].plot(zoom, 100 * close[key], color=color, ls=ls, lw=max(width, 1))
    axes[0].set(xscale="log", xlim=(1, 48), ylim=(0, 102),
                xlabel="Beacon period (h)", ylabel="Geometric phase coverage (%)")
    axes[0].set_xticks([1, 2, 4, 8, 12, 24, 48], labels=[1, 2, 4, 8, 12, 24, 48])
    axes[0].xaxis.set_minor_formatter(NullFormatter())
    axes[0].legend(fontsize=7.2, loc="upper right", handlelength=2)
    axes[0].set_title("(a) Full 1–48 h interval", loc="left", fontsize=9)
    axes[1].set(xlim=(23.8, 24.4), ylim=(0, 11.3), xlabel="Beacon period (h)")
    axes[1].set_title("(b) Near one sidereal day", loc="left", fontsize=9)
    axes[1].set_xticks([23.8, 24.0, 24.2, 24.4])
    for x, label, y, color in ((SIDEREAL_H, "Fixed notch", 2.2, SCHEDULE_COLORS[0]),
                               (alias, "Sweep notch", 3.6, SCHEDULE_COLORS[1])):
        axes[1].annotate(label, xy=(x, 100 * VISIT_H / x), xytext=(x + .018, y),
                         fontsize=7, color=color, rotation=0,
                         arrowprops={"arrowstyle": "-", "lw": .6, "color": color})
    schedule_save_figure(fig, output, "figure3_policies")
    summary = {}
    for key, times in policies.items():
        summary[key] = {"mean_coverage_pct": float(100 * exact[key].mean()),
```

```
            "evaluation_min_coverage_pct": float(100 * exact[key].min()),
            "evaluation_min_period_h": float(evaluation[np.argmin(exact[key])]),
            "evaluation_median_coverage_pct": float(100 * np.median(exact[key])),
            "evaluation_fraction_below_half_greedy_pct": float(100
                * np.mean(exact[key] < .5 * exact["greedy"])),
            "at_sidereal_period_pct": float(100 * schedule_phase_coverage(times,
                [SIDEREAL_H])[0]),
            "at_sweep_alias_pct": float(100 * schedule_phase_coverage(times, [alias])[0]),
            "zoom_min_pct": float(100 * close[key].min()),
            "zoom_max_pct": float(100 * close[key].max()),
            "scan_starts_hours": times.tolist()}
        mask = exact[key] < .5 * exact["greedy"]
        summary[key]["own_median_in_below_half_greedy_trough_pct"] = (
            float(100 * np.median(exact[key][mask])) if mask.any() else None)
    return {"figure2": {"same_phase_eight_pct": same[-1],
                        "distinct_phase_eight_pct": spread[-1], "station_minutes": 80},
            "figure3": {"visibility_hours": schedule_visibility_hours(),
                "sweep_alias_hours": alias,
                        "n_visits": 30, "station_minutes": 150, "random_seed": SEED,
                        "sidereal_greedy_minus_fixed_pp": (
                            summary["greedy"]["at_sidereal_period_pct"]
                            - summary["fixed"]["at_sidereal_period_pct"]),
                        "sidereal_greedy_over_fixed": (
                            summary["greedy"]["at_sidereal_period_pct"]
                            / summary["fixed"]["at_sidereal_period_pct"]),
                        "design_periods": 160, "evaluation_log_bin_midpoints": 160000,
                        "greedy_tie_tolerance": 1e-12, "policies": summary}}

def planning_values():
    k_b, pc, light, freq = 1.380649e-23, 3.0856776e16, 299792458., 1.42e9
    omega, radius = 7.2921159e-5, 6.371e6
    table = []
    for diameter, efficiency, temperature in [(1.2,.6,100),(3,.6,100),(6,.6,100),
                                              (25,.65,35),(100,.7,25)]:
        area = efficiency * math.pi * diameter**2 / 4
        sefd = 2 * k_b * temperature / area * 1e26
        flux = 10 * sefd * 1e-26 / math.sqrt(2 * 300)
        eirp = 4 * math.pi * (10 * pc)**2 * flux
        table.append({"diameter_m": diameter, "aperture_efficiency": efficiency,
            "system_temperature_K": temperature, "SEFD_Jy": sefd,
            "beam_deg": math.degrees(1.2 * light / freq / diameter),
            "EIRP_10pc_W": eirp, "ratio_to_Arecibo_2e13_W": eirp / 2e13})
    latitude, separation = np.deg2rad([40, 15])
    def delta_frequency(h, dec):
        return (freq * omega * radius / light * math.cos(latitude) * math.cos(dec)
                * (math.sin(h - separation / 2) - math.sin(h + separation / 2)))
    changes = {}
    half_beam = math.radians(2.4)
    for deg in (0, 45):
        dec = math.radians(deg)
        deviations = [delta_frequency(dh, dec + dd) - delta_frequency(0, dec)
                      for dh in (-half_beam, 0, half_beam)
                      for dd in (-half_beam, 0, half_beam)]
        changes[f"declination_{deg}_deg"] = {"midpoint_signed_offset_Hz": delta_frequency(0,
            dec),
                "direction_box_offset_change_Hz": [min(deviations), max(deviations)]}
    base = delta_frequency(0, 0)
    small = table[1]["EIRP_10pc_W"]
    return {"table2_link_budget": table,
        "planning": {"three_m_EIRP_W_at_pc": {str(d): small * (d / 10)**2 for d in (10,100,
            1000)},
                "three_m_to_10Jy_ratio_dB": 10 * math.log10(table[1]["SEFD_Jy"] / 10),
                "SEFD_ratio_1p2_to_3m": table[0]["SEFD_Jy"] / table[1]["SEFD_Jy"],
                "threshold_multiplier_150s": math.sqrt(2),
                "threshold_multiplier_capture_0p7": 1/.7,
                "threshold_multiplier_1200s": .5},
        "doppler": {"prefactor_Hz": freq * omega * radius / light,
            "symmetric_equatorial_offset_amplitude_Hz": abs(base),
            "differential_drift_amplitude_Hz_per_s": abs(base) * omega,
            "midpoint_instantaneous_drift_Hz_per_s": 0.,
            "oscillator_1e_minus_11_Hz": freq * 1e-11, **changes},
        "ideal_matched_duration": {"gain_vs_300s": {str(b): math.sqrt(300/b) for b in (10,
            20,60,150,300)},
            "illustrative_trials": 48, "illustrative_q10_with_trials": math.sqrt(100+2
                *math.log(48)),
```

```
        "qualification": ("Ideal gains and illustrative Gaussian trials only; "
                          "native archive bins constrain feasible windows.")}}

def self_test():
    # Independent interval splitting/merging validates wraparound and the
    # circular-gap implementation, including long bursts and empty scans.
    def split_union(times, period, visit, burst):
        tau = min(visit/2, burst/2)
        width = visit + burst - 2*tau
        if len(times) == 0:
            return 0.
        if width >= period:
            return 1.
        pieces = []
        for time in times:
            left = (time + tau - burst) % period
            pieces.extend([(left, min(period, left+width))])
            if left + width > period:
                pieces.append((0., left+width-period))
        total, end = 0., 0.
        for a, b in sorted(pieces):
            total += max(0., b-max(end,a))
            end = max(end,b)
        return total / period
    rng = np.random.default_rng(113)
    for _ in range(200):
        period = rng.uniform(.1, 30)
        burst, visit = rng.uniform(.01, period, 2)
        times = rng.uniform(-1000, 1000, rng.integers(0, 15))
        actual = schedule_phase_coverage(times, [period], visit, burst)[0]
        assert abs(actual - split_union(times, period, visit, burst)) < 5e-12
    assert np.isclose(schedule_phase_coverage([0], [24])[0], 5/1440)
    assert np.isclose(schedule_phase_coverage([0, 24], [24])[0], 5/1440)
    assert np.isclose(schedule_phase_coverage([0, 1/6, 1/3], [24], burst=1)[0], 80/1440)
    assert np.isclose(schedule_visibility_hours(), 8.441184114788822)
    assert (np.isclose(planning_values()["table2_link_budget"][1]["EIRP_10pc_W"],
        3.180277170233066e14))

def build_schedule_artifacts(output=None):
    output = Path(output or Path(__file__).parent)
    output.mkdir(parents=True, exist_ok=True)
    self_test()
    results = {**figures(output), **planning_values(),
        "self_tests": ("passed: 200 randomized union comparisons "
                       "plus analytic checks")}
    (output / "schedule_results.json").write_text(json.dumps(results, indent=2)+"\n")
    with (output / "table2_link_budget.csv").open("w", newline="") as handle:
        rows = results["table2_link_budget"]
        writer = csv.DictWriter(handle, fieldnames=rows[0].keys())
        writer.writeheader(); writer.writerows(rows)
    return results


def main():
    import argparse
    parser = argparse.ArgumentParser(description=__doc__)
    parser.add_argument('--out', default='onesky_results', help='Output directory')
    output = Path(parser.parse_args().out)
    output.mkdir(parents=True, exist_ok=True)
    run_coverage_tests()
    coverage = make_coverage_outputs(output)
    scheduling = build_schedule_artifacts(output)
    results = {**coverage, **scheduling}
    (output / 'numerical_results.json').write_text(json.dumps(results, indent=2) + '\n')
    print(json.dumps(results, indent=2))

if __name__ == '__main__':
    main()
```